

THE GOLDEN GATE BRIDGE - DEDICATED MAY 27, 1937
CHIEF ENGINEER JOSEPH B. STRAUSS
APC International Publishing Corporation

Beam Theory: Architecture for Cells and Skeletons

10

Overview: In which we apply the theory of elastic rods to a diverse array of biological filaments

Many of the macromolecules of living organisms are filamentous. Not only are they a striking visual feature within cells, but their structure is intimately tied to the ways in which such molecules are used in cells. The representation of geometric structures as networks of one-dimensional elements is a perspective of great power and applicability. That part of mechanics which has grown up around this approximation is known traditionally as “beam theory.” Historically, the study of the structural mechanics of beams has been equally rewarding whether applied to the flying buttresses of Notre Dame or to the wings of the jumbo jet that carried the tourist to Paris. The theme of the present chapter is the observation that the biological world has put one-dimensional beams to similar architectural uses. The physical ideas introduced here all amount to arrangements of springs (that is, Hooke’s law) and permit us to examine a range of seemingly unrelated problems such as DNA bending during transcriptional regulation, packaging of DNA in viruses and in the eukaryotic nucleus, and the properties of the cytoskeleton.

“I demolish my bridges behind me - then there is no choice but forward.”

Fridtjof Hansen

10.1 Beams Are Everywhere: From Flagella to the Cytoskeleton

One-Dimensional Structural Elements Are the Basis of Much of Macromolecular and Cellular Architecture

When we reflect on structures such as trees, animal skeletons, the cytoskeletal networks of cells, or the molecules that make up Crick’s “two great polymer languages,” these structures are all based upon elements with one dimension that is much larger than the other two.

Geometric structures with this property are known as beams (or rods) and have interesting mechanical features that affect their biological function. Figure 10.1 shows a variety of different examples of structures that can be viewed mechanically as one-dimensional beams. Figure 10.1(A) is a reminder of the role of beams as structural elements in conventional architecture. A much smaller example that has been the basis of impressive single-molecule studies of biological molecules is the atomic-force microscope based upon cantilevers like that shown in Figure 10.1(B).

In this chapter, we focus on the interesting biological examples of structural elements that can be thought of as elastic beams. In particular, Figures 10.1(C–G) reveal a number of examples of this kind. Figure 10.1(C) illustrates the network of one-dimensional filaments that make up the cytoskeleton of a typical eukaryotic cell. Another beautiful example is that of the hair cells of the mammalian inner ear as shown in Figure 10.1(D). These hair cells have slender protrusions, tens of microns in length, known as stereocilia, which vibrate in response to sound. Figure 10.1(E) suggests that the flagella that propel cells ranging from *E. coli* to eukaryotic sperm can be viewed as elastic rods. Flagella like that shown in the figure can be as much as 10 times longer than the cell bodies that they are used to propel, while their cross-section can be orders of magnitude smaller. Figures 10.1(F) and (G) show the way in which individual polymers (microtubules and DNA, respectively) have geometries that are very long in comparison with their width.

In the sections to follow, we develop the quantitative framework for thinking about the structure and energetics of biological structures that can be approximated as one-dimensional beams. Two of the most important case studies involve the deformations that DNA

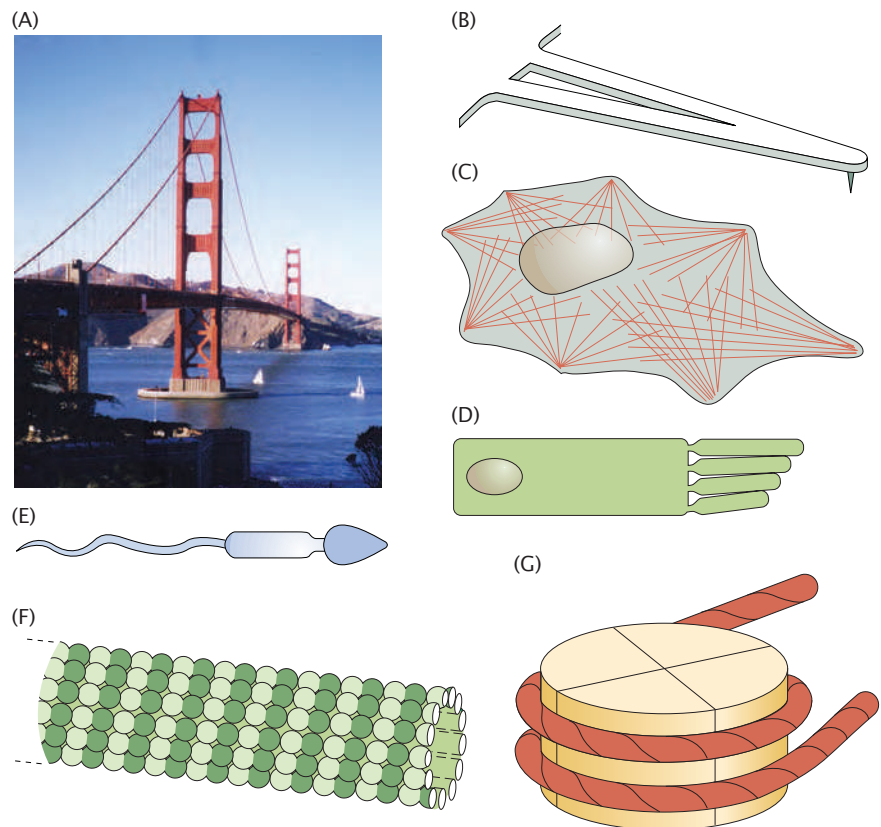


Figure 10.1: Diverse examples of the way in which structures can be interpreted using beam theory. (A) A bridge as a collection of beams and cables. (B) A small-scale cantilever used in an atomic-force microscope. (C) The cytoskeleton in a eukaryotic cell. (D) Stereocilia on an inner ear hair cell. (E) The flagellum of a sperm. (F) An individual microtubule. (G) A representation of DNA as an elastic rod in the context of the nucleosome.

is subjected to both in the gene regulatory setting as a result of interactions with DNA-binding proteins and in the DNA-packaging context. DNA packaging, whether in viruses, prokaryotes, or eukaryotes, involves compaction of huge DNA molecules into confined spaces with an attendant energy cost associated with elastic bending. A second class of case studies will involve the mechanics of the cytoskeleton when considered as a collection of elastic beams. One of the interesting observations offered by physical biology is that problems that may seem very distantly related when viewed from the biological perspective are very close when viewed from the physical biology perspective. In this case, the study of elastic deformations of beams will lead us to speak of transcriptional regulation, DNA packaging, and cytoskeletal mechanics using one common language.

10.2 Geometry and Energetics of Beam Deformation

10.2.1 Stretch, Bend, and Twist

Biological filaments are characterized by one dimension (the length) that is much greater than their transverse dimensions. For example, in the case of the bacterial flagellum, the structure has a length in excess of microns with a diameter that is measured in only tens of nanometers. As was illustrated in Figure 2.27 (p. 64), the same can be said of tobacco mosaic virus, actin, and microtubules, which similarly have a characteristic length of order microns with cross-sectional dimensions measured in nanometers. Because of this geometric asymmetry, it is possible to invoke key simplifying assumptions that permit us to write the energy of deformations of these asymmetric structures very simply.

Beam Deformations Result in Stretching, Bending, and Twisting

We begin with a qualitative discussion of the geometric character of the three key independent modes of deformation to which a beam may be subjected. In particular, the deformation of beams can be described in terms of extension, bending, and torsion as shown in Figure 10.2. Extensional deformations have already been discussed in Section 5.4.1 (p. 216) and correspond to simple extension (or compression) of the beam along its long axis from length L_0 to length $L_0 + \Delta L$. In this chapter, we complement the earlier discussion by thinking about two other modes of deformation that are also important in the mechanics of biological systems. As seen in Figure 10.2(B), the bending of a beam can, at its simplest, be thought of as a deformation that takes a straight beam and bends it into an arc of a circle. The other key mode of deformation corresponds to twisting the beam about its long axis as shown in Figure 10.2(C).

A Bent Beam Can Be Analyzed as a Collection of Stretched Beams

The state of deformation that will animate the majority of our discussion in this chapter is that of bending. Bending is of interest in many of the examples already revealed in Figure 10.1. For example, when determining the free energy associated with the wrapping of DNA around the histone octamer as shown in Figure 10.1(G), the energy

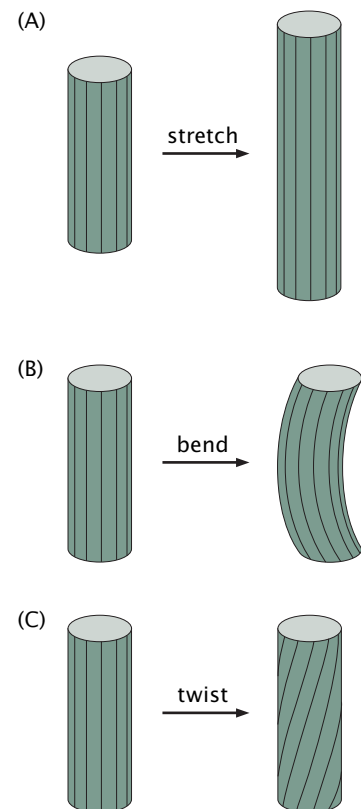
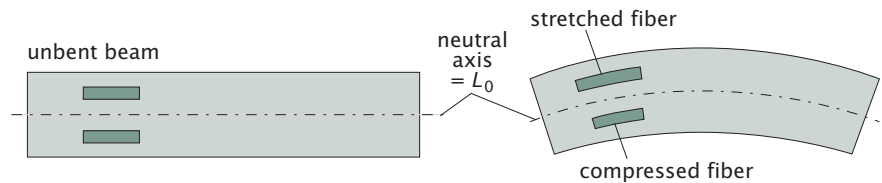


Figure 10.2: Three modes of deformation of a beam: (A) stretching of a beam, (B) bending of a beam, and (C) twisting of a beam.

Figure 10.3: Stretching of material elements in a bent beam. The length of the dotted line representing the neutral axis is unchanged after bending. Material elements above the neutral axis are stretched and those below it are compressed.



of bending the DNA serves as one of the two key ingredients in the overall free energy and competes with the DNA–protein interaction energy to determine the overall free energy of nucleosome formation.

We begin by examining the nature of the geometric assumptions that are made about the state of deformation of a bent beam. Indeed, the first thing we will say about the geometry of a bent beam is that there is a certain axis, known as the neutral axis, that is neither stretched nor compressed during the act of beam bending, as shown in Figure 10.3. The key observation embodied in the figure is that material elements above the neutral axis are stretched, while those below the neutral axis are compressed.

The geometric characterization of a beam subjected to some complicated (that is, nonuniform) state of bending as shown in Figure 10.4 is analyzed by recourse to a divide-and-conquer strategy that will be taken up repeatedly throughout the book. The idea of this strategy is that we divide the beam up into a collection of small segments,

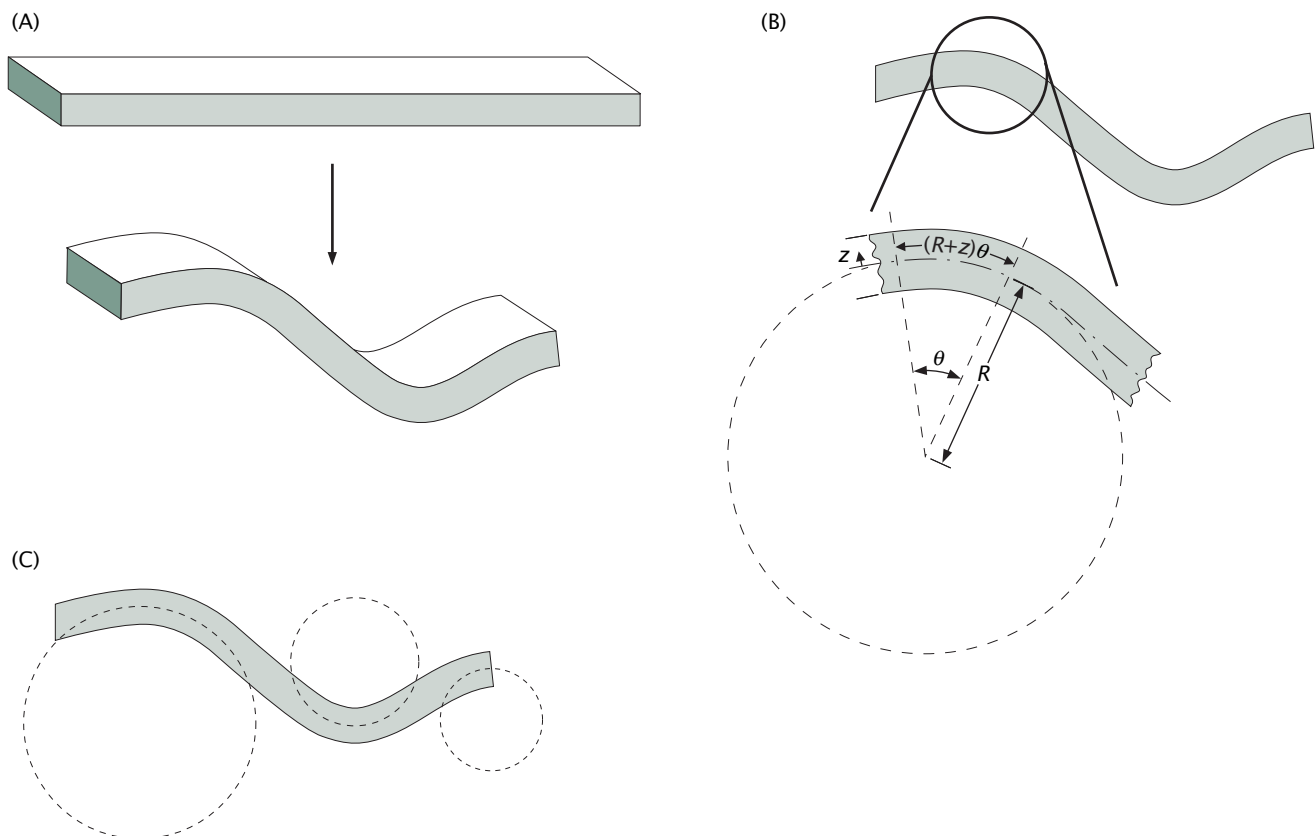


Figure 10.4: Curvature of a beam. The schematic shows the way in which a beam can be thought of as a collection of segments, each of which is bent locally into an arc of a circle. (A) Nonuniform bending of a beam. (B) Determination of curvature at a given point along the beam. (C) Illustration of different curvature at different points along the beam.

each of which can be thought of as being a part of an arc of a circle. We will argue below that for beams bent into circular arcs, it is a straightforward exercise to compute both the curvature ($1/R$) and the resulting deformation energy. Consequently, when trying to evaluate the bending energy of the entire beam, we just add up the contributions from each and every segment (by mathematical integration). The implementation of this geometric idea is illustrated in Figure 10.4. The figure invites us to ignore the dimension perpendicular to the plane of the page since by symmetry all beam elements along this direction are equivalent. The second key point illustrated in the figure is that, regardless of the complexity of the bending, we may think of the beam as a collection of circular arcs all glued to each other. That is, each cross-section along the length of the beam can be fitted onto a circle with a particular radius, this radius corresponding to the *local* radius of curvature of the beam itself.

Once the radius of curvature is in hand, we are prepared to compute the state of strain within the beam. As was shown in Section 5.4.1 (p. 216), given the strain for a beam that has been stretched, we can use Hooke's law to compute the strain energy of bending. The strain (for the simple cases of interest here) is a measure of the fractional extension ($\Delta L/L_0$) of a particular little element of the material. In particular, we note that the extent of compression or extension is a linear function of the perpendicular distance of a given material point from the neutral axis (that is, $\Delta L(z) = \alpha z$). On geometric grounds, we may determine the unknown constant α through an examination of Figure 10.4, where we note that all parts of the beam subtend the angle θ as shown in the figure. As a result, the fiber at distance z from the neutral axis is of overall length $L(z) = (R + z)\theta$, which may be rewritten as $L(z) = (R + z)L_0/R$, where we have used the fact that $\theta = L_0/R$. As a result, we can write the change in length of the fiber at distance z from the neutral axis as

$$\Delta L(z) = L(z) - L_0 = (R + z)\frac{L_0}{R} - L_0 = \frac{zL_0}{R}. \quad (10.1)$$

Continuing along these same lines, we conclude that the extensional strain $\varepsilon(z)$ at a distance z from the neutral axis is given by

$$\varepsilon(z) = \frac{\Delta L(z)}{L_0} = \frac{z}{R}. \quad (10.2)$$

Recall that z is measured from the neutral axis and hence material elements above this axis in Figure 10.4 are stretched ($\varepsilon > 0$) while those below this axis are compressed ($\varepsilon < 0$).

The Energy Cost to Deform a Beam Is a Quadratic Function of the Strain

Given the state of strain within a beam described above, it is now possible to assess the energy stored in the beam by virtue of such deformation. The beautiful idea is to use what we already know about the energetics of stretching beams to work out the energy cost to *bend* a beam. The principle is illustrated in Figure 10.5, which is a microscopic impression of the idea that the bonds between atoms above the neutral axis are stretched while those below the neutral axis are compressed. Though the detailed forces between atoms in macromolecules are complicated, in the case of small deformations the energetics of bond bending and stretching is simply quadratic.

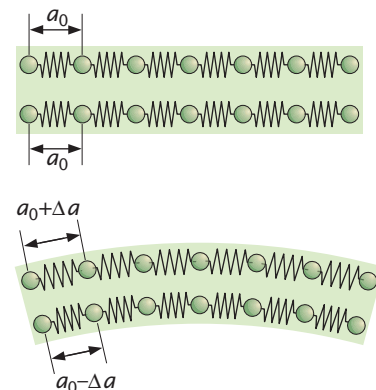


Figure 10.5: Microscopic interpretation of the energetics of beam bending. When a beam is subjected to bending, the bonds above the neutral axis are stretched while those below the neutral axis are compressed.

The energy associated with the extension of a beam can be *estimated* easily by ignoring the Poisson effect (essentially the effect that if we pull in one direction, the material contracts in the others) and noting that the strain energy density (that is, energy per unit volume) is given by

$$W(\varepsilon) = \frac{1}{2} E \varepsilon^2 = \frac{1}{2} E \left(\frac{\Delta L}{L_0} \right)^2, \quad (10.3)$$

where we remind the reader that E is the Young modulus of the material of interest (see Section 5.4.1 on p. 216 for a refresher). Given this estimate for the strain energy density, reckoning the total stored elastic energy in the little volume of deformed material requires multiplying the energy density by the volume of the little volume element.

A more precise treatment of the energy of deformation accounts for the fact that the state of strain depends upon where we are within the deformed beam. In particular, we invoke Equation 10.2, which tells us how the strain depends upon the perpendicular distance from the neutral axis. From an energetic point of view, this means that we have to add up the contributions from each little material element separately. In light of the z dependence of the strain itself, the total strain energy stored in the beam is written as

$$E_{\text{bend}} = L_0 \int_{\partial\Omega} dA \frac{E}{2R^2} z^2, \quad (10.4)$$

where the integration is over the cross-sectional area $\partial\Omega$ perpendicular to the beam axis. The strain energy integral is shown schematically in Figure 10.6.

More generally, our result may be written as

$$E_{\text{bend}} = \frac{EIL}{2R^2}, \quad (10.5)$$

where we have introduced the geometric moment

$$I = \int_{\partial\Omega} z^2 dA, \quad (10.6)$$

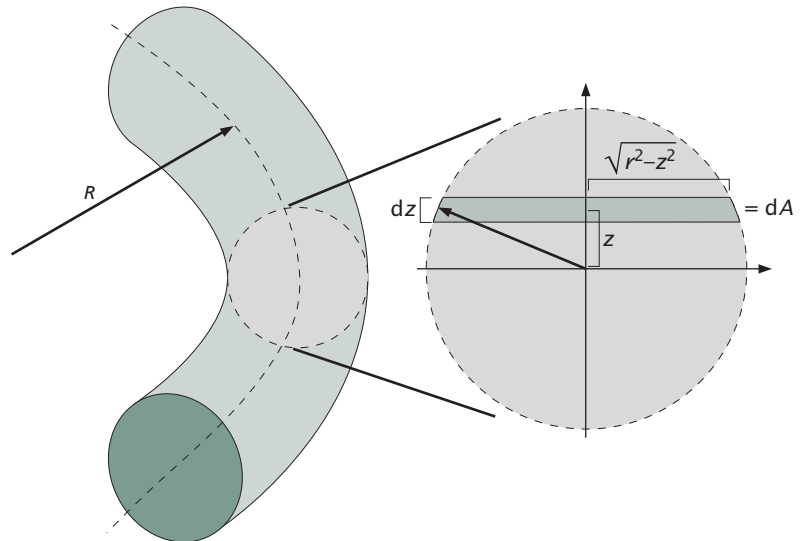


Figure 10.6: Computation of the bending energy. The bending energy is obtained by integrating over the cross-sectional area of the beam.

where once again the area element is perpendicular to the beam axis, and the quantity z is a measure of the perpendicular distance from the neutral axis. Equation 10.5 is one of the more important equations to be considered in this book, since it tell us the bending energy associated with bending a beam into a circular arc of length L with radius of curvature R . We may take that result even further through the recognition that for a circular loop with radius of curvature R , its length is $L = 2\pi R$ and the bending energy is

$$E_{\text{loop}} = \frac{\pi EI}{R}. \quad (10.7)$$

The energetics of simple circular loops will be applied in numerous places throughout the remainder of the chapter and especially in our consideration of transcriptional regulation and DNA packaging.

