INSIGHTS INTO THE STRUCTURE AND FUNCTION OF THE AGGREGATE-REACTIVATING MOLECULAR CHAPERONE CLPB

by

MARIA NAGY

B.S., Babes-Bolyai University, Cluj-Napoca, Romania, 1997
M.S., Babes-Bolyai University, Cluj-Napoca, Romania, 1998

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Graduate Biochemistry Group
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

Abstract

ClpB is a bacterial heat-shock protein that disaggregates and reactivates strongly aggregated proteins in cooperation with the DnaK chaperone system. ClpB contains two ATP-binding AAA+ modules, a linker coiled-coil domain, and a highly mobile N-terminal domain. It forms ring-shaped hexamers in a nucleotide-dependent manner. The unique aggregation reversing chaperone activity of ClpB involves ATP-dependent translocation of substrates through the central channel in the ClpB ring. The initial events of aggregate recognition and the events preceding the translocation step are poorly understood. In addition to the full-length ClpB95, a truncated isoform ClpB80, that is missing the whole N-terminal domain, is also produced *in vivo*.

Various aspects of the structure and function of ClpB were addressed in this work. The thermodynamic stability of ClpB in its monomeric and oligomeric forms, as well as the nucleotide-induced conformational changes in ClpB were investigated by fluorescence spectroscopy. Equilibrium urea-induced unfolding showed that two structural domains-the small domain of the C-terminal AAA+ module and the coiled-coil domain-were destabilized in the oligomeric form of ClpB, which indicates that only those domains change their conformation or interactions during formation of the ClpB rings. Several locations of Trp-fluorescence probes were also found to respond to nucleotide binding.

The biological role of the two naturally-occurring ClpB isoforms was also investigated. We discovered that ClpB achieves optimum chaperone activity by

synergistic cooperation of the two isoforms that form hetero-oligomers. We found that ClpB95/ClpB80 hetero-oligomers form preferentially at low protein concentration with higher affinity than homo-oligomers of ClpB95. Moreover, hetero-oligomers bind to aggregated substrates with a similar efficiency as homo-oligomers of ClpB95, do not show enhanced ATPase activity over that of the homo-oligomers, but display a strongly stimulated chaperone activity during the reactivation of aggregated proteins. We propose that extraction of single polypeptides from aggregates and their delivery to the ClpB channel for translocation is the rate-limiting step in aggregate reactivation and that step is supported by the mobility of the N-terminal domain of ClpB. We conclude that the enhancement of the chaperone activity of the hetero-oligomers is linked to an enhancement of mobility of the N-terminal domains.

INSIGHTS INTO THE STRUCTURE AND FUNCTION OF THE AGGREGATE-REACTIVATING MOLECULAR CHAPERONE CLPB

by

MARIA NAGY

B.S., Babes-Bolyai University, Cluj-Napoca, Romania, 1997
M.S., Babes-Bolyai University, Cluj-Napoca, Romania, 1998

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Graduate Biochemistry Group
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

Approved by:

Major Professor
Michal Zolkiewski

Copyright

.

Abstract

ClpB is a bacterial heat-shock protein that disaggregates and reactivates strongly aggregated proteins in cooperation with the DnaK chaperone system. ClpB contains two ATP-binding AAA+ modules, a linker coiled-coil domain, and a highly mobile N-terminal domain. It forms ring-shaped hexamers in a nucleotide-dependent manner. The unique aggregation reversing chaperone activity of ClpB involves ATP-dependent translocation of substrates through the central channel in the ClpB ring. The initial events of aggregate recognition and the events preceding the translocation step are poorly understood. In addition to the full-length ClpB95, a truncated isoform ClpB80, that is missing the whole N-terminal domain, is also produced *in vivo*.

Various aspects of the structure and function of ClpB were addressed in this work. The thermodynamic stability of ClpB in its monomeric and oligomeric forms, as well as the nucleotide-induced conformational changes in ClpB were investigated by fluorescence spectroscopy. Equilibrium urea-induced unfolding showed that two structural domains-the small domain of the C-terminal AAA+ module and the coiled-coil domain-were destabilized in the oligomeric form of ClpB, which indicates that only those domains change their conformation or interactions during formation of the ClpB rings. Several locations of Trp-fluorescence probes were also found to respond to nucleotide binding.

The biological role of the two naturally-occurring ClpB isoforms was also investigated. We discovered that ClpB achieves optimum chaperone activity by

synergistic cooperation of the two isoforms that form hetero-oligomers. We found that ClpB95/ClpB80 hetero-oligomers form preferentially at low protein concentration with higher affinity than homo-oligomers of ClpB95. Moreover, hetero-oligomers bind to aggregated substrates with a similar efficiency as homo-oligomers of ClpB95, do not show enhanced ATPase activity over that of the homo-oligomers, but display a strongly stimulated chaperone activity during the reactivation of aggregated proteins. We propose that extraction of single polypeptides from aggregates and their delivery to the ClpB channel for translocation is the rate-limiting step in aggregate reactivation and that step is supported by the mobility of the N-terminal domain of ClpB. We conclude that the enhancement of the chaperone activity of the hetero-oligomers is linked to an enhancement of mobility of the N-terminal domains.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost I would like to express my deepest gratitude to my advisor Dr. Michal Zolkiewski for the continuous support, encouragement, and the opportunities provided during my graduate studies. I still remember the first science-related conversation we had several years ago that greatly influenced my decision to change my field of study from physics to biochemistry and to start pursuing my PhD degree at KSU. I thank him for the possibility to work under his guidance, and for sharing his knowledge and experience with me with great patience.

I want to acknowledge members of my committee, Dr. Lawrence Davis, Dr. Ramaswamy Krishnamoorthi and Dr. Paul Smith for their willingness to accept this responsibility and for their valuable suggestions, as well as their interest in my work. I would also like to thank Dr. Weiqun Wang from the Department of Human Nutrition, for serving as the outside chair during my defense.

I would like to thank Dr. Maria I. Zavodszky from Michigan State University for helping us with the homology modeling and the normal mode analysis; Dr. Dan Boyle from the Department of Biology for his help with TEM imaging; and Dr. John Tomich for generously allowing us to use the spectrofluorimeter for many of our studies.

My thanks to former and current members of the Zolkiewski Lab, Dr. Vladimir Akoev, Dr. Micheal E. Barnett, Dr. Vekalet Tek, Dr. Sabina Kedzierska, Dr. Zhonghua Liu, Hui-Chuan Wu (Arleen) and Ting Zhang for their help and friendship. I am especially grateful to Mike Barnett for being such a patient and effective teacher during

my first years in the lab. Mike was never to busy to answer the many questions I encountered and was a wonderful friend throughout. Also, the last two years would not have been the same without Arleen and the coffee breaks we spent together discussing our research.

I would also like to express my appreciation to the entire faculty and fellow graduate students of the Department of Biochemistry for creating a wonderful and stimulating environment. Many thanks to the biochemistry office staff members for their assistance throughout the years, particularly Crystal Sapp and Rebecca Darkow.

I want to convey my gratitude to Maria and Peter Zavodszky for their friendship, support and encouragement. Thank you for believing in me and motivating me to pursue this academic path.

Most importantly, I would like to express my deepest gratitude to my family. All this would not have been possible without the many sacrifices made by my parents and my grandmother. I am also thankful to my brothers István, Szabolcs and Levente for their unconditional love and support.

Finally, I would like to thank my husband Albert for his love, understanding constant support and optimism that made this great journey possible and enjoyable. I hope I can provide the same support during his endeavors.

Dedication

I dedicate this work to my husband Albert, to my wonderful and brave mother Mária, and to the loving memory of my beloved father Árpád and grandmother Pépel Anna.

CHAPTER 1 - Introduction

Protein misfolding and aggregation

Proper cell function and viability depends not only on the presence of correctly folded, functional proteins, but also on the efficiency of the protein folding process and the ability of the already folded proteins to remain in a functional state. Cells have evolved elaborate strategies to control the protein folding process and to ensure that the correctly folded proteins remain in a functional state throughout the lifespan of the organism.

Protein folding has been long considered one of the most exciting and challenging processes of modern biology. Major advances in the understanding of the mechanism of the folding process have been made since the famous experiment of Anfinsen on the *in vitro* folding of ribonuclease A (Anfinsen, 1973). The original observation - that the information required for correct folding is encoded in the primary amino acid sequence of the protein - still remains one of the fundamental principles that govern protein folding. However, protein folding is a complex process whether it happens in the natural setting of the living cells or under laboratory conditions. *In vitro* refolding experiments revealed that some smaller proteins (with less than 100 residues) fold, and thus find their lowest-energy/most stable structure, by following simple two-state kinetics. In contrast, folding of larger proteins, especially the ones composed of multiple domains, frequently involves partly folded intermediate states on the way to the native conformation. These partly folded intermediates expose regions of the polypeptide

(mostly hydrophobic) that would normally become buried in the native conformation. As a result, these non-native intermediate states populated during the folding or unfolding process are prone to misfolding and have a high tendency to associate with each other and form aggregates with low solubility (Fig. 1.1) (Horwich, 2002). This non-native self-association process is even more complex, but at the same time more relevant, in the crowded environment of living cells. It is important to mention that the misfolding aggregates described above are different from protein precipitates that retain the native folded conformation like the ones formed by salting out or isoelectric precipitation (Fink, 1998).

Protein aggregation can also become a major hassle in biotechnology and in research laboratories. Inclusion body formation during protein overexpression in bacteria, aggregation during protein purification or during protein refolding from denaturant as well as during the protein lyophilization process, are probably the most familiar examples. But most importantly, protein misfolding and aggregation represents a major threat for living organisms, as exemplified by the large number of pathological conditions associated with severe human diseases.

Protein misfolding diseases

Some protein misfolding diseases arise as a result of defective folding leading to a decrease in the functional protein (e.g. cystic fibrosis (Thomas et al., 1995), and some types of cancer (Bullock and Fersht, 2001)) or improper trafficking (e.g. α_1-antitrypsin deficiency (Thomas et al., 1995)). However, the majority of the protein misfolding-related human diseases-commonly referred to as protein deposition diseases-are associated with the abnormal deposition of proteins or protein fragments within cells or

in the extracellular space. In general, when the aggregates are formed extracellularly, as it happens with the majority of the cases, they are referred to as amyloid deposits or plaques and the corresponding diseases are known as amyloidoses. When the amyloid-like aggregates are localized inside of the cell, they are often called intracellular inclusions. Aggresomes and Lewy or Russell bodies are well known examples of intracellular inclusion bodies. Disease-associated aggregates can accumulate in the brain of affected individuals as in the case of the neurodegenerative disorders: Alzheimer's disease, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis (ALS). Amyloid-like aggregate accumulation can also occur in a tissue other than the brain; in certain diseases it can occur in a single type of tissue as in the case of non-neuropathic localized diseases, or aggregation can be detected even in multiple tissues as in non-neuropathic systemic disorders. Cataract, type II diabetes, and inclusion-body myositis are examples of localized diseases, while senile systemic amyloidosis, lysozyme amyloidosis, and fibrinogen amyloidosis are examples of systemic non-neuropathic diseases (Chiti and Dobson, 2006).

The amyloid deposits associated with the overwhelming majority of the protein deposition diseases are characterized by highly stable and organized thread-like filamentous structures named amyloid fibrils (or fibers) made up of several laterally associated protofilaments. The amyloids associated with the various diseases are all enriched in β-sheets, and stabilized by intermolecular hydrogen bonds between them. Also, it is widely accepted that they all share a core "cross-β" architecture, where the β-strands are oriented perpendicular to the main axis of the fiber, while the resulting β-sheets run parallel to it (Sunde and Blake, 1997; Serpell, 2000). This core structure is

shared by different pathological deposits in spite of the fact that each disease involves a different protein or peptide as the main component of the aggregates. Those various proteins, all soluble in their native form, lack common features in primary amino acid sequence and range from globular proteins or protein fragments to natively unfolded proteins. Since the majority of pathological conditions are associated with the formation of amyloid fibers, often the main focus is on these ordered aggregates even though there are a number of diseases that involve non-fibrillar amorphous deposits (e.g. inclusion body myositis, cataracts).

In spite of its great significance, little is known about the exact mechanism of amyloid formation and the underlying cause of the clinical symptoms associated with the different pathological conditions. The process of amyloidosis is extremely complex and probably differs from type to type. However, it is widely accepted that in at least the majority of the cases, the amyloid aggregation is triggered by the accumulation of partly folded intermediates (Chiti and Dobson, 2006). Most importantly, for the majority of the deposition diseases, old age is considered a major risk factor. Some familiar amyloid diseases originate as a result of genetic mutation in the disease-related protein, others as a result of aberrant or incomplete proteolysis. In sporadic amyloid diseases, which constitute the majority of the protein deposition diseases, the proteins involved have the correct wild type sequence and the aggregation occurs due to changes in the cellular environment (Dobson, 2003). The prion diseases are unique since they can be transmitted by ingestion of pre-formed aggregates through a seeding mechanism (Prusiner, 1997).

In general, it is believed that fiber formation happens through a “nucleated growth mechanism” (Merlini and Bellotti, 2003) characterized by a lag phase followed by an exponential growth. The assembly process involves soluble, low-molecular-weight oligomeric precursors, the so called prefibrillar aggregates or protofibrils. The prefibrillar species can have different morphologies; they can form non-fibrillar spherical beads, amorphous small aggregates, and ring-shaped structures, although all seem to be enriched in β-sheet structures (Stefani and Dobson, 2003). Contrary to the original belief, which assumed that the mature amyloid fibrils are the toxic species, recent data suggest that the early aggregates are the major pathogenic species, at least in case of the neurodegenerative diseases. The toxicity is attributed to the ability of amyloid aggregates to interact with cellular components including membranes. Since the mature fibers are usually inert, it was even suggested that amyloid formation can be considered a protective mechanism used by cells to sequester the toxic small aggregates (Dobson, 2003; Chiti and Dobson, 2006). Nevertheless, in systemic disorders that are known to involve accumulation of large quantities of amyloid deposits (even kilograms), the pathological conditions can arise simply due to physical disruption of the affected tissues and organs (Merlini and Westermark, 2004).

Generic nature of protein aggregation

Amorphous and globular aggregates formed during recombinant protein production in bacteria and during *in vitro* refolding experiments, or thermal aggregation, are considered distinct from the highly ordered amyloid fibrils, although both are enriched β-sheet structures. It has been shown that some inclusion bodies and thermal

aggregates are able to bind the dyes typically used for detection of amyloid fibers (Goloubinoff et al., 1999; Ismail et al., 1992).

In general the ability to form both amyloid-like or amorphous aggregates is not restricted to the few proteins associated with clinical disorders, but is also observed in case of other non-disease related proteins under specific *in vitro* conditions (which destabilize the native state) (Fandrich et al., 2001; Stefani and Dobson, 2003). Recently, it has been proposed that the ability to form aggregates is a generic feature of the polypeptide chains that is rooted in the intrinsic physico-chemical properties of each polypeptide backbone (Dobson, 1999). Also, the same forces that determine globular protein structures also favor aggregation. However, the conditions under which a protein can be converted to this alternative conformation still depend on the properties of the individual polypeptide (amino acid composition and primary sequence). Indeed, it has been shown that hydrophobicity, secondary structure propensities and net charge of the protein sequences involved, are all playing and important role in the aggregation process. Aggregation is favored when the polypeptide sequence has high hydrophobicity, enhanced propensity for β-sheet formation and a low net charge (Chiti and Dobson, 2006; Chiti et al., 2003). Thus, while the amyloid structure is an extremely stable alternative conformation available to polypeptides, under normal physiological conditions the native conformation is still favored over the amyloid form.

