Title
Dissecting the molecular mechanisms of the ClpXP protease, one molecule at a time

Permalink
https://escholarship.org/uc/item/9m9149j4

Author
Sen, Maya

Publication Date
2014

Peer reviewed|Thesis/dissertation
Dissecting the molecular mechanisms of the ClpXP protease, one molecule at a time

by

Maya Sen

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

Committee in charge:

Professor Carlos Bustamante, Chair
Associate Professor Jamie Cate
Assistant Professor Andreas Martin

Fall 2014
Dissecting the molecular mechanisms of the ClpXP protease, one molecule at a time

Copyright 2014
by
Maya Sen
Abstract

Dissecting the molecular mechanisms of the ClpXP protease, one molecule at a time

by

Maya Sen

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Carlos Bustamante, Chair

The cell is a fundamental unit of life, and in order for survival, maintaining a stable intracellular environment is crucial. ATP-dependent protease complexes regulate protein quality and abundance to ensure homeostasis in the cell. They use chemical energy to power processes necessary for regulating the intracellular environment including protein unfolding, polypeptide translocation, and targeted degradation of abnormal and short-lived regulatory proteins.

ClpXP, a well-studied ATP-dependent protease complex from Escherichia coli, is an assembly of homohexameric ClpX coaxially stacked onto a barrel-shaped protease ClpP. ClpX utilizes the energy from ATP hydrolysis to bind the appropriately tagged polypeptide substrates, denature and translocate them into the degradation cavity of ClpP. There have been long-standing questions in the field, specifically whether ClpX can generate force to unfold proteins, how do the six ClpX subunits communicate, and coordinate their ATPases cycles to generate force.

Optical tweezers are used as a powerful single-molecule technique to characterize the mechanochemical properties of biomolecules in the nanometer and piconewton range. After a decade of research, we have developed a novel optical tweezers assay to monitor ClpX as it binds, unfolds and translocates various green fluorescent protein (GFP) fusion substrates in real time under a range of ATP concentrations.

We characterized the general properties of the motor, which are described in chapter 2. From the analysis of these single-molecule trajectories, we have determined several unique properties of the motor previously undetectable in bulk. First, ClpX can generate up to 20 pN of force to translocate and unfold proteins, eventually being stalled in movement as the opposing force approaches 20 pN.

We subsequently explored the communication within the hexameric ring of wild-type ClpX at various ATP, ADP, and Pi concentrations. Our results, described in chapter 3, provided the first direct evidence of a force-generation mechanism and that the phosphate release is coupled to the force-generating step of ClpX. We demonstrated that two to four subunits of the ClpX hexamer actively participates in bursts of translocation and that between the
translocation bursts, the motor has a mean constant dwell duration independent of ATP concentration. Our analysis revealed defined a new archetype of motor coordination, which is critical as a fail-safe mechanism to prevent the motor disengagement from its substrate.
To my mom and dad
Contents

1 Introduction 1
  1.1 The role of force in cellular processes ................................. 1
  1.2 Molecular machines of the ASCE family ............................... 3
  1.3 Structure of AAA+ ClpXP protease .................................... 5
  1.4 Functional role of ClpXP .................................................... 7
  1.5 Probing the Mechanochemistry and Coordination of ClpXP .............. 8

2 Characterizing the AAA+ molecular machine ClpXP 10
  2.1 Summary ................................................................. 10
  2.2 Introduction ............................................................. 10
  2.3 Results ................................................................. 11
  2.4 Discussion .............................................................. 26
  2.5 Methods ................................................................. 29
  2.6 Extended Experimental Procedures ..................................... 30

3 Studying mechanisms of force-generation and intersubunit coordination 38
  3.1 Summary ................................................................. 38
  3.2 Introduction ............................................................. 38
  3.3 Results ................................................................. 40
  3.4 Discussion .............................................................. 53
  3.5 Methods ................................................................. 58
  3.6 Extended Experimental Procedures ..................................... 59

4 Conclusions 65
  4.1 Commentary on Data Analysis .......................................... 65
  4.2 Concluding Remarks and Future directions ............................. 70
List of Figures

1.1 Optical trapping a dielectric bead ............................................. 2
1.2 The ASCE division is composed of RecA-like and AAA+ subfamilies ......... 4
1.3 Crystal structure of ClpX bound to ATPγS ................................... 5
1.4 Crystal structure of the peptidase component ClpP ............................ 6
1.5 ClpXP uses ATP to degrade its substrates ................................... 7

2.1 Unfolding and Translocation of GFP-Titin Fusion Proteins by ClpX and ClpXP 12
2.2 Motor Properties and Mechanochemical Cycle of ClpX(P) ...................... 14
2.3 ClpXP Translocation Stepping Displays Coordination between Subunits .... 16
2.4 GFP Unfolding by ClpX(P) Results in a Well-Defined Contour Length Change, Related to Figure 2.5 ........................................... 19
2.5 Unraveling of GFP by ClpX(P) Shows a Well-Defined Structural Transition and Reveals a Single-Unfolding Intermediate .......................... 20
2.6 Side-by-Side Comparison of ClpX (Red Bars) and ClpXP(Orange Bars) Unfolding of GFP, Related to Figure 2.5 ......................................... 22
2.7 A Schematic Representation of GFP Unfolding via a Short-Lived Intermediate . 23
2.8 ClpXP Is a More Robust Motor with Higher Unfolding Efficiency and Lower Slipping Frequency than ClpX Alone .................................... 25
2.9 Model for the Mechanical Unfolding of GFP by ClpX(P) ....................... 26
2.10 Protein Modifications and dsDNA-Handle Attachment to Protein Substrates, Related to Figure 2.1 ........................................... 33
2.11 Related to Figure 2.3 and Table 2.1 ........................................... 35
2.12 Pulling Control Experiments in ATP-γ-S Stall Buffer, Related to Figure 2.1 ... 36

3.1 Single-Molecule Experimental Constructs ................................... 41
3.2 Nucleotide dependence of ClpX translocation velocity, related to Figure 3.3 ... 42
3.3 Effects of ATP, ADP, and P_i on Translocation ................................ 43
3.4 ATPγS Induces Long-Lived Pauses ........................................... 46
3.5 Rate of ClpXP ATPγS Hydrolysis Measured by Thin-Layer Chromatography, Related to Figure 3.4 ........................................... 48
3.6 Translocation Burst Size is Dictated Mostly by the Geometry of ClpX and Not by Regularities of the Polypeptide Substrate, Related to Figure 3.7 .... 49
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7 Effect of ATP Concentration on Burst Size and Dwell Duration</td>
<td>50</td>
</tr>
<tr>
<td>3.8 Detection and Characterization of Intermediates During GFP Unfolding, Related to Figure 3.9</td>
<td>51</td>
</tr>
<tr>
<td>3.9 GFP Unfolding Mechanism</td>
<td>53</td>
</tr>
<tr>
<td>3.10 Minimal Mechanochemical Model of ClpXP Translocation</td>
<td>56</td>
</tr>
<tr>
<td>4.1 Filtering ClpX traces with different methods</td>
<td>66</td>
</tr>
<tr>
<td>4.2 Pause detection using the SIC</td>
<td>68</td>
</tr>
<tr>
<td>4.3 Step detection tools for ClpX’s variable step</td>
<td>69</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Pause Density and Duration as a function of force, related to Figure 2.2D. . . . 31
2.2 ATPase activity of ClpX free in solution and immobilized on 2.1 um streptavidin-coated polystyrene beads (ClpX-SAPB). . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
Acknowledgments

I would first like to thank my advisor, Dr. Carlos Bustamante for his support and mentorship. I am grateful for such a rewarding and challenging opportunity to study science in a rigorous manner. He has been incredibly supportive and available in times of need. I would also to our close collaborator Dr. Andreas Martin for support and insightful discussions. Special thanks for Dr. Jamie Cate for his advice and optimistic support. Last but not least to Nacho for sitting through my group meetings and providing advice on finer details during presentations.

The research presented in this dissertation was made possible in a collaboration of great minds. I am grateful for the intense and thoughtful discussions with my current and former subgroup members: Rodrigo Maillard, Maurizio Righini, Jiongyi Tan, Kris Nyquist, and Piere Rodriguez. Chapter 2 was the initial characterization of ClpXP protease using optical tweezers and working with my colleagues: Rodrigo A. Maillard, Gheorghe Chistol, Maurizio Righini, Jiongyi Tan, Christian Kaiser, and Courtney Hodges. It was a honor to work on characterizing the first protein translocase with Rodrigo Maillard. Chapter 3 was an attempt for a more detailed understanding of the motor and with a team of Rodrigo A. Maillard, Kristofor Nyquist, and Piere Rodriguez-Aliaga. All of the research was done in close collaboration with Dr. Andreas Martin. Thanks to the phage group-Gheorge Chistol, Craig Hetherington, and Shixin Liu- for their helpful discussions and in particular, thanks to Gheorghe for building the initial core of ClpX analysis code. Thanks to my lab colleagues for helpful discussions in no particular order: Bibiana Onoa, Yves Coello, Charlie, Yara Meija, Courtney Hodges, Errol Watson, Ninning Liu, Phillip Elms, Ariel Kaplan, Tingting Liu, Yuanbo Cui, Antony Lee, Varsha Desai, Lisa Alexander, Natalie Thung, Melina Winterton, Sanghyuk Lee, Jae Yen Shin, Cesar Diaz, Filip Frank, Liang Wee, Toyotaka Ishibashi, Lacra Bintu, and Luis Ramirez. Special thanks to my fellow chemistry graduate students, Daniel Goldman and Pim Dangulwanich.

Last but not least, I would like to acknowledge my parents, brother, sister-in-law, grandparents, cousins, aunts, and uncles for their unconditional support and great sense of humor. My parents and brother have been my first role models. My father taught me that the study of science is a privilege. I can only hope that I have the same passion and curiosity about science that he still has. My mother taught me perseverance and optimism can go hand in hand. She is the most creative person I know. Special thanks to Ilana Somasunderam Mitchell Flax, and Rathi Srinivas. And to you, the reader who downloaded it and read this long document.
Chapter 1

Introduction

1.1 The role of force in cellular processes

Force plays an important role in cellular tasks ranging from macroscopic to microscopic length scales. Tasks of DNA replication, protein synthesis, and cell motility are all driven by forces. In recent years, understanding the role of force in biological systems have become increasing significant due to the advent of new techniques with better spatial and temporal resolution. These tools have provided detailed mechanistic insight on individual biomolecules as they can exert forces on biological systems that require sensitivity in force and displacement measurements.

Optical tweezers is widespread tool to understand force in cellular processes

In the 1970s, Arthur Ashkin first proposed the idea of optically trapping microscopic objects. He demonstrated that a focused beam of light could optically trap colloidal particles in both air and liquids (Ashkin, 1970). These important findings eventually led to the advent of a single-beam gradient force optical trap or optical tweezers, which has led to applications in biology, physics and chemistry (Ashkin and Dziedzic, 1971; Ashkin et al., 1986; Ashkin and Dziedzic, 1987; Ashkin et al., 1987). In particular, the use of optical tweezers has been widespread in understanding mechanical properties of biopolymers and molecular machines at a single-molecule level.

Single-molecule optical tweezers is a powerful tool in studying molecular machines. The advantage of this technique is the ability to monitor dynamic motion of biological processes one molecule at a time in real time with with nanometer and pico-Newton precision. Using this tool adds another layer of information as tracking of the dynamics of an individual molecule prevents bias from ensemble averaging of bulk studies and static snapshots from crystal structures. Such information can provide more details about mechanistic properties of an individual molecule.
Figure 1.1: (A) In this scenario of a stable trap, the light is focused and produces a 3D intensity gradient profile. A focused beam of light with radial intensity can stably trap dielectric bead. Two rays of light (black) are refracted upon hitting the bead however the change in momentum ($\Delta P$) results in a net force towards the focus. The lateral forces (blue arrows) cancel each other out and the scattering force balances the axial forces so that the bead will always be drawn back to its equilibrium position. (B) When a bead is illuminated with rays of light of different intensities, the change in momentum results in a net force that is greater on top ($F_1 > F_2$). Thus, the bead moves upward and is not stably trapped (Neuman and Block, 2004).
CHAPTER 1. INTRODUCTION

The principle behind optical tweezers is that a tightly focused beam of light with an high numerical aperture objective lens can optically trap a microscopic object (such as a dielectric polystyrene bead). Two types of forces are involved in optical traps that will act on the object: gradient and scattering forces. A gradient force pulls the object along the spatial gradient of light intensity, whereas a scattering force pushes an object along the direction of light propagation. In order to stably trap a bead, the gradient force pulling the bead toward the highest point of light intensity must exceed the scattering force pushing the bead away from the focal region. As a result of the push-pull between the forces, the optical trap acts as a Hookean spring (see Figure 1). The literature regarding optical tweezers is significant, and several previous members have written detailed reviews and theses on the principles, building, alignment, and calibration of optical tweezers (Neuman and Block, 2004; Moffitt et al., 2006; Bustamante et al., 2009; Block, 1992; Berg-Sørensen and Flyvbjerg, 2004; Lang et al., 2002).

1.2 Molecular machines of the ASCE family

In cells, many biomechanical processes rely on molecular machines, which are multisubunit complexes that couple chemical energy to accomplish cellular tasks, such as transcription, translation, intracellular trafficking, membrane transport, and proteolysis of proteins. A large number of these are oligomeric, ring-shaped NTPases that belong to a broad superfamily, known as the Additional strand conserved E (catalytic glutamate). The motors of the ASCE family share the core nucleotide binding folds, however can be categorized further into two subfamilies RecA and AAA+ (ATPases Associated with various cellular Activities) due to distinct features in topology (Lyubimov et al., 2011; White and Lauring, 2007b; Erzberger and Berger, 2006; Hanson and Whiteheart, 2005).

The subunit of these oligomeric complexes contain a Walker A and B motif with nucleotide binding site at the interface between neighboring subunits. The conserved Walker A and B motifs bind to the β and γ moieties of the nucleotide and a Mg$^{2+}$ cofactor. The catalytic glutamate, resides either within the Walker B or close by, plays an important role in hydrolysis through coordinating the Mg$^{2+}$ cation to direct a nucleophilic attack on the β - γ bond of the nucleotide. As the nucleotide binds to the interface between neighboring subunits, each subunit contains an arginine finger, which is another defining feature of the ASCE division. This conserved arginine plays a role in the inter-subunit communication and the hydrolysis of the nucleotide. The basis of most biochemical actions in ASCE family is to couple the free energy of NTP hydrolysis to conformational changes within the ring or in other complexes (Wendler et al., 2012; Ogura and Wilkinson, 2001; Ogura et al., 2004; Aravind et al., 2004; Lyubimov et al., 2011; Erzberger and Berger, 2006; Iyer et al., 2004).
Figure 1.2: Topology diagrams on the classification and features of the ASCE division with the core ASCE fold in green, and other structural elements Walker A, B, Arginine Finger (A, B and R), sensor I and II motif in red (Adapted from Lyubimov et al., 2011; Erzberger and Berger, 2006; Burroughs et al., 2007).
1.3 Structure of AAA+ ClpXP protease

ATP-dependent proteases of the AAA+ family regulate protein quality and abundance in the cell by degrading damaged, abnormal, or short-lived regulatory proteins. A plethora of proteases are responsible for such tasks in bacteria, such as Clp, Lon, HsIUV, and FtSH. The members of Clp protease family (ClpAP, ClpCP, ClpXP, HsIUV) consist of two components: An ring-shaped ATPase chaperone that co-axially stacks upon a proteolytic core. In the presence of nucleotide, the chaperone or unfoldase recognizes the tagged substrate for degradation, unfolds and translocates the polypeptide chain through its central pore into the protease component, where the proteolytic sites are located. Many of these chaperones can also function without its peptidase component to disassemble protein complexes. The focus of my thesis is on understanding how the model system ClpXP (Caseinolytic protease complex XP) couples chemical energy to perform mechanical work.

ATP-dependent translocase ClpX

The ClpX ring is made of six identical subunits, with each subunit consisting of a large AAA+ domain, a small AAA+ domain, and a family-specific N domain. The large domain is against the clockwise adjacent small subunit, with the nucleotide-binding pocket at the interface between the small and large AAA+ domain. ClpX has an asymmetric structure with a diameter of $\sim 135\text{Å}$ with a pore size of 20-30 Å (Glynn et al., 2012; Hwang and Lang, 2013; Baker and Sauer, 2012; Glynn et al., 2009).

Figure 1.3: (A) Crystal structure of single-chain ClpX (pdb:4I81) bound to ATP\textsubscript{γ}S molecules. The structure shows that ClpX can bind a maximum of four nucleotides, hence four ClpX subunits are loadable and two are unloadable due to the strain in the ring. (B) Side view of the same crystal is shown here for dimension purposes. ClpX has a diameter of $\sim 10$ nm with a height of $\sim 5$ nm.
Crystal structures also show that ClpX is viewed as a single-rigid body due to the strong contacts between adjacent subunits, however the hinge between the small and large AAA+ domain allows for the ring to be more flexible. In the presence and absence of nucleotide, the ClpX subunits are in either two different conformations: loadable (L) and unloadable (U). In the loadable state, ATP can bind at the interface between the small and large AAA+ domains. After four subunits have bound ATP, the remaining two subunits can no longer bind ATP since the small AAA+ domain blocks the cleft for nucleotide binding, most likely due to the ring strain from the other nucleotide-bound subunits. Mutations within ClpX subunits have been shown to abolish ATP binding and disrupt formation of the ClpXP complex, which strongly suggests an intersubunit coordination amongst the ATPase cycles within the ring. The conformational changes imposed on the ATP binding could be enhanced through the hexameric ring, which indicates that the subunits within the ring coordinate their ATP hydrolysis cycles in order to unfold and translocate substrate (Glynn et al., 2012; Hwang and Lang, 2013; Baker and Sauer, 2012; Glynn et al., 2009; Martin et al., 2005).

Peptidase ClpP

ClpP is the proteolytic component of the ClpXP molecular machine that is highly conserved in bacteria and eukaryota. ClpP is composed of 14 identical subunits arranged into two heptameric rings with proteolytic active sites enclosed in the spherical cavity (diameter $\sim 50\text{Å}$). The mature ClpP protein (21.5 kDa) is active after the in vivo cleavage of the first 14 residues. Each subunit contains the catalytic triad (Ser97-His122-Asp171) that are spaced $\sim 25\text{ Å}$ apart (easily spanned by a 7 amino acid peptide), predicting an effective local concentration of $\sim 350\text{ mM}$ for the active sites (Baker and Sauer, 2012; Kress et al., 2009; Zolkiewski, 2006; Yu and Houry, 2007).

Figure 1.4: The crystal structure of ClpP shown in a side view (A), top view (B), and the internal cavity with the serine active sites in red (C) (pdb:1Y6).
The molecular architecture of ClpP (with pore diameter of 10 Å) allows smaller polypeptide chains of 6 amino acids to freely diffuse into the degradation chamber however the axial channel of ClpP forms a stem-loop structure that restricts larger peptides (Kress et al., 2009). Therefore, ClpP must complex with an associate ATPase to thread larger polypeptides through its central pore to reach the proteolytic sites. Substrates that enter the degradation cavity bind to the hydrophobic grooves at the equatorial plane of ClpP. The polypeptide chain is then broken down into small peptide fragments and later into free amino acids by exopeptidases. Biochemical studies have reported that ClpP peptide cleavage is on the order of \( \sim 10000 \text{ min}^{-1} \). And with a cleavage product of 10 residues, this rate should be sufficient to prevent the chamber from suffering due to overfilling of polypeptide chain (Baker and Sauer, 2012).

### 1.4 Functional role of ClpXP

![ClpX and ClpP diagram](image)

In the presence of ATP, ClpX recognizes the tagged substrate for degradation, unfolds and translocates the polypeptide chain through its central pore into ClpP.

In order to degrade tagged protein substrates, ClpX pore aligns with the ClpP pore to form a continuous channel that allows access of polypeptides into the degradation chamber of ClpP (Baker and Sauer, 2012). Interestingly, a symmetry mismatch exists between the hexameric ClpX ring and heptameric ClpP ring.

