Dynamics of Viral Packaging: Single-Molecule Observations in Multiple Dimensions

by

Craig Lee Hetherington

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Physics in the Graduate Division of the University of California, Berkeley

Committee in charge:
Professor Carlos Bustamante, Chair
Professor Jan Liphardt
Professor David E. Wemmer

Fall 2011
Dynamics of Viral Packaging: Single-Molecule Observations in Multiple Dimensions

Copyright © 2011

Craig Lee Hetherington
Abstract

Dynamics of Viral Packaging: Single-Molecule Observations in Multiple Dimensions

by

Craig Lee Hetherington
Doctor of Philosophy in Physics
University of California, Berkeley

Professor Carlos Bustamante, Chair

During the self-assembly of bacteriophage φ29, the viral genome is packaged into a pre-formed capsid by a molecular motor. The packaging motor is a complex of several oligomers including a dodecameric connector ring and a pentameric ATPase ring. These rings coordinate with each other, generating high forces in order to compact the dsDNA genome into the capsid at high pressures.

The connector was proposed to engage and directly perform work on the DNA during packaging. Consideration of the symmetry of the connector and the capsid predicts that the connector must rotate relative to the capsid as part of the mechanism for translocating the DNA. An experiment designed to directly measure rotation of the connector is discussed in Chapter 2 of this dissertation. A combination magnetic tweezers and total internal reflection microscope is used to track the polarization of single fluorescent dyes attached to the connector. No evidence of polarization changes were found, indicating that the connector does not rotate at the expected rates. This further suggests that the connector does not directly perform mechanical work on the DNA during packaging.

Viral packaging can be observed in optical tweezers by monitoring the length of the DNA as it is drawn into the capsid. Past studies have revealed many details of the packaging mechanism by following the dependence of the packaging velocity on factors like ATP concentration and applied load. In Chapter 3, I propose an experimental design intended to measure the effect of the packaging motor on the angle of the DNA in addition to its length and thus recover the full three-dimensional trajectory of the DNA as it passes into the capsid. In addition, this scheme can be used to apply torque and thus provides an additional tool with which to probe the packaging mechanism.
In Chapter 4, we find that during packaging the downstream DNA is twisted in an underwinding direction, with a magnitude that depends on the extent to which the capsid is filled. The change in twisting can be attributed to cumulative looping of the DNA within the capsid, and the data predicts that the loops formed in the last kilobasepair of packaging are as small as 4 nm in radius. In addition, a non-lethal method of rupturing the viral capsids prior to packaging was discovered. Observations of DNA twisting by those packaging complexes revealed that, in the absence of internal pressure, the DNA is twisted by -1.2 °/bp. This number suggests that one of the packaging motor’s five subunits makes contact with the DNA every ten basepairs, and that the cycle repeats with that subunit performing the same function every time. Such a strict functional segregation, in addition to the catalytic segregation revealed in previous high-resolution optical tweezers experiments, is an important part of the mechanism by which the motor packages DNA against high forces.
For mom and dad
## Table of Contents

Chapter 1. Introduction to Single-Molecule Techniques and Viral Packaging........... 1  
1.1 Why biophysics?..................................................................................................................... 1  
1.2 Introduction to viral packaging........................................................................................................... 4  
1.3 Structure and organization of the packaging complex................................................................. 5  
1.4 Functional role of the motor components......................................................................................... 11  
1.5 Mechanism of ATP hydrolysis and force generation........................................................................... 13  
1.6 Optical tweezers experiments on φ29 packaging............................................................................. 16  
1.7 Building a model of translocation..................................................................................................... 21  
1.8 Organization and physics of the genome during and after packaging........................................... 22  
Figures ............................................................................................................................................ 26  

Chapter 2. Experimental Test of Connector Rotation During Packaging...................... 38  
2.1 Abstract................................................................................................................................... 38  
2.2 Introduction............................................................................................................................... 38  
2.3 Results..................................................................................................................................... 40  
2.4 Discussion................................................................................................................................. 45  
2.5 Materials and Methods............................................................................................................... 49  
Figures ............................................................................................................................................ 53  
Appendix......................................................................................................................................... 59  

Chapter 3. Approaches to Single-Molecule Measurement of DNA Angle ............... 70  
3.1 Abstract................................................................................................................................... 70  
3.2 Theoretical considerations in measuring DNA angle ................................................................. 70  
3.3 Measuring DNA angle with optical tweezers........................................................................... 72  
3.4 Combined Optical and Magnetic Manipulation and Observation Tweezers (COMMOT)..................................................................................................................................... 75  
3.5 Development of Novel DNA Modification Strategies I: Peptide Nucleic Acids... 78  
3.6 Development of Novel DNA Modification Strategies II: Terminal labeling with terminal transferase .................................................................................................................................................... 84  
3.7 Use of PNA to study packaging by bacteriophage φ29....................................................... 87  
Figures ........................................................................................................................................... 89  
Appendix......................................................................................................................................... 99
Chapter 4. Observations of DNA Twist During Packaging Reveal Details of Motor Mechanism and the Organization of DNA Within the Capsid................................. 112

4.1 Abstract ................................................................................................................................. 112
4.2 Introduction .......................................................................................................................... 112
4.3 Materials and Methods ................................................................................................. 113
4.4 Results ................................................................................................................................... 115
4.5 Discussion ............................................................................................................................ 118
4.6 Conclusions ......................................................................................................................... 125
Figures .......................................................................................................................................... 127
Appendix ....................................................................................................................................... 138

Chapter 5. Afterword .................................................................................................................... 152
References ........................................................................................................................................ 156
Preface

I first conceived of the experiments presented in this dissertation in February 2005, at the Biophysical Society annual meeting, upon seeing electron micrographs of DNA packaged into phage heads. Little did I know that I would have the opportunity to execute those ideas over the next six years, and little did I know that others in our lab had already been working on them. While the data and analyses shown here are my own, this work took place in a rich and supportive environment in the Bustamante lab. I am deeply grateful to everyone who has contributed to this work, either intentionally or inadvertently, as well as to others who have touched my life during the past eight years.

First, I thank my advisor, Carlos Bustamante, who gave me the freedom to do what I thought best even when he disagreed, and who never stopped believing that these projects would eventually bear fruit even as I lost hope. His “Latin passion” has been critical in pushing these experiments to the end. And his sharp and unsatisfied mind is my model for how a scientist should think.

The bona fide φ29 virologists, Shelley Grimes and Paul Jardine of the University of Minnesota, were constantly supportive and enthusiastic about this work. At a moment’s notice they provided reagents, wisdom, and humor. The concept and realization of this work was developed in close collaboration with them. I will miss our conversations.

When I saw those first EM images in 2005, I immediately called Aathavan Karunakaran, the old man of the phage. He patiently listened as I explained exactly how I would measure the spooling of DNA inside the capsid, and then he told me that it was a great idea, and that he was working on it already. Credit and blame for almost everything I know about molecular biology and φ29 in particular is due to Aathi. He also taught me much about life, and the importance of occasionally pushing its boundaries.

Jeff Moffitt, the eye of the tiger, was a constant companion and champion in the φ29 quest. In those countless hours of conversation, I learned that it’s OK to be wrong, but it’s always wrong to stop. Thorsten Hugel introduced me to φ29 and single-molecule fluorescence. In turn, I like to think that I introduced Gheorghe Chistol and Shixin Liu to the system. They are a great team that will surely discover even more mysteries of viral packaging. The entire φ29 crew should be acknowledged: Adam Politzer, Ariel Kaplan, Yann Chemla, and Marta Kopaczynska. I had the pleasure of working with Tae-Hyung Chris Lee and his baseball cap during his undergrad years. Jörg Schnitzbauer brought some German efficiency and dance to the phage group. Other transient companions were Dylan Reid, Marissa Lee, Anna Labno, Jascha Sohl-Dickstein, Jeff Tang, Lea Witkowski, and Antony Lee.

Yara Mejia’s surefooted balance helped keep me on the right path, scientifically and spiritually, throughout all of grad school. Arthur Edelstein constantly challenged me scientifically and taught me that the most productive times are often those days you don’t do any real work.
Many other labmates have contributed scientifically and otherwise. In no particular order, thanks are due to Jeff Gore, Shirley Mihardja, Sophie Dumont, Wei Cheng, Jay Quinlan, Jens Michaelis, Steve Smith, Seok-Cheol Hong, Oren Levy, Marcelo Nollmann, Paul Pease, Maumita Mandal, Courtney Hodges, Lacra Bintu, Ryan Case, Errol Watson, Rodrigo Maillard, Christian Kaiser, Jae Yen Shin, Yuanbo Cui, Ninning Liu, Tingting Liu, Toyotaka Ishibashi, Pim Dangkulwanich, Maya Sen, Dan Goldman, Jesse Dill, Phillip Elms, Melissa Strycharska, Killian Strycharska, Daniel Guerra, Claudio Rivetti. In particular, outdoor adventures with Jay, Jae Yen, Rodrigo, Aathi, and Courtney were crucial for maintaining sanity.

Luis Comolli and Andy Spakowitz had numerous discussions with me regarding the arrangement of DNA within the capsid. Marc Morais shared cryo-EM data on the packaging motor. Rockney Atz provided bioinformatics insight into the ATPase. Wayne Falk helped with fluorescence bulk experiments. Dwight Anderson discovered φ29. George Oster and Jin Yu shared their spectacular ideas about the microscopic mechanism of DNA rotation and always provoked debate. Doug Smith and James Tsay shared their data on φ29 and lambda packaging. Hagar Zohar introduced me to PNA, provided materials, and collaborated on the PNA work, and started a journal club. Josh Kittleson engineered plasmids and provided cell strains. Stephen Morris discussed the subtleties of fluid mechanics and reassured me regarding the validity of the drag assumptions used.

My many housemates throughout the years have put up with my whining: Drew Carlson, Dave Gorin, Zack Subin, David Hembry, Josh Kittleson, Chuck Krois, Marcus Carr.

One should try to attend as many talks as possible, not only because one might learn something. I had the good fortune to meet Shannon Yan at one such lecture. Grad school would not have been the same without her presence. It’s time to repay some of the confidence and faith that I have borrowed the past few years.

Finally, I would like to thank my family. My brothers, Kevin and Forrest, have grown up to be role models for me, somehow. My parents, Ann Lee Hetherington and William Morley Hetherington, gave me the curiosity to want to discover the small things that make up the big things. They have supported me in everything I have done. I can only hope that I inherit the same inner strength that has made it possible for them to accept my long path through school.
Chapter 1. Introduction to Single-Molecule Techniques and Viral Packaging

1.1 Why biophysics?

The application of physical concepts and techniques to the traditional disciplines of biochemistry, molecular biology, and cellular biology is transforming our understanding of the processes that sustain life. Just as the invention of the optical microscope allowed Robert Hooke to discern the spatial organization of tissues and discover the cell, so have modern tools like force spectroscopy and fluorescence microscopy enabled us to look within the cell and study the organization, mechanics, and dynamics of its constituent parts. Simple consideration of thermal fluctuations reveals that the cell, its organelles, and even the catalytic foci within cannot efficiently operate as diffusion-limited reaction vessels. Rather, the cell resembles a factory, containing a coordinated network of specialized reaction centers, each executing specific tasks, connected by well-defined transport pathways. In this emerging picture of the cell, mechanical processes that spatially and temporally coordinate biochemical reactions are ubiquitous and essential. The agents of these mechanical processes are called molecular motors.

Molecular motors

As many of the fundamental building blocks of the cell, such as DNA, RNA, proteins, and saccharides – as well as the superstructures formed by assembly of those components – take the form of polymers, a basic mechanical phenomenon is directed motion along a periodic track. Such translocation is essential for a wide variety of functions both in and out of the cell, ranging from exertion of extracellular force to maintenance and remodeling of cellular structure to synthesis, processing, and dismantling of DNA, RNA, and proteins.

Processive and directed motion requires energy. Because the track is periodic, the system’s enthalpy is symmetric under unit translations, so the energetic cost for directed motion is, in principle, bounded only by entropy. A lower bound on this energy input can be estimated by considering the entropic cost of unidirectional motion as opposed to a random walk. Given two allowed steps – forward or backward – each non-diffusive forward step along a one-dimensional periodic track by a motor reduces the entropy of the system by $k_B T \cdot \ln(2)$. So what is the thermal energy on the scale of a biological molecular motor? At room temperature, $k_B T$ is $4.14 \cdot 10^{-21}$ J. Typical proteins are on the order of 1-10 nm in length, so a convenient system of units involves nanometers and picoNewtons. This also provides a clue as to the typical interaction strength between a molecular motor and its track – given
a 1 nm periodicity to the track, the motor must be able to grip its track with forces in excess of 4 pN in order to preclude Brownian motion. In fact, individual molecular motors do typically exert forces in the range of 1-100 pN. One might expect that the basic unit of energy exchange for molecular motors, providing the energy for a single step, would be on the order of 10 \( k_B T \), and this too is true. Biological motors are powered by the hydrolysis of a phosphate bond, a reaction which, under typical conditions, liberates 10-40 \( k_B T \).

**Single-molecule tools**

The task of the molecular biophysicist, then, is to measure forces in the pN range and distances in the nanometer range. Recent astounding advances in optical imaging far beyond the diffraction limit (Bates et al., 2007; Betzig et al., 2006; Donnert et al., 2006; Yildiz et al., 2003) have reached a resolution of about a few nanometers. Imaging with X-rays or electrons can reach higher resolution still but is destructive to the biological sample, rendering truly time-resolved measurements impossible. There exist a number of other, non-imaging-based biophysical methods for inferring spatial information at the nanometer scale, such as nuclear magnetic resonance and X-ray crystallography, as well as biochemical assays. These methods, though powerful, suffer a common limitation in that they are ensemble measurements of the average of a large (typically >10^6) number of individual molecules. Such measurements miss the full variation in the properties of the object and also, because dynamics on the nanoscale are stochastic, cannot provide much time-resolved information.

One solution to these inadequacies has come, in the past decade or two, in the form of single-molecule measurement and manipulation. Tracking of the dynamics of individual particles allows the full distribution of temporal and spatial behavior to be measured, hence revealing new and important details about the underlying mechanisms. In turn, manipulation of single molecules provides an additional dimension of experimental control which can expose new dynamics. Of course, because these experiments measure single particles, they will have significant and essential limitations – in particular, a low signal-to-noise ratio and low experimental throughput. These two unavoidable facts make it difficult to measure full distributions of properties, and great care must be taken in selecting the proper experimental technique and protocol for a specific biological question. The two techniques that are discussed in this dissertation are single-molecule fluorescence microscopy and optical tweezers.

In single-molecule fluorescence measurements, a biological object of interest is labeled by a single (or small number of) fluorophores that absorb light and fluorescently emit light in a manner that may be dependent on the local
environment. As will be discussed in chapter 2, the polarization of emitted light provides information about the orientation of the dye’s electric dipole moment (Adachi et al., 2000). Another example is that of FRET, or Forster Resonant Energy Transfer: if two fluorescent moieties are within a few nanometers of each other, energy can be transferred from one fluorophore to the other with an efficiency that depends sensitively on the inter-dye distance, hence revealing very small conformational changes in a biological nanostructure (Stryer, 1978). Recent experiments have utilized single-molecule fluorescence to study a myriad of biological questions ranging from the internal conformation of a single biomolecule up to the organization and mechanics of whole cells, and even between cells.

Optical tweezers were originally conceived as a means of manipulating single particles (Ashkin, 1970), but they too can provide precise spatial information at the nanometer scale. The principle of optical tweezers is simply that a small dielectric particle, such as a micron-sized polystyrene sphere, can be trapped by a focused spot of light (Moffitt et al., 2008; Svoboda and Block, 1994). The position of that trapped particle can then be controlled in three dimensions by moving the focus. In the ray-optics description, the sphere acts as a lens and deflects off-axis light further away from the optical axis. By conservation of momentum, the bead must feel an impulse toward the optical axis, hence being trapped at the focus. The scattered light contains information on the position of the sphere and the forces exerted. However, for particles small enough to be trapped – namely, particles of the same size as the wavelength of trapping light – the ray-optics equations in the so-called Mie regime are rather difficult to solve, making this explanation unsatisfying. The alternative explanation is that the bead material is polarizable, being a dielectric, and it hence feels an attractive interaction with an electric field. Typically, the attractive force is proportional to the bead’s distance from the light focus. Optical traps can exert and measure forces in the picoNewton range and measure the position of the bead to an accuracy of picometers. However, because the trapped particle diffuses within the potential well formed by the focal spot, the resolution at biologically relevant timescales is usually at the nanometer scale.

In order to measure a biological sample, the sample must be “tethered” between two optically trapped beads, or between one optically trapped bead and a fixed object. Force is applied in the form of tension on that tether, and the experiment is designed such that the signal of interest takes the form of a change in tension or a change in the extension of the tether. A molecular motor that moves along DNA, for example, would cause the length of a tether to change if the DNA is attached to one bead and the motor to the other bead. For this reason, optical tweezers are exquisitely well-suited for the study of molecular motors that move along DNA and other polymer tracks.
DNA translocation

A sample of the many known DNA-processing phenomena in the prokaryotic cell reveals the ubiquity of directed motion mediated by molecular motors. When DNA is replicated prior to cell division, the bits of information on one strand are read and copied, one by one, directionally and processively, by a molecular motor called DNA polymerase (Wuite et al., 2000), with the assistance of various strand-separating motors called helicases (Dumont et al., 2006). After replication, the two DNA copies must be segregated into daughter cells, a task performed by a DNA pump called FtsK (Pease et al., 2005). The information in DNA is processively transcribed into RNA by RNA polymerase (Davenport et al., 2000), and the transcription process can be terminated by another motor, Rho helicase, which walks up the nascent RNA transcript and deactivates the elongating complex (Thomsen and Berger, 2009). Because replication and transcription tend to twist DNA, other motors like topoisomerases twist the DNA back into the correct form (Stone et al., 2003). If the DNA is found to be defective, it can be repaired by the homologous recombination system, which aligns two nearly-identical sequences and exchanges strands in order to facilitate error correction, utilizing the RecBCD and RuvAB motor complexes (Galletto et al., 2006). DNA is also continually patrolled by Type I and Type III restriction enzymes like EcoR124I which walk along it, looking for chemical signs of viral invasion (Stanley et al., 2006).

1.2 Introduction to viral packaging

Viruses, which are the most abundant biological entities on the planet, utilize molecular motors, both encoded by the virus and hijacked from the host cell. A motor of particular interest is the viral packaging motor, which plays an essential role in viral self-assembly. Some viruses, including the majority of prokaryotic viruses (Calendar and Abedon, 2005), first assemble a purely proteinaceous capsid, devoid of nucleic acid, and then physically pump their genetic material into this empty shell. This energetically unfavorable task is accomplished by a remarkable molecular machine called the packaging motor, which is comprised of several components that coordinate the conversion of chemical energy from nucleoside triphosphates (NTPs) into the needed mechanical work. Among the strongest and most efficient molecular motors known, the packaging motors have been subjected to four decades of research, and in the past few years our single-molecule experiments have greatly extended our understanding of principles underlying the mechanism of these complex motors (reviewed in (Catalano, 2005; Grimes et al., 2002; Jardine and Anderson, 2006; Kainov et al., 2006; Rao and Feiss, 2008; Sun et al., 2008)).
Bacteriophage φ29 is a model viral packaging system that has been studied to unprecedented detail with optical tweezers and other single-molecule methods (Aathavan et al., 2009; Chemla et al., 2005; Comolli et al., 2008; Grimes et al., 2002; Hugel et al., 2007; Jardine and Anderson, 2006; Moffitt et al., 2009; Rickgauer et al., 2008; Smith et al., 2001; Yu et al., 2010). See Figure 1 for a schematic of the viral assembly process. The mature viral particle contains a 19.3-kilobasepair double-stranded DNA genome (with contour length of about 6.5 µm and a persistence length of 50 nm) within a shell, or “capsid,” less than 50 nm on a side. The virus injects its genome into a target cell in the first step of its life cycle. In the interior of an infected cell, the viral genome is both replicated and expressed, leading to the production of multiple copies of the viral genome and viral “proheads.” The proheads consist of an empty capsid and a unique dodecameric structure that encloses a channel into the capsid, called the “portal” or “connector.” At initiation of packaging, a viral genome is specifically recognized by the packaging machinery, via terminal protein covalently linked to the DNA, and associates with a capsid. The packaging motor then translocates the DNA into the capsid while ATP is hydrolyzed by a molecular motor. The packaging complex consists of several obligate coaxial rings which encircle and possibly engage the DNA: a ring of ATPase proteins, called the “packaging ATPase”; the dodecameric connector; and a pentameric ring of RNA molecules called the “pRNA” that bridges the ATPase to the connector (Figure 2). At the conclusion of packaging, the ATPase and pRNA rings dissociate from the prohead and are replaced by a “tail,” which stabilizes the complete and mature particle and later serves to inject the genome into the next target cell.

1.3 Structure and organization of the packaging complex

Φ29 is not particularly unique among the bacteriophages. Different tailed bacteriophage species, with little or no sequence identity and targeting highly different hosts, employ functionally and structurally similar macromolecular tools to translocate dsDNA (reviewed in (Catalano, 2005; Jardine and Anderson, 2006). The universal elements of this toolkit – the ATPase and the connector/portal – appear to be modular, having been assorted independently according to phylogenetic studies (Burroughs et al., 2007). Hence it is likely that determination of the structure or mechanism of one component in one viral species will be highly relevant for its role in other species, and synthesizing a proposed mechanism for DNA packaging requires integrating information from many different viral systems. And although the φ29 family is the only one that features an RNA component in the packaging motor, it is likely that the functional role of that RNA ring will be filled by some other, proteinaceous, component in the other viral systems.

ATPase ring

ATPase ring
The packaging ATPase forms an oligomer that associates with the DNA and with the prohead (Figure 2). The DNA is then translocated into the capsid while adenosine triphosphate (ATP), are hydrolyzed into adenosine diphosphate (ADP) and inorganic phosphate (Pi). Sequence analysis has shown that all the tailed bacteriophage packaging ATPases belong to the Additional Strand Conserved E (ASCE) division of ATP-hydrolyzing proteins, a large protein group responsible for a wide variety of cellular activities (Burroughs et al., 2007; Iyer et al., 2004a; Iyer et al., 2004b). The relationships between the packaging motors and other ATPases are shown in Figure 3. The ASCE core is characterized by its secondary structure – a sheet formed of alternating beta strands and alpha helices – and by the amino acids found in certain highly conserved motifs such as Walker A, Walker B, catalytic glutamate, sensor, and arginine finger, all of which are required for the ATP binding and hydrolysis cycle. The details of these and other motifs in the packaging motors will be discussed below in the context of the DNA translocation mechanism. In terminase viruses – a family somewhat distinct from φ29 – the terminase large subunit is the packaging ATPase; it is composed of two domains – an N-terminal ATPase domain, and a C-terminal nuclease domain (Burroughs et al., 2007; Kanamaru et al., 2004). X-ray crystallography of the T4 terminase large subunit has confirmed the presence of a RecA-like canonical ATP-binding fold but indicates some degree of divergence from the other RecA-like proteins (Sun et al., 2008; Sun et al., 2007). In contrast, the φ29-like bacteriophages – as well as some eukaryotic viruses like the poxviruses – utilize an ATPase closely related to the HerA/FtsK family of dsDNA translocases (Burroughs et al., 2007). The sequence homology of the viral ATPases to known molecular motors strongly suggests that these ATPases are also responsible for DNA translocation.

Packaging ATPases, like many ASCE proteins, take the form of an oligomeric ring in the packaging complex (Figure 2). The exact oligomeric state may vary between phage types: five subunits are visible in cryo-electron microscopy 3D reconstruction (cryo-EM) of φ29 (Morais et al., 2008) and T4 (Sun et al., 2008), whereas λ may have four subunits (Maluf et al., 2006). The pentameric packaging motor of φ29 has a different oligomeric state than its hexameric cousins in the HerA/FtsK family, further suggesting that the exact number is not a conserved feature required for DNA translocase activity. Circular geometry may have important functional consequences such as enabling communication and coordination between the subunits.

pRNA ring

The φ29-like bacteriophages contain a unique factor not yet found in other packaging complexes: the prohead RNA, or pRNA (reviewed in (Grimes et al., 2002; Guo, 2002)). pRNA originates as a 174-base transcript from the left end of the viral genome. After transcription, the pRNA oligomerizes on the
prohead (Guo et al., 1991b) in the final step of morphogenesis, with no requirement for the presence of the ATPase. According to biochemical and cryo-EM studies, the pRNA oligomer is positioned as a bridge between the ATPase ring and the capsid (Figure 2) (Koti et al., 2008; Morais et al., 2008; Simpson et al., 2000; Sun et al., 2006; Xiao et al., 2008). Functionally, pRNA helps to initiate packaging on the left end of the DNA (Grimes and Anderson, 1989; Wichitwechkarn et al., 1989), stimulates hydrolysis by the ATPase (Grimes and Anderson, 1989), and is required for assembly and activity of the packaging motor. After the completion of packaging, the pRNA, along with the associated ATPase, is released as the tail components are added to complete the mature viral particle (Chen and Guo, 1997).

The positioning of the pRNA relative to the capsid and the ATPase suggests that it plays a structural role in the complex. This hypothesis is borne out by cryo-EM reconstructions which show that the major helix of the pRNA extends down from the connector, interdigitating with and grasping the subunits of the ATPase pentamer (Morais et al., 2008). Furthermore, sequence mutations to the helix that do not change its geometry have little effect on packaging, unless they simultaneously abolish ATPase binding. Finally, mutations which change the angle of the helices can modulate packaging efficiency (Zhang et al., 1997) (Zhao et al., 2008a). However, the pRNA may not be a static scaffold; some experiments suggest that flexibility at specific points in the pRNA is crucial for packaging (Grimes et al., 2002; Guo, 2002). This flexibility may be required by some models of DNA translocation, to be described below, which involve large (nanometer-scale) conformational changes of the ATPase ring (Sun et al., 2008) and possibly the connector (Simpson et al., 2000) as they engage and move the DNA. The pRNA, providing structural support for the ATPase, may itself need to adjust and reorient in order to accommodate or even induce such conformational changes.

As it is likely that the pRNA plays an important role in organizing and positioning the ATPase, much effort has been devoted to determining the organization of the pRNA oligomer (reviewed in (Grimes et al., 2002; Guo, 2002)). A set of clever experiments utilized the fact that compensatory mutations can be made in the intermolecular pseudoknots that bind adjacent pRNA subunits, in order to generate hetero-oligomeric rings of pRNA. It was determined that pRNA mutants that cannot form dimers also cannot support packaging; likewise, pRNA incapable of trimerization also failed to support packaging, indicating that the relevant oligomeric state is a multiple of two and three. These and related experiments using analytical centrifugation suggest that the pRNA assembles as a hexamer. A recent single-molecule fluorescence experiment provided additional evidence in support of the hexamer model (Shu et al., 2007). Stalled complexes
containing Cy3-tagged pRNA were examined in TIRF microscopy and identified as fluorescent spots; the number of RNA molecules at a spot was determined by counting the number of photobleaching steps. When the finite labeling efficiency was considered, the data were best fit by a binomial model in which six pRNA molecules are bound to a prohead.

In contrast, the most recent and advanced cryo-EM studies of proheads reveal five pRNA subunits symmetrically arranged in a ring, bound to the prohead (Simpson et al., 2000). Imaging of highly symmetric particles like viruses can contain artifacts due to the computational imposition of symmetry during the averaging process. However, new computational and biochemical techniques introduced to control for these artifacts have confirmed that the pRNA density exhibits fivefold symmetry (Morais et al., 2005; Morais et al., 2008; Morais et al., 2001; Zhao et al., 2008b). One potential solution to the discrepancy between these experimental approaches is that six pRNA molecules are present, but only five are symmetrically organized in the packaging complex. The most recent single-molecule fluorescence experiments provide circumstantial evidence of this hypothesis: it was found that pRNA incapable of supporting packaging nonetheless can bind to proheads with weak affinity, yielding one or zero photobleaching steps (Xiao et al., 2008) thus suggesting a systematic overestimate of the pRNA number. Nonspecific binding of one pRNA to the prohead is detectable with fluorescence but probably not with averaged cryo-EM reconstruction. Moreover, if an additional pRNA molecule is nonspecifically bound to a prohead, as we suggest, it may not be functionally relevant and may not contribute to the organization of the ATPase oligomer.

Connector

The DNA enters the capsid via a channel called either the portal or the connector. Both names reflect the multifaceted role of this protein oligomer: it is a passageway, structural platform (Simpson et al., 2000), and signal transducer (Oliveira et al., 2006), and it has been proposed to interact directly with the DNA during translocation (Hendrix, 1978; Oliveira et al., 2006), among other functions (reviewed in (Rao, 2009; Valpuesta and Carrascosa, 1994)). Across all known packaging dsDNA bacteriophage families, the portal is a dodecameric ring located at a unique fivefold vertex of the icosahedral capsid (Agirrezabala et al., 2005; Guasch et al., 2002; Lander et al., 2006; Lebedev et al., 2007; Simpson et al., 2000). The crystal structure of the SPP1 portal (Lebedev et al., 2007) is the most complete structure available, and the crystal structure of the φ29 connector has revealed nearly identical folds despite little sequence similarity (Simpson et al., 2000) (Guasch et al., 2002). The overall shape of the connector is a funnel, with a wide aperture nested within the capsid and a narrow end that extends outside (Figure 4a). At the wide end are SH3 domains likely
responsible for oligomerization and contact with the capsid (the wing). Also within the capsid, but closer to the central pore, is the crown domain thought to contact DNA when packaging is complete and, in some phages, aid in the termination of packaging. The central domain of the connector consists of multiple alpha-helices that run along its length (the stem), forming a near-continuous barrel. The external domain of the connector, called the clip domain, appears to be the assembly point for the rest of the packaging motor (Atz et al., 2007; Morais et al., 2008).

At its narrowest point, the inner diameter of the crystal structure of SPP1 is 27 Å (Ångstroms). However, the portal crystallized as a tridecamer, which differs from the known dodecameric state of the prohead-bound portal. Cryo-electron microscopy and subsequent modeling of the high-resolution structure predicts that in the prohead the tunnel inner diameter is as little as 18 Ås – too small for DNA to pass. The tunnel is constricted by the “tunnel loop,” a flexible but structured region that contains a conserved acidic residue (Figure 4b). In addition to its flexibility, the tunnel loop may be mobile because it is at the end of an alpha-helical lever. Based on these two facts, it was suggested that the loops can move in a coordinated manner during packaging, hence permitting DNA passage and perhaps actively driving the DNA (Lebedev et al., 2007). Even allowing for significant tunnel loop flexibility, the small size of the opening seems to be at odds with recent observations that φ29 can package both loops (Rickgauer et al., 2006) and bulges (Aathavan et al., 2009), as can λ (Pearson, 1988), and possibly also T4 (Ray et al., 2010). Since packaging of those DNA structures may require threading four ssDNA strands through the channel, it seems likely that the connector as a whole must be capable of considerable conformational flexibility. We will discuss the functional implications of these experiments in the next section.