Protein sequences evolved to minimize aggregation

It is widely accepted that protein sequences have evolved not only to generate a wide array of specific biological functions but also to avoid unwanted aggregation (Dobson, 2002). Sequences of intrinsically unfolded proteins provide an excellent

example of how evolution managed to produce proteins that are devoid of tertiary structure under physiological conditions and yet soluble. These proteins, when compared to folded proteins, possess higher net charge and lower overall hydrophobicity (Uversky, 2002).

Computer analysis of protein databases indicated that clusters of consecutive hydrophobic residues (Schwartz et al., 2001), but also alternating patterns of hydrophobic and hydrophilic residues (known to promote β-sheet formation like the ones in amyloid fibers) occur less frequently in naturally occurring protein sequences than it would be expected on a random basis (Broome and Hecht, 2000). It is undisputable though that both hydrophobic interactions and β-structures are major determinants of the native folds and globular structures.

It has been found that in general short aggregation-prone sequences occur often in globular proteins; however they are found more often in prokaryotes than in eukaryotes (Rousseau et al., 2006). Interestingly, the lower percentage of aggregation-prone residues in eukaryotic proteomes correlates with the higher prevalence of intrinsically unfolded proteins (characterized by lower aggregation propensity) in these systems. Since aggregation-prone sequences cannot be eliminated altogether, their subsequent aggregation is minimized by the so called gatekeeper residues. Thus strongly aggregating sequences across proteomes are often capped by prolines and glycines, but also by positively charged residues like lysine and arginine, which oppose aggregation (Rousseau et al., 2006; Otzen et al., 2000; Parrini et al., 2005).

Moreover, it has been also shown that the edges of different β-structures found in native structures are protected by “negative design” (i.e. contain strategically placed

charged residues, prolines, or β-bulges, are short or covered with loops or α-helices) in order to avoid unwanted intermolecular interaction between the β-sheet edges (Richardson and Richardson, 2002).

The charge of the residues is also a key parameter in modulating aggregation. It has been noted that the net charge of the overwhelming majority of native, folded proteins is within a narrow range (-10 to +10) (Lawrence et al., 2007). In general a net charge close to zero favors aggregation, but interestingly a large number of proteins with opposite net charges could also be aggregation-prone (Konno, 2001; Goers et al., 2003). Finally, as expected, there seems to be an evolutionary link between the aggregation propensities of polypeptide sequences and the expression levels of the corresponding genes. Indeed, it has been shown recently that a lower expression level of several human genes corresponds to a higher aggregation propensity of the protein (Tartaglia et al., 2007). This observation, although not unexpected, emphasizes the role of the cellular environment during the careful selection of protein sequences. The authors make an interesting point that the sequences are finely “tuned“ to stay soluble at a concentration that is required for the proper functioning of an organism, and there is no point for further perfecting those sequences (Tartaglia et al., 2007).

In summary, in a correctly functioning living organism, proteins are able to avoid aggregation and to fold into biologically active native forms in a well controlled cellular environment. This remarkable achievement can be attributed in part to the careful evolutionary selection of protein sequences, but also to the evolution of protein quality control network.

Molecular chaperones

In an attempt to keep the protein folding process under control and to prevent accumulation of harmful protein aggregates, all living systems developed quality control networks composed of molecular chaperones and proteases. Whereas molecular chaperones enhance the efficiency of *de novo* protein folding and promote reactivation of misfolded proteins, energy-dependent proteases are responsible for the degradation of irreversibly damaged, unsalvageable proteins (Lee and Tsai, 2005). As the prevalence of the protein deposition diseases illustrates, under certain extreme conditions the crucial safeguarding capability of the quality control system can be overwhelmed by the large amount of misfolded proteins. However, under physiological conditions the combined effects of these powerful tools ensure regular cell function and viability. Moreover, since protein misfolding is more likely to occur after heat shock and other cellular stress, the expression of molecular chaperones and proteases is increased transiently following various stress conditions. This sharp increase in the synthesis of a small set of proteins, the so called heat-shock proteins (Hsps), combined with a concomitant decrease in the synthesis of the rest of the cellular proteins is due to the activation of the heat-shock response (Smith et al., 1998; Lindquist, 1986).

A quite large and diverse group of proteins are known to function as molecular chaperones. The major evolutionarily conserved chaperone families, sorted according to the approximate molecular weight of the family members (in kDa) are the small heat-shock proteins (small Hsps), Hsp 40, Hsp60, Hsp70, Hsp90 and Hsp100 families. Often the chaperones are further divided based on their mode of action as “holders” and

"folders" (sometimes unfolders) (Dougan et al., 2002). Although in general they are known to recognize exposed hydrophobic regions in misfolded protein or folding intermediates, the individual chaperone families employ different mechanisms to ensure productive folding. For the most part, the different chaperone families and their functions are well characterized in bacterial systems. In higher organisms the homologous family members are more diverse and can be located in different cellular compartments (Fink, 1999).

Small Hsps (IbpA/IbpB in *E. coli*) are considered "holder" chaperones that form large oligomeric assemblies and bind non-native proteins at their surface. Thus they are able to passively prevent aggregation by transiently shielding the aggregation-prone regions, but they are unable to reactivate misfolded proteins (Fink, 1999). The refolding of the bound proteins requires assistance of other chaperones. Indeed, it has been shown that sHsps often cooperate with the ATP-dependent chaperones like Hsp70 and Hsp100 (Mogk et al., 2003a; Haslbeck et al., 2005). As opposed to the passive holding activity of sHsp family members, "folder" chaperones, like Hsp60 and Hsp70 are able to actively promote reactivation of misfolded proteins in an ATP-dependent manner.

The Hsp60 family member from *E. coli*, GroEL, also referred to as chaperonin, is the best described of all chaperones and provides a unique window into the intricate mechanism used by molecular chaperones. GroEL forms a large barrel-shaped complex, consisting of two back-to-back stacked rings, the central cavity of which recognizes and binds non-native substrates. Folding of the GroEL-bound polypeptide occurs inside of the enclosed chamber formed upon additional binding of ATP and the lid-like cochaperonin GroES. The lifetime of this complex is controlled by ATP

hydrolysis, as both the GroES and the encapsulated polypeptide are released upon ATP hydrolysis. Depending on substrates, multiple binding and release cycles may be necessary to achieve the correct native fold (Bukau and Horwich, 1998; Horwich, 2002).

Hsp70 is a very large family of chaperones, with most organisms containing multiple isoforms. Besides having a major role in protein folding, Hsp70 family members are involved in several other cellular functions like membrane translocation of secretory and organellar proteins, control of regulatory proteins like transcription factors (Fink, 1999). Just as in case of Hsp60s the best understood of the different Hsp70 members is DnaK from *E. coli*.

DnaK is composed of an N-terminal nucleotide-binding domain (ATPase domain) and a substrate-binding domain connected by a short linker. The C-terminal substrate-binding domain is composed of two sub-domains: a β-sandwich sub-domain that forms the substrate binding cleft, and a flexible α-helical region. Small hydrophobic segments exposed by non-native proteins interact with the β-sandwich sub-domain in an extended conformation, while the α-helical region functions as a lid and can open or close the substrate binding pocket. The movement of the lid is determined by the nucleotide state of DnaK. When ATP is bound to the nucleotide-binding domain the substrate-binding pocket is open and DnaK has low affinity for substrate. In the ADP-bound state the lid is closed over the substrate and DnaK has an increased affinity for the substrate. Two additional proteins regulate the cycling of DnaK between the low and the high affinity states: the Hsp40 cochaperone DnaJ and a nucleotide exchange factor, GrpE. DnaJ binds to the ATPase domain of DnaK and triggers ATP hydrolysis but it can also recruit substrates to DnaK, whereas GrpE mediates substrate release by catalyzing ADP

exchange to ATP (Bukau and Horwich, 1998; Mayer and Bukau, 2005; Hartl and Hayer-Hartl, 2002). The above described functional cycle suggests that the repeated binding and release of the misfolded proteins by the DnaK chaperone system (DnaK/DnaJ/GrpE) facilitates folding to the native conformation. However, the exact mechanism is still not elucidated although several models were proposed to explain the folding or refolding activity DnaK (Goloubinoff and De Los, 2007).

GroEL/GroES chaperonin and the DnaK system are perhaps the two best known and described examples of "conventional" chaperone systems that play a crucial role in preventing aggregation by facilitating folding of the newly synthesized proteins (co-and-posttranslatianally) and active refolding of the misfolded proteins.

For a long time it was thought that once misfolded proteins accumulate and start to aggregate the process is irreversible and those proteins cannot be reactivated. However, a milestone discovery by Lindquist and co-workers disproved this belief by showing that the large aggregates formed in yeast cells as a result of a non-lethal heat shock can be resolubilized and, most importantly, reactivated. These remarkable properties were observed only in *S. cerevisiae* cells expressing a major thermotolerance conferring protein named Hsp104, a member of the Hsp100 protein family (Parsell et al., 1994; Sanchez and Lindquist, 1990; Sanchez et al., 1992). Further *in vitro* studies demonstrated that Hsp104 in cooperation with the Hsp70 chaperone system from yeast (Ssa1 and Ydj1) is capable to dissociate and fully reactivate aggregated proteins (Glover and Lindquist, 1998) (also see Fig. 1.1).

Soon after the discovery of this bi-chaperone system in yeast, it was revealed that the bacterial homolog ClpB and the DnaK/DnaJ/GrpE system cooperate during the

disaggregation and reactivation of aggregates substrates (Zolkiewski, 1999; Goloubinoff et al., 1999; Motohashi et al., 1999). Surprisingly, ClpB/Hsp104 is found only in bacteria, plants and fungi.

AAA+ family of ATPases

The *E. coli* molecular chaperone ClpB (as well as the yeast Hsp104) belongs to the Hsp100/Clp subgroup of the broader AAA+ superfamily of ATPases associated with various cellular activities (Neuwald et al., 1999). AAA+ ATPases are a large group of proteins, often described as "molecular machines", which use the energy from ATP hydrolysis to induce conformational changes on macromolecular target substrates. This activity is at the center of many important biological processes. Indeed, as their name suggests AAA+ family members play an essential role in cellular processes as diverse as: protein quality control, membrane fusion and vesicular transport, organelle biogenesis, cytoskeletal regulation and DNA replication, repair and transcription regulation (Ogura and Wilkinson, 2001). The unifying feature of AAA+ family members is a structurally conserved 230-250 residue-long ATP-binding domain, named AAA+ module, with signature Walker A (GxxxxGK[T/S], x= any residue), Walker B (hhhhDE, h=hydrophobic residue), sensor-1 (T/N), arginine finger and sensor-2 (R/K) sequence motifs. This common AAA+ domain is built up from a core ATPase domain with the characteristic α/β Rossman fold and a smaller C-terminal mostly α-helical region. Based on the number of the AAA+ modules, AAA+ ATPases are further divided in class I and class II members. Class I ATPases contain two conserved AAA+ modules, while class I ATPases have only one. The attachment of various extra domains before or between the AAA+ modules, as well as interactions with different adaptor or partner proteins,

contributes to the structural and functional diversity of the AAA+ family members. Generally, AAA+ ATPases assemble into ring-or cylinder-shaped oligomers, mostly hexamers, with a narrow channel located at the center and with the nucleotide-binding sites located at the interfaces between subunits (Hanson and Whiteheart, 2005; Neuwald et al., 1999).

ClpB structural organization

ClpB is often described best, both structurally and functionally, when compared to the other well known bacterial Clp ATPases. In *E. coli,* the closest sequence homologs of ClpB are: ClpA, ClpX and HslU. Among them, from the structural point of view ClpB is more similar to ClpA, as they are both class I AAA+ members and contain a tandem repeat of two AAA+ modules, while ClpX and HslU belong to class II and contain only one AAA+ module (Fig. 1.2). However, only ClpB takes part in protein disaggregation, the other homologs function in protein degradation (Zolkiewski, 2006).

ClpB is a 95 kDa protein that, aside from the two nucleotide binding AAA+ domains (AAA-1 or D1 and AAA-2 or D2), contains an N-terminal domain as well as a quite large and unique middle domain (M domain) that links D1 to D2 (Fig. 1.2). The other three homologous Clp ATPases also possess additional domains either at the N-terminus, as in case of ClpA and ClpX, or as is the case of HslU, inserted into the AAA+ module. These accessory domains are involved in binding adaptor or partner proteins, and in recognition of substrates.

The recently solved crystal structure of the *Thermus thermophilus* ClpB (Lee et al., 2003) in the presence of a nonhydrolyzable ATP analog confirmed that each of the two AAA+ modules assume the conserved three-dimensional fold characteristic to

AAA+ ATPases, with the nucleotide binding site enclosed between the large α/β Rossman fold (D1-large; D2-large) and the small α-helical domain (D1-small; D2-small). It also showed that the N-domain is located on top of the AAA+ modules and is highly mobile. But most importantly, the structure revealed for the first time that the middle domain inserted into the first AAA+ module forms a long and mobile coiled coil (Fig. 1.3).

As bona fide AAA+ family members, the bacterial Clp ATPases including ClpB function as hexameric ring-shaped oligomers. Biochemical techniques like gel filtration and analytical untracentrifugation as well as imaging with TEM confirmed that ClpB in the presence of nucleotides forms hexameric ring-shaped oligomers (Zolkiewski et al., 1999; Akoev et al., 2004). However, high-resolution structural information is not available for the functionally relevant hexameric ClpB. In fact HslU is the only Clp ATPase that was successfully crystallized as a hexamer (Bochtler et al., 2000; Sousa et al., 2000).

The hexameric model of *T. thermophilus* ClpB was built based on the cryoEM image of the ClpB hexamer and the crystal structure of the ClpB monomer. This model further revealed that the D1 and D2 are oriented head-to-tail and give rise to the characteristic double ring structure with the narrow central pore. The inside of the channel defined by this pore is lined by the large sub-domains of D1 and D2, while the small sub-domains and located on the outside of the hexamer. Interestingly, the unique middle domains protrude axially from the outer surface of the oligomeric ring and thus resemble the shape of a “two-bladed propeller” (Lee et al., 2003; Lee et al., 2004). Although, this first reconstruction does not include the N-terminal domains, based on a

yet another hexameric model of ClpB (Diemand and Lupas, 2006) the N-terminal domains occupy the top of the double-ringed cylinder and thus surround the central pore (Fig. 1.4).

Mechanism of ClpB-mediated protein disaggregation

Even with structural information and extensive biochemical characterization at hand, the mechanism of ClpB mediated protein disaggregation is not fully understood. Recent evidence suggests that the mechanism of ClpB-mediated protein disaggregation involves unfolding of substrate proteins through ATP-dependent translocation across the narrow central channel of the ClpB oligomeric ring (Weibezahn et al., 2004). This threading mechanism is similar to the mechanism employed by the homologous Clp ATPases that function as regulatory particles of proteasome–like proteolytic systems. These hexameric ring-forming Clp ATPases (ClpA, ClpX and HslU) associate with the large barrel-shaped multimeric peptidases (ClpP and HslV) and control the passage of carefully selected substrates to the interior of the proteolytic chamber where degradation occurs in a secluded environment (Sauer et al., 2004; Mogk et al., 2004). The central channel of Clp ATPases as well as the entry portal into the peptidase chamber are too narrow to allow passage of native folded proteins. This implies that proteins destined for degradation are able to reach the active sites of the peptidase chamber only through the coaxially associated channel of the ring-shaped ATPases and only in an unfolded form. Indeed, it has been directly demonstrated that both ClpA and ClpX are able to completely unfold and translocate stable folded substrates directly to the associated ClpP peptidase for degradation (Weber-Ban et al., 1999; Singh et al., 2000; Reid et al., 2001). The translocation of polypeptides through the central channel

is facilitated by the ATP-binding and hydrolysis induced conformational changes of the oligomeric rings. Unstructured loops located inside of the narrow channel have been identified in many AAA+ proteins, and it is suggested that nucleotide-dependent movements of these loops may be enough to drive translocation (Hinnerwisch et al., 2005; DeLaBarre and Brunger, 2005; Wang et al., 2001).

