OPTICAL METHODS FOR MEASURING DNA FOLDING

ADAM D. SMITH*, OBINNA A. UKOGU*, LUKA M. DEVENICA*, ELIZABETH D. WHITE, and ASHLEY R. CARTER†

Physics Department, Amherst College, Merrill Science Center, Amherst, MA 01002, USA
†acarter@amherst.edu

Received Day Month Day
Revised Day Month Day

One of the most important biological processes is the dynamic folding and unfolding of deoxyribonucleic acid (DNA). The folding process is crucial for DNA to fit within the boundaries of the cell, while the unfolding process is essential for DNA replication and transcription. To accommodate both processes, the cell employs a highly active folding mechanism that has been the subject of intense study over the last few decades. Still, many open questions remain. What are the pathways for folding or unfolding? How does the folding equilibrium shift? And, what is the energy landscape for a particular process? Here, we review these emerging questions and the *in vitro*, optical methods that have provided answers, introducing the topic for those physicists seeking to step into biology. In addition, we discuss two iconic experiments for DNA folding, the tethered particle motion (TPM) experiment and the optical tweezers experiment.

Keywords: DNA folding, optical trap, optical tweezers, tethered particle motion, protamine, histone.

*Authors who contributed equally to this manuscript.

†Corresponding author.

1. Introduction

All living organisms are faced with the challenge of storing large amounts of deoxyribonucleic acid (DNA). This is a challenge for two reasons. The first challenge is that the DNA must fit within the physical boundaries defined by the organism. For example, humans store 2 m of DNA in a nucleus that is 6 μm in diameter,[1] a 300,000-fold difference. This is like compressing a string half the height of Mt. Everest into a pocket watch! The second challenge is that the DNA must also be locally unfolded so that it is accessible. Thus, the question is, if the DNA needs to be folded, how can it be unfolded?

The answer to this question is that the cell employs dynamic mechanisms to fold and unfold the DNA on command.[1-6] Specifically, some DNA regions are folded and remain inactive, while other DNA regions are unfolded and are available for binding of the cellular machinery for copying, reading, or repairing of the DNA.[7-9] Dynamic folding and unfolding of the DNA allows for switching between the active and inactive states. Thus, the physics behind how the DNA folds is an important question in biology, creating an interesting research niche for the biophysicist.

In this review, our goal is to provide a first look at research on DNA folding for physicists unfamiliar with biology. We will begin with a brief tutorial on DNA structure

and natural folding pathways, discussing some of the important research questions in the field (Section 2). Then, we will broadly examine the *in vitro*, optical methods that have been used to provide answers to these questions (Section 3). Finally, we will describe two widely-used optical methods, the tethered particle motion (TPM) assay (Section 4) and the optical tweezers assay (Section 5). At the beginning of each section of the review, we will point out resources for further investigation of this exciting interdisciplinary field.

2. Natural DNA Folding Pathways

DNA is stored inside of every animal, plant, and bacterium on the planet. The DNA in every organism has the same basic structure and is subject to the same physical laws. However, the way in which DNA folds within different organisms may vary. Here, we will describe this incredible variation and identify some of the unanswered research questions in the field.

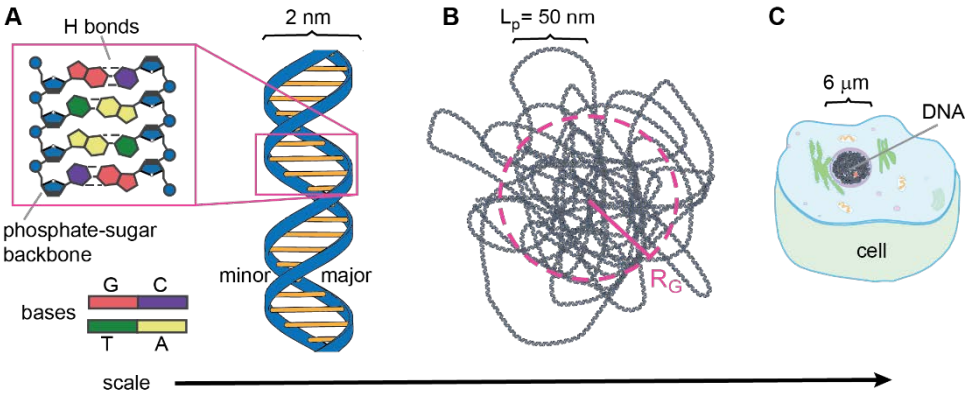

Fig. 1. DNA structure at different length scales. *A*) At the nanoscale, the structure of DNA consists of two strands, each composed of a series of nitrogenous bases (adenine, A; thymine, T; cytosine, C; and guanine, G) and a sugar-phosphate backbone. The two strands are bound together by hydrogen bonds and form a double helix with a rise of 0.34 nm per base pair and a width of 2 nm. The turns of the helix are uneven, creating a major and a minor groove. *B*) At the molecular scale, DNA collapses into a globule due to entropy. The length over which the DNA is straight and rod-like is about 50 nm and is referred to as the persistence length, L_p. Typically, the length of a DNA molecule—the contour length—is much longer than the persistence length. The approximate radius of the globule is given by the radius of gyration, R_G, which depends on the persistence length and contour length of the molecule. *C*) At the cellular scale, DNA is folded into the physical boundaries of the nucleus. In humans, the fully extended DNA molecule is 2 m in length and must be compacted into the cellular nucleus, which has a diameter of only 6 μm.

2.1. *DNA Structure Common to all Organisms*

The DNA within all organisms has a basic structure (Fig. 1A) that is perfectly suited to store the genetic code.[1, 10-12] At the nanoscale, DNA consists of two strands that are each made up of a series of nucleotides. Each nucleotide has a phosphate group, a five-carbon sugar (deoxyribose), and a nitrogeneous base: either adenine (A), cytosine (C), guanine (G), or thymine (T). These nucleotides are covalently bound together into a single DNA strand through bonds which link the phosphate of one nucleotide to the sugar of the next

nucleotide, creating a sequence of bases with a phosphate-sugar backbone. This sequence of bases (*e.g*. ACTCGT) forms the genetic code. The two strands of DNA are linked together through a series of hydrogen bonds between the bases: two hydrogen bonds link A with T, while three hydrogen bonds link C with G. Interestingly, this creates two copies of the genetic code: one copy on the "sense" strand containing the ACTCGT sequence and one copy on the "anti-sense" strand containing the complementary sequence, TGAGCA. Hydrogen bonds between the two strands allow for easy strand separation, which is required in order for the cell to copy, repair, or read the genetic code. When the two strands of DNA are bound together through hydrogen bonds, they wrap around each other forming a double helix, which looks like a twisted ladder. This double helix has a pitch of 0.34 nm per base pair (bp), a diameter of 2 nm, and two asymmetric grooves (the major and minor grooves). The major groove exposes the bases for access by the cellular machinery, while the minor groove exposes the phosphate-sugar backbone at the surface. Thus, the double helix structure is ideal for storing the genetic code, while also allowing the cell access to the code.

On the molecular scale, DNA does not remain as a long, rigid double helix, and instead, collapses into a globule due to entropy.[13-15] Typically, the effective radius of the globule is approximated by the radius of gyration, R_G, which is roughly the average distance between all of the subunits along the molecule to the center of mass.[13] To determine the radius of gyration, we can model the DNA as a long, one-dimensional polymer (Fig. 1B).[13-15] As a polymer, we assume that the DNA takes on a conformational state that resembles a three-dimensional random walk. In this case, the average displacement between the two DNA ends for many molecules, $\bar{R}$, would be zero, but the mean squared displacement, $\overline{R^2}$, would be equal to the number of steps, N, multiplied by the square of the step size, δ,

$$\overline{R^2} = N\delta^2. \quad (1)$$

To relate this mean squared displacement to the radius of gyration for the globule only requires a few more steps, namely determining the step size and number of steps for the polymer. The step size for a polymer will vary with the rigidity of the polymer. One way to characterize rigidity is by determining the persistence length, L_p, for a polymer. The persistence length is essentially the average length of a straight section along the polymer. More mathematically, the persistence length is the decay length for the slope of the tangent vector to the polymer. The slope of the tangent should remain the same if the polymer is straight, however if the polymer begins to bend, this slope will change. Specifically, the slope of the tangent will decay according to e^{-s/L_p}, where s is the distance along the polymer and the persistence length is the decay length.[16] In many polymers, the persistence length of the polymer is close to the size of a subunit if each subunit in the polymer is able to freely rotate. However, in the case of DNA, base pairs are not allowed to freely rotate since they are bound into a double helix. This sets the persistence length of DNA at about 50 nm or 150 bp,[16, 17] two orders of magnitude larger than the size of a subunit (0.34 nm or 1 bp). Changes in temperature, viscosity, or salt content of the surrounding medium can impact the persistence length of the DNA, leading to measurements of 35-130 nm.[18-24] Still, this variation is small compared to the overall length of the DNA, which can be thousands to millions of base pairs. Using the measured persistence length, we can

approximate the step size for a DNA molecule undergoing a random walk. Typically, this is estimated at twice the persistence length.[13] The number of steps for the polymer will be given by the length of the molecule along the polymer chain or the contour length, L_c, divided by the step size. Substituting these two values into Eq. 1 yields,

$$\overline{R^2} = \left(\frac{L_c}{2L_p}\right)\left(2L_p\right)^2 = 2L_cL_p. \quad (2)$$

If $L_c >> L_p$, as is the case for DNA, the MSD can be related to the radius of gyration of the molecule, R_G, by the equation,[13, 15]

$$\overline{R_G^2} \approx \frac{\overline{R^2}}{6} \approx \frac{L_cL_p}{3}. \quad (3)$$

Thus, we see that the radius of gyration and therefore the effective radius of the DNA globule is set by the persistence length and the contour length of the molecule.

Whether or not the DNA will require extra folding to fit within the physical boundaries set by the organism will then depend on the size of the boundaries and the length of the DNA. In humans, there are 46 DNA molecules (chromosomes) and some have >200 million base pairs.[25] A single DNA molecule with 200 million base pairs would have an R_G of approximately 30 μm, which is an order of magnitude larger than the average diameter of the human nucleus (6 μm).[1] Hence, human cells require some form of higher level DNA folding to fit all 46 DNA molecules into the nucleus (Fig. 1C). The goal in this review is look at the methods that answer questions like: how does this folding occur? What are the pathways? And, what are the thermodynamic and kinetic parameters for each pathway?

2.2. *Folding Forces Common to all Organisms*

To fold DNA or other polymers within an organism, there are several physical factors at play, including electrostatic forces, hydration forces, entropic interactions, hydrophobic interactions, pi-pi interactions, and Van der Waals forces.[15, 26-29]

Electrostatic forces are important for DNA folding due to the negatively charged phosphate backbone of DNA. As the DNA folds onto itself, the negative charges increasingly repel one another, preventing folding. The binding of positively charged, multivalent cations or proteins to the DNA neutralizes most of this negative charge, mitigating the repulsive force.[30] Interestingly, experiments show that DNA condensation occurs when about 90% of the charge on the DNA has been neutralized.[31] Thus, this binding of DNA to positively charged proteins or molecules is integral to DNA folding.

Hydration forces are forces that can be either repulsive or attractive,[32, 33] and may aid or inhibit DNA folding. Hydration forces occur when a hydrating surface, like DNA, interacts with water to form hydrogen bonds, polarizing the local water molecules.[33, 34] This perturbation then propagates through successive layers of water molecules, due to local water-water hydrogen bonds, creating an ordered hydration cloud. When two DNA regions approach each other, their hydration clouds overlap. If the DNA regions have the same charge, then their hydration clouds are polarized in opposite directions, and the hydrogen bond network is disrupted, creating a repulsive force. Conversely, if the DNA regions are oppositely charged (usually due to the presence of multivalent cations),[33, 35] then the

hydration clouds would be polarized in the same direction. This would reinforce the hydrogen bond network, creating an attractive force.

Entropic factors are also important. Entropy drives the collapse of the linear polymer molecule onto itself, setting the overall size of the DNA.[13-15] In addition, entropic factors due to molecular crowding can also aid folding.[36] The intracellular or nuclear environments of most organisms are replete with macromolecules, such as DNA or proteins. The presence of large amounts of these molecules produce a significant crowding effect that tends to favor more compact configurations which take up less volume in solution.[36-39] Thus, we see that in both molecular collapse and crowding, entropy favors folded conformations.

Finally, other factors will contribute as well. Hydrophobic interactions, for example, cause the DNA to kink or fold if a series of nucleotides are methylated.[40] Pi-pi orbital interactions (or stacking) between the aromatic rings of the DNA bases,[41] create an attraction that is important for stabilization of the double helix.[42] Additionally, Van Der Waals forces create attractions between molecules that must be programmed into folding simulations to get reasonable results.[43]

These physical interactions are available in all organisms to aid folding. However, different organisms will employ different physical methods in different situations. Before exploring the numerous ways natural DNA folding can occur, let's consider the biological requirements of the cell that affect DNA folding.

2.3. *Biological Requirements that Affect DNA Folding*

While the size constraints of the cell necessitate DNA folding, the biological requirements of the cell affect the dynamics of that folding. Specifically, biology requires that the genetic information stored in the DNA be available when needed. This information is used by the cell for one main purpose: to create proteins. In this section, we will introduce proteins and the biological requirements that affect DNA folding, summarizing much of the introductory biology found in textbooks.[44-46] This brief review should allow physicists without any biology training to follow the main points of the discussion.