Once packaging is complete, the connector plays a crucial role in stabilizing the pressurized particle. When the end of the DNA has been reached, the motor (in the presence of nucleotide) continues to hold onto the DNA and the capsid until a tail complex displaces it (Chemla et al., 2005). While the pRNA and ATPase are dissociating from the capsid, it is thought that the connector prevents the DNA from being ejected. After the tail assembles and binds, a short length of DNA slips out from the capsid into the tail (Tang et al., 2008), where the terminal protein is bound by the tail, yielding a stable, mature particle.

In some viral species, the connector even plays a crucial role in terminating DNA translocation when the head is filled. These phages – for example T4, P22, and SPP1 – package approximately, but not exactly, one genome length of DNA before cleavage. The exact length of DNA packaged varies from particle to particle, with the average packaging length dependent on the size
of the capsid (Casjens and Weigele, 2005), suggesting that the packaging complex somehow recognizes that the capsid is full before cleaving. This phenomenon is called headful termination. If the headful mechanism is not activated, multiple DNA molecules can be packaged by T4 (Leffers and Rao, 1996). Mutational screens found that, in P22 and SPP1, the portal is responsible for headful sensing. Single-residue substitutions in the portal can increase or decrease packaging length (Casjens et al., 1992) (Tavares et al., 1992) and the effect of mutations seems to be additive. A structural basis for headful sensing has become apparent in recent years. Comparison of 18 Å structures of isolated wild-type and headful-defective portals of SPP1 showed significant differences in the crown domain (Orlova et al., 1999). The recent cryo-EM reconstruction of the portal of mature P22 particles (Lander et al., 2009; Lander et al., 2006) clearly shows that the portal makes contact with the encapsidated DNA, via those same crown domains. Moreover, comparison of this structure with the free portal reveals major structural differences in the crown region – differences that are attributed to the interaction with the highly pressurized DNA (Figure 4c). Thus the portal crown domain is likely to be the headful sensor; indeed, φ29, which has no documented headful cleavage, has a significantly reduced crown domain compared to P22 (Xiang et al., 2006). Because the distal end of the portal is positioned to contact the terminase, it was proposed that conformational changes in the connector crown may propagate through the protein and ultimately to the terminase nuclease domain in order to actuate the cleavage. This hypothesis is corroborated by the fact that a specific mutation to the ATPase-binding clip domain – not the headful-sensing crown – also changes the length of packaged DNA (Oliveira et al., 2006). Thus, the portal is intimately involved in all stages of viral assembly, packaging, and infection (Isidro et al., 2004a).

**Capsid**

After passing through the portal, the DNA reaches its destination, the capsid. As it defines the volume into which the genome must be compressed, the capsid is an important determinant of packaging behavior. The dsDNA phages share a basic icosahedral shape, typically about 50 nanometers across, that can vary in radius and length depending on species (reviewed in (Hendrix, 2005; Reddy and Johnson, 2005)). Located at one fivefold vertex of the head is the dodecameric connector (Figure 2) through which DNA passes during packaging. Between the initial capsid assembly and the completion of packaging, the capsid undergoes significant structural rearrangement involving processes such as proteolysis, (Conway et al., 1995) crosslinking (Lata et al., 2000), departure of the scaffold (Morais et al., 2003), and volume expansion (Conway et al., 1995), with different bacteriophages displaying different combinations of these maturation events. Some of these
rarrangements occur during and are, perhaps, triggered by packaging itself. For example, the scaffold proteins of φ29 exit during packaging (Bjornsti et al., 1983), and many phages undergo expansion during packaging (Feiss and Catalano, 2005; Serwer, 2005).

1.4 Functional role of the motor components

Because packaging requires the involvement of all the components of the complex – the capsid, connector, and ATPase – biochemical experiments have not conclusively identified which component is responsible for performing work on the DNA as part of packaging. The ATPase is responsible for ATP binding and hydrolysis; it may directly drive the DNA, or it may induce conformational changes in the connector/portal which in turn engage and move the substrate.

The connector regulates the ATPase and is intimately involved in translocation

The connector is intimately involved in the packaging process: it encircles the DNA during packaging; it binds DNA; it modifies the hydrolysis rate of the ATPase; and it is responsible for headful sensing in packaging termination (Lander et al., 2006). The position and organization of the connector are clearly compatible with DNA engagement, and connectors of packaged T4, T7, and φ29 particles were shown to bind about 40 bp of a linear DNA molecule in footprinting assays of mature phage, proving a DNA-binding capability, albeit in a different part of the life cycle (Zachary and Black, 1992). A possible structural basis for this interaction was revealed by crystal structures of the SPP1 portal, which showed that the most constricted part of the channel – 18 Å measured perpendicular to the DNA axis – is defined by the “tunnel loops,” 16-residue interhelical regions whose positioning depends on the bending of an adjacent alpha helix (Lebedev et al., 2007) (Figure 4b). In order to explore the nature of the connector-DNA interaction, the effect of portal mutations on packaging and hydrolysis by SPP1 proheads was studied (Isidro et al., 2004b; Oliveira et al., 2010; Oliveira et al., 2006). Three intriguing mutations were examined in great detail: two in the clip domain responsible for ATPase binding, and one located in the tunnel loop predicted to contact the DNA. None of the mutations were predicted to change secondary structure, and indeed none had an effect on prohead formation or ATPase binding. In contrast, all three mutations significantly reduced both ATP hydrolysis by the ATPase and the efficiency of packaging. This apparently regulatory function was interpreted as “cross-talk” between the portal and the ATPase: the portal may assume one of several states, and that state has an effect on the ATPase. Moreover, the location of the tunnel loops involved in this coupling suggests that the DNA is also a participant in the communication.
Following up on these experiments was the discovery that flexibility in the portal is essential for packaging (Cuervo et al., 2007). Using the crystal structure as a guide, cysteines were introduced into the stem helices of the SPP1 portal in order to permit reversible crosslinking of neighboring subunits. In oxidizing conditions the portals formed covalently linked high-molecular weight oligomers. Interestingly, proheads with crosslinked portals are incapable of fully packaging DNA in a stable manner; however, packaging could be restored simply by exposing the complexes to reducing conditions, proving that removal of the crosslinks is sufficient for normal activity. Abrogation of packaging was not simply due to blocking of the channel, as particles crosslinked after maturation exhibited normal ejection kinetics. Hence it was proposed that crosslinking inhibits conformational flexibility in the portal that is essential for packaging itself. Integrating the conclusions of the crosslinking and mutational studies, it is clear that connector conformation and dynamics can affect both packaging and ATP hydrolysis. However, there is no evidence to date of communication in the opposite direction – no evidence that the ATPase can regulate, stimulate, or actuate the connector. This kind of communication would be essential in any kind of ATP-fueled packaging by the connector.

As early as 1978 – before the symmetry state of the connector was known – Hendrix, in a remarkable chain of inductive logic, proposed that the connector exhibits sixfold symmetry, that the connector engages the DNA, and that the symmetry mismatch between the connector and capsid could allow the connector to rotate while translocating DNA into the capsid (Hendrix, 1978). The first two of these hypotheses were substantiated, and the third idea, unproven, became a paradigm of viral packaging, spawning a number of more detailed proposals (Dube et al., 1993; Guasch et al., 2002; Lebedev et al., 2007; Nummela and Andricioaei, 2009; Simpson et al., 2000; Valpuesta and Carrascosa, 1994). If the connector is, in fact, the DNA-engaging part of the motor, simple symmetry considerations make strong predictions as to the rotation of the connector relative to the DNA during packaging. This hypothesis is directly tested in Chapter 2 of this dissertation.

The ATPase hydrolyzes ATP and likely binds the DNA

There is no direct evidence that purified ATPase protein induces translocation of DNA in the absence of the prohead. All packaging motors exhibit ATPase activity that is stimulated by DNA, and hydrolysis is maximized in the presence of all packaging complex components (Kondabagil et al., 2006) (Baumann and Black, 2003; Gual et al., 2000; Kondabagil et al., 2006; Mitchell et al., 2002; Oliveira et al., 2006). DNA binding has been documented in many packaging ATPases, although this could be mediated through the endonuclease domain present in the terminases. The terminase
of λ has been credited with helicase activity but the strand separation is not processive nor clearly dependent on ATP (Feiss and Catalano, 2005; Yang et al., 2009), so this activity does not imply that the terminase is a self-sufficient translocase. The most intriguing evidence of packaging-related binding comes from the T4 system: double-stranded-specific binding could be localized to the ATPase domain (Alam and Rao, 2008), and there is evidence of a correlation between the state of the binding site and the state of the ATP-binding pocket. There are two possible functions of this DNA binding ability – engaging the DNA during translocation and/or facilitating cleavage by the adjacent nuclease domain – both of which are compatible with the data. The ϕ29 ATPase, which does not have a nuclease domain, also binds DNA and induces supercoiling of DNA loops, but the ATP dependence of this process is not clear (Grimes and Anderson, 1997), so this does not prove that the ATPase directly pushes on the DNA.

The dsDNA packaging ATPases fall within a larger class of proteins, comprising 5-10% of the predicted gene products (including G-proteins, some kinases, and some cytoskeletal proteins, among others) among all genomes, termed the P-loop NTPases because they contain a conserved phosphate-binding loop (Iyer et al., 2004a). The ASCE division itself is highly diverse, including proteins that transport a wide variety of nucleic acid, protein, and small-molecule substrates (Iyer et al., 2004a; Iyer et al., 2004b). Sensitive bioinformatics studies have found that about 15% of the packaging phages, including ϕ29, possess an ATPase of the FtsK/HerA clade of dsDNA translocases (Fig. 3) (Burroughs et al., 2007). FtsK, the clade prototype, is a homohexameric ring motor responsible for chromosomal segregation during bacterial cell division (Massey et al., 2006). so it seems likely that these phage ATPases perform a similar DNA pumping function in viral replication.

1.5 Mechanism of ATP hydrolysis and force generation

ATP Binding

The ASCE family shares a common core secondary structure consisting of five parallel beta strands. The loops at the ends of the beta strands bear residues that are crucial for ATP binding and hydrolysis, such as the phosphate-binding Walker-A motif, the metal-binding Walker-B motif, a catalytic acidic group, a polar Sensor or C-motif or motif III, and an arginine finger (reviewed in (Enemark and Joshua-Tor, 2008; Erzberger and Berger, 2006; Rao and Feiss, 2008; Thomsen and Berger, 2008)). Typically, nucleotide is bound at the interface between two subunits, in a groove formed by the arginine finger of one subunit and the other critical residues from the partner subunit. ATP hydrolysis begins with nucleotide and metal
ion positioning, followed by nucleophilic attack of a water molecule on the gamma-phosphate, guided by the catalytic acid. The sensor or C-motif is thought to sense the presence of nucleotide and couple nucleotide binding to conformational changes elsewhere in the protein; likewise, the arginine finger may communicate a change in nucleotide state to the adjacent subunit. In addition to a host of mutational studies, the T4 terminase crystal structure – the only extant structure of a dsDNA viral packaging ATPase – confirms this general structure for the ATP-binding site, although the arginine finger comes from the same subunit (Sun et al., 2008; Sun et al., 2007).

Global conformational changes

Local conformational changes associated with ATP hydrolysis must propagate through the protein to the DNA. Several different mechanisms for coupling hydrolysis to translocation have been proposed based on crystal structures of various nucleic acid translocases (reviewed in (Enemark and Joshua-Tor, 2008; Rao and Feiss, 2008; Singleton et al., 2007)). The T4 terminase was crystallized with very close contacts between the ATPase domain and the DNA-binding nuclease domain whereas in cryo-EM the domains appear to be spatially well-separated, leading to a proposal that a change in nucleotide state is coupled to switching between “tensed” and “relaxed” states, a movement equivalent to two base pairs (Sun et al., 2008) (Figure 5a). This hypothesis was supported by experiments showing that mutations to the interface between domains reduced packaging but not packaging-independent ATP hydrolysis, presumably because switching between the tensed and relaxed states was disrupted (Sun et al., 2008). Moreover, the fact that the T4 terminase appears to utilize an intramolecular arginine finger suggests that interdomain, rather than intermolecular, motions accompany nucleotide turnover. This “inchworm” model for DNA translocation is very similar to one inferred from crystal structures of FtsK, a close relative of the φ29 packaging motor, which suggested a hinge motion between the alpha and beta domains of the protein due to change in nucleotide state (Massey et al., 2006). However, the φ29 packaging motor contains neither an FtsK-like alpha domain nor a T4-like nuclease domain, hence ruling out such an inchworm mechanism in this virus, and casting doubt on the functional relevance of the hinge motion in FtsK translocation. Moreover, a strict inchworm mechanism, making identical contacts after each step, could not explain the 2.5-bp steps that correspond to single ATP turnovers in φ29 (Moffitt et al., 2009), nor the sensitivity of the motor to every tenth phosphate (Aathavan et al., 2009), as will be discussed in the following section about previous single-molecule experiments.

Although some other molecular motors have been proposed to act as domain-hinge inchworms, in particular the so-called monomeric SF1 and
SF2 helicases (Singleton et al., 2007), there exist other structurally-motivated models. It was recently proposed that φ29 stores elastic energy in the beta sheet of the ASCE core (Yu et al., 2010), based on analogies with F1-ATPase (Abrahams et al., 1994). As an alternative example, the conformational changes in the Rho helicase are transmitted through a network of salt bridges within one subunit and between adjacent subunits (Thomsen and Berger, 2009).

Bioinformatics has implicated the C-motif in linking product release to conformational change (Draper and Rao, 2007). Some C-motif mutants of T4 are capable of binding and hydrolyzing ATP, but not releasing product, which can be explained if coordination between DNA binding and nucleotide state had been disrupted (Draper and Rao, 2007). Combining this idea with an optical tweezers result that implicates the adenine-proximal Q-motif in force generation (Tsay et al., 2009), it seems that the ATP binding pocket – a dense network of hydrogen and ionic bonds – is a tightly coupled system that reacts in specific ways in response to a change in nucleotide state. Conversely, the ATP binding site itself must catalyze hydrolysis or change affinity in response to DNA translocation. In some molecular motors, the “power stroke” – the force-generating conformational change – coincides with ATP binding, largely due to the zippering of ionic bonds around the phosphate groups (Oster and Wang, 2000; Yasuda et al., 2001). However, in the φ29 packaging motor, the power stroke is not ATP binding, but rather, most likely, post-hydrolysis phosphate release (Chemla et al., 2005). Interestingly, the rotary motor F1-ATPase combines the two models: it performs two strokes per ATP turnover, corresponding to ATP binding and phosphate release (Yasuda et al., 2001) (Adachi et al., 2007).

**DNA Interaction**

The conformational changes in the protein must be efficiently passed on to the DNA through DNA-binding residues. The T4 terminase has been shown to bind DNA through the ATPase domain (Alam and Rao, 2008) but these contacts are not evident from the crystal structure (Sun et al., 2008). Although the terminase was crystallized in the absence of DNA, modeling revealed likely contacts between the nuclease domain and the DNA via loops containing specific arginines and lysines (Sun et al., 2008). These contacts were distinct from the binding sites used in cleavage. Such an electrostatic mode of contact could explain the phosphate interaction observed in φ29 (Aathavan et al., 2009) and is also is consistent with previously published observations of nucleic acid binding by single-stranded translocases such as phage φ12 protein P4 (Mancini et al., 2004), bovine papillomavirus E1 helicase (Enemark and Joshua-Tor, 2006), and Rho helicase (Thomsen and Berger, 2009), although the structures and proposed mechanisms of those motors are quite different. The latter two proteins were crystallized as
substrate-bound oligomeric rings with subunits in each of several different nucleotide states, so they may provide the most biologically relevant information. The DNA-binding residues are typically located in a flexible loop, making specific contacts with the backbone, and with the adjacent residues sometimes helping to stabilize the complex via side-chain and main-chain hydrogen bonds (Chen et al., 2008; Thomsen and Berger, 2008). The DNA binding loops of the subunits in the ring form a “spiral staircase” that tracks the substrate backbone, such that several subunits simultaneously make contact with the substrate. The position of a binding loop is dependent on nucleotide state, and the spiral of binding loops is thought to move as a wave, pushing the DNA as hydrolysis propagates around the ring (Figure 5b). The specific relationship of nucleotide state and loop state gives rise to the polarity of translocation. (Thomsen and Berger, 2009) Moreover, stepping by the motor is coordinated also because the affinity of the loop for DNA varies with nucleotide state. For the φ29 ATPase, as well as for a number of other RecA-like hexameric nucleic acid translocases, the ATP-bound state has highest affinity for the polymer while the ADP-bound and apo states have much lower affinity (Chemla et al., 2005; Thomsen and Berger, 2008).

The wave mechanism used by the related ring ATPases requires a high degree of coordination between subunits in order to enable efficient translocation. This communication in the hexameric helicases, among other members of the ASCE division, is mediated by nucleotide bound at the interfaces between subunits (Enemark and Joshua-Tor, 2008; Erzberger and Berger, 2006; Thomsen and Berger, 2008). Therefore, the coordinated and sequential firing of the DNA binding loops in the hexameric helicases should be accompanied by a change in nucleotide state that proceeds, sequentially, around the ring. However, the T4 terminase structure provides no clear mechanism for intersubunit communication, because it lacks an intermolecular arginine finger (Sun et al., 2007). If there is coordination, as there must be if each subunit has only one DNA-binding site, the signal may be transmitted through the DNA itself.

1.6 Optical tweezers experiments on φ29 packaging

In order to observe packaging with optical tweezers, the packaging complex, consisting of the prohead, packaging motor, and DNA, is stretched between two micron-sized polystyrene beads (Fig. 6a). The DNA can be attached to a streptavidin-coated bead by modifying the distal DNA end with biotin while the other end of the packaging complex – the prohead – is affixed to another bead via antibodies to the phage capsid protein. Both beads can be held in separate optical traps (or one in an optical trap and the other affixed via
suction to a micropipette) allowing the tension and extension of the DNA to be monitored in real time as the DNA is packaged.

As packaging progresses, the DNA tether shortens. By converting the bead-to-bead distance (the extension) to DNA contour length using the wormlike chain model of DNA elasticity, the maximum packaging velocity was determined to be about 120 bp/s (Smith et al., 2001). The packaging velocity depends on various conditions such as the applied force, or load, and the concentration of ATP.

Energy landscape

A unique capability of optical tweezers is that the application of force allows the energy landscape for translocation to be controllably modified due to an additional term, $F\Delta x$, which results from the mechanical work required to move a distance $\Delta x$ against a force $F$. The dependence of packaging on force thus contains information about the energy landscape in the direction of packaging (Fig. 6b). First, at no force does the motor move backward, meaning that stepping is kinetically irreversible and that the energy drop from one cycle to the next must be large. Given a defined energy available due to ATP hydrolysis, the force at which packaging velocity drops to zero (which has been estimated at 110 pN (Rickgauer et al., 2008)), provides an upper bound for the step size of about 1 nm. In addition, the fact that velocity decreases with force above about 10 pN means that the physical step of the molecular motor is rate-limiting at those forces. Moreover, fitting the rate of stepping to a Kramers model with a force-independent attempt rate and a force-dependent activation energy,

$$v = A e^{-\frac{\Delta G + F\Delta x}{k_B T}}$$

reveals a distance to transition state $\Delta x$ which represents the distance over which the system moves before the remainder of the motion is spontaneous. In φ29, the distance to transition state for stepping is 1.1 Angstroms, which is the size of a single hydrogen atom (Chemla et al., 2005). In this way, optical tweezers can be used to infer high-resolution details about the conformational changes that contribute to DNA translocation.

Coupling of chemistry and mechanics

Other details of the sequence of events that contribute to a cycle of the packaging motor can be discerned from packaging as a function of reactant (ATP) and product (ADP and Pi) conditions (Chemla et al., 2005). Φ29’s packaging velocity has a Michaelis-Menten dependence on ATP concentration with a $V_{\text{max}}$ of 120 bp/s and a $K_m$ of 30 µM. Both $K_m$ and $V_{\text{max}}$ depend on force, but their ratio, interestingly, does not; this implies that ATP
binding is not the mechanical step, nor is it connected to the mechanical step through a pathway of reversible kinetic transitions (Figure 7). Notably, fitting the curves to a more general Hill model reveals a Hill coefficient of 1, suggesting that ATP binding by the five subunits is not cooperative or simultaneous. Only one subunit can bind ATP at a given moment in time. Other experiments showed that ADP binds reversibly to the motor, meaning the ADP release process cannot trigger stepping (since stepping is irreversible.) In contrast, increasing the concentration of inorganic phosphate does not slow packaging, meaning that phosphate release is irreversible – and thus could coincide with the stepping transition.

*Ultrahigh-resolution observations*

Given an ATP consumption rate measured in bulk of one ATP per every two basepairs packaged, along with a velocity on the order of 100 bp/s, it was proposed that the packaging motor takes a 2-bp step on average every 20 ms. In order to directly observe these steps, an ultrahigh-resolution optical tweezers was constructed (Moffitt et al., 2006). Data collected by these optical tweezers have an order of magnitude less noise than that previously possible, largely thanks to two innovations: decoupling the tether from the sample chamber by forming two optical traps, and recording only asymmetric (anti-correlated) fluctuations in the bead positions.

*Stepping is coordinated and adds up to 10 bp*

At low forces, it was observed that DNA was packaged in increments of 10.0 ± 0.2 bp, independent of the concentration of ATP (Moffitt et al., 2009). See Figure 8a. Between each 10-bp increment, the motor dwells for a short, ATP-dependent time. Closer inspection of these packaging traces revealed that the 10 bp packaging events were actually bursts of multiple smaller steps. Because applied force should slow the rate of stepping according to the Arrhenius-like model shown above, the steps could be resolved at sufficiently high opposing force. Each burst contains four steps, and, remarkably, each step covers a non-integer number of base pairs, 2.4 ± 0.1 (Figure 8b). A step size that does not correspond to the periodicity of the DNA track suggests that the contacts between the motor and DNA during the burst may not be chemically specific. However, the fact that the cycle resets every ten basepairs leaves open the possibility of specific chemical contacts at the beginning and end of the burst.

*The motor binds four ATPs in a coordinated manner*

Careful examination of the dwells that separate the translocation bursts revealed more about the chemical changes that take place during a cycle. The shape of the distribution contains information about the number of rate-
limiting steps that take place during the dwell. A single stochastic rate-limiting step would generate an exponential dwell time distribution, whereas a large number of equally-rate-limiting steps would, by the central limit theorem, generate a Gaussian-like distribution with a mean and a standard deviation much less than the mean. The ratio of the squared mean to the variance, then, is a measure of the number of rate-limiting steps (Svoboda et al., 1994); it has been termed $n_{\text{min}}$ (Moffitt et al., 2010). Applying this fluctuation methodology to the dwell time distribution reveals that, at limiting ATP concentrations, the motor must bind at least four ATP molecules before beginning a burst. These observations strongly suggest that each full mechanochemical cycle of the motor is composed of two phases: (i) a dwell phase in which four of the five subunits bind ATP sequentially, each delaying the utilization of this molecule until the ring is loaded and (ii) then a burst phase in which these subunits cooperatively generate a series of four 2.5 bp steps in quick succession (Figure 8c).

In this way, high-resolution measurements have revealed three essential asymmetries of this ring-shaped ATPase. First is a mechanical asymmetry – despite having five subunits, the motor moves the DNA in four consecutive increments before dwelling to bind ATP. Then, the motor binds either four or five ATPs, with each ATP being bound with a different catalytic efficiency – a kinetic asymmetry. And finally, individual ATP binding events and DNA stepping events are strictly separated in time, taking place in an ordered manner. Such profound asymmetries between subunits reveal that the subunits are not so identical as they appear. As seen in figure 8c, there must be one “special” subunit that does not translocate DNA during one 10-bp burst, and may not bind ATP at all. These asymmetries can be seen within an individual cycle. How the 5-fold symmetry of the motor is broken, and whether it is broken differently from one cycle to the next, remains unknown. One possible origin of the symmetry breaking is the interaction of the motor with the DNA.

**Backbone phosphates improve processivity and regulate the chemical cycle**

To address these issues in the packaging motor of φ29, Aathavan et al. (Aathavan et al., 2009) challenged the packaging motor with a DNA molecule that had been locally modified and monitored packaging as the complex encountered the modification with an optical tweezers assay. To probe the importance of ionic contacts with the charged phosphate backbone, a neutrally-charged, structurally-identical mimic of DNA was tested. The packaging motor successfully packaged neutralized DNA of various lengths from 5 – 30 bp with relatively high efficiency (greater than 10%). By reintroducing charge to each strand selectively, it was found that the most important phosphate contacts are made with adjacent phosphates on the 5’-3’ strand, in the direction of packaging, every 10 bp. The absence of a
phosphate charge on that strand results in a significant increase in the frequency with which the motor slips backward on the DNA, revealing that phosphate contacts are necessary for mechanical stability and high enzymatic processivity. In addition, in high resolution observations, it was found that the rate at which the motor attempts to package the DNA decreased 10-fold in the absence of the phosphate charge. This reduced attempt rate was largely due to a single slow kinetic event. Thus, in the φ29 packaging motor, contacts with phosphate charges on the 5'-3' strand every 10 bp provide the motor with its high processivity and appear to play a role in the regulation of the chemical cycle.

*Other contacts are not specific to DNA*

Other modifications to the DNA also provided insight into the functional role of certain motor-DNA contacts. A wide variety of chemical modifications were tried, including abasic DNA, single stranded gaps, unpaired bulges, and even a polyethylene glycol oligomer with no resemblance to DNA. The motor was capable of packaging all of these modifications with relatively high efficiency. The only conclusion is that none of the individual chemical features of DNA—phosphates, sugars, and bases—are absolutely required for packaging. By performing a multi-variate logistic regression Aathavan *et al.* were able to quantify the “importance” of each of the different chemical moieties of DNA (Aathavan *et al.*, 2009). The results can be summarized in a heat-map of the interactions between DNA and the motor (Figure 9). The picture that emerges then is that important phosphate contacts are made by the motor every 10 bp on the 5'-3' strand oriented in the direction of packaging. These periodic contacts are presumably made during the dwells during which ATP binds to the motor. During the burst, on the other hand, no specific contacts are absolutely required, although they do improve the probability of a successful step. It seems reasonable to propose that the motor is making steric (repulsive) contacts with the substrate during the burst; alternatively, it is possible that no contacts are truly necessary. In the latter model, the motor would step like an inchworm, making contact only every 10 bp. Mutational studies of the related ASCE motor, simian virus 40 hexameric helicase, provide some support for the necessity of non-specific, perhaps steric contacts (Shen *et al.*, 2005).

*Motor-DNA contacts change through the cycle*

By integrating what we know about the importance of the motor-DNA contacts along with the intersubunit coordination during ATP binding and translocation, we can infer the timing of the motor-DNA contacts within the mechanochemical cycle. During the dwell phase of the cycle, as the motor loads four ATPs, phosphate contacts on the 5'-3' strand provide the necessary strong contacts to keep the motor bound to the DNA against
opposing loads. The fact that removing these contacts dramatically changes the chemical cycle suggests that there is a single chemical transition during this dwell phase which is strongly dependent on the proper positioning of the DNA. This kinetic step might serve as a mechanochemical checkpoint, properly synchronizing the completion of the packaging from the previous cycle with the loading of nucleotides for the next cycle. Since neutral stretches of DNA shorter than 10 bp were readily packaged, phosphate contacts are not nearly as important during the burst phase, when the DNA is packaged in four 2.5-bp steps, as during the dwell phase. During this phase, it is likely that more transient, non-nucleic-acid-specific contacts are what actually drive packaging. Indeed, nonspecific contacts provide a natural explanation for the noninteger step size. While such promiscuous contacts present a significant challenge to the prevailing view of specific motor-DNA contacts from an energetic perspective, they may also provide one solution to a basic problem. During active packaging strong contacts are needed to transduce large forces, but these contacts must also be short-lived since quick packaging requires them to break a few milliseconds after they are made. Non-specific, potentially steric contacts may satisfy this requirement. Similar periodic contacts have been inferred from nucleotide-analog-interference studies with the related Rho translocase (Schwartz et al., 2009).

1.7 Building a model of translocation

We now outline a potential model for packaging that integrates data from the packaging motors with inferences from other ATPase molecular machines (Figures 7, 8, and 9) (for a more detailed model, see (Aathavan et al., 2009; Yu et al., 2010)). The packaging ATPase motor is a five-membered nearly-planar ring that encircles the DNA and binds ATP at the interfaces between the subunits. In this state, the DNA-binding loops or residues are positioned such that they have high affinity for the substrate, although each may not actually contact the DNA at this time. Once at least four ATP molecules are bound, one of the subunits hydrolyzes its ATP and subsequently releases inorganic phosphate while moving the DNA-binding motif over a distance of about 2.5 basepairs. This DNA-binding motif, utilizing any or all kinds of interaction – electrostatic attraction, electrostatic repulsion, hydrophobic interaction, hydrogen bonding, van der Waals bonding, etc. – pushes the DNA by about the same distance. At the end of the stroke, the change in nucleotide state has two effects: it causes the adjacent subunit to hydrolyze its ATP, release phosphate, and undergo a stroke, and it also reduces the affinity of the first subunit for DNA, hence allowing the DNA to move with the second subunit. The chain of hydrolysis proceeds around the ring until four steps, adding to ten basepairs, have been taken. The four ADP
molecules must then be exchanged for ATP while the motor does not stroke; exchange is probably sequential, with the energetic cost of releasing ADP from one binding pocket being balanced by the energetic benefit of binding ATP to another (Moffitt et al., 2009; Yu et al., 2010). During the extended stationary period (including ATP binding and at least four other steps), the fifth subunit, being in the high-affinity ATP-bound state, grips the DNA tightly through an electrostatic interaction with the phosphate backbone. Because this bond is ionic, the total distance packaged in a full cycle is rectified to be exactly ten basepairs even if slight misalignments or slips take place during the burst. Once the four other subunits bind ATP, the hydrolysis cycle can begin again – triggered by the fifth subunit, which is making the phosphate contact that is known to regulate entry into the burst phase. The entire process is likely allosterically activated by the connector, which also acts as a one-way valve to prevent backward slipping.

This model explains the data we have reviewed here for φ29, but it raises some important biophysical questions. For example, what is the extent of coordination in the system – are hydrolysis, phosphate release, or ADP release also synchronized and temporally separated? What is the role and physical origin of the unique fifth non-translocating subunit? How is the symmetry of the motor broken, and is it always broken in the same way? The fivefold rotational symmetry of the motor and the 10-bp periodicity of its bursting on DNA are intriguing in the context of the DNA’s geometry. B-form DNA has a mean pitch of 10.4 basepairs per turn, so a 10-bp burst would cause the DNA to be out of alignment with the motor. How is this symmetry mismatch rectified? This question will be directly addressed in Chapter 4.