These proteolytic systems provided a convenient set-up for testing the threading mechanism of ClpB. Although ClpB shares the basic ring-shaped architecture of ClpX and ClpA, ClpB is unable to interact with the ClpP peptidase since it is missing a conserved tripeptide that is known to facilitate such interaction. However, Bukau and co-workers engineered a ClpA-like ClpB variant, referred to as BAP, simply by replacing a helix-loop-helix loop motif located on the bottom surface of the ClpB oligomers with the corresponding region from ClpA that contains the conserved tripeptide essential for ClpP recognition (Weibezahn et al., 2004). Subsequently, the BAP variant of ClpB was able to associate with ClpP and moreover the resulting complex was proteolytically active. Thus the BAP-ClpP complex efficiently degraded soluble ClpB substrates but also aggregated proteins (Fig. 1.5). However, degradation of aggregated proteins required the additional presence of the DnaK system. Additionally, these experiments showed that the disaggregation and the degradation of a model aggregated substrate by the DnaK/BAP/ClpP are kinetically coupled events. Previous experiments confirmed that the disaggregation and the reactivation of model substrates also occur with similar kinetics during the ClpB/DnaK–mediated disaggregation (Schlieker et al., 2004a). Both observations are in agreement with the threading model of protein disaggregation, and

also emphasize that the continuous extraction of single polypeptides is coupled with their forced unfolding by translocation through the ClpB channel.

Cooperation of ClpB with the DnaK system

The experiment with the BAP variant of ClpB clearly demonstrated that the substrate threading through the narrow central pore of the oligomeric ClpB is an integral part of the disaggregation process. While this represents a major progress towards the understanding of the unique mechanism of disaggregation, it is only one piece of the puzzle. Clearly, the cooperation of ClpB with the DnaK chaperone system has an important, but so far not clearly defined role in protein solubilization. The fact that degradation of aggregated proteins by the BAP/ClpP complex still remains dependent on the DnaK system, while unfolded model ClpB substrates are successfully degraded by the same complex, implies that the DnaK system acts during the first steps of the disaggregation process, before substrate threading takes place (Weibezahn et al., 2004). Indeed, recent data confirmed that while unfolded substrates are able to bind to conserved channel loops that define the entrance to the ClpB channel, interaction of aggregated proteins with this site is observed only in presence of the DnaK system. Thus DnaK system might somehow facilitate the extraction of single polypeptide chains from the large aggregates and their insertion into the ClpB threading channel (Schlieker et al., 2004b; Haslberger et al., 2007).

This still does not rule out an additional role downstream of the translocation step; DnaK could take over the translocated proteins after their exit from the ClpB channel and might facilitate their refolding. It has been also noted that functional cooperation between ClpB and DnaK system occurs only between chaperones from the

same species, as the yeast Hsp104 is not able to cooperate with the bacterial DnaK system, and vice versa (Glover and Lindquist, 1998; Krzewska et al., 2001). This specificity of function points to a simultaneous action of the bi-chaperone system during the disaggregation process.

ClpB middle domain function

Another piece of the puzzle is the role of the middle domain (M domain) of ClpB in the disaggregation process. This ~120 amino-acid long domain forms the mobile coiled-coil structure that is located on the outer surface of the ClpB oligomeric ring, adjacent to the first AAA+ domain (Fig. 1.4). These propeller like domains are not only mobile in AMPPNP (adenosine 5`-(β,γ-imido)triphosphate)-bound state, but additionally they undergo large conformational changes during the different cycles of ATP binding and hydrolysis (Lee et al., 2003; Lee et al., 2007). The presence of this unique middle domain is the most important fingerprint of the ClpB structure. Indeed, the closely related ClpA is missing this domain and is still able to drive protein translocation and unfolding across its central channel. This suggests that perhaps this domain is not required for the basal translocation activity. However, when the M domains are deleted (Kedzierska et al., 2003; Mogk et al., 2003b) or their mobility is restricted (Lee et al., 2003), ClpB loses the ability to disaggregate proteins. The presence of the mobile M domain led to the formulation of yet another mechanism for ClpB-mediated protein disaggreagation. Besides the threading mechanism, it was proposed that the mobile M domains could pry apart the large protein aggregates or remodel their surface by a “crowbar” mechanism (Glover and Lindquist, 1998; Lee et al., 2003). So far there is no direct experimental evidence that would confirm such a mechanism. Due to their

positioning, M domains (located in close proximity to the large D1 sub-domain of the neighboring subunit) were proposed to facilitate nucleotide-dependent cooperative conformational changes in the ClpB hexameric ring (Lee et al., 2007).

Recent work that followed the effect of conserved single amino acid mutations within the M domain of ClpB, found that the M domain mutants were defective in activities that require assistance from the DnaK system, such as interaction of substrates with the ClpB pore sites and disaggregation of aggregated proteins. Accordingly, it was proposed that the M domains coordinate the cooperation of the ClpB and the DnaK chaperone systems, and are required for the efficient delivery of substrates to the central channel of the oligomeric ring (Haslberger et al., 2007). So far, attempts to detect a direct interaction between DnaK and ClpB failed.

ClpB N-terminal domain function

Thus the M domain of ClpB is a major contributor to the unique chaperone activity of ClpB, but also truly exemplifies how adding an attachment domain to an otherwise conserved core threading machine can achieve functional diversity. The N-terminal domain of ClpB is the other attachment domain that could contribute to the specificity of ClpB. However, the role of the N-terminal domain of ClpB is not fully understood and the experimental data available so far is controversial. Some reports indicated that the N-terminal domain of ClpB is dispensable for the *in vitro* protein disaggregation activity (Mogk et al., 2003b; Beinker et al., 2002) or for the development of thermotolerance *in vivo* (Clarke and Eriksson, 2000). At the same time other published reports have pointed out that this domain is relevant for ClpB activity, as the deletion of this domain, or single amino acid substitutions in the N-terminal domain

greatly influenced the disaggregation activity of ClpB as well as substrate binding and stimulation of ATPase activity by model substrates (Park et al., 1993; Barnett et al., 2000; Li and Sha, 2003; Liu et al., 2002; Tanaka et al., 2004).

Interestingly, a truncated version of ClpB that is lacking the whole N-terminal domain of the full-length ClpB is also produced *in vivo* (Park et al., 1993; Squires et al., 1991). The biological significance of production of two ClpB isoforms constitutes another mystery.

Even prior to the availability of a high resolution structure for ClpB, it was shown that the 149 amino acid-long N-terminal region of ClpB behaves as an independently folded stable domain that does not form strong contacts with the ClpB core AAA+ domains (Tek and Zolkiewski, 2002). The crystal structure of *T. thermophilus* ClpB revealed that the highly mobile N-terminal domains are connected to the first AAA+ domain through a short flexible linker and possess a surprisingly large degree of rotational freedom (Lee et al., 2003) (Fig. 1.3). The functional advantage of having such mobile domains located on top of the AAA+ core domains remains to be elucidated. It has been proposed that the N-terminal domains might play a role in controlling substrate specificity of ClpB and providing additional substrate binding sites (Barnett et al., 2000; Li and Sha, 2003; Barnett et al., 2005; Liu et al., 2002).

Mechanism of substrate recognition by ClpB

Although ClpB employs a similar translocation mechanism as the closely related Clp ATPases, they recognize different substrates: ClpB recognizes aggregated substrates while ClpA, ClpX and HslU recognize irreversibly damaged or misfolded proteins. While most proteins destined for degradation are tagged with specific

recognition signals, aggregated proteins are not tagged, but they are expected to contain a specific blend of amino acids on their surface that is not observed in native proteins. Although no such “disaggregation signal” was identified on aggregated proteins, peptide-binding studies indicated that ClpB binds preferentially to positively charged residues (Lys, Arg), and to a lower extent to aromatic residues (Phe,Tyr,Trp) (Schlieker et al., 2004b). The same peptide-binding studies indicated that the central pore that defines the entrance to the ClpB translocation channel (located in the first AAA+ domain) binds to the model peptide substrates. In particular, three residues, a tyrosine and two glutamic acids, were identified to support substrate binding to ClpB. However, these residues are located on the flexible loops that line the central pore and are conserved in many AAA+ ATPases that do not take part in protein disaggregation (Zolkiewski, 2006). Also as mentioned before, aggregated substrates are able to bind these loop regions only with help from the DnaK system. Thus the conserved residues may not be directly involved in the initial step of substrate recognition, instead they most probably play an important role in initiating the substrate threading through the central channel. The N-terminal domain of ClpB is a main candidate for the location of the substrate recognition sites. Indeed, recent results from our lab indicate that conserved residues from this region support binding of aggregated substrates (Barnett et al., 2005).

Goals of current research

The availability of a high resolution crystal structure of ClpB and the direct demonstration of the threading mechanism greatly enhanced our understanding of the unique aggregation reversing capability of the ClpB-DnaK chaperone system. However,

several major questions still remain unanswered. The majority of those questions are related to the initial steps of the disaggregation process, namely substrate recognition, binding and initiation of substrate threading, as well as the role of the DnaK chaperone system and the ClpB “attachment” domains throughout these processes.

The overall goal of current research was to gain a better understanding of the *E. coli* ClpB and of the unique process of protein solubilization and reactivation assisted by this chaperone. Throughout this work several aspects related to the structure and function of ClpB were investigated. Major emphasis was placed on the functional interplay between the two naturally occurring isoforms of ClpB during the disaggregation process.

References

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

Anfinsen,C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

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

Barnett,M.E., Zolkiewska,A., and Zolkiewski,M. (2000). Structure and activity of ClpB from *Escherichia coli*. Role of the amino- and -carboxyl-terminal domains. J. Biol. Chem. *275*, 37565-37571.

Beinker,P., Schlee,S., Groemping,Y., Seidel,R., and Reinstein,J. (2002). The N terminus of ClpB from *Thermus thermophilus* is not essential for the chaperone activity. J. Biol. Chem. *277*, 47160-47166.

Bochtler,M., Hartmann,C., Song,H.K., Bourenkov,G.P., Bartunik,H.D., and Huber,R. (2000). The structures of HsIU and the ATP-dependent protease HsIU-HsIV. Nature *403*, 800-805.

Broome,B.M. and Hecht,M.H. (2000). Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloidogenesis. J. Mol. Biol. *296*, 961-968.

Bukau,B. and Horwich,A.L. (1998). The Hsp70 and Hsp60 chaperone machines. Cell *92*, 351-366.

Bullock,A.N. and Fersht,A.R. (2001). Rescuing the function of mutant p53. Nat. Rev. Cancer *1*, 68-76.

Chiti,F. and Dobson,C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. *75*, 333-366.

Chiti,F., Stefani,M., Taddei,N., Ramponi,G., and Dobson,C.M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature *424*, 805-808.

Clarke,A.K. and Eriksson,M.J. (2000). The truncated form of the bacterial heat shock protein ClpB/HSP100 contributes to development of thermotolerance in the cyanobacterium *Synechococcus sp*. strain PCC 7942. J. Bacteriol. *182*, 7092-7096.

DeLaBarre,B. and Brunger,A.T. (2005). Nucleotide dependent motion and mechanism of action of p97/VCP. J. Mol. Biol. *347*, 437-452.

Diemand,A.V. and Lupas,A.N. (2006). Modeling AAA+ ring complexes from monomeric structures. J. Struct. Biol. *156*, 230-243.

Dobson,C.M. (1999). Protein misfolding, evolution and disease. Trends Biochem. Sci. *24*, 329-332.

Dobson,C.M. (2002). Getting out of shape. Nature *418*, 729-730.

Dobson,C.M. (2003). Protein folding and misfolding. Nature *426*, 884-890.

Dougan,D.A., Mogk,A., and Bukau,B. (2002). Protein folding and degradation in bacteria: to degrade or not to degrade? That is the question. Cell Mol. Life Sci. *59*, 1607-1616.

Fandrich,M., Fletcher,M.A., and Dobson,C.M. (2001). Amyloid fibrils from muscle myoglobin. Nature *410*, 165-166.

Fink,A.L. (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. Fold. Des *3*, R9-23.

Fink,A.L. (1999). Chaperone-mediated protein folding. Physiol Rev. *79*, 425-449.

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

Goers,J., Uversky,V.N., and Fink,A.L. (2003). Polycation-induced oligomerization and accelerated fibrillation of human alpha-synuclein *in vitro*. Protein Sci. *12*, 702-707.

Goloubinoff,P., Mogk,A., Zvi,A.P., Tomoyasu,T., and Bukau,B. (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.

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

Hartl,F.U. and Hayer-Hartl,M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. Science *295*, 1852-1858.

Haslbeck,M., Franzmann,T., Weinfurtner,D., and Buchner,J. (2005). Some like it hot: the structure and function of small heat-shock proteins. Nat. Struct. Mol. Biol. *12*, 842-846.

Haslberger,T., Weibezahn,J., Zahn,R., Lee,S., Tsai,F.T., Bukau,B., and Mogk,A. (2007). M domains couple the ClpB threading motor with the DnaK chaperone activity. Mol. Cell *25*, 247-260.

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

Horwich,A. (2002). Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. J. Clin. Invest *110*, 1221-1232.

Ismail,A.A., Mantsch,H.H., and Wong,P.T. (1992). Aggregation of chymotrypsinogen: portrait by infrared spectroscopy. Biochim. Biophys. Acta *1121*, 183-188.

Kedzierska,S., Akoev,V., Barnett,M.E., and Zolkiewski,M. (2003). Structure and function of the middle domain of ClpB from *Escherichia coli*. Biochemistry *42*, 14242-14248.

Konno,T. (2001). Amyloid-induced aggregation and precipitation of soluble proteins: an electrostatic contribution of the Alzheimer's beta(25-35) amyloid fibril. Biochemistry *40*, 2148-2154.

Krzewska,J., Langer,T., and Liberek,K. (2001). Mitochondrial Hsp78, a member of the Clp/Hsp100 family in *Saccharomyces cerevisiae*, cooperates with Hsp70 in protein refolding. FEBS Lett. *489*, 92-96.

Lawrence,M.S., Phillips,K.J., and Liu,D.R. (2007). Supercharging proteins can impart unusual resilience. J. Am. Chem. Soc. *129*, 10110-10112.

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

Lee,S., Sowa,M.E., Choi,J.M., and Tsai,F.T. (2004). The ClpB/Hsp104 molecular chaperone-a protein disaggregating machine. J. Struct. Biol. *146*, 99-105.

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

Lee,S. and Tsai,F.T. (2005). Molecular chaperones in protein quality control. J. Biochem. Mol. Biol. *38*, 259-265.

Li,J. and Sha,B. (2003). Crystal Structure of the *E. coli* Hsp100 ClpB N-Terminal Domain. Structure. (Camb.) *11*, 323-328.

Lindquist,S. (1986). The heat-shock response. Annu. Rev. Biochem. *55*, 1151-1191.

