2018

Examination Of The Dynamic Assembly Mechanism Of E. Coli Clpa

Elizabeth C. Duran
University of Alabama at Birmingham

EXAMINATION OF THE DYNAMIC ASSEMBLY MECHANISM OF *E. COLI* CLPA

by

ELIZABETH C. DURAN

AARON L. LUCIUS, COMMITTEE CHAIR
CHRISTIE BROUILLETTE
SUZANNE E. LAPI
PETER E. PREVELIGE
CHARLES L. TURNBOUGH JR.

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2018

EXAMINATION OF THE DYNAMIC ASSEMBLY MECHANISM OF *E. COLI* CLPA

ELIZABETH DURAN

CHEMISTRY

ABSTRACT

E. coli ClpA is an AAA+ (ATPase Associated with diverse cellular Activities) chaperone that catalyzes the ATP-dependent unfolding and translocation of substrate proteins targeted for degradation into a protease, ClpP. ClpA, like many other AAA+ proteins, assembles into a hexameric ring competent for binding polypeptide substrate clients in the presence of ATP. Each ClpA protomer contains two nucleotide binding domains, NBD1 and NBD2. Recently, our lab applied hydrodynamic techniques to quantify the self-assembly mechanism of a closely related AAA+ chaperone, *E. coli* ClpB. Here we apply the techniques and analysis from that work to investigate the nucleotide-linked self-assembly mechanism of ClpA. In this work, sedimentation velocity studies are used to quantitatively examine the ClpA self-assembly mechanism in the absence of nucleotide for wild type ClpA ($ClpA_{WT}$) and various ClpA variants shown to be deficient at hydrolyzing ATP at one or both NBDs. We observe differences in the resulting self-assembly equilibrium constants obtained for $ClpA_{WT}$ relative to that obtained for each ClpA variant. In that work we showed that changes to the primary sequence of the proteins could perturb the assembly mechanism. Sedimentation velocity studies were also performed on ClpA in the presence of ATP using variants deficient in ATP hydrolysis at both NBDs. The stoichiometry and affinity of nucleotide binding to NBD1 and NBD2 are revealed by examining the dependence of the apparent association equilibrium constants

on nucleotide concentration. In this work we show that in the presence of ATP, ClpA resides in a distribution of monomers, dimers, tetramers, hexamers, and dodecamers. Average stoichiometries and affinity for ATP binding to each oligomer are revealed from modeling the apparent equilibrium constant for dimerization, tetramerization, hexamerization, and dodecamerization of ClpA as a function of free nucleotide concentration.

Keywords: ClpA, ClpB, AAA+ Molecular Chaperones, Sedimentation Velocity Experiments, Self-Assembly Mechanism

DEDICATION

To my mother, who brought me to this country as a child with the promise of pursuing my dreams and never let me lose sight of them.

ACKNOWLEDGEMENTS

I would first like to thank Dr. Aaron L. Lucius. Thank you for training and mentoring me well beyond the academic pursuits herein described. In addition to your invaluable scientific training, thank you for graciously sharing your personal advice on how to navigate the work-life balancing act of during these formative early career years. I would also like to thank my co-workers in the Lucius lab for making work a fun place to go regardless of the scientific hurdles. A special thanks to JiaBei Lin, Tao Li, Nate Scull, Zachary Ingram, and Clarissa L. Durie for all of their advice and friendship throughout my graduate school years. A special recognition to Clarissa for enthusiastically giving me her vote of confidence for just about all the things I dreamt up while in the program, as well as for being my confidence buddy, partner in crime, favorite collaborator, and dear friend. My thanks also to my graduate committee, Christie Brouillette, Suzanne Lapi, Peter E. Prevelige, and Charles L. Turnbough Jr. for their time and intellectual contributions to advancing me towards independence throughout my graduate career.

I would like to also acknowledge and thank my family and friends, whose daily support gave me the healthy balance and perspective needed to push through the challenging times. First, my deep and sincere gratitude to my loving husband, Darek D. Krolczyk, for upholding and amplifying the best parts of myself, and for all of his patience and kindness throughout this pursuit. Lastly, I thank my Birmingham family, Bill Miller,

Allison Argall, Matthew Thomas, Elizabeth Thomas, Haley Staton, and Shaun Staton for keeping me grounded and keeping me laughing regardless of what life throws our way.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table *Page*

LIST OF FIGURES

Figures *Page*

COMPARATIVE ANALYSIS OF THE STRUCTURE AND FUNCTION OF AAA+ MOTORS CLPA, CLPB, AND HSP104: COMMON THREADS AND DISPARATE FUNCTIONS

ATP HYDROLYSIS INACTIVATING WALKER B MUTATION PERTURBS *E. COLI* CLPA SELF-ASSEMBLY ENERGETICS IN THE ABSENCE OF NUCLEOTIDE

EXAMINATION OF THE NUCLEOTIDE LINKED ASSEMBLY MECHANISM OF *ESCHERICHIA COLI* CLPA

LIST OF ABBREVIATIONS

AAA+	ATPase associated with various cellular activities
ATP	Adenosine 5'-triphosphate
ATPγS	Adenosine 5'-[gamma-thio]-triphosphate
Clp	Caseinolytic peptidase
CY5	Cyanine 5 Maleimide dye
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
Hsp	Heat shock protein
2-ME	2-mercaptoethanol (β-mercaptoethanol)
NBD	Nucleotide binding domain
NLLS	nonlinear-least-squares
TNP-ATP	2'-(or-3')-*O*-(2,4,6-tinitrophenyl) adenosine 5'-triphosphate

CHAPTER 1

INTRODUCTION

ATPases associated with diverse cellular activities (AAA+ proteins) carry out myriad regulatory and post-translational quality control functions necessary for proper proteome maintenance in all organisms (1-3). The functions catalyzed by many AAA+ enzymes center around the regulation of protein folding pathways to prevent or correct protein misfolding. In the densely packed environment within cells, proteins can unfold or misfold in response to cellular stress, such as exposure to extremes of temperature, pH, or chemical toxins. Left unregulated, the exposed hydrophobic regions of unfolded proteins can nonspecifically interact with various components of the crowded cellular environment leading to the formation of protein aggregates.

In humans, protein aggregates have been linked to a number of intractable diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease, and type II diabetes (4). To reverse or prevent the formation of toxic protein aggregates, cells of various organisms have evolved a number of protein quality control mechanisms involving AAA+ molecular chaperones. Detailed characterization of these molecular mechanisms is in large part spurred by an interest in developing effective therapeutics to prevent or delay the onset of protein aggregation-linked diseases.

In this dissertation, we focus on the relevance of characterizing the self-association mechanisms of a subclass of AAA+ molecular chaperones to understand the molecular

basis for their protein quality control mechanisms of action. To this end, we first review the structure and function of three closely related AAA+ molecular chaperones: ClpA, ClpB, and Hsp104. These three proteins belong to a subclass of the AAA+ superfamily of proteins known as Clp/Hsp100 (5, 6). In this review, we identify the challenges inherent in studying the molecular basis of the two major protein quality control functions catalyzed by these molecular chaperones: protein unfolding and protein remodeling.

As described in Chapter 2, characterizing the population of chaperone-associated complex that catalyzes substrate unfolding or remodeling is one important aspect of designing experiments to study ClpA, ClpB, and Hsp104. These proteins function as hexameric rings and often work with various binding partners (for review see Chapter 2 and (3, 7, 8)). In addition, ATP hydrolysis drives both the oligomeric assemblies of these AAA+ chaperones as well as the substrate unfolding or remodeling reaction catalyzed by the chaperone. Therefore, in order to measure the concentration of the complex catalyzing substrate unfolding/remodeling, we need to determine the equilibrium constants for the protein-protein interactions involved in forming the complex. Specifically, this requires the equilibrium constant for the formation of the biologically active oligomer of the AAA+ protein and the equilibrium constant for the binding between that oligomer and any associated co-chaperones.

In the second part of this dissertation, Chapters 3 and 4, we present strategies for quantifying the dynamic self-assembly mechanism of ClpA in the absence or presence of ATP. In addition, we examine whether changes to the primary sequence of the protein, that is introducing mutations often used to study these AAA+ molecular chaperones, perturbs the assembly mechanism. The hydrodynamic techniques and analysis described can be

broadly applicable to similarly complex assembly mechanisms of other AAA+ molecular chaperones. To understand the relevance of these pursuits, we briefly review the overall mechanisms of AAA+ unfoldases and disaggregases in the context of protein quality control.

AAA+ Molecular Chaperone Catalyzed Protein Quality Control

AAA+ molecular chaperones involved in protein quality control mechanisms function in two major pathways that either prevent or reverse the formation of protein aggregates. One of these is the targeted substrate unfolding and subsequent degradation by energy dependent molecular chaperone systems (7). In these systems, the AAA+ chaperone associates with a compartmentalized protease. Multiple rounds of ATP binding and hydrolysis catalyzes allows to chaperone to apply physical force on the substrate for the purpose of unfolding it (9-11). In the presence of the partner protease, unfolded substrates are handed off from the molecular chaperone to the protease for degradation. This is achieved through a conserved structural homology in which the AAA+ motor protein flanks one of both ends of the proteolytic component.

In prokaryotes the AAA+ proteins ClpA, ClpX, and ClpC interact with the protease ClpP forming ClpAP, ClpXP, and ClpCP complexes (12). Polypeptide substrates recognized and bound by the AAA+ motor are translocated into a central cavity of the ClpP protease for degradation. The same structural organization is found in archea where the AAA+ motor proteins PAN and Cdc48 interact with the 20S protease, and in the much more complex eukaryotic energy dependent protease, the 26S proteasome (13-15).

An alternative strategy for regulating the formation of protein aggregates, is achieved by AAA+ molecular chaperones that function as disaggregation machines capable of rescuing proteins from an aggregated state (16). The first of these disaggregases to be identified was the yeast molecular chaperone Hsp104 (ClpB in bacteria) (17, 18). Hsp104 in collaboration with the Hsp70 and Hsp40 co-chaperone system (DnaKJE co-chaperone system in bacteria), has been shown to solubilize proteins from an aggregated state and restore their function once freed from the aggregate (17).

Despite extensive research since the discovery of molecular chaperones involved in energy dependent protein degradation and disaggregation in the 1990s, the underlying molecular mechanisms remain incompletely understood. Key questions regarding polypeptide substrate recognition, the coordination of nucleotide binding and hydrolysis to polypeptide substrate processing, and the effects of co-chaperones on those activities remains unclear.

One question that remains contested in the field is, what are the differences in the molecular mechanisms between substrate unfolding and substrate disaggregating chaperones? The structural similarities between AAA+ Hsp100 unfoldases (ClpA, ClpX, Cdc48) and the disaggregases (ClpB, Hsp104) have formed the basis for hypothesizing that protein disaggregation and unfolding proceed through a common mechanism whereby polypeptide substrate is completely threaded through the axial channel of the ring-like hexameric structure of the chaperone (5, 19). Yet, recently emerging asymmetric spiral hexamer structures captured with higher resolution techniques are revealing the possibility that the disaggregases undergo conformational switching between axial and asymmetric ring structures (20, 21). These findings may lend support for partial (non-processive)

threading as opposed to complete (processive) threading models for disaggregation. Findings from our lab support a model of non-processive translocation for ClpB catalyzed substrate processing. In this model, ClpB takes one to two steps on the polypeptide substrate before falling off (22). This mechanism would allow for the possibility of falling off and rebinding on several parts of an aggregate to achieve disaggregation. A similar non-processive translocation mechanism is emerging from our lab for the disaggregation catalyzed by Hsp104 (Weaver, et. al., manuscript in preparation). The new structural data have spurred interest in re-examining of the complete threading model and created a need to develop experimental designs to test newly emerging hypothesis.

Importance of Characterizing the Macromolecular Assembly State of AAA+ Molecular Chaperones

Incomplete characterization of the assembly state of AAA+ chaperones remains a barrier to answering many of the remaining questions for understanding their molecular mechanisms. In previous studies on ClpB assembly, our lab has shown that ClpB resides in a dynamic equilibrium of various oligomers in the absence and presence of ATP (23, 24). Using absorbance sedimentation velocity techniques, we similarly observe that ClpA resides in a mixture of oligomeric states in the absence of nucleotide (25, 26). Yet, most of the studies designed to characterize the molecular mechanism of polypeptide substrate recognition and processing by these AAA+ molecular chaperones do not account for the complex macromolecular assembly state.

Previous characterizations of ClpA assembly have largely relied on gel filtration assays to characterize the formation the hexameric population. For example, Kress and co-workers designed experiments to examine the ATPase activity of each NBD during ClpA

catalyzed substrate unfolding (27). For this purpose, they designed ClpA variants shown to be inactive in ATP hydrolysis at NBD1 or NBD2. In order to assess whether the mutations incorporated at each NBD disrupts ClpA assembly, the authors observe that the ClpA variants form hexamers capable of interacting with ClpP using analytical gel filtration assays. The problem with this analysis is that the observation of oligomers at a fixed concentration of protein and nucleotide does not preclude the possibility that the assembly is perturbed by the mutations. In this and a similar study designed to test the differences in ATPase activity of each ClpA NBD in the absence and presence of ClpP, the hexamer concentration is calculated by dividing the monomeric concentration by six (28). If the concentration of ClpA does not depend linearly on the concentration of ATP, then this method of determining the concentration would lead to an overestimate for the hexameric population. This in turn would lead to errors in the reported Michaelis-Menten parameters resulting from the steady state ATPase analysis. In order to investigate the nature of the ATP dependence for oligomerization of ClpA, we perform a thermodynamically rigorous examination on the self-assembly of ClpA in the absence and presence of ATP (Chapter 3 and Chapter 4, respectively).

CHAPTER 2

COMPARATIVE ANALYSIS OF THE STRUCTURE AND FUNCTION OF AAA+ MOTORS CLPA, CLPB, AND HSP104: COMMON THREADS AND DISPARATE FUNCTIONS

by

ELIZABETH C. DURAN, CLARISSA L. WEAVER, AARON L. LUCIUS

Frontiers in Molecular Biosciences, Volume 4, Article 54

Format adapted for dissertation

ABSTRACT

Cellular proteostasis involves not only the expression of proteins in response to environmental needs, but also the timely repair or removal of damaged or unneeded proteins. AAA+ motor proteins are critically involved in these pathways. Here, we review the structure and function of AAA+ proteins ClpA, ClpB, and Hsp104. ClpB and Hsp104 rescue damaged proteins from toxic aggregates and do not partner with any protease. ClpA functions as the regulatory component of the ATP dependent protease complex ClpAP, and also remodels inactive RepA dimers into active monomers in the absence of the protease. Because ClpA functions both with and without a proteolytic component, it is an ideal system for developing strategies that address one of the major challenges in the study of protein remodeling machines: how do we observe a reaction in which the substrate protein does not undergo covalent modification? Here, we review experimental designs developed for the examination of polypeptide translocation catalyzed by the AAA+ motors in the absence of proteolytic degradation. We propose that transient state kinetic methods are essential for the examination of elementary kinetic mechanisms of these motor proteins. Furthermore, rigorous kinetic analysis must also account for the thermodynamic properties of these complicated systems that reside in a dynamic equilibrium of oligomeric states, including the biologically active hexamer.

INTRODUCTION

The central dogma of molecular biology tells us that proteins are constantly being produced by the cell upon exposure to environmental stresses, nutrients, and metabolites. For example, if we expose cells to a source of lactose we know that synthesis of all of the proteins responsible for lactose metabolism will be upregulated in response. However, the central dogma does not address what happens to those gene products when the lactose is gone. Indeed, the cytosol is a protein rich environment. However, every protein that was produced to respond to stimuli cannot persist in the cytosol when the stimuli are removed and the protein is no longer needed. Rebinding of repressors and removal of the mRNA are two aspects of this. Yet stemming the flow of nascent protein does not address the manner in which they are removed when new and different proteins are needed.

Generally, longer-lived proteins are sequestered into lysosomes for degradation. Shorter-lived proteins are degraded in the cytosol. The presence of a PEST region (region rich in proline, glutamate, serine, and threonine) has been associated with shorter protein half-lives.(1) The N-end rule, proposed in the 1980s and expanded upon since then, proposes that certain amino terminal residues promote ubiquitination in eukaryotes and proteolysis, two ATP-dependent processes occurring within the cytosol.(2) Over all, cytosolic proteins can have half-lives ranging from minutes, to hours, to days (for reviews, see(3, 4)).

Proteolysis in the cytosol is a potentially dangerous activity for the cell, so removal of proteins that are no longer required presents a challenge. The cell cannot have unregulated proteolysis running rampant in the cytosol. Unregulated proteolysis in the

cytosol would deplete necessary, active proteins. In fact, because dysregulation of cytosolic proteases is deadly to cells, it has been explored as an antibacterial strategy (5, 6).

The challenge of regulating proteolysis in the cytosol is met by ATP dependent proteases, for review see (7). However, what is the requirement for ATP in ATP dependent proteolysis? Peptide bond cleavage is exergonic. Proteases do not require an energy source to catalyze proteolysis. For example, serine proteases, cysteine proteases, aspartic proteases, etc. simply bind to a polypeptide chain and cleave the peptide bond. AAA+ (ATPases associated with a variety of cellular activities) motors and ATP serve as the regulators of proteolytic activity in the protein rich environment of the cytosol.

Across species, ATP dependent proteases are composed of a barrel shaped protease with proteolytic active sites lining the interior cavity (for review see (8, 9)). These active sites are accessible by a pore on each end of the barrel that is too small for folded proteins to enter without first being unfolded. Certain AAA+ hexameric ring motors associate with each end of the barrel and couple the energy from ATP binding and hydrolysis to processive translocation of a polypeptide chain through the axial channel of the hexameric ring and into the proteolytic cavity of the protease. Thus, the energy source in an ATP dependent proteolytic reaction serves to both unfold the protein and processively translocate the unfolded polypeptide chain into the proteolytic chamber.

The 26S proteasome in humans and bacterial ClpAP are examples of ATP dependent proteases. ClpA is a AAA+ motor protein that contains two ATP binding sites per monomer and assembles into hexameric rings. These hexameric rings bind to one or both ends of the tetradecameric serine protease ClpP to form ClpAP. ClpA catalyzes

protein unfolding and translocation of the polypeptide chain into the proteolytic cavity of ClpP.

Like proteases in the cytosol, enzyme catalyzed protein unfolding in the cytosol is potentially dangerous for the cell. However, this function emerges, or putatively emerges, in many biological contexts. For example, both ClpA and ClpX, another AAA+ motor that associates with ClpP, catalyze "protein remodeling" reactions in the absence of the proteolytic component, ClpP. ClpA remodels an inactive dimer of RepA into two active monomers (10) and ClpX remodels the highly salt-stable MuA transposase (11, 12) to induce dissociation from DNA. More recently, mitochondrial ClpX was reported to partially unfold ALA synthase in a tranlocation-depedent mechanism to facilitate pyridoxal phosphate cofactor binding during heme biosynthesis(13). Although it is well established that both ClpA and ClpX processively translocate a substrate into ClpP for the purposes of proteolytic degradation, it is not clear if the motors fully translocate a substrate during protein remodeling reactions. Thus, the question remains; do the motors need to fully translocate a substrate to catalyze such protein remodeling reactions? Furthermore, do they use the same elementary mechanisms to translocate substrates for proteolytic degradation as they do for protein remodeling reactions?

The AAA+ motors Katanin (14) and Spastin (15) catalyze microtubule severing. Microtubule severing could also be classified as a protein remodeling reaction. It is thought that Katanin and Spastin catalyze this reaction by binding to unstructured tails on α– and β-tubulin (16). Then, using the energy from ATP, they either fully or partially translocate

the tubulin molecule through their axial channel. Once a monomer of tubulin is removed from the microtubule, a severing event occurs.

The N-ethylmaleimide-sensitive fusion protein (NSF) is a AAA+ motor involved in vesicle fusion (17-20). Specifically, the protein is responsible for disassembly of tightly associated SNARE proteins. NSF may also catalyze partial or complete unfolding/translocation in the process of dissembling the SNARE complex.

The AAA+ motors bacterial ClpB and yeast Hsp104 have the unique ability to recognize and disrupt protein aggregates *in vivo*. It has been hypothesized that these enzymes processively translocate a polypeptide chain out of a protein aggregate and through their hexameric ring structure (21, 22). However, more recent results suggest that complete translocation may not be the case (23).

One common thread among Katanin, Spastin, NSF, ClpB, and Hsp104 is that they do not interact with a protease and they are not, themselves, proteases. Thus, they do not covalently modify the substrate on which they operate. This lack of proteolytic activity leads to a technical barrier in addressing the question of whether these enzymes pass a polypeptide chain through their axial channels fully or partially. This is, in part, because unfolding alone is not evidence for complete passage. A number of studies have used GFP and its variants to examine the unfolding reaction (24, 25). However, it remains unclear how much of the GFP tertiary structure needs to be unfolded before the fluorescence is extinguished. Thus, loss of fluorescence does not allow one to conclude that complete translocation has occurred.

Complete proteolytic degradation catalyzed by ClpP is the evidence for complete translocation catalyzed by ClpA and ClpX. Much of what has been learned about translocation catalyzed by ClpA and ClpX has been determined from observing proteolytic degradation catalyzed by the protease, ClpP, in ClpAP and ClpXP, respectively. However, this leads to the question; do the motors catalyze processive translocation the same way in the absence of the proteolytic component as they do in its presence? Determining the mechanism of complete translocation catalyzed by ClpA or ClpX without covalent modification of the substrate presents the same technical difficulties as those articulated for any of the other AAA+ motors mentioned so far, i.e. the substrate on which they operate is not covalently modified.

This review is focused on efforts to examine polypeptide translocation catalyzed by AAA+ motors in the absence of proteolytic degradation. We have sought to develop a set of tools that would allow us to use transient state kinetics to examine the elementary kinetic mechanism of enzyme catalyzed protein unfolding and translocation. Specifically, we sought to determine the elementary rate constants as well as the step-size (distance per step) that define the elementary mechanism of translocation. To this end, the work began with developing strategies to examine ClpA since it was known to be a processive translocase. The work has continued by applying these approaches to the protein disaggregating machines ClpB/Hsp104. However, the work quickly revealed that in order to fully interpret the kinetic mechanistic observations a number of questions regarding the energetics of assembly and ligand binding required attention. These issues are discussed below, building on an overview of the structure of these proteins.

STRUCTURAL FEATURES OF CLPA, CLPB, AND HSP104

Primary through tertiary structure

ClpA, ClpB, and Hsp104 share similarities that have formed the basis for their classification. They are members of the AAA+ superfamily that are further classified as Hsp100 proteins for their roles in coupling ATPase activity to changes in the folding and/or assembly of substrate clients (Schirmer et al., 1996; Neuwald et. al., 1999). Hsp100 members are partitioned into two classes based on the number of nucleotide binding domains (NBDs) contained per monomeric unit. Class I proteins, such as ClpA, ClpB, and Hsp104, contain two NBDs while Class II proteins, such as ClpX, contain a single NBD per monomer. In the presence of ATP, these proteins assemble into homohexameric ring-like structures that perform their chaperone activity. ATP binding and hydrolysis occur at canonical Walker A and B motifs contained within each nucleotide binding domain (26).

The protomer structures of ClpA, ClpB, and Hsp104 have been reported from various organisms in various nucleotide-bound states. In the case of ClpA, the monomer structure has been reported from *Escherichia coli* ClpA in the ADP-bound state (27). For the disaggregases, atomic resolution crystal structures have been reported for *Thermus thermophilus* ClpB in the AMPPNP-bound state (28) and for *Chaetomium thermophilum* Hsp104 in complex with ADP (29). Comparison of the available protomer structures (Figure 1A) as well as the primary sequences of the three motors, highlights their shared structural features. In general, each monomer is made up of an N-domain, nucleotide binding domains 1 (NBD1) and 2 (NBD2) joined by a linker region, and a C-terminal domain. The residues that separate the Walker A and Walker B motifs in each NBD have been modeled to form a loop that extends into the axial channel of the hexameric ring

structures. Evidence from multiple studies has implicated conformational changes of these residues with ATP hydrolysis at each NBD in the mechanism of polypeptide substrate translocation by ClpA, ClpB/Hsp104, as discussed below in Sections 3 and 4, as well as other AAA+ motors (21, 30-35). As shown in Figure 1B, single particle reconstructions of the hexamer structures for the three motors predict similar arrangements of each domain within the quaternary structure. This structural similarity, in part, formed the basis for the hypothesis that ClpA, ClpB, and Hsp104 operate on substrate proteins through a shared mechanism.

ClpB and Hsp104 share an important structural feature that ClpA lacks. There is a middle domain (MD) located within the C-terminal end of NBD1 (Figure 1 A). In the tertiary structure, this region adopts a coiled-coil fold, made up of four α-helices, that extends approximately 85 Å from NBD1 (Figure 1 B). This domain is flexible and restriction of this flexibility has been shown to decrease disaggregation activity (28). MD flexibility has made its position and orientation within the hexamer difficult to assign in the multiple ClpB/Hsp104 structures available. The variable MD orientations in hexameric models have led to the hypothesis that nucleotide driven conformational switching of the MD may be an important part of the ClpB/Hsp104 disaggregation mechanism (36-38). Various studies have also shown the MD to be the binding target of ClpB/Hsp104 co-chaperones, DnaK/Hsp70 (37-44).

One other structural element distinguishes the protein translocase, ClpA, from the protein disaggregases ClpB/Hsp104: the presence or absence of a tripeptide motif requisite for the assembly with ClpP. ClpA hexamers interact with the protease ClpP through a conserved IGL/F motif nestled in a helix-loop-helix region near the C-terminal end of

NBD2 (45). ClpB and Hsp104 lack that IGL/F motif, and accordingly, do not naturally associate with ClpP or any known protease.

Quaternary structure and nucleotide-linked self-assembly

In the presence of nucleotide, ClpA, ClpB, and Hsp104 oligomerize to form homo-hexamers that interact with client substrates and partner proteins. Structural models of the hexameric state have been reported for all three motors, in various nucleotide-bound states (27-29, 46-51). In most cases, the hexameric state is reported to be a planar, ring-like structure with a central axial channel as shown in Figure 1 C. In these models and single-particle reconstructions, the NBDs from each protomer align side-by-side around the hexamer, forming a NBD1 tier and a NBD2 tier. Hexamer models that capture the orientation of the flexible N-domain, have a third N-domain tier above NBD1, as seen for the hexameric single particle reconstructions in Figure 1 C. ClpB and Hsp104 hexamers additionally have the MD protruding from the NBD1 tier.

Recently, an alternative asymmetric spiral structure has been reported for the Hsp104 hexamer in the AMPPNP-bound state, and in the ATPγS bound state with casein bound as a substrate (50, 52) spurring interest and speculation about its structural implications for the disaggregation mechanism. Similarly, Ripstein *et al.* recently reported images of another AAA+ protein, VAT, which threads protein substrates through its axial channel into the proteasome for degradation, in transient, asymmetric conformations.(53) These asymmetric hexameric structures observed by cryo-EM are similar to the extended spirals reported previously in crystallographic studies (27-29)). These provocative asymmetric structures invite further investigation. Biochemical assays will be key in

determining how the asymmetric Hsp104 spiral structure fits into the disaggregation mechanism. This and other efforts to discern the mechanistic details of substrate processing by ClpA, ClpB, and Hsp104, will require the ability to precisely quantify the concentration of hexamers competent for polypeptide substrate binding.

Many studies have established that ClpA and ClpB reside in a distribution of oligomers in the absence of nucleotide (54-58). Hydrodynamic studies from Maurizi and co-workers concluded that ClpA resides in a distribution of monomers and dimers in the absence of nucleotide and that ATP is required for assembly into hexamers (54). In later work, Kress *et al.* report that ClpA hexamerization occurs through a transient tetramer intermediate (59). Using hydrodynamic and thermodynamic techniques, it was later shown that ClpA resides in a distribution of monomers, dimers, and tetramers in the absence of nucleotide (60, 61) thereby showing that the tetramer was not a transient intermediate on the pathway to assembly but was significantly populated at thermodynamic equilibrium independent of path. Notably, in the presence of excess nucleotide, ClpA hexamers as well as lower order oligomers remain in solution (61, 62). However, a complete quantification of the nucleotide linked assembly reaction is still needed.

On the other hand, the energetics of ClpB self-assembly in the absence and presence of nucleotide has been quantified (63, 64). ClpB, like ClpA, resides in a distribution of monomers, dimers, tetramers, and hexamers. An important distinction between the two motors is the observation that ClpB, unlike ClpA, forms hexamers in the absence of nucleotide. A rigorous, in-depth investigation of the self-assembly of Hsp104 is currently lacking in the field, however recent results suggest that, similar to ClpB, Hsp104 populates hexamers and lower order oligomers in both the absence and presence of nucleotide (65).

Taken together, these quantitative investigations of ClpA and ClpB self-assembly reveal that macromolecular assembly is thermodynamically linked to nucleotide binding. This has fundamental implications for the driving forces that tune the population of each oligomer in solution.

Specifically, two thermodynamic driving forces govern the self-assembly of these enzymes into hexamers: the free monomer concentration and the free nucleotide concentration. As a result, assays performed on these enzymes in which the concentrations of protein or nucleotide change throughout the experiment, must account for the changing distribution of oligomers. Failure to do so can lead to conclusions about nucleotide processing at each NBD and NBD1-NBD2 interdependence that could otherwise be explained by changes in the macromolecular state.

In much of the published work on ClpA, ClpB, and Hsp104 it has been generally assumed that in the presence of 1-2 mM nucleotide concentrations, all of the protein is in the hexameric state. This assumption is generally supported with size exclusion chromatography (SEC). However, SEC is a non-equilibrium technique, meaning that the equilibrium is perturbed by running the sample through the column. That is to say, the chemical potential of both the protein and nucleotide are changing throughout the experiment and therefore the distribution of oligomeric states is changing throughout the experiment. Moreover, the observation of hexamers in SEC does not rule out the presence of smaller oligomers. Further, it does not rule out the possibility that the self-association equilibrium has been perturbed upon introduction of a mutation in the protein.

It is clear from the self-association and polypeptide binding properties of ClpA and ClpB that smaller order oligomers do persist at saturating concentrations of nucleotide (61, 62, 64, 66). For example, Figure 2 shows the fraction of ClpB oligomers populated in the presence of 100 μM and 2 mM nucleotide as a function of total [ClpB], simulated from the reported energetic parameters for ClpB assembly (63, 64). In the presence of 100 μM nucleotide (Figure 2 A), a 1 μM ClpB sample would be made up of approximately 6 % hexamers, while 94 % of the population would reside in a mixture of monomers, dimers, and tetramers. In the presence of 2 mM nucleotide (Figure 2 B), the same sample would reside in a distribution made up of 74 % hexamers and 26 % lower order oligomers. In fact, under these conditions, even at 10 μM ClpB, the hexameric state is not fully populated, with hexamers making up ~89 % of the total [ClpB]. This fact severely limits the ability to draw conclusions about the ATPase activity of the hexamer from steady state kinetic measurements where a different distribution of oligomers is present at each substrate concentration.

The simplest explanation for why the assumption that all motor protein is in the hexameric state is problematic, is that the Michaelis-Menten equation is scaled linearly by the total enzyme concentration, i.e. $V_{max} = k_{cat} \times E_0$. Recall, E_0 is the total amount of enzyme in the experiment, which is controlled by the experimentalist, whereas, E is the free (unbound) enzyme concentration at any given time and its concentration is unknown by the experimentalist. Thus, the maximum velocity is measured at saturating substrate concentration and divided by the known total enzyme concentration, E_0, and the k_{cat} is reported.

It is important to recall that $V_{max} = k_{cat}$ x E_0 emerges from two assumptions in the derivation of the Michaelis-Menten equation. The first is that the substrate is in large excess over the enzyme. The total substrate concentration relative to the total enzyme concentration is controlled by the experimentalist but, mathematically, it results in being able to assume $[S]k_1$ is a constant in the first differential equation given by Equation (1) for Scheme 1.

$$\frac{d[ES]}{dt} = [E][S]k_1 - (k_2 + k_3)[ES] \tag{1}$$

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

Scheme 1

The second assumption, which is based on the first is that because the substrate is in large excess of the enzyme, the concentration of *ES* is considered constant or "in the steady-state". If *ES* is constant, then the differential equation above is set to zero and solved algebraically for *ES*. However to do this the free enzyme term must be replaced with $E_0 - ES$. This assumption is valid if and only if the [*ES*] is constant, which is our underpinning assumption. Thus, under constant *ES* conditions the conservation of mass equation can be rearranged to $E = E_0 - ES$.

The assumption that the total enzyme, E_0, is equal to the free enzyme, *E*, plus the bound enzyme, *ES*, only holds for a non-dissociating macromolecule. Understandably, this was not pointed out by Michaelis and Menten. However, we have not seen it expressly stated since.

The assumptions hold for self-associating systems that do not reside in dynamic equilibria, for example, if *E* forms only dimers and does not dissociate into monomers. Alternatively, if the experimentalist can maintain the concentration of the enzyme in large excess over the dimerization dissociation equilibrium constant then it may be possible to assume only dimers reside in solution. However, one has to be certain that by doing this they do not simultaneously violate the assumption that the substrate is maintained in large excess over the enzyme concentration.