Once ClpX recognizes the protein substrate via a recognition tag, three key structural elements play a role in translocation and protein unfolding: RKH, pore-1, and pore-2 loops. RKH loops line the narrow pore on the top of the ClpX. These loops play an important role in substrate recognition and are not to be coupled to ATP hydrolysis. Whereas the pore-1 loops (with a conserved GYVP motif) that line the central pore of ClpX are coupled to ATP hydrolysis and responsible for contacting the polypeptide chain. The tyrosine residue in the
GYVG-loop of Clp family is essential for unfolding and translocation. This conserved residue plays an important role in the grip that is required to tug on the substrate for unfolding. Removal of the tyrosine results in a large ATPase rates, suggesting that ClpX incorrectly grips the substrate and burns ATP more frequently in order to accomplish the task.

The pore-2 loop is located at the bottom of the central channel of ClpX and interacts with N-terminal ClpP residues. This interaction results in a ATPase suppression of the wild-type ClpX. Pore-2 loop has been also shown to depend on the nucleotide state of individual subunits and assists in substrate unfolding. Interestingly, the ClpX$_2$P complex has only one chaperone translocating substrate at a given time, indicating that these two unfoldases, not only has its own intersubunit coordination, must communicate and coordinate their mechanical tasks through ClpP (Baker and Sauer, 2012; Kress et al., 2009; Martin et al., 2008a; Martin et al., 2007).

Previous single-chain ClpX biochemical studies have also led to important findings on the intersubunit coordination within the ring. The ClpX subunits are covalently-linked and lack a N terminal domain, yet still retain the ability to bind to ClpP and degradation function. Using different combinations of wild type and mutant subunits (both abolish ATP binding, but E185Q can adopt an ATP-bound state and R370K adopts an ‘empty’ or ADP state) in the single-chain, these results demonstrated that ClpX power stroke is generated by ATP hydrolysis in a single subunit, indicating a probabilistic sequence of hydrolysis events around the ring is favored (Martin et al., 2005).

1.5 Probing the Mechanochemistry and Coordination of ClpXP

Over several decades, there have been significant breakthroughs in understanding how this protease functions, however several key mechanisms by which this protease coordinates its six subunits to convert into work remain unclear. We would specifically like to understand the mechanochemistry and intersubunit of this motor using optical tweezers.

In chapter 2, we will characterize functional properties of ClpXP protease as an ATP-dependent molecular machine by asking the following questions.

- What are the mechanical properties of the motor?
- What is the velocity of this motor under tension?
- Does ClpX(P) exert force and is the motor a brownian ratchet or power stroke?
- What is the step size of the motor?
- How does ClpP affect ClpX’s ability as an unfoldase? How does ClpX handle a mechanical obstacle such as GFP?
In chapter 3, we will link the ATPase cycle of ClpX to its mechanical tasks, and further understand the communication between ClpX subunits.

- How does ClpX coordinate its chemical cycle (ATP binding, hydrolysis, and product release) to mechanical work (translocation and protein unfolding)?
- What is the communication amongst all six subunits?
- How do all or a subset of subunits coordinate hydrolysis to translocation and protein unfolding events?
Chapter 2

Characterizing the AAA+ molecular machine ClpXP


2.1 Summary

AAA+ unfoldases denature and translocate polypeptides into associated peptidases. We report direct observations of mechanical, force-induced protein unfolding by the ClpX unfoldase from E. coli, alone, and in complex with the ClpP peptidase. ClpX hydrolyzes ATP to generate mechanical force and translocate polypeptides through its central pore. Threading is interrupted by pauses that are found to be off the main translocation pathway. ClpX’s translocation velocity is force dependent, reaching a maximum of 80 aa/s near-zero force and vanishing at around 20 pN. ClpX takes 1, 2, or 3 nm steps, suggesting a fundamental step-size of 1 nm and a certain degree of intersubunit coordination. When ClpX encounters a folded protein, it either overcomes this mechanical barrier or slips on the polypeptide before making another unfolding attempt. Binding of ClpP decreases the slip probability and enhances the unfolding efficiency of ClpX. Under the action of ClpXP, GFP unravels cooperatively via a transient intermediate.

2.2 Introduction

ATP-dependent proteases of the AAA+ (ATPases associated with various cellular activities) superfamily power the degradation of abnormal, denatured, or otherwise damaged polypeptides, as well as the removal of short-lived regulatory proteins (King et al., 1996). The hydrolytic active sites of these proteases are sequestered in the internal chamber of a
barrel-shaped peptidase complex, preventing the diffusion and nonspecific degradation of folded or even large unfolded polypeptides (Wang et al., 1997). To facilitate specific protein degradation, the peptidases pair with energy-dependent hexameric AAA+ unfoldases that recognize appropriately tagged protein substrates and utilize the energy from ATP hydrolysis to unfold and translocate the polypeptide into the associated peptidase chamber for degradation (Baker and Sauer, 2006). It has been suggested previously that AAA+ unfoldases may exert mechanical force to unravel the tertiary and secondary structures of protein substrates. Even though there have been initial single-molecule fluorescence studies of ClpX (Shin et al., 2009), direct evidence for the generation of force and a detailed characterization of the mechanochemistry of these molecular machines are still lacking. Here we investigate the motor properties of ClpX, a homohexameric AAA+ ATPase from *Escherichia coli* that recognizes proteins with a C-terminal ssrA tag and uses cycles of ATP hydrolysis to unfold and translocate the substrates into its associated peptidase, ClpP (Gottesman et al., 1998). We use a single-molecule optical tweezers-based assay to demonstrate that ClpX generates mechanical force to unfold its substrates. We characterize the dynamics of ClpX as it encounters a folded substrate such as GFP and translocates unfolded polypeptides through its central processing pore. Furthermore, we investigate the effects of ClpP on the translocation activity and unfolding efficiency of ClpX. These studies thus provide important new insight into the general operating principles used by energy-dependent proteases to unfold and degrade protein substrates inside the cell.

### 2.3 Results

#### ClpX Unfolding Trajectories

A dual-trap optical tweezers geometry was used to monitor real-time trajectories of individual ClpX hexamers or ClpXP complexes as they unfold and translocate GFP-titin fusion substrates. In these substrates, either one GFP molecule or two GFP moieties in tandem separated by either a short or a long unstructured linker were C-terminally fused to a permanently unfolded ssrA-tagged titin I27 module (Ti\textsubscript{cm}) and attached by their N termini to a dsDNA handle (Figure 2.1, see Extended Experimental Procedures). The DNA-tethered substrates and a single-chain variant of ClpX were immobilized on different polystyrene beads coated with anti-digoxigenin antibody and streptavidin, respectively. ClpX was allowed to bind and engage the ssrA-tagged substrate by bringing the two beads into close proximity in the presence of saturating ATP and an ATP regeneration system (Figure 2.1).

After successful substrate engagement, we monitored ClpX unfolding and translocation activities by measuring the changes in extension between the two beads in passive mode (i.e., the trap positions are fixed and the force load on the motor is allowed to vary with motor activity). Traces for the three different substrates displayed sudden extension gains (rips) followed by a slower decrease in extension, together resembling a saw-tooth-like pattern (Figure 2.1B-D). We assigned rips either to ClpX-induced unfolding events of GFP or to...
Figure 2.1: The geometry of our single-molecule assay in dual-trap optical tweezers: ClpX(P) complexes were immobilized on a streptavidin polystyrene bead (SA) via a biotin tag on ClpX. The GFP-titin fusion substrate is covalently linked to a 3 kbp dsDNA handle with a Dig tag that binds to an antibody-coated polystyrene bead (AD). All substrates included one or two GFP molecules (green) fused to a Ti\textsubscript{cm}-ssrA moiety (red and black). The blue flexible linker corresponds to the ybbR tag. (B-D) GFP unfolding (red arrowheads), motor slips (black arrowheads) and translocation trajectories for three different substrates obtained at 7 pN. (B) Single GFP-titin substrate. (C) Double GFP substrate with a short linker (10 aa) between the GFP molecules. (D) Double GFP substrate with a long flexible linker of 200 aa corresponding to two Ti\textsubscript{cm} domains (red) between the GFP molecules. (E) Basic features of our single-molecule trajectories. The rip corresponding to a GFP unfolding is preceded by a pause (orange). After GFP is unfolded, ClpX translocates the unfolded polypeptide chain (green) with occasional pauses (orange). In all cases the raw data were filtered and decimated to 1000 Hz (in gray) or 2.5 Hz (in red). For the protocol describing the attachment of the dsDNA handle to the fusion substrates, see Figure 2.10. In the presence of ATP-\gamma-S, none of the unfolding or translocation events described in (B)−(E) were observed (Figure 2.12).
temporary disengagement of the motor along the polypeptide backbone track (Figure 2.1,
red and black arrowheads, respectively). The continuous decrease in extension after a rip
was identified as the translocation of the unfolded polypeptide chain through the ClpX
pore. As seen in Figure 2.1B-E, translocation is interspersed by pauses of various lengths. In
> 95% of all trajectories, translocation events were not observed before the first rip. We
therefore conclude that the motor had already translocated the unfolded Ti_cm-ssrA segment
by the time we started recording ClpX unfolding activity. The changes in extension (in nm)
associated with rips and translocation can be expressed in terms of numbers of amino acids
(aa) using the worm-like chain (WLC) force-extension formula (Bustamante et al., 1994)
(see Extended Experimental Procedures).

**ClpX Translocation and Force Dependence**

The characteristics of the translocation process of ClpX are fundamentally distinct from
those of other molecular motors that have been studied under external force (Bustamante
and Cheng, 2011; Yildiz and Selvin, 2005). For molecular motors moving along stiff tracks
such as dsDNA or microtubules, the track persistence length (P) is orders of magnitude larger
than the step-size of the motor (for example, P_{dsDNA} = \sim 50 \text{ nm} compared to motor step-
sizes of just 0.34 and 0.85 nm for RNA polymerase and \phi 29 packaging motor, respectively
[Abbondanzieri et al., 2005; Bustamante et al., 2009]; and P_{microtubule} = \sim 1 \text{ mm} compared to
a step-size of 8 nm for kinesin [Svoboda et al., 1993]). We observe ClpX(P) taking 1 nm steps
(see below), which are larger than the persistence length of the polypeptide track (P_{peptide} =
0.65 nm). As a result, on a length scale comparable to the step-size, ClpX(P) translocates
a flexible, irregular polymer, whereas motors such as kinesin or dsDNA translocases move
along much more rigid, periodic tracks.

The highly elastic nature of unfolded polypeptides and the unique characteristics of
this AAA+ unfoldase compelled us to define two distinct quantities: contour velocity (in
aa/s) and extension velocity (in nm/s). Extension velocity refers to the geometric length
(in nm) of the translocated polypeptide chain, reflecting the size of the motor step, whereas
contour velocity measures how many amino acids pass through the ClpX pore per unit time.
For example, 1 nm/s of extension velocity corresponds to \sim 8 aa/s of contour velocity at
4 \text{ pN} external force and only \sim 4 aa/s at 13 \text{ pN}. To illustrate this point, imagine ClpX
as an ant walking on a rubber band stretched under force. At a low force the ant will
traverse the entire length of the rubber band faster than at a high force, even though the ant
maintains a constant step-size and stepping rate in both cases. The product of the step-size
and the stepping rate in the ant analogy corresponds to the extension velocity (nm/s) of
ClpX, whereas the amount of rubber traveled by the ant corresponds to the contour velocity.
Although contour velocity (aa/s) has been widely used in bulk studies (zero-external force),
the extension velocity (nm/s) is necessary to properly characterize polypeptide translocation
under external forces. Because all of our single-molecule experiments were performed under
external force, most of our discussions refer to the extension velocity (nm/s). Our analysis
yielded a pause-free extension velocity of 8.2 \pm 0.3 nm/s (mean \pm standard error of the mean
Figure 2.2: (A) Extension translocation velocity (mean±SEM in nm/s) as a function of external force for ClpX (red symbols) and ClpXP (blue symbols). The dotted curves represent phenomenological fits to the data. (B) Contour length translocation velocity (mean±SEM in aa/s) as a function of external force for ClpX (red symbols) and ClpXP (blue symbols). The dotted lines extrapolate velocity to zero-external force. (C) Histograms of extension velocity for ClpX (red) and ClpXP (blue) at forces below 13 pN, where velocity is force independent. (D) The pause density (or frequency) during ClpX(P) translocation (mean±SEM) as a function of force. (E) Correlation plot between ClpX(P) pause density (mean±SEM) versus translocation velocity, when velocity is force independent (<13 pN). (F) The proposed minimal mechanochemical cycle for ClpX(P) during translocation, including associated force-dependent rates. X(P)_{n-1}, X(P)_n, and X(P)_{n+1} refer to consecutive steps of ClpX(P) during translocation. The translocation rate (k_f) and pause entry (k_p), reflected by the pause-free velocity and pause density, respectively, are force dependent whereas the exit from a pause (k_p) is independent of force. For additional information on pause density and duration, see Table 2.2.
CHAPTER 2. CHARACTERIZING THE AAA+ MOLECULAR MACHINE CLPXP

(SEM)) for ClpX at opposing forces between 5-13 pN (Figure 2.2A, red symbols), whereas the extension velocity for the ClpXP complex in the same force range was 6.0 ± 0.3 nm/s (mean ± SEM, Figure 2.2A, blue symbols). This 27% decrease in velocity can be explained by the repression in ATP hydrolysis activity of ClpX when bound to ClpP (Table 2.1). We used the contour velocity (aa/s) to estimate near-zero force velocities of ∼80 aa/s and ∼60 aa/s for ClpX and ClpXP, respectively (Figure 2.2B). The latter value is in excellent agreement with bulk estimates for ClpXP translocation (Martin et al., 2008a).

Interestingly, a rather broad distribution of the mean pause-free translocation velocities for ClpX and ClpXP was seen regardless of the force range (Figure 2.2C), revealing an intrinsic heterogeneity in the activity of individual ClpX hexamers. Although similar heterogeneity has been described previously for other molecular motor enzymes (Neuman et al., 2003), it is possible that this dispersion in part reflects the chemical and physical heterogeneity of the unfolded polypeptide track.

ClpXP: Allosteric or Force-Generating Enzyme

A long-standing question about AAA+ unfoldases is whether these enzymes in fact exert mechanical force or just use an allosteric binding mechanism to unfold their substrates. To investigate the unfolding mechanism used by ClpX, we monitored the response of the motor to an opposing force while translocating a substrate. The force versus velocity plot shows that ClpX is capable of working against and therefore generating mechanical force (Figure 2.2A, red symbols). This plot also reveals that the ClpX translocation velocity is nearly constant up to ∼13 pN, implying that within this force range and at saturating ATP concentrations, chemical steps (hydrolysis or product release) are rate limiting. At opposing external forces exceeding 13 pN, the pause-free velocity decreases monotonically, indicating that conformational changes driving translocation become rate limiting. We were able to place a lower bound of ∼20 pN for the stall force of ClpX (Figure 2.2A). ClpXP’s stall force and force-velocity dependence is very similar to that of ClpX (Figure 2.2A, blue symbols). It has been proposed that protein unraveling by a AAA+ unfoldase results from the enzyme’s attempts to translocate the folded structure through its narrow central pore. Thus, the stall force for translocating a resisting polypeptide gives a measure of the maximum mechanical force that ClpX can apply to unfold a substrate.

At high opposing loads, it was possible to observe well-defined translocation steps for ClpXP (Figure 2.3). The pairwise distribution analysis of these traces revealed a remarkable stepping periodicity of 1, 2, and 3 nm (Figure 2.3B and 2.3F). The observed 1 nm step is in good agreement with high-resolution structural data of ClpX (Figure 2.3C), which show distinct conformations of subunits in different nucleotide states and indicating potential transitions and loop movements of ∼1 nm per ClpX subunit that might lead to substrate translocation (Glynn et al., 2009). It is assumed that conformational changes of the pore 1 loops with their highly conserved Tyr residues (Figure 2.3C, red loops) together with rigid-body movements of subunits in the ClpX hexamer propel the substrate through the central pore (Martin et al., 2008a). Based on mutational studies, it was suggested that subunits
Figure 2.3: (A) Fragment of a ClpXP translocation region at 10 pN displaying stepping periodicity of 1.0 nm. Raw data filtered and decimated to 500 Hz are shown in gray, and the raw data boxcar-filtered to 20 Hz are shown in blue. (B) The pairwise distance distribution for the trace in (A) shows a 1 nm periodicity (black arrows). (C) Side view of the nucleotide-bound ClpX hexamer (Glynn et al., 2009) with three subunits removed to allow visualization of the pore-1 loops (red) with the critical Tyr153 (stick representation) shows distinct staggering relative to the pore axis. The loop of the nucleotide-free subunit 1 is close to the top of the pore, whereas the loops of the nucleotide-bound subunits 2 and 3 are in an intermediate and bottom position, respectively. The distances between Tyr153 in each of these staggered positions are 1 nm. (D and E) Fragment of a ClpXP translocation region at 14 pN displaying a stepping periodicity of 2.2 and 3.0 nm, respectively. (F) The pairwise distance distribution for the traces in (D) and (E) shows the 2.2 and 3.0 nm periodicity (upper and lower panels, respectively). For data comparing the ATP hydrolysis rates of ClpX and ClpXP, see Table 2.1 and Figure 2.11.
contribute additively to ClpX activity and that ATP hydrolysis in one subunit at a time drives the conformational changes for substrate translocation (Martin et al., 2005). Our measured extension velocities of 6 and 8 nm/s, along with the smallest observed step-size of 1 nm, predict hydrolysis rates of 360 and 480 ATP min\(^{-1}\) for ClpXP and ClpX, respectively (assuming 1 ATP consumed per step). These values are in good agreement with our bulk measurements of the ATPase activities during translocation of permanently unstructured substrates (Table 2.1 and Figure 2.11). Our observation of distinct 2 and 3 nm steps (Figure 2.3D-F) suggests a coordinated, near-simultaneous stepping of two or three motor subunits that cannot be resolved in our measurements.

**Power Stroke versus Brownian Ratchet Model**

Two general models of motor operation have been proposed for AAA+ molecular machines. In the Brownian ratchet model, the motor uses ATP binding/hydrolysis or product release to rectify its Brownian motion, cross the energy barrier, and move in one direction (Astumian, 1997). In the power-stroke model, the motor uses the energy of ATP binding/hydrolysis or product release to directly drive the motion. Given a lower bound of \(\sim 20\) pN for the stall force and a step-size of 1 nm, the work performed by a single ClpX subunit near stall is \(\Delta W_{\text{subunit}} = 20\) pN\cdot1 nm = 5 \(k_B T\) per hydrolyzed ATP. Similar calculations for *E. coli* RNA polymerase (Brownian ratchet) and \(\phi 29\) ATPase (power-stroke motor) yield near-stall work values of \(\sim 2\) \(k_B T\) and \(\sim 10\) \(k_B T\), respectively (Moffitt et al., 2009; Neuman et al., 2003; Smith et al., 2001.) Because the maximum work done by a ClpX subunit is \(>5\) \(k_B T\), we favor the power-stroke model over the Brownian ratchet for ClpX. Moreover, we estimated that the free energy of hydrolyzing one ATP molecule in our buffer is \(\sim 55\) pN\cdotnm. Therefore, the maximum thermodynamic efficiency of ClpX per step is 20 pN\cdotnm/55 pN\cdotnm = \(\sim 35\)\%, within the efficiency range of other power-stroke motors (Smith et al., 2001). Future experiments, in which the concentration of ATP and hydrolysis products are changed systematically under varying external force, will be required to identify the force-generating step of the mechanochemical cycle and will help confirm a power-stroke mechanism for ClpX.