Liu,Z., Tek,V., Akoev,V., and Zolkiewski,M. (2002). Conserved amino acid residues within the amino-terminal domain of ClpB are essential for the chaperone activity. J. Mol. Biol. *321*, 111-120.

Mayer,M.P. and Bukau,B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol. Life Sci. *62*, 670-684.

Merlini,G. and Bellotti,V. (2003). Molecular mechanisms of amyloidosis. N. Engl. J. Med. *349*, 583-596.

Merlini,G. and Westermark,P. (2004). The systemic amyloidoses: clearer understanding of the molecular mechanisms offers hope for more effective therapies. J. Intern. Med. *255*, 159-178.

Mogk,A., Deuerling,E., Vorderwulbecke,S., Vierling,E., and Bukau,B. (2003a). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. Mol. Microbiol. *50*, 585-595.

Mogk,A., Dougan,D., Weibezahn,J., Schlieker,C., Turgay,K., and Bukau,B. (2004). Broad yet high substrate specificity: the challenge of AAA+ proteins. J. Struct. Biol. *146*, 90-98.

Mogk,A., Schlieker,C., Strub,C., Rist,W., Weibezahn,J., and Bukau,B. (2003b). 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.

Motohashi,K., Watanabe,Y., Yohda,M., and Yoshida,M. (1999). Heat-inactivated proteins are rescued by the DnaK.J-GrpE set and ClpB chaperones. Proc. Natl. Acad. Sci. U. S. A *96*, 7184-7189.

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

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

Otzen,D.E., Kristensen,O., and Oliveberg,M. (2000). Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly. Proc. Natl. Acad. Sci. U. S. A *97*, 9907-9912.

Park,S.K., Kim,K.I., Woo,K.M., Seol,J.H., Tanaka,K., Ichihara,A., Ha,D.B., and Chung,C.H. (1993). Site-directed mutagenesis of the dual translational initiation sites of the clpB gene of *Escherichia coli* and characterization of its gene products. J. Biol. Chem. *268*, 20170-20174.

Parrini,C., Taddei,N., Ramazzotti,M., Degl'Innocenti,D., Ramponi,G., Dobson,C.M., and Chiti,F. (2005). Glycine residues appear to be evolutionarily conserved for their ability to inhibit aggregation. Structure. *13*, 1143-1151.

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

Prusiner,S.B. (1997). Prion diseases and the BSE crisis. Science *278*, 245-251.

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

Richardson,J.S. and Richardson,D.C. (2002). Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc. Natl. Acad. Sci. U. S. A *99*, 2754-2759.

Rousseau,F., Serrano,L., and Schymkowitz,J.W. (2006). How evolutionary pressure against protein aggregation shaped chaperone specificity. J. Mol. Biol. *355*, 1037-1047.

Sanchez,Y. and Lindquist,S.L. (1990). HSP104 required for induced thermotolerance. Science *248*, 1112-1115.

Sanchez,Y., Taulien,J., Borkovich,K.A., and Lindquist,S. (1992). Hsp104 is required for tolerance to many forms of stress. EMBO J. *11*, 2357-2364.

Sauer,R.T., Bolon,D.N., Burton,B.M., Burton,R.E., Flynn,J.M., Grant,R.A., Hersch,G.L., Joshi,S.A., Kenniston,J.A., Levchenko,I., Neher,S.B., Oakes,E.S., Siddiqui,S.M., Wah,D.A., and Baker,T.A. (2004). Sculpting the proteome with AAA(+) proteases and disassembly machines. Cell *119*, 9-18.

Schlieker,C., Tews,I., Bukau,B., and Mogk,A. (2004a). Solubilization of aggregated proteins by ClpB/DnaK relies on the continuous extraction of unfolded polypeptides. FEBS Lett. *578*, 351-356.

Schlieker,C., Weibezahn,J., Patzelt,H., Tessarz,P., Strub,C., Zeth,K., Erbse,A., Schneider-Mergener,J., Chin,J.W., Schultz,P.G., Bukau,B., and Mogk,A. (2004b). Substrate recognition by the AAA+ chaperone ClpB. Nat. Struct. Mol. Biol. *11*, 607-615.

Schwartz,R., Istrail,S., and King,J. (2001). Frequencies of amino acid strings in globular protein sequences indicate suppression of blocks of consecutive hydrophobic residues. Protein Sci. *10*, 1023-1031.

Serpell,L.C. (2000). Alzheimer's amyloid fibrils: structure and assembly. Biochim. Biophys. Acta *1502*, 16-30.

Singh,S.K., Grimaud,R., Hoskins,J.R., Wickner,S., and Maurizi,M.R. (2000). Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. Proc. Natl. Acad. Sci. U. S. A *97*, 8898-8903.

Smith,D.F., Whitesell,L., and Katsanis,E. (1998). Molecular chaperones: biology and prospects for pharmacological intervention. Pharmacol. Rev. *50*, 493-514.

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

Squires,C.L., Pedersen,S., Ross,B.M., and Squires,C. (1991). ClpB is the *Escherichia coli* heat shock protein F84.1. J. Bacteriol. *173*, 4254-4262.

Stefani,M. and Dobson,C.M. (2003). Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. J. Mol. Med. *81*, 678-699.

Sunde,M. and Blake,C. (1997). The structure of amyloid fibrils by electron microscopy and X-ray diffraction. Adv. Protein Chem. *50*, 123-159.

Tanaka,N., Tani,Y., Hattori,H., Tada,T., and Kunugi,S. (2004). Interaction of the N-terminal domain of *Escherichia coli* heat-shock protein ClpB and protein aggregates during chaperone activity. Protein Sci. *13*, 3214-3221.

Tartaglia,G.G., Pechmann,S., Dobson,C.M., and Vendruscolo,M. (2007). Life on the edge: a link between gene expression levels and aggregation rates of human proteins. Trends Biochem. Sci. *32*, 204-206.

Tek,V. and Zolkiewski,M. (2002). Stability and interactions of the amino-terminal domain of ClpB from *Escherichia coli*. Protein Sci. *11*, 1192-1198.

Thomas,P.J., Qu,B.H., and Pedersen,P.L. (1995). Defective protein folding as a basis of human disease. Trends Biochem. Sci. *20*, 456-459.

Uversky,V.N. (2002). Natively unfolded proteins: a point where biology waits for physics. Protein Sci. *11*, 739-756.

Wang,J., Song,J.J., Seong,I.S., Franklin,M.C., Kamtekar,S., Eom,S.H., and Chung,C.H. (2001). Nucleotide-dependent conformational changes in a protease-associated ATPase HslU. Structure. (Camb.) *9*, 1107-1116.

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

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

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.

Zolkiewski,M. (2006). A camel passes through the eye of a needle: protein unfolding activity of Clp ATPases. Mol. Microbiol. *61*, 1094-1100.

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

Figures

Figure 1.1 Model of protein folding and aggregation.

Partly folded intermediates can be populated during *de novo* protein folding or upon stress-induced unfolding of native proteins. The partly folded intermediates have a high tendency to associate with each other and form aggregates that have low solubility. Classical molecular chaperones (like GroEL/ES and DnaK in bacteria) promote productive folding and prevent aggregation. Chaperones from Hsp100 family are able to revere the aggregation process.

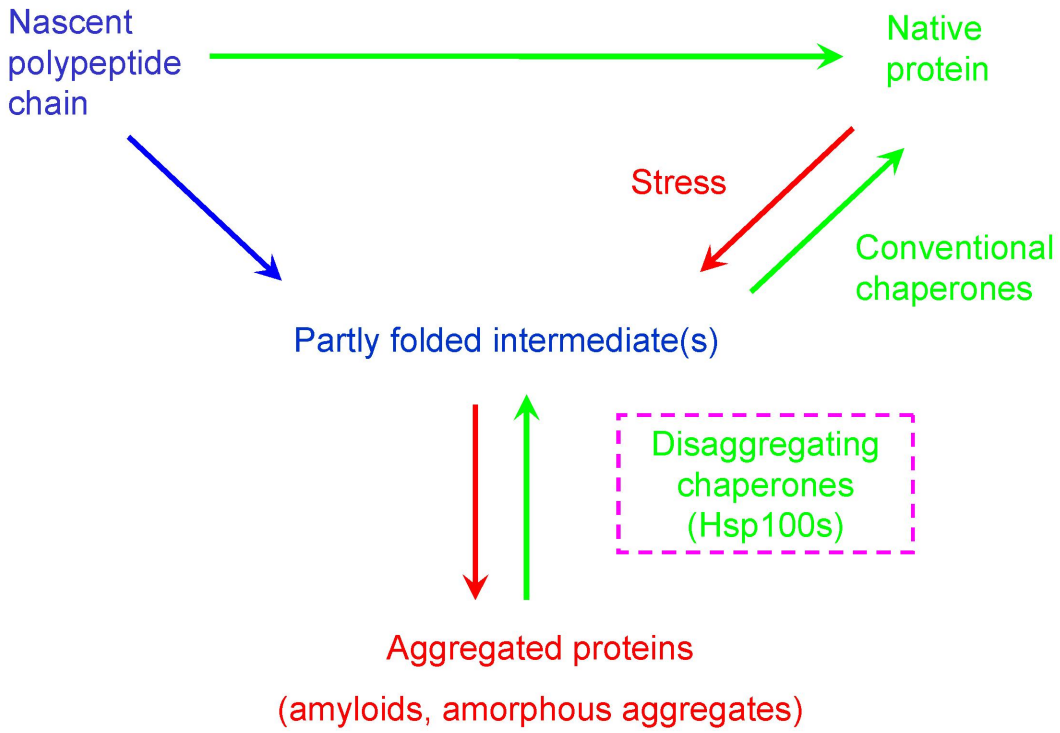


Figure 1.2 Domain organization of major bacterial Clp ATPases.

ClpB and ClpA contain two conserved AAA+ modules (AAA-1 and AAA-2) while ClpX and HslU (ClpY) contain only one AAA+ module that resembles the second AAA+ module of ClpA and ClpB. The less-conserved N-terminal (N) domains and the unique middle domain of ClpB (M), as well as the insertion domain of HslU (I) are also shown.

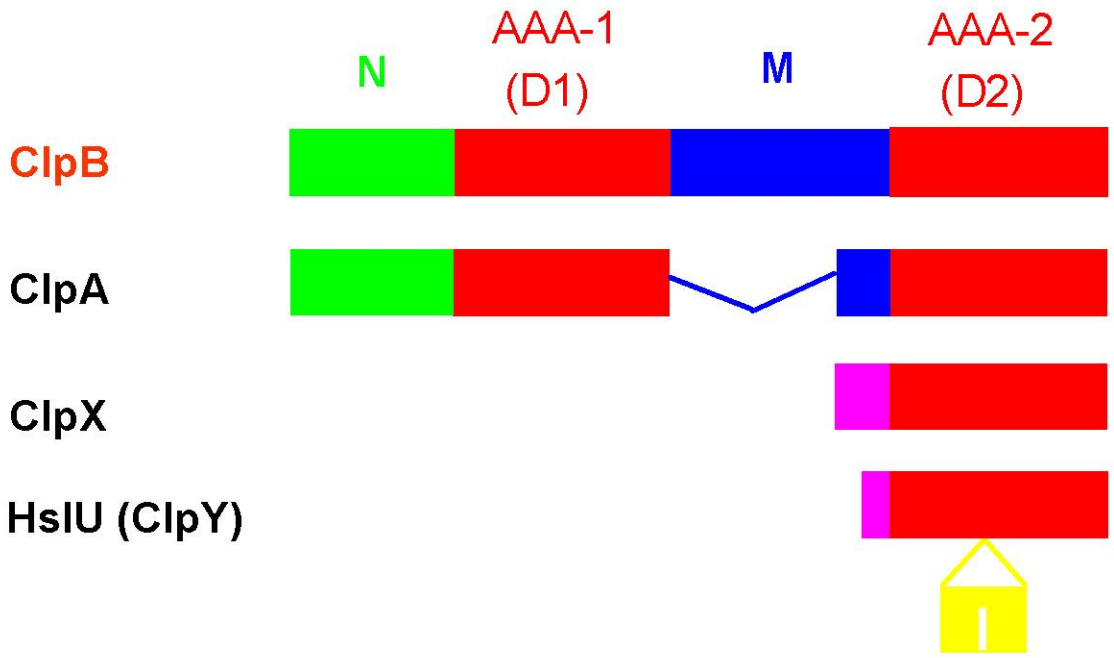


Figure 1.3 Crystal structure of ClpB from *Thermus thermophilus.*

The ribbon drawing of the ClpB monomer (PDB code: 1qvr) (Lee et al., 2003). The N-terminal domain is located on top of the AAA+ modules (ND, colored in green), the large α/β (D1-large, D2-large) and the small α-helical (D1-small, D2-small) sub-domains of the conserved AAA+ modules are shown in orange and magenta, respectively. The unique coiled coil middle domain (MD) is inserted into the small sub-domain of the first AAA+ module (D1-small) and is shown in blue. AMPPNP is shown as CPK model.

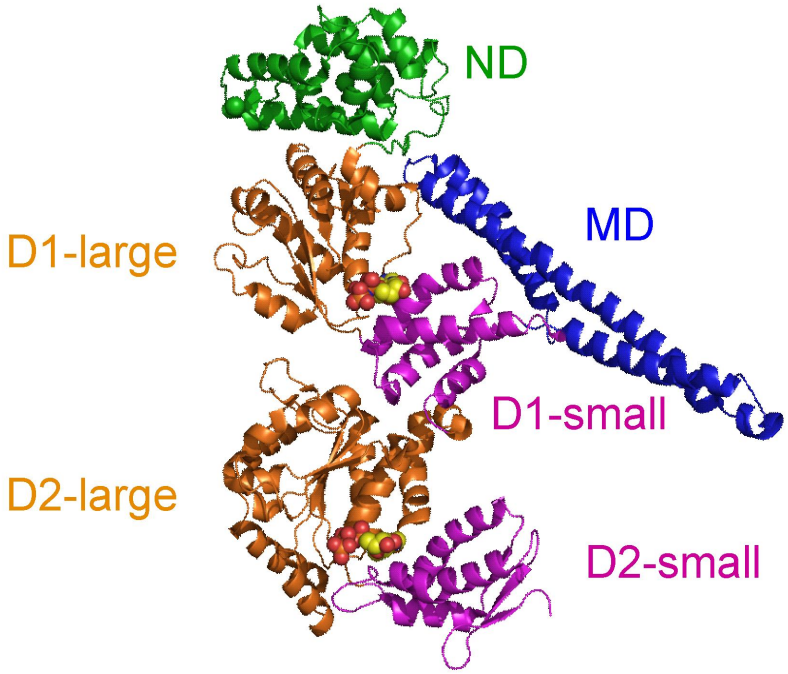


Figure 1.4 Hexameric model of *Thermus thermophilus* ClpB.

Side-view (A) and top-view (B) of the hexameric model of ClpB is colored by subunits (Diemand and Lupas, 2006). The two AAA+ domains are oriented head-to-tail (A) and form the core of the oligomeric ring and define the narrow central channel (B). The N-terminal domains are located on top of the double-ringed cylinder right above the first AAA+ modules (A) and surround the channel entrance. The coiled coil middle domains protrude axially from the outer surface of the oligomeric ring and resemble the shape of a two-bladed propeller.