If the dimer exists in a dynamic equilibrium between monomers and dimers, and concentrations of enzyme below the dimerization equilibrium constant are used, then the assumption is violated. The issue is made much more complicated for ClpA and ClpB where we, and others, have shown that both enzymes reside in a mixture of monomers, dimers, tetramers, and hexamers(56, 61, 63, 64). Moreover, the populations of these species are governed by the free concentration of the substrate (nucleotide). Consequently, the Michaelis-Menten equation will not be scaled by a simple relationship like k_{cat} x E_0. This is because the simplest relationship that one can write down that relates the known total monomer concentration to the species that reside in solution for a system such as ClpA or ClpB is given by Equation (2).

$$E_0 = E + 2E_2 + 4E_4 + 6E_6 + \sum_{i=1}^{2} ES_i + 2\sum_{i=1}^{4} E_2S_i + 4\sum_{i=1}^{8} E_4S_i + 6\sum_{i=1}^{12} E_6S_i \qquad (2)$$

where the subscript on *E* represents the oligomeric state and the subscript on *S* represents the number of nucleotides bound to that oligomer, represented with the counting index, *i*, There is no simple algebraic way to express Equation (2) to replace *E* in the

differential equation given by Equation (1). Indeed, if no other oligomers are in solution then Equation (2) simplifies to Equation (3)

$$E_0 = 6E_6 + 6\sum_{i=1}^{12} E_6 S_i \tag{3}$$

Equation (4) is typically applied to the analysis of steady-state ATPase experiments on ClpA and ClpB. The total monomer concentration is divided by six, the V_{max} is measured and divided by $E_0/6$ and a k_{cat} is reported.

$$\frac{E_0}{6} = E_6 + \sum_{i=1}^{12} E_6 S_i \tag{4}$$

But what does this parameter mean when we know that the system resides in a dynamic equilibrium and the total enzyme concentration is actually given by Equation (2) ? The answer may be that the k_{cat} is not that meaningful because it has been acquired by dividing V_{max} by a concentration that does not reflect the true hexamer concentration. However, the V_{max} itself contains meaningful information. Contained within the V_{max} is information about the self-association equilibrium constants and the nucleotide binding constants. This is because the concentration terms in Equation (2) can be replaced with the appropriate self-association equilibrium constants and the nucleotide binding constants given by Equation(5).

$$E_0 = [E]P_1 + 2L_{2,0}[E]^2 P_2 + 4L_{4,0}[E]^4 P_4 + 6L_{6,0}[E]^6 P_6 \tag{5}$$

where $L_{2,0}$, $L_{4,0}$, and $L_{6,0}$ represent the self-association equilibrium constants for the formation of dimers, tetramers, and hexamers in the absence of nucleotide, respectively. The first subscript represents the oligomeric state and the second subscript represents the number of nucleotide bound, E and E_0 are as above, and P_1, P_2, P_4, and P_6 are the partition

functions for nucleotide binding to the monomer, dimer, tetramer, and hexamer, respectively. Each of the partition functions are functions of the nucleotide binding equilibrium constants and the free nucleotide concentration. Although there are many forms that the partition functions could take, one example for binding to the monomer could be given by Equation (6),

$$P_1 = \left(1 + K_1\left[ATP\right] + K_1K_2\left[ATP\right]^2\right) \tag{6}$$

where K_1 and K_2 would represent the equilibrium constants for binding to NBD 1 and 2, respectively. This leads to the conclusion that if one observes differences in the V_{max} for various point mutations in the enzyme, especially mutations in the ATPase active site, then there are three potential explanations. The first is that the activity has been affected, which is the typical interpretation. However, the second and third explanation are that the nucleotide binding affinity or the self-association equilibrium has been affected by the mutation. If the mutation has perturbed the self-association equilibrium and/or the nucleotide binding affinity, then a series of comparisons on ATPase activity between variants and wild type enzymes at the same fixed protein concentration are not reporting on the same concentrations of hexamers catalyzing ATP turnover. Again, showing that hexamers still form upon introduction of mutation does not show that the self-association reaction has not been perturbed.

The resolution to this problem is to employ a thermodynamically rigorous technique that would allow one to measure the equilibrium constants and accurately predict the concentration of the active species in solution (63, 64, 67). In other words, define the thermodynamic parameters in Equation (5) and use them to interpret the

kinetic/mechanistic data. In general, the apparent self-association constant for the ligand linked assembly of ClpB would be given by Equation(7),

$$L_{6,app} = \frac{\left[ClpB_6\right] + \sum_{i=1}^{12}\left[ClpB_6 ATP\gamma S_i\right]}{\left(\left[ClpB_1\right] + \sum_{i=1}^{2}\left[ClpB_1 ATP\gamma S_i\right]\right)^6} = \frac{\left\{ClpB_6\right\}}{\left\{ClpB_1\right\}^6} \tag{7}$$

where the numerator represents the summation of all of the nucleotide ligation states of hexameric ClpB in solution and the denominator represents all of the nucleotide ligation states in the monomeric state. The curly braces on the right hand side of equation (7) are used as a shorthand notation for the summation on the left. Equation (7) can be simplified to Equation (8)

$$L_{6,app} = L_{6,0} \cdot \frac{P_6}{\left(P_1\right)^6} \tag{8}$$

where $L_{6,0}$ is as above, the hexamerization equilibrium constant in the absence of nucleotide, and P_6 and P_1 are the partition functions for nucleotide binding to the hexamer and the monomer, respectively. We showed, for ClpB(64), that the apparent hexamerization equilibrium constant is given by Equation (9)

$$L_{6,app} = L_{6,0} \cdot \frac{\left(1 + \kappa_6 \cdot [\mathrm{ATP}\,\gamma S]_f\right)^{m_6}}{\left(\left(1 + \kappa_1 \cdot [\mathrm{ATP}\,\gamma S]_f\right)^{m_1}\right)^6} \tag{9}$$

Where the partition functions for nucleotide binding to the hexamer and the monomer in equation (8) are given by the partition functions for the n-independent and identical sites model, a model that is commonly used to analyze ITC data and was applied to ITC data for ClpB binding ADP (46). In this model κ_1 and κ_6 are the average step-wise

equilibrium constants for nucleotide binding, m_1 and m_6 are the stoichiometries of binding to monomers and hexamers, respectively. In a thermodynamically rigorous and model independent analysis of our data we showed that 12 ATPγS molecules were bound to hexameric ClpB and one ATPγS was bound to the monomer. $L_{6,0}$ was determined in an analysis of assembly in the absence of nucleotide (63) and from an analysis of the dependence of $L_{6,app}$ on ATPγS we determined κ_6 and κ_1 (64).

What is most striking, telling, and predictive about equations (8) and (9) is that they are the simple product of two terms, the hexamerization equilibrium constant in the absence of nucleotide multiplied by the ratio of partition functions for nucleotide binding. If one seeks to introduce a mutation into a protein like ClpB then these are the parameters to interrogate. The mutation would have the ability to influence $L_{6,0}$, which represents the intrinsic propensity of the protein to assembly into hexamers. However, more likely, introduction of a mutation, especially one in the ATPase active site is likely going to influence the affinity for nucleotide. It seems highly unlikely that the affinity for nucleotide binding to the hexamer, κ_6, would not change upon introduction of a mutation in the ATP binding site. Whether the intrinsic propensity of the enzyme to assemble or the nucleotide binding affinity is perturbed Equation (9) predicts that the concentration of hexamers in solution will be effected.

The unanswered question we now seek to address is how do partner proteins influence this equilibrium? A hallmark of AAA+ protein unfoldases is that they interact with partner proteins. ClpA interacts with the protease, ClpP and various adaptor proteins. ClpB interacts and collaborates with the KJE system and Hsp104 collaborates with Hsp70

and Hsp40. Equation (9) predicts that if these protein-protein interactions perturb the nucleotide binding by either modulating the stoichiometry or affinity then this will perturb the hexamerization equilibrium constant and thereby the concentration of hexamers present in solution. It is tempting to assert that partner proteins like ClpP and the KJE system would stabilize the hexamers. However, for a ligand linked assembling system, Equation (9) informs us that the interaction could stabilize or destabilize. In fact, since the nucleotide concentration in the cell is well above the affinity constant, here we hypothesize that the ability of partner proteins to modulate the nucleotide binding affinity allows for fine control over the concentration of hexamers present and available to do work. With a detailed analysis of ClpB assembly, we now stand poised to determine how the KJE system influences self-association.

Similarly, several groups have reported that the steady-state ATP hydrolysis rate for ClpA is reduced in the presence of ClpP (68, 69). In addition to ClpP exerting allosteric control over the rate of ATP hydrolysis, again, Equation (9) predicts that this phenomenological observation could be due to many factors. Our transient state kinetics experiments have suggested that ClpA uses only the NBD2 ATPase sites to catalyze processive translocation when associated with ClpP (70, 71). This observation does not rule out the possibility that NBD1 is still binding to ATP. However, when combined with the predictions from Equation 9 it does suggest that if the system goes from a stoichiometry of binding of 12 to 6 then this would perturb the hexamer concentration. Thus, the reduction in the steady-state ATPase rate could be due to a two-fold reduction in the binding stoichiometry and thereby a reduction in the concentration of free hexamers. Alternatively, if ClpP does stabilize the hexameric form then one would have to conclude

that the elevated rate of ATP hydrolysis observed in the absence of ClpP must be due to a significant population of monomers, dimers, and tetramers rapidly hydrolyzing ATP.

The coordination of NBD1 and NBD2 has been, and continues to be, an area of great interest in the field. The use of these and other similar variants, abolishing ATP binding (Walker A) or hydrolysis (Walker B or Sensor 1) have been used by many groups to investigate the coordination of the 12 ATP binding and hydrolysis sites within a the ClpA hexamer, as well as for the ClpB and Hsp104 hexamer. One common strategy is "mutant doping," in which a variant is added to wild type protein in known ratios (42, 72-75). Many conclusions have been drawn regarding sequential, probabilistic, or concerted ATP hydrolysis mechanisms. Although the statistical distribution of the number of mutant protomers contained within a hexamer is valid, it may not hold if the mutation perturbs the assembly equilibrium. Many of these studies suffer from the assumption that the entire population of protein resides in the hexameric state. The most convincing among them are experiments where the signal is only sensitive to the hexameric form. For example, it seems clear that ClpX invokes a stochastic model since the studies used a linked hexamer (76, 77). Consequently, the issues surrounding assembly have been removed.

MECHANISMS OF POLYPEPTIDE TRANSLOCATION BY CLPA AND CLPAP

ClpA mechanism in the absence of ClpP

Horwich and coworkers showed that ClpAP could catalyze global unfolding of an SsrA tagged GFP construct (24). This was done by incorporating the eleven amino acid SsrA tag, which is a known binding sequence for ClpA and ClpX, at the carboxy-terminus of GFP (78). When the GFP-SsrA construct was presented to ClpA in the presence of ATP, a slight decrease in fluorescence was observed. However, when the construct was presented to ClpAP in the presence of ATP, a near complete loss of fluorescence was observed. This was interpreted to mean that when ClpA unfolded the GFP in the absence of a protease, GFP was allowed to spontaneously refold. However, in the presence of the proteolytic component, GFP was degraded and thus complete loss of fluorescence was observed.

To examine directional translocation catalyzed by ClpA, Horwich and coworkers developed a FRET based assay (79). In this design, a donor fluorophore was placed in the central cavity of ClpP and an acceptor at various positions on model substrates all containing the SsrA sequence at the carboxy-terminus. If ClpA translocates the polypetide chain into the ClpP cavity from the SsrA sequence at the carboxy-terminus directionally to the amino-terminus, then FRET time courses would reveal this. FRET time courses were consistent with processive translocation from the carboxy-terminus to the amino-terminus. The results clearly showed that ClpA drives translocation of a polypeptide chain into the proteolytic chamber of ClpP.

Until recently, the elementary kinetic parameters governing this translocation reaction had not been reported. Moreover, most of the mechanistic investigations available

were performed in the presence of ClpP. Thus, the critically important elementary kinetic mechanism for polypeptide translocation catalyzed by ClpA was missing from the field. Determining this mechanism required the development of techniques that would be sensitive to the elementary steps in polypeptide translocation in the absence of proteolytic degradation. Such approaches could then be broadly applied to a variety of enzymes that do not associate with proteases (see examples in the Introduction). This kinetic mechanism would include the elementary rate constants governing the reaction, kinetic step-size (amino-acids translocated between two rate-limiting steps), processivity (probability the enzyme will translocate vs. dissociate), and directionality (C to N vs. N to C).

A single-turnover fluorescent stopped flow assay was developed to elucidate these kinetic parameters (80, 81). Figure 3 shows a generalized schematic representation of this rapid mixing assay. Synthetic polypeptide substrates containing the eleven amino acid SsrA binding sequence at the carboxy-terminus and a single cysteine at the amino-terminus were constructed. The sequence of the polypeptide was based on the Titin I27 domain because the long term goal was to move to full length tandem repeats of I27 as had been done for ClpX (82, 83). The cysteine was labeled with fluorescein-5-maleimide. ClpA was bound to the SsrA sequence in the presence of the slowly hydrolysable ATP analogue, ATPγS. Upon ClpA binding, fluorescence quenching was observed. Fluorescence quenching has since been observed for binding by both ClpB and Hsp104 to their respective substrates (23, 65) This sample was then loaded into one syringe of the stopped-flow fluorometer, see Figure 3. In the other syringe was loaded a large excess of ATP and unlabeled SsrA peptide to serve as a trap for ClpA, i.e. any free ClpA would rapidly bind to SsrA and not the fluorescently modified polypeptide (81). The large excess of trap

ensures single-turnover conditions with respect to the complex of ClpA bound to fluorescently labeled peptide.

In the single-turnover fluorescence assay, the two solutions are rapidly mixed within 2 ms in a stopped-flow fluorometer and fluorescence is observed as a function of time. Fluorescence was observed to increase with time indicating that ClpA dissociated from the polypeptide chain. The question is; do the kinetic time courses yield information on translocation before ClpA dissociates? In principle, if ClpA is taking multiple steps before dissociating then the observed kinetic time courses should reflect the number of steps the enzyme takes before dissociation. Thus, if the length of the peptide is increased, the number of steps the enzyme takes before reaching the end should also increase. That is to say, if the time courses are sensitive to processive translocation, then the time courses should depend upon substrate length.

To test the substrate length dependence of the kinetic time courses, time courses were collected as a function of polypeptide substrate length ranging from 30 to 50 amino acids (81). Observed was a lag (constant fluorescence) followed by an increase in fluorescence. This lag was observed to increase in duration with increasing substrate length indicating that ClpA remained on the polypeptide for an increasing amount of time with increasing substrate length. This observation is interpreted to indicate that ClpA is taking more steps with each increase in substrate length. Therefore, the single-turnover fluorescence stopped-flow assay is sensitive to processive translocation.

To elucidate the elementary rate constants using transient state kinetics one needs to perturb the system. Variables like temperature, salt concentration and type, pH, etc. can

be used for this perturbation. For a molecular motor that couples ATP binding and hydrolysis to repeated rounds of translocation, the simplest perturbation is to vary the ATP concentration. The initial experiments are usually carried out at excess ATP so that it can be assumed that ATP binding is not rate-limiting. As the [ATP] is reduced, the observed rate constant will reflect ATP binding, or a step coupled to ATP binding. Importantly, because the motor-peptide complex is preassembled prior to rapidly mixing with ATP (Figure 3), the signal is insensitive to the changing population of ClpA hexamers throughout the ATP range assayed. The kinetic time courses were collected as a function of ATP from ~125 μM to 5 mM. As the [ATP] was reduced, the observed kinetic rate constant decreased. This is further evidence that the time courses are reporting on translocation since simple dissociation would not be predicted to be ATP concentration dependent.

The kinetic time courses were subjected to global non-linear-least-squares (NLLS) analysis (80, 84). For ClpA, the enzyme translocated with a repeating rate constant, k_t = (1.39 ± 0.06) s^{-1} and an overall rate of (19 ± 1) AA s^{-1} at saturating ATP with a kinetic step-size of (14 ± 2) AA $step^{-1}$. It is important to note that the kinetic step-size represents the average number of amino acids translocated between two rate-limiting steps and may or may not represent physical stepping. While similar strategies have been successfully used to examine helicase catalyzed DNA unwinding and single strand DNA translocation (85-88), this was the first step-size reported for a polypeptide translocase (80, 81).

The processivity is quantitatively defined as the rate constant for translocation divided by the summation of the rate constants for translocation and dissociation. For example, a translocating enzyme following the mechanism shown in Figure 4, where $E{\cdot}P$

represents the enzyme pre-bound to a peptide of length L, the enzyme can proceed forward with rate constant k_t or dissociate with rate constant k_d. $I_{(L-m)}$ represents the first intermediate that has been translocated by some distance m (step-size).

The processivity is the probability given by equation (10) (80, 84).

$$P = \frac{k_t}{k_t + k_d} \tag{10}$$

When $k_d = 0$, then $P = 1$ and every enzyme that binds will translocate to the end without dissociation. On the other hand, as k_d increases, P approaches zero, which would describe an enzyme with low processivity (an enzyme that has a higher propensity to dissociate than reach the end of the polypeptide chain). The processivity described as a probability, P, can be related to processivity expressed in terms of the average number of amino acids translocated per binding event, N, given by Equation (11) (for a complete derivation of Equation (11) see Appendix B of (84).

$$P = e^{-(m/N)} \tag{11}$$

It is tempting to assume that a hexameric ring motor that encircles the linear lattice on which it translocates would be highly processive. However, this is not always true. For example, the hexameric ring helicase, DnaB exhibits a processivity of P ~0.89 (89). The proposed model is that the ring opens and substrate can 'escape' thereby resulting in a dissociation event. However, this primary replicative helicase likely exhibits much higher processivity in the context of the fast moving replication fork, likely due to interactions with other proteins. With respect to ClpB and Hsp104, both enzymes have been proposed to be in 'rapid subunit exchange' (72). Thus, loss of a subunit in a hexameric ring could also result in a dissociation event. Moreover, like DnaB, partner proteins are likely to

influence the processivity. Regardless of the mechanism, there is a dearth of quantitative measurements of processivity for polypeptide translocases.

In the initial examination of ClpA catalyzed polypeptide translocation with synthetic peptides, a measureable dissociation rate constant, k_d, was not detected above 500 μM ATP. However, at 300 μM ATP and below, a measureable dissociation rate constant was observed, allowing for the calculation of processivity. The processivity was determined to be $P = (0.876 \pm 0.006)$ at low [ATP]. Using equation (11) a processivity of ~ 100 amino acids per binding event is predicted, which is 2-fold larger than the longest polypeptide used in this study. Thus, this is a preliminary estimate of the processivity at limiting [ATP] and methods allowing the examination of longer polypeptides are needed to rigorously test the processivity for this and related enzymes. Qualitatively, the findings support the idea that ClpA is highly processive, confirming that reported by Maurizi and coworkers (90).

Effect of ClpP on the translocation mechanism catalyzed by ClpA

With a method in hand that is sensitive to polypeptide translocation in the absence of proteolytic degradation the question could be asked; does ClpAP translocate using the same mechanism as ClpA alone? A qualitative assessment of stopped-flow time courses had been reported previously that concluded ClpAP translocated faster than ClpA alone but rate constants were not reported (91).

The single turnover stopped-flow method described above was employed to examine polypeptide translocation catalyzed by ClpAP. However, upon building a complex of polypeptide bound by ClpAP, a number of questions emerge. Hexameric ClpA

can bind to either apical surface of ClpP forming a 1:1 complex, or to both apical surfaces of ClpP forming a 2:1 complex, see Figure 5. Should the experimental design conditions examine 1:1 or 2:1 hexameric ClpA to tetradecameric ClpP? Similarly, if the 2:1 complex is examined, should both sides of the enzyme be bound with peptide?

Based on activity measurements, Maurizi and coworkers reported an affinity for ClpA hexamer binding to ClpP tetradecamer to be ~4 nM (54). However, the fact that ClpA resides in a distribution of oligomers was not taken into account. ClpA resides in a distribution of monomers, dimers, and tetramers in the absence of nucleotide (58, 60). However, even at concentration of nucleotide above 1 mM there remains a distribution of oligomeric states (61, 62). Thus, it cannot be assumed that all of the ClpA present in solution is in the hexameric state.

For a macromolecule with two binding sites, one can be certain to ligate only one of the binding sites if the two-site macromolecule is maintained in large excess over the ligand. Thus, whether 1:1, 2:1 or a mixture of the ClpAP complexes are present in solution, by maintaining the complex in excess over the polypeptide only one peptide can be bound to any given ClpAP complex in the ensemble.

To build a peptide pre-bound complex, 86 nM tetradecameric ClpP and 1 μM monomer of ClpA were used in the presence of 150 μM ATPγS. Note that, unlike ClpA, ClpP forms stable tetradecamers (54)(E. Duran unpublished data). However, the question is; how much hexameric ClpA is present at 1 μM monomer? To address this question, sedimentation velocity experiments measured the concentration of hexameric ClpA in the presence of 150 μM ATPγS at 1 μM total ClpA monomer concentration. Under these

conditions, the hexameric concentration was determined to be 130 nM. It is important to note that if the 1 μM total monomer concentration is divided by six, i.e. assume only hexamers are in solution, then one would predict 170 nM hexamers, an over estimate by 30 % of the hexameric ClpA population. Under these conditions, a mixture of 1:1 and 2:1 complexes is predicted. With that in mind, binding the complex to 20 nM peptide maintains ClpAP (whether 1:1 or 2:1 complex) in large excess over the peptide. Keeping the ClpAP complex in excess over the peptide concentration ensures that peptide is only bound to one ClpA hexamer in a given ClpAP molecule. That is to say, it would be thermodynamically unfavorable to have a doubly peptide ligated 2:1 ClpAP complex.

Subjecting ClpAP to the same analysis as performed on ClpA alone revealed that, indeed, ClpAP does translocate with a faster overall rate of ~35 AA s^{-1} (71). This is approximately 1.5 times faster than the ~20 AA s^{-1} observed for ClpA alone. The overall rate is the product of the step size and the elementary rate constant governing that step. One of the strengths of the transient state kinetic approaches used is that it is sensitive to these two additional parameters. Interestingly, the kinetic step size for ClpAP was observed to be ~5 AA $step^{-1}$ in stark contrast to the ~14 AA $step^{-1}$ measured for ClpA alone (81). Further, the rate constant governing translocation was found to be ~7 s^{-1}, which is ~5-fold faster than the ~1.4 s^{-1} measured for ClpA (71).

As stated above, the kinetic step-size does not necessarily represent physical movement. However, a recent single-molecule examination of ClpAP translocation reports steps of ~ 1 nm (92), which was reported to be consistent with the 5 AA $step^{-1}$ reported from the single turnover experiments described above (71). A single molecule experiment that would be sensitive to mechanical movement has not been performed on ClpA alone.

Such an experiment would either confirm or refute the measured ~14 AA step^{-1}. Additional testing is necessary to determine whether or not this kinetic step-size represents mechanical movement.

All in all, it is clear that ClpP exerts an allosteric influence on ClpA catalyzed polypeptide translocation. Thus, ClpA and ClpAP should be considered to be two different enzymes that translocate with two different mechanisms. Moreover, questions remain regarding the activities of the 2:1 and 1:1 complexes.

The Walker A and Walker B motifs that form the ATP binding pocket are separated by a loop that extends into the axial channel of ClpA (27). It has been proposed that the loop cycles up and down as the ATP binding site cycles through bound ATP to bound ADP + P_i and then release of ADP and P_i. This up and down motion is thought to drive translocation. Hinnerwisch and coworkers showed through crosslinking studies that polypeptide substrate crosslinked with the NBD2 loop in the central channel of ClpA (32). From these observations, Hinnerwisch and coworkers proposed that the NBD2 loop was responsible for mechanical pulling on the substrate polypeptide being translocated. They proposed a cycle of translocation to consist of ATP binding at NBD2 with the NBD2 loop in the up conformation, followed by ATP hydrolysis that drives movement of the NBD2 loop to the down conformation and concurrent movement of the polypeptide substrate that is bound to the NBD2 loop. Consistently, synchrotron footprinting data showed that the NBD2 loop proceeds through a nucleotide-dependent conformational change (93).

From examination of the ATP concentration dependence of the kinetic step-size and rate constant for ClpAP, the observed step immediately follows ATP binding (71).

Coupling this observation with the Hinnerwisch model, the step detected in the single-turnover experiments could be either ATP hydrolysis or a conformational change; a conformational change that may represent movement of the NBD2 loop. Since a single repeating step was detected in each cycle of translocation, loop movement may represent movement by ~5 amino acids.

If the measured kinetic step-size for ClpAP truly represents mechanical movement by ~5 amino acids then why does ClpA alone exhibit a different kinetic step-size of ~14 AA $step^{-1}$? A potential answer to this question lies in the dependence of the overall translocation rate on [ATP] for ClpA and ClpAP. The translocation rate constant for ClpA alone exhibited a sigmoidal dependence on ATP. The isotherm could not be described by a simple rectangular hyperbola. Rather, it required analysis using a Hill model with a hill coefficient of ~2.5. In contrast, the translocation rate constant for ClpAP did not exhibit a sigmoidal dependence. Since ClpA contains two ATP binding sites per monomer and the single-turnover kinetic time courses are sensitive only to bound hexamer, the observation of a sigmoidal dependence suggests that there is cooperativity between multiple ATP binding sites that are involved in polypeptide translocation. On the other hand, since ClpAP did not exhibit any cooperativity, this indicates that the presence of ClpP relieves the cooperative interactions.

With these observations in mind, Figure 6 illustrates a working model for both ClpA and ClpAP polypeptide translocation, incorporating known structural information and various biochemical/biophysical studies. Figure 6A illustrates ClpA, in the absence of ClpP, with both the NBD1 and NBD2 loops in the up conformation and ATP bound to both domains. The polypeptide substrate is shown in black and is making contact with both the

NBD1 and NBD2 loops. Crosslinking studies have shown that contacts between polypeptide substrate and ClpA were only observed with the NBD2 loop, but many various single site mutations throughout the NBD1 loop abolished translocation activity (32). Moreover, recent work indicates that both ATPase sites are involved in translocation catalyzed by ClpA in the absence of ClpP (81). These two observations implicate the NBD1 loop in translocation. The next step would be for NBD1 to hydrolyze ATP and cause the NBD1 loop to move down and translocate (push) the substrate by up to 14 amino acids creating a polypeptide loop inside the axial channel of ClpA. The loop in the substrate can be accommodated in ClpA since it has been shown that ClpA forms a cavity between the NBD1 and NBD2 loops (94, 95). NBD1 would contain ADP and P_i in the ATP binding site and therefore the NBD1 loop would have a reduced affinity for the polypeptide, which would allow for rebinding by another NBD1 loop loaded with ATP in a neighboring subunit in the hexamer (61, 96). The NBD2 loop would cycle through multiple rounds of ATP hydrolysis coupled to translocation of the substrate by 2 – 5 amino acids per cycle with a rate constant of ~4 s^{-1}. This will occur several times thereby shortening the loop inside the cavity of ClpA before NBD1 translocates another ~14 amino acids of the polypeptide into the cavity with a rate constant of 1.4 s^{-1}.

Figure 6B illustrates the working model for how ClpA translocates when associated with ClpP. Since the ATP concentration dependence of the rate of ClpAP catalyzed polypeptide translocation suggests reduced cooperativity between ATP binding sites, it is hypothesized that NBD2 drives translocation in the ClpAP complex. Repeating cycles of ATP binding and hydrolysis could occur at NBD1, but they do not limit the observed translocation. Therefore, this model predicts repeating cycles of ATP binding and

hydrolysis at NBD2 would lead to translocation of the substrate by distances of 2 – 5 aa step^{-1}.

The working model predicts that in the absence of ClpP, NBD1 should hydrolyze ATP with a rate constant of (1.39 ± 0.06) s^{-1} and NBD2 should hydrolyze ATP with a rate constant of (7.9 ± 0.2) s^{-1} in the presence of polypeptide substrate. Kress *et al.* examined the steady state rate of ATP hydrolysis catalyzed by ClpA both in the presence and absence of ClpP (69). Further, they made two variants of ClpA that are deficient in ATP hydrolysis at either NBD1 or NBD2, which allow for the examination of ATP hydrolysis at each domain in the absence of hydrolysis at the other domain, and in the presence or absence of ClpP and SsrA substrate. Interestingly, in the absence of ClpP and the presence of GFP-SsrA, NBD1 hydrolyzes ATP with a rate constant of (0.8 ± 0.2) s^{-1}, which is comparable to the rate constant determined for translocation of (1.39 ± 0.06) s^{-1}determined using the single-turnover stopped flow experiments. Similarly, in the presence of ClpP and GFP-SsrA, NBD2 hydrolyzes ATP with a rate constant of (6.3 ± 0.5) s^{-1}, which is similar to the estimate of (7.9 ± 0.2) s^{-1} (71).

MECHANISM OF TRANSLOCATION BY CLPB/HSP104

As stated above, ClpB/Hsp104 shares many structural characteristics with ClpA (see Figure 1) and therefore has been hypothesized to share a similar translocation mechanism. One important difference is the absence of an IGF/L loop in ClpB/Hsp104, necessary in ClpA for binding the protease ClpP. This structural difference intimates an important functional difference; ClpB/Hsp104 does not partner with any known protease (97).

A disaggregase such as ClpB/Hsp104 does not covalently modify its protein substrate. Disaggregation has been measured by monitoring changes in turbidity, solubility, and various staining techniques *in vitro,* thermotolerance development studies *in vivo,* and enzyme reactivation *in vivo* or *in vitro* (41, 74, 98-107). These macroscopic observations, while informative, do not report on the molecular level events involved in the mechanism. How can the molecular events in the translocation or disaggregation mechanism be studied in the absence of a covalent modification to the protein substrate? Early investigations of the ClpB/Hsp104 disaggregation mechanism addressed this challenge by building upon the structural similarities between ClpB/Hsp104 and *E. coli* ClpA. As discussed above, ClpA processively translocates protein substrates through its axial channel and into the protease, ClpP. The similarities in sequence, tertiary structure, and quaternary structure lead the Bukau group to engineer the IGF/L loop onto the C terminal surface of ClpB and Hsp104. This loop allows a non-native interaction with ClpP, resulting in degradation of the substrate, a measurable covalent modification (21, 22). The rationale was that if they could 'force' ClpB (Hsp104) to interact with ClpP and they observed proteolytic degradation, then this must mean that ClpB, like ClpA was translocating a substrate through the axial channel and into ClpP for proteolytic degradation.

In these studies, the Bukau group showed that the non-native BAP (ClpB-ClpA-P loop) -ClpP or HAP (Hsp104-ClpA-P loop) -ClpP complex was indeed able to degrade substrate proteins. This observation was interpreted as evidence that BAP and HAP, and therefore ClpB and Hsp104, processively translocate entire proteins through the axial channel and into ClpP, just as is done by the processive translocase ClpA (21, 22). Notably,

additional studies of BAP-ClpP in which only portions of a substrate were unfolded lead the Bukau group to conclude, “partial threading of the unfolded substrate moiety through the central channel of ClpB is sufficient for efficient protein disaggregation in a physiologically relevant context” and that “partially threaded polypeptide chains are released from ClpB to be refolded” (108). Since these publications, however, many researchers in the field have often interpreted or summarized the Bukau results with less nuance, carrying forward only the “complete threading” model of polypeptide translocation.

The current prevailing hypothesis in the field is that the BAP-ClpP and HAP-ClpP findings, together with the structural similarities to ClpA, are evidence of complete threading or processive translocation by ClpB and Hsp104. The dominant mechanistic model is the translocation of an entire full-length protein pulled out of an aggregate through the axial channel of the disaggregating motor. The exclusive portrayal of this complete threading/processive translocation mechanism for these disaggregases has been schematized throughout the literature (39, 109). Other primary research has also been interpreted as consistent with the complete threading model based largely on the BAP/HAP–ClpP results (104, 110). It should be noted, however, that some researchers in the field do point out the possibility of both complete and partial threading mechanisms (111).

Another important challenge to the findings using BAP and HAP with ClpP is that recent work has shown that BAP-ClpP degrades α-casein in both the absence and the presence of ATP (23). Thus, the degradation observed in this experimental design does not report strictly on the ATP-dependent translocation mechanism. Nakazaki and Watanabe’s

findings from their study of various mutations of *T*BAP-ClpP were interpreted as passive threading, independent of ATP hydrolysis (110). However these results could alternatively be understood to show that the *T*BAP-ClpP construct does not report exclusively on ATP-dependent translocation (threading) since they found "no correlation between ATPase activities and degradation rates" (110).

A complementary approach to the BAP-ClpP degradation experiments, one in which there is no forced interaction with a protease, is needed. The stopped-flow fluorometer experimental design, described above (Figure 3), developed for the study of ClpA in the absence of ClpP is one such complementary approach (81). Using this design, Li *et al*. demonstrated that ClpB is a non-processive translocase, taking only one or two kinetic steps before releasing the polypeptide substrate (23). This finding is at odds with the prevailing model of complete threading, by which one polypeptide chain is extracted from an aggregate. However, the Li *et al* conclusion is in good agreement with previous results of observed partial threading (108). Additional studies are needed to expand this work into Hsp104.

Though Hsp104 and ClpB are both structurally and functionally similar, important differences have been observed. For example, both Hsp104 and ClpB can resolve disordered aggregates, however only Hsp104, not ClpB, can also resolve more structured amyloid aggregates (42). Hsp104 also has an additional function in prion curing not observed for ClpB (103). What mechanistic differences give rise to these observations?

One possible contribution to the differences between the disaggregases is the differing roles of the two NBDs. The interplay between the NBDs within a hexamer is

complex and cooperative. Still, some distinctions between NBD1 and NBD2 have been drawn. Notably, nucleotide binding at NBD1 is necessary for stabilization of ClpB hexamers (74, 100, 112, 113). This role of NBD1 in oligomerization is conserved between ClpB and ClpA. Surprisingly, in Hsp104, nucleotide binding in NBD2 is required for stabilization of hexamers (114, 115).