1.8 Organization and physics of the genome during and after packaging

The ultimate purpose of the packaging phenomena discussed thus far is to compact the viral genome inside the capsid. Compaction of DNA to the near-crystalline densities found in mature viral particles requires overcoming a significant energetic barrier, the size of which depends crucially on the physics of condensed DNA (reviewed in (Nurmemmedov et al., 2007; Petrov and Harvey, 2008)). Moreover, it has been suggested that the high energy of packaged DNA assists the ejection of DNA from the particle during the subsequent infection process; thus, the stored energy can have crucial biological consequences (Koster et al., 2009).

Packaging dynamics at high internal filling

Φ29 packages at a maximum velocity of about 120 bp s⁻¹, but it slows down as more DNA is encapsidated (Smith et al., 2001), similar to the decrease in packaging that arises if an external force is applied to the DNA being
packaged (Figure 5b). This observation suggests that there is an internal force, resisting packaging, that builds in the capsid as the DNA is packaged. With some assumptions, Smith et al. were able to use the dependence of the packaging velocity on externally applied force to estimate this internal force from the measured packaging velocity (Rickgauer et al., 2008). The total effective force – the sum of the external load and the inferred internal force – at which packaging velocity drops to zero is 110 pN. When the motor is unopposed by an external force, it reaches these slow packaging velocities only at the end of packaging its full 19.3 kbp genome, suggesting that ~100 pN is the internal force resisting packaging when the entire φ29 genome is packaged. It is experimentally difficult to measure the physics of DNA compressed to such high densities; thus, this indirect estimate of the mechanical energy needed to package the genome has provided a useful basis for improving theoretical models of the physics of DNA confinement (Petrov and Harvey, 2008). Similar arguments have been used to estimate the energetic cost of DNA confinement inside the capsid for T4 and λ (Fuller et al., 2007a; Fuller et al., 2007b), and the energy scales and internal forces are all comparable. T4, λ, and φ29 have significantly different genome sizes, but their capsid volumes scale proportionally, suggesting that, perhaps, the physics of DNA confinement constrains the co-evolution of genome size and capsid volume.

Capsid dynamics at high internal filling

Measurements of packaging velocity as a function of capsid filling have also revealed structural changes in the capsid during packaging. Volume increase concomitant with capsid expansion (Conway et al., 1995) should reduce the free energy of the confined DNA and thus reduce the energetic cost of packaging more DNA, thereby lowering the resistance to packaging. In λ, the velocity of packaging is a constant 600 bp/s until the genome is about 20% packaged (Figure 5c), at which point the velocity begins to drop (Fuller et al., 2007b). The drop is interrupted by a sudden increase in velocity at about 30% filling, which coincides with biochemical observations of the timing of capsid expansion (Feiss and Catalano, 2005). In simulations of packaging, the energy of packaged DNA is particularly sensitive to the volume and geometry of the capsid (Petrov et al., 2007) suggesting that even a minor change in the capsid structure could have an effect on packaging velocity. A second structural change was also apparent in the single molecule experiments: at about 90% filling, many packaging complexes suddenly accelerated to the maximum velocity, which can be explained if the DNA were no longer strongly confined. Because the in singulo experiment lacked the capsid-stabilizing accessory protein gpD, the acceleration was interpreted as capsid rupture. Moreover, the “internal pressure” at which the capsid ruptures is an indication of the capsid strength. In this way, optical
tweezers can be used as a tool to probe both the dynamics and structural mechanics of the capsid (Fuller et al., 2007b; Rickgauer et al., 2008).

**Energetic of DNA confinement**

The energy of DNA confined in the capsid is determined by contributions from electrostatic self-repulsion, bending strain, twist strain, volume exclusion, and entropy (reviewed in (Petrov and Harvey, 2008)). The dominant factor is thought to be electrostatic, although entropy may also be significant (Locker et al., 2007). In order to experimentally assess the electrostatic component of the confinement energy, Fuller et al. examined packaging in a variety of different buffer conditions containing different species and concentrations of cations (Fuller et al., 2007c). After controlling for the chemical effects of the different cations on turnover of the packaging motor itself, the internal force could be calculated. This internal force generally decreased with cation valency and concentration, as would be expected due to electrostatic shielding, but this force was uniformly much higher than the quantitative predictions of existing models (Purohit et al., 2005). This discrepancy points to an incomplete understanding of either the mechanism by which packaging is slowed, or the physics of DNA itself.

The physics of the confined DNA can be inferred by studying another step in the viral life cycle: DNA ejection. In ejection, the internal forces that drive the DNA out of the capsid must work against some external forces such as osmotic pressure (reviewed in (Castelnovo and Evilevitch, 2007; Nurmemmedov et al., 2007)). By tuning the osmotic pressure difference between the interior of the capsid and the outside environment (by varying the concentration of polyethylene glycol), and monitoring the dynamics of ejection, those “internal pressures” can be estimated (Evilevitch et al., 2004; Evilevitch et al., 2008; Evilevitch et al., 2005; Evilevitch et al., 2003; Grayson et al., 2006; Lof et al., 2007a, b). The inferred pressures are in the range of a few tens of atmospheres. It should be noted that the functional equivalence of polymer osmotic pressure and DNA internal forces has not been proven. In order to address this concern, isothermal titration calorimetry of ejection by λ has been used to directly measure the heat of ejection (Jeembaeva et al., 2010). Interestingly, although the heat released is consistent with theoretical estimates, the contribution of entropy is, surprisingly, opposed to DNA ejection. The effect was attributed to a decrease in entropy of solvation as the DNA leaves the crowded capsid interior and acquires a water layer. This counterintuitive finding highlights the limitations of coarse-grained models of DNA compaction. More experimental and theoretical work is needed to fully understand the physical nature of the packaged viral particle.

**Organization of packaged DNA**
The energy of the packaged DNA is also expected to depend on its arrangement inside the capsid. Minimization of bending energy suggests that the DNA, having a persistence length of 50 nm (Bustamante et al., 1994), would prefer to bend as gradually as possible inside the capsid, forming loops the size of the capsid inner diameter (Kindt et al., 2001) (Odijk, 1998). At the same time, electrostatic energy is likely to be minimized by a hexagonal close-packed configuration (Odijk, 1998). Images of fully assembled particles seem to confirm this intuition: When viewed along the connector/portal axis, a rotationally symmetric, radially periodic pattern in the DNA electron density can be seen (Figure 10) (Comolli et al., 2008) (Cerritelli et al., 1997; Chang et al., 2006; Comolli et al., 2008; Fang et al., 2008; Fokine et al., 2004; Jiang et al., 2006; Lander et al., 2006) and these patterns were initially interpreted as spools. However, on closer inspection of some phages, this interpretation has been challenged. The periodic pattern in φ29 is visible from all angles and, more significantly, the pattern does not emerge until greater than 90% of the genome has been packed (Comolli et al., 2008). Hence this local order need not reflect any long-range organization such as spooling (except for phages that have an axial core such as T7) (Comolli et al., 2008; Fang et al., 2008). Indeed, DNA inside P2 and P4 phages is capable of forming a wide variety of knots (Arsuaga et al., 2002; Liu et al., 1981) when the cohesive ends are allowed to anneal, suggesting a great deal of heterogeneity in DNA organization between particles. This heterogeneity suggests that the internal forces, and thus the packaging dynamics, among an ensemble of single particles may vary. Interestingly, when treated with a crosslinker while in the capsid and later released, the genome of lambda displays a degree of crosslinking (Virrankoski-Castrodzeza et al., 1982), which suggests a nonzero linking number for the packaged DNA. That linking number can be stored in the form of twist or writhe, or both. In summary, while the packaging of DNA inside the particles may lead to some close packing of the local segments of the DNA, models in which the DNA is an orderly spool have not been experimentally validated. The first experimental evidence of long-range order within the capsid will be presented in Chapter 4 of this dissertation.
Figures

Figure 1-1. Outline of viral packaging pathway, components not drawn to scale. (a) Bacteriophage Φ29 DNA, labeled with the terminal proteins (orange balls), forms a lariat and is bound by the ATPase (blue) as well as a prohead, composed of the capsid (yellow), connector (green at base of capsid), and pRNA (magenta). Packaging proceeds from the left terminal protein until the right terminal protein is reached. The ATPase and pRNA then dissociate and the tail complex (gray) attaches.
Figure 1-2. Structure of the φ29 packaging complex. The dodecameric connector (cyan), with a channel large enough to permit dsDNA, is ensconced in the capsid (gray). The pentameric pRNA (magenta) binds the connector at one end and the ATPase (blue) at the other. Figure courtesy of M. Morais and M. Rossmann and reprinted from (Morais et al., 2008).
Figure 1-3. Phylogenetics of the ASCE division. (a) Topology showing the highest-order relationships proposed by Burroughs, Iyer, and Aravind (Burroughs et al., 2007; Iyer et al., 2004a; Iyer et al., 2004b). Strands and helices comprising the ASCE core are shown and numbered; important ATP binding motifs are shown in magenta (A = Walker A, B = Walker B, R = arginine finger). If a motif is not shown for one family, it has not been
conclusively located. Where two identical motifs appear in one family, the motif has been found in both places. The RecA family includes the ssRNA packaging phage ATPases. The other families noted in this figure are: PilT, the pilus retraction ATPase; RecA, the filament-forming oligomeric ATPase involved in homologous recombination; SF1/SF2 helicases, non-ring-forming ATPases that manipulate nucleic acids; ABC, the ATP-binding cassette transmembrane transporters; KAP, a largely uncharacterized phyletic group likely to function in assembly of membrane-associated signaling complexes; STAND, another poorly understood group that participates in signal transduction during programmed cell death; AAA+, ATPases associated with various cellular activities. Figure inspired by (Burroughs et al., 2007).
Figure 1-4. Structure and conformational changes of portal/connector. (a) Crystal structure of the SPP1 portal. The free portal crystallized as a 13-mer but is known to be a 12-mer in the capsid. (b) Domains of the portal. The wing domain (red) contacts the capsid; the crown domain (green) faces the capsid interior; the stem domain (cyan) leads to the clip domain (magenta) which faces the ATPase. The tunnel loops (black) are proposed to contact DNA at the interior of the tunnel, and they may be coupled to changes in the kinking of the helix at the intersection of the tunnel loop and wing domains. (c) Conformational change associated with completion of packaging in P22 phage. At the top is a cryo-EM image reconstruction of free portal; below is a reconstruction of the portal component of a fully assembled virion. A ring of DNA (orange) can be seen surrounding the portal crown domain. Figures (a) and (b) are from PDB accession code 2JES (Lebedev et al., 2007). Figure (c) courtesy of G. Lander and J. Johnson from (Lander et al., 2006).
Figure 1-5. Proposed mechanisms of DNA translocation. (a) Upon nucleotide binding to the ATPase domain, the terminase large subunit of T4 is thought to undergo domain motions allowing the nuclease domain (brown) to pull DNA by 7 Angstroms, or 2 basepairs. After one translocation step, a new motor-DNA contact must be made by subdomain II (gray), or by another subunit in the oligomer. (b) The RNA helicase Rho forms a ring (colored balls) with nucleotides bound at the interfaces (not shown); the ring encircles an RNA polymer (gray balls). The six RNA binding loops shown here form a spiral staircase tracking the RNA helix, with the height of each lever and the extent of interaction with the RNA corresponding to the state of its ATP-binding region (T = ATP, T* = ATP tightly bound in preparation for hydrolysis, D = ADP, E = empty). Although the phage ATPases are 5-membered rings that translocate DNA, they may operate via a similar mechanism. Figure (a) courtesy of J. Tainer and adapted from (Williams et al., 2008). Figure (b) inspired by (Patel, 2009).
Figure 1-6. Optical tweezers observations of phage packaging. (a) Experimental setup. The capsid is immobilized onto a micron-sized polystyrene bead via antibodies while the far end of the DNA is attached to another bead via a biotin-streptavidin linkage. Each bead is held in a focused laser spot, allowing measurement and application of forces and positions. Some experiments use only one optical trap, with the other bead immobilized by suction onto a micropipette. (b) Hypothetical example of how force modifies the energy landscape of a molecular motor. In this example, the motor takes steps (as indicated by the distance between minima in the potential $\Delta x$). The horizontal distance between a minimum and the nearest maximum is the distance to the transition state $\Delta x^t$, and the energy difference between minimum and maximum is the activation energy of a step $\Delta G^t$. When a resisting force, or load, is applied to the motor, the motor must perform additional mechanical work $F\Delta x$ in order to move against the force. Likewise, the height of the barrier to stepping is increased by a mechanical term $F\Delta x^t$. 
Figure 1-7. Force dependence of packaging velocity. (a) Packaging velocity as a function of external force under saturating [ATP], normalized to the value at 5 pN. (b) Michaelis-Menten parameters, $V_{\text{max}}$ and $K_M$, as a function of force for the packaging motor of φ29. (c) The force dependence of the ratio of these two parameters. Data in (a) previously published in (Fuller et al., 2007a; Fuller et al., 2007b; Rickgauer et al., 2008). Data in (b) and (c) reproduced from (Chemla et al., 2005). This figure is reproduced from (Moffitt, 2009).
Figure 1-8. Intersubunit coordination in the packaging motor of φ29. (a) Schematic diagram of the two phase cycle superimposed on a actual packaging trace. During the dwell phase the motor loads four ATPs while holding the DNA at constant length. During the burst phase the motor uses these ATPs to package the DNA in four 2.5-bp steps. (b) Sample packaging traces under high opposing loads. The fast 2.5-bp steps are not observable under low forces, but under high forces they are clear. (c) Schematic picture of the full mechanochemical cycle of the packaging motor. During the dwell phase, a single subunit is capable of binding ATP (green) and once it does (red) it activates the adjacent subunit. Once fully loaded the motor enters the burst phase where, triggered by Pi release, it takes a series of 2.5-bp steps. Blue subunits correspond to subunits that may contain product ADP and the purple subunit signifies the fifth, distinct subunit. This “special” subunit may change each cycle. A careful analysis of fluctuations in the dwell phase duration reveals that each of the four subunits must have a catalytic efficiency (ratio of $k_{cat}$ to $K_M$) that must vary by roughly a factor of four from
the previous. While depicted as increasing here, the data cannot uniquely
determine whether the catalytic efficiencies decrease or increase around the
ring. Figure from (Moffitt et al., 2009).
Figure 1-9. Heat map of importance of interactions between DNA and the packaging motor of φ29. Contact importance is defined as the reciprocal of the number of base pairs of a given chemical moiety (phosphate charge or sugars/bases) that must be removed to decrease the probability of packaging this lesion to 50%. The most important contacts are with adjacent phosphates on the 5'-3' strand in the direction of packaging (down in this image). These contacts are likely made during the dwell phase. During the burst phase, contacts are made with all other chemical moieties on both strands. Figure reproduced from (Aathavan et al., 2009).
Figure 1-10. 3D cryo-EM reconstruction of mature φ29 particles reveals DNA organization. A periodic pattern is apparent in the DNA density of cross-sections from the top (a) and side (b) of the reconstructed particle. 3D views of the reconstruction (c) and (d) reveal individual layers of DNA, in addition to rings of DNA close to the connector (arrows). Figure provided by L. Comolli, A. Spakowitz, and K. Downing, and reproduced from (Comolli et al., 2008).
Chapter 2. Experimental Test of Connector Rotation During Packaging

This contents of this chapter have been published in PLoS Biology (Hugel et al., 2007).

2.1 Abstract

The bacteriophage φ29 generates large forces to compact its double-stranded DNA genome into a protein capsid by means of a portal motor complex. Several mechanical models for the generation of these high forces by the motor complex predict coupling of DNA translocation to rotation of the head-tail connector dodecamer. Putative connector rotation is investigated here by combining the methods of single-molecule force spectroscopy with polarization-sensitive single-molecule fluorescence. In our experiment, we observe motor function in several packaging complexes in parallel using video microscopy of bead position in a magnetic trap. At the same time, we follow the orientation of single fluorophores attached to the portal motor connector. From our data, we can exclude connector rotation with greater than 99% probability and therefore answer a long-standing mechanistic question.

2.2 Introduction

As part of its viral infection cycle, the Bacillus subtilis bacteriophage φ29 packages its double-stranded DNA genome into a preformed capsid shell, or prohead, by means of a powerful molecular motor (Grimes et al., 2002; Smith et al., 2001). The DNA-packaging motor is situated at a unique 5-fold vertex of the prohead and is a complex assembly of multiple components. At the core of the motor is the dodecameric head-tail connector, gene product 10 (gp10). Associated with the connector is a ring of RNA molecules (prohead RNA or pRNA), which is required for packaging. A ring of ATPases (gp16) interacts with the pRNA to complete the packaging machinery. gp16 belongs to the Her A, FtsK superfamily of ATPases (Burroughs et al., 2007; Iyer et al., 2004b). Hydrolysis of ATP powers the motor and drives viral DNA into the prohead.

While numerous biochemical, structural, biophysical, and theoretical studies have elucidated many details of the packaging process (Chemla et al., 2005; Grimes et al., 2002; Guasch et al., 2002; Jardine and Anderson, 2006; Kainov et al., 2006; Kindt et al., 2001; Lee and Guo, 2006; Morais et al., 2001; Purohit et al., 2005; Purohit et al., 2003; Robinson et al., 2006; Simpson et al., 2000; Smith et al., 2001; Tzlil et al., 2003) a complete
mechanistic understanding of how the components of the portal motor force the DNA into the capsid has not been presented. In particular, many theoretical models for the function of the connector have been proposed (Cuervo et al., 2007; Dube et al., 1993; Fujisawa and Morita, 1997; Guasch et al., 2002; Hendrix, 1978; Lebedev et al., 2007; Simpson et al., 2000). Most of these models include a rotation, either passive or active, of the connector with respect to the prohead shell—an idea first introduced by Hendrix in 1978 (Hendrix, 1978). Recently, a study of DNA packaging in T4, using cross-linking of bulky domains to the connector that could interfere with connector rotation, provided indirect evidence that the connector of T4 does not rotate during packaging (Baumann et al., 2006). However, no direct structural, biochemical, or biophysical experiments have been published that address the rotation hypothesis. Here, we directly test this hypothesis using single-molecule fluorescence polarization (SMFP) spectroscopy in combination with single-molecule force spectroscopy.

Single-molecule force spectroscopy has proven to be a powerful method for studying the movement of motor proteins. In recent years, a wealth of different systems has been studied, such as actin- and microtubule-based molecular motors (Finer et al., 1994; Svoboda et al., 1993); motors that move along DNA, like DNA polymerase (Wuite et al., 2000), RNA polymerase (Adelman et al., 2002; Forde et al., 2002; Perkins et al., 2003; Yin et al., 1995), exonuclease (Perkins et al., 2003; van Oijen et al., 2003), and DNA pumps (Pease et al., 2005; Smith et al., 2001); as well as motors that move pili (Maier et al., 2002) or whole bacteria (Block et al., 1989). Here, we use a magnetic tweezers set-up to observe the packaging of many DNA-packaging complexes simultaneously.

Single-molecule fluorescence spectroscopy has been used recently to study conformational changes of single-motor complexes (Adachi et al., 2000; Borsch et al., 1998; Churchman et al., 2005; Forkey et al., 2003; Kapanidis et al., 2005; Rasnik et al., 2004; Sosa et al., 2001). In particular, detection of changes in the polarization of emission by a single-dye molecule is well suited to investigate conformational changes involving rotation events (Adachi et al., 2000; Forkey et al., 2003; Sosa et al., 2001). We utilize this method to investigate the putative rotation of the connector protein with respect to the prohead shell.

To test the hypothesis of connector rotation, we track orthogonal polarization components of fluorescent light emitted by single-dye molecules attached to the connector. Several issues must be overcome to ensure experimental fidelity: (i) DNA packaging by the labeled motor complex must be observed simultaneously with the fluorescence detection; (ii) dye labeling must be specific to the connector; (iii) the dye molecule must stay at a fixed angle relative to the connector; and (iv) the prohead itself must be immobilized.
without possibility of rotation. In order to fulfill all these requirements, we designed and implemented a combined SMFP and magnetic tweezers packaging assay as described below. The results of our experiment allow us to rule out connector rotation during packaging with more than 99% probability.

2.3 Results

Single-molecule packaging

Figure 1 shows a schematic of the experimental geometry. Packaging is initiated in bulk as described previously (Smith et al., 2001), and complexes stalled with non-hydrolyzable ATP analog (ATP-γS) and enriched for active particles are bound to streptavidin-coated superparamagnetic beads via a biotin modification to the distal end of the DNA (see Materials and Methods for details). Antibodies against the capsid protein gp8 are used to anchor the stalled complexes via the prohead shell to a quartz slide. The experiment is observed through a glass cover slip, which together with the quartz slide forms a fluid chamber, using a 1.2-numerical aperture (NA) water-immersion objective. Magnets placed on either side of the objective pull the beads away from the quartz surface, stretching the unpackaged DNA under a force of about 0.1 pN. We monitor the beads via video microscopy and can calculate the length of the DNA tether through either Brownian motion (Charvin et al., 2004) or change of focus of the bead image. Upon exchange of buffers to remove ATP-γS and reintroduce ATP, the complexes resume packaging as confirmed by gradual reduction of tether length. The packaging is monotonic and ATP-dependent. In a typical experiment, about 80% of stalled, tethered complexes package to completion at rates consistent with previous bulk and single-molecule measurements (Chemla et al., 2005).

Fluorescent labeling

Conformational changes that occur during the enzymatic cycle of a motor can be efficiently probed by attaching dye molecules as local reporters to specific proteins of a large multi-subunit complex. Single-dye molecules are sufficiently small so as not to interfere sterically with biological activity in most cases. The simplest way to attach a dye molecule as a reporter to a specific site of a protein is to make use of the high specificity of a cysteine-maleimide reaction. To this end, one needs to make point mutations in order to remove native cysteines, which could be inadvertently labeled, and introduce new cysteines (one at a time) to desired exposed locations. A dye molecule with a reactive maleimide group can then be covalently attached to that site.
For our experiments, we replaced the two native cysteines of the connector with serines by site-directed mutagenesis and introduced new cysteines at various positions at the inside and outside of the connector. Figure 2 shows the X-ray crystallographic structure of the connector highlighting the position of the amino acids that were mutated to cysteines (see Materials and Methods). Since current techniques are unable to assemble the free connector protein gp10 into a prohead in vitro, it was necessary to label the connector in the presence of the complete capsid, including capsid (gp8), head fiber (gp8.5), and scaffold proteins (gp7) (none of which contain cysteines). Figure 3a shows, as an example, the fluorescence image of a denaturing gel (SDS-PAGE) of the proheads with a cysteine at amino acid 170 (170C) of the connector, labeled with a Cy3-maleimide. About 80% of the total fluorescence signal is in the single band of the connector, while capsid protein (which makes up more than 95% of the total mass), head fibers, scaffolding protein, and residual contaminants of the Escherichia coli particle expression system show only very weak Cy3 fluorescence. Furthermore, the labeled proheads package DNA in vitro with the same efficiency as unlabeled particles, as can be seen by bulk packaging experiments (see Figure 3b and Materials and Methods).

Having established a specific labeling scheme in a complex that is competent for in vitro packaging, we also need to ensure that the dye molecule can act as a reporter for a rotational movement of the connector. Several criteria must be satisfied: the dye must be bright enough to allow integration times less than one rotational period; the time before photobleaching of the dye molecule must be long enough to observe several rotations; the dye must be attached at angles relative to the connector axis and relative to the objective optical axis such that changes in the dye dipole moment can be seen with our instrument; and the dye molecule must be attached rigidly. To address the latter concern, we measured the fluorescence anisotropy of an ensemble of labeled proheads. The resulting anisotropies of the labeled mutant particles were typically $r \sim 0.3$. The common interpretation is that the dye can freely rotate within a cone that is defined by steric limitations; in this case, the measured anisotropy corresponds to a cone half-angle of around 25°. However, the bulk anisotropy measurements only test the freedom of the dye molecules to reorient on the time scale of the fluorescence lifetime (a few nanoseconds). In order to measure the time evolution of the orientation of a connector monomer (a few seconds) and to probe the time before photobleaching of the dye, as well as its brightness and orientation, we had to perform single-molecule fluorescence experiments.

*Single-molecule fluorescence observations of stalled complexes*
The long-term rotational rigidity of the dye, the rigidity of immobilization of the packaging prohead complexes, and the suitability of the dye orientation relative to the rotation axis can all be studied in single-molecule experiments (Ha et al., 1996). To this end, we measured the polarization of the emitted fluorescence from proheads labeled with single-dye molecules. The labeled, stalled prohead complexes are attached to the surface of the micro-fluidic chamber and illuminated in prism-type total internal reflection geometry, with a green laser (\(\lambda = 532\) nm). A schematic of the experimental setup is shown in Figure 4 (see also Materials and Methods). The fluorescence signal is collected by the objective and spatially separated into two perpendicular polarization components that are simultaneously imaged on a charge-coupled device camera. An overlay of the two polarization images yields a time-dependent fluorescence signal for each fluorophore in two orthogonal polarizations, thus allowing us to track the changes in relative orientation of a single dye in real time.

\(\phi 29\) packaging complexes stalled by incubation with ATP-\(\gamma\)S and therefore incapable of enzymatically driven rotation were used to assess the dye suitability. Figure 5 shows the fluorescence signal of a single fluorophore attached to a stalled complex, with the excitation polarization being rotated between the two orthogonal directions, henceforth called horizontal and vertical, with a frequency of 0.7 Hz. Fluorescence signals were integrated for 75 ms, and a typical time before photobleaching of the dye molecule was 10 s. The measured intensities in the two emission channels (perpendicular polarizations) are shown in black and red. The change in excitation polarization can be seen through an oscillation of the intensity in each channel. We observed a correlated signal (which is a very clear indication for a stable orientation of the dye) in more than 50% of the test experiments. Furthermore, it is important to note that the average signal intensity in the two channels remains almost constant over the lifetime of the dye, which shows that the dye keeps a stable orientation on timescales larger than the integration time. In contrast, a free diffusive rotation of the dye on the timescale of the integration time would lead to anti-correlation of the vertical and horizontal fluorescence signals. Anti-correlation (and therefore free rotation of the dye molecule) is likely caused by imperfections in the surface attachment of the stalled complexes. Finally, the fluorescence disappears in a single step at around time \(t = 22\) s, indicating the presence of a single dye that was photobleached. These experiments (and an additional control discussed in the Appendix to this chapter), demonstrate that the polarization of the emitted fluorescence is an accurate reporter of the position of the connector protein, and that our instrument is indeed capable of detecting the changes in the fluorescence polarization, and hence the connector orientation, due to rotation on the single-fluorophore level. Constrained by the properties of single dyes and the camera, we can measure connector
rotation rates from about 0.1–2.5 Hz, which corresponds to actual signal frequencies of 0.2–5 Hz due to the fact that the dipole emission has a 2-fold rotational symmetry. Current models for rotation predict a frequency within our detection bandwidth.

*Single-molecule fluorescence observations of packaging complexes*

Simpson *et al* (Simpson et al., 2000) have proposed a model for connector rotation based on the symmetry of the capsid and the connector. The connector is a homododecamer that sits at the 5-fold portal vertex of the icosahedral capsid. This 12/5 symmetry mismatch dictates that the relative alignment between connector and capsid is recapitulated with every 12°-rotation between the two structures. Furthermore, in vitro packaging experiments have measured the DNA-packaging rate at saturating ATP to be ~100 bp/s (Smith et al., 2001), and biochemical studies have shown that the step-size of the motor is 2 bp per ATP (Chemla et al., 2005; Guo et al., 1987). Given these figures, this rotation hypothesis predicts a rotation frequency of 1.65 Hz. Such rotation would result in a frequency of 3.3 Hz for the measured fluorescence signal in our experiment. Alternatively, if the connector were to track the helicity of the DNA (Hendrix, 1978), the rotation would be 36° per basepair and the resulting signal frequency would be ~20 Hz. In our experiments, the packaging velocity was reduced 2- to 4-fold by simply reducing the ATP concentration (Chemla et al., 2005) in order to reach a signal frequency below 5 Hz, which is within our experimental time resolution.

For actual packaging experiments, we illuminated the dye with homogeneously polarized light by time-sharing two different excitation polarizations. The emitted fluorescence was then detected in two orthogonal detection channels (see Materials and Methods for details). If the dye were to rotate in the plane parallel to the chamber surface around the DNA being packaged, the intensity of horizontally polarized light would oscillate according to a sine-squared function, while the vertically polarized intensity would oscillate in the same manner with a phase shift of 90°. The two channels, therefore, would show an anti-correlated modulation of the fluorescence intensity. (If the rotation axis is not perpendicular to the chamber surface, other phase shifts might be observed. We performed extensive simulations that suggested various limitations to our detection ability, and these are discussed below.)

Figure 6 shows typical examples of the single-molecule fluorescence signal during DNA packaging by the φ29 motor complex. Simultaneously, packaging activity is observed using a magnetic bead attached to the free end of the viral DNA. The red and black time traces show the fluorescence
intensity detected in the two perpendicular polarization directions—horizontal and vertical, respectively. Figure 6a shows the signal for the mutant 170C. At $t = 0.5$ s, the signal of two single fluorophores can be seen. The first fluorophore bleaches after about 4 s, the second after about 19 s. This multi-step bleaching demonstrates our ability to quantify the number of dyes observed, and in rotation experiments only single-fluorophore traces were analyzed. The ratio of the intensities in the black (vertical) and red (horizontal) channel indicates that the first fluorophore is aligned almost horizontally, while the second fluorophore is at an angle of about 45° in between the two polarization directions. After about 100 s, scattered light from the magnetic bead, which is attached to the free end of the DNA, becomes visible in the trace. As the prohead continues packaging the DNA, the bead is pulled further into the evanescent field of the excitation by the green laser until it touches the surface at about $t = 165$ s. The packaging, assuming an approximately 10-kb tether, was therefore about 60 bp/s, consistent with bulk and optical tweezers measurements at this ATP concentration. It should be noted that we did not observe beads that were slowly pulled toward the packaging complexes in control experiments without ATP. Therefore, this behavior can be, without doubt, identified with the ATP-dependent DNA packaging of the motor complex. The center of the bead and the center of the initial single-dye fluorescence signal are within one pixel from each other, demonstrating the colocalization between the fluorophore and the packaging complex. By considering the density of fluorescent spots on the surface when using highly overlabeled proheads, we estimate that over 98% of such colocalized events are due to a tethered bead and dye molecule attached to the same packaging prohead (see Appendix for details). We have recorded more than 50 of these single dye/bead colocalized traces from six different mutants (see Appendix for a complete list).