As yet, our analysis has been for the case in which the entire beam of length L is uniformly deformed into a single arc of a circle with radius R . More generally, we will be interested in states of deformation in which the local curvature differs from one point to the next. In this case, we invoke a *locality* assumption in which we pretend that the energy of a given material particle is that of the beam described above and for which the curvature is constant. The total energy is gotten by adding up the energy material particle by material particle and is of the form

$$E_{\text{bend}} = \frac{K_{\text{eff}}}{2} \int_0^L ds \frac{1}{R(s)^2}, \quad (10.8)$$

where we have introduced the flexural rigidity $K_{\text{eff}} = EI$, which embodies both material parameters (that is, the Young modulus E) and the geometric shape (through the geometric moment I). This equation is the mathematical embodiment of the cartoon shown in Figure 10.4, where we show how each point on the beam can be thought of locally as part of a circle. There are several other convenient ways of writing Equation 10.8 that will arise in the remainder of the book. First, we note that $\kappa(s) = 1/R(s)$, where we have defined the curvature $\kappa(s)$. This implies that we can rewrite the equation as

$$E_{\text{bend}} = \frac{K_{\text{eff}}}{2} \int_0^L \left| \frac{d\mathbf{t}}{ds} \right|^2 ds, \quad (10.9)$$

where we use the fact that the curvature can itself be written as the derivative of the tangent vector.

10.2.2 Beam Theory and the Persistence Length: Stiffness is Relative

Thermal Fluctuations Tend to Randomize the Orientation of Biological Polymers

One intriguing theme that we will return to repeatedly throughout the book (already introduced in Section 5.1.1 on p. 189) is the idea of a competition between thermal effects and deterministic forces. For example, in Figure 9.12 (p. 369) we introduced the Bjerrum length as the length scale over which charges can wander from a protein without offending the Coulomb interaction too significantly. Here we examine another example of this same kind of argument in which a biological polymer such as DNA is kicked around into various different

orientations as a result of interactions with the surrounding fluid. In this case, it is the elastic forces rather than the Coulomb forces that set the length scale over which such fluctuations are tolerated. In particular, the intuition corresponding to this idea is that if the polymer is too short, it will be indifferent to thermal fluctuations. By way of contrast, for very long polymers, the orientation at one extremity is completely indifferent to that at the other.

The Persistence Length Is the Length Over Which a Polymer Is Roughly Rigid

The competition described above between thermal fluctuations and the energetic cost associated with beam bending is succinctly captured in the emergence of a single length scale, namely, the persistence length. The persistence length is a measure of the competition between the entropic parts of the free energy, which tend to randomize the orientation of the polymer, and the energetic cost of bending. Throughout physical biology, there are a variety of different length scales that arise that, at the deepest level, reflect the interplay between energy and entropy. Generally, length scales that reflect the competition between thermal and deterministic energies can be estimated by equating the deterministic energy cost for the particular mechanism of interest to $k_B T$, the thermal energy. When elastic energies are competing with thermal fluctuations, the relevant comparison is

$$k_B T \approx \frac{EIL}{2R^2}, \quad (10.10)$$

where L is the length of the fragment of interest. Roughly speaking, the persistence length ξ_p is that length of polymer for which the radius of curvature is equal to the length of polymer itself. Hence, if we set L and R both equal to ξ_p , we see that our estimate for the persistence length is given by

$$\xi_p \approx \frac{EI}{2k_B T}. \quad (10.11)$$

The Persistence Length Characterizes the Correlations in the Tangent Vectors at Different Positions Along the Polymer

An alternative view of the persistence length is to think of biological polymers from the standpoint of the geometry of space curves. There are a number of different ways of characterizing the mathematics of space curves. One way is via correlation functions, which measure the extent to which the geometry of one part of the polymer is correlated with some other part.

To be concrete, consider a space curve in parametric representation given by the vector $\mathbf{r}(s)$, where s is the arclength parameter. The function $\mathbf{r}(s)$ is a rule that assigns a position vector for every value of the arclength s . Such a configuration is shown in Figure 10.7. The question we pose concerning a given space curve is the nature of the tangent-tangent correlation function defined as $g(s) = \langle \mathbf{t}(\tau + s) \cdot \mathbf{t}(\tau) \rangle$. Our notation is built around the unit tangent vector, $\mathbf{t}(s)$, which is defined as the unit vector tangent to the curve at arclength s . The tangent-tangent correlation function as defined above is thus far a purely mathematical notion and provides a measure of the relation

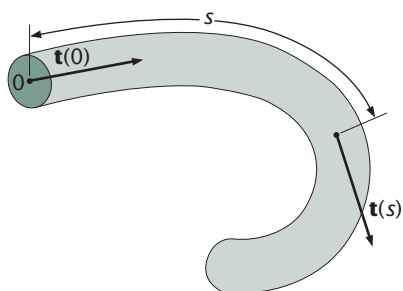


Figure 10.7: Tangent vectors on a fluctuating polymer chain. The parameter s measures the distance along the polymer. The tangent vector at contour length s is represented as $\mathbf{t}(s)$.

between the tangents at arclength τ and arclength $\tau + s$. When the correlation function is 1, the tangent vectors are parallel.

This idea was made explicit in Figure 8.5 (p. 319), where we showed the genome of a bacterium that had been released from the cell. In the inset to that figure, we labeled the persistence length with small white arrows, each with a length equal to that of the persistence length of DNA. These arrows reveal that on the scale of tens of nanometers, the tangents are correlated while on the scale of microns they are not.

The Persistence Length Is Obtained by Averaging Over All Configurations of the Polymer

Qualitatively, we see that the actual numerical value of the persistence length depends upon the stiffness of the filament. Microtubules and DNA do not have the same persistence length, with DNA characterized by a persistence length of roughly 50 nm while microtubules have a persistence length of the order of 6 mm. To compute the persistence length, we average over all configurations of the polymer, with each such configuration assigned a Boltzmann weight that depends upon the bending energy. Before we plunge into the mathematics of the tangent–tangent correlation function

$$g(s) = \langle \mathbf{t}(s) \cdot \mathbf{t}(0) \rangle, \quad (10.12)$$

it is useful to think about its limits. First, since the tangent vector is of unit length, $g(0) = 1$. On the other hand, for s much larger than the persistence length, we expect the two tangent vectors to be independent and $g(s) \rightarrow 0$. A simple function with these properties is an exponential function

$$g(s) = e^{-s/\xi_p} \quad (10.13)$$

where ξ_p is the persistence length. The mathematical proof that $g(s)$ is indeed exponential is somewhat subtle. It is based on the property that two adjacent segments of an elastic beam are buffeted by independent thermal forces and hence the tangent–tangent correlation function over both segments is the product of the individual ones, that is, $g(s_1 + s_2) = g(s_1)g(s_2)$; s_1 and s_2 are the lengths of the two segments (further details can be found in Nelson 2004).

We seek a relation between the persistence length, as defined by Equation 10.13, and the flexural rigidity $K_{\text{eff}} = EI$ of an elastic beam. To this end, we take a short beam of length $s \ll \xi_p$ and compute $g(s)$. For a short beam, thermal forces can only bend the beam slightly and the shape at any instant can be approximated by an arc of a circle of radius R . The energy of such a configuration is given by Equation 10.5, which can also be written as

$$E_{\text{bend}} = \frac{EI}{2s} \theta^2, \quad (10.14)$$

with $\theta = s/R$. Taking the tangent of one end of the beam to point along the z -direction, the tangent–tangent correlation function becomes

$$g(s) = \langle \cos \theta(s) \rangle, \quad (10.15)$$

with $\theta(s)$ the angle between the tangent vector at the other end and the z -direction. Given our assumption of small deflection angles ($\theta(s) \ll 1$) appropriate for a short beam, the cosine function can be expanded

into a Taylor series (see The Math Behind the Models on p. 215) to yield a simplified expression

$$g(s) = \left\langle 1 - \frac{\theta^2(s)}{2} \right\rangle. \quad (10.16)$$

The thermal average in the above equation is computed by summing over all possible orientations of the tangent vector at s , which in three dimensions traces out the surface of a unit sphere. Therefore,

$$\langle \theta^2(s) \rangle = \frac{1}{Z} \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \theta^2 e^{-(EI/2k_B Ts)\theta^2}, \quad (10.17)$$

where

$$Z = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta e^{-(EI/2k_B Ts)\theta^2}, \quad (10.18)$$

is the normalization factor (partition function). To simplify calculations we note that the integrand in Equation 10.17 differs from the integrand for Z by the factor θ^2 , which we can pull down from the Boltzmann factor in Equation 10.18 by differentiating with respect to E as introduced in Section 6.9 (p. 241). Then we can write

$$\langle \theta^2(s) \rangle = \frac{1}{Z} \left(-\frac{2k_B Ts}{I} \frac{\partial Z}{\partial E} \right) = -\frac{2k_B Ts}{I} \frac{\partial \ln Z}{\partial E}, \quad (10.19)$$

and all that remains to be calculated is the integral in Equation 10.18.

We compute the integral in question by making the substitution $\sin \theta \approx \theta$, valid for small angles, and by making use of a change of variables $u = (EI/2k_B Ts)\theta^2$,

$$Z = \frac{2\pi k_B Ts}{EI} \int_0^\infty du e^{-u} = \frac{2\pi k_B Ts}{EI}. \quad (10.20)$$

Note that the upper integration bound for the variable u tends to infinity in the limit when s is much smaller than the persistence length. Using the result $\partial \ln Z / \partial E = -1/E$ in Equation 10.19, and substituting the value for $\langle \theta^2(s) \rangle$ obtained in this way in Equation 10.16, we arrive at the result

$$g(s) = 1 - \frac{k_B T}{EI} s. \quad (10.21)$$

Comparing this with Equation 10.13 in the $s \ll \xi_p$ limit, we conclude

$$\xi_p = \frac{EI}{k_B T}. \quad (10.22)$$

Note that the scaling is precisely that found in Equation 10.11, which was based on a simple comparison of the thermal energy and the bending energy of an elastic beam bent into a circular arc of 1 radian. This result also allows us to rewrite the flexural rigidity as $EI = \xi_p k_B T$.

10.2.3 Elasticity and Entropy: The Worm-Like Chain

The Worm-Like Chain Model Accounts for Both the Elastic Energy and Entropy of Polymer Chains

So far in the book, we have used two different physical models to characterize the structure and free energy of polymers. In Chapter 8,

we examined the behavior of polymers as random walks. These ideas permitted us to consider experiments such as force–extension measurements that probe the entropic forces tending to keep polymers collapsed. In this chapter, we have taken a complementary view that considers the energetic cost of deforming polymers. Evidently, both the bending energy and entropic contributions to the overall free-energy budget are important. To that end, a class of models known as worm-like chain models has been set forth in which one evaluates the partition function associated with different polymer configurations, where each such configuration is appropriately weighted by its corresponding elastic energy cost. The concept of these models is shown in Figure 10.8.

The competition between chain entropy and bending energy is captured by the partition function, which for the worm-like chain model reads

$$Z = \int \mathcal{D}\mathbf{t}(s) \exp\left(-\frac{\xi_p}{2} \int_0^L \left|\frac{d\mathbf{t}}{ds}\right|^2 ds\right). \quad (10.23)$$

The mathematics expressed by the above equation translates into a simple algorithm:

- (1) draw a curve of length L representing a possible DNA configuration;
- (2) evaluate its bending energy

$$E_{\text{bend}} = \frac{\xi_p k_B T}{2} \int_0^L \left|\frac{d\mathbf{t}}{ds}\right|^2 ds$$

and the corresponding Boltzmann factor $e^{-E_{\text{bend}}/k_B T}$; and

- (3) repeat (1) and (2) for all possible curves and sum all the Boltzmann factors to obtain Z .

The sum over curves is the celebrated Feynman path integral, written here as $\int \mathcal{D}\mathbf{t}(s)$, where $\mathbf{t}(s)$ is the tangent vector of the curve, parameterized by the arclength s . For the present purposes, Equation 10.23 should just be thought of as a fancy, shorthand notation for the partition function that reminds us that there are an infinite number of different ways of configuring the polymer.

In order to compute the force–extension curve of the worm-like chain model, we consider a force F applied to one end of the chain in the z -direction. The energy of the worm-like chain then acquires an additional $-F \int_0^L t_z ds$ term, where t_z is the z -component of the tangent vector \mathbf{t} . This is equivalent to the energy of the freely jointed chain model in the presence of an applied force, discussed previously, and amounts to accounting for the lowering of weights when the polymer is elongated as shown schematically in Figure 8.24 (p. 341). For fixed F , the chain extension $z = \int_0^L t_z ds$ averaged over all configurations is

$$\langle z \rangle = \frac{1}{Z(f)} \int \mathcal{D}\mathbf{t}(s) z \exp\left(-\frac{\xi_p}{2} \int_0^L \left|\frac{d\mathbf{t}}{ds}\right|^2 ds + f \int_0^L t_z ds\right), \quad (10.24)$$

where $Z(f)$ is the partition function in the presence of the applied force $F = f k_B T$. Note that the reduced force f has units of inverse length. Equation 10.24 can be rewritten as

$$\langle z \rangle = \frac{d \ln Z(f)}{df}, \quad (10.25)$$

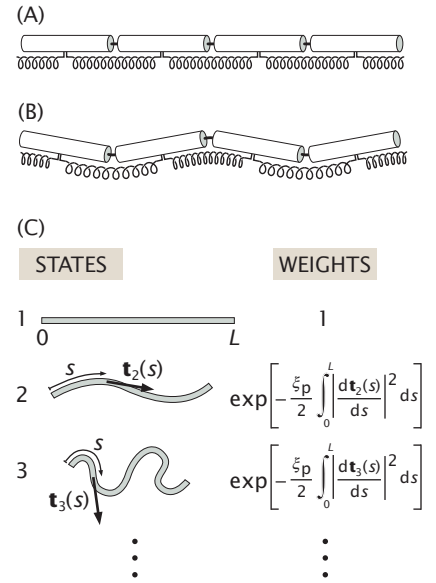


Figure 10.8: The worm-like chain concept. As in Chapter 8, chains are represented by cylinders connected by flexible links. (A) The undeformed configuration showing that the springs are unstretched, but the links are deprived of entropy because there is only one such possible arrangement of the segments. (B) A deformed configuration showing that there is an energetic cost to bend the chain, but there are more configurations of the system. (C) States and weights for the worm-like chain model in the absence of applied force. There are an infinite number of different states corresponding to all the different space curves that can be made with the polymer, but to find the energy of each such configuration, we simply compute the bending energy of the conformation of interest.

owing to the fact that z is the thermodynamic conjugate of the force. The problem of computing the force–extension curve therefore reduces to computing the partition function $Z(f)$. This is a rather difficult mathematical problem, since one has to make sense of summing over all curves. Fortunately, this problem greatly simplifies in the low- and high-force limits. In Section 10.7 (the appendix at the end of the chapter), we examine both the low- and high-force limits, which permit a determination of the force–extension properties of the worm-like chain model in these limits.

The outcome of these calculations can be summarized in an interpolation formula

$$f\xi_p \approx \frac{z}{L} + \frac{1}{4(1 - z/L)^2} - \frac{1}{4} \quad (10.26)$$

that permits an evaluation of the extensions suffered by a worm-like chain over the entire range of applied forces. The high-force limit in this formula is very different from the freely jointed chain result for the force–extension curve (see Section 8.3.2 on p. 340), which predicts a $1/f\xi_p$ approach to full extension $z = L$. The slower approach to full extension in this case is a consequence of the finite bending rigidity at scales below the Kuhn length in the worm-like chain model.

10.3 The Mechanics of Transcriptional Regulation: DNA Looping Redux

The physical properties of DNA as a deformable polymer complement its well-known properties as the storage medium for genetic information. Indeed, there are many cases where the interplay between the informational and physical properties of the DNA molecule have intriguing biological consequences. One example of great importance is that of gene regulation. There are a variety of examples in which transcriptional regulation is carried out by the binding of proteins, which have the effect of forming loops in the DNA. An example of this effect was introduced in Figure 1.11 (p. 19) and the entropic contribution to the free energy of looping was discussed in Section 8.2.4 (p. 333). The aim of this section is to take stock of how the ideas on beam bending developed in this chapter can be applied to examine DNA deformations in transcriptional regulation.

10.3.1 The *lac* Operon and Other Looping Systems

As was already emphasized in Section 4.4.3 (p. 157), one of the most celebrated examples of gene regulation is the *lac* operon. This genetic network oversees the metabolism of the sugar lactose in *E. coli*. Historically, this system served as one of the focal points resulting in the emergence of a picture of how genes are controlled. In the present setting, we consider repression of transcription in the *lac* operon, which is mediated by the formation of loops. The argument of this section is that such loops can be considered from the perspective of the elastic theory of beams. As a brief reminder, we recall that the *lac* operon is subject to both positive and negative control. In particular, RNA polymerase is “recruited” to the promoter by an activator protein known

as CAP. Negative control of transcription is dictated by the presence of the DNA-binding protein Lac repressor.

Transcriptional Regulation Can Be Effected by DNA Looping

Lac repressor, as shown in Figure 8.19 (p. 334), is a tetrameric protein that binds two distinct DNA sites (one of which is in the vicinity of the promoter) and loops the intervening DNA. In particular, there are a total of three specific binding sites for Lac repressor, denoted O1, O2, and O3, with O2 being 401 base pairs downstream of O1 and O3 being 92 base pairs upstream of O1. Full repression by the repressor molecule demands that the tetrameric repressor molecule bind O1, whose center is located 11 base pairs downstream from the RNA polymerase transcription start site, and one of the auxiliary operators simultaneously, thus forming a loop with the intervening DNA. Depending upon which operators are linked, the looped region can either be 401 or 92 base pairs in length. Note that the 92 base pair loop appears to involve substantial bending of a DNA fragment that is notably smaller than the persistence length.

In addition to the example of the *lac* operon, there are a host of other intriguing prokaryotic and eukaryotic examples of regulatory regions involving DNA looping. In prokaryotes, another important example is the *ara* operon, which is associated with metabolism of the sugar arabinose. More generally in eukaryotes, there are a wide variety of examples of *cis*-regulatory regions that control developmental processes that also involve DNA looping. To give an impression of the complexity of *cis*-regulatory regions associated with eukaryotic promoters, we consider one of the well-characterized regulatory networks associated with sea urchin development. Figure 10.9 shows the regulatory region associated with the *cyIIIa* gene in sea urchin development as well as electron micrographs of the looped segments of DNA resulting from the binding of particular transcription factors.