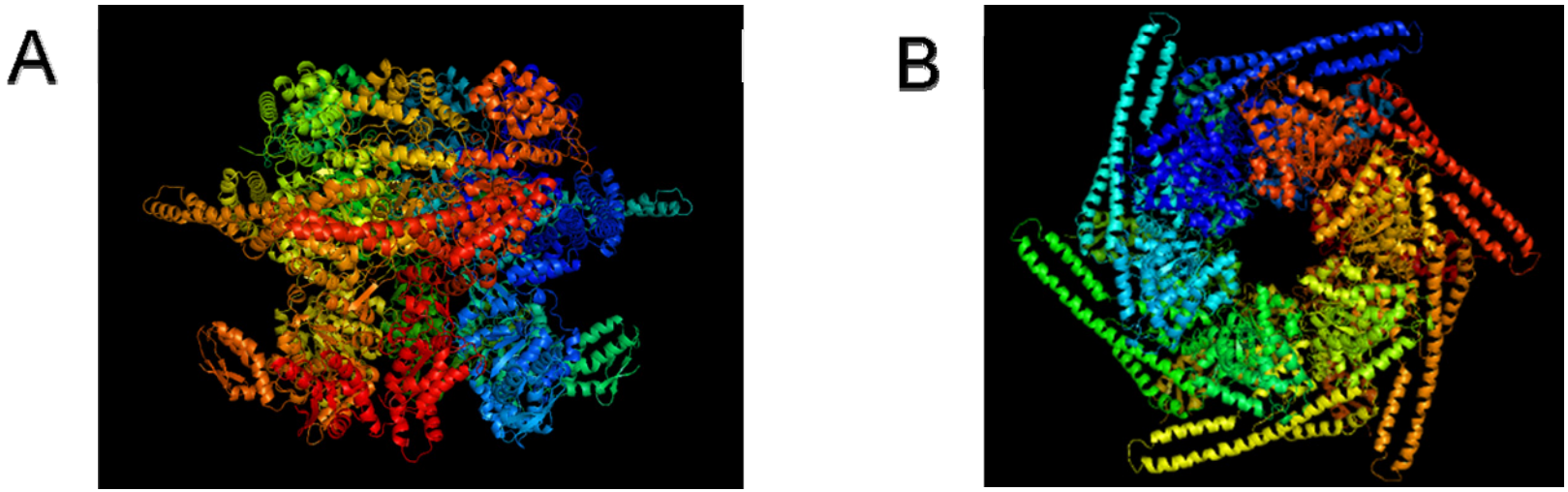
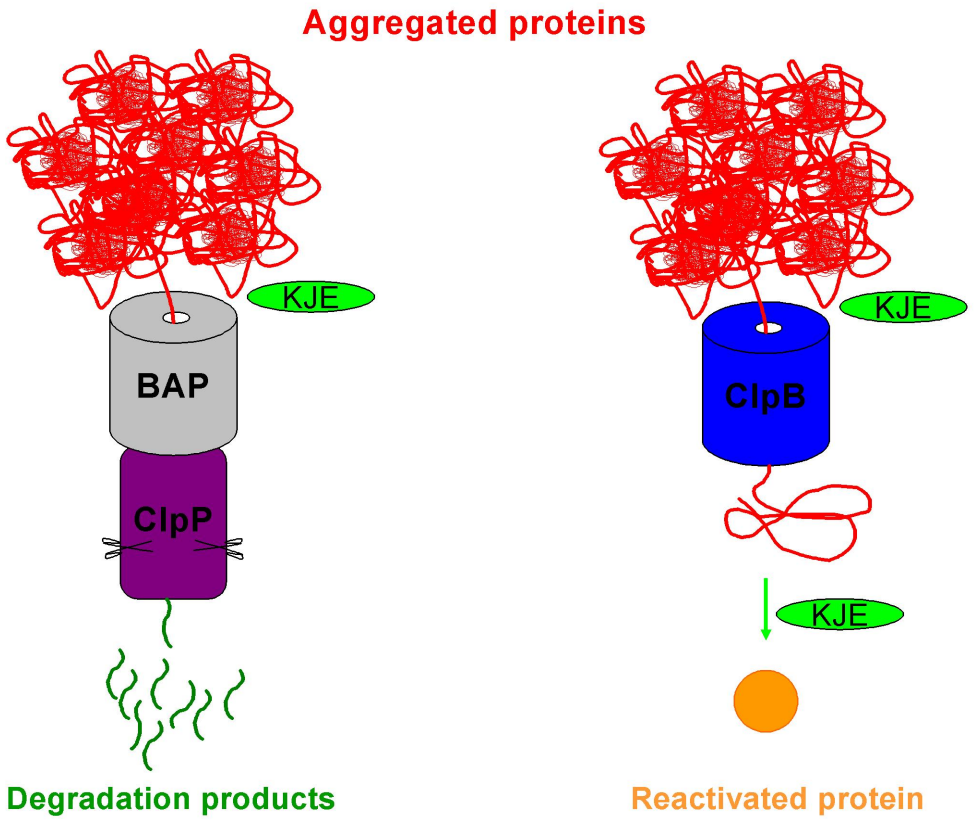


Figure 1.5 Threading model of the ClpB-mediated protein disaggregation. Left side: set up used by Bukau and coworkers (Weibezahn et al., 2004) for the demonstration of the threading mechanism of ClpB. An engineered ClpB with a ClpA-like ClpP binding loop (BAP) associated with the ClpP peptidase and degraded aggregated substrates in additional presence of DnaK/DnaJ/GrpE (KJE). Right side: Mechanism of ClpB-mediated protein disaggregation involves extraction of single polypeptides from aggregates and their unfolding by ATP-dependent translocation/threading across the narrow central channel of the ClpB oligomeric ring.



CHAPTER 2 - Homology modeling of *E. coli* ClpB

This part of my research was performed in collaboration with Dr. Maria I. Zavodszky, Department of Biochemistry and Molecular Biology, Michigan State University

Introduction

It is indisputable that the availability of three-dimensional structural information greatly benefits the functional characterization of proteins. AAA+ ATPases function as large ring-shaped complexes (Neuwald et al., 1999), thus NMR spectroscopy-based structure determination is not an option for the proteins belonging to this superfamily. Instead, X-ray crystallography was successfully used for the determination of the three-dimensional structures of many AAA+ family members. These proteins often crystallized as monomers or non-physiological oligomers; only very few AAA+ family members crystallized in the functionally relevant oligomeric ring form.

High resolution crystal structures are available for proteins related to ClpB, such as the bacterial Clp ATPases ClpA (PDB code1ksf, (Guo et al., 2002)) and ClpX (PDB code 1um8, (Kim and Kim, 2003)) as well as for HslU(ClpY) (PDB code 1do0, (Bochtler et al., 2000); PDB code 1q3i, (Sousa et al., 2000)). Among them, only HslU crystallized in the hexameric form. In the case of ClpB, crystal structures were reported for the isolated N-terminal domain (PDB code 1khy, (Li and Sha, 2003)) as well as for the large sub-domain of the first AAA+ module (D1 large) from *E. coli* (PDB code1jbk, (Li and

Sha, 2002)). Structure of the full-length ClpB is only available for the *Thermus thermophilus* ClpB (PDB code 1qvr, (Lee et al., 2003)).

The structure of the isolated N-terminal domain (amino acids 1-148) of *E. coli* ClpB was determined at 1.9 Å resolution (Li and Sha, 2003). The crystallographic asymmetric unit contained four identical N-domain monomers, each consisting of eight α helices and two small β strands that form a single domain (Fig. 2.1). It was also noted that the structure of the N domain possesses a pseudo two-fold symmetry where the two sub-domains are connected via a flexible region. The flexible linker region is characterized by high temperature factors and, indeed, parts of the linker region were not resolved (amino acids 75-81) in the crystal structure. The two structurally similar sub-domains correspond to the previously identified sequence repeats in the N-domain of the Clp ATPases (Lo et al., 2001). However, the two sequence repeats within the N-domain do not define separate folding domains.

The second *E. coli* ClpB fragment with crystal structure deposited in the Protein Data Bank contains the large subunit of the first AAA+ module of ClpB (amino acids 159-351) (Li and Sha, 2002). This isolated fragment crystallized as a monomer and is missing a short, probably disordered segment (between amino acids 287 and 292). This structure was solved at 1.8 Å resolution and was found to assume the conserved three-dimensional α/β Rossmann fold of the AAA+ ATPases (Fig. 2.2). Since the crystallized fragment only contains the large D1 subunit and a small part of the first α helix of the small D1 subunit, it is not surprising that the crystal structure does not include the bound nucleotide. An important feature of the structure is that it contains the conserved pore

loop (residues 243-254) known to line the entrance of the channel in the hexameric AAA+ rings.

The crystal structure of the full-length *T. thermophilus* ClpB in complex with the non-hydrolysable ATP analog AMPPNP was solved at 3 Å resolution (Lee et al., 2003). Three independent ClpB monomers were identified in the asymmetric unit of the crystal, each containing two AMPPNP molecules (one per AAA+ domain). Even though all three monomers were present in the same nucleotide-bound state, they were found to assume different conformational states in the crystallographic asymmetric unit. Superposition of the backbones of the three monomers showed the largest differences between the conformations of the N-terminal domains, suggesting these domains are highly mobile and can undergo rotational motions up to 120˚ (Fig. 2.3). The crystal structures provided several other pieces of useful information regarding domain organization of ClpB, as described in the introduction section of this dissertation. Interestingly, the non-physiological trimers did not assemble into a complete hexamer, instead they formed a helical structure. A hexameric model of *T. thermophilus* ClpB was built by fitting the crystal structure of one of the three representations of the monomeric ClpB into the three-dimensional envelope of the ClpB hexameric ring obtained by cryo-EM (Lee et al., 2003). In this reconstituted hexamer, the subunits are organized in such a way that the D1 large sub-domain is positioned right above the D2 large sub-domain of the same monomer. However, this hexameric model does not contain the N-terminal domains of ClpB, since these domains were not visible in the single-particle reconstruction. Also, the conserved pore loops of D1 (residues 235-245) and D2 (residues 637-650) AAA+ domains, as well as an additional short segment (residues

270-290) are missing from the model since they were not resolved in the crystal structure.

Recently a new approach was developed for the modeling of the AAA+ type ring complexes from monomeric structures (Diemand and Lupas, 2006). This semi-automated modeling procedure relies on structural constraints collected from the atomic structures of AAA+ family members that crystallized in the functional oligomeric form, and combines this information with sampling algorithm and protein docking until an acceptable oligomeric model is reached. This new procedure was used to build a hexameric model of the full-length ClpB starting from a monomeric crystal structure derived (chain B from 1qvr) from the non-physiological trimer of the *T. thermophilus* ClpB (Lee et al., 2003). This model is also missing the conserved pore loops (Fig. 2.4), but includes the N-terminal domains and proposes a different arrangement for the two AAA+ rings relative to the previous model. Based on this modeling, a staggered arrangement was proposed where the D1 domain of one subunit is above the D2 domain of the neighboring subunit. However, this model was built by separately docking the D1 and D2 domains into hexameric rings before building the characteristic double-ring structure and as a result the structure contains a chain-discontinuity site at the interface of the two rings (residue 543-544) as shown is Fig. 2.4. This procedure is available online under iMolTalk server (http://i.moltalk.org/), thus we are going to refer to this hexameric model of ClpB (Diemand and Lupas, 2006) as the iMolTalk hexameric model.

Aside from the structure of the isolated N-terminal domain and the D1 large sub-domain fragment, there is no high-resolution structure available for the full-length *E.coli*

ClpB. However, the high sequence identity between *E. coli* ClpB and the *T. thermophilus* and the availability of the crystal structures of the two segments mentioned above make *E. coli* ClpB a good candidate for homology modeling.

Programs

Sequences of *E. coli* ClpB and *T. thermophilus* ClpB were aligned using the BOXSHADE program (www.ch.embnet.org). The percent of identical residues was calculated manually. The atomic coordinates of the iMolTalk hexameric model of *T.thermophilus* ClpB were downloaded from http://i.moltalk.org/, while the coordinates of the N-terminal domain of *E. coli* ClpB (PDB code 1khy), the large D1 domain of *E. coli* ClpB (PDB code 1jbk), and of the non-physiological trimer of *T.thermophilus* ClpB (PDB code1qvr) were downloaded from the Protein Data Bank. *E. coli* ClpB with the NCBI accession number P63284 was used as a target sequence for modeling. Target and template sequences were aligned by using the ClustalW server (www.ch.embnet.org), while PyMOL v0.99 (DeLano Scientific LLC) was used for the visualization of the 3D structures. Restraint-based homology modeling was performed using MODELLER version 9.1 (Sali and Blundell, 1993; Sali and Overington, 1994; Fiser et al., 2000). This program was run on a 64-bit Linux desktop computer with dual Xeon processors.

Homology modeling with MODELLER

Comparative or homology modeling is known to give a fairly accurate three-dimensional model of a target protein if it has at least 30 % sequence identity to the template sequence whose atomic structure is available. In general, the quality of the model increases with the increasing sequence similarity between the target and template, but decreases as the number and the length of the gaps in the aligned

sequences increases. A model based on greater than 50% identity usually has about 90% of the backbone in the right conformation with only some minor errors in loop conformations and side-chain packing. The quality of such model is comparable to a low-resolution crystal structure (above 3 Å) or a medium-resolution NMR structure (Sali and Blundell, 1993). These models proved to be useful, for example, for designing mutations to test hypotheses regarding protein function (Muller et al., 2006), rationalizing the effects of mutations (Marabotti and Facchiano, 2005), locating binding sites (Indarte et al., 2008), identifying the determinants of substrate specificity (Sauder et al., 2000), structure-based drug development (Takeda-Shitaka et al., 2004), and predicting protein-protein interactions (Davis et al., 2007).

E. coli ClpB shares 56 % of its sequence with ClpB from *T.thermophilus,* as calculated from the alignment shown in Fig. 2.5. The sequence identity is highest in the AAA+ domains and somewhat lower in the N-terminal region and in the M domain. To increase the accuracy of the model in regions of lower sequence identity, instead of relying solely on the crystal structure of the full-length *T. thermophilus* ClpB, we also included the other two existing crystal structures of the *E. coli* ClpB fragments as templates for the modeling. There are several completely automated web servers available to perform homology modeling. However we used the downloaded, command-line version of MODELLER that allows model building based on multiple template structures and *ab initio* loop modeling. Since all three template crystal structures were missing short disordered segments, we performed loop refinement on those segments with the automated loop modeling module of MODELLER (Fiser et al., 2000).

As the first step in the model building, the *E. coli* ClpB target sequence was aligned with the template sequences. Although MODELLER has its own module for sequence alignment, we obtained the best results (highest sequence identity and fewest gaps) when ClustalW was used to perform pairwise alignments of the target with the three template sequences (*T. thermophilus* full-length sequence; *E. coli* N-terminal domain sequence; *E. coli* large D1 domain sequence). The pairwise alignments were then combined to get the multiple sequence alignment in a PIR database format. Further manual adjustments were done to remove the residues from the template sequences whose atomic coordinates are not resolved in the crystal structures.

In addition to the sequence alignment of the target and three templates, the three PDB files containing the 3D coodinates of the template atoms were also used as input to MODELLER. For the full-length *T. thermophilus* ClpB, the coordinates of the ClpB monomer were extracted from the iMolTalk hexamer. The PDB file containing the D1 large sub-domain of *E. coli* ClpB (PDB code 1jbk) was used without modifications, but the PDB file that contained the *E.coli* ClpB N-terminal fragment (PDB code 1khy) was modified by retaining the coordinates of only one of the four chains.