**Mechanochemistry and Minimal Kinetic Cycle of ClpX**

To better understand the mechanochemistry of ClpX, we analyzed the motor’s pause density (number of pauses per amino acid translocated) and pause duration during translocation of the unfolded polypeptide in two opposing force regimes, 5-12 pN and 12-20 pN. We found that higher force loads increased the probability of ClpX entering a pause state by a factor of two from 0.025 \(\pm 0.005\) nm\(^{-1}\) to 0.045 \(\pm 0.008\) nm\(^{-1}\) (\(p = 0.008\), Table 2.2). On the other hand, the pause duration (which reflects the probability to exit a pause state) was not affected by high opposing force load (\(p = 0.99\), Table 2.2). The dissimilar effect of the external force on pause entry versus exit can be explained if the transition state, \(x^\ddagger\), is located very close to the pause state (Tinoco and Bustamante, 2002). For a displacement (\(\Delta x\)) between the active and pause states, this means that \(\Delta x - x^\ddagger \approx 0\). Calculation of \(x^\ddagger\) based on the pause
density distribution (Figure 2.2D) allowed us to estimate $\Delta x \approx x^\dagger = 1.7 \pm 0.1 \, \text{Å}$ (mean ± standard deviation [SD], Table 2.2). Analysis of pause durations shows that they are distributed according to a single exponential ($k = 2.3 \pm 0.6 \, \text{s}^{-1}$, $R^2 = 0.99$), indicating that exiting from the pause state involves a single kinetic event.

In order to establish whether or not pauses are states off the main translocation pathway of ClpX, we analyzed the natural fluctuations of the pause-free velocity in a force range that does not affect the rate of motor translocation and calculated the correlation between the translocation rate and the probability of entering a pause (pause density in units of 1/nm). Between 5-12 pN, the translocation rate and pause density were negatively correlated ($R^2 = -0.5$, Table 2.2). Thus, we observe an increase in the number of pauses as the pause-free velocity decreases (Figure 2.2E), indicating that pausing and translocation compete kinetically. This type of kinetic competition is expected when pauses are states off the main translocation pathway. These results also explain the increase in pause density for forces between 12-20 pN, as the force-induced reduction of the motor velocity increases the pause entry probability. The analysis of pause density, pause duration, and their force dependence was statistically indistinguishable between ClpX and the ClpXP complex (Table 2.2). Based on these results, we propose the kinetic cycle shown in Figure 2.2F for the translocation of unfolded polypeptide by ClpXP, as well as the effects of mechanical forces on the various steps.

**ClpX Unfolds GFP via a Well-Defined Intermediate**

Another fundamental question about AAA+ unfoldases is whether denaturation of single-domain substrates is primarily determined by the protein’s energy landscape and occurs in a single cooperative unfolding transition or proceeds through several unfolding events depending on the presence of distinct mechanical barriers along the unfolding trajectory. To address this question, we analyzed the extensions of all rips for the three different GFP-titin substrates (Figure 2.1). Unfolding trajectories of the fusion construct with a single GFP molecule showed mainly two or three rips, each one followed by translocation of unfolded polypeptide chain (Figure 2.1B). On average, we observed twice as many rips throughout the unfolding trajectory for substrates with two GFP molecules (Figure 2.1C and 2.1D). From all these trajectories, we clearly recognized identical rip extensions that indicate a common unfolding signature of GFP (Figure 2.1, red arrow-heads). In fact, a histogram of the rip extensions (in number of amino acids) obtained from all GFP substrates revealed a higher probability for a transition centered at 207 ± 2 aa (mean ± SEM, Figure 2.5A, dark red bars). In addition to this 207 aa rip, we observed rips whose extensions are less regular and vary between 20 and 300 aa, with the highest probability at 37 ± 6 aa (mean ± SEM, Figure 2.5A, pink bars). We interpret these irregular extensions as slipping events of ClpX backward along an already unfolded polypeptide (black arrowheads, Figure 2.1).

Given the extension of the folded GFP molecule ($X^F$), and the experimentally observed extension change upon GFP unfolding ($\Delta X_{\exp(F-U)}$), the true extension of the unfolded GFP ($X^U$) can be calculated from $\Delta X_{\exp(F-U)} = X^U - X^F$ (Figure 2.7, upper portion).
Based on the crystal structure (Ormø et al., 1996; Yang et al., 1996), the extension of folded GFP in our experimental geometry is 2.4 nm, equivalent to 13 ± 1 aa in the force range of 6-10 pN. Thus, the $\Delta X_{\exp}(F-U) = 207$ aa rip corresponds to the unfolding of $220 ± 3$ aa (mean ± SEM). This result is in excellent agreement with the number of amino acids that show well-defined secondary structures in GFP (residues 5 to 227, Figure 2.9A). To corroborate the assignment of GFP unfolding events in our trajectories, we used the specific extension signature of the construct with two Ti<sub>cm</sub> domains inserted between two GFP molecules. This long stretch of unfolded polypeptide (~200 aa) served as an independent internal marker to identify the unfolding event of the first C-terminal GFP before the long translocation (Figure 2.1D, red arrowheads). When we analyzed the size distribution of those rips immediately preceding the translocation of the long unfolded polypeptide, we observed a peak at 207 aa (Figure 2.4), thus corroborating our previous structural assignment.

Close inspection of the GFP unfolding events showed that the 207 aa rip consisted of two steps, separated by a transient intermediate with a lifetime of ~180 ms (Figure 2.5B, black arrow). In fact, this rip-transition-rip signature (Figure 2.5C) was present in > 70% of all events of the 207 aa peak. The presence of a well-defined transition indicates that the ClpX-induced mechanical unfolding of GFP from the C terminus proceeds via a short-lived intermediate state. Furthermore, a plot of the size distribution of each rip segment revealed that the first rip is 107 ± 2 aa, whereas the second portion consists of 100 ± 2 aa (mean ± SEM, Figure 2.5D and 2.5F, respectively).

In order to estimate the secondary structures involved in such an intermediate, we mapped the first and second segments of the 207 aa rip onto the tertiary structure of GFP. Given the dimension of fully native GFP ($X^F$), the dimension of the folded portion of the unfolding intermediate ($X^I$), and the experimentally observed extension change upon the unfolding of the first, C-terminal GFP segment ($\Delta X_{\exp}(F-I)$), we can compute the true extension ($X^{U1}$) corresponding to the first rip using the following relation: $\Delta X_{\exp}(F-I) = X^{U1} - (X^F - X^I)$ (Figure 2.7). Then $X^{U1}$ can easily be converted into amino acids via the WLC formalism. Because the dimensions of the remaining folded structure (i.e., the

![Figure 2.4](image_url): The total rip size distribution peaked at ~210 aa in agreement with the expected contour length change upon GFP unfolding.
Figure 2.5: (A) The distribution of rip sizes (in aa) for all observed GFP unfolding events (red, n = 107) and slips (light red, n = 250) in ClpX and ClpXP traces. (B) ClpX trace containing the unfolding of the first GFP in the double GFP construct with the long linker (Figure 2.1D). The unfolding event is followed by the translocation of the unfolded GFP and the long linker. Raw data filtered and decimated to 800 Hz. (C) A detailed view of the short-lived GFP unfolding intermediate (black arrow in panel B). (D) Histogram of the first rip size during GFP unfolding. (E) The distribution of GFP unfolding intermediate lifetimes is well described by a single exponential. Dashed lines represent 95% confidence interval of the fit. (F) Histogram of the second rip size during GFP unfolding. For the rip size distribution before the long translocation marker (panel B), see Figure 2.4, and for side-by-side comparison of ClpX and ClpXP unfolding of GFP, see Figure 2.6.
unfolding intermediate) are not known, we defined the lower and upper bounds to $X^I$ as 2.4 and 4.2 nm, which correspond to the short and long axes of the folded GFP. In the force range of 6-10 pN, these numbers translate into 13 and 23 aa, respectively. Therefore, the first rip of 107 aa corresponds to the unfolding of a C-terminal GFP segment anywhere between 97 and 107 residues, leaving between 120 and 130 N-terminal residues still folded (see Figure 2.9A).

Based on the topology map of GFP, an unfolding intermediate with the N-terminal residues 1-120 still folded would require part of $\beta$ strand 6 to be unstructured. We do not favor this intermediate boundary because unfolding of $\beta$ strands is highly cooperative, as observed in previous mechanical unfolding experiments (Marszalek et al., 1999). In contrast, residue 130 is located within a long loop, near the end of $\beta$ strand 6. Thus, we surmise that ClpX unfolds 97 aa from the C terminus corresponding to $\beta$ strands 7-11, generating a GFP unfolding intermediate with $\sim$130 N-terminal residues still structured ($\beta$ strands 1-6).

In order to corroborate the structural elements assigned to the unfolding intermediate at the N terminus of GFP, we analyzed the second segment of the 207 aa rip (Figure 2.5C). As described above, we estimated the dimensions of the folded portion of the GFP intermediate (XI) to be $\sim$4.2 nm or 23 aa at 6-10 pN. Calling $X^{U2}$ the true extension corresponding to the second rip (the size of the unfolding intermediate when unfolded), we can calculate the observed change in extension upon unfolding of the intermediate $\Delta X_{\exp(I-U)}$ using the following expression: $\Delta X_{\exp(I-U)} = X^{U2} - X^I$ (Figure 2.7). Thus, $X^{U2} = 100$ aa + 23 aa = 123 aa. Unfolding 123 amino acids starting at residue 130 would end at position 7 (as counted from the N terminus), which is in excellent agreement with structural data for GFP that show Glu5 to be the first residue involved in secondary structures (Figure 2.9A). Analysis of the structural transitions of GFP when unfolded by the ClpXP protease is essentially indistinguishable from that of ClpX alone (Figure 2.6).

In addition to providing a structural assignment for the unfolding of GFP, we also sought to investigate the dynamic components during GFP unraveling by ClpX. The distributions of the dwell times preceding the first and second rip segments are well described by a single exponential (Figure 2.8A and 2.5E, respectively). However, the time constant for the first rip is nine times longer than for the second one (1.7 ± 0.3 s versus 0.18 ± 0.03 s). We interpret the time constant of the first rip dwell as the average time required by ClpX to destabilize and unravel the fully folded GFP molecule. In contrast, the second rip time constant corresponds to the time required for the spontaneous unfolding of the 130 N-terminal residues of GFP. During the 180 ms lifetime of this N-terminal intermediate, ClpX would be able to translocate only about 15 residues of the chain unraveled in the first step of GFP unfolding. Therefore, by the time the N-terminal portion of GFP unfolds, ClpX is still several nm away and unlikely to play a major role in this second stage of GFP unfolding.

Interestingly, for the ClpXP complex, the time constant for the first stage of GFP unfolding was about five times longer compared to that for ClpX alone (9.1 ± 1.4 s versus 1.7 ± 0.3 s, Figure 2.8B and 2.8A, respectively). The former value is similar to the previously reported time constant $t = 5.6$ s for the loss of GFP fluorescence in single-turnover degradation (Martin et al., 2008b). These degradation experiments had also suggested that ClpX
initiates GFP unfolding by extracting the C-terminal β strand 11 and trapping it through at least four subsequent translocation steps. This requirement for rapid translocation leads to a strictly nonlinear dependence of GFP unfolding on the rate of ClpX ATP hydrolysis. In fact, reducing the ATP hydrolysis rate by 30% was found to decrease GFP degradation ~3-fold (Martin et al., 2008b). Therefore, our observed 5-fold difference in unfolding rate between ClpX and ClpXP is consistent with the ~30 percent lower translocation velocity of ClpXP at all forces observed in our experiments (Figure 2.2A), which is, in turn, a consequence of
Figure 2.7: (F), (I), and (U) denote the folded, intermediate, and unfolded conformations of GFP. ∆X quantities correspond to experimentally observed extension changes (obs). $X^F$ and $X^I$ are the dimensions of the folded and intermediate states of GFP, estimated from structural data (struct). $X^{U_1}$ and $X^{U_2}$ are the true extensions (true) corresponding to the unfolding of the first and second portions of GFP (shown as flexible chains in green and purple, respectively). The true extension of the entire unfolded GFP molecule is $X^{U_1} + X^{U_2}$. The table provides reference values estimated for 6-10 pN. Bottom right panel displays a characteristic rip-transition-rip during GFP unfolding.
the reduced ATP hydrolysis rate and pulling frequency of ClpX when bound to ClpP.

**ClpP Enhances the Unfolding Activity of ClpX**

Besides rips corresponding to successful GFP unfolding, the trajectories for all fusion substrates showed frequent slippage events, in which ClpX apparently failed to unfold GFP, disengaged the substrate, and moved backward along the polypeptide track (Figure 2.1, black arrowheads). Such behavior is consistent with previous degradation studies that have shown that hard-to-unfold substrates with a short ssrA tag are frequently released and re-bound by ClpXP before successful unfolding (Kenniston et al., 2005). We can imagine two possible mechanisms by which ClpX can resume tugging at the substrate after a failed unfolding attempt. Either it can remain engaged to the substrate, making an immediate new unfolding attempt, or it can completely disengage the polypeptide and diffuse backward for a short time before re-engaging the substrate. The first scenario is beyond the spatiotemporal resolution of the present experiments. The second scenario should manifest itself as rips of several amino acids due to the applied opposing load. We therefore sought to investigate in greater detail the motor slips observed during GFP unfolding.

The probability distribution of slip extensions for ClpX peaked at \( \sim 30-40 \) aa but also displayed longer slips well beyond 50 aa (Figure 2.8C, lower panel). Because ClpX alone cannot hydrolyze polypeptides, it can slip backward up to the entire length of the already translocated polypeptide chain. Remarkably, the observed slip distribution for the ClpXP complex also peaked between 30-40 aa; however, it lacked longer slips (Figure 2.8C, upper panel). This absence of longer slips for ClpXP is expected, as the ClpP peptidase constantly trims the polypeptide chain inside its proteolytic cavity down to \( \sim 38 \) aa, a length sufficient to span the distance between the ClpX-pore entry and the ClpP active sites (Martin et al., 2008b). Slips of ClpXP longer than 38 aa result, therefore, in tether rupture. The fact that the distribution of slip sizes peak between 30-40 aa for both ClpX and ClpXP indicate that the time ClpX takes to re-engage the substrate is the same with or without ClpP. The size of slips depends on the product of the substrate re-engagement time by the motor and the speed at which the polypeptide is dragged out of the pore. Thus, the presence of an external force in our experiments greatly amplifies the sizes of the slips observed here relative to those that occur in the cell and in bulk assays (Martin et al., 2008a).

Importantly, we found that ClpXP complexes are much less prone to slipping compared to ClpX hexamers alone: for ClpX, 70% of the rips observed correspond to slips and the rest to unfolding events, whereas for ClpXP, this number is only 27%. Thus, binding of ClpP dramatically decreases the slip frequency of ClpX, potentially due to additional contacts between the polypeptide and the extended processing pore of ClpXP. These additional interactions may prevent substrate release after an unsuccessful unfolding attempt.

We also analyzed the slip entry rate as a function of force for ClpX alone and in complex with ClpP. The slip entry rate is defined as the inverse of the average waiting time before a slip occurs in front of a mechanical barrier. We found that the slip entry rate for ClpXP is 5-10 times lower than that of ClpX at all forces (Figure 2.8D). For instance, when using
Figure 2.8: (A and B) The dwell time before each ClpX- or ClpXP- mediated unfolding event is exponentially distributed. Dashed lines represent 95% confidence interval of the fit. (C) Slip size histograms for ClpX (red) and ClpXP (blue). (D) The slip entry rate (mean ± SEM) for ClpX (red) and ClpXP (blue). Slip entry rate is the inverse of the average waiting time before a slip occurs.

the double GFP + long linker substrate, ClpXP was able to process both GFP molecules in 45 % of all traces, whereas ClpX alone was successful in only 10 %-15% of the cases. These numbers are in good agreement with the probabilities of successful unfolding versus slips determined above. For ClpX, the probability of unfolding both GFP molecules within a single trace is \(0.27 \cdot 0.27\) (~10%), whereas for ClpXP it is \(0.70 \cdot 0.70\) (~50%). Because in the cell there is no opposing force on the substrate, we surmise that the probabilities described above would correspond to a lower bound for the successful unfolding and translocation of this tandem substrate.
CHAPTER 2. CHARACTERIZING THE AAA+ MOLECULAR MACHINE CLPXP

2.4 Discussion

Our studies provide direct experimental evidence that ClpX is able to generate mechanical force to induce protein denaturation. Moreover, the force dependence of the motor velocity
favors a power-stroke mechanism for ClpX. Although the stall force of the motor is just 20 pN, this force should be sufficient to unfold most cellular proteins, given the low loading rates at which AAA+ unfoldases may operate in the cell. The loading rate is a measure of the speed at which the force is applied to the protein substrate. Because protein unfolding always involves the stochastic thermal crossing of an energy barrier, pulling at lower loading rates gives the substrate more time to unfold at low forces. The most likely force at which a protein unfolds scales as the log of the loading rate (Bustamante et al., 2004). We estimate that the loading rate at which ClpX pulls the folded substrate in our experimental conditions is $\sim 0.15$ pN/s (Extended Experimental Procedures) and is five orders of magnitude smaller than the loading rate in mechanical protein. Given that typical protein unfolding forces in AFM experiments are between 50-150 pN, we expect that the corresponding unfolding forces should be at least five times smaller under the loading rates applied by ClpXP. Although the motor is hence strong enough to unfold most proteins mechanically, it is possible that ClpX also uses allosteric contacts to distort the local structure of its substrates and thereby weaken their thermodynamic stabilities.

The single-molecule trajectories obtained here allow us to address two additional important questions about the motor properties of ClpX, namely, the mechanisms involved in motor slow-down and motor stall under force. First, why does the motor pause-free velocity decrease at forces above 13 pN. One possible explanation is that the external force slows down the mechanochemical cycle of the motor and decreases the frequency of pulling events. Alternatively, high forces could cause the translocation loops of individual subunits of ClpX to slip on the substrate, resulting in futile translocation attempts that fail to move the substrate through the central pore. We favor the former explanation because in $\geq 98\%$ of all traces we do not observe any small rips that may be indicative of motor slippage during translocation. In fact, even if we could not resolve these rips, their presence should manifest itself as an increased noise in velocity at high force, which we do not observe.

The second question concerns the mechanism of motor stalling. Is the maximum force generated by the motor equivalent to its thermodynamic stall force (Bustamante et al., 2001) or an operational stall force at which the motor is rendered unable to translocate on its track (due, for example, to mechanical unfolding of the motor itself, slippage on the track, etc.). At forces around 20 pN, near the stall, a large fraction of the traces ($>80\%$) are interrupted by tether ruptures that appear to be caused by backsliding of the entire, already translocated polypeptide out of the ClpX(P) motor pore. In support of this interpretation we observed that the motor slip entry rate increases significantly from $\sim 1 \text{ s}^{-1}$ at forces $< 13 \text{ pN}$ to $\sim 3.5 \text{ s}^{-1}$ at forces near the stall (Figure 2.8D, $p = 0.006$, Kolmogorov-Smirnov test).

Thus, it seems likely that the reduction of the motor’s ATPase rate and frequency of pulling events at higher opposing loads increases the chance of complete loss of grip and disengagement of the substrate, causing an operational stall.

ClpX(P) maintains a constant extension velocity (nm/s) up to forces of 13 pN (Figure 2.2A). At low forces ClpX translocates $\sim 8$ aa per 1 nm step compared to only $\sim 4$ aa/step at 13 pN. Because the number of residues translocated per step changes as a function of external force, the 1 nm step-size of ClpX must be dictated by the relevant conformational
change of the motor during the power stroke rather than any spatial periodic features of the substrate. Moreover, because of the chemical heterogeneity of the track, it is likely that nonspecific steric contacts between ClpX pore loops and the substrate are more important than specific chemical interactions when threading an unfolded polypeptide through the central pore. These nonspecific steric contacts might be key to ClpX’s ability to translocate its irregular and diverse polypeptide substrates in either C-to-N or N-to-C direction (Barkow et al., 2009; Martin et al., 2008a).