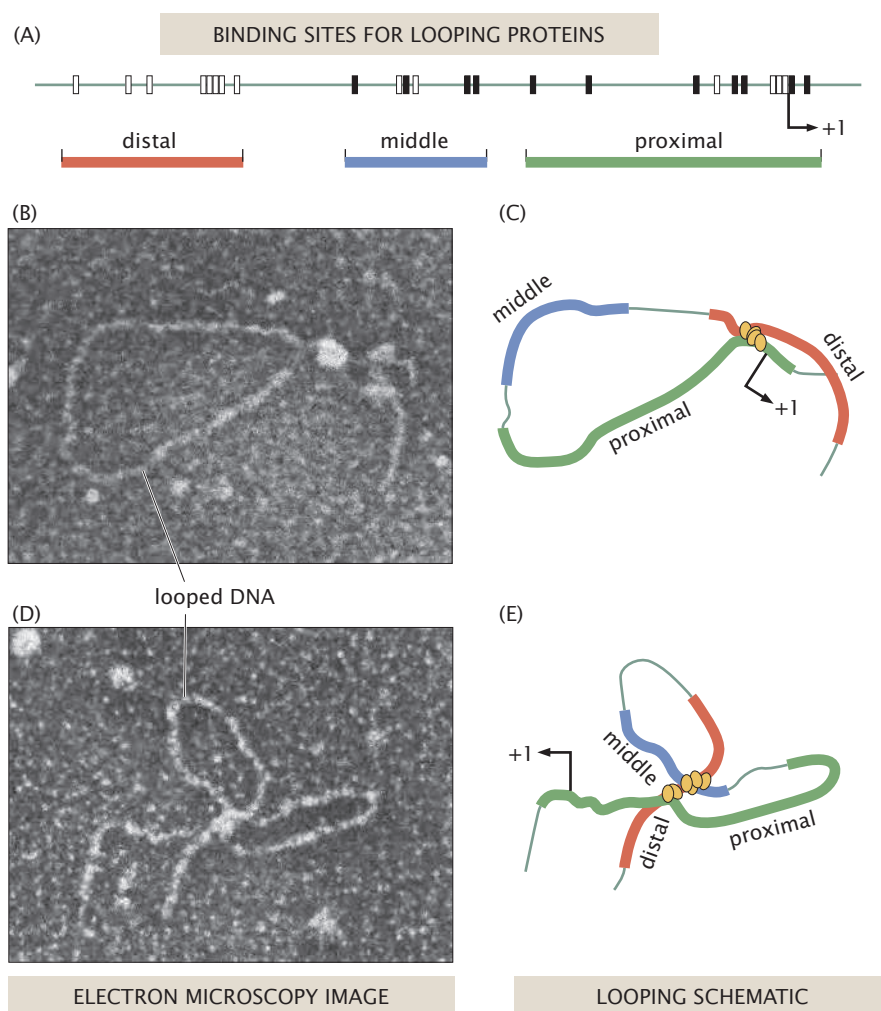
10.3.2 Energetics of DNA Looping

In Section 8.2.4 (p. 333) we examined the entropic consequences of DNA looping. That discussion now must be complemented by an evaluation of the energetic consequences of looping. We begin by estimating the amount of elastic energy that has to be paid in order for the repressor molecule to effect its binding. From the perspective of elasticity, a simple estimate for the energetics of DNA looping can be obtained on the assumption that the loops are circular. In particular, once we accept this geometric picture, the energetics of a loop of radius R is given by $E_{\text{loop}} = \xi_p \pi k_B T / R$, as was deduced in Equation 10.7. A more useful way of writing this is in terms of the number of base pairs involved in the loop. In particular, if we recall that the length of a given fragment of DNA is given by $L \approx 0.34 N_{\text{bp}}$ nm, then we can write the loop energy more directly as

$$\frac{E_{\text{loop}}}{k_B T} = \frac{2\pi^2}{N_{\text{bp}}} \left(\frac{\xi_p}{\delta} \right) \approx \frac{3000}{N_{\text{bp}}}, \quad (10.27)$$

where we have introduced the parameter $\delta \approx 0.34$ nm, which is the length of DNA per base pair, and to obtain the numerical expression we have used a persistence length $\xi_p = 50$ nm. This expression for the loop formation energy is plotted in Figure 10.10.

Figure 10.9: Gene regulation and DNA looping. (A) Schematic of the DNA *cis*-regulatory region associated with the *cyIIIa* cytoskeletal actin gene. (B) Electron micrograph of looped DNA. (C) Schematic showing which regions of the DNA shown in (A) join together to form a loop. (D) Electron micrograph of a second looped configuration. (E) Schematic showing which regions of the DNA shown in (A) join together to form the loop. (Adapted from R. W. Zeller et al., *Proc. Natl Acad. Sci. USA* 92:2989, 1995.)



10.3.3 Putting It All Together: The J-Factor

The simple model of the energy to form a DNA circle can be married with the estimate of the looping entropy described in Section 8.2.4 (p. 333) in order to compute the overall free energy of looping. The idea is that for long fragments, the free-energy cost of looping is dominated by the entropic cost of depriving the DNA of configurations as a result of forming the loop. On the other hand, at very short DNA lengths, the energetic cost of bending the DNA dominates the free energy. We are now ready to examine the interplay of these different effects and to compute an expression for the probability of looping that respects both the energetic and entropic contributions to the free energy.

Equation 10.27 reveals the elastic energy cost to form a circular loop of DNA and can be rewritten as

$$\Delta E_{\text{loop}} \approx 3000 k_B T / N_{\text{bp}}, \quad (10.28)$$

where only the bending energy is taken into account. As shown in Chapter 8, in the simplest random walk model of a long polymer chain, the probability of loop formation, p_c , is proportional to $N_{\text{bp}}^{-3/2}$ (see Equation 8.60 on p. 336). This implies that the entropy loss due to

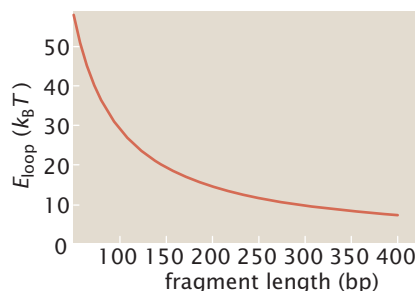


Figure 10.10: Energy of DNA loops as a function of the number of base pairs in the DNA. The calculation assumes that the loops are perfect circles.

loop formation is

$$\Delta S_{\text{loop}} = k_B \ln p_o = k_B \left(-\frac{3}{2} \ln N_{\text{bp}} + \text{const} \right), \quad (10.29)$$

where the constant term does not depend on the length of the chain. This follows because $\Delta S_{\text{loop}} = S_{\text{loop}} - S_{\text{total}} = k_B \ln(W_{\text{loop}}/W_{\text{total}})$, and $W_{\text{loop}}/W_{\text{total}}$ is nothing more than p_o , the probability to form a loop. Our results for the looping energy and entropy can be assembled to form an approximate free energy for loop formation of the form

$$\Delta G_{\text{loop}} = \Delta E_{\text{loop}} - T\Delta S_{\text{loop}} \approx k_B T \left(\frac{3000}{N_{\text{bp}}} + \frac{3}{2} \ln N_{\text{bp}} + \text{const} \right). \quad (10.30)$$

This free energy is plotted in Figure 10.11. One notable feature of this curve is that there is an optimal DNA length at which the looping free energy is at its minimum, slightly under 2000 bp.

The simple estimates described above are amenable to experimental verification through *in vitro* studies of DNA cyclization (see Figure 10.12). In particular, in a series of now-classic experiments, the probability of looping could be assessed explicitly and compared with the results of the simple linear elastic model described above. The idea of such experiments is to trap spontaneous thermal fluctuations of the DNA fragments into the looped configuration when the two ends of the linear DNA fragment are in close proximity. To do this, the DNA of interest is prepared with complementary sticky ends. When this DNA is examined at a given concentration, two key outcomes are possible: either the DNA will form circles or it will form dimers. An enzyme called ligase that sews together fragments of DNA that have complementary sticky ends is used to lock the DNA into one or the other of these conformations and the relative concentrations of these two species are measured. The probability of loop formation directly measured in these cyclization experiments has a maximum closer to 500 bp. The quantitative discrepancy is not surprising, since our approximate model fails at intermediate chain

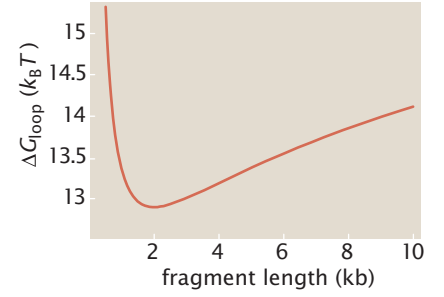


Figure 10.11: Looping free energy. The free energy to form a circular loop as a function of the number of base pairs in the loop. This result is based upon a toy model in which the energy is computed using a simple elastic model of rod bending and an entropy based upon the random walk model.

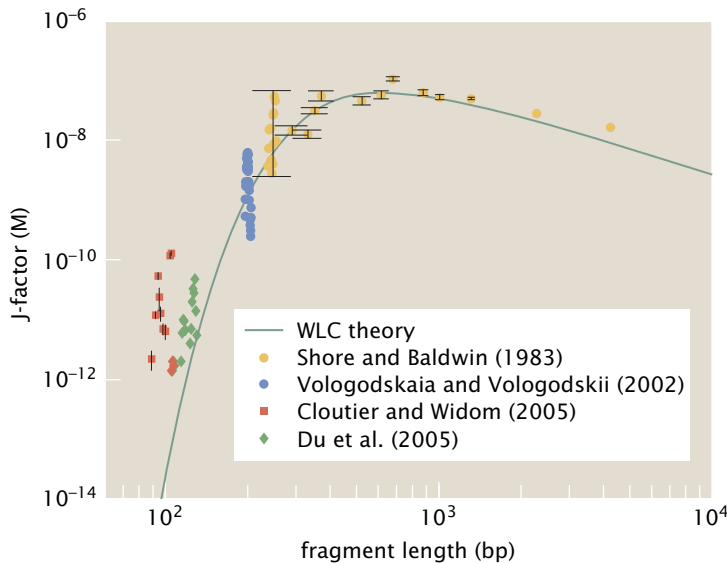


Figure 10.12: J-factor as a function of DNA fragment length. The curve shows the result of using the worm-like chain model to describe the cyclization. This model is more quantitatively realistic than the toy model considered in this chapter, since rather than only considering the elastic energy of circular conformations at small DNA segment lengths, it considers all shapes that close upon themselves and their corresponding entropies. (Data adapted from T. E. Cloutier and J. Widom, *Proc. Natl Acad. Sci. USA* 102:3645, 2005; Q. Du et al., *Proc. Natl Acad. Sci. USA* 102:5397, 2005; D. Shore and R. L. Baldwin, *J. Mol. Biol.* 170:957, 1983; M. Vologodskaja and A. Vologodskii, *J. Mol. Biol.* 317:205, 2002.)

lengths when entropic and energetic contributions to the free energy of looping are comparable.

In previous chapters, we have shown how the binding of proteins to their substrates can be characterized both in terms of binding energies and in terms of dissociation constants. Similarly, there are two different languages that can be used to describe DNA mechanics and probabilities of DNA adopting particular configurations. Cyclization experiments are often reported using a quantity known as the J-factor, which has units of concentration. This magnitude can be interpreted as the effective concentration of one end of the DNA molecule in the vicinity of the other. The J-factor is proportional to $e^{-\beta \Delta G_{\text{loop}}}$, where ΔG_{loop} is the looping free energy and the proportionality factor, which has units of concentration, depends on the microscopic model employed. The reader is invited to consider this in detail in Problem 10.4.

The results of a variety of different measurements of the J-factor are shown in Figure 10.12. The trends are consistent with the model described here, namely, the long length limit is entropy-dominated, while the short-length limit is dictated by the energy of bending.

10.4 DNA Packing: From Viruses to Eukaryotes

The Packing of DNA in Viruses and Cells Requires Enormous Volume Compaction

One of the many remarkable features of DNA is its length relative to the linear dimensions of the volumes within which it is usually confined. For example, in the case of a bacteriophage, in excess of $10\ \mu\text{m}$ of DNA is packed inside a capsid that is roughly $50\ \text{nm}$ across. To put this in terms of everyday dimensions, this is like putting $500\ \text{m}$ of cable (with a diameter of the order of $10\ \text{cm}$) from the Golden Gate Bridge into the back of a delivery truck. There are several ways to view the extent to which a particular genome is compacted. The simplest measure, though misleading, is, as was suggested above, to compare the linear dimensions of the DNA of interest with those of the region within which that DNA is confined. For example, in the case of “typical” eukaryotic DNA packaging, the linear dimension of the DNA is of the order of centimeters, while the dimension of the nucleus where that DNA is stored is measured in microns. A much more useful measure of the degree of packing of DNA in different settings is to consider the ratio of the volume taken up by the DNA (when viewed as a long solid cylinder of diameter $2\ \text{nm}$) to the volume of the region where the DNA is stored. In particular, this leads us to consider the dimensionless ratio

$$\nu = \frac{\Omega_{\text{genome}}}{\Omega_{\text{container}}} \approx \frac{N_{\text{bp}}}{\Omega_{\text{container}}}, \quad (10.31)$$

where the volume of the container, $\Omega_{\text{container}}$, is measured in cubic nanometers and we have used the rule of thumb that the volume per base pair for dsDNA is roughly $1\ \text{nm}^3$.

Estimate: The DNA Packing Compaction Ratio. Equation 10.31 permits us to examine the degree of relative compaction of the DNA in the different examples shown in Figure 10.13. As a concrete example of a bacteriophage, we consider the famed lambda phage. The capsid of lambda phage can be thought of as a sphere of radius 27 nm that holds a 48,500 bp genome. Note that there is a slight subtlety due to the thickness of the capsid itself which we will ignore. The packing ratio for lambda is given by

$$\nu_{\text{lambda}} = \frac{N_{\text{bp}} \text{ nm}^3}{\frac{4}{3}\pi R^3 \text{ nm}^3} \approx \frac{5 \times 10^4}{4 \times 27^3} \approx 0.6. \quad (10.32)$$

This result reveals that the DNA inside a bacteriophage is packed to nearly crystalline densities, a fact that has far-reaching consequences for bacteriophage infectivity and lifestyle. In particular, as will be shown in the remainder of this section, as a result of these high packing densities, there is a large free-energy cost to packing the genomic DNA, which requires the services of an ATP motor. It has also been speculated that these large free energies of packing are responsible for assisting the translocation of the genetic material during the infection process itself.

The second example of interest is that of the bacterial nucleoid. In this case, while apparently less structured than in the viral context, the DNA is still highly localized. We idealize the nucleoid as a sphere of radius 0.25 μm . Given the genome length of *E. coli* of 5×10^6 bp, this implies a packing ratio

$$\nu_{\text{bacterium}} \approx \frac{5 \times 10^6 \text{ nm}^3}{4 \times (250 \text{ nm})^3} \approx 0.1. \quad (10.33)$$

In the case of a eukaryotic sperm cell, the DNA is highly compacted, as was already revealed in Figure 10.13. In particular, the sperm head (which we treat as a sphere) is roughly 5 μm in diameter, resulting in a compaction ratio

$$\nu_{\text{sperm}} \approx \frac{10^9 \text{ nm}^3}{4 \times (2500 \text{ nm})^3} \approx 0.02. \quad (10.34)$$

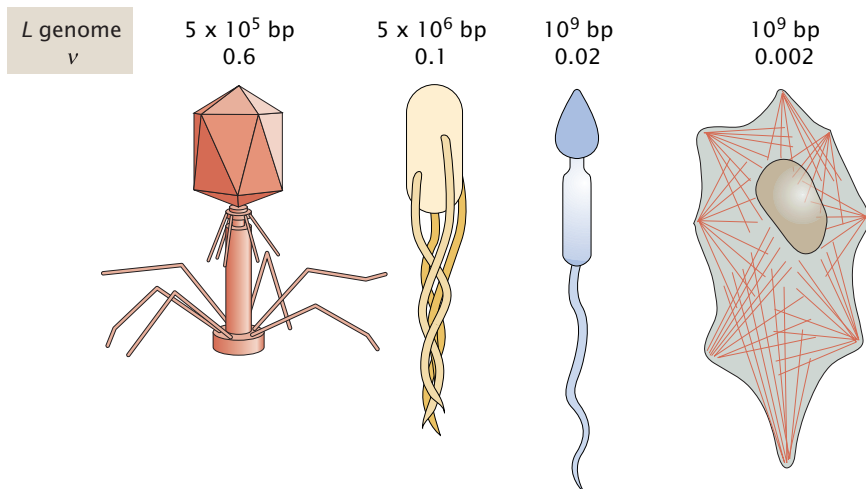


Figure 10.13: Schematics of some of the different scenarios associated with DNA packaging in lambda phage, an *E. coli* cell, a human sperm cell, and a human fibroblast. In each case, the genome length and the packing ratio ν (see Equation 10.31) are shown.

The final case we will consider is that of the eukaryotic nucleus. In this case, we can think of a genome with an approximate length of 10^9 bp housed in a nucleus with a radius of roughly $5\text{ }\mu\text{m}$. These dimensions imply a packing ratio

$$\nu_{\text{nucleus}} \approx \frac{10^9 \text{ nm}^3}{4 \times (5000 \text{ nm})^3} \approx 0.002. \quad (10.35)$$

The problem of DNA packaging is intriguing not only on the grounds of sheer geometric crowding, but also because of the recognition that the regions within which DNA is packaged (such as in a viral capsid) have linear dimensions that are comparable to the persistence length of the DNA. The claim that animates the remainder of our discussion of DNA packaging is that there is an elastic energy cost to be paid to effect such packaging, which complements the electrostatic contributions introduced in Chapter 9. Two examples to be examined presently both concern the packaging of genomic data, on the one hand, in the context of bacterial viruses and, on the other, the structure of eukaryotic chromosomes. In both cases, the smallness of the packaging region in comparison with the size of the persistence length indicates the presence of elastic (and other) free-energy penalties to pay. Indeed, the packaging of viral DNA exacts such a high cost that there is a molecular motor that carries out the necessary work to compress the viral DNA.

10.4.1 The Problem of Viral DNA Packing

In rapid succession, a number of new experimental insights into the way DNA in viruses is packaged and ejected have come into focus. As the basis for this discussion, we remind the reader of the character of the phage life cycle already shown in Figure 3.26 (p. 122). For a bacteriophage such as lambda phage, the virus attaches to the host cell receptor (the outer membrane protein LamB in the case of the *E. coli* host of lambda phage) and ejects its ~ 50 kb genome in roughly a minute. Over the next 10–20 minutes, the viral genome is replicated and the proteins coded for by that genome are synthesized. Once a sufficient concentration of proteins is available, the capsids spontaneously assemble, followed by a fascinating packaging process in which the portal motor consumes ATP and pushes the DNA into the capsid. A quantitative analysis of molecular motors will be carried out in Chapter 16. Upon completion of packaging, the capsid assembly is finished with the addition of the tails, and the mature virions are released from the ravaged bacterium to repeat their evil deeds elsewhere.

Structural Biologists Have Determined the Structure of Many Parts in the Viral Parts List

Structural biologists have begun to determine the structures of many of the molecular participants in the viral life cycle. For example, the structure of part of the portal motor that packs the DNA in the capsid has been determined, as has the structure of the membrane-puncturing device that leads to the delivery of the T4 viral genome. More important for the present discussion, there has also been a great deal of progress in determining the structures of the viral capsids

themselves as well as the DNA packaged within them. Figure 2.29 (p. 66) shows a famous gallery revealing the structures of a host of different viral capsids. One of the most notable features of these capsids is their impressive icosahedral symmetry. Another important feature is the observation that, to within a factor of 2 or so, viral capsids are all roughly the same size. Just as there are databanks that serve as a repository for protein structures, there is a databank of viral structures known as VIPER that provides atomic-level structures of a range of different capsids in various states during their maturation.

A second key structural question concerns the geometry of the packaged genome itself. Cryo-electron microscopy experiments as well as X-ray scattering experiments have revealed that the DNA within viral capsids is highly ordered. In particular, these experiments reveal a definite spacing between adjacent strands of the packed DNA. An example of such results is shown in Figure 10.14 for phage T7, while Figure 10.15 is a gallery of such structures showing both the beautiful symmetry of the phage capsids and their packaged DNA. As a result of these kinds of experiments, a number of structural hypotheses for the packaged DNA have been advanced, with one of the most favored structural models being that the DNA is packed in concentric arrangements within viral capsids as shown in Figure 10.16.

In addition to data obtained using electron microscopy, X-ray diffraction experiments have also shed light on the structure of the packed DNA. For example, measurements have been made on the spacing between the DNA segments as a function of the fractional filling of the capsid as shown in Figure 10.17. These data can be interpreted using a simple geometric model of the packaging. As shown in Figure 10.16, if the spacing between DNA strands is uniform and equal to d_s , then the volume taken up by the genome is approximately Ld_s^2 , with L being the length of the genome packed. Since the total volume taken up by the genome is roughly equal to the capsid volume, that is, $V_{\text{capsid}} \approx Ld_s^2$, it follows that the spacing between DNA strands $d_s \approx \sqrt{V_{\text{capsid}}/L}$. The predicted scaling, $d_s \sim L^{-1/2}$, seems to hold true for the experimental data, as shown in Figure 10.17, albeit over a fairly small range in genome length.

