Title
A Single-molecule Approach to Study Multimeric Molecular Motors and Optimal Thermodynamic Length

Permalink
https://escholarship.org/uc/item/9nw2s918

Author
Tafoya, Sara

Publication Date
2017

Peer reviewed|Thesis/dissertation

A Single-molecule Approach to Study Multimeric Molecular Motors and Optimal
Thermodynamic Length

By

Sara Tafoya

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Carlos J. Bustamante
Professor Eva Nogales
Professor Susan Marqusee
Professor David V. Schaffer

Fall 2017

Abstract

A Single-molecule Approach to Study Multimeric Molecular Motors and
Optimal Thermodynamic Length

By

Sara Tafoya

Doctor of Philosophy in Biophysics

University of California, Berkeley

Professor Carlos J. Bustamante, Chair

Single molecule techniques are uniquely informative for kinetic processes. As a result, in recent years they have become the methods of choice to interrogate many complex biomolecular systems (Bustamante & Tafoya 2017). During my PhD, I used optical tweezers, a technique for single-molecule manipulation, to study various biological processes.

First, I revisited the high internal pressure built inside the viral capsid of the bacteriophage ϕ29 during genome encapsidation (Liu et al. 2014b). During assembly of double-stranded DNA bacteriophages, the viral genome is encapsidated by a DNA packaging motor. High internal pressure builds up inside the viral capsid as a result of entropic and electrostatic repulsive forces resulting from DNA confinement. Previous single-molecule studies have determined the value of the internal pressure to be as high as 110 pN towards the end of DNA packaging. However, this value seemed overly high based on theoretical calculations. Using higher resolution data than in previous studies, my colleagues and I showed that the internal pressure reaches $\sim 20 \pm 7$ pN at 100% capsid filling, which is in better agreement with previous theoretical models.

Second, I determined the molecular mechanism for inter-subunit coordination in a viral ring ATPase. Subunits in multimeric ring motors must coordinate their enzymatic activity to perform their function (Tafoya et al. 2017). The bacteriophage ϕ29 DNA packaging motor is a pentameric ring ATPase whose subunits have been shown to operate in a highly coordinated manner. Therefore, this system is ideal to investigate how global subunit coordination can arise from stochastic processes and local molecular interactions. Using single-molecule optical tweezers and targeted mutagenesis, I showed that coordination arises from inter-subunit enzymatic regulation. The subunits use their arginine finger to promote nucleotide exchange and to activate ATP hydrolysis in their neighbors. These regulatory processes

1

display similar features to those observed in small GTPases.

Third, in light of what I learned about the $\phi 29$ DNA packaging motor's operation, I reviewed various mechanisms of small GTPase-like regulation in different motor proteins (Tafoya & Bustamante 2017). In particular, I highlighted the fact that all these mechanisms share a general feature: the motor's function is controlled by stimulation or repression of its ATPase activity, which is regulated allosterically by different factors.

Finally, I tested a prediction from fluctuation theorems to minimize the thermodynamic length in a process out of equilibrium (Tafoya et al. 2017b). Genome encapsidation by the $\phi 29$ DNA packaging motor is only an example of the multiple non-equilibrium processes occurring in the cell. In fact, to maintain their organization, biological systems must operate far from equilibrium, continuously utilizing and dissipating energy. Non-equilibrium theory is underdeveloped, but recent work has approximated the excess work in processes out of equilibrium. I tested this theory's predictions performing pulling experiments on a DNA hairpin. I found that the predicted minimum-dissipation protocols indeed require significantly less work than naive ones across a wide span of driving velocities.

Dedicated to Daniel, Lupyta and David

I discovered science while playing with you

Acknowledgements

I want to thank my PhD advisor, Carlos Bustamante, for the opportunity to work in his lab. Besides the various scientific skills that I learned from him, he also showed me how to navigate the challenges that a person doing science has to confront everyday. I will greatly miss his charm and knowledge about books, food, music and movies. I look forward to having him as a mentor and friend in the future. I also want to thank all the people that helped me along my doctoral studies. I want to give special thanks to Gheorghe Chistol and Shixin Liu, who initially took me under their wings and trained me in the studies of the $\phi29$ DNA packaging motor. In particular, I thank Shixin Liu for his friendship and advice on life, which I kept in mind well after he left the Bustamante Lab. I acknowledge the tremendous help that I got from Juan Pablo Castillo during the last two years. His critical view and constructive comments about my work improved it immensely. I thank him also for his valuable friendship in the hardest moments during my PhD, and for his sense of humor, which made many situations a lot more bearable. I acknowledge the constant support that I got from my collaborators: Paul Jardine, Shelley Grimes, David Sivak and Steve Large. In particular, I will take with me the memory of Shelley Grimes, who sadly passed away this year; her love and approach to science is an example of the scientist in its purest form, one that is almost lost in today's highly competitive scientific world. I also thank the advice of other members of the Bustamante Lab, who along the years have contributed to my career in one way or another: Rodrigo Maillard, Jae Yen Shin, Luis Ramirez, Shannon Yan and Antony Lee. I acknowledge all the members from the Bustamante Lab from 2011 to 2017 for their continuos scientific advice, as well as former members: Aathavan Karunakaran, Lacramoira Bintu and Sophie Dumont, for sharing with me their trajectory in science. I also want to acknowledge the program UC-Mexus, sponsored by the national council for science in Mexico (CONACYT) and the UC system, which provided me with a doctoral fellowship for 5 years. I want to thank Natalie Thung and Kate Chase for dealing with the many complications of having multiple sources of funding for my tuition. I thank Eva Nogales and Susan Marqusee for many conversations and advice; knowing them inspires me to continue my career as a woman in science. I want to thank the Biophysics program for all the incredible people that brought to my life and all the many activities that kept me motivated during these years. I want to give special thanks to my friends Itziar Ibarlucea Benitez, Yasin Ezber and Jacques P. Bothma for becoming my support network in this country. I also want to thank my parents, Jose and Susana, because it was their support as a child that provided me with more educational opportunities. I want to thank my siblings, Daniel, Lupyta and David for constantly challenging me to become a better person, a better thinker; and to my husband, Michael R. Lomnitz, for making me laugh every time I was struggling with work, and for being the best career counselor and the best friend that I have ever had.

Contents

Chapter 1

Introduction

1.1 Single-molecule Methods

Structure and function are the two faces of the biological coin; out of the fundamental tendency of matter to self-organize, the mechanism of evolution sculpts the structure to fulfill the function. Biologists have made an enormous effort to uncover the structures of proteins, nucleic acids, and their complexes, with the ultimate hope of understanding their function. Indeed, few advances in biology have been more impressive than those of the methods of structural determination—X-ray crystallography and NMR—during the last four decades. Because of these efforts, the Protein Data Bank contains today more than 100,000 proteins whose structures are represented at atomic or near-atomic resolution. In many cases, the structure itself provides sufficient hints as to its function, but more often than not such insight is not enough. Function directs not only the structure of the basic parts, but their interaction in larger and more complex molecular organizations and the structures can at best provide only a static picture.

In contrast, methods to follow the dynamics of these structures have developed comparatively more slowly. The reason can be traced to the nature of dynamics itself in molecular systems. The molecule is an open thermodynamic system that can exchange energy and matter with a large thermal bath kept at a temperature T. Each degree of freedom of the molecule can exchange energy with this bath. This constant exchange of energy, in turn,

is responsible for inducing transitions across energy barriers that appear as fluctuations. Furthermore, any kinetic process involved is stochastic. Each step in a reaction has a different rate; each intermediate species has a different lifetime. For each single kinetic step, the lifetimes of each species are distributed exponentially and can take any given value in an unpredictable way. Thus, signals in molecular processes are small, rapidly varying and highly fluctuating magnitudes. How can we best capture this information?

Biochemical methods are robust because the instantaneously measured signals are proportional to the number of molecules present. However, the averaged bulk signal provides little or no information about rare states and fluctuations. In contrast, fluctuations of the signal can be easily captured in single-molecule experiments, and tiny signals can now be measured by highly specialized instrumentation. Because single-molecule techniques are uniquely informative for kinetic processes, over the last two decades they have often become the methods of choice to interrogate many complex biomolecular systems.

During my PhD, I took advantage of the power of one of these techniques: single-molecule manipulation with optical tweezers, to obtain kinetic information regarding various molecular processes. Particularly, I focused my efforts to study the operation of a multimeric ring-shaped ATPase that encapsidates the genome of the bacteriophage ϕ29, and to determine energy dissipation in non-equilibrium DNA hairpin unfolding/refolding experiments. The results that I present here illustrate the type of information that can be obtained using single-molecule manipulation in the study of biomolecular processes.

1.2   Ring ATPases

A variety of essential biological processes, such as ATP production, protein degradation, membrane remodeling and cell division, are carried out by proteins belonging to the $\underline{A}$dditional $\underline{S}$trand $\underline{C}$onserved Glutamat$\underline{E}$ (ASCE) clade (Hanson & Whiteheart 2005). To perform their biological task, these proteins transform the energy released in the hydrolysis of nucleotide triphosphates (NTPs) into mechanical work (Figure 1.1A and 1.1B). Members of

the ASCE clade share a common fold that includes two conserved NTP binding motifs, known as Walker A (W-A, G/A-xxxx-GK-T/S, where x is any residue) and Walker B (W-B, zzzzDE, where z is any hydrophobic residue) motifs, an additional β-strand that lies in between the two binding motifs, and the presence of a conserved acidic residue that lies in the W-B motif (Erzberger & Berger 2006) (Figure 1.1C).
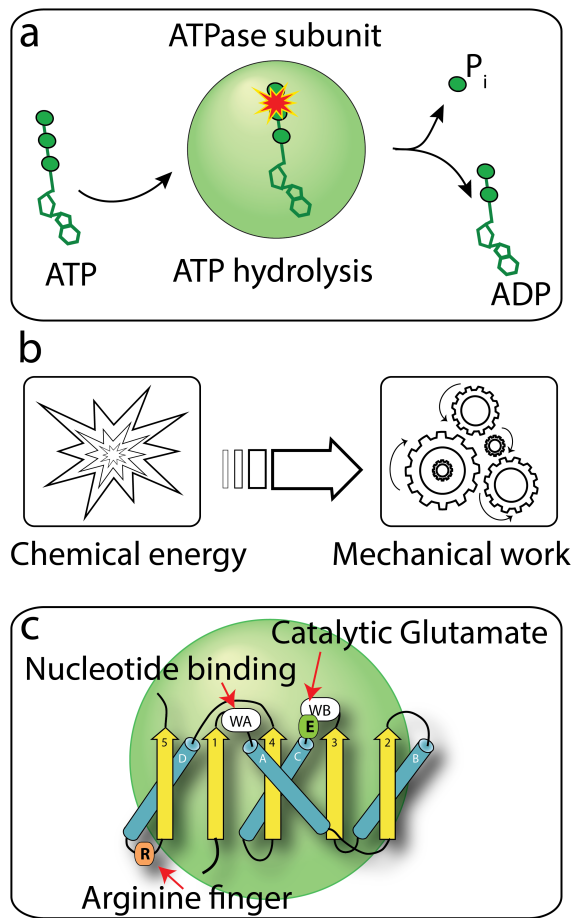

Figure 1.1: **ASCE proteins share a common catalytic fold.**
a) Schematic representation of an ASCE protein binding ATP, hydrolyzing the nucleotide and releasing the products of the reaction: ADP and P_i.
b) Many ASCE proteins transform chemical energy into mechanical work.
c) Schematic of the super-secondary structure of ASCE protein's catalytic core.

The clade divides into two superfamilies (Erzberger & Berger 2006): the RecA-like family, which includes close structural relatives of RecA, a protein involved in homologous recombination, and the family of <u>A</u>TPases with <u>A</u>dditional cellular <u>A</u>ctivities (AAA+), which includes proteins involved in DNA strand separation, such as SV40 and E1 helicases, and protein degradation such as ClpX, among many others. Most ASCE NTPases utilize ATP and are active only as oligomeric assemblies in which pairs of adjacent subunits form the catalytic pockets. The oligomeric assemblies form ring-shaped structures (referred later as ring ATPases or ring motors) which are typically pentamers, hexamers or heptamers, although tetramers have been reported as well (Maluf et al. 2006). The orientation and distance of the conserved ASCE fold relative to the central axis of the ring-shaped structure varies from one protein type to another. However, the spatial position of the nucleotide-interacting residues is highly conserved among all members of the clade. The remarkable functional diversity across all members of the ASCE clade can be rationalized by the presence of a variety of factors and accessory domains associated to each protein type. Moreover, the multiple subunits are known to coordinate their enzymatic activity in schemes that are thought to have arisen to optimize the biological function of the oligomeric protein (Abbondanzieri & Zhuang 2009).

The scheme of coordination in a ring ATPase refers more specifically to the different nucleotide states of the individual subunits across the chemical cycle of the oligomeric complex (Figure 1.2). Some schemes of coordination previously reported are: *i) Concerted*, in which all the subunits simultaneously bind ATP, hydrolyze ATP and release ADP, i.e. at any given time the subunits mirror each other, *ii) Sequential*, in which the different nucleotide states, ATP-bound, ADP-bound and Apo (empty pocket) states are followed by each other spatially and temporarily, *iii) Semi-stochastic*, in which there is little or no apparent correlation between the nucleotide state of one subunit and its neighbor, and *iv) Bi-phasic*, in which a sequential scheme of coordination is partially combined with a concerted scheme, such that the nucleotide states are segregated in two phases. In the first phase, all subunits are in the ADP-bound subunit. The subunits turn to the ATP bound state, one by one and in a sequential manner until all of them are ATP-bound. Then, in the second phase, the ATP-bound subunits turn to the ADP-bound state sequentially by hydrolyzing ATP and releasing phosphate, one subunit

4

at a time and in a sequential manner.
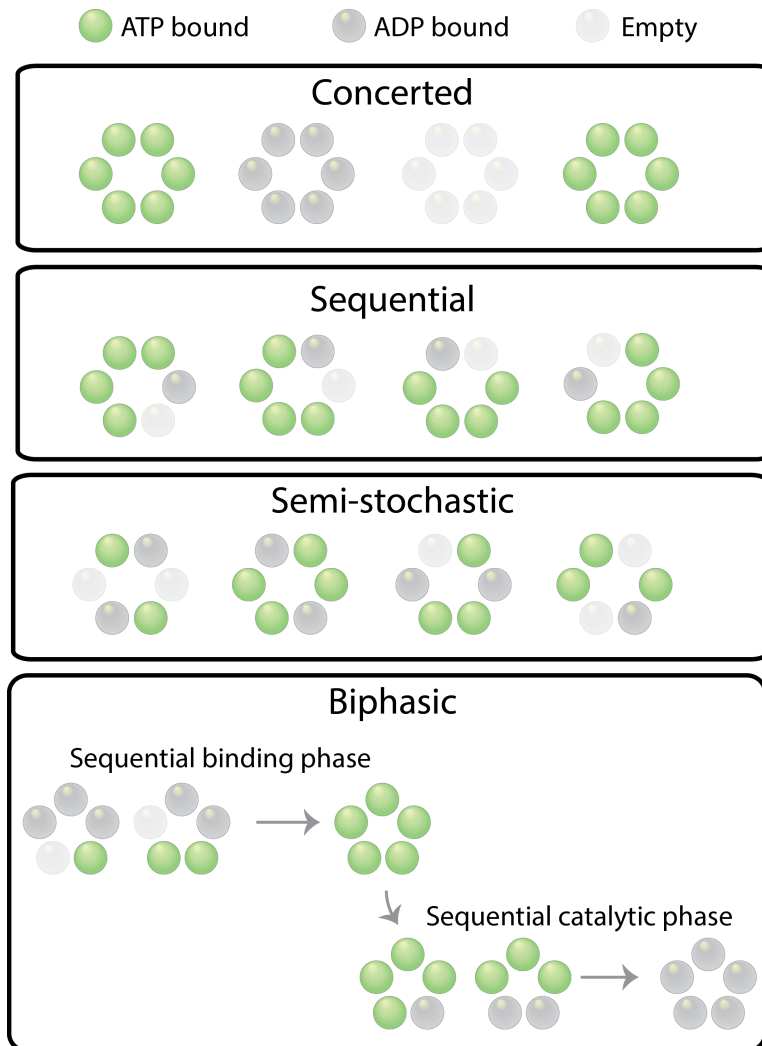
Figure 1.2: **Coordination schemes in ring ATPases.**

As evidenced by the schemes of coordination just described (see also Figure 1.2), the ATPase subunits must rely on a mechanism to communicate their nucleotide state to their neighbors, and therefore, they require structural elements that mediate such process at the molecular level. Structural studies of ring ATPases have provided detailed information regarding the molecular interactions at the interface between two adjacent subunits. From

these structural studies, it is known that the binding pocket of a given sub-unit is completed by residues contributed by their neighbor. Among those contributed residues, a highly-conserved arginine, commonly known as the *arginine finger*, is the best candidate to mediate inter-subunit communication (Ogura et al. 2004). Different roles have been assigned to the arginine finger in various biological systems. Biochemical studies have shown that the highly conserved arginine is a *trans-acting* residue that is necessary for ATP hydrolysis. Structural studies show that in many biological systems the guanidinium group forms hydrogen bonds with the γ and β phosphates of the nucleotide (distances typically $\sim$ 2.6 - 3.1 Å) in the presence of ATP or nucleotide analogs mimicking the hydrolysis transition state (Wendler et al. 2012). Thus, the hydrogen bonds established by the arginine are thought to polarize the negative charge of the phosphate groups, and facilitate the nucleophilic attack required for the hydrolysis reaction. Thus, these observations support a trans-acting role of the arginine to stabilize the hydrolysis transition state. However, roles other than catalysis have also been suggested for the arginine finger.

Structural studies have also shown that in the presence of ADP, the arginine finger often interacts with other residues far away from the nucleotide (distances from the nearest phosphate $\sim$ 4.7 to 9.7 Å) (Wendler et al. 2012). In those cases, a superposition of the ATP-bound and ADP-bound states of the nucleotide binding pocket, reveals that the change in nucleotide state is also accompanied by significant conformational changes that propagate to distant elements in the two adjacent subunits. Such observations suggest that the arginine finger might be part of an allosteric sensor that couples a change in nucleotide state to a change in conformation (Chen et al. 2010). Additionally, mutagenesis studies have shown that in some biological systems arginine finger mutant subunits have reduced nucleotide affinity or cannot bind the nucleotide. In some extreme cases, these mutant subunits cannot form oligomers (Wendler et al. 2012). Based on those observations it has been proposed that the role of the arginine finger is to modulate nucleotide affinity and to promote the subunits' oligomerization.

Although these studies have provided valuable information regarding the role of the arginine finger in *local* communication between two adjacent sub-units, how *global* coordination around the ring is attained, and how different schemes of coordination arise from elementary stochastic processes remains unknown. A more comprehensive understanding requires information regard-

ing the transition between nucleotide states, and whether the arginine finger plays different roles at different stages of the motor's cycle. Given that the dynamics of the protein would be averaged using ensemble methods, this information is not accessible through bulk studies. Molecular dynamic simulations are a powerful tool for studying the protein dynamics. For instance, a simulation of ATP unbinding from F1-ATP synthase provides a temporal map of the residues interacting with the nucleotide as it leaves the binding pocket (Antes et al. 2003). Interestingly, the arginine finger forms one of the strongest and lasting (last-to-break) hydrogen bonds with the γ phosphate of ATP. The reverse order of those molecular interactions could provide information regarding the process of ATP binding in F1-ATPase, the ring ATPase that catalyzes the reverse reaction. In that case, the map of interactions from the dynamic simulation would suggest that, during ATP binding, one of the earliest and strongest hydrogen bonds is established between the arginine finger and the γ phosphate of ATP. However, even if those conclusions turn out to be sensitive, we are still far from being able to understand *global* inter-subunit coordination using molecular dynamic simulations; the current computational power only enables simulations of processes taking place in the μs timescale in small proteins (e.g. 15 kDa), or fast processes (e.g. 15 ns) for large molecular complexes (few MDa). This is still far from the required computational power to recapitulate the coordination of all the subunits in the ring ATPase (few MDa) throughout the chemical cycle (in the order of 100 ms).

Optical tweezers are a powerful tool to manipulate single molecules and to track their mechanical operation with high spatial and temporal resolution (Bustamante et al. 2014). The experimental setup allows to naturally follow the trajectory of proteins as they displace over their substrate, as is the case of motors that participate in intracellular transport (e.g. Kinesin and Dynein proteins), or as they translocate their substrate, as is the case of DNA and peptide translocases (e. g. SpoIIIE and ClpXP). In particular, the DNA packaging motor of the bacteriophage ϕ29 is an extremely well-suited system to investigate the operation and inter-subunit coordination of ring ATPases using optical tweezers for various reasons. First, being a component of one of the smallest bacteriophages ever reported, the number of components necessary to have a fully reconstituted system *in vitro* is minimal compared to other ring ATPases from bacteria and eukaryotic cells that require additional factors or the interaction with various proteins in the cell (Anderson et al.

1966). Second, unlike other viral ring motors, the packaging trajectories of the ϕ29 DNA packaging motor display enough signal-to-noise ratio that we can easily resolve individual motor cycles, each trajectory containing thousands of those events (Liu et al. 2014). Finally, the DNA packaging motor is among the most coordinated ring ATPases reported to date, displaying multiple layers of coordination and regulation (Chistol et al. 2012). Since the motor's mechano-chemical cycle has been extensively characterized, and all the chemical transitions have been located within the cycle, it provides a particularly well-suited experimental platform to dissect the physical basis that enable precise subunit coordination.

1.3   Viral Packaging Motors

The ϕ29 bacteriophage has a double-stranded DNA (dsDNA) genome. Double-stranded DNA bacteriophages and some human pathogens, such as herpes virus and adenovirus, encapsidate one copy of their viral genome into a pre-formed protein shell during replication (Casjens 2011). This process of genome encapsidation has a high energy cost, which includes the associated energy to bend dsDNA, which is a relatively stiff polymer, the entropic penalty resulting from polymer confinement, and the energy necessary to overcome the large electrostatic repulsive forces that result from confining the negative charges of the DNA backbone. In the case of the bacteriophage ϕ29, its 19.3 kb genome ($\sim$ 6.3 μM) must be confined to near crystalline density inside a small protein capsid whose dimensions—54 nm $\times$ 45 nm (Anderson et al. 1966)—are similar to the reported persistence length of the dsDNA ($\sim$ 50 nm). Consequently, genome encapsidation is carried out by one of the most powerful nano-machines reported to date, capable of exerting large forces ($>$ 60 pN) to overcome the high internal pressures that build as the capsid fills up.

The DNA packaging motor of the bacteriophage ϕ29 has three components (Morais 2012): the connector, the prohead RNA (pRNA) and the ring ATPase (Figure 1.3). The connector is a dodecameric complex of gp10 (gene product 10) that sits at the portal vertex of the icosahedral prolate (elongated) viral capsid. The DNA is thread through the pore of the connector, a channel that has multiple loops exposed containing positively charged residues, which are thought to interact with the DNA phosphate backbone.
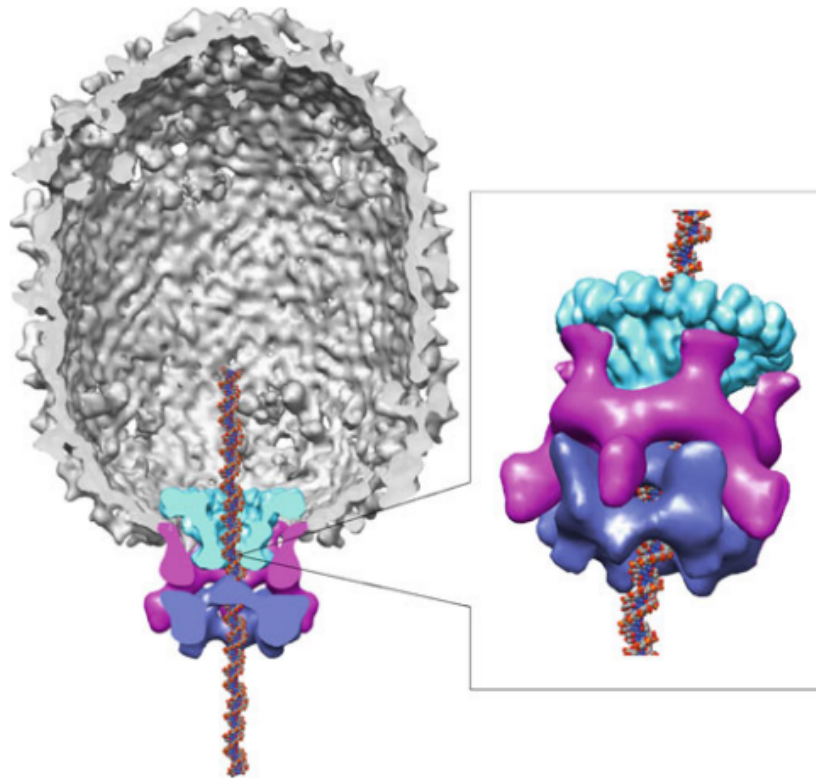
Figure 1.3: **Three-dimensional reconstruction of the ϕ29 DNA packaging motor.** The particle is shown in a cutaway view to facilitate visualization. The capsid is shown in grey, the connector in green, the pRNA in magenta, and the ATPase in blue. The dsDNA was shown as spheres. Adapted from (Morais 2012).

The connector has been shown to work as a one-way valve that allows movement of the DNA towards the inside of the capsid but not towards the outside. The mechanism by which this process takes place is not well understood. However, a recent cryo-EM reconstruction of the mature virion shows that once the DNA is packaged, the highly-pressurized head is sealed by a toroidal (donut) shape structure made of DNA and located at the exit of the connector pore, thus, clogging the exit. The pRNA is a ring-shaped RNA structure formed via intermolecular base-pairing interactions by 5 identical copies of a 174 nt transcript (Harjes et al. 2012). The pRNA assembles onto the capsid by interacting specifically with the five-fold symmetric capsid, and protrudes away 5 arms or superhelices that serve as a scaffold to bind the

ATPase subunits and facilitate their assembly into another ring structure. A 120 nt transcript lacking the 54 nt 3' domain II is fully competent for DNA packaging and phage assembly, but lacks directionality to package the left hand of the genome first, as it occurs with the full 174 nt transcript. The ring ATPase sits just below the portal vertex of the capsid and feeds the DNA into the connector channel. It is comprised of five identical *gp*16 (gene product 16) subunits. Out of the three components of the motor, this is the only one with ATPase activity.

Unlike the connector and the ring ATPase, which are conserved across all DNA packaging motors, the pRNA is found only in ϕ29-like bacteriophages, such as ϕ15, M2, B103, BS32, Nf and PZA, which infect *bacilus subtilis* and related species. Other bacteriophages, such as T4 and λ (which infect *E. coli*), have evolved larger ATPase subunits that comprise the large terminase, a ring-shaped structure that is thought to have incorporated together the role of the ring ATPase and the pRNA. The large terminase subunits also have nuclease activity to cut the concatemeric viral DNA after one full copy of the genome has been encapsidated in the phage head. The viral DNA of ϕ29-like phages is replicated into individual genome units, and their ATPase subunits do not possess nuclease activity. The packaging rates of T4 and λ ring motors are faster relative to ϕ29 (770 ± 300 bp/s and 590 ± 118 bp/s vs 145 ± 15 bp/s, respectively) (Chemla & Smith 2012). These rates seem to scale to finalize genome packaging (genome sizes are 171 kb and 48.5 kb for T4 and λ, respectively) within a limited window of time $\sim$ 2-5 minutes and complete the infection cycle in 20-30 minutes. Because of the fast packaging rates of T4 and λ ring motors, with the current technologies it is not possible to resolve the individual translocation events by the packaging motors of these phages. In contrast, at the relatively slow packaging rate of the ϕ29 packaging motor, multiple translocation events can be observed using high-resolution optical tweezers.

1.4   The ϕ29 DNA Packaging Motor

In early single-molecule studies of the ϕ29 DNA packaging motor (Moffitt et al. 2009), it was found that this molecular machine operates using a two-

phase cycle: a "dwell" phase during which the DNA remains stationary, followed by a "burst" phase during which 10 base-pairs (bp) of DNA are rapidly translocated into the head (Figure 1.4A). This observation immediately suggested a model in which each burst is composed of five 2 bp translocation sub-steps, one sub-step per subunit. This mechanism would be consistent with the number of base pairs (two) packaged per hydrolyzed ATP derived from bulk experiments. Surprisingly, packaging trajectories obtained subsequently with high-resolution optical tweezers, revealed that the "burst" was comprised actually of four 2.5 bp sub-steps (Figure 1.4B). This observation indicated that one of the subunit does not have a mechanical role, a result completely inaccessible through ensemble measurements. In this same study, the authors showed that decreasing the concentration of ATP lengthened the dwells, indicating that nucleotide binding takes place during this phase of the cycle. It had been previously shown (Chemla et al. 2005) that the only chemical transition capable of supporting the work done by the motor near its stall force (~ 57 pN) is the release of inorganic phosphate after ATP hydrolysis.