In both ClpB and Hsp104, like in ClpA, the tyrosines in the pore loops of both NBDs are important for substrate processing (21, 22, 31, 50, 52, 116). As the ATP hydrolysis cycle is carried out in either NBD, the pore loop is thought to move through space due to conformational changes induced by the nucleotide ligation state. The relatively large, planar surface of the tyrosine residue is thought to interact with the polypeptide substrate, pushing or pulling the polypeptide through the central channel. It's possible that differences in nucleotide binding/hydrolysis induced pore loop conformational changes account for the functional differences that exist between ClpB and Hsp104 catalyzed protein disaggregation. Experimental designs that report on the molecular level events involved in ClpB/Hsp104 polypeptide substrate processing, in particular those sensitive to the coordination between pore loop movement and nucleotide ligation state during disaggregation, will be key in testing this hypothesis.

Effect of DnaK/Hsp70 on ClpB/Hsp104 mechanism

ClpB and Hsp104 were initially observed to disaggregate clients only in the presence of co-chaperones. These disaggregating motors are far more potent in collaboration with co-chaperones, although conditions have since been found in which ClpB and Hsp104 have innate disaggregation abilities. The co-chaperone system for *E. coli*

ClpB is made up of DnaK, DnaJ, and the nucleotide exchange factor GrpE (termed the KJE system). Yeast Hsp104 collaborates with the co-chaperones Hsp70 (analogous to DnaK) and Hsp40 (analogous to DnaJ). Like ClpB/Hsp104, DnaK/Hsp70 is an ATPase and a disaggregase that can function independently of co-chaperones. The full systems, ClpB/KJE and Hsp104/70/40, have ATPase and disaggregase activity greater than the sum of the components' activities. There are three proposed possibilities that could explain this enhanced activity: (1) DnaK modifies the aggregate making a better binding site for ClpB, (2) DnaK accepts substrate from ClpB after the substrate has been completely translocated, or (3) the ClpB-DnaK complex has greatly amplified disaggregation activity relative to ClpB alone, possibly through a fundamentally different mechanism.

Early attempts to identify which component of the system acted upon an aggregate or client first resulted in divergent findings. The Liberek group identified DnaK as the first actor. They found that DnaK, with DnaJ and ATP, remodeled aggregates to facilitate ClpB-catalyzed disaggregation. Neither a transient tertiary complex with ClpB or additional roles for DnaK downstream of ClpB's action were ruled out (117, 118). On the other hand, early work from the Bukau group concluded that ClpB acted first. Specifically, ClpB was observed to expose a substrate's hydrophobic regions, which could then be recognized by the KJE system (107). The development of a ClpB trap mutant (double Walker B variant, able to bind but not hydrolyze ATP) also revealed that $ClpB_{trap}$ outcompeted DnaK for binding to a model substrate and inhibited DnaK activity (101).

Over time, the idea of a ClpB-DnaK (Hsp104-Hsp70) complex has come into favor. One compelling observation in support of this finding is that the activity of the co-chaperones is species specific. ClpB works with DnaK but not Hsp70. Hsp104 works with

Hsp70 but not DnaK. This suggests a direct interaction between the chaperones. Furthermore, both the Wickner and Tsai groups engineered sets of chimeras in which domains from ClpB were replaced by the analogous domain from Hsp104 and vice versa. Both groups found that the M domain dictates which species formed a productive cochaperone partnership. For example, Hsp104 with the M domain from ClpB partnered effectively with the KJE system, not the Hsp70 system. This finding is consistent with the identification of the M domain as the binding site for DnaK (38, 39, 44).

Binding affinities of 17 µM and 25 µM have been reported for *T. thermophilus* ClpB and DnaK (38, 102). For *E. coli* ClpB and DnaK, the K_d has been estimated in the range of 7-30 µM (119). Notably, while the ClpB-DnaK complex has been observed by co-elution assays (102, 120) and NMR (38), the ternary complex of ClpB-DnaK-client has not been observed. Furthermore, despite the K_d measurements and estimates in the range of ~20 µM, biochemical assays are often carried out with nanomolar to low micromolar concentrations of DnaK, conditions in which a significant population of the ClpB-DnaK complex is not expected (21, 37, 38, 42, 44, 108, 111). Nevertheless, in these cases observations are attributed to the interplay between ClpB and DnaK. Though the existence of a DnaK-ClpB or Hsp70-Hsp104 complex has become widely accepted, the role of co-chaperones upstream and/or downstream of that complex remains under investigation. The convergence of evidence suggests that DnaK acts on the aggregate first, possibly targeting the client to ClpB, and then DnaK binds ClpB unleashing the disaggregating power of ClpB (21, 37, 39, 41). DnaK may also have additional roles in the proper refolding of the client after release from ClpB.

CONCLUSIONS

The reviewed studies reveal important considerations for design and implementation of the experiments needed to address outstanding questions about ClpA and ClpB/Hsp104 catalyzed protein translocation, degradation, and disaggregation mechanisms, respectively. One major aspect of assay design is the ability to predict the population of degradation/disaggregation active complex present under the chosen experimental conditions. As work on ClpB has revealed, these proteins persist in a distribution of oligomers even at high nucleotide concentrations (Figure 2). Therefore, dividing the monomeric protein concentration by six will yield overestimates for the hexameric population present and available to interact with partner proteins and substrates in solution. Instead, quantification of the active hexamer population in a given assay will require a thermodynamically rigorous characterization of the energetics governing nucleotide-linked self-assembly. Although this work has been done for ClpB, the mechanisms of ClpA and Hsp104 ligand linked self-assembly remain to be examined.

A related consideration in assay design is the effect of mutations on AAA+ motor self-assembly. Because the propensity of a protein to oligomerize is in part driven by its primary sequence, mutations of the sequence will have effects on its self-assembly. If unaccounted for, assay readout changes resulting from up- or downregulation of the hexamer population as a result of mutations, could be misinterpreted as up- or downregulation of “activity” in ATPase, reactivation, or other assays. Thus, when designing experiments for AAA+ motors and their corresponding variants, it is important to know whether the signal being monitored reports on events that could be controlled by

changes in the assembly state. Interpretations of those results should be tempered by possible contributions from variability in the assembly state.

Single turnover translocation experiments have been designed to yield information about the molecular level events governing AAA+ motor activity without rigorous quantification of the self-assembly mechanism (23, 81). However, this was possible, in part, because only hexamers are bound to the polypeptide substrate. If smaller oligomers contributed to the translocation signal then measures would have to be taken to account for this. For example, as soon as ClpP is introduced to ClpA then one has to start asking how the distribution of 1:1 and 2:1 hexameric ClpA to tetradecameric ClpP influences the signal. Similar techniques are being adopted to investigate the molecular level events governing the mechanism of ClpB/Hsp104 catalyzed disaggregation in the absence and presence of partner co-chaperones. As work on ClpA and ClpAP revealed, ClpP induces a major change in the mechanism of ClpA catalyzed polypeptide translocation. It's reasonable, then, to expect cochaperones like DnaK/Hsp70 to similarly affect the disaggregation mechanism of ClpB/Hsp104. Implementation of these transient state kinetic techniques will prove powerful in the deconvolution of cochaperone contributions to the disaggregation activities of ClpB/Hsp104 and functional differences between ClpB and Hsp104.

By definition, motor proteins use an energy source to perform mechanical work. ClpA and ClpB/Hsp104 use the energy from ATP binding/hydrolysis to perform this mechanical work. For any translocase there is interest in how far the translocase moves on its lattice, how much energy is required to make this movement, and how much force is exerted. For ClpA we reported the first kinetic step-size for any AAA+ protein translocase

to be ~14 amino acids per step (81). Similarly, we showed that ClpAP translocated with a reduced kinetic step-size of ~5 amino acids per step (71). Consistently, a single molecule optical tweezer measurement reported a step-size of ~5 amino acids per step for ClpAP (92). Similarly, single molecule optical tweezer experiments showed that ClpXP translocated in 5 – 8 amino acid steps (121, 122). In many cases, single-molecule and single turnover kinetics experiments can get around the limitations on the interpretation imposed by macromolecular assembly. Thus, going forward, the combination of single-molecule and transient state kinetic experiments are going to be essential for addressing detailed mechanistic questions on AAA+ motors.

REFERENCES

1. Rogers, S., R. Wells, and M. Rechsteiner. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science 234:364-368.
2. Bachmair, A., D. Finley, and A. Varshavsky. 1986. In vivo half-life of a protein is a function of its amino-terminal residue. Science 234:179-186.
3. Dice, J. F. 1987. Molecular determinants of protein half-lives in eukaryotic cells. FASEB J 1:349-357.
4. Varshavsky, A. 1996. The N-end rule: functions, mysteries, uses. Proc Natl Acad Sci U S A 93:12142-12149.
5. Brotz-Oesterhelt, H., D. Beyer, H. P. Kroll, R. Endermann, C. Ladel, W. Schroeder, B. Hinzen, S. Raddatz, H. Paulsen, K. Henninger, J. E. Bandow, H. G. Sahl, and H. Labischinski. 2005. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. Nat Med 11:1082-1087.
6. Hinzen, B., S. Raddatz, H. Paulsen, T. Lampe, A. Schumacher, D. Habich, V. Hellwig, J. Benet-Buchholz, R. Endermann, H. Labischinski, and H. Brotz-Oesterhelt. 2006. Medicinal chemistry optimization of acyldepsipeptides of the enopeptin class antibiotics. ChemMedChem 1:689-693.
7. Sauer, R. T., D. N. Bolon, B. M. Burton, R. E. Burton, J. M. Flynn, R. A. Grant, G. L. Hersch, S. A. Joshi, J. A. Kenniston, I. Levchenko, S. B. Neher, E. S. Oakes, S. M. Siddiqui, D. A. Wah, and T. A. Baker. 2004. Sculpting the proteome with AAA(+) proteases and disassembly machines. Cell 119:9-18.
8. Sauer, R. T., and T. A. Baker. 2011. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annual review of biochemistry 80:587-612.
9. Olivares, A. O., T. A. Baker, and R. T. Sauer. 2016. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat Rev Microbiol 14:33-44.
10. Wickner, S., S. Gottesman, D. Skowyra, J. Hoskins, K. McKenney, and M. R. Maurizi. 1994. A molecular chaperone, ClpA, functions like DnaK and DnaJ. Proc Natl Acad Sci U S A 91:12218-12222.
11. Kruklitis, R., D. J. Welty, and H. Nakai. 1996. ClpX protein of Escherichia coli activates bacteriophage Mu transposase in the strand transfer complex for initiation of Mu DNA synthesis. The EMBO journal 15:935-944.
12. Levchenko, I., L. Luo, and T. A. Baker. 1995. Disassembly of the Mu transposase tetramer by the ClpX chaperone. Genes & development 9:2399-2408.
13. Kardon, J. R., Y. Y. Yien, N. C. Huston, D. S. Branco, G. J. Hildick-Smith, K. Y. Rhee, B. H. Paw, and T. A. Baker. 2015. Mitochondrial ClpX Activates a Key Enzyme for Heme Biosynthesis and Erythropoiesis. Cell 161:858-867.
14. McNally, F. J., and R. D. Vale. 1993. Identification of katanin, an ATPase that severs and disassembles stable microtubules. Cell 75:419-429.

15. Hazan, J., N. Fonknechten, D. Mavel, C. Paternotte, D. Samson, F. Artiguenave, C. S. Davoine, C. Cruaud, A. Durr, P. Wincker, P. Brottier, L. Cattolico, V. Barbe, J. M. Burgunder, J. F. Prud'homme, A. Brice, B. Fontaine, B. Heilig, and J. Weissenbach. 1999. Spastin, a new AAA protein, is altered in the most frequent form of autosomal dominant spastic paraplegia. Nat Genet 23:296-303.

16. Roll-Mecak, A., and F. J. McNally. 2010. Microtubule-severing enzymes. Curr Opin Cell Biol 22:96-103.

17. Dalal, S., M. F. Rosser, D. M. Cyr, and P. I. Hanson. 2004. Distinct roles for the AAA ATPases NSF and p97 in the secretory pathway. Molecular biology of the cell 15:637-648.

18. Block, M. R., B. S. Glick, C. A. Wilcox, F. T. Wieland, and J. E. Rothman. 1988. Purification of an N-ethylmaleimide-sensitive protein catalyzing vesicular transport. Proc Natl Acad Sci U S A 85:7852-7856.

19. Fleming, K. G., T. M. Hohl, R. C. Yu, S. A. Muller, B. Wolpensinger, A. Engel, H. Engelhardt, A. T. Brunger, T. H. Sollner, and P. I. Hanson. 1998. A revised model for the oligomeric state of the N-ethylmaleimide-sensitive fusion protein, NSF. J Biol Chem 273:15675-15681.

20. Zhao, C., E. C. Smith, and S. W. Whiteheart. 2012. Requirements for the catalytic cycle of the N-ethylmaleimide-Sensitive Factor (NSF). Biochimica et biophysica acta 1823:159-171.

21. Weibezahn, J., P. Tessarz, C. Schlieker, R. Zahn, Z. Maglica, S. Lee, H. Zentgraf, E. U. Weber-Ban, D. A. Dougan, F. T. Tsai, A. Mogk, and B. Bukau. 2004. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. Cell 119:653-665.

22. Tessarz, P., A. Mogk, and B. Bukau. 2008. Substrate threading through the central pore of the Hsp104 chaperone as a common mechanism for protein disaggregation and prion propagation. Molecular microbiology 68:87-97.

23. Li, T., C. L. Weaver, J. Lin, E. C. Duran, J. M. Miller, and A. L. Lucius. 2015. Escherichia coli ClpB is a non-processive polypeptide translocase. The Biochemical journal 470:39-52.

24. Weber-Ban, E. U., B. G. Reid, A. D. Miranker, and A. L. Horwich. 1999. Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. Nature 401:90-93.

25. Kim, Y. I., R. E. Burton, B. M. Burton, R. T. Sauer, and T. A. Baker. 2000. Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. Mol Cell 5:639-648.

26. Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. The EMBO journal 1:945-951.

27. Guo, F., M. R. Maurizi, L. Esser, and D. Xia. 2002. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. J Biol Chem 277:46743-46752.

28. Lee, S., M. E. Sowa, Y. H. Watanabe, P. B. Sigler, W. Chiu, M. Yoshida, and F. T. Tsai. 2003. The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell 115:229-240.

29. Heuck, A., S. Schitter-Sollner, M. J. Suskiewicz, R. Kurzbauer, J. Kley, A. Schleiffer, P. Rombaut, F. Herzog, and T. Clausen. 2016. Structural basis for the disaggregase activity and regulation of Hsp104. eLife 5.

30. Yamada-Inagawa, T., T. Okuno, K. Karata, K. Yamanaka, and T. Ogura. 2003. Conserved pore residues in the AAA protease FtsH are important for proteolysis and its coupling to ATP hydrolysis. J Biol Chem 278:50182-50187.

31. Schlieker, C., J. Weibezahn, H. Patzelt, P. Tessarz, C. Strub, K. Zeth, A. Erbse, J. Schneider-Mergener, J. W. Chin, P. G. Schultz, B. Bukau, and A. Mogk. 2004. Substrate recognition by the AAA+ chaperone ClpB. Nature structural & molecular biology 11:607-615.

32. Hinnerwisch, J., W. A. Fenton, K. J. Furtak, G. W. Farr, and A. L. Horwich. 2005. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell 121:1029-1041.

33. Martin, A., T. A. Baker, and R. T. Sauer. 2008. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. Nature structural & molecular biology 15:1147-1151.

34. Biter, A. B., S. Lee, N. Sung, and F. T. Tsai. 2012. Structural basis for intersubunit signaling in a protein disaggregating machine. Proc Natl Acad Sci U S A 109:12515-12520.

35. Zeymer, C., T. R. Barends, N. D. Werbeck, I. Schlichting, and J. Reinstein. 2014. Elements in nucleotide sensing and hydrolysis of the AAA+ disaggregation machine ClpB: a structure-based mechanistic dissection of a molecular motor. Acta crystallographica. Section D, Biological crystallography 70:582-595.

36. Oguchi, Y., E. Kummer, F. Seyffer, M. Berynskyy, B. Anstett, R. Zahn, R. C. Wade, A. Mogk, and B. Bukau. 2012. A tightly regulated molecular toggle controls AAA+ disaggregase. Nature structural & molecular biology 19:1338-1346.

37. Seyffer, F., E. Kummer, Y. Oguchi, J. Winkler, M. Kumar, R. Zahn, V. Sourjik, B. Bukau, and A. Mogk. 2012. Hsp70 proteins bind Hsp100 regulatory M domains to activate AAA+ disaggregase at aggregate surfaces. Nature structural & molecular biology 19:1347-1355.

38. Rosenzweig, R., S. Moradi, A. Zarrine-Afsar, J. R. Glover, and L. E. Kay. 2013. Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. Science 339:1080-1083.

39. Miot, M., M. Reidy, S. M. Doyle, J. R. Hoskins, D. M. Johnston, O. Genest, M. C. Vitery, D. C. Masison, and S. Wickner. 2011. Species-specific collaboration of heat shock proteins (Hsp) 70 and 100 in thermotolerance and protein disaggregation. Proc Natl Acad Sci U S A 108:6915-6920.

40. Lee, J., J. H. Kim, A. B. Biter, B. Sielaff, S. Lee, and F. T. Tsai. 2013. Heat shock protein (Hsp) 70 is an activator of the Hsp104 motor. Proc Natl Acad Sci U S A 110:8513-8518.

41. Sielaff, B., and F. T. Tsai. 2010. The M-domain controls Hsp104 protein remodeling activity in an Hsp70/Hsp40-dependent manner. J Mol Biol 402:30-37.

42. DeSantis, M. E., E. H. Leung, E. A. Sweeny, M. E. Jackrel, M. Cushman-Nick, A. Neuhaus-Follini, S. Vashist, M. A. Sochor, M. N. Knight, and J. Shorter. 2012. Operational plasticity enables hsp104 to disaggregate diverse amyloid and nonamyloid clients. Cell 151:778-793.

43. Desantis, M. E., E. A. Sweeny, D. Snead, E. H. Leung, M. S. Go, K. Gupta, P. Wendler, and J. Shorter. 2014. Conserved distal loop residues in the Hsp104 and ClpB middle domain contact nucleotide-binding domain 2 and enable Hsp70-dependent protein disaggregation. J Biol Chem 289:848-867.

44. Doyle, S. M., S. Shastry, A. N. Kravats, Y. H. Shih, M. Miot, J. R. Hoskins, G. Stan, and S. Wickner. 2015. Interplay between E. coli DnaK, ClpB and GrpE during protein disaggregation. J Mol Biol 427:312-327.

45. Kim, Y. I., I. Levchenko, K. Fraczkowska, R. V. Woodruff, R. T. Sauer, and T. A. Baker. 2001. Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. Nature structural biology 8:230-233.

46. Carroni, M., E. Kummer, Y. Oguchi, P. Wendler, D. K. Clare, I. Sinning, J. Kopp, A. Mogk, B. Bukau, and H. R. Saibil. 2014. Head-to-tail interactions of the coiled-coil domains regulate ClpB activity and cooperation with Hsp70 in protein disaggregation. eLife 3:e02481.

47. Wendler, P., J. Shorter, C. Plisson, A. G. Cashikar, S. Lindquist, and H. R. Saibil. 2007. Atypical AAA+ subunit packing creates an expanded cavity for disaggregation by the protein-remodeling factor Hsp104. Cell 131:1366-1377.

48. Wendler, P., J. Shorter, D. Snead, C. Plisson, D. K. Clare, S. Lindquist, and H. R. Saibil. 2009. Motor mechanism for protein threading through Hsp104. Mol Cell 34:81-92.

49. Lee, S., B. Sielaff, J. Lee, and F. T. Tsai. 2010. CryoEM structure of Hsp104 and its mechanistic implication for protein disaggregation. Proc Natl Acad Sci U S A 107:8135-8140.

50. Yokom, A. L., S. N. Gates, M. E. Jackrel, K. L. Mack, M. Su, J. Shorter, and D. R. Southworth. 2016. Spiral architecture of the Hsp104 disaggregase reveals the basis for polypeptide translocation. Nature structural & molecular biology.

51. Effantin, G., T. Ishikawa, G. M. De Donatis, M. R. Maurizi, and A. C. Steven. 2010. Local and global mobility in the ClpA AAA+ chaperone detected by cryo-electron microscopy: functional connotations. Structure 18:553-562.

52. Gates, S. N., A. L. Yokom, J. Lin, M. E. Jackrel, A. N. Rizo, N. M. Kendsersky, C. E. Buell, E. A. Sweeny, K. L. Mack, E. Chuang, M. P. Torrente, M. Su, J. Shorter, and D. R. Southworth. 2017. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. Science.

53. Ripstein, Z. A., R. Huang, R. Augustyniak, L. E. Kay, and J. L. Rubinstein. 2017. Structure of a AAA+ unfoldase in the process of unfolding substrate. eLife 6.

54. Maurizi, M. R., S. K. Singh, M. W. Thompson, M. Kessel, and A. Ginsburg. 1998. Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. Biochemistry 37:7778-7786.

55. Zolkiewski, M., M. Kessel, A. Ginsburg, and M. R. Maurizi. 1999. Nucleotide-dependent oligomerization of ClpB from Escherichia coli. Protein Sci 8:1899-1903.

56. del Castillo, U., C. Alfonso, S. P. Acebron, A. Martos, F. Moro, G. Rivas, and A. Muga. 2011. A quantitative analysis of the effect of nucleotides and the M domain on the association equilibrium of ClpB. Biochemistry 50:1991-2003.

57. Akoev, V., E. P. Gogol, M. E. Barnett, and M. Zolkiewski. 2004. Nucleotide-induced switch in oligomerization of the AAA+ ATPase ClpB. Protein Sci 13:567-574.

58. Veronese, P. K., R. P. Stafford, and A. L. Lucius. 2009. The Escherichia coli ClpA Molecular Chaperone Self-Assembles into Tetramers. Biochemistry 48:9221-9233.

59. Kress, W., H. Mutschler, and E. Weber-Ban. 2007. Assembly Pathway of an AAA+ Protein: Tracking ClpA and ClpAP Complex Formation in Real Time. Biochemistry 46:6183-6193.

60. Veronese, P. K., and A. L. Lucius. 2010. Effect of Temperature on the Self-Assembly of the Escherichia coli ClpA Molecular Chaperone. Biochemistry 49:9820-9829.

61. Veronese, P. K., B. Rajendar, and A. L. Lucius. 2011. Activity of Escherichia coli ClpA Bound by Nucleoside Di- and Triphosphates. Journal of molecular biology 409:333-347.

62. Li, T., and A. L. Lucius. 2013. Examination of Polypeptide Substrate Specificity for E. coli ClpA. Biochemistry 52:4941-4954.

63. Lin, J., and A. L. Lucius. 2015. Examination of the dynamic assembly equilibrium for E. coli ClpB. Proteins 83:2008-2024.

64. Lin, J., and A. L. Lucius. 2016. Examination of ClpB Quaternary Structure and Linkage to Nucleotide Binding. Biochemistry 55:1758-1771.

65. Weaver, C. L., E. C. Duran, K. L. Mack, J. Lin, M. E. Jackrel, E. A. Sweeny, J. Shorter, and A. L. Lucius. 2017. Avidity for Polypeptide Binding by Nucleotide-Bound Hsp104 Structures. Biochemistry 56:2071-2075.

66. Li, T., J. Lin, and A. L. Lucius. 2015. Examination of polypeptide substrate specificity for Escherichia coli ClpB. Proteins 83:117-134.

67. Lin, J., and A. L. Lucius. 2015. Analysis of Linked Equilibria. Methods Enzymol 562:161-186.

68. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112:5377-5382.

69. Kress, W., H. Mutschler, and E. Weber-Ban. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. J Biol Chem 284:31441-31452.

70. Miller, J. M., and A. L. Lucius. 2014. ATP-gamma-S Competes with ATP for Binding at Domain 1 but not Domain 2 during ClpA Catalyzed Polypeptide Translocation. Biophysical chemistry 185:58-69.

71. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425:2795-2812.

72. Werbeck, N. D., S. Schlee, and J. Reinstein. 2008. Coupling and dynamics of subunits in the hexameric AAA+ chaperone ClpB. J Mol Biol 378:178-190.

73. Hoskins, J. R., S. M. Doyle, and S. Wickner. 2009. Coupling ATP utilization to protein remodeling by ClpB, a hexameric AAA+ protein. Proc Natl Acad Sci U S A 106:22233-22238.

74. del Castillo, U., J. A. Fernandez-Higuero, S. Perez-Acebron, F. Moro, and A. Muga. 2010. Nucleotide utilization requirements that render ClpB active as a chaperone. FEBS letters 584:929-934.

75. Yamasaki, T., Y. Oohata, T. Nakamura, and Y. H. Watanabe. 2015. Analysis of the Cooperative ATPase Cycle of the AAA+ Chaperone ClpB from Thermus thermophilus by Using Ordered Heterohexamers with an Alternating Subunit Arrangement. J Biol Chem 290:9789-9800.

76. Martin, A., T. A. Baker, and R. T. Sauer. 2005. Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. Nature 437:1115-1120.

77. Cordova, Juan C., Adrian O. Olivares, Y. Shin, Benjamin M. Stinson, S. Calmat, Karl R. Schmitz, M.-E. Aubin-Tam, Tania A. Baker, Matthew J. Lang, and Robert T. Sauer. 2014. Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell 158:647-658.

78. Levchenko, I., C. K. Smith, N. P. Walsh, R. T. Sauer, and T. A. Baker. 1997. PDZ-like domains mediate binding specificity in the Clp/Hsp100 family of chaperones and protease regulatory subunits. Cell 91:939-947.

79. Reid, B. G., W. A. Fenton, A. L. Horwich, and E. U. Weber-Ban. 2001. ClpA mediates directional translocation of substrate proteins into the ClpP protease. Proc Natl Acad Sci U S A 98:3768-3772.

80. Lucius, A. L., J. M. Miller, and B. Rajendar. 2011. Application of the Sequential n-Step Kinetic Mechanism to Polypeptide Translocases. Methods Enzymol 488:239-264.

81. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399:665-679.

82. Kenniston, J. A., T. A. Baker, J. M. Fernandez, and R. T. Sauer. 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. Cell 114:511-520.

83. Kenniston, J. A., T. A. Baker, and R. T. Sauer. 2005. Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing. Proc Natl Acad Sci U S A 102:1390-1395.

84. Lucius, A. L., N. K. Maluf, C. J. Fischer, and T. M. Lohman. 2003. General methods for analysis of sequential "n-step" kinetic mechanisms: application to single turnover kinetics of helicase-catalyzed DNA unwinding. Biophys J 85:2224-2239.

85. Lucius, A. L., and T. M. Lohman. 2004. Effects of temperature and ATP on the kinetic mechanism and kinetic step-size for E.coli RecBCD helicase-catalyzed DNA unwinding. J Mol Biol 339:751-771.

86. Lucius, A. L., C. Jason Wong, and T. M. Lohman. 2004. Fluorescence stopped-flow studies of single turnover kinetics of E.coli RecBCD helicase-catalyzed DNA unwinding. J Mol Biol 339:731-750.

87. Fischer, C. J., and T. M. Lohman. 2004. ATP-dependent translocation of proteins along single-stranded DNA: models and methods of analysis of pre-steady state kinetics. J Mol Biol 344:1265-1286.

88. Fischer, C. J., N. K. Maluf, and T. M. Lohman. 2004. Mechanism of ATP-dependent translocation of E.coli UvrD monomers along single-stranded DNA. J Mol Biol 344:1287-1309.

89. Galletto, R., M. J. Jezewska, and W. Bujalowski. 2004. Unzipping mechanism of the double-stranded DNA unwinding by a hexameric helicase: quantitative analysis of the rate of the dsDNA unwinding, processivity and kinetic step-size of the Escherichia coli DnaB helicase using rapid quench-flow method. J Mol Biol 343:83-99.

90. Thompson, M. W., S. K. Singh, and M. R. Maurizi. 1994. Processive degradation of proteins by the ATP-dependent Clp protease from Escherichia coli. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis. J Biol Chem 269:18209-18215.

91. Kolygo, K., N. Ranjan, W. Kress, F. Striebel, K. Hollenstein, K. Neelsen, M. Steiner, H. Summer, and E. Weber-Ban. 2009. Studying chaperone-proteases using a real-time approach based on FRET. J Struct Biol 168:267-277.

92. Olivares, A. O., A. R. Nager, O. Iosefson, R. T. Sauer, and T. A. Baker. 2014. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nature structural & molecular biology.

93. Bohon, J., L. D. Jennings, C. M. Phillips, S. Licht, and M. R. Chance. 2008. Synchrotron protein footprinting supports substrate translocation by ClpA via ATP-induced movements of the D2 loop. Structure 16:1157-1165.

94. Beuron, F., M. R. Maurizi, D. M. Belnap, E. Kocsis, F. P. Booy, M. Kessel, and A. C. Steven. 1998. At sixes and sevens: characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. Journal of structural biology 123:248-259.

95. Guo, F., L. Esser, S. K. Singh, M. R. Maurizi, and D. Xia. 2002. Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. J Biol Chem 277:46753-46762.

96. Farbman, M. E., A. Gershenson, and S. Licht. 2007. Single-Molecule Analysis of Nucleotide-Dependent Substrate Binding by the Protein Unfoldase ClpA. J Am Chem Soc.

97. Woo, K. M., K. I. Kim, A. L. Goldberg, D. B. Ha, and C. H. Chung. 1992. The heat-shock protein ClpB in Escherichia coli is a protein-activated ATPase. J Biol Chem 267:20429-20434.

98. Parsell, D. A., Y. Sanchez, J. D. Stitzel, and S. Lindquist. 1991. Hsp104 is a highly conserved protein with two essential nucleotide-binding sites. Nature 353:270-273.

99. Zolkiewski, M. 1999. ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from Escherichia coli. J Biol Chem 274:28083-28086.

100. Mogk, A., C. Schlieker, C. Strub, W. Rist, J. Weibezahn, and B. Bukau. 2003. Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. J Biol Chem 278:17615-17624.

101. Weibezahn, J., C. Schlieker, B. Bukau, and A. Mogk. 2003. Characterization of a trap mutant of the AAA+ chaperone ClpB. J Biol Chem 278:32608-32617.

102. Schlee, S., P. Beinker, A. Akhrymuk, and J. Reinstein. 2004. A chaperone network for the resolubilization of protein aggregates: direct interaction of ClpB and DnaK. J Mol Biol 336:275-285.

103. Shorter, J., and S. Lindquist. 2004. Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers. Science 304:1793-1797.

104. Schaupp, A., M. Marcinowski, V. Grimminger, B. Bosl, and S. Walter. 2007. Processing of proteins by the molecular chaperone Hsp104. J Mol Biol 370:674-686.

105. Parsell, D. A., A. S. Kowal, M. A. Singer, and S. Lindquist. 1994. Protein disaggregation mediated by heat-shock protein Hsp104. Nature 372:475-478.

106. Glover, J. R., and S. Lindquist. 1998. Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. Cell 94:73-82.

107. Goloubinoff, P., A. Mogk, A. P. Zvi, T. Tomoyasu, and B. Bukau. 1999. Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. Proc Natl Acad Sci U S A 96:13732-13737.

108. Haslberger, T., A. Zdanowicz, I. Brand, J. Kirstein, K. Turgay, A. Mogk, and B. Bukau. 2008. Protein disaggregation by the AAA+ chaperone ClpB involves partial threading of looped polypeptide segments. Nature structural & molecular biology 15:641-650.

109. Doyle, S. M., O. Genest, and S. Wickner. 2013. Protein rescue from aggregates by powerful molecular chaperone machines. Nat Rev Mol Cell Biol 14:617-629.

110. Nakazaki, Y., and Y. H. Watanabe. 2014. ClpB chaperone passively threads soluble denatured proteins through its central pore. Genes Cells 19:891-900.

111. Aguado, A., J. A. Fernandez-Higuero, Y. Cabrera, F. Moro, and A. Muga. 2015. ClpB dynamics is driven by its ATPase cycle and regulated by the DnaK system and substrate proteins. The Biochemical journal 466:561-570.

112. Kim, K. I., G. W. Cheong, S. C. Park, J. S. Ha, K. M. Woo, S. J. Choi, and C. H. Chung. 2000. Heptameric ring structure of the heat-shock protein ClpB, a protein-activated ATPase in Escherichia coli. J Mol Biol 303:655-666.

113. Watanabe, Y. H., K. Motohashi, and M. Yoshida. 2002. Roles of the two ATP binding sites of ClpB from Thermus thermophilus. J Biol Chem 277:5804-5809.

114. Schirmer, E. C., C. Queitsch, A. S. Kowal, D. A. Parsell, and S. Lindquist. 1998. The ATPase activity of Hsp104, effects of environmental conditions and mutations. J Biol Chem 273:15546-15552.

115. Parsell, D. A., A. S. Kowal, and S. Lindquist. 1994. Saccharomyces cerevisiae Hsp104 protein. Purification and characterization of ATP-induced structural changes. J Biol Chem 269:4480-4487.

116. Lee, S., J. M. Choi, and F. T. Tsai. 2007. Visualizing the ATPase cycle in a protein disaggregating machine: structural basis for substrate binding by ClpB. Mol Cell 25:261-271.