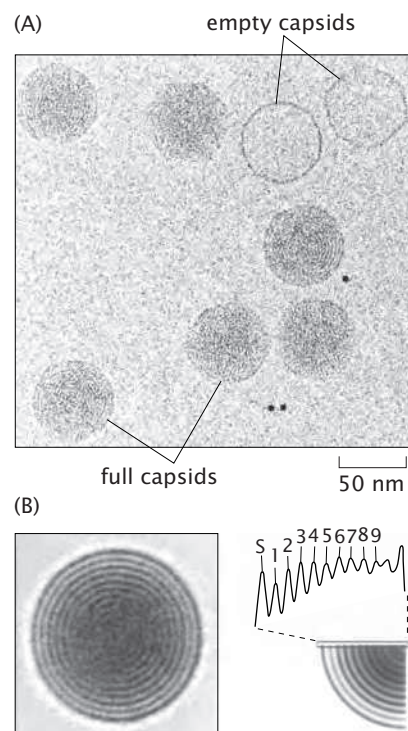


Figure 10.14: Cryo-electron microscopy images of packaged DNA. (A) Image showing several T7 capsids with their complement of packed DNA. (B) High-resolution view of a single capsid, which demonstrates the rings of ordered DNA within the capsid. On the right, the series of peaks show the alternating density as a function of radius. The numbering corresponds to the different shells of DNA, while the label "S" refers to the protein shell of the capsid itself. (Adapted from M. E. Cerritelli et al., *Cell* 91:271, 1997.)

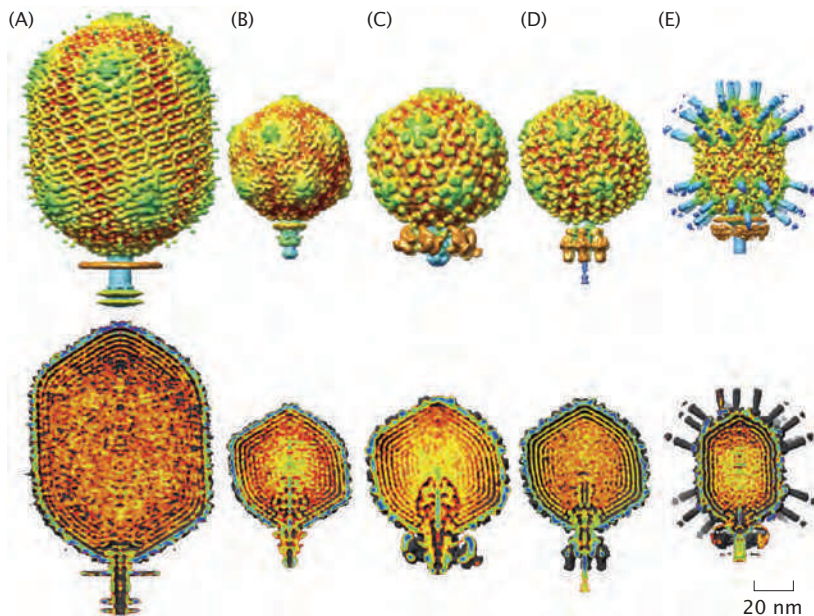


Figure 10.15: Structure of bacteriophage capsids and their packaged DNA. Cryo-electron microscopy images of the capsid structure (top row) and packaged DNA (bottom row) for bacteriophages (A) T4, (B) T7, (C) epsilon 15, (D) P22, and (E) ϕ 29. (Adapted from J. E. Johnson and W. Chiu, *Curr. Opin. Struct. Biol.* 17:237, 2007.)

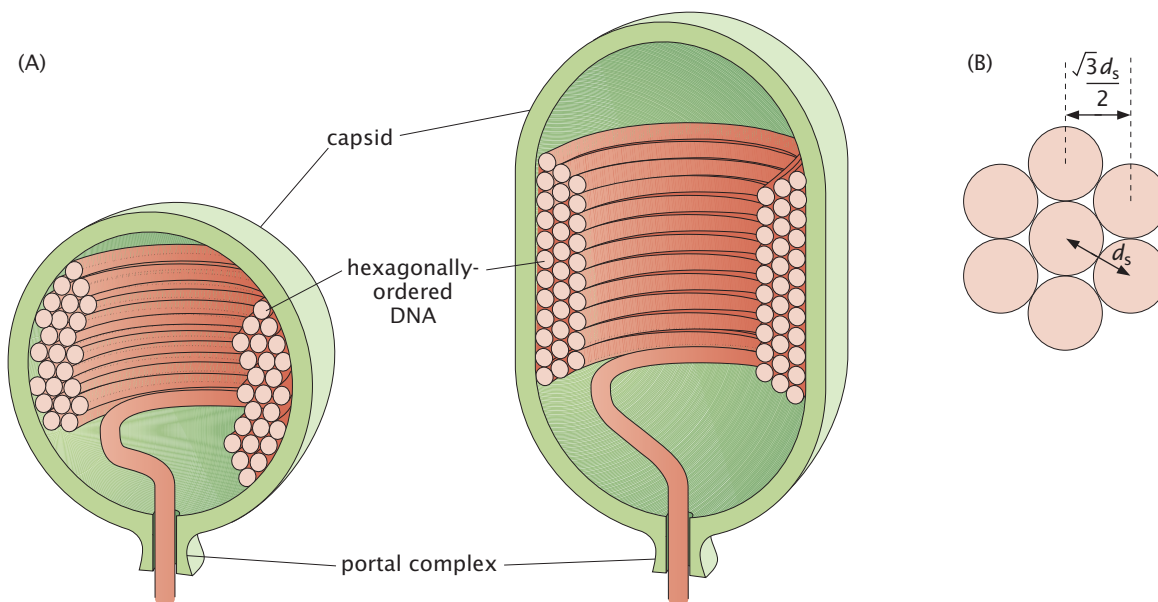


Figure 10.16: Two simplified models used to illustrate the calculation of free energy associated with packaged DNA. (A) The spherical capsid on the left is a simplified version of T7 or lambda phage. The elongated capsid on the right is a simplified version of the $\phi 29$ capsid. (B) Close up of cross-sections of DNA packed inside the capsids. Note that a given segment of DNA is surrounded by six neighboring segments in a hexagonal arrangement.

The Packing of DNA in Viruses Results in a Free-Energy Penalty

These structural insights have been complemented by single-molecule experiments on the DNA packaging process as well as *in vitro* studies of DNA ejection. The single-molecule-packaging experiments measure the force exerted by the packaging motor during the process of viral packaging itself, as shown schematically in Figure 10.18. These techniques make it possible to grab the viral DNA while it is being packaged into the viral capsid and to measure the force exerted by the portal motor as it overcomes the increasing stored energy of the packed DNA. The kind of data that emerge from these measurements is shown in Figure 10.19, which illustrates both the packaging rate and the forces that build up during packaging as a function of the percentage of the genome that is actually inside of the capsid. All told, the combination of structural insights like those described above in conjunction with biochemical and single-molecule experiments has led to a picture of the viral life cycle that can be examined quantitatively.

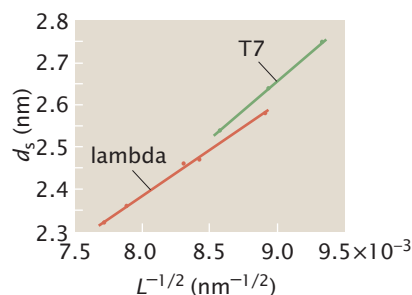


Figure 10.17: DNA spacing as a function of the packaged length. Data on the ordered packaging of DNA in two distinct bacteriophages (T7 and lambda) have been measured using X-rays and cryo-electron microscopy. (Adapted from P. K. Purohit et al., *Biophys. J.* 88:851, 2005.)

The aim of this section is to take stock of the mechanical forces that come into play during viral packaging and to reckon these forces explicitly in closed form in a simple model of DNA elasticity and electrostatic repulsions. These forces are then compared with those measured in the fascinating single-molecule experiments on DNA packaging mentioned above. The energetics of viral packaging is characterized by a number of different factors, including (i) the entropic spring effect, which favors a spread-out configuration for DNA in solution; (ii) the energetics of elastic bending, which results from inducing curvature in the DNA with a scale that is considerably smaller than the persistence length of ~ 50 nm; and (iii) those factors related to the presence of charge both on the DNA itself and in the surrounding solution. The entropic contribution is smaller by a factor of 10 or more relative to the bending energies and those mediated by the charges on the DNA and the surrounding solution. Our strategy is to examine the

elastic and charge-related interaction energies separately and then to assemble them to obtain a complete estimate of the energetics of viral DNA packaging.

In particular, we seek a free energy of packaged DNA of the form

$$G_{\text{tot}}(d_s, L) = G_{\text{bend}}(d_s, L) + G_{\text{charge}}(d_s, L), \quad (10.36)$$

where d_s is the spacing of DNA in the capsid (defined in Figure 10.16) and L is the length of DNA that is packed. With this quantity in hand we can compute the equilibrium spacing for the DNA as a function of the amount of genome packed, by setting $\partial G_{\text{tot}}/\partial d_s = 0$ and solving for d_s . Substituting the function $d_s(L)$ into G_{tot} , we can compute the force that resists packaging as $F = -dG_{\text{tot}}(d_s(L), L)/dL$. In doing this, we make the implicit assumption that the packaging process is quasistatic and the DNA remains in equilibrium. As already discussed in detail in Chapter 6, the use of equilibrium assumptions is a matter of the various time scales that are in play. In this case, the rate of packaging and the DNA relaxation time are the key variables of interest. Since roughly 100 bp are packaged each second, this leaves sufficient time for the DNA within the capsid to rearrange itself.

A Simple Model of DNA Packing in Viruses Uses the Elastic Energy of Circular Hoops

The elastic estimate that we obtain for the forces associated with viral packaging is predicated upon the most naive usage of the linear elastic theory of beams developed earlier in the chapter. In particular, we neglect the accumulation of stored elastic energy as a result of twist, for which little is known in the context of these viral DNA packing problems, and concentrate instead only on the contribution of the bending to the stored energy. Within this approximation, the elastic contribution to the free energy can be obtained by using Equation 10.7.

Though we imagine the viral DNA to be packed in the form of a helix as shown in schematic form in Figure 10.16(A), from the perspective of our elastic energy functional, the geometry may be thought of as a stacking of hoops of radius R . The key point is that although the actual radius of curvature is given by $R_c = R(1 + p^2/4\pi R^2)$, where p is the helical pitch, for the geometries of interest here, $p \approx 2$ nm while $R \approx 20$ nm and hence the parameter $p^2/4\pi R^2 \ll 1$ and can be neglected without compromising on the key features of the analysis. In light of this approximation and using the fact that for circular hoops $\kappa = 1/R$, the elastic energy can be written as

$$G_{\text{bend}} = \pi \xi_p k_B T \sum_i \frac{N(R_i)}{R_i}, \quad (10.37)$$

where $N(R_i)$ is the number of hoops that are packed at the radius R_i . The presence of this term reflects the fact that, due to the shape of the capsid, as the radius becomes smaller the DNA can pack higher up into the capsid, thus increasing the number of allowed hoops.

To make analytic progress with the expression for the stored elastic energy given above, we convert it into an integral of the form

$$G_{\text{bend}} = \frac{2\pi \xi_p k_B T}{\sqrt{3} d_s} \int_R^{R_{\text{out}}} \frac{N(R')}{R'} dR'. \quad (10.38)$$

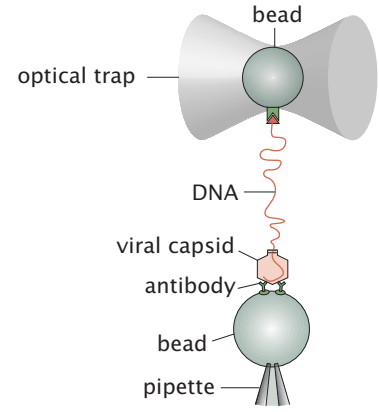


Figure 10.18: Optical tweezers measurement of the forces during DNA packaging. The viral capsid is attached to one bead using antibodies and the viral genome is attached to a second bead. This second bead is held in an optical trap and the forces are monitored as the DNA is reeled into the capsid by the ATP-consuming portal motor. (Adapted from D. E. Smith et al., *Nature* 413:748, 2001.)

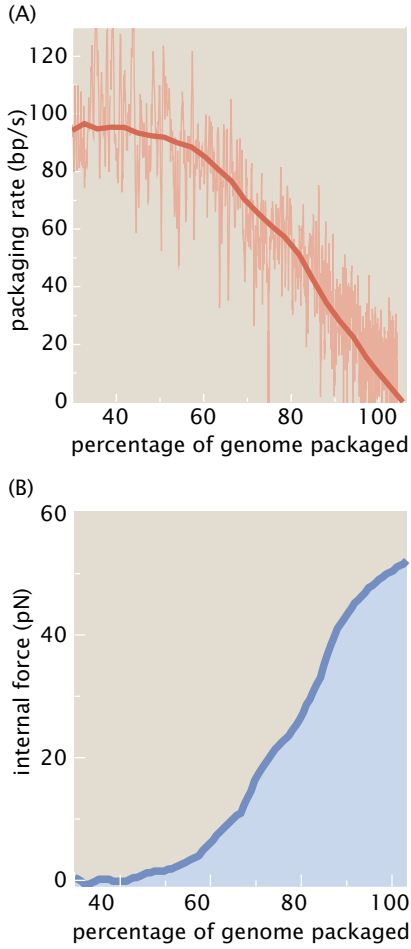


Figure 10.19: Data on DNA packaging in bacteriophage $\phi 29$. (A) Rate of DNA packaging and (B) force resisting further packaging as a function of the amount of genome present in the capsid. (Adapted from D. E. Smith et al., *Nature* 413:748, 2001.)

The summation \sum_i has been replaced by an integral $\int 2 dR' / \sqrt{3} d_s$, where $(\sqrt{3}/2)d_s$ is the center-to-center distance between adjacent strands of the DNA along the radial direction; see Figure 10.16(B). Replacing a discrete sum by an integral would in general lead to some error, but this error diminishes as the number of hoops becomes large. This calculation will yield the elastic energy in terms of the radius of the packed DNA. It is necessary to convert it to an expression in terms of the total length packed for purposes of comparison with experimental results. The length packed is generally given as

$$L = \frac{2}{\sqrt{3}d_s} \int_R^{R_{\text{out}}} 2\pi R' N(R') dR'. \quad (10.39)$$

To illustrate how to explicitly reckon the elastic energy, we first consider this energy for the simplest (and perhaps unrealistic) case of a cylindrical capsid with DNA in the inverse spool configuration. Such a geometry may be obtained, for instance, by neglecting the conical caps of an icosahedral capsid. In particular, this cylinder is characterized by the geometric parameters z (the height) and R_{out} (the radius). In this case, $N(R) = z/d_s$ since the height of the cylinder is z and the spacing between two adjacent DNA strands is d_s , and the corresponding elastic energy gotten by using Equation 10.38 is given by

$$G_{\text{bend}}(R) = \frac{2\pi \xi_p k_B T z}{\sqrt{3}d_s^2} \ln\left(\frac{R_{\text{out}}}{R}\right). \quad (10.40)$$

It follows from Equation 10.39 that the packed length is given by

$$L(R) = \frac{2\pi z}{\sqrt{3}d_s^2} (R_{\text{out}}^2 - R^2) \quad (10.41)$$

or, equivalently, $R = R_{\text{out}} \sqrt{1 - \sqrt{3}d_s^2 L / 2\pi z R_{\text{out}}^2}$.

The elastic energy may now be rewritten as a function of the packed length of DNA,

$$G_{\text{bend}} = -\frac{\pi \xi_p k_B T z}{\sqrt{3}d_s^2} \ln\left(1 - \frac{\sqrt{3}d_s^2 L}{2\pi z R_{\text{out}}^2}\right). \quad (10.42)$$

This result may be used in turn to compute the force associated with the accumulation of elastic energy. In particular, differentiating the energy obtained above with respect to the length of packed DNA, we find

$$f(L) = -\frac{dG_{\text{bend}}}{dL} = -\frac{\xi_p k_B T / 2 R_{\text{out}}^2}{1 - \sqrt{3}d_s^2 L / 2\pi z R_{\text{out}}^2}. \quad (10.43)$$

Though the algebra is messier for other capsid shapes, these same basic steps may be imitated in each case to obtain the elastic contribution to the energy. Performing these calculations is left to the problems at the end of the chapter.

DNA Self-Interactions Are also Important in Establishing the Free Energy Associated with DNA Packing in Viruses

Simple as it is, our model thus far really does not do justice to all of the competing energies in the problem of DNA packaging. In particular, the elastic contributions must be supplemented by energetic

terms related to the presence of charges both on the DNA and on the counterions in the surrounding solution. Under physiological conditions, these charge interactions manifest themselves as a repulsion between nearby segments of DNA. Rather than attempting a “first principles” treatment of this interaction, we appeal instead to experiments that attempt to measure that interaction directly. The idea of the experiments was introduced in Chapter 6 and involves the use of osmotic pressure to squeeze an ordered arrangement of DNA as shown in Figure 6.23 (p. 266) and to simultaneously measure the spacing between adjacent strands using X-rays. The osmotic pressure is built up by the presence of polyethylene glycol in solution. What emerges from such experiments is a relation between the osmotic pressure and the spacing between adjacent strands as shown in Figure 6.24 (p. 266). To this piece of empirical evidence we add the assumption that parallel strands interact through a pair potential $v(d_s)$ per unit length and that interactions are limited only to the first nearest neighbors.

Once the pressure has been determined, it is possible to compute the interaction energy, in turn, and from that energy to obtain the pair potential itself. To that end, consider N parallel strands of length l , each packed in a hexagonal array with a spacing d_s . The total energy stored in these strands by virtue of their interactions is

$$G_{\text{charge}} = 3Nlv(d_s), \quad (10.44)$$

where the factor 3 appears because each strand interacts with 6 nearest neighbors (ignoring surface effects) and we multiply by 1/2 to take care of double counting. The total volume of the assemblage is

$$V = N \frac{\sqrt{3}}{2} d_s^2 l. \quad (10.45)$$

The pressure is $p(d_s) = -dG_{\text{charge}}/dV$, where $dV = Nl\sqrt{3}d_s dd_s$. As a result, we may write the force in terms of the pressure as

$$f(d_s) = -\frac{dv(d_s)}{dd_s} = \frac{1}{\sqrt{3}}p(d_s)d_s, \quad (10.46)$$

where $f(d_s)$ is the force per unit length of the strands. We now substitute $p(d_s) = F_0 e^{-d_s/c}$ (the result of a fit to the data of Figure 6.24 on p. 266) and solve the resulting differential equation that determines the pair potential, remembering that $v(\infty) = 0$. This procedure leads to the potential

$$v(d_s) = \frac{1}{\sqrt{3}}F_0(c^2 + cd_s)e^{-d_s/c}. \quad (10.47)$$

With the interaction potential in hand, we can write the interaction contribution to the free energy as

$$G_{\text{charge}} = \sqrt{3}F_0(c^2 + cd_s)L e^{-d_s/c}, \quad (10.48)$$

where $L = Nl$ is the total length of the strands. An interesting fact about the interaction energy is that it can be controlled by the experimenter to a certain extent by changing the ionic strength of the buffer solution, a topic that was discussed in detail in Chapter 9. This, in turn, our simple model would predict, will change the pressure with which DNA is packaged inside the viral capsid, possibly affecting the ability of the virus to inject its DNA into the host cell.