The data were analyzed in various ways. In short: First we used a normalized sliding correlation function that measures the correlation of the two perpendicular signals over a window of variable size. If the motor rotated during packaging, our simulations (including noise) suggest that one out of four traces should give an average correlation coefficient less than −0.3 for many seconds, assuming a random dye orientation. We never observed such correlation coefficients for an extended period of time (several seconds) in any of the data collected. Second, we checked for periodicities in the channels by looking at Fourier transforms and cross-correlations. Our simulations predict that motor rotation generates a periodicity in more than 90% of the traces in the accessible frequency range, assuming random orientation of the dye molecule relative to the axis of rotation and random orientation of this rotation axis with respect to the optical axis. We did not observe this periodicity in a single trace.
2.4 Discussion

We have studied a possible rotation of the bacteriophage φ29 portal motor protein with respect to the capsid on the single-molecule level during DNA packaging. With six different connector mutants, we did not observe a single trace resembling a signal expected for connector rotation. These results permit us to rule out with very high probability (see below) the compressive-ratchet model for connector rotation proposed by Simpson et al. (Simpson et al., 2000) that involves a rotation by 6° per basepair. At the experimental ATP concentration of 25–50 µM, this rotation would lead to a rotation frequency in the fluorescence signal of about 1–2 Hz, which is easily detectable with our instrument. Similarly, a rotation model in which the DNA is wrapped around the external surface of the connector, with rotation providing an indirect translocating force (Turnquist et al., 1992), would generate a rotation rate of 3° per basepair, also within the detection range of this experiment. A rotation rate below our detection limit is very unlikely, as a frequency of 0.1 Hz (our lower detection limit) would already correspond to as few as 0.6° per basepair—which would not fit any current model for rotation. On the other hand, we would be capable of observing a rotation of the connector if it were to follow the DNA double helix pitch in a nut and bolt fashion (10.5 bp per 360°), which would yield a rotation frequency in the fluorescence signal of 5 Hz at 25 µM ATP. While we can rule out complete rotation of the connector relative to the shell, we cannot rule out partial rotation over a small angle followed by return to the original position. In order to detect such transient rotation that would be consistent with models of connector flexure, polarization sensitive fluorescence correlation spectroscopy experiments would have to be performed.

There remains a small uncertainty about packaging motor rotation due to the unknown orientation of the dye molecule with respect to the putative rotation axis of the connector and due to the lack of absolute labeling specificity. Simulations of dye emission (unpublished data) show that there are orientations of the dye where neither a correlated nor an anti-correlated signal can be observed, even if the connector is rotating. Given the random orientation of the rotation axis with respect to the substrate and of the dye axis with respect to the rotation axis, this situation should happen in about one out of ten traces for the signal levels shown. Here, we reduced the likelihood of an unfavorable orientation of the dye molecule with respect to the putative rotation axis by investigating six different mutants. We cannot rule out that all mutants result in unfavorable dye molecule orientations, although we consider this possibility highly unlikely.
Considering the labeling specificity, there is currently no method to separate the connector protein (gp10) from the capsid protein (gp8) and re-assemble them again. Therefore, we have to label the connector in intact prohead particles, which might allow some dyes to attach to the nonrotating capsid. However, the fact that the capsid does not contain cysteines allowed us to achieve a specificity of more than 80% of dye on the connector of the 170C as observed on denaturing (SDS-PAGE) gels (Figure 3). We can also rule out that our selection of actively packaging complexes leads to a selection of nonlabeled complexes, since we have shown that labeling does not affect packaging efficiency or speed (Figure 3b and unpublished data). As a result, about four out of five of the observed dyes can be expected to be attached to the connector.

While no rotation of the connector was observed experimentally, does that mean that the connector does not rotate? In order to answer this question, we have performed a mathematical analysis of the remaining uncertainties. This analysis is described in great detail in the appendix. In brief, if one includes the uncertainty of unfavorable orientation, colocalization, labeling specificity, and rigidity of attachment, one can rule out connector rotation with more than 99% probability.

In our experimental design, we were able to eliminate several limitations of previous efforts to combine single-molecule fluorescence and magnetic tweezers (Shroff et al., 2005). Magnetic beads scatter a great deal of light and are therefore not compatible with most single-molecule fluorescence experiments. However, as demonstrated by our results, if one uses highly localized excitation fields, like an exponentially decaying fluorescence excitation used here, they can be combined with single-molecule fluorescence. For a processive motor like the bacteriophage φ29 packaging complex, the disadvantage of using spatially separated trapping and fluorescence detection can be overcome by the colocalization of single-molecule fluorescence and bead fluorescence after the bead has been pulled close to the surface. Our experiments have shown that beads kept at a distance of larger than ~1 µm do not introduce a significant scattered signal at the Cy3 fluorescence wavelength. They do block part of the emitted fluorescence light, but for tethers longer than 1 µm and bead sizes of 1 µm, the detected dye fluorescence is calculated to decrease by less than 10%. Another advantage of this setup is the possibility of parallel observation. In some preparations, we could observe the packaging of more than five complexes simultaneously.

In previous single-molecule fluorescence studies on biological systems, the fluorescence signal itself was the only evidence of biologically relevant activity. Therefore, there had to be a characteristic and expected feature in
the single-molecule fluorescence signal and sufficient statistics to confirm that the biological system is the cause for the observed signal. Obtaining these statistics can necessitate the observation of tens of thousands of fluorophores (Adachi et al., 2000). Here, we have presented a method that overcomes this problem. We can select for fluorophores that are attached to independently monitored active biological systems and observe their single-molecule fluorescence. In the present application, we can colocalize a fluorophore with a translocated bead. This leads to a 98% confidence that the observed fluorescence signal originates from a packaging prohead. The setup, therefore, opens exciting opportunities for the study of a number of different systems, such as RNA polymerase transcription initiation or elongation complexes, ribosomes, spliceosomes, or as shown here, nucleic acid translocation or packaging complexes in real time and with high resolution.

In conclusion, we were able to test the connector rotation hypothesis, a long-standing prediction of several DNA packaging models. Our single-molecule experiments exclude with very high probability (more than 99%) the predominant model that the connector rotates with respect to the capsid. Having established that the connector does not rotate during packaging, it is important to ask how DNA is driven into the capsid. A model consistent with all experimental data was proposed recently by Chemla et al. (Chemla et al., 2005). In this model, ATP binding, hydrolysis, and release of products induce conformational changes in the ATPases that are directly involved in the translocation of the DNA. Specifically, the translocation step of the DNA is triggered by, or performed by, the ring of ATPases via a conformational change that follows release of phosphate after ATP hydrolysis. Here, we add to this model the idea that the connector could function as a valve to prevent DNA from leaking out.

Although the connector does not rotate, the aforementioned symmetry mismatch between the DNA and the packaging complex may still require rotation – that is, rotation of the DNA relative to the rest of the complex. Alternatively, perhaps the DNA is not rigidly coupled to the portal. A single-molecule mechanical experiment supports this idea. Atomic force microscopy of free connectors (Muller et al., 1997) found that the connector is elastic at a force of 100-150 pN – the same scale as the forces involved in packaging (Rickgauer et al., 2008; Smith et al., 2001). Such conformational flexibility, which may be the same flexibility that is required for packaging (Cuervo et al., 2007) may allow the connector’s DNA-binding domains to accommodate the mismatches, and hence avoid accumulating any angular misalignment (Roos et al., 2007; Yu et al., 2010). Finally, the possibility that translocation could be driven by a (nonrotating) compression/extension ratchet mechanism is an intriguing idea, but one for which there is no direct
experimental evidence to date and that is distinct from the mode of action established for certain AAA+ -related ATPases.

If the connector is not tightly coupled to the DNA, what is its role in packaging? At the very least, the connector is essential for proper assembly of both the capsid (Fu and Prevelige, 2009; Guo et al., 1991a) and the packaging motor. In φ29, the connector N-terminus is known to bind to the pRNA (Atz et al., 2007; Robinson et al., 2006; Sun et al., 2006; Xiao et al., 2005) and may simply serve as a temporary platform for pRNA binding. After packaging is complete, the rest of the packaging motor disassembles, leaving the connector responsible for preventing DNA ejection after departure of the ATPase and before docking of the tail. This may explain the portal mutant phenotypes in which the number of stably packaged particles is drastically reduced (Cuervo et al., 2007; Isidro et al., 2004b). Mutations homologous to the tunnel loop mutations of SPP1 were produced in φ29 and single packaging events were monitored in optical tweezers. The characteristics of packaging such as velocity and force sensitivity were unchanged, suggesting that the mutations must disrupt the process at a later stage, such as in preventing post-packaging DNA leakage (Rockney Atz, Shuhua Ma, K. Aathavan, Carlos Bustamante, Jiali Gao, Dwight L. Anderson and Shelley Grimes, in preparation). The connector may assume a similar duty during packaging itself: It could act as a one-way valve, allowing DNA to be packaged into the capsid but restraining its exit (Hugel et al., 2007). The spring-like shape of the connector suggests, indeed, that through compression and expansion, the connector may act as a “Chinese finger trap” (George Oster, personal communication) allowing the passage of the DNA in one direction during packaging but preventing its exit in the reverse. In fact, a recent paper by Jing et al. (Jing et al., 2010) demonstrated that DNA driven by an electric potential can pass in only one direction through connectors inserted into a lipid bilayer membrane – the direction corresponding to packaging into a prohead – and provided evidence that, in partially packaged particles, the connector alone can prevent DNA escape from the prohead. Such a one-way function would assist the packaging motor in maintaining processivity when packaging against high loads or high internal pressures as the capsid fills.

A complete understanding of the coupling that occurs between the ATPase, the pRNA, and the connector substructures is needed to refine our picture of the molecular mechanism of this powerful motor. Given the apparent universality of connector structure, the elusive details of connector function in φ29 may well prove to be fundamental to a wide variety of viruses.
2.5 Materials and Methods

Surface preparation

For all experiments, we used quartz slides (Finkenbeiner Inc.) cleaned in piranha at 60 °C overnight, rinsed with purified water (Barnstead, E-pure), sonicated in 2% Hellmanex (Hellma), rinsed again, sonicated, and stored in pure water. Slides were blown dry with nitrogen immediately before being placed in 1 ml vectabond (Vector Laboratories) and 100 ml acetone for 5–10 min. The slides were then washed in 100 ml water, slowly pulled out of the water bath such that no water remained on the hydrophobic surface, and placed in a wet box.

100 µl of a mixture of 3 mg Biotin-PEG-NHS (Mw3400, Nektar Therapeutics), 80 mg mPEG-SPA (Mw5000, Nektar Therapeutics), and 550 µl 0.1 M bicarbonate buffer was then placed on each slide and kept in the dark for 3 h. Afterward, the slides were rinsed thoroughly with water, dried with nitrogen, and assembled into a flow chamber by placing a Nescofilm gasket in between the quartz slide and a cover slip and heating for 3 min at 100 °C.

The assembled chamber was then rinsed with 1 ml phosphate-buffered saline and incubated with 0.2 mg/ml streptavidin (Roche) for about 20 min. After being rinsed again with phosphate-buffered saline and incubated with biotinylated antibodies against the capsid protein, gp8, (0.1 mg/ml) for 25 min, the chamber was rinsed with 1 ml 0.5× TMS (25 mM Tris-HCl, 50 mM NaCl, 5 mM MgCl₂) and then with 250 µl of buffer XS (800 µl of 0.5× TMS, 10 µM ATP, 10 µM ATP-γS, 0.2 mg/ml BSA, 2 mg/ml glucose, 1% w/v beta-mercaptoethanol, 0.02 mg/ml catalase, 0.1 mg/ml glucose oxidase, and 0.8 µl of RNase inhibitor (SuperaseIn, Ambion).

Protein preparation, mutation, and labeling

With the exception of prohead particles, components for the φ29 in vitro packaging system (DNA-gp3, gp16, 120-base pRNA) were produced as previously described (Grimes and Anderson, 1997; Reid et al., 1994).

Prohead particles were produced in E. coli by overexpression of prohead structural proteins in HMS(DE3)pAR 7-8-8.5-9-10 (Guo et al., 1991a). Two wild-type cysteines (C76 and C265) in the φ29 connector, gp10, were replaced by serines using standard site-directed mutagenesis to produce a cysteine-free clone, and individual amino acids in gp10 were replaced with cysteines to produce a library of single-cysteine mutants. Particles were produced by induction of mid-log cultures grown in LB media with 0.5 mM IPTG for 2 h. Cells were pelleted and re-suspended in a lysis buffer.
containing 50 mM Tris HCl (pH 8.0), 20 mM NaCl, 2 mM EDTA, 2 mM DTT, and 10 mg/ml lysozyme. After a 20-min incubation to produce spheroplasts, MgCl$_2$ was added to 4 mM final concentration and DNase was added to a final concentration of 10 µg/ml to digest cellular DNA. Complete lysis was achieved with the addition of deoxycholate to 0.25% w/v. Extracts were clarified by centrifugation, and prohead particles were purified on 10%–40% w/v sucrose zonal gradients (45 kilo-rotations per min, 45 min, 20 °C) in a SW55 rotor (Beckman) buffered with 1× TMS buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl$_2$ [pH7.8]). Particles were collected, concentrated by pelleting, and re-suspended in 1× TMS.

Labeling was conducted by adding an equal volume of Cy3-maleimide in H$_2$O to prohead samples to provide the appropriate molar amount of dye with respect to available connector cysteine. After labeling for 1 h at room temperature, particles were purified away from free dye by sucrose gradient sedimentation (5%–20% w/v sucrose in 1× TMS, 45 K, 30 min, 20 °C). Particles were pelleted and re-suspended as above. Particles were quantified and checked for purity and labeling efficiency by SDS-PAGE and for DNA-packaging activity (Figure 3).

**Stalled complex preparation**

Labeled proheads were reconstituted with 120-base pRNA at a ratio of 10 pRNAs/prohead by mixing 2 µl of pRNA (0.07 mg/ml) with 2 µl of labeled proheads (1 mg/ml) for 15 min in 0.5× TMS. Reconstituted proheads were then added to a packaging reaction containing 2 µl of biotinylated DNA (0.44 mg/ml) and 2 µl of ATPase gp16 (0.025 mg/ml) in a final volume of 18 µl of buffered in 0.5× TMS (for a final ratio of two proheads : one DNA-gp3:1 gp16). After 5 min, packaging was initiated with 4 µl of ATP (250 µM). After 60 s, 2 µl of ATP-γS (1 mM) was added to stall the reaction. Magnetic beads (MyOne Dynabeads, Invitrogen) were prepared by washing three times in 0.5× TMS, then blocked by adding 2 µl of beads to 25 µl of 2 mg/ml BSA in 0.5× TMS. Blocked beads were then treated with 0.5 µl of RNAse inhibitor. Freshly prepared stalled complexes from above (4 µl) were bound to beads by mixing 0.1 µl of Superase Inhibitor, 3 µl of BSA (10 mg/ml), 0.5 µl of ApaLI (10 U/µl) in 80 µl of buffer X (0.5× TMS, 10 µM ATP, 10 µM γ-S ATP). After gentle mixing, the sample was incubated for 1 h for the ApaLI restriction digestion which cleaves at both the extreme right and left ends of the DNA, to reduce the presence of biotinylated DNA-gp3 on the bead surface not associated with stalled-packaging complexes, which can form nonspecific tethers due to the stickiness of gp3. (Left ends of packaged DNA-gp3 are in the prohead and protected from digestion.) Using a magnet, we washed the magnetic beads three times with buffer X. These washes removed free and cut DNA-gp3 ends, all free dye, and proheads that did not initiate packaging.
Finally, the beads were flowed into the chamber and incubated for 10 min. To restart packaging, 0.5× TMS buffer containing 0.2 mg/ml BSA, 2 mg/ml glucose, 1% w/v beta-mercaptoethanol, 0.02 mg/ml catalase, 0.1 mg/ml glucose oxidase, and 50 µM ATP was flushed into the chamber.

**Instrument design**

During SMFP, the sample is illuminated in prism type total internal reflection geometry with a green laser (532 nm, CrystaLaser). There are two ways for a dye molecule to report on the rotation of a macromolecule: First, one could excite the dye molecule with linearly polarized light. The emitted fluorescence intensity would then be proportional to the square of the scalar product of dipole orientation and polarization direction. For a rotating dye molecule in the plane of the evanescent field, the emitted intensity would oscillate between a maximum and minimum within a given polarization as the molecule changes its orientation. A second option for using a dye molecule as a direction sensor is to illuminate with both horizontal and vertical polarization (with equal intensities, i.e., homogeneous polarization) but detect the polarization state of the emitted fluorescence. Here, the angle between emission dipole and polarizer in the detection path becomes important. For a molecule rotating in the polarization plane, the result would be an anti-correlated signal between the vertical and horizontal polarization. We chose to use the second option, since an anti-correlated signal cannot be confused with other events such as blinking or changes in molecular brightness. In order to achieve an illumination with homogeneous polarization, the light of the laser is coupled into an electro optical modulator (Linos Photonics, Incorporated) that switches between two perpendicular polarization directions with a frequency of 10 kHz. This switch is orders of magnitude faster than the integration time during the experiment, and therefore, only homogenous polarization is observed. The two perpendicular polarizations are then split by a polarizing beam splitter (PBS). The resulting beams are both brought to the chamber in s-polarization, but from perpendicular directions—one from the side and one from the top. For this reason, we used a custom-made prism with two input ports. We checked the light intensity from both directions by comparing the signal scattered by beads attached to the flow chamber surface and adjusted the intensities to be equal in the center of the field of view; the intensities varied by less than 50% across the field of view.

The fluorescence light is collected by a high NA objective (Nikon, 1.2 NA) and separated in two perpendicular polarization components by a PBS. The two beams are spatially offset and recombined by another PBS. The two beams are then focused on a CCD camera (Cascade 512B, Photometrics, Roper Scientific), such that two images, one for each polarization, can be read out.
simultaneously. One pixel on the camera (physical size 14 µm) corresponds to about 400 nm. Because of the huge dilution of phages, and therefore dyes on the substrate (less than one dye per 10 µm²), we used hardware binning by 3 × 3 pixels. Furthermore, since the fluorescent spot was not always centered on one point, we added up to four adjacent points for signal optimization. Two band-pass filters (580BP50, Omega Optical) are used to separate excitation and LED illumination from the single-molecule fluorescence signal. At the same time, the sample is illuminated with a red LED to observe the magnetic beads that are pulled away from the surface by two magnets. The red light scattered from the magnetic beads is separated from the fluorescence light with a dichroic mirror (630 DCLP, Omega Optical) and detected on a separate CCD camera (Watec 902C; Watec). In addition, an epi illumination can be used to focus onto the chamber surface without illuminating (and therefore bleaching) the surface.
Figures

Figure 2-1. Schematic of experimental geometry (not to scale). Stalled-packaging complexes are attached to the surface of the flow-chamber via biotinylated antibodies to the prohead major capsid protein, gp8. The dye molecule, attached to the connector, is excited via an evanescent wave using total internal reflection. The biotinylated free end of the DNA is attached to streptavidin-coated superparamagnetic beads that are pulled away from the fluorescence excitation by a magnetic field gradient that is created by a pair of magnets next to the objective. The magnetic beads are illuminated for video microscopy using a red LED; both the signals for bright-field and fluorescence images are collected by a high NA microscope objective.
Figure 2-2. Connector structure with position of dye labels. A schematic of the structure of the connector, based on the crystal structure by Simpson et al. (Simpson et al., 2000) is shown. The positions of the residues that were mutated to cysteines and investigated with single-molecule fluorescence are indicated.
Figure 2-3. Dye labeling and DNA packaging of φ29 proheads (A) SDS-PAGE of 170C-connector-mutant proheads. Protein stain of proheads shows the structural components gp8 (capsid), gp8.5 (fiber), gp10 (connector), and gp7 (scaffold) in lane (a). Fluorescence scan of the gel showing labeled proheads with various amounts of dye per gp10 monomer used in labeling reaction: 1 dye per gp10, lane (b); 0.5 dyes per gp10, lane (c); 0.25 dyes per gp10, lane (d); 0.125 dyes per gp10, lane (e); 0.0625 dyes per gp10, lane (f); and no dye, lane (g). The bands in the fluorescence scan with no match in the protein stain originate from highly reactive but quantitatively minor *E. coli* proteins. (B) DNA packaging tested by nuclease (EcoRI) protection assay using the labeled proheads from (A). Lane (a) shows input DNA-gp3; lane (b) shows a negative (no ATP) control. Packaged DNA is protected from nuclease digestion. Packaging activity is unaffected by dye labeling when compared to a 193C packaging control, lane (c). (B) Shows labeled proheads from (A), ranging from 1 dye per gp10, lane (d); 0.5 dyes per gp10, lane (e); through to no dye, lane (i).
Figure 2-4. Schematic of the experimental setup. Combined SMFP and magnetic tweezers setup. M, mirror; P, pinhole; λ/2, λ/2 plate; PBS, polarizing beam splitter; EOM, electro optical modulator; S, shutter; F, filter; Pol, polarizer; DC, dichroic mirror. Dashed components can be removed and are solely used for alignment purpose.
Figure 2-5. Fluorescence polarization studies of dye-labeled, stalled-packaging complexes. Dye-labeled, stalled-packaging complexes were attached to the surface of a flow chamber and excited using the total internal reflection microscope. The excitation polarization was rotated between s- and p-polarization with a frequency of 0.7 Hz. The emitted fluorescence was separated into s- and p-polarization, respectively, and simultaneously detected (black and red). The dye bleached after 22 s. The integration time per data point was 75 ms. a.u., arbitrary units.
Figure 2-6. Fluorescence signal from packaging complexes. The graphs show the fluorescence intensity of vertical (black) and horizontal (red) polarization detected simultaneously. One example for each of the six investigated mutants is shown. The data were recorded with an integration time of 75 ms and three-points smoothing was applied. The camera background (closed shutter) was subtracted. The signal was normalized using the fluorescence intensity of the bead after packaging was completed. The traces are vertically shifted for clarity. The inset shows a zoom of the data. Here, the unfiltered data is displayed (scatter) together with the three-point sliding average. (A) Fluorescence recorded for a single complex of mutant 170C. The shutter is opened after about half a second. After \( t = 4 \) s the first dye bleaches and after \( 19 \) s the second. After about \( t = 100 \) s the fluorescence signal again starts to increase, since the magnetic bead is pulled into the evanescent field as the prohead reels in the DNA. The signal increases in an exponential fashion as the fluorescent bead samples the intensity profile of the evanescent wave. Similar behavior was observed for (B) 97C; (C) 168C; (D) 189C; (E) 190C; and (F) 260C.
Appendix

Numerical modeling of fluorescence polarization signal of packaging

The dye is fixed on the putatively rotating connector with some angle relative to the connector axis. The connector is coaxial with the capsid, which has its own orientation relative to the laboratory frame. (Figure A1). Some combination of angles may not generate an excitation-polarization-dependent signal. For example, if the dye dipole is pointed in the same direction as the connector axis, there will be no change. Other combinations may generate a correlated signal, rather than an anti-correlated one. The expected signal was computed for all possible combinations of attachment angles; a realistic (Poissonian) noise model was included. For a few examples, see Figure A2. We developed analysis algorithms to identify rotating connectors based on this simulated data. We found that over 90% of the angle combinations produced a clear, periodically-varying signal. To our surprise, about 75% of the simulated traces were correlated – that is, the two channels rise and fell together. Only about 25% produced an anti-correlated signal.

Various sources of noise such as dye blinking, laser power fluctuations, and photon counting noise would generate signals that are uncorrelated or positively correlated. In contrast, the only feasible cause of a negatively correlated signal is the rotation or reorientation of the dye molecule. An anti-correlated signal over a certain period of time is a sufficient indicator of dye rotation. To check for anti-correlation we applied a standard sliding correlation function to our data. The correlation function is given in Equation A1.

\[ Corr_{12} = \frac{1}{n \sigma_1 \sigma_2} \sum_{i=1}^{n} (R_{1i} - \mu_{1i})(R_{2i} - \mu_{2i}) \]  

(A1)

\( R_{1i} \) and \( R_{2i} \) are the measured values in the two polarization directions, respectively.

\( \mu_{1i} \) and \( \mu_{2i} \) are the mean values of the measured values and \( \sigma_1 \) and \( \sigma_2 \) are the standard deviations. This function was calculated for a sliding window of about 30 points which corresponds to 2.25 seconds. In addition, we applied a threshold criteria which was necessary in order to differentiate anti-correlation from random signal fluctuations. Anti-correlation was defined as an average correlation coefficient of less than -0.3 for a duration exceeding two seconds. The probability to detect a rotating dye-molecule with this
function is already about 25%, for 18 traces the probability to detect at least one trace displaying anti-correlation would be larger than 99%.

Column A in Figure A3 shows the sliding correlation function of the six traces shown in Figure 6 of the main paper, namely examples of mutants 170C, 97C, 168C, 189C, 190C and 260C (from top to bottom). The sliding correlation function does not exceed the anti-correlation threshold in any of these traces. Furthermore, we did not find a single packaging complex that showed anti-correlation. Moreover, to check for any other phase shift, we calculated the cross-correlation function for the two perpendicular polarization directions (Equation A2) and its power spectrum.

\[
XCorr_i = \sum_{t=-n}^{n} R_{i:t} R_{i:t+1} \tag{A2}
\]

The results for the six traces from Figure 6 of the main paper are shown in column B and C of Figure A1, respectively. The number of points sampled during the lifetime of the dye is often too small for computing a meaningful power spectrum. Therefore, we concatenated four repeats of the measured signal of each dye molecule in order to calculate the power spectrum. This procedure yields a good sampling of the frequency range of interest, but in turn produces peaks at very low frequency (around 0.05 Hz). Those peaks in trace one, four and five from the top correspond to a period of about 20 seconds, which is longer than the lifetime of the dye and therefore an artifact of the catenation. The peak from the mutant 168C (third trace from top) has a period of 8 seconds, which corresponds very well to the lifetime of the dye and is therefore also no indication of rotation.

While no signal of rotation was observed for any packaging complexes, the statistical analysis of the control experiments (compare to Figure 5 and description in the main text of Chapter 2) demonstrate the viability of the method (Figure A4). Here, again we show the sliding correlation function, the cross-correlation as well as the power spectrum of the cross-correlation in A, B and C, respectively. The excitation polarization was rotated with a frequency of about 0.7 Hz; this frequency is found very clearly in the power-spectrum of the cross-correlation function (please note the almost one order of magnitude larger scale of the power-spectrum, as compared to the scale in figure A3C).
The rigidity of the dye-molecule attachment is crucial to the experiment, since the dye-molecule is used as a probe for the connector orientation. Additional evidence for this rigidity is presented in Figure A5. Here, the fluorescence signal of a phage labeled with two dye molecules is shown. We used conditions for the connector labeling resulting in rather low labeling efficiencies of on average less than one dye per prohead. However, we did observe a handful of events with a two-step bleaching, indicating the presence of two dye molecules.

The two dye molecules have different contributions to the two perpendicular polarizations, indicating that the orientation of the molecules is both, distinct and fixed on the timescale of the experiments (see Figure caption for details).

**Estimate for the probability of non-rotation**

The probability for non-rotation is estimated from the following four arguments:

1.) Co-localization is defined as an observed dye (the single molecule signal at the beginning of a fluorescence trace) being attached to the same phage that packages the DNA attached to the observed bead (the exponential rising signal at the end of a fluorescence trace). Due to the fact that not all complexes are labeled, as well as that not all complexes are packaging, there is a possibility that such co-localization arises not from a single complex, but from a pair of complexes. One of the complexes would have to be labeled, but not packaging, and the other would have to be not-labeled, but packaging. The probability for this co-localization is estimated in the following way:

We assume that the distance between phage proheads on the surface follows a Gaussian distribution. The average distance between proheads is estimated by using heavily labeled proheads (about 10 dye molecules per prohead) following the same experimental procedures as for the experiments with under-labeled particles. These experiments yielded an average distance of proheads of 5 points (a point is 3x3 pixels – see Instrument Design in main text for details) and a standard deviation of 1.5 points (square root of half the average distance). As a bead will not be co-localized if its final distance to the dye is more than 2 points, we obtain a confidence of 98% with the Normalized Gauss Distribution given in Equation A3 ($x_1=0$, $x_2=2$, $\mu=5$, $\sigma =1.5$).

$$p(t) = \frac{1}{\sigma \sqrt{2\pi}} \int_{x_1}^{x_2} e^{-x^2/2\sigma^2} \, dt$$  \hspace{1cm} (S3)
With $t_i = (x_i - \mu) / \sigma$ and the standard deviation $\sigma$ and the mean $\mu$.

2.) Given the experimentally determined signal to noise levels and assuming a random orientation of the dye molecule with respect to the rotation axis, 90% of the molecules should show a periodically varying signal in one or both polarizations. This result was obtained in extensive simulations, which will be published elsewhere.

3.) The labeling specificity of the dye to the connector is not 100%. We estimated the labeling specificity using SDS Page electrophoresis and a high resolution gel imager (see main text for details). The labeling specificity varied between the mutants investigated (see Table A6 for details). For the 170C mutant the specificity was 80% (see Figure 3A in main text).

4.) Another important factor in estimating the overall fidelity of the experiment is the rigidity of the attachment. For this estimate, we attached stalled packaging complexes to the surface of the flow-chamber and rotated the excitation polarization. In 50% of the traces we find a correlated signal between the two detected perpendicular polarization directions, which is expected for a fixed dye molecule. In about 50% of the traces we see an anti-correlated or uncorrelated signal which can be explained by a dye molecule that is free to orientationally diffuse or rotate faster than the time scale of the integration time, but hindered on the time scale of the fluorescence lifetime. Most likely this sub-population is caused by imperfections in the surface attachment, such as only single antibody attachments, yielding complexes that are free to re-orient during the time scale of the experiment. We conclude from this control that the probability that a dye molecule has a fixed orientation when the phage is not packaging is 50%.

In order to find an estimate of overall fidelity, we assume that each aforementioned probability is independent. Multiplying the four probabilities outlined above gives a fidelity for non-rotation of $Q(\text{non-rot}_{\text{event}}) \sim 35\%$ for the mutant 170C. We did not observe such a signal in any of the experiments. In other words, the probability that the connector rotates despite the absence of a periodically varying signal for any trace is $P(\text{rot}_{\text{event}}) = 1 - Q(\text{non-rot}_{\text{event}}) \sim 65\%$. Therefore, in order to estimate the probability that the connector might rotate in spite of our experimental data one simply needs to compute the overall probability:
\[ P(\text{rot}_{\text{total}}) = [P(\text{rot}_{\text{event}})]^\text{number of events} \]

For the case of the 170C mutant we observed 18 co-localized packaging events and therefore \(P(\text{rot}_{\text{total}})_{170C} < 0.001\). Therefore we can exclude rotation with a probability \(Q(\text{non-rot}_{\text{total}}) = 1 - P(\text{rot}_{\text{total}}) > 99\%\).