117. Zietkiewicz, S., J. Krzewska, and K. Liberek. 2004. Successive and synergistic action of the Hsp70 and Hsp100 chaperones in protein disaggregation. J Biol Chem 279:44376-44383.

118. Zietkiewicz, S., A. Lewandowska, P. Stocki, and K. Liberek. 2006. Hsp70 chaperone machine remodels protein aggregates at the initial step of Hsp70-Hsp100-dependent disaggregation. J Biol Chem 281:7022-7029.

119. Kedzierska, S., L. S. Chesnokova, S. N. Witt, and M. Zolkiewski. 2005. Interactions within the ClpB/DnaK bi-chaperone system from Escherichia coli. Arch Biochem Biophys 444:61-65.

120. Barnett, M. E., M. Nagy, S. Kedzierska, and M. Zolkiewski. 2005. The amino-terminal domain of ClpB supports binding to strongly aggregated proteins. J Biol Chem 280:34940-34945.

121. Maillard, R. A., G. Chistol, M. Sen, M. Righini, J. Tan, C. M. Kaiser, C. Hodges, A. Martin, and C. Bustamante. 2011. ClpX(P) generates mechanical force to unfold and translocate its protein substrates. Cell 145:459-469.

122. Aubin-Tam, M. E., A. O. Olivares, R. T. Sauer, T. A. Baker, and M. J. Lang. 2011. Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. Cell 145:257-267.

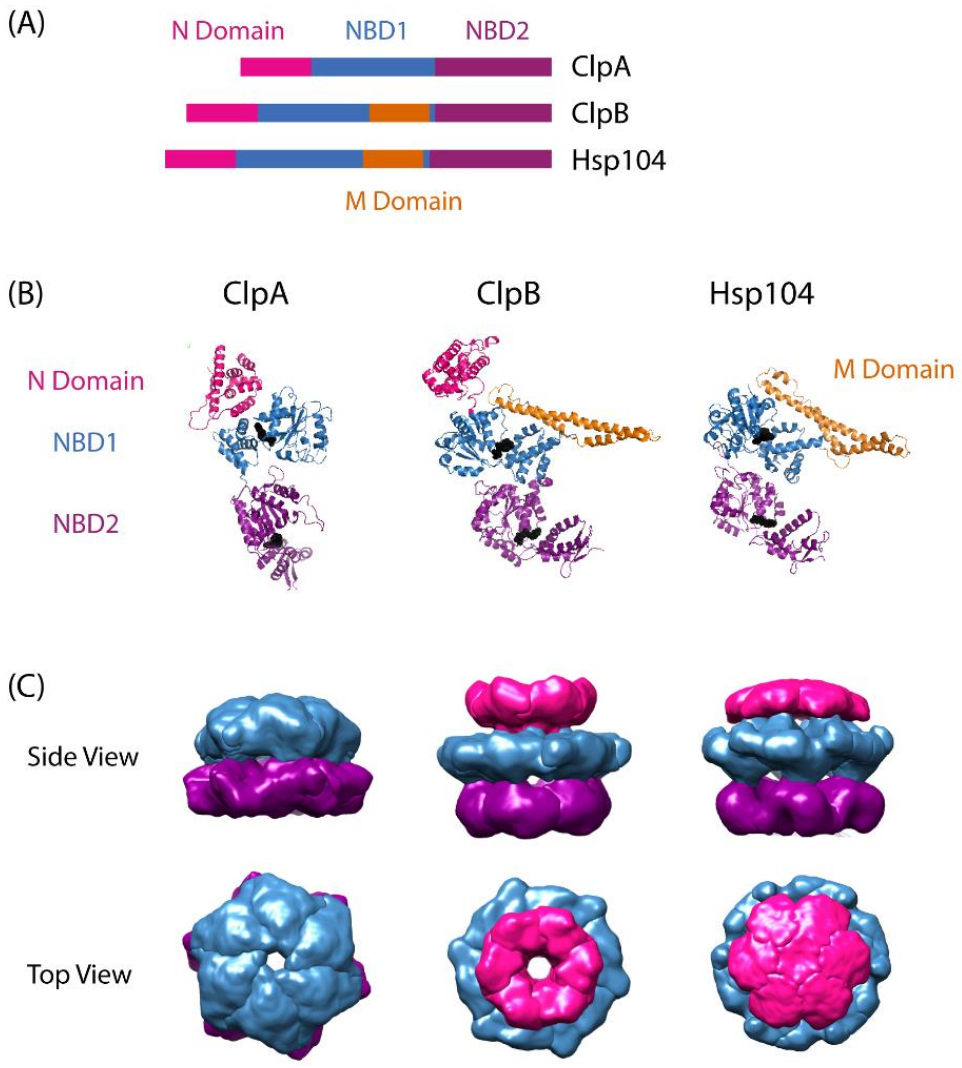


Figure 1. Structural comparison of ClpA, ClpB, and Hsp104. (A) Sequence alignment showing relative organization of N domain, NBD1, NBD2, and M domain in the AAA+ protomers compared. (B) Protomer crystal structures of E. coli ClpA (PDB ID code 1ksf) (Guo et al., 2002b), T. thermophilus ClpB (PDB ID code 1qvr– chain C) (Lee et al., 2003), and C. thermophilum Hsp104 (PDB ID 5d4w – chain A) (Heuck et al., 2016). E. coli ClpA and T. thermophilus ClpB N domains are shown in pink. C. thermophilum Hsp104 also has an N terminal domain, however its electron density was not resolved, likely due to flexibility. Nucleotide Binding Domain 1 (NBD1) is shown in blue for each protomer. In ClpB and Hsp104, the Middle Domain (M Domain) is shown in gold, extending in a coiled-coil from within NBD1. Nucleotide Binding Domain 2 (NBD2) is shown in purple. Bound nucleotide is shown as black spheres. These images were prepared using PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC (Schrodinger, 2015a,b,c). Adaptation of structural comparison presented in Doyle and Wickner (2009). (B) Single particle reconstructions of E. coli ClpA (EMD-1673) (Effantin et al., 2010), E. coli ClpB (EMD-2563) and S. cerevisiae Hsp104 (EMD-2561) (Carroni et al., 2014), hexameric rings from cryo-electron microscopy. ClpA and Hsp104 models were built from images of the

motor protein bound to ClpP. For Hsp104, this required use of HAP, the variant designed by the Bukau group to interact with ClpP. Top row shows views from the side. Note that the N terminal domain of ClpA was not defined in the electron density map, likely due to flexibility, similar to the observation from the crystallographic study of Hsp104. Bottom row shows views from the top, looking down through the axial channel. These images were prepared using UCSF Chimera (Computer Graphics Laboratory, University of California, San Francisco).

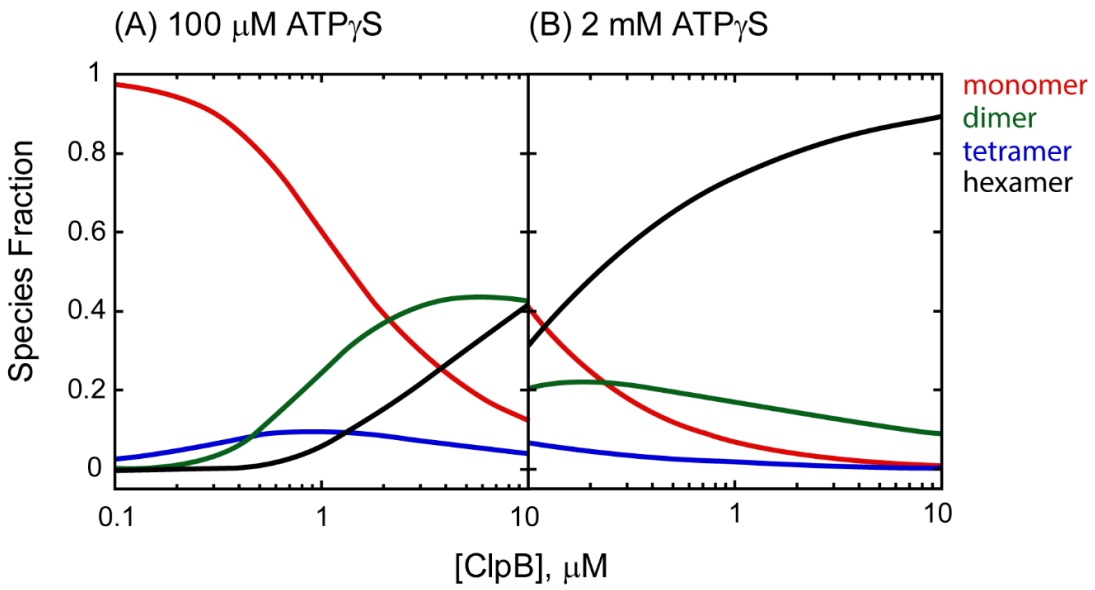


Figure 2. Species fraction plot as a function to total [ClpB] in monomer units. Species fractions were simulated using apparent equilibrium constants for the oligomerization of each ClpB n-mer (Ln, app) predicted in the presence of (A) 100 μM and (B) 2 mM ATPγS, as well as ClpB n-mer nucleotide binding constants, and stoichiometries reported in Lin and Lucius (2016). The equilibrium constants for the ClpB n-mer oligomerization in the absence of nucleotide (Ln,0) were used as reported in Lin and Lucius (2015b). The solid lines represent the fraction of monomer (red), dimer (green), tetramer (blue), and hexamer (black) present as a function of total [ClpB] in μM monomer.

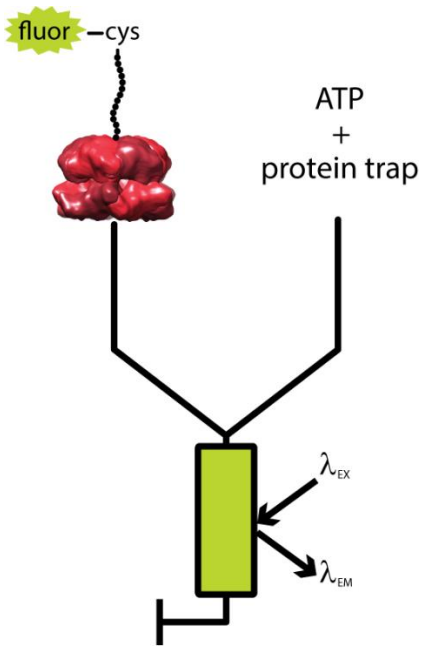


Figure 3. Schematic of single turnover fluorescence stopped-flow experiment. ATPγS-bound ClpA is pre-assembled with a fluorescently labeled, unstructured polypeptide substrate, fluor-peptide. The fluor-peptide bound ClpA complex (left) is then rapidly mixed with a solution of ATP and a non-fluorescent peptide (protein trap, right) held in large excess over the fluorescently modified peptide concentration. Upon mixing, any ClpA hexamers that dissociate from the fluorescently modified peptide will be swiftly bound by protein trap, ensuring the reaction monitored is single-turnover with respect to the fluor-peptide bound ClpA complex. The mixture is excited at a specified fluorophore excitation wavelength (λ_{EX}), and fluorescence emission at an indicated fluorophore emission wavelength (λ_{EM}) is monitored as a function of time.

$$
\begin{tikzcd}
(E \bullet P)_L \arrow[r, "k_t"] \arrow[d, "k_d"] & I_{(L-m)} \arrow[r, "k_t"] \arrow[d, "k_d"] & \ldots I_{(L-im)} \ldots \arrow[r, "k_t"] \arrow[d, "k_d"] & E + P \\
P & P & P &
\end{tikzcd}
$$

Figure 4. General scheme of a translocating enzyme mechanism. Translocating enzyme (E) in complex with a peptide (P) of length L, (E·P)L, will either translocate the peptide through a translocation rate constant (k_t) to form an peptide intermediate translocated by a some distance m, I(L−m), or dissociate from the peptide through a dissociation rate constant (k_d). The translocase proceeds through multiple translocation steps of a given step-size (m) until the peptide is fully translocated.

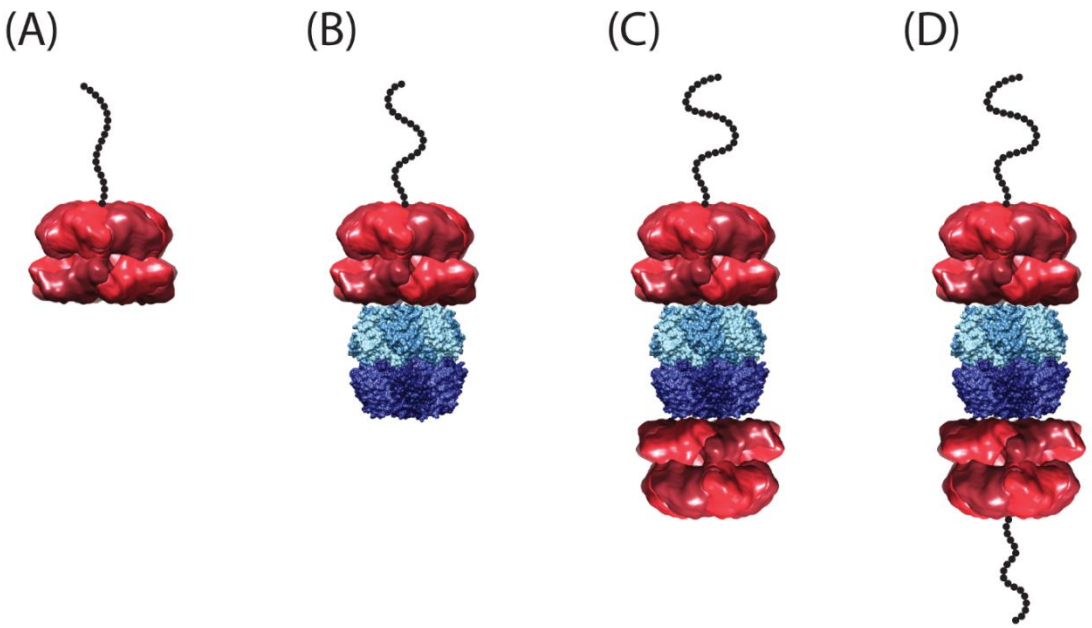


Figure 5. Structural models of ClpA and ClpAP complex in various states of peptide ligation. Models of (A) hexameric ClpA, (B) 1:1 ClpAP, and (C) 2:1 ClpAP bound by one peptide or (D) bound by two peptides. Structures are shown as side views in complex with a cartoon of an unstructured polypeptide substrate (black). The single particle reconstruction of E. coli ClpA hexamer (EMD-1673) (Effantin et al., 2010) is shown with monomers colored in alternating red shades. In the ClpAP complexes, a molecular surface from the crystal structure of E. coli ClpP tetradecamer (PDB-2FZS) (Szyk and Maurizi, 2006) is shown with protomers in the each heptameric ring colored in alternating shades of either light blue (top) or dark blue (bottom). The models shown here are not energy minimized. Images were prepared using UCSF Chimera (Computer Graphics Laboratory, University of California, San Francisco).

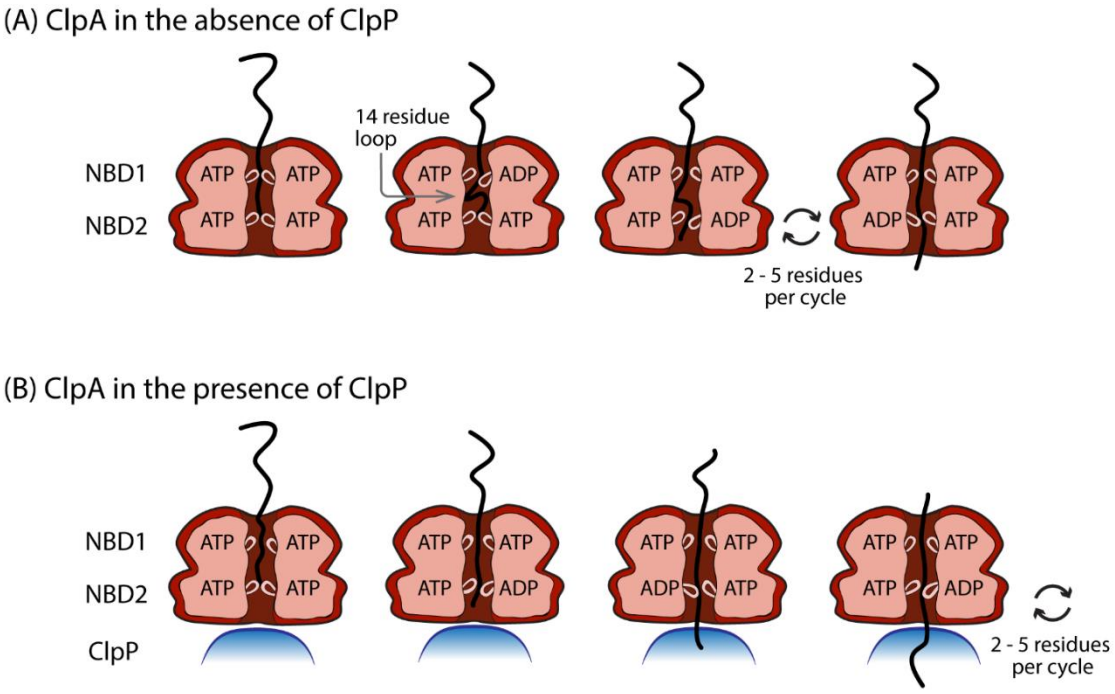


Figure 6. Proposed model of the movement of pore loops in NBD1 and NBD2 of ClpA in polypeptide translocation. (A) In the absence of ClpP, conformational changes in the pore loops of both NBD1 and NBD2 contribute to the translocation of polypeptide substrate through the ClpA axial channel. ATP binding and hydrolysis at NBD1 results in a pore loop conformational change that moves the incoming polypeptide substrate ~14 amino acids down the axial channel toward NBD2. This results in the formation of a polypeptide substrate loop in the axial space between NBD1 and NBD2. This loop is moved through the axial channel by multiple rounds of ATP hydrolysis cycles at NBD2 that lead to the translocation of 2–5 amino acids per cycle by the NBD2 pore loops. (B) In the presence of ClpP, polypeptide translocation is driven by NBD2 ATP hydrolysis induced conformational changes. Cycles of ATP binding and hydrolysis at NBD1, do not result in conformational changes that limit the observed ClpAP catalyzed translocation rate.

CHAPTER 3

ATP HYDROLYSIS INACTIVATING WALKER B MUTATION PERTURBS *E. COLI* CLPA SELF-ASSEMBLY ENERGETICS IN THE ABSENCE OF NUCLEOTIDE

by

ELIZABETH C. DURAN AND AARON L. LUCIUS

In submission to *Biophysical Chemistry Journal*

Format adapted for dissertation

ABSTRACT

E. coli ClpA is an AAA+ (ATPase Associated with diverse cellular Activities) chaperone that catalyzes the ATP-dependent unfolding and translocation of substrate proteins for the purposes of proper proteome maintenance. Biologically active ClpA hexamers contain two nucleotide binding domains (NBD) per protomer, D1 and D2. Despite extensive study, complete understanding of how the twelve NBDs within a ClpA hexamer coordinate ATP binding and hydrolysis to polypeptide translocation is currently lacking. To examine nucleotide binding and coordination at D1 and D2, ClpA Walker B variants deficient in ATP hydrolysis at one or both NBDs have been employed in various studies. In the presence of ATP, it is widely assumed that ClpA Walker B variants are entirely hexameric. However, a thermodynamically rigorous examination of the self-assembly mechanism has not been obtained. Differences in the assembly due to the mutation can be misattributed to the active NBD, leading to potential misinterpretations of kinetic studies. Here we use sedimentation velocity studies to quantitatively examine the self-assembly mechanism of ClpA Walker B variants deficient in ATP hydrolysis at D1, D2, and both NBDs. We found that the Walker B mutations had clear, if modest, effects on the assembly. In the absence of nucleotide, the assembly of each Walker B variant examined is best described by a monomer, dimer, tetramer, hexamer equilibrium. These data suggest that the Walker B mutation stabilizes a hexamer population in the absence of nucleotide, that is not present for analogous concentrations of wild type ClpA. Our results indicate that Walker B mutants, widely used in studies of AAA+ family proteins, require additional characterization as the mutation affects not only ATP hydrolysis, but also ligand

linked assembly of these complexes. This linkage must be considered in investigations of unfolding or other ATP dependent functions.

INTRODUCTION

Hsp100/Clp proteins are members of the AAA+ superfamily (ATPases Associated with diverse cellular Activities). This superfamily includes molecular motors that play crucial and diverse roles in protein quality control pathways (1, 2). Some of the processes undertaken by AAA+ motor proteins to ensure proper proteome maintenance include the remodeling of toxic protein aggregates formed during cellular stress, regulation of protein complex assembly, and unfolding and degradation of substrates targeted for destruction (3, 4). In these pathways, the AAA+ protein serves as the motor component, often within a larger protein assembly, coupling the energy from ATP binding and hydrolysis into physical movements of the substrate for a particular function. For example, the bacterial AAA+ motor ClpB (Hsp104 in yeast) works in collaboration with the DnaKJE system (Hsp70 and Hsp40 in yeast) to resolubilize protein aggregates. In a different pathway, the bacterial motors ClpA and ClpX associate with the ClpP protease to form energy-dependent protein degradation machines with architectural similarity to the eukaryotic 26S proteasome (4-8).

The presence of one or two nucleotide binding domains (NBD) within Clp proteins have formed the basis for their subclassification into one of two classes (1). Class I AAA+ proteins such as ClpA, ClpB, ClpC, and Cdc48, contain two NBDs, D1 and D2, per monomer (9-12). In contrast, Class II AAA+ proteins, such as ClpX, HslU, Lon, and FtsH have a single NBD per monomer (13-16). Each NBD contains the canonical Walker A and B motifs that mediate nucleotide binding and hydrolysis, respectively (17). ATP binding is known to drive the oligomerization of Class I and II AAA+ motors into homohexameric rings with an axial channel.

E. coli ClpA is a representative, well studied, Class I AAA+ molecular chaperone. In the presence of ATP, hexameric ClpA couples the energy obtained from ATP binding and hydrolysis to the translocation of polypeptide substrates down its axial channel. The tugging and pulling force of translocation causes the unfolding of polypeptide clients. ClpA hexamers translocate polypeptide clients in the absence of, or in complex with ClpP (ClpAP). Polypeptide substrates translocated by ClpAP are transferred from the axial channel of hexameric ClpA into a cavity formed by ClpP tetradecamers for proteolytic degradation. Multiple groups have reported that ClpAP translocates polypeptide substrates with a faster overall rate than ClpA alone (4, 18, 19). Despite this, it's not completely understood how the twelve NBD sites within ClpA hexamers coordinate and couple ATP hydrolysis to polypeptide translocation in the absence and presence of ClpP.

In addition to protein unfolding and translocation, ClpA requires ATP binding to assemble into biologically active hexamers (20). Thus, a detailed understanding of how nucleotide binding and hydrolysis by ClpA is coordinated to polypeptide substrate unfolding and translocation requires the complete characterization of its assembly in the absence and presence of nucleotide. In the absence of nucleotide, ClpA has been shown to reside in a distribution of monomers, dimers, and tetramers (21, 22). More recent work shows that in addition to ClpA hexamers, lower order ClpA oligomers persist even at elevated nucleotide concentrations. Consequently, characterization of the biologically active hexamer population in the presence of nucleotide depends on quantification of the nucleotide linked assembly energetics.

ATP hydrolysis at each ClpA NBD has been studied using ClpA Walker B variants deficient in ATP hydrolysis at D1 or D2 (23, 24). From steady state ATPase assays using

ClpA Walker B variants, Weber-Ban and co-workers observe a lower ATPase activity at D1 relative to that of D2 (23). Moreover, in the presence of ClpP they observe polypeptide substrate-dependent differences in the rates of ClpA ATP hydrolysis at D1 and D2. More structurally ordered substrates required fully active D1 and D2 for degradation, while the ATPase activity at D1 was sufficient for degradation of less stable substrates.

More recently, Baytshtok et. al. use ClpA Walker B variants to examine the energetic contribution of each NBD at unfolding a model substrate in the absence and presence of ClpP (24). They observed a decrease in the degradation activity of the model substrate using a ClpA variant with hydrolysis inactivated at D1, relative to that of ClpA with unaltered D1 and D2. Inactivating ATP hydrolysis at D2 resulted in total elimination of the substrate degradation activity. Although such differences are attributed to the effect of ClpP on the ATPase activity of each NBD, it's also possible that changes in the assembly state due to the presence of the mutations are contributing to the observed differences. In all studies testing the steady state ATPase activity of ClpA Walker B variants, the concentration of ClpA hexamers is calculated by simply dividing the monomer ClpA concentration by six. That is, the lower order oligomers are not accounted for. If either the Walker B mutation or the presence of ClpP perturbs the assembly state of ClpA, differences in the population of hexamers catalyzing substrate turnover could be misinterpreted as differences in the ATPase activity of D1 and D2.

Here we asked whether there are changes to the ClpA assembly state as a consequence of introducing an ATP hydrolysis inactivating mutation in the Walker B motif of D1, D2, or both D1 and D2. Using thermodynamically rigorous sedimentation velocity experiments, we describe the self-assembly mechanism of wild type ClpA ($ClpA_{WT}$) and

each ClpA Walker B variant in the absence of nucleotide. We observe differences between the self-assembly mechanisms of $ClpA_{WT}$ and the Walker B variants. Our findings show that introduction of the E286A and E565A mutations, which have been shown to eliminate ATP hydrolysis but not binding in D1 and D2, respectively, perturb the intrinsic propensity of the protein to self-assemble.

MATERIALS AND METHODS

Buffers and Reagents

Reagent grade materials were used in the preparation of all buffers. Buffers were prepared using ultrapure water obtained from a Purelab Ultra System (Evoqua Water Technologies). ATP was obtained from Sigma-Aldrich. Buffer H500 contains 25 mM HEPES, pH 7.5 at 25 °C, 10 mM $MgCl_2$, 2 mM 2-mercapthoethanol, 500 mM NaCl, and 10 % (v/v) glycerol.

Strains, plasmids, and proteins

The $ClpA_{WT}$ gene was cloned into pET 30a plasmid ($pClpA_{WT}$) under control of the T7 promoter and the protein was isolated as previously described (21). Plasmids containing $ClpA_{E286A}$ and $ClpA_{E565A}$ genes ($pClpA_{E286A}$ and $pClpA_{E565A}$, respectively) were created by inserting the indicated single point mutations to the ClpA gene within $pClpA_{WT}$. The $ClpA_{E286A/E565A}$ plasmid ($pClpA_{E286A/E565A}$) was created by inserting an E565A mutation into the ClpA gene of $pClpA_{E286A}$. All single point mutations were done using the QuickChange II site-directed mutagenesis kit (Agilent Technologies). Primers for single point mutations were purchased from Integrated DNA Technologies. All mutations were verified by DNA sequencing. Cells containing $pClpA_{WT}$, $pClpA_{E286A}$, $pClpA_{E565A}$, or

$pClpA_{E286A/E565A}$ were selected by growing in LB Miller growth media (Fischer Scientific) containing 30 µg/mL kanamycin.

$ClpA_{WT}$ knockout BL21(DE3) cells, $\Delta ClpA_{WT}$-BL21(DE3), were constructed using recombineering (25). An ampicillin cassette was inserted in place of the $ClpA_{WT}$ gene in the BL21(DE3) genome. The knockout cells were selected using 25 µg/mL ampicillin in growth media.

Electrocompetent stocks of $\Delta ClpA_{WT}$-BL21(DE3) cells were transformed with $pClpA_{E286A}$, $pClpA_{E565A}$, or $pClpA_{E286A/E565A}$ using electroporation. ClpA Walker B variants were overexpressed in $\Delta ClpA_{WT}$-BL21(DE3) cells by adding 1 mM IPTG during mid-log growth. Each Walker B variant was isolated using the same protocol as used for the isolation of $ClpA_{WT}$, previously described (21).

Sedimentation velocity experiments

Sedimentation velocity experiments on 6, 10, and 20 µM $ClpA_{WT}$, $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$ were performed using an ProteomeLab XL-I analytical ultracentrifuge (Beckman Coulter, Brea, CA). For each experiment, 390 µL of sample or reference was loaded into each corresponding sector of a dual channel, Epon charcoal filled centerpiece. Samples were subjected to an angular velocity of 40,000 rpm and interference boundaries were collected every 30 seconds for the duration of sedimentation. Protein and reference solutions incubated at 25 °C for 2 hours before centrifugal force was applied.

Interference scans were corrected for time stamp errors using REDATE Version 0.1.7 (Chad Brautigam, University of Texas Southwestern Medical Center) (26, 27). Time corrected data was subsequently subjected to c(s) analysis using Sedfit version 14.4f (Peter

Schuck, NIH) as previously described (22, 28). The resulting c(s) distributions were integrated over the range of 1 – 25 S to obtain weight average sedimentation coefficients. Sedimentation coefficients, s, are corrected to $s_{20,w}$ to reflect standard conditions of infinite dilution in water at 20 °C, as previously described (22), using Eq. (1):

$$s_{20,w} = \frac{\left(1-\rho_{20,w}\bar{v}\right)}{\left(1-\rho\bar{v}\right)} \cdot \frac{\eta}{\eta_{20,w}} \cdot s \tag{1}$$

where ρ is the solution density, η the viscosity, and $\bar{v}$ the partial specific volume of the protein. ClpA partial specific volume was determined from its primary protein sequence to be 0.7405 mL g^{-1} using Sednterp (David Hayes, Magdalen College, Tom Laue, University of New Hampshire, and John Philo, Alliance Protein Laboratories) (29). A corrected value of $\bar{v}$ = 0.7438 mL g^{-1} was used to correct for the presence of 10 % (v/v) glycerol in our buffer, as previously done (22, 28). This results in a sedimentation coefficient correction factor of 1.501.

ClpA oligomer sedimentation coefficients used in global analysis of time difference curves

The sedimentation coefficient for ClpA monomer and hexamer in buffer H500 were determined experimentally. Sedimentation coefficient for other ClpA oligomers was approximated using Eq. (2):

$$s_n = s_1 \left(n\right)^{2/3} \tag{2}$$

where *n* represents the number of protomers in a ClpA oligomer. Importantly, this approximation is limited by the assumption that all ClpA *n*-mers have the same frictional ratio (30). Using the experimentally determined s_1 = 2.8 S leads to a calculated s_6 = 9.2 S

from Eq. (2). This is 2 % different compared to the experimentally determined s_6 = 9.4 S. Fits done using values for the s_2 and s_4 within a difference of 2 % do not change the equilibrium constants determined from global analysis of time difference curves. Table 1 shows the sedimentation coefficients for each oligomer constrained in the global fitting.

Global Analysis of sedimentation velocity time difference curves

SedAnal was used to transform sedimentation velocity data into time difference curves (31). Time difference curves for 6, 10, and 20 µM of $ClpA_{WT}$, $ClpA_{E286A}$, $ClpA_{E565A}$, or $ClpA_{E286A/E565A}$ were subsequently subjected to global NLLS analysis to determine thermodynamic information from the self-assembly of each protein. For each concentration, every scan collected in the first 3 hours of sedimentation was included in the global analysis.

A set of path independent equilibrium equations were used to describe the data with each model tested (Table 2). The numbers used in the name of each model describe the oligomers present at stoichiometric equilibrium. For example, the 1-2-4 model consists of the path-independent stoichiometric reactions in Eq. (3) – (4):

$$2ClpA_1 \xrightleftharpoons{L_{2,0}} ClpA_2 \tag{3}$$

$$4ClpA_1 \xrightleftharpoons{L_{4,0}} ClpA_4 \tag{4}$$

$$6ClpA_1 \xrightleftharpoons{L_{6,0}} ClpA_6 \tag{5}$$

where $L_{n,0}$ is the stoichiometric binding constant for the formation of a ClpA dimers and tetramers. Similarly, the 1-2-4-6 model consists of reactions in Eq. (3) – (5) describing the

association of ClpA monomers to form dimers, tetramers, and hexamers. Equilibrium constants and loading concentrations were floated parameters in the analysis.

The goodness of fit between two models tested on the same data set was compared using the *F*-test, as previously described (32). A calculated *F* value, $F_{calculated}$, was determined from Eq. (6):

$$F_{calculated} = \frac{RMSD_2^{\,2}}{RMSD_1^{\,2}} \tag{6}$$

where the analysis using a model that resulted in the lowest RMSD ($RMSD_1$) was compared to the analysis using a different model ($RMSD_2$). The $F_{calculated}$ was compared to a critical *F* value, $F_{critical}$, of 1.00002. An $F_{calculated}$ value larger than $F_{critical}$ indicates a significantly worse fit compared to the best fit.

Oligomer sedimentation coefficients, molecular weights, extinction coefficients, and density increments were constrained parameters in the analysis. Table 1 lists the sedimentation coefficients and molecular weight used. The fringe extinction used for ClpA and each variant is 3.192 fringes per mg/mL. Density increment was calculated using Eq. (7).

$$1-\bar{v}\rho \tag{7}$$

RESULTS & DISCUSSION

ClpA Walker B variants exhibit dynamic assembly in the absence of nucleotide

Figure 1 A shows the location of the Walker B motif in each NBD of the ClpA monomer crystal structure. Residues 281-287 (ILFIDEI) make up the Walker B motif of

NBD1 and in NBD2 it is residues 561-567 (LLLDEIE). Replacement of the catalytic glutamate within the Walker B motif of either NBD (Figure 1 A) with alanine abolishes ATP hydrolysis but maintains binding activity (23). Previous studies have established that nucleotide binding, but not hydrolysis, is required to assemble ClpA into hexamers (33-35). Consistently, ClpA Walker B variants, have been shown to bind nucleotide and assemble into hexamers capable of binding ClpP. Multiple groups have examined the ATPase activity of ClpA Walker B variants deficient in ATP hydrolysis in one NBD in order to describe the ATPase activity of the functional NBD.