Our analysis of the ClpX-induced unfolding of GFP in the C-to-N terminal direction revealed a well-defined, short-lived intermediate (Figure 2.9). This intermediate, comprising the N-terminal 130 residues that form β strands 1 thru 6, is in good agreement with previous mechanical unfolding experiments of GFP (Bertz et al., 2008; Perez-Jimenez et al., 2006). Importantly, ClpX and ClpXP lead to the formation of the same transient intermediate (Figure 2.5E and 2.6H), indicating that the unfolding mechanism is determined largely by the energy landscape of the substrate and the presence of cooperative folding units that unravel after ClpX or ClpXP disrupt critical interactions at the C terminus. AAA+ unfoldases target structurally and functionally diverse proteins in all cells. Moreover, their client proteins are found not only in a folded, soluble conformation, but also in hyperstable misfolded or aggregated states (Horwich, 1999). These molecular machines must therefore utilize efficient mechanisms to unravel proteins with a wide range of thermodynamic stabilities, topologies, and sequence characteristics. The present study shows that ClpX(P) is able to generate and apply mechanical forces sufficient to unfold most target proteins.

ClpX processes substrates in a linear fashion, applying force and overcoming only the local mechanical barriers encountered along the unfolding trajectory of the protein. As a result, it is the linear profile of these barriers, as defined by the pulling end and the topology of the substrate, but not the global protein stability that determines the kinetics of substrate processing. Because the chance of slipping on the polypeptide track increases significantly near the stall force, ClpX will work on a hard-to-unfold substrate much longer, repeatedly tugging and slipping until an unfolding attempt is successful. Mechanical unfolding ultimately involves the thermally induced crossing of an energy barrier. By maintaining a constant tugging on a hard-to-unfold substrate, ClpX decreases the magnitude of this barrier while increasing the chance that sooner or later a thermal fluctuation within the protein substrate will allow its crossing. When facing a high mechanical barrier, the motor will thus require more time and consume larger amounts of ATP before such spontaneous crossing occurs. In this way, the motor is able to process proteins with a wide range of thermodynamic and mechanical stabilities.

As the ClpX motor shares its basic design and operating principles with other AAA+ unfoldases, including the prokaryotic ClpA, ClpB, HslU, FtsH, or Lon and the eukaryotic 26S proteasome, it is conceivable that all these enzymes utilize very similar mechanisms to generate mechanical force and disrupt the secondary, tertiary, and quaternary structures of their protein substrates. However, it remains to be determined how differences in the rate of ATP hydrolysis, the length of the central processing channel, or the heterohexameric versus homohexameric architecture of the AAA+ unfoldases affect pulling forces, translocation
velocities, and the frequency of pausing or slipping on hard-to-unfold substrates.

2.5 Methods

Protein Modifications and dsDNA-Handle Attachment to Protein Substrates

Single-chain ClpX hexamers and GFP-titin I27 fusion proteins were expressed and purified as described previously (Martin et al., 2005; Martin et al., 2008a). In all fusion substrates, the titin I27 domains were permanently unfolded (Ti_{cm}) by carboxymethylation (Martin et al., 2008a). Single-chain ClpX hexamers included an avi tag that was biotinylated using purified BirA in vitro (Chen et al., 2005). We covalently attached a 3 kbp dsDNA handle to the N termini of the GFP-Ti_{cm} fusion proteins by utilizing the ybbR tag/Sfp system (Yin et al., 2005). A detailed protocol for the dsDNA-handle attachment to the protein substrates is described in Extended Experimental Procedures (Figure 2.10A and 2.10B).

Single-Molecule Sample Preparation

All single-molecule unfolding trajectories were obtained in ClpX-100 buffer (25 mM HEPES-KCl, pH 7.4, 20 mM MgCl2, 100 mM KCl, and 1 mM EDTA), 5 mM ATP and ATP regeneration system (16 mM creatine phosphate and 32 mg/ml of creatine phosphokinase) (Kenniston et al., 2005). Before an experiment, the buffer was passed through a 0.22 mm pore filter and degassed thoroughly. In the presence of ATP-\gamma-S, tethers between ClpX and the DNA-tethered substrate were obtained but no rips or continuous decrease in extension were observed. Details of the control experiments with ATP-\gamma-S are described in Extended Experimental Procedures (Figure 2.12). Tethers in the absence of ATP or ATP-\gamma-S were not observed. Experiments conducted with ClpP contained 500 nM in all chambers to ensure formation of the ClpXP complex (Kd = 90 nM) (Joshi et al., 2004).

Single-Molecule Data Collection

Data was collected in a dual-trap instrument with differential detection (Moffitt et al., 2006). Raw single-molecule data were acquired at 2000 Hz. Sudden extension changes and translocation (in nm) were converted into polypeptide contour length (in amino acids) in a two-step procedure. We first removed the extension contribution of the 3 kbp dsDNA handle using the WLC formalism (Extended Experimental Procedures). We then calculated the unfolded polypeptide contour length using the WLC model and P_{peptide} = 0.65 nm (Cecconi et al., 2005).
Single-molecule Data Analysis

During data acquisition, we monitored the dsDNA handle + polypeptide extension. For our 3 kbp dsDNA handle, the expected extension is between 930-1000 nm in the force range of our experiments (4-20 pN, respectively). Any tethers with extensions out of the expected range were discarded. During data analysis we had an additional internal control-the characteristic signature of GFP unfolding by ClpX(P): a double rip with a short-lived intermediate in between (Figure 2.5C). If the two rips did not add up to the expected ∼200 aa change in contour length, we discarded the trace. A total of 101 ClpX traces and 62 ClpXP traces passed our screening and were used in further analysis. We analyzed three distinct events (Figure 2.1E): the time before a sudden extension gain occurs, the length of the sudden change in extension (rip) caused by substrate unfolding or motor slippage, and the continuous decrease in length due to translocation of the unfolded polypeptide (translocation velocity). In the analysis of pause durations before a sudden rip, we only included pauses after the first rip is observed; that is, pauses preceding rip events 2, 3, and so on within a single trajectory. This is due to the uncertainty in the time at which a tether between ClpX and the substrate was formed and the time before the first rip was observed. In order to locate the occurrence of sudden rips and their extensions, we analyzed our traces using the Student’s t test (Figure 2.12). For velocity calculation, data were filtered and decimated to 2.5 Hz, and any pauses longer than 1 s were removed. Velocity data for a given force range were computed for each molecule. Velocity values corresponding to different molecules were used to compute the mean and standard error. The calculated velocity was essentially the same for filter bandwidths ranging from 1 Hz to 10 Hz.

2.6 Extended Experimental Procedures

Calculations Related to the Dissimilar Effect of Opposing Force on Pause Entry and Pause Exit

For a transition state $x^\dagger$ located closer to the pause state, the pause-entry rate ($k_p$) shows a stronger force-dependence than the pause-exit rate ($k_{-p}$). This is illustrated in first-order kinetics, considering the dependence of the rate constant on applied force (Tinoco and Bustamante, 2002):

$$\frac{d}{dF}(\ln k) = \frac{x^\dagger}{k_B T} \quad \text{Eq.1}$$

If we consider a low and a high force (5 and 15 pN, respectively) and $\Delta x$ the displacement between the active and paused states, we obtain the following:

$$\text{PauseEntry} : \frac{k_p(F_{\text{low}})}{k_p(F_{\text{high}})} = \exp\left(-\frac{F_{\text{low}} - F_{\text{high}} \cdot x^\dagger}{k_B T}\right) = \exp\left(-\frac{10 \cdot x^\dagger}{k_B T}\right)$$

$$\text{PauseExit} : \frac{k_{-p}(F_{\text{low}})}{k_{-p}(F_{\text{high}})} = \exp\left(-\frac{F_{\text{low}} - F_{\text{high}} \cdot (\delta x - x^\dagger)}{k_B T}\right) = \exp\left(-\frac{10 \cdot (\Delta x - x^\dagger)}{k_B T}\right)$$
Thus, in order to observe force-dependence in pause density \( \left( k_p \right) \) but not in pause exit \( \left( k_{-p} \right) \), as it is observed for ClpX (Table 2.2), we require that \( \Delta x - x^\ddagger \approx 0 \). In order to estimate the distance to the transition state between the active and paused configurations, \( x^\ddagger \), we replaced the pause entry rate, \( k_p \), into Eq. 1 (see above) for forces above 12 pN. The pause entry rate, \( k_p \), was estimated by multiplying the pause-free velocity (proportional to the translocation rate constant or \( k_T \)) times the pause density.

<table>
<thead>
<tr>
<th>ClpX</th>
<th>Pause density (1/nm)</th>
<th>Pause duration (s)</th>
<th>Spearman Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-12 pN</td>
<td>0.0025 ± 0.005 (n=170)</td>
<td>0.69 ± 0.19 (n=63)</td>
<td>( R^2 = -0.5 )</td>
</tr>
<tr>
<td>12-20 pN</td>
<td>0.045 ± 0.008 (n=36)</td>
<td>0.65 ± 0.19 (n=20)</td>
<td>—</td>
</tr>
<tr>
<td>( P )-value</td>
<td>0.008</td>
<td>0.99</td>
<td>( 10^{-9} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ClpXP</th>
<th>Pause density (1/nm)</th>
<th>Pause duration (s)</th>
<th>Spearman Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-12 pN</td>
<td>0.024 ± 0.003 (n=133)</td>
<td>0.60 ± 0.03 (n=89)</td>
<td>( R^2 = -0.6 )</td>
</tr>
<tr>
<td>12-20 pN</td>
<td>0.057±0.013 (n=38)</td>
<td>0.67±0.04 (n=26)</td>
<td>—</td>
</tr>
<tr>
<td>( P )-value</td>
<td>0.005</td>
<td>0.99</td>
<td>( 10^{-13} )</td>
</tr>
</tbody>
</table>

Table 2.1: Effect of opposing force on pause density and duration during polypeptide translocation. The Kolmogorov-Smirnov test was used to determine the statistical significance difference between the pause density and duration at low and high force loads.

**Attachment of a dsDNA handle to the protein substrates harboring a ybbR tag. Related to Figure 2.10A and 2.10B.**

This handle attachment process includes two major steps. First, the enzyme-catalyzed reaction by Sfp covalently links a CoA moiety (modified with a short dsOligonucleotide of 24 bp) to the first serine residue in the ybbR tag (sequence: DSLEFIASKLA). Second, the dsOligo-CoA-substrate is ligated to a complementary 3kbp dsDNA modified at its opposite terminus with a 5’ digoxigenin. Below we explain in detail the covalent attachment of a 3kbp dsDNA handle to the N terminus of the GFP-Ti\( _{cm} \) fusion protein utilizing the ybbR tag/Sfp system (Yin et al., 2005). Identical results were obtained for the other two substrates used in this study (Figure 2.1).

** Modification of a short dsOligo with a CoA moiety**

A Coenzyme-A (CoA) moiety was introduced at the 5’-terminus of a 24-nucleotide long oligonucleotide by chemical crosslinking. This CoA-modified oligonucleotide was annealed to a complementary oligonucleotide with a 5’-overhang of 4 nucleotides. In the resulting
double-stranded oligonucleotide, one 5’-terminus harbors a CoA moiety; the other provides a ”sticky” end of 4 nucleotides.

Specificity of CoA attachment to the ybbR tag in our substrates

We ensured that the enzymatic linkage of a CoA moiety to the different protein substrates was specific at the ybbR tag and not to any other part in the protein sequence. To do so, we performed control experiments in which we used the same exact protein substrate but without the ybbR tag and did not see any attachment (data not shown).

dsDNA Oligo-CoA attachment to protein substrates

We then proceeded to covalently link the dsOligo (modified at the 5’ end with a CoA moiety) to the ybbR tag on our different protein substrates. This short oligo contains a restriction site that is also present in a 3-kbp dsDNA, which was used as molecular handle. We monitor the yield of the attachment reaction by changes in shift mobility in SDS-PAGE (Figure 2.10-A).

Ligation of a 3-kbp dsDNA handle to the oligoCoA-substrate and degradation test of the DNA-tethered substrate by ClpXP

After the specific covalent attachment of the oligo-CoA to the ybbR tag on our substrates, we ligated (using T4 Ligase, NEB) this oligo-substrate to a 3-kbp dsDNA handle. This handle contained a 5’ digoxigenin at the opposite end of the ligation site for bead attachment (Figure 2.10-B). To analyze the success of the ligation reaction, we monitored the differential mobility of the ligated product vs. the free DNA handle. We increased the differential mobility between the ligation product and the unligated DNA by truncating the DNA handle with Hind III to 785 bp (Figure 2.10B). We observed that the ligation product migrated at 1.1 kbp, while the DNA-handle migrated to its expected size of 785 bp. Moreover, we also tested that the dsDNA-tethered substrate is efficiently degraded by the ClpXP protease in ATP (5 mM) and an ATP regeneration system (16 mM creatine phosphate and 32 µg/ml of creatine phosphokinase). We observed that upon addition of ClpXP and ATP, the intensity of the 1.1-kbp band decreased by more than 80 % (detected by Ethidium Bromide or by observing the fluorescence of GFP) (Figure 2.10-B). Moreover, we noticed that the band corresponding to the ClpXP degradation product is larger than the free DNA handle fragment (Figure 2.10-B, blue arrow). This observation indicates that the substrate chain was not fully degraded. This suggests that ClpX cannot translocate DNA, leaving a short polypeptide tail that corresponds to the distance between the entrance of the ClpX pore and the proteolytic sites inside the ClpP protease.
Figure 2.10: (A) (Top) Schematic representation of the short oligo-CoA and the ybbR tag in the single GFP-Ti\textsubscript{cm} fusion substrate. The compact green structure represents GFP, the red is the permanently unfolded Titin (Ti\textsubscript{cm}) and in blue and black are the ybbR and ssrA tags, respectively. (Bottom) Modification of the protein substrate with the dsOligo-CoA. We monitored the degree of the modification by shift mobility in a 10% SDS-PAGE at time 0 and 90 min after adding Sfp (at 37°C). For reference, we included loading controls of Sfp and the protein substrate. Approximately 80% of the protein substrate was modified with the oligo-CoA. (B) (Top) Schematic representation of the DNA-tethered substrate and the position of the HindIII restriction site in the DNA handle. (Bottom) Ligation of the dsOligo-CoA-substrate to the 3-kbp dsDNA handle detected by shift mobility of the DNA (left) or by GFP fluorescence (right). The ligation product migrated at 1.1 kbp (green arrow), while the (truncated) DNA handle migrated at 785 bp. Degradation of the ligation product by ClpXP was confirmed by a decrease in the band intensity (lane 2).

**Activity Assays of Single-Chain ClpX Hexamers Immobilized on 2.1 µm Polystyrene Beads**

We performed bulk ATPase and substrate degradation control experiments to ensure that the biotinylated ClpX hexamers immobilized on the surface of streptavidin-coated beads (Spherotech, Inc.) were structurally intact and active. After incubating biotinylated ClpX and streptavidin-coated beads for 10 min, we washed the bead suspension thoroughly with ClpX-100 buffer (25 mM HEPES-KCl, pH 7.4, 20 mM MgCl\textsubscript{2}, 100 mM KCl, and 1 mM EDTA) to remove unbound ClpX molecules (final dilution factor of unbound ClpX molecules after washing was \( \approx 10^6 \)). We then added identical amounts of ClpX beads into a solution containing a NADH-coupled regeneration system (3 U/ml pyruvate kinase, 3 U/ml lactate dehydrogenase, 1mM NADH, and 7.5 mM phosphoenol pyruvate). The ClpX ATPase activity
was measured in the presence of ClpP, Ti$_{cm}$, and both ClpP and Ti$_{cm}$ by monitoring changes in NADH absorbance (340 nm) at 30°C. We observed that changes in ClpX ATPase activity upon addition of ClpP and/or Ti$_{cm}$ were consistent with previous reports (Martin et al., 2005; Martin et al., 2008b), indicating that the immobilized ClpX were fully active in hydrolysis, ClpP binding and substrate processing (Figure 2.11A and Table 2.1). In order to test for substrate degradation activity, we mixed immobilized ClpX with ClpP (µM), ATP (5 mM), and an ATP regeneration system (16 mM creatine phosphate and 32 mg/ml of creatine phosphokinase) (Kenniston et al., 2003; Kenniston et al., 2005) in ClpX-100 buffer, equilibrated for 5 min at room temperature, and then added the GFP-Ti$_{cm}$ fusion substrate. We followed substrate degradation by monitoring the decrease of GFP fluorescence in a native gel (Figure 2.11B). Finally we determined the fraction of active ClpX hexamers on the surface of the beads. To do so, we first experimentally determined the number of ClpX moles per bead mass and obtained a value of 0.075 pmoles of ClpX/mg bead (2.1 µm streptavidin-coated beads). By measuring the bead mass (using light scattering at 600 nm) we estimated the number of pmoles in a sample, and therefore the ClpX concentration for a given reaction volume. Table 2.1 summarizes the ATPase activity of immobilized ClpX under different conditions (i.e., in the presence of ClpP, protein substrate, or both), corrected by the experimentally determined ClpX concentration. We also included control experiments with ClpX free in solution (not immobilized). After correcting for the ClpX concentration in samples containing beads, we observed that the ATPase rates (per minute, per ClpX hexamer) are essentially indistinguishable between free ClpX and ClpX-SAPB, indicating that the majority of the single-chain ClpX hexamers immobilized on the surface of the beads are active.

<table>
<thead>
<tr>
<th></th>
<th>Free ClpX</th>
<th>ClpX-SAPB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>ATP (min$^{-1}$ X$^{-1}$)</td>
</tr>
<tr>
<td>ClpX</td>
<td>68</td>
<td>1.0X</td>
</tr>
<tr>
<td>+ClpP</td>
<td>31</td>
<td>0.5X</td>
</tr>
<tr>
<td>+Ti$_{(cm)}$-ssra</td>
<td>189</td>
<td>2.8X</td>
</tr>
<tr>
<td>+ClpP &amp; Ti$_{(cm)}$-ssra</td>
<td>117</td>
<td>1.7x</td>
</tr>
</tbody>
</table>

Table 2.2: ATPase activity of ClpX free in solution and immobilized on 2.1 um streptavidin-coated polystyrene beads (ClpX-SAPB).

Single-Molecule Control Experiments with ATP-$\gamma$-S

We performed control experiments in order to establish that any sudden gain or continuous decrease in extension, observed in our single-molecule trajectories, is due exclusively to the activity of individual single-chain ClpX hexamers in ATP. First, we formed a tether between
Figure 2.11: (A) ATP hydrolysis activity of ClpX monitored by a NADH-coupled assay under different conditions: X, XP, XP-S, and X-S denote curves for ClpX alone, ClpXP, ClpXP substrate, and ClpX substrate, respectively. (B) (Left) Degradation of GFP-Ti\textsubscript{cm}-ssrA by ClpX immobilized on streptavidin-coated beads (ClpX-SAPB), ClpP and ATP (lane labeled + ATP). In the absence of ATP, the GFP fluorescence is comparable to the substrate loading control (lanes labeled Background and No ATP, respectively). In the right part of the same gel, positive control experiment for ClpXP degradation activity were included, in which we added ClpX (not immobilized), ClpP, and ATP at different ClpX concentrations.

ClpX and the DNA-tethered substrate in a buffer containing a non-hydrolyzable ATP analog (5 mM ATP-\(\gamma\)-S) and collected force-extension data within our working force range (4-20 pN, loading rates 1-10 pN/s). We did not observe any rips that would indicate partial or complete unfolding of the GFP-Ti\textsubscript{cm} fusion proteins(Figure 2.12A). In the second control experiment, we held the ClpX-substrate tether in a buffer containing 5mM ATP-\(\gamma\)-S for several minutes at various forces between 4 and 16 pN. We did not observe any sudden changes in extension (Figure 2.12B). Thus, we concluded that the data in all our traces correspond specifically to ClpX activity in ATP.

Extension Contribution of the 3 kbp dsDNA Handle

We collected force-extension data with the 3 kbp dsDNA handle used in our experiments and measured a persistence length (P) of 25 ± 4 nm. While this value is smaller than the 50nm value often quoted in the literature, it is known that magnesium ions reduce the dsDNA persistence length(Baumann et al., 1997). In addition, it has been shown that the effective persistence length of dsDNA decreases for short tethers (Seol et al., 2007).