As mentioned in the main text, we cannot exclude that the dye on the mutant 170C is oriented in an unfavorable direction. Therefore, we investigated five more mutants at various positions and orientations in the connector (see Figure 2 of main text). An overview of the data and probabilities for each single mutant investigated is given in Table S4. None of the traces showed any hint towards rotation. The probability for the five mutants other than 170C is only lower because of the smaller sample size. We thus conclude that the connector of the \(\Phi 29\) portal packaging motor does not rotate during packaging.
Figure 2-A1. Consideration of capsid attachment and dye attachment geometry. Two angles describe the orientation of the capsid relative to the laboratory frame. Likewise, two angles define the dye orientation relative to the capsid.
Figure 2-A2. Simulation of expected signal for two randomly chosen angle combinations. On the left is a correlated signal; on the right is an anticorrelated one.
Figure 2-A3. Statistical analysis of the single-molecule data. Normalized sliding correlation function with n=30 (A), cross-correlation (B) and power spectrum (C) for the traces from Figure 6 in the main paper. The traces are from top to bottom of mutants 170C, 97C, 168C, 189C, 190C and 260C.
Figure 2-A4. Statistical analysis of the control. Normalized sliding correlation function with $n=30$ (A), cross-correlation (B) and power spectrum (C) for the control data from Figure 5 in the main paper.
Figure 2-A5. Fluorescence signal of packaging complex labeled with two dye molecules. A) Fluorescence signal of a packaging phage (260C) that was labeled with two dyes. The two dyes bleach at different times – the first at about 7s and the second at about 11s. The dye that bleached first can be seen to have contributed mostly to the intensity in the vertical polarization (black), whereas the longer-lived dye yields comparable intensities in both channels. Such an example shows that we can discriminate between different dye orientations. B) The sliding correlation function shows no sign of rotation.
<table>
<thead>
<tr>
<th>Mutant</th>
<th># of traces</th>
<th>label specificity</th>
<th>$Q_{\text{non-rot event}}$</th>
<th>$Q_{\text{total}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>170C</td>
<td>18</td>
<td>80</td>
<td>0.33</td>
<td>&gt; 99%</td>
</tr>
<tr>
<td>189C</td>
<td>9</td>
<td>80</td>
<td>0.31</td>
<td>98%</td>
</tr>
<tr>
<td>190C</td>
<td>9</td>
<td>80</td>
<td>0.26</td>
<td>98%</td>
</tr>
<tr>
<td>168C</td>
<td>8</td>
<td>70</td>
<td>0.31</td>
<td>95%</td>
</tr>
<tr>
<td>260C</td>
<td>7</td>
<td>70</td>
<td>0.31</td>
<td>93%</td>
</tr>
<tr>
<td>97C</td>
<td>5</td>
<td>60</td>
<td>0.26</td>
<td>80%</td>
</tr>
</tbody>
</table>

Table 2-A6. Number of co-localized single molecule traces (# of traces), estimated labeling specificity (label specificity), probability for one event showing rotation if the connector would rotate ($Q_{\text{non-rot event}}$) and total probability for non-rotation ($Q_{\text{total}}$) for all the different investigated mutants. Note: The probability for the five mutants other than 170C is only considerably lower because of the smaller sample size.
Chapter 3. Approaches to Single-Molecule Measurement of DNA Angle

3.1 Abstract

The previous chapter discussed a fluorescence-based experiment to detect rotation of a protein component of the packaging complex. Another possible site of rotation in the packaging complex is between the motor and the DNA; if the prohead is fixed, the DNA should rotate. Detecting rotation of a soft polymer like DNA imposes certain experimental constraints that are best satisfied by a tweezers setup, as will be explained below. These constraints motivated an instrumental advance in the form of a hybrid magnetic-optical tweezers that can manipulate with force and torque simultaneously. In order to accommodate the biochemical parameters of viral initiation and packaging, several new strategies for attaching handles to DNA were also pursued.

3.2 Theoretical considerations in measuring DNA angle

Introduction to DNA Mechanics

The relationship between the tension applied to a DNA molecule and the resulting molecular end-to-end extension can be evaluated using the wormlike chain model, in which the energetic cost of bending the polymer is parameterized by the extensional persistence length $P$. This persistence length describes the length scale over which the polymer's direction is correlated. The persistence length is related to the bending modulus $B$ in the following way:

$$ P = \frac{B}{k_B T} $$

However, DNA is not one-dimensional, so other modes contribute to the state of the molecule. In particular, two other energy terms must considered: stretching, parameterized by the stretch modulus $S$, and twisting, characterized by the torsional modulus $C$ (Moroz and Nelson, 1998).

$$ E = k_B T \int_{\text{contour}} \left( \frac{B}{2} \left( \frac{\partial t}{\partial s} \right)^2 + \frac{C}{2} \theta^2 + \frac{S}{2} \psi^2 \right) ds $$

In this equation, $s$ is the contour parameter, $t$ is the tangent vector at $s$, $\theta$ is the local twist rate, and $\psi$ is the local stretch rate. Accepted values of the parameters at physiological conditions are: $B = 230$ pN nm$^2$, $C = 440$ pN
nm², \( S = 1200 \) pN. The ratio of the torsional modulus \( C \) to the DNA length \( L \) is a measure of the stiffness of the DNA under torque. The angular displacement \( \theta_{\text{DNA}} \) from equilibrium within a DNA molecule across a length \( L \) is related to the resulting torque in the molecule:

\[
N = -\frac{C}{L} \theta_{\text{DNA}}
\]

Using the equipartition theorem, the variance in the azimuthal angle \( \theta \) of the DNA about its mean angle \( \theta_0 \) is given by:

\[
\frac{1}{2L}((\theta - \theta_0)^2) = \frac{1}{2}k_B T
\]

For the characteristic length scale of a packaging experiment, where \( L \sim 1 \) µm (and decreases to zero in 30 seconds during packaging), the variance of \( \theta \) is about \( 12\pi \) radian².

**Hydrodynamics of the DNA tether system**

The dynamics of these fluctuations depend not only on the elastic properties of DNA but on the hydrodynamics of the entire system including the label used to report the angle of the DNA. Reynolds pointed out that the ratio of inertial forces to viscous forces characterizes the relative importance of those two physical effects, and it became known as the eponymous Reynolds number \( Re \). See the Appendix of Chapter 4 for a more detailed discussion of hydrodynamics. For water at 25°C, the ambient temperature for these experiments, the density \( \rho \) is about \( 10^3 \) kg/m³ and the dynamic viscosity \( \eta = 10^{-3} \) Ns/m². For objects of characteristic size 1 µm and characteristic velocity 1 µm/s, the Reynolds number \( Re = \rho v L / \eta \sim 10^{-11} \) so inertial effects can be neglected in comparison to viscous effects and DNA twists as a severely overdamped harmonic oscillator. In this system, then, dynamics are dominated by hydrodynamic drag, which is characterized by the drag coefficient \( \gamma \), the proportionality constant relating the velocity of an object to the drag forces resisting its motion:

\[
\vec{F}_{\text{drag}} = -\gamma \vec{v}
\]

If the system experiences a hydrodynamic drag \( \gamma \), the autocorrelation time for fluctuations of the system is given by

\[
\tau = \frac{L}{C}
\]

For a long cylinder rotating about its axis, the drag coefficient is given by Perrin’s formula (Levinthal and Crane, 1956; Thomen et al., 2002) where \( \eta \) is the dynamic viscosity of water and \( a \) the radius of the cylinder:
\[ \gamma = 4 \pi \eta L a^2 \]

For the phage packaging system, \( \eta = 10^{-3} \text{ Pa s} \), \( L \sim 1 \mu \text{m} \), and \( a = 1.2 \text{ nm} \), the crystallographic radius of B-form DNA. Let us consider tracking fluorescence from a small dye molecule. The drag of the dye is negligible compared to the DNA, so the autocorrelation time is tens of microseconds or less. This is considerably less than the integration time of a CCD camera, so any signal from a packaging complex will be averaged over the entire distribution of dye positions and orientations. A fast photon detector such as a photomultiplier tube or avalanche photodiode, on the other hand, is not spatially resolved and thus is not well suited for packaging experiments. Averaging over the entire distribution of dye orientations within a single datapoint is problematic, because the DNA angle appears the same modulo 2\( \pi \) radians, and the distribution, as we have seen, is typically that wide. This fact was not a problem in the connector-tracking experiment described in the previous chapter because the rotating object – the connector – was very stiff, so the Boltzmann distribution of angles was small.

The difficulty can be rectified by avoiding averaging over the angular fluctuations. This is possible if the autocorrelation time is greater, as is the case if the hydrodynamic drag is increased significantly.

### 3.3 Measuring DNA angle with optical tweezers

**Rotor bead assay**

The “rotor bead” experiments developed by Bryant et al (Bryant et al., 2003; Gore et al., 2006a; Gore et al., 2006b) take advantage of this fact in order to measure the twist state of DNA. In these experiments (Figure 1a) a third bead, called a rotor bead, is attached to the middle of the DNA. This bead is “torsionally coupled” to one bead and “torsionally uncoupled” from the other due to a nick in the otherwise double-stranded DNA, so that the bead’s position is a direct measure of the angle of the DNA base to which it is attached. In this experimental setup, the drag of the system is dominated by the rotor bead, whose radius is on the order of 1 micron. The drag is given by (see Appendix for derivation) (Happel and Brenner, 1965; Landau and Lifshitz, 1969):

\[ \gamma = 14 \pi \eta r^3 \]

The characteristic autocorrelation time for the rotor bead system is about \( 10^2 \) seconds. Hence the DNA’s angle is not averaged away during a single observation (30 ms at video rate); the mean DNA angle can be determined by simply monitoring the bead angle for a sufficiently long period of time. If the
experimentalist seeks to detect angle changes on faster time scales, a shorter autocorrelation time may be necessary. The autocorrelation time, which is equivalent to the time resolution of the experiment, can be improved simply by utilizing a smaller rotor bead or shorter DNA because $\tau \sim L r^3$.

The rotor bead method is particularly powerful because forces can be applied while DNA angle is measured. It was used to determine several mechanical properties of DNA, such as the torsional modulus (Bryant et al., 2003) and the stretch-twist coupling (Gore et al., 2006a), and to map out a phase diagram for DNA under tension and torque (Bustamante et al., 2003). In addition, the activity of DNA gyrase, an enzyme that twists DNA in units of 4$\pi$, was studied with this assay (Gore et al., 2006b).

While it is easy to measure DNA angle with this technique, it is rather difficult to apply twist. In previous experiments, the tether was twisted by turning the micropipette. Although this occasionally works, the technique is generally unfeasible because any noise in the motor, which is attached to the base of the pipet, is amplified over the ~4 cm length of the pipet, causing large (>5 micron) jitter. Such fluctuations are usually fatal to the tether.

Optical torque wrench

Torque can be applied directly to the optically trapped bead if there is a way of transferring light’s angular momentum to the bead. This has been accomplished in a number of different ways (Friese et al., 1998; O’Neil and Padgett, 2002; Paterson et al., 2001); the most successful biological implementation of this was devised by Wang’s group (La Porta and Wang, 2004). A micron-sized cylinder of quartz is microfabricated and optically trapped; because the quartz is birefringent, the momentum transferred to the particle depends on the angle of the trapping light polarization relative to the quartz fast axis. By rotating the polarization of the trap light, therefore, a torque is applied (see Figure 1b). Likewise, torque can be gauged by carefully measuring the polarization of the scattered trapping light. From the torque, the DNA twist can be calculated through the torsional spring equation.

Although this technique is quite powerful, it requires the microfabrication of unique particles. Current scales of production are on the order of millions of particles, which is much less than are typically used for experiments on biological systems.

Magnetic tweezers

Magnetic tweezers are well suited to applying twist. Magnetic tweezers consist merely of a magnet placed a few millimeters away from a sample chamber (see Figure 1c). Micron-sized ferromagnetic particles are attracted up the magnetic field gradient, so a tether between the bead and the
chamber surface will be stretched under a specific force. Because the distance from the chamber surface to the magnet is much greater than the tether’s contour length, the force is effectively constant regardless of tether extension. The bead can be monitored optically, with a resolution on the order of 50 nm, as a measure of the tether extension. In addition, because the bead has a fixed magnetic moment, it will align with the magnetic field, hence allowing the bead’s angle to be controlled. If the DNA is torsionally coupled to the bead and the bead is manipulated by turning an external magnet, the DNA can be twisted. Like optical tweezers, magnetic tweezers have been used to measure a number of DNA properties (Bustamante et al., 1994; Charvin et al., 2004; Fu et al., 2010; Fu et al., 2011).

DNA twist can also be measured in magnetic tweezers under certain circumstances (Charvin et al., 2004). Using magnetic tweezers, the force can be held constant while DNA is twisted. For low twist, the extension does not change as twist increases, but at some point the torque on the DNA will cause it to buckle and form a loop rather than remain twisted. As the DNA is twisted beyond this point, each twist is converted into a loop; the loop formed in this way is called a plectoneme. These loops must reduce the end-to-end extension of the DNA tether, so the extension drops in proportion to the new twist introduced. However, this method cannot be used for tensions greater than 3 pN, because at this force the energetic cost of plectoneme formation is always greater than that of twisting; plectonemes cannot form. Moreover, detecting DNA twist by using plectonemes requires that the DNA be pre-twisted, which may interfere with normal packaging dynamics. Nonetheless, magnetic tweezers have been used in this way to estimate DNA twisting by the DNA translocase FtsK (Saleh et al., 2005) and the restriction enzyme EcoR124I (Stanley et al., 2006). In these studies, the twisting rate and the translocation rate are measured separately, and the ratio of the rates is a measure of the trajectory of the molecular motor along and around the DNA.

*Introduction to DNA topology*

An experiment that infers twist from plectoneme formation is predicated on the topological equivalence of twisting and looping. A general quantity that characterizes the topology of a pair of curves, such as the two strands of a DNA double helix is the linking number $L_k$. The linking number is the number of times the strands wrap around each other. Equivalently, if the ends of each strand are fixed, the linking number is the number of times that one of the strands must physically pass through the other in order to separate the strands. A straight segment of length $L$ basepairs of dsDNA has a linking number of $L/10.5$ because the helical pitch is 10.5 bp/turn. If the DNA segment is not straight, the linking number also depends on its path. A
Theorem proven by Calugareanu and White (Adams, 1994) identifies the linking number as the sum of the twist and the writhe of the curves.

\[ Lk = Tw + Wr \]

In the case of dsDNA, the writhe is the number of times that the DNA’s central axis wraps around itself. If the DNA forms a loop, the writhe changes by one. If the linking number is held constant by fixing the ends of both strands – a condition called torsional constraint – any change in plectoneme number must reflect an equal and opposite change in twist number.

In the case of straight but torsionally relaxed dsDNA, because the twist is fixed at \( Tw_0 = L/10.5 \), we will define the intrinsic linking number \( Lk_0 = Tw_0 \).

It is often more convenient to speak of the relative linking number density, normalized by the intrinsic linking number, such that relaxed DNA has a relative linking number density of zero:

\[ \sigma = \frac{Lk - Lk_0}{Lk_0} \]

3.4 Combined Optical and Magnetic Manipulation and Observation Tweezers (COMMOT)

I designed an experimental scheme that combines the twisting capability of magnetic tweezers with the twist-sensing capability of a rotor bead setup. Combined Optical and Magnetic Manipulation and Observation Tweezers (COMMOT) uses a standard three-bead setup in which the rotor bead is ferromagnetic. If a magnet is placed near the sample chamber (Manipulation Mode), it can be used to spin the rotor bead and hence twist the DNA. In Observation Mode, the magnet is removed and the DNA will relax. The angle of the rotor bead, the extension of the DNA, and the tension on the DNA can all be monitored simultaneously. In Observation Mode, DNA angle is monitored; alternatively, Manipulation Mode can be used to apply a specific twist to the DNA in order to study the effect of torque on the system of interest.

Materials

COMMOT work was performed on a counterpropagating dual-beam optical tweezers (Smith et al., 2003). Tether force was determined by direct measurement of the momentum change of the trapping light. The position of the rotor bead was determined by video microscopy. In order to improve the resolution of video microscopy, LED emission at 400 nm was used for bright-field illumination (LED from Superbright LEDs, Inc.). For magnetic manipulation, a pair of neodymium magnets (K&J Magnetics) in antiparallel
orientation (see Figure 2a) were placed about 3 cm above the tip of the micropipette. For other instrumental details, see the Appendix.

A tether with three distinct hapten patches was constructed (Figure 2b). See the Appendix for details of the DNA design. At one end is a 200-bp patch of DNA, amplified with PCR, containing fluorescein-dUTP (Roche). This patch is ligated to a 6-kbp segment of plasmid DNA. The plasmid DNA is ligated to a 150-bp patch containing biotin-dUTP (Roche). That patch, in turn, is ligated to a 3100-bp DNA of plasmid origin that has been modified at the far end with digoxigenin-dUTP (Roche) (see the section entitled Terminal Labeling Of DNA later in this chapter). The total DNA molecule is double-stranded with the exception of a nick between the biotin patch and the 3100-bp section, and the dig-containing tail.

The DNA is incubated with 2.88-µm anti-fluorescein-coated Protein G beads (antibody from Roche; beads from Spherotech) and introduced into an optical tweezers flow chamber. A 2.1-µm anti-dig-coated Protein G bead (antibody from Roche; beads from Spherotech) is sucked onto a micropipette in the chamber. Tethers are formed between an optically trapped anti-fluorescein bead and the suctioned anti-digoxigenin bead. One a tether is formed, a transverse flow of 10-20 pN is started. A 2-µm streptavidin-coated ferromagnetic bead (available from either Estapor Inc or Micromod GMBH) is then optically trapped and fished for an attachment to the middle of the flow-extended DNA. (Micron-sized magnetic beads are often marketed as “superparamagnetic,” but that term properly describes a bulk property of a suspension of beads. Individual beads are ferromagnetic; if they were paramagnetic, they could not be manipulated by moving the applied magnetic fields.) Once the attachment point is found, the optical trap is turned off, moved to the anti-fluorescein bead, and re-activated. The flow is then switched off.

**Applying force with COMMOT**

Force can be applied to the tether via the optically trapped bead. However, it is possible that the tether feels a force due to a magnetic force on the rotor bead. The force is given by the gradient of the potential energy of the magnetic bead in the magnetic field. Where $\mu$ is the magnetic dipole moment of the bead and $B$ is the magnetic field, and assuming that the bead aligns with the magnetic field,

$$\vec{F} = -\nabla U = -\nabla (\vec{B} \cdot \vec{\mu}) = -\mu \nabla \vec{B}$$

Assuming that the magnetic fields in this configuration are similar to that of a dipole (this is true because the magnet size is small compared to the
distance between the magnet and the pipet tip), the field strength decays as $d^{-3}$ where $d$ is the distance from the dipole. The magnet is specified to have a surface field of about 5000 Gauss at the surface, so the magnetic field in the sample plane can be estimated as 1 Gauss. However, the magnetic moment of the individual beads is not known and probably varies from bead to bead.

An upper bound on the magnetic force can be obtained experimentally. When the magnet is brought into position, a magnetic bead in the sample chamber moves toward the magnet at a velocity on the order of 1 µm/s. This drift velocity corresponds to a force of 0.05 pN. This can also be used to calculate a rough estimate for the magnetic moment of an individual bead: on the order of $10^5$ pN nm Gauss$^{-1}$.

**Applying twist with COMMOT**

The tether is extended by moving the pipet away from the optical trap and holding at a force between 5 and 20 pN. In order to apply twist, the magnet is moved to a position just above the sample chamber, about 3 cm above the tip of the pipet. The magnet is turned at a rate of 1-5 Hz, causing the rotor bead to turn around the DNA axis. The torque applied by the magnet on the rotor bead is given by

$$\vec{N} = \vec{\mu} \times \vec{B}$$

Using the bead magnetic moment estimate from above and the magnetic field estimate, we can put an upper limit to the torque that can be applied to the bead: $10^4$ pN nm. In practice, we rarely need to apply large torques because DNA denatures under torques in excess of a critical value (40 pN nm for overwinding and 10 pN nm for underwinding).

**Measuring twist and torque with COMMOT**

The rotor bead’s position is determined by video tracking. In order to avoid motion blur, the duration of a single frame must be much less than the autocorrelation time of the bead-DNA system as described in the previous section. This is easily satisfied by video rate recording (30 Hz). Also, determination of the mean rotor bead position requires tracking it over a period several times longer than the autocorrelation time, which could be hundreds of seconds for these 2-µm particles.

As an example of the capability of COMMOT, see Figure 2c. A magnetic bead was attached to the DNA as a rotor bead. An external magnet was brought nearby and revolved; the magnetic bead rotated in response. After 40 turns were introduced, the magnet was removed and the DNA relaxed over a period of ~50 seconds.
Torque can be inferred from the angular velocity of the rotor bead. At low Reynolds number, the drag torque $N_d$ is related to the rotor bead angular velocity $\omega_{rb}$ through the drag coefficient of the rotor bead $\gamma_{rb}$ as mentioned above:

$$N_d = -\gamma_{rb}\omega_{rb} = -14\pi \eta r^3 \omega_{rb}$$

Again, torques greater than +40 pN nm or less than -10 pN nm cannot be measured with a rotor bead, because the DNA undergoes a phase transition, changing its twist at constant torque.

Finally, COMMOT has an advantage over other methods of optically manipulating magnetic particles. Optical trapping of an iron-containing bead causes rapid heating, melting, or even explosion. Because this method uses a low (~12%) magnetite content particle, heating is lessened. And secondly, because the magnetic bead is not held in the optical trap during the actual experiment, heating does not affect the biological activity of interest.

### 3.5 Development of Novel DNA Modification Strategies I: Peptide Nucleic Acids

The following section contains data and text that was published in Nano Letters (Zohar et al., 2010).

Rotor-bead measurements require a unique DNA substrate that must be carefully designed and constructed. The DNA molecules used in earlier studies – containing a fluorescein-containing patch (produced by PCR), a DNA of interest, a biotin-containing patch, a DNA spacer, and a digoxigenin-containing patch – required four high-efficiency ligations in order to connect five independent DNA molecules. Given that a typical ligation efficiency is about 90%, and given that multiple ligation products are generally possible, the yield of intended product can be below 10%. Often, the nonspecific products will have two hapten moieties, meaning that they can bind to both an end bead and the rotor bead and hence interfere with rotation. Moreover, packaging initiation by bacteriophage φ29 is significantly inhibited by the presence of nonspecific DNA molecules. For these reasons, traditional DNA construction techniques were inadequate and new methods were required. Specifically, I sought to reduce the number of required ligation steps by using alternative methods of labeling DNA.

In order to remove two ligation steps, methods of internally labeling the DNA were investigated. At first, a number of nonspecific DNA binding molecules were tried. In optical tweezers experiments, positively-charged dimethylamino-coated beads (Spherotech) were found to stick to DNA, but
then slide along it in response to small forces (data not shown). This behavior can be generalized to any purely electrostatic DNA-binding mode because the potential of the interaction should be constant along the DNA length. A more specific DNA-binding molecule – anti-DNA antibodies (Millipore) – were then coated onto Protein-G-coated polystyrene beads (Spherotech) and used to fish for the middle of a DNA molecule in the tweezers. The same general behavior, with the bead sticking but then sliding under low loads was observed, suggesting that the antibody-antigen interaction is also weak and not specific. Inadvertently, one particular batch of beads was found that sticks nonspecifically but irreversibly to DNA (2.1 µm streptavidin-coated magnetic beads from Micromod, batch date unknown). The mode of binding remains unknown but is likely “dirt” on the bead surface as the binding function is abrogated by sonication for 30 minutes. The large size of these beads, however, makes them less useful for twist measurement, because the relaxation time is typically on the order of minutes.

Ideally, then, internal labeling will also exhibit a degree of sequence specificity, because a typical optical tweezers tether will contain thousands or tens of thousands of possible binding sites. Sequence-specific DNA binding is a property of many molecules such as some proteins and nucleic acids. A review of the many options can be found in (Zohar and Muller, 2011). One particularly exciting DNA-binding technology is a synthetic hybrid of proteins and nucleic acids called a Peptide Nucleic Acid, or PNA (reviewed in (Nielsen, 1999, 2001)). Composed of nucleic acid bases and a peptide-like backbone, PNAs bind sequence-specifically to DNA (Demidov et al., 1994). Moreover, because the backbone is neutrally charged, binding between PNA and its complementary ssDNA sequence is actually stronger than between two complementary ssDNA strands.

Materials and Methods

Different PNA molecules are known to bind to dsDNA through a variety of different modes. We designed a bis-PNA which is expected to invade dsDNA and form a triplex-like structure around one strand. See Figure 3a. The sequence used, from N-terminus to C-terminus, was termed twPNA and consisted of

TAMRA-OO-lys-lys-TTT CTC TT-OOO-TTJ TJT TT-lys-OO-lys-biotin,

where TAMRA is carboxytetramethylrhodamine, O is 8-amino-3,6-dioxaoctanoic acid, a flexible hydrophilic tether, lys is Lysine, and J is pseudoisocytosine. The PNA sequence was modified from those used in earlier work (Chan et al., 2004). Biotin allows for conjugation to Avidin-coated beads, and TAMRA provides a spectroscopic means of identifying bound PNA. The lysines provide positive charges to stabilize the PNA-DNA
interaction. The molecule targets the DNA sequence AAGAGAAA. The PNA was synthesized by Bio-Synthesis, Inc.

**DNA construct for optical tweezers experiments.** DNA molecules containing both dig and a twPNA were prepared for study in the optical tweezers. A digoxigenin-modified DNA construct, called DNA-dig, was first formed by the following procedure. A 510-bp fragment from λ-DNA, comprised of base pairs 24241 to 24741, was generated via PCR using Taq polymerase (New England Biolabs). The PCR buffer contained 200 µM dATP, dCTP, and dGTP (Fermentas); 133 µM dTTP; and 66 µM dig-dUTP (Roche). After cleanup, the PCR product was digested with XbaI (New England Biolabs) and shrimp Antarctic phosphatase (New England Biolabs). Simultaneously, λ-DNA (New England Biolabs) was digested with XbaI (New England Biolabs) in NEBuffer 2 at 37 °C. The two DNA fragments were ligated overnight using T4 ligase (Fermentas) with a 2-fold excess of PCR product over digested λ-DNA. The ligation product, called DNA-dig, was then dialyzed into 10 mM Tris, 0.1 mM EDTA, pH 8.0 through a 25 nm VSWP nitrocellulose filter (Millipore) for one hour at room temperature. This DNA construct was then incubated with PNA under the conditions described below.

Another construct, the control molecule biotin-DNA-dig, was made to serve as a basis of comparison for optical tweezer experiments performed with the twPNA-DNA-dig construct described below. Biotin-DNA-dig simply has a biotin modification on one end and a dig modification on the other end. It was produced by treating λ-DNA with Klenow fragment, exo minus (New England Biolabs) in the presence of 70 µM dGTP and dTTP; and 30 µM biotin-dCTP and biotin-dATP (Invitrogen), thus filling the cohesive ends. This was then digested with XbaI and ligated to the dig patch as above, yielding biotin-DNA-dig. It was then subjected to the same thermal and salt conditions as the twPNA-DNA-dig complex and labeled with spheres as described below for the twPNA-DNA-dig complex.

**PNA-DNA binding conditions.** Complexes of twPNA and DNA containing the target sequence were formed in a 50 µL binding reaction (pH 8) consisting of 10 mM Tris, 0.1 mM EDTA, 5 nM DNA, and 1 µM PNA that was incubated at high temperature (37-65 °C) for 5 hours, then cooled to 4 °C over 30 minutes. The NaCl concentration was brought to 50 mM to inhibit any further binding and the reaction volume was heated to 50 °C for 10 minutes to enhance binding specificity by reducing nonspecific binding. Binding conditions had been optimized in bulk on lambda DNA, which contains a target sequence, and assessed by a shift in electrophoretic mobility of a SalI restriction fragment containing the target sequence. Binding conditions were optimized as a function of incubation temperature and salt concentration. See Figure 3b for an example.
**Force-ramp experiments in optical tweezers.** The twPNA-DNA-dig complexes were pre-bound to 2.1 µm anti-digoxigenin-Protein-G-coated polystyrene spheres and introduced into an optical tweezers flow chamber in a buffer that consisted of 50 mM Tris, 50 mM NaCl, and 5 mM MgCl₂, pH 8.0. The twPNA-DNA-dig complex was then stretched between an optically trapped anti-digoxigenin bead and a 2.1 µm streptavidin-coated polystyrene bead sucked onto a micropipette (Figure 3c). The pipette was moved away from the optical trap at a rate of 175 nm per second, starting from a force <2 pN, while the tension and the extension of the molecule were monitored. If the force reached 90 pN without a rupture event, the ramping process was reversed and then repeated.

**Results**

**Optimizing binding conditions.** Electrophoretic mobility assays indicated that the highest efficiency of specific binding of PNA to the DNA resulted from incubation at 65°C in the absence of salt for 5 hours. The force-extension curve of >90% of the control tethers (biotin-DNA-dig) prepared in this way had the general form shown in Figure 5a, with no apparent hysteresis. The curve displays several major differences from the force-extension curve of normal dsDNA (see Figure 5c). The low-force regime (below about 10 pN) the behavior is generally similar; at higher forces the DNA is significantly less stiff than normal dsDNA. These differences can be attributed to alterations in the DNA resulting from the long incubation period at high temperature. Under such conditions, even long DNA molecules will melt and it is possible that subsequence steps did not promote re-annealing of the strands; however, the DNA ran normally on a native agarose gel. Also, the force-extension curve is not consistent with ssDNA and, moreover, that the pattern continues for several cycles suggests that the DNA is not in a kinetically-trapped, mis-paired state. It is likely that the change in DNA mechanical properties, therefore, result from covalent changes to the DNA, possibly from extensive depurination due to partial melting, high temperatures, and low ionic strength, all of which increase the depurination rate (Lindahl, 1993).

In order to reduce DNA damage, the incubation temperature was reduced to 37°C. Under these conditions, the force-extension curve of biotin-DNA-dig returned to normal (Figure 5b). The curve of twPNA-DNA-dig prepared in the same way was indistinguishable (Figure 5c). However, the efficiency of labeling was below 30% (data not shown).

**Rupture Force Distribution.** During the force ramp, the control tether (biotin-DNA-dig) never broke, whereas the twPNA-DNA-dig tether often would break. This break must be due to a loss of the PNA-DNA interaction. A distribution of the rupture forces is shown in Figure 6a. The mean rupture force of this
tether is comparable to the rupture force of a tether anchored through a single dig, indicating that PNA can withstand the force typical in an optical tweezers experiment (see, for example, Figure 9c).

The suitability of PNA in tweezers experiments can be quantified through the expected lifetime of a tether at a given force. The rupture rate as a function of force can be inferred from the rupture force distribution using a transformation devised by Dudko (Dudko et al., 2008). Where $p(f)$ represents the probability density of the rupture force and $\dot{F}$ is the loading rate, the force-dependent lifetime is estimated as

$$\tau(F) = \frac{1}{\dot{F}(F)p(F) \int_{F}^{\infty} p(f) df}$$

This analysis assumes that the breaking kinetics are single-exponential. Before transforming, the rupture force distribution was smoothed using kernel density estimation. The result is shown in Figure 6b, which indicates that at forces below 20 pN, a typical tether will survive hundreds of seconds, plenty of time for most optical tweezers experiments.

However, the rupture-force distribution is rather broad and is symmetric, in contrast to the rupture-force distribution for the single-dig-anchored tether. One explanation for such a wide rupture-force distribution is a very high pulling rate. (See (Dudko et al., 2008) for examples). A useful figure of merit for pulling rate is the ratio of the standard deviation of the rupture force distribution to the mean lifetime at mean rupture force. If the pulling rate is less that that ratio, the force-dependent rupture rate does not change much over the lifetime of a tether, so the rupture can be said to be “spontaneous,” resulting in a left-tailed distribution (see Figure 9c for an example). In contrast, if the pulling rate is much higher than the ratio, the rupture rate does increase significantly over the life of the tether, so the distribution should be more symmetric. For this tether, the loading rate near the rupture is on the order of 10 pN/s, while the figure of merit is about 200 pN/s, suggesting that another explanation for the distribution’s symmetry is needed.