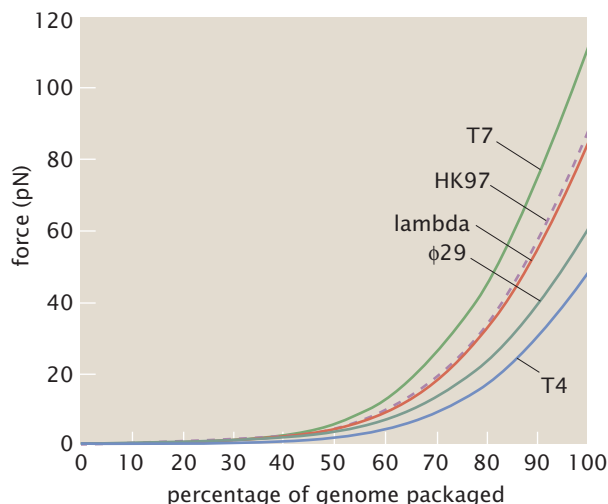
DNA Packing in Viruses Is a Competition Between Elastic and Interaction Energies

Our calculations of the elastic and interaction terms can now be assembled together to provide an expression for the total free energy associated with the packaged DNA. These two terms result in a competition that sets the length scale of the hoop spacing in the capsid. In particular, the interaction terms favor keeping the strands as far apart as possible. On the other hand, keeping the strands too far apart means that the loops will have to adopt configurations with small radii of curvature. Using the cylindrical capsid as a concrete example, the total free energy can be written as

$$G_{\text{tot}} = G_{\text{bend}} + G_{\text{charge}} = -\frac{\pi \xi_p k_B T z}{\sqrt{3} d_s^2} \ln \left(1 - \frac{\sqrt{3} d_s^2 L}{2\pi z R_{\text{out}}^2} \right) + \sqrt{3} F_0 (c^2 + c d_s) L e^{-d_s/c}. \quad (10.49)$$

Note that these equations reflect the simplifying assumption that the bending energy and the energy associated with the charges are independent, though, in fact, the persistence length (and hence bending energy) does depend upon the charges that are present. In order to compute the forces associated with the packaged DNA, we now minimize this free energy, at fixed L , to find the optimal spacing of the DNA at that particular packaged length. Similarly, we can find the resistive force that is set up as a result of the packaged DNA by evaluating $F(L) = -\partial G_{\text{tot}}/\partial L$, while using the optimal d_s at each L . The results of this model are shown in Figure 10.20; the results for $\phi 29$ should be compared with Figure 10.19(B). The reader is invited to make that comparison explicitly in the problems at the end of the chapter. More important is that we can now use the same model on a number of different bacteriophages and predict what the resistive force that builds up during packaging would be if the same laser-tweezers experiment were performed on the new phage; see Figure 10.20. Note that for T7 we predict a maximum force in excess of 100 pN, suggesting either a more powerful motor than the one in $\phi 29$, or that in the buffer conditions used in the $\phi 29$ experiment this packaging reaction would not go to completion. Either way, the model suggests that measuring the force build-up during packaging in T7 would yield further insight into this fascinating biological process.

Figure 10.20: Packing forces for different viruses. Different bacteriophages have different capsid sizes and genome lengths, resulting in different overall packaging energies. These curves show the predicted forces for a number of different bacteriophages, (T7, HK97, lambda, and $\phi 29$) for the case in which the solution conditions are the same as those used in the optical-tweezers experiment on phage $\phi 29$. (Adapted from P. K. Purohit et al., *Biophys. J.* 88:851, 2005.)



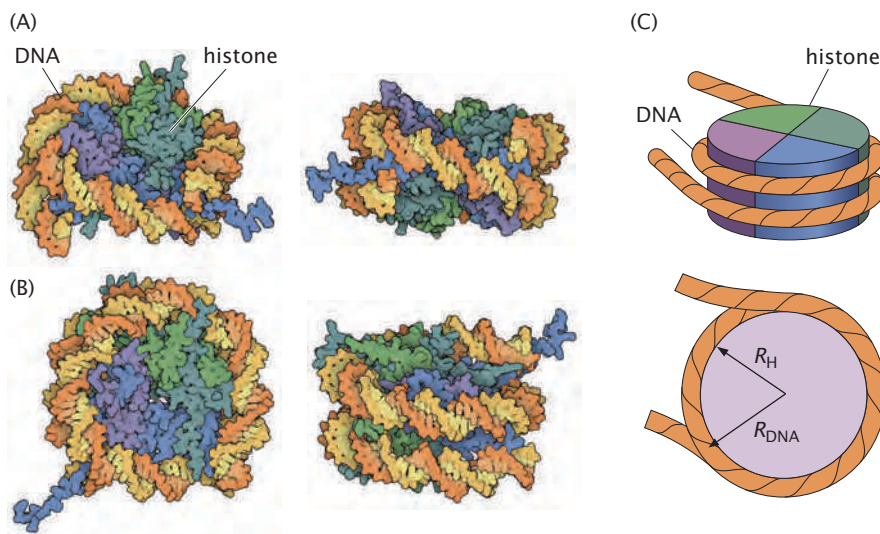


Figure 10.21: Structure of the nucleosome. (A) Atomic-level depiction of the nucleosome core particle revealing both the histone octamer and the encircling DNA. (B) Alternative atomic-level view of the nucleosome. (C) Cartoon representation of the nucleosome core particle where the histone and DNA radii (R_H and R_{DNA} , respectively) are shown. (A, B, courtesy of D. Goodsell.)

10.4.2 Constructing the Nucleosome

The second key example of DNA packaging that we aim to examine in quantitative detail from the point of view of beam theory is the eukaryotic nucleus. DNA packaged in the eukaryotic nucleus is arranged in a hierarchical structure that at the smallest scales involves the wrapping of DNA around a protein assembly known as the histone octamer to form the nucleosome already introduced in Figure 2.26 (p. 63). At the next scale up, individual nucleosomes are arranged to make the 30 nm chromatin fiber in which the nucleosomes form an orderly array like beads on string and are further condensed into a helical coil. In this discussion, we consider the smallest link in the chain of structures, namely, the nucleosome core particle. To shed light on this structure with a level of detail in excess of that given in Figure 2.26, Figure 10.21 shows an atomic-level depiction of the wrapping of DNA around the histone octamer. The relevant orders of magnitude to bear in mind are that the histone octamer is roughly 7 nm in diameter and 6 nm high, and will be treated presently as a tiny cylinder. The DNA wraps around this cylindrical particle roughly 1.75 times with 147 bp.

Estimate: Sizing Up Nucleosomes To get a rough impression of the packaging of eukaryotic DNA, we note that there are 147 bp assigned to each histone, with intervening linker regions of roughly 50 bp. As a result, we make the simple estimate that there are roughly 200 bp per histone. For a genome with 3×10^9 bp, this implies that the number of histone octamers is roughly

$$N_{\text{octamer}} = \frac{3 \times 10^9 \text{ bp}}{200 \text{ bp}} \approx 10^7. \quad (10.50)$$

This estimate should be seen as an underestimate because some fraction of the histones will probably be free in solution, for example.

It is also of interest to estimate the fraction of the nuclear volume that is taken up by histones. Given that we think of the histone octamer as a cylinder of diameter roughly 7 nm and



ESTIMATE

height 6 nm, each such histone has a volume of

$$\Omega_{\text{histone}} = \pi r^2 h \approx 225 \text{ nm}^3. \quad (10.51)$$

Hence, the total volume taken up by the 10^7 histones is $\sim 2.3 \times 10^9 \text{ nm}^3$, which should be compared with the roughly $5 \times 10^{11} \text{ nm}^3$ of volume available within the nucleus of $5 \mu\text{m}$ radius.

Eukaryotic DNA is packed in a hierarchical structure, with much of the DNA sequestered within nucleosomes. This fact is biologically mysterious since many of the most essential transactions involving DNA require proteins to bind to specific target sites on the DNA. For example, we have already discussed the observation that the process of transcription is mediated by batteries of transcription factors (activators and repressors) that bind the DNA and modulate transcription. As a result, it is of great interest to understand the accessibility of nucleosomal DNA. How easily can target sites be accessed as a function of their distance from the unwrapped ends of the nucleosomal DNA? To examine this question, we begin by considering the free energy of nucleosome assembly, followed by a statistical mechanics treatment of the equilibrium accessibility of nucleosomal DNA.

Nucleosome Formation Involves Both Elastic Deformation and Interactions Between Histones and DNA

One of the goals of this section is to examine the free-energy balance associated with the assembly of the nucleosome core particle. We recall that the persistence length for DNA is of the order of 50 nm. By way of contrast, note that the radius of curvature associated with the DNA wrapped around the histone octamer is roughly 4.5 nm. Like with the viral packaging example considered in the previous section, there are clearly elastic consequences to such highly deformed DNA fragments. In addition to the cost of elastic deformation, there are other energy contributions to consider as well. In particular, we need to consider the favorable electrostatic interactions between the DNA and the histone octamer since the DNA itself is negatively charged and the histone octamer surface is covered with lysine and arginine residues that present a compensating positive charge. Within this simple picture, the free energy of formation for a nucleosome can be written as

$$G_{\text{nucleosome}} = G_{\text{bend}} + G_{\text{DNA-histone}}. \quad (10.52)$$

To construct a preliminary estimate of the deformation energy associated with the formation of the nucleosome core particle, we treat the DNA as a featureless rod subject to a uniform state of deformation. In light of Equation 10.8, the energy stored in each turn of the DNA by virtue of its deformation is given by

$$G_{\text{bend}} = \frac{\pi \xi_p k_B T}{R_{\text{DNA}}}, \quad (10.53)$$

where R_{DNA} is the radius of curvature of the wrapped DNA.

To proceed with our estimate, we need to compute the curvature associated with the 147 bp segment when it is wrapped around the histone octamer. As shown in Figure 10.21, the DNA wraps around the histone octamer in a helical pattern (we are not speaking here of the DNA double helix itself) with a small pitch. For simplicity of calculation, we use an approximate expression for the curvature in which the helical pitch is neglected. In addition, we assume that the DNA is wrapped fully around the histone octamer two times, resulting in a total bending free energy cost of $G_{\text{bend}} = 2\pi\xi_p k_B T / R_{\text{DNA}} \approx 70 k_B T$.

The second key contribution to the free energy of formation of the nucleosome is dictated by the interactions between the DNA and the positively charged residues on the histones. As a simplest model, we characterize this interaction energy via an adhesive energy γ_{ad} , which has units of energy/length. In light of this model, the adhesive contribution to the total free energy is

$$G_{\text{DNA-histone}} = 4\pi R_{\text{DNA}} \gamma_{\text{ad}}, \quad (10.54)$$

where we have assumed that the DNA is twice wrapped fully around the histone octamer. In the next section, we show how this adhesive energy can be deduced from experimental data on the equilibrium accessibility of nucleosomes.

10.4.3 Equilibrium Accessibility of Nucleosomal DNA

As noted above, DNA molecules in eukaryotic cells wind around histone octamers to form the nucleosome. In this state, the DNA is not directly accessible to regulatory proteins since their binding sites are occluded by the nucleosome. We extend the discussion of nucleosome assembly given above to investigate the statistical mechanics of binding of regulatory proteins to DNA in the nucleosome. Key experiments on nucleosome accessibility were performed *in vitro* by assessing the susceptibility of particular sites on the DNA to cleavage by restriction enzymes as a function of the distance of these sites from the unwrapped ends of the nucleosomal DNA. These restriction enzymes are proteins that cleave DNA at specific recognition sites and served as a convenient readout for assessing nucleosome accessibility. Effectively, these experiments provide a position-dependent equilibrium constant that depends upon the distance of the site of interest from the unwrapped ends of the nucleosomal DNA.

The Equilibrium Accessibility of Sites within the Nucleosome Depends upon How Far They Are from the Unwrapped Ends

The model put forward to interpret these results envisions the binding of a DNA-binding protein (for the experiment in question, restriction enzymes were used as the protein of interest) to its target site as a two-step process: first the DNA unwraps from the histones simply as a result of thermal fluctuations, and then the restriction enzyme binds to its specific site, which is no longer occluded by the nucleosome. This process is shown schematically in Figure 10.22. In Figure 10.23 we present the results for the equilibrium constant of DNA unwrapping from the nucleosome. The simple observation is that this equilibrium constant decreases the further away the binding site is from the unwrapped ends of the nucleosome. To address these data in a quantitative way, we turn once again to statistical mechanics and

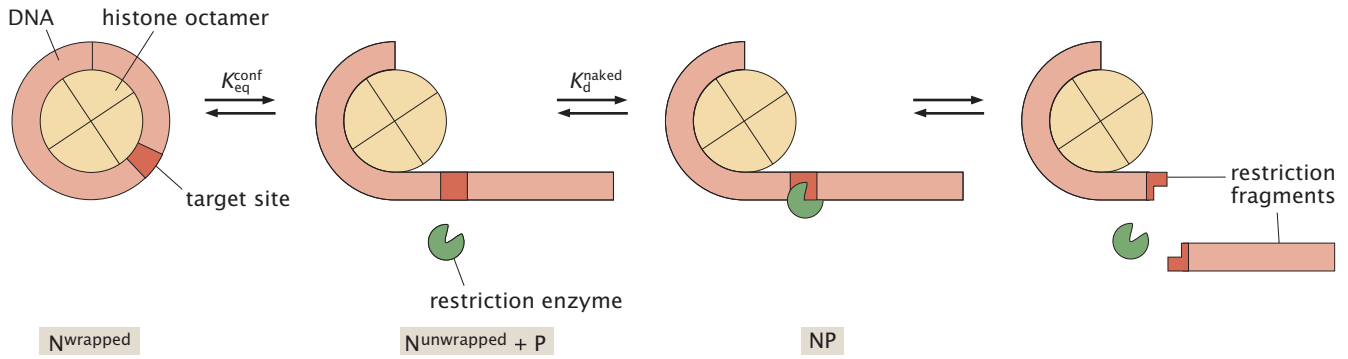


Figure 10.22: Experiment to measure equilibrium accessibility of nucleosomes. Nucleosomal DNA is prepared with a binding site for a restriction enzyme. A wrapped nucleosome, N^{wrapped} can transiently unwrap ($N^{\text{unwrapped}}$) and interact with a restriction enzyme, P, upon exposure of its target binding site forming the NP complex. A measurement is made of the probability of restriction digestion as a function of the distance of the target site from the unwrapped ends of the nucleosomal DNA.

Figure 10.23: Equilibrium constant for site exposure as a function of the distance of the site from the unwrapped end of the nucleosomal DNA. The approximate positions of the binding sites are shown. The dyad is the center of symmetry of the DNA molecule when it is wrapped around the histone octamer. (Adapted from K. J. Polach and J. Widom, *J. Mol. Biol.* 254:130, 1995.)

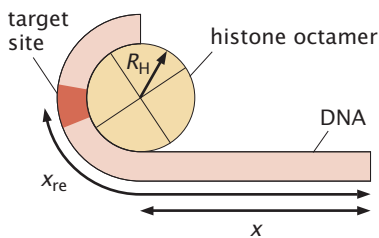
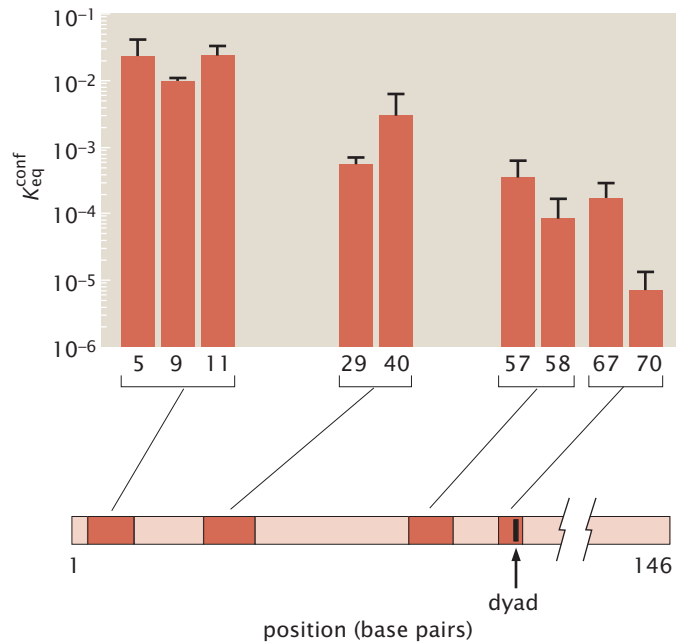


Figure 10.24: Geometry of site accessibility of the nucleosome. The coordinate x refers to how much the DNA is unwrapped and x_{re} refers to the minimum distance that the DNA needs to be unwrapped in order to access the target site.

compute the probability that an enzyme is bound to the target site as a function of the target site location.

The simple model we consider has two degrees of freedom as shown in Figure 10.24. One of the relevant degrees of freedom is the length of DNA unwound from the nucleosome, x , which is continuous and takes values between 0 and $L \approx 25$ nm; this length is half of the total DNA length wound in a typical nucleosome. The other degree of freedom is discrete, and it simply keeps track of whether a restriction enzyme is bound or not to its specific site. The binding site for the enzyme we take to be located at position x_{re} , which is somewhere along the DNA length between 0 and L . The quantity we wish to compute is the probability p_{bound} that the restriction enzyme is bound to its site, as a function of the restriction site location.

The probability that an enzyme is bound is readily computed in the grand canonical ensemble introduced in Section 7.2.1 (p. 289). The

grand partition function is

$$\mathcal{Z} = \underbrace{\int_0^L e^{\frac{1}{k_B T}(\gamma_{ad} - \gamma_{bend})(L-x)} \frac{dx}{a}}_{\text{no enzyme bound}} + \underbrace{e^{\frac{\mu}{k_B T}} e^{\frac{\varepsilon_{bind}}{k_B T}} \int_{x_{re}}^L e^{\frac{1}{k_B T}(\gamma_{ad} - \gamma_{bend})(L-x)} \frac{dx}{a}}_{\text{enzyme bound}}. \quad (10.55)$$

The first term in \mathcal{Z} is a sum over all possible values for the length x of DNA unwound from the nucleosome, with no restriction enzyme bound. The sum over discrete base pairs is for convenience replaced by an integral over the DNA length, in which case a is the distance between consecutive base pairs; this guarantees that the number of states, $\int_0^L dx/a$, is equal to the number of base pairs of DNA in the nucleosome. The statistical weight for a given x is determined by the energy per unit length for DNA binding to histones ($-\gamma_{ad}$) and the energy per unit length associated with bending the DNA (γ_{bend}).

The second term accounts for unwound states of the DNA when there is a restriction enzyme bound to its specific binding site. For this binding to occur, the DNA must unwind at least by an amount given by the position of the binding site x_{re} , which sets the lower bound of the integral. If the DNA is unwound by less than this amount, the binding site remains inaccessible. The weight associated with a given x now has two additional factors, one that takes into account the chemical potential of the restriction enzymes in solution, μ , and the other arising from the binding energy of the enzyme to its specific site on the DNA, $-\varepsilon_{bind}$.

Now, the probability that the enzyme is bound to its restriction site is the ratio of the second term in the partition function in Equation 10.55 to the total partition function,

$$p_{\text{bound}} = \frac{\frac{[P]}{[P]_0} e^{\frac{\varepsilon_{bind}}{k_B T}} \left(e^{\frac{1}{k_B T}(\gamma_{ad} - \gamma_{bend})(L-x_{re})} - 1 \right)}{\frac{[P]}{[P]_0} e^{\frac{\varepsilon_{bind}}{k_B T}} \left(e^{\frac{1}{k_B T}(\gamma_{ad} - \gamma_{bend})(L-x_{re})} - 1 \right) + \left(e^{\frac{1}{k_B T}(\gamma_{ad} - \gamma_{bend})L} - 1 \right)}, \quad (10.56)$$

where we have made use of the ideal-solution result for the chemical potential, $e^{\mu/k_B T} = [P]/[P]_0$. Here $[P]$ is the concentration of proteins (restriction enzymes) in solution, while $[P]_0$ is the concentration at which the chemical potential is zero.

To make contact with the experimental results, we should translate the above expression into the language of dissociation constants. The probability that a restriction enzyme is bound to DNA can be expressed in terms of the equilibrium dissociation constant

$$K_d = \frac{[P][N]}{[NP]} \quad (10.57)$$

for the reaction $P + N \rightleftharpoons NP$. Here P denotes free restriction enzymes, N are nucleosomes free of restriction enzymes, and NP are nucleosome-protein complexes.

The probability that a restriction enzyme is bound can now be expressed in terms of the ratio of the concentration of restriction enzyme-nucleosome complexes and the total concentration of all nucleosomes,

$$p_{\text{bound}} = \frac{[NP]}{[NP] + [N]}. \quad (10.58)$$

Using Equation 10.57, this leads to the expression

$$p_{\text{bound}} = \frac{[P]/K_d}{1 + [P]/K_d}. \quad (10.59)$$

Finally, comparing this equation with Equation 10.56, we deduce

$$K_d = \frac{[P]_0}{e^{\frac{\epsilon_{\text{bind}}}{k_B T}}} \frac{e^{\frac{1}{k_B T}(\gamma_{\text{ad}} - \gamma_{\text{bend}})L} - 1}{e^{\frac{1}{k_B T}(\gamma_{\text{ad}} - \gamma_{\text{bend}})(L - x_{\text{re}})} - 1} \quad (10.60)$$

for the equilibrium dissociation constant.