The hexamerization equilibrium constant for a protein that requires nucleotide binding to form hexamers is given by Eq. (8)

$$L_{6,app} = L_{6,0} \frac{P_6}{P_1^6} \tag{8}$$

Where $L_{6,app}$ denotes an apparent hexamerization constant that depends upon the nucleotide concentration. $L_{6,0}$ is the hexamerization equilibrium constant in the absence of nucleotide, where the 6 represents hexamers and the zero represents zero bound nucleotides, and P_1 and P_6 are the partition functions for nucleotide binding to the monomeric and hexameric state, respectively. The partition functions are functions of the nucleotide binding constant to each nucleotide binding site and the free nucleotide concentration. For example, if both sites in the monomer bind with identical binding constants, then P1 would be given by Eq. (9), where K is the binding constant for binding to either D1 or D2, i.e. two independent and identical sites.

$$P_1 = \left(1 + K\left[ATP\right]\right)^2 \tag{9}$$

Upon inspection of Eq. (8) one can predict that introduction of a mutation in the nucleotide binding site will influence the hexamerization constant, $L_{6,app}$, if the nucleotide binding constant, K, is affected by the mutation. Even though the Walker B variants still bind nucleotide, it seems unlikely that in the presence of the mutation the affinity would be unchanged. Further, $L_{6,0}$ represents the ability of the protein to assemble in the absence of nucleotide. We refer to this as the intrinsic propensity of the protein to assemble. It also seems possible that the introduction of a mutation could perturb the intrinsic propensity of the protein to assemble, i.e. $L_{6,0}$.

Indeed, it has been frequently, shown using size exclusion chromatography at a fixed nucleotide concentration, that these variants do form hexamers (23). However, the observation of hexamers under one set of conditions does not reveal the impact on the self-association equilibrium. Using a combination of hydrodynamic techniques, we previously showed that ClpA resides in a dynamic equilibrium of monomers, dimers, and tetramers in the absence of nucleotide (21, 22). If the self-assembly energetics are different for each Walker B variant, then a comparison, for example, of the steady state ATPase activity as a function of protein concentration would be reporting on different populations of oligomers turning over ATP at each fixed total ATP concentration. To determine if introduction of the mutation influences the intrinsic propensity of the protein to assemble, Walker B variants deficient in ATP hydrolysis at NBD1 ($ClpA_{E286A}$), NBD2 ($ClpA_{E565A}$), or both NBDs ($ClpA_{E286A/E565A}$) were overexpressed and isolated as described (see Materials and Methods).

As a necessary precursor to investigating whether abolishing ATP hydrolysis at either one or both nucleotide binding domains of ClpA perturbs assembly, we first

examined the association of wild type ClpA ($ClpA_{WT}$) in the absence of nucleotide. Although we have done and reported this, our previous hydrodynamic studies on the self-assembly of $ClpA_{WT}$ were carried out in the presence of 300 mM NaCl. However, under the solution conditions used in this study (25 mM HEPES, pH 7.5, 2 mM 2-ME) all Walker B variants exhibited poor solubility at 300 mM NaCl; a first indication that the conformation has been changed due to introduction of point mutations. Consequently, all hydrodynamic studies presented here on $ClpA_{WT}$ and Walker B variants were performed in the presence of 500 mM NaCl as described in Materials and Methods.

Sedimentation velocity experiments were performed on 6, 10, and 20 μM $ClpA_{WT}$ monitoring interference. Figure 1 B shows a representative set of interference fringe boundaries collected for 6 μM $ClpA_{WT}$ by monitoring changes in fringe displacement as a function of radial position and time. Boundaries collected for 6, 10, and 20 μM $ClpA_{WT}$ were subjected to c(s) analysis in SedFit (Peter Schuck, NIH). Figure 1 C shows an overlay of the resulting c(s) distributions. Inspection of the c(s) distributions (Figure 1 C) reveals the presence of multiple peaks for each concentration, and a shift towards higher sedimentation coefficients with increasing protein concentration. These data are consistent with previous findings showing that ClpA resides in a distribution of oligomers in the absence of nucleotide.

With these data in hand we asked the question, does the equilibrium distribution of ClpA oligomers change as a result of introducing the Walker B mutation at one or both NBDs? That is, do ClpA Walker B variants exhibit different assembly energetics relative to $ClpA_{WT}$? To test this, analogous sedimentation velocity experiments were performed on 6, 10, and 20 μM $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$. As a first level of analysis,

weight average sedimentation coefficients ($\bar{s}$) were obtained and compared for $ClpA_{WT}$ and each ClpA Walker B variant. C(s) distributions for each concentration of $ClpA_{WT}$ (Figure 1 C) were integrated over the entire sedimentation coefficient range to obtain a weight average sedimentation coefficient for each concentration. The same analysis was repeated on the c(s) distributions for 6, 10, and 20 μM of $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$. Figure 1 D shows the resulting weight average sedimentation coefficients for $ClpA_{WT-500}$ (green), $ClpA_{E286A}$ (purple), $ClpA_{E565A}$ (blue), and $ClpA_{E286A/E565A}$ (pink) as a function of protein concentration. From Figure 1 D, it is clear that there are differences in $\bar{s}$ for Walker B variants relative to $ClpA_{WT}$.

The population of each oligomer at any given protein concentration is defined by the equilibrium constants in the self-assembly mechanism. Consequently, one explanation for the observed differences in Figure 1 D is that the indicated mutations are perturbing the energetics of ClpA assembly. To examine this, we next sought to characterize the self-assembly mechanism of $ClpA_{WT}$ and each Walker B variant.

ATP binding drives assembly of $ClpA_{E286A/E565A}$

In order to analyze the energetics of assembly for ClpA Walker B variants in the absence of nucleotide, the sedimentation coefficients of $ClpA_{E286A/E565A}$ monomer and hexamer were experimentally determined. Sedimentation velocity experiments were performed on 2 μM $ClpA_{E286A/E565A}$ in the absence of nucleotide and in the presence of 1 mM ATP (Figure 2). The sedimentation velocity boundaries were subjected to c(s) analysis using SedFit. Figure 2 A shows the resulting c(s) distribution of 2 μM $ClpA_{E286A/E565A}$ in the absence of nucleotide. In the absence of nucleotide, $ClpA_{E286A/E565A}$ sediments as a single ideal species with $s_{20,w}$ = (4.3 ± 0.2) S. The reported error is the sedimentation

coefficient standard deviation obtained by integrating the c(s) distribution of three replicates. Veronese, et. al. report a sedimentation coefficient of $s_{20,w}$ = (4.5 ± 0.1) S for $ClpA_{WT}$ monomers in the presence of 300 mM NaCl and 10 mM $MgCl_2$ (see Table 3, (21)). Thus, the peak present in Figure 2 A is consistent with $ClpA_{E286A/E565A}$ monomers. The apparent decrease in sedimentation coefficient for $ClpA_{E286A/E65A}$ relative to that reported for $ClpA_{WT}$ could be due to differences in the hydration of the sedimenting monomers in the presence of the higher salt used in this study.

To characterize the sedimentation coefficient of the $ClpA_{E286A/E565A}$ hexamer in buffer supplemented with 500 mM NaCl, sedimentation velocity experiments were performed on various concentrations of $ClpA_{E286A/E565A}$ in the presence of 1 mM ATP. Figure 2 B shows an overlay of the resulting c(s) distributions of 2 μM (pink), 6 μM (blue), and 10 μM (orange) $ClpA_{E286A/E565A}$ in the presence of 1 mM ATP. In each case, there is a ~4.3 S peak, consistent with the monomer population as well as a peak with $s_{20,w}$ = (14.6 ± 0.2) S. The ~14.6 S peak does not shift with increasing $ClpA_{E286A/E565A}$ concentration, indicating that it represents the largest oligomer populated in the presence of ATP. Various studies show that nucleotide binding, in the absence of hydrolysis, drives the formation of ClpA hexamers capable of interacting with ClpP (23, 35). We previously reported a sedimentation coefficient of 15.5 S for peptide-bound ClpA hexamers populated in the presence of 1 mM ATPγS (28). Thus, the peak at ~14.6 S observed in Figure 2 B is consistent with a $ClpA_{E286A/E565A}$ hexamer population. Consequently, a hexamer sedimentation coefficient value of $s_{20,w}$ = 14.6 S.

Global analysis of $ClpA_{WT}$ sedimentation velocity data

$ClpA_{WT}$ has been reported to reside in a distribution of monomers, dimers, and tetramers in the absence of nucleotide in buffer supplemented with 300 mM NaCl (21, 22). It is possible that some of the intermediates populated in the presence of 300 mM NaCl are not significantly populated at the elevated salt concentration of 500 mM NaCl, examined here. To test this, the sedimentation velocity boundaries collected for $ClpA_{WT}$ in buffer supplemented with 500 mM NaCl were globally analyzed using the time difference curve method (see Materials and Methods).

The difference curves that result from sedimentation velocity experiments collected at 6, 10, and 20 µM $ClpA_{WT}$ in buffer H500 are shown in Figure S1. These data were subjected to global NLLS analysis to the stoichiometric binding models listed in Table 2. The first model tested assumes that monomers, dimers, and tetramers (1-2-4 in Table 2) are significantly populated at thermodynamic equilibrium. In this analysis the equilibrium constants for the path independent formation of dimers ($L_{2,0}$) and tetramers ($L_{4,0}$) is allowed to float, where the 2 and 4 in the subscript indicate dimer and tetramer, respectively, and the zero indicates no bound nucleotide. This model describes the data reasonably well, with an RMSD = 2.699×10^{-2} (Figure S1, red traces). The equilibrium constants result in values of $L_{2,0} = (1.51 \pm 0.05) \times 10^{4}\ M^{-1}$ and $L_{4,0} = (8.9 \pm 0.9) \times 10^{13}\ M^{3}$ (average ± SD, n=3). Importantly, this model does not preclude the existence of other intermediates. For example, because ClpA forms hexamers in the presence of nucleotide, it is possible that a hexamer population exists, but is not sufficiently populated to detect in these experiments.

To determine if the analysis can detect additional intermediates, the difference curves shown in Figure S1 were subjected to global analysis to three additional models that

assume the presence of a hexamer population: a monomer-dimer-hexamer model (1-2-6), a monomer-tetramer-hexamer model (1-4-6), and a monomer-dimer-tetramer-hexamer model (1-2-4-6). The data are not well described by globally fitting to either the 1-2-6 or 1-4-6 models based on resulting RMSD values of 1.654×10^{-1} and 5.866×10^{-2}, respectively. Results from the F-test shown in Table 2 comparing each of these models to the 1-2-4 model supports that the 1-2-6 and 1-2-4-6 are significantly worse models at describing the observed difference curves. Global NLLS fit to the 1-2-4-6 model results in an RMSD = 2.704×10^{-2}. Notice that this is only 0.00005 larger than the resulting RMSD for fitting the same data to a 1-2-4 model. Although both the 1-2-4 and 1-2-4-6 models describe the data well, the 1-2-4 model describes the data significantly better based on the RMSD and the F-test shown in Table 2.

Table 3 shows the equilibrium constants that result from global NLLS analysis of $ClpA_{WT}$ fit to either a 1-2-4 or a 1-2-4-6 model ($L_{n,0}$ values reported as average ± SD, n=3). The resulting values of $L_{2,0}$ and $L_{4,0}$ from fitting to either model are within error of one another. Fitting to the 1-2-4-6 model additionally results in an equilibrium constant for hexamer formation, $L_{6,0} = (4 \pm 5) \times 10^{22}\ M^{-5}$. Judging from the large uncertainty, this resulting parameter is not well constrained from globally fitting three replicate concentration dependent data sets. This suggests that there is not enough information in the data to describe the energetics of hexamerization, perhaps due to a very low population of hexamers present at the concentrations tested. Thus, a monomer-dimer-tetramer model is the simplest model to sufficiently describe the experimental observations for $ClpA_{WT}$ in the presence of 500 mM NaCl.

Determination of the self-assembly mechanism for ClpA Walker B variants

Sedimentation velocity data collected for 6, 10, and 20 μM $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$ were subjected to NLLS analysis using the difference curve method. Figure 3 (black) shows one representative set of difference curves for 6, 10, and 20 μM $ClpA_{E286A/E565A}$. These difference curves were subjected to global NLLS analysis using the models previously tested on $ClpA_{WT}$ (Table 2). For each Walker B variant, based on the fitting RMSD and subsequent F-test shown in Table 2, the data are significantly better described by the 1-2-4-6 model compared to all other models tested. This observation suggests that hexamers are significantly populated for each of the variants, which is not the case for WT.

Figure 3 (red traces) shows the difference curves from global NLLS modeling of the $ClpA_{E286A/E565A}$ data using the 1-2-4-6 model. Similarly, global NLLS analysis using the 1-2-4-6 model is shown for ClpAE286A (Figure S2) and ClpAE565A (Figure S3). In all cases, the experimental data are well described by the model.

The stoichiometric binding constants ($L_{n,0}$) resulting from NLLS global analysis of $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$ difference curves to the 1-2-4-6 model are summarized in Table 3. Note that three replicate sets of 6, 10, and 20 μM protein were subjected to global analysis for each variant. Consequently, the uncertainties reported in Table 3 for each parameter represent the standard deviation from three global fits. Table 3 shows that all proteins examined, have similar dimerization equilibrium constants of $L_{2,0}$ ~ (1.1 – 3.4) x 10^4 M^{-1}. The resulting $L_{4,0}$ for each protein also narrowly range within an order of magnitude difference of each other. The largest difference is found in the resulting values for $L_{6,0}$. For example, the analysis of $ClpA_{E286A/E565A}$ results in a hexamerization

binding constant of (1.1 ± 0.5) x 10^{23} M^{-5}. This is about an order of magnitude different from the resulting $L_{6,0}$ for $ClpA_{E565A}$ of (1.4 ± 0.5) x 10^{24} M^{-5}. In fact, the binding constants for the formation of dimers, tetramers, and hexamers is larger for the $ClpA_{E565A}$ variant, than either of the other two variants ($ClpA_{E286A}$ and $ClpA_{E286A/E565A}$).

To understand the impact of these differences, the equilibrium constants shown in Table 3 were used to model the weight average sedimentation coefficients as a function of the concentration of free ClpA ($[ClpA]_{free}$), as done previously (21, 30). The dependence of the weight average sedimentation coefficient on $[ClpA]_{free}$ is described by Eq. (10) where s_n is the sedimentation coefficient of each ClpA *n*-mer (Table 1). The total ClpA ($[ClpA]_{total}$) concentration is related to $[ClpA]_{free}$ by the conservation of mass expression shown in Eq. (11). The $\bar{s}$ data shown in Figure 1 D were subjected to NLLS analysis by implicitly fitting to Eq. (10) and Eq. (11) using Micromath Scientist 3.0 (Micromath, St. Louis, MO).

$$\bar{s} = \frac{s_1 + 2s_2L_{2,0}[ClpA_f] + 4s_4L_{4,0}[ClpA_f]^3 + 6s_6L_{6,0}[ClpA_f]^5}{\left(1 + 2L_{2,0}[ClpA_f] + 4L_{4,0}[ClpA_f]^3 + 6L_{6,0}[ClpA_f]^5\right)} \tag{10}$$

$$ClpA_t = ClpA_f + 2L_{2,0}[ClpA_f]^2 + 4L_{4,0}[ClpA_f]^4 + 6L_{6,0}[ClpA_f]^6 \tag{11}$$

Figure 1 D shows the resulting binding isotherms (solid traces) overlaid with the $\bar{s}$ values obtained experimentally. The binding isotherms describe the experimentally determined $\bar{s}$ values obtained at each fixed $[ClpA]_{total}$. As shown in Figure 1 D, differences in $L_{n,0}$ values (Table 3) between $ClpA_{WT}$ and each variant causes the binding curves to shift along the $[ClpA]_{total}$ axis relative to $ClpA_{WT}$. For example, $ClpA_{E286A}$ has larger $L_{2,0}$ and $L_{6,0}$ values relative to the corresponding values of $ClpA_{WT}$, and a $L_{4,0}$ that is within error of

$ClpA_{WT}$. These differences result in a binding isotherm for $ClpA_{E286A}$ (Figure 1 D, solid purple trace) that is shifted to the left on the $[ClpA]_{total}$ axis relative to $ClpA_{WT}$ (Figure 1 D, solid green trace). Consistently, the $ClpA_{E565A}$ binding isotherm is shifted further left compared to $ClpA_{WT}$ and $ClpA_{E286A}$ as a result of the larger $L_{n,0}$ values that describe $ClpA_{E565A}$ self-assembly. In all cases, the indicated Walker B mutations perturb the self-assembly energetics of $ClpA_{WT}$. Consistently, identical total concentrations of each protein will reside in a distribution of states made up of different fractions of each oligomer.

To comparatively examine the impact of the Walker B mutation on the $ClpA_{WT}$ self-assembly, Eq. (3) – (5) were used to simulate the fraction of each oligomer populated as a function of $[ClpA]_{total}$. Figure 4 shows the fractions of monomers (red), dimers (blue), tetramers (green), and hexamers (black) populated as a function of total concentration of $ClpA_{WT}$, $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$ (Figure 4, panels A – D, respectively). These species fractions were simulated using the self-association binding constants reported for each protein in Table 3. The concentration dependent oligomer fractions of $ClpA_{WT}$ shown in Figure 4 A, are also plotted in Figure 4 B – D (broken traces) for comparison.

Figure 4 C – D shows clear differences in the distribution of oligomers populated for each ClpA Walker B variant (solid traces) relative to $ClpA_{WT}$ (broken traces). In general, larger deviations between the fractions of oligomers populated occur with increasing ClpA concentrations. For example, a sample containing 15 μM $ClpA_{WT}$ would reside in a distribution made up of 63% monomers, 18 % dimers, and 19 % tetramers (Figure 4 A). In contrast, the same concentration of $ClpA_{E286A}$ contains a 6 % lower population of monomers, a 9 % greater population of dimers, a 10 % lower population of tetramers, and

a 6 % hexamer population (Figure 4 B, contrast solid and broken traces). At low concentrations (< 5 μM $ClpA_{total}$) there is less deviation in the fractions of populated oligomers for the ClpA Walker B variants relative to $ClpA_{WT}$. However, deviations in the fraction of monomers and dimers (red and blue traces, respectively) persist at all concentrations simulated.

CONCLUSIONS

ATP hydrolysis by each NBD during polypeptide substrate unfolding and degradation has been examined with steady state ATPase experiments using the ClpA Walker B variants examined here (23, 24, 36). An inherent assumption of these ensemble ATPase assays is that the entire population catalyzing ATP turnover is in the hexameric state. Specifically, the concentration of hexamers is determined by simply dividing the total protein concentration by six. The justification for doing this is the observation of ClpA Walker B variant hexamers in analytical gel filtration assays done in the presence of nucleotide. However, that does not indicate the extent to which the self-association equilibrium has been perturbed. Because ATP binding drives the self-assembly of ClpA, it stands to reason that changes in the nucleotide binding pocket may perturb the enzyme's intrinsic propensity to form hexamers in the absence of nucleotide. To test whether this is the case, we examined the self-assembly mechanism of $ClpA_{WT}$ and three ClpA Walker B variants in the absence of nucleotide (Figure 1).

Our findings show that the self-assembly of $ClpA_{WT}$ is best described by a monomer-dimer-tetramer equilibrium (Figure S1). In contrast, the self-assembly of each ClpA Walker B variant tested in the same conditions is best described by a monomer-dimer-tetramer-hexamer equilibrium (Figures 1, S2, and S3). The fact that hexamerization

must be incorporated to describe the self-association of the Walker B variants supports the hypothesis that the Walker B mutation changes the inherent propensity of the protein to oligomerize.

Our results support a model by which the hydrolysis inactivating ClpA Walker B mutation upregulates hexamerization of ClpA. Table 3 shows the results from the global analysis of $ClpA_{WT}$ difference curves using the 1-2-4 and 1-2-4-6 models. Fitting to the 1-2-4-6 model results in a worse fit and yields an unconstrained $L_{6,0}$. Despite this, we cannot preclude the existence of $ClpA_{WT}$ hexamers under the conditions tested. However, if $ClpA_{WT}$ hexamers are present, we can conclude from the analysis that they are not significantly populated to be quantified by the methods used here. In contrast, the $L_{6,0}$ that results from global analysis of each Walker B variant is about an order of magnitude larger compared to $ClpA_{WT}$. Since the only change between $ClpA_{WT}$ and each Walker B variant is insertion of the indicated single point mutations (E286A or E565A), we conclude that the observed increase in $L_{6,0}$ is a direct result of the indicated change to the protein's primary sequence.

The ClpA hexamer population is controlled by the concentration of ClpA and the chemical potential of ATP in solution, that is, the free nucleotide concentration. Given the results presented here, it is unlikely the case that nucleotide binding affinity is not also changed by primary sequence changes in the ATP binding pocket. If this were the case, the population of ClpA hexamers available to unfold polypeptide substrates for the D1 versus the D2 ClpA Walker B variant would be different. As a result, differences in the apparent ATPase activity would reflect changes in the assembly state, and not inherent differences in the ATP hydrolysis rate of each NBD ring of ClpA. Thus, a thorough examination of the

energetics of nucleotide linked assembly should be sought to correctly interpret the molecular basis of observed differences in the ATPase rates of ClpA Walker B variants.

ACKNOWLEDGEMENTS

This investigation was supported by the NSF (grant MCB-1412624 to ALL).

REFERENCES

1. Schirmer, E. C., J. R. Glover, M. A. Singer, and S. Lindquist. 1996. HSP100/Clp proteins: a common mechanism explains diverse functions. Trends in biochemical sciences 21:289-296.

2. Neuwald, A. F., L. Aravind, J. L. Spouge, and E. V. Koonin. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome research 9:27-43.

3. Ogura, T., and A. J. Wilkinson. 2001. AAA+ superfamily ATPases: common structure--diverse function. Genes Cells 6:575-597.

4. Olivares, A. O., T. A. Baker, and R. T. Sauer. 2016. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat Rev Microbiol 14:33-44.

5. Striebel, F., W. Kress, and E. Weber-Ban. 2009. Controlled destruction: AAA+ ATPases in protein degradation from bacteria to eukaryotes. Curr Opin Struct Biol 19:209-217.

6. Matyskiela, M. E., and A. Martin. 2013. Design principles of a universal protein degradation machine. Journal of molecular biology 425:199-213.

7. Barthelme, D., and R. T. Sauer. 2012. Identification of the Cdc48*20S proteasome as an ancient AAA+ proteolytic machine. Science 337:843-846.

8. Yedidi, R. S., P. Wendler, and C. Enenkel. 2017. AAA-ATPases in Protein Degradation. Front Mol Biosci 4:42.

9. Guo, F., M. R. Maurizi, L. Esser, and D. Xia. 2002. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. J Biol Chem 277:46743-46752.

10. Lee, S., M. E. Sowa, Y. H. Watanabe, P. B. Sigler, W. Chiu, M. Yoshida, and F. T. Tsai. 2003. The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell 115:229-240.

11. Carroni, M., K. B. Franke, M. Maurer, J. Jager, I. Hantke, F. Gloge, D. Linder, S. Gremer, K. Turgay, B. Bukau, and A. Mogk. 2017. Regulatory coiled-coil domains promote head-to-head assemblies of AAA+ chaperones essential for tunable activity control. Elife 6.

12. Banerjee, S., A. Bartesaghi, A. Merk, P. Rao, S. L. Bulfer, Y. Yan, N. Green, B. Mroczkowski, R. J. Neitz, P. Wipf, V. Falconieri, R. J. Deshaies, J. L. Milne, D. Huryn, M. Arkin, and S. Subramaniam. 2016. 2.3 A resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. Science 351:871-875.

13. Glynn, S. E., A. Martin, A. R. Nager, T. A. Baker, and R. T. Sauer. 2009. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell 139:744-756.

14. Sousa, M. C., C. B. Trame, H. Tsuruta, S. M. Wilbanks, V. S. Reddy, and D. B. McKay. 2000. Crystal and solution structures of an HslUV protease-chaperone complex. Cell 103:633-643.

15. Lin, C. C., S. C. Su, M. Y. Su, P. H. Liang, C. C. Feng, S. H. Wu, and C. I. Chang. 2016. Structural Insights into the Allosteric Operation of the Lon AAA+ Protease. Structure 24:667-675.

16. Langklotz, S., U. Baumann, and F. Narberhaus. 2012. Structure and function of the bacterial AAA protease FtsH. Biochimica et biophysica acta 1823:40-48.

17. Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. The EMBO journal 1:945-951.

18. Kolygo, K., N. Ranjan, W. Kress, F. Striebel, K. Hollenstein, K. Neelsen, M. Steiner, H. Summer, and E. Weber-Ban. 2009. Studying chaperone-proteases using a real-time approach based on FRET. J Struct Biol 168:267-277.

19. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425:2795-2812.

20. Seol, J. H., S. H. Baek, M. S. Kang, D. B. Ha, and C. H. Chung. 1995. Distinctive roles of the two ATP-binding sites in ClpA, the ATPase component of protease Ti in Escherichia coli. J Biol Chem 270:8087-8092.

21. Veronese, P. K., R. P. Stafford, and A. L. Lucius. 2009. The Escherichia coli ClpA Molecular Chaperone Self-Assembles into Tetramers. Biochemistry 48:9221-9233.

22. Veronese, P. K., and A. L. Lucius. 2010. Effect of Temperature on the Self-Assembly of the Escherichia coli ClpA Molecular Chaperone. Biochemistry 49:9820-9829.

23. Kress, W., H. Mutschler, and E. Weber-Ban. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. J Biol Chem 284:31441-31452.

24. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112:5377-5382.

25. Sharan, S. K., L. C. Thomason, S. G. Kuznetsov, and D. L. Court. 2009. Recombineering: a homologous recombination-based method of genetic engineering. Nat Protoc 4:206-223.

26. Zhao, H., R. Ghirlando, G. Piszczek, U. Curth, C. A. Brautigam, and P. Schuck. 2013. Recorded scan times can limit the accuracy of sedimentation coefficients in analytical ultracentrifugation. Analytical biochemistry 437:104-108.

27. Ghirlando, R., A. Balbo, G. Piszczek, P. H. Brown, M. S. Lewis, C. A. Brautigam, P. Schuck, and H. Zhao. 2013. Improving the thermal, radial, and temporal accuracy of the analytical ultracentrifuge through external references. Analytical biochemistry 440:81-95.

28. Veronese, P. K., B. Rajendar, and A. L. Lucius. 2011. Activity of Escherichia coli ClpA Bound by Nucleoside Di- and Triphosphates. Journal of molecular biology 409:333-347.

29. Laue, T. M., Shah, B.D., Ridgeway, T.M., Pelletier, S.L. 1992. Computer-aided interpretation of analytical sedimentation data for proteins. In Analytical Ultracentrifugation in Biochemistry and Polymer Science. A. J. R. S.E. Harding, J.C. Horton, editor. Royal Society of Chemistry, Cambridge.

30. Correia, J. J. 2000. Analysis of weight average sedimentation velocity data. Methods Enzymol 321:81-100.

31. Stafford, W. F., and P. J. Sherwood. 2004. Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. Biophysical chemistry 108:231-243.

32. Li, T., J. Lin, and A. L. Lucius. 2015. Examination of polypeptide substrate specificity for Escherichia coli ClpB. Proteins 83:117-134.

33. Kessel, M., M. R. Maurizi, B. Kim, E. Kocsis, B. L. Trus, S. K. Singh, and A. C. Steven. 1995. Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome. J Mol Biol 250:587-594.

34. Maurizi, M. R., S. K. Singh, M. W. Thompson, M. Kessel, and A. Ginsburg. 1998. Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. Biochemistry 37:7778-7786.

35. Maurizi, M. R. 1991. ATP-promoted interaction between Clp A and Clp P in activation of Clp protease from Escherichia coli. Biochem Soc Trans 19:719-723.

36. Olivares, A. O., A. R. Nager, O. Iosefson, R. T. Sauer, and T. A. Baker. 2014. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nature structural & molecular biology.

Table 1

Sedimentation coefficient and MW for ClpA oligomers used in global analysis of sedimentation velocity data

		WT	E286A & E565A	E286A/E565A
n	s (S)	MW (Da)	MW (Da)	MW (Da)
1	2.8	84206	84149	84091
2	4.4	168412	168298	168182
4	7.1	336824	336596	336364
6	9.4	505236	504894	504546

Table 2

Comparison of models tested to globally describe ClpA assembly

Model	WT_{H500}		E286A		E565A		E286A/E565A	
	RMSD	$F_{calculated}$	RMSD	$F_{calculated}$	RMSD	$F_{calculated}$	RMSD	$F_{calculated}$
1-2, 1-4	0.02699	1.00000	0.04153	1.76582	0.04986	2.27607	0.04497	2.69402
1-2, 1-6	0.16535	37.52683	0.03331	1.13593	0.04014	1.47515	0.03359	1.50236
1-4, 1-6	0.05866	4.72289	0.05226	2.79636	0.05081	2.36347	0.04498	2.69416
1-2, 1-4, 1-6	0.02704	1.00383	0.03125	1.00000	0.03305	1.00000	0.02740	1.00000

All models represent stoichiometric binding reactions. An $F_{calculated}$ value larger than an $F_{critical}$ = 1.0002 indicates a significantly worse fit compared to the best fit.

Table 3

$L_{n,0}$ results from global analysis of difference curves

1-2-4	$L_{2,0}$ (M^{-1})	$L_{4,0}$ (M^{-3})
WT	$(1.51 \pm 0.05) \times 10^4$	$(8.9 \pm 0.9) \times 10^{13}$

1-2-4-6	$L_{2,0}$ (M^{-1})	$L_{4,0}$ (M^{-3})	$L_{6,0}$ (M^{-5})
WT	$(1.7 \pm 0.3) \times 10^4$	$(8.4 \pm 0.4) \times 10^{13}$	$(4 \pm 5) \times 10^{22}$
E286A	$(2.8 \pm 0.3) \times 10^4$	$(7 \pm 1) \times 10^{13}$	$(4 \pm 1) \times 10^{23}$
E565A	$(3.4 \pm 0.4) \times 10^4$	$(2.4 \pm 0.3) \times 10^{14}$	$(1.4 \pm 0.5) \times 10^{24}$
E286A/E565A	$(1.1 \pm 0.4) \times 10^4$	$(4 \pm 1) \times 10^{13}$	$(1.1 \pm 0.5) \times 10^{23}$

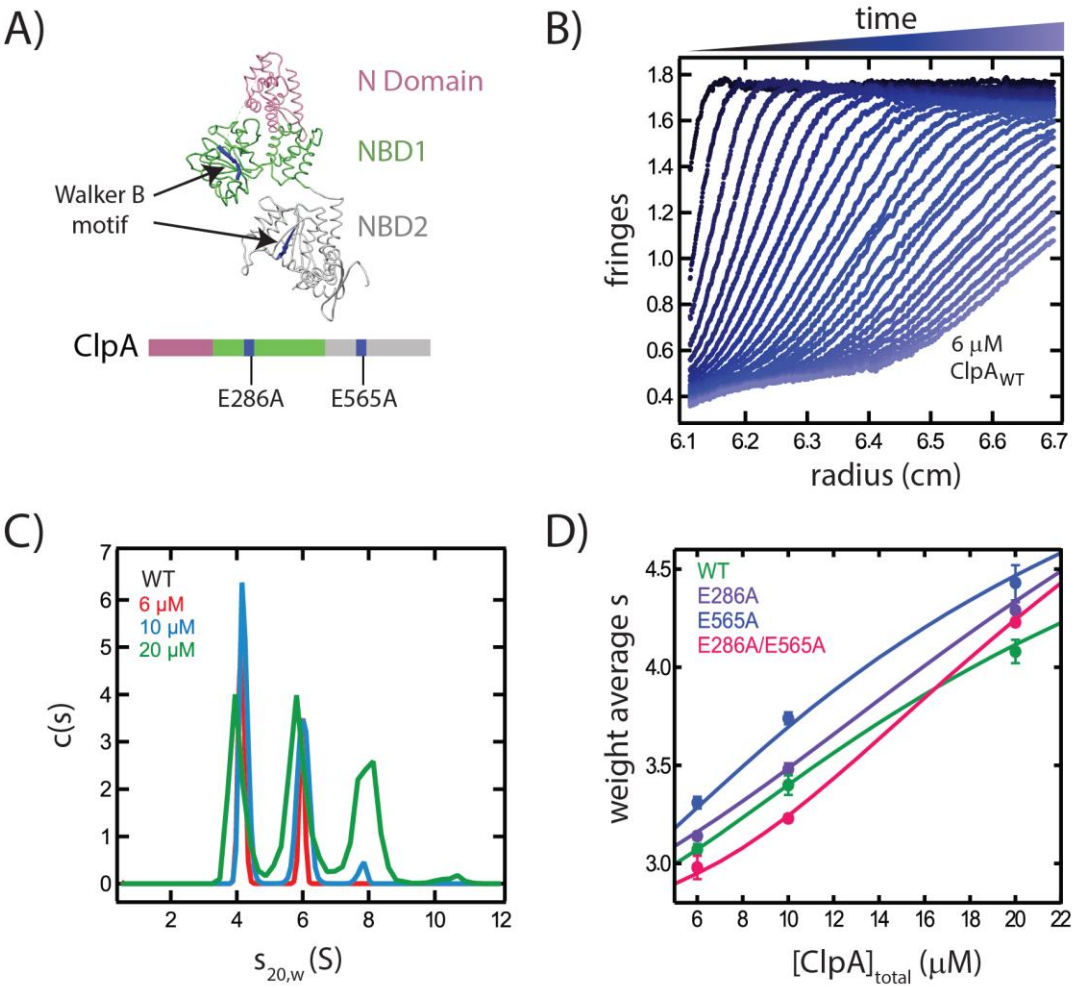


Figure 1. Sedimentation velocity experiments of ClpA$_{WT}$ and ClpA Walker B variants. (A) ATP hydrolysis occurs in the Walker B motif within each ClpA nucleotide binding domain (NBD) shown in green and grey in the ClpA monomer crystal structure (PDB ID: 1ksf). Three ATP hydrolysis inactive ClpA variants were constructed containing an alanine replacement for the Walker B glutamate in NBD1 (ClpA$_{E286A}$), NBD2 (ClpA$_{E565A}$), or both NBDs (ClpA$_{E286A/E565A}$). Sedimentation velocity experiments monitoring fringe displacement as a function of radial position and time were collected for 6 (B), 10, and 20 μM ClpA$_{WT}$, ClpA$_{E286A}$, ClpA$_{E565A}$, and ClpA$_{E286A/E565A}$. The resulting sedimentation velocity data was subjected to c(s) analysis. (C) An overlay of the c(s) distributions for various concentrations of ClpA$_{WT}$ shows a concentration dependent distribution of oligomers. Weight average sedimentation coefficients were obtained by integrating c(s) distributions across the entire c(s) range shown. (D) The resulting weight average sedimentation coefficients for ClpA$_{WT}$ (green), ClpA$_{E286A}$ (purple), ClpA$_{E565A}$ (blue) and ClpA$_{E286A/E565A}$ (pink) are plotted as a function of protein concentration. The solid traces represent simulation of weight average sedimentation coefficient using Eq. (10) - (11) and the equilibrium constants in Table 3 for each protein, respectively.