Another possibility is that the rupture kinetics are not single exponential. Such kinetics would result from multiple binding modes or binding sites. There are clues of this scenario in the distribution of tether contour lengths (Figure 7), which is quite broad, with peaks at the expected tether lengths of 4000 and 5000 nm (±500 nm). All other rupture lengths indicate that a PNA was bound at a site other than the target sequence. Specifically bound PNAs accounted for less than 50% of all PNAs detected.
When only tethers of the expected length were considered, the mean rupture force increased considerably, from about 40 pN to 57 pN. Specifically-bound PNAs, forming a minimal-energy complex with the target sequence, are more strongly bound than other PNAs which may be missing one or several basepairing interactions.

Discussion

The strength of the PNA-DNA bond compares favorably with the dig-anti-dig bond commonly used in tweezers experiments. These PNA sequences also exhibit a modicum of specificity, shown by the ~40% of tethers that were the expected length. However, the specificity that was observed in the optical tweezers is somewhat worse than what was seen in two different single-molecule studies. First, Zohar (Zohar et al., 2010) studied PNA specificity by binding a fluorescent bead to the DNA-bound PNA, trapping the fluorescently-labeled DNA in a stagnation-point flow cell, and imaging both the bead and the DNA. In that experiment, which used a different sequence, specificity near 100% was found. In another experiment, using a sequence identical to the twPNA used in my work, Chan (Chan et al., 2004) used single-molecule fluorescence imaging of the Tamra dye to determine that specificity ranged from 50-80%. Both specificity and kinetics changed when a PNA labeled with two Tamra dyes was used. It seems that minor changes to the PNA can have significant effects in ways that are not well understood.

Given that there are some 48,000 8-basepair sequences in lambda DNA, these PNA molecules had a 24,000-fold preference for the target sequence. This preference corresponds to a free-energy difference of 10 k_BT. It is likely that specificity is kinetically controlled, because specificity is enhanced by a short incubation at 50°C in 50 mM NaCl after the initial binding step. Hence, the energy difference likely corresponds to the height of the barrier to detaching the PNA from the DNA. These kinds of experiments can easily be continued to reveal more about the energetics of PNA-DNA binding. Such research will be important if PNAs are to become a useful tool, because it is clear that designing a PNA for a specific task is fraught with unexpected difficulties.

Perhaps the biggest obstacle to the use of PNA in a wider variety of experiments is the dependence of its binding rate and binding specificity on salt concentration (Tomac et al., 1996). For example, the ideal conditions for binding the PNA used in this project – high temperature and no salt – promoted irreversible damage to the DNA. If protocols for PNA design are standardized and are successful in overcoming this issue, PNA can be used to attach almost any functional group to any DNA site. In this work, the mechanical strength of the PNA-DNA bond is shown to be strong enough to manipulate DNA. Given that PNAs have already been employed in in vivo
experiments, they are a viable candidate for coupling to specific DNA sequences in vivo mechanical tools such as optical tweezers. For example, the in vivo strength of nucleosome binding to a gene subject to epigenetic regulation could be measured if that gene could be isolated with PNA. These experiments demonstrate the potential of PNAs as a molecular handle for the manipulation of DNA.

### 3.6 Development of Novel DNA Modification Strategies II: Terminal labeling with terminal transferase

In order to eliminate another ligation step, I sought to directly label the end of a DNA molecule with a hapten other than biotin. Because of the long duration of the packaging experiments – often > 10 minutes – the ideal tether would contain multiple labels at that end in order to strengthen the attachment. An additional requirement unique to rotor bead experiments is that one end – the end opposite from the molecular motor of interest – must be free to swivel. Single bonds typically permit such swiveling unless it is prevented by significant steric hindrance of the adjacent chemical groups. Single-stranded DNA thus is not torsionally constrained, whereas double-stranded DNA is constrained due to topology. A natural way of generating an unconstrained multiple-hapten handle, therefore, is for the haptens to be part of a single-stranded region.

Such a terminal single-stranded “tail” is produced by the enzyme terminal transferase (TdT). Terminal transferase is a human enzyme used during antibody gene recombination to generate antibody diversity and acts as a template-free polymerase that extends DNA from the 3’ end (Roychoudhury et al., 1976). It is commonly used to label DNA in vitro, in fixed whole cells or tissues, most notably to identify double-stranded breaks in DNA, often as a marker of apoptosis (Li et al., 1996). High-efficiency tailing of DNA requires an unpaired 3’ end as a primer.

The use of terminal transferase to generate a single-stranded region has the unexpected benefit of allowing the first single-molecule application of the new hapten bromo-deoxyuridine (BrdU). BrdU can be incorporated into DNA by common polymerases both in vitro and in vivo (Leuner et al., 2010). Single-stranded DNA containing BrdU can be visualized in a Western blot-like procedure if the DNA is denatured before being mixed with anti-BrdU antibodies (Vanderlaan and Thomas, 1985). Likewise, single-stranded BrdU tails on dsDNA can be generated by terminal transferase and then detected via anti-BrdU in an apoptosis assay (Darzynkiewicz et al., 2008). In this section I will present several small experiments. First I discuss the development of a protocol for tailing dsDNA with bromo-deoxyuridine or digoxigenin-deoxyuridine nucleotides. Second, I demonstrate the first single-
molecule use of BrdU, and I evaluate the strength of tethers formed via BrdU in comparison to tethers formed via digoxigenin-dU.

Materials and Methods

Generating a 3’ tail on plasmid DNA. A 6294 bp plasmid, termed “DD,” containing a BglI site, was cloned into DH5-alpha E. coli cells and extracted (Qiagen). See the Appendix for details regarding the plasmid. For the in vitro experiments, a 402-bp PCR product containing that site was amplified from that plasmid. For the tweezers experiment, the plasmid was used as is. The DNA was digested with BglI (New England Biolabs), which leaves a 3’ overhang, and cleaned up (Qiagen).

Extending the 3’ tail using TdT. All terminal transferase reactions were performed in NEB TdT buffer plus 0.25 mM CoCl2, plus 500 µM dNTP or modified dNTP.

Assaying reaction product. After TdT treatment, the yield on PCR DNA was assayed by running the resulting fragments on a 1.5% agarose gel in TBE (Figure 8). The yield on plasmid DNA was assayed by first digesting with EarI and then running on the gel.

Preparing a DNA tether with biotin at one end and dig or BrdU at the other. The TdT-treated plasmid DNA was digested with BamHI (New England Biolabs) resulting in two different fragments, of length 3165 and 3129. The two are not distinguishable in optical tweezers experiments so they were both denoted “D.” The DNA was then treated with DNA Polymerase I, Klenow fragment, exo-minus (New England Biolabs) in the presence of 25 µM biotin-dCTP (Invitrogen) and 25 µM dATP, dGTP, and dTTP, in order to fill the EcoRI site with biotin-labeled nucleotides.

Extension-ramp experiments to assay tether strength. D-dig molecules were mixed with 2.1-µm anti-digoxigenin-coated Protein-G beads (anti-dig from Roche; beads from Spherotech). D-BrdU molecules were mixed with 2.1-µm anti-BrdU-coated Protein-G beads (anti-BrdU from Invitrogen; beads from Spherotech). A streptavidin-coated bead was sucked onto the micropipette and the DNA-coated beads were optically trapped and fished. Once a tether was formed, the pipet was moved away from the optical trap at a rate of 60 nm/s and the force and extension were recorded until the tether ruptured.

Results

Terminal transferase activity depends on substrate

3’ tail extension is highly processive in dNTPs or BrdUTP. Figure 8c shows the result of TdT treatment of the PCR product. In the presence of dTTP, a tail can be added, producing a mobility shift and smear (lane 3); in the
presence of BrdUTP, the shift is greater (lane 4). After cleanup to remove nucleotides, protein, and buffer, the reaction could be repeated. This produced a second shift for both nucleotides (lanes 5 and 6). This proves that multiple BrdU nucleotides can be added by terminal transferase. Because the tail is single-stranded, the number of nucleotides added is not apparent from the gel. However, based on the increase in $A_{260}$, the number of nucleotides added in one round was about 300.

A similar experiment using dig-dUTP as the nucleotide yields quite different results (Figure 8d). In the presence of dig-dUTP, the shift is rather small and the band is not a smear (lane 4). After a second round of extension, there is not a second shift (lanes 6 and 7). Thus TdT is “poisoned” by a dig-modified tail. This poisoning effect indicates that the number of dig groups added was less than ten or so, and almost certainly one.

BrdU-labeled tethers are less strong than dig-dU-labeled tethers

In the optical tweezers, the strength of a tether was assessed by measuring the rupture force distribution while moving the pipet away from the optical trap at a constant rate of 60 nm/s. The rupture force distribution for the tether biotin-D-BrdU is shown in Figure 9a and for biotin-D-dig in Figure 9c. The mean rupture force in these conditions is about 25 pN for the BrdU tether and 55 pN for the dig tether.

A more relevant parameter for optical tweezers experiments, however, is the lifetime of the tether at a particular force. In the planned rotor bead experiments on packaging phage complexes, a force of about 10 pN will be applied for more than ten minutes, with a brief (< 1 minute) period of increased force due to flow during buffer exchange. Due to the serial nature of optical tweezers experiments, it is impractical to directly measure, with statistical significance, lifetimes greater than a few tens of seconds. In order to estimate the lifetime in the range 10-20 pN, the transformation due to Dudko et al. was applied to the rupture force distribution after smoothing by kernel density estimation (Figure 9b,d). The lifetime of a BrdU-containing tether at 10 pN, at $\sim 10^2$ seconds, is too short for the intended packaging experiment. In contrast, the lifetime of the dig-mediated attachment, at $10^3 - 10^4$ seconds, is sufficient.

As noted earlier in the chapter, the shape of the rupture force distribution contains information about the kinds of bonds that are broken during a rupture event. The left-tailed distribution for the dig tether is consistent with a single rate-limiting event, suggesting that a single dig-anti-dig interaction is broken. In contrast, the tether containing BrdU yielded a much wider distribution, implying the involvement of multiple rupture modes, namely multiple BrdU-anti-BrdU bonds. This interpretation is reinforced by the observation that the BrdU tether, when held at a constant force of 10 pN,
undergoes dynamic ~ 10 nm changes in extension, which are probably due to breakage of individual BrdU-anti-BrdU interactions (data not shown). A batch of biotin-D-BrdU that had been treated twice with BrdUTP exhibited a rupture force distribution that was not significantly different from the batch that had been treated once. This BrdU-labeling procedure, therefore, did introduce multiple BrdU moieties but these additional groups do not improve tether strength to the level of a dig-containing tether. In contrast, the single dig label that is added by terminal transferase is sufficiently robust so as to form a reliable attachment for a phage packaging experiment.

3.7 Use of PNA to study packaging by bacteriophage φ29

Materials and Methods

In order to study viral packaging by bacteriophage φ29 in the presence of PNA, a 28-kbp DNA fusion product was engineered by ligating a 13-kbp PCR-amplified DNA containing the PNA target site to a 15-kbp gp3-DNA restriction product, as indicated in Figure 10. See the Appendix for details of the DNA construction. For details on the initiation, packaging, and stalling protocol, see Chapter 4. The PCR-amplified DNA had been modified with digoxigenin using terminal transferase.

Results and Discussion

In bulk DNase protection assays, a segment of DNA equal in length to the wild-type viral genome was packaged, but the protected DNA could have originated from any of the three ligation products with length equal to or greater than the wild-type genome. Fluorescence from the PNA could not be detected in the gel, indicating that no PNA entered the capsid during packaging. Two possible explanations for this are that the proheads are not capable of packaging φ29-DNA, or that the packaging complex strips PNA from the substrate.

In optical tweezers, tethers could be formed between anti-φ29-coated beads and anti-digoxigenin beads, indicating that packaging complexes had initiated on the desired DNA molecule. However, attempts to attach a streptavidin-coated rotor bead to the middle of the tether failed (>100 attempts), showing that the DNA did not contain the biotinylated PNA. Intriguingly, tethers could be formed between anti-φ29 beads and streptavidin beads, but those tethers did not bind anti-digoxigenin beads (~20 attempts). These facts imply that φ29 does not package DNA that is labeled with PNA.
There are two conceivable mechanistic explanations for this observation; both require an interaction of the prohead with DNA prior to encapsidation. The first is that the packaging complex actively strips PNA from the DNA before the DNA is packaged into the capsid, sometime during or immediately after the initiation event. The second possibility is that the packaging complex is capable of sensing the presence of PNA on downstream DNA. The mechanism for this early DNA interaction is yet to be determined, but it is consistent with the hypothesis that packaging is initiated by looping and supercoiling of the gp3-DNA as seen in electron micrographs (Turnquist et al., 1992). There is also single-molecule evidence of gp3-DNA lariat formation during initiation (Rickgauer et al., 2006). If initiation requires supercoiling by gp16, the gp16 may also clear the DNA of roadblocks during that time, allowing encapsidation to proceed without difficulty. A “wire-stripping” ability has already been documented in in vitro assays on FtsK, the close relative of the φ29 packaging motor (Levy et al., 2005). This model fits the optical tweezers data, but more conclusive proof will be needed. Perhaps this effect can be seen in atomic force microscope images of packaging initiation.
Figures

Figure 3-1. Methods of measuring DNA twist. (a) Direct observation of twist using a rotor bead in an optical tweezers; torque is inferred from angular velocity. Figure reproduced from (Bryant et al., 2003). (b) Direct observation of torque in an optical torque wrench optical tweezers; twist is inferred from elasticity. Figure reproduced from (La Porta and Wang, 2004). (c) Indirect observation of twist in a magnetic tweezers by a change in extension due to plectoneme formation with increasing twist. Figure reproduced from (Gore, 2005).
Figure 3-2. (a) COMMOT experimental setup (not to scale). A magnet placed ~3 cm above the pipette (outside of the chamber) is used to manipulate a magnetic rotor bead. (b) DNA design for COMMOT. The upper 6 kbp, including the biotin patch, is torsionally constrained. The nick and single-stranded tail ensure that the lower 3 kbp are not torsionally connected to the rotor bead. (c) Demonstration of magnetic manipulation. The external magnet was rotated 40 times and removed. The DNA then relaxed, returning the rotor bead to its original position. The decay time is the autocorrelation time of the system.
Figure 3-3. Mode of PNA binding. (a) This bis-PNA is expected to invade duplex DNA and form a triplex with one strand. (b) PNA binding causes a well-defined gel shift that can be used to assess binding efficiency and specificity. PNA binding to this SalI fragment of DNA does not depend strongly on stoichiometry.
Figure 3-4. Experimental design for PNA pulling experiments. The tether was a 24-kbp fragment of lambda DNA with digoxigenin at one end. There were two PNA binding sites – 15100 and 11600 bp away from the digoxigenin.
Figure 3-5. Force-extension curve of tether via PNA. (a) High-temperature, low-salt binding conditions cause an abnormal force-extension curve that resembles DNA unzipping. (b) 37 C binding condition improves the force-extension curve. (c) Force-extension curve of the control tether, which is held through covalently linked biotin. This particular molecule exhibited an unusual rip around 40 pN.
Figure 3-6. Mechanical stability of PNA binding to DNA. (a) Rupture force distribution under a constant loading rate of 160 nm/s. (b) Lifetime of tether as a function of force, obtained by Dudko analysis of the rupture force distribution in (a).
Figure 3-7. Binding specificity assessed in optical tweezers. The distribution of the tether contour lengths reveals the range of locations at which the PNA was bound to different DNA molecules. Specific binding would have generated a contour length of either 4000 or 5000 nm, depending on which of two specific binds sites were occupied. About 50% of tethers were deemed to be specific, that is, had a contour length of between 3500 and 5500 nm. The peak at 8000 nm corresponds to PNA bound to the end of the DNA.
Figure 3-8. (a) Experimental setup. (b) DNA design. (c) Terminal transferase activity with different nucleotides. Gel shift with BrdUTP is greater than with dTTP (lanes 3 & 4). After cleanup and a second treatment, the shift is greater still (lanes 5 & 6). (d) Terminal transferase does not extend form a dig-dUMP tail. After one round of treatment in the presence of dig-dUTP, there is a shift (lane 4). After cleanup and a second round in the presence of dig-dUTP or normal dNTPs, there is no further shift, showing that the enzyme is not active on a tail that already contains dig-dUMP.
Figure 3-9. Mechanical characteristics of antibody-BrdU bonds and antibody-dig bonds. (a) Rupture force distribution for BrdU-labeled DNA at a loading rate of 65 nm/s. (b) Force-dependent lifetime for the BrdU-containing tether. The shoulder above 40 pN derives from the long tail in the rupture force distribution and likely indicates that multiple antibody attachments can modestly increase the strength of the tether. (c) Rupture force distribution for dig-containing tether. (d) Force-dependent lifetime of the dig-containing tether.
Figure 3-10. Experimental setup for proposed φ29 packaging on PNA-labeled DNA.
Appendix

**COMMOT proof-of-principle DNA design**

The DNA used in the COMMOT experiments was a ligation of four different pieces of DNA, called alpha, B, C, and D.

*Fluorescein patch alpha*

The fluorescein-modified patch, “alpha,” was generated by PCR of a fragment containing the SpeI site of the φ29 genome. PCR utilized Taq polymerase (New England Biolabs) in standard Taq buffer, supplemented with 200 µM dATP, 200µM dCTP, 200 µM dGTP, 150 µM dTTP (all New England Biolabs), and 50 µM fluorescein-dUTP (Roche). The amplified section had the following sequence.

```
tatgtgtgccgacggaagtgtttcgctatatattacatggttaaacctcaggtttgatgtgtgtaaggagtctattaacctcagcaggtatcggaggaaacgtaacaggcaccat"ntacatatatgtgattttcgcgtgaaccaggataatactcgggacgcaactacattgtggttaaggagtaccaaggagcccttaataaacttgggagatatgaaaatgatagcggtgagatgcaacact
```

After digesting with SpeI, the mixture of two fragments (lengths 292 and 162) was cleaned up (Qiagen Qiaquick).

*6 kbp segment B*

Segment B is of two possible lengths: 5869 and 5682 basepairs. Each fragment comprises one half of a 11551-basepair plasmid with SpeI and BamHI sites at opposite poles. The plasmid was transformed into E. coli strain jtk164 rbs 24, which was necessary because the plasmid contains a *pir* phagemid origin of replication (gift of J. Kittleson and J. C. Anderson).

The plasmid was first digested with SpeI, then treated with Antarctic Phosphatase (New England Biolabs), then digested with BamHI (New England Biolabs), yielding the two fragments of nearly-identical length. The sequence of the entire plasmid is below.

```
GAATTGATCAGATCTtaaatctatcaccgcaagggataaatctaatcaacccccgtgatgcttggtgcgttatatttactctggcggtgataatgttgccGGATCTTtagctactagagaagaagaggaatatactagatgcggtaaagccgaaagcttgtgggtgaaactggatggtgatgtcaacggtcataatctttttcgtgcggcaggatgctgcgttaacgggacagtacagetggatatgaggagacttgaaggaatgtgggggatgagca
```
cgtacaatcacgtcgagaatgtaaatgagggagatcatacctgttgaataaaccccagctcccaggccccagttagtgcagccgatcttgcatgtgtgtctgagcatccgaggttcgcatcttggtcgtggtggcagcatggtatggatgaaactgtactaatgaatgattcctgatggctgacttggtaaccgcagaactttgctatctgctatggtctgttacttctgatctacatgtctttcagttgaaatctgagcttataatgtttaatattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Biotin patch C

Segment C is a 162 or 111 basepair sequence, comprising half of a PCR product. The template for the PCR is a 2-kbp plasmid. The amplified section contains a central BamHI site and two EcoRI sites, each close to each end. The PCR was performed using Taq polymerase (New England Biolabs) in the presence of standard Taq buffer, 200 µM dATP, dCTP, and dGTP (New England Biolabs), 150 µM dTTP, and 50 µM biotin-dUTP (Roche). After cleanup, the PCR product was digested with EcoRI (New England Biolabs) and BamHI (New England Biolabs) and subsequently treated with Antarctic Phosphatase. Finally, the 162 and 111-bp segments were isolated by Qiaquick gel purification (Qiagen) in 2% low-melt agarose gel (TBE). The sequence of the PCR product is below.

Spacer DNA and dig handle D

Segment D is of length 3131 or 3142 bp. Its source is a 6273-bp plasmid with EcoRI and BglI sites at opposite poles. The plasmid was transformed into DH5-alpha cells (New England Biolabs). It was first digested with BglI and cleaned up. It was then treated with terminal transferase (New England Biolabs) in the presence of 0.25 mM CoCl$_2$ and 500 µM digoxigenin-dUTP (Roche). After cleanup, it was digested with EcoRI. The complete plasmid sequence is below.
Ligation

The ligation of the complete DNA construct was sequential. First, C and D were ligated together with an excess of C, with T4 DNA ligase (New England Biolabs) overnight and then digested with Proteinase K (New England Biolabs). The ligation product, CD, was then isolated by gel extraction from a 1% TAE agarose gel using QiaEXII (Qiagen). CD was then treated with T4 polynucleotide kinase (New England Biolabs) in ligase buffer (New England Biolabs) and heat inactivated. CD was then ligated to B overnight. After ligation, the enzyme was heat inactivated; then, ATP was supplemented to 1 mM and the DNA was treated with T4 PNK followed by heat inactivation. Finally, segment alpha was added and the final overnight ligation was performed.

DNA design for viral packaging in the presence of PNA

The DNA used for these experiments was a ligation product of two NcoI fragments. The first fragment was the 15 kbp NcoI fragment from the φ29 genome. The other segment was a 12,252-bp PCR product out of lambda containing an NcoI site close to one end and a single PNA-binding site 5213 bp away from the NcoI site. The primer for the distal end contained an inosine at the 19\textsuperscript{th} base, replacing an adenosine. The PCR was performed using the TripleMaster PCR Extender System (Five-Prime, Inc.) with Tuning Buffer. After amplification and cleanup, the PCR product was digested with Endonuclease V (New England Biolabs), which cleaves the 5’-3’ backbone two bases 3’ of an inosine. The result was a 3’ overhang at one end of the DNA. Next, the DNA was treated with terminal transferase (New England Biolabs) in the presence of 500 µM dig-dUTP. The DNA was then digested with NcoI (New England Biolabs), cleaned up with the Qiagen Qiaquick kit, and then ligated overnight with the 15-kbp φ29 genome fragment. The complete sequence of the PCR product is below.

tttgtcccactccctgccctclgtcatacagatacagtgtgcgtgcagctttatgccccgagaagatgttgacgcctggaacatcgtcttcagagctcgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatgtgagctgtgcaagattcagggactgctgtagttttgatg
Chapter 4. Observations of DNA Twist During Packaging Reveal Details of Motor Mechanism and the Organization of DNA Within the Capsid

4.1 Abstract

Previous optical tweezers experiments have focused on how φ29 moves along DNA during packaging. In this chapter, I discuss a rotor bead experiment for direct observation of DNA twisting which reveals the full, three-dimensional trajectory of the packaging motor along and around its substrate. When the capsid is empty, the motor slightly underwinds the downstream DNA; as the capsid fills, the degree of twisting increases. Observations at low filling reveal that the same subunit of the packaging motor makes contact with the DNA at the same point in each 10-bp cycle. In addition, from the twist at late stages of packaging, the degree of supercoiling inside the capsid of a packaged virus can be estimated.

4.2 Introduction

Optical tweezers have proven to be powerful tools for the investigation of microscopic mechanical phenomena and the discovery of hidden chemical processes. However, traditional optical tweezers that utilize two tethered beads are limited in that they do not directly measure the force or extension of a biomolecule. Rather, the measurement is of the projection of that quantity onto the vector connecting the two beads. In the case of DNA translocases, where the convenient coordinate system is cylindrical, only the axial forces and extensions can be measured; mechanical processes in the radial or angular direction are hidden. Using the rotor bead assay discussed in the previous chapter, we will track the angular motion of the packaging motor of φ29 and hence reconstruct its complete, three-dimensional trajectory. This measurement should provide valuable information about supramolecular coordination within the packaging motor to generate high forces and processive motion.

High-resolution optical tweezers experiments have revealed many details about the mechanism of packaging (Moffitt et al., 2009). The motor translocates DNA in a burst of 10 basepairs followed by a dwell during which several events transpire, including binding of at least four ATP’s. Under high load, the 10-bp burst breaks up into four 2.5-bp steps, but ATP binding must wait until the four steps are complete. The chemical and mechanical cycles of the packaging motor, therefore, reset every 10 basepairs (Figure 1-8 in Chapter 1).
Other tweezers experiments regarding packaging of modified DNA revealed details of the interaction between the motor and the DNA (Aathavan et al., 2009). First, packaging is sensitive to the phosphates on the 5'-3' strand. If ten or fewer consecutive phosphates are neutralized (without significantly perturbing the DNA structure), packaging proceeds as normal, but if 11 phosphates are neutralized, probability of packaging that region drops dramatically. This was interpreted to mean that normal packaging behavior depends on every 10th phosphate. High-resolution experiments showed that if 10 or more phosphates are removed, the motor’s stepping rate drops significantly, and when it does step, it is likely to slip backward. Other features of the DNA, such as 3'-5' phosphates, bases, and sugars, were much less important to packaging. These studies demonstrated that the motor has a strong interaction with the DNA, with a spatial periodicity of about 10 basepairs (Figure 1-9 in Chapter 1).

It remains unknown what motor components interact with the DNA at 10-bp intervals. Given that there seems to be one important site on the DNA which must be bound, which subunit binds to it? The structurally unique DNA-binding subunit may be special in other ways. For example, it is not known how the motor’s fivefold symmetry is broken so as to generate four steps and to slowly bind four ATP’s. This unique subunit, if it exists, should be identifiable from the angle-resolved trajectory of phage packaging. Because DNA is helical with a pitch of about 10.5 bp/turn, a 10-bp step along the DNA does not correspond to a full turn around the DNA. This symmetry mismatch suggests that the phage should rotate around the DNA as it packages; alternatively, from the prohead’s point of view, the DNA should twist as it is pulled in.

4.3 Materials and Methods

Configuring the experiment

Rotor-bead tracking experiments were performed in the COMMOT instrument described in the previous chapter, using a polystyrene rotor bead rather than a magnetic bead. See Figure 1A. Prior to the experiment, proheads, gp16, and a DNA construct were mixed and incubated for 5 minutes, at the ratios described in Chapter 2. Packaging was initiated by adding 250 µM ATP, allowed to proceed for 30 s, and then stalled by adding 250 µM ATP-γS. Stalled complexes were then mixed with 2.88-µm anti-gp8-coated Protein G beads (Spherotech), the restriction enzyme ApaLI, and BSA. After 45 minutes, the stalled complexes were diluted into a “stall buffer” consisting of buffer 1/2X TMS (50 mM Tris, 50 mM NaCl, and 5 mM MgCl2, pH 8.0), and supplemented with 50 µM ATP and 50 µM ATP-γS. The experiment took place in a standard optical tweezers chamber which
contained, initially, the stall buffer. A 2.1-µm anti-digoxigenin-coated Protein-G bead (Spherotech) was drawn onto a micropipette via suction; after suction was released, the bead remained fixed in place. A bead containing stalled complexes was optically trapped and then briefly rubbed against the pipet bead until a tether was formed. On average, 5 to 10 bead pairs were fished before a single tether was formed. Next, the optical trap was released and a horizontal flow of 10-20 pN was used to stretch the tether between the two beads. A third “rotor” bead, usually a 0.9-µm streptavidin-coated bead (Spherotech) was then optically manipulated along the outstretched tether until an attachment was formed. Nearly all two-bead complexes were capable of binding a rotor bead. Once the streptavidin bead stuck, the optical trap was released, flow was stopped, the trap was reactivated in order to manipulate the phage bead, and the system was arranged into the configuration shown in Figure 1a.

Viral packaging

The tether was held at a constant tension of 10 pN. Buffers in the chamber were exchanged by flow, removing the ATP-γS and replacing it with 500 µM ATP, which is saturating. The flow was stopped immediately after packaging was detected; residual flow in the chamber was less than 1 µm/s. While maintaining constant force by computer-controlled feedback, the extension of the tether was monitored using the optical tweezers’ light lever while the position of the rotor bead was recorded on video as the tether shortened due to packaging. The observation ended due to one of three events: packaging halt; tether breakage; or collision of the rotor bead with the phage bead.

DNA design

The DNA construct used in these experiments is depicted in Figure 1b. At one end is a gp3 protein, required for packaging initiation. The terminal protein is followed by 21 kbp of DNA, then a short biotinylated patch, then a 3 kbp section of DNA, and finally a tail containing digoxigenin. For details on its assembly, see the Appendix. Briefly, a 15-kbp SpeI restriction fragment of viral gp3-DNA was ligated onto a 6-kbp plasmid DNA. The plasmid DNA was ligated through an EcoRI site to a short biotinylated patch. The patch was ligated to a 3-kbp segment of plasmid DNA, through a BamHI site, to provide a spacer between the rotor bead and the end of the DNA. That plasmid DNA had been modified at the far end with dig-dUTP by the action of terminal transferase. Because the dig groups were located on a single-stranded region, the rotor bead was not torsionally coupled to that end. Hence any twisting of the DNA will be relaxed by motion of the rotor bead, so the rotor bead’s motion will equal any twist imposed by the packaging complex (excepting hydrodynamic effects).

Viral materials
Wild-type proheads, consisting of capsid, connector, and pRNA, were extracted and purified from infected *B. subtilis* cultures by our collaborators, Paul Jardine and Shelley Grimes of the Department of Diagnostic and Biological Sciences at the University of Minnesota. Gp16 protein was expressed in *E. coli* and purified in Minnesota. Viral gp3-DNA was extracted from mature viral particles in Minnesota.

*Trepanated proheads*

In order to trepan the proheads, the proheads were first digested with RNase and then purified and resuspended in 1X TMS. These pRNA-free heads were repeatedly frozen at -20 C and thawed to room temperature for a total of 10 cycles. The heads were subsequently mixed with pRNA molecules produced by *in vitro* transcription in order to reconstitute the pRNA ring.

**4.4 Results**

*DNA rotation by wild-type packaging complexes*

The rotor bead moved around the DNA axis during packaging by wild-type complexes. The direction of motion of the bead was toward overwinding, which means that the packaging motor was underwinding the DNA (hence forcing the DNA to re-wind itself), so the rotor bead angular velocity is given a negative sign. Bead rotation is ATP-dependent and correlated with packaging; pauses in packaging are accompanied by a pause in rotation (Figure 2a,b,c), indicating that twist is induced by the motor. However, the rotor bead lags behind packaging, as seen in the figure; this is expected due to hydrodynamic drag.