In a later study (Chistol et al. 2012), the authors used non-hydrolyzable nucleotide analogs (ATPγS and AMP-PNP) to stop the motor at specific chemical transitions of the catalytic cycle. They were able to show that ATP binding and ADP release occur in an interlaced manner during the dwell phase, while ATP hydrolysis and Pi release occur in the burst phase (Figure 1.4D). Moreover, ATP hydrolysis and Pi release are also interlaced, with the release of inorganic phosphate coupled to DNA movement (the power stroke). Furthermore, this study provided insight into the role of the non-translocating subunit. Experiments in which one subunit was temporarily inactivated by binding to the non-hydrolyzable analog, ATPγS, showed that the non-translocating subunit must hydrolyze ATP to fulfill a critical regulatory role in the dwell-burst cycle (Figure 1.4D). Thus, this study resolved the apparent contradiction between single molecule and bulk studies on the number of base pairs packaged per hydrolyzed ATP: four are used to performed a mechanical task and one to perform a regulatory one.

At this point, however, the origin of symmetry breaking among motor subunits remained unknown. In a previous single-molecule study, the packaging motor was challenged to translocate a variety of modified DNA substrates
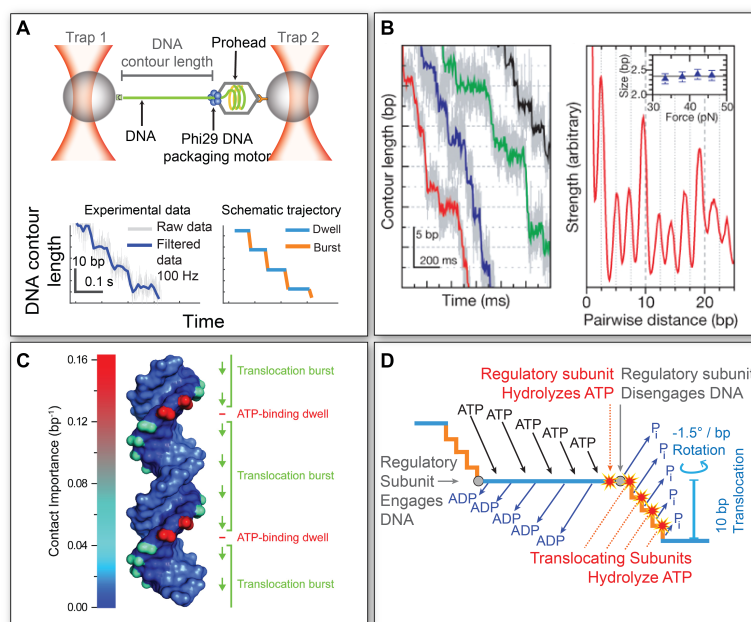
Figure 1.4: **Results from previous single-molecule studies of the ϕ29 DNA packaging motor.**
a) (Top) Single-molecule experimental geometry and (Bottom) experimental data and schematic representation.
b) (Left) Representative packaging traces collected with external loads of 40 pN at saturating [ATP]. Data in light grey are plotted at 1.25 kHz whereas data in color are boxcar-filtered and decimated to 100 Hz. (Right) Average pairwise distribution of packaging traces. Adapted from (Moffitt et al. 2009).
c) Relative importance of the motor-DNA contacts. Units reflect the distance at which the transversal probability of the packaging motor in a given DNA construct drops to 50%. Adapted from (Aathavan et al. 2009).
d) Model of the mechano-chemical cycle indicating the location of all chemical transitions in the regulatory and translocating subunits, as well as DNA rotation. Adapted from (Chistol et al. 2012).

(Aathavan et al. 2009). These experiments helped to determine the nature of the motor-DNA interactions throughout the mechano-chemical cycle (Figure 1.4C). During packaging, the motor makes specific electrostatic contacts every 10 bp with a pair or phosphates in the DNA backbone during the dwell. In contrast, the motor makes mostly nonspecific, steric contacts to propel the DNA during the four power strokes of the burst phase. In view of these results, Chistol et al. proposed in their study a model in which the subunit

that makes these specific contacts is the non-translocating subunit and that the binding to the DNA phosphate is the event that confers it its regulatory role (Figure 1.4D). However, a question that remained unanswered is: how does the motor deal with the difference between the 10-bp motor burst size and the 10.4-bp DNA helical pitch?

In a later study (Liu et al. 2014b), the authors developed a 'rotor bead' assay capable of measuring the three-dimensional trajectory of DNA during packaging. These experiments provided the first experimental demonstration that the motor rotates the DNA as it packages it. To answer whether rotation is dependent on DNA organization inside the capsid, the authors used "trepanated" proheads (wherein the DNA does not accumulate due to the perforation of the proheads). Significantly, in experiments performed with "trepanated" proheads, the DNA was still observed to rotate by about -1.5°/bp, indicating that rotation is intrinsic to the motor mechanism. This figure accounts for the mismatch between the 10 bp burst and the 10.4 bp DNA helical pitch, as it is the exact amount of rotation ($\sim$ 15 degrees in 10 bp) necessary for the same subunit to contact the DNA phosphates ever 10 bps and to retain its regulatory role over consecutive cycles (Figure 1.4D). Furthermore, it was long known that packaging slows down as the head fills with DNA as the internal pressure opposing the motor increases. However, the mechanism by which head filling regulates motor velocity was unclear. Therefore, Liu et al. also examined ϕ29 packaging dynamics at different stages of head filling and found that multiple aspects of the motor's mechanochemical cycle are altered in the process (see also Chapter 2). First, the motor displays long-lived pauses. Second, the motor burst size is reduced from 10 bp to 9 bp, yet still taking four steps per burst. Remarkably, the decrease in the burst size observed with head filling is accompanied by a corresponding increase in the magnitude of DNA rotation in the precise amount needed for the regulatory subunit to engage the DNA phosphate and retain its identity at the end of each burst. In other words, DNA rotation changes to preserve the coordination between the regulatory and translocating subunits.

The results described above evidence the importance of the motor's inter-subunit coordination to complete its biological task, and invite an obvious question: how does such coordination emerge from stochastic, elementary molecular interactions? During my PhD, I performed a series of single-

molecule experiments to identify the structural elements and interactions that enable this coordination. In particular, I studied extensively the highly-conserved trans-acting arginine residue that is thought to mediate communication between two adjacent subunits in other ring ATPases. The goal of those experiments was to understand the role of this trans-acting residue in establishing *global* inter-subunit coordination from *local* interactions (see Chapter 3).

1.5 Non-equilibrium Processes in Biology

Genome encapsidation by the ϕ29 DNA packaging motor is an example of the multiple non-equilibrium processes occurring in the cell. In fact, to maintain their organization, biological systems must operate far from equilibrium (Schrödinger 1992), continuously utilizing and dissipating energy. Presumably, however, they must also avoid the unnecessary waste of energy. Recent studies have revealed that molecular machines, e.g. F1 ATPase, several Kinesin and Myosin systems (Yasuda et al. 1998, Howard 2001), etc., are extremely efficient thermodynamically when compared to their macroscopic counterparts. There are also tantalizing hints of machines displaying non-equilibrium energetic efficiency (Liu et al. 2014b, Berndsen et al. 2014). However, the principles that govern non-equilibrium efficiency are not well understood, and there exists today no generally accepted metrics to quantify efficiency of systems operating far from equilibrium (Brown & Sivak 2017).

Reversible heat engines operating according to the Carnot cycle do not dissipate energy; their energetic efficiency is limited only by the entropy increase of the surroundings associated with the transfer of heat from a hot to a cold reservoir. The extra nonequilibrium energy cost associated with a process carried out at a finite rate further reduces its efficiency (Callen 2014). This is the case for biological machines that must operate under the time constraint of the cell cycle. For instance, as I discussed before, the viral genome of dsDNA bacteriophages must be encapsidated within 2-5 min. Similarly, the 4.6×10^6 base pairs of DNA in *E. coli* must be replicated in about 40 minutes (Alberts 2002), while during sporulation in *B. subtilis*, the DNA translocase SpoIIIE, localized at the asymmetric septum between mother cell and pre-spore, must transfer between them two thirds of the 4.2×10^6

base pairs of DNA in only 15 minutes (Grainge 2008).

Finite-time operation by these machines involves energy dissipation and extra work. The parameter that governs the near-equilibrium dissipation during a finite rate process is the friction coefficient. Recently, an approximate framework to quantify this (in general intractable) thermodynamic quantity (Sivak & Crooks 2012) has been proposed. These ideas have been explored theoretically and numerically (Zulkowski et al. 2012, 2013, Zulkowski & DeWeese 2014, Bonanca & Deffner 2014, Zulkowski & DeWeese 2015b,a, Rotskoff & Crooks 2015, Sivak & Crooks 2016, Rotskoff et al. 2017), but experimental tests were still missing. During my PhD, I performed the first test of these theoretical predictions using the mechanical unfolding and refolding of single DNA hairpins (see Chapter 5).

Chapter 2

High Internal Pressure in the Capsid of Bacteriophage $\phi 29$

2.1 Summary

During assembly of double-stranded DNA bacteriophages, the viral genome must be encapsidated by a DNA packaging motor. High internal pressure builds up inside the viral capsid as a result of entropic and electrostatic repulsive forces. Previous single-molecule studies have determined the value of the internal force resisting DNA confinement to be as high as 110 pN towards the end of genome encapsidation. However, previous theoretical estimations could not account for such high internal pressure. In this chapter, I revisit this value using single-molecule experiments. Using higher resolution data than in previous studies, it is shown that the internal pressure reaches $\sim 20 \pm 7$ pN at 100% capsid filling, which is in better agreement with theoretical models.

2.2 Background

The life-cycle of double-stranded DNA bacteriophages depends on a remarkable mechanical reaction: infection of the host bacterium begins with the injection of a fraction of the viral genome inside the host cell. The energy required for this physical process—*genome ejection*, from the perspective of

the phage—is stored as potential energy in the form of internal pressure built inside the viral prohead. Such potential energy is accumulated at the time of phage assembly by the ATP-powered DNA packaging motor during the process of genome encapsidation. Indeed, in early single-molecule studies of the $\phi 29$ DNA packaging motor, it was shown that the rate of packaging gradually slows down as the capsid fills with DNA (Figure 2.1A) (Smith et al. 2001), but the mechanism responsible for this phenomenon was not understood. The reduction in packaging velocity was explained by a model in which an internal pressure builds up inside the capsid, slowing down and eventually stalling the motor toward the end of packaging (Smith et al. 2001, Rickgauer et al. 2008). The sharp increase in internal force with filling during the latter stages of packaging was in qualitative accord with many subsequent theoretical calculations. In those calculations the free energy was obtained as the sum of the bending energy and an empirically determined inter-strand interaction energy that accounts for electrostatic self-repulsion and hydration effects; then, the energy was minimized by balancing those two terms. The predictions of these theoretical calculations, which had been estimated for the specific case of $\phi 29$ (21), were in qualitative agreement with the experimental findings. Nevertheless, none of the existent models predicted forces as high as those *measured* in the early single-molecule studies (up to 110 pN) (Smith et al. 2001, Rickgauer et al. 2008).

As I will discuss in this chapter, the early *measurements* of internal pressure relied on the untested assumption that the effect of DNA confinement on the motor dynamics is equivalent to that of an opposing force applied externally with optical tweezers. Nevertheless, such estimations neglected whether other aspects of motor operation that are not related to force generation, such as nucleotide binding and inter-subunit communication, are also modified by the increasing capsid filling, producing a slowing down of the packaging rate. In this chapter, I present the experiments that my colleagues and I performed (Liu et al. 2014b) to reassess the internal pressure built during genome encapsidation in the $\phi 29$ bacteriophage.

2.3 Experimental results

Revisiting High-Internal Pressure in the Bacteriophage $\phi 29$

To understand what causes the previously reported decrease in packaging velocity towards the end of genome encapsidation (Smith et al. 2001), we used dual-trap optical tweezers (see Figure 1.4A) to monitor DNA translocation at various filling levels with base pair-scale resolution (see Appendix A for a detailed description of sample preparation and experimental procedures). We found that the biphasic dwell-burst structure is retained at high filling, i.e., ATP hydrolysis events remain segregated from nucleotide-binding events throughout the entire packaging process (Figure 2.1B). However, the amount of time that the motor spends on each packaging cycle increases significantly, with the median cycle time rising from 0.15 s at 40% filling to 0.66 s at 100% filling (Liu et al. 2014b). These results indicate that the changes in the $\phi 29$ motor's operation at high filling are not caused by disruptions of the overall motor coordination. Instead, they are caused by modulations of certain kinetic transitions within the same mechano-chemical framework, resulting in lengthening of the dwell phase, emergence of long-lived pauses, lengthening of the burst phase, reduction in the burst size, and increased frequency of slipping (Liu et al. 2014b). Among these modifications, the burst phase of the cycle—when DNA translocation occurs and the force generating steps occur—has only a minor effect on the overall velocity (Figure 2.1B).

It was previously proposed that the reduced packaging velocity at high capsid filling was caused by an internal force resisting packaging due to DNA confinement, F_{int}, whose effect was equivalent to the slowing down of the motor by an opposing force applied externally, F_{ext} (Smith et al. 2001, Rickgauer et al. 2008). Because of the assumed equivalence between F_{int} and F_{ext}, it was thought that the magnitude of the former could be inferred from the observed velocity's dependence on capsid filling using the F_{ext}-velocity calibration. In this way, it was estimated that F_{int} reaches 60 pN at 100 % capsid filling (Smith et al. 2001), a value later revised to 110 pN (Rickgauer et al. 2008). The capability of resolving individual dwell-burst cycles enabled us to dissect the effect of external force and that of capsid filling on motor operation in greater detail. We found that external force increases
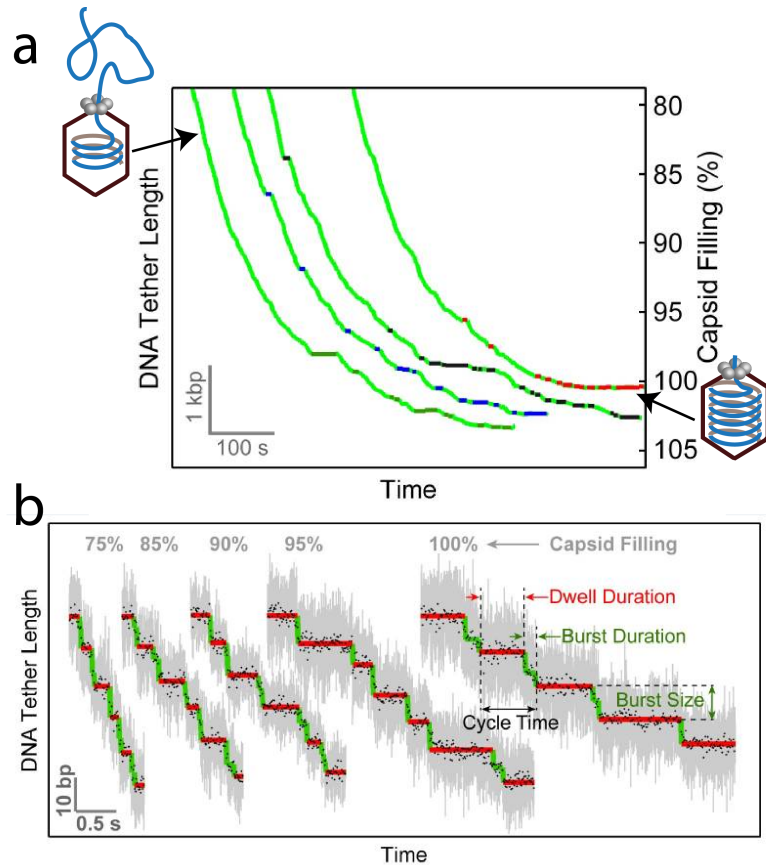
Figure 2.1: **Sample packaging trajectories of the ϕ29 DNA packaging motor at high internal pressure.** Adapted from (Liu et al. 2014b).
a) Sample traces displaying individual packaging cycles at various levels of capsid filling (schematics on the side indicate low filling—top left—and high filling—bottom right). Raw 2,500 Hz data are shown in gray, and down-sampled 100 Hz data are in black. Stepwise fit to the data highlights dwells and bursts in red and green, respectively.
b) Sample traces highlighting the locations of long-lived pauses (LLPs). For clarity, only LLPs longer than 5 s are shown. LLPs from different packaging complexes are colored green, blue, black, and red.

only the duration of the DNA-translocating burst phase but has little effect on the duration of the stationary dwell phase (Figures 2.2A and 2.2B). In contrast, capsid filling prolongs both dwell and burst durations (Figures 2.1B and 2.2C).
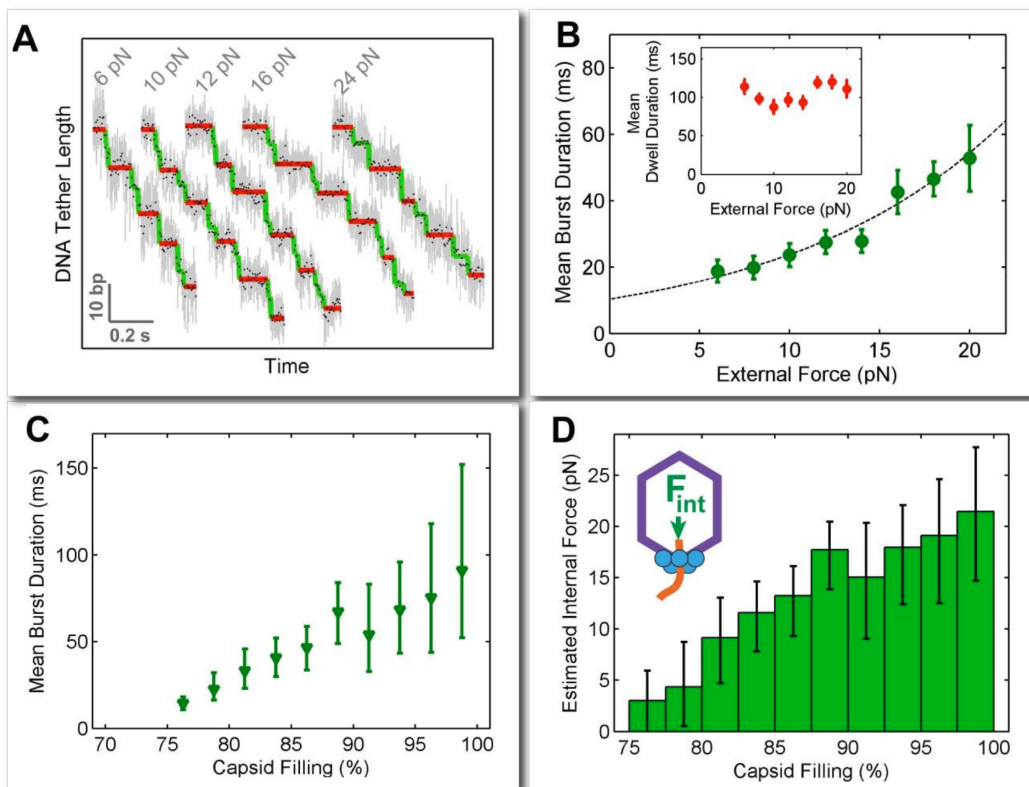
Figure 2.2: **Estimation of Internal Force Based on Burst Duration as a Function of Capsid Filling.** Adapted from (Liu et al. 2014b).
a) Sample packaging traces collected at low capsid filling (15% - 30%) and various external loads. Raw 2,500 Hz data are shown in gray, and down-sampled 100 Hz data are in black. Stepwise fit to the data highlights dwells and bursts in red and green, respectively.
b) Mean burst duration versus external force at low capsid filling (15% - 30%). The data are fit to an Arrhenius-type equation: $\tau_{burst}(F) = \tau_{burst}(0) \times e^{F\Delta x/kT}$ (dashed curve, $\Delta x = 0.33 \pm 0.08$ nm from the fit). Inset shows mean dwell duration versus external force.
c) Mean burst duration versus capsid filling.
d) The magnitude of the internal force as a function of capsid filling, obtained by applying the τ_{burst} - Force curve **(b)** to the τ_{burst} - filling dependence **(c)**.

To revise the F_{int} estimate, we assumed that the burst duration depends only on the opposing force acting on the motor through DNA, either externally or internally. We can then use the F_{ext}-burst duration calibration (Figure 2.2B) to estimate the magnitude of F_{int} responsible for lengthening the burst during capsid filling (Figure 2.2C). This revised analysis yields an internal force of only 20 ± 7 pN at 100% filling (Figure 2.2D). Notably, this value is three to

21

five times lower than those calculated from prior single-molecule experiments and is in better agreement with predictions from analytical modeling and numerical simulation studies (Kindt et al. 2001, Purohit et al. 2005).

Chapter 3

Mechanisms of Inter-subunit Coordination in a Viral Ring Motor

3.1 Summary

Subunits in multimeric ring-shaped motors must coordinate their activities to ensure correct and efficient performance of their mechanical tasks. Here, we study wild-type and arginine finger mutants of the pentameric bacteriophage ϕ29 DNA packaging motor. Our results reveal the molecular interactions required to coordinate ADP-ATP exchange during the dwell phase and ATP hydrolysis during the burst phase of the motor's mechano-chemical cycle. We show that two distinct regulatory mechanisms determine this coordination. In the first, an arginine finger in each subunit promotes ADP-ATP exchange in—and stimulates ATP hydrolysis of—its neighbor. Accordingly, we suggest that the subunits perform the roles described for GDP exchange factors and GTPase-activating proteins observed in small GTPases. In the second mechanism, the DNA up-regulates a single subunit, transforming it into a global regulator that initiates the dwell and burst phases. We propose that these mechanisms are fundamental to inter-subunit coordination and are likely present in other ring ATPases.

3.2   Background

Many biological processes such as protein degradation, ATP synthesis, DNA strand separation, and chromosomal segregation, require the mechanical operation of the Additional Strand Conserved Glutamate (ASCE) division of ring-shaped Nucleotide Triphosphatases (NTPases) (Erzberger & Berger 2006). These ring motors perform specific mechanical tasks, such as exerting large forces in the case of DNA packaging viral motors (Smith et al. 2001) and translocating DNA at high speeds during cell division by proteins such as FtsK and SpoIIIE (Liu et al. 2015). Various schemes of global inter-subunit coordination in ring NTPases have been reported and are thought to have arisen to optimize the motor's mechanical performance (Liu et al. 2014).

As discussed in chapter 1, ATPase subunits of the ASCE protein division rely on highly-conserved structural elements to bind nucleotide (Walker A motif), catalyze hydrolysis (Walker B motif) and coordinate the enzymatic activity at the interface of two adjacent subunits (trans-acting arginine finger) (Wendler et al. 2012). The arginine finger is known to mediate communication between adjacent subunits (Ogura et al., 2004), the mechanism underlying the global coordination of the ring is still not understood. Moreover, given that all ring ATPases appear to rely on the same structural elements for inter-subunit communication, it is of interest to understand how different coordination schemes arise from these conserved features.

The bacteriophage ϕ29 DNA packaging motor is a pentameric ring ATPase whose mechano-chemical cycle has been comprehensively dissected through single-molecule studies (see section *The ϕ29 DNA packaging motor* in chapter 1). This motor's subunits have been shown to operate in a highly coordinated manner, making it an ideal system to investigate how global motor coordination can arise from elementary molecular interactions. Its mechano-chemical cycle is comprised of a dwell phase in which the motor exchanges ADP for ATP in an interlaced manner, and a burst phase in which nucleotide hydrolysis coupled to DNA translocation occur sequentially one subunit at a time. However, experimental evidence for the underlying molecular interactions that enable this coordination is lacking. Moreover, the current data indicate that the pentameric ring displays symmetry breaking during its operation: only 4 of the 5 identical subunits are seen to translocate DNA by 2.5

base pairs (bp) each, resulting in a 10 bp burst (Moffitt et al. 2009), while the remaining subunit is thought to play a regulatory role (Figure 3.4B). It has been proposed that this role is adopted by the subunit that presumably contacts two consecutive DNA phosphates every 10 bp during packaging (Aathavan et al. 2009, Chistol et al. 2012). However, experimental evidence further supporting this model is still required. The regulatory subunit is thought to initiate the beginning of the dwell and burst phases cycle after cycle (Liu et al. 2014b), but the mechanism by which the subunit making the periodic contact with the DNA regulates the motor's cycle remains unknown.

In this chapter, I present the experiments that I performed to elucidate how the 'clockwork' operation of the $\phi29$ DNA packaging motor arises from a series of otherwise stochastic molecular processes. To this end, I use optical tweezers (Figure 3.4A) and targeted mutagenesis of the conserved and recently identified arginine finger, R146 (Mao et al., 2016). Then, I combined this data with cryo-EM reconstructions from my collaborator, Marc Morais. My colleagues and I show that higher-order coordination of the motor results from two mechanisms of regulation. In the first mechanism, each subunit uses R146 to facilitate ADP-ATP exchange and to stimulate ATP hydrolysis in the neighboring catalytic pocket. In the second mechanism, a single subunit contacts the DNA cycle after cycle. We propose that this contact induces the release of ADP by the bound subunit, initiating the dwell phase. It also transforms the bound subunit into a more efficient enzyme that can hydrolyze ATP faster than the remaining subunits, enabling it to initiate the burst phase. Cryo-EM reconstructions presented here support the proposed mechanism. The merging of these two mechanisms of regulation results in a highly coordinated scheme that allows the motor to spend a large fraction of the time ($\sim 90\%$) bound with high affinity to the DNA (during the dwell), and to be only transiently detached from its substrate while it packages the polymer (during the burst). We propose that such coordination scheme enables the motor to work against the large internal pressure building inside the viral capsid as it fills up with DNA.

3.3 Experimental results

Bulk Screening and Characterization of Arginine Finger Mutants

We generated mutants affecting the arginine finger, R146, to dissect the communication tasks mediated by this trans-acting residue. The packaging activities of homomeric ring motors carrying the substitution R146K or R146A, have been previously characterized by bulk DNase I protection assays (Mao et al. 2016). These studies have shown that both types of homomeric mutant rings are incapable of full-genome encapsidation, showing either minimal ATPase activity (R146K) or no activity (R146A). The DNase I protection assays report only the fraction of motors capable of fully packaging the viral genome inside the capsid. To determine if R146K or R146A homomeric rings retained at least partial DNA packaging activity, we performed single-molecule experiments (Figure 3.4A). In these experiments, packaging is initiated in bulk, and only those ring motors that packaged an initial amount of DNA can be tethered in the single-molecule setup. This approach naturally selects for active ring motors. We did not observe any tethers with homomeric rings of R146K or R146A subunits, indicating that such rings do not package DNA or do not assemble into functional motors.

To obtain active mutant motors in the single-molecule setup, we use two different strategies. First, we determined if rings containing a limited number of mutant subunits display packaging activity. To this end, we mixed wild-type (WT) with either R146K or R146A subunits in various proportions (Figure 3.4C). Second, we designed a substitution of the residue adjacent to the arginine finger, F145I, to moderately perturb R146K's local environment. Homomeric rings of F145I mutants displayed near-WT ATPase and intermediate packaging activities in bulk assays. Following these two approaches, we were able to obtain single-molecule packaging trajectories of hybrid (R146K/WT and R146A/WT) and of homomeric (F145I) mutant rings. These constructs provided the experimental framework to investigate the role of arginine R146 in inter-subunit coordination.
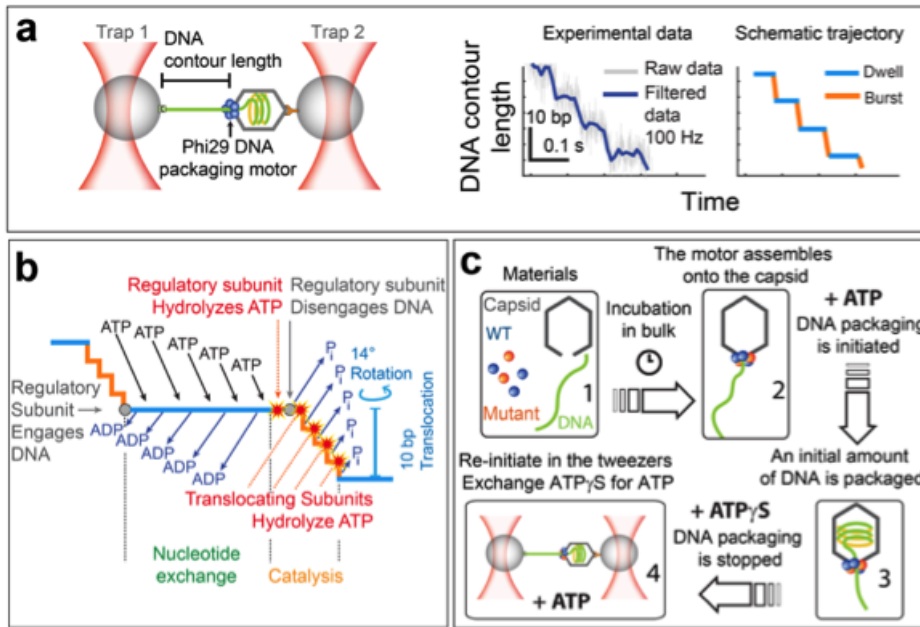
Figure 3.1: **Single Molecule Experiments and Mechano-Chemical Cycle of the ϕ29 DNA Packaging Motor.**
a) Single-molecule experimental geometry (left) and data (right).
b) Model of the mechano-chemical cycle (see main text for description and (Chistol et al. 2012).
c) Experimental methodology of subunit mixtures.