We now wish to connect our model with the quantity measured experimentally, namely the equilibrium constant for nucleosomal accessibility, $K_{\text{eq}}^{\text{conf}}$. It is defined as the ratio of the concentration of nucleosomes whose DNA is sufficiently unwrapped so as to allow the restriction enzymes to bind to the concentration of nucleosomes whose DNA is wrapped such that no DNA binding is allowed. We call these concentrations $[N^{\text{unwrapped}}]$ and $[N^{\text{wrapped}}]$, respectively as shown in Figure 10.22. Our claim is that $K_{\text{eq}}^{\text{conf}}$ is embedded as a part of the total K_d . In order to see this, we multiply and divide K_d by $[N^{\text{unwrapped}}]$, resulting in

$$K_d = \frac{[N][P]}{[NP]} \cdot \frac{[N^{\text{unwrapped}}]}{[N^{\text{unwrapped}}]}. \quad (10.61)$$

Note that $[NP]$ is the same as $[N^{\text{unwrapped}}P]$, since in order to form the nucleosome–restriction-enzyme complex, we need the nucleosomes to be sufficiently unwrapped. With this in mind, we can express the dissociation constant as

$$K_d = \underbrace{\frac{[N^{\text{unwrapped}}][P]}{[N^{\text{unwrapped}}P]}}_{K_d^{\text{naked}}} \frac{[N]}{[N^{\text{unwrapped}}]}, \quad (10.62)$$

where $K_d^{\text{naked}} = [P]_0 / e^{\epsilon_{\text{bind}}/k_B T}$ is the dissociation constant associated with naked DNA, where the equilibrium between restriction enzymes bound to the DNA and those in solution is established in the absence of histones. The last term in Equation 10.62 can be written in terms of the wrapped and unwrapped nucleosome species. In order to do that, we remind ourselves that $[N] = [N^{\text{unwrapped}}] + [N^{\text{wrapped}}]$, leading to

$$\frac{[N]}{[N^{\text{unwrapped}}]} = \frac{[N^{\text{unwrapped}}] + [N^{\text{wrapped}}]}{[N^{\text{unwrapped}}]} = 1 + \underbrace{\frac{[N^{\text{wrapped}}]}{[N^{\text{unwrapped}}]}}_{(K_{\text{eq}}^{\text{conf}})^{-1}}, \quad (10.63)$$

which results in a dissociation constant as a function of K_d^{naked} and $K_{\text{eq}}^{\text{conf}}$,

$$K_d = K_d^{\text{naked}} \left(1 + \frac{1}{K_{\text{eq}}^{\text{conf}}} \right). \quad (10.64)$$

Comparing this with Equation 10.60, we arrive at the expression

$$K_{\text{eq}}^{\text{conf}} = \frac{e^{\frac{1}{k_B T}(\gamma_{\text{ad}} - \gamma_{\text{bend}})L} - e^{\frac{1}{k_B T}(\gamma_{\text{ad}} - \gamma_{\text{bend}})x_{\text{re}}}}{e^{\frac{1}{k_B T}(\gamma_{\text{ad}} - \gamma_{\text{bend}})L} \left(e^{\frac{1}{k_B T}(\gamma_{\text{ad}} - \gamma_{\text{bend}})x_{\text{re}}} - 1 \right)}, \quad (10.65)$$

which depends only on the position of the restriction binding site along the DNA sequence, and not on the strength of the binding site,

ϵ_{bind} .

We are now in the position to compare the theoretical results with the measurements discussed earlier. Figure 10.25 shows both the experimental data points and a best fit to these points using the adhesive energy as a free parameter. We conclude that $K_{\text{eq}}^{\text{conf}}$ decreases by four orders of magnitude as the position of the binding site shifts from the end of the wound DNA to the dyad at $x = 70 \text{ bp} \approx 25 \text{ nm}$, in good agreement with the measurements of Polach and Widom (1995). (Note that the sharp increase at $x_{\text{re}} = 0$ seen in Figure 10.25 is an artifact of the continuum approximation assumed by the model; it leads to an infinite $K_{\text{eq}}^{\text{conf}}$ when $x_{\text{re}} \rightarrow 0$. Deriving the discrete model that remedies this pathology of the continuous model is left as a problem for the reader at the end of the chapter.) Although these experiments address the unwrapping of short DNA segments from histones *in vitro*, it is likely that similar energetic considerations apply to DNA wrapping and unwrapping for chromosomes in the nuclei of living eukaryotic cells.

10.5 The Cytoskeleton and Beam Theory

The mechanical behavior of cells is largely determined by the material properties and architectural arrangements of structural elements that tend to look like beams or networks of beams, collectively known as the cytoskeleton. A particularly dramatic illustration of the centrality of these filamentous structural elements in determining cell shape and mechanics is afforded by the neuron. In the late nineteenth century, the Spanish anatomist Santiago Ramón y Cajal keenly observed neurons and then drew spectacular pictures to reflect his observations, as shown in Figure 10.26. The most striking feature of this cell is its enormously elaborate and yet tightly controlled morphology. No simple physical process such as surface tension or bending can account for this kind of elaborate structure; it must be carefully constructed from within the cell.

Eukaryotic Cells Are Threaded by Networks of Filaments

If we look inside a cell, we can see that it is filled with filamentous structures that appear to provide the structural support for these elaborate and complex shapes. These filaments are typically made of protein, but it is striking that they are extremely large, much larger than any individual polypeptide chain could actually be. In fact, these filaments are made of many identical copies of protein subunits that assemble in helical aggregates, as already shown schematically in Figure 2.27 (p. 64). Thousands or even tens of thousands of subunits come together to form a single filament. If we experimentally disrupt the association of these subunits, for example using drugs or mutations, then cell shape is horrifically compromised. Quantitatively, how can we think about the ways in which the mechanical properties of these filaments are responsible for the structure and organization of the cell? From a mechanical perspective, the most striking feature of these filaments is their aspect ratio: their length is orders of magnitude larger than their cross-sectional dimension, an observation we also made of DNA. As a result, we can use a similar

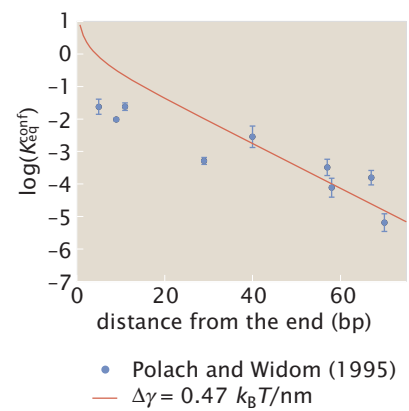
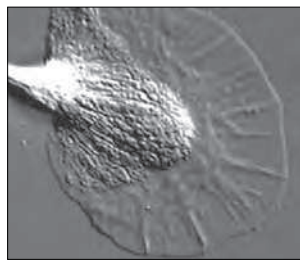


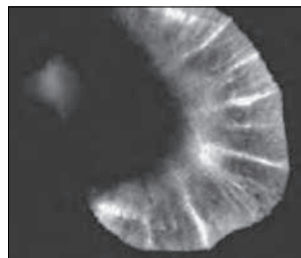
Figure 10.25: Equilibrium constant for DNA accessibility. The data points result from measurements described in this section. The curve is the result of the model worked out in Section 10.4. We define the parameter $\Delta\gamma = \gamma_{\text{ad}} - \gamma_{\text{bend}}$. Using a persistence length of 50 nm implies $\gamma_{\text{bend}} = 1.2 k_B T/\text{nm}$. (Data adapted from K. J. Polach and J. Widom, *J. Mol. Biol.* 254:130, 1995.)



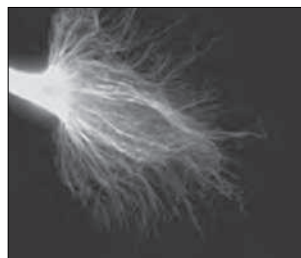
Figure 10.26: Drawing of a neuron. This figure shows one of the many impressive drawings from Ramón y Cajal of a neuron. In particular, this figure shows a human Purkinje cell and illustrates the morphological complexity dictated by cytoskeletal proteins. (Adapted from S. Ramón y Cajal, *Histology of the Nervous System*. Oxford University Press, 1995.)



(A)



(B)



(C)

10 μ m

Figure 10.27: Cytoskeletal filaments in a neural growth cone. (A) Differential interference contrast image of an *Aplysia* (sea slug) neuron growth cone. (B) Fluorescence image of actin in the growth cone. (C) Fluorescence image of microtubules in the growth cone. The distribution of intermediate filaments (neurofilaments) is similar to the distribution of microtubules shown here. (Courtesy of P. Forscher.)

mechanical treatment to describe the properties of these filaments as structural elements.

10.5.1 The Cellular Interior: A Structural Perspective

If we take the neuron shown in Figure 10.26, and separate out the molecularly distinct filament types that it contains, we will find that three predominate. One group of these filaments are microtubules, which tend to be straight filaments about 25 nm in diameter. We will also find actin filaments, which are roughly 8 nm in diameter, and finally, intermediate filaments, so-called because they are intermediate in size between actin and microtubules. These different filaments have distinct locations within the neuronal cell as depicted in Figure 10.27, which shows actin and microtubules within the growth cone found at the tip of a developing neuron. The various filaments also have different mechanical properties and perform distinct biological functions.

Microtubules Microtubules have been implicated in a variety of key cellular processes, serving as a conduit for the motion of molecular motors such as kinesin and dynein as well as presiding over the organization of cell division. Several layers in the structural hierarchy of microtubule organization are shown in Figure 10.28. These filaments are made up of individual tubulin subunits with a characteristic scale of about 8 nm. These subunits bind to each other in a head-to-tail fashion, forming strings called protofilaments. A test tube containing purified tubulin subunits will spontaneously assemble to form microtubules under appropriate conditions. These microtubules are hollow cylindrical structures in which a cross-section contains 13 tubulin subunits arranged in a ring. Looking along the side of a microtubule, we can see that each individual subunit forms two kinds of protein-protein contacts with its neighbors. Around the ring, these subunits have lateral interactions, while along their length, they form head-to-tail contacts such that all individual subunits are pointing in the same direction along the microtubule lattice. Within cells, microtubules are frequently found organized into bundles. For example, microtubules in the mitotic spindle are responsible for separating duplicated chromosomes during cell division. Each chromosome is anchored to a bundle of approximately 30 microtubules that are all oriented in the same direction and cross-linked together. As we will discuss below, bundling of cytoskeletal filaments alters their mechanical properties in ways that are important for cell function.

Actin Filaments Like microtubules, actin filaments are helical assemblies of globular subunits. Although the individual subunits are comparable in size (5 nm for actin versus 8 nm for microtubules), the filaments are smaller because they are only 2 protofilaments across in cross-section rather than 13, as shown in Figure 10.29. Inside cells, actin filaments are also frequently found in bundles. Actin serves a wide variety of structural and functional roles in cells, with some of its most familiar activities related to motility. As a result of their ability to hydrolyze ATP, these filaments can perform useful work resulting in motile functions such as the creation of protrusions in motile cells (see Figure 15.2 on p. 576) and the motion of bacterial pathogens (see Figure 15.3 on p. 577). Together with

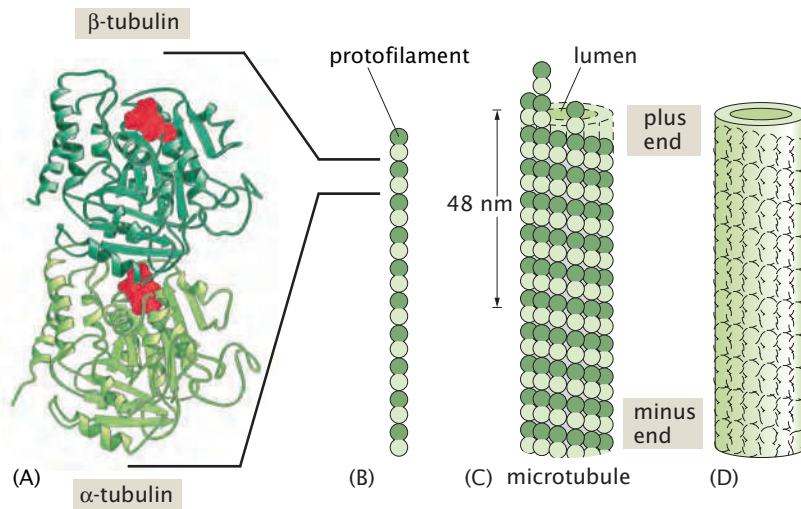


Figure 10.28: Structure of a microtubule. (A) Ribbon diagram depiction of a tubulin subunit, consisting of a dimer of the two proteins α -tubulin and β -tubulin. Bound GTP molecules are shown in red. (B) Individual tubulin subunits assemble in a head-to-tail fashion. A single string of subunits is called a protofilament (such structures are unstable on their own). (C) Microtubules are hollow cylinders of 13 protofilaments. Because of the intrinsic polarity of the tubulin subunit, the microtubule itself is also polarized, with all β -tubulin proteins exposed at the plus end and all α -tubulin proteins exposed at the minus end. (D) For idealized representations, we will treat the microtubule as a uniform hollow cylinder with a diameter of 25 nm. (Adapted from B. Alberts et al., *Molecular Biology of the Cell*, 5th ed. Garland Science, 2008.)

myosin motors, actin filaments are also required for muscle contraction. The development of the complex network of actin filaments seen in cells is tightly controlled in both space and time by an array of actin-related proteins, as will be discussed in more detail in Chapter 15. One of the outcomes of the activity of these actin-related proteins is the ability to form higher-order structures such as branched networks.

Intermediate Filaments The subunits of intermediate filaments are very different from either actin or tubulin subunits in that they are elongated proteins rather than globular. Assembly of intermediate filaments involves lateral bundling and twisting of these subunits into coiled-coil structures. Their functional role is, at least in part, to provide mechanical resistance to stretching of the cells in tissues. These coiled-coil dimers join to form antiparallel tetramers that are,

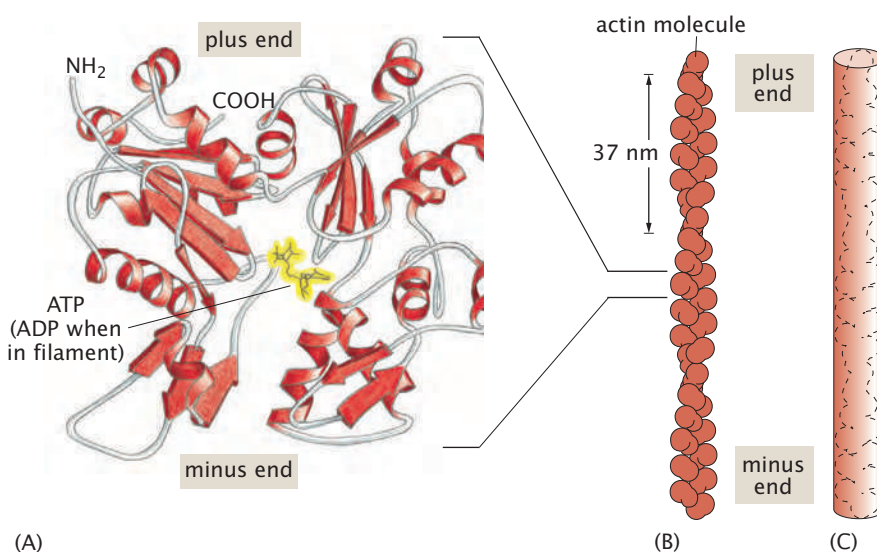


Figure 10.29: Structure of actin. (A) Ribbon diagram depiction of a single G-actin monomer. Each actin monomer is a globular protein 5.4 nm in diameter that can bind a single nucleotide (ATP or ADP). (B) F-actin filaments are assembled from two protofilaments that wrap around each other with a 37 nm repeat. All actin subunits within the filament are oriented in the same direction. Because the filament consists of two parallel protofilaments, the addition of a single subunit extends the overall filament length by 2.7 nm. (C) For idealized treatments of actin filaments, we will regard them as uniform cylinders with a diameter of 8 nm. (Adapted from B. Alberts et al., *Molecular Biology of the Cell*, 5th ed. Garland Science, 2008.)

in turn, wound together in a rope-like structure that lends them their mechanical strength.

Prokaryotic Cells Have Proteins Analogous to the Eukaryotic Cytoskeleton

The assembly of large helical filaments from small protein subunits appears to be a universally conserved feature of cellular organization. In particular, actin and tubulin are found in nearly all cells on the planet, including bacteria and archaea, as well as eukaryotes. In bacteria, for historical reasons, their names are different (for example, the tubulin homolog in bacteria and archaea is called FtsZ and various actin homologs go by names such as MreB and ParM; often, bacterial proteins are named based on the genetic screen in which they were discovered rather than on their structure) but it is clear from protein crystal structures that the proteins are extraordinarily similar. The protein structure of these organizational molecules has been assiduously conserved over three billion years of evolution. In all cells, cytoskeletal filaments contribute to the determination of cell shape and the mechanical processes associated with cell division. Their exact roles vary substantially from cell to cell. However, the basic theme that they provide large-scale organization by self-assembly is universal.

10.5.2 Stiffness of Cytoskeletal Filaments

The Cytoskeleton Can Be Viewed as a Collection of Elastic Beams

We have asserted that these cytoskeletal filaments are responsible for the organization and mechanical properties of cells. This implies that they can provide mechanical support against applied loads and overcome competing forces that arise from a variety of mechanisms such as membrane elasticity, binding reactions, surface tension, and so on. A prerequisite to deciphering the role of architectural elements is to characterize their response as individual mechanical elements. In particular, as was already shown for DNA, a critical measure of beam stiffness is captured by the persistence length or flexural rigidity.

For cytoskeletal elements, three different factors determine the effective stiffness of structures within the cell. The first, of course, is the intrinsic flexibility of the filaments themselves. The second is the presence or absence of filament binding proteins that modulate filament stiffness, and the third is larger-scale organization of filaments into bundles, networks, or gels. Multiple techniques have been developed to measure stiffness at each of these scales. At the level of single filaments, thermal motions are sufficient to induce filament bends over a spatial scale that can be directly observed in a light microscope. Therefore, the most straightforward and arguably the least invasive way to measure filament stiffness has been by direct observation of filaments either fluorescently labeled so that individual filaments can be visualized or, in the case of microtubules, using enhanced contrast microscopy, which permits visualization even of unlabeled microtubules. Other measurements use mechanical manipulation of filaments so that their deformation can be observed under a specific applied load. These kinds of measurements may use either calibrated microneedles or beads held in optical traps to deliver known forces to the ends of captured filaments. A dramatic example of this approach is

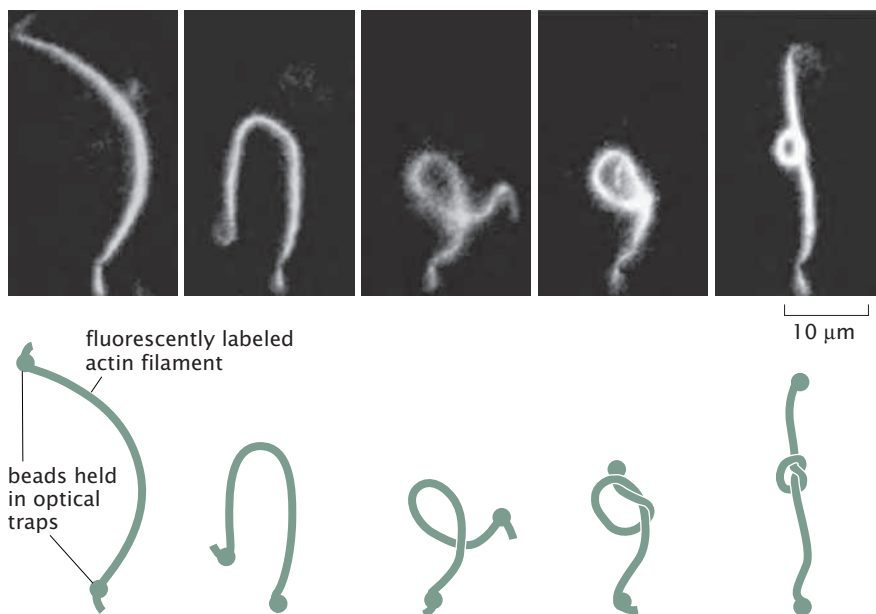


Figure 10.30: Actin tied in a knot. The flexibility of the actin filament is vividly illustrated in this experiment where two beads are attached to opposite ends of a filament and manipulated using optical traps to tie a tiny knot. (Adapted from Y. Arai et al., *Nature* 399:446, 1999.)

shown in Figure 10.30, where optical tweezers are used to tie a single actin filament into a knot.