The packaging motor twists the DNA, and the torque stored in the twisted DNA rotates the rotor bead through the viscous fluid. In order to correct for drag torques, we can use the Stokes relation to determine the torque on the bead and hence the angular displacement which exerts that torque (Bryant et al., 2003). Where $N_d$ is the drag torque, and $N_e$ is the elastic torque, $\gamma$ is the drag, $r$ is the rotor bead radius, $C$ is the torsional modulus, $\theta$ is the angle, $\omega$ is the angular velocity, and the subscripts $rb$ and DNA refer to the rotor bead and the DNA, respectively (Happel and Brenner, 1965; Landau and Lifshitz, 1969),

\[
N_d = -\gamma_{rb}\omega_{rb} = -14\pi\eta r^3 \omega_{rb}
\]

\[
-N_d = N_e = -\frac{C}{L}(\theta_{rb} - \theta_{DNA})
\]
Hence the bead’s angular velocity is proportional to the error $\theta_{rb} - \theta_{DNA}$. Using these equations, the DNA angle $\theta_{DNA}$ – the angle imposed by the packaging motor – is inferred from the bead angle $\theta_{rb}$. See Figure 2d,e,f for a comparison of bead angle and DNA angle.

Note that these equations are valid only for torques below the buckling torque of DNA, which for underwinding is -10 pN nm. A larger error would be hidden, because the DNA could not turn the rotor bead faster. Torques near -10 pN nm were rarely seen in the data and when such a torque was observed, the data was not considered for the subsequent analysis.

Twisting is a function of the length of DNA already packaged

The number of basepairs packaged varied from trace to trace because the complexes had been pre-stalled at a random location. In order to compare the different traces, the ratio $<\rho>$ of total DNA twist to total DNA length packaged was computed.

$$\langle \rho \rangle = \frac{\theta_{DNA}(t = \infty) - \theta_{DNA}(t = 0)}{L(t = \infty) - L(t = 0)}$$

This ratio varied considerably from trace to trace, with a mean of -2.4 °/bp and a standard deviation of 1.3 °/bp. Despite this variation, the ratio was never positive. Interestingly, as can be seen in Figure 3a, there is a correlation between the number of basepairs packaged and the twist ratio. This trend suggests that the twist ratio $\rho$ may change as the capsid fills.

In order to test the idea that the twist ratio changes with filling, the DNA twist can be plotted opposite the length of the tether (Figure 3b). If the ratio is constant during packaging, the twist should depend linearly on the tether length. The slight curvature that is observed means that the ratio increases in magnitude as the tether length decreases. To make this more quantitative, each packaging trace was divided into 1-kbp segments and the local twist ratio within each 1-kbp bin was computed. The data were then averaged across all traces, one bin at a time, as seen in Figure 4, where the mean within each bin is represented by a circle and the standard error of the mean is represented by the error bars. In order to avoid systematic errors due to residual flow at the beginning of the trace or to beads colliding at the end of the trace, those bins were removed from the data. A clear trend is visible: at the intermediate stage of packaging, when 15 kbp have been encapsidated, the local mean twist ratio $\rho = -2.5$ °/bp, and the ratio increases in magnitude toward the end of packaging.

Development of prohead trepanation
What about earlier stages of packaging? DNA rotation data was collected from about 25 different wild-type packaging complexes. The median tether length for the complexes before they were restarted was about 9 kbp, which means that the median complex had already packaged 15 kbp before being stalled. This distribution did not change significantly if the pre-stalling time was reduced. (This may be because of the ability of φ29 to start packaging from the middle of the DNA, presumably by quickly packaging a loop in the initiation phase (Rickgauer et al., 2006).) In order to observe DNA rotation during packaging under conditions of lower filling, a different strategy was needed.

Wild-type packaging complexes slow down as they fill (Rickgauer et al., 2008; Smith et al., 2001), an effect attributed to the increasing energetic cost of packing DNA into a diminishing, and increasingly charged, volume. Occasionally, however – a few percent of the time – a phage will not slow down and can package a length of DNA greater than its genome at constant velocity. These proheads must permit the DNA to be released so as to reduce the internal pressure, probably through a perforation in the head; we have termed them “trepanated heads.”

We discovered that alternately freezing and thawing wild-type proheads for 10 or more cycles causes the heads to become trepanated about 25% of the time, as judged by their packaging velocity (see inset to Figure 5a). However, initiation by freeze-thawed proheads is quite inefficient, with the number of tethers less than 10% of normal. Moreover, most tethers that did form were rather weak, rupturing quickly under a 10 pN load. These freeze-thawed proheads were not usable in rotor bead experiments. Fortuitously, it was found that reconstituting the pRNA ring with fresh pRNA raised to nearly wild-type levels both the initiation efficiency and robustness to force.

These findings indicate that freezing causes damage to both the capsid and to the pRNA. Damage to the capsid renders it unable to contain the DNA that is pulled in, resulting in the disappearance of forces resisting packaging. Damage to the pRNA causes the motor-DNA complex to be less stable and less resistant to force. It is likely that the ATPase ring does not bind well to a defective pRNA ring, which would explain the lower initiation efficiency. And even if a prohead does grab DNA, the motor is not stably bound to the pRNA (or the pRNA is not stably bound to the capsid) and ordinary force can dislodge it.

**Packaging by trepanated proheads**

Three packaging events have been recorded with a rotor bead assay in which the linear velocity remained constant throughout; in all three cases, the phage bead and the rotor bead collided within 45 seconds. See Figure 5a. Because the packaging velocity was much higher than in the wild-type
prohead experiments, the angular velocity of the rotor bead was considerably higher as well. Correction of the bead rotation to account for drag was therefore especially important. It was not possible to completely account for drag, because the distance between the phage bead and rotor bead dropped to zero, causing the apparent torque to diverge a few seconds before the end of the trace. Neglecting the one or two turns that were missed because of this effect, the three traces all exhibited an identical twist ratio $\rho = -1.2(+0.1, -0.4)$ (Figure 5b).

**Torque resistance of stalled complexes**

Using COMMOT in manipulation mode, the DNA of a stalled wild-type complex was twisted. When the magnet was removed, the rotor bead relaxed exponentially back to its original position (data not shown). This indicates that the stalled complex is torsionally constrained.

### 4.5 Discussion

**Reconstructing the full 3-dimensional trajectory of the system**

In this experiment, the viral capsid is fixed while the DNA moves up. As the DNA is known to move in 10-bp bursts, each burst results in an upward translocation of 10 bp, and a left-handed rotation of between 12 and 50 degrees depending on filling (Figure 6a). The system can also be viewed in a frame where the DNA is stationary and the packaging motor is moving down. Again, the motor takes a 10-bp burst down and rotates around the DNA to the left. The packaging motor takes a left-handed helical trajectory on the right-handed helical track. A few other DNA translocases have been studied for their DNA-twisting properties: the helicase EcoR124I (Stanley et al., 2006) and the close φ29 relative FtsK (Saleh et al., 2005). The helicase was, not surprisingly, found to track the helix. FtsK turned out to positively – not negatively – supercoil the downstream DNA at about $+2.4^\circ$/bp. The rather significant difference in twisting by FtsK must be due to mechanistic divergence, as will be discussed below.

**Interpretation of twisting by trepanated heads: motor mechanism**

DNA rotation by a trepanated prohead takes place in the absence of internal forces which could change the mechanism. Also, the external force of 10 pN has a minor effect on a non-rate-limiting, irreversible step in the mechanochemical cycle (Chemla et al., 2005; Moffitt et al., 2009). For this reason, this condition can be considered to be the “native” state of the packaging motor.
The twist induced by a trepanated prohead can be explained by considering the geometry of the motor and the geometry of the DNA (Figure 6). The highest-resolution structures of the packaging motor (Morais et al., 2008), obtained through cryo-EM, indicate that the ATPase and the pRNA are rings with fivefold rotational symmetry, shown in Figure 6b as DNA-binding motifs at multiples of 72° around a circle. B-form DNA in solution has a helical pitch of 10.4-10.5 bp per turn, so each basepair step subtends an angle of 34.3° – 34.6° (Wang, 1979; Wynveen et al., 2008). Contacts are made with a backbone phosphate group during the ATP-binding dwell, and the motor then bursts across 10 basepairs before binding the 10th downstream phosphate and resetting its chemical cycle. Across a 10-bp burst, the DNA backbone travels 343° - 346°. The DNA rotation that we have observed in our experiment, in this reference frame, is \( -\rho = +12° (-1°, +4°) \) per 10 basepairs, meaning that the DNA rotates over to make nearly perfect contact with the same subunit that it had contacted one cycle earlier.

Models in which the motor-DNA alignment is enforced in other ways, such as a 2.5-bp step, all predict significantly different rotation values. Although there may be more complicated models, in which the \( n^{th} \) subunit of an \( m\)-fold ring makes contact with the \( p^{th} \) basepair, that fall within error of the data, the simplest model is one in which the alignment is enforced only by one subunit, every ten basepairs. Note that the geometric considerations discussed here are not unique to backbone phosphates; any feature of the DNA that repeats every 10 bp could be responsible for registering the DNA with the motor. The data means that the motor makes the same contacts with the DNA every cycle, and the functional role of at least one subunit – one of the subunits that binds DNA –is identical with every cycle.

The unique subunit may be unique in other ways beyond its DNA-alignment function as well. The superficially fivefold-symmetric motor displays broken symmetry during the burst phase of the mechanochemical cycle. The motor moves in four steps that comprise a 10-bp burst. The subunit responsible for DNA phosphate binding in a particular cycle is functionally unique in two ways: it sustains a strong electrostatic contact and it appears to regulate the rate of transitions in the chemical cycle. It is appealing to propose that the unique subunit responsible for DNA-motor alignment is identical to the phosphate-binding subunit. The electrostatic interaction provides a mechanism for pulling the DNA into alignment. Proper alignment can also be a checkpoint that is required before the motor can take another step. This would explain the sensitivity of the chemical catalysis rate to the charge of the phosphate group (Aathavan et al., 2009).

There are two schemes for the mechanical motions in the ring, motivated by what we already know about coordination and DNA interaction, that are consistent with the twisting data. The first is shown in Figure 7a. The unique
subunit binds a phosphate during a dwell. During a burst, the unique subunit releases the phosphate and each of the other four subunits sequentially moves a DNA-contacting lever by 2.5 bp. The DNA makes contact with each of four subunits during the stroke, and the stroke contacts are steric rather than electrostatic in nature. After the burst, the unique subunit binds a phosphate again. It is possible that the unique subunit also strokes during the burst, but, if so, one of the other subunits must not stroke. The second model is shown in Figure 7b; in this scenario, the burst motion is a collective stroke of the entire ring. During the dwell, one subunit binds the DNA though a phosphate; the ring opens, breaking one intersubunit interface; one arm of the ring reaches 10 bp along the DNA; and then the arm retracts by 10 bp; and then the DNA is transferred back to the original subunit. In this model, the four 2.5-bp steps are the product of relative motion at four intersubunit interfaces. The motor contacts the DNA at only two places, and the other three subunits do not need to interact with the DNA at all, explaining why only the 10th phosphates are essential for packaging.

Comparison to other motors

This stepping pattern falls into the general category of inchworm mechanisms. Although the other four subunits of the motor also hydrolyze ATP and undergo significant conformational changes to help translocate the DNA, the motor binds to the DNA in exactly the same way, with the same subunit, in each cycle. The packaging motor can thus be considered a macromolecular inchworm. Unlike the rotary inchworm models of DNA translocation in FtsK and T4 packaging motor (Massey et al., 2006; Sun et al., 2008), in which each subunit individually inchworms, for φ29 the entire ring constitutes the inchworm.

A close comparison of the φ29 packaging motor and FtsK is warranted because FtsK is a close homolog, it has been crystallized, and its translocation and supercoiling properties have been studied. Structurally, the two proteins are closely related (see Figure 5 in Chapter 1), with two major differences (Massey et al., 2006). First, the φ29 ATPase does not contain a domain which is thought to be crucial for translocation by FtsK based on the crystal structure; the rotary inchworm model for FtsK cannot be valid for φ29. Secondly, FtsK has sixfold rotational symmetry while the φ29 packaging motor is a pentamer. FtsK was found to positively twist the downstream DNA by Δσ = +0.07 in a magnetic tweezers plectonemic assay (Saleh et al., 2005). In terms of the twist ratio, this is ρ = 2.4°/bp. This was rationalized with a 13-bp step and DNA contact precessing to the right by one subunit per step. Such a scheme fit nicely with the sequential rotary inchworm mechanism that had been proposed earlier; however, the authors do note that other models fit the observed rotation. It is clear, nonetheless,
that FtsK does not make periodic contacts with the same subunit each cycle, so it proceeds by a different mechanism than does φ29. The exact value of supercoiling introduced by FtsK had an elegant teleological justification: it is identical to the degree of supercoiling that naturally exists in *E. coli*, so that the enzyme’s activity does not disturb the natural state in the cell. Can such a biological explanation be found for the twist induced by φ29? The answer may lie within the capsid.

*Complete twist vs. filling curve*

Packaging wild-type proheads provided data for the later stages of filling; packaging by trepanated heads reveals the twist induced in the absence of internal pressure. Using the trepanated heads as the zero-filling point, the relationship between filling and twist become clear. As can be seen in Figure 5b, at zero filling the twist is less in magnitude that at the later stages of packaging. The twist ratio seems to be nearly constant over the first 50% of packaging but then changes sharply toward the end.

What could cause the increase in twist as the capsid fills? There are several possibilities. The first is that the motor mechanism may change with filling, and the twist is a function of motor mechanism. The second possibility is that the external twist that is measured in this experiment is a proxy for the internal state of the DNA. If the compacted DNA acquires a nonzero linking number due to motor-induced twisting or to the organization within the capsid, that linking number may be relaxed through the external DNA. Cryo-EM experiments suggest that the internal DNA configuration is not identical across different particles (Comolli et al., 2008), and optical tweezers experiments on the internal forces at the end of packaging also reveal heterogeneity. Hence it is likely that the linking number that evolves during packaging is not the same for all complexes.

Indeed, there is considerable variation in the twist ratio between particles. Figure 4b shows the same data as Figure 4a, but with the error bars representing standard deviations rather than standard errors. This variation is far greater than the experimental error, which can be conservatively estimated as ±0.4°/bp, so it is true variation in DNA twist. In addition to this static disorder – variation between particles – the system may also display dynamic disorder – variation within one trace – although this disorder is obscured by the thermal fluctuations of the bead.

Because the ratio is averaged across a 1000-bp region, it is unlikely that momentary changes in the packaging mechanism could have such disparate effects on ρ. For this reason, neither the static nor dynamic disorder can be explained by a packaging-mechanism-dependent twist. The DNA twist, at least toward the end of packaging, must contain information about the linking number of the enclosed genome.
Inferring the conformation of the packed DNA

If the encapsidated DNA is under strain due to nonzero linking number, perhaps the linking number can be measured if the DNA is decoupled from the motor. On two occasions, the packaging motor slipped backward on the DNA before re-engaging it (for example, Figure 2c). A few seconds later, the motor slipped again, this time catastrophically. Those few seconds were enough for the rotor bead’s position to be observed. In both cases, the rotor bead spun in the opposite direction from the usual, and the ratio of turns to slip length was, within error, equal to the ratio for normal packaging. The event can be rationalized if the internal DNA bears a linking number density $\sigma(s)$ where $s$ is a coordinate along the DNA contour; when the motor slips and releases a length $L$, the total linking number $Lk_{ext}$ stored in that length $L$ is also released to the external tether.

$$Lk_{ext} = \int_0^L \sigma(s) ds$$

Because the DNA is under tension due to the optical trap, the linking number in the newly-released DNA is converted entirely to twist, and the rotor bead rotates by an equivalent amount. Thus, the slip allows us to measure the local linking number density for the DNA that was released. Most importantly, this data is proof that the confined DNA does have a nonzero linking number.

Is the twist measured during packaging precisely equal and opposite to the linking number of the DNA within the capsid? If true, this would require that the internal end of the DNA is always constrained in both writhe and twist. This is the case at the end of packaging, as we have just seen. The reason for this constraint is not clear. Perhaps the DNA, at high densities, becomes conformationally “jammed,” which prevents relaxation of twist. However, the DNA is certainly not jammed at the beginning of packaging when the capsid is empty. It may be possible that at the beginning, the DNA is already constrained, since the gp3 is a “sticky” protein and may adhere to the inner surface of the capsid. For these reasons, the total DNA twist measured during packaging is not necessarily a quantitative measure of the total linking number of the stored DNA. However, it is illuminating to consider what the external DNA twist does tell us about the internal DNA conformation.

First, every packaging complex rotated the downstream DNA in a negative direction. Hence the internal linking number must be positive; there is a net chirality to the path of the DNA chain. Relative to the velocity vector of DNA entering the capsid through the portal, the path is right-handed.
To understand the internal DNA conformation more quantitatively, we must address how the linking number of the confined DNA is partitioned into twist and writhe. By minimizing the total energy (bending plus twisting), it can be shown that at zero force the ratio of twist to writhe is simply the ratio of the bending modulus to the twisting modulus (Charvin et al., 2004).

\[ \sigma_{total} = \sigma_{wr} + \sigma_{tw} \]
\[ \frac{\sigma_{tw}}{\sigma_{wr}} = \frac{B}{C} \]

This continuum model is valid as long as DNA is an isotropic, elastic rod at equilibrium. Under extreme forces or torques, the DNA response is nonlinear. For example, overwound DNA denatures at \( \sigma_{tw} = 0.05 \). Also, the dimensions of the capsid require that the DNA bend fairly sharply, with a radius of curvature less than 25 nm. DNA bent into circles less than about 8 nm in radius may be anomalously flexible due to local kinking (Cloutier and Widom, 2004, 2005; Crick and Klug, 1975; Mathew-Fenn et al., 2008; Peters and Maher, 2010; Wiggins et al., 2006), so the elastic bending model may no longer be entirely valid for a moderately full capsid. There is no consensus as to this effect (Peters and Maher, 2010).

Let us consider the measured twist ratio at the beginning of packaging, about -1.2°/bp, which corresponds to an internal linking number density of \( \sigma = +0.03 \), less than the denaturing torque. It would be expected that writhing of the DNA chain would constitute \( C/(C+B) = 2/3 \) of the total linking number; this writhing corresponds to a loop radius of 27 nm, about the radius of the longest axis of the capsid. This remarkably good correspondence to the structural data suggests that the packaging motor is somehow optimized to lay down DNA in an organized, spooling fashion inside the capsid. If the end of the DNA is constrained, the linking number change induced by the motor is optimal for spooling the DNA in a chiral manner. On the other hand, if the DNA end is not constrained, the boundaries of the capsid will force the DNA into loops of similar size, but without chiral preference.

At the end stages of packaging, when the observed twist ratio is about -5°/bp, the local internal linking number density must be \( \sigma = +0.15 \). However, the continuum model may no longer apply. The hypothetical onset of kinking takes place at a writhe density of 0.06 or so, meaning that writhing becomes energetically favorable and that any additional linking number takes the form of writhe, not twist. If it is true that the DNA softens, then, the partition would be \( \sigma_{tw} = 0.03 \) and \( \sigma_{wr} = 0.12 \). On the other hand, if DNA does not soften and the continuum model still applies, we find \( \sigma_{tw} = 0.05 \) and \( \sigma_{wr} = 0.10 \). Writhe will constitute something between 2/3 and 4/5 of the total internal linking number.
Using this partition model, we can predict the size of loops within the capsid as a function of filling, based on the externally measured twist (Figure 8a). At the beginning of packaging, the loops are large. If the linking number is entirely in the form of writhe, the loops will have radius 17 nm, while if there is some twist the loop size increases to about 25 nm. At the end of packaging, for both models, the loops are necessarily small, about 3.5 - 4 nm in radius. Note that there must be variation between individual packaging complexes as the DNA is stuffed into whatever voids stochastically form in the crowded intra-capssid environment.

This model can be made more concrete by comparing to recent structural data on the arrangement of DNA packaged into the φ29 capsid. Comolli and Spakowitz et al (Comolli et al., 2008) used cryo-EM to map the DNA density inside the prohead at different stages of packaging, ranging from 0 to 100% filled (Figure 9). At high filling, a highly periodic pattern was visible in the density for individual particles, as well as after averaging over thousands of particles. With a spacing of 2.4 nm, the DNA appears to be in a close hexagonally-packed structure. Interestingly, the averaged pattern looks the same from all sides, suggesting that the DNA is not arranged in a coaxial spool. Moreover, the pattern appeared only at the very last stages of packaging, when more than 90% of the genome was inside; this implied that the arrangement became rigid only at the end. The authors interpreted this to mean that, while there may be local organization like close-packing, there need not be any long-range organization such as spooling.

Given our evidence for spooling, it is worthwhile to re-examine the 3D reconstruction of the density in the fully packaged particle. The largest complete shell of density has a radius of 16.5 nm, which corresponds to \( \sigma_{\omega} = 0.035 \), or 1.2 degrees per bp. The smallest well-resolved shell has radius 7 nm, but there is also a ring visible only horizontally, with radius 4 nm, corresponding to \( \sigma_{\omega} = 0.14 \) and 5 degrees per bp. Surprisingly, the DNA density observed could account for only \( \sim 85\% \) of the genome (Comolli, personal communication). About 15% of the DNA density is “lost” during the averaging process, being disordered enough that it contributes only to the background. Although the DNA density cannot be traced as a direct measure of writhe, the general parameters are consistent with the measured external DNA twist. This further validates the idea that the external twist is coupled to the internal state.

Assuming that the missing density is distributed uniformly across the genome, we can predict the external twist if that twist is entirely generated by internal DNA spooling (Figure 10). Two models are considered here. One is a coaxial model in which the DNA is always spooled horizontally with maximum radius; the arrangement resembles a solenoid (Figure 10b). The second is a model in which the DNA takes the largest possible loop either
vertically or horizontally but not diagonally, so as to minimize its bending energy (Figure 10c). The relationship between length packaged and loop size for these models is shown in Figure 8b. The largest-loop model—which minimizes the bending enthalpy—does not fit as well as the coaxial model, which suggests that the bending energy is not minimized; other factors such as entropy, DNA-DNA repulsion, twist-bend coupling, or kinetic trapping could all play a role in determining the DNA configuration. See the Appendix for a discussion of the missing DNA density and an alternative way of accounting for it; it does not change the qualitative conclusions presented here.

The highly idealized model discussed above seems to predict the general behavior of the external twist. This partitioning model had been used to explain tweezers data concerning the extension of DNA molecules as a function of force and torque. There are some additional caveats when applying it to confined DNA. The DNA within the capsid is in a quite different extensional regime, akin to “negative tension” because the end-to-end distance is much less than the length of a freely jointed chain. The conformation of encapsidated DNA is much more analogous to plasmids or DNA minicircles—the molecules for which apparent softening was first noticed (Cloutier and Widom, 2004, 2005). Moreover, the dense environment of the capsid interior contains a highly negatively charged polymer so the interactions between different DNA segments, and condensed cations, cannot be neglected.

Interestingly, this experiment is a direct test of a prediction regarding the organization of packaged DNA. Spakowitz and Wang predicted that the DNA would tend to spool if it were twisted by the packaging motor, whereas it would form folds if a twist were not imposed by the motor (Spakowitz and Wang, 2005). Moreover, spools of DNA tended to align perpendicular to the axis of entry, while folds tended to align parallel to that axis. The organization most consistent with the data presented here is one with coaxial spools. If the DNA minimized its curvature by aligning with the long capsid axis, the resulting twist would have been somewhat less than was measured.

### 4.6 Conclusions

We have found that bacteriophage φ29 twists the downstream DNA by about -1.2 °/bp under conditions of low filling, and that the twisting increases as the capsid fills. The low-filling twist ratio is consistent with a mechanical model for coordination of the motor subunits in which one subunit makes contact with the DNA at the same point in each cycle, and the DNA is rotated to be aligned with that one subunit. At higher filling, the twisting is an
indication of the linking number of the already-packaged DNA, and it may be a direct measure of the topology of the path taken by the DNA.
Figure 4-1. (a) Experimental setup. The packaging complex is affixed to an optically trapped bead and the DNA stretched by attachment to a bead sucked onto a micropipette. During packaging in force-feedback mode, the motor will pull the DNA in and the instrument will move the pipet up in order to keep the force constant. At the same time, if the packaging motor is rotating the DNA, the DNA will in turn rotate the rotor bead. (b) DNA construct. The DNA starts with the terminal protein gp3, required for initiation. 21.5 kbp downstream is a patch of biotinylated DNA, followed by a nick, then 3.1 kbp of normal DNA, and then a single-stranded tail including digoxigenin. The rotor bead binds to the biotin patch and is thus torsionally coupled to the upper 21.5 kbp of DNA, but not to the bottom segment.
Figure 4-2. (a-c) Representative packaging traces. Green is the tether length; blue is the position of the rotor bead in revolutions. The tether length is constant at the beginning of the experiment as he buffer is exchanged; the rotor bead cannot turn because of flow. The motor in (c) slipped backward and briefly re-engaged the DNA. (d-f) Corrections for drag. Blue traces are
the position of the rotor bead; green is the inferred angle of the DNA, calculated by a correction for hydrodynamic drag of the rotor bead. In trace (f), the angular velocity of the counter-rotating bead during the slip was too great to permit recovery of the DNA angle after the slip.
Figure 4-3. (a) Relationship between the total amount of DNA packaged in a trace and the mean twist ratio across that trace. (b) For the trace in figure 3a, the angle-position curve. The curvature indicates that the local twist ratio changes with position.
Figure 4-4. (a) Local twist ratio depends on how much DNA is inside the capsid. Error bars are standard error of the mean. The two datapoints less than 5000 bp all come from one trace. Many traces covered the region 12kbp – 18kbp, so the standard error is much smaller than the standard deviation there. (b) The same data, but error bars are standard deviations.
Figure 4-5. (a) Representative trace of a trepanated prohead. Packaging proceeds at a constant rate of ~100 bp/s, not stopping around 7000 bp like a wild-type prohead (c.f. Figure 2). Note that the rotor bead collided with and stuck onto the phage bead around t = 160s. Inset: the fraction of traces that did not slow after packaging 18 kbp of DNA (not a rotor bead experiment). (b) Local twist ratio as a function of filling, including the trepanated heads as a zero-filling point (magenta square).
Figure 4-6. (a) The trajectory of the DNA required to bring the 10th downstream 5’-3’ phosphate into register with the packaging motor. An upward translation of the DNA is accompanied by a left-handed rotation. (b) Essential geometry of packaging motor and DNA. This is a top view of the packaging complex, with the DNA moving out of the page. The pentameric motor, colored in ochre, has five potential DNA-pushing motifs, indicated by the vertices of the pentagon. The DNA has a pitch of about 34.5 degrees per basepair, so the 10th basepair downstream is located at about 345°. A 15° rotation every 10 bp, therefore, would bring the DNA into register with the same subunit (the bottom vertex of the pentagon).
Figure 4-7. Models for the functional role of the subunit that makes the periodic DNA contact. The DNA-contacting subunit stabilizes the motor on the DNA during the dwell. (a) The other four subunits hydrolyze ATP and translocate the DNA, each making nonspecific contacts during a 2.5-bp stroke. At the end of the cycle, the original subunit re-engages the DNA. (b) One of the intersubunit interfaces opens and the ring becomes a shallow helix. The helical oligomer reaches 10 basepairs along the DNA, engages it, and then retracts in four steps. The intervening subunits need not contact the DNA. At the end of the cycle, the contact is transferred to the original subunit.
Figure 4-8. (a) Predicted size of loops within capsid based on the measured twist. The blue dashed line is for the case where the internal writhe is identical to the external measured twist. The magenta dashed line is for the case where two-thirds of the linking number change is devoted to internal writhe, and one-third to internal twist. (b) Comparison of the predicted loop size with two possible models for DNA organization. The green line is for a coaxial spool; the black line is for a model in which the DNA always forms the longest possible loop in a horizontal or vertical direction.
Figure 4-9. DNA density within a packaged phage head seen by cryo-electron microscopy. (a,b) After averaging over many particles, a periodic pattern in the density is visible from all directions. The periodicity is about 2.3 nm, consistent with close-packed DNA. (a) The density across a radial section (blue line on the left). The values indicate the radius of the density ring. (b) The density across a transverse section (blue box on the left). (c) 3D reconstruction reveals shells of density. Figure courtesy of Luis Comolli, Andrew Spakowitz and Ken Downing (Comolli et al., 2008).
Figure 4-10. External twist ratio predicted by two models for DNA organization. The magenta solid line corresponds to the coaxial spooling model; the black dashed line is for the maximum-loop-size model. (b) Schematic of the coaxial spool model. (c) Schematic of the maximum-loop-size model. Neither schematic is exactly to scale.
Appendix

Detailed materials

DNA Construct

The DNA used in packaging experiments was a product of ligating DNA segments of four different lengths, named A, B, C, and D. See the Appendix of Chapter 3 for the sequences of the different segments. Segment A is a 14,928-bp sequence formed by digesting gp3-DNA of viral origin by SpeI (New England Biolabs) and then purifying with phenol-chloroform extraction. Segment B is of two possible lengths: 5869 and 5682 basepairs. Each fragment comprises one half of a 11551-basepair plasmid with SpeI and BamHI sites at opposite poles. The plasmid was transformed into E. coli strain jtk164 rbs 24, which was necessary because the plasmid contains a pir phagemid origin of replication (gift of J. Kittleson and J. C. Anderson). The plasmid was first digested with SpeI, then treated with Antarctic Phosphatase (New England Biolabs), then digested with BamHI (New England Biolabs), yielding the two fragments of nearly-identical length. Segment C is a 162 or 111 basepair sequence, comprising half of a PCR product. The template for the PCR is a 2-kbp plasmid. The amplified section contains a central BamHI site and two EcoRI sites, each close to each end. The PCR was performed using Taq polymerase (New England Biolabs) in the presence of standard Taq buffer, 200 µM dATP, dCTP, and dGTP (New England Biolabs), 150 µM dTTP, and 50 µM biotin-dUTP (Roche). After cleanup, the PCR product was digested with EcoRI (New England Biolabs) and BamHI (New England Biolabs) and subsequently treated with Antarctic Phosphatase. Finally, the 162 and 111-bp segments were isolated by Qiaquick gel purification (Qiagen) in 2% low-melt agarose gel (TBE). Segment D is of length 3131 or 3142 bp. Its source is a 6273-bp plasmid with EcoRI and BglII sites at opposite poles. The plasmid was transformed into DH5-alpha cells (New England Biolabs). It was first digested with BglII, then treated with terminal transferase (New England Biolabs) in the presence of 0.25 mM CoCl$_2$ and 500 µM digoxigenin-dUTP (Roche). After cleanup, it was digested with EcoRI.

The ligation of the complete DNA construct was sequential. First, C and D were ligated together with T4 DNA ligase (New England Biolabs) overnight and then digested with Proteinase K (New England Biolabs). The ligation product, CD, was then isolated by gel extraction from a 1% TAE agarose gel using QiaEXII (Qiagen). CD was then treated with T4 polynucleotide kinase (New England Biolabs) in ligase buffer (New England Biolabs) in the presence of 1 mM and the DNA
was treated with T4 PNK followed by heat inactivation. Finally, segment A was added and the final overnight ligation was performed.