Calculation of the Loading Rate of ClpX

The loading rate is the product of the pulling velocity times the stiffness (\(\kappa\)) of the system. We estimated the effective stiffness of the combined polypeptide, dsDNA handles and optical
Figure 2.12: (A) Force-extension curves of tethers obtained between ClpX and two different substrates (single and double GFP) in the presence of 5mM ATP-γ-S. (B) GFP does not unfold at 4, 8, or 14 pN when ClpX is stalled in a 5 mM ATP-γ-S buffer (double GFP substrate was used). The ClpX-substrate tether was held at a given force for over 40 sec, which is longer than the duration of a typical ClpX(P) experiment in ATP. (C) The Student’s t test was used to reliably identify rips in ClpX(P) traces. Dotted black lines mark sudden raw extension changes in a sample ClpX trace. The raw data was filtered and decimated to 500 Hz (gray).
trap by calculating the instantaneous change in force divided by the instantaneous change in extension in our unfolding trajectories. We obtained a value \( \kappa = 0.05 \pm 0.01 \text{ pN/nm} \) between forces in the range of 5-13 pN. We used 3 nm/sec as the pulling velocity of ClpX before a structural barrier, which corresponds to the maximum instantaneous velocity of ClpX before a sudden rip occurs. Thus, ClpX loading rate is: \((0.05 \text{ pN/nm}) \cdot (3 \text{ nm/s}) = 0.15 \text{ pN/s}\).
Chapter 3

Studying mechanisms of force-generation and intersubunit coordination

Published as:

3.1 Summary

ATP-dependent proteases are vital to maintain cellular protein homeostasis. Here, we study the mechanisms of force generation and intersubunit coordination in the ClpXP protease from *E. coli* to understand how these machines couple ATP hydrolysis to mechanical protein unfolding. Single-molecule analyses reveal that phosphate release is the force-generating step in the ATP-hydrolysis cycle and that ClpXP translocates substrate polypeptides in bursts resulting from highly coordinated conformational changes in two to four ATPase subunits. ClpXP must use its maximum successive firing capacity of four subunits to unfold stable substrates like GFP. The average dwell duration between individual bursts of translocation is constant, regardless of the number of translocating subunits, implying that ClpXP operates with constant “rpm” but uses different “gears.”

3.2 Introduction

ATP-dependent proteases of the AAA+ family play crucial roles for the maintenance of cellular protein homeostasis, including the clearance of misfolded, aggregated, or damaged proteins, as well as the disassembly of large macromolecular complexes (King et al., 1996; Hanson and Whiteheart, 2005; White and Lauring, 2007a). These multimeric molecular
machines encounter unique chemical and mechanical challenges during their task of protein unfolding and polypeptide translocation. First, the motor must translocate along a heterogeneous and flexible polypeptide track composed of amino acids with highly diverse chemical and physical properties. Second, the motor must unravel a diverse array of folded protein domains with a range of stabilities that represent mechanical obstacles along the track. Understanding how ATP-dependent proteases perform these tasks will shed light on their general principles of operation and their mechanochemistry, specifically how they convert chemical energy from ATP hydrolysis to mechanical work and how individual subunits are coordinated.

*Escherichia coli* ClpXP is a well-characterized model system to investigate the operating principles of ATP-dependent proteases (Baker and Sauer, 2012). Homohexameric rings of the ClpX ATPase associate with the peptidase ClpP to form the ClpXP protease. ClpP is made up of two coaxially stacked homoheptameric rings with proteolytic active sites sequestered inside an internal cavity (Gottesman et al., 1998). To achieve specific protein degradation, ATP-bound ClpX recognizes protein substrates with certain degradation tags (Baker and Sauer, 2012) and uses cycles of ATP hydrolysis to unfold and translocate these substrates through its central pore into ClpP for proteolysis (Figure 3.1A) (Baker and Sauer, 2006). Conserved loops (pore-1 loops) that protrude from every ClpX subunit into the central pore have been proposed to directly contact the substrate (Martin et al., 2008a), and ATP-dependent conformational changes of these subunits are thought to unravel folded domains and propel the polypeptide through the central channel (Glynn et al., 2009).

Previous mutagenesis studies have suggested that the ClpX subunits contribute additively to substrate processing and that the power stroke for translocation may be generated by ATP hydrolysis in one subunit at a time (Martin et al., 2005), supporting a probabilistic mechanism of ring subunit coordination (Glynn et al., 2009; Stinson et al., 2013). Using optical tweezers, we and others provided direct demonstration that ClpX transforms the energy of ATP hydrolysis into mechanical force, and that polypeptide translocation occurs in cycles composed of a dwell phase, during which the substrate does not move, and a burst phase, during which ClpXP near-instantaneously translocates the polypeptide by a certain length (Aubin-Tam et al., 2011; Maillard et al., 2011).

Our previous findings motivated us to perform a mechanochemical characterization of ClpXP and to address the following questions: do all ATPase subunits participate during substrate translocation and what is the mechanism of coordination within the hexameric ring? Is the coordination among subunits different during protein unfolding versus processive translocation of an unstructured polypeptide? Which transitions determine the timing of the mechanochemical cycle? How is the chemical energy from ATP hydrolysis coupled to the mechanical cycle that drives translocation?

To address these mechanistic questions, we utilized single-molecule optical tweezers, which allow us to probe the motor’s mechanochemical coupling by applying external forces while simultaneously perturbing the chemical transitions of the ATPase cycle. These studies provided us with several important findings. We determined that to stall the motor during polypeptide translocation ATP hydrolysis in at least three of the six subunits must be inhib-
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

We found that a process not coupled to ATP binding sets the dwell duration between translocation bursts and that the burst size depends on the number of hydrolyzing ClpX subunits in the hexamer. This number distributes between two, three, and four subunits and their relative occurrence changes as ATP is varied from Km to saturation. During a burst, the near-simultaneous firing and translocation by two, three, or four subunits occurs in a coordinated fashion before the hexamer starts a new mechanochemical cycle. We find that these highly coordinated power strokes occur upon phosphate release and that they play a crucial role in the ability of ClpXP to denature kinetically stabilized protein substrates like GFP. Contrary to previously proposed probabilistic models (Martin et al., 2005; Glynn et al., 2009; Stinson et al., 2013), our results establish a high degree of coordination between ATP-bound subunits in the ClpX hexamer. ClpX seems to employ a novel mechanism of translocation that significantly deviates from canonical motor mechanisms, demonstrating how specialized molecular machines have been optimized to carry out their specific tasks.

3.3 Results

Single-molecule Assay

We used dual-trap optical tweezers in passive mode (constant trap position, variable force) to monitor a single ClpXP complex as it unfolds and translocates protein substrates in an ATP-hydrolysis-dependent manner. ClpXP was immobilized on one polystyrene bead and the ssrA-tagged protein substrate was attached to another. Each bead was held in an optical trap, and a tether formed between the beads once ClpXP engaged its substrate (Figure 3.1B) (Maillard et al., 2011). Two fusion substrates of the green fluorescent protein (GFP) and a permanently unfolded variant of the I27 domain of titin (titin$^{\text{CM}}$) were used (C to N terminus): ssrA-titin$^{\text{CM}}$-GFP-(titin$^{\text{CM}}$)$_2$-GFP and ssrA-(titin$^{\text{CM}}$)$_4$-GFP (Figure 3.1C and 3.7D, respectively). Substrate unfolding was measured as a sudden gain in extension of the tether (rip), whereas polypeptide translocation was monitored as the gradual decrease of extension with time. Because ClpXP cannot efficiently unfold GFP at ATP concentrations below 200$\mu$M, the second substrate allowed the measurement of extended translocation of titin domains without the requirement for GFP unfolding.

Phosphate Release Is the Force-Generating Step in the ClpXP ATPase Cycle

To probe the relationship between the generation of mechanical work and the ATPase cycle, which includes initial ATP binding, tight binding, hydrolysis, and the release of ADP and inorganic phosphate (see scheme in Figure 3.3A), we studied ClpXP in the presence of various concentrations of ATP, ADP, and inorganic phosphate ($P_i$). First, we explored the effect of the ATP concentration on translocation and thereby focused solely on the titin$^{\text{CM}}$ regions of the fusion substrate in order to minimize potential effects of amino acid sequence
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

Figure 3.1: (A) Cartoon depicting protein unfolding and polypeptide translocation by ClpXP. (B) Experimental geometry of dual-trap optical tweezers assay (not to scale). Biotinylated ClpX was immobilized on streptavidin-coated beads (SA). The DNA-tethered substrate was immobilized on the surface of beads coated with anti-digoxigenin antibodies (AD). (C and D) Single-molecule trajectories of substrate processing by ClpXP at 1 mM ATP and forces ranging from 6 to 12 pN. GFP unfolding events are indicated by arrows and followed by translocation of unfolded polypeptide. Substrates are composed of GFP moieties (green) fused to titin\textsuperscript{CM} and a C-terminal ssrA tag (black and red, respectively), as well as an N-terminal ybbR tag (light blue) for attachment to the bead. Raw data (2.5 kHz in gray) were filtered and decimated to 100 Hz (green, black, and blue lines).
differences between titin\textsuperscript{CM} and GFP (Figure 3.3B). Translocation was punctuated by rare pauses that are typically longer than 1 to 2 s (Extended Experimental Procedures). Pauses were previously shown to be in kinetic competition with translocation and therefore off the main translocation pathway (Maillard et al., 2011). Pause-free translocation as a function of the ATP concentration followed a general Michaelis-Menten behavior

$$v = \frac{V_{\max}[ATP]}{[ATP] + K_m}$$

(Figure 3.5A; see Extended Experimental Procedures).

---

Figure 3.2: (A) Pause-free velocity is shown as a function of [ATP] at 7.5 pN resisting force. Each data point represents the average of the pause-free velocities from a group of individual trajectories (n=8-50) with SEM error bars. Data were fit to a simple Michaelis-Menten equation with $K_m=31 \pm 6 \mu M$ (SEM) and $V_{\max}=8.7 \pm 0.4 \text{ nm/s (SEM)}$. (B) Pause-free velocity (mean ± SEM) plotted against [ATP] at [ADP] =10 $\mu M$ (black) or [ADP] = 50 $\mu M$ (purple), with [Pi] held constant at 5 $\mu M$. Each point is the mean velocity from a set of individual traces (n= 7-30) measured at 7.5 pN resisting force. The data fit well to a Michaelis-Menten equation (solid lines in black and purple).

To determine whether ATP binding is the force-generating step, we examined how the translocation velocity depends on force at various ATP concentrations. If ATP binding is coupled to force generation, then conditions in which ATP binding becomes rate limiting should make the conformational changes that drive translocation also rate limiting. In this case, the motor velocity would be highly sensitive to the applied external load. However, we found that the translocation velocity of ClpX is largely insensitive to opposing mechanical forces at low ATP concentrations (Figure 3.3C, black symbols). In contrast, at saturating ATP concentrations ($\leq 500 \mu M$) and opposing forces between 12 and 20 pN, the force-generating step has become rate limiting and ClpX translocation is force sensitive (Figure 3.3C, red
Figure 3.3: (A) General scheme depicting a motor (M) that binds to one ATP molecule (T), undergoes a tight binding transition, and hydrolyzes ATP, followed by the release of inorganic phosphate (Pi) and ADP. (B) Representative trajectories for translocation of the titin\textsuperscript{CM} moiety of the fusion substrates measured between 6 and 12 pN at different ATP concentrations with ATP regeneration system (ATP/RS). The trajectories are offset for clarity. (C) Pause-free velocity of translocation (mean ± SEM) as a function of external force at 5 mM ATP (red symbols) and 35 mM ATP (black symbols). (D) $K_m$ (blue) and $V_{\text{max}}$ (green) are plotted against force. Inset: $K_m/V_{\text{max}}$ ratio plotted for forces between 5 to 15 pN. Error bars are from the fits (SEM). (E) $K_m$ (blue) and $V_{\text{max}}$ (green) plotted against ADP concentration at 7.5 pN. Error bars are from the fits (SEM). (F) Pause-free velocity of translocation (mean ± SEM) plotted as a function of phosphate concentration [Pi] at 7.5 pN with a fixed [ATP] (see also Figure 3.2).
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

44

symbols). These results clearly indicate that ATP binding does not power substrate translocation.

Next, we analyzed the force dependence of the Michaelis-Menten parameters $V_{\text{max}}$ and $K_m$ to determine where the force-generating step may be located in the nucleotide-hydrolysis cycle. The Michaelis-Menten fits to our data revealed that both $V_{\text{max}}$ and $K_m$ decrease with the applied load (Figure 3.3D), but the $K_m/V_{\text{max}}$ ratio remains force insensitive (Figure 3.3D, inset). A force-independent $K_m/V_{\text{max}}$ is consistent with our conclusion that ATP binding does not power translocation and indicates that the force-generating step must occur after the first irreversible transition connected to ATP binding (Chemla et al., 2005; Keller and Bustamante, 2000; Visscher et al., 1999). This irreversible transition is most likely the tight binding of ATP, as previously seen in other ring ATPases such as the F1-ATPase and the φ29 DNA packaging motor (Adachi et al., 2007; Chemla et al., 2005; Oster and Wang, 2000).

The results above leave only ATP hydrolysis, ADP release, or $P_i$ release as possible transitions that may couple to the force-generating step (Figure 3.3A). ATP hydrolysis is unlikely to power translocation, because the small rotation of the terminal phosphate upon hydrolysis does not provide the free energy change required for the power stroke (Oster and Wang, 2000). To investigate the role of product release in the mechanochemical cycle, we varied the concentrations of ADP and $P_i$. The effect of increasing ADP concentrations on the pause-free velocity of ClpXP obeyed a Michaelis-Menten model of competitive inhibition (Figure 3.2B). The apparent $K_m$ for ATP was seen to increase linearly with $[ADP]$ according to

$$K_m = K_m^0 \times (1 + [ADP]/K_i)$$

whereas $V_{\text{max}}$ remained constant (Figure 3.3) (Segel, 1975). Our results are consistent with the known role of ADP as a competitive inhibitor of the ClpX motor (Burton et al., 2003). In contrast, increasing $[P_i]$ from 5µM to 10 mM did not affect the translocation velocity (Figure 3.3F), indicating that $P_i$ release is a largely irreversible transition.

To discriminate between the possible roles of ADP and $P_i$ release in the mechanism of force generation, we estimated the free energy changes of these events based on their respective dissociation constants and compared them to the observed maximum work performed by ClpXP. Using an estimated step size of 1 nm (Glynn et al., 2009; Aubin-Tam et al., 2011; Maillard et al., 2011) and a stall force of at least 20 pN, ClpX subunits perform a near-maximum work of 20 pN·1 nm or $\Delta G = 4.8$ k$_B$T when taking 1 nm steps near stall. Using $K_i = 33$ µM as a dissociation constant for ADP (Figure 3.1E; Extended Experimental Procedures), the change in free energy from ADP release is $\Delta G_D \sim 1.8$ k$_B$T (at $[ADP] = 5$ µM) and thus is insufficient to account for the work performed by ClpXP. In contrast, phosphate release is considered irreversible because the velocity is unaffected even at high $[P_i]$ (10 mM). Therefore, the dissociation constant for phosphate release must be $K_d \gg 10$ mM, with a corresponding change in free energy of $\Delta G_P \gg 7.6$ k$_B$T (for $[P_i] = 5$ µM). Phosphate release would thus provide sufficient energy to power the work produced by ClpXP in every translocation step and is the most likely candidate for the force-generating step of the motor.
ATPγS Dependence of Pause Density and Pause Duration Reveals High Intersubunit Coordination during Translocation

Effective unfolding and/or translocation of protein substrates by ClpX may require the participation of multiple subunits in the hexamer (Martin et al., 2008a). We therefore sought to characterize how individual subunits coordinate their ATP hydrolysis activities around the ring. To this end, we slowed down the hydrolysis in a given subunit by using the ATP analog ATPγS, which ClpX hydrolyzes \( \sim 90 \) times slower than ATP (Figure 3.5; Extended Experimental Procedures) and determined how binding of this analog affected the polypeptide translocation by the remaining ATP-bound subunits in the ring. We held [ATP] fixed at 500 \( \mu \)M and varied [ATPγS] from 0 to 250 \( \mu \)M. In the presence of ATPγS, we observed pauses longer than 1 s, which were extremely rare in the presence of ATP alone, and we attributed these long pauses to ATPγS-bound ClpX subunits (Figure 3.4A). Furthermore, the relationship between the average translocation velocity and the ATPγS concentration decreased in a nonlinear fashion due to these long pauses. The trend could be fitted well to a modified Hill equation with a Hill coefficient \( n_{ATPγS} = 1.5 \pm 0.3 \) (SEM) (Figure 3.4B), indicating that more than one ATPγS binding to the ring is necessary to induce a long pause during translocation.

We observed that the pause density (PD), i.e., the number of pauses per nm of translocated polypeptide, increased with the ATPγS concentration, indicating that the entry into a pause was caused by the binding of ATPγS to the ring (Figure 3.3C; Extended Experimental Procedures). The maximum PD, PD_{max}, reflects the existence of an ATPγS concentration at which the motor has nearly 100% probability of binding the minimum number of ATPγS molecules required to stall translocation. Accurate pause detection became difficult at high concentrations of ATPγS (greater than \( > 200 \) to 250 \( \mu \)M). To calculate PD_{max}, we plotted the inverse of the ATPγS concentration against the inverse of PD and estimated a PD_{max} of \( \sim 0.5 \) nm\(^{-1}\) (Figure 3.4C). Thus, the motor has a 50% probability of entering a pause at [ATPγS] = 200 \( \mu \)M (Figure 3.4C, inset). Using a \( K_{m,ATPγS} = 29 \) \( \mu \)M as an upper bound for the ATPγS dissociation constant (because \( k_{cat} \) has little contribution), we calculated that the probability of having three or more ATPγS bound to the ring at [ATPγS] = 200 \( \mu \)M is \( \sim 0.44 \) (see Extended Experimental Procedures), very close to the observed 50% probability of the motor entering into an analog-induced pause (Figure 3.4C, inset). We conclude that at least three molecules of ATPγS are required to bind to the ring in order to stall translocation and induce a pause. Consequently, intersubunit coordination around the ring does not require the involvement of all ATP-binding-competent subunits.

One of two different processes is likely to dictate the kinetics of exit from a pause: (1) the dissociation or (2) the hydrolysis of ATPγS. To distinguish between these alternatives, we compared the ATPγS hydrolysis rate with the average duration of the ATPγS-induced pauses. Hydrolysis of a single ATPγS molecule takes \( \sim 10 \) s in the presence of titin\(^{CM}\)-ssrA (\( k_{cat} \) of \( \sim 6 \) min\(^{-1}\) hexamer\(^{-1}\)), which is \( \sim 90 \) times slower than the rate of ATP hydrolysis under identical conditions (Figure 3.5; Extended Experimental Procedures). Because the observed mean pause durations are significantly shorter than the time for ATPγS hydrolysis,
Figure 3.4: (A) Representative trajectories for forces between 6 and 12 pN were measured with increasing [ATPγS] at fixed [ATP]. Trajectories are offset for visual clarity. (B) Translocation rate (mean ± SEM) plotted against [ATPγS], with the fit shown in red. (C) Inverse density of ATPγS-induced pauses (mean ± SEM) plotted against the inverse of [ATPγS], with the linear fit shown in red. Inset: pause density (mean ± SEM) plotted as a function of [ATPγS]. (D) Pause duration (mean ± SEM) as a function of [ATPγS] (see also Figure 3.5).
we conclude that they primarily reflect the off-rate of ATPγS. Interestingly, the mean pause durations increased from \( \sim 1.5 \) to \( \sim 2.5 \) s as the ATPγS concentration was increased from 50 to 250 \( \mu M \) (Figure 3.4D; see Extended Experimental Procedures). Thus, the exit from a pause takes longer as the motor loads with an increasing number of ATPγS molecules, most likely reflecting the extra time required to eject multiple ATPγS molecules until just two ATPγS remain bound to the ring.