Proteins are biological molecules that are integral to the life of an organism, carrying out many essential processes. Within an organism, there are four main groups of macromolecules: nucleic acids, carbohydrates, lipids, and proteins. Each one of these molecular groups has a role, but proteins take on the lion's share. Nucleic acids, like DNA or ribonucleic acid (RNA), either store the genetic code or interact with the genetic code. Though recent years have shown a larger functional role for RNA.[44] Carbohydrates or sugar molecules can be stored within the cell as an energy source and are typically used as components in other molecules, like DNA. Lipids or fats have hydrophobic and hydrophilic regions that allow them to coalesce within a cell, creating structural membranes or inert droplets that store energy. Yet, there are so many other tasks carried out by the cell. It is the role of proteins to perform these other tasks. Proteins act as enzymes to facilitate reactions; as scaffolds to provide the cell with structure; as signaling molecules to cue cellular changes; as motors to drive cellular growth, division, and movement; as antibodies

to provide immunity; as hormones to control metabolism; and as nutrients, transporters, and regulators. For this reason, proteins are essential for cellular health.

The incredible functional diversity of proteins is due to their structural variety. Proteins are polymers that are made by bonding amino acids together. Each amino acid contains an amine group ($-NH_2$) bonded to a central carbon atom which is then bonded to a carboxyl group (-COOH). The amine and carboxyl group on each amino acid allow for a series of amino acids to be linked together in a linear fashion. The central carbon atom in the amino acid is important because it is bound to what is known as a side-chain molecule. These side-chain molecules differ, creating 20 different amino acids with varying charges, shapes, hydrophobicities, and flexibilities. Variation of the amino acid sequence creates different structural motifs that provide different functions. For example, a protein with many positively charged amino acids, might bind tightly to negatively charged DNA. Repeating this amino acid sequence in multiple proteins would allow each to bind the DNA. Other sequences in the proteins would then create functional variety, allowing one protein to perhaps bind and fold DNA and another protein to bind and cut DNA. In this way, proteins are able to perform a stunning array of actions, creating complex cells and organisms. The cost of this diverse functionality is storing the code to make each protein.

DNA is the molecule responsible for storing the code to make all of the proteins for a particular organism. The code for a specific protein or set of proteins is called a gene. Within a gene, there are sequences of three base pairs that code for each of the 20 amino acids, providing a blueprint for the protein. This blueprint is constantly being copied (replicated), repaired, or read (transcribed), imposing several biological requirements on DNA folding. During DNA replication, the DNA must be completely unfolded at each location so that it can be copied for the next generation. During DNA repair, when the DNA is accessed to fix any breaks, mutations, or base pair mismatches, the DNA again must be unfolded. Finally, during DNA transcription, when the double stranded DNA is transcribed into RNA and later translated into a sequence of amino acids to make proteins, the DNA must also be unfolded. In fact, to carry out all of these processes requires a slew of proteins to bind the DNA, unwind the DNA, process the DNA, regulate each step, and supply energy for the necessary reactions. These processes cannot be completed if the DNA is folded, requiring unfolding of the particular region of interest.

It is worth noting that not all organisms have the same level of folding or the same need for processing the DNA, creating a diverse set of folding pathways across different organims. These pathways can be broadly classified by looking at folding in three organisms: viruses, bacteria, and eukaryotes (organisms that have cells with nuclei), as we will see in the next section.

2.4. *Variations of DNA Folding Pathways Between Organisms*

DNA folding mechanisms vary across organisms (Fig. 2), particularly between viruses, bacteria, and eukaryotes.[47]

A virus consists mainly of nucleic acids enveloped by a protein coat called a capsid (Fig. 2, "Virus").[1, 48] A virus does not make its own proteins, instead, it hijacks the machinery of the host cell.[49] Therefore, viral DNA does not need to be unfolded or remain

accessible while in the capsid, it just needs to be condensed enough so that virus is able to invade the host. There are several ways DNA is condensed into the capsid. In some bacteriophages (viruses that infect bacteria), DNA automatically fills the immature capsid by interacting with the positively charged proteins in the capsid and compacting through a series of random steps.[50, 51] Other methods include using a motor protein to pack the DNA inside the immature capsid.[52-54] Exactly how this packing occurs is a matter of study.

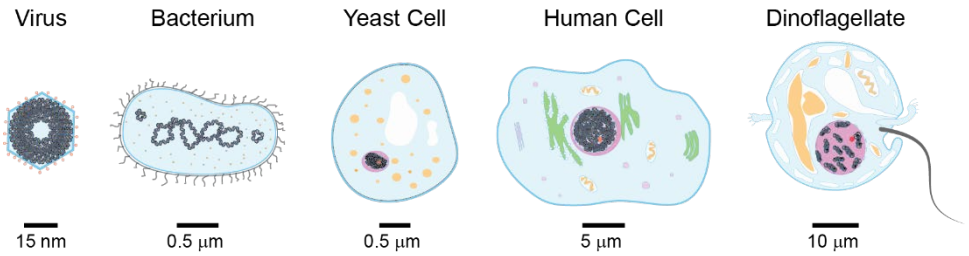

Fig. 2. Variation in DNA condensation between organisms. DNA (black) is condensed to almost crystalline packing levels in viruses, but remains relatively uncondensed in bacteria. In most eukaryotic cells, like yeast and humans, the DNA is condensed to fit into the nucleus (pink), but not to the packing levels seen in viruses. One notable exception is dinoflagellates, which have packing levels similar to viruses.

In addition, whether or not the viral genome has a generalized folded structure across different viruses has been a matter of contention.[55] Over the last 50 years several models have been proposed for the structure of packaged DNA within viruses – the coaxial spool model,[56] the folded toroid model,[57] and the liquid crystal model.[58] So far, the evidence seems to eliminate the liquid crystal model, and at least partially discredit the folded toroid model on energetic grounds.[51] Thus, it may be the case that there is no unique structure for compacted DNA in viruses.[55] However, in general, it can be said that the DNA is packaged within the capsid into a coaxial globule with hexagonally packed outer concentric layers and a highly disordered central core.[50, 57]

DNA condensation in bacteria is quite distinct from condensation in viruses. Bacterial DNA is located in the cytoplasm and is not condensed into a capsid or a nucleus (Fig. 2, "Bacterium"). Rather, the genome is localized into a well-defined area of the cell, called the nucleoid, where it remains fairly uncondensed in a series of supercoils.[47, 59, 60] The folding requirements for DNA in the nucleoid are lax,[61] creating almost the opposite biological conditions from a virus. In a virus, physical space is limited and DNA processing is negligible, while in bacteria, space is abundant and DNA transcription occurs almost uninhibited.[61] Interestingly, bacteria do possess nucleoid associated proteins (NAPs) that are known to regulate transcription[62] and facilitate DNA condensation.[63, 64] However, evidence suggests that the most important condensing force in bacteria is macromolecular crowding.[65] Comparisons of extracted nucleoids from wild-type bacterial cells and mutant bacterial cells with deleted NAP binding sites indicate that the absence of NAPs has no significant structural effect on nucleoid structure.[66]

DNA condensation in eukaryotes, like yeast and humans, is distinct from both viruses and bacteria (Fig. 2, "Yeast Cell" and "Human Cell").[1-6] In eukaryotes, DNA must be condensed into a nucleus, but also must remain accessible for processing.[67] This means that

the DNA needs to be folded, like in viruses, but also unfolded, like in bacteria. To solve this problem, eukaryotic cells locally fold and unfold their DNA in a dynamic fashion using a complex system of proteins. This allows precise unfolding of a particular DNA region while keeping other regions folded.

However, not all DNA folding in eukaryotic organisms is the same, as is the case for dinoflagellates (Fig. 2). Dinoflagellates are marine plankton that are often bioluminescent and give the waters off of Cape Cod in the United States and Mosquito Bay in Puerto Rico a sparkling glow. Dinoflagellates are an interesting case because they have the largest known genomes among living organisms,[68] but contain an extremely low ratio of DNA-condensing proteins (mass ratio of 1:10 proteins to DNA as compared to 1:1 for other eukaryotes).[68-70] Despite this low concentration of DNA-condensing proteins, dinoflagellates package their DNA into liquid crystalline chromosomes (LCCs) that are thought to be some of the most condensed DNA structures that have been observed in nature.[68] Exactly how this DNA condensation occurs is still largely unknown.

Figuring out how DNA folding occurs in viruses, bacteria, and eukaryotes is an ongoing research question. What forces or proteins are involved in each case? How condensed is the folding? Is the folding dynamic? And, what conditions change the folding state? Interestingly, these questions about folding not only apply across organisms, but also within an organism as some organisms will have multiple folding pathways.

2.5. *Variations of DNA Folding Pathways within an Organism*

While we have seen that DNA folding varies across organisms, DNA folding also varies within organisms. For multicellular organisms, DNA folding can vary across cell type and with cell cycle. Here, we will focus on the differences in DNA condensation pathways within multicellular organisms, particularly on the difference between somatic cells (cells within the body of the organism) and sperm cells (Fig. 3).

In human somatic cells, DNA is condensed into chromatin through the action of positively-charged histone proteins (Fig. 3, *top*).[44, 71-74] In the first step of chromatin condensation, the DNA wraps around an octamer of histones (two each of H2A, H2B, H3, and H4) to form a nucleosome core particle. A series of nucleosome core particles is separated by regions of unbound DNA (20-90 bp long), called linker DNA. The nucleosome core particle and the linker DNA form a nucleosome, and a succession of nucleosomes form a structure referred to as "beads on a string". In this way, DNA is condensed by wrapping the linear DNA molecule with a 2 nm width around a nucleosome core particle with an 11 nm width, trading length for width. Interestingly, much is known about this first step of chromatin condensation, including the fact that the DNA wraps around the histone octamer in a two-step process.[75] Still, many questions remain. Does DNA wrapping around histones change in the presence of chromatin remodeling proteins or with DNA sequence? What conditions shift the equilibrium or kinetics of the process? Future DNA folding experiments will be needed to solve these questions.

In the second step of chromatin condensation, the H1 histone binds to the nucleosome, causing the nucleosomes to wrap around one another into a 30-nm-diameter fiber.[44, 72, 73, 76, 77] However, the exact structure of this 30-nm fiber is unknown. Research indicates that

there are at least two distinct 30-nm fibers that can exist simultaneously in a nucleus: the "solenoid" and the "zig-zag".[76] The solenoid fiber contains helical twists of nucleosomes, while the zig-zag fiber contains a series of nucleosomes folded back and forth onto each other. The cellular conditions that give rise to each type of structure, as well as the possible existence of other structures besides the solenoid and zig-zag, have yet to be confirmed experimentally.

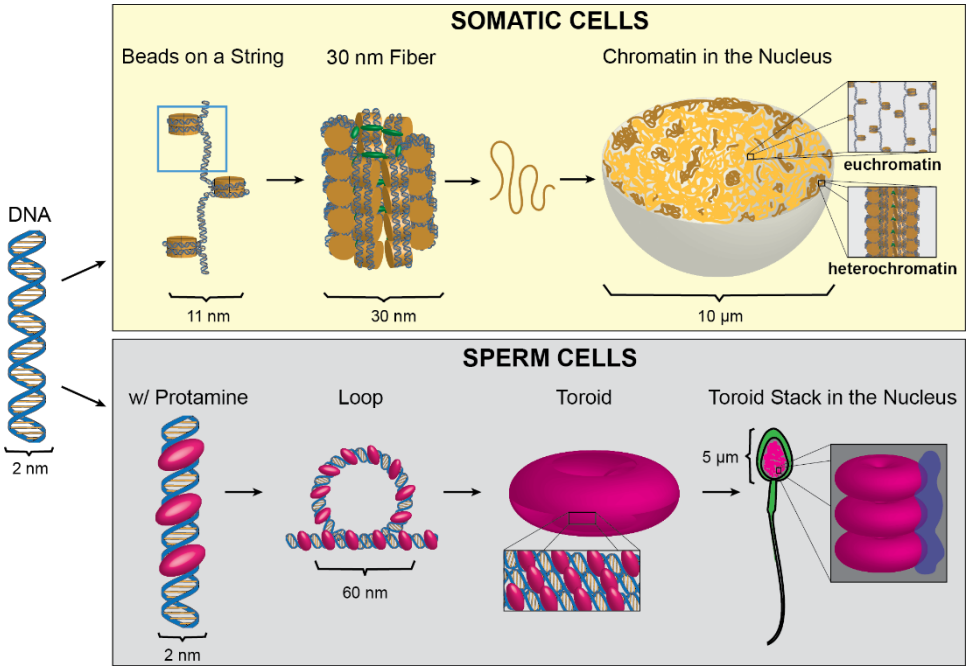

Fig. 3. Variation in DNA condensation within a multicellular organism. In humans, the somatic cells—the cells within the body—condense DNA into chromatin using histone proteins (*top*). At the lowest level of condensation, DNA (*blue*) wraps into nucleosomes (*blue box*) which consist of unbound linker DNA and DNA bound to an octamer of histone proteins (*brown cylinder*). A series of nucleosomes is referred to as beads on a string. At the next level of condensation, another histone protein (*green*) binds the DNA causing the nucleosomes to coil into a fiber with a diameter of 30 nm. Further condensation of the 30 nm fiber can occur by folding the fiber (*brown wavy line*). However, during most of the cell cycle, chromatin within the nucleus remains as either beads on a string (typically in euchromatin regions, *light brown*) or as the 30 nm fiber (typically in heterochomatin regions, *dark brown*). In sperm cells (*bottom*), DNA is folded by protamine proteins. Protamines (*pink*) bind to the major groove of the DNA and coordinate the looping of DNA into a toroid. Multiple toroids are held together through nuclear scaffolding proteins (*purple*) forming a stack within the nucleus. DNA remains condensed for the life of the sperm.