It should be noted that the 8 possible sequences of segment BCD have identical length to within 200 basepairs; they could not be distinguished in the optical tweezers. Segment A was contaminated with an equal number of 4.3-kbp SpeI fragment from the viral genome. These segments could not form tethers in the optical tweezers because the proheads could not initiate packaging on them.

**Theoretical and numerical treatment of rotor bead dynamics**

*Basic physics*

In this experiment, the packaging motor twists the downstream DNA, engendering a torque stress. This torque then pulls the rotor bead around, hence reducing the strain in the helix. If $\theta_p$ is the angle determined by the packaging phage and $\theta_b$ is the present angle of the rotor bead, the torque $N$ on the DNA is given by the Hookean spring equation

$$N = -\frac{C}{L} (\theta_p - \theta_b)$$

Here, $C$ is the torsional modulus and $L$ is the length of the rotationally coupled segment of DNA. The torque on the rotor bead is equal and opposite to the torque on the DNA. For objects of this size (1 um) in water, the Reynolds number $Re = \rho v L / \eta \sim 10^{-11}$ so inertial effects can be neglected in comparison to viscous effects. Under these conditions, an object moving with angular velocity $\omega_b$ experiences an instantaneous drag torque proportional to the angular velocity.

$$N_d = -\gamma_b \omega_b = -\gamma_b \dot{\theta}_b$$

The proportionality constant is the drag coefficient $\gamma$, which is particular to the trajectory of the object. In the case of the rotor bead, which is rotating about a point which is fixed to the DNA, the drag coefficient is

$$\gamma_b = 14\pi\eta r^3$$

This result has been proven rigorously (Happel and Brenner, 1965; Landau and Lifshitz, 1969). However, the following argument will appeal to the biophysical experimentalist. It follows the general approach of Happel and Brenner.

*Derivation of the drag coefficient of a sphere rotating about an axis tangent to its surface*
Here we will show that the drag force for rotation about the DNA axis can be written as a linear combination of pure translational drag and pure spinning (revolution about the sphere center) drag. In general, calculating the hydrodynamic coupling of a particle with the fluid in which it is embedded involves first calculating the fluid flow field in the presence of the particle and other boundary conditions, and then integrating the resulting fluid stresses over the surface of the particle, yielding the total force on the particle.

The Navier-Stokes equation and the continuity equation are used to compute the fluid flow and pressure. The relevant physical parameters are the fluid density $\rho$ and dynamic viscosity $\mu$. Assuming a structureless, Newtonian fluid, the equations are as follows.

\[
\rho \frac{D}{Dt} \vec{q} = \rho \left( \frac{\partial \vec{q}}{\partial t} + \vec{q} \cdot \nabla \vec{q} \right) = -\nabla p + \mu \nabla^2 \vec{q}
\]

\[\nabla \cdot \vec{q} = 0\]

The terms on the left-hand side of the Navier-Stokes equation describe the rate of change of momentum of an infinitesimal fluid element; the right-hand side describes the forces that act on the fluid element.

Although the general solution to the Navier-Stokes equation is not known, several assumptions can be made which considerably simplify the equation. Both assumptions relate to the Reynolds number.

First, let us consider the term on the left-hand side that describes the inertia of a fluid element, $\rho \vec{q} \cdot \nabla \vec{q}$, and compare it to the viscous term of the right side, $\mu \nabla^2 \vec{q}$. For a particle of characteristic length $L$ and velocity $V$, the inertial term should be proportional to $\rho c L^2$ while the viscous part should be proportional to $\mu V L$. Reynolds pointed out that the ratio of these terms characterizes the relative importance of those two physical effects, and it became known as the eponymous Reynolds number $Re = \rho VL/\mu$. For water at 25°C, the ambient temperature for these experiments, the density is about $10^3$ kg/m$^3$ and the dynamic viscosity 0.001 Ns/m$^2$. For the particles observed in these experiments, $L \sim$ 1 micron and $V \sim$ 1 micron per second, yielding $Re \sim 10^{-6}$. The assumption that the inertial term can be entirely neglected is known as the Stokes, or “creeping motion” assumption and will be employed for the remainder of this section. The treatises of Oseen and, later, Proudman and Pearson, demonstrated that the exact higher-order dependence of the hydrodynamic coupling on $Re$ depends on the boundary conditions. For linear translation of a sphere, the lowest-order correction to the drag force is linear in $Re$, a very small correction indeed:
The second simplifying assumption is that of a steady flow, namely, that $\rho \left( \frac{\partial \vec{q}}{\partial t} \right)$ can also be neglected. Happel and Brenner show that nondimensionalizing the Navier-Stokes equation yields

$$Re_{rot} \frac{\partial \vec{q}}{\partial t} + Re_{trans} \vec{q} \cdot \nabla \vec{q} = -\nabla p + \nabla^2 \vec{q}$$

Where $Re_{trans}$ is the usual Reynolds number and $Re_{rot} = l^2 \omega \rho / \mu$ is the rotational Reynolds number, also $\sim 10^{-6}$ for a typical rotor bead experiment. Under this rescaling, the right-hand side is of order one, so the left-hand side can be safely neglected.

Thus, in this time-independent Stokes approximation, the equations describing the hydrodynamic flow of an incompressible, viscous fluid are now tractable. Flow is denoted by $q$ and hydrodynamic pressure by $p$; $\eta = \rho / \mu$ is the kinematic viscosity.

$$\eta \nabla^2 \vec{q} = \nabla p$$
$$\nabla \cdot \vec{q} = 0$$

The solutions for flow and pressure in these partial differential equations are uniquely specified by the boundary condition assumption of no slipping between the fluid and the object at the object’s surface. That is, the velocity of the fluid at a point on the object surface must be identical to the velocity of the surface itself. Given the geometry and trajectory of the sphere, therefore, the flow and pressure are determined.

In general, the coupling of an object’s motion to the flow is evaluated by integrating the hydrodynamic pressure tensor $\Pi$ over the object surface. The pressure tensor is defined:

$$\Pi = -lp + \mu [\nabla \vec{q} + (\nabla \vec{q})^T]$$

The boundary condition is as follows. Where $v$ is the velocity of an element of the bead surface,

$$\vec{q}_{surface} = \vec{v}_{surface}$$

The boundary conditions of a pure linear translation of a sphere specify a unique solution $q_{trans}$, $p_{trans}$. Likewise, revolution of the sphere about an axis containing its center yields $q_{rev}$, $p_{rev}$. The motion of interest (rotation about the DNA axis) is a linear combination of translation and revolution, since the differentiation is distributive, as we can show explicitly at the bottom of this
section. The boundary condition for rotation is also a linear combination of the boundary conditions for translation and revolution:

\[
\vec{v}_{\text{surf,rot}} = \vec{v}_{\text{surf,trans}} + \vec{v}_{\text{surf,rev}} \\
q_{\text{surf,rot}} = \vec{v}_{\text{surf,rot}}
\]

The flow solution for rotation, \( q_{\text{rot}} = q_{\text{trans}} + q_{\text{rev}}, \) will consequently satisfy the boundary condition for rotation, because \( q_{\text{surf,trans}} = \nu_{\text{surf,trans}} \) and \( q_{\text{surf,rev}} = \nu_{\text{surf,rev}} \). Because the Stokes equations are linear, \( p_{\text{rot}} = p_{\text{trans}} + p_{\text{rev}} \) as well.

Likewise, the pressure tensor \( \Pi \) is linear in \( p \) and \( q \), so the integral of \( \Pi \) over the sphere surface will be a linear combination of the integrals for pure translation and pure revolution. This integral, which we call the drag force, represents the total force on the sphere due to hydrodynamics.

Hence, if the trajectory of the sphere’s surface can be described as a linear combination of translations and revolutions, the total force required to move the sphere through that trajectory is simply the linear combination of the force required for the translation plus the force required for the revolution.

For a sphere of radius \( r \) undergoing a pure translation, the drag force generally has the form

\[
\vec{F} = -\gamma_{\text{trans}} \vec{v} \\
\gamma_{\text{trans}} = 6\pi\eta r
\]

Similarly, for pure revolutions, the drag torque is

\[
\vec{T} = -\gamma_{\text{rev}} \vec{\omega} \\
\gamma_{\text{rev}} = 8\pi\eta r^3
\]

The rotation of the rotor bead around the DNA axis can be written as a linear combination of translation and revolution. This can be seen explicitly in the following. The motion of the rotor bead is such that the point of contact between the bead and the DNA does not move or slip. The center of the bead is moving with a trajectory (letting radius \( r = 1 \))

\[(x_c, y_c) = (\cos(\omega t), \sin(\omega t))\]

Relative to the bead center, a point \( \alpha \) is moving with trajectory
\[(x_a', y_a') = (\cos(\omega t + a), \sin(\omega t + a))\] (prime denotes relative to bead center)

Hence the motion of point \(a\) in the Cartesian frame is
\[(x_a, y_a) = (x_c, y_c) + (x_a', y_a') = (\cos(\omega t + a) + \cos(\omega t), \sin(\omega t + a) + \sin(\omega t))\]

Differentiating,
\[
\left(\frac{\dot{x}_a}{\dot{y}_a}\right) = \left(\frac{\dot{x}_c}{\dot{y}_c}\right) + \left(\frac{\dot{x}_a'}{\dot{y}_a'}\right)
\]
\[
\frac{1}{\omega}(\dot{x}_a, \dot{y}_a) = (-\sin \omega t - \sin \omega t \cos a - \cos \omega t \sin a, \cos \omega t + \cos \omega t \cos a - \sin \omega t \sin a)
\]

We can identify the velocity of the surface point \(a\) as being a linear combination of the velocity of the reference frame and the velocity of \(a\) in that frame. Now, in order to translate the bead center in a circle, the bead center velocity must be \(r \omega\) and any forces on the bead body are exerted at a distance \(r\) away from the axis of rotation. Hence we have, for a bead rotating about an axis tangent to its surface,

\[\vec{T}_{trans} = -\gamma_{trans} r^2 \vec{\omega}\]

And, thus, for this motion,
\[\vec{T}_{total} = \vec{T}_{trans} + \vec{T}_{rev} = 14\pi \eta r^3 \vec{\omega}\]

**Analytical solution**

The basic equations of this system are those of an overdamped harmonic oscillator where the spring constant depends on the length of the DNA. As the phage packages, the length of the DNA decreases, hence stiffening the spring. If the phage also twists the DNA during packaging, the angle \(\theta_p\) changes. For now, we will assume that the amount of twist added per basepair packaged, \(\rho\), is a constant. The differential equation describing the rotor bead, then, is

\[\dot{\theta}_b = \frac{C}{\gamma(L_0 - vt)(\rho vt - \theta_b)}\]

Where \(L_0\) is the initial length of the DNA and \(v\) is the (constant) packing velocity. The solution to this is
\[ \theta(t) = \frac{\rho v}{C - \gamma v} \left\{ C t - \gamma L_0 + \gamma L_0 \left( 1 - \frac{vt}{L_0} \right)^C \right\} \]

The last two terms are the transient response of the bead, as seen by its exponential decay. During the transient time, the phage has twisted the DNA but, because the spring is very soft, the rotor bead is not moving as quickly. The first term is the steady-state rotation of the bead. It is important to note that for large times, the bead’s angular velocity \( \dot{\theta}_b = \frac{\rho v}{1 - \frac{1}{\gamma v}} \) is slightly greater than the phage’s angular velocity \( \rho v \). During this phase, the rotor bead is catching up to the phage. This could introduce some bias into the measurement. If the induced twist per basepair \( \rho \) is constant, the measured twist per basepair \( \rho_b \) will be an underestimate at short times and an overestimate at long times. The fractional error in the steady state is \( \frac{1}{1 - \frac{1}{\gamma v}} = 1.7 \) at the maximum packaging velocity of 100 bp/s. In Figure A1a, the exact solution is plotted for an initial tether length of about 20 kbp and an induced twist ratio \( \rho \) of -1.2 °/bp and a rotor bead radius of 450 nm.

However, the bias is not known quantitatively under the experimental conditions. In particular, the packaging velocity will decrease as the capsid fills, potentially mitigating the aforementioned bias. There may be other effects, such as nonlinearities due to denaturing of the DNA at high torque. Finally, the bead will be subject to considerable Brownian fluctuations which could obscure the signal. In order to model the behavior of the rotor bead under realistic conditions, a different approach was needed.

**Langevin dynamics simulation**

A Langevin equation for the system will include thermal fluctuations in addition to the elastic and dissipative forces mentioned earlier. The torque on the bead will include a fluctuating torque \( N_f \):

\[ \gamma \dot{\theta}_b = \frac{C}{L} (\theta_b - \theta_p) + N_f(t) \]

This fluctuating torque is due to random collisions with water molecules and should satisfy the following:

\[ \langle N_f(t) \rangle = 0 \]
\[ \langle N_f(t)N_f(t') \rangle = 2\gamma k_B T \delta(t - t') \]

These equations can easily be used in a Brownian dynamics simulation. Because the time scale of the bead’s fluctuations is on the order of seconds or greater, a non-microscopic timestep, such as 33 ms, can be used yet
capture the essential dynamics. The simulations also include DNA
denaturing, nonconstant velocity, and nonconstant hydrodynamic drag.

In Figure A1, simulations of several different scenarios are shown. In all
cases, packaging was simulated to within 1000 basepairs of the end. Figure
A1b shows a trajectory at a constant packaging velocity and constant \( \rho \) (twist
per unit length packaged). Figure A1c shows a trajectory that slows as it fills
according to the data in Rickgauer et al (Rickgauer et al., 2008). It should be
noted that the slowdown in packaging reduces the difference between the
observed angle and the actual DNA angle. However, the rotor bead still lags
behind the DNA at early times. In order to solve the problem of lag, we can
include a correction: the angular velocity of the bead is proportional to the
torque exerted on the bead, which is proportional to the difference between
the bead angle and DNA angle. This correction is shown in Figure A1d.
Although this correction nearly eliminates the curvature, it introduces some
additional noise into the angle. The extra noise results from the angular
velocity, which is the derivative of the already-noisy bead angle data.

**Hydrodynamics of a bead between two other beads**

The above formula for the drag of a bead rotating about an axis tangent to its
surface is true in an unbounded fluid, where the boundary conditions are
specified on the bead surface and at infinity. However, in this experimental
geometry, there are nearby no-slip surfaces that must be considered. At the
beginning of a packaging experiment, the center of the 0.9 um-diameter
rotor bead is a couple of microns from the optically trapped bead but only 1
\( \mu \)m from the pipet bead. By the end of the experiment, the distance between
the rotor bead and the optically-trapped bead has also dropped to less than
1 um. The relevant parameter governing the hydrodynamics is the radius of
the rotor bead to the inter-bead distance; it is close to unity so significant
changes in drag are expected.

The drag coefficient of this system has not, to my knowledge, been treated
either analytically or numerically. To understand the hydrodynamics, we can
break the system into two smaller problems: the hydrodynamics of a bead
close to another bead, and the hydrodynamics of a bead between two
surfaces.

The first problem – orbital motion of a bead near another bead – has not
been solved either. However, it can be simplified into two cases – pure
rotations and pure translations – and the results superposed analogously to
the previous section. Because of the linearity of the equations, this result
should be exact. The form of motion that we are interested in can be
decomposed into: (1) translation of a bead perpendicular to the line
connecting it to another bead, (2) rotation of a bead about the axis
connecting it to another bead, and (3) rotation of a bead about an axis perpendicular to the line connecting it to the other bead.

The case of two unequal spheres translating and rotating perpendicular to their line of centers (1) was solved in the form of an infinite series by Davis (Davis, 1969). By weighting the values in Davis by the component of the orbit in each of those directions, we can calculate their contribution to the drag torque. At a ratio $R = 0.1$ (the inter-bead distance divided by the rotor bead radius), which is a lower bound for the experiment, the additional drag torque is 0.5 times the unbounded $(8\pi\eta r^3)$ and the additional translational drag is 0.3 (times $6\pi\eta r$). The case of two unequal spheres rotating about the axis connecting the two was treated by Jeffery (Jeffery, 1915). In this case, the drag torque for pure rotations is seen to increase by 0.1 at the same distance. Combining these three results, the net effect of the proximity of the rotor bead to another bead at this distance is about 1.5. This correction is approximated by a phenomenological fit,

$$
\gamma = \gamma_0 (1 + \frac{1.1}{1 + \exp\left(\frac{L}{500}\right)})
$$

Where $L$ is the distance between the rotor bead and the other bead, measured in nm. This fit has no physical basis other than that it reproduces the expected drag increase at several points.

There is no exact solution for the drag coefficient at an arbitrary position between two walls, let alone at an arbitrary position between two spheres. In order to estimate the drag of the rotor bead when surrounded by obstacles both above and below, we will simply superpose the contributions of each nearby obstacle. This assumption was first invoked by Oseen (Swan and Brady, 2010). This assumption cannot be justified based on linearity, because the boundary conditions used to solve for each individual obstacle are different from the boundary conditions describing the complete system. A recent computational paper by Swan and Brady, looking at the drag between two plane walls, showed that the errors in this assumption are about 5% when the bead is close to one wall, but that superposition is a 60% overestimate midway between the two walls (Swan and Brady, 2010). Given all the uncertainties in estimating the drag in the experimental setup used in this chapter, I sought to assess the effect the errors would have after propagation into the data analysis.

A set of forty runs of a Brownian dynamics simulation comprised a dataset that was analyzed using the approach described in the main text of Chapter 4. In these simulations, the drag coefficient varied with tether length as described above. The analysis was then performed using several different weights for the additional drag, ranging from 0 to 5 times the weight imposed
during the simulation. As can be seen in Figure A2, the results did not change. The reason is that, at the times that the additional drag becomes significant, the packaging velocity is already quite low, allowing the rotor bead to keep pace. These simulations indicate that errors in the hydrodynamic drag do not introduce significant bias in the measurement. The increasing magnitude of twist as the capsid fills cannot be attributed to errors in the drag.

**Two additional models of DNA organization**

Cryo-EM has been used to observe patterns of DNA density within a fully packaged φ29 head (Comolli et al., 2008). The authors found that maxima of the density took the form of five shells spaced by 2.4 nm. They further state that the pattern can be explained if the DNA is, on average, hexagonally close-packed with a strand-to-strand center-to-center distance of 2.8 nm. However, the volume comprised by DNA in the 3D reconstruction is about 30,000 nm$^3$, while the volume of 19.3 kbp of DNA, assuming a strand radius of 1.4 nm, is 40,000 nm$^3$. For the packing models that were discussed in Chapter 4, it was assumed that the “lost” DNA comes equally from across the genome, that its bending is identical to that of its neighborhood in the genome, and that the lost DNA is uniformly distributed throughout the capsid.

One possible resolution of the volume discrepancy is a re-assignment of the outermost density peak. It had been assumed that the outermost shell, with a maximum at 18.4 nm, was capsid; however, it is possible that the density is in fact DNA and the small shoulder around 21 nm could be the capsid (Luis Comolli, personal communication). If this is the case, the total volume available for DNA increases to 44,000 nm$^3$, which fits very well with the 90% efficiency of a hexagonal packing arrangement.

Assuming that the DNA can form 18.4-nm loops allows us to refine the two models for DNA organization already discussed. With the new dimensions, both models can fit the entirety of the genome within the capsid. The resulting DNA writing and twisting follows the same trends that were already mentioned in the main text. See figure A3. Both models invoke somewhat less writhing at low filling than the data suggest. At higher filling, the coaxial spooling model generally fits the data better than the maximum-loop-size model. Again, this suggests that the DNA is packed without minimizing its bending energy; other energy terms may be important, or it may become kinetically trapped in higher-writhe states.
Figure 4-A1. Expected data. All plots here assume 20 kbp of packaging with a twist ratio of -1.2 degrees per basepair and a rotor bead of radius 450 nm. (a) Analytical solution for the bead angle (blue) due to a constant twist ratio during packaging at constant velocity (magenta dotted line). The bead underestimates the twist at early times and overestimates it at later times. (b) Brownian dynamics simulation of the same system as (a). (c) The same as (b), but with a velocity that decreases as the capsid fills. (d) Correction of the bead angle (blue curve) using its angular velocity (green curve) is successful in recovering the original DNA angle (magenta dashed line).
Figure 4-A2. Higher-order corrections to hydrodynamics drag do not introduce significant bias in the data analysis. 40 Brownian dynamics simulations were performed including the phenomenological approximation for the drag of the rotor bead near the pipet bead and the optically trapped bead. These simulations were then analyzed for the rotation in each 1000-basepair region using the algorithms described in Chapter 4. (a) Analysis assuming that the drag of the bead was equal to the unbounded-fluid value and was constant regardless of inter-bead distance. (b) Analysis assuming that the drag of the bead depends on distance using the same phenomenological approximation that was used in generating the simulated data. (c) Analysis assuming a drag coefficient that changed 5 times faster than the actual values used in the simulation. The fact that all three results are statistically indistinguishable means that considerable uncertainty in the drag does not have a significant effect on the quantitative conclusions.
Figure 4-A3. Comparison of DNA packing models with the observations. (a) Loop size for the coaxial spooling model, assuming a 18.4-nm radius capsid, is in green. Loop size for the maximum-loop-size model, allowing for vertical loops, is in black. The data is shown in blue circles (assuming all external twist corresponds to internal writhe) and magenta squares (assuming 2/3 of external twist is converted in internal writhe). (b) The predicted external twist ratio for the two models shows that, at high filling, the coaxial spooling model (magenta) fits the data better than the maximum-loop-size model (black dashed line).
Chapter 5. Afterword

In this dissertation I have presented the first two-dimensional single-molecule observations of a DNA translocase in action. In each case, two variables were monitored simultaneously: the position of the molecular motor along the DNA, and a quantity related to the angle between the molecular motor and another component of the packaging complex. The ability to record in two dimensions opens up a wide variety of new experiments which can reveal new details about the packaging motor. Moreover, the ability to apply torques as well as forces provides an additional axis of control with which to probe the mechanochemical landscape.

Experiments on trepanated proheads will elucidate details of motor mechanism

Trepanated proheads twist DNA during packaging, which implies that the twist is an integral aspect of the motor mechanism. Given that the motor’s cycle is known to be a series of chemical steps – ATP binding, ATP hydrolysis, Pi release, and ADP release – it is not immediately clear which step corresponds to twisting of the DNA. Previous experiments showed that linear movement of the DNA coincides with Pi release (Chemla et al., 2005); if the power stroke of the motor is not exactly parallel to the DNA, perhaps twisting is required for the linear translocation step of 2.5 basepairs. However, the twisting that has been observed is consistent with motor-DNA alignment every 10 basepairs, and not consistent with alignment at a 2.5-bp scale. Considering the regulatory role of electrostatic interactions every 10 basepairs (Aathavan et al., 2009), it is possible that proper motor-DNA alignment must be achieved in order to trigger a new round of translocation – that is, the twist may occur prior to linear translocation. These questions can be addressed in several different ways.

Observations of packaging under varying torque conditions will reveal the energy landscape of twisting. A torque between -9 pN nm and +36 pN nm can be applied to the downstream DNA using the COMMOT scheme in manipulation mode. First, we can study the linear velocity of the motor as a function of applied torque. The effect of torque will be to slow the kinetic transition that corresponds to the rotation of the DNA. From transition state theory, the rate of this kinetic transition, \(k\), will have a roughly exponential dependence on torque \(N\), \(k(N) \sim e^{-(\Delta \theta N/kBT)}\) just as the linear translocation step rate depends exponentially on force. This angle to transition state \(\Delta \theta\) is intimately connected to the mechanics of the packaging motor: it provides an upper limit for the rotational step size and also for the width of the potential energy well in which the motor fluctuates. When combined with the linear distance to transition state, 1.1 Å, measured previously (Chemla et al., 2005), these measurements will define the two-dimensional configuration of the transition state. Interestingly, because the sign of the torque can be changed, the setup permits us to apply a torque that either “opposes” or “assists” packaging. Tilting the packaging energy landscape downward along the twist coordinate, could, counterintuitively, actually increase the velocity for movement along the orthogonal translocation coordinate.
Utilizing the capacity of COMMOT to independently vary force and torque, the coupling of linear and twisting motion can be assessed. The temporal relationship between the linear step and the twisting step will be apparent from the torque-force-velocity manifold. If force and torque are generated in the same step, the velocity will always be exponentially dependent on both parameters. But if they are different steps, there will be conditions under which one step can be made rate-limiting, rendering the velocity insensitive to the other parameter. The same analysis can be extended to place torque generation into the entire chemical cycle using the approach of (Chemla et al., 2005), in which $V_{\text{max}}$ and $K_m$ as a function of ADP, force, and torque reveal details of the structure of the cycle.

Ultrahigh-resolution optical tweezers have revolutionized the investigation of mechanochemistry because they are capable of resolving individual turnovers of the packaging motor (Moffitt et al., 2009). But such techniques are close to their fundamental resolution limit and, at the catalytic rate and step size of the packaging motor, the steps can barely be resolved. In general, the signal-to-noise ratio will go as the inverse square root of the bandwidth of the measurement, so a motor that takes smaller steps or is faster will be difficult to track with optical tweezers. Interestingly, angular measurements may be capable of greater time resolution.

Direct angular measurements such as rotor bead tracking have a unique advantage over optical tweezers for DNA translocases. A measure of the quality of a measurement is the signal-to-noise ratio, which is proportional to the signal size $\Delta \theta$ and the stiffness $K$, and inversely proportional to the square root of the drag coefficient $\gamma$, the thermal energy $k_B T$, and the measurement bandwidth $B$ (Moffitt et al., 2006):

$$SNR \sim \frac{K}{\sqrt{k_B T \gamma B}} \Delta \theta$$

For angular measurements, the signal is an angle change $\Delta \theta$ and the stiffness $K = C/L$, the twist stiffness; for a linear measurement, the signal is an extension change $\Delta x$ and the stiffness is a complicated function, strongly dependent on force. The first advantage of angular measurements, then, is that the SNR does not depend on force, thus permitting measurements under less-invasive conditions. The second advantage relates to the drag coefficient. For the rotor bead, $\gamma = 14 \pi \eta r^3$, whereas for linear translocation, $\gamma = 6 \pi \eta r$. The much steeper dependence of rotational drag on bead radius means that, for sufficiently small bead size, the SNR for a rotor bead system can greatly exceed that of a conventional optical tweezers. Note that the bead size for the optical tweezers is limited to that which can be trapped, but there is no lower bound to the size of a rotor bead.

The angular step size for the known DNA translocases falls into the range of ~12°/cycle (for φ29) to ~35°/cycle (for helix-tracking motors). For the φ29 system, making a 12° turn every 10 bp at about 10 Hz, the rotor bead measurement is superior to the linear measurement for a rotor bead size below about 100 nm – only 4.5 times smaller than that used for this study. Note that for helix trackers like RNA
polymerase, because the angular step size is larger and the linear step size is smaller, the bead need not be so small.

Of course, the rotor bead angle must be observed with ~ 10° instrumental accuracy at 10 Hz, so the current video-rate tracking of a micron-size bead will be insufficient. The current setup, using near-ultraviolet bright-field illumination, is capable of resolving beads down to a diameter of about 600 nm; smaller beads simply do not scatter enough light. This can be overcome in several ways. The simplest solution would be to increase the signal by using a bright rotor particle, in the form of a fluorescent bead, nanoparticle, or quantum dot. The centroid of such particles can be tracked to ~20 nm resolution, sufficient for this purpose. In order to reduce background, alternative illumination techniques can be easily employed, such as epi-illumination or dark-field illumination. Another possibility would be phase contrast microscopy, which would allow the use of non-emitting rotor beads.

Experiments on wild type proheads will reveal the mechanics of the confined DNA

The above experiments will provide information about motor mechanism in the absence of stresses due to internally confined DNA. A different set of experiments has the potential to elucidate the physics of the DNA itself. Twisting of the DNA outside the capsid is a measure of the local linking number of the DNA that has entered the capsid. The variation in linking number from complex to complex suggests that the DNA, at least at the end stages of packaging, does not reach one equilibrium state. Either a wide variety of states have equal energy, or the system becomes kinetically trapped in higher-energy states. These two scenarios can be differentiated by looking at the twist rate as a function of packaging rate. Slower packaging would allow the confined DNA more time to equilibrate, and hence should change the average linking number. For this reason, a measurement of twist rate as a function of [ATP] will provide insight to the energetics and dynamics of DNA within the capsid.

The aforementioned experiments – tracking the twist of the outside DNA – are an indirect measure of the internal linking number. In order to more directly measure the internal linking number, the motor must be removed from the system. Occasionally, a packaging motor will slip off the DNA – but then the tension in the tether drops to zero, and any torque can be relieved as writhe rather than through the rotor bead. If a motor were to slip without resulting in tether loss, on the other hand, the rotor bead’s trajectory will reveal precisely the linking number within the segment of released DNA. This could possibly be accomplished in a number of different ways. It is known that high concentrations of ADP promote slipping, but this may have an effect on motor mechanism as well. Another possibility – one capable of revealing the linking number of the entire genome – would be to crosslink the internal DNA end to the capsid. In this case, a packaged complex could be stored in a stall buffer, assembled into a rotor bead geometry, and then ruptured under high force. As the DNA is released, all of the stored twist and writhe would be relieved through the rotor bead. There are several possible ways of performing the crosslinking; click chemistry, introduced genetically into the capsid and terminal protein, will likely be the simplest, most specific, and most biocompatible (Sletten and Bertozzi, 2009).
Experiments on motor mutants will reveal the protein responsible for DNA twist

While observations of twisting by trepanated heads strongly suggest that the same motor subunit makes contact at the same point in every cycle, they do not uniquely identify which protein species is involved. The pentameric ATPase, though likely responsible for the translocation steps, may not make the periodic 10-bp contact, as long as a different part of the complex always makes the periodic contacts. For example, the connector may make the 10-bp contact that aligns the DNA with the same subunit. This would be consistent with the connector’s putative one-way valve function: the strong electrostatic contact with a phosphate group could prevent slipping of the DNA while the motor reloads ATP. Motor variants in which the possible DNA-contacting residues have been mutated can be used to expose the functional role of the different motor components. Deletions of the connector loop, though known to cause little change in linear packaging velocity (Rockney Atz, Shuhua Ma, K. Aathavan, Carlos Bustamante, Jiali Gao, Dwight L. Anderson and Shelley Grimes, in preparation), may result in a different twist rate simply because the DNA may rotate around to a different subunit more easily. Likewise, mutations to the ATPase that disable one subunit may not change the DNA twist if the twist is enforced by the connector only.

Ten years of φ29 packaging at the single molecule level have revealed many unexpected facts about the physics that underlie viral biology. But each new discovery brings a monotonic increase in the number of unanswered questions. When does the motor twist the DNA? How is the DNA arranged? Why does the DNA appear to not be in the minimal-energy configuration? Answering these questions will be difficult. The potential of multi-dimensional observation, however, is clear. In the years ahead, these techniques will surely help to unravel the complicated knot that is viral packaging.