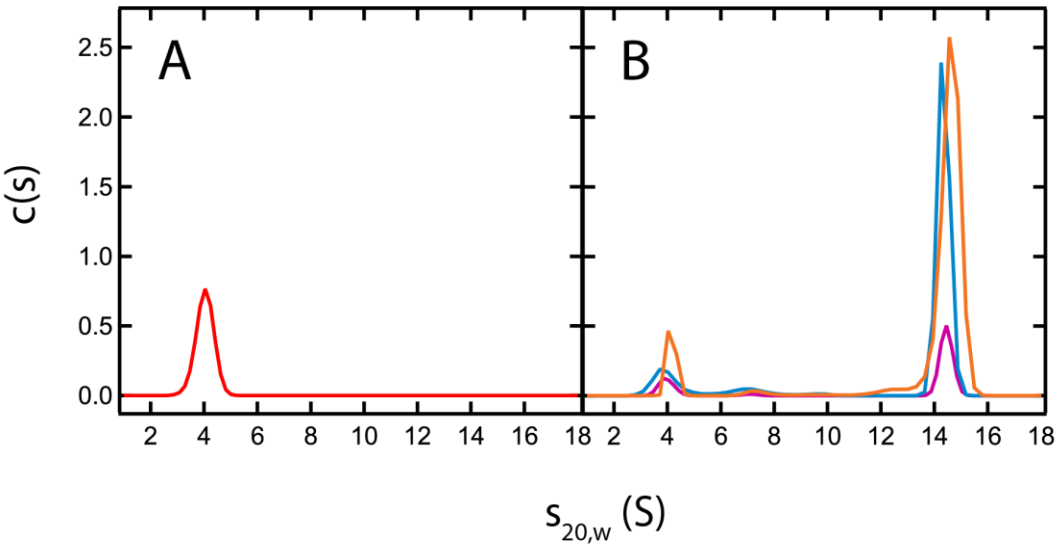


Figure 2. C(s) analysis of ClpA$_{E286A/E565A}$ in the absence and presence of 1 mM ATP. The c(s) analysis from sedimentation velocity experiments performed on (A) 2 µM ClpA$_{E286A/E565A}$ in the absence of ATP and (B) on 2 µM (pink), 6 µM (blue), and 10 µM (orange) ClpA$_{E286A/E565A}$ in the presence of 1 mM ATP.

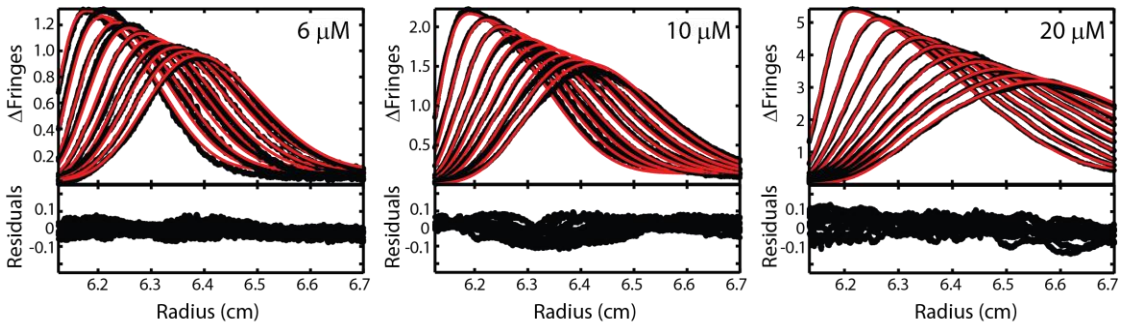


Figure 3. Global analysis of difference curves from sedimentation velocity data of 6, 10, and 20 μM ClpA$_{E286A/E565A}$ using the 1-2-4-6 model. Sedimentation velocity boundaries collected every 30 seconds for the first 3 hours of sedimentation were included in the fit. Every 16th difference curve, Δ fringes as a function of radial position (cm), is shown. The data are shown in black and the fit is shown in red. The fitting RMSD is 0.0274. Fitting residuals randomly distributed about zero, plotted below the difference curves for each concentration, indicate the data are well described by the model.

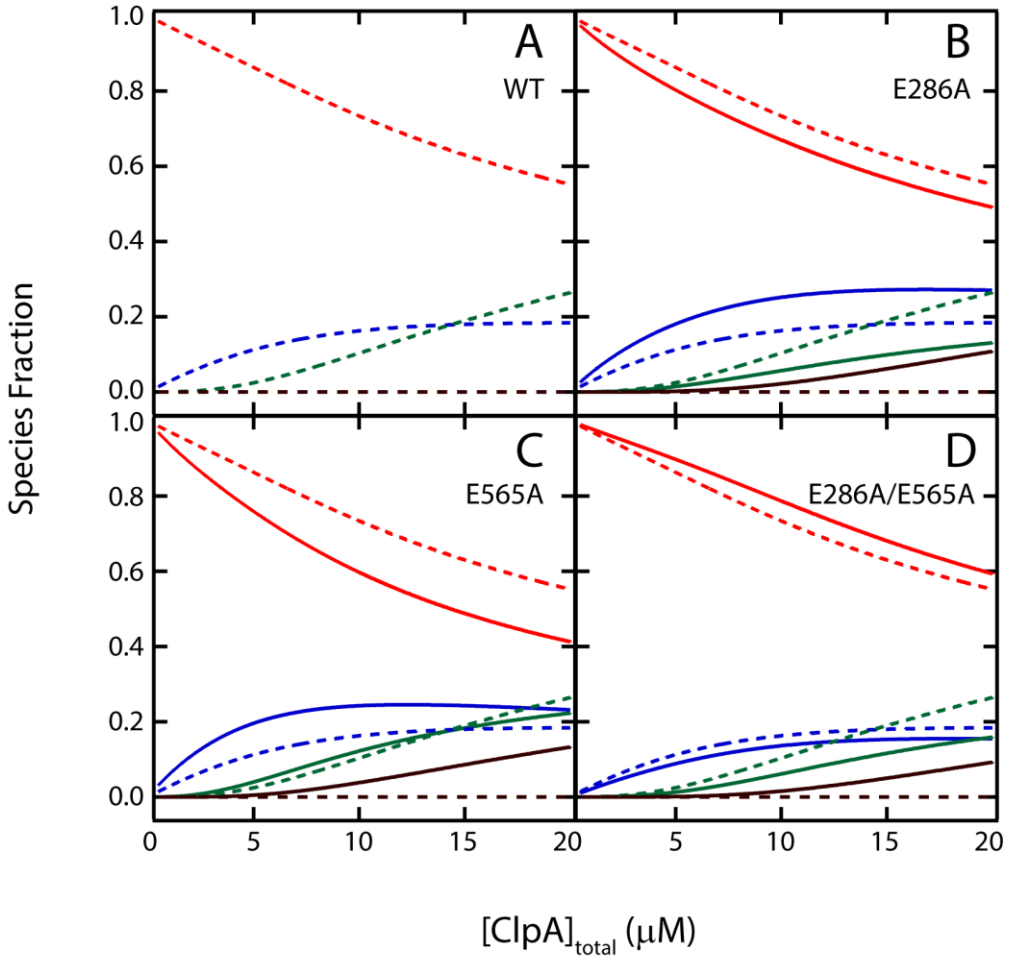
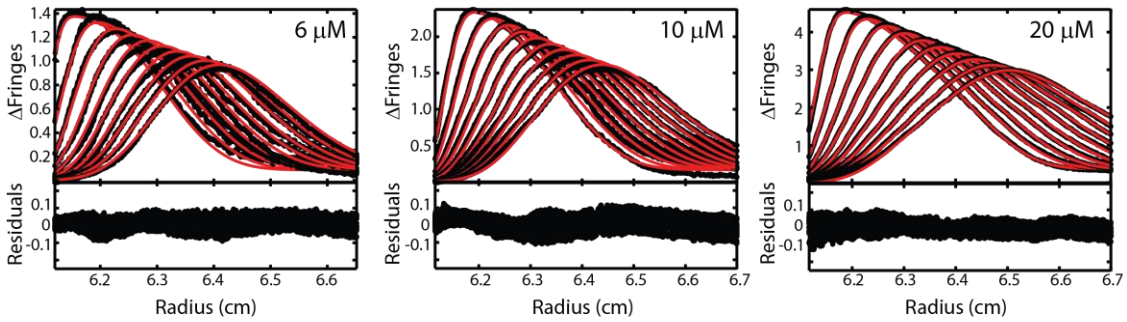
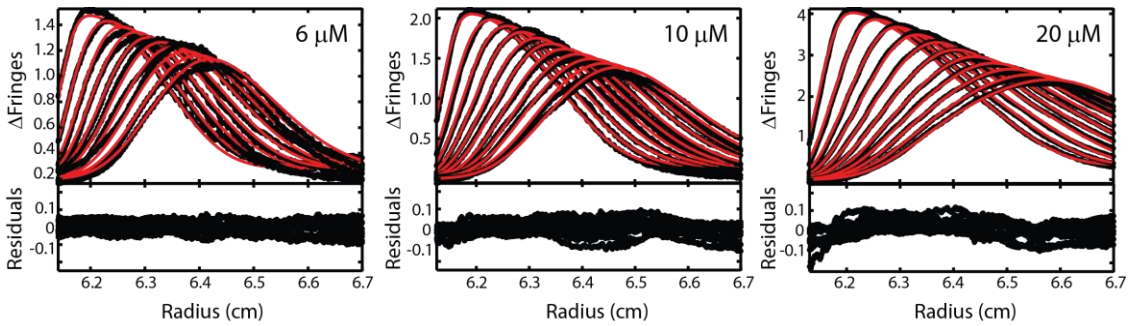


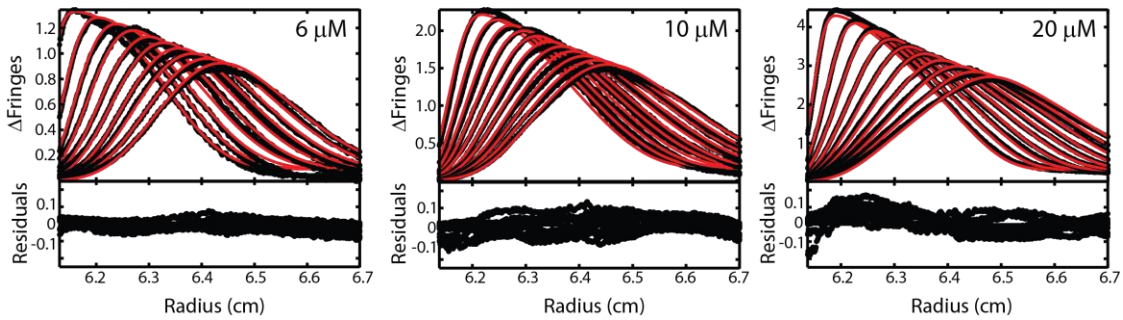
Figure 1: Species fraction simulation of $ClpA_{WT}$, $ClpA_{E286A}$, $ClpA_{E565A}$, and $ClpA_{E286A/E565A}$ as a function of total protein concentration. Species fractions were simulated using the $L_{n,0}$ parameters shown in Table 3. Species fractions for the oligomers of each protein were simulated as a function of total ClpA concentrations. Fractions of monomers (red), dimers (blue), tetramers (green), and hexamers (black) are shown as a function of total (A) [$ClpA_{WT}$], (B) [$ClpA_{E286A}$], (C) [$ClpA_{E565A}$], and (D) [$ClpA_{E286A/E565A}$] (μM). The species fraction simulations for $ClpA_{WT}$ shown in panel A (broken traces) are also shown in panels B – D to compare differences in the populated distribution of oligomers between $ClpA_{WT}$ and the indicated variant.



Supplementary Figure S1. Global analysis of difference curves from sedimentation velocity data of 6, 10, and 20 μM $ClpA_{WT}$ using the 1-2-4 model. Sedimentation velocity boundaries collected every 30 seconds for the first 3 hours of sedimentation were included in the fit. Every 16th difference curve, Δfringes as a function of radial position (cm), is shown. The data are shown in black and the fit is shown in red. The fitting RMSD is 0.02699. Fitting residuals randomly distributed about zero, plotted below the difference curves for each concentration, indicate the data are well described by the model.



Supplementary Figure S2: Global analysis of difference curves from sedimentation velocity data of 6, 10, and 20 μM ClpA$_{E286A}$ using the 1-2-4-6 model. Sedimentation velocity boundaries collected every 30 seconds for the first 3 hours of sedimentation were included in the fit. Every 16th difference curve, Δ fringes as a function of radial position (cm), is shown. The data are shown in black and the fit is shown in red. The fitting RMSD is 0.03125. Fitting residuals randomly distributed about zero, plotted below the difference curves for each concentration, indicate the data are well described by the model.



Supplementary Figure S3: Global analysis of difference curves from sedimentation velocity data of 6, 10, and 20 μM ClpA$_{E565A}$ using the 1-2-4-6 model. Sedimentation velocity boundaries collected every 30 seconds for the first 3 hours of sedimentation were included in the fit. Every 16th difference curve, Δ fringes as a function of radial position (cm), is shown. The data are shown in black and the fit is shown in red. The fitting RMSD is 0.03305. Fitting residuals randomly distributed about zero, plotted below the difference curves for each concentration, indicate the data are well described by the model.

CHAPTER 4

EXAMINATION OF THE NUCLEOTIDE LINKED ASSEMBLY MECHANISM OF *ESCHERICHIA COLI* CLPA

by

ELIZABETH C. DURAN AND AARON L. LUCIUS

In preparation to *Biophysical Journal*

Format adapted for dissertation

ABSTRACT

Energy-dependent molecular chaperones are ubiquitous molecular machines involved in protein degradation and remodeling processes essential to cellular vitality. *E. coli* ClpA is a AAA+ (ATPase Associated with diverse cellular Activities) chaperone that catalyzes the ATP-dependent unfolding and translocation of substrate proteins targeted for degradation into a protease, ClpP. ClpA hexamers associate with one or both ends of ClpP tetradecamers to form either 1:1 or 2:1 ClpAP complexes based on the relative concentrations of each component. ClpA contains two nucleotide binding sites per monomer, NBD1 and NBD2, and self-assembly into hexamers is thermodynamically linked to nucleotide binding. Although extensive solution studies have been done on ClpAP catalyzed substrate degradation, the field is currently unable to quantitatively predict the concentration of ClpA hexamers available to interact with ClpP for any given nucleotide and total ClpA concentration. In this work, sedimentation velocity studies are used to quantitatively examine the self-assembly of a ClpA Walker B variant in the presence of ATP. In addition to the hexamerization, we observe the formation of previously unreported ClpA dodecamer in the presence of ATP. Apparent equilibrium constants for the formation of each ClpA oligomer was obtained from direct boundary modeling of the sedimentation velocity data. The energetics of nucleotide binding to NBD1 and NBD2 are revealed by examining the dependence of the apparent association equilibrium constants on free nucleotide concentration. This analysis is the first step in a detailed quantitative understanding of how the twelve nucleotide binding and hydrolyses sites within the hexameric ring coordinate ATP hydrolysis and coupling to polypeptide translocation.

INTRODUCTION

ATP-dependent proteases play critical and diverse roles in the proteome maintenance of prokaryotes and eukaryotes. Diverse cellular pathways such as heat shock response (1), post-translational protein triaging (2), and regulation of gram positive bacteria pathogenesis (3), depend on the activity of ATP-dependent chaperone systems. These processes involve the coordination of a AAA+ (ATPases Associated with diverse cellular Activities) molecular chaperone and a compartmentalized protease. The AAA+ chaperones uses cycles of ATP binding and hydrolysis to unfold targeted protein substrates and translocate them into a proteolytic compartment of the associated protease for degradation (4-6).

E. coli ClpA is the AAA+ molecular chaperone component of the bacterial energy dependent chaperone system, ClpAP. It is sub-classified as a Class I Hsp100/Clp molecular chaperone due to the presence of two nucleotide binding domains (NBDs) within each ClpA protomer **(7)**. Each ATP binding pocket of each protomer contains canonical Walker A and B motifs, where ATP binding and hydrolysis occur (8). ATP binding has been observed to drive the formation of ClpA hexamers, the functional oligomer of ClpA capable of interacting with polypeptide substrate and ClpP (9, 10). Structurally, ClpA hexamers have been modeled to be planar rings containing an axial channel where polypeptide substrates are translocated through for unfolding (8, 11).

ATP binding and hydrolysis by ClpA is coupled to both oligomeric assembly and the translocation of polypeptide substrates through the axial channel of the hexameric ring. In the presence of ATP, Kress and co-workers showed that ATP binding but not hydrolysis drives the oligomerization of ClpA into hexamers ($ClpA_6$) capable of binding polypeptide

substrate and tetradecameric ClpP ($ClpP_{14}$) **(9)**. In addition to driving oligomerization, ATP binding and hydrolysis drives the processive translocation of polypeptide substrates down the axial channel of ClpA hexamers (12). Repeated rounds of ATP binding and hydrolysis are proposed to modulate the conformational switching of a loop that connects the Walker A and Walker B motifs within each NBD that extends into the axial channel of ClpA hexamers (13). Using crosslinking studies Hinnerwisch and co-workers observe the NBD2 axial loops to be in direct contact with a polypeptide substrate and find that translocation activity can be abolished by mutations in the NBD1 axial loops. Despite structural understanding of the translocation mechanism, how nucleotide is hydrolysis is coordinated to substrate unfolding within the 12 NBDs of a hexamer remains poorly understood.

ClpA hexamers catalyze the unfolding and translocation of polypeptide substrates in the absence and presence of ClpP (14, 15). Multiple studies have shown that ClpP has an effect on the translocation rate of ClpA. Using a single turnover polypeptide translocation assay we observe that in the absence of ClpP, ClpA translocates polypeptide substrates at a rate of 20 aa s^{-1} (12). In contrast, when ClpP is present, we observe that ClpA translocates polypeptide substrates with a faster overall rate of 36 aa s^{-1} (16). Using a FRET assay to monitor ClpA catalyzed polypeptide unfolding, Baytshtok et. al. similarly observed differences for the catalytic rate of unfolding by ClpA in the absence and presence of ClpP.

The molecular basis for the observations of ClpP modulating ClpA catalytic activities is not well understood. One proposed model for these observed differences is that ClpP changes the rate of the physical up and down movements of the ClpA axial loops. However, we hypothesize that an alternative and not necessarily mutually exclusive,

explanation is that ClpP perturbs the ClpA assembly state. Changes to the binding affinity or hydrolysis rate of ATP at each ClpA NBD could lead to differences in the population of ClpA hexamers available to catalyze polypeptide translocation in the absence or presence of ClpP.

We have shown that in the absence of ATP, ClpA resides in distribution of monomers, dimers, and tetramers (17, 18). In addition, we observe that mutations in the binding pocket of ClpA perturb the self-association equilibrium constants for oligomerization in the absence of nucleotide (E.C. Duran, in submission). This leads to differences in the populations of oligomers present at equivalent concentrations of the ClpA variants examined. Given these observations, it's possible that interaction between ClpA and ClpP, could lead to changes in either the protein-protein interaction between ClpA oligomers or the binding affinity of ATP at each NBD that perturb the ClpA assembly state. Testing whether ClpP modulates the macromolecular assembly state of ClpA requires the deconvolution of the affinity constants for the formation of hexamers and the affinity of ATP binding to the hexamer in the absence of ClpP.

Having recently examined the self-association of ClpA in the absence ATP, in this study we sought to quantify the macromolecular assembly of ClpA in the presence of ATP. Surprisingly, we observe that the hexamer is not the highest order oligomer populated in the presence of nucleotide. We find that $ClpA_{E286A/E565A}$ resides in a distribution of monomers, dimers, tetramers, hexamers, and dodecamers in the presence of ATP. By quantifying the self-assembly equilibrium constants for the formation of each oligomer in the presence of varying ATP concentrations, we constructed binding isotherms that allowed us to separate ATP binding affinity from macromolecular assembly. The binding

isotherms for the formation of each oligomer as a function of ATP concentration were analyzed to determine the affinity and stoichiometry for ATP binding to each oligomer. We report binding constants for ATP binding each ClpA oligomer and observe that on average ClpA monomers, dimers, tetramers, hexamers, and dodecamers bind 1, 3, 6, 10, and 20 ATP molecules, respectively.

MATERIALS AND METHODS

Buffers and reagents

All buffers used in the experiments described were prepared using reagent grade materials and ultrapure water obtained from a Purelab Ultra System (Evoqua Water Technologies). The ATP used was purchased from Sigma-Aldrich. All experiments with purified protein were carried out in buffer H500 containing 25 mM HEPES, pH 7.5 at 25 °C, 10 mM $MgCl_2$, 2 mM 2-mercapthoethanol, 500 mM NaCl, and 10 % (v/v) glycerol.

Protein isolation

$ClpA_{E286A/E565A}$ was overexpressed in *E. coli* $\Delta ClpA_{WT}$-BL21(DE3) cells as previously described (Duran, et. al. in submission). After overexpression, the protein was isolated following the same isolation protocol as that used to isolated wild type ClpA, previously described (17). Purified stocks of $ClpA_{E286A/E565A}$ were in buffer supplemented 25 mM Tris base (pH 7.5 at 4 °C), 100 µM EDTA, 2 mM 2-mercaptoethanol, 1 M NaCl and 50% (v/v) glycerol for long term storage at -80 °C. Prior to performing experiments, protein samples were dialyzed against buffer H500 at 4 °C. After dialysis, protein concentration was determined spectroscopically using a molar extinction coefficient of 31,000 M^{-1} cm^{-1} at 280 nm.

Equilibrium dialysis

Equilibrium dialysis experiments were conducted to gradually equilibrate $ClpA_{E286A/E565A}$ samples with ATP. The experiments were carried out using a 3-chamber micro-equilibrium dialyzer system purchased from The Nest Group. In the 3-chamber system, a central chamber is separated from two adjacent chambers by a dialysis membrane. The central chamber was loaded with 500 µL of the protein solution and separated from the other two chambers with a 25 KDa molecular weight cut off (MWCO) cellulose membrane. Both external chambers were supplemented with 500 µL aliquots of an ATP solution. In this experimental design, the central chamber containing $ClpA_{E286A/E565A}$ was permeable to the exchange of ATP. On the other hand, ClpAE286A/E565A, with a molecular weight of ~84,100 g/mol per protomer, was kept within the central chamber by the 25 KDa MWCO dialysis membrane.

After protein and ATP solutions were placed in equilibrium dialyzers, the system was left to equilibrate at 25 °C with gentle shaking at 50 rpm. After a 14-hour incubation, the samples were removed from each chamber. The solutions taken from each external chamber of a given equilibrium dialyzer were combined and its ATP concentration was determined spectroscopically. This was taken to be the concentration of ATP_{free} for a given experiment. Similarly, a spectroscopic measurement of the protein and nucleotide solution recovered from the central chamber was used to determine the concentration of ATP_{total} (ATP_{free} + $ATP_{ClpA\text{-}bound}$). The central chamber solutions containing ATP and protein were loaded into the sample sector of a dual sector analytical ultracentrifugation cell assembly. The matched ATP_{free} solution for a given sample solution was loaded into the reference sector.

Sedimentation velocity experiments

Interference sedimentation velocity experiments were collected for 6 or 10 μM $ClpA_{E286A/E565A}$ in the presence of various ATP_{free} concentrations. Importantly, each sedimentation velocity experiment was carried out on one fixed protein and nucleotide concentration. Experiments were collected in a ProteomeLab XL-I analytical ultracentrifuge (Beckman Coulter, Brea, CA). Protein and reference solutions were loaded into the respective sector of a dual channel, Epon charcoal filled centerpiece. Approximately 390 μL of either protein or reference solution was loaded into each sector. Interference boundaries were collected every 30 seconds at an angular velocity of 40,000 rpm for the duration of sedimentation as previously described.

Absorbance sedimentation velocity experiments were collected for 6 μM ClpAE286A/E565A in the presence of 30 μM TNP-ATP. Boundaries were collected monitoring absorbance at 408 nm as a function of radial position and time. Samples were subjected to an angular velocity of 40,000 rpm and absorbance boundaries were collected every 4 minutes.

Time corrected sedimentation velocity was subjected to c(s) analysis in SedFit (Peter Schuck, NIH) or difference curve analysis SedAnal, as previously described (19-21). Sedimentation coefficient values expressed in terms of $s_{20,w}$, reflect standard conditions of infinite dilution in water at 20 °C as defined in Eq. (1),

$$s_{20,w} = \frac{\left(1-\rho_{20,w}\bar{v}\right)}{\left(1-\rho\bar{v}\right)} \cdot \frac{\eta}{\eta_{20,w}} \cdot s \qquad (1)$$

where ρ indicates the solution density, η the solution viscosity, and $\bar{v}$ the partial specific volume of the protein. The partial specific volume of $ClpA_{E286A/E565A}$ was determined from its primary protein sequence to be 0.7405 mL g^{-1} using Sednterp (David Hayes, Magdalen College, Tom Laue, University of New Hampshire, and John Philo, Alliance Protein Laboratories) (22). A correction of 0.0033 mL g^{-1} was subsequently added to account for the presence of 10 % (v/v) glycerol included in the experimental buffer (H500) (18, 23). From Eq. (1), a correction factor of 1.501 is obtained to calculate $s_{20,w}$ from uncorrected sedimentation coefficient values.

Analysis of sedimentation velocity time difference curves

Sets of time difference curves collected at fixed protein and ATP concentrations were model in SedAnal. Protein self-association was modeled using the monomer-dimer-tetramer-hexamer-dodecamer stepwise assembly model (1-2-4-6-12 model) shown in Eq. (2) – (5),

$$2\{ClpA_1\} \underset{k_{r2}}{\overset{k_{f2}}{\rightleftharpoons}} \{ClpA_2\} \tag{2}$$

$$2\{ClpA_2\} \underset{k_{r4}}{\overset{k_{f4}}{\rightleftharpoons}} \{ClpA_4\} \tag{3}$$

$$\{ClpA_2\} + \{ClpA_4\} \underset{k_{r6}}{\overset{k_{f6}}{\rightleftharpoons}} \{ClpA_6\} \tag{4}$$

$$2\{ClpA_6\} \underset{k_{r12}}{\overset{k_{f12}}{\rightleftharpoons}} \{ClpA_{12}\} \tag{5}$$

where the curly bracket notation represents all ATP ligation states of a ClpA *n*-mer, while $k_{f,n}$ and $k_{r,n}$ represent the forward and reverse rate constant for *n*-mer formation, respectively. The forward and reverse rate constants are related to the stepwise equilibrium

constant (K_n) as shown in Eq. (6). Difference curves were analyzed by floating either both K_n and $k_{r,n}$ or by floating K_n while constraining $k_{r,n}$.

$$K_n = \frac{k_{f,n}}{k_{r,n}} \tag{6}$$

Sedimentation coefficients were constrained in the analysis of difference curves. Sedimentation coefficient values were previously obtained for the monomer (2.8 S), dimer (4.4 S), tetramer (7.1 S), and hexamer (9.4 S) as described (Duran, et. al., in submission). The sedimentation coefficient for the dodecamer was approximated to be 14.4 S ($s_{20,w}$ = 21.6 S) using Eq. (7) as previously described (24). Note that this prediction is in agreement with the ~22 S population observed in c(s) distributions obtained in for 6 μM $ClpA_{E286A/E565A}$ in the presence of ATP (Figure 1, blue and purple traces). Importantly, the sedimentation coefficients constrained in the analysis of difference curve were not standardized to $s_{20,w}$ values.

$$s_n = s_1\left(n^{2/3}\right) \tag{7}$$

RESULTS

$ClpA_{E286A/E565A}$ exhibits dynamic assembly in the presence of ATP

We have previously shown that ATPγS binding will drive the formation of ClpA hexamers competent for binding to peptide (23). Although AMPPNP will drive the formation of hexamers, we found that substantially higher concentrations compared to ATPγS were required to do so, indicating weaker binding of AMPPNP. Moreover, AMPPCP, ADP, and ADP-BeF did not drive the formation of hexamers with polypeptide binding activity. With these results in mind, we initiated a study to examine the ligand

linked assembly equilibrium using ATPγS and wild-type ClpA (data not shown). However, due to slow ATPγS hydrolysis these experiments are both irreversible and poorly reproducible. To overcome this problem, we chose to examine the ligand linked assembly process using $ClpA_{E286A/E565A}$, a Walker B variant that binds but does not hydrolyze ATP. Since this ClpA variant does not hydrolyze ATP we initiated an examination of the ligand linked assembly process using hydrolysable ATP.

We recently examined the self-assembly mechanism and hydrodynamic properties of $ClpA_{E286A/E565A}$ in the absence of nucleotide. We observe that in the absence of nucleotide, $ClpA_{E286A/E565A}$ resides in a distribution of monomers, dimers, tetramers, and hexamers (E.C. Duran, in submission). As a next step, here we sought to quantify the ATP-linked self-association reaction of $ClpA_{E286A/E565A}$ using sedimentation velocity experiments.

Sedimentation velocity experiments were performed on 6 μM $ClpA_{E286A/E565A}$ in the absence of ATP or in the presence of 500 μM or 1 mM ATP. Figure 1 A shows a representative set of sedimentation boundaries collected for 6 μM $ClpA_{E286E/E565A}$ in the presence of 1 mM ATP monitoring fringe displacement as a function of radial position and time. The sedimentation boundaries collected in the absence and presence of ATP were subjected to c(s) analysis in SedFit (Peter Schuck, NIH). An overlay of the resulting c(s) distributions are shown in Figure 1 B.

In the absence of nucleotide (Figure 1 B, black), the c(s) distribution of 6 μM $ClpA_{E286A/E565A}$ is dominated by a single species at $s_{20,w}$ = (4.1 ± 0.2) S. In our recent examination of the hydrodynamic properties of ClpA Walker B variants, we observe a sedimentation coefficient for $s_{20,w}$ = (4.3 ± 0.2) S for $ClpA_{E286A/E565A}$ monomers. Consistent

with those findings, Figure 1 B shows that in the absence of nucleotide, the distribution of 6 μM $ClpA_{E286A/E565A}$ is largely in the monomeric state.

When ATP is added to 6 μM $ClpA_{E286A/E565A}$, Figure 1 B (blue and purple traces) there is a clear shift in the c(s) distribution towards larger sedimentation coefficients. To examine the impact of nucleotide addition on the c(s) distribution of 6 μM $ClpA_{E286A/E565A}$, weight average sedimentation coefficients ($\bar{s}_{20,w}$) were obtained by integrating the c(s) distributions shown in Figure 1 B over the entire sedimentation coefficient range shown. In the absence of nucleotide a $\bar{s}_{20,w}$ ~ 4.5 S is obtained. In contrast, when ATP is present (Figure 1 B, blue and purple traces), the weight average sedimentation coefficient shifts to $\bar{s}_{20,w}$ ~ 13.1 S and 14.9 S for 500 μM ATP and 1 mM ATP, respectively. This is a clear indication that ATP binding drives the assembly of $ClpA_{E286A/E565A}$ towards higher order oligomers.