Heteromeric R146K/WT and Homomeric F145I Rings Exhibit Slower DNA Packaging Activity

We first assayed samples generated by mixing R146K and WT subunits in the proportion 1:3 (25% in Figure 3.2A and 3.2B). Motors from these samples displayed a broad distribution of packaging velocities, indicating the formation of different types of active rings comprised of various combinations of WT and mutant subunits (Figure 3.2A). We found that increasing the proportion of R146K subunits in the mixture shifted the packaging velocity distribution monotonically towards lower values (Figure 3.2B), and decreased the tether efficiency, presumably reflecting a reduced number of active motors. We detected packaging up until R146K and WT were mixed in proportions of 2:3 (40% in Figure 3.2B). In contrast, F145I homomeric rings showed a nar-

row and single-peak velocity distribution with a mean velocity ($\sim$65 bp/s) reduced relative to WT ($\sim$100-125 bp/s) (Figure 3.2C).
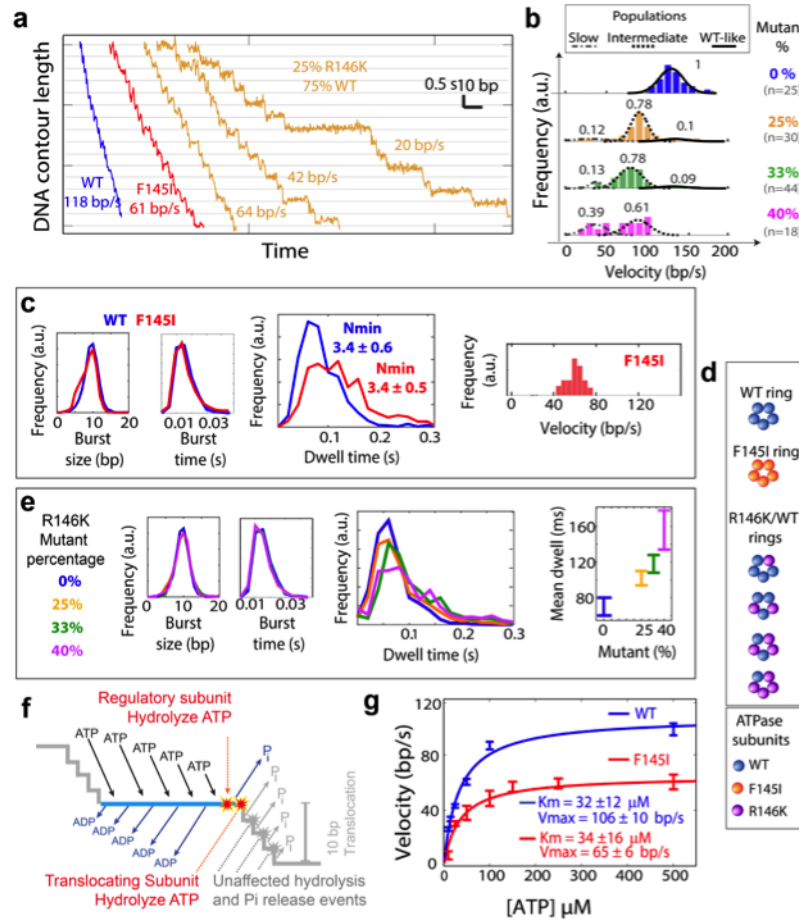

Figure 3.2: **Modifications of the Arginine Finger Slow Down Nucleotide Exchange.**
a) High resolution packaging trajectories from WT, F145I and R146K/WT hybrid rings.
b) Velocity distribution of R146K/WT mixtures for different mixing ratios (see section 3.4 for derivation of fractional populations).
c) High resolution analysis of the dwell and burst phases of F145I rings.
d) Schematic of homomeric and hybrid ring configurations.
e) High resolution analysis of the dwell and burst phases for all mixing conditions of R146K/WT hybrid rings.
f) Chemical transitions during the dwell phase that could be affected by the mutations (highlighted in color).
g) Estimation of Michaelis-Menten parameters of F145I rings by ATP titration.

Packaging velocity of R146K/WT can be clustered in three main groups: fast (WT-like at ~125 bp/s), intermediate (~85 bp/s) and slow (~45 bp/s) (Figure 3.2B). This observation rules out the simplest scenario in which the motor can tolerate only 1 mutant subunit, displaying the same phenotype whether it affects the translocating and the regulatory subunit, for such scenario would result in only two populations (WT and slower motors). Next, we considered the case in which a single mutant subunit is tolerated in the ring but affects differentially the translocating and the regulatory subunits, a scenario that would explain the three peaks observed. In this case, we would expect that as the ratio of mutant-to-WT subunits is increased in the mixture, the heights of the two slower peaks should grow at the expense of the WT peak but their relative proportion should not change. However, the results in Figure 3.2B indicate that the size of the two slower populations change relative to one another, clearly ruling out also this scenario. Therefore, we conclude that the motor can tolerate 2 or more R146K subunits. We used the binomial distribution to estimate the fractions of motors in each group (WT-like, intermediate or slow) that should be observed for all different mixing conditions (1:3, 1:2 and 2:3), for each scenario in which the ring tolerates 1, 2, 3 or 4 mutant subunits. The results of this analysis (see section 3.4 and Table 3.A.1) further support the conclusion that the motor can tolerate 2 or more R146K subunits in the ring without complete loss of activity (Figure 3.2D).

The Arginine Finger Mediates Nucleotide Exchange

We analyzed in detail single-molecule trajectories of R146K and F145I mutants. Since nucleotide exchange occurs only during the dwell and catalysis mainly during the burst (Figure 3.4B, (Chistol et al. 2012), we sought to establish whether the dwell, the burst or both phases are affected by the mutations. First, we compared the burst size and burst duration distributions of F145I homomeric ring motors (Figure 3.2C) and the different R146K/WT mixtures (Figure 3.2E) with those of WT rings. Neither mutant displayed changes in the size or duration of the burst phase relative to WT, indicating that the chemical events—ATP hydrolysis and phosphate release—and coordinated DNA translocation are unaffected or affected below our detection capabilities. The somewhat surprising result that the substitution R146K does not affect ATP hydrolysis may be rationalized by the observation that

29

in other ring ATPases a trans-acting lysine residue plays the role of the arginine finger (Greenleaf et al. 2008). Likewise, although the arginine's environment is modified by the F145I substitution, the subunits in the homomeric ring still retain the trans-acting arginine, R146. Second, we found that the distribution of dwell times of F145I homomeric ring motors is displaced towards larger values relative to WT (Figure 3.2C); the same is observed with increasing R146K/WT mixing ratios (Figure 3.2E). We conclude that both substitutions affect the dwell phase.

The following processes are thought to occur during the dwell: nucleotide exchange, hydrolysis by the regulatory subunit to initiate the burst and hydrolysis by the first translocating subunit (Figure 3.2F). Since the R146K and F145I substitutions do not affect any of the hydrolysis reactions during the burst, and the local chemical environment required for hydrolysis is conserved in these mutant subunits, we consider it improbable that they would affect only the first two hydrolyses events in the dwell. Instead, we assert that the R146K and F145I mutations lengthen the dwell duration because they slow down the rate of nucleotide exchange.

The Arginine Finger Facilitates ADP Release and Modulates Nucleotide Affinity

During the nucleotide exchange cascade in the dwell, all five ADPs generated in the previous cycle are replaced sequentially by new ATP molecules. The lengthening of the dwell times in R146K and F145I mutants indicates that either ADP release, ATP binding, or both are slow down relative to WT. To identify which specific kinetic rates might be affected by the F145I or R146K substitutions, we derived an analytical expression for the packaging velocity (see kinetic scheme i in Figure 3.3E and section 3.4 for derivation). The overall velocity of the motor's cycle can be written as:

$$v = \frac{V_{max} * [ATP]}{[ATP] + K_m} \cdot 10 \text{ bp} \tag{3.1}$$

where

$$V_{max} \approx \frac{k_{ADPr}}{4 \cdot \left(1 + \frac{k_{ADPr}}{k_{ATP-tight}}\right)} \tag{3.2}$$

and

$$K_m \approx \frac{k_{ADPr} \cdot \left(1 + \frac{k_{ATPr}}{k_{ATP-tight}}\right)}{k_{ATPb} \cdot \left(1 + \frac{k_{ADPr}}{k_{ATP-tight}}\right)} \tag{3.3}$$

Where k_{ADPr}, k_{ATPb}, k_{ATPr} and $k_{ATP-tight}$ are the kinetic rates for ADP release, ATP loose binding or 'docking', ATP release, and ATP tight binding of each subunit, respectively.

Equation 3.11 indicates that the reduced V_{max} observed with mutant ring motors is due to a slower rate of ADP release, slower rate of ATP tight binding, or both. In WT motors, k_{ADPr} determines V_{max} since ADP release by the 4 translocating subunits are the slowest events during the dwell phase (Chistol et al. 2012). Accordingly, for WT motors, the statistical parameter N_{min} (defined as $< \tau >^2 / (< \tau^2 > - < \tau >^2)$, where τ is the dwell time), which represents a lower bound to the number of rate limiting events underlying the dwell time distribution (Moffitt & Bustamante 2014), equals 3.4 (Figure 3.2C). The substitution F145I does not change the value of N_{min} (Figure 3.2C), indicating that the number of rate limiting events is not affected. The invariance in N_{min} is easily explained if k_{ADPr} remains rate-limiting in F145I, but reduced relative to WT to account for the lengthening of the dwell in these mutants. In contrast, if $k_{ATP-tight}$ alone is reduced in F145I, its value would have to be comparable to k_{ADPr} ($k_{ATP-tight} \approx 1.5 \cdot k_{ADPr}$, according to equation 3.11) in order to produce the observed reduction in V_{max}. However, in this case, the number of rate limiting events in F145I and N_{min} value would be higher than in WT (as shown by Monte Carlo simulations, see Table 3.3.2), which we do not observe (Figure 3.2C). Thus, we conclude that k_{ADPr} is affected by the F145I substitution. Interestingly, an ATP titration shows that although V_{max} is reduced in F145I, K_m is invariant (Figure 3.2G). According to equation 3.12, to retain K_m invariance, the reduced ADP release rate in F145I must be accompanied by a similar reduction in ATP docking rate, k_{ATPb}, if we assume that the rates for ADP release and ATP undocking remain comparable, $k_{ADPr} \approx k_{ATPr}$; the latter
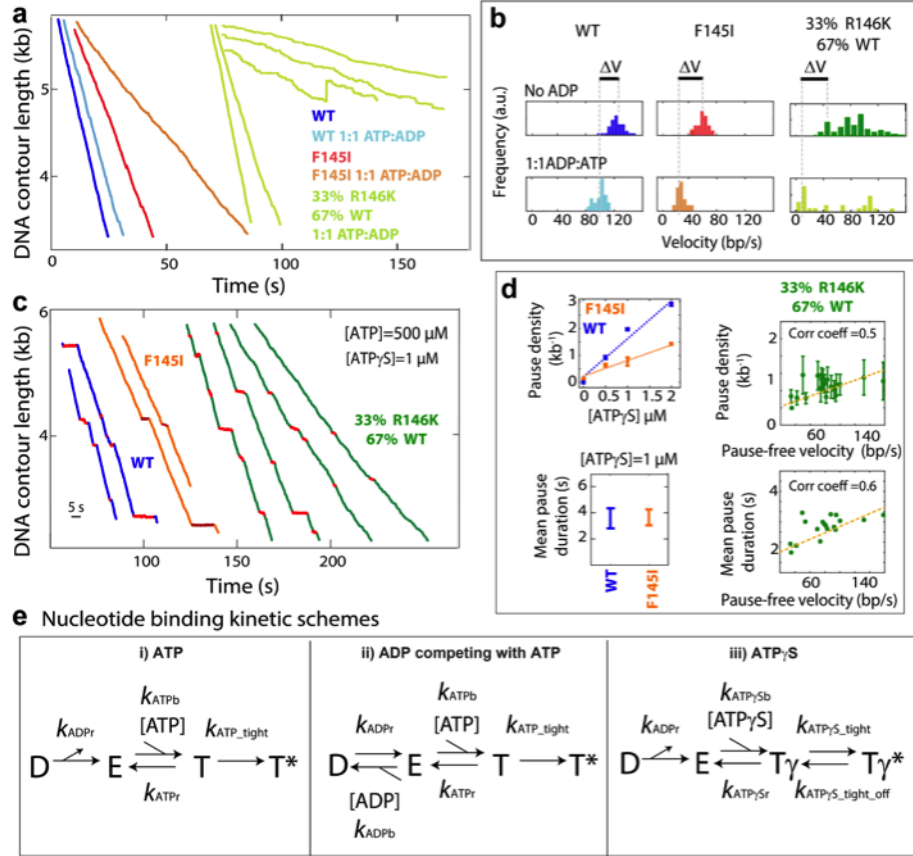
Figure 3.3: **The Arginine Finger Facilitates ADP Release and Locks ATP in the Tight Bound State.**
a) Sample packaging trajectories for WT, F145I and R146K/WT hybrid motors alone and in competition experiments with ADP ([ATP]=500 μM; [ADP]=500μM). **b)** Analysis of ADP inhibition of homomeric WT and F145I rings and of the slowest population of R146K/WT hybrid rings. **c)** Sample packaging trajectories for WT, F145I and R146K/WT hybrid motors in the presence of ATPγS. **d)** Pause density and mean pause duration of ATPγS-induced pauses in packaging trajectories of WT and F145I (average over many molecules and various [ATPγS]), and of R146K/WT hybrid motors (measured in each molecule, [ATPγS]=1μM). **e)** Kinetic schemes for competition experiments with nucleotide analogs. Given that ATP molecules are hydrolyzed much faster than $k_{ATP_{tight-off}}$, ATP tight binding transition is considered irreversible in schemes i and ii.

is a reasonable assumption, given that both nucleotides are likely to undo similar contacts with the motor during release. A Monte-Carlo simulation (see section 3.4) confirms that equally reduced k_{ADPr} and k_{ATPb} decrease

V_{max} without affecting K_m and N_{min}. This simulation also predicts that the mutant rings would be inhibited more in the presence of ADP than WT rings (Table 3.3.2). We tested this prediction by measuring the activities of F145I homomeric rings and R146K/WT hybrid rings obtained from a 1:2 mixture of R146K and WT subunits (33% in Figure 3.3A and 3.3B) in the presence of equal amounts (500 μM) of ADP and ATP (Figure 3.3A). Indeed, ADP inhibited F145I motors more than WT rings ($\sim$50% reduction in mean velocity for F145I vs $\sim$16% for WT, Figure 3.3B). Moreover, in the presence of ADP, the slowest population in the R146K mixture (rings with the highest number of mutant subunits) was also more inhibited by ADP than WT motors ($\sim$66% reduction in mean velocity, Figure 3.3B). Thus, both mutations affect nucleotide exchange in a similar manner. These results strongly support a model in which the trans-acting arginine, R146, mediates the communication between adjacent subunits during the nucleotide exchange phase (dwell) by modulating ADP release and ATP docking transitions.

The Arginine Finger Locks ATP in the Tight Bound State

Next, we sought to determine if the arginine finger plays a role in the ATP tight binding transition. To this end, we performed competition experiments with the non-hydrolyzable ATP analog, ATPγS. The rationale of these experiments is as follows: we previously showed that binding of a single non-hydrolyzable molecule (ATPγS or AMP-PNP) halts the activity of WT motors during the catalytic burst phase, inducing a pause in the packaging trajectory (Chistol et al. 2012). Since ATPγS is not hydrolyzed in the time scale of the pause duration, the motor can resume activity only after analog dissociation from the tight bound state (Chistol et al. 2012). Accordingly, the resulting pause duration is controlled by the rate of dissociation of the analog from the tight-bound state. ATPγS-induced pauses in R146K/WT hybrid rings display are shorter than in WT, whereas the length of these events is unaffected in F145I rings (Figure 3.3C and Figure S1). This observation indicates that chemical moieties of the arginine side chain stabilize the ATP tight bound state (see kinetic scheme iii in Figure 3.3E and section 3.4).

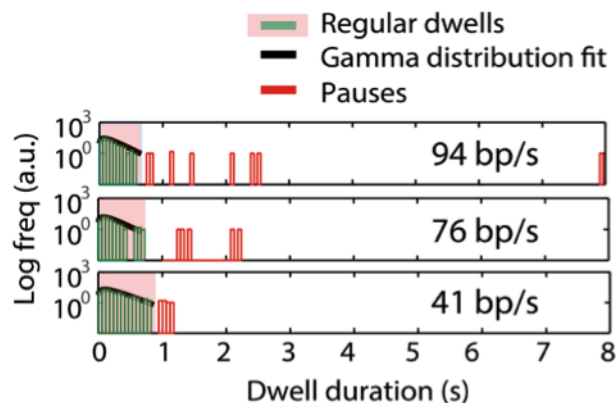Moreover, for any given [ATPγS], the probability of entry into an ATPγS-

Figure 3.4: **Identification of ATPγS-induced Pauses in R146K/WT Hybrid Motors.** Sample distributions of dwell times from individual packaging trajectories with different velocities. Dwell times were fitted to a Gamma distribution (black solid line). Events accounted by the fit (pink shade area) were considered as regular dwells (green). The long events unaccounted by the fit were considered pauses (orange).

induced pause (pause density) depends on the rate at which the nucleotide analog docks and is tight bound for a later hydrolysis attempt (see kinetic scheme iii in Figure 3.3E). Figure 3.3D shows that the ATPγS-induced pause densities in R146K/WT and F145I ring motors are reduced relative to WT, consistent with having a reduced ATP docking rate, a slower ATP tight binding rate, or both. Next, we proceeded to investigate if the arginine finger also performs a catalytic task.

The R146A Substitution is Tolerated Only at the Regulatory Subunit's Interface

Homomeric mutant rings harboring the substitution R146A, have been shown to be incompetent for ATPase activity in bulk assays (Mao et al. 2016). To obtain a phenotype for the R146A substitution at the single-molecule level, we applied the subunit mixing strategy used previously to study R146K hybrid rings. A mixture containing R146A and WT subunits in a proportion of 1:4 (20% in Figure 3.5B) yielded two clearly distinguishable phenotypes: WT-like activity (> 100 bp/s) and very slow motors (∼10 bp/s) (Figure 3.5A). This observation is in drastic contrast with the R146K phenotype,

34

which shows a multimodal velocity distribution and monotonic decrease in mean packaging velocity with increasing ratios of mutants in the mixture. The observation that a small proportion (20%) of R146A mutant subunits in the mixture greatly decreased the tether efficiency, suggests that a small number of R146A subunits in the rings is sufficient to render them inactive for packaging. A mixture containing a slightly larger proportion of mutant subunits (1:3 or 25% in Figure 3.5B) resulted in even fewer tethers, all of which displayed only the slow packaging behavior (Figure 3.5B). No tethers formed in samples with larger proportions of R146A in the mixture, consistent with R146A being much less tolerated than R146K. Notably, analysis of the slow packaging traces showed that the average dwell time is 10 times longer (mean dwell duration $\sim$720 ms) than in WT (mean dwell duration $\sim$72 ms) (Figure 3.5C), whereas the size and duration of the burst phase remains unaffected (Figure 3.5C).

Since we previously showed that each subunit powers the translocation of 2.5 bp (Moffitt et al. 2009), the conserved 10 bp burst size in R146A/WT hybrid rings indicates that 4 subunits are fully competent for DNA translocation and, therefore, also for ATP hydrolysis. This observation, together with the bi-modal velocity distribution (Figure 3.5B), strongly suggests that the slow phenotype in R146A/WT mixtures corresponds to rings with one R146A and 4 WT subunits (Figure 3.5D). Moreover, the mutation must be located at the regulatory subunit interface. Interestingly, a similar behavior was observed in WT motors in the presence of ATPγS (Chistol et al. 2012). Specifically, a fraction of motors that bound one single ATPγS were seen to package DNA with bursts of 10 bp, albeit separated by very long dwell times. In that study, a model was proposed in which motors displayed a single pause if the ATPγS had bound to any of the 4 translocating subunits, whereas they exhibited a cluster of long pauses separated by 10 bp if the non-hydrolyzable nucleotide was bound to the regulatory one (Chistol et al. 2012). The striking similarity between the packaging activities of R146A/WT hybrid motors and WT motors displaying pause clusters in the presence of ATPγS (Figure 3.5A), provides evidence that in both cases the subunit compromised is the regulatory one (Figure 3.5D). Presumably, motor configurations in which the single R146A substitution was located at a translocating subunit were paused and were unable to form tethers.
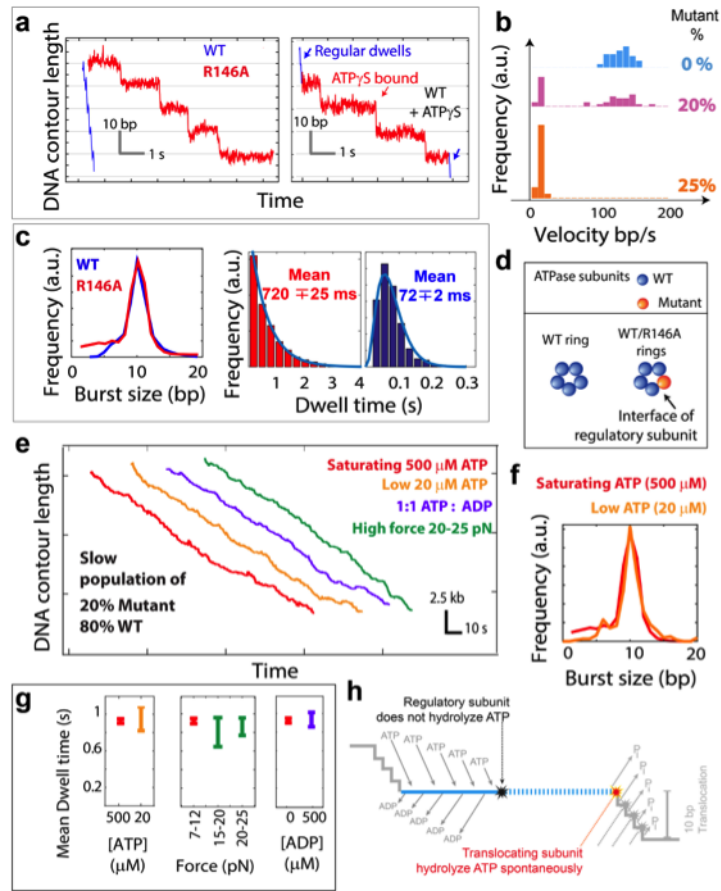
Figure 3.5: **R146A Substitution is Only Tolerated at the Regulatory Interface and is Necessary to Trigger DNA Translocation.**
a) High resolution packaging trajectories of WT and slow R146A/WT hybrid motors (left). Sample packaging trajectory of WT motors in the presence of the non-hydrolyzable nucleotide, ATPγS (right).
b) Velocity distribution of R146A/WT mixtures for different mixing ratios.
c) Analysis of the burst and dwell phases o f slow R146A/WT hybrid rings (red) vs WT motors (blue).
d) Schematic of homomeric and hybrid motors with packaging activity in R146A/WT mixtures.
e) Sample packaging trajectories of R146A/WT hybrid motors packaging in different conditions ([ADP]=500 μM).
f) Comparison of the burst phase of R146A/WT hybrid motors at saturating and low [ATP].
g) Mean dwell times of R146A/WT hybrid motors in different conditions (same color code as in e).
h) Schematic representation of the chemical transition affected by the R146A substitution.

36

R146-mediated ATP hydrolysis at the Regulatory Interface is Necessary for Timely Initiation of the Burst Phase

Next, we proceeded to determine what processes in the dwell phase are affected by the R146A substitution. The dwell times of R146A/WT hybrid motors follow a single exponential distribution with a time constant of $\sim$0.72 s, indicating that a single rate-limiting step controls the end of this phase (Figure 3.5C). In contrast, we previously showed that in WT motors the average dwell duration is $\sim$0.072 s (Figure 3.5C) and that 4 ADP release events that occur in this phase are rate-limiting (k_{ADPr}^{-1} $\sim$0.02 s), see section 3.4 and (Chistol et al. 2012). In view of the results presented in the previous section, the step that has become rate-limiting by the R146A substitution must be either the ADP release, the ATP binding, or the ATP hydrolysis by the regulatory subunit, all taking place during the dwell (Figure 3.4B). To test whether ATP binding and/or ADP release are affected, we collected packaging trajectories of R146A/WT hybrid motors at low [ATP] (20 μM) alone or at high [ATP] (500 μM) but in competition with ADP (500 μM). None of these conditions showed a change in mean dwell time, indicating that neither transition—ATP binding nor ADP release—is the single rate-limiting event in these motors (Figure 3.5E and 3.5G). Accordingly, the chemical transition slowed down or abolished by the alanine substitution must be the ATP hydrolysis at the regulatory interface. Because homomeric rings of R146A subunits do not show any ATPase activity in bulk assays (Mao et al. 2016), we conclude that the dwells are lengthened because ATP hydrolysis is completely abolished by the substitution R146A.

According to the current model (Figure 3.4B), the first ATP hydrolysis, which is catalyzed by the regulatory subunit, initiates the cascade of hydrolysis in the remaining translocating subunits. In R146A/WT hybrid rings, ATP hydrolysis is abolished at the interface of the regulatory subunit but packaging still proceeds, albeit very slowly. R146A/WT slow molecular trajectories (Figure 3.5A) imply that, after a long enough dwell, the translocating subunits must be able to bypass the absence of the regulatory signal and spontaneously initiate the burst. Therefore, the mean of the dwell time distribution of R146A/WT hybrid rings ($\sim$0.72 s) yields the spontaneous rate of hydrolysis by a translocating subunit ($\sim$1.4 s^{-1}). Because R146A/WT motors display consistently 10 bp bursts (Figure 3.5C), the subunit that hydrolyzes

ATP spontaneously hydrolyzing must be the first immediately following the regulatory one (Figure 3.5D), and not the second, third or fourth, as those cases would result in 7.5, 5.0, and 2.5 bp bursts, respectively (Figure 3.5H). This observation also implies that the rate of spontaneous hydrolysis of the remaining translocating subunits is slower than ~ 1.4 s^{-1}. It is not clear at this point what are the physical or structural bases behind the faster spontaneous hydrolysis rate of the first translocating subunit. As shown below, however, asymmetric cryo-EM reconstructions (Figure 3.7E) of packaging motors suggest that such difference can be originated by more extensive contacts established between the regulatory and the first translocating subunits than between any other pair.

R146A/WT rings provide us with the means to investigate the mechanism by which the regulatory subunit initiates the burst in WT motors. At the end of the dwell, the regulatory subunit is tight bound to the DNA through electrostatic interactions (Aathavan et al. 2009), and ATP hydrolysis by this subunit is thought to be necessary to 1) undo this interaction in WT motors and 2) trigger the cascade of ATP hydrolysis (Chistol et al. 2012). However, whether these two events are independent of each other, or if either of them is sufficient to allow the cycle to continue is not known. The application of high external forces in R146A/WT hybrid rings (up to 25 pN, enough to break the motor-DNA interaction as shown by the increased terminal tether breaking in Figure 3.6), did not lead to a change in dwell time duration (Figure 3.5E and 4G), which would have indicated a premature initiation of the burst. These results show that in the absence of ATP hydrolysis by the regulatory subunit, breaking of the motor-DNA interaction is not sufficient to initiate the burst. It follows that ATP hydrolysis by the regulatory subunit is the necessary signal that triggers the hydrolysis cascade in the remaining subunits in a timely manner.

The Regulatory Subunit Possesses Both Spontaneous and Stimulated ATPase Activities that Control the End of the Dwell

The results in the previous section indicate that, under extreme conditions and a long enough idle interval, the first translocating subunit can spontaneously hydrolyze its ATP without the signal from the regulatory subunit.
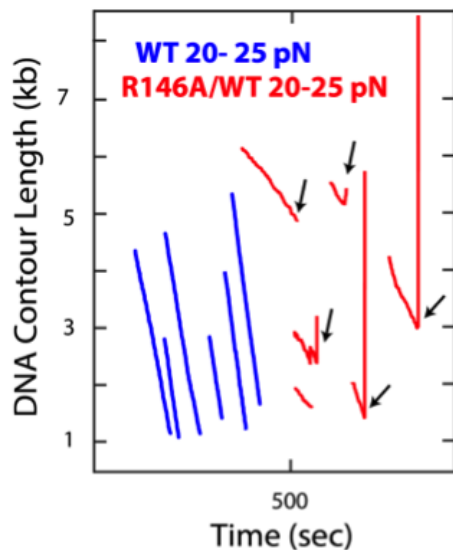
Figure 3.6: **Tethers of R146A/WT Hybrid Motors Display Higher Rupture Probability than WT at High Force (20-25 pN).** Packaging trajectories of WT motors (blue) are processive and continue to package DNA for stretches of 2-4 kilo-bases of DNA. In contrast, the slow motors from a mixture containing R146A mutant subunits in a 1:4 (WT:R146A) ratio display fatal rupture often before packaging 2 kb of DNA . Fatal rupture events (black arrows) were observed as sudden change in extension followed by tether breaking.

