# BE/APh 161: Physical Biology of the Cell

Justin Bois

Caltech

Winter, 2019

This document was prepared at Caltech with financial support from the Donna and Benjamin M. Rosen Bioengineering Center.

# 1 What is Physical Biology of the Cell?

You are enrolled in a class entitled *Physical Biology of the Cell*. An obvious first question to ask is, What is *Physical Biology of the Cell*? Perhaps this is best answered by considering an example.

## 1.1 Nucleosome wrapping

In eukaryotes, DNA is packaged in the nucleus in chromosomes. The chromosomes consist of condensed chromatin fibers. The chromatin fibers are made of packed nucleosomes. A nucleosome consists of an octameric histone and DNA wrapped around it. The histone is about 8 nm in diameter, and the DNA wraps around it almost twice. Approximately 147 base pairs of DNA are wrapped around this complex.

These facts about nucleosomes raise important questions. How does wrapping happen? How stable are the wrapped structures? What are the dynamics of the wrapped DNA on the histone (i.e., how does it "breathe?")? What do we need to know to be able to answer these questions?

- What is the energetic costs of bending the DNA?
- What is the interaction energy between the DNA and the nucleosome core complex?
- What is the magnitude of the electrostatic repulsion of the DNA?
- What is the geometry of the contacts?
- What other factors (such as reader-writer and code-reader complexes) may be in play *in vivo*?

These are all physical questions and they demand physical approaches.

Let's talk about some of the approaches we could take. First, we could do a full all-atom molecular dynamics simulation. If we have a big enough computer, good enough force fields, and accurate enough structural information, we can just integrate equations of motion and get at the dynamics. This was done in Ettig, et al., *Biophys. J.*, **101**, 1999–2008, 2011. While impressive, the limits of computation mean that we can only simulate 10s of nanoseconds. Furthermore, electrostatics are hard to treat.

As an alternative approach, we could do a simulation in which the nucleotides and amino acids in the proteins are coarsened into beads that interact with each other via harmonic and Morse potentials. This was done in Voltz, et al., *Biophys. J.*, **102**, 849–858, 2012. They found some long-lived DNA detachments and discovered that the so-called H3 tail of the histone was an important player in these detachments.

### 1.1.1 A coarsened approach to DNA bending

While these computational approaches are useful and yield insight, we might want to zoom out some more to get a broader picture of nucleosome construction. Let's treat the DNA as a **semiflexible filament**. That is, it is an elastic rod that can bend, but resists doing so.[1] With this coarsened treatment, we can use some of the knowledge about elastic rods that structural engineers have known for a long

[1] We will talk about semiflexible filaments in much greater depth later in the course.

time. Specifically, the bending energy of the segment of DNA of length $\ell$ around the histone is given by

$$\frac{E_{\text{bend}}}{\ell} = \frac{1}{2} EIC^2. \tag{1.1}$$

Here, $E$ is the Young's modulus, $I$ is the moment of inertia, and $C$ is curvature. We will discuss these terms in depth later in the course, but for now we mention that a larger Young's modulus means that the rod is harder to bend, and the moment of inertia is a function of the cross-sectional geometry of the filament. The curvature is the inverse of the radius of curvature, $R$. We often define the **flexural rigidity** $K$ as $K = EI$. Then, we have

$$\frac{E_{\text{bend}}}{\ell} = \frac{1}{2} \frac{K}{R^2}. \tag{1.2}$$

What is the flexural rigidity of DNA? This is sometimes easier thought of in terms of a **persistence length**. The persistence length, $\xi_p$, can be defined as the length scale where the bending energy is of the same magnitude as the thermal energy $kT$.

$$E_{\text{bend}} \approx kT = \frac{K}{\xi_p}, \tag{1.3}$$

which gives $K = \xi_p kT$. So,

$$\frac{E_{\text{bend}}}{\ell} = \frac{1}{2} \frac{\xi_p kT}{R^2}. \tag{1.4}$$

So, we now have a way to estimate the bending energy. We only need to know the persistence length of DNA, which was measured by elegant single molecule experiments in the early 90s to be about 50 nm. The base stack height in DNA, known from crystal structures, is 0.34 nm. Since 127 base pairs are bent around the histone (147 base pairs go around the histone in total, but the ten base pairs on each end are straight), we have 43 nm of bent wrapped DNA. The radius of curvature is $R \approx 4$ nm, since the histone core is about 8 nm in diameter. Finally, the thermal energy $kT$ is about 4.1 pN-nm (piconewton-nanometers) at physiological temperatures. We have all the pieces we need.