Further inspection of Figure 1 B shows that in the presence of ATP, the distributions of 6 μM $ClpA_{E286A/E565A}$ contain peaks spanning a range of $s_{20,w}$ ~ 4 – 22 S. As expected, the peak at ~4.1 S consistent with ClpA monomers, decreases when ATP is present, relative to the c(s) distribution obtained in the absence of ATP. In addition, a broad reaction boundary at $s_{20,w}$ ~ 6 – 11 S and a peak at ~15 S emerge. The broad reaction boundary is consistent with the previously reported sedimentation coefficients of $ClpA_{E286A/E565A}$ dimers and tetramers. This likely indicates the presence of a mixture of dimers and tetramers at both ATP concentrations tested. Similarly, the ~15 S peak is consistent with our previously observed sedimentation coefficient of 15.5 S for peptide-bound wild type ClpA ($ClpA_{WT}$) hexamers in the presence of ATPγS (23). These data support previous observations that ATP binding drives the hexamerization of ClpA. Yet we additionally

observe that oligomers smaller than the hexamer persist at thermodynamic equilibrium under conditions where ATP is in large excess over the macromolecule concentration. Thus quantifying the population of hexamers present at any given nucleotide concentration requires examination of the nucleotide linked assembly equilibrium, as previously done for the closely related AAA+ molecular chaperone, ClpB (25).

Surprisingly, the hexamer is not the largest oligomer observed in the c(s) distributions determined for ATP-bound $ClpA_{E286A/E565A}$ shown in Figure 1 B. In addition to the features described above, there is a peak that emerges at $s_{20,w}$ ~22.5 S. A similar 23 S species has been observed before from absorbance sedimentation velocity experiments performed on $ClpA_{WT}$ in the presence of ADP (23). Although the authors speculated this to be a ClpA dodecamer, there were not enough constraints in the absorbance sedimentation velocity data and sedimentation equilibrium data collected to model dodecamer formation. In absorbance sedimentation velocity experiments, sedimenting boundaries are collected on the order of minutes. As a result, boundary spreading due to the movement of sedimenting oligomers on that timescale leads to lower resolution data relative to interference boundaries. Thus, it's possible that with greater temporal resolution in the sedimenting boundaries collected here (see Materials and Methods) the data contain enough information to characterize the energetics of dodecamerization.

ClpAE286A/E565A resides in a monomer-dimer-tetramer-hexamer-dodecamer equilibrium in the presence of ATP

To investigate the ATP-linked assembly mechanism of $ClpA_{E286A/E565A}$, we first sought to determine the ATP concentration dependence of the formation of each ClpA oligomer in the assembly pathway. To do this, sedimentation velocity experiments were

collected for 6 and 10 µM $ClpA_{E286A/E565A}$ at fixed nucleotide concentrations ranging from approximately 20 µM to 1 mM ATP. The resulting sedimentation boundaries were directly modeled using the time difference curve method (see Materials and Methods). From this analysis, apparent equilibrium constants ($L_{n,app}$) were obtained for the formation of each $ClpA_{E286A/E565A}$ oligomer as a function of ATP. Here $L_{n,app}$ is defined as Eq. (8)

$$L_{n,app} = L_{n,0} \frac{P_n}{(P_1)^n} \tag{8}$$

where $L_{n,0}$ represents the equilibrium constant for ClpA *n*-mer formation in the absence of nucleotide and P_n represents the partition function for ClpA *n*-mer binding ATP. Importantly, the partition function is only a function of free ligand, in this case $[ATP]_{free}$. Thus, we devised an experimental strategy to be able to measure $[ATP]_{free}$.

Prior to subjecting $ClpA_{E286A/E565A}$ to sedimentation velocity, protein samples and ATP were equilibrated using equilibrium dialysis. For each reaction, protein and nucleotide samples were placed in separate chambers of an equilibrium dialyzer separated by a membrane permeable to nucleotide but not protein as described (see Materials and Methods). After a 14 hour incubation at 25 °C, the protein chamber contains the total concentration of ATP (ATP_{total}) which is equal to the sum of the concentration of free and ClpA-bound ATP. Since all chambers are permeable to ATP, the chambers impermeable to protein contain equilibrium concentrations of ATP_{free}. The contents of each chamber were separated and the $[ATP]_{free}$ was measured spectroscopically. In addition to allowing us to directly measure the $[ATP]_{free}$, gradually mixing in ATP and allowing the system to equilibrate over 14 hours was an experimentally rigorous way to achieve a reversible equilibrium.

The protein-nucleotide sample from the equilibrium dialysis experiment was then subjected to sedimentation velocity. The free ATP concentration, measured from the equilibrium dialysis experiments to be in the range of 17 µM to 1.1 mM remained constant throughout each sedimentation velocity experiment. Time difference curves were subsequently obtained from sedimenting boundaries collected at fixed protein and nucleotide concentrations. These data were subjected to NLLS analysis using a model that assumes the presence of monomers, dimers, tetramers, hexamers, and dodecamers (1-2-4-6-12 model) at thermodynamic equilibrium (Eq. (2) – (5)). Figure 2 shows representative sets of difference curves for 6 and 10 µM $ClpA_{E286A/E565A}$ in the presence of 62.3 µM and 68.5 µM ATP_{free} respectively, analyzed using the 1-2-4-6-12 model (Figure 2, green traces). For difference curves collected in the presence of ≥200 µM ATP_{free}, the measured $[ATP]_{free}$ was approximately equal $[ATP]_{total}$. In these cases, difference curves for 6 and 10 µM $ClpA_{E286A/E565A}$ at one fixed $[ATP]_{free}$ were globally analyzed using the 1-2-4-6-12 model. In contrast, sets of time difference curves collected at one fixed protein and free ATP concentrations <200 µM were individually modeled using the 1-2-4-6-12 model.

Stepwise equilibrium constants, K_2, K_4, K_6, and K_{12} were obtained at each concentration of $ATP_{free,}$ from the analysis of difference curves using the 1-2-4-6-12 model. These were used to calculate stoichiometric equilibrium constants, $L_{2,app}$, $L_{4,app}$, $L_{6,app}$, and $L_{12,app}$ using Eq. (9) – (12), where $ClpA_n$ represents a ClpA *n*-mer. Table 1 shows average $L_{n,app}$ and standard deviations from the independent analysis of three experimental replicates.

$$L_{2,app} = \frac{\{ClpA_2\}}{\{ClpA_1\}^2} = K_2 \tag{9}$$

$$L_{4,app} = \frac{\{ClpA_4\}}{\{ClpA_1\}^4} = K_2{}^2 K_4 \tag{10}$$

$$L_{6,app} = \frac{\{ClpA_6\}}{\{ClpA_1\}^6} = K_2{}^3 K_4 K_6 \tag{11}$$

$$L_{12,app} = \frac{\{ClpA_{12}\}}{\{ClpA_1\}^{12}} = K_2{}^6 K_4{}^2 K_6{}^2 K_{12} \tag{12}$$

In sedimentation velocity experiments, dissociation rates are measurable if they occur in the range of ~10^{-2} – 10^{-5} s^{-1} (26). The $L_{n,app}$ values in Table 1 were obtained assuming a dissociation rate constant of 0.01 s^{-1}. This assumes that all oligomers exhibit instantaneous dissociation on the time scale of sedimentation. To test whether oligomer dissociation occurs on the timescale of sedimentation, the kinetic rate constants for the dissociation of each oligomer, $k_{r,n}$ (Eq. (2) – (5)) were floated as fitting parameters in the analysis of difference curves.

Table S1 shows $L_{n,app}$ and $k_{r,n}$ values obtained when equilibrium and dissociation rate constants are floated parameters in the analysis of difference curves. The error reported in the table for each parameter is the standard deviation obtained from the independent analysis of three replicates. On average the reverse rate constants reported in Table S1 fall outside of the measurable range. This observation suggests that the oligomers either undergo instantaneous dissociation or are infinitely stable on the timescale of sedimentation (26).

To examine the impact of assuming $k_{r,n} \geq 0.01$ s^{-1} on $L_{n,app}$, the apparent equilibrium constants reported in Table 1 and Table S1 are plotted as a function of $[ATP]_{free}$ in Figure

S1. Figure S1 shows that the $L_{n,app}$ reported in Table 1 and Table S1 are within error. This observation suggests that there is no additional information obtained about the energetics of assembly by floating $k_{r,n}$ terms as fitting parameters. Thus all $k_{r,n}$ were constrained to 0.01 s^{-1} in the analysis of difference curves.

Model independent determination of the ClpA$_{E286A/E565A}$ ATP binding density

The ATP binding stoichiometry of each ClpA$_{E286A/E565A}$ oligomer was determined by model independent Wyman analysis (27). This strategy, has been previously employed by us to determine the ATPγS binding density of ClpB oligomers (25). As shown previously, Eq. (13) is obtained by first taking the log of Eq. (8) and subsequently taking the first derivative of the resulting expression with respect to the log([ATP]$_{free}$).

$$\frac{\partial \log\left(L_{n,app}\right)}{\partial \log\left(\left[ATP\right]_{free}\right)} = \overline{X}_n - n\overline{X}_1 \tag{13}$$

$\bar{X}_n$ represents the extent of binding to the *n*-mer. Extent of binding is defined as the ratio of the concentration of ligand bound to the concentration of total macromolecule (28). In this analysis, $\bar{X}_n$ is the extent of binding to a ClpA$_{E286A/E565A}$ oligomer ([ATP]$_{bound}$/[A$_n$]$_{total}$) and $\bar{X}_1$ represents the extent of binding to ClpA$_{E286A/E565A}$ monomer ([ATP]$_{bound}$/[A$_1$]$_{total}$). The difference expressed in Eq. (13) is equal to the slope of a plot of log($L_{n,app}$) as a function of the log($[ATP]_{free}$), known as a Wyman plot (27).

Four Wyman plots were constructed using the $L_{n,app}$ values shown in Table 1. Figure 3 (A – D) shows the linear region of each plot fit to a line. The slopes obtained from the linear regression analysis, can be expressed as the system of linear equations shown in Eq. (14) – (17).

$$\overline{X}_2 - 2\overline{X}_1 = 1.2 \pm 0.1 \quad (14)$$

$$\overline{X}_4 - 4\overline{X}_1 = 2.4 \pm 0.3 \quad (15)$$

$$\overline{X}_6 - 6\overline{X}_1 = 4.7 \pm 0.4 \quad (16)$$

$$\overline{X}_{12} - 12\overline{X}_1 = 9 \pm 1 \quad (17)$$

Equations (14) – (17) represent a system of four linear equations with five unknowns. The simplest way to solve this system of equations is to make assumptions about the monomer extent of binding and solve for the others. The reason for making assumptions about the extent of binding for the monomer is that each ClpA monomer contains two nucleotide binding sites. Consequently, there are only three possibilities for the stoichiometry of ATP binding the monomer. By constraining $\bar{X}_1$ equal to either zero, one, or two, we can easily solve Eq. (14) – (17) to determine the extent of binding to dimers, tetramers, hexamers, and dodecamers.

Table 2 shows the extent of binding to each oligomer obtained assuming a monomer extent of binding of zero, one, or two. Each ClpA oligomer contains 2*n* nucleotide binding sites. In the case of $\bar{X}_1 = 2$, the extent of binding predicted for each oligomer is greater than the number of nucleotide binding sites present in each oligomer. From this analysis we can rule out a monomer extent of binding of two. In contrast, assuming a monomer extent of binding equal to either zero or one results in extent of binding values for all other oligomers that are less than the number of nucleotide binding sites present in each oligomer. With this in mind, we devised a sedimentation velocity experiment to determine whether monomers bind ATP.

$ClpA_{E286A/E565A}$ monomers bind ATP

To determine whether ClpA monomers bind nucleotide, sedimentation velocity experiments were performed on $ClpA_{E286A/E565A}$ in the presence of TNP-ATP. Here TNP-ATP was used as a chromophore and its absorbance at 408 nm was monitored. Any peaks present in the resulting c(s) distributions reflect TNP-ATP bound species.

Figure 4 shows the c(s) distributions from 6 µM $ClpA_{E286A/E565A}$ in the presence of 30 µM TNP-ATP (red). For comparison the distribution from 6 µM $ClpA_{E286A/E565A}$ in the absence of nucleotide (Figure 1, black) is also shown in Figure 4. In the presence of TNP-ATP, Figure 4 (red) shows a distribution at ~4 S consistent with ClpA monomers. Populations at ~7.5 S and ~14.6 S are also present when TNP-ATP is added, consistent with ClpA tetramers and hexamers. The ~4 S peak overlays well with the monomer population that dominates the distribution in the absence of nucleotide (Figure 4 and Figure 1, black traces). Since the signal collected in the presence of TNP-ATP is only sensitive to absorbance from TNP-ATP, we conclude that ATP must bind monomers. Importantly, the additional peaks at ~7.5 S and ~14.6 S indicate that TNP-ATP drives assembly of $ClpA_{E286A/E565A}$ into higher order oligomers, analogous to ATP.

The results shown in Figure 4 allow us to rule out a monomer extent of binding equal to zero. A monomer extent of binding of two has also been ruled out from the model independent analysis discussed above. From data shown in Figure 4 and the results from the Wyman analysis, we conclude a monomer extent of binding equal to one. From Table 2, we obtain extent of binding values of ~3, ~6, ~10, and ~21 for ATP binding to dimers, tetramers, hexamers, and dodecamers, respectively. With estimates for the stoichiometry

of ATP binding to each oligomer in hand, we next sought to determine the average affinity of ATP binding to monomers, dimers, tetramers, hexamers, and dodecamers.

NLLS analysis of $L_{n,app}$ as a function of $[ATP]_{free}$

In order to determine the ATP binding affinity to each oligomer, binding isotherms were constructed by plotting log($L_{n,app}$) as a function of ATP_{free}. Figure 5 shows the binding isotherms for $L_{2,app}$, $L_{4,app}$, $L_{6,app}$, and $L_{12,app}$ on $[ATP]_{free}$. In order to model the binding isotherms globally, we need a partition function that describes ATP binding to the monomer, dimer, tetramer, hexamer, and dodecamer, respectively (Eq. (8)).

The simplest way to model ATP binding to each oligomer is with a *n*-independent and identical sites model. This model assumes that ATP binding to each site within an oligomer is independent of the other sites and can be described by the same ATP binding constant. Equation (18) describes the general form of the partition function for a ClpA *n*-mer binding ATP, assuming *n*-independent, identical sites.

$$P_n = \left(1 + \kappa_n \left[ATP\right]_{free}\right)^{m_n} \tag{18}$$

In Eq. (18), κ_n describes the average stepwise equilibrium constant for ATP binding a ClpA *n*-mer and m_n represents its ATP binding stoichiometry. Substitution of Eq. (18) into Eq. (8), we arrive at a general model for $L_{n,app}$ shown in Eq. (19).

$$L_{n,app} = L_{n,0} \frac{\left(1 + \kappa_n \left[ATP\right]_{free}\right)^{m_n}}{\left(\left(1 + \kappa_1 \left[ATP\right]_{free}\right)^{m_1}\right)^n} \tag{19}$$

The binding isotherms shown in Figure 5 were subjected to global NLLS analysis using the logarithmic form of Eq. (19) for each $L_{n,app}$, shown in Eq. (20) – (23).

$$\log\left(L_{2,app}\right)=\log\left(L_{2,0}\right)+\log\left(1+\kappa_2\left[ATP\right]_{free}\right)^{m_2}-2\log\left(1+\kappa_1\left[ATP\right]_{free}\right)^{m_1} \quad (20)$$

$$\log\left(L_{4,app}\right)=\log\left(L_{4,0}\right)+\log\left(1+\kappa_4\left[ATP\right]_{free}\right)^{m_4}-4\log\left(1+\kappa_1\left[ATP\right]_{free}\right)^{m_1} \quad (21)$$

$$\log\left(L_{6,app}\right)=\log\left(L_{6,0}\right)+\log\left(1+\kappa_6\left[ATP\right]_{free}\right)^{m_6}-6\log\left(1+\kappa_1\left[ATP\right]_{free}\right)^{m_1} \quad (22)$$

$$\log\left(L_{12,app}\right)=\log\left(L_{12,0}\right)+\log\left(1+\kappa_{12}\left[ATP\right]_{free}\right)^{m_{12}}-12\log\left(1+\kappa_1\left[ATP\right]_{free}\right)^{m_1} \quad (23)$$

The $L_{n,0}$ terms were constrained in the NLLS analysis of $L_{n,app}$ ATP binding isotherms. The self-association equilibrium constants, $L_{2,0}$, $L_{4,0}$, and $L_{6,0}$ were recently determined for $ClpA_{E286A/E565A}$ in the absence of nucleotide (Duran et. al., in submission). In the absence of nucleotide, we do not observe $ClpA_{E286A/E565A}$ dodecamers and therefore could not experimentally measure $L_{12,0}$. Instead we calculated an upper limit of $L_{12,0} \leq 1 \times 10^{50}$ M^{-11}. We obtained this limit by generating species fractions for the formation of each ClpA oligomer using the experimentally determined $L_{2,0}$, $L_{4,0}$, and $L_{6,0}$ while constraining $L_{12,0}$ to finite values. A $L_{12,0} = 1 \times 10^{50}$ M^{-11} was determined to be the largest value for which a dodecamer would not be populated at 20 μM $ClpA_{E286A/E565A}$, the highest concentration used to determine the energetic parameters for the self-association in the absence of nucleotide. Thus a value of $L_{12,0} = 1 \times 10^{50}$ M^{-11} was constrained in the global analysis of $L_{n,app}$ binding isotherms (Figure 5).

The binding stoichiometry terms (m_2, m_4, m_6, m_{12}) and average ATP binding constants (κ_1, κ_2, κ_4, κ_6, κ_{12}) were floated as fitting parameters in this analysis. Note that m_1 and κ_1 are global parameters, while m_2, m_4, m_6, m_{12}, κ_2, κ_4, κ_6, and κ_{12} are local parameters in the analysis. Given that we conclude a binding stoichiometry of 1 the monomer from the

results discussed above (Table 2 and Figure 4), m_1 was constrained to 1 in the NLLS analysis of the binding isotherms. When all other m_n and κ_n are floated parameters, the data are well described by the fit (Figure 4, red traces). The parameters obtained from NLLS analysis are summarized in Table 3.

The error on each fitted parameter in the analysis was determined from the standard deviation of Monte Carlos simulations (1000 iterations). From this analysis we find that m_4, m_6, m_{12}, κ_2, κ_4, κ_6, and κ_{12} are well constrained (Table 3). However, we also find the binding stoichiometry of the dimer to be unconstrained. This is not surprising given that the $L_{2,app}$ does not have a clear ATP dependence for ATP_{free} concentrations < 200 µM (see Figure 5 and Table 2). That is likely due to the low population of dimers present at low ATP concentrations.

With this in mind, the data were modeled again, constraining $m_2 = 3$, the stoichiometry for ATP binding the dimer obtained from Wyman analysis (see Table 2). The resulting fitting parameters for m_4, m_6, m_{12}, κ_2, κ_4, κ_6, and κ_{12} (Table 3) are within error of that obtained when m_2 is floated in the analysis, as expected given that m_2 is not a well constrained parameter in the analysis. Comparing the resulting stoichiometries, the NLLS analysis done with m_2 constrained results in stoichiometries for ATP binding dimers, tetramers, and hexamers that agree with the values obtained from the Wyman analysis (compare Tables 2 and 3).

The next level of complexity for modeling the $L_{n,app}$ binding isotherms would be to incorporate cooperativity between the binding sites. A hallmark of cooperative binding is a sigmoidal dependence on ligand concentration and steep binding isotherms that span fewer than two log units. As shown in Figure 4, this analysis does not result in

steep binding isotherms that would indicate the ATP binding sites are cooperative. There is no justification in the data to warrant incorporating complexity to the model. From this we conclude that the simplest model of *n*-independent, identical sites is sufficient to describe the data.

DISCUSSION

ATP binding has been shown to drive the formation of ClpA hexamers competent for binding to ClpP and polypeptide substrates (9, 12, 29). Thus, quantifying the hexamer population and the affinity of ATP binding to ClpA hexamers is an important aspect of determining the molecular basis for ClpA catalyzed substrate unfolding and translocation in the absence and presence of ClpP. Various groups have observed differences in substrate unfolding and/or translocation rate by ClpA in the absence versus the presence of ClpP (14, 16, 29). We hypothesize that one explanation for these observed differences is that ClpP can tune the ATP binding affinity of ClpA hexamers thereby altering the population of hexamers catalyzing substrate translocation in the presence of ClpP relative to that in the absence of ClpP. With this in mind, we examined the ATP-driven assembly mechanism of ClpA.

In the current study, we chose to use a ClpA variant ($ClpA_{E286A/E565A}$) previously shown to be ATP hydrolysis inactive to examine the assembly state in the presence of ATP (29). Previously determined equilibrium constants for the association of $ClpA_{E286A/E56A}$ in the absence of nucleotide were constrained in this analysis. Using an ATP hydrolysis deficient ClpA variant allowed us to slowly equilibrate the reaction of $ClpA_{E286A/E565A}$ with ATP over a 14 hours period without having to account for any changes to the [ATP] due

to ATP hydrolysis. This slow equilibration was achieved using equilibrium dialysis to ensure reversibility.

The self-association of $ClpA_{WT}$ in the presence of ATPγS has been previously examined using absorbance sedimentation velocity studies performed at 4 °C (9). In that study, a single $ClpA_{WT}$ oligomer, modeled as a hexamer, was observed in the presence of saturating [ATPγS]. Based on that observation, and the observation of ClpA hexamers in gel filtration assays performed at excess nucleotide concentrations, ClpA is often assumed to be completely hexameric in the presence of high nucleotide concentrations. In contrast to this, using a sedimentation velocity experiments with higher temporal resolution, we observe a distribution of ClpA oligomers at present at thermodynamic equilibrium for the range of $[ATP]_{free}$ tested.

The results in Figure 1 B show that while a dodecameric species is populated (~22 S), it is not the dominant species. The distributions observed from c(s) analysis and predicted from the equilibrium constants reported in Table 1, predict that the hexamer is the most populated oligomer in the presence of ATP. However, in order to correctly calculate the concentration of ClpA hexamers at a given [ATP], we must account for the population of dodecamer and all other lower order oligomers present. The equilibrium constants determined from the analysis presented here (Table 1) allow us to do that.

Analysis of the sedimentation velocity data obtained shows that in the presence of ATP, ClpA resides in a dynamic equilibrium of monomers, dimers, tetramers, hexamers, and dodecamers (Figure 2, Table 1). The populations of each oligomer depend on both the free nucleotide concentration and the protein concentration, as predicted from Eq. (8). We observe that the equilibrium constant for oligomer formation is not linearly dependent on

$[ATP]_{free}$ (Figure 5). The concentration of ClpA hexamers therefore cannot be determined by dividing the concentration of monomeric ClpA by six. Doing so would lead to an overestimation of the concentration of ClpA hexamers present and available to interact with polypeptide substrate and/or ClpP in solution.

Steady state kinetic studies designed to examine the ATPase rate and coordination of ClpA NBDs must account for the assembly state. Kress and co-workers investigated the steady state ATPase activity of each ClpA NBD using Walker B variants inactive in ATP hydrolysis at NDB1 or NBD2 (29). In this study they observe that ATPase activity at NBD2 increases in the presence of ClpP and decreases in the presence of another ClpA co-chaperone, ClpS. However, this study does not account for the changing distribution of oligomers present at the various ATP concentrations used to measure the ATPase rates. Thus, while the rate of ATP hydrolysis at NBD2 may indeed be altered when ClpP or ClpS is present, it's not possible to determine the effect of the co-chaperone without accounting for potential perturbations to the macromolecular assembly state.

The methods used here allowed us to determine the binding stoichiometry and affinity of ATP for each ClpA oligomer observed in the self-assembly equilibrium. To our knowledge, this is the first report of these parameters for ClpA. With this information in hand, we can start examining the effect of binding partners such as ClpP on the ATP binding affinity and assembly state of ClpA. Moreover, these methods are broadly applicable to other complex AAA+ chaperones with ligand linked assembly mechanisms.

REFERENCES

1. Wickner, S., S. Gottesman, D. Skowyra, J. Hoskins, K. McKenney, and M. R. Maurizi. 1994. A molecular chaperone, ClpA, functions like DnaK and DnaJ. Proc Natl Acad Sci U S A 91:12218-12222.

2. Gottesman, S., S. Wickner, and M. R. Maurizi. 1997. Protein quality control: triage by chaperones and proteases. Genes & development 11:815-823.

3. Butler, S. M., R. A. Festa, M. J. Pearce, and K. H. Darwin. 2006. Self-compartmentalized bacterial proteases and pathogenesis. Mol Microbiol 60:553-562.

4. Matyskiela, M. E., and A. Martin. 2013. Design principles of a universal protein degradation machine. Journal of molecular biology 425:199-213.

5. Olivares, A. O., T. A. Baker, and R. T. Sauer. 2016. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat Rev Microbiol 14:33-44.

6. Yedidi, R. S., P. Wendler, and C. Enenkel. 2017. AAA-ATPases in Protein Degradation. Front Mol Biosci 4:42.

7. Schirmer, E. C., J. R. Glover, M. A. Singer, and S. Lindquist. 1996. HSP100/Clp proteins: a common mechanism explains diverse functions. Trends in biochemical sciences 21:289-296.

8. Guo, F., M. R. Maurizi, L. Esser, and D. Xia. 2002. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. J Biol Chem 277:46743-46752.

9. Maurizi, M. R., S. K. Singh, M. W. Thompson, M. Kessel, and A. Ginsburg. 1998. Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. Biochemistry 37:7778-7786.

10. Piszczek, G., J. Rozycki, S. K. Singh, A. Ginsburg, and M. R. Maurizi. 2005. The molecular chaperone, ClpA, has a single high affinity peptide binding site per hexamer. J Biol Chem 280:12221-12230.

11. Effantin, G., T. Ishikawa, G. M. De Donatis, M. R. Maurizi, and A. C. Steven. 2010. Local and global mobility in the ClpA AAA+ chaperone detected by cryo-electron microscopy: functional connotations. Structure 18:553-562.

12. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399:665-679.

13. Hinnerwisch, J., W. A. Fenton, K. J. Furtak, G. W. Farr, and A. L. Horwich. 2005. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell 121:1029-1041.

14. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112:5377-5382.

15. Olivares, A. O., A. R. Nager, O. Iosefson, R. T. Sauer, and T. A. Baker. 2014. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nature structural & molecular biology.

16. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425:2795-2812.

17. Veronese, P. K., R. P. Stafford, and A. L. Lucius. 2009. The Escherichia coli ClpA Molecular Chaperone Self-Assembles into Tetramers. Biochemistry 48:9221-9233.

18. Veronese, P. K., and A. L. Lucius. 2010. Effect of Temperature on the Self-Assembly of the Escherichia coli ClpA Molecular Chaperone. Biochemistry 49:9820-9829.

19. Zhao, H., R. Ghirlando, G. Piszczek, U. Curth, C. A. Brautigam, and P. Schuck. 2013. Recorded scan times can limit the accuracy of sedimentation coefficients in analytical ultracentrifugation. Analytical biochemistry 437:104-108.

20. Ghirlando, R., A. Balbo, G. Piszczek, P. H. Brown, M. S. Lewis, C. A. Brautigam, P. Schuck, and H. Zhao. 2013. Improving the thermal, radial, and temporal accuracy of the analytical ultracentrifuge through external references. Analytical biochemistry 440:81-95.

21. Stafford, W. F., and P. J. Sherwood. 2004. Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. Biophysical chemistry 108:231-243.

22. Laue, T. M., Shah, B.D., Ridgeway, T.M., Pelletier, S.L. 1992. Computer-aided interpretation of analytical sedimentation data for proteins. In Analytical Ultracentrifugation in Biochemistry and Polymer Science. A. J. R. S.E. Harding, J.C. Horton, editor. Royal Society of Chemistry, Cambridge.

23. Veronese, P. K., B. Rajendar, and A. L. Lucius. 2011. Activity of Escherichia coli ClpA Bound by Nucleoside Di- and Triphosphates. Journal of molecular biology 409:333-347.

24. Correia, J. J. 2000. Analysis of weight average sedimentation velocity data. Methods Enzymol 321:81-100.

25. Lin, J., and A. L. Lucius. 2016. Examination of ClpB Quaternary Structure and Linkage to Nucleotide Binding. Biochemistry 55:1758-1771.

26. Lin, J., and A. L. Lucius. 2015. Analysis of Linked Equilibria. Methods Enzymol 562:161-186.

27. Wyman, J., Jr. 1964. Linked Functions and Reciprocal Effects in Hemoglobin: A Second Look. Advances in protein chemistry 19:223-286.

28. Wyman, J., and S. J. Gill. 1990. Binding and linkage : functional chemistry of biological macromolecules. University Science Books, Mill Valley, Ca.

29. Kress, W., H. Mutschler, and E. Weber-Ban. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. J Biol Chem 284:31441-31452.

Table 1

Equilibrium constants as a function of $[ATP]_{free}$

$[ATP]_{free}$ (μM)	$\log(L_{2,app})$	$\log(L_{4,app})$	$\log(L_{6,app})$	$\log(L_{12,app})$
17.0	4.0 ± 0.1	15.1 ± 0.1	23 ± 1	51 ± 1
17.9	4.0 ± 0.4	15.7 ± 0.1	26 ± 0.2	57 ± 3
26.8	4.0 ± 0.1	16.1 ± 0.2	26.3 ± 0.5	57 ± 3
27.0	4.5 ± 0.5	16.1 ± 0.3	27.1 ± 0.3	59.5 ± 0.8
62.3	5.0 ± 0.8	17.2 ± 0.5	28 ± 1	63 ± 2
68.5	4.2 ± 0.4	16.9 ± 0.2	27.7 ± 0.5	61.3 ± 0.8
120.3	4.6 ± 0.5	17.3 ± 0.1	29.1 ± 0.1	63 ± 2
120.3	4.0 ± 0.4	17.3 ± 0.2	29.0 ± 0.3	63.0 ± 0.7
203.2	3.88 ± 0.08	17.7 ± 0.3	29.6 ± 0.5	64.3 ± 0.8
536.2	4.8 ± 0.1	19 ± 0.3	32 ± 0.4	69.2 ± 0.8
1100	5.1 ± 0.5	19 ± 1	32 ± 1	69 ± 2

Table 2

Extent of ATP binding to each ClpA oligomer predicted from Wyman analysis

	$\overline{X}_1 = 0$	$\overline{X}_1 = 1$	$\overline{X}_1 = 2$
$\overline{X}_2$	1.2	3.2	5.2
$\overline{X}_4$	2.4	6.4	10.4
$\overline{X}_6$	4.7	10.7	16.7
$\overline{X}_{12}$	9	21	33

Table 3

Results from NLLS analysis of $L_{n,app}$ binding isotherms

floating parameters	m_1 constrained	m_1 and m_2 constrained
m_1	1*	1*
m_2	5 ± 1106	3*
m_4	5.3 ± 0.4	5.7 ± 0.3
m_6	9.1 ± 0.6	9.6 ± 0.5
m_{12}	20 ± 1	21 ± 1
κ_1 (M^{-1})	(2.5 ± 0.9) x 10^3	(4 ± 2) x 10^3
κ_2 (M^{-1})	(5 ± 3) x 10^3	(2.0 ± 0.6) x 10^3
κ_4 (M^{-1})	(1.2 ± 0.1) x 10^5	(1.3 ± 0.2) x 10^5
κ_6 (M^{-1})	(8 ± 1) x 10^4	(1.1 ± 0.2) x 10^5
κ_{12} (M^{-1})	(8 ± 1) x 10^4	(9 ± 2) x 10^4

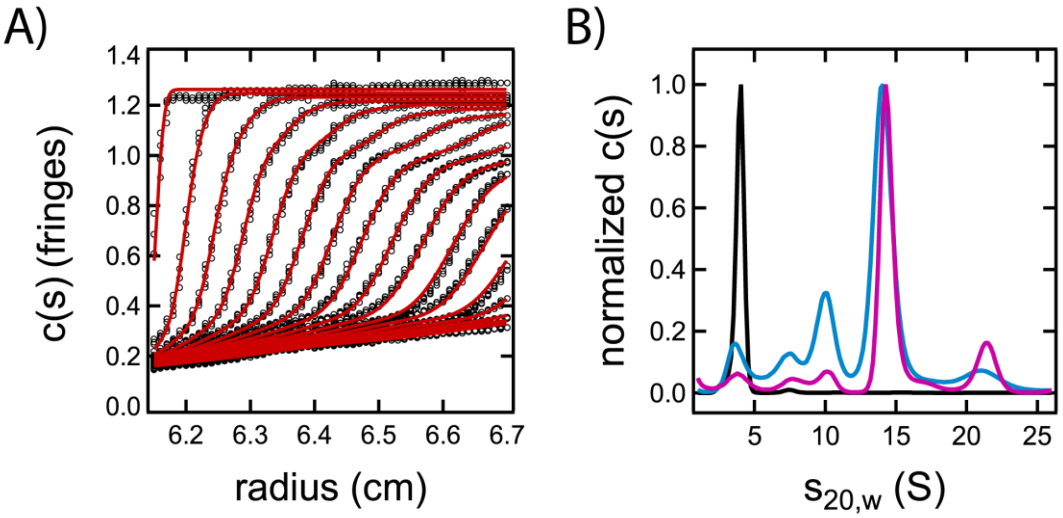


Figure 1. C(s) distributions for ClpA$_{E286A/E565A}$ in the absence and presence of ATP. (A) Sedimentation velocity experiments performed on 6 µM$_{E286A/E565A}$ monitoring fringe displacement as a function of radial position and time. Sets of sedimenting boundaries collected (black, open circles) were subjected to c(s) analysis (red, solid traces) to obtain c(s) distributions. (B) c(s) distributions are shown for 6 µM ClpA$_{E286A/E565A}$ in the absence of ATP (black) and in the presence of 500 µM (blue) and 1 mM ATP (purple).