The behavior of WT motors packaging DNA at saturating [ATP] suggests that this signal is in turn triggered by the completion of ATP binding by all five subunits (Figure 3.4B). This scenario begs the question: can the regulatory subunit also hydrolyze ATP without the signal from the last subunit that binds ATP? To answer this question, we next performed experiments at very low [ATP] designed to delay the last binding event and with it the signaling to the regulatory subunit to hydrolyze its nucleotide. We monitored the activity of WT motors in the presence of 20 μM and 10 μM [ATP]—well below the K_m value of 35 μM—(Figure 3.7A). Remarkably, as ATP becomes more limited, the average burst size decreases from 10 bp at 500 μM to 7.6 bp at 20 μM and to 6.1 bp at 10 μM (Figure 3.7B). Because each subunit translocates 2.5 bp of DNA, these observations indicate that at low enough [ATP] fewer than 4 subunits on average are involved in translocation per cycle, implying that the motor can initiate the burst before all subunits have bound ATP.

39

Do subunits under these extreme conditions operate independently of each other? We note that even for the most limiting [ATP] conditions tested (10 μM), the bursts are generally larger or equal to 5 bp (Figure 3.7B). This observation suggests that low [ATP] the subunits retain a degree of coordination in nucleotide binding and ATP hydrolysis. In fact, analysis of motor trajectories operating at very low [ATP] in the presence of ATPγS (Figure S3A), shows that a single bound analog suffices to halt the motor (see section 3.4), similarly to what is observed at nucleotide concentrations above K_m (Chistol et al. 2012). Moreover, we found that even at low [ATP], ADP still behaves as competitive inhibitor (Figure S3B) and that the Hill coefficient for ATP, nH, still equals 1 (Figure S3C), consistent with the idea that only one site continues to be available for ATP binding during the dwell in these conditions (see section 3.4). Therefore, we conclude that even at very low [ATP] nucleotide exchange takes place in a coordinated sequential manner, and the smaller bursts observed result from the coordinated sequential firing by fewer adjacent subunits.

Next, we asked: what event causes the end of the dwell when the ring is not fully saturated with ATP. Importantly, we observed that at limiting [ATP] the duration of the dwell phase correlates positively with the size of the following burst (Figure 3.7C): longer dwells ($\sim$350 ms at 20 μM ATP) result in $\sim$10 bp translocation events, whereas shorter dwells ($\sim$200 ms at 20 μM ATP) tend to result in smaller bursts. In contrast, at saturating [ATP] no correlation is observed between dwell duration and burst size (Figure 3.7C) and the mean dwell time is shorter ($\sim$72 ms). These observations confirm the idea that at high [ATP] hydrolysis by the regulatory subunit is promptly activated upon ring saturation, whereas at very low [ATP] a different, stochastic mechanism must trigger the end of the dwell. Such spontaneous mechanism could be the thermally activated hydrolysis by the regulatory subunit or by one of the translocating subunits. However, as shown in the previous section, the rate of spontaneous hydrolysis of a translocating subunit ($\sim$1.4 s^{-1}, leading to an average dwell duration of $\sim$720 ms) is too slow to account for the average dwell durations observed at very low [ATP] ($\sim$250 ms, corresponding to a rate of 4 s^{-1}). Therefore, we propose that the event that triggers the end of the dwell at very low [ATP] is the spontaneous hydrolysis by the regulatory subunit. If our interpretation is correct, any other process that delays the nucleotide saturation of the ring is expected to produce a similar spontaneous initiation of the burst. In the presence of ADP, ATP binding
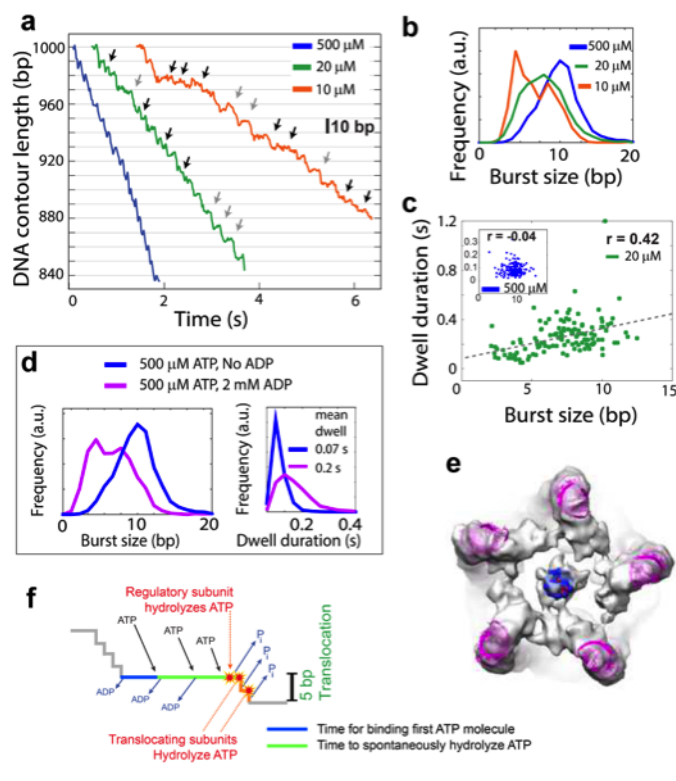
Figure 3.7: **Spontaneous ATP Hydrolysis Triggers DNA Translocation at Low [ATP] in WT Motors.**
a) Sample of packaging trajectories of WT motors at different [ATP]. Grey arrows indicate 10-bp bursts. Black arrows indicate <10 bp bursts.
b) Burst size distribution at saturating and sub-K_m ATP conditions.
c) Correlation of dwell duration with the following burst size. r is the *Pearson* correlation coefficient.
d) Analysis of burst and dwell phases of WT motors in the absence (blue) and presence of high [ADP] (magenta).
e) Asymmetric cryo-EM reconstruction. Electron density of viral particles stalled during DNA packaging with ATPγS. ATPase subunits and DNA viewed as a bottom view (from the outside) of the bacteriophage. Atomic resolution structures of pRNA (magenta) and DNA (dark blue) were fitted to the electron density.
f) Schematic representation of the chemical cycle at limiting [ATP].

is delayed due to the competition between the two nucleotides for the reversible docking state, resulting in longer dwells. Indeed, in 4-fold excess of [ADP] over [ATP], we observed a comparable average dwell duration of ∼200 ms and translocation events with a size <10 bp (Figure 3.7D). Again, these

41

observations are consistent with a scenario in which the regulatory subunit spontaneously triggers the hydrolysis cascade after a long enough dwell even though not all 5 subunits have bound ATP. Together, these results indicate that the regulatory subunit in WT motors exhibits both spontaneous and stimulated ATPase activities, the latter resulting from saturation of the ring with ATP.

ATP Hydrolysis by the Regulatory Subunit is Stimulated by DNA Interaction

The results presented above show that the spontaneous rate of the first translocating subunit (~ 1.4 s^{-1}) is slower than the spontaneous rate of the regulatory subunit (~ 4 s^{-1}), which raises the question: what is the origin of such difference between otherwise identical subunits? We previously proposed that an electrostatic contact between one of the subunits and DNA phosphates every 10 bp (Aathavan et al. 2009) is the event that breaks the symmetry of the motor (Chistol et al. 2012). According to this model, the subunit that binds the DNA adopts the regulatory role. To detect a possible symmetry breaking interaction between the DNA and the regulatory subunit, we imaged packaging motor complexes, stalled with ATPγS, using cryo-electron microscopy. We avoided imposing five-fold symmetry constraints on the data as done previously (Mao et al. 2016) such that structural differences between individual subunits could be visualized. In this way, an asymmetric reconstruction of the stalled particles was obtained. Based on the current model (Liu et al. 2014b), this structure of the motor, stalled with ATPγS, likely represents a state in the dwell just prior to the burst phase (Figure 3.4B), in which the motor has not yet detached from the DNA phosphates. In the resulting ~ 12 $\mathring{A}$ electron density map (Figure 3.7E), one of the five subunits is seen to interact more extensively with the DNA than the rest. In addition, the data suggest more extensive contacts between the regulatory and the first translocating subunits than between any other pair. The symmetry breaking observed in this cryo-EM reconstruction provides support to the idea that the spontaneous activity of a subunit is stimulated by its periodic contact with the DNA to become the regulatory one. This interpretation is consistent with observations made in bulk showing that DNA stimulates the ATPase activities of the motor subunits (Todd et al. 2012). Although contact with the DNA might increase the spontaneous ATPase

activity of the regulatory subunit over its translocating counterparts ($\sim$4 s^{-1} vs $\sim$1.4 s^{-1}, respectively), a $\sim$3 times faster rate would not ensure that such subunit will always hydrolyze ATP first, and cannot solely account for the termination of the dwell and the initiation of the burst after only 72 ms. In fact, the rate of hydrolysis of the regulatory subunit, which occurs in the dwell, ought to be much faster than the rate of ADP release—the rate-limiting event during that phase ($\sim$50 s^{-1}, see section 3.4 and (Chistol et al. 2012)—and it is likely to be at least as fast as the stimulated hydrolysis of the translocating subunits during the burst ($\sim$330 s^{-1}, assuming that 3 hydrolysis events must be completed in the $\sim$0.01 s duration of this phase). Thus, another event must occur to further increase the ATPase activity of the regulatory subunit. At this point we do not know what this event is. In the following section we speculate what its nature could possibly be.

3.4   Discussion

Discussion: Molecular Basis for Inter-subunit Coordination

It was previously shown that the ATPase subunits of the ϕ29 DNA packaging motor are highly coordinated (Figure 3.4B) (Chistol et al. 2012). However, the molecular mechanism for such coordination has not been understood. In this paper, we have presented evidence of the role played by the trans-acting arginine finger R146 in mediating the communication between adjacent subunits in the motor to bring about its global coordination.

First, we have shown that the trans-acting arginine finger, R146, of a subunit locks the ATP in the tight bound state in the binding pocket of the adjacent subunit. Moreover, we have shown that R146 facilitates the release of ADP. These data are consistent with a model in which tight binding of ATP to a subunit—stabilized by an arginine finger (Figure 3.8 *ii* and *iii*, colored in black)—induces, in turn, the release of ADP from the adjacent pocket by a mechanism facilitated by the next arginine finger (Figure 3.8 *ii* and *iii*, colored in green). Together, these two events give rise to the sequential exchange of nucleotide during the dwell phase (Figure 3.8, green dashed arrow). We speculate that the first ADP release event—responsible for initiating nucleotide exchange—takes place at the regulatory subunit when it

43

contacts the DNA phosphates at the beginning of the dwell (Figure 3.8 i, grey up arrow).

Second, we have shown that the motor subunits possess spontaneous and stimulated ATPase activities. Indeed, experiments with R146A/WT hybrid motors containing a single mutant subunit, show a 10-fold lengthening of the dwell ($\sim$720 ms). We propose that this phenomenon occurs because the mutation prevents the regulatory subunit to initiate its own hydrolysis and activate the hydrolysis of the first translocating subunit (Figure 3.5H). Accordingly, the lengthened dwell reflects the spontaneous hydrolysis rate of the first translocating subunit ($\sim$1.4 s^{-1}) in the absence of stimulation by its regulatory counterpart. Similarly, experiments at low [ATP] show that the burst cascade initiates spontaneously—before the ring is fully saturated with ATP—after $\sim$250 ms, which we propose reflects the spontaneous rate of hydrolysis of the regulatory subunit ($\sim$4 s^{-1}).

We argue that the ability of the subunits to switch from spontaneous to stimulated ATPases, is fundamental to the biphasic operation of the WT packaging motors. During normal motor operation, none of the subunits hydrolyze ATP throughout the nucleotide exchange process (Figure 3.8 $i - iv$), indicating that both the regulatory and the translocating subunits display their spontaneous ATPase activities throughout most of the dwell. As a result, towards the end of the dwell, all subunits have managed to bind ATP without hydrolyzing it (Figure 3.8 iv). At this point, the regulatory and first translocating subunits must quickly hydrolyze ATP to initiate the burst. The cryo-EM data presented here is consistent with a model in which the regulatory subunit is enhanced by the DNA contact, making it a slightly faster ATPase subunit that tends to hydrolyze the nucleotide first. However, its spontaneous rate ($\sim$4 s^{-1}) is not large enough to quickly initiate the burst (at a rate > 50 s^{-1}, the rate of ADP release). Thus, the regulatory subunit must be stimulated additionally through a second process. Since the burst phase is triggered soon after all the subunits complete ATP binding (Moffitt et al. 2009), we propose that the last subunit to bind ATP is responsible for the additional stimulation of the regulatory subunit (Figure 3.8 vi). We also show that the ATPase activity of the regulatory subunit stimulates the activity of the first translocating subunit by a factor > 200. Finally, each of the remaining translocating subunits must be stimulated by its predecessor to quickly hydrolyze ATP at a rate of $\sim$330 s^{-1} to complete the burst phase in $\sim$10 ms (Figure 3.4B).
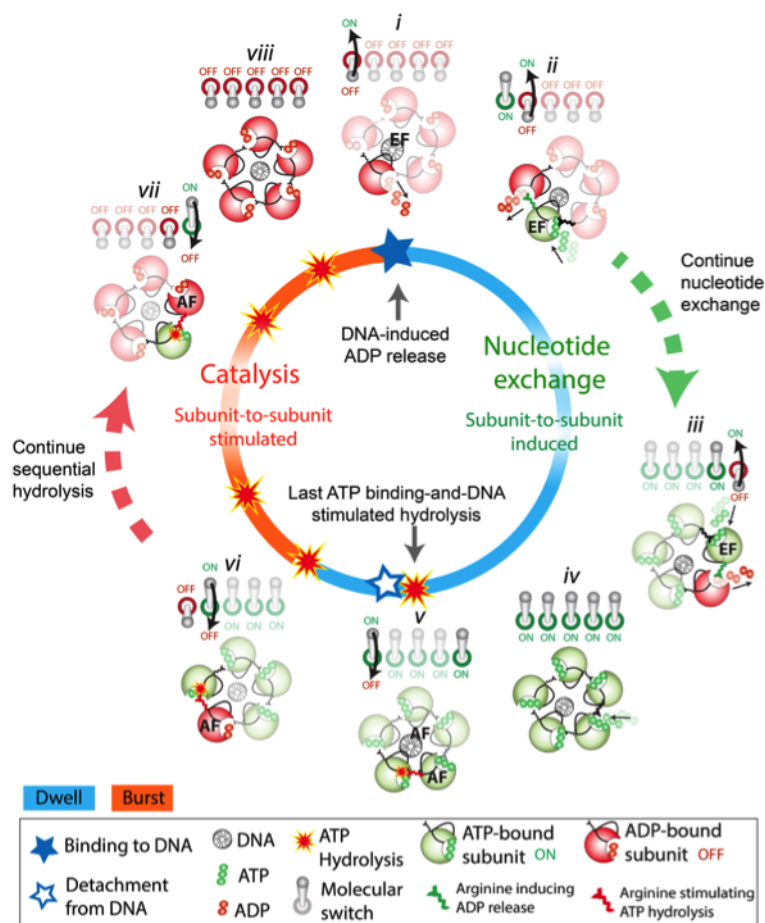
Figure 3.8: **Model for Inter-subunit Coordination and Regulation.**
Schematic representation of the ATPase ring in a bottom view configuration. See Movie S1. (Clockwise) (i) At the beginning of the dwell one subunit contacts the DNA (blue solid star). The DNA adopts the role of an exchange factor, facilitating ADP release by the regulatory subunit. (ii) R146-mediated (black) ATP tight binding at the regulatory subunit converts it into a nucleotide exchange factor for the next subunit via its arginine finger (green). The resulting binding pocket is empty and is now able to bind ATP. The process of nucleotide exchange repeats around the ring (green dotted arrow and iii). (iv) Towards the end of the dwell all subunits are bound to ATP or in the 'on' state. (v) The DNA and the last subunit to bind ATP adopt the role of activating factors, both stimulating ATP hydrolysis at the regulatory subunit (red). (vi) ATP hydrolysis and phosphate release in the catalytic pocket of the regulatory subunit converts it into an activating factor for the next subunit via its arginine finger (red). The process repeats around the ring creating the cascade of events observed during the burst (red dashed arrow and vii). (viii) At the end of the burst all 5 subunits are ADP-bound or in the 'off' state. The cycle begins again.

45

The pseudo-atomic model of the packaging motor (Mao et al. 2016) suggests the physical bases for the sequential ATPase stimulation and DNA translocation (Figure 3.8 and Movie S1). The subunits' catalytic sites are adjacent to the putative DNA translocating loops, which are, therefore, well positioned to perform a power stroke following ATP hydrolysis. Moreover, normal mode analysis of the ATPase subunit indicates that such movement of the DNA translocating loop is coupled to a corresponding change in conformation of the arginine finger at the opposite side of the same subunit (Mao et al. 2016). We propose that the last ATP binding event induces the re-positioning of the arginine finger at the regulatory subunit's catalytic site, which stimulates its ATPase activity (Figure 3.8 *vi*). Similarly, ATP hydrolysis in any subunit causes a power stroke by the DNA translocating loop, which in turn re-positions the linked arginine finger to stimulate ATP hydrolysis in the following catalytic site (Figure 3.8 *vi* and *vii*). These events result in the sequential hydrolysis and translocation observed during the burst (Figure 3.8, red dashed arrow).

Discussion: The Coordination in the ϕ29 DNA Packaging Motor Arises from Two Mechanisms of Regulation

Here we have shown that throughout the motor's cycle, the subunits switch from 'poor' (during the dwell) to 'efficient' (during the burst) ATPase activity, a process that requires a mechanism of enzymatic regulation. A well-studied example of NTPase activity regulation is found among small-GTPases, which are 'poor' GTPases and are known as molecular switches because they can be set 'on' or 'off' in a regulated manner by certain protein factors (Cherfils and Zeghouf, 2013). Some of these factors—GTPase-activating proteins (GAPs)—enhance the enzymatic activity of small-GTPases using a trans-acting arginine residue to generate their GDP-bound or 'off' state (Ahmadian et al. 1997). Others—GDP-exchange factors (GEFs)—promote nucleotide exchange by facilitating GDP release through an allosteric competitive mechanism and generate the GTP-bound or 'on' state (Guo et al. 2005). We have shown that the subunits of the ϕ29 DNA packaging motor can be set 'on' or 'off' in a regulated manner and, as a result, the ring can be thought of as a set of 5 molecular switches in a closed configuration (Figure 3.8). According to this view, the subunits adopt first GEF—and then GAP—like functions to regulate their adjacent partners during the dwell and

burst phases, respectively. Moreover, we propose that the motor/DNA contact provides an additional mechanism of regulation for nucleotide exchange and ATP hydrolysis on the regulatory subunit. As a result, the DNA-bound subunit becomes a master regulator that initiates both the cascade of nucleotide exchange and the cascade of ATP hydrolysis. We suggest that these two mechanisms of regulation enable the precise global coordination that results in the 'clockwork' operation of the bacteriophage ϕ29 DNA packaging motor (See Movie S1).

Discussion: Molecular-switch-like Regulation in Other Ring ATPases

The coordination scheme of the ϕ29 DNA packaging motor evolved to successfully translocate DNA against the large internal forces generated during packaging (Smith et al. 2001). As a result, the motor displays an effective strategy that allows it to spend a proportionally large fraction of its time ($\sim$90%) in a high-affinity state for DNA during nucleotide exchange (dwell), and to be only transiently detached from its substrate during packaging (burst) (Chistol et al. 2012). Other ring ATPases perform their mechanical tasks on different substrates and contexts, and accordingly, they display alternative coordination schemes that may be better suited for their biological function (Liu et al. 2014b). Nonetheless, these ring ATPases all possess the conserved arginine at the interfacial active sites. Thus, it is likely that to coordinate their activity, their ATPase subunits employ small GTPase-like regulatory mechanisms similar to those proposed here for the ϕ29 packaging motor, although to different degrees, depending on the nature of their substrate, its length, and its periodicity. For instance, in F1-ATPase the β subunit contains the catalytic pocket, whereas the α subunit is known to both stimulate the ATPase activity (Komoriya et al. 2012) and facilitate ADP release by the β subunit (Yukawa et al. 2015) via its arginine finger. The enzymatic activity of the β subunit is greatly enhanced in the presence of the γ subunit (the 'substrate' of this rotary motor) (Iino & Noji 2013), pointing to a second mechanism of regulation possibly required for the high efficiency of the F1-ATPase motor.

Small GTPase-like regulation has been proposed to underlie the operation of other essential proteins in a variety of biological contexts. For instance,

the similarity between the small-GTPase regulation and the substrate-to-subunit regulation in transport motors, such as kinesin and myosin, has been previously discussed (Vale 1996). Similarly, the human and bacterial base pair mismatch recognition complexes have been proposed to operate as DNA-regulated molecular switches (Gradia et al. 1997, Lebbink et al. 2010). It is tempting to speculate that the two mechanisms of regulation described here might be present in other ring ATPases, as well as in other biological contexts.

3.A   Appendices

Appendix: Estimated Population Fractions for Various Numbers of R146K Mutants Tolerated in the Ring

In the R146K/WT mixing experiments, the motors' velocity clustered roughly in three main groups: fast ($\sim$125 bp/s, WT-like), intermediate ($\sim$ 85 bp/s) and slow ($\sim$ 45 bp/s) (Figure 2B). For each mixing condition (0%, 25%, 33% or 40%), each group was fitted to a Gaussian distribution. The superposition of the Gaussian functions within one mixing condition was normalized, and the area under each Gaussian was considered to be the fraction of motors belonging to either group (WT-Like, intermediate or slow) in a given mixing condition. Next, we proceeded to estimate the fractions of ring motors with 1, 2, 3 and 4 R146K subunits predicted by the binomial distribution and by the four R146K/WT mixing ratios (0%, 25%, 33% and 40% of the subunits were R146K), assuming that either subunit, mutant or WT, has the same affinity to be incorporated in the ring —a reasonable assumption given that motor subunits bind to a pRNA scaffold to form the ring configuration, the pRNA binding site in the subunit is far from R146 and pRNA-subunit binding is not affected by the presence of nucleotide, ATP or AMP-PNP (Ding et al. 2011). Next, we considered individually the scenarios of allowing the motor to tolerate up to 1, 2, 3 or 4 R146K subunits. In each scenario, motors with more mutant subunits than the maximum allowed were considered inactive, and the fractions were normalized after discarding the portion of inactive motors for each scenario. We observed that, in all mixing conditions the estimated number of WT motors (0 mutant subunits in the ring) was higher compared to the fraction of WT-like traces observed experimentally. We rationalize this observation considering that the mixture samples were
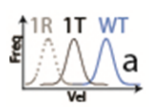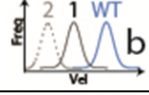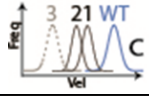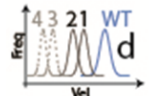
Cases:	Population fractions for 1:3 mixing ratio – (25%)			Population fractions for 1:2 mixing ratio – (33%)			Population fractions for 2:3 mixing ratio – (40%)			Variance (deviation from experiment)
	WT-like	Inter-mediate	Slow	WT-like	Inter-mediate	Slow	WT-like	Inter-mediate	Slow	
Experimental	0.09	0.78	0.128	0.09	0.78	0.13	0	0.6	0.39	0
a	0.37	0.5	0.125	0.29	0.57	0.14	0.23	0.61	0.15	0.35
b	0.26	0.44	0.26	0.16	0.41	0.41	0.11	0.37	0.5	0.45
c	0.24	0.67	0.08	0.16	0.69	0.14	0.085	0.66	0.25	0.08
d	0.23	0.65	0.1	0.2	0.66	0.13	0.07	0.61	0.31	0.07

Table 3.1: **Estimated Population Fractions for Rings Tolerating Various Numbers of R146K Mutant Subunits.** Compare cases across rows. Experimental: fractions were obtained as the area under the Gaussian fit for each group. Case a, b, c and d: fractions were calculated based on the binomial distribution and the mixing ratio (1:3, 1:2 or 2:3). Case a: the ring tolerates 1 mutant subunit, but the substitution affects differentially the translocating and regulatory subunits. In the schematic velocity distribution (first column), 1R refers to 1 mutant regulatory subunit and 1T to 1 mutant translocating subunit. Case b, c and d: the ring tolerates a maximum of 2, 3 and 4 mutants in the ring, respectively. In the schematic velocity distributions (first column): 1 = rings with one mutant subunit, 2 = rings with 2 mutant subunits, 3 = rings with 3 mutant subunits and 4 = rings with 4 mutant subunits.

allowed to pre-package 30 s longer than WT samples before stalling the viral complexes. This extra time was necessary to favor the packaging initiation of the slower mutant rings, but in turn, it allowed WT motors (packaging rate at ∼125 bp/s) to package ∼3.7 kb of DNA. In WT samples, motors are often tethered as they are packaging the last 3 kb piece of DNA. Thus, it is likely that a fraction of WT-like motors in the mixture samples had already completed packaging by the time the motors could be tethered. We next proceeded to assign the slow and intermediate groups to rings with different number of mutants, or a combination of them, accordingly to the maximum number of mutant subunits tolerated in each scenario. In the scenario where

rings tolerate only 1 mutant subunit (case a), we assigned the slow group to rings with 1 mutant regulatory subunit, and the intermediate group to rings with 1 mutant translocating subunit (see Table S1). In scenarios where rings could tolerate up to 2 (case b), 3 (case c) or 4 (case d) R146K mutants in the ring, the intermediate group was assigned to rings with 1 mutant subunit in case b, and the superposition of rings with 1 and 2 mutant subunits in cases c and d (see Table S1). Similarly, the slow group was assigned to rings with 2 mutants in case b, 3 mutants in case c and 3 and 4 mutants in case d. With these assignments we obtained for each case an estimated fraction of motors in each group (WT-like, intermediate and slow) for each mixing condition. Next, we estimated the goodness of each case by measuring the variance relative to the experimental distribution. The variance for a given case x is defined as follows:

$$Var_x = \sum_{i,j}(F_{T_{ij}} - FE_{ij})^2 \qquad (3.4)$$

Where $F_{T_{ij}}$ is the estimated fraction of group i in the mixing condition j, and $F_{E_{ij}}$ is the experimental fraction of group i in the mixing condition j. As shown by the variance value, the estimated fractions were closer to the experimental ones in cases c and d. Thus, according to this analysis then, rings can tolerate 3 or more R146K subunits without complete loss of packaging activity.