$$E_{\text{bend}} \approx \frac{\ell}{2} \frac{\xi_p kT}{R^2} \approx \frac{43 \text{ nm}}{2} \frac{50 \text{ nm} \cdot 4.1 \text{ pN-nm}}{(4 \text{ nm})^2} \approx 275 \text{ pN-nm} \approx 65\, kT. \tag{1.5}$$

So, we already know that the binding energy between the histone and the DNA filament must be more than 65 $kT$, or 275 pN-nm.

### 1.1.2 The energetics of DNA-histone interactions

From crystal structures, we know that the DNA contacts the histone in the minor groove when it faces inwards. Thus, we have histone-DNA contact every helical twist, or about every 10 base pairs. There are then 14 total contacts. Assuming they are all about the same, we have 14 times the interaction energy of a single contact contributing to the DNA-histone binding energy.

Polach and Widom did an ingenious experiment to measure this binding energy. They wrapped sequences of DNA around a histone and then put the nucleosome complexes in a solution with restriction enzymes that cleave the DNA at a specific point if that point is not bound to the histone. They then looked at the probability of a segment of DNA being unwound, which is possible because

this probability is proportional the rate of cleavage. This, in turn, is proportional to $e^{-nE_{\text{net}}/kT}$, where $E_{\text{net}}$ is the free energy of binding of one site of the DNA to the histone and $n$ is the number of points that need to be unbound to unwind the DNA. Note that this is the *net* binding energy that includes all energetics, including the DNA bending energy. They found that the rate of cleavage for cut points in the middle of the wrapped DNA were about 4 or 5 orders of magnitude greater than for cut points at the ends. So, we can compute the free energy of binding a single site, knowing that a single site becomes unbound for cleavage at the end of the wrapped DNA and seven sites become unbound for a site at the center.

$$\frac{\mathrm{e}^{-E_{\text{net}}/kT}}{\mathrm{e}^{-7E_{\text{net}}/kT}} = 10^{-4} \text{ or } 10^{-5}. \tag{1.6}$$

This gives $E_{\text{net}} \approx 1.5\text{–}2\,kT \approx 6\text{–}8$ pN-nm. Multiplying this by the 14 sites gives the total binding energy as

$$E_{\text{tot}} \approx 28\,kT \approx 112 \text{ pN-nm}. \tag{1.7}$$

Therefore, the binding energy, exclusive of DNA bending is about $65\,kT + 28\,kT \approx 90\,kT$, or about $6\,kT$ per binding site. So, more than 2/3 of the binding energy is spent just bending the DNA around the histone. Furthermore, the total binding energy is only $6\,kT$ per site, so thermal fluctuations can result in significant "breathing" of the complex.

## 1.2 This is Physical Biology of the Cell

What we just did is physical biology of the cell. We took a biological problem and took a quantitative, physical approach to solving it. We build a simple, treatable model that is sufficient for our question (a rigid beam wrapped around a cylinder with point-contacts). The model allows use to measure parameters.

This approach allows us to:

- Define and measure parameters.
- Generate falsifiable predictions.
- Bridge concepts across length and time scales.
- Deal with complexity. We choose appropriate models for the desired level of detail.
- Illuminate life's constraints. Life, like everything, is bound by the laws of physics.
- Make synthetic biology possible. The ability to make predictions about how systems will behave, as physics can do, enables *engineering*.

## 1.3 What Physical Biology of the Cell is not

What will we *not* be doing in this class?

- Biology with some math. All too often, a biologist will do some great work and then throw in some math because it is sexy and will get their paper in a better journal. We are only interested in learning about biology if and when physical models are necessary.
- Biology-inspired physics. *Physical* questions often arise from biology. These are often very interesting to physicists and useful for exposing undiscovered physical principles, but do not explain biological phenomena. These are great studies, but not what we will do here. We want to learn about *biology*.

- Model-making for understood systems. This is kind of like "biology with some math." We do not need to invent models unless we are trying to gain a *deeper* understanding. Making a model that happens to match already measured and understood results is superfluous (and often part of research teams hunting for a more prestigious journal for their paper).