Cells build mechanical structures by combining individual cytoskeletal filaments in ways that alter their large-scale mechanical properties. This is most easily envisioned in the common situation where cytoskeletal filaments are bundled together to create stronger elements that resist bending. Figure 10.31 shows bundles of actin filaments forming protrusions of the cell membrane on the surface of human intestinal epithelial cells. The function of these protrusions is to vastly increase the surface area through which nutrients can be absorbed. As we will explore in detail in Chapter 11, bending the cell plasma membrane, particularly in such a tortuous manner, requires energy input. Looked at from another perspective, the membrane exerts a force on these bundles of actin filaments. The persistence of

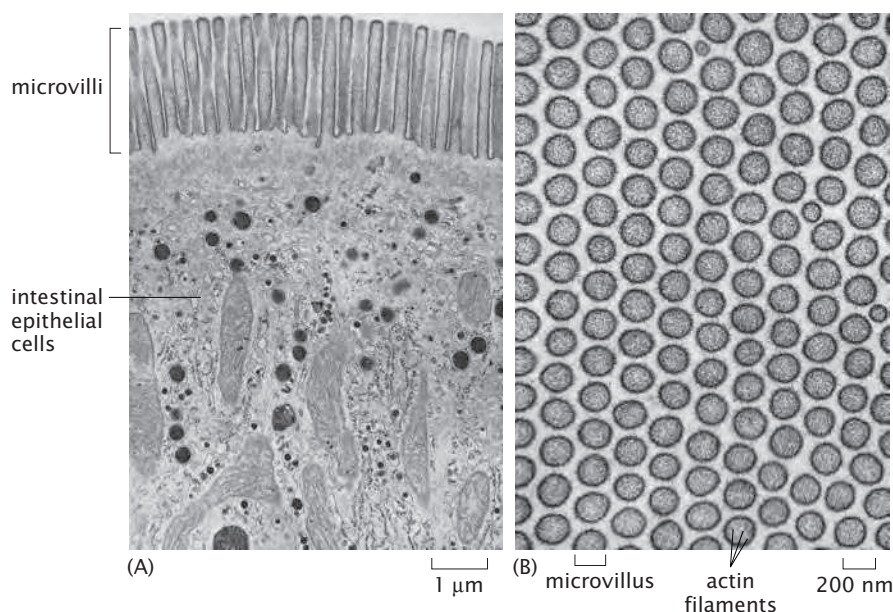


Figure 10.31: Intestinal brush border. Two electron micrographs show the organization of actin filaments and cell membrane at the absorptive surface of human intestinal epithelial cells. (A) Section of intestinal epithelial cells seen from the side with the absorptive microvilli at the top. (B) Glancing cross-section across the top of the cell showing the close packing and organization of the microvilli with their internal bundles of actin filaments. (Adapted from D. W. Fawcett, *The Cell, Its Organelles and Inclusions: An Atlas of Fine Structure*. W. B. Saunders, 1966.)

the bundle structures at the surface of epithelial cells demonstrates that the force exerted by the membrane is not sufficient to bend or buckle these filament bundles. In the next section, we will make an estimate showing that a single filament with length comparable to a microvillus cannot support the forces applied by the membrane and will buckle. Therefore, it seems reasonable to conclude that the biological function of bundle formation in this case is to increase the stiffness of the composite filament beyond that of single filaments. In a human, a typical microvillus includes roughly 30 actin filaments packed together in a hexagonal lattice.

A very different and even more elegant use of the enhanced stiffness of actin-filament bundles is found in the vertebrate inner ear hair cell, the cell type responsible for converting motions of air into perceived sound. Hair cells within the inner ear lie adjacent to a tectorial membrane as shown in Figure 10.32, which forms one wall of a fluid-filled chamber that vibrates in response to sound waves hitting the ear drum. At the top surface of the hair cells where they contact the tectorial membrane are a series of rigid projections called stereocilia that, like microvilli, are constructed of a bundle of actin filaments closely wrapped by the cell plasma membrane. Unlike microvilli, which are all approximately the same length, the stereocilia within a bundle on an individual hair cell present a graded series of lengths resembling a staircase. Movements of the tectorial membrane cause the longest of these stereocilia to lean relative to the surface of the cell. Because the individual stereocilia are so stiff, the bundles do not bend. So instead, as the stereocilia lean over, the shorter ones slide relative to the longer ones. This leads to the opening of an ion channel that converts the mechanical signal to an electrical signal that can be perceived by the brain. The extreme stiffness of the stereocilia is a necessary element of this mechanical signal transduction event and could not be achieved using a single cytoskeletal filament.

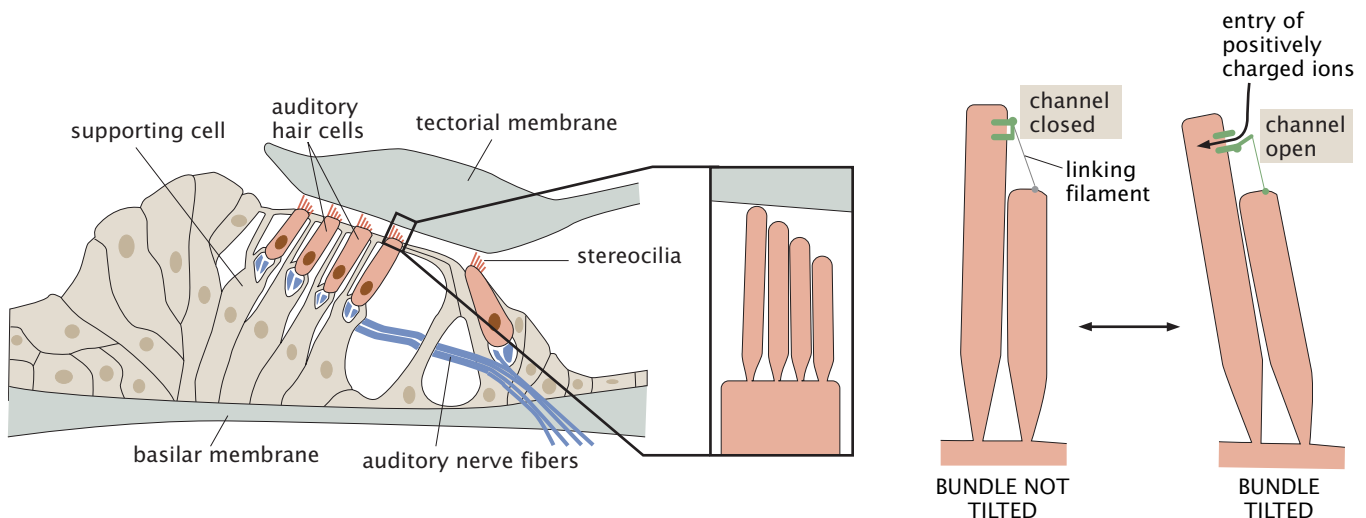


Figure 10.32: Beam mechanics in hearing. Sound is detected and converted into an electrical signal in the nervous system by hair cells. The upper surface of hair cells has a projecting bundle of stereocilia that make contact with a structure in the inner ear called the tectorial membrane. The tectorial membrane vibrates at specific points in response to sounds. Vibration of the membrane pushes on the stereocilia bundle. Neighboring stereocilia within the bundle are linked to one another by elastic filaments that connect the top of one stereocilium to the side of a taller one. When the bundle is tilted, stretching of the linking filaments leads to opening of ion channels. (Adapted from B. Alberts et al., *Molecular Biology of the Cell*, 5th ed. Garland Science, 2008.)

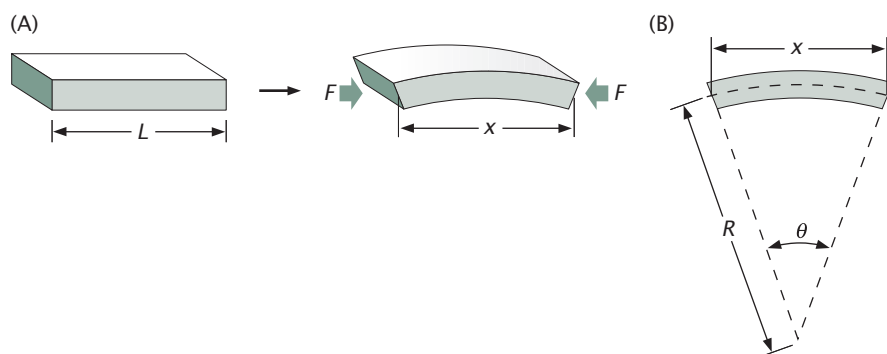


Figure 10.33: Schematic of the buckling process. (A) A beam of length L is loaded at the ends with compressive forces. At sufficiently large forces, the beam undergoes buckling. (B) Geometric parameters used to characterize the buckling process.

10.5.3 Cytoskeletal Buckling

A Beam Subject to a Large Enough Force Will Buckle

Our previous analysis of the deformation of beams has been restricted to those cases in which the material is subject to deformations that can be considered “small.” Of course, from an experimental perspective, such restrictions are entirely artificial and we are free to push and pull on beams with forces that engender deformations that exceed those characterized by the theory of elasticity. One of the ways in which materials can deform other than the simple deformations considered thus far is buckling. An example of the type of process we have in mind is shown in Figure 10.33. In the discussion to follow, we show how a simple estimate of the buckling force can be determined by examining the energetics required to deform the beam into either a semicircular arc or a sinusoidal profile.

Cytoskeletal filaments can exert forces during their polymerization process. One especially intriguing class of processes is that in which the leading edge of polymerizing actin filaments is responsible for deforming the membrane of a cell, such as in the filopodia of motile cells on surfaces. An intriguing question that arises in this context is how much force such filaments can take before they suffer a buckling deformation. Indeed, in a series of *in vitro* experiments, polymerizing microtubules were grown on a surface into a groove that would stall their progress and induce buckling. A schematic of such an experiment is shown in Figure 10.34. We can begin to interpret this experiment with simple ideas on elastic buckling.

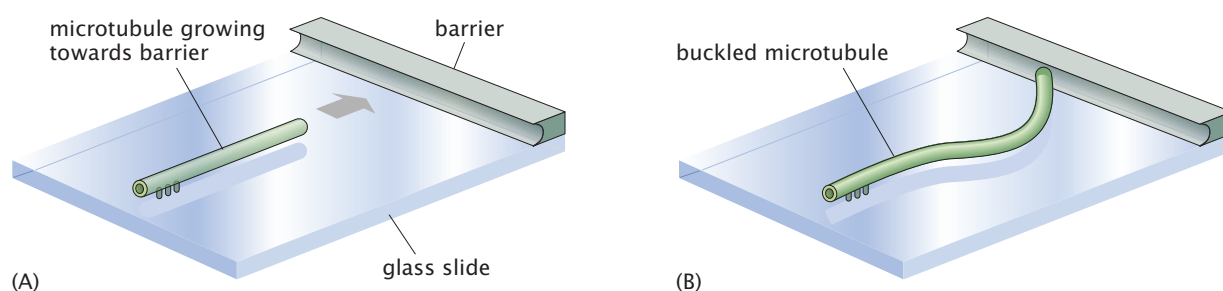


Figure 10.34: Microtubule buckling during polymerization. (A) At the start of the experiment, a microtubule is tightly bound to a glass slide, oriented such that its plus end is growing towards a barrier wall with a small overhang. With the addition of tubulin subunits, the microtubule grows freely until it contacts the wall. (B) Small thermal fluctuations in the position of the tip of the growing microtubule are thought to allow sufficient space for insertion of tubulin subunits occasionally, even when the microtubule tip is in contact with the wall. The force built up at this interface causes the microtubule to bend and then to buckle. (Adapted from M. Dogterom and B. Yurke, *Science* 278:856, 1997.)

10.5.4 Estimate of the Buckling Force

We consider a long, thin, elastic beam such as a microtubule with a force F applied at the two ends as shown in Figure 10.33. Experience tells us that if we increase the force applied to the beam, it will buckle at some point. Here we investigate this phenomenon in the context of beam theory. To provide an estimate of the critical load for buckling, we once again adopt the simple geometry of a circular arc already used earlier in the chapter. In particular, we assume a simple shape for the buckled beam, an arc of a circle of radius R and length L ; $\theta = L/R$ is the angle made by the arc. We are interested in determining the critical force for which $\theta \neq 0$ is the energy-minimizing value. That is, as the beam buckles, the energy associated with the applied force decreases while the bending energy increases. If the bending energy increase dominates, the beam will remain straight ($\theta = 0$); this will always be the case at sufficiently small forces. At very high forces, the opposite will be true and the beam will buckle. Making use of the circular geometry for the buckled beam, we can easily estimate the value of the critical force, F_{crit} .

For the present situation, we compute the total energy as a sum of the beam bending energy and the the energy associated with the applied force, resulting in

$$E_{\text{tot}} = \underbrace{\frac{\xi_p k_B T}{2} \frac{L}{R^2}}_{\text{elastic energy}} - \underbrace{F(L - x)}_{\text{loading device}}, \quad (10.66)$$

where

$$x = 2R \sin \frac{\theta}{2} \quad (10.67)$$

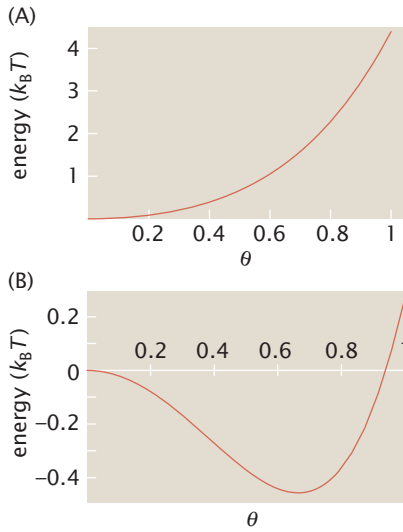


Figure 10.35: Energetics of beam buckling. Energy of a buckled beam as a function of the angle θ for (A) $F < F_{\text{crit}}$ and (B) $F > F_{\text{crit}}$. The parameters used in this plot are $L = 20 \mu\text{m}$ and $\xi_p k_B T = 30 \text{ pN } \mu\text{m}^2$, which correspond to the numbers appropriate to the experiment indicated schematically in Figure 10.34.

is the distance between the two ends of the beam. As elsewhere in the book, the loading device term captures the energy relaxation associated with the device applying the force F . In terms of the angle θ , the total energy of the buckled beam is

$$\frac{E_{\text{tot}}}{k_B T} = \frac{\xi_p}{L} \frac{\theta^2}{2} - \frac{FL}{k_B T} \left(1 - \frac{2}{\theta} \sin \frac{\theta}{2} \right). \quad (10.68)$$

Our strategy is to search for a buckling instability by seeing if at large enough forces the lowest-energy state corresponds to a nonzero value of θ . To see this, we plot the energy as a function of the angle θ for different values of the force. In Figure 10.35, we show two plots, one of the energy for $F < F_{\text{crit}}$ and the other when $F > F_{\text{crit}}$. In the first case, the minimum of the energy is at $\theta = 0$ and no buckling occurs, the beam prefers to stay straight. In the second case, the minimum of the energy is at $\theta \neq 0$, indicating that the beam prefers to buckle. The value of the preferred energy in the second case is negative, indicating that the energy associated with the applied force, which is negative, wins over the positive bending energy.

Beam Buckling Occurs at Smaller Forces for Longer Beams

We can take this result further and obtain an estimate for the critical value of the force. In particular, by expanding Equation 10.68 in small θ , which is legitimate if the applied force is close to its critical value,

we arrive at

$$\frac{E_{\text{tot}}}{k_B T} = \frac{\xi_p}{L} \frac{\theta^2}{2} - \frac{FL}{k_B T} \frac{\theta^2}{24}. \quad (10.69)$$

From this expression, we see that for $F < F_{\text{crit}}$, where

$$F_{\text{crit}} = 12 \frac{k_B T \xi_p}{L^2}, \quad (10.70)$$

the straight-beam solution with zero energy is the energy-minimizing solution. As soon as the force rises above F_{crit} this is no longer the case and the beam buckles. More to the point, for $F > F_{\text{crit}}$, the sign of the energy for small θ goes from positive to negative, indicating that the solution at $\theta = 0$ goes from being stable to unstable. Note that the critical force for buckling scales as L^{-2} , meaning that longer beams can be buckled more easily than shorter beams. For the experiment indicated schematically in Figure 10.34, the length of the microtubules is roughly $20 \mu\text{m}$ and they are characterized by a flexural rigidity $\xi_p k_B T = 30 \text{ pN } \mu\text{m}^2$, corresponding to a critical force of 0.9 pN .

The more sophisticated approach to this problem would be to relax the assumption of a circular shape for the buckled beam and allow any shape. Then the minimizing procedure would have to be performed over all shapes, which, mathematically speaking, corresponds to a variational problem. Such an approach would yield a sinusoidal shape as the energy minimizer and the expression for the critical force would differ from our Equation 10.70 only by a numerical factor, with π^2 replacing our 12.

10.6 Summary and Conclusions

Although DNA, microtubules, and large bones such as femurs are very different in nearly every way in their biological function, their size scale, and their material nature, nonetheless, many aspects of their mechanical behavior in biological systems can be quantitatively described using the same simple set of linear elasticity models. Indeed, this demonstrates the assertion that problems that are quite distinct in a biological sense can be seen as quite similar when viewed from the physical biology perspective. This chapter developed the relevant tools for using linear elasticity to model elastic beams and illustrated how these ideas can be used for a host of distinct and concrete problems. The consequences of elasticity can be found in such unexpected places as the efficiency of gene expression and the development of technology to create cantilever-based biosensors that can serve as artificial noses.

10.7 Appendix: The Mathematics of the Worm-Like Chain

In Section 10.2.3, we introduced the worm-like chain model as a way of uniting the chain statistics introduced in Chapter 8 and the elasticity of beam bending introduced in this chapter. However, evaluation of the force–extension properties of this model is considerably more complicated than the freely jointed chain model introduced in Section 8.3.2 (p. 340). We now examine the worm-like chain model in a more detailed fashion than presented in the chapter.

First, we examine the low-force limit defined by $f\xi_p \ll 1$; for dsDNA this becomes $F \ll k_B T/\xi_p = 0.08$ pN, a small force indeed. In this limit, the partition function can be expanded in powers of $f\xi_p$,

$$Z(f) = \int \mathcal{D}\mathbf{t}(s) \left\{ \exp\left(-\frac{\xi_p}{2} \int_0^L \left|\frac{d\mathbf{t}}{ds}\right|^2 ds\right) \left[1 + f \int_0^L t_z(s) ds + \frac{f^2}{2} \int_0^L t_z(s) ds \int_0^L t_z(u) du + O((f\xi_p)^3) \right] \right\}, \quad (10.71)$$

and we can safely retain only the first three terms in the expansion. The approximate expression that we get in this way can be rewritten as

$$Z(f) = Z(0) \left[1 + f \int_0^L \langle t_z(s) \rangle_0 ds + \frac{f^2}{2} \int_0^L \int_0^L ds du \langle t_z(s) t_z(u) \rangle_0 \right]. \quad (10.72)$$

Here $\langle \dots \rangle_0$ is the average evaluated using the zero-force partition function, Equation 10.23. The second term does not contribute, since the average extension in the zero-force case is zero. The third term in the expansion gives a non-vanishing contribution, which can be evaluated with the help of the tangent–tangent correlation function $\langle \mathbf{t}(s) \cdot \mathbf{t}(u) \rangle_0 = e^{-|s-u|/\xi_p}$. Recall that the tangent–tangent correlation function was defined in Section 10.2.2. Since $\mathbf{t}(s) \cdot \mathbf{t}(u) = t_x(s)t_x(u) + t_y(s)t_y(u) + t_z(s)t_z(u)$, and since the worm-like chain energy is invariant under rotations, $\langle t_z(s)t_z(u) \rangle_0 = \frac{1}{3} \langle \mathbf{t}(s) \cdot \mathbf{t}(u) \rangle_0$ follows. Substituting this into Equation 10.72, and using $\int_0^L \int_0^L ds du e^{-|s-u|/\xi_p} = 2L\xi_p$, which holds in the $L \gg \xi_p$ limit, gives

$$Z(f) = Z(0) \left(1 + \frac{f^2 L \xi_p}{3} \right). \quad (10.73)$$

Finally, making use of the relation given in Equation 10.25 we arrive at

$$\frac{\langle z \rangle}{L} = \frac{2f\xi_p}{3}, \quad (10.74)$$

which is the same as for the freely jointed chain as long as we take the length of the Kuhn segment to be twice the persistence length.