Effect of [ATP] on Burst Sizes and Dwell Durations

The observation that the motor can maintain its operation even when one or two analog molecules are bound to the ClpX ring led us to investigate in greater detail the mechanism by which subunits communicate with each other around the ring. Previous single-molecule studies showed that ClpXP translocates polypeptide in bursts of 1, 2, and 3 nm (Maillard et al., 2011). Based on the ClpX crystal structure that shows an \( \sim 1 \) nm displacement between the pore-1 loops in ATP-bound and empty states (Glynn et al., 2009), a 1 nm step was thought to reflect the basic power stroke of a single ClpX subunit (Aubin-Tam et al., 2011; Maillard et al., 2011). Accordingly, the 2 and 3 nm bursts were interpreted as the near-simultaneous firing of two and three ClpX subunits, respectively (Maillard et al., 2011). An important remaining question is what mechanism individual ATPase subunits use to coordinate their translocation activities at any given ATP concentration. To address this point, we analyzed how different ATP concentrations affected the distributions of burst sizes and dwell times (Figure 3.7A).

At [ATP] \( \gg K_m \), burst sizes of 2-4 nm were observed at opposing forces between 6 and 15 pN (Figures 3.7B and 3.6; Extended Experimental Procedures). The distribution of burst sizes has a maximum at 3 nm, with a correspondingly lower number of \( \sim 2 \) and \( \sim 4 \) nm bursts. In contrast, when ATP binding is near rate limiting ([ATP] = 35 \( \mu M \)), ClpXP translocates mostly in 2 or 3 nm bursts, and bursts of 4 nm are completely absent (Figure 3.7B, red histogram).

Next, we analyzed the duration of the cycle time (the sum of the dwell and burst phase duration) at various ATP concentrations to better understand the translocation mechanism of this motor. Surprisingly, we observed that the mean cycle time has no apparent dependence on [ATP]. We found that the mean duration of the dwell phase \( < \tau > \) is 350 \( \pm 20 \) ms in the range between 35 \( \mu M \) and 5 mM (Figure 3.7C). In addition, the duration of the burst phase contributes to less than \( \sim 3\% \) of the cycle time and has a mean duration of less than \( \sim 10 \) ms. In this range of nucleotide concentrations, the motor translocation rate approximately doubles from 5 to 9 nm/s. Consequently, the observed change in translocation rate is not due to changes in the mean cycle time, but rather to a systematic increase in motor burst size with increasing [ATP].
Figure 3.5: (A) Thin-layer chromatography assay of $^{35}$S-labeled ATP$\gamma$S hydrolysis by ClpXP. $^{35}$S-ATP$\gamma$S was incubated with 0.3 µM ClpX and 1.5 µM ClpP (top) or buffer only (bottom) for the time indicated, before being quenched with 2.5 volumes of stop buffer and spotted on TLC plates (see Extended Experimental Procedures). The positions of ATP$\gamma$S and PO$_4$S are indicated. (B) Rate of ClpXP-mediated hydrolysis of ATP$\gamma$S as a function of [ATP$\gamma$S] in the presence (red) and absence (blue) of titin$^{CM}$-ssrA. Conditions were as in (A) with the addition of 10 µM titin$^{CM}$-ssrA as indicated. Fitting the data to the Michaelis-Menten equation provided the values of $k_{cat}$ = 6.3 ± 0.9 min$^{-1}$ ClpX$^{-1}$ and $K_m$ = 29 ± 10 µM in the presence of titin$^{CM}$-ssrA, and $k_{cat}$ = 2.5 ± 0.2 min$^{-1}$ ClpX$^{-1}$ and $K_m$ = 6.2 ± 2.0 µM in the absence of titin$^{CM}$-ssrA. All fitted values are mean ± SEM. (C) Rate of ATP hydrolysis by ClpXP is shown as a function of [ATP] in the presence (red) and absence (blue) of titin$^{CM}$-ssrA. Fitting the data to the Michaelis-Menten equation provided the values of $k_{cat}$ = 497 ± 20 min$^{-1}$ hexamer$^{-1}$ and $K_m$ = 57.2 ± 6.1 µM in the presence of titin$^{CM}$-ssrA, and $k_{cat}$ = 111 ± 13 min$^{-1}$ hexamer$^{-1}$ and $K_m$ = 58.9 ± 2.0 µM in the absence of titin$^{CM}$-ssrA. All fitted values are mean ± SEM.
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

Figure 3.6: (A and B) Distribution of burst sizes occurring between 6-10 pN (red line) and 10-14 pN (blue line). (A) Burst size is shown in units of extension (nm), reflecting the end-to-end distance of polypeptide translocated during a single burst phase. The distributions of burst sizes in nanometers within each force range were determined to be identical with $p=0.72$ (two-sample Kolmogorov-Smirnov test, null hypothesis: both distributions identical). (B) Burst size is shown in units of contour length in amino acids, reflecting the number of amino acids that pass through the ClpX pore during a single burst phase. The distribution of burst sizes in amino acids within each force range was determined to be different with $p=1.5e-6$ (two-sample Kolmogorov-Smirnov test, null hypothesis: both distributions identical).

To obtain insight into the molecular processes that occur during the dwell time, we calculated the kinetic parameter $n_{\text{min}}$, which is defined as the ratio of the squared mean of dwell times over the variance of the dwell times,

$$n_{\text{min}} = \frac{\langle \tau \rangle^2}{\langle \tau^2 \rangle - \langle \tau \rangle^2}$$

(Moffitt et al., 2006; Schnitzer and Block, 1995). It has been shown that this parameter provides a strict lower bound to the number of rate-limiting events during the dwell phase (Moffitt et al., 2006). We sought to understand how the number of rate-limiting events ($n_{\text{min}}$) changes with the ATP concentration. At saturating and near-$K_m$ concentrations of ATP, we measured $n_{\text{min}} = 2.1 \pm 0.4$ and $2.0 \pm 0.6$ (SEM), respectively, suggesting that for both conditions there are at least two rate-limiting transitions in the dwell leading up to the burst phase. Similar values of $n_{\text{min}} = 1.9 \pm 0.6$ were obtained at the intermediate ATP concentrations of 100 and 200 $\mu$M. Therefore, at least two processes-not associated with ATP binding-control the duration of the dwell in the [ATP] range between 35 $\mu$M (near $K_m$) and 5 mM (saturating).

Altogether, these results suggest a mechanism of translocation that is strikingly different from those of other motors. The well-characterized DNA packaging motor of bacteriophage $\phi$29 exhibits a variable, ATP-concentration-dependent dwell time followed by a constant, ATP-concentration-independent burst size of 10 bp, which reflects a high degree
of coordination among the ring subunits that must all load ATP before the motor can initiate translocation (Chistol et al., 2012; Moffitt et al., 2009). In sharp contrast, ClpXP exhibits an ATP-concentration-independent cycle time, during which a variable number of ATP molecules bind to the motor, resulting in a distribution of burst sizes.

**C**lpXP Requires Four Highly Coordinated Power Strokes for Successful Unfolding of GFP

Our findings that ClpX translocates with variable burst sizes and that the burst-size distribution is ATP concentration dependent raises two important questions about the mechanism by which the motor successfully unfolds protein substrates. First, how are the power strokes of individual subunits coordinated during protein unfolding? And second, what is the kinetic competition between the motor translocation bursts and the substrate’s resistance to unravel and its tendency to refold? To answer these questions, we quantified how variations in ATP concentration (and therefore in translocation burst sizes) affect the ability of the motor to unfold GFP. We found that reducing the ATP concentration decreased the probability of GFP unfolding by ClpXP nonlinearly from a maximum of 0.6 when [ATP] ≥500 µM to less than 0.1 when [ATP] <50 µM (Figure 3.9A).
Figure 3.8: (A) The rips during GFP-unfolding revealed three well-defined transitions, indicating the presence of two unfolding intermediates (I and II) at saturating [ATP]. (B) Distribution of lifetimes for the folded state (F), I and II during GFP unfolding by ClpXP. (C) Plot of force versus change in extension for the transitions of F→I (blue), I→II (red), II→U (green), and F→U (black). Shift in the orientation of the folded structure removed for clarity (Eqn. 3.14). (D) Root-mean-square-deviation of the measured extension changes as a function of linker residue position for the transition to I (left) and II (right) with GFP topology below. (E)-(G) Histograms of contour-length change for the transitions: (E) F→II, (F) II→U, and (G) F→U. (H)-(I) Lifetime distribution of II. (J) and (K) Distributions of the change in contour length and the mean time constant for unfolding and refolding events of β11. (J) Distribution of changes in extension upon unfolding and refolding of β11. (K) β11 refolding time fits to a single-exponential.
The nonlinear relationship between GFP unfolding probability and ATP concentration suggests that GFP unfolding requires the coordinated and near-simultaneous ATP hydrolysis of multiple ClpX subunits in the ring. In order to elucidate the coordination mechanism by which ClpXP successfully unfolds GFP, we characterized the intermediates observed during GFP unfolding. Unraveling of GFP from the C terminus proceeded via two transient intermediates with mean lifetimes of 45 ± 10 and 130 ± 15 ms (Figure 3.9B 5B, 3.8A, and 3.8B; Extended Experimental Procedures). By using the wormlike chain (WLC) model of polymer elasticity (Bustamante et al., 1994), we estimated that the transition from the folded state “F” to the first intermediate “I” (F→I) has a contour length increase ∆Lc^F→I of 8.3 ± 0.4 nm (SEM) corresponding to the extraction of β strand 11 (β11) from the GFP barrel. The second transition from “I” to the second intermediate “II” (I→II) has a ∆Lc^I→II of 31.2 ± 0.8 nm and most likely corresponds to the unfolding of β strands 10 through 7. The last transition from “II” to the unfolded state “U” (II→U) has a ∆Lc^II→U of 42.2 ± 0.8 nm and reflects the unraveling of the remaining six β strands, as described previously (Maillard et al., 2011) (Figure 3.9C and 3.8C-I; Extended Experimental Procedures). The total contour length increase ∆Lc^F→U ( = ∆Lc^F→I + ∆Lc^I→II + ∆Lc^II→U) of 82.8 ± 3.2 nm is in agreement with the expected value for the complete unfolding of GFP (Figure 3.8C; Extended Experimental Procedures).

At [ATP] ≤ 200 µM, we detected small unfolding and refolding events before ClpXP completely unraveled GFP (Figure 3.9D). The change in contour length during the reversible unfolding and refolding events is similar to ∆Lc^F→I, suggesting that these events most likely correspond to the extraction and quick refolding of β11. Analysis of these reversible transitions revealed that β11 snaps back into the GFP barrel with a mean refolding time constant ∼ 240 ms at forces between 7 and 9 pN (Figure 3.8J and 3.8K; Extended Experimental Procedures). These results provide direct experimental evidence of the molecular tug-of-war between the motor, attempting to unravel folded structures, and a substrate with a strong tendency to refold. Hence, protein-unfolding machines have to perform not only the thermodynamic function of mechanically destabilizing the native state, but also the kinetic task of quickly capturing the unstructured polypeptide before it can refold. When [ATP] > K_m, ClpXP is able to move in bursts of 4 nm during a single translocation cycle. Such coordinated translocation is sufficient to trap most of the dislodged β11 and prevent its refolding. These observations indicate that the efficient unfolding of GFP by ClpXP requires not only a 4 nm burst, but also this burst to occur faster than the refolding time of β11 (< 240 ms). Using the distribution of burst size and dwell duration determined here, we estimated the probability of ClpXP taking a 4 nm burst in less than 240 ms to be on average 0.031 (0.018-0.046). This result indicates that ClpXP must pass through ∼33 translocation cycles before it can unfold GFP and trap the first intermediate for subsequent unfolding. Multiplying 33 translocation cycles by the mean duration of the dwell (∼0.35 s) predicts a mean GFP unfolding time of ∼11.5 s, which agrees very well with the mean time constant obtained from the distribution of GFP unfolding times, < τ > = 11.8 ± 0.9 s (Figure 3.8B; Extended Experimental Procedures), as well as from previous single-turnover GFP degradation measurements (Martin et al., 2008b).
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

Figure 3.9: (A) Probability of GFP unfolding as a function of [ATP] and ATPase rate (inset). Error bars are ± SEM. (B) GFP unfolding events display two intermediates at 300 Hz. (C) Mechanism of GFP unfolding by ClpXP at [ATP] ≥ K_m. (D) Trajectory at [ATP] = 200 mM illustrating the ClpXP-induced unfolding and refolding of β11 (see also Figure 3.8).

At [ATP] ~K_m, ClpXP moves in bursts of at most 3 nm, which may be sufficient to promote the extraction of β11 from the GFP barrel (ClpX succeeds in carrying out the thermodynamic task) but are too small to prevent the refolding of β11 (ClpX fails to accomplish the kinetic task). Thus, under subsaturating ATP conditions, the kinetic competition between ClpXP attempts to translocate the unfolded region and the tendency of that region to refold greatly reduces the unfolding efficiency of the motor.

3.4 Discussion

Force Generation in ClpXP Is Coupled to Phosphate Release

Here, we propose a mechanochemical model for ClpXP that identifies the force-generating step in the chemical cycle of ATP hydrolysis (Figure 3.10A). As shown above, our results revealed that all transitions reversibly connected to ATP binding up to and including the first irreversible step are not involved in force generation. Furthermore, the first irreversible transition in other molecular motors has been identified as the tight binding of ATP (Adachi et al., 2007; Chemla et al., 2005). Indeed, the duration of the ATP_γS-induced pauses makes the tight binding of ATP a likely candidate for the first irreversible transition in the mechanochemical cycle of ClpXP. Because tight ATP binding is apparently not rate limiting during translocation, its rate constant (k_{TB}) can be estimated by using a lower bound that
corresponds to the translocation rate of the motor, $k_{\text{cat}} = 9 \, \text{s}^{-1}$. The reverse transition from tight to loose binding, $k_{\text{TB}}$, is given as the inverse of the mean ATPγS-pause duration, $\sim 0.6 \, \text{s}^{-1}$. Thus, we obtain $k_{\text{TB}} / k_{\text{-TB}} = 15$ and a corresponding free energy change associated with tight binding $\Delta G_{\text{TB}} > 2.7 \, \text{kJ/mol}$. The tight binding of ATP can therefore be considered the first irreversible transition following ATP binding. As described above, our results exclude the possibility of ATP hydrolysis and ADP release being coupled to the force-generating step and reveal instead that force generation likely occurs upon $P_i$ release. Interestingly, in this aspect ClpX resembles other members of the ASCE family, such as the φ29 DNA-packaging motor and F1-ATPase, harnessing $P_i$ release as a force-generating step despite the distinct architectures and functions within this large family of motors (Chemla et al., 2005).

Intersubunit Coordination Determines the ClpX Translocation Mechanism

Titrations with the slowly hydrolyzable analog ATPγS as well as the analyses of burst sizes and dwell times have revealed several aspects of the intersubunit coordination in the ClpX hexamer. The pause-density dependence on ATPγS concentration suggests that ClpX maintains operation even when one or two ATPγS molecules bind to the ring. At least three ATPγS molecules must bind to the ring in order to stall the motor. Based on previous biochemical and structural studies, at most four subunits in the hexamer can bind nucleotide (Hersch et al., 2005; Glynn et al., 2009; Stinson et al., 2013). Because binding of three analog molecules stalls the motor, the remaining fourth subunit may bind ATP but would still be unable to drive translocation. This analysis suggests that the minimal operational unit for substrate translocation by ClpX involves at least two subunits.

That it takes binding of three analog molecules to stall translocation by the ClpX hexamer rules out models of strict intersubunit coordination described previously for other molecular motors (Lyubimov et al., 2011). For instance, it excludes models of concerted hydrolysis, where the subunit power strokes occur simultaneously after all subunits have been loaded with nucleotide. It also contradicts strictly sequential hydrolysis models, in which the power stroke of one subunit occurs only after the power stroke of its neighbor. If one of these models of intersubunit coordination were to apply to ClpX, a single ATPγS binding event would be sufficient to stall the motor. Furthermore, our data are also inconsistent with stochastic or probabilistic hydrolysis models, where all subunits act independently from each other. In a stochastic scenario, all the active subunits would have to bind ATPγS in order to stall translocation. Thus, we find that the ATPase cycles of individual ClpX subunits are neither strictly coordinated nor completely independent from each other. Instead, as will be discussed below, we find that two, three, or four subunits can coordinate their activity in each cycle, depending on the number of ATPs bound to the motor.

The analysis of burst sizes provides direct evidence for how many subunits participate in translocation during a single burst phase. The ClpX crystal structure reveals a distance of 1 nm between the pore loops of adjacent subunits in different nucleotide states, suggesting
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

a 1 nm power stroke per hydrolyzing subunit during substrate translocation (Glynn et al., 2009). Furthermore, dividing the maximum translocation velocity, \( V_{\text{max}} = 8.5 \) nm/s, by the rate of ATP hydrolysis, \( k_{\text{cat}} = 8.3 \) ATP s\(^{-1}\), yields \( 1.02 \pm 0.03 \) nm per hydrolyzed ATP. Thus, if one ATP is hydrolyzed per ClpX subunit, the fundamental translocation step size must be \( \sim 1 \) nm, consistent with previous single-molecule reports (Aubin-Tam et al., 2011; Maillard et al., 2011). Consequently, the variable burst sizes of 2, 3, or 4 nm observed at various ATP concentrations reflect the near-instantaneous, coordinated firing of two, three, or four subunits around the ring. The maximum burst size of 4 nm suggests that up to four subunits in the hexamer participate during a single translocation cycle. The smallest observed burst size as well as the minimum number of ATP\( \gamma \)S molecules required to stall the motor allow us to propose that active translocation requires a minimum number of two coordinated subunits.

Surprisingly, our data reveal that bursts of 2 or 3 nm occur much more frequently than bursts of 4 nm, even at saturating concentrations of ATP, suggesting that the most relevant translocation cycle of the ClpX motor involves the coordinated hydrolysis and conformational change of two or three subunits. These results thus provide direct evidence for the operational flexibility of ClpX during polypeptide translocation. Previous bulk biochemical assays and crystal structures of ClpX (Hersch et al., 2005; Glynn et al., 2009; Stinson et al., 2013) revealed hexamers bound to four nucleotides. These findings most likely reflect the state of the ring preceding a 4 nm burst, a state that, in view of our results, is only one of a larger ensemble of nucleotide-bound configurations. Future structural studies in the presence of substrate may capture ClpX loaded with two or three ATP molecules, which according to our data correspond to conformational states that are more probable during polypeptide translocation.

Our results show that the number of subunits participating in a single translocation cycle depends on the availability of ATP. When the ATP concentration was lowered from saturating to near \( K_m \), we observed a redistribution of burst sizes from 3 and 4 nm to 2 and 3 nm. The predominant population of 2 nm bursts and the almost complete absence of 1 nm bursts indicates that, even at partly rate-limiting ATP concentrations ([ATP] \( \sim \) 35 \( \mu \)M), at least two of the subunits must bind ATP in order to initiate a translocation cycle. Therefore, the \( K_m \) of these two subunits that bind ATP first must be significantly lower than the average \( K_m \) of 35 \( \mu \)M. The lack of 4 nm bursts at ATP concentrations near 35 \( \mu \)M indicates that the remaining subunits have a higher \( K_m \) than the first two high-affinity sites and thus contribute to translocation only at high ATP concentrations.

Two different classes of ATP binding sites with correspondingly different binding affinities have been previously identified in ClpX using traditional nucleotide competition assays (Hersch et al., 2005). Furthermore, recent mutational studies of ClpX have shown strong evidence for a dynamic mechanism of subunit switching, whereby nucleotide binding can affect the affinities of the remaining subunits within the ring (Stinson et al., 2013). Our data shed light onto how the coexistence of subunits with high and low ATP affinity affects the dynamics of the translocation cycle, allowing the burst phase to be initiated with either two ATPs (both binding to high-affinity subunits), three ATPs (two high- and one low-affinity
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

subunits), or four ATPs (two high- and two low-affinity subunits) bound to the motor.