During cell division, the chromatin is further condensed into pairs of chromosomes, the familiar microscopic structures that look like an "X".[44, 72, 73, 78] Compaction into chromosomes allows the safeguarding of DNA during cellular division and facilitates distribution of the two copies of DNA to the resultant cells.[44] At this level of chromatin condensation, the exact structures, global organization, and folding mechanisms are still the subject of ongoing research.[9, 79]

During the rest of the cell cycle, the chromatin is organized into the nucleus as regions of heterochromatin (tightly compacted DNA) and euchromatin (loosely compacted DNA).[44, 80, 81] Heterochromatin is predominantly located near the exterior of the nucleus and contains DNA that is mostly folded into the 30 nm fiber.[78] In contrast, euchromatin is concentrated in the interior of the nucleus and contains DNA that is folded into a beads on a string structure.[82] The two levels of folding (30-nm fiber and beads on a string) in the two DNA regions lead to a striking contrast in gene transcription, with genes within euchromatin regions being more actively transcribed.[82] Interestingly, cells with similar functions will have similar regions of active DNA and are able to pass on to future generations not only copies of their DNA, but also the pattern that the DNA is folded into.[83] This folding pattern is set by molecular tags (e.g. acetyl, methyl, or phosphate groups) attached to the histone proteins. Exactly how these molecular tags lead to a particular folding pattern or how the tags get passed along through successive generations is the subject of ongoing research.[83, 84] Interestingly, inherited information that is not genetic, like these molecular tags, is termed epigenetic.[44, 72, 81] Epigenetic factors can be influenced by the environment, and are one of the reasons why twins or cloned animals can be different in appearance.[44] Understanding DNA folding then leads to understanding how these epigenetic factors are passed down to the next generation.

In contrast to DNA folding in somatic cells, DNA folding in sperm cells (Fig. 3, *bottom*) is much less dynamic with most of the DNA existing in a highly compact form. Tighter compaction of the DNA in the sperm head ensures that the cell can swim efficiently to the egg[85] and protects the DNA from the higher amounts of UV radiation outside of the body.[86, 87] In addition, gene transcription in sperm is almost nonexistent,[88] eliminating requirements for cellular access. Due to these incredibly different requirements, the DNA within sperm cells is folded by a different mechanism than somatic cells.[89]

The first step of DNA condensation in sperm involves the formation of DNA-protamine loops. Protamine is a positively charged protein of approximately 50 amino acids that is able to bind the major groove of the DNA, coating the double helix.[85, 90] The coated DNA then forms loops of approximately 50 nm in diameter[91] with loop size varying between organisms and with salt concentration.[92] Protamine-induced DNA looping has great potential for further research, since so many questions remain unanswered. For example, what is the mechanism for loop formation? Is the mechanism dependent on DNA sequence or protamine structure? Is loop formation reversible? And, is the first loop that forms different from subsequent loops? Answering these questions would give insight into the first step in the DNA folding pathway in sperm.

The second step of DNA condensation in sperm is the wrapping of protamine-DNA loops into a toroid. To form a toroid, loops bind to one another in an ordered hexagonal fashion.[57, 92-94] Similar to loops, research suggests that toroid dimensions and structural integrity are a function of the surrounding salt concentration[92] and the amount of positive charge of the constituent protamine.[95] Ongoing research hopes to address how loops assemble into toroids, how the size of the toroid is determined, and if there are any structural or organizational differences between toroids made up of one DNA molecule or several molecules.

Further DNA condensation occurs as toroids stack on top of one another in the sperm head and are held together by scaffolding proteins.[86] These scaffolding proteins ensure the positional stability of the toroid stack.[96] However, the complete function of the scaffolding protein, including its mechanical properties, is the subject of ongoing scientific inquiry.[97]

From this brief review, we see that there are several different pathways for natural DNA folding within an organism or between organisms. There are also other natural folding pathways, such as folding into supercoils, hairpins, or junctions,[98, 99] that are beyond the scope of this review. However, many of the open questions for these pathways are similar. What are the DNA folding states for a particular pathway? What are the kinetic or thermodynamic parameters of these states? And, how does changing the DNA sequence, the folding protein, or the kinetic or thermodynamic parameters of the states affect the folding pathway? To answer these questions, researchers have employed optical methods with much success.

3. Overview of Measuring DNA Folding with Optical Methods

One method for measuring natural DNA folding is to use an *in vitro* assay or experimental setup. In the *in vitro* assay, the biological folding pathway is isolated, allowing for measurements of each constituent part. If the *in vitro* assay is also a single molecule assay where the folding pathway of an individual molecule is tracked, then variations in different folding states for different molecules can be easily visualized. The problem with a single molecule, *in vitro* assay is directly tracking the folding of an individual, dynamic DNA molecule at the nanoscale. The solution to this problem has been to image single molecules of DNA indirectly using reporters (*i.e.* beads, fluorescent molecules) and optical methods. Here we will review these optical methods, describing the reporters commonly used (Section 3.1), the methods to image them (Section 3.2), the ways to apply force to unfold or fold the DNA molecule (Section 3.3), and the resulting measurements (Section 3.4).

3.1. *Reporters of DNA Folding*

Since dynamic DNA folding cannot be directly visualized with optical methods due to the size of the molecule, these optical methods rely on having a reporter – a particle or dye that can be detected. Reporters must be fluorescent or scatter light and can range in size from nanoscale organic dyes and fluorescent proteins to micron-sized beads.[100-103] The amount of light produced by the fluorophore or scattered by the bead affects the signal to noise ratio of the measurement, with larger and brighter reporters producing a greater signal.[104] However, larger diameter reporters will also have larger drag forces that may affect the dynamics of the biological system.[104] In addition, attachment of reporters to the DNA requires biochemical linkage,[101, 102] which may be easier in some cases than others. Due to these limitations, there are many different types of reporters to choose from.

There are three common fluorescent reporters: organic dyes, quantum dots, and fluorescent proteins (Fig. 4, *top*).[105] Organic dyes are ~1-nm-diameter, fluorescent molecules that either intercalate between the DNA base pairs, or bind tightly to one of the DNA grooves.[106] Binding is often nonspecific along the DNA, allowing for complete coverage and visualization of the entire DNA molecule. The downside to these dyes is that

they have limited photostability (~1 s),[107] and their binding may affect protein interactions.[108] An alternative to organic dyes are quantum dots.[108, 109] Quantum dots are semiconductor nanocrystals that are brighter and more photostable than organic dyes, but are also larger (10-30 nm diameter) and need to be bound to the DNA through a biochemical linkage (*e.g.* antibody-antigen interaction).[108] Still, another choice is a fluorescent protein,[101, 107] like green fluorescent protein (GFP). Fluorescent proteins are smaller than quantum dots with a diameter of ~2 nm, but are dimmer than quantum dots and have a lower photostability (~100 ms)[107] than organic dyes. In addition, fluorescent proteins, like quantum dots, require biochemical linkage to the DNA. The advantage of fluorescent proteins is that it is easy to link them to other proteins.[101, 107, 110] Thus, researchers might choose to link the protein that folds the DNA with a fluorescent protein to monitor the folding protein as it binds the DNA; another reporter would still be needed to visualize the DNA folding.

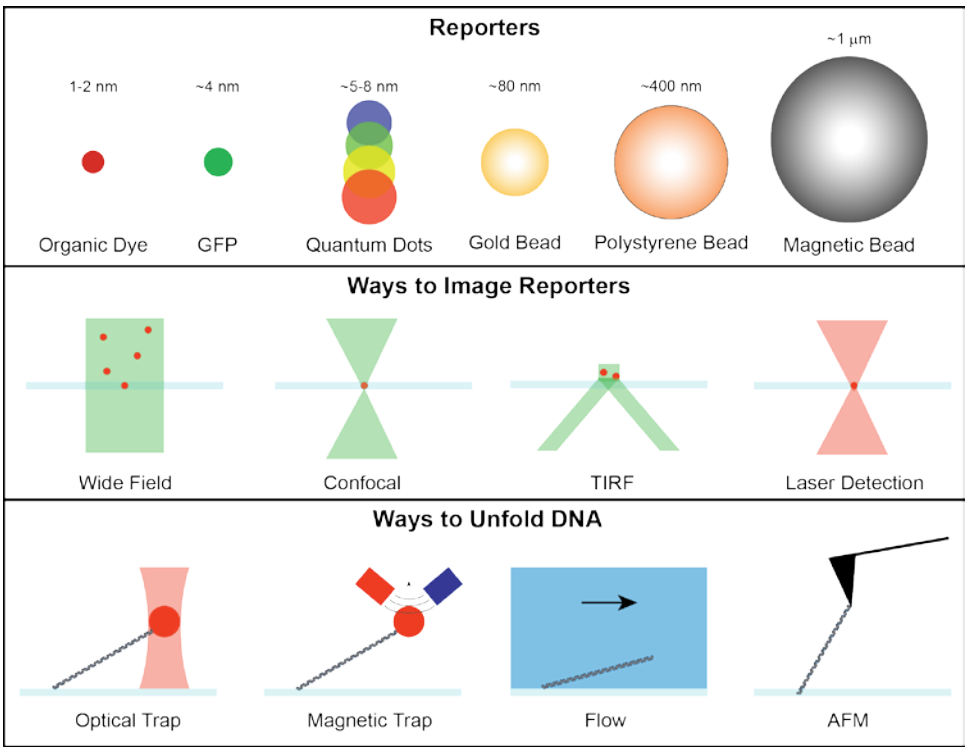

Fig. 4. Optical methods to measure DNA folding and unfolding. DNA is not visible optically and requires a reporter to visualize (*top*). Reporters can be fluorescent like organic dyes, fluorescent proteins (*e.g.* green fluorescent protein, GFP), or quantum dots. Larger diameter (50 nm – 5 μm), non-fluorescent beads can also be used as reporters if they scatter enough light. If the reporter is biochemically linked to the DNA, it allows for optical visualization. Reporters can be imaged through a variety of microscope techniques (*middle*) including bright field microscopy, confocal microscopy, total internal reflection fluorescence (TIRF) microscopy, or direct laser detection. Once the reporter is imaged, force can be used to unfold the DNA (*bottom*). Force can be applied using an optical trap, a magnetic trap, a flow, or an atomic force microscope (AFM).

There are also three common non-fluorescent reporters: metallic beads, dielectric beads, and magnetic beads (Fig. 4, *top*). Since non-fluorescent reporters must scatter light, they are typically larger (40 nm – 5 μm in diameter) than their fluorescent counterparts.[102] Increasing the size of the reporter increases the drag force on the DNA (see Section 3.3) and limits the data rate for the measurement, since larger beads require more time to average out their Brownian motion.[111] However, beads have the advantage that they can also be used in optical tweezers[103, 112, 113] or magnetic tweezers[114-116] assays to apply force (see Section 3.3).

Between the three types of beads there are some differences. Metallic beads are typically smaller (20-80 nm in diameter)[102] and scatter more light for their size than dielectric beads, but may cause significant heating.[117, 118] Metallic beads are often used in dark field imaging applications[19] and in situations where large drag forces are problematic.[119, 120] Polystyrene beads or other dielectric beads (*e.g.* silica) can be purchased with an array of biochemical labels already attached to the beads, including fluorescent proteins, antibodies, amine groups, or carboxyl groups. These biochemical labels facilitate attachment to the DNA. In addition, polystyrene beads can be made in an array of sizes with coefficients of variation <10%, which reduces errors on measurements of DNA length.[121] Magnetic beads are often made by coating polystyrene particles in ferromagnetic or paramagnetic layers, and are used in magnetic tweezers assays.[116]

3.2. *Methods to Image Reporters*

To image a reporter requires an optical method, typically microscopy (Fig. 4, *middle*). The choice of optical method depends on the type of reporter, the biological assay under study, and the signal-to-noise requirements for the measurement. [102, 120, 122-124]

For fluorescent reporters, a fluorescent microscope that has a lamp or laser to excite the fluorescent transition of interest and optics to collect the emission of the fluorophore is necessary, though the type of microscope can vary. The simplest system is a wide field microscope[125] where the entire sample is illuminated with a collimated beam of light, exciting fluorescent reporters that are in the sample plane and ones that are not (Fig. 4, *middle*, "Wide Field"). When fluorescent reporters outside of the sample plane or other structures fluoresce, this produces unwanted emission that increases the background for a measurement.[126] To eliminate this background emission and improve the signal-to-noise ratio, researchers commonly use either confocal microscopy or total internal reflection fluorescence (TIRF) microscopy. In confocal microscopy,[126] a pinhole is placed in a confocal plane to the sample plane, limiting the excitation light to a small diffraction-limited spot (~250 nm) at the sample (Fig. 4, *middle*, "Confocal"). This only excites the fluorescent reporter of interest and removes most of the background. However, the downside of this technique is that taking a two-dimensional image requires the time-consuming step of scanning the pinhole or sample. On the other hand, if only a single fluorescent reporter is to be imaged, then scanning is not necessary and a single photodiode could collect the emitted light from the one location on the sample, allowing for GHz data rates.[127] If a larger field of view is needed, then TIRF microscopy might be the right choice. In TIRF microscopy,[128-130] a beam of light is reflected off the sample at a high angle of

incidence, causing total internal reflection (Fig. 4, *middle*, "TIRF"). In total internal reflection, there is no transmission through the sample. However, the boundary conditions for Maxwell's equations cause an evanescent field at the surface of the sample, which can excite fluorophores up to ~300 nm in depth.[128] This reduces the background signal present in wide field microscopy,[131] and eliminates the need for scanning associated with confocal microscopy, allowing for multiple molecules to be imaged at once. The downside for TIRF is that it requires the biological assay be bound to the surface of the sample. Wide field and confocal microscopy can image at the surface or in solution. Thus, there are advantages and limitations of each imaging method; the appropriate method will depend on the experiment of interest.