Appendix: Derivation of an Analytical Expression for the Motor's Packaging Velocity

The packaging velocity is the ratio of the length of DNA packaged in the burst (10 bp) to the total duration of the motor's cycle. The latter is the inverse of the cycle rate, k_{cycle}, which depends on the rate at which the motor completes nucleotide exchange and hydrolysis and phosphate release in all 5 (4 translocating and 1 regulatory) subunits, $k_{nucx,cycle}$ and $k_{h,cycle}$, respectively. The rate for the overall k_{cycle} can then be written as:

$$k_{cycle} = \frac{k_{nucx,cycle} \cdot k_{h,cycle}}{k_{nucx,cycle} + k_{h,cycle}} \qquad (3.5)$$

50

Here, $k_{h,cycle} = k_h/5$, with k_h being the rate of hydrolysis and phosphate release rate per subunit, and we have neglected the difference in these parameters between the regulatory and translocating subunits. We previously showed that nucleotide exchange takes place sequentially, one subunit at a time (Chistol et al. 2012). Moreover, we showed that the kinetic events of nucleotide exchange in each subunit are separated from those of the preceding subunit by a largely irreversible transition (Moffitt et al. 2009). Furthermore, that transition is likely to be ATP tight binding, which locks the nucleotide in the bound state and provides the energy to open the binding pocket of the following subunit for ADP release (Moffitt et al. 2009). The nucleotide exchange scheme for one subunit can be written as:

$$D \xrightarrow{k_{ADPr}} E \underset{k_{ATPr}}{\overset{\overset{\displaystyle k_{ATPb}}{[ATP]}}{\rightleftarrows}} T \xrightarrow{k_{ATP\_tight}} T^*$$
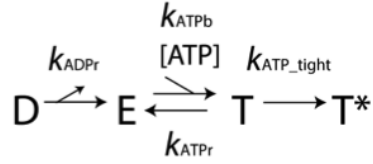
In this scheme, **D**, **E**, **T** and **T**∗ indicate ADP-bound, Empty, ATP-bound, ATP-tight bound states, respectively, and k_{ADPr}, k_{ATPb}, k_{ATPr} and $k_{ATP-tight}$ are the kinetic rates for ADP release, ATP binding, ATP release and ATP tight binding per subunit, respectively. We proceeded to determine an expression for $k_{nuc,x}$ for a ring motor with 5 subunits sequentially exchanging nucleotide according to the previous scheme (2). The rate of nucleotide exchange by the regulatory subunit is much faster than the corresponding rate of the 4 translocating subunits (Chistol et al. 2012, Moffitt et al. 2009). Thus, considering only the slower translocating subunits, the rate of nucleotide exchange of the whole ring is:

$$k_{nucx,cycle} = \frac{\frac{k_{ADPr} \cdot k_{ATP-tight}}{k_{ADPr} + k_{ATP-tight}} \cdot [ATP]}{4 \cdot \left([ATP] + \frac{k_{ADPr}}{k_{ATPb}} \cdot \left(\frac{k_{ATP-tight} + k_{ATPr}}{(k_{ATP-tight} + k_{ADPr})}\right)\right)} = \frac{V_{nucx} \cdot [ATP]}{4 \cdot ([ATP] + K_{nucx})} \tag{3.6}$$

where

$$V_{nucx} = \frac{k_{ADPr} \cdot k_{ATP-tight}}{k_{ADPr} + k_{ATP-tight}} \tag{3.7}$$

51

and

$$K_{nucx} = \frac{k_{ADPr}}{k_{ATPb}} \cdot \frac{k_{ATP-tight} \cdot k_{ATPr}}{k_{ATP-tight} + k_{ADPr}} \tag{3.8}$$

denote the Michaelis-Menten parameters per subunit. Accordingly, the packaging velocity of the motor is:

$$v = \frac{\frac{V_{nucx}}{V_{nucx} \cdot \left(\frac{1}{k_{h,cycle}}\right) + 4} \cdot [ATP]}{[ATP] + \frac{K_{nucx}}{V_{nucx} \cdot \left(\frac{1}{4 \cdot k_{h,cycle} + 1}\right)}} * 10bp \tag{3.9}$$

Since the rate of the catalytic phase is faster ($k_{h,cycle}$ is ~ 100 s^{-1} based on the mean burst duration ~ 0.01 s) than the rate of ADP release per subunit (k_{ADPr} is ~ 50 s^{-1} assuming 4 ADP release events during a mean dwell duration ~ 0.072 s, see Monte Carlo simulation in SI and (Chistol et al. 2012)), and therefore, $4 \cdot k_{h,cycle}$ ¿¿ k_{ADPr}, it can be shown that $4 \cdot k_{h,cycle}$¿¿ V_{nucx}, and thus the overall velocity of the cycle can be approximated as:

$$v = \frac{V_{cycle} * [ATP]}{[ATP] + K_{cycle}} \cdot 10bp \tag{3.10}$$

where

$$V_{cycle} \approx \frac{k_{ADPr}}{4 \cdot \left(1 + \frac{k_{ADPr}}{k_{ATP-tight}}\right)} \tag{3.11}$$

and

$$K_{cycle} \approx \frac{k_{ADPr} \cdot \left(1 + \frac{k_{ATPr}}{k_{ATP-tight}}\right)}{k_{ATPb} \cdot \left(1 + \frac{k_{ADPr}}{k_{ATP-tight}}\right)} \tag{3.12}$$

The contribution of kh,cycle to Vcycle and Kcycle is only about 10% in equations (5) and (6). In the main text V_{cycle} and K_{cycle} will be referred to as Vmax and Km, respectively.

Appendix: Monte-Carlo Simulation of the Motor's Mechano-Chemical Cycle and its Inhibition by ADP

The Monte Carlo algorithm uses the kinetic schemes of Figure 3.3E—(i) for WT scenario and (ii) ADP competition experiments—to compute the duration of each cycle assuming that all 5 subunits of the ring must exchange nucleotide sequentially one subunit at a time. The 5 subunits must bind ATP before the first hydrolysis event is allowed, and the catalytic phase is coupled to a 10 bp translocation event. The cycle is repeated 1000 times (equivalent to packaging 10 kb). The packaging velocity is estimated as the amount of packaged DNA over the total time it takes to complete the 1000 cycles.

	WT scenario		F145I - 5 mutant subunits				R146K/WT- 3* mutant subunits in the ring			
	Exp. data	M.C.	Exp. data	k_{ATP_tight} 1/50	k_{ADPr} 1/23	k_{ADPr} -1/23 k_{ATPb} -0.69	Exp. data	k_{ATP_tight} 1/18	k_{ADPr} 1/12	k_{ADPr} -1/12 k_{ATPb} -0.36
V_{max}	125	128	65	64.5	66.5	65.5	45	45.5	46.5	45.5
K_m	32	33.5	35	31	18.5	37	N/M	--	--	--
N_{min}	3.4	3.6	3.5	7.2	3.7	3.3	N/M	--	--	--
$V_{ADP\_1:1}$	105	103	30	50.5	52	36	15	35.5	37.5	18
% ΔV	16%	19.5%	54%	22%	22%	45%	67%	22%	19%	60.5%

Table 3.2: **Summary of Results from Monte-Carlo Simulation** Compare across columns. In the WT scenario, experimental values are displayed in green. The rate constants in the Monte Carlo (M.C.) simulation were selected to match the experimental values, starting from previous values (Chistol et al. 2012). See Extended Experimental Procedure for all rate constants. In the F145I and R146K/WT scenarios, the modified rate constants (indicated in the top of each column) were selected to match the experimental velocity corresponding to either case (shown in green). The remaining rate constants were kept as in the WT scenario.
$V_{ADP1:1}$ is the velocity of the motor in the presence of equal amounts of ATP and ADP ([500 μM]).
$\%\Delta V$ is the percentage reduction in velocity due to the ADP inhibition.
Exp. Data = Experimental Data.
N/M= Not measured.
*According to our analysis, the number of mutant subunits in the slowest population is 3 or 4.

The kinetic parameters that reproduced the behavior of WT motors at saturating [ATP] conditions (500 μM) are: $k_{ATPb} = 1.1$ μM^{-1}s^{-1}, $k_{ATPr} =50$ s^{-1}, $k_{ADPr} = 50$ s^{-1}, $k_{ATP-tight} = 5000$ s^{-1}; $k_{h,cycle}=10$ s^{-1}. The amount

of ADP inhibition observed in WT motors ([ADP]=500 μM) is well reproduced by the following kinetic rate: $k_{ADPb} = 0.5$ $\mu M^{-1}s^{-1}$. According to equation (5), a reduction in V_{max} must derive from changes in k_{ADPr} and/or $k_{ATP-tight}$. Maintaining other rates equal, the following reduction in kinetic rates reproduced the velocity of F145I rings (65 bp/s): i) $k_{ATP-tight} = 50$ s^{-1} alone or ii) $k_{ADPr} = 23$ s^{-1} alone, and the slowest population in the R146K/WT mixtures (45 bp/s): i) $k_{ATP-tight} = 18$ s^{-1} alone or ii) k_{ADPr} $= 12$ s^{-1} alone. In either case, reducing simultaneously both rates resulted in the same behavior as reducing only k_{ADPr}. According to equation (6), for K_m to remain invariant, a reduction in k_{ADPr} must be accompanied by the proportional change in k_{ATPb}. Thus, the following rate constants reproduced the velocity and K_m values of F145I rings: k_{ADPr}=23 s^{-1} and k_{ATPb}=0.69 μM^{-1}s^{-1}, and of R146K/WT hybrid rings: k_{ADPr}=12 s^{-1} and k_{ATPb}=0.36 μM^{-1}s^{-1}. Table S2 summarizes the results of this simulation.

Appendix: Sequential ATP Hydrolysis and Interlaced Nucleotide Exchange is Preserved at Limiting [ATP]

Packaging trajectories of WT motors at very low ATP ([ATP]=20 μM) in the presence of small amounts of ATPγS (500:1 molecular ratio, i.e. in 100 motor cycles one of the 5 subunit is likely to bind ATPγS) exhibited the distinct pauses previously observed at saturating [ATP] (Chistol et al. 2012). The duration of these pauses is single exponentially-distributed showing that a single event defines the duration of the pause (Figure 3.9A). The pause density increases linearly with [ATPγS] (the highest [ATPγS] tested was 125:1 molecular ratio, meaning that in 25 motor cycles one of the 5 subunits is likely to bind ATPγS), suggesting that a single molecule is enough to halt the motor (Figure 3.9A bottom inset). The pause duration does not change with [ATPγS], consistent with a mechanism in which the residence time of a single ATPγS molecule in the subunit defines the duration of the pause (Figure 3.9A top inset). Previously, it was shown that at saturating [ATP] the hydrolysis events are strictly sequential, and as result, binding of a single ATPγS molecule is enough to interrupt the catalytic phase (Chistol et al. 2012). Our results show that one ATPγS molecule is also enough to halt the motor at limiting [ATP], implying that sequential ATP hydrolysis continues to take place in these limiting conditions. We conclude that the subunits' coordination during the catalytic phase is preserved even when the

ring is not fully saturated with ATP. Additionally, a careful ATP titration at sub K_m [ATP] shows that the motor's velocity is best described by the Hill equation with the coefficient, $n_H=1$ (Figure 3.9B). This result could be interpreted simply as lack of cooperative binding (Copeland 2002). However, we previously showed that a scheme of sequential ATP binding at saturating [ATP] exhibits the binding dynamics of a single enzyme (or $n_H=1$) since at any given time only one pocket is available to bind the nucleotide (Moffitt et al. 2009). Thus, we conclude that ATP binding continues to take place sequentially at low [ATP]. Moreover, competition experiments with ADP show that the mean dwell duration increases linearly as a function of [ADP] (Figure 3.9C). Such linear behavior is in agreement with the competitive ADP inhibition observed at saturating [ATP] (Figure 3.9C inset). Competitive ADP inhibition within a sequential ATP binding scheme is only possible if the ADP release events and the ATP binding events are interlaced—see (Chistol et al. 2012) for alternative scenarios. We conclude that coordination during nucleotide exchange is preserved at limiting [ATP].
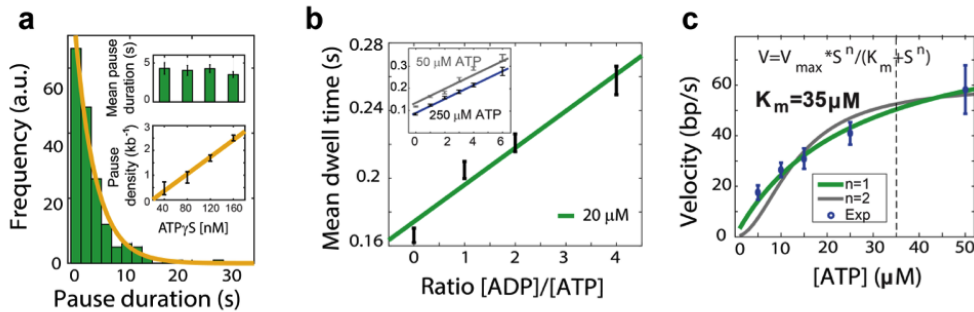
Figure 3.9: **Coordination is Preserved at Limiting [ATP]**
a) Analysis of pauses induced by the non-hydrolyzable ATP analog, ATPγS. [ATP]= 20 μM.
b) ADP inhibition at limiting [ATP]. [ATP]=20 μM.
c) ATP titration below K_m conditions. Packaging velocity as a function of [ATP] fits well to the hill equation with coefficient, $n_H = 1$.

Abbreviated methods

Sample Preparation in Bulk WT and mutant ATPase gp16, genomic DNA and fiberless proheads were isolated as described previously (Mao et al.

2016). To avoid high filling conditions, the 19.3 kb genomic DNA was digested with BstEII (New England Biolabs) to produce a 12.5 kb piece of DNA. The 12.5 kb piece of DNA was biotinylated using Klenow exo- (New England Biolabs) to fill in the overhang with biotinylated nucleotides (Invitrogen). Proheads were partially packaged (30-60s) with biotinylated 12.5 kb genomic DNA and stalled with ATPγS (Roche). The packaging buffer (1/2 x TMS) contained 25 mM Tris-HCl, pH 7.8, 50 mM NaCl, 5 mM MgCl$_2$ with various concentrations of nucleotides and nucleotide analogs as specified in the main text (see chapter A for extended details).

Optical Trapping   High-resolution packaging measurements were conducted on a dual-trap instrument using a solid-state 1064-nm laser, as described previously (Liu et al. 2014b). Traps were calibrated as previously described (Chistol et al. 2012, Moffitt et al. 2009). Tethers were formed between a 0.90-μm-diameter streptavidin-coated bead and a 0.88-μm-diameter anti-capsid-antibody-coated bead (Spherotech) held in separate optical traps. Packaging was restarted in an ATP-containing buffer, and DNA translocation by individual motor complexes was determined from the decrease in the bead-to-bead distance. An oxygen scavenging system (100 μg ml^{-1} glucose oxidase, 5 mg ml^{-1} dextrose (Sigma-Aldrich) and 20 μg ml^{-1} catalase, Calbiochem) was included in the buffer to prevent the formation of reactive singlet oxygen. All packaging experiments were conducted as described before in a semi-passive mode such that tension applied to the motor was kept between 7-12 pN.

Data Analysis   Raw 2.5-kHz data was collected and filtered to 100 - 250 Hz for further analysis. A modified Schwarz Information Criterion (SIC) algorithm was used to find steps in the packaging traces as described previously (Liu et al. 2014b) in a custom written Matlab script. Mean values of the dwell duration distributions were computed by bootstrapping to normalize the distribution and compute confidence intervals (displayed as errors in the figures). Velocity distributions were built as histograms of the motors' velocities and were normalized to compare across different sets of data. The velocity of each motor was determined by fitting a straight line to its packaging trajectory.

Cryo-EM Imaging   The DNA packaging motor was assembled as described in the sample preparation. The reaction was stalled by adding ATPγS after three minutes of initiating packaging with ATP (Roche) (ATP concentration is 500μM). The sample was incubated for 2 minutes. 1 unit of RQ DNase I (Promega) was added and incubated at room temperature for 10 minutes. The sample was placed on ice until grid preparation for cryoEM imaging. A prohead:DNA ratio of 2:1 was used to maximize packaging efficiency. Three minutes packaging corresponds to 75% head-full. Grid preparation and image reconstruction was performed as described previously (Mao et al. 2016).

Chapter 4

Molecular switches and allostery in protein motors

4.1 Summary

Motor proteins are powered by nucleotide hydrolysis and exert mechanical work to carry out many fundamental biological tasks. To ensure their correct and efficient performance, the motors' activities are allosterically regulated by additional factors that enhance or suppress their NTPase activity. Here, I review two highly conserved mechanisms of ATP hydrolysis activation and repression operating in motor proteins—the glutamate switch and the arginine finger—and their associated regulatory factors. I examine the implications of these regulatory mechanisms in proteins that are formed by multiple ATPase subunits. I argue that the regulatory mechanisms employed by motor proteins display features similar to those described in small-GTPases, which require external regulatory elements, such as dissociation inhibitors, exchange factors and activating proteins, to switch the protein's function 'on' and 'off'. Likewise, similar regulatory roles are taken on by the motor's substrate, additional binding factors, and even adjacent subunits in multimeric complexes. However, in motor proteins, more than one regulatory factor, and the two mechanisms described, often underlie the machine's operation. Furthermore, ATPase regulation takes place throughout the motor's cycle, which enables a more complex function than the binary 'active' and 'inactive' states.

4.2 Background

As described in previous chapters, motor proteins are macro-molecular complexes that transform the energy released in ATP hydrolysis into mechanical work to perform fundamental biological processes (Liu et al. 2014, Gennerich & Vale 2009). Although motor proteins are ubiquitous, their function is necessary only at relevant locations in the cell and at distinct stages of its cycle (Verhey & Hammond 2009). Therefore, in order to ensure adequate spatial and temporal control of their function, motor proteins must be switched between 'active' and 'inactive' states in a regulated manner (Schumacher et al. 2004).

The current mechanistic picture of regulation in motor proteins has been partially drawn from the detailed understanding of function activation in small GTPases (Figure 4.1). Small GTPases activate signaling pathways in the GTP-bound state—the 'on' state—and do not display such activation in the GDP-bound state—the 'off' state (Goitre et al. 2014). The active state of small GTPases is not required at all times and, therefore, the protein's activity is switched 'on' and 'off' throughout the cell cycle in a regulated manner (Figure 4.1). Such regulation is performed by specialized protein factors that either prevent or promote ADP release—known as Guanine Dissociation Inhibitors (GDIs) or Guanine Exchange Factors (GEFs), respectively—or that stimulate GTP hydrolysis—known as GTPase-activating proteins (GAPs) (Cherfils & Zeghouf 2013). Likewise, the activities of motor proteins are regulated by additional factors, such as the motor's substrate, associated protein factors, and subunits adjacent to each other in multimeric complexes, which take on regulatory roles analogous to those described in small GTPases (Figure 4.1).

In the light of what I learned of the $\phi29$ DNA packaging motor's operation (see chapter 3), in this chapter, I review various mechanisms of regulation in different motor proteins. I highlight the fact that all these mechanisms share a general feature: the motor's function is controlled by stimulation or repression of its ATPase activity, which is regulated allosterically by different factors. I show that regulatory factors can control the motor's ATPase activity at two distinct stages of the enzymatic cycle: nucleotide exchange or ATP hydrolysis. First, I discuss specific cases in which these factors control nucleotide exchange by preventing or facilitating ADP release. Then, I describe in detail the functions of two highly conserved elements of ATP hy-

drolysis activation: the arginine finger and the glutamate switch. I show how the action of two repressing or two activating elements can combine resulting in a heavily inhibited or heavily activated motor activity. Throughout this chapter, I consider the implications of these mechanisms in multimeric ring motors.
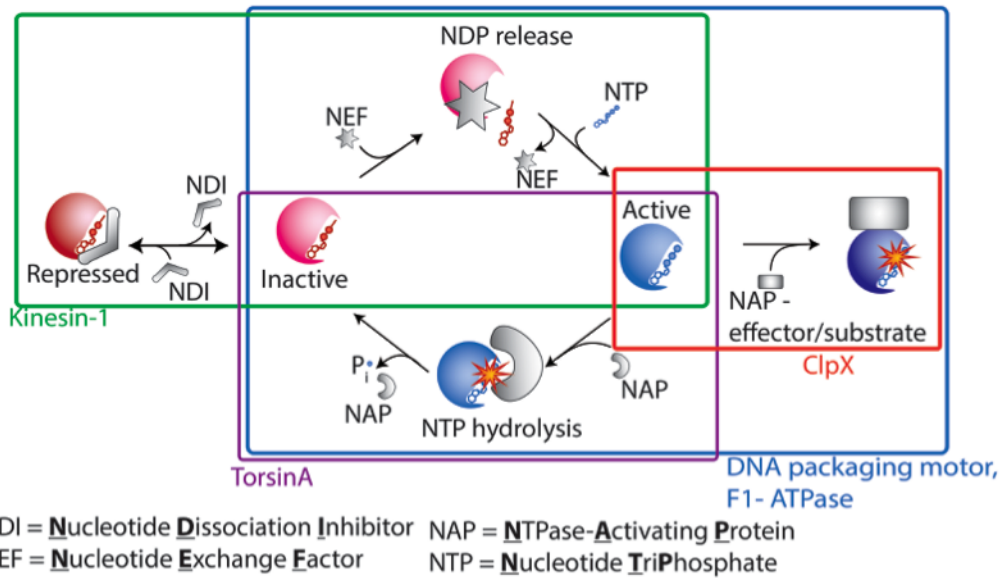
Figure 4.1: **Small GTPase-like Regulation.** The activity of small-GTPases is regulated by additional binding factors. Different motor proteins display distinct features of these regulatory mechanisms. The specific features exhibited by each motor are indicated in boxes with the same color code as the name of the protein.

4.3 Review

Dissociation Inhibitors and Exchange Factors in Motor Proteins

Kinesin-1 is an essential protein that transports vesicles and organelles towards the (+) end of microtubules (Hirokawa et al. 2009). For kinesin to achieve the velocity of $\sim$900 nm s^{-1} observed during transport (Verbrugge

61

et al. 2009), its ATPase cycle —which includes ADP release and other chemical transitions—must take place at a rate of at least $\sim$112.5 s^{-1} (assuming an 8 nm step size). Interestingly, un-loaded kinesin-1 remains mostly in its ADP-bound state, displaying ADP release rates of about 0.00048 s^{-1} (Hackney & Stock 2008), whereas in the presence of its cargo and microtubules, these rates are increased to $\sim$300 s^{-1} (Cross 2004). Kinesin-1 switches between these heavily repressed and transport-competent states through the combined action of two regulatory factors that control its nucleotide exchange: kinesin-1's own tail and microtubules. Indeed, in the absence of cargo, Kinesin-1 remains in the heavily repressed state (Figure 4.2A) by using its tail domain as a built-in inhibition factor that prevents ADP dissociation (Figure 1). Deletion of the tail domain is known to activate the ATPase activity without need of cargo binding, and inhibition is re-established by the addition of exogenous tail peptide (Coy et al. 1999). The crystal structure shows that the tail folds and cross-links the motor domains, preventing the movement of the ATPase domains and, presumably, the allosteric signal necessary to promote ADP release (Kaan et al. 2011). However, even after binding to its cargo, kinesin-1 still binds tightly to ADP (ADP release rates of $\sim$0.005 s^{-1}, Figure 4.2A, (Cheng et al. 1998). Only in the presence of microtubules, cargo-bound-kinesin-1 displays ADP release rates competent for regular transport, indicating that the polymer track functions as an exchange factor that facilitates ADP release (Figure 4.2A). Indeed, crystal and biochemical studies suggest an allosteric path that promotes ADP release in kinesin-1 after binding to microtubules (Atherton et al. 2014). In particular, the interaction between kinesin-1 ATPase head and residue E415 in α-tubulin is thought to produce conformational changes that are propagated to the nucleotide-binding site to promote the release of ADP (Uchimura et al. 2010).

Kinesin-1 depends on its exchange factor to increase its ATPase activity. Instead, in multimeric ring motors, exchange factors are necessary to ensure timely release of ADP at appropriate times during the motor's cycle. The subunits and, in some cases, the subunit's subdomains, can adopt the exchange factor role by promoting ADP release in the neighboring nucleotide binding pocket (Komoriya et al. 2012, Tafoya et al. 2017, Franzmann et al. 2011). Next, we will briefly describe three ring motors that rely on exchange factors for their proper operation.

F1-ATPase, a subdomain of ATP-synthase, is formed by three dimers of α
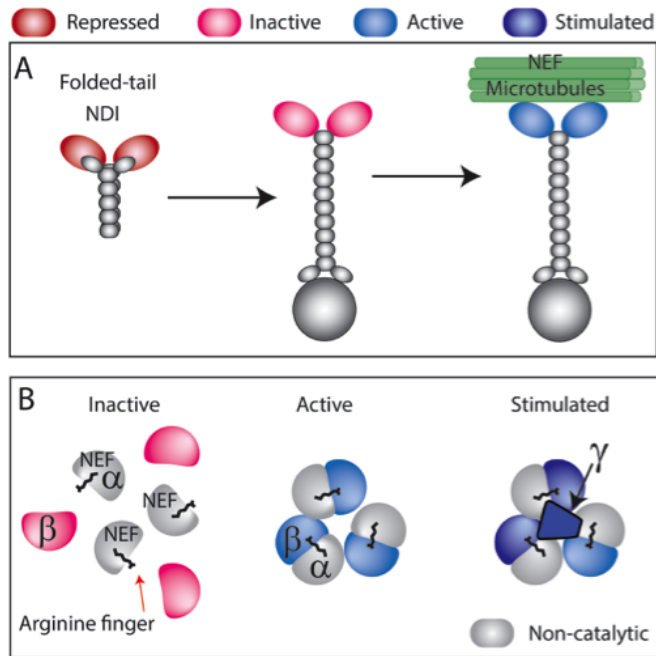
Figure 4.2: **Multiple Mechanisms of Regulation Operate in a Single Motor Protein.**
a) The tail of Kinesin-1 serves as a built-in dissociation inhibitor (left). Tail-inhibition is released upon cargo-binding (middle). The microtubules play the role of an exchange factor by accelerating ADP release (right).
b) In F1-ATPase the α-subunits serve the role of the exchange factor. In solution, the beta subunits do not release ADP (left). In the ring configuration, the β-subunits are released from the ADP-inhibited by the arginine finger of the α-subunit (middle). The activity of the ring motor is further enhanced by the presence of the γ-subunit.

and β-subunits ($\alpha 3\beta 3$), and one copy of the γ-subunit, a central protein that rotates as the motor's chemical cycle proceeds (Noji et al. 1997). The catalytic site at the interface of each α-β dimer is formed by the nucleotide binding pocket of the β-subunit, and several residues contributed by the non-catalytic α-subunit, including a highly conserved trans-acting arginine residue, α-R373 (Abrahams et al. 1994) (Figure 4.2B). It was shown that ATP binding to the non-catalytic α-subunit site is anti-cooperative and promotes the release of ADP from the adjacent β-subunit (Amano et al. 1999). Interestingly, in motors containing mutants of the highly conserved arginine finger, α-R373, the β-subunits is known to remain in the ADP-inhibited

state (Komoriya et al. 2012). Additionally, replacing the arginine finger for an unnatural amino-acid analog of lysine, Lyk—which has the same length as the arginine—prevented the β-subunit ADP-inhibited state (Yukawa et al. 2015), indicating that the length of the trans-acting side chain is sufficient for nucleotide exchange. These observations suggest that the ATP-bound α-subunit uses its arginine finger to facilitate ADP release from its neighboring β-subunit (Figure 4.2B).

A surprisingly similar mechanism seems to underlie the $\phi29$ DNA packaging motor. The $\phi29$ DNA packaging motor is a pentameric ring ATPase that encapsidates one copy of the viral genome into a pre-formed protein shell. During motor's operation, the subunits are known to undergo ADP-ATP exchange one at a time. ATP binding in one subunit induces the release of ADP in the neighboring catalytic pocket (Moffitt et al. 2009). ADP release is the slowest process during nucleotide exchange and, thus, the subunits tend to stay bound to the nucleotide unless they are induced to release it (Chistol et al. 2012). The single-molecule experiments that I conducted show that mutants of the highly conserved arginine finger, R146K, display even slower ADP release rates, demonstrating that the arginine residue is necessary for normal ADP release (Tafoya et al. 2017). Based on the published structure (Mao et al. 2016), the catalytic site and the arginine finger are located at opposite sides of each subunits but are connected through a direct peptide linkage (Figure 4.3B). These observations are consistent with a mechanism in which ATP binding in one subunit triggers an allosteric signal that is propagated via this linkage to its arginine finger that, in turn, facilitates the release of ADP in the adjacent catalytic pocket.

The previous examples suggest a conserved allosteric process connecting ATP binding in one subunit to ADP release in the adjacent binding pocket via the arginine finger. Although experimental evidence about the details of this allosteric pathway is still needed, communication between adjacent binding pockets is also thought to underlie nucleotide exchange in other ring ATPases. An interesting variation of this mechanism is seen in the dissagregase Hsp104, a hexameric ring ATPase from yeast that targets protein aggregates to resolve them (Franzmann et al. 2011). The subunits of Hsp104 contain two binding domains: NBD1 and NBD2. A mutagenesis study shows that ATP binding to each protomer, NBD2 or NBD1, increases ATP turnover by almost 10 fold in the other. ADP release is known to be the rate-limiting step for the hydrolysis reaction, and thus, the observed increase in ATPase rate is thought to result

from facilitating ADP release in one binding domain upon binding of ATP by the other (Franzmann et al. 2011). Although, the allosteric regulation has been shown to occur in both directions, NBD2 is thought to be the main engine of Hsp104, whereas, NBD1 is thought to play more of a regulatory role.

The examples presented above show that nucleotide factors operate in many motor proteins to regulate their ATPase activity. Mechanisms for nucleotide exchange often operate in combination with factors that modulate ATP hydrolysis rates. In what follows I will review some of these activating factors and their associated mechanisms to stimulate or suppress ATP hydrolysis.