MODELLER builds 3D models of the target sequence by a method that relies on the satisfaction of spatial restraints extracted from the template structures, followed by an optimization method to minimize the violations of these restraints (Sali and Blundell, 1993). The resulting models are evaluated and scored. We kept the model of *E. coli* ClpB with the best score (i.e. the lowest value of the MODELLER objective function) for further refinement. Since all three template crystal structures were missing short segments, we performed loop refinement on those segments with the *de novo* loop

modeling module of MODELLER. This extra step was especially important in order to optimize the conformation of the D2 pore loop of ClpB (residues 647-660) since the full-length *T. thermophilus* ClpB template structure was missing the conserved pore loops. The model of full-length *E. coli* ClpB monomer with the best score after the loop refinement was used to build the hexameric model of the *E. coli* ClpB. This was done by superimposing the model of the monomeric *E. coli* ClpB over each subunit of the iMolTalk hexamer of *T. thermophilus* ClpB. Optimal packing at the subunit interfaces was achieved by full-length backbone superposition, which was performed in PyMOL. The coordinates of these superimposed monomers were merged into a hexamer.

The resulting hexameric model was further optimized by a short (500-step) energy minimization. This was done with the CHARMM force field using the Generalized Born (GB) solvent model. Following the energy minimization, no major van der Waals overlaps were detected at the subunit interfaces.

Results and Discussion

The cartoon representation of the final model of the *E. coli* ClpB in the hexameric form is presented in Fig. 2.6. The monomer is shown superimposed over chain A of the iMolTalk model of *T. thermophilus* ClpB in Fig. 2.7. The backbones of the two models are in very similar conformations, with a root mean square deviation of 1.96 Å between them. The quality of the model was evaluated using the PROCHECK program (Laskowski et al., 1993) by generating the Ramachandran plot of backbone phi-psi angles. This is a simple and sensitive way to asses the quality of a structural model in the absence of experimental data. For a good quality structure, about 90% of the residues are expected to lie in the most favorable regions of phi-psi values. The

Ramachandran plot of the *E. coli* ClpB homology model after the energy minimization is shown in Fig. 2.8. The final model had more than 88% of the residues in the most favorable regions, while another 11% fell into the additionally allowed regions. Only 1% of the residues had unfavorable backbone conformations with backbone phi-psi values in the disallowed regions (Fig. 2.7 and Fig. 2.8). Unexpectedly, these residues were scattered throughout the 3D structure of the monomer and were mainly located in various surface loops as shown in Fig. 2.7. For some of the residues the unfavorable conformation arises simply because the corresponding residues from the template are also in an unfavorable backbone conformation (as for Arg119 and Ala589) or are missing from the crystal structure used as a template (as for Lys250 and Asp290). In other cases considerable differences between the length and the conformation of short regions of the template and the model could be the reason for the unfavorable backbone conformation of the residues that fall into the disallowed region of the Ramachandram plot (as for Ser45, Asp696, Glu731, Glu768). Ironically this could be attributed to the high sequence identity between the *E.coli* and the *T. thermophilus* ClpB (Fig. 2.5). As the two amino acid sequences have large regions with very high identity and only short connecting sequence regions with large differences, the MODELLER program tried to fit very closely the regions with high sequence identity; this resulted in tension or unfavorable backbone conformation for some phi-psi angles of the residues located in the connecting regions that are different in template and query sequence. The short energy minimization intended to remove major van der Waals overlaps following the model building could not get the residues from the disallowed region over the energy barrier to a more favorable conformation. A long molecular dynamics (MD)

simulation would be needed to solve this problem, but it would be computationally very expensive and time consuming.

Overall this hexameric model of *E.coli* ClpB is a structure of reasonable quality. As a comparison, models built from crystal structures solved at least at 2 Å resolution and an R-factor of not more than 20%, are expected to have ~ 90% of the residues in the most favored regions (Laskowski et al., 1993).

In our new model of the hexameric *E. coli* ClpB the D1 and D2 AAA+ domains are arranged in a staggered conformation similar to the conformation observed in case of the iMolTalk hexameric model (the only full length template structure used for modeling) and different from the conformation proposed for the N-terminally truncated *T. thermophilus* hexameric model where the D2 in right below the D1 in each subunit (Lee et al., 2003). Staggered arrangement of the two AAA+ rings was also proposed previously for the ClpA hexameric model (Guo et al., 2002).

Compared to the iMolTalk hexamer of *T. thermophilus* ClpB, our model of the *E. coli* hexamer does not contain the chain-discontinuity at the interface of the two AAA+ rings. The other advantage of this model is that the pore loop regions of both the D1 and D2 AAA+ domains are included in the structure. The conserved Tyr residues, Tyr251 and Tyr653 located in the highly flexible pore loops of each of the six subunits were shown to directly contact the substrate proteins (Weibezahn et al., 2004). Tyr251 lines the entrance to the ClpB channel while Tyr653 is located further down in the channel. Interestingly, in the hexameric model, the upper six D1 pore loops are not in a completely symmetrical arrangement, as shown by the orientation of the six Tyr251 residues (Fig. 2.9). This is the result of the final refinement and it is in agreement with

the fact that these loops are extremely mobile. Indeed, it was proposed that nucleotide-dependent movements of the flexible loops located inside of the ClpB channel drive translocation of substrates through the ClpB channel (Park et al., 2005; Hinnerwisch et al., 2005). In our model, it looks like the approximately 14 residue long D2 loops (residues 647-660 in each subunit) are in upward positions, very close to the D1 ring (Fig. 2.10). It has been suggested previously that ATP hydrolysis by the D2 ATPase domain can induce a downward movement of this loop which in turn may drive the translocation of the polypeptides through the narrow pore of the oligomer (Hinnerwisch et al., 2005).

Finally, the diameter of the pore at the channel entrance defined by the Tyr251 residues is approximately 15 Å (measured in PyMOL). This value is consistent with the previously published diameter of the nucleotide-bound oligomeric ring approximated from the TEM images (Akoev et al., 2004; Lee et al., 2003). The weakness of our model is that it does not contain bound nucleotides. The full-length *T. thermophilus* ClpB template contained ATP analogs (AMPPNP), but the structure of the large D1 domain of *E. coli* ClpB was solved without a bound nucleotide. Overall we obtained a good quality model of the hexameric full-length *E. coli* ClpB, which will be useful in designing mutants to explore the mechanism of this protein and in interpreting experimental data.

References

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

Bochtler,M., Hartmann,C., Song,H.K., Bourenkov,G.P., Bartunik,H.D., and Huber,R. (2000). The structures of HsIU and the ATP-dependent protease HsIU-HsIV. Nature *403*, 800-805.

Davis,F.P., Barkan,D.T., Eswar,N., McKerrow,J.H., and Sali,A. (2007). Host pathogen protein interactions predicted by comparative modeling. Protein Sci. *16*, 2585-2596.

Diemand,A.V. and Lupas,A.N. (2006). Modeling AAA+ ring complexes from monomeric structures. J. Struct. Biol. *156*, 230-243.

Fiser,A., Do,R.K., and Sali,A. (2000). Modeling of loops in protein structures. Protein Sci. *9*, 1753-1773.

Guo,F., Maurizi,M.R., Esser,L., and Xia,D. (2002). Crystal Structure of ClpA, an Hsp100 Chaperone and Regulator of ClpAP Protease. J. Biol. Chem. *277*, 46743-46752.

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

Indarte,M., Madura,J.D., and Surratt,C.K. (2008). Dopamine transporter comparative molecular modeling and binding site prediction using the LeuT(Aa) leucine transporter as a template. Proteins *70*, 1033-1046.

Kim,D.Y. and Kim,K.K. (2003). Crystal structure of ClpX molecular chaperone from *Helicobacter pylori*. J. Biol. Chem. *278*, 50664-50670.

Laskowski,R.A., MacArthur,M.W., Moss,D.S., and Thornton,J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography *26*, 283-291.

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

Li,J. and Sha,B. (2002). Crystal structure of *E. coli* Hsp100 ClpB nucleotide-binding domain 1 (NBD1) and mechanistic studies on ClpB ATPase activity. J. Mol. Biol. *318*, 1127-1137.

Li,J. and Sha,B. (2003). Crystal Structure of the *E. coli* Hsp100 ClpB N-Terminal Domain. Structure. (Camb.) *11*, 323-328.

Lo,J.H., Baker,T.A., and Sauer,R.T. (2001). Characterization of the N-terminal repeat domain of *Escherichia coli* ClpA-A class I Clp/HSP100 ATPase. Protein Sci. *10*, 551-559.

Marabotti,A. and Facchiano,A.M. (2005). Homology modeling studies on human galactose-1-phosphate uridylyltransferase and on its galactosemia-related mutant Q188R provide an explanation of molecular effects of the mutation on homo- and heterodimers. J. Med. Chem. *48*, 773-779.

Muller,T.A., Zavodszky,M.I., Feig,M., Kuhn,L.A., and Hausinger,R.P. (2006). Structural basis for the enantiospecificities of R- and S-specific phenoxypropionate/alpha-ketoglutarate dioxygenases. Protein Sci. *15*, 1356-1368.

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

Park,E., Rho,Y.M., Koh,O.J., Ahn,S.W., Seong,I.S., Song,J.J., Bang,O., Seol,J.H., Wang,J., Eom,S.H., and Chung,C.H. (2005). Role of the GYVG pore motif of HslU ATPase in protein unfolding and translocation for degradation by HslV peptidase. J. Biol. Chem. *280*, 22892-22898.

Sali,A. and Blundell,T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779-815.

Sali,A. and Overington,J.P. (1994). Derivation of rules for comparative protein modeling from a database of protein structure alignments. Protein Sci. *3*, 1582-1596.

Sauder,J.M., Arthur,J.W., and Dunbrack,R.L., Jr. (2000). Modeling of substrate specificity of the Alzheimer's disease amyloid precursor protein beta-secretase. J. Mol. Biol. *300*, 241-248.

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

Takeda-Shitaka,M., Takaya,D., Chiba,C., Tanaka,H., and Umeyama,H. (2004). Protein structure prediction in structure based drug design. Curr. Med. Chem. *11*, 551-558.

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

Figures

Figure 2.1 Crystal structure of the N-terminal domain of *E. coli* ClpB.
Ribbon representation shows the four identical monomers found in the asymmetric unit of the crystal (PDB structure 1khy). The four chains are represented in different colors.

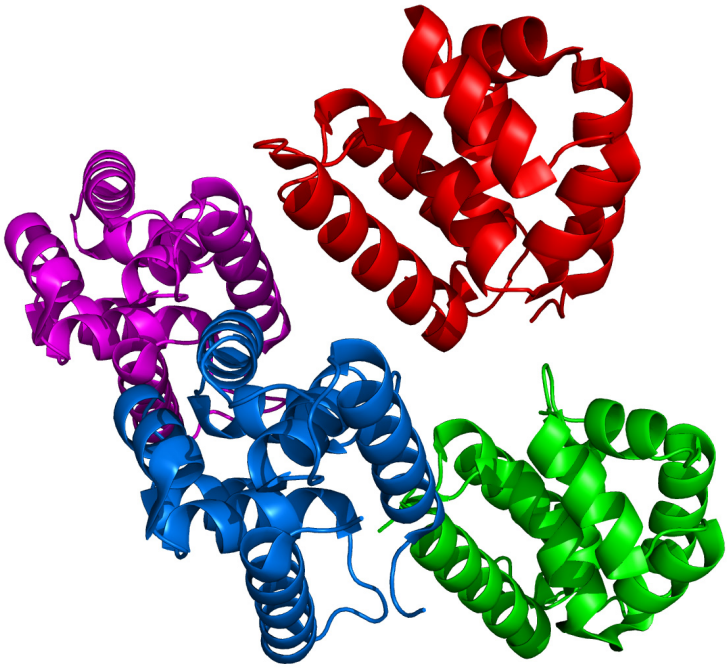


Figure 2.2 Crystal structure of D1 large sub-domain of *E. coli* ClpB.

Ribbon representation of the isolated D1 large sub-domain of *E. coli* ClpB from PDB structure 1jbk colored by secondary structure. The conserved D1 pore loop and the Walker A loop (P-loop) are colored in blue and cyan, respectively. The N-terminus and C-terminus of the molecule are labeled.

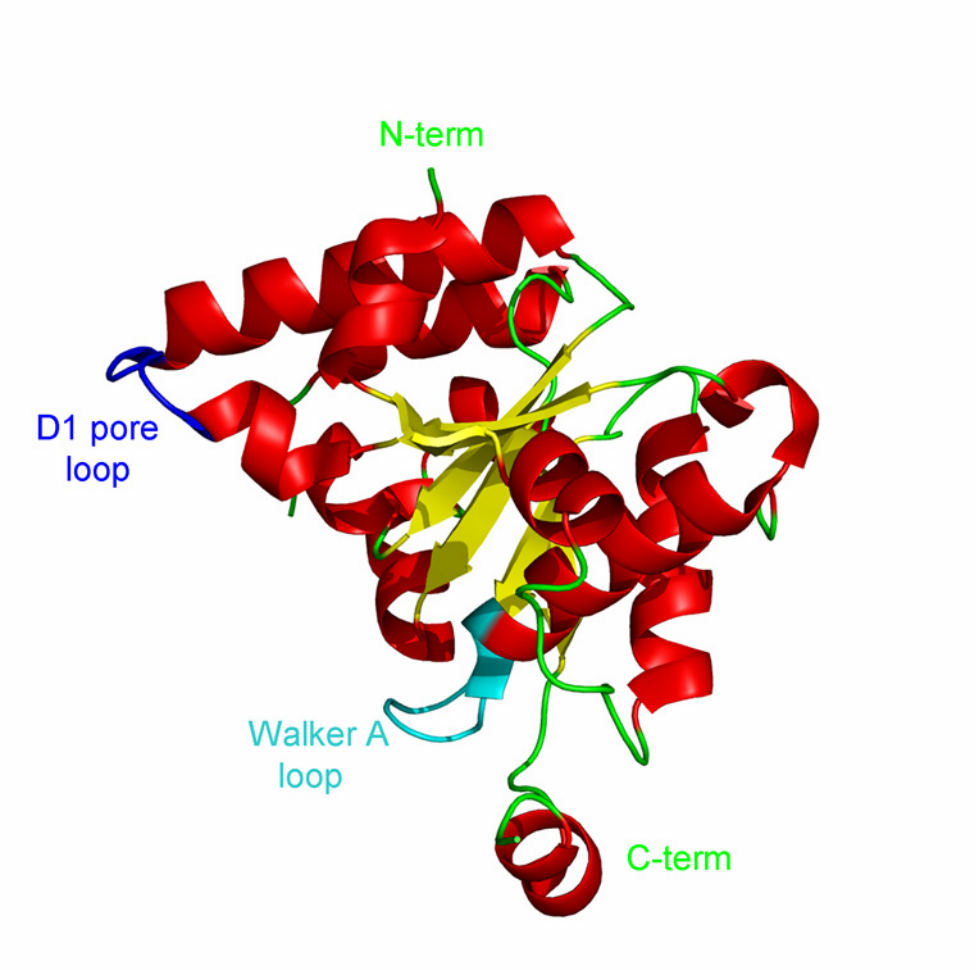


Figure 2.3 Superimposed chains of the non-physiological trimer of the *T. Thermophilus* ClpB.

Chains A, B, C from the 1qvr structure are colored red, blue and green, respectively. The N-terminal domain is located at the top and the C-terminus is at the bottom on panel A. Only the N-terminal domains of the three chains are shown in panel B rotated 90 degrees compared to panel A. The N-terminal residues are shown as spheres to illustrate the different orientation of the N-terminal domain of chain C relative to chains A and B.