For non-fluorescent reporters like beads, the imaging system must collect the light scattered by the bead. One method is to use a microscope (*e.g.* a bright field microscope[125]) that illuminates the entire field of view with a collimated beam and collects scattered light of the same wavelength. Since the beads are typically larger than the wavelength of the illumination beam, they can be directly resolved, but experiments on DNA folding require the beads only to be localized. Localization of the bead using a camera is typically on the order of 1-10 nm[102] with data rates of 10-40 Hz for the entire field of view.[132, 133] To improve data rate and localization precision, beads can be imaged using laser detection. One form of laser detection (Fig. 4, *middle*, "Laser Detection"), is back-focal-plane interferometry, [89, 113, 134-137] which requires illuminating the bead with a laser focused to a diffraction-limited spot and either collecting the forward scattered or back scattered signal. Back-focal-plane interferometry has a theoretical precision that is less than a picometer,[138] though noise typically limits the localization of beads to 100 pm in three dimensions.[134, 135] Data rates are also higher than video microscopy and can be up to several hundred kHz.[135] Another imaging method besides bright field microscopy and laser detection is dark field microscopy. [19, 125, 139, 140] In dark field microscopy, the illumination light hits the sample at a high incidence angle and only light scattered by the bead is collected, improving the signal to noise of the system.[141] This is advantageous for *in vivo* assays,[142] where scattering from the cell or organism can distort the signal.

The choice of imaging method may also depend on the type of assay. Assays that track protein binding and folding may have more than one reporter,[115] requiring imaging at multiple wavelengths. Assays that look for intermediate states in folding transitions may require imaging methods with a high data rate or precision.[143] Thus, finding the right imaging method will depend on the assay, the reporter, and the precision of the measurement.

3.3. *Methods to Apply Force*

In many DNA folding assays, force is required to physically unfold the DNA. There are four common ways to apply force: fluid flow, optical tweezers, magnetic tweezers, or atomic force microscopy (AFM) (Fig. 4, *bottom*).[103, 144] Use of a particular method may limit the imaging system or the type of reporter used in the assay.

The most straightforward way to apply force is by fluid flow.[19, 21, 131, 145] When there is a relative velocity difference between the DNA molecule and the surrounding fluid, the

DNA experiences a drag force in the direction of the flow, stretching the DNA. Fluid flow can be be created by a piezo stage that moves the sample relative to the fixed fluid[137] or by using a hydraulic device that flows the fluid through a fixed sample.[146] To estimate the drag force on the DNA due to the fluid flow, we first assume that the flow is laminar. Laminar flow will occur when the Reynold's number, *Re*, is less than 2000.[147] The Reynold's number depends on the density of the fluid ρ, the relative velocity of the fluid flow *v*, the effective radius of the DNA molecule *r*, and the viscosity of the medium η,

$$Re = \frac{\rho v r}{\eta}. \quad (4)$$

For a DNA molecule with an effective radius of 0.3 μm (contour length of 5 μm or 16 kb according to Eq. 3) in a fluid flow of 1000 μm/s, *Re* is quite small at 0.0003. Since the flow at this low Reynold's number is laminar, we can use the Stokes equation to find the drag force. In the Stokes equation, the magnitude of the drag force, *F*, is proportional to the relative speed of the fluid flow, *v*, and the drag coefficient, γ. Since the drag coefficient is a function of the viscosity of the medium and the effective radius of the object in the fluid, the equation for the force is,

$$F = \gamma v = 6\pi\eta r v. \quad (5)$$

For fluid flows of 1-2000 μm/s and an effective DNA radius of 0.3 μm, the drag force is 0.005-10 pN. This is enough force to overcome entropy and stretch out the DNA molecule (~90% extended at 1 pN)[16] or to unwrap DNA from a single nucleosome core particle (first stage of unwrapping occurs at 3 pN)[148] However, for assays that require larger forces (10-1000 pN) or more precise force measurements, other methods are required.

A more precise method to apply force is to use an optical tweezers system.[113, 137, 149-151] Optical tweezers are capable of producing forces from 0.1-100 pN[103, 121] and are often used in DNA folding experiments.[16, 20, 152-155] In these experiments with optical tweezers, a bead is attached to the DNA and optically trapped. Moving the optical trap, moves the bead like a pair of tweezers, applying force to the bead and therefore the DNA. The force on the bead in the optical trap is due to the intensity gradient of the focused laser. Specifically, if the bead is larger than the wavelength of the laser, then the laser light will refract through the bead, changing the momentum of the light. This momentum change will correspond to a force, which by Newton's third law has to have an equal and opposite reaction on the bead. Since the bead is spherical, light on one side of the bead will refract in the opposite direction as light on the other side, essentially giving the bead the unique ability to sense the spatial intensity profile of the laser. If the bead is at the center of the laser focus, the intensity of the light refracted by the bead is the same on both sides and there is no net force. However, if the bead is offset from the laser focus, then the intensity of the light on one side of the bead is greater than the other side. This causes a larger momentum change on one side of the bead versus the other side and a net force on the bead toward the laser focus. In this way, the intensity profile of the laser essentially creates a potential energy landscape for the bead. The greater the intensity of the laser, *I*, the lower the potential energy, *U*, as

$$I \propto -U. \quad (6)$$

For beads that are much smaller than the wavelength of light, this relationship also holds since the potential energy for the bead depends on the induced dipole moment, $\vec{p}$, and the electric field, $\vec{E}$,

$$U = -\vec{p} \cdot \vec{E} = -\alpha E^2 = -\alpha I. \quad (7)$$

Since the induced dipole moment is just the polarizability of the bead, α, times the electric field, the intensity of the laser, E^2, is again proportional to potential energy. The gradient of the intensity profile of the laser then sets the force on the bead. In the case of a focused laser, the intensity profile is Gaussian, creating a potential energy well that is approximately parabolic and a linear gradient force on the bead in the x direction, F, of

$$F = kx_{\mathrm{bd}}. \quad (8)$$

This linear force increases with the distance of the bead from the center of the trap, x_{bd}, or with the stiffness of the trap, k. Bead position is measured by using video microscopy or laser detection techniques (Section 3.2). The stiffness of the trap is set by the laser intensity and requires an initial calibration when the trap optics are first installed.[137] Optical traps can exert forces laterally, by moving the laser focus in the sample plane,[20] or axially, by moving the objective that focuses the laser.[156] Typically, optical tweezers systems work well for *in vitro* assays but may damage *in vivo* preparations or fluorescent reporters due to the increased presence of oxygen radicals. In addition, nontransparent or thick samples degrade trap performance, and forces are limited to the picoscale. A more detailed discussion of optical trapping instrumentation is given in Section 5.

Magnetic tweezers systems[103, 116, 157-159] operate under a similar principle as optical tweezers systems, but can exert higher forces of 100 pN with sub-piconewton precision[116] and do not create oxygen radicals. However, in simple, one-pole, magnetic tweezers systems, force is generated in one dimension, limiting the technique to surface-coupled assays with forces in the axial direction.[103] Recent multipole systems have been developed to overcome this limitation, but require more complex instrumentation.[157, 160, 161] In addition, bead position is generally measured using video microscopy, limiting localization precision, though laser based detection is possible.[160] In magnetic tweezers systems, a magnetic field gradient is used to create a trapping force, rather than an intensity gradient.[103, 116, 157, 158] Briefly, a superparamagnetic bead in a magnetic field has a potential energy, U, that is given by the dot product between the magnetic dipole moment of the bead, $\vec{m}$, and the magnetic field vector, $\vec{B}$, such that

$$U = -\vec{m} \cdot \vec{B}. \quad (9)$$

The corresponding force on the bead would then be equal to the gradient of this potential energy. In the simple magnetic tweezers system,[158] there are two magnetic fields that are applied to the superparamagnetic bead. The first is a larger, constant magnetic field that creates a maximum magnetization in the bead, M_{max}. The other is a smaller magnetic field that has some spatial variation. In this case, the magnitude of the force, F, would be dependent on the magnetization of the bead, the volume V of the bead, and the gradient in the magnetic field, say in the x direction,

$$F = M_{\max} V \frac{dB}{dx}. \quad (10)$$

Often, researchers set the magnetic field gradient such that a linear force is produced. The stiffness of this linear force can be calibrated by measuring the force on the bead as the magnetic field gradient increases,[158] similar to optical tweezers experiments. In addition, magnetic traps also apply a torque to the superparamagnetic bead, which allows for rotation of the bead and application of torque onto the DNA.[137] Optical traps can only apply a torque to birefringent beads.[162]

Finally, another method to apply an unfolding force to the DNA is to use an AFM.[163-165] An AFM can apply large forces (10-1000 pN),[166] but typically has reduced force precision when compared to optical or magnetic traps.[167] Ultrastable AFM setups,[167, 168] however, are capable of sub-picoNewton forces.[167] In AFM, a micron-sized, conical tip is attached to the end of a long cantilever hundreds of microns in length and is mechanically manipulated with a piezo stage. Typically, the tip has a radius of 5-10 nm and can be functionalized to provide a biochemical linkage to a DNA molecule, which is perhaps already attached to the sample surface. Mechanical movement of the cantilever in the axial dimension, then exerts a force on the DNA, F, which is dependent on the stiffness of the cantilever, k, and the axial deflection of the tip, z,

$$F = kz. \quad (11)$$

Like the optical tweezers and magnetic tweezers systems, measuring force with an AFM requires calibration of the stiffness of the cantilever, which has an accuracy of 5-20%.[103] Deflection of the AFM tip, z, is measured by reflecting a laser off the backside of the cantilever and measuring the resulting positional change in the laser. This ensures high data rates (kHz) and positional precision (sub-nanometer).[163-165] In addition, to force measurements, AFM can also be used to image DNA samples with atomic-scale precision adding to their versatility.[163, 164]

All the methods—fluid flow, optical tweezers, magnetic tweezers, and AFM—have advantages and drawbacks. The right choice for a particular experiment will depend on the geometry of the biological assay, the force requirements, the type of reporter, and the imaging method. While fluid flow is compatible with many reporters, imaging methods, and assay geometries, the force range (0.005-10 pN) is limited. Optical tweezers systems can exert larger forces (0.1-100 pN) and are compatible with a variety of assay geometries and imaging methods, but may damage fluorescent reporters and biological samples. Magnetic tweezers can access a large force range (10-100 pN) with sub-piconewton precision, but have typically been more limited in choice of reporter (superparamagnetic bead), assay geometry (surface-coupled system with forces exerted in the axial direction), and imaging method (video microscopy). Finally, AFM can exert the largest forces (10-1000 pN), but has reduced force precision, requires the use of a cantilever with laser based detection, and is limited to surface-coupled geometries.

3.4. *DNA Folding Measurements Obtained with Optical Methods*

The fundamental question in DNA folding is *how* the DNA folds. To answer this question, researchers would like to identify the long-lived and intermediate states in the DNA folding pathway and measure the energetic parameters for those states.

To accomplish this goal, researchers have used a variety of reporters, imaging methods, and biological assays. The common thread for all of these experiments is that the experiment is able to measure the topology of the DNA (the length or conformational state of the DNA) as it folds or unfolds. For example, in one assay, an organic dye is bound directly to a DNA molecule stretched by a force.[131, 169, 170] When folding proteins are added to the system, the DNA condenses, decreasing the length of the visible rod of DNA. Direct measurements of DNA length then lead to visualization of the DNA topology. In a different example, two fluorophores are biochemically attached to different locations on a DNA molecule and fluorescence resonance energy transfer (FRET) is used to determine DNA topology.[171] In FRET,[107, 124, 130, 171, 172] one fluorophore is excited by the illumination light, but not the other. As the DNA folds, the first fluorophore might come into proximity (1-10 nm) with the other fluorophore. If the emission of the first fluorophore is matched to the absorption of the second fluorophore, then a non-radiative transfer of energy will occur and the second fluorophore will begin to emit light. The intensity of this emission is related to the distance between the fluorophores. Thus, measurements of fluorophore intensity then allow for visualization of the DNA topology. Finally, in another experiment a DNA molecule may be stretched between two optically trapped beads.[173, 174] The distance between the two beads at a particular force then ultimately gives the DNA topology. In all of these cases, the methods may be different, but the overall goal of measuring DNA conformational state and topology are the same.

Once the folding states have been identified using these measurements of DNA topology, the next step is to measure the kinetic and thermodynamic parameters for these states. Again, methods will vary depending on the experiment. In this brief review, our goal will be to look at this question for the case of a TPM assay (Section 4) and the case of an optical tweezers assay (Section 5).

4. Measuring DNA Folding with the TPM Assay

One way to measure DNA folding is to use a tethered particle assay.[103, 113, 137, 145, 175] In the tethered particle assay, a particle, such as a micron-sized polystyrene bead, is tethered to the sample surface by a DNA molecule. The particle undergoes Brownian motion, but cannot diffuse away from the location where it is tethered. If there is no force on the bead, the tethered particle assay is called a TPM assay.[19, 111, 122, 176-184]

To measure DNA folding in a TPM assay (Fig. 5A), the movement of the bead is tracked in the x and y directions over time (Fig. 5B) using laser detection[185] or video microscopy.[181] If the x and y bead locations are plotted against each other (Fig. 5C-D), then the points on the graph form a two dimensional Gaussian distribution with a standard deviation that is related to the length of the DNA tether.[178] Measurements of the standard deviation with time can then track DNA length or conformation. The downside to using the TPM assay is the low spatiotemporal precision of the assay, which is on the order of

~10 nm at 1 Hz for a 3477-bp-long DNA tether attached to a 480-nm-diameter bead.[178] This low spatiotemporal precision is due to the Brownian motion of the bead.