Next, we consider the limit of very large forces ($f\xi_p \gg 1$), which, as remarked earlier, for dsDNA corresponds to stretching forces in excess of 0.08 pN. In this limit, chain configurations that contribute appreciably to the partition function $Z(f)$ have their tangent vectors pointing roughly in the direction of the force. In other words, the components of the tangent vector in the x - and y -directions can be considered small, and

$$\mathbf{t} \approx (t_x, t_y, 1 - \frac{1}{2}(t_x^2 + t_y^2)). \quad (10.75)$$

This approximate expression for the tangent vector turns the formula for the energy into a quadratic form in t_x and t_y given by

$$E_{\text{tot}} = \frac{\xi_p k_B T}{2} \int_0^L ds \left[\left(\frac{dt_x}{ds} \right)^2 + \left(\frac{dt_y}{ds} \right)^2 \right] + \frac{fk_B T}{2} \int_0^L ds (t_x^2 + t_y^2) - fk_B TL. \quad (10.76)$$

For quadratic energies such as this, the path integral for the partition function turns into a Gaussian integral, which can be evaluated explicitly.

Using Equation 10.75, the average extension in the high-force limit can be written as

$$\langle z \rangle = L - \frac{1}{2} \int_0^L ds \langle t_x^2 + t_y^2 \rangle, \quad (10.77)$$

where $\langle \dots \rangle$ is the average with respect to $Z(f)$. To compute $\langle t_x^2 + t_y^2 \rangle$, we first express the energy functional in terms of the Fourier components of the tangent vector,

$$t_\alpha(s) = \sum_w e^{iws} t_\alpha(w) \quad (\alpha = x, y), \quad (10.78)$$

where the frequencies are defined by $w = 2\pi j/L$ with j an integer. In Fourier space, the energy takes on the form of the potential energy of a collection of harmonic oscillators, two for each value of the frequency w and given by

$$E_{\text{tot}} = \frac{Lk_B T}{2} \sum_w (\xi_p w^2 + f) (|t_x(w)|^2 + |t_y(w)|^2). \quad (10.79)$$

This observation allows us to compute the average $|t_\alpha(w)|^2$ without explicitly computing the path integral. Namely, we remind ourselves of the equipartition theorem of equilibrium statistical mechanics, which states that the average energy for every quadratic degree of freedom is $k_B T/2$. Therefore,

$$\left\langle \frac{Lk_B T}{2} (\xi_p w^2 + f) |t_\alpha(w)|^2 \right\rangle = \frac{k_B T}{2} \quad (\alpha = x, y), \quad (10.80)$$

and

$$\frac{\langle z \rangle}{L} = 1 - \frac{1}{L} \sum_w \frac{1}{\xi_p w^2 + f} \quad (10.81)$$

follows immediately from Equation 10.77 if we replace the integral in that equation by a sum over Fourier modes. Now the sum in Equation 10.81 can be evaluated by taking a continuum limit $\sum_w \rightarrow (L/2\pi) \int_{-\infty}^{+\infty} dw$; the remaining integral over w gives

$$\frac{\langle z \rangle}{L} = 1 - \frac{1}{2\sqrt{f\xi_p}}. \quad (10.82)$$

The results obtained here in the low-force and the high-force limit can be used to construct an approximate formula for the force-extension relationship at all forces. If we rewrite Equation 10.74 as

$$f\xi_p = \frac{3}{2} \frac{\langle z \rangle}{L} \quad (10.83)$$

and Equation 10.82 as

$$f\xi_p = \frac{1}{4(1 - \langle z \rangle/L)^2}, \quad (10.84)$$

then these two limiting forms can be combined to obtain Equation 10.26. This interpolation formula has the correct behavior at low

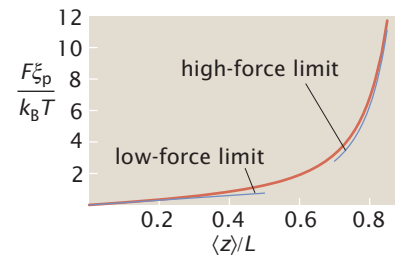


Figure 10.36: Force-extension curve for a worm-like chain. The red line is the interpolation formula, Equation 10.26, while the two blue lines are the limiting results obtained in the low- and high-force limits.

force and at high force, as shown in Figure 10.36, and it deviates at most by 10% from the exact result, which can be obtained numerically using a number of different schemes.

10.8 Problems

Key to the problem categories: • **Model refinements and derivations** • **Data interpretation** • **Model construction**

• 10.1 Persistence length and Fourier analysis

In the chapter, we computed the tangent–tangent correlation function for a polymer, which we modeled as an elastic beam undergoing thermal fluctuations. The calculation was carried out in the limit of small fluctuations and it led to an expression for the persistence length in terms of the flexural rigidity of the polymer. Here we reexamine this problem for a two-dimensional polymer, but without the assumption of small fluctuations.

(a) For a polymer confined to a plane, the tangent vector $\mathbf{t}(s)$ can be written in terms of the polar angle $\theta(s)$ as $\mathbf{t}(s) = (\cos \theta(s), \sin \theta(s))$. Rewrite the beam bending energy, Equation 10.9, in terms of the polar angle $\theta(s)$.

(b) Expand the polar angle $\theta(s)$ into a Fourier series, taking into account the boundary conditions $\theta(0) = 0$ and $d\theta/ds = 0$ for $s = L$. The first boundary condition comes about by choosing the orientation of the polymer so that the tangent vector at $s = 0$ is always along the x -axis. Convince yourself that the second is a consequence of there being no force acting on the end of the polymer.

(c) Rewrite the bending energy in terms of the Fourier amplitudes $\tilde{\theta}_n$, introduced in (b), and show that it takes on the form equivalent to that of many independent harmonic oscillators. Use equipartition to compute the thermal average of each of the Fourier amplitudes.

(d) Make use of the identity $\langle \cos X \rangle = e^{-X^2/2}$, which holds for a Gaussian distributed random variable X , to obtain the equation for the tangent–tangent correlation function, $\langle \mathbf{t}(s) \cdot \mathbf{t}(0) \rangle = e^{-\langle \theta(s)^2 \rangle / 2}$. Then compute $\langle \theta(s)^2 \rangle$ by using the Fourier series representation of $\theta(s)$ and the average values of the Fourier amplitudes $\tilde{\theta}_n$ obtained in (c). Convince yourself either by plotting or Fourier analysis that on the interval $0 < s < L$, $\langle \theta(s)^2 \rangle = s/\xi_p$.

(e) How does the persistence length in two dimensions compare with the value obtained in three dimensions? Explain why the tangent–tangent correlation function decays slower in two dimensions. What is the situation in one dimension?

• 10.2 Flexural rigidity of biopolymers

(a) Recall from p. 389 that when treating macromolecules as elastic beams it is the combination $K_{\text{eff}} = EI$, the flexural rigidity, that dictates the stiffness of that molecule. The flexural rigidity is a product of an energetic (E , Young modulus) and geometric (I , areal moment of inertia) factor. Reproduce the argument given in the chapter that culminated in Equation 10.8 for the bending energy of a beam and show that the flexural rigidity enters as claimed above.

(b) Using what you know about the geometry of DNA, actin filaments, and microtubules, determine the areal moment of inertia I for each of these molecules. Be careful and

remember that microtubules are hollow. Make sure that you comment on the various simplifications that you are making when you replace the macromolecule by some simple geometry.

(c) Given that the elastic modulus of actin is 2.3 GPa, take as your working hypothesis that E is universal for the macromolecules of interest here and has a value 2 GPa. In light of this choice of modulus, compute the stress needed to stretch both actin and DNA with a strain of 1%. Convert this result into a pulling force in piconewtons.

(d) Using the results from (b) and (c), compute the persistence lengths of all three of these molecules.

(e) Given that the measured persistence length of DNA is 50 nm, and using the areal moment of inertia you computed in (b), compute the Young modulus of DNA. How well does it agree with our 2 GPa rule of thumb from above?

• 10.3 Twisting DNA

For small torques exerted on a beam of elastic material, there is a linear relationship between the torque and the torsional strain, $\tau = Cd\theta/dz$. This is analogous to the bend elasticity worked out in the chapter. Single-molecule experiments using magnetic or optical tweezers allow measurements of the twist elasticity of DNA molecules by systematically winding up the tethered DNA molecule and examining the build up of torque as indicated schematically in Figure 10.37.

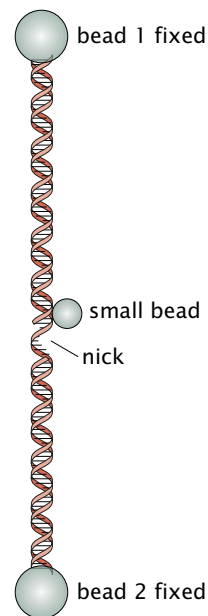


Figure 10.37: DNA twist elasticity experiment. A DNA molecule is constrained rotationally by two beads at both ends. A third bead is used to monitor the unwinding when a nick is introduced. (Adapted from Z. Bryant et al., *Nature* 424:338, 2003.)

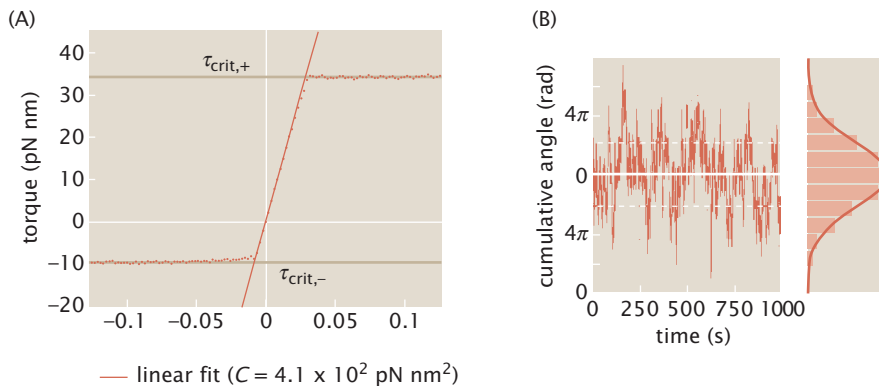


Figure 10.38: Experimental data from DNA twisting experiment. (A) Experimental results for the relation between the torque and the angular deflection of the DNA molecule. (B) Distribution of angular fluctuations in the bead position. (Adapted from Z. Bryant et al., *Nature* 424:338, 2003.)

(a) Use the data shown in Figure 10.38(A) to estimate the twist modulus.

(b) Estimate the torque needed to wind a 14.8 kb DNA molecule by 10 rotations and by 100 rotations.

(c) Write an expression for the energy stored in the twisted DNA molecule by virtue of its twist deformation. Using this expression and the data given in Figure 10.38(A), estimate the energy per base pair stored in the DNA molecule of 14.8 kb in length after the molecule has been subjected to 50 complete revolutions.

(d) An alternative way to measure the twist modulus (as with the bending modulus) is to probe the thermal fluctuations of the molecule. Figure 10.38(B) shows the angular excursion of the molecule (as reported by the motion of a bead stuck to the DNA in the middle of the molecule as shown in Figure 10.37) due to thermal fluctuations. Derive an expression linking the twist fluctuations to the stiffness and use the data to estimate the twist modulus. In particular, find the probability distribution $p(\theta)$ by computing the partition function. Show that $\langle (\Delta\theta)^2 \rangle = k_B TL/C$ and use this result and the data to find C .

• 10.4 The J-factor connection

Derive the relation between the J-factor and the free energy of cyclization described in Section 10.3.3. In particular, construct a lattice model of the cyclization process by imagining a box of Ω lattice sites each with volume v . We consider three species of DNA: monomers with sticky ends that are complementary to each other; dimers, which reflect two monomers sticking together; and DNA circles in which the two ends on the same molecule have stuck together; The number of molecules of each species is N_1 , N_2 , and N_c , respectively.

(a) Use the lattice model to write down the free energy of this assembly and attribute an energy ϵ_b to the binding of complementary ends, ϵ_{loop} to the looped configurations, and ϵ_{sol} as the energy associated with DNA-solvent interactions when the length of the DNA corresponds to one monomer. Further, assume that the solvent energy for a dimer is $2\epsilon_{\text{sol}}$ and for a looped configuration is identical to that of a monomer. Use a Lagrange multiplier μ to impose the constraint that $N_{\text{tot}} = N_1 + 2N_2 + N_c$.

(b) Minimize the free energy with respect to N_1 , N_2 , and N_c to find expressions for the concentrations of the three species. Then use the fact that J is the concentration of monomers at which $N_2 = N_c$ to solve for the unknown Lagrange multiplier μ and to obtain an expression for the J-factor in terms of the looping free energy.

• 10.5 Packing free energy for a spherical virus

Repeat the calculations of the energy of DNA packing for a spherical capsid. Contrast this result with that obtained in the chapter where it was assumed that the capsid is a cylinder. Use the experimental data provided on the book's website and compare the model with the data shown in Figure 10.19(B).

Relevant data for this problem is provided on the book's website.

• 10.6 Entropy of packing $\phi 29$

(a) Estimate the total work that the portal motor of $\phi 29$ needs to perform in order to overcome the entropy loss of the DNA after it has been packaged. (Hint: To count states, assume that the conformations of the DNA out of the capsid can be modeled by a freely jointed chain with 90° angles between the segments.)

(b) What is the total work of the motor when packaging the $\phi 29$ DNA, as measured by Smith et al. (2001)? What conclusion do you draw from comparing this value with the one obtained in (a)?

Relevant data for this problem is provided on the book's website.

• 10.7 Nucleosomes in a box

Repeat the derivation of nucleosome accessibility using the canonical distribution. This means that you should imagine a nucleosome in a box with L DNA-binding proteins and then work out the probability that the nucleosome will be occupied by one of those proteins as a function of their concentration and as a function of the position of the binding site on the DNA.

• 10.8 Nucleosome formation and assembly

(a) Repeat the derivation given in the chapter for the nucleosome formation energy, but now assuming that there is a discrete number of contacts ($N = 14$) between the DNA and the histone octamer.

(b) Use the discrete model to calculate the equilibrium accessibility of binding sites wrapped around nucleosomes. Apply this model to the data by Polach and Widom (1995) and fit the adhesive energy per contact γ_{discrete} .

(c) Reproduce Figure 10.25 and compare your results for the equilibrium accessibility versus burial depth from (a) and (b) with the continuum model.

(d) Look at some of the binding affinities of different DNA sequences to histones reported by Lowary and Widom

(1998). Once again, assume that the electrostatic interaction between the histone and the different DNA molecules does not vary, that is, it is not sequence-dependent. This is equivalent to saying that the difference between each sequence lies in its flexibility, in its bending energy. What would one expect the difference in their persistence lengths to be?

(e) Model the case of having two binding sites for the same DNA-binding protein on a DNA molecule that is wrapped

around a histone octamer. How does the equilibrium accessibility depend on the protein concentration and the relative position of the binding site? How does the problem change if the two binding sites correspond to two different DNA-binding proteins?

Relevant data for this problem is provided on the book's website.

10.9 Further Reading

Calladine, CR, Drew, HR, Luisi, BF, & Travers, AA (2004) *Understanding DNA: The Molecule and How It Works*, 3rd ed., Academic Press. This book gives an in-depth description of many topics we have only touched on in this chapter.

Gordon, JE (1978) *Structures or Why Things Don't Fall Down*, Penguin Books. A beautiful description of the relevance of beam theory to the world of macroscopic structures is to be found in Chapter 11, which is entitled "The advantage of being a beam."

Alberts, B, Johnson, A, Lewis, J, et al. (2008). *Molecular Biology of the Cell*, 5th ed., Garland Science. We will make reference to Alberts et al. repeatedly in this book for a compelling reason: their book is a treasure trove. Chapter 16 on the cytoskeleton is a useful starting point.

Amos, LA, & Amos, WB (1991) *Molecules of the Cytoskeleton*, The Guilford Press. An interesting source on the cytoskeleton.

Bray, D (2001) *Cell Movements: From Molecules to Motility*, Garland Science. Bray's unique book is full of great stuff ranging from the nature of the cytoskeleton to the nature of bacterial motion.

Boal, D (2012) *Mechanics of the Cell*, 2nd ed., Cambridge University Press. Boal's treatment of the elastic theory of beams and rods is excellent.

Marko, JF, & Siggia, E (1995) Stretching DNA, *Macromolecules* **28**, 209. A classic reference on force-extension in the worm-like chain model.

Riemer, SC, & Bloomfield, VA (1978) Packaging of DNA in bacteriophage heads: some considerations on energetics, *Biopolymers* **17**, 785. A beautiful paper describing the free-energy contribution to DNA packaging in bacteriophages.

Kindt, J, Tzliil, S, Ben-Shaul, A, & Gelbart, WM (2001) DNA packaging and ejection forces in bacteriophage, *Proc. Natl Acad. Sci. USA* **98**, 13671.

Fuller, DN, Rickgauer, JP, Jardine, PJ, et al. (2007) Ionic effects on viral DNA packaging and portal motor function in bacteriophage $\phi 29$, *Proc. Natl Acad. Sci. USA* **104**, 11245. This paper raises questions about the ability of the simple model presented in the chapter to account for measurements of the packaging forces during viral packaging.

Schiessel, H (2003) The physics of chromatin, *J. Phys.: Condens. Matter* **15**, R699. This excellent article describes the structure and energetics of nucleosomes.

Dogterom, M, & Yurke, B (1997) Measurement of the force-velocity relation for growing microtubules, *Science* **278**, 856. This paper illustrates how buckling of microtubules as shown in Figure 10.34 is used to measure the relation between force and velocity for growing microtubules.

10.10 References

Alberts, B, Johnson, A, Lewis, J, et al. (2008) *Molecular Biology of the Cell*, 5th ed., Garland Science.

Arai, Y, Yasuda, R, Akashi, K-I, et al. (1999) Tying a molecular knot with optical tweezers, *Nature* **399**, 446.

Bryant, Z, Stone, MD, Gore, J, et al. (2003) Structural transitions and elasticity from torque measurements on DNA, *Nature* **424**, 338.

Cerritelli, ME, Cheng, N, Rosenberg, AH, et al. (1997) Encapsidated conformation of bacteriophage T7 DNA, *Cell* **91**, 271.

Cloutier, TE, & Widom, J (2005) DNA twisting flexibility and the formation of sharply looped protein-DNA complexes, *Proc. Natl Acad. Sci. USA* **102**, 3645.

Du, Q, Smith, C, Shiffeldrim, N, et al. (2005) Cyclization of short DNA fragments and bending fluctuations of the double helix, *Proc. Natl Acad. Sci. USA* **102**, 5397.

Fawcett, DW (1966) *The Cell, Its Organelles and Inclusions: An Atlas of Fine Structure*, W. B. Saunders.

Johnson, JE, & Chiu, W (2007) DNA packaging and delivery machines in tailed bacteriophages, *Curr. Opin. Struct. Biol.* **17**, 237.

Lowary, PT, & Widom, J (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning, *J. Mol. Biol.* **276**, 19.

Nelson, P (2004) *Biological Physics: Energy, Information, Life*, W. H. Freeman.

Polach, KJ, & Widom, J (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation, *J. Mol. Biol.* **254**, 130.

Purohit, PK, Inamdar, MM, Grayson, PD, et al. (2005) Forces during bacteriophage DNA packaging and ejection, *Biophys. J.* **88**, 851.

Ramón y Cajal, S (1995) *Histology of the Nervous System* (translated by Neely Swanson and Larry W. Swanson), Oxford University Press.

Shore, D, & Baldwin, RL (1983) Energetics of DNA twisting. I. Relation between twist and cyclization probability, *J. Mol. Biol.* **170**, 957.

Smith, DE, Tans, SJ, Smith, SB, et al. (2001) The bacteriophage $\phi 29$ portal motor can package DNA against a large internal force, *Nature* **413**, 748.

Vologodskaya, M, & Vologodskii, A (2002) Contribution of the intrinsic curvature to measured DNA persistence length, *J. Mol. Biol.* **317**, 205.

Zeller, RW, Griffith, JD, Moore, JG, et al. (1995) A multimerizing transcription factor of sea urchin embryos capable of looping DNA, *Proc. Natl Acad. Sci. USA* **92**, 2989.