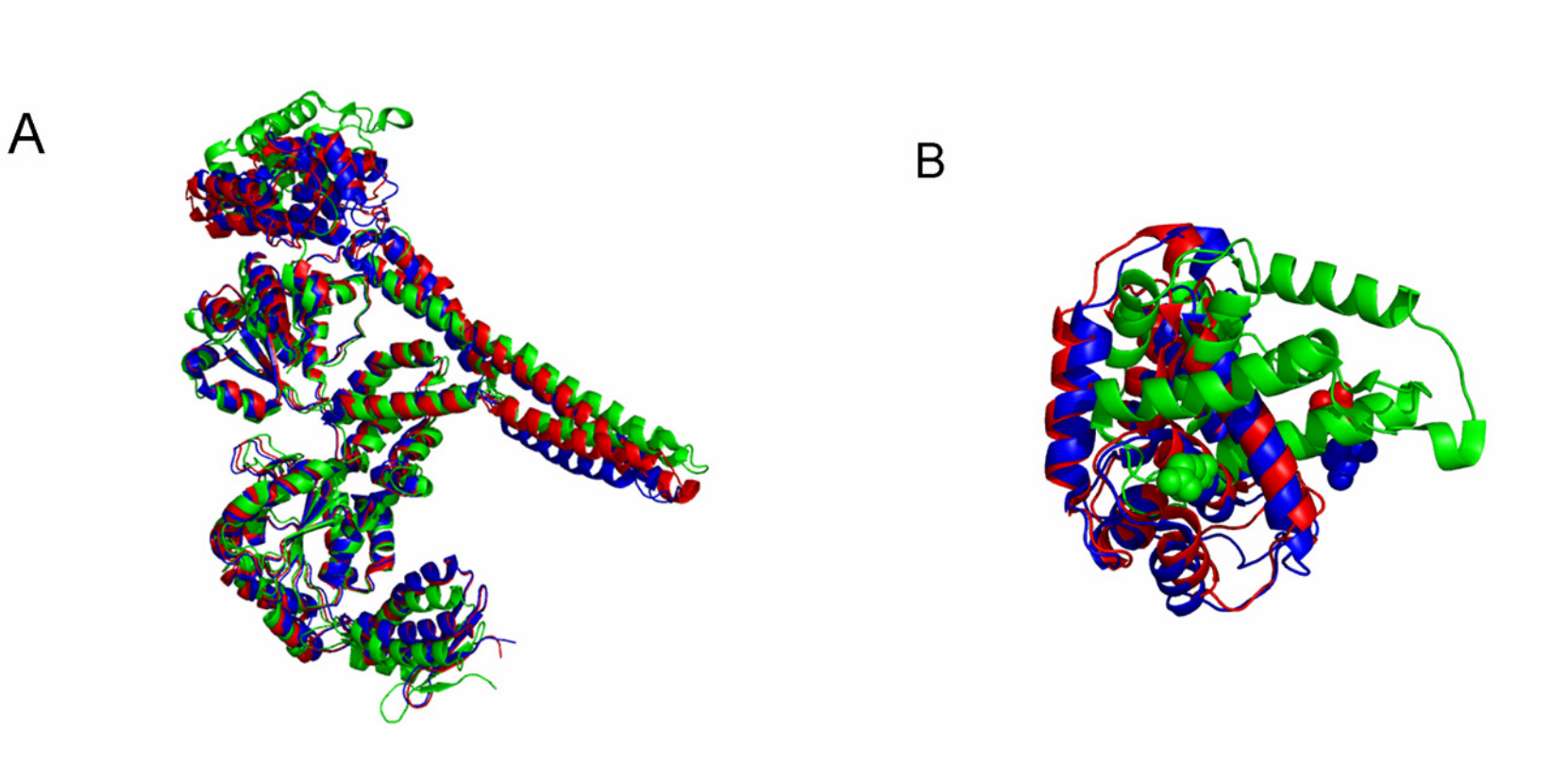


Figure 2.4 iMolTalk hexameric model of *T. thermophilus* ClpB.

The top view (left) and the side view (right) of the hexamer with each monomer in a different color. The left panel illustrates the entrance to the ClpB channel. The mobile D1 loop (residues 235-245) and D2 loop (637-650) are missing from each subunit. The arrow in the right panel indicates the chain-discontinuity site between the residues 543-544.

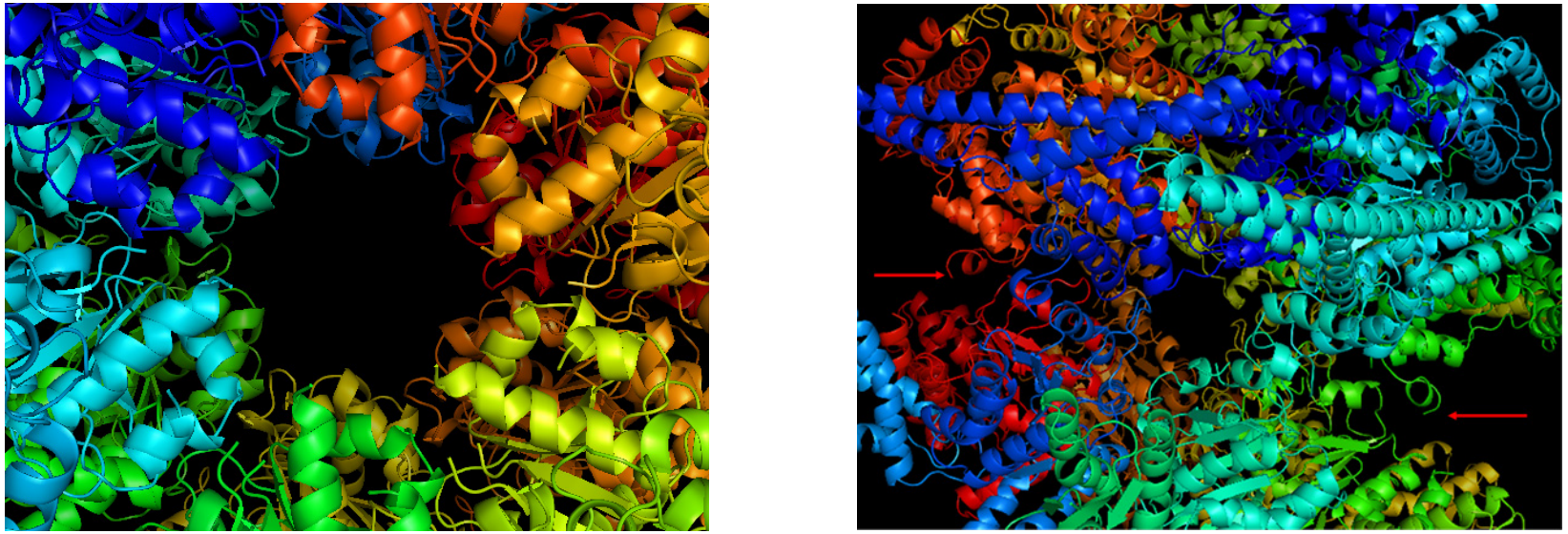


Figure 2.5 Sequence alignment of *E. coli* and *T. thermophilus* ClpB.

The sequence of *E. coli* ClpB (EC_ClpB) and *T. thermophilus* ClpB (TT_ClpB) was aligned using BOXSHADE. The background is black for identical residues and gray for similar ones.

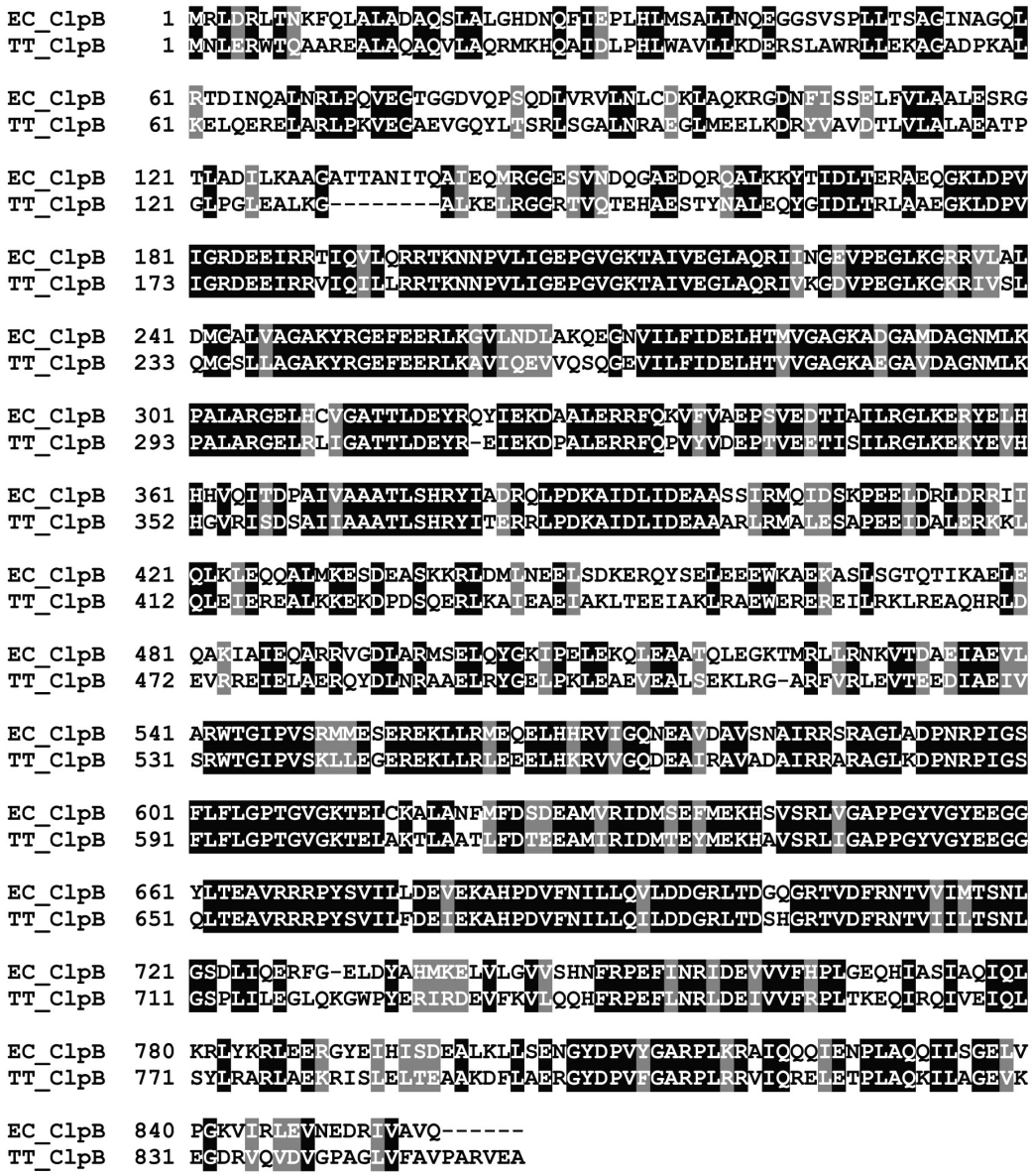
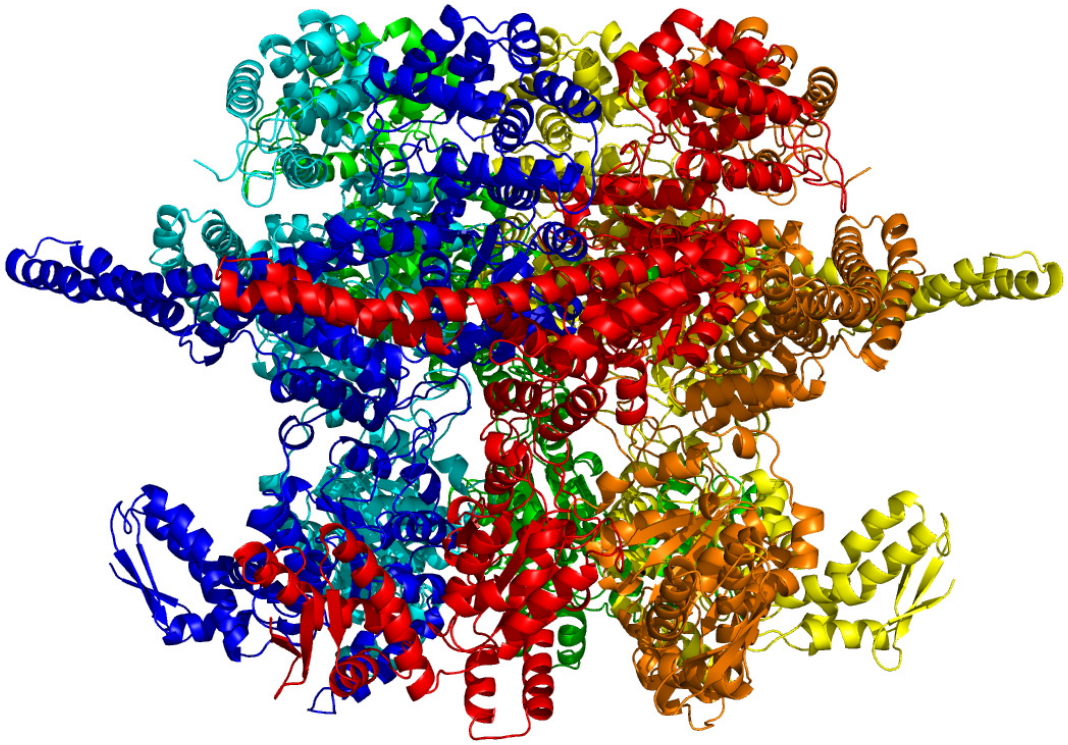


Figure 2.6 Hexameric model of the *E. coli* ClpB.

Panel A: Side-view of the hexameric model of the *E. coli* ClpB colored by subunits. The N-terminal domains are at the top. Panel B: Top-view of the hexameric model of the *E. coli* ClpB colored by subunits. The previously unresolved flexible loops are protruding into the central pore.