ATPase Modulation Via the Arginine Finger

Most ASCE proteins reach high steady state ATPase activity only in their oligomeric form (Wong et al. 1996, Adzuma & Mizuuchi 1991). In contrast, as monomers, the subunits bind nucleotide with low affinity and support low rates of ATPase hydrolysis, preventing futile energy consumption. The high ATPase activity observed after oligomerization results from catalytic pocket sites at the interface of each pair of adjacent subunits (Wendler et al. 2012)(Figure 4.3). ATP hydrolysis rates in the catalytic sites are enhanced by trans-acting arginine residues that stabilize the transition state (Figure 4A). In heteromeric ring motors, ATP hydrolysis in the catalytic subunits—containing both, the nucleotide binding pocket and catalytic residues —are activated by the non-catalytic subunits—contributing the arginine residues (Figure 4.3A). The catalytic and non-catalytic subunits have different sequences, and often differ significantly in their tertiary structures. In homomeric ring motors, ATP hydrolysis in the catalytic sites is activated by arginine residues provided by identical neighboring subunits (Figure 4.3B). In each subunit, the nucleotide binding pocket and catalytic residues are located at the opposite side of the arginine finger. Below I review several heteromeric and homomeric ring ATPases where subunits take on the role of activating factors to stimulate ATP hydrolysis in the neighboring catalytic pocket. We will also discuss allosteric processes that enable sequential subunit stimulation.
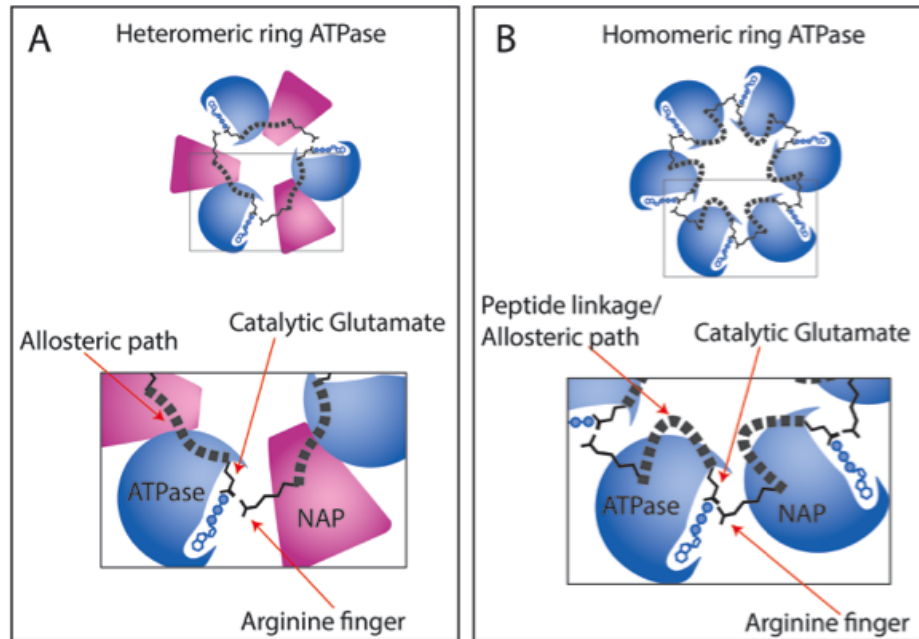
Figure 4.3: **Catalytic Pockets in Heteromeric and Homomeric Ring Motors.**
a) In heteromeric ring motors, two different proteins form the catalytic pocket: the ATPase subunit (blue) and an activating protein with no intrinsic ATPase activity (pink). Signals are transmitted by the activating protein to the adjacent catalytic pocket through allosteric pathways (grey dashed line).
b) In homomeric ring motors, the catalytic pocket is formed by two identical subunits. Signals are transmitted to the next catalytic pocket through peptide linkages connecting the nucleotide-binding site and the arginine finger (grey dashed line).

The Arginine Finger in Heteromeric Rings Motors

The $\alpha3\beta3$ ring complex of F1-ATPase exhibits its maximum ATPase activity in the presence of the γ-subunit (Figure 4.2B), but also displays significant ATPase activity alone ($\sim$25% of the full F1-ATPase)(Digel et al. 1996). The $\alpha3\beta3$ intrinsic ATPase activity is supported by three catalytic sites formed at the interface of each α-β dimer (Figure 4.2B). Early biochemical studies showed that catalytic β-subunits alone do not display significant ATPase activity and require the presence of the non-catalytic α-subunits to attain detectable enzymatic activity (Dunn & Futai 1980). Moreover, mutants of the highly conserved arginine finger were shown to decrease the ATPase activity

by a factor of 10^3 (Komoriya et al. 2012). Although some of this reduction derives from larger ADP-inhibition of the β-subunit (see previous section), single molecule studies have shown that the transition most affected by the substitution is ATP hydrolysis (Komoriya et al. 2012). These studies, together with molecular dynamic simulations, show that the α-subunit plays an activating role by contributing its trans-acting arginine to the catalytic site (Figure 4.3A and 4A). Interestingly, recent studies based on high-speed atomic force microscopy show that rotor-less $\alpha3\beta3$ complexes exhibit ATP-induced dynamic conformational changes and binding asymmetry similar to those observed during the operation of the full F1-ATPase complex (Uchi-hashi et al. 2011). This observation indicates that the α and β subunits are fully competent to propagate signals between catalytic sites without the concourse of the γ subunit (Figure 4.3A). The fact that mutants of the arginine finger abolish multi-site catalysis (Soga et al. 1989), suggested that the α-subunit, via its arginine finger, plays an important role in transmitting signals between catalytic sites. Isolated β-subunits are known to undergo large conformational changes upon nucleotide binding (Yagi et al. 2009) that closely resemble the open-to-closed conformation transition. In contrast, isolated α-subunits exhibit much more limited conformational changes upon ligand binding in bulk studies (Shirakihara et al. 1997). However, molecular dynamic simulations show that the open-to-closed conformational transition in α-subunits is essentially barrierless (Hahn-Herrera et al. 2016), implying that the non-catalytic subunit could easily mirror the β-subunit's conformational changes upon induction. Moreover, mutagenesis studies have identified a set of residues in the vicinity of the arginine finger in the α-subunit that are necessary for nucleotide binding cooperativity and multisite hydrolysis but that have no effect on unisite catalysis—the hydrolysis rate when only one out of the three catalytic sites is occupied with nucleotide (Soga et al. 1989). Based on the crystal structure of the full F1-ATPase complex, this set of residues is thought to amplify conformational changes occurring in the arginine finger during the ATP hydrolysis transition state (Weber & Senior 1997). Although more experimental evidence is required, the inherent flexibility of the α-subunit suggests an allosteric path that allows long-range communication between β-subunit catalytic pockets (Figure 4.3A).

β-subunits of F1-ATPase pair exclusively with α-subunits to form complete catalytic pockets. By contrast, other ATPase proteins can pair with multiple activating proteins, a strategy that possibly evolved to regulate the function
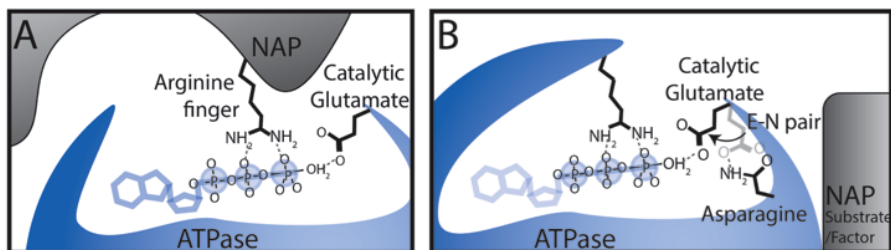
Figure 4.4: **Molecular Mechanisms for ATP Hydrolysis Activation.**
a) An arginine finger interacts with the β and γ phosphates to stabilize the transition state during the water nucleophilic attack.
b) The catalytic Glutamate pairs with an Asparagine when the protein is in the inactive state. The residue is released from the pair by external activating factors to activate ATP hydrolysis.
NAP= NTPase-Activating Protein.

of the motor by different factors at distinct locations in the cell. TorsinA, a member of the AAA+ branch of NTPases, related to proteases and Hsp proteins, illustrates this behavior (Chase et al. 2017). TorsinA is found in the endoplasmic reticulum and nuclear envelope of higher eukaryotes, and does not display ATPase activity in isolation. However, ATP hydrolysis by TorsinA subunits is induced upon association with LAP1 at the nuclear envelope, or with LULL1 at the endoplasmic reticulum. Both activating proteins lack intrinsic ATPase activity. EM studies show that TorsinA and LAP1 or LULL1 subunits form alternating hetero-hexameric rings in which the activators donate an arginine finger to complete TorsinA's ATPase active site (Figure 4.3A) (McCullough & Sundquist 2014). Many questions remain to be answered about this system, including whether LAP1 and LULL1 can mediate long-range communication between TorsinA subunits in a manner similar to α-subunits in F1-ATPase. Nonetheless, the ATPase activation observed in TorsinA raises the possibility that other proteins are also activated by multiple activating factors, each forming a different heteromeric ring depending on their location in the cell.

The Arginine Finger in Homomeric Ring Motors

The mechanism for ATP hydrolysis activation in homomeric ring motors (Figure 4.3B) has been well characterized in a model ring ATPase, the ϕ29 DNA packaging motor. The operation of this motor is known to be segregated in two phases while the subunits display high degree of coordination (Chistol et al. 2012); in the first phase all 5 subunits in the ring sequentially exchange ADP for ATP. Then, saturation of the ring with ATP is a signal that activates the first hydrolysis event. This event in turn initiates a cascade of hydrolysis by the remaining subunits (Chistol et al. 2012). The single-molecule experiments that I conducted show that the observed high degree of coordination results from the ability of the subunits to switch between spontaneous ('poor') and stimulated ('efficient') ATPase activity during the motor' s operation (Tafoya et al. 2017). Throughout the nucleotide exchange stage, the subunits display their basal low ATPase activity, allowing all subunits to bind nucleotide. During the catalytic phase, ATP hydrolysis at the catalytic sites is activated by the neighboring subunit via the trans-acting arginine finger, R146 (Figure 4A). According to the recently published crystal structure of the monomer (Mao et al. 2016), the conserved catalytic Glutamate, E119, is connected with the putative DNA binding loop which prolongs to the Arginine finger, R146 (Figure 4.3B). Thus, an allosteric mechanism can be envisioned for the activation of the first ATP hydrolysis by the last ATP binding event and for the following sequential ATP hydrolysis; the last ATP binding event is sensed by residues in the catalytic pocket. These residues propagate a conformational change that repositions the arginine finger in the next catalytic pocket to stimulate ATP hydrolysis. ATP hydrolysis in that subunit induces the repositioning of its arginine finger to stimulate ATP hydrolysis in the next catalytic pocket. By repetition of this process, ATP hydrolysis stimulation is propagated sequentially around the ring. Many ring ATPases are known to display similar sequential coordination of ATP hydrolysis, such as Rho transcription factor and the chaperonin CCT/TRiC (Thomsen et al. 2016, Gruber et al. 2017). The mechanism described here for the ϕ29 DNA packaging motor possibly underlies the operation of similar molecular machines.

As seen above, ATP hydrolysis stimulation via the arginine finger is a highly conserved mechanism employed by many ring ATPases (Figure 4A). Additionally, the ATPase activities of most ring motors are repressed or further

stimulated upon binding other elements, such as the motor's substrate or additional protein factors. In the next section, we will describe a possible mechanism underlying the operation of these activating/repressing factors.

ATPase Modulation via the 'Glutamate-switch' Mechanism

In multimeric ring motors, ATPase activation and repression upon binding additional elements—such as the motor's substrate or additional protein factors—has been widely reported (Davies et al. 2014, Besprozvannaya et al. 2013, Roy et al. 2012). While these additional elements are often thought as ligands that regulate the motors' activities, the molecular mechanism that couples ligand binding to the change in protein's ATPase activity is still not well understood.

A comparative structural study of multiple motor proteins belonging to the AAA+ superfamily of the ASCE division (Zhang & Wigley 2008), provides important insights into this ligand-mediated regulation. This study includes the structural maps of 50 different proteins (including F1-ATPase, PspF, ORC1, HslU, RFC and SV40 among others) found in different states, such as ADP-bound, ATP-bound, and in the presence or absence of their regulatory ligand. In the absence of the regulatory ligand, the catalytic Glutamate side chain was seen to be well positioned to interact with the missing γ phosphate in the ADP-bound state, but systematically rotated away from this position by approximately 100° in the ATP bound state. In the rotated configuration, the Glutamate forms a hydrogen bond with another residue located in the vicinity, typically an Asparagine (Figure 4B). The systematic formation of the Glutamate-Asparagine (E-N) pair suggests an explanation for the weak ATPase activity exhibited by many proteins in the absence of their regulatory ligand: the state of the protein corresponds to the orientation of the catalytic Glutamate. In other words, the protein is inactive when the E-N pair is formed and active when the Glutamate is released from it. This study also shows that, in most cases, there is a direct peptide linkage between the E-N pair and the regulatory ligand-binding region. This linkage suggests an allosteric mechanism that couples the binding event to changes in the catalytic Glutamate's orientation. In the proposed mechanism, the regulatory ligand binding induces the release of the catalytic Glutamate from the E-N

pair to stimulate—or promotes the formation of the E-N pair to repress—the protein's ATPase activity (Figure 4B). Additional structural maps of proteins in the presence of their regulatory ligands are still required. However, mutagenesis studies show that the substitution of the Asparagine required to form the E-N pair renders the protein irresponsive to the presence of the regulatory ligand (Joly et al. 2008). Although further experimental evidence is needed, the Glutamate-switch model suggests a compelling mechanism for the activation and repression of the motors' ATPase activities by additional binding factors. In the following section, we will consider the implications of this activating/repressing mechanism in multimeric ring motors.

Asymmetric ATPase Activity in Ring Motors

In multimeric ring motors, the individual subunits must coordinate their operation to perform a single biological task (Abbondanzieri & Zhuang 2009). However, during motor' s operation, the subunits might not bind simultaneously, all at once, to the substrate or associated factors. Thus, only those subunits that bind to the motor' s regulatory ligand will be stimulated or repressed, resulting in symmetry breaking of the ATPase activity around the ring complex. Moreover, as the motor's cycle proceeds, the subunits-regulatory ligand interactions change, and the ATPase activity of the individual subunits are expected to change accordingly. Here, we discuss two ring motors that display asymmetric ATPase activity: the ϕ29 DNA packaging motor and the protease ClpXP.

The ϕ29 DNA packaging motor is known to translocate 10 bp of DNA per cycle in a burst of 4 power strokes, each 2.5 bp in size, interspersed by dwell times of about 80 ms on average at saturating [ATP] (Moffitt et al. 2009). Because the ϕ29 packaging motor is a homo-pentamer, the 4 power strokes reveal a crucial symmetry breaking in its operation: only 4 out of the 5 subunits perform DNA translocation. It has been proposed that the remaining subunit performs a regulatory function in the ring. In fact, temporary inactivation of the fifth subunit with ATPγS results in greatly lengthened dwell times (Moffitt et al. 2009) followed by 10 base pair burst, suggesting that proper turnover of ATP by the fifth subunit is necessary to initiate the translocation cascade by the other four subunits in a timely manner. Because the motor is known to contact two DNA phosphates every 10 bp, it was proposed that the observed functional difference between otherwise iden-

tical subunits results from the periodic contact of the fifth subunit with the motor' s substrate (Liu et al. 2014b). The single-molecule experiments that I conducted combined targeted mutagenesis and cryo-EM reconstruction, and provided further insights into the mechanism that enables division of labor among identical subunits (Tafoya et al. 2017). In my study, a substitution that abolishes the subunit's ATPase activity is shown to be tolerated in the ring only if it inactivates the regulatory subunit. Such mutant motors also display greatly lengthened dwell times separated by exactly 10 bp bursts. From these experiments, it was possible to extract the spontaneous ATP hydrolysis rates of the translocating subunits and that of the regulatory subunit. The data showed that the regulatory subunit has a 3-fold increase in ATP hydrolysis rate relative to its translocating counterparts. The asymmetric cryo-EM reconstruction shows that, indeed, only one of the five subunits establishes extensive contacts with the DNA prior to translocation. These observations are consistent with a mechanism in which the periodic DNA contact enhances the regulatory subunits' ATP hydrolysis rates. In agreement with this interpretation, previous bulk measurements have shown that DNA stimulates the subunits' ATPase activities in solution (Todd et al. 2012). Moreover, the pseudo-atomic structure of the motor's subunit (Mao et al. 2016) shows that the putative DNA binding loop is adjacent to the catalytic Glutamate, E119, which provides a direct peptide linkage that connects DNA contact to the stimulation of the subunit' s ATPase activity. As discussed in the previous section, in the $\phi29$ DNA packaging motor, the first ATP hydrolysis (by the regulatory subunit) takes place quickly after saturation of the ring with ATP. Thus, the data suggest that the DNA-bound regulatory subunit is primed to be further activated by the last ATP binding event. Moreover, the different hydrolysis rates suggest that other rates could be different in the subunit contacting the DNA. In particular, it has been speculated that ADP release takes place much faster at the regulatory subunit than in any other subunit to initiate nucleotide exchange, again activated by the binding of that subunit to the regulatory ligand. Thus, functional symmetry breaking between the subunits of the $\phi29$ DNA packaging motor, provide a mechanism by which one of the subunits triggers, in turn, first the beginning of the nucleotide exchange and then the beginning of the hydrolysis cascade.

Similar mechanisms possibly operate in the hexameric protease ClpX. ClpX recognizes ssrA-tagged proteins, unfolds them, and feeds the polypeptide

through its central pore into the proteolytic chamber, ClpP, for degradation (Baker & Sauer 2012). The ATPase activities of ClpX subunits are known to be highly asymmetrical and consist of two classes defined by the orientation between the proteins' subdomains: ATP-unloadable (U) and ATP-loadable (L) (Olivares et al. 2016). Among the loadable type, some sites release ATP rapidly, whereas others release ATP slowly. The mechanism that produces such asymmetry is not well understood. However, the ATPase activities of ClpX subunits are known to be repressed in the presence of the proteolytic chamber, ClpP, and stimulated by ssrA-tagged substrates (Baker & Sauer 2012). Two types of ClpX lumenal loops are important for this ATPase regulation (Martin et al. 2008). The first type, known as pore-1 or GYVG loops, are located in the middle of the pore and propel substrates forward along the ring channel through hydrophobic interactions. Pore-1 loops are known to influence the subunits' ATP-hydrolysis rates, although the mechanism of ATPase modulation is not yet clear. The second type, known as pore-2 loops, are heavily populated with charged residues and are located at the interface between the ring ATPase and the proteolytic chamber, ClpP. Pore-2 loops extend directly from the Walker B motif, which harbors the catalytic Glutamate, E185, involved in ATP hydrolysis (Glynn et al. 2009), and thus, are thought to translate motor's interactions into enhancement or suppression of ATP hydrolysis by controlling the configuration of the glutamate residue at the catalytic site. Indeed, pore-2 loops are known to be necessary for the repression of the ATPase activity by ClpP (Joshi et al. 2004). Furthermore, pore-1 and pore-2 loops form a helical inner surface running parallel to each other Glynn et al. (2009), and thus, it is conceivable that the motor's substrate serves as an allosteric platform that supports communication between the two types of loops. Given the chemical and conformational heterogeneity of the unfolded polypeptide chain, it is possible that the asymmetric ATPase activity observed in ClpX stems from the highly irregular and continuously changing substrate-subunit interactions.

4.4   Concluding Remarks

As shown throughout this chapter, different motors have adopted distinct allosteric processes to regulate distinct stages of their enzymatic cycle, specifically nucleotide exchange and ATP hydrolysis. This regulation takes place ei-

ther through selective interactions with activators, repressors, and nucleotide exchange factors or, as observed in ring ATPases, through interactions with adjacent partners. Such regulatory interactions recall the control of activity seen in small GTPases. The mechanisms reviewed above employ highly conserved structural elements and, thus, likely underlie the operation of many other motor proteins. It is fascinating to verify how, through discrete intermolecular interactions a global, deterministic, machine-like behavior emerges from purely stochastic molecular events.

Chapter 5

Optimal protocols for non-equilibrium processes in biology

5.1 Summary

Macromolecular complexes that transform chemical energy into mechanical work dissipate energy as they perform their biological function. Thus, their energetic efficiency is determined by the underlying non-equilibrium processes that they conduct. Non-equilibrium theory is underdeveloped, but recent work has approximated the excess work of processes out of equilibrium. This theory offers minimum dissipation protocols that maximize energetic efficiency based on equilibrium behavior. In this chapter, I test the predictions of this theory with folding/unfolding experiments of DNA hairpins. I find that the theoretically predicted minimum-dissipation protocols indeed require significantly less work than naive ones across a wide span of driving velocities. These experiments validate a simple method to predict

minimum-dissipation protocols, promising insights into biological design and experimental practice.

5.2   Background

A fundamental problem in modern thermodynamics is how a motor protein performs mechanical work, while operating away from thermal equilibrium without excessive energy dissipation. Recently developed theoretical frameworks have enable the derivation of a generalized friction tensor that predict minimum-dissipation protocols (Sivak & Crooks 2012). Experiments *in silico* have shown that the friction coefficient of the driving protocols is sharply peaked at the interface between metastable regions, which leads to minimum-dissipation protocols that drive rapidly within a metastable basin, but then linger longest at the interface, giving thermal fluctuations maximal time to kick the system over the barrier (Sivak & Crooks 2016). Thus, the theory provides a design principle for the construction of thermodynamically efficient protocols. However, experimental tests of these predictions were still missing.

In this chapter, I present the first test of these theoretical predictions using the mechanical unfolding and refolding of single DNA hairpins. This system is ideally suited for this test, as their sequence can be varied to tune the separation, free energy difference, free energy barrier, and transition rates between folded and unfolded states (Woodside et al. 2006). The experiments confirm the predictions of the theory: my colleagues and I showed that it is possible to design unfolding and refolding protocols that minimize the amount of energy dissipated during the extension/relaxation cycle. In these minimum-dissipation protocols, molecules are seen to unfold and refold earlier—and hence, closer to equilibrium—than in the naive equivalents. We showed that the difference in energy dissipated between minimum-dissipation and naive protocols and the breakdown of the approximate theory, change with the protocol duration according to the theory.

Theory

Near equilibrium, the power dissipated by a system taken from an initial to a final state according to a protocol Λ is determined by a generalized friction coefficient ζ (Sivak & Crooks 2012):

$$\langle P_{\mathrm{ex}}(t)\rangle_\Lambda \approx \zeta(\lambda)\dot{\lambda}^2. \tag{5.1}$$

ζ can be computed from the time integral of the force autocorrelation function $\langle \delta f(0)\,\delta f(t)\rangle_\lambda$, which can be decomposed into the force variance $\langle \delta f^2\rangle_\lambda$ and force relaxation time $\tau_{\mathrm{relax}}(\lambda) \equiv \int_0^\infty \langle \delta f(0)\,\delta f(t)\rangle_\lambda \mathrm{d}t / \langle \delta f^2\rangle_\lambda$:

$$\zeta(\lambda) \equiv \beta \int_0^\infty \langle \delta f(0)\,\delta f(t)\rangle_\lambda \mathrm{d}t \tag{5.2a}$$

$$= \beta \langle \delta f^2\rangle_\lambda \tau_{\mathrm{relax}}(\lambda) \, , \tag{5.2b}$$

where $f(t)$ is the force at time t. It can be shown (Sivak & Crooks 2012) that minimum-dissipation driving protocols (for both forward and reverse protocols) proceed with a velocity $\dot{\lambda}_{\mathrm{MD}} \propto \zeta(\lambda)^{-1/2}$, proportional to the inverse square root of the friction coefficient ζ. The proportionality is fixed by the time taken to perform the protocol, so changing the protocol time corresponds to a global rescaling of all velocities. In this approximation, excess work scales inversely with protocol time, $W_{\mathrm{ex}} \propto \tau^{-1}$.

5.3   Experimental Results

Equilibrium sampling

We first monitored the force fluctuations at various fixed extensions of the optical trap. This equilibrium sampling (Figure 5.1a) gives results expected for a bistable system: far from the hopping regime, for very short or very long extension (Figure 5.1b), the force fluctuates around a single mean value (left and right panels). For intermediate extension (Figure 5.1b), the force fluctuates between two different values, corresponding to the open and closed DNA hairpin conformations (middel panel). We calculated the force autocorrelation function $\langle \delta F(0)\delta F(t)\rangle$ (Figure 5.1c); as expected, in the hopping

regime the force variance is larger and fluctuations decay slower. More quantitatively, from this autocorrelation function we calculated the force variance (Figure 5.2d) and the force relaxation time (Figure 5.2e), for each fixed trap position. The force variance peaks at an intermediate extension, which corresponds to the trap extension at which the hairpin splits roughly equal time between folded and unfolded states (henceforth denoted $X_{1/2}$). Likewise, the force relaxation time peaks at the same $X_{1/2}$ intermediate extension. This maximum reflects the fact that the hairpin must relax across the barrier which separates folded and unfolded states in order to equilibrate. The friction coefficient (Figure 5.2f), which has been shown to be the product of the those two terms (see section *Theory* in this chapter), also peaks at $X_{1/2}$.

The theory indicates that the minimum-dissipation velocity scales as the inverse square root of the friction coefficient (see section *Theory* in this
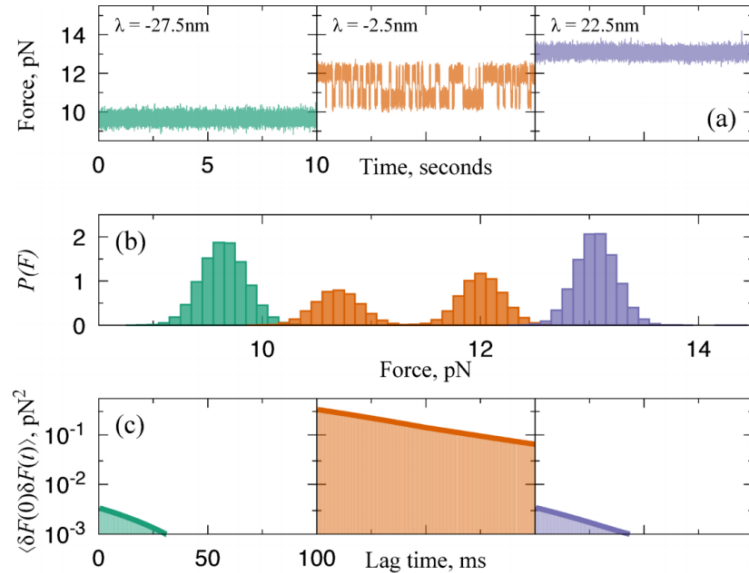

Figure 5.1: **Equilibrium sampling reveals the friction coefficient peaks strongly at the hopping regime.**
a) Sample force traces as function of time, for folded hairpin (left, green), hopping hairpin (middle, orange), and unfolded hairpin (right, purple).
b) Distributions of equilibrium force distributions for corresponding optical trap extensions (same color code as in **a**).
c) Force correlation as a function of time (same color code as in **a**).
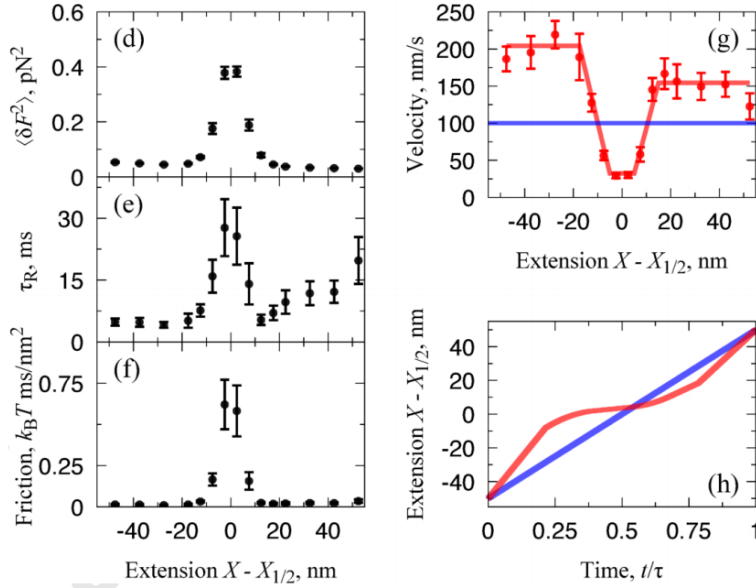
78

Figure 5.2: **Equilibrium sampling reveals the friction coefficient peaks strongly at the hopping regime.**
a) The friction coefficient $\zeta(X)$ (bottom) as a function of fixed optical trap extension beyond the extension $X_{1/2}$, and its factors: the force relaxation time $\tau_{\text{relax}}(X)$ (middle) and the force variance $\langle \delta f^2 \rangle_X$ (top). **b)** The minimum-dissipation velocity $\mathrm{d}X/\mathrm{d}t \propto \zeta^{-1/2}$ (red points) with best fit of model that minimizes Akaike information criterion (red curve), compared to naive velocity (blue line). This sample protocol duration is 1 second.
c) Minimum-dissipation (red) and naive (blue) protocols as function of protocol time t/τ.

chapter). Thus, we sought to determine an approximation of the minimum-dissipation velocity as a function of distance to $X_{1/2}$ that can be easily implemented experimentally in the optical trap instrument. To this end, we fitted a simple piecewise constant acceleration protocol (Figure 5.2g). The resulting minimum-dissipation protocols differ substantially from naive ones, which correspond to constant velocity protocols completing in the same elapsed time (Figure 5.2h). In particular, minimun-dissipation protocols proceed more rapidly than a constant-velocity protocol when the hairpin only populates either the folded or unfolded states, and more slowly in the hopping regime. Intuitively, a slow velocity around $X_{1/2}$ provides more time for thermal fluctuations to induce the opening or closing of the DNA hairpin without requiring additional work input, and thus, is expected to decrease the amount of work required to drive the DNA hairpin to either conformation (Sivak &

Crooks 2016).