A surprising finding is that, regardless of the number of subunits participating in the translocation cycle, the mean cycle time remains constant for ATP concentrations between \( \sim K_m \) and saturating. The mean cycle time is expected to increase when ATP binding becomes rate limiting, and, to probe this behavior, we would have to monitor translocation in the presence of ATP concentrations at or below the \( K_m \) for the high-affinity subunits \( (K_m \leq 35 \mu M) \). However, no translocation activity was observed for ATP concentrations below \( \sim 29 \mu M \). Within the accessible ATP-concentration range, we found that the dwell duration is governed by at least two non-ATP-binding events \( (n_{\text{min}}=2) \). The nonbinding events in the dwell could correspond to conformational changes within the ClpX ring that either (1) result from or (2) are completely independent of ATP binding.

Figure 3.10: (A) The pathway of ATP hydrolysis for a single subunit of ClpX. An empty subunit \( (E, \text{red sphere}) \) binds ATP \( (T, \text{orange}) \) and undergoes a tight binding of ATP \( (T^*, \text{green}) \). Then, the subunit hydrolyzes ATP to ADP and \( P_i \) \( (DP, \text{blue}) \), the force-generation step occurs upon phosphate release \( (D, \text{purple}) \), and ADP dissociates, leaving an empty subunit \( (E, \text{red sphere}) \). (B) Schematic depiction of intersubunit coordination at saturating \( (\text{green box}) \) and limiting ATP concentrations \( (\text{red box}) \) for one possible scenario depicting sequential ATP binding. The subunits in gray correspond to those that do not bind ATP. During the dwell phase, at least two ATPs are bound to the high-affinity subunits \( (T, \text{blue outline}) \), and additional ATPs can bind to the low-affinity ClpX subunits \( (T, \text{green outline}) \), depending on \([\text{ATP}]\). During the burst phase, the motor hydrolyzes all bound ATPs, releases phosphate, and translocates the substrate by 2, 3, or 4 nm into the central pore.
An “Internal Clock” Triggers Polypeptide Translocation

Our results provide a model to rationalize the observed invariant dwell-time distribution and the variable burst-size distribution of ClpX as a function ATP concentration. The dwell duration is largely determined by two slow, non-ATP-binding events. By the time these slow transitions occur, the two high-affinity subunits are ATP bound and, depending on the ATP concentration, one or two of the low-affinity subunits are occupied with ATP as well. Due to the constrained ring geometry, only four of the six subunits can be ATP bound (Hwang and Lang, 2013), which provides an explanation for why we do not observe bursts of more than 4 nm even at saturating ATP concentrations. After the motor hydrolyzes all of the bound ATP molecules, Pi release and concomitant subunit power strokes occur near simultaneously around the ring, resulting in a burst size that is proportional to the number of hydrolyzed ATP molecules. The essence of this model is that the mean duration of the dwell phase is constant and set by an “internal signal” or “clock”, which may or may not follow ATP binding and could, for instance, correspond to the reaching of a strain threshold in the ring or the hydrolysis of the first-bound ATP. In contrast to this constant average dwell phase, the burst size is variable and proportional to the number of subunits bound to ATP before the clock triggers the initiation of a translocation cycle around the ring. As a result, the ClpXP motor operates at a constant frequency (fixed “rpm”) and a variable burst size (different “gears”). Although our current model, depicted in Figure 3.10B, suggests a spatial and temporal order of ATP-docking events, future single-molecule studies of ClpX mutants will be required to definitively establish the order of ATP-binding and hydrolysis events around the hexameric ring.

Biological Implications of the ClpX Translocation Mechanism

The model proposed here provides a framework to understand how ClpXP successfully unfolds stable protein substrates. The ability of the motor to bind four ATP molecules allows it to translocate in large bursts and thus destabilize and rapidly trap partially unfolded intermediates. For instance, to successfully unfold GFP, ClpX subunits must near-simultaneously take a 4 nm burst, which results in the extraction and translocation of β11 from GFP before this strand can refold onto the β barrel. At saturating concentrations of ATP, the time required for ClpX-induced unfolding of GFP is determined by the time that passes before the motor makes a 4 nm burst. In contrast, ClpX rarely unfolds GFP at ATP concentrations near K_m because hydrolysis under these conditions is always triggered before four ATP molecules can bind to the motor. In support of our finding, previous biochemical studies have shown a nonlinear decrease in GFP degradation with decreasing ATPase rate (Martin et al., 2008b), whereas a linear correlation emerges for circular permutants of GFP that do not form a stable unfolding intermediate upon extraction of their C-terminal strand (Nager et al., 2011). Our results reveal that the decreased probability of the motor taking a 4 nm burst (from ~20% at saturating [ATP] to ~0% at [ATP] near K_m) is responsible for the observed nonlinear decrease of GFP unfolding probability with the ATPase rate.
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

Based on bulk biochemical studies, it has recently been suggested that the ClpX motor may operate by a partially probabilistic mechanism during unfolding and translocation (Stinson et al., 2013). Our results show that ClpX translocates substrates using a highly coordinated mechanism in which, regardless of the number of translocating subunits, the average dwell duration is constant and followed by a variable burst size that reflects the firing of a different number of subunits in rapid succession. Although there is a stochastic element during the loading of ATP, we observe a high degree of intersubunit coordination during the rapid burst phase of translocation.

Such a mechanism of a constant cycle time (resulting in a variable burst size) allows flexibility for ClpX to successfully overcome unique chemical and mechanical obstacles during polypeptide translocation. A constant cycle time governed by an internal clock provides the motor with a fail-safe mechanism to drive translocation even when some subunits are not loaded with nucleotide. This characteristic will allow the motor to prevent substrate disengagement or prolonged periods of stalling during unfolding attempts when a subset of subunits is unable to maintain grip on the polypeptide.

Our results thus provide an archetype of molecular-motor operation that differs significantly from those described previously for other ring-shaped motors. For instance, the translocation mechanism of the φ29 DNA packaging ATPase requires that periodic contacts be made by the motor every 10 bp along the helical pitch of double-stranded DNA. In order to achieve a fixed burst size of 10 bp, the motor must wait during the dwell until all subunits are loaded with ATP. This mode of operation may have been optimized to keep the motor in register with the symmetry of its DNA substrate and provides a fascinating contrast to our results on ClpX. Because a polypeptide track is aperiodic, ClpX cannot rely on contacting a regularly repeating motif, and its variable burst size (different ‘gears’) and constant cycle time (constant ‘rpm’) may have arisen as a flexible mechanism that optimizes its efficiency for robust kinetic trapping of unfolding intermediates. The distinction between these mechanisms offers clear evidence for the evolutionary constraints imposed by the motor’s substrates to favor certain mechanisms of operation. The mechanisms described here thus provide important insights into the operating principles of ATP-dependent proteases and may have critical implications for the understanding of other ring-shaped ATPases of the AAA+ and RecA families in general.

3.5 Methods

Sample Preparation

Biotinylated ClpX single-chain hexamers, GFP-titin\textsuperscript{CM} I27 fusion proteins, and 3 kbp dsDNA handle for protein attached via ybbR tag/Sfp system were prepared as described previously (Maillard et al., 2011; Martin et al., 2005; Martin et al., 2008a). Tethers were assembled in a buffer (25 mM HEPES-KCl [pH 7.4], 20 mM MgCl\textsubscript{2}, 100 mM KCl, and 0.5 mM EDTA) supplemented with (1) \([\text{ATP}] = 35 \text{ mM, 45 mM, 100 mM, 200 mM, 0.5 mM, 1 mM, and 5}\]
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

mM with ATP regeneration system (Kenniston et al., 2003), (2) [ADP] = 10 mM, 20 mM, 50 mM, 100 mM, and 250 mM in the presence of [ATP] = 100 mM, 200 mM, 500 mM, 1,000 mM and [P\textsubscript{i}] = 5 mM, (3) 5 mM, 5 mM, 10 mM NaPO\textsubscript{4} (Sigma), (4) [ATP\textsubscript{γS}] = 25 mM, 50 mM, 100 mM, 175 mM, and 250 mM (Roche) in the presence of [ATP] = 500 mM. All single-molecule experiments required 500 nM of ClpP for the formation of the ClpXP complex (K\textsubscript{d} = 90 nM) (Joshi et al., 2004).

Data Collection

Two different dual-trap optical trapping instruments with 1,064 nm laser were employed (Moffitt et al., 2006). The unfolded polypeptide contour length was calculated as previously described (Bustamante et al., 1994; Cecconi et al., 2005; Maillard et al., 2011). Using 0.9 mm polystyrene beads, an oxygen scavenging system was added to prevent the formation of the reactive species singlet oxygen (100 mg/ml glucose oxidase, 20 mg/ml catalase, 5 mg/ml dextrose; Sigma-Aldrich) (Moffitt et al., 2009).

Analysis

Pauses were removed from velocity using a previously described modified Kalafut-Visscher (KV) algorithm with a pause threshold and penalty (Chen et al., 2008; Chistol et al., 2012). A cutoff threshold was calculated by taking three SDs of a gamma-fitted distribution. Velocity was calculated as the end-to-end distance \( \Delta x/\Delta t \). To calculate pause-free velocity, the selected pauses were removed in the \( \Delta t \) component. Furthermore, we analyzed the pause density and frequency of ATP\textsubscript{γS} using the modified KV algorithm. ATP\textsubscript{γS}-induced pauses were extracted by removing ATP-only dwells using a double exponential fit with one time constant fixed to the mean dwell duration measured at saturating [ATP]. Steps and dwells were analyzed using pairwise distribution and t test (Moffitt et al., 2009). Data were filtered to 15-25 Hz and binned into 0.3 and 0.4 nm for the pairwise distributions. The unfolding events and its related measurements were measured by a previously described method (Maillard et al., 2011).

3.6 Extended Experimental Procedures

Measurements of ATP-hydrolysis rate

The ATP hydrolysis rate of ClpXP was measured using an NADH-coupled ATP-regeneration system as previously described (Martin et al., 2008b; Maillard et al., 2011). Assembled hexamers of ClpX (0.3 \( \mu \)M) were mixed with ClpP (1.5 \( \mu \)M) in a ClpX-100 buffer (25 \( \mu \)M HEPES pH 7.6, 20 mM MgCl\textsubscript{2}, 100 mM KCl, and 0.5 mM EDTA) containing an NADH-coupled regeneration system (3 U/mL pyruvate kinase, 3 U/mL lactate dehydrogenase, 1 mM NADH, and 7.5 mM phosphoenolpyruvate). The ATP-hydrolysis rate of ClpX was measured
both in the presence and absence of 10 µM titin\textsuperscript{CM}-ssrA by monitoring the absorbance of NADH (340 nm) at 30 °C (Figure 3.2). The observed ATPase rates were consistent with previous studies (Martin et al., 2005; Aubin-Tam et al., 2011).

**Measurements of ATP$\gamma$S-hydrolysis rate**

The ATP$\gamma$S-hydrolysis rate of ClpXP was measured as previously described (Burton et al., 2003) using thin-layer chromatography of $^{35}$S-labeled ATP$\gamma$S. Hydrolysis was measured at room temperature by mixing 0.3 µM single-chain ClpX hexamer and 1.5 µM ClpP with variable concentrations of ATP$\gamma$S (5-800 µM) and trace amounts (70 nM) of $^{35}$S-labeled ATP$\gamma$S (Perkin-Elmer Inc.) in ClpX-100 Buffer (see above), with or without ssrA-titinCM (10 µM). One-microliter aliquots from each 20 µL reaction were removed at different times and immediately quenched with 2.5 µL of stop buffer (50 mM Tris-HCl pH 7.6, 100 mM EDTA, 20 mM ATP$\gamma$S, and 20 mM Na$_{3}$PO$_{4}$S). Each quenched sample was spotted onto a plastic-backed PEI-cellulose sheet (J.T. Baker Inc.) and chromatographed at room temperature in 1.5 M formic acid and 0.4 M LiCl. After the sheet was dried, radioactivity in each spot was quantified using a Molecular Dynamics Phosphorimager. In the control experiment with only $^{35}$S-labeled ATP$\gamma$S, we observed that the single radioactive spot, corresponding to the un-hydrolyzed $^{35}$S-ATP$\gamma$S, hardly migrated from its loading position. In the presence of ClpXP, we observed a second, fast-migrating radioactive spot whose intensity increased in a time-dependent fashion (Figure 3.5B). This spot corresponds to PO$_{3}^{35}$S that is released by ClpXP upon $^{35}$S-ATP$\gamma$S hydrolysis, and its absence in samples without ClpX indicates that spontaneous hydrolysis of ATP$\gamma$S is negligible under our experimental conditions.

**Motivation for fitting ATP$\gamma$S-induced pauses to the Hill Equation**

The reaction pathway for an ATP$\gamma$S-induced pause can be written as

\[ M_a + nATP\gamma S \leftrightarrow M_p \]  

where $M_a$ is the actively translocating state of the motor, $M_p$ is the ATP$\gamma$S-induced paused state of the motor bound to the analog, and $n$ is the number of ATP$\gamma$S molecules required to induce a pause. The apparent dissociation constant $K_d$ of $n$ ATP$\gamma$S molecules binding to the motor can be expressed as

\[ K_d = (P(M_a)[ATP\gamma S]^n)/(P(M_p)) \]  

where $P(M_a)$ is the probability of the motor being in an actively translocating state and $P(M_p)$ is the probability of the motor being in an ATP$\gamma$S-induced paused state. The pause density $PD([ATP\gamma S])$, or number of pauses per unit length of translocated polypeptide, should be directly proportional to $P(M_p)$. Since intrinsic pauses in the absence of ATP$\gamma$S are infrequent- on average one pause per 50 nm of translocated polypeptide (Maillard et al., 2011)- the approximation is appropriate.

\[ P(M_p) + P(M_a) = 1 \]  

(3.3)
CHAPTER 3. STUDYING MECHANISMS OF FORCE-GENERATION AND INTERSUBUNIT COORDINATION

Using Eqn. 3.3 in Eqn. 3.2 to eliminate $P(M_a)$ results in

$$PD([\text{ATP}\gamma\text{S}]) \propto P(M_p) = [\text{ATP}\gamma\text{S}]^n/(K_d + [\text{ATP}\gamma\text{S}]^n)$$

Hence, $PD([\text{ATP}\gamma\text{S}])$ takes the form of the Hill equation with $n$ reflecting the number of ATPγS molecules that must bind to the motor in order to induce a pause.

Calculating the fraction of ATPγS molecules bound to the hexamer

For a ClpX motor subunit $M$ present in solution with a mixture of ATP and ATPγS (denoted as $A$ and $B$, respectively), the pathway for the subunit binding either nucleotide can be written as

$$MB \leftrightarrow B + M + A \leftrightarrow MA$$

The total number of motor subunits in the ensemble is constant and equals

$$[M]_T = [M] + [MA] + [MB]$$

Thus, the fraction of $A$ and $B$ molecules bound to $M$ with their corresponding dissociation constants rewritten in terms of $f_A$ and $f_B$:  

$$f_A = [MA]/[M]_T; K_d^A = [M][A]/[MA] = (1 - f_A - f_B)[A]/f_A$$

and

$$f_B = [MB]/[M]_T; K_d^B = [M][B]/[MB] = (1 - f_A - f_B)[B]/f_B$$

Equations 3.7 and 3.8 constitute a system of equations that can be used to obtain expressions for $f_A$ and $f_B$ in terms of $[A]$, $[B]$, and the dissociation constants $K_d^A$ and $K_d^B$:


and


From our biochemical studies of ATP and ATPγS hydrolysis in the presence of substrate, $K_m$ can be used as an upper-bound estimate for the $K_d$, since $k_{cat}$ has a small contribution. From the hydrolysis experiments described above (Figure 3.5), the $K_m$ values were determined to be 57 $\mu$M and 29 $\mu$M for ATP and ATPγS, respectively. Therefore, we can use the $f_A$ to estimate the probability of $n$ATPγS molecules bound to the available six active sites in the hexamer with the binomial distribution: 

$$\binom{n}{k} p^n q^{6-n},$$

where $p = f_A$, and $q = 1-f_A$ which includes both populations bound to ATP and not bound to any nucleotide.
Spatial/Temporal resolution of the instrument

The lifetime distributions of each intermediate in GFP unfolding (Figure 3.8B) follow a single exponential decay function. To calculate the probability of observing an intermediate with duration \( \tau \) we must first transform this function into its corresponding probability density function, which is expressed as

\[
f(x) = ae^{-at}
\]

(3.11)

where \( a \geq 0 \) and corresponds to the inverse of the time constant in seconds, and \( x=[0, \infty) \).

The probability of observing a GFP-unfolding intermediate whose lifetime is between 0 and \( \tau \) can be calculated by obtaining the area under the curve given by the following expression:

\[
P(0 \leq x \leq \tau) = \int_{0}^{\tau} ae^{-ax} \, (dx)
\]

(3.12)

Using the time constants from the exponential fit of the distributions in Figure 3.8B, we can use this expression to test if the apparent absence of the first or the second intermediate in some traces is because i) they may be shorter in duration than the detection limit of the instrument, or ii) GFP unfolding may follow multiple pathways where the first or second intermediates may not be obligatory transitions in this process.

To address the first scenario, we established the spatial/temporal detection limits of our instrument. We calculated the fluctuations of the noise at different forces and different bandwidths through the analysis of the power spectrum of a trace where ClpXP translocation has been stalled with 100% ATP. Since we analyzed the GFP-unfolding intermediates at 300 Hz (i.e., 300 data-points per second), we found that fluctuations of the noise at this bandwidth are approximately 3 nm. With this value, we calculated the signal-to-noise ratio (SNR), which corresponds to the size of the signal divided by the size of fluctuations of the noise. We included this SNR in the following expression:

\[
\Delta l \geq \sqrt{\frac{2KTN\gamma}{<\tau>\kappa_{tether}}} \cdot SNR
\]

(3.13)

which predicts the minimal duration of a step that is required to be observed given its size. Here, \( \Delta l \) corresponds to the mean size of the step, \( K \) is the Boltzmann constant, \( T \) is the temperature, \( N \) is the number of uncorrelated measurements per average dwell, \( \gamma \) is the drag coefficient of the beads, \( \kappa_{tether} \) is the mean duration of a dwell, \( \kappa_{tether} \) is the stiffness of the tether, and SNR is the signal-to-noise ratio that we have calculated (Moffitt et al., 2008).

Using this expression for a burst size of 8.3 nm (corresponding to the mean extension of the first rip), we found that its corresponding mean dwell duration (first intermediate duration) should be at least of 30 ms to be observed at 300 Hz. With this threshold, we calculated the theoretical probability of not observing an intermediate due to limitations of the instrument by calculating the probability \( P(0 \leq x \leq 30) \) using the expression in Eqn. 3.12. Finally, we compared this theoretical probability with the experimental probability of not observing the first or second intermediate in our traces. In other words, based on their lifetimes,
we estimated the probability of missing the first and second intermediates in the unfolding trajectories to be 0.52 (0.44-0.65, Confidence Interval 95%) and 0.15 (0.13−0.19, Confidence Interval 95%), respectively. The predicted numbers were in good agreement with the fraction of single-molecule traces in which we did not observe either intermediate (0.65 ± 0.11 and 0.15 ± 0.03 for the first and second intermediate, respectively). Thus, the simplest conclusion is that the four different unfolding scenarios described in Figure 3.8A result from a single ClpXP-mediated GFP unfolding pathway that has two obligatory intermediates rather than a branched unfolding pathway. In conclusion, a subset of GFP unfolding intermediates was not detected due to instrument limitations.

Identifying the structures of ClpXP-mediated GFP unfolding intermediates

The crystal structure of GFP (PDB ID: 1GFL) was used to map the observed rip size of each unfolding transition in our traces to the corresponding extraction of GFP structural elements using a previously established methodology (Dietz and Rief, 2004; Maillard et al., 2011; Ornö et al., 1996; Yang et al., 1996). The observed change in contour length $\Delta L_C$ could be written as

\[ \Delta L_C = m_i \cdot L_{aa} - (X^F_{m_i+1:N} - X^F_{m_i-1+1:N}) \]  

(3.14)

The first term ($L_C$) describes the actual contour length of each unfolding transition, where $m_i$ is the number of residues that unfold during the transition to the $i^{th}$ state, and $L_{aa}$ is the contour length of one amino acid (0.365nm/aa). The second term comprises the shift in the distance between the last structured residue at the N terminus (residue position 1) and the last structured residue at the C terminus as the transition occurs. Both distances $X^F_{m_i+1:N}$ and $X^F_{m_i-1+1:N}$ are determined from the reported crystal structure coordinates.