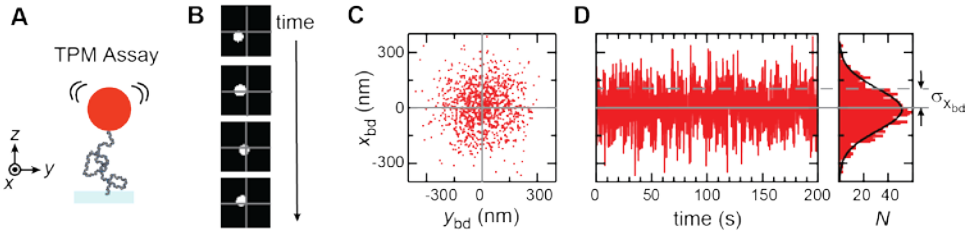

Fig. 5. Tethered particle motion (TPM) assay. *A*) In a TPM assay, a particle (*red*) is tethered to the surface by a DNA molecule. The particle moves due to Brownian motion, but this motion is restricted by the DNA tether. *B*) Here, we simulate data for a 560-µm-diameter polystyrene bead tethered to the surface by a 400-nm-long DNA molecule. Video microscopy images are taken at 200 ms and thresholded such that the bead is white and the background is black. The bead position (x_{bd}, y_{bd}) relative to the origin (*gray lines*) is calculated from these thresholded images. *C*) Plot of bead position (*red*) for every frame in the simulation shows that the average bead position is zero. *D*) However, the standard deviation of the bead position, σ_{xbd}, is nonzero and is related to the length of the DNA tether. Standard deviation, and therefore DNA length, can be calculated from a trace of bead position or by fitting a Gaussian (*black*) to a histogram of the bead position.

A more detailed look at the TPM assay can be found in this section as we outline how to prepare the assay (Section 4.1), how to detect bead position (Section 4.2), how to calibrate a TPM assay (Section 4.3), and finally, how to analyze DNA folding data for a TPM measurement (Section 4.4).

4.1. *Assay Preparation*

TPM assays involve tethering one end of a DNA molecule to a bead, and the other end to the sample chamber. Attachment is typically achieved by biochemical linkage. Specifically, in our assays we generate DNA using a biochemical procedure called polymerase chain reaction (PCR).[178, 180] Using PCR, researchers can make DNA of a particular length and can label it with specific proteins. In our experiments, we typically make ~1-µm-long DNA with a digoxigenin protein at one end and a biotin protein at the other end. These proteins will then attach to antidigoxigenin proteins on the surface and streptavidin proteins on the bead, respectively.[77] The attachment of proteins to the surface is often non-specific and occurs through electrostatic interactions with the negatively charged cover slip.[186] These interactions may fail after 24 hours or so, limiting the lifetime of the assay. Beads, on the other hand, can be purchased with covalent linkages to a number of different proteins and are viable for about a year.[187]

Preparation of the assay requires a sample chamber. Sample chambers can be made by attaching a microscope cover glass to a microscope slide with epoxy. A spacer between the cover glass and slide creates a chamber.[121, 133, 136] Solutions in the sample chamber can vary, but most involve a buffer to set the correct pH, a salt to screen charge (*i.e.* the negative charge on the backbone of the DNA or the negative charge on the glass surface), an oxygen scavenging system to remove free oxygen radicals, and lubricants to prevent unwanted nonspecific sticking and aggregation.[178, 180, 184, 188] In our experiments, we use wash buffer

(WB), which includes 25 mM Tris-HCl (pH = 7.5) as the buffer, 1 mM $Mg(CH_3COO)_2$ and 1 mM NaCl as the salts, 1 mM dithiothreitol (DTT) as the oxygen scavenging system, and 0.4% Tween and 3 mg/mL bovine serum albumin (BSA) as lubricants.[121, 189, 190]

To prepare the assay, there are three main steps. The first step is to incubate the beads and DNA molecules together in a test tube to facilitate biochemical linkage. In our experiments, we incubate 30 μL of streptavidin-coated beads (Spherotech, Lake Forest, IL) diluted in WB solution to 900 pM with 30 μL of the biotin-labeled DNA diluted in WB at a ratio of 9 beads to 1 DNA molecule for one hour at room temperature.[121] The large ratio ensures that there is only 1 DNA molecule per bead. The second step is to non-specifically bind our biochemical linking proteins to the surface of the sample chamber. In our experiments, we add a solution of 50 μL of 20 μg/mL antidigoxigenin (diluted in 100 mM Na-Phos, pH = 7.5) to the sample chamber and incubate for one hour at room temperature to get binding to the surface.[121] The higher salt concentration in the Na-Phos buffer facilitates non-specific binding of the antidigoxigenin. The sample chambers are then washed twice with 200 μL of WB[121] to remove any unbound antidigoxigenin. Finally, the last step is to add the bead-DNA solution to the sample chamber and incubate at room temperature for about an hour. This facilitates binding of the proteins on the DNA to the proteins on the surface. After incubation, the sample chamber can be washed again to remove any unbound bead-DNA complexes and is now ready for experiments.

4.2. *Detecting Bead Movement*

One way to detect bead movement in a TPM assay is to use video microscopy.[102, 123, 133, 157, 179, 191] In video microscopy, the first step is for a camera to image the focal plane of the objective, capturing images of all of the beads tethered to the surface in the field of view. If the full field of view is imaged, then the frame rate is fairly slow (5-30 Hz),[133, 182] but faster rates are possible when smaller regions of the field of view are used.[179] The second step is to post-process the images with image analysis software (*e.g.* ImageJ, Matlab),[102, 133] which can detect the locations of the beads in each frame. There are a number of algorithms to do this,[111, 145, 178, 180, 182, 192] but the simplest is to threshold the image at a particular intensity value, turning the background black and the beads that are in the focal plane white. The centroid of the bead, (x, y), is then calculated by averaging the x or y locations of all of the white pixels.[145, 178] Another method is to fit a two-dimensional Gaussian to the intensity profile of the bead.[111, 192]

Another method to detect bead movement is to use laser-based detection. In one type of laser-based detection, back-focal-plane interferometry,[89, 113, 134-137] a focused laser is scattered off of a bead of interest and this scattered light is used to identify the bead location. Specifically, in a plane that is confocal to the back-focal-plane of the objective, there will be an interference pattern created by the superposition of the scattered light with either the transmitted[135, 138, 193] or reflected[134, 136] light depending on the location of the confocal plane of interest. This interference pattern can be measured by a detector that detects both light intensity and position (*e.g.* a quadrant photodiode or position sensitive detector). When the bead moves relative to the focused laser, the interference pattern changes, changing the readout of the detector. Laser detection allows for positional

measurements in all three dimensions at the sub-nanometer scale at a data rate of hundreds of kHz.[134, 135] However, only one bead can be detected at once, instead of the multi-bead detection that occurs in video microscopy. And, the range of the detection is typically only hundreds of nanometers,[156] which is problematic in TPM assays where the motion of the bead can be thousands of nanometers.[178] The range of the detection can be increased by increasing the spot size of the focused beam, but this sacrifices positional precision.[137] In addition, the detector must be calibrated by moving the bead relative to the laser focus in known increments,[155] which can be accomplished with the optical trap in an optical tweezers system, but is difficult in a TPM assay. Thus, laser detection is typically used with optical tweezers systems and video microscopy is typically used in TPM assays.

Once bead location has been determined by the detection system, these locations are stored in a trace of either x or y location with time. Further data processing of the x and y traces is often necessary. For example, beads that are non-specifically bound to the surface of the sample chamber or those bound by multiple DNA molecules need to be identified and removed from the analysis. In a TPM assay, these beads can be identified by two characteristics.[182] First, may have traces with a smaller standard deviation. Second, a plot of the x trace versus y trace may exhibit a higher degree of eccentricity. Cutoffs on these two variables will yield traces that are likely to correspond to a single bead tethered to the surface by a single DNA molecule and filter out the rest.[177, 184] Traces may also need to be de-drifted to remove any gradual, non-stochastic movement in a particular direction, perhaps caused by mechanical settling, temperature fluctuations, or fluid evaporation. One simple method is to fit the long-time-scale, correlated movement of all the beads on the slide and then to subtract this correlated movement from the trace.[178] Other passive or active methods are possible as well.[121, 134, 135] Once the traces have been filtered and de-drifted, they are ready for analysis.

4.3. *TPM Calibration*

In analyzing data for a TPM experiment, researchers often want to calculate DNA length from the motion of the tethered bead. This requires a calibration. Specifically, a calibration equation is needed to turn the standard deviation of the movement of the bead (*e.g.* σ_{xbd}) into DNA length (L). To relate standard deviation of the bead motion to DNA length there are two methods.

In the first method, a theoretical model is used to relate standard deviation to length. In this model, the average movement of the bead is calculated by adding the movement of the bead due to Brownian motion to the movement of the bead due to the tensional force exerted by the DNA molecule.[175, 177, 180, 183, 184] The displacement of the bead in a single dimension due to Brownian motion is essentially the root mean squared displacement, x_{diff}, which was theorized by Einstein in 1905.[194] This root mean squared displacement increases with time lag, τ, and the diffusion coefficient, D, via the relationship

$$x_{\mathrm{diff}} = \sqrt{2D\tau}. \tag{12}$$

To calculate the displacement of the bead due to the tensional force of the DNA requires first modeling this tensional force. The tensional force due to the DNA, F, can be approximated as an entropic spring,[19, 184, 195]

$$F \approx -kx_{\text{diff}}, \tag{13}$$

where k is the spring constant of the DNA given by the worm-like chain (WLC) model.[196] In the WLC model, each subunit in the polymer is not able to freely rotate, which is the case for double-stranded DNA. Instead, there is cooperativity between subunits so that the polymer is only semi-flexible and requires energy to bend. In this case, the spring constant is approximately equal to,[19]

$$k \approx \frac{3k_{\text{B}}T}{2L_{\text{p}}L_{\text{c}}}. \tag{14}$$

Here $k_{\text{B}}T$ is the thermal energy, L_{p} is the persistence length of the DNA, and L_{c} is the contour length of the DNA. From this equation, one can see that the tensional force increases as the diffusion of the bead increases or as the contour length of the DNA decreases. This tensional force will create a drift velocity, v_{drift}, that will oppose the diffusive movement. This drift velocity has been derived elsewhere as,[16, 155, 197]

$$v_{\text{drift}} \approx \frac{-kx_{\text{diff}}}{k_{\text{B}}T} D. \tag{15}$$

Using this drift velocity, we can then calculate an average drift displacement, x_{drift}, which is a displacement in the opposite direction of the diffusion of the bead due to the tensional force,

$$x_{\text{drift}} \approx \frac{-kx_{\text{diff}}}{k_{\text{B}}T} D\tau. \tag{16}$$

Adding this movement to the diffusive movement then gives the total root-mean-square movement of the bead. This is equal to the standard deviation of the bead movement, σ_{xbd}, when the mean is zero,

$$\sigma_{\text{xbd}} \approx x_{\text{diff}} + \frac{-kx_{\text{diff}}}{k_{\text{B}}T} D\tau. \tag{17}$$

From this equation, we see that the standard deviation of the bead motion should decrease if the spring constant for the DNA increases. This occurs when the length of the DNA decreases, as we expect. However, this equation is not immediately useful for determining a calibration of standard deviation to DNA length. To get a useful equation, computer simulations are used to generate data, and then a polynomial fit to the simulated data gives the desired standard deviation to length calibration curve.[183]

In the second method, an experimental model is used to relate the standard deviation of the bead motion to the DNA length.[176-178] Specifically, the standard deviation of the bead is measured for various tethers of known length. These data points are then fit with a polynomial to get the desired calibration curve. In practice, both the theoretical and experimental methods are used to measure the calibration curve to make sure there are no discrepancies.

4.4. *TPM Measurements*

TPM experiments measure standard deviation of bead motion with time, which can then be related to the overall DNA length. Thus, an important measurement in DNA folding assays with TPM is measuring this length parameter with time. This gives the states in the folding pathway and the sequence of those states. In addition, if the folding process is reversible, then measurement of length vs. time will ultimately yield the kinetic and thermodynamic parameters for the folding pathway as well.

To illustrate this point, we have created a simulation for a TPM assay with a 560-nm-diameter bead attached to a 400-nm-long DNA molecule (Fig. 6A). This DNA molecule forms a 200-nm-circumference loop when it folds, perhaps due to the addition of protamine proteins. Using these input parameters, we first generate 5 Hz data for bead position vs. time using Gaussian noise with the appropriate standard deviation and exponentially distributed dwell times for both the folded and unfolded state. Then, we calculate the standard deviation of the bead motion and smooth (average over consecutive overlapping intervals) to 0.05 Hz. Applying the standard deviation to length calibration recovers the original DNA length (Fig. 6B). From this data, we can identify two distinct states and can see that the molecule goes from one state to the other reversibly. We did not simulate any intermediate states and do not see any intermediate states in the length vs. time plot.