Nonequilibrium protocols

We next proceeded to determine the force-extension curves from minimum-dissipation and naive protocols with durations of 4, 2, 1, 0.5, 0.25 and 0.125 s. The force extension curves display significant differences depending on which protocol type was implemented during the unfolding and refolding of the DNA hairpin (Figure 5.3a and b).
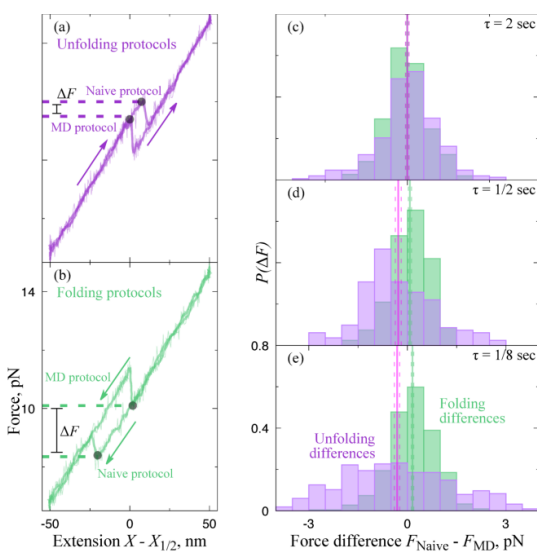

Figure 5.3: **Minimum-dissipation protocols consistently unfold at lower force and refold at higher force.**
(a,b) Sample force-extension curves from a sample molecule for protocol time $\tau = 1/8$ second, highlighting the unfolding/folding event (black dots) and the corresponding forces (dashed lines). (a) The minimum-dissipation protocol typically unfolds at a lower force than the corresponding naive protocol. (b) The minimum-dissipation protocol typically folds at a higher force than the corresponding naive protocol.
(c-e) Distributions of differences between naive and minimum-dissipation forces for unfolding (purple) and folding (green) protocols, with means and standard errors (solid and dashed lines, respectively).

As expected theoretically, when unfolding was performed using the minimum-dissipation protocols, the DNA hairpin unfolded, on average, at lower forces

than when the same process was performed using the naive protocols (Figure 5.3a). Similarly, the DNA hairpin refolded, on average, at higher force in experiments using the minimum-dissipation protocols than when using the naive ones (Figure 5.3b). Taken together, unfolding and refolding trajectories, the results indicate a lower hysteresis associated to the minimum-dissipation protocols (Figure 5.4a and b).

To quantify this results, we compared the force at which the DNA hairpin unfolds/refolds using the minimum-dissipation protocols versus the naive ones in multiple realizations and for many molecules. These results are quantified in Figure 5.3c, d and e, where the distribution of unfolding force differences $f_{\text{naive}}^{\text{u}} - f_{\text{opt}}^{\text{u}}$ has an average above zero and the distribution of refolding force differences $f_{\text{naive}}^{\text{f}} - f_{\text{opt}}^{\text{f}}$ has an average below zero. As the protocols speed up, the system is driven farther away from equilibrium, and thus, is expected to display higher hysteresis. Consistently, the absolute mean value of the force difference distributions increase as the protocols speed up (from top to bottom), indicating that minimum-dissipation protocols have less hysteresis associated to them than the naive ones.
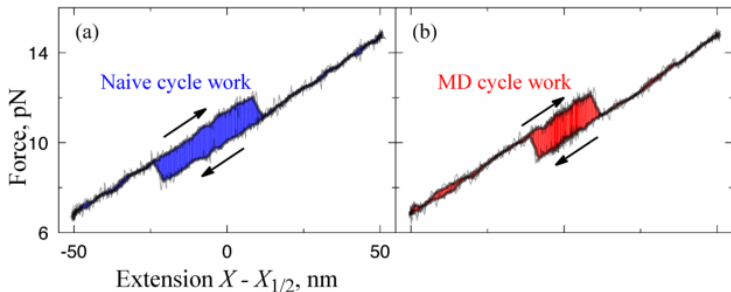
Figure 5.4: **Hysteresis in force-extension curves for minimum-dissipation and naive protocols.**
(a,b) Example force-extension curves showing, shaded in blue (a) and red (b), the cycle work from the sum of works to unfold and fold for naive (a) and minimum-dissipation (b) protocols.

In agreement with those observations, during an unfolding/folding cycle, the minimum-dissipation cycle work (hysteresis) is significantly less than the naive cycle work (Figure 5.4a and b). The cycle work means and variances increase as the protocols speed up (top to bottom), indicating higher hys-

teresis. The difference between the work cycle using the naive and minimum-dissipation (Figure 5.5f, g and h) shows that as protocol time decreases (and hence driving velocity increases), both minimum-dissipation and naive protocol ensembles get further from equilibrium. Moreover, as driving velocity increases, the difference increases between the mean works required for minimum-dissipation and for naive protocols.
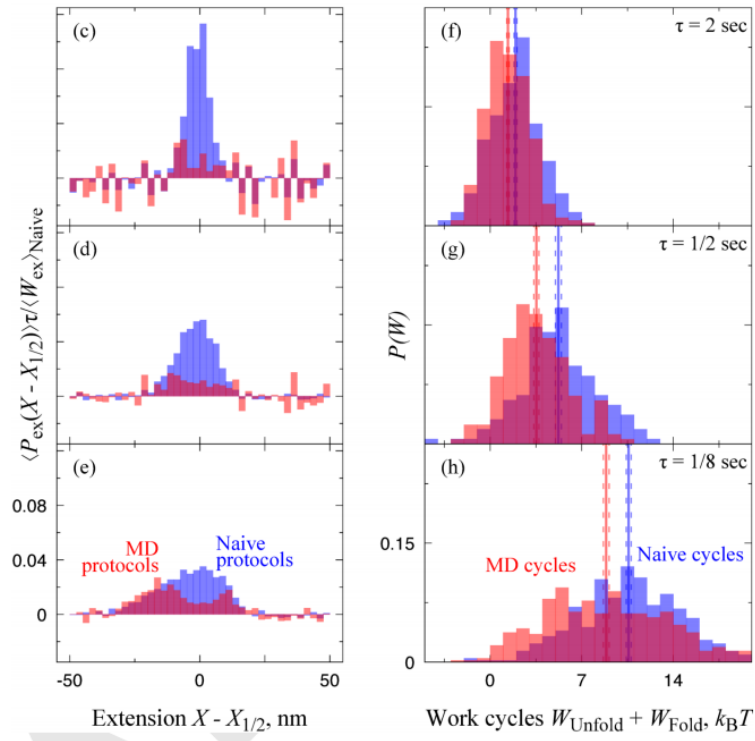
Figure 5.5: **Minimum-dissipation protocols consistently require less work than corresponding naive protocols, and produce less cycle hysteresis.**
(c-e) Excess work per protocol interval, for naive (blue) and MD (red) protocols.
(f-h) Distributions of cycle works $W_{\text{Unfold}} + W_{\text{Fold}}$, for naive (blue) and minimum-dissipation (red) protocols, for protocols ranging from slow (top) to fast (bottom), with means and standard errors (solid and dashed lines, respectively).

5.4 Discussion

Equilibrium sampling of force fluctuations in DNA hairpins confirms that these molecules display the dynamics of a two-state system: they remain folded for a short optical trap extension, unfolded for a long extension, and hopping in between the two states at an intermediate extension (Figure 5.1a, and b). According to linear response theory, these equilibrium fluctuations imply energetic costs due to the fast driving of the system between the two states. In other words, the force fluctuations measured at equilibrium (various fixed positions of the traps) lead to a prediction in near-equilibrium conditions: the greatest non-equilibrium cost accumulates as the system is driven across the hopping regime. We confirmed these predictions with non-equilibrium experiments by rapidly unfolding and refolding DNA hairpins (Figure 5.5c, d and e). Therefore, we conclude that linear response theory determines a driving schedule (which we have implemented with different velocities) that minimizes the work required for a fixed-time protocol based on equilibrium fluctuations, and predicts accurately that the excess work in minimum-dissipation protocols scales linearly for different fixed protocol times.

In our experiments, we explored driving protocols whose average velocity vary by a factor of 32. In all cases, the presumed minimum-dissipation protocols are indeed more energetically efficient than the naive ones that proceed at a constant pulling speed (Figure 5.5f, g and h). Thus, this near-equilibrium theory gives meaningful and quantitative guidance to the dynamical out-of-equilibrium thermodynamics of biomolecular systems. These results have immediate applications in the streamlining of single-molecule experiments: the more dissipation in a given single-molecule protocol, the more repetitions needed for a given precision of free-energy difference inference (Gore et al. 2003). Thus initial sampling of equilibrium fluctuations could be used to craft nonequilibrium protocols that dissipate significantly less energy, and hence achieve significantly greater precision.

The agreement of theory and experiment suggest obvious follow-ups in more complex contexts. For example, ATP synthase is known to be remarkably efficient at stall forces; one could estimate the minimum-dissipation protocol to drive rotation of F1 at 100 Hz (ATP synthase natural rate) and measure
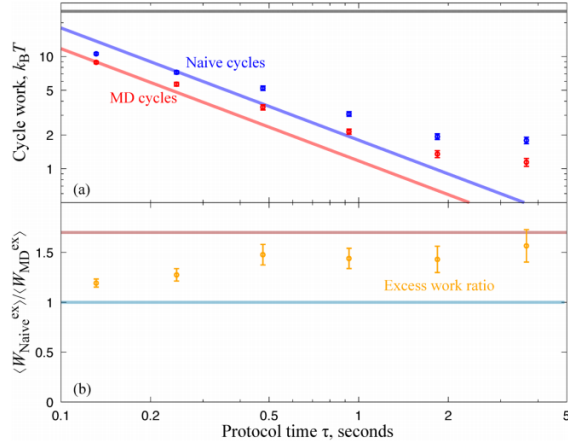
Figure 5.6: **a)** Scaling with protocol time of cycle works for naive (red) and minimum-dissipation (blue) protocols with empirical power-law fits (dashed lines), compared to theoretical near-equilibrium predictions (solid lines).
b) Scaling with protocol time of cycle work ratio $\langle W_{\text{naive}}^U + W_{\text{naive}}^F \rangle / \langle W_{\text{MD}}^U + W_{\text{MD}}^F \rangle$, compared to theoretical near-equilibrium prediction with no free parameters (brown dashed line).

the minimal excess work (and the savings compared with naive driving). In this sense, the excess work required by the minimum-dissipation protocol at a given speed provides a scale for judging the non-equilibrium performance of molecular machines that must turnover on that timescale. However, molecular machines do not implement deterministic driving protocols, and thus, further extensions of this theory are already quantifying the costs associated with a stochastic protocol/ensemble of protocols (Large & Sivak 2017). In the long run, insights from this framework promise to guide the design of synthetic machines.

5.5 Methods

Basic optical trap setup

High-resolution force-extension measurements were conducted on a dual-trap instrument using a solid-state 1064-nm laser, as described previously. Moffitt et al. (2009) Traps were calibrated as previously described. Chis-

tol et al. (2012). DNA tethers were formed between a 0.90-μm-diameter streptavidin-coated bead and a 1-μm-diameter anti-digoxigenine-coated bead (Spherotech) held in separate optical traps (see Chapter A). An oxygen scavenging system (100 μg ml^{-1} glucose oxidase, 5 mg ml^{-1} dextrose (Sigma-Aldrich) and 20 μg ml^{-1} catalase, Calbiochem) was included in the buffer to prevent the formation of reactive singlet oxygen. Including an oxygen scavenging system increased the lifetime of DNA tethers (see Chapter A).

DNA molecules

Hairpin DNA sequences were selected to display hopping dynamics such that determining $X_{1/2}$ was accessible experimentally—very fast hopping dynamics were difficult to distinguish from noise, and very slow dynamics required long periods of data acquisition and laser exposure prior to pulling experiments. Minimizing laser exposure avoids molecule photo-damage. All data in Supplementary Material is from sequence 1:GAGTCCTGGATCCT-GTTTTTTTTCAGGATCCAGGACTC, was previously characterized and exhibited appropriate hopping dynamics ($t_{1/2} \approx 0.24$ s Woodside et al. (2006)). All data in the main text is from sequence 2: TACCTGATCAGGTGC TTTTTTTTTGCACCTGATCAGGTA, resulting from modifying sequence 1 to increase high GC-content at the loop neck. This change in sequence is expected to facilitate nucleation of the native conformation and to avoid molecule mis-folding Li et al. (2007).

Equilibrium sampling

Each molecule is initially probed to find $X_{1/2}$: the distance between the traps is increased gradually until the residence time at open and closed conformations is $\sim 50\%$. The global shift in reported extension is due to variation in XXX. For each molecule, each extension is sampled for 30 seconds, in order from smallest ($X_{1/2} - 50$ nm) to largest extension ($X_{1/2} + 50$ nm), at 10 nm spacing far from $X_{1/2}$ and 5 nm spacing near $X_{1/2}$ to more precisely resolve the friction variation at the hopping regime. Changes in extension are instructed to be performed instantaneously but are limited by the response of the mirror controlling the steering trap (~ 2 ms). Equilibrium force fluctuations at each of several fixed extensions were measured independently in

each of 20 different molecules. From these fluctuations the generalized friction coefficient was estimated using (5.2a). At each extension, we jackknife resampled from the set of 20 friction estimates to calculate the mean generalized friction and standard error Wasserman (2004).

We fit several piecewise-constant-acceleration profiles of protocol velocity to the minimum-dissipation one $\frac{d\lambda}{dt}_{\text{MD}} \propto [\zeta(\lambda)]^{-1/2}$ predicted from the empirically determined generalized friction $\zeta(\lambda)$. Each model velocity profile has constant velocity (zero acceleration) far away from $X_{1/2}$ and in the immediate vicinity of $X_{1/2}$. Constant-acceleration regions interpolate between these constant-velocity regions. The model parameters are the region boundaries and the constant velocities. Different model velocity profiles impose different symmetries, such as inversion symmetry about $X_{1/2}$, thus reducing the number of free parameters. We used the velocity profile (Fig. 5.1) that minimized the Akaike Information Criterion Akaike (1998).

Naive and minimum-dissipation protocols

We estimate the work for a given trajectory by numerically integrating the raw force-extension curve of each non-equilibrium protocol. We require that the contribution to the total work from the (unknown) free energy difference between the initial and final states (at $\lambda = X_{1/2} \pm 50$nm) is zero. The cycle work (hysteresis) $W_{\text{Cycle}} = W_{\text{Fw}} + W_{\text{Rev}} = W_{\text{ex,Fw}} + W_{\text{ex,Rev}}$ sums the forward and reverse realizations of a protocol. The work difference $W_{\text{Diff}} = W_{\text{Naive}} - W_{\text{MD}} = W_{\text{ex,Naive}} - W_{\text{ex,MD}}$ subtracts the work during a minimum-dissipation protocol from the work during a naive protocol for either the forward or reverse direction. We investigate 6 different protocol times, ranging from 1/8 second to 4 seconds. For each protocol time, we calculate the work along ~ 1200 individual realizations, ~ 300 of each of the 4 protocol types: minimum-dissipation/naive forward/reverse. To theoretically predict the average excess work at a given protocol time, we numerically evaluate Equation (5.2b) using the empirically determined friction coefficients and the applied minimum-dissipation or naive protocols. To estimate the unfolding (refolding) force in a given force-extension curve, we first smooth the force trace using a second-order Savitsky-Golay filter with window width $\sim 0.4ms$. We report the unfolding (refolding) force as the maximum (minimum) force before the final unfolding (refolding) event takes

place. We control for intermolecular variation by analyzing the difference between unfolding/refolding forces along naive and minimum-dissipation protocols, instead of raw unfolding/refolding forces.

Appendix A

DNA packaging motor protocols

In this chapter, we describe protocols for the single-molecule DNA packaging assay with the $\phi 29$ motor, which includes bulk activity assessment, microfluidic chamber construction, sample preparation, instrument operation, data acquisition, and data analysis. It is worth noting that the protocols for the $\phi 29$ system have been adapted to study other viral packaging motors and ring ATPases, yielding many interesting features of motor dynamics and providing a broad panorama of the diversity of operation of these important cellular machines [9].

A.1   Bulk DNA Packaging Assay

The *in vitro* DNA packaging activity of the motor is evaluated by measuring the amount of packaged DNA inside the phage capsid that is resistant to DNase digestion [10].

Materials:

1. Viral components: $\phi 29$ proheads and ATPase gp16 (store in small aliquots at -80°C), $\phi 29$ genomic DNA with terminal gp3 protein (DNA-gp3, store at 4°C). Purified as described in [10]. See Note1.

2. 0.025-μm membrane filter (Millipore).

3. 1 M Tris-HCl (pH 7.8).

4. 0.5× TMS buffer: 25 mM Tris-HCl (pH 7.8), 50 mM NaCl, and 5 mM MgCl$_2$.

5. 100 mM ATP. Store at -20°C.

6. DNase I (Calbiochem).

7. 0.5 M EDTA.

8. Proteinase K (New England Biolabs).

Methods:

1. DNA-gp3, isolated from phage-infected *B. Subtilis* cells or purified phages, is dialyzed on a 0.025-μm membrane filter against 10 mM Tris-HCl (pH 7.8) for 45 min.

2. ϕ29 proheads (1×10^{11} copies) are mixed with DNA-gp3 (5×10^{10} copies) and gp16 [$(1.2 - 1.5) \times 10^{12}$ copies] in $0.5 \times$ TMS buffer in a total volume of 18 μL. The mixture is incubated for 5 min at room temperature.

3. Add 2 μL of 5 mM ATP and incubate the mixture for 15 min at room temperature.

4. Unpackaged DNA is digested by adding DNase I to 1 μg/mL. Incubate for 10 min at room temperature.

5. To deactivate DNase I and release the packaged DNA from viral capsids, the mixture is treated with 25 mM EDTA (final concentration) and 500 μg/mL Proteinase K (final concentration) for 30 min at 65 C.

6. DNA packaging efficiency is evaluated by running a 1% agarose gel.

A.2 DNA Preparation for Single-Molecule Packaging Assay

The $\phi29$ genomic DNA is 19.3-kbp in length, with one copy of the terminal protein gp3 covalently bound to each 5' terminus. To systematically investigate the effect of capsid filling level on the motor's packaging behavior, we use DNA substrates of various lengths in the single-molecule packaging assay (Table A.1) [8].

Materials:

1. DNA-gp3 (see above).

2. 1 M Tris-HCl (pH 7.8).

3. Selected restriction enzyme ClaI, XbaI, BstEII, or NcoI with its respective 10× buffer (New England Biolabs).

4. Klenow fragment exo- (New England Biolabs).

5. Biotinylated deoxyribonucleotides (Invitrogen).

6. T4 DNA ligase and 10× T4 ligase buffer (New England Biolabs).

7. 100 mM ATP (Roche)

8. PCR thermocycler.

Methods:

1. After dialysis in 10 mM Tris-HCl, DNA-gp3 is digested with one selected restriction enzyme (Table A.1). Use 1 unit of enzyme to digest 1 μg of DNA-gp3 for 1 h. Choose the optimal buffer and temperature according to the manufacturer's protocol.

2. The 5' overhang from the restriction cut is filled in with biotinylated nucleotides using the Klenow fragment of DNA polymerase I (exo- mutant). Use 1 unit of Klenow fragment and 100 pmol of nucleotides for every 1 μg of DNA-gp3. Set the reaction at 37°C for 30 min, then 75°C for 15 min to deactivate the enzyme.

3. Dialyze the solution on a 0.025-μm membrane filter against 10 mM Tris-HCl (pH 7.8) for 45 min. Store the DNA substrate at 4°C. See Note3.

4. To generate DNA substrates longer than the ϕ29 genome length, a biotinylated DNA piece is ligated to the enzyme-digested DNA-gp3. For example, to create a 21-kbp DNA substrate, first generate a 6-kbp DNA piece that is PCR amplified from lambda DNA using a biotinylated primer and cut with NcoI; then ligate it to NcoI-cut DNA-gp3. Use 5 molar excess of 6-kbp DNA to DNA-gp3. Use New England Biolabs' standard T4 DNA ligase protocol.

5. After the ligation reaction, dialyze the mixture in 10 mM Tris-HCl (pH 7.8) for 45 min. Store the product at 4°C.

Capsid filling level	DNA length	Restriction enzyme used	Remaining overhang
32	6147	ClaI	5' CG
46	8929	XbaI	5' CTAG
65	12,466	BstEII	5' GTCAC
78	15,023	NcoI	5' CATG

Table A.1: Summary of the different DNA lengths used in the single-molecule packaging experiments

A.3 Bead Preparation

In the single-molecule experiment, a prohead-ATPase-DNA complex is tethered between two beads held in two laser traps (Fig. 1a). The viral capsid is attached to a bead coated with antibodies against the major capsid protein gp8. The biotinylated distal end of the DNA substrate is attached to a streptavidin-coated bead.

Materials:

1. 1× phosphate buffered saline (PBS buffer).

2. 0.5× TMS buffer (see above).

3. Polystyrene beads stock solution: 1% (w/v) 0.88-μm Protein G-coated beads, 1% (w/v) 0.90-μm streptavidin-coated beads (Spherotech). Store at 4°C.

4. Vortexer.

5. Rotator.

6. Anti-phage antibodies stock solution (1 mg/mL) (produced by the Jardine and Grimes Lab, University of Minnesota). Store in 20 μL aliquots at -80°C.

Methods:

Preparing Antibody-Coated Beads:

1. Pipette 40 μL of 1% (w/v) 0.88-μm Protein G-coated beads in a 1.5-mL microcentrifuge tube.

2. Add 1× PBS buffer to a total volume of 200 μL.

3. Resuspend the beads by vortexing the solution on high speed for 30 s. Spin the beads down at 10,000×g for 2 min in a benchtop centrifuge.

4. Remove the supernatant.

5. Repeat steps 2-4 twice.

6. Resuspend the pellet in 30 μL of 1× PBS buffer.

7. Add 20 μL of anti-phage antibodies (1 mg/mL; purified from rabbit antisera prepared against ϕ29 proheads). Gently tumble the mixture for 4 h in a tube rotator at room temperature.

8. Wash the beads by repeating steps 2-4 three times.

9. Resuspend the beads in 60 μL of 1× PBS buffer. Store at 4°C.

Preparing Streptavidin-Coated Beads:

1. Pipette 30 μL of 1% (w/v) 0.90-μm streptavidin-coated beads in a 1.5-mL microcentrifuge tube.

2. Add 0.5× TMS buffer to a total volume of 200 μL.

3. Resuspend the beads by vortexing the solution on high speed for 30 s. Spin the beads down at $10,000\times$g for 2 min in a benchtop centrifuge.

4. Remove the supernatant.

5. Repeat steps 2-4 twice.

6. Resuspend the pellet in 60 μL of 0.5× TMS buffer. Store at 4°C.

A.4 Microfluidic Chamber Construction

The design of the microfluidic chamber is shown in Fig. A.1.

Materials:

1. Cover glass (VWR, size #1, 24 × 60 mm).

2. Nescofilm (Karlan).

3. Laser cutter (Universal Laser Systems).

4. Glass dispenser tube (King Precision Glass, 0.1-mm diameter).

5. Heat block (100°C).

6. PE20 polyethylene tubing (BD Intramedic).

Methods:

1. Drill six holes on a cover glass using a laser cutter.

2. Make a three-channel pattern on a piece of Nescofilm using a laser cutter.
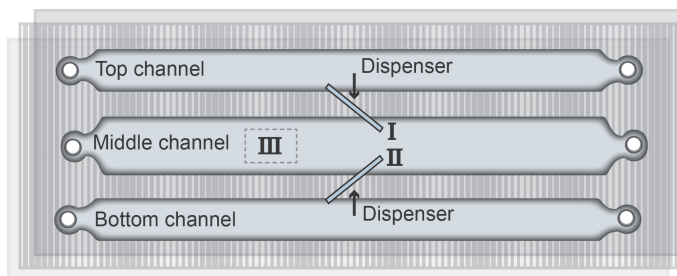
Figure A.1: Microfluidic chamber design. The solutions flow from left to right

3. Lay the patterned Nescofilm on a second cover glass. Use two glass dispenser tubes to connect the channels. Then put the drilled cover glass on top of the Necsofilm.

4. Put the chamber on a 100°C heat block for 30 s. Gently press the chamber to seal the two cover glasses. Inspect for any air bubbles.

5. Mount the chamber onto a metal frame. Assemble the inlet/outlet tubings. Then place the chamber between the two objectives of the optical tweezers instrument.

6. Wash the channels with 1 mL of water and then 1 mL of 0.5× TMS buffer before each experiment.

A.5 Single-Molecule Packaging Assay

DNA packaging is initiated in a 1.5-mL microcentrifuge tube by feeding DNA substrates to reconstituted prohead/ATPase complexes in the presence of ATP. Packaging is then stalled by adding the non-hydrolyzable analog ATPγS. The stalled packaging complexes are delivered to the microfluidic chamber and restarted in an ATP-containing solution. See Note4.

Materials:

1. 10× TMS buffer: 500 mM Tris-HCl (pH 7.8), 1 M NaCl, and 100 mM MgCl2.

2. BSA (20 mg/mL) (New England Biolabs).

3. RNaseOUT (40 units/μL) (Invitrogen).

4. 100 mM ATP (Roche)

5. 100 mM ATPγS (Roche).

6. Oxygen-scavenging system: 100 μg/mL glucose oxidase, 20 μg/mL catalase, and 5 mg/mL dextrose (Sigma-Aldrich).

7. ApaLI (10 units/μL) (New England Biolabs).

8. 1-mL syringes (BD).

9. Needles (BD PrecisionGlide, 26 G×½ in.).

10. High-resolution dual-trap optical tweezers. See Note2.

Methods:

Bead Passivation

1. Add 2 μL of stock streptavidin-coated or antibody-coated beads and 1 μL of 20 mg/mL BSA to 20 μL of 0.5× TMS buffer.

2. Vortex at high speed for 45 min at room temperature. Then put the beads on ice.

3. Unused passivated beads are stored at 4°C. Vortex again before using them the next day.

Assembling Stalled Packaging Complexes

1. Add in order: 4.5 μL of H2O, 1 μL of 10× TMS buffer, 0.5 μL of RNaseOUT, 2 μL of biotinylated DNA-gp3 (2.5×10^{10} copies), and 4 μL of proheads (1×10^{11} copies). Mix gently.

2. Add 4 μL of gp16 (2.5×10^{12} copies). Mix gently. Incubate the mixture for 5 min at room temperature.

3. Packaging is initiated by adding 2 μL of 5 mM ATP. Mix well and incubate for 30 s.

4. Packaging is stalled by adding 2 μL of 5 mM ATPγS. Mix well.

5. The stalled complexes are stored on ice and must be used within the same day. See Note5.

Making Solutions for the Three Channels of the Microfluidic Chamber

1. Top channel solution. 4 μL of passivated streptavidin-coated beads are diluted in 1 mL of 0.5× TMS buffer.

2. Middle channel solution. 1 mL of 0.5× TMS buffer supplemented with the oxygen scavenger system (to prevent the formation of reactive singlet oxygen that would damage the tether) and ATP. A typical saturating ATP concentration is 500 μM.

3. Bottom channel solution. Mix in order: 10 μL of 0.5× TMS buffer, 2 μL of 5 mM ATP, 2 μL of 5 mM ATPγS, 0.5 μL of RNaseOUT, 0.5 μL of ApaLI, 4 μL of passivated antibody-coated beads, and 1 μL of stalled complexes. Incubate for 20 min at room temperature. Then dilute the mixture in 1 mL of 0.5× TMS buffer containing 100 μM ATP and 100 μM ATPγS. See Note6.

Forming Tethers and Recording Packaging Trajectories

1. Transfer the solutions above from 1.5-mL microcentrifuge tubes to 1-mL syringes. Connect the syringes to the inlet tubings of the microfluidic chamber via 26 G×½ needles.

2. Push 50 μL of the top channel solution into the top channel. Streptavidin-coated beads are delivered to Position I of the middle channel through the dispenser tube (Fig. A.1). Catch a single bead in Trap A (Fig. 1a, left).

3. Push 50 μL of the bottom channel solution into the bottom channel. Antibody-coated beads, which are conjugated to stalled complexes, are

delivered to Position II of the middle channel through the dispenser tube. Catch a single bead in Trap B (Fig. 1a, right).

4. Bring Trap A and Trap B close to each other, while quickly moving them to Position III (within 10 s). See Note7.

5. Move the two traps apart. If the force reading increases as the traps are being separated, it is an indication that a tether has formed. See Note8.