Equation 3.14 provides an iterative means to determine which structural elements of GFP were extracted during each intermediate transition (Figure 3.8C). The force at which each unfolding transition occurred was plotted as a function of its corresponding change in extension in nanometers. Using the worm-like-chain model of polymer elasticity (Bustamante et al., 1994), a force extension curve was generated to calculate the $L_C$ of an unfolding transition occurring at each amino acid position from Eqn. 3.14 (iterate $m_i$). Selecting the contour length ($L_C$) that minimized the root-mean-square deviation (RMSD) from the WLC fit allowed us to establish optimal location of the unfolding segment for each intermediate (Figure 3.8D).

This approach provided an unambiguous identification of the first intermediate, namely, a state with $\beta$-strand 11 ($\beta11$) extracted from the GFP barrel. For the second intermediate, however, the RMSD versus amino acid position yielded two minima with similar RMSD values: one solution corresponded to the extraction of $\beta10-7$ and a second solution to the extraction of $\beta10-6$. A simple line of reasoning allowed for distinguishing between these two possibilities. When GFP unfolding occurred in the absence of the first small intermediate (Figure 3.8A, right-top and right-bottom panels), the RMSD versus amino acid position had
a single minimum corresponding to the extraction of $\beta_{11-7}$. Importantly, the distribution of intermediate lifetimes was identical to the distribution of second-intermediate lifetimes when the first intermediate was observed (Figure 3.8H-I), suggesting also similar structural stability of the unfolded intermediate. Thus, the most parsimonious explanation is that the second unfolding transition both in presence and absence of the first one corresponds to the same structural intermediate, with $\beta_{11-7}$ unraveled from the GFP barrel.

### Prediction of GFP Unfolding Time Based on Burst Size Distribution and Dwell Durations

Using the distribution of burst size and dwell duration (Figure 3.4), we estimated the probability of ClpXP taking a 4 nm burst in less than 240 ms (mean refolding time of $\beta_{11}$ from Figure 3.8K) to be on average $0.031$ ($0.018 - 0.046$, confidence interval [C.I.] to 95%). This result indicates that ClpXP must pass through on average $\sim 33$ translocation cycles before it can unfold GFP and trap the first intermediate for subsequent unfolding. Multiplying 33 translocation cycles by the mean duration of the dwell ($\sim 0.35$ s) predicts a mean GFP unfolding time of $11.5$ s ($7$ s - $214$ s, C.I. 95%), which agrees very well with the mean time constant obtained from the distribution of GFP unfolding times, $< \tau > = 11.8 \pm 0.9$ s (Figure 3.8B, rightmost panel).
Chapter 4

Conclusions

4.1 Commentary on Data Analysis

Below is a detailed outline of the analysis performed for the ClpXP system. Different techniques may be better suited for other biological systems, however the goal is to provide documentation on what was done for the ClpXP subgroup in the early analysis. A word of advice is to be extremely cautious when analyzing data (avoid bias) and use several different methods to test the robustness of certain step finding algorithms.

Downsampling Raw Data

There are several methods to filter the raw data (2kHz or 2.5kHz depending the instrument). In time series, the filtering is a function that will take an input of data points and provide an outpoint of a reduced number of data points. Filtering is dependent on the resolution of the instrument, the polystyrene bead size, the stiffness of the trap, and inherent noise from the biological data. Analysis on the raw data is generally computationally expensive, so filtering is a trade-off to reduce the noise and lower computation time for step detection algorithms. As there are many different types of filtering, I would suggest to begin with the moving average and the Savitzky-golay filtering. Extracting signal from noisy processes is an established field in mathematics, statistics, electrical and computer engineering; there is plenty of algorithms that may maximize the information extracted from the data.

We originally used the Savitzky-Golay, which performs a least-squares fit to a set of points using a polynomial (to the nth degree) to obtain filter coefficients. Although some information may be lost, the advantage of the Savitzky-Golay is that this algorithm preserves high-frequency components of the data. The main disadvantage is that it is not as a successful in rejecting noise compared to a moving average and smoothing alone has only cosmetic value (Guiñón et al., 2007; Savitzky and Golay, 1964). Then, we decided to use a moving average and then decimate to reduce the noise and number of points. A moving average smooths the raw data by taking the average of an number of consecutive points within a span, which is equivalent to a lowpass filtering (Guiñón et al., 2007; Doucet and Johansen, 2009).
CHAPTER 4. CONCLUSIONS

Figure 4.1: (A)-(C) are of the same single-molecule trace, which had three different filters applied: (A) Savitzy-Golay algorithm with a polynomial of 4 and span of 101, (B) a simple moving average was used with a span of 101, and (C) a moving average and decimation filter was applied with a span of 101. Both Savitzy-Golay and moving average retain the same number of points, while the moving average and decimation is reduced by $\sim 100$ fold in the points.

As seen in Figure 4.1, the Savitz-Golay and moving average smoothing with a span will still maintain an output of same number points in the time series but now each point is weighted by a function of the neighboring points within this span. Only decimating without smoothing the time series by a factor of 10 would mean that the points in output would be 10, 20, 30, 40, 50, ..., 1000. It would just reduce the point out and the noise would not be reduced. However, if smoothing was done without decimation, the noise will be reduced, giving the impression that there is more information. The simple moving average with decimation is recommended to avoid over extraction of detailed dynamics through advanced signal processing algorithms.

Pause-free Velocity

The velocity of a motor is the change in distance over time. There are two basic ways used in analysis for ClpXP. We either 1) take the average of the instantaneous velocity or 2) take the distance and time endpoints of the trajectories over a certain force range. In order to calculate the instantaneous velocity, we used the function `diff` in Matlab. The main disadvantage to calculating the instantaneous velocity is the computation time to calculate in the certain force bin for every molecule. While the instantaneous velocity shows differences in consecutive points, which can capture some more dynamics (such as slips). Traditionally with this method, the instantaneous velocity is plotted as a histogram with pauses should be distributed around 0 nm/s, which allow the user to calculate a threshold to separate the pauses from the pause-free velocity. Unfortunately, the calculation of pause-free velocity becomes difficult when all of the instantaneous velocities have a broad histogram distribution, as in the case of ClpX. Selecting a threshold to separate the populations then becomes
tricky and lacks robustness to inaccurately distinguish the pauses from the actual speed of the motor. That led our focus to shift using the endpoints to calculate the velocity, which would more accurately calculate the slope and standard deviation of the single-molecule trajectories. Then we coupled this with a step-finder known as the Schwartz Information Criterion (SIC) to detect regions where there were considered pauses (Figure 4.2).

SIC is a fitted model that assesses whether this model is a balance between the 'goodness of fit' and parsimony, ideally corresponding to a model which is most plausible by the data. The SIC is an asymptotic approximation to the transformation of the Bayesian posterior probability of a candidate model. The Bayesian posterior probability model is starting from observed events and a model, it gives the probability of the hypotheses that may explain the observed data (i.e. of the unknown model parameters). The data should be a case of independent, identically distributed observations and linear models under the assumption that the likelihood is from the regular exponential family. Let say that X is our observed data. Our goal is to have X be described through model Y from a candidate models (Y1 to YL), with each model being unique in its vector parameterization $\theta$ (in our case, the mean and variances of the data) and n (the number of data in the set) this space. Let p denote the number of independent parameters in $\theta$ that are estimated. Let $L$ represent the likelihood for $\theta$ on X and denote the estimator of $\theta$ obtained by maximizing the likelihood over. Thus, the SIC is defined as SIC = $-2 \log L(\theta) + p \log n$ and adjusts the number of fitted parameters with penalty that depends on size of the sample n. The SIC is computed for each of the models and the model corresponding to the minimum value of SIC is chosen (Schwarz et al., 1978; Cavanaugh and Neath, 1999; Chen et al., 2008). The SIC has applications in a variety of different systems, however the SIC is sensitive to noise and should be used with care in analyzing single-molecule trajectories. There are pros and cons of every pause-detection code, but the best hope is to find a couple algorithms that will accurately and robustly detect pauses.

**Dwells and Step size of the motor**

Decreasing the polystyrene bead size from 2 to 0.9 μm allowed for more consistent visualization and detection of the motor’s steps. As stated before, the ClpXP translocation is composed of a dwell and burst phase, during which the motor remains stationary on its substrate and the motor moves along its substrate a certain distance, respectively. Characterizing the dwell-burst phase behavior is essentially a change-point problem, which estimates the changes in mean and variance of the signal within the time series. Now, the nontrivial aspect of these problem is the data composed of steps can be hidden by noise. In the case of ClpXP, there were several analysis performed on the step size: pairwise correlations combined with student T-test and modified SIC with kernel density (Figure 4.3) (Bustamante et al., 2009; Chistol et al., 2012). All robustly confirmed that ClpXP takes 2-4 nm steps with limited 1 nm steps.

There were several benchmarks we employed in our single-molecule trajectories: 1) Fitting to the entire single-molecule trajectories is haphazard and potentially dangerous de-
Figure 4.2: (A) Using the SIC, we were able to generate distributions of dwell and pause regions. (B) The distribution of dwells including pauses were plotted as a histogram. (C) Then, the distribution of dwells including pauses were fit to a gamma function and the threshold to remove pauses was calculated as the mean of the distribution plus three standard deviations (∼1-2 depends on ATP concentration). To test the robustness of this threshold, we calculated the pause-free velocity and average velocity. We found that a threshold of 1-2 seconds more significantly affected the standard error of the mean.

Mechanochemical studies of the motor

Chemical titrations are a natural method to perturb certain aspects regarding the mechanochemical cycles of a motor (ATP binding, ADP and P$_i$). In single-molecule experiments, the titration of a certain nucleotide in the buffer condition can lead to interesting results. Based on the Chemla paper, we titrated with various concentrations of ATP, ADP, P$_i$, AMP-PNP and ATP-γS and measured the extension as a function of force (5-18 pN). Traces beyond 18 pN become difficult to collect due to fatal slippages of the motor. We then calculated the average and pause-free velocity of each nucleotide or analog concentration at several force

...
Figure 4.3: (A) This panel represents the pairwise combined with the Student T-test code that will detect constant stepping regions of ClpX. (B) This panel uses the SIC with a penalty factor of 3 (Chistol et al., 2012). While each code will detect clean steps of data shown here, there are advantages and disadvantages to each algorithm. There are pros and cons of each step-finding algorithm. While the pairwise distribution and T-test are not well-suited to a variable step such as ClpXP, the algorithm is better for motors like \( \phi 29 \). In addition, the long dwell time of ClpX actually skews the correlation in distance. While SIC may be a better tool for a motor that takes a variable step, there are some implicit problems as the step-finder is sensitive to correlations found in the noise.

bins (data was filtered to 100 Hz). Typically, the several different bin sizes were chosen, varying from 1 to 5 pN. Initially, the mean and variance of the velocity were plotted in 1 pN bins, however if they were not statistically significant from the prior or subsequent bin, they were pooled together in order to obtain a significant \( n \) (number of molecules) > 20 per bin. In the case of ATP and ADP titrations, one can extract Michaelis-Menten parameters by plotting the velocities as a function ATP concentration for one force bin at a time. The Matlab curve fitting tool and Origin predicted similar Michaelis-Menten parameters with a Hill coefficient of 1 (the Hill coefficient was left unbounded). Origin has a superior plotting method than the Matlab curve fitting tool as the underlying code uses a combination of the Levenberg-Marquadt (L-M) and the Gauss-Newton to numerically iterate for the calculations.

As for the ATP\( \gamma \)S experiments, we titrated surprisingly high amounts of the analog and observed analog-induced pauses. The data was filtered to 100 Hz and fit to a modified SIC
previously used in the lab at a certain force bin (Chistol et al., 2012). In order to avoid overfitting of the pauses, the fits matched the center of clustered Gaussians to ensure they belonged to the same pause or were different. The fits were plotted in a histogram and fit to a biexponential, which extracted the mean pause duration and density. The effect of ATP$_\gamma$S on the single-molecule ClpX is dramatic and potentially more analysis could be provide insight in the coordination of the motor. AMP-PNP experiments were performed in hopes of finding an analog that ClpXP would not be able to hydrolyze, however high amounts of this analog were required ($> 1$ mM AMPPNP with 0.5 mM ATP) to even begin observe changes in the traces.

4.2 Concluding Remarks and Future directions

Single-molecule experiments on the ClpXP system has provided us a unique tool to inves-
tigate the inner workings of a protease. ClpXP was the first protease characterized to be characterized in single-molecule by our group and others. In chapter 2, we learned we have determined several unique properties of the motor previously undetectable in bulk. First, ClpX translocates unfolded polypeptide at a maximum velocity of $35 \pm 11$ amino acid/second ($\mu \pm \sigma$) with an opposing force between 5 to 20 pN, which is only half of the reported velocity in bulk studies under no force. As the opposing force approaches 20 pN, the motor eventually stalls movement. Since the force applied is equal but opposite to the motor processing the substrate, ClpX can generate up to 20 pN of force to translocate and unfold proteins. Second, ClpX unfolds the well-characterized, single-domain GFP in a discrete manner, and here we monitor well-defined unfolding intermediate that GFP may undergo for disassembly in the cell. Last, ClpX either pauses or slips backwards by losing grip on the protein when it encounters a high mechanical barrier or experiences a high opposing force. Slippage may be a potential fail-safe mechanism, allowing ClpX to release substrates that are difficult to process.

In chapter 3, we characterized the mechanisms by which the six ClpX subunits use chemical energy of ATP to generate force and communicate with each other to success-
fully degrade substrates. Our results provided the first direct evidence of a force-generation mechanism and that the phosphate release is coupled to the force-generating step of ClpX. Furthermore, we demonstrated that two to four subunits of the ClpX hexamer actively participates in bursts of translocation and that between the translocation bursts, the motor has a mean constant dwell duration independent of ATP concentration. Our analysis revealed defined a new archetype of motor coordination, which is critical as a fail-safe mechanism to prevent the motor disengagement from its substrate. It has been truly remarkable to work on this system, and the story does not end here, but rather a host of interesting questions remain to be discovered. I look forward to the future and learning even more about the ClpXP protease.
Linking the remaining events in the ATPase cycle to the dwell-burst behavior of ClpXP

Our studies on the ClpXP system have naturally led to more questions regarding the ATPase cycle of the ClpX subunits. In last paper, we were able to determine that the $P_i$ release occurs during the burst. Although there were hints of the ATP binding occurring in the dwell, we still do not know the location of ATP binding, ATP hydrolysis, and ADP release. In addition, could one of the steps in the ATPase cycle be linked to the 'internal' signal, such as the conformational changes associated with ATP tight binding? Other motors predict that these events would be located during the dwell (the stationary phase in our single-molecule trajectories). However, the barrier in our studies is the inability to probe robustly below the $K_m$ of the system. This leads us to another curiosity behind the motor’s function or lack of in this regime of ATP.

Recently, our lab has characterized the finer details of the ATPase cycle of φ29 DNA packaging motor (Chistol et al., 2012). In their studies, they were able to locate that ATP hydrolysis occurs within the dwell phase through ATPγS titration experiments. Furthermore, experiments with orthovanadate (a molecule that inhibits ADP release) demonstrated that the ADP release event occurs within the dwell and were interlaced with ATP binding. These careful titration experiments and analysis could prove fruitful for understanding how ClpXP works under such conditions.

Intersubunit coordination of ClpX

While our studies demonstrated a novel mechanism of translocation for ClpXP, the underlying ATP hydrolysis order around the ring still remains unclear. We do not know whether the coordinated firing of 2, 3, and 4 subunits occur in a strict sequential, sequential or stochastic manner. Here, the use of the single-chain ClpX will be a powerful tool for single-molecule studies. Previous studies on the mutagenesis with single-chain ClpX have already demonstrated some fundamental mechanisms on intersubunit ClpX coordination. The ATPase and degradation activities of mutants and wild-type were studied in various arrangements within the ring. E185Q (E) and R370K (R) were mutated in the Walker B and sensor-2 motifs. The E185Q mutation blocked hydrolysis but allowed conformational changes associated to ATP bindings, whereas the R370K mutation blocked hydrolysis and does not allow for the conformational changes coupled to ATP hydrolysis (such as ClpP and recognition tag binding) (Martin et al., 2005). The findings indicated that the power stroke is coupled to ATP hydrolysis in a single subunit, suggesting evidence for a probabilistic sequence of firing events.

Single-molecule studies on different arrangements of mutants and wild-type (W) would allow us to obtain high resolution information on ClpX intersubunit coordination. Currently, the Sauer group has reported on a RWERWE mutant that takes 1-3 nm with a predominante peak at 2 nm (Cordova et al., 2014). Further studies on different combinations of R and E
may finally resolve the firing order within the ring in terms of the dwell-burst behavior and whether or not the firing order is sequential or stochastic.

**ClpXP pore-1 loop couples ATP hydrolysis and mechanical work**

Each ClpX subunits contains a protruding conserved aromatic hydrophobic dipeptide in its central pore. The aromatic hydrophobic dipeptide is critical in substrate binding, regulating ATPase rates, and transmitting force to denature substrates. The ATPase rates and unfoldase ability of Y153A, V154A, and V154F pore-1 loops mutants were characterized. Mutants with decreased volume in their side chains had increase ATPase rates, where the mutation with the larger side chain had lower ATPase rates. These studies suggested that tighter packing of residues in the pore may result in slower conformational changes and that these transitions could be the slow step in the ATPase cycle. Furthermore, the tyrosine residue in the conserved GVYP loop couples ATPase rates to the mechanical work of protein unfolding and translocation. The tyrosine is important in the the motor’s ability to grip the substrate, as its deletion results in high ATPase rates, slips and inability to denature substrates (Martin et al., 2008a).

These studies bring forth interesting questions about the fundamental conversion of chemical energy to mechanical work. The force is transmitted through this conserved structural sequence and mutations in this region lead to a drastic change in the consumption of fuel. Given this information, how would the general properties of the motor be affected (the velocity, stall force, and unfolding ability)? And this study suggest, would the conformational changes of the pore-1 loops be the slow step in the cycle of ATP binding, hydrolysis and product release? Is the internal signal coupled to the conformational changes associated with these pore-1 loops?

**The interplay between ClpX and ClpP**

A central question regarding these ATP-dependent machines is how the ATPase and peptidase coordinate their mechanochemical activities to carry out tasks of protein unfolding and degradation. The interactions between ClpX-ClpP have several functional roles: there are high affinity contacts near the periphery of ClpX and ClpP rings that position the mismatched rings correctly and help increase the substrate degradation activity. These static interactions are critical to maintain a stable ClpXP complex. The relatively weak axial contacts between loops at the bottom of the ClpX central channel and N-terminal loop of ClpP vary dynamically with the nucleotide state of individual ClpX subunits, control ATP-hydrolysis rates, and facilitate efficient protein unfolding (Martin et al., 2007).

Given our current model of the ClpXP translocation mechanism, how does ClpP impact this mechanochemical cycle during translocation and motor properties of ClpX? We have initial evidence that ClpP lengthens the dwell duration is $<\tau> = 351 \pm 11$ ms with burst sizes between 2 to 4 nm, whereas in the absence of ClpP, the dwell duration is slightly shorter with $<\tau> = 260\pm7$ ms and maintains variable burst sizes of 2 to 4 nm with a small pop-
ulation at 5 nm. The duration of the burst appears not to change with ATP concentration. It would be interesting to study these individual components of the ClpXP protease as ClpX has the ability to work independently. Specific details on their communication is critical in understanding the mechanochemistry of this protease.
References


[40] Greg L Hersch et al. “Asymmetric Interactions of ATP with the AAA+ ClpX6 Unfoldase: Allosteric Control of a Protein Machine”. In: Cell 121.7 (July 2005), pp. 1017–1027.


REFERENCES