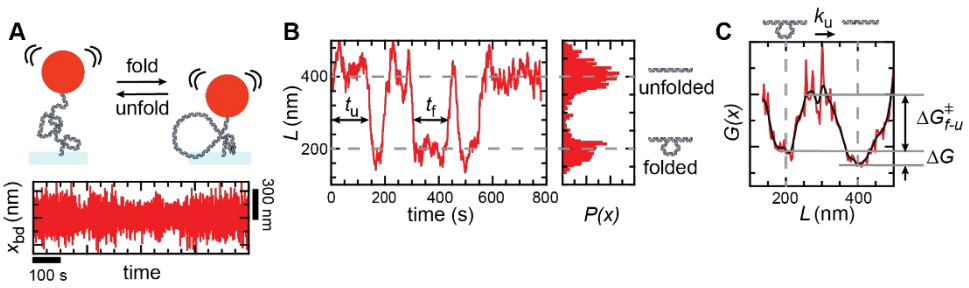

Fig. 6. TPM measurements of DNA folding. *A*) Cartoon of a TPM assay with a 560-nm-diameter polystyrene bead tethered to the surface by a 400-nm-long DNA molecule. When the DNA folds, it forms a 200-nm-circumference loop. Cartoon not to scale. Inset shows simulated bead motion, x_{bd}, at a data rate of 5 Hz as the DNA folds and unfolds. *B*) The length of the DNA (L) calculated from the standard deviation of the bead motion sharply decreases upon folding the DNA into a loop. Data smoothed to 0.05 Hz. Unfolding the DNA causes an increase in the DNA length back to the initial level. Measurements of the dwell time in the folded or unfolded state (t_f or t_u, respectively) allow for calculation of the kinetic parameters (k_f or k_u, respectively). A normalized histogram of the data shows the probability density, *P(x)* and allows for calculation of the free energy *G(x)*. *C*) The energy landscape (*red*) with a reaction coordinate of DNA length shows the energy difference between the states (ΔG) and the height of the energy barrier from the folded or unfolded state ($\Delta G^{\ddagger}_{f\text{-}u}$ or $\Delta G^{\ddagger}_{u\text{-}f}$, respectively). Smoothing the data (*black*) reduces noise.

Since the data is reversible, we can also measure thermodynamic and kinetic parameters for our simulation. To measure the thermodynamic parameters, we create a normalized histogram of the length vs. time trace (Fig. 6B), which is equal to the probability density, $P(x)$. This probability density is related to the relative free energy of the states, $G(x)$, through the thermal energy, k_BT,[143]

$$G(x) = -k_BT \ln[P(x)]. \qquad (18)$$

Once the graph of *G(x)* is known (Fig. 6C), we can then measure the energy difference between the two states (ΔG) and the height of the energy barrier ($\Delta G^{\ddagger}$). To measure the kinetic parameters, we calculate the average time the molecule spends in each state, the dwell time. The average dwell time then gives the inverse of the rate constant for transitioning from one state to the other.

Thus, we see that TPM assays are useful for measuring DNA length with time, allowing for quantification of DNA folding without an optical tweezers system. This is important in cases where DNA folding occurs at low forces, such that even a small force of 1 pN, might skew the energy landscape enough to remove folding states from being visualized by the assay.[143] For example, DNA loop formation in sperm cells due to protamine binding[89] might be a good candidate for TPM assays since intermediates in the nucleation of the loop may require an assay without force. However, if force can be applied to an assay, then the addition of an optical tweezers system allows for a greater variety of measurements.

5. Measuring DNA Folding with Optical Tweezers

Optical tweezers[20, 77, 103, 113, 137, 143, 146, 151, 152, 155, 198-205] are a versatile method that can be used to directly unfold DNA, allowing for measurements of the folded states, the pathway, and the kinetic and thermodynamic variables. To measure DNA unfolding with an optical tweezers system, a focused laser is used to optically trap a bead with an intensity gradient (see Section 3.3). Typically, this bead is tethered to the surface by a DNA molecule in a tethered particle assay, however other geometries are possible.[113, 198] Moving the laser focus moves the equilibrium position of the optical trap, allowing for tweezers-like manipulation of the bead in the sample plane. To unfold the DNA, the optical trap moves relative to the tether point at the surface, applying a tensional force that stretches the DNA. As the DNA stretches, it unfolds, sequentially transitioning through a series of folded states. Once the molecule is completely unfolded, the force can be relaxed to allow for refolding. Two measurements that are often taken with optical tweezers systems are force-extension curves and measurements of DNA conformation with time.

A more detailed look at the optical tweezers instrument can be found in this section as we outline how to set up an optical trap (Section 5.1), how to calibrate the stiffness of the trap (Section 5.2), and finally how to measure DNA unfolding with the instrument (Section 5.3).

5.1 Experimental Setup for an Optical Tweezers System

The instrumentation for an optical tweezers system has been reviewed previously. [103, 113, 137, 143, 198, 199] Here our goal will be to discuss the main considerations for the system and the general experimental outline.

The top consideration for the optical trap is laser intensity, though wavelength is also important. The higher the intensity of the trap laser, the higher the value of the trap stiffness, and therefore, the trapping force (~1 pN per 10 mW).[137] Due to the high intensity requirement, the trap laser is typically a solid-state laser. These lasers usually have output intensities of ~10 W. The other consideration for the optical trap is the wavelength of the laser. The wavelength needs to be chosen such that there is reduced photon absorption by

the surrounding fluid, limiting oxygen radicals.[206] Often wavelengths in the near infrared work well, such as the 1064 nm wavelength for a neodymium-doped yttrium aluminum garnet (Nd:YAG) crystal laser. Also, when choosing a wavelength for the trap laser, the detection system for the bead will need to be considered. If a detection laser will be used, then the two lasers will need to have different wavelengths so that they can both be coupled into the same beam path by a dichroic mirror. If the Nd:YAG laser is chosen for the trap laser, then the wavelength of the detection laser is often in the red or near infrared (700-950 nm). In contrast to the trap laser, the detection laser, has a lower intensity (~100 mW) and is often a much cheaper ($1k versus $30k), fiber-coupled, diode laser. The fiber coupling removes the elliptical mode of the laser diode in favor of a Gaussian mode.[150] However, this spatial filtering also converts mode noise into intensity noise, which may need to be removed in precision assays.[121] Selection of the trap and detection lasers will depend on the intensity needed for the assay and the availability of the optics for the chosen wavelength.

The general experimental outline to produce a working optical trap is to couple the trap laser into a high numerical aperture objective (1.2-1.4 NA). The objective focuses the trap laser to a diffraction-limited spot in the sample plane, creating the intensity gradient for the trap. The tighter the focus for the trap laser, the higher the intensity gradient available for trapping beads. Movement of the trap in x and y is controlled by a piezo-electric mirror or other beam steering optic[137] placed in the laser path at a location confocal to the back-focal-plane of the objective. Rotations of the laser in this conjugate plane lead to translations at the sample plane, with minimal clipping on the objective aperture. The axial position, z, of the trap is manually controlled by moving lenses in a telescope system in the laser path that changes the collimation of the laser. If a detection laser is used, then both the detection and trap lasers are collocated in the sample plane so that the detection laser measures the position of the bead relative to the center of the optical trap. Movements of the detection laser relative to the trap laser are controlled by a separate beam-steering optic and lens system in the detection laser path. Movements of the tether point relative to both lasers are accomplished by incrementing a piezo-electric stage that holds the sample.

5.2 Calibrating an Optical Tweezers System

To calibrate an optical trap, researchers measure trap stiffness as a function of laser intensity.[137, 193, 198, 207, 208] The reason for this calibration is so that a program can easily set the desired stiffness by modifying the laser intensity on the fly. Typically, researchers use three methods to calibrate the trap: the drag method, the equipartition theorem method, and the power spectrum method. Here our goal will be to briefly describe each method and to present the limitations and advantages of each.

In the drag method, stiffness is determined by measuring the displacement of the bead from the trap center at various forces. The slope of this force vs. displacement data is the trap stiffness, using Eq. 8. To apply a known force, researchers create a drag force by moving the sample with a piezo-electric stage or flowing fluid through the sample (see Section 3.3). Increasing the relative speed of the fluid increases the drag force as in Eq. 5. Measuring the one-dimensional stiffness (*e.g.* k_x) can then be accomplished by determining

the drag coefficient, γ, the relative speed of the fluid, v, and the bead displacement (*e.g.* x_{bd}),

$$k_{\mathrm{x}} = \frac{dF}{dx_{\mathrm{bd}}} \approx \frac{\gamma \Delta v}{\Delta x_{\mathrm{bd}}} \approx 6\pi\eta r_{\mathrm{bd}} \frac{\Delta v}{\Delta x_{\mathrm{bd}}}. \quad (19)$$

Repeating the entire procedure at different laser intensities gives the stiffness as a function of laser intensity. One limitation of this method is that the drag coefficient must be known, which is problematic if the trap laser causes significant heating of the medium.[117] Another limitation is that researchers must calibrate the detector that measures bead position. However, the advantage of this method is that the relationship between trapping force and bead position is measured directly and is not assumed to be linear.

In the equipartition theorem method, the equipartition theorem is used to equate the potential energy of the bead in the trap in one dimension, U, to one-half the thermal energy, $k_{\mathrm{B}}T$,

$$U = \frac{k_B T}{2}. \quad (20)$$

Since the trap is harmonic, the potential energy of the bead in the trap in one dimension depends on the trap stiffness and the square of the standard deviation of the bead position (e.g. σ_{x}). Substituting this value into Eq. 21 and rearranging gives,

$$k_x = \frac{k_B T}{\sigma_x^2}. \quad (21)$$

The advantage of this method is the ease of the measurement and that the equation makes no assumptions about the drag coefficient of the surrounding fluid. However, the equipartition theorem method relies on the assumption that the bead position is calibrated, the temperature of the medium is known, and the trap force is linear.

Finally, in the power spectrum method, the Fourier transform of the motion of the bead in the trap can be used to determine stiffness. The motion of the bead in the trap over time, t, is described via the overdamped Langevin equation,

$$k_x x_{bd} + \gamma \frac{dx_{bd}}{dt} = F_x. \quad (22)$$

Here, k_{x} is trap stiffness, γ is the drag coefficient, x_{bd} is the one-dimensional bead position, and F_{x} is the trapping force, whose time average is zero since the bead is trapped in a stable equilibrium. The Fourier transform of the Langevin equation into frequency space, f, yields a one dimensional Lorentz-like power spectrum, S_{x},[137]

$$S_x(f) = \frac{k_B T}{\gamma \pi^2 (f_c^2 + f^2)}. \quad (23)$$

The characteristic frequency of the spectrum, f_{c}, is related to the trap stiffness by,

$$k_x = 2\pi\gamma f_c. \quad (24)$$

Measurement of the power spectrum of the bead motion at each laser intensity then gives the stiffness as a function of laser intensity. The limitations of the power spectrum method are that the drag coefficient must be known and that the trap is assumed to be linear to apply the Langevin equation. However, the advantage is that the method does not rely on calibrating bead position, since only the frequency, rather than the absolute bead displacement is of relevance.

For reliable results, all three methods are typically used to calibrate the instrument and others may be used as well.[207, 209] One of these other methods is to drive the stage at a known velocity and simultaneously measure the power spectrum of the bead as it is displaced from the trap center, removing the requirement for measuring the drag coefficient and the calibration for bead position.[209] The choice of method and how many methods to perform will depend on the precision of the calibration needed and the trouble shooting required.

5.3 Optical Tweezers Measurements

Two typical optical tweezers measurements for DNA unfolding are force-extension curves andDNA conformation with time.[20, 113, 137, 143, 152, 210, 211] Both measurements give information about the topology of the DNA as it unfolds and both may be used to determine the energy parameters of the folding process.

To illustrate these measurements, we have created a simulation of an optical tweezers experiment with a tethered particle assay. In this simulation, we have a DNA tether that contains a folded region consisting of one loop (Fig. 7A). When the optical trap is moved, the bead in the optical trap moves too, applying a stretching force (F) to the DNA. As the stretching force increases, the loop unfolds increasing the end-to-end distance or extension (x) of the DNA.

One particular measurement of interest is to plot the force-extension curve for this scenario. Typically, the extension of the DNA at a particular stretching force will be given by the WLC model for polymers.[15, 196, 211, 212] Using this model, the approximate force on the DNA is related to the thermal energy, $k_{\mathrm{B}}T$ (4 pN-nm), the persistence length, the contour length, and the extension of the molecule,[196]

$$F = \frac{k_{\mathrm{B}}T}{L_{\mathrm{p}}}\left(\frac{1}{4\left(1-\frac{x}{L_{\mathrm{c}}}\right)^2} - \frac{1}{4} + \frac{x}{L_{\mathrm{c}}}\right). \qquad (25)$$

This equation works for moderate forces (0.1-10 pN), but corrections are needed at other force regimes.[196, 213] If the DNA molecule is just condensed into a globule due to entropy and doesn't have any folds, then a fit of the WLC curve to the force-extension curve should produce the contour length of the DNA. However, if the DNA is folded, then the apparent contour length of the molecule is smaller, and the force-extension curve follows a WLC curve with a smaller apparent contour length. In our simulated force-extension curve (Fig. 7B), we see that our simulated data follows two WLC curves, indicating the presence of two states, and no intermediates. The fit with a smaller contour length corresponds to the folded DNA state, while the fit with the larger contour length corresponds to the unfolded state. Interestingly, at the moment when the DNA unfolds, there is a characteristic "pop" in the force-extension curve. This characteristic pop shows a dramatic decrease in the force with a concurrent increase in the extension. This allows for measurement of the unfolding force. Hence, using the force-extension curve allows researchers to measure the unfolding force (F_{unfold}), the length change for the fold (ΔL), and the sequence of the various states in the unfolding pathway.[153-155, 197, 214, 215]

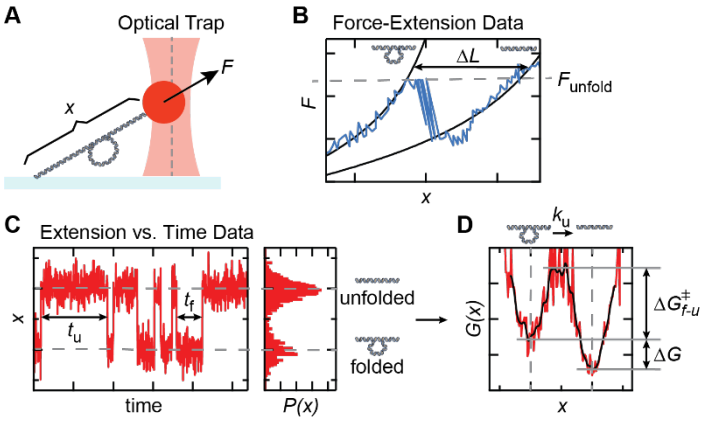

Fig. 7. Optical trapping measurements of DNA unfolding. *A*) Optical trapping assay where bead (*red*) is trapped with a trap laser (*pink*) and tethered to the surface by a DNA molecule. Moving the trapped bead away from the tether point exerts a force on the DNA (*F*) causing it to stretch to a particular extension (*x*). *B*) A simulated force-extension curve for a DNA molecule with one loop shows that as the DNA is stretched, the force and extension increase. At a particular force (F_{unfold}), the loop in the DNA unfolds, dramatically increasing the extension and decreasing the force. If stretching continues, the force and extension again increase. Fitting of the force-extension curve to a WLC model (*black*) before and after the unfolding event allow for measurement of the contour length of the molecule. The difference in the contour length (ΔL) gives the length of DNA folded in the loop. *C*) Simulated extension vs. time data at a force close to F_{unfold} shows that the extension increases upon unfolding. Measurement of the dwell times in each state give the kinetic parameters. Histogram of the data gives the probability density and allows for calculation of the free energy. *D)* The energy landscape (*red*) for the reaction coordinate of extension. Variables and color scheme same as in Fig. 6.