6. Start recording the positions of the two beads and the trap separation at 2.5-kHz bandwidth. See Note9.

7. The packaging experiment is typically conducted in a semi-passive mode, in which the distance between the two optical traps is adjusted periodically so that the force applied to the motor is kept within a small range (e.g., 7-12 pN). See Note10.

A.6 Data Analysis

Materials:

1. Custom LabView software.

2. Custom MATLAB software.

Methods:

1. Trap stiffness and detector response are calibrated by fitting a modified Lorentzian to the fluctuation power spectrum of a trapped bead [11, 12].

2. The optical force F is determined by $F = \kappa d$, where κ is the trap stiffness, and d is the displacement of the bead from the trap center achieved by back focal plane interferometry at the position-sensitive photodetectors.

3. The extension of the tether is calculated by subtracting the bead displacements and the bead radii from the trap separation.

4. The DNA tether's contour length is calculated from the measured force and tether extension using the worm-like chain model of double-stranded DNA elasticity [13], using a persistence length of 53 nm and a stretch modulus of 1200 pN. Length in nm is then converted to base pairs (bp) using an average B-form DNA rise of 0.34 nm/bp.

5. Raw 2.5-kHz data are filtered to 100-250 Hz for further analysis. A modified Schwarz Information Criterion (SIC) algorithm is used to find steps in the packaging traces (Fig. 3). See Notes11 and 12.

A.7 Notes

1. ϕ29-like phages have a unique and essential RNA component, known as the prohead RNA (pRNA), in their packaging motor complexes. Some experiments involve the usage of truncated or mutated versions of pRNA. In these cases, purified proheads are first treated with RNase to remove the wild-type pRNA. These RNA-free proheads are then incubated with fresh pRNA molecules prior to use.

2. The single-molecule packaging experiments are conducted on a home-built high-resolution dual-trap optical tweezers instrument. Detailed information on the concept, design, and use of this instrument can be found in [14].

3. Despite the fact that two biotinylated DNA-gp3 species are generated by this procedure, it was shown that the left end of the ϕ29 genome is preferentially packaged into the prohead (pRNA likely plays a key role in such selection) [15]. Therefore, it is not necessary to separate these two species before mixing them with the proheads in a single-molecule packaging experiment.

4. Packaging can also be initiated in situ without prepackaging and stalling in the tube [6, 16]. In this case, the biotinylated DNA is bound to a streptavidin-coated bead, and the prohead/ATPase complex is attached to an antibody-coated bead. Packaging is then initiated by bringing the two beads into close contact in the presence of ATP. This procedure allows for the detection of very early stages of packaging.

99

5. The quality of the stalled complexes is essential for the outcome of the single-molecule packaging experiment. Once prepared, the stalled complexes can be used for the entire day. However, we notice that the efficiency of forming active tethers slowly drops with time, perhaps due to residual packaging and/or disassembly of the stalled complexes in the tube. Thus it is advised to prepare a fresh stalled complex sample every 4-5 h.

6. The ApaLI cutting site is located near the left end of the $\phi29$ genome (214 bp from the left terminus). DNA is protected from ApaLI cleavage if packaging is properly initiated. Therefore we add ApaLI to the mixture in order to avoid tethering with inactive prohead/ATPase/DNA complexes that did not initiate packaging.

7. Recording of DNA packaging activity is performed at Position III, away from the dispenser tubes opposite to the direction of flow. This is to avoid accidentally capturing additional beads into the traps during data collection. This region also has reduced buffer turbulence, which helps lower data noise.

8. The likelihood of forming a tether is dependent on the density of stalled complexes on the bead. Too high a density causes multiple tethers between the bead pair, whereas too low a density makes experiments time-consuming. We empirically adjust the ratio of bead concentration to stalled complex concentration, such that on average one tether forms every three to four bead pairs. Under this condition most tethers are single tethers, which is desired.

9. During data recording, we sample the voltages proportional to the position of the light centroid in x and y directions at the two position-sensitive photodetectors, and the voltage proportional to the total amount of light at each detector. We also record the voltages proportional to the horizontal and vertical angle of the piezo mirror that controls the position of the steerable trap. These eight voltage signals are acquired at a rate of 500 kHz, or 62.5 kHz per channel. They are then averaged to 2.5-kHz bandwidth before saving.

10. After the packaging process has completed, the tether is intentionally broken by applying a high force (~30 pN). Two additional calibration files are collected with the same bead pair. First, the positions

100

of the two beads are recorded at 62.5 kHz to determine their fluctuation power spectra. Second, the two beads (untethered) are slowly brought together and the voltage signal as a function of trap separation is recorded. Residual force calculated from this baseline signal is subtracted from the force measured during packaging to correct for the interference pattern between the two traps.

11. The SIC algorithm is an iterative procedure that fits a series of steps to the data and assesses the fit quality for every round of fitting. In the original algorithm [17], the quality of the fit is determined via the formula: $SIC(j_1, \ldots, j_k) = (k+2)log(n) + nlog\left(\hat{\sigma}^2_{j_1,\ldots,j_k}\right)$, where n is the number of data points, k is the number of steps, and $\hat{\sigma}^2_{j_1,\ldots,j_k}$ is the maximum likelihood estimator of variance when k steps are fitted to the data. We find that the original SIC algorithm over-fits experimental data containing colored noise (Fig. 3). We therefore introduce an additional penalty factor (PF): $SIC(j_1, \ldots, j_k) = PF(k+2)log(n) + nlog\left(\hat{\sigma}^2_{j_1,\ldots,j_k}\right)$ [7]. Optimal stepwise fits can usually be achieved using PF values of 3-5. Steps assigned by this method are validated using a residence time histogram analysis (Fig. 3).

12. Optimal resolution of an optical tweezers measurement is achieved when the DNA tether length is less than 2 kbp, since longer tethers are intrinsically noisier. Thus, in order to probe the stepping behavior of the motor at different capsid filling levels, DNA substrates of various lengths are used (Table 1).

Bibliography

Aathavan, K., Politzer, A. T., Kaplan, A., Moffitt, J. R., Chemla, Y. R., Grimes, S., Jardine, P. J., Anderson, D. L. & Bustamante, C. (2009), 'Substrate interactions and promiscuity in a viral DNA packaging motor', *Nature* **461**(7264), 669–673.

Abbondanzieri, E. A. & Zhuang, X. (2009), 'Molecular biology: Concealed enzyme coordination', *Nature* **457**(7228), 392–393.

Abrahams, J. P., Leslie, A. G., Lutter, R. & Walker, J. E. (1994), 'Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria', *Nature* **370**(6491), 621–628.

Adzuma, K. & Mizuuchi, K. (1991), 'Steady-state kinetic analysis of ATP hydrolysis by the B protein of bacteriophage mu. Involvement of protein oligomerization in the ATPase cycle', *J. Biol. Chem.* **266**(10), 6159–6167.

Ahmadian, M. R., Stege, P., Scheffzek, K. & Wittinghofer, A. (1997), 'Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras', *Nat. Struct. Biol.* **4**(9), 686–689.

Akaike, H. (1998), *A New Look At The Statistical Model Identification*, Springer New York.

Alberts, B. (2002), *Molecular Biology of the Cell*, Garland Science.

Amano, T., Matsui, T., Muneyuki, E., Noji, H., Hara, K., Yoshida, M. & Hisabori, T. (1999), 'alpha3beta3gamma complex of F1-ATPase from thermophilic Bacillus PS3 can maintain steady-state ATP hydrolysis activity depending on the number of non-catalytic sites', *Biochem. J.* **343 Pt 1**, 135–138.

Anderson, D. L., Hickman, D. D. & Reilly, B. E. (1966), 'Structure of Bacillus subtilis bacteriophage phi 29 and the length of phi 29 deoxyribonucleic acid', *J. Bacteriol.* **91**(5), 2081–2089.

Antes, I., Chandler, D., Wang, H. & Oster, G. (2003), 'The unbinding of ATP from F1-ATPase', *Biophys. J.* **85**(2), 695–706.

Atherton, J., Farabella, I., Yu, I. M., Rosenfeld, S. S., Houdusse, A., Topf, M. & Moores, C. A. (2014), 'Conserved mechanisms of microtubule-stimulated ADP release, ATP binding, and force generation in transport kinesins', *Elife* **3**, e03680.

Baker, T. A. & Sauer, R. T. (2012), 'ClpXP, an ATP-powered unfolding and protein-degradation machine', *Biochim. Biophys. Acta* **1823**(1), 15–28.

Berndsen, Z. T., Keller, N., Grimes, S., Jardine, P. J. & Smith, D. E. (2014), 'Nonequilibrium dynamics and ultraslow relaxation of confined DNA during viral packaging', *Proc. Natl. Acad. Sci. U.S.A.* **111**(23), 8345–8350.

Besprozvannaya, M., Pivorunas, V. L., Feldman, Z. & Burton, B. M. (2013), 'SpoIIIE protein achieves directional DNA translocation through allosteric regulation of ATPase activity by an accessory domain', *J. Biol. Chem.* **288**(40), 28962–28974.

Bonanca, M. V. & Deffner, S. (2014), 'Optimal driving of isothermal processes close to equilibrium', *J Chem Phys* **140**(24), 244119.

Brown, A. I. & Sivak, D. A. (2017), Toward the design principles of molecular machines. arXiv:1701.04868v2.

Bustamante, C. J., Kaiser, C. M., Maillard, R. A., Goldman, D. H. & Wilson, C. A. (2014), 'Mechanisms of cellular proteostasis: insights from single-molecule approaches', *Annu Rev Biophys* **43**, 119–140.

Bustamante, C. & Tafoya, S. (2017), 'Biochemistry and Biophysics in singulo: When Less is More.', *Proc. Pont. Acad. Sci. Vat. (In press)* .

Callen, H. (2014), *Thermodynamics and Introduction to Thermastatistics*, Wiley.

Casjens, S. R. (2011), 'The DNA-packaging nanomotor of tailed bacteriophages', *Nat. Rev. Microbiol.* **9**(9), 647–657.

Chase, A. R., Laudermilch, E. & Schlieker, C. (2017), 'Torsin ATPases: Harnessing Dynamic Instability for Function', *Front Mol Biosci* **4**, 29.

Chemla, Y. R., Aathavan, K., Michaelis, J., Grimes, S., Jardine, P. J., Anderson, D. L. & Bustamante, C. (2005), 'Mechanism of force generation of a viral DNA packaging motor', *Cell* **122**(5), 683–692.

Chemla, Y. R. & Smith, D. E. (2012), 'Single-molecule studies of viral DNA packaging', *Adv. Exp. Med. Biol.* **726**, 549–584.

Chen, B., Sysoeva, T. A., Chowdhury, S., Guo, L., De Carlo, S., Hanson, J. A., Yang, H. & Nixon, B. T. (2010), 'Engagement of arginine finger to ATP triggers large conformational changes in NtrC1 AAA+ ATPase for remodeling bacterial RNA polymerase', *Structure* **18**(11), 1420–1430.

Cheng, J. Q., Jiang, W. & Hackney, D. D. (1998), 'Interaction of mant-adenosine nucleotides and magnesium with kinesin', *Biochemistry* **37**(15), 5288–5295.

Cherfils, J. & Zeghouf, M. (2013), 'Regulation of small GTPases by GEFs, GAPs, and GDIs', *Physiol. Rev.* **93**(1), 269–309.

Chistol, G., Liu, S., Hetherington, C. L., Moffitt, J. R., Grimes, S., Jardine, P. J. & Bustamante, C. (2012), 'High degree of coordination and division of labor among subunits in a homomeric ring ATPase', *Cell* **151**(5), 1017–1028.

Copeland, R. (2002), 'Cooperativity in Enzyme Catalysis', *Enzymes* pp. 367–384.

Coy, D. L., Hancock, W. O., Wagenbach, M. & Howard, J. (1999), 'Kinesin's tail domain is an inhibitory regulator of the motor domain', *Nat. Cell Biol.* **1**(5), 288–292.

Cross, R. A. (2004), 'The kinetic mechanism of kinesin', *Trends Biochem. Sci.* **29**(6), 301–309.

Davies, B. A., Norgan, A. P., Payne, J. A., Schulz, M. E., Nichols, M. D., Tan, J. A., Xu, Z. & Katzmann, D. J. (2014), 'Vps4 stimulatory element of the cofactor Vta1 contacts the ATPase Vps4 7 and 9 to stimulate ATP hydrolysis', *J. Biol. Chem.* **289**(41), 28707–28718.

Digel, J. G., Hightower, K. E. & McCarty, R. E. (1996), 'Subunit movement during catalysis by F1-F0-ATP synthases', *J. Bioenerg. Biomembr.* **28**(5), 439–442.

Ding, F., Lu, C., Zhao, W., Rajashankar, K. R., Anderson, D. L., Jardine, P. J., Grimes, S. & Ke, A. (2011), 'Structure and assembly of the essential RNA ring component of a viral DNA packaging motor', *Proc. Natl. Acad. Sci. U.S.A.* **108**(18), 7357–7362.

Dunn, S. D. & Futai, M. (1980), 'Reconstitution of a functional coupling factor from the isolated subunits of Escherichia coli F1 ATPase', *J. Biol. Chem.* **255**(1), 113–118.

Erzberger, J. P. & Berger, J. M. (2006), 'Evolutionary relationships and structural mechanisms of AAA+ proteins', *Annu Rev Biophys Biomol Struct* **35**, 93–114.

Franzmann, T. M., Czekalla, A. & Walter, S. G. (2011), 'Regulatory circuits of the AAA+ disaggregase Hsp104', *J. Biol. Chem.* **286**(20), 17992–18001.

Gennerich, A. & Vale, R. D. (2009), 'Walking the walk: how kinesin and dynein coordinate their steps', *Curr. Opin. Cell Biol.* **21**(1), 59–67.

Glynn, S. E., Martin, A., Nager, A. R., Baker, T. A. & Sauer, R. T. (2009), 'Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine', *Cell* **139**(4), 744–756.

Goitre, L., Trapani, E., Trabalzini, L. & Retta, S. F. (2014), 'The Ras superfamily of small GTPases: the unlocked secrets', *Methods Mol. Biol.* **1120**, 1–18.

Gore, J., Ritort, F. & Bustamante, C. (2003), 'Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements', *Proc. Natl. Acad. Sci. U.S.A.* **100**(22), 12564–12569.

Gradia, S., Acharya, S. & Fishel, R. (1997), 'The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch', *Cell* **91**(7), 995–1005.

Grainge, I. (2008), 'Sporulation: SpoIIIE is the key to cell differentiation', *Curr. Biol.* **18**(18), R871–872.

Greenleaf, W. B., Shen, J., Gai, D. & Chen, X. S. (2008), 'Systematic study of the functions for the residues around the nucleotide pocket in simian virus 40 AAA+ hexameric helicase', *J. Virol.* **82**(12), 6017–6023.

Gruber, R., Levitt, M. & Horovitz, A. (2017), 'Sequential allosteric mechanism of ATP hydrolysis by the CCT/TRiC chaperone is revealed through Arrhenius analysis', *Proc. Natl. Acad. Sci. U.S.A.* **114**(20), 5189–5194.

Guo, Z., Ahmadian, M. R. & Goody, R. S. (2005), 'Guanine nucleotide exchange factors operate by a simple allosteric competitive mechanism', *Biochemistry* **44**(47), 15423–15429.

Hackney, D. D. & Stock, M. F. (2008), 'Kinesin tail domains and Mg2+ directly inhibit release of ADP from head domains in the absence of microtubules', *Biochemistry* **47**(29), 7770–7778.

Hahn-Herrera, O., Salcedo, G., Barril, X. & Garcia-Hernandez, E. (2016), 'Inherent conformational flexibility of F1-ATPase -subunit', *Biochim. Biophys. Acta* **1857**(9), 1392–1402.

Hanson, P. I. & Whiteheart, S. W. (2005), 'AAA+ proteins: have engine, will work', *Nat. Rev. Mol. Cell Biol.* **6**(7), 519–529.

Harjes, E., Kitamura, A., Zhao, W., Morais, M. C., Jardine, P. J., Grimes, S. & Matsuo, H. (2012), 'Structure of the RNA claw of the DNA packaging motor of bacteriophage ϕ29', *Nucleic Acids Res.* **40**(19), 9953–9963.

Hirokawa, N., Noda, Y., Tanaka, Y. & Niwa, S. (2009), 'Kinesin superfamily motor proteins and intracellular transport', *Nat. Rev. Mol. Cell Biol.* **10**(10), 682–696.

Howard, J. (2001), *Mechanics of Motor Proteins and the Cytoskeleton*, Sinauer Associates Inx.

Iino, R. & Noji, H. (2013), 'Intersubunit coordination and cooperativity in ring-shaped NTPases', *Curr. Opin. Struct. Biol.* **23**(2), 229–234.

Joly, N., Burrows, P. C. & Buck, M. (2008), 'An intramolecular route for coupling ATPase activity in AAA+ proteins for transcription activation', *J. Biol. Chem.* **283**(20), 13725–13735.

Joshi, S. A., Hersch, G. L., Baker, T. A. & Sauer, R. T. (2004), 'Communication between ClpX and ClpP during substrate processing and degradation', *Nat. Struct. Mol. Biol.* **11**(5), 404–411.

Kaan, H. Y., Hackney, D. D. & Kozielski, F. (2011), 'The structure of the kinesin-1 motor-tail complex reveals the mechanism of autoinhibition', *Science* **333**(6044), 883–885.

Kindt, J., Tzlil, S., Ben-Shaul, A. & Gelbart, W. M. (2001), 'DNA packaging and ejection forces in bacteriophage', *Proc. Natl. Acad. Sci. U.S.A.* **98**(24), 13671–13674.

Komoriya, Y., Ariga, T., Iino, R., Imamura, H., Okuno, D. & Noji, H. (2012), 'Principal role of the arginine finger in rotary catalysis of F1-ATPase', *J. Biol. Chem.* **287**(18), 15134–15142.

Large, S. & Sivak, D. A. (2017), Extensions to the optimal protocol theory. Personal communication.

Lebbink, J. H., Fish, A., Reumer, A., Natrajan, G., Winterwerp, H. H. & Sixma, T. K. (2010), 'Magnesium coordination controls the molecular switch function of DNA mismatch repair protein MutS', *J. Biol. Chem.* **285**(17), 13131–13141.

Li, P. T., Bustamante, C. & Tinoco, I. (2007), 'Real-time control of the energy landscape by force directs the folding of RNA molecules', *Proc. Natl. Acad. Sci. U.S.A.* **104**(17), 7039–7044.

Liu, N., Chistol, G. & Bustamante, C. (2015), 'Two-subunit DNA escort mechanism and inactive subunit bypass in an ultra-fast ring ATPase', *Elife* **4**.

Liu, S., Chistol, G. & Bustamante, C. (2014), 'Mechanical operation and intersubunit coordination of ring-shaped molecular motors: insights from single-molecule studies', *Biophys. J.* **106**(9), 1844–1858.

Liu, S., Chistol, G., Hetherington, C. L., Tafoya, S., Aathavan, K., Schnitzbauer, J., Grimes, S., Jardine, P. J. & Bustamante, C. (2014b), 'A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills', *Cell* **157**(3), 702–713.

Maluf, N. K., Gaussier, H., Bogner, E., Feiss, M. & Catalano, C. E. (2006), 'Assembly of bacteriophage lambda terminase into a viral DNA maturation and packaging machine', *Biochemistry* **45**(51), 15259–15268.

Mao, H., Saha, M., Reyes-Aldrete, E., Sherman, M. B., Woodson, M., Atz, R., Grimes, S., Jardine, P. J. & Morais, M. C. (2016), 'Structural and Molecular Basis for Coordination in a Viral DNA Packaging Motor', *Cell Rep* **14**(8), 2017–2029.

Martin, A., Baker, T. A. & Sauer, R. T. (2008), 'Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding', *Nat. Struct. Mol. Biol.* **15**(11), 1147–1151.

McCullough, J. & Sundquist, W. I. (2014), 'Putting a finger in the ring', *Nat. Struct. Mol. Biol.* **21**(12), 1025–1027.

Moffitt, J. R. & Bustamante, C. (2014), 'Extracting signal from noise: kinetic mechanisms from a Michaelis-Menten-like expression for enzymatic fluctuations', *FEBS J.* **281**(2), 498–517.

Moffitt, J. R., Chemla, Y. R., Aathavan, K., Grimes, S., Jardine, P. J., Anderson, D. L. & Bustamante, C. (2009), 'Intersubunit coordination in a homomeric ring ATPase', *Nature* **457**(7228), 446–450.

Morais, M. C. (2012), 'The dsDNA packaging motor in bacteriophage 29', *Adv. Exp. Med. Biol.* **726**, 511–547.

Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. (1997), 'Direct observation of the rotation of F1-ATPase', *Nature* **386**(6622), 299–302.

Ogura, T., Whiteheart, S. W. & Wilkinson, A. J. (2004), 'Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and intersubunit interactions in AAA and AAA+ ATPases', *J. Struct. Biol.* **146**(1-2), 106–112.

Olivares, A. O., Baker, T. A. & Sauer, R. T. (2016), 'Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines', *Nat. Rev. Microbiol.* **14**(1), 33–44.

Purohit, P. K., Inamdar, M. M., Grayson, P. D., Squires, T. M., Kondev, J. & Phillips, R. (2005), 'Forces during bacteriophage DNA packaging and ejection', *Biophys. J.* **88**(2), 851–866.

Rickgauer, J. P., Fuller, D. N., Grimes, S., Jardine, P. J., Anderson, D. L. & Smith, D. E. (2008), 'Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage phi29', *Biophys. J.* **94**(1), 159–167.

Rotskoff, G. M. & Crooks, G. E. (2015), 'Optimal control in nonequilibrium systems: Dynamic Riemannian geometry of the Ising model', *Phys Rev E Stat Nonlin Soft Matter Phys* **92**(6), 060102.

Rotskoff, G. M., Crooks, G. E. & Vanden-Eijnden, E. (2017), 'Geometric approach to optimal nonequilibrium control: Minimizing dissipation in nanomagnetic spin systems', *Phys Rev E* **95**(1-1), 012148.

Roy, A., Bhardwaj, A., Datta, P., Lander, G. C. & Cingolani, G. (2012), 'Small terminase couples viral DNA binding to genome-packaging ATPase activity', *Structure* **20**(8), 1403–1413.

Schrödinger, E. (1992), *What is Life?*, Cambridge University Press, Cambridge.

Schumacher, J., Zhang, X., Jones, S., Bordes, P. & Buck, M. (2004), 'ATP-dependent transcriptional activation by bacterial PspF AAA+protein', *J. Mol. Biol.* **338**(5), 863–875.

Shirakihara, Y., Leslie, A. G., Abrahams, J. P., Walker, J. E., Ueda, T., Sekimoto, Y., Kambara, M., Saika, K., Kagawa, Y. & Yoshida, M. (1997), 'The crystal structure of the nucleotide-free alpha 3 beta 3 subcomplex of F1-ATPase from the thermophilic Bacillus PS3 is a symmetric trimer', *Structure* **5**(6), 825–836.

Sivak, D. A. & Crooks, G. E. (2012), 'Thermodynamic metrics and optimal paths', *Phys. Rev. Lett.* **108**(19), 190602.

Sivak, D. A. & Crooks, G. E. (2016), 'Thermodynamic geometry of minimum-dissipation driven barrier crossing', *Phys Rev E* **94**(5-1), 052106.

Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L. & Bustamante, C. (2001), 'The bacteriophage straight phi29 portal motor can package DNA against a large internal force', *Nature* **413**(6857), 748–752.

Soga, S., Noumi, T., Takeyama, M., Maeda, M. & Futai, M. (1989), 'Mutational replacements of conserved amino acid residues in the alpha subunit change the catalytic properties of Escherichia coli F1-ATPase', *Arch. Biochem. Biophys.* **268**(2), 643–648.

Tafoya, S. & Bustamante, C. (2017), 'Molecular switch-like regulation in motor proteins', *Phil. Trans. Royal Soc. B (In press. DOI: 10.1098/rstb.2015.0615)* .

Tafoya, S., Large, S. J., Liu, S., Sivak, D. A. & Bustamante, C. (2017b), 'Non-Equilibrium protocols with Minimal thermodynamic length in DNA hairpin pulling experiments.', *In preparation* .

Tafoya, S., Liu, S., Castillo, J. P., Atz, R., Morais, M., Grimes, S., Jardine, P. & Bustamante, C. (2017), 'Molecular switch-like regulation enables global subunit coordination in a viral ring ATPase', *Submitted for publication* .

Thomsen, N. D., Lawson, M. R., Witkowsky, L. B., Qu, S. & Berger, J. M. (2016), 'Molecular mechanisms of substrate-controlled ring dynamics and substepping in a nucleic acid-dependent hexameric motor', *Proc. Natl. Acad. Sci. U.S.A.* **113**(48), E7691–E7700.

Todd, J., Thielman, B. & Wendell, D. (2012), 'Detailed kinetic analysis of the Phi29 DNA packaging motor providing evidence for coordinated intersubunit ATPase activity of gp16', *Virology* **432**(2), 370–375.

Uchihashi, T., Iino, R., Ando, T. & Noji, H. (2011), 'High-speed atomic force microscopy reveals rotary catalysis of rotorless F1-ATPase', *Science* **333**(6043), 755–758.

Uchimura, S., Oguchi, Y., Hachikubo, Y., Ishiwata, S. & Muto, E. (2010), 'Key residues on microtubule responsible for activation of kinesin ATPase', *EMBO J.* **29**(7), 1167–1175.

Vale, R. D. (1996), 'Switches, latches, and amplifiers: common themes of G proteins and molecular motors', *J. Cell Biol.* **135**(2), 291–302.

Verbrugge, S., van den Wildenberg, S. M. & Peterman, E. J. (2009), 'Novel ways to determine kinesin-1's run length and randomness using fluorescence microscopy', *Biophys. J.* **97**(8), 2287–2294.

Verhey, K. J. & Hammond, J. W. (2009), 'Traffic control: regulation of kinesin motors', *Nat. Rev. Mol. Cell Biol.* **10**(11), 765–777.

Wasserman, L. (2004), *All of Statistics: a Concise Course in Statistical Inference*, Springer-Verlag.

Weber, J. & Senior, A. E. (1997), 'Catalytic mechanism of F1-ATPase', *Biochim. Biophys. Acta* **1319**(1), 19–58.

Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. (2012), 'Structure and function of the AAA+ nucleotide binding pocket', *Biochim. Biophys. Acta* **1823**(1), 2–14.

Wong, I., Moore, K. J., Bjornson, K. P., Hsieh, J. & Lohman, T. M. (1996), 'ATPase activity of Escherichia coli Rep helicase is dramatically dependent on DNA ligation and protein oligomeric states', *Biochemistry* **35**(18), 5726–5734.

Woodside, M. T., Behnke-Parks, W. M., Larizadeh, K., Travers, K., Herschlag, D. & Block, S. M. (2006), 'Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins', *Proc. Natl. Acad. Sci. U.S.A.* **103**(16), 6190–6195.

Yagi, H., Kajiwara, N., Iwabuchi, T., Izumi, K., Yoshida, M. & Akutsu, H. (2009), 'Stepwise propagation of the ATP-induced conformational change of the F1-ATPase beta subunit revealed by NMR', *J. Biol. Chem.* **284**(4), 2374–2382.

Yasuda, R., Noji, H., Kinosita, K. & Yoshida, M. (1998), 'F1-ATPase is a highly efficient molecular motor that rotates with discrete 120 degree steps', *Cell* **93**(7), 1117–1124.

Yukawa, A., Iino, R., Watanabe, R., Hayashi, S. & Noji, H. (2015), 'Key chemical factors of arginine finger catalysis of F1-ATPase clarified by an unnatural amino acid mutation', *Biochemistry* **54**(2), 472–480.

Zhang, X. & Wigley, D. B. (2008), 'The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins', *Nat. Struct. Mol. Biol.* **15**(11), 1223–1227.

Zulkowski, P. R. & DeWeese, M. R. (2014), 'Optimal finite-time erasure of a classical bit', *Phys Rev E Stat Nonlin Soft Matter Phys* **89**(5), 052140.

Zulkowski, P. R. & DeWeese, M. R. (2015*a*), 'Optimal control of overdamped systems', *Phys Rev E Stat Nonlin Soft Matter Phys* **92**(3), 032117.

Zulkowski, P. R. & DeWeese, M. R. (2015*b*), 'Optimal protocols for slowly driven quantum systems', *Phys Rev E Stat Nonlin Soft Matter Phys* **92**(3), 032113.

Zulkowski, P. R., Sivak, D. A., Crooks, G. E. & DeWeese, M. R. (2012), 'Geometry of thermodynamic control', *Phys Rev E Stat Nonlin Soft Matter Phys* **86**(4 Pt 1), 041148.

Zulkowski, P. R., Sivak, D. A. & DeWeese, M. R. (2013), 'Optimal control of transitions between nonequilibrium steady states', *PLoS ONE* **8**(12), e82754.