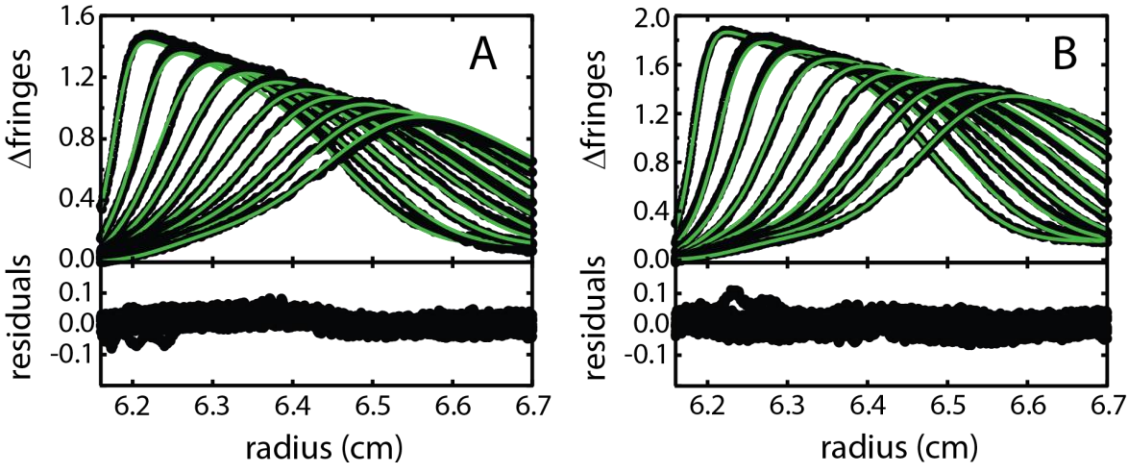


Figure 2. Analysis of representative time difference curves fit to a 1-2-4-6-12 stepwise assembly model. Time difference curves obtained for (A) 6 µM $ClpA_{E286A/E565A}$ in the presence of 62.3 µM ATP and (B) 10 µM $ClpA_{E286A/E565A}$ in the presence of 68.5 µM ATP_{free}, were subjected to NLLS analysis using the 1-2-4-6-12 model. Sedimentation boundaries were collected every 30 seconds and all boundaries collected in the first 2 hours of sedimentation were included in the analysis. Every 10th boundary is shown. The data are shown in black open circles and the fit is shown in green. Fitting residuals randomly distributed about zero, shown underneath each plot, and fitting RMSD of (A) 0.0268 and (B) 0.0186, respectively indicate the data are well described by the 1-2-4-6-12 assembly model.

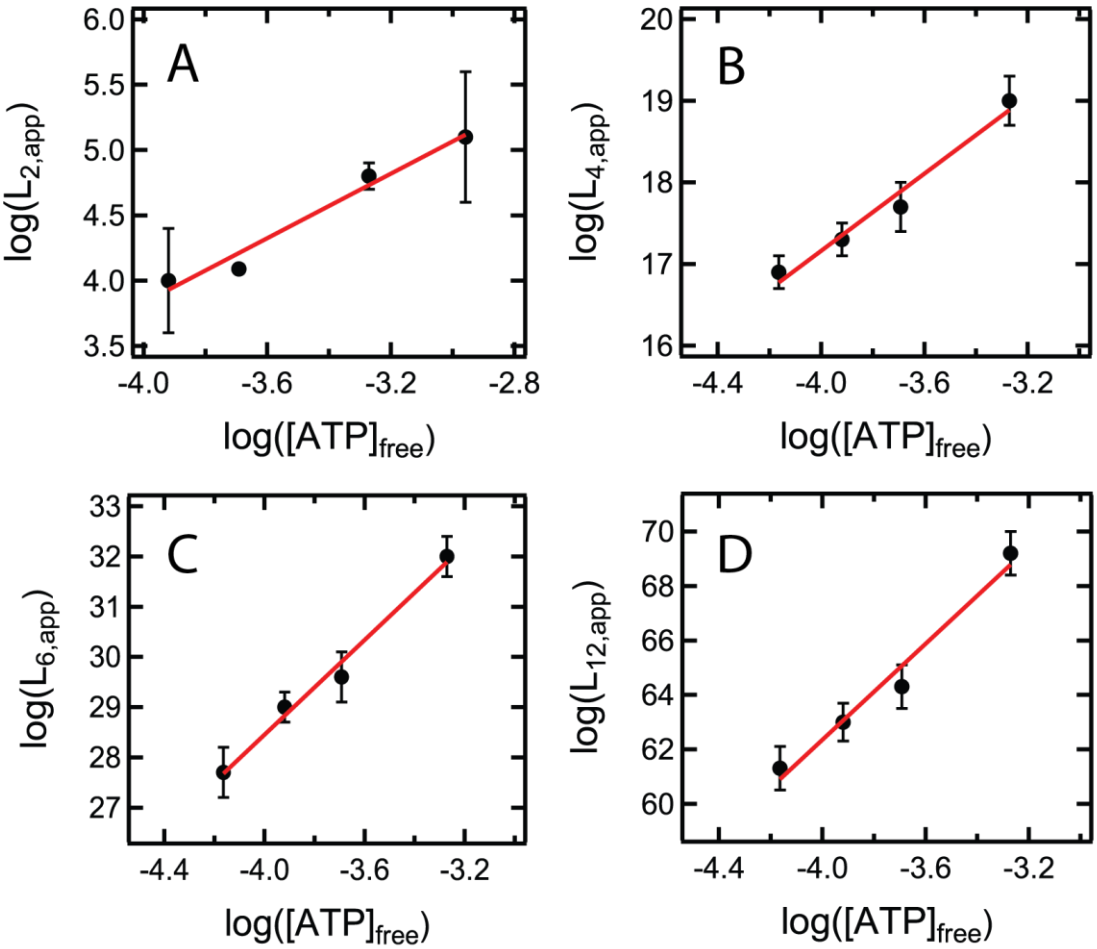


Figure 3. Wyman plots of log($L_{n,app}$) as a function of log[ATP]$_{free}$. The linear region of each plot of log of (A) $L_{2,app}$ (B) $L_{4,app}$, (C) $L_{6,app}$, and (D) $L_{12,app}$ as a function of log[ATP]$_{free}$. The data in each plot were individually fit to a linear regression (red) resulting in slopes reported in Eq. (14) – (17).

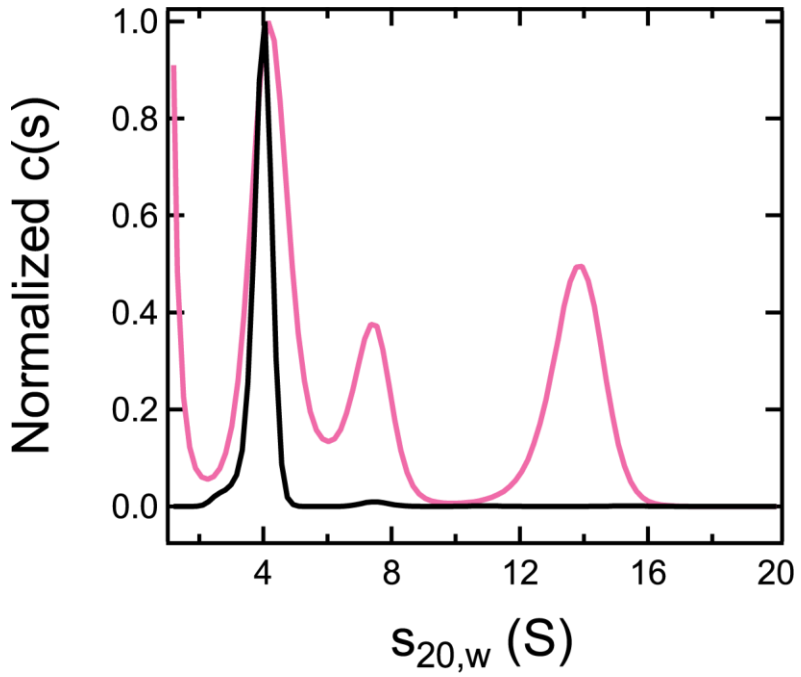


Figure 4. C(s) distributions of 6 µM ClpAE286A/E565A in the absence and presence of TNP-ATP. Sedimentation velocity experiments were collected for 6 µM ClpAE286A/E565A in the absence of ATP monitoring interference (black) and in the presence of TNP-ATP monitoring absorbance at 408 nm (pink). Because protein does not absorb in the visible region of the UV-vis spectrum, peaks present in the distribution collected monitoring absorbance at 408 nm represent TNP-ATP bound $ClpA_{E286A/E565A}$ species.

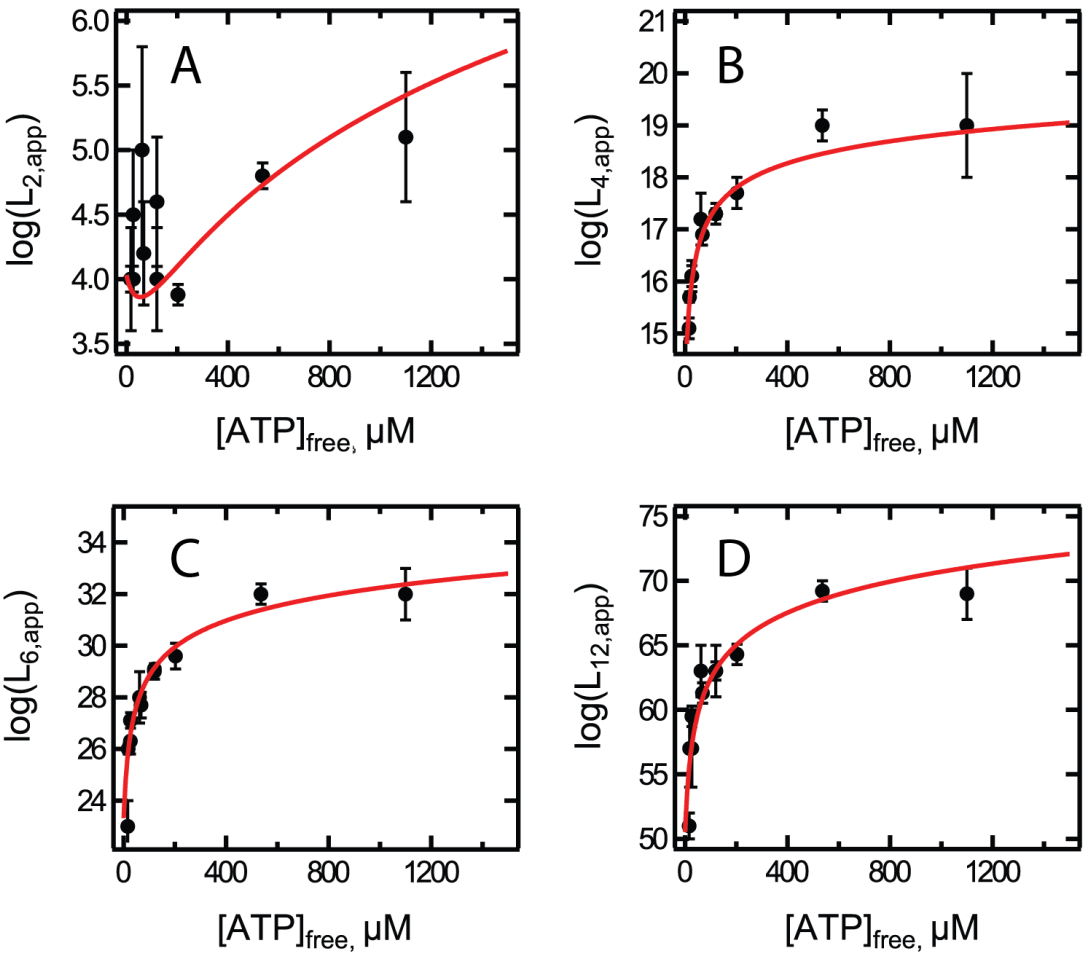


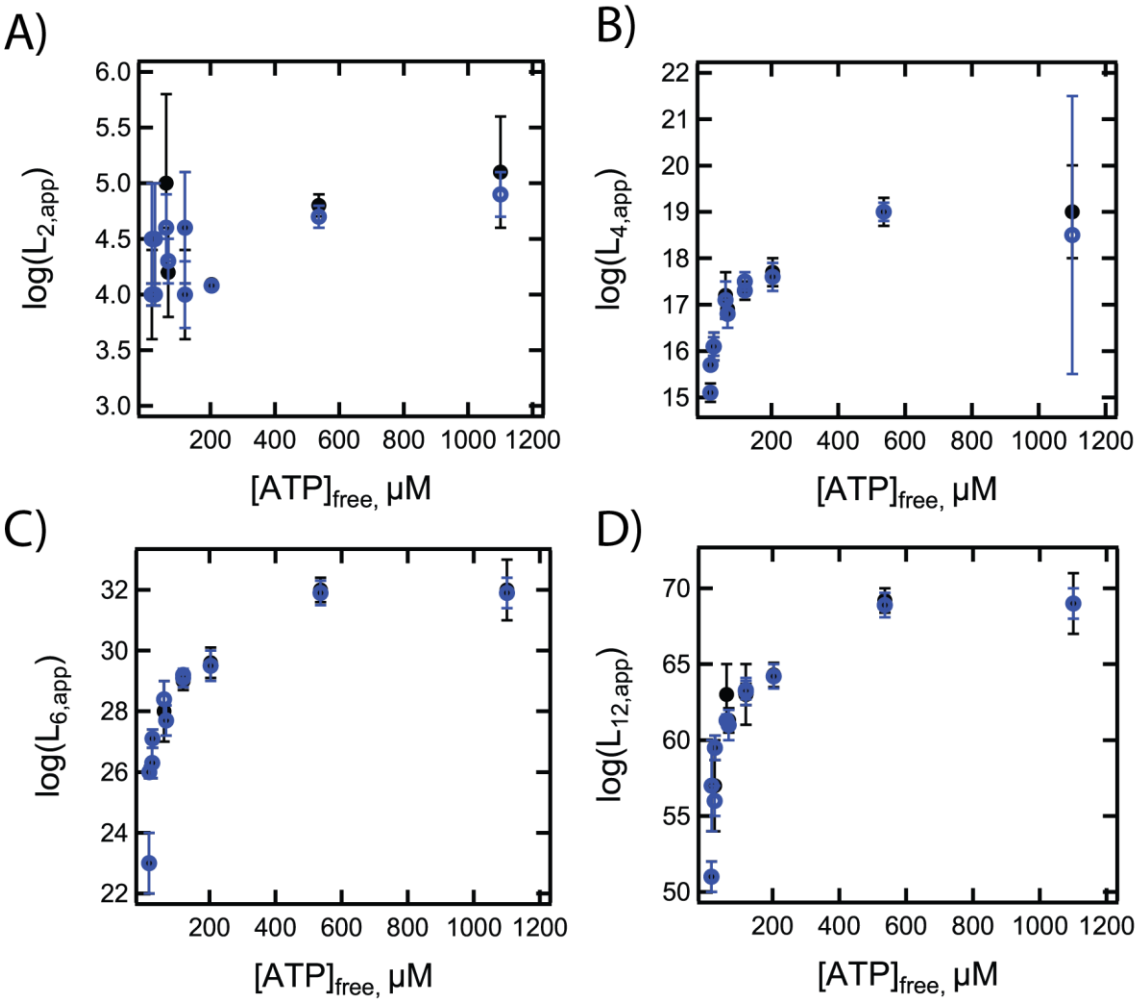
Figure 5. Global analysis of $L_{n,app}$ binding isotherms. (A) Log($L_{2,app}$), (B) log(L4,app), (C) log(L6,app), and (D) log(L12,app) as a function of [ATP]free were subjected to global NLLS analysis using Eq. (20) – (23). The black closed circles represent the data presented in Table 1. The solid red lines represent the result from NLLS analysis obtained by constraining m_1 = 1 and m_2 = 3. The best fit values obtained for m_4, m_6, m_{12}, κ_1, κ_2, κ_4, κ_6, and κ_{12} are summarized in Table 3.

Supplementary Table S1

Assembly equilibrium ($L_{n,app}$) and reverse rate ($k_{r,n}$) constants as a function of $[ATP]_{free}$

$[ATP]_{free}$ (μM)	$\log(L_{2,app})$	$\log(L_{4,app})$	$\log(L_{6,app})$	$\log(L_{12,app})$
17.0E-6	4.0 ± 0.1	15.1 ± 0.1	23 ± 1	51 ± 1
17.9E-6	4.5 ± 0.5	15.7 ± 0.1	26 ± 0.2	57 ± 3
26.8E-6	4.0 ± 0.1	16.1 ± 0.2	26.3 ± 0.5	56 ± 1
27.0E-6	4.5 ± 0.5	16.1 ± 0.3	27.1 ± 0.3	59.5 ± 0.8
62.3E-6	4.6 ± 0.3	17.1 ± 0.4	28.4 ± 0.6	61.3 ± 0.4
68.5E-6	4.3 ± 0.2	16.8 ± 0.3	27.7 ± 0.5	61 ± 1
120.3E-6	4.6 ± 0.5	17.3 ± 0.1	29.2 ± 0.2	63.2 ± 0.9
120.3E-6	4.0 ± 0.3	17.5 ± 0.2	29.1 ± 0.3	63.3 ± 0.6
203.2E-6	3.08 ± 0.03	17.6 ± 0.3	29.5 ± 0.5	64.2 ± 0.8
536.2E-6	4.7 ± 0.1	19.0 ± 0.2	31.9 ± 0.4	68.9 ± 0.8
1.1E-3	4.9 ± 0.2	18.5 ± 3	31.9 ± 0.5	69 ± 1

$[ATP]_{free}$ (μM)	$k_{r,2}$	$k_{r,4}$	$k_{r,6}$	$k_{r,12}$
17.0E-6	$(2 \pm 1) \times 10^{-2}$	$(2 \pm 2) \times 10^{-2}$	$(0.9 \pm 1) \times 10^{-6}$	1131 ± 1960
17.9E-6	$(3 \pm 4) \times 10^{-2}$	$(4 \pm 6) \times 10^{-2}$	$(0.8 \pm 1) \times 10^{-2}$	$(5 \pm 4) \times 10^{-3}$
26.8E-6	$(3 \pm 5) \times 10^{-2}$	$(3 \pm 3) \times 10^{-3}$	$(1 \pm 0.9) \times 10^{-2}$	$(1 \pm 2) \times 10^{7}$
27.0E-6	3 ± 5	$(8 \pm 8) \times 10^{-3}$	10 ± 14	6 ± 11
62.3E-6	0.6 ± 1	0.7 ± 0.9	$(2.9 \pm 0.5) \times 10^{-4}$	$(2 \pm 3) \times 10^{-2}$
68.5E-6	6 ± 5	0.9 ± 1	$(1 \pm 2) \times 10^{-2}$	$(3 \pm 3) \times 10^{-3}$
120.3E-6	1.2 ± 0.8	2 ± 2	$(4.4 \pm 0.7) \times 10^{-4}$	$(1.4 \pm 0.8) \times 10^{-3}$
120.3E-6	0.2 ± 0.3	0.07 ± 0.06	$(3 \pm 1) \times 10^{-4}$	$(3 \pm 3) \times 10^{-4}$
203.2E-6	$(7.0 \pm 0.4) \times 10^{-3}$	1.1 ± 0.8	$(2.4 \pm 0.5) \times 10^{-5}$	$(4 \pm 2) \times 10^{-4}$
536.2E-6	$(1.0 \pm 0.7) \times 10^{-2}$	2 ± 0.1	$(1.3 \pm 0.8) \times 10^{-5}$	$(1 \pm 1) \times 10^{-5}$
1.1E-3	0.4 ± 0.4	0.2 ± 0.3	$(0.6 \pm 1) \times 10^{-9}$	$(2 \pm 1) \times 10^{-4}$



Supplementary Figure S1. Comparison of $L_{n,app}$ binding isotherms collected by constraining $k_{r,n}$ and floating $k_{r,n}$ in the analysis of time difference curves. Binding isotherms constructed from equilibrium constants listed in Table 1 as a function of $[ATP]_{free}$ are shown as black closed circles. These were obtained by constraining oligomer dissociation rate constants ($k_{r,n}$) in the analysis of difference curves. For comparison, $L_{n,app}$ values obtained by floating the equilibrium constants and the reverse rate constants in the analysis of difference curves are also shown (blue, open circles). The apparent equilibrium constants obtained from each fitting strategy are within error or one another.

CHAPTER 5

CONCLUSION

In this dissertation we examine the importance of characterizing the nucleotide linked assembly mechanism of AAA+ molecular chaperones with respect to understanding the mechanism of polypeptide unfolding and remodeling. First we examine the structural similarities of three closely related Hsp100 AAA+ molecular motors: ClpA, ClpB, and Hsp104. In this review, we identify key considerations for designing experiments to examine the translocation mechanisms of these three, and closely related, AAA+ molecular chaperones. Chief amongst these considerations is accounting for the ATP-linked self-assembly mechanism in the experimental design. In the second part of this work, we apply the analytical techniques discussed to examine the assembly mechanism of ClpA in the absence and presence of ATP.

Challenges of Investigating Molecular Mechanisms of AAA+ Chaperones

In Chapter 2, we highlight the challenges in studying the mechanisms of AAA+ motors and present a review of the work that has been done to meet these challenges with respect to ClpA, ClpB, and Hsp104. These proteins pose a unique challenge to the experimentalist because the product of the reaction they catalyze, either substrate unfolding or substrate remodeling, is not covalently modified in the course of the reaction. In addition, ATP binding and hydrolysis is coupled to oligomerization of the AAA+ motor and substrate unfolding or remodeling activities.

One solution to meet these challenges is to design assays to characterize the transient state kinetic mechanism of polypeptide substrate unfolding or remodeling. AAA+ molecular motors use ATP to perform exert physical force on the polypeptide substrate. Therefore, designing assays that relate the dependencies of elementary rate constants on nucleotide concentration can lead to direct insights about the substrate unfolding or remodeling mechanism. For example, by using a single turnover transient state kinetic assay our lab reported the kinetic step size for ClpA catalyzed translocation in the absence and presence of ClpP (29, 30). Similarly, single turnover experiments like single-molecule techniques, can measure transient state kinetic parameters without convolution from coupling to macromolecular assembly (31-34).

Another key consideration is to have a measure for the population of the AAA+ motor complex catalyzing substrate (ATP or polypeptide) turnover. This often requires characterization of the ligand linked assembly state in the absence and presence of ATP. As we've shown in Chapter 2 on the work reviewed for ClpB assembly, and in the current work on ClpA assembly presented in Chapter 3 and 4, these proteins exhibit a dynamic assembly mechanism in the absence and presence of ATP (23, 24). Thus the characterizing the energetics that govern the concentration of the active oligomer and the effect of partner proteins and/or mutations is crucial to interpreting the ATP-dependent parameters being assayed in any experiment.

ClpA Assembly in the absence of nucleotide and presence of ATP

In the second part of this dissertation, we perform a thermodynamically rigorous examination of the assembly of ClpA in the absence (Chapter 3) and presence (Chapter 4) of ATP. Here we employ hydrodynamic techniques and analysis used to similarly

investigate the assembly of ClpB in the absence and presence of nucleotide (23, 24, 35). We find that ClpA, like ClpB, resides in a distribution of states in the absence and presence of nucleotide.

In the absence of nucleotide, we observe that wild type ClpA ($ClpA_{WT}$) resides in a distribution of monomers, dimers, and tetramers. This is in agreement with previously published hydrodynamic study on $ClpA_{WT}$ using lower resolution techniques (25, 26). Here we report equilibrium constants for the formation of each oligomer obtained using higher resolution techniques.

In the same study (Chapter 3) we additionally examine the effect of mutations on the self-assembly mechanism of ClpA in the absence of nucleotides. Specifically, we studied the effect of three mutations known to abolish the ATP hydrolysis activity at NBD1 ($ClpA_{E286A}$), NBD2 ($ClpA_{E565A}$), and both NBDs ($ClpA_{E286A/E565A}$). We observe that the assembly data collected for each of the variants, was best described by a monomer-dimer-tetramer-hexamer model. In addition, we observe differences in the self-assembly equilibrium constants determined for dimerization, tetramerization, and hexamerization of each variant. Together, these results highlight the importance of characterizing the effect of mutations on the assembly state.

In Chapter 4, we examine the thermodynamic mechanism of ClpA assembly in the presence of ATP. In this analysis, we use equilibrium dialysis in the sample preparation to rigorously achieve a thermodynamically reversible equilibrium. From this approach we additionally obtain direct measure of the concentration of free nucleotide, the driving force for oligomerization. From direct boundary analysis of time difference curves, we obtain equilibrium constants for the formation of ClpA oligomers populated for a range of

measured $[ATP]_{free}$. The apparent oligomerization equilibrium constants were modeled as a function of ATP to obtain binding stoichiometry and binding affinity for ATP binding to each ClpA oligomer.

In the presence of ATP, we observe that ClpA resides in a distribution of monomers, dimers, tetramers, hexamers, and dodecamers. Although the hexameric population dominates the distribution at all nucleotide concentrations examined, the other oligomers must be accounted for in order to avoid overestimates for the population of ClpA hexamers in solution.

The resulting binding isotherms for apparent dimerization, tetramerization, hexamerization, and dodecamerization as a function of $[ATP]_{free}$ were modeled to determine the average stoichiometry and ATP affinity for each oligomer. From this analysis we observe that ATP binding is sufficiently described by an n-independent and identical sites model. In addition, we report stoichiometries of 1, 3, 6, 10, and 20 for ATP binding monomers, dimers, tetramers, hexamers and dodecamers, respectively.

Future Directions

One future direction of this work, with a thermodynamically rigorous characterization of the ClpA assembly state, is the examination of the binding affinity of ClpA hexamers to ClpP tetradecamers. Previous analysis been done to characterize the affinity of ClpA for ClpP. Maurizi and co-workers report a binding constant of ~4 nM for the association of ClpA hexamers to ClpP tetradecamers (36). Yet conflicting results emerge in a later studies regarding whether the binding of ClpA to ClpP is cooperative or not (37, 38). One potential explanation for these discrepancies in the literature is that none

of these studies account for the distribution of oligomers in the ClpA assembly state. The concentration of hexamers is simply calculated by dividing the monomeric ClpA concentration by 6. With the assembly state energetics provided in this dissertation, we can now rigorously determine the affinity of ClpA for ClpP.

Determination of the affinity constant for ClpA binding to ClpP will allow us to explore more interesting questions about this ATP-dependent protease system. For example, one outstanding question regarding ClpAP catalyzed substrate translocation is, can two-to-one ClpAP complexes translocate polypeptide substrate from one or both ends of the complex? In addition, being able to calculate a species fraction simulation for the distribution of single and doubly ClpA-ligated ClpP will allow us to preferentially populate each of these states and examine implications of the resulting observations.

GENERAL REFERENCES

1. Gottesman, S., M. R. Maurizi, and S. Wickner. 1997. Regulatory subunits of energy-dependent proteases. Cell 91:435-438.

2. Bukau, B., J. Weissman, and A. Horwich. 2006. Molecular chaperones and protein quality control. Cell 125:443-451.

3. Olivares, A. O., T. A. Baker, and R. T. Sauer. 2016. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat Rev Microbiol 14:33-44.

4. Chiti, F., and C. M. Dobson. 2017. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. Annu Rev Biochem 86:27-68.

5. Schirmer, E. C., J. R. Glover, M. A. Singer, and S. Lindquist. 1996. HSP100/Clp proteins: a common mechanism explains diverse functions. Trends in biochemical sciences 21:289-296.

6. Neuwald, A. F., L. Aravind, J. L. Spouge, and E. V. Koonin. 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome research 9:27-43.

7. Yedidi, R. S., P. Wendler, and C. Enenkel. 2017. AAA-ATPases in Protein Degradation. Front Mol Biosci 4:42.

8. Balchin, D., M. Hayer-Hartl, and F. U. Hartl. 2016. In vivo aspects of protein folding and quality control. Science 353:aac4354.

9. Schmidt, M., J. Hanna, S. Elsasser, and D. Finley. 2005. Proteasome-associated proteins: regulation of a proteolytic machine. Biological chemistry 386:725-737.

10. Sauer, R. T., and T. A. Baker. 2011. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annual review of biochemistry 80:587-612.

11. Matyskiela, M. E., and A. Martin. 2013. Design principles of a universal protein degradation machine. Journal of molecular biology 425:199-213.

12. Grimaud, R., M. Kessel, F. Beuron, A. C. Steven, and M. R. Maurizi. 1998. Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. J Biol Chem 273:12476-12481.

13. Benaroudj, N., P. Zwickl, E. Seemuller, W. Baumeister, and A. L. Goldberg. 2003. ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. Molecular cell 11:69-78.

14. Chen, S., J. Wu, Y. Lu, Y. B. Ma, B. H. Lee, Z. Yu, Q. Ouyang, D. J. Finley, M. W. Kirschner, and Y. Mao. 2016. Structural basis for dynamic regulation of the human 26S proteasome. Proceedings of the National Academy of Sciences of the United States of America 113:12991-12996.

15. Barthelme, D., J. Z. Chen, J. Grabenstatter, T. A. Baker, and R. T. Sauer. 2014. Architecture and assembly of the archaeal Cdc48*20S proteasome. Proceedings of the National Academy of Sciences of the United States of America 111:E1687-1694.

16. Mogk, A., B. Bukau, and H. H. Kampinga. 2018. Cellular Handling of Protein Aggregates by Disaggregation Machines. Mol Cell 69:214-226.

17. Parsell, D. A., A. S. Kowal, M. A. Singer, and S. Lindquist. 1994. Protein disaggregation mediated by heat-shock protein Hsp104. Nature 372:475-478.

18. Lindquist, S., and G. Kim. 1996. Heat-shock protein 104 expression is sufficient for thermotolerance in yeast. Proceedings of the National Academy of Sciences of the United States of America 93:5301-5306.

19. Doyle, S. M., O. Genest, and S. Wickner. 2013. Protein rescue from aggregates by powerful molecular chaperone machines. Nat Rev Mol Cell Biol 14:617-629.

20. Deville, C., M. Carroni, K. B. Franke, M. Topf, B. Bukau, A. Mogk, and H. R. Saibil. 2017. Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. Sci Adv 3:e1701726.

21. Gates, S. N., A. L. Yokom, J. Lin, M. E. Jackrel, A. N. Rizo, N. M. Kendsersky, C. E. Buell, E. A. Sweeny, K. L. Mack, E. Chuang, M. P. Torrente, M. Su, J. Shorter, and D. R. Southworth. 2017. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. Science 357:273-279.

22. Li, T., C. L. Weaver, J. Lin, E. C. Duran, J. M. Miller, and A. L. Lucius. 2015. Escherichia coli ClpB is a non-processive polypeptide translocase. The Biochemical journal 470:39-52.

23. Lin, J., and A. L. Lucius. 2015. Examination of the dynamic assembly equilibrium for E. coli ClpB. Proteins 83:2008-2024.

24. Lin, J., and A. L. Lucius. 2016. Examination of ClpB Quaternary Structure and Linkage to Nucleotide Binding. Biochemistry 55:1758-1771.

25. Veronese, P. K., and A. L. Lucius. 2010. Effect of Temperature on the Self-Assembly of the Escherichia coli ClpA Molecular Chaperone. Biochemistry 49:9820-9829.

26. Veronese, P. K., R. P. Stafford, and A. L. Lucius. 2009. The Escherichia coli ClpA Molecular Chaperone Self-Assembles into Tetramers. Biochemistry 48:9221-9233.

27. Kress, W., H. Mutschler, and E. Weber-Ban. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. J Biol Chem 284:31441-31452.

28. Baytshtok, V., T. A. Baker, and R. T. Sauer. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc Natl Acad Sci U S A 112:5377-5382.

29. Rajendar, B., and A. L. Lucius. 2010. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. J Mol Biol 399:665-679.

30. Miller, J. M., J. Lin, T. Li, and A. L. Lucius. 2013. E. coli ClpA Catalyzed Polypeptide Translocation is Allosterically Controlled by the Protease ClpP. Journal of Molecular Biology 425:2795-2812.

31. Olivares, A. O., A. R. Nager, O. Iosefson, R. T. Sauer, and T. A. Baker. 2014. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nature structural & molecular biology.

32. Martin, A., T. A. Baker, and R. T. Sauer. 2005. Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. Nature 437:1115-1120.

33. Cordova, Juan C., Adrian O. Olivares, Y. Shin, Benjamin M. Stinson, S. Calmat, Karl R. Schmitz, M.-E. Aubin-Tam, Tania A. Baker, Matthew J. Lang, and Robert T. Sauer. 2014. Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell 158:647-658.

34. Smith, D. M., H. Fraga, C. Reis, G. Kafri, and A. L. Goldberg. 2011. ATP binds to proteasomal ATPases in pairs with distinct functional effects, implying an ordered reaction cycle. Cell 144:526-538.

35. Lin, J., and A. L. Lucius. 2015. Analysis of Linked Equilibria. Methods Enzymol 562:161-186.

36. Maurizi, M. R., S. K. Singh, M. W. Thompson, M. Kessel, and A. Ginsburg. 1998. Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. Biochemistry 37:7778-7786.

37. Kress, W., H. Mutschler, and E. Weber-Ban. 2007. Assembly Pathway of an AAA+ Protein: Tracking ClpA and ClpAP Complex Formation in Real Time. Biochemistry 46:6183-6193.

38. Kolygo, K., N. Ranjan, W. Kress, F. Striebel, K. Hollenstein, K. Neelsen, M. Steiner, H. Summer, and E. Weber-Ban. 2009. Studying chaperone-proteases using a real-time approach based on FRET. J Struct Biol 168:267-277.