Another measurement of interest is to plot the extension vs. time graph for this scenario (Fig. 7C). In this case, we set the force close to the measured unfolding force and just monitor the extension. In our simulation, the extension moves back and forth between two values, indicating that the process is reversible between the two detected states. There are no intermediates. Using this data, we can then calculate the kinetic and thermodynamic parameters (see Sec. 4.4) for the unfolding pathway (Fig. 7C-D). If the process was not reversible, then calculation of these parameters would require more sophisticated methods.[143] In addition, the extension of the DNA can be related to the DNA length for measurement of ΔL, and the force can be changed to shift the energy landscape.

Thus, there are a striking array of measurements that can be done to elucidate a DNA folding pathway with an optical tweezers system. These measurements have been used to look at the unfolding of DNA from chromatin,[155] the unwrapping of DNA around nucleosomes,[148, 216] the looping of DNA into toroids,[93] and the unfolding of DNA from hairpin structures,[173] and will be important for future assays as well.

6. Future Directions

Currently, DNA folding is measured using a large variety of reporters, imaging methods, and ways to apply force. The TPM assay and optical tweezers system are just two of the methods in use. Yet, despite this bounty of experimental variation, future assays will undoubtedly encounter new technical difficulties that will drive more innovation. Recently, researchers have worked to combine methods, including fluorescence and optical

tweezers,[217] laser-based detection and AFM,[168] multi-color detection and TIRF,[218] and super resolution imaging and optical tweezers[219] to provide even more options. In the future, more technological advances will need to be made, including perhaps the ability to multiplex using "lab on a chip" platforms.[220, 221]

In addition, single molecule biology in recent years has been moving toward more *in vivo* experiments.[222-224] For DNA folding, this might mean answering questions like: How does folding change with cell cycle or development? Are the folding dynamics different in the crowded environment of the cell? Or, how do multiple protein complexes within the cell interact to hinder or promote the folding process? These questions would require *in vivo* experiments, but would still need to be measured using the individual dynamics of single molecules.

Finally, DNA folding is not just limited to the natural folding pathways we discussed. Biomaterials research has been exploding in recent decades with pushes to engineer DNA into different types of nanostructures, including 3D shapes, nanomachines, and large macroscale materials.[225-228] Understanding DNA folding for engineering purposes will only continue to become more important in the future.

Acknowledgements

This work was supported by a Research Corporation Cottrell Science Award (ARC, 23239), a National Science Foundation CAREER award (ARC, 1653501), and Amherst College.

References

1. B. Alberts, Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P., *Molecular Biology of the Cell*, 4th ed. (Garland Science, New York, 2002)."Chapter 4. DNA and Chromosomes."
2. W. Fischle, et al., *Curr. Opin. Cell Biol.* **15** (2003) 172-183.
3. A. Shilatifard, in *Annu. Rev. Biochem.* (Annual Reviews, Palo Alto, 2006), Vol. 75, pp. 243-269.
4. V. W. Zhou, et al., *Nat. Rev. Genet.* **12** (2011) 7-18.
5. K. Luger, et al., *Nat Rev Mol Cell Biol* **13** (2012) 436-447.
6. T. Jenuwein and C. D. Allis, *Science* **293** (2001) 1074-1080.
7. E. I. Campos and D. Reinberg, *Annu Rev Genet* **43** (2009).
8. A. J. Bannister and T. Kouzarides, *Cell Res* **21** (2011).
9. J. Mateos-Langerak, et al., *Arxiv* (2007).
10. J. D. Watson and F. H. Crick, *Nature* **171** (1953) 737-738.
11. M. H. F. Wilkins, Stokes, A.R. and Wilson, H.R., *Nature* **171** (1953) 738-740.
12. R. a. G. Franklin, R. G., *Nature* **171** (1953) 740-741.
13. R. Phillips, et al., *Physical Biology of the Cell, Second Edition*. (Taylor & Francis Group, 2012)."Physical Biology of the Cell, Second Edition."
14. M. D. Frank-Kamenetskii, *Physics Reports* **288** (1997) 13-60.
15. H. B. Bohidar, *Fundamentals of Polymer Physics and Molecular Biophysics*. (Cambridge University Press, 2015)."Fundamentals of Polymer Physics and Molecular Biophysics."
16. C. Bustamante, et al., *Current opinion in structural biology* **10** (2000) 279-285.
17. P. J. Hagerman, *Annu Rev Biophys Biophys Chem* **17** (1988) 265-286.
18. C. Bustamante, et al., *Nature* **421** (2003) 423-427.
19. S. Brinkers, et al., *The Journal of chemical physics* **130** (2009) 215105.
20. M. D. Wang, et al., *Biophysical Journal* **72** (1997) 1335-1346.
21. S. B. Smith, et al., *Science* **258** (1992) 1122-1126.
22. T. R. Strick, et al., *Science* **271** (1996) 1835-1837.
23. Y. Lu, et al., *Biopolymers* **61** (2001) 261-275.
24. S. Geggier, et al., *Nucleic Acids Research* **39** (2011) 1419-1426.
25. S. G. Gregory, et al., *Nature* **441** (2006) 315-321.
26. K. A. Dill, *Biochemistry* **29** (1990) 7133-7155.
27. K. A. Dill and J. L. MacCallum, *Science* **338** (2012) 1042.
28. V. A. Bloomfield, *Current Opinion in Structural Biology* **6** (1996) 334-341.
29. G. C. Wong and L. Pollack, *Annu Rev Phys Chem* **61** (2010) 171-189.
30. S. S. Cohen, *A Guide to the Polyamines*. (Oxford University Press, 1998)."A Guide to the Polyamines."
31. V. A. Bloomfield, *Biopolymers* **44** (1997) 269-282.
32. D. C. Rau, et al., *Proceedings of the National Academy of Sciences* **81** (1984) 2621-2625.
33. D. C. Rau and V. A. Parsegian, *Biophysical Journal* **61** (1992) 246-259.
34. N. A. M. Besseling, *Langmuir* **13** (1997) 2113-2122.
35. J. Li, et al., *Nanomaterials* **5** (2015) 246.
36. S. B. Zimmerman and L. D. Murphy, *FEBS letters* **390** (1996) 245-248.
37. A. P. Minton, *Current Opinion in Biotechnology* **8** (1997) 65-69.
38. D. Hall and A. P. Minton, *Biochim Biophys Acta* **1649** (2003) 127-139.

39. A. P. Minton, *Structural and Organizational Aspects of Metabolic Regulation.* (1990) 291-306.
40. R. L. Jernigan, et al., *Journal of Biomolecular Structure and Dynamics* **4** (1986) 41-48.
41. C. A. Hunter and J. K. M. Sanders, *Journal of the American Chemical Society* **112** (1990) 5525-5534.
42. P. Yakovchuk, et al., *Nucleic Acids Research* **34** (2006) 564-574.
43. T. Schlick and W. K. Olson, *Journal of Molecular Biology* **223** (1992) 1089-1119.
44. B. Alberts, et al., *Molecular Biology of the Cell*, 6 ed. (Garland Science, Taylor & Francis Group, LLC, New York, 2015)."Molecular Biology of the Cell."
45. S. R. Mishra, *Biomolecules*. (Discovery Publishing House, 2003)."Biomolecules."
46. J. M. Berg, et al., *Biochemistry*. (W. H. Freeman, 2007)."Biochemistry."
47. P. Carrivain, et al., *Soft Matter* **8** (2012) 9285-9301.
48. M. G. Mateu, *Structure and Physics of Viruses: An Integrated Textbook*. (Springer Netherlands, 2013)."Structure and Physics of Viruses: An Integrated Textbook."
49. N. E. Davey, et al., *Trends Biochem Sci* **36** (2011) 159-169.
50. L. R. Comolli, et al., *Virology* **371** (2008) 267-277.
51. P. J. Jardine and D. L. Anderson, *The bacteriophages* **2** (2006) 49-65.
52. S. Sun, et al., *Current opinion in structural biology* **20** (2010) 114-120.
53. D. E. Smith, et al., *Nature* **413** (2001) 748-752.
54. D. N. Fuller, et al., *Journal of Molecular Biology* **373** (2007) 1113-1122.
55. L. W. Black and J. A. Thomas, *Advances in experimental medicine and biology* **726** (2012) 469-487.
56. M. E. Cerritelli, et al., *Cell* **91** (1997) 271-280.
57. N. V. Hud and K. H. Downing, *Proceedings of the National Academy of Sciences* **98** (2001) 14925-14930.
58. J. Lepault, et al., *The EMBO journal* **6** (1987) 1507-1512.
59. M. Thanbichler, et al., *Journal of cellular biochemistry* **96** (2005) 506-521.
60. P. A. Wiggins, et al., *Proceedings of the National Academy of Sciences* **107** (2010) 4991-4995.
61. K. Struhl, *Cell* **98** 1-4.
62. S. C. Dillon and C. J. Dorman, *Nature reviews. Microbiology* **8** (2010) 185-195.
63. R. T. Dame, *Molecular microbiology* **56** (2005) 858-870.
64. S. Y. Lee, et al., *Scientific Reports* **5** (2015) 18146.
65. J. Stavans and A. Oppenheim, *Physical biology* **3** (2006) R1.
66. R. Brunetti, et al., *Biochimie* **83** (2001) 873-882.
67. R. M. Martin and M. C. Cardoso, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **24** (2010) 1066-1072.
68. M. H. Chow, et al., *Eukaryotic cell* **9** (2010) 1577-1587.
69. E. Kellenberger, *Trends in Biochemical Sciences* **12** 105-107.
70. E. Kellenberger and B. Arnold-Schulz-Gahmen, *FEMS Microbiology Letters* **100** (1992) 361.
71. R. Balhorn, *The Journal of cell biology* **93** (1982) 298-305.
72. H. F. Lodish, *Molecular cell biology*. (W.H. Freeman and Co., New York, 2013)."Molecular cell biology."
73. D. E. Sadava, *Life : the science of biology*. (Sinauer Associates, Sunderland, MA, 2014)."Life : the science of biology."
74. H. J. Szerlong and J. C. Hansen, *Biochem Cell Biol* **89** (2011) 24-34.

75. L. S. Shlyakhtenko, et al., *Biochemistry* **48** (2009) 7842-7848.
76. S. A. Grigoryev and C. L. Woodcock, *Experimental Cell Research Special Review Issue: Chromosome Biology, 2012* **318** (2012) 1448-1455.
77. F. Ritort, et al., *Physical review letters* **96** (2006) 118301.
78. K. Mekhail and D. Moazed, *Nature Reviews, Molecular and Cell Biology* **11** (2010) 317-328.
79. M. Baribieri, et al., *Nucleus* **4** (2013) 267-273.
80. V. B. Teif and K. Bohinc, *Progress in Biophysics and Molecular Biology* **105** (2011) 208-222.
81. M. M. Cox and M. O'Donnell, *Molecular biology : principles and practice*. (W.H. Freeman and Company, New York, 2015)."Molecular biology : principles and practice."
82. D. W. Van de Vosse, et al., *Wiley Interdiscip Rev Syst Biol Med* **3** (2011) 147-166.
83. A. M. Brunner, et al., *Epigenetics & Chromatin* **7** (2014) 2.
84. P. G. Adenot, et al., *Development* **124** (1997) 4615-4625.
85. R. Balhorn, *Genome biology* **8** (2007) 1.
86. W. S. Ward, *Molecular human reproduction* **16** (2010) 30-36.
87. S. Gonzalez-Rojo, et al., *PLOS ONE* (2014) 1-21.
88. S. Lambard, et al., *Molecular Human Reproduction* **10** (2004) 535-541.
89. L. H. Cree, et al., *Protein and peptide letters* **18** (2011) 802-810.
90. D. Poccia, *International review of cytology* **105** (1986) 1-65.
91. N. V. Hud and I. D. Vilfan, *Annu. Rev. Biophys. Biomol. Struct.* **34** (2005) 295-318.
92. M.-L. Ainalem and T. Nylander, *Soft Matter* **7** (2011) 4577-4594.
93. L. R. Brewer, *Integrative biology : quantitative biosciences from nano to macro* **3** (2011) 540-547.
94. R. Golan, et al., *Biochemistry* **38** (1999) 14069-14076.
95. R. Balhorn, et al., *Molecular Reproduction and Development* **56** (2000) 230–234.
96. T. Jenuwein, et al., *Nature* **385** (1997) 269-272.
97. H. H. Q. Heng, et al., *Journal of Cell Science* **117** (2004) 999–1008.
98. D. Bikard, et al., *Microbiology and Molecular Biology Reviews* **74** (2010) 570-588.
99. L. Finzi and W. K. Olson, *Biophysical Reviews* **8** (2016) 1-3.
100. S. Weiss, *Science* **283** (1999) 1676-1683.
101. D. Alcor, et al., *The European journal of neuroscience* **30** (2009) 987-997.
102. C. Manzo and M. F. Garcia-Parajo, *Reports on Progress in Physics* **78** (2015) 124601.
103. K. C. Neuman and A. Nagy, *Nature Methods* **5** (2008) 491-505.
104. M. P. Clausen and B. C. Lagerholm, *Current protein & peptide science* **12** (2011) 699-713.
105. G. Y. Wiederschain, *Biochemistry (Moscow)* **76** (2011) 1276-1276.
106. G. Cosa, et al., *Photochemistry and photobiology* **73** (2001) 585-599.
107. T. Ha and P. Tinnefeld, in *Annual Review of Physical Chemistry, Vol 63*, edited by M. A. Johnson and T. J. Martinez (Annual Reviews, Palo Alto, 2012), Vol. 63, pp. 595-617.
108. A. Crut, et al., *Nucleic Acids Res* **33** (2005) e98.
109. W. C. W. Chan and S. Nie, *Science* **281** (1998) 2016.
110. T. Kanda, et al., *Current Biology* **8** (1998) 377-385.
111. J. Milstein, et al., *Biopolymers* **95** (2011) 144-150.
112. R. K. Montange, et al., *Optics express* **21** (2013) 39-48.

113. J. R. Moffitt, et al., *Annu. Rev. Biochem.* **77** (2008) 205-228.
114. S. H. Leuba, et al., *Proc. Natl. Acad. Sci. U. S. A.* **100** (2003) 495-500.
115. J. M. T. Thompson, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* **464** (2008) 2811-2829.
116. I. De Vlaminck and C. Dekker, in *Annual Review of Biophysics, Vol 41*, edited by D. C. Rees (Annual Reviews, Palo Alto, 2012), Vol. 41, pp. 453-472.
117. Y. Seol, et al., *Opt. Lett.* **31** (2006) 2429-2431.
118. M. Iwaki, et al., *Nano Letters* **15** (2015) 2456-2461.
119. H. Isojima, et al., *Nature chemical biology* **12** (2016) 290-297.
120. K. Ritchie and A. Kusumi, *Methods Enzymol* **360** (2003) 618-634.
121. A. R. Carter, et al., *Biophysical Journal* **96** (2009) 2926-2934.
122. N. Pouget, et al., *Nucleic Acids Research* **32** (2004) e73-e73.
123. J.-H. Spille, Massachusetts Institute of Technology, 2014.
124. C. Joo, et al., *Annu. Rev. Biochem.* **77** (2008) 51-76.
125. D. B. Murphy, *Fundamentals of Light Microscopy and Electronic Imaging*. (Wiley, 2002)."Fundamentals of Light Microscopy and Electronic Imaging."
126. R. L. Price and W. G. J. Jerome, *Basic Confocal Microscopy*. (Springer New York, 2011)."Basic Confocal Microscopy."
127. C. Zander, et al., *Appl. Phys. B-Lasers Opt.* **63** (1996) 517-523.
128. K. N. Fish, *Current protocols in cytometry* **Chapter 12** (2009) Unit12.18.
129. A. Trache and G. A. Meininger, *Current protocols in microbiology* **Chapter 2** (2008) Unit 2A.2.1-2A.2.22.
130. A. Gust, et al., *Molecules (Basel, Switzerland)* **19** (2014) 15824-15865.
131. E. C. Greene, et al., *Methods in enzymology* **472** (2010) 293-315.
132. T. Fujiwara, et al., *The Journal of Cell Biology* **157** (2002) 1071-1082.
133. M. A. Catipovic, et al., *American Journal of Physics* **81** (2013) 485-491.
134. A. R. Carter, et al., *Optics express* **15** (2007) 13434-13445.
135. A. R. Carter, et al., *Applied optics* **46** (2007) 421-427.
136. F. B. Shipley and A. R. Carter, *Optics express* **20** (2012) 9581-9590.
137. K. C. Neuman and S. M. Block, *Review of Scientific Instruments* **75** (2004) 2787-2809.
138. F. Gittes and C. F. Schmidt, *Opt Lett* **23** (1998) 7-9.
139. H. Ueno, et al., *Biophys J* **98** (2010) 2014-2023.
140. F. Verpillat, et al., *Optics express* **19** (2011) 26044-26055.
141. H. Harutyunyan, et al., *Nano Lett* **10** (2010) 5076-5079.
142. W. Qian, et al., *BIOMEDO* **15** (2010) 046025-046025-046029.
143. M. T. Woodside and S. M. Block, *Annual review of biophysics* **43** (2014) 19.
144. Y. Seol and K. C. Neuman, *Cell* **153** 1168-1168.e1161.
145. D. Song, et al., *American Journal of Physics* **83** (2015) 418-426.
146. P. Gross, et al., *Methods Enzymol* **475** (2010) 427-453.
147. J. R. Taylor, *Classical Mechanics*. (University Science Books, 2005)."Classical Mechanics."
148. S. Mihardja, et al., *Proceedings of the National Academy of Sciences* **103** (2006) 15871-15876.
149. M. Dienerowitz, et al., *J. Nanophotonics* **2** (2008) 32.
150. P. H. Jones, et al., *Optical Tweezers: Principles and Applications*. (Cambridge University Press, 2015)."Optical Tweezers: Principles and Applications."

151. D. Kishan, et al., *Physics World* **15** (2002) 31.
152. U. Bockelmann, et al., *Biophysical Journal* **82** (2002) 1537-1553.
153. F. T. Chien and J. van Noort, *Current pharmaceutical biotechnology* **10** (2009) 474-485.
154. Y. Cui and C. Bustamante, *Proc. Natl. Acad. Sci. U. S. A.* **97** (2000) 127-132.
155. M. L. Bennink, et al., *Nat Struct Biol* **8** (2001) 606-610.
156. S. Yehoshua, et al., *Biophysical Journal* **108** (2015) 2759-2766.
157. C. Gosse and V. Croquette, *Biophysical Journal* **82** (2002) 3314-3329.
158. C. Haber and D. Wirtz, *Review of Scientific Instruments* **71** (2000) 4561-4570.
159. Y. Seol and K. C. Neuman, *Methods in molecular biology (Clifton, N.J.)* **783** (2011) 265-293.
160. J. K. Fisher, et al., *The Review of scientific instruments* **77** (2006) nihms8302.
161. A. H. B. de Vries, et al., *Biophysical Journal* **88** 2137-2144.
162. M. E. J. Friese, et al., *Nature* **394** (1998) 348-350.
163. Y. L. Lyubchenko, et al., *Methods in molecular biology (Clifton, N.J.)* **1117** (2014) 367-384.
164. Y. L. Lyubchenko and L. S. Shlyakhtenko, *Methods* **47** (2009) 206-213.
165. W. Kalle and P. Strappe, *Micron* **43** (2012) 1224-1231.
166. F. Oesterhelt, et al., *New Journal of Physics* **1** (1999) 6.
167. A. B. Churnside and T. T. Perkins, *FEBS Letters* **588** (2014) 3621-3630.
168. G. M. King, et al., *Nano Letters* **9** (2009) 1451-1456.
169. T. T. Perkins, et al., *Science* **264** (1994) 819.
170. L. Brewer, et al., *Journal of Biological Chemistry* **277** (2002) 38895-38900.
171. R. Roy, et al., *Nat Meth* **5** (2008) 507-516.
172. R. M. Clegg, *Curr Opin Biotechnol* **6** (1995) 103-110.
173. M. T. Woodside, et al., *Science* **314** (2006) 1001-1004.
174. M. T. Woodside, et al., *Proceedings of the National Academy of Sciences* **103** (2006) 6190-6195.
175. C. Monico, et al., *International journal of molecular sciences* **14** (2013) 3961-3992.
176. J. F. Beausang, et al., *Biophysical Journal* **92** (2007) L64-L66.
177. L. Han, et al., in *Mathematics of DNA structure, function and interactions* (Springer, 2009), pp. 123-138.
178. P. C. Nelson, et al., *The Journal of Physical Chemistry B* **110** (2006) 17260-17267.
179. K. B. Towles, et al., *Physical biology* **6** (2009) 025001.
180. F. Vanzi, et al., *Nucleic acids research* **34** (2006) 3409-3420.
181. H. Yin, et al., *Biophysical Journal* **67** (1994) 2468-2478.
182. D. A. Schafer, et al., *Nature* **352** (1991) 444-448.
183. J. F. Beausang, et al., *American Journal of Physics* **75** (2007) 520-523.
184. S. Kumar, et al., *Biophysical Journal* **106** (2014) 399-409.
185. S. Yamada, et al., *Biophysical Journal* **78** (2000) 1736-1747.
186. O. Limanskaya and A. Limanskii, *Gen. Physiol. Biophys* **27** (2008) 322-337.
187. I. Bangs Laboratories, (2016), pp. 2.
188. J. Sambrook and D. W. Russell, *Molecular Cloning*. (Cold Spring Harbor Laboratory, New York, 2001)."Molecular Cloning."
189. A. R. Carter, et al., *Nucleic Acids Res* **44** (2016) 5849-5860.
190. T. T. Perkins, et al., *Biophysical Journal* **86** (2004) 1640-1648.

191. M. Capitanio, et al., *The European Physical Journal B-Condensed Matter and Complex Systems* **46** (2005) 1-8.
192. K. E. Dunn, et al., *Nature* **525** (2015) 82-86.
193. K. Visscher, et al., *IEEE Journal of Selected Topics in Quantum Electronics* **2** (1996) 1066-1076.
194. A. Einstein, (1956).
195. C. S. Epstein and A. J. Mann, *MIT Department of Physics* (2012) 5.
196. J. F. Marko and E. D. Siggia, *Macromolecules* **28** (1995) 8759-8770.
197. M. L. Bennink, et al., *Single Mol* **2** (2001) 91-97.
198. T. T. Perkins, *Laser & Photonics Reviews* **3** (2009) 203-220.
199. J. E. Molloy and M. J. Padgett, *Contemporary Physics* **43** (2002) 241-258.
200. M. Capitanio and Francesco S. Pavone, *Biophysical Journal* **105** (2013) 1293-1303.
201. A. Ashkin, et al., *Opt. Lett.* **11** (1986) 288-290.
202. A. A. Almaqwashi, et al., *Nucleic Acids Research* **44** (2016) 3971-3988.
203. F. Kilchherr, et al., *Science* **353** (2016).
204. K. Sakata-Sogawa, et al., *European biophysics journal* **27** (1998) 55-61.
205. S. Suei, et al., *Biochemical and Biophysical Research Communications* **466** (2015) 226-231.
206. K. C. Neuman, et al., *Biophysical Journal* **77** (1999) 2856-2863.
207. K. Berg-Sørensen and H. Flyvbjerg, *Review of Scientific Instruments* **75** (2004) 594-612.
208. Y. Jun, et al., *Biophysical Journal* **107** (2014) 1474-1484.
209. S. F. Tolić-Nørrelykke, et al., *Review of scientific instruments* **77** (2006) 103101.
210. O. K. Dudko, et al., *Biophysical journal* **92** (2007) 4188-4195.
211. O. K. Dudko, et al., *Proceedings of the National Academy of Sciences* **105** (2008) 15755-15760.
212. G. R. Strobl, *The Physics of Polymers: Concepts for Understanding Their Structures and Behavior*. (Springer Berlin Heidelberg, 2007)."The Physics of Polymers: Concepts for Understanding Their Structures and Behavior."
213. C. Bouchiat, et al., *Biophysical Journal* **76** (1999) 409-413.
214. G. J. Gemmen, et al., *Journal of Molecular Biology* **351** (2005) 89-99.
215. M. Kruithof, et al., *Nat Struct Mol Biol* **16** (2009) 534-540.
216. B. D. Brower-Toland, et al., *Proc. Natl. Acad. Sci. U. S. A.* **99** (2002) 1960-1965.
217. K. D. Whitley, et al., in *Optical Tweezers: Methods and Protocols*, edited by A. Gennerich (Springer New York, New York, NY, 2017), pp. 183-256.
218. J. Larson, et al., *Nat. Protocols* **9** (2014) 2317-2328.
219. I. Heller, et al., *Nat Meth* **10** (2013) 910-916.
220. K. Frykholm, et al., *Small* **10** (2014) 884-887.
221. N. Pires, et al., *Sensors* **14** (2014) 15458.
222. A. Gahlmann and W. E. Moerner, *Nat Rev Micro* **12** (2014) 9-22.
223. B. Schuler and H. Hofmann, *Current opinion in structural biology* **23** (2013) 36-47.
224. T. Xia, et al., *Annual review of physical chemistry* **64** (2013) 459-480.
225. C. Zhang, et al., *Faraday Discussions* **143** (2009) 221-233.
226. S. Zhang, *Nature biotechnology* **21** (2003) 1171-1178.
227. U. Feldkamp and C. M. Niemeyer, *Angewandte Chemie (International ed. in English)* **45** (2006) 1856-1876.
228. S. M. Douglas, et al., *Nature* **459** (2009) 414-418.