A



B

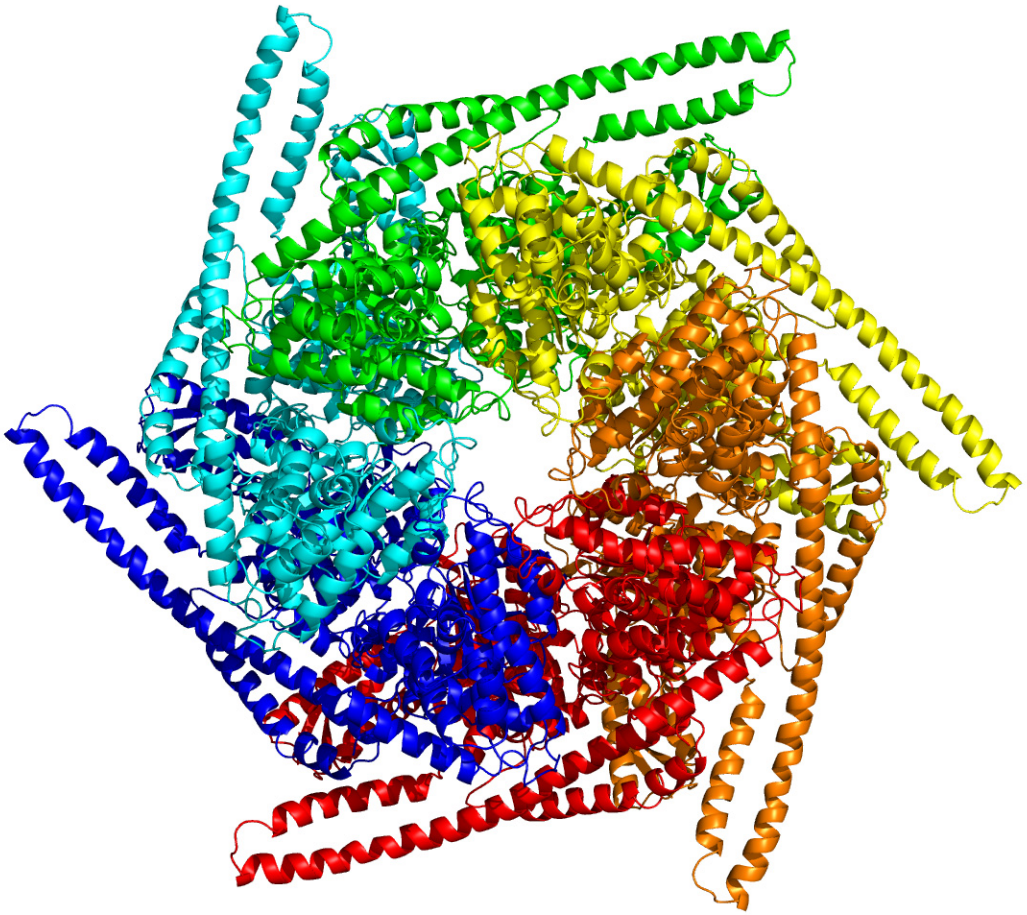


Figure 2.7 Superposition of the *E. coli* and *T. thermophilus* ClpB.
Cartoon representation of the *E. coli* ClpB homology model (shown in green) superimposed over chain A of the *T. thermophilus* hexameric ClpB model from iMolTalk (shown in red). Residues with backbone phi-psi angles in the disallowed regions (Ser45, Arg119, Lys250, Asp290, Ala589, Asp696, Glu731, Glu768) are shown as spheres.

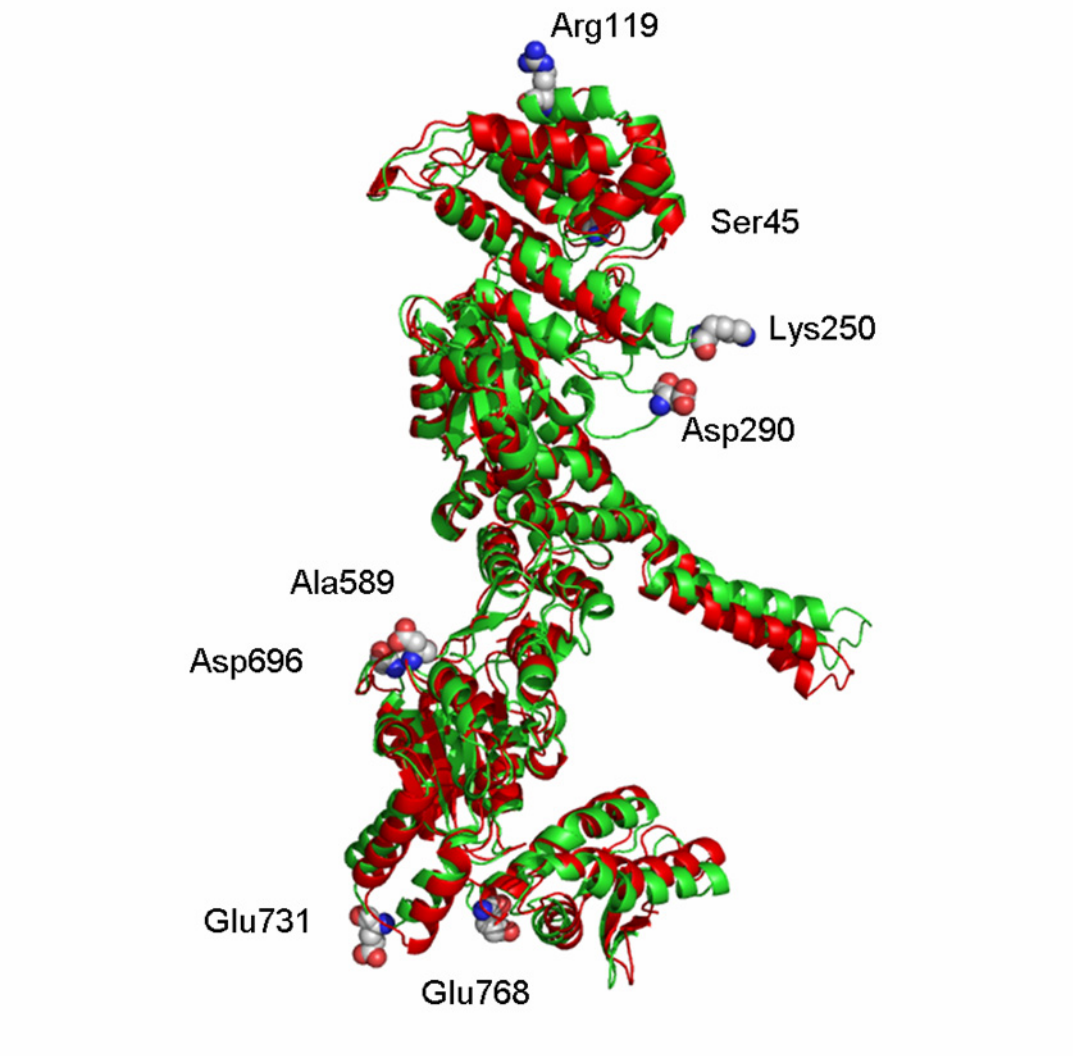
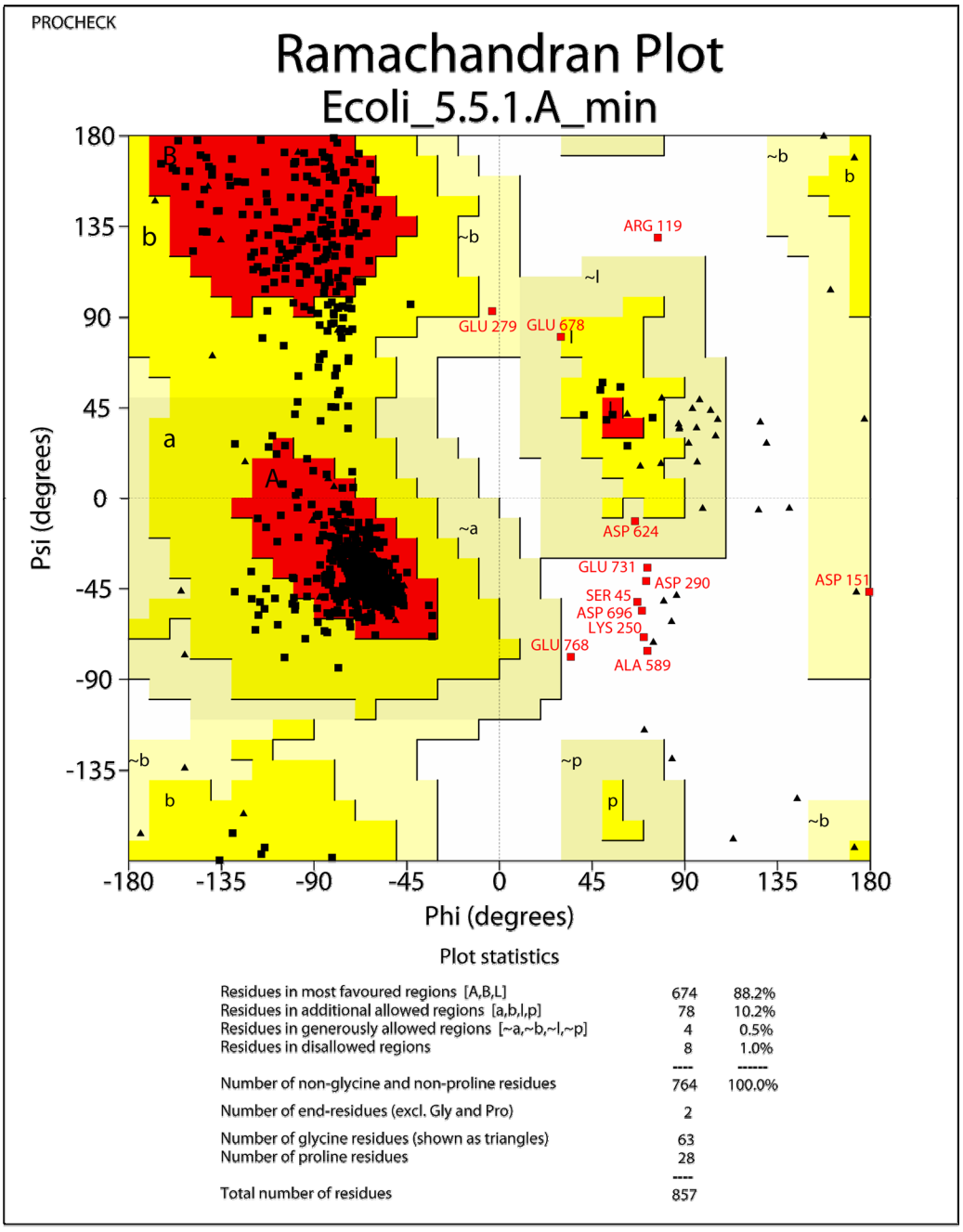


Figure 2.8 Ramachandran plot.

The Ramachandran plot of one of the monomers from the final *E. coli* ClpB hexameric model generated with PROCHECK. Three-letter codes and numbers of the residues in the disallowed region are shown in red.



Ecoli_5.5.1.A_min_01.ps

Figure 2.9 Flexible pore loops in hexameric ClpB.

Top view of the hexameric model of the *E. coli* ClpB. Conserved Tyr residues Tyr251 (red) and Tyr653 (blue) located on the D1 (Tyr251) and D2 (Tyr653) flexible loops inside of the ClpB channel.

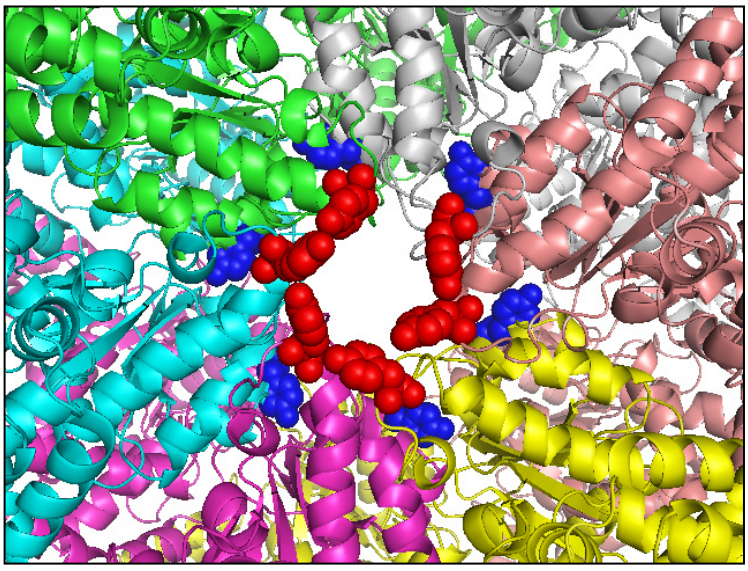
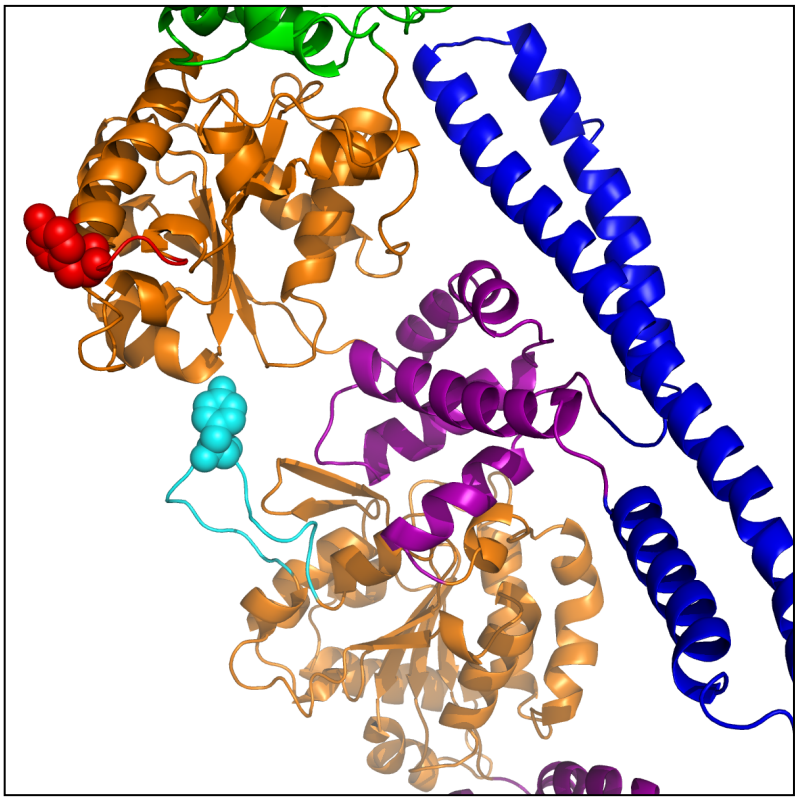


Figure 2.10 Position if the D1 and D2 conserved pore loops.

Close up side-view of full-length *E. coli* ClpB monomer model with the D1 and D2 flexible loops. The conserved Tyr251 of the D1 loop is shown as red sphere, while the conserved Tyr653 of the D2 loop is shown in cyan. The large and small sub-domains of D1 and D2 are colored in orange and magenta, respectively. The M domain is blue and the N domain is green (only parts of the domains are visible).



CHAPTER 3 - Domain stability and nucleotide-induced conformational changes in ClpB

Introduction

The ring-shaped oligomeric form and the ATP-dependent remodeling activity, attributed to the AAA+ superfamily members, are tightly coupled. This is because the nucleotide binding sites are located both between the large and the small sub-domains of the AAA+ modules of a given subunit and at the interfaces between the subunits of the oligomeric ring. The nucleotide-binding pocket is formed by the conserved Walker A and Walker B sequence motifs of the large AAA+ sub-domain. Conserved residues that belong to these motifs are required for ATP binding and hydrolysis. Specifically, the classical Walker A motif directly interacts with the phosphates of ATP and is also refereed to as the phosphate binding loop, or P-loop. Also, the conserved Glu residue of the Walker B motif is believed to act as a catalytic base for the hydrolysis reaction, while the Asp residue coordinates the Mg^{2+} required for the ATP hydrolysis. Besides the classical Walker motifs AAA+ ATPases contain conserved motifs that “sense” the presence or the absence of nucleotides, transmit this information to distal places, and coordinate ATP hydrolysis and conformational changes between the subunits of the oligomer. Sensor-1 polar residue (N/T) is located between the Walker A and Walker B motifs in the 3D structure and interacts both with the Walker B residues and the γ-phosphate of the bound ATP. Similarly the conserved sensor-2 Arg residues located in the small AAA+ sub-domain is positioned so it can directly contact the γ-phosphate of

the bound ATP and mediate conformational changes in the small subunit. Finally, the "arginine finger" residue from one subunit of the oligomer contributes to the ATP binding site of the neighboring subunit in the oligomer and could transmit the ATP hydrolysis-driven conformational changes to the neighboring subunit (Ogura and Wilkinson, 2001; Neuwald et al., 1999; Hanson and Whiteheart, 2005).

In case of ClpB, it has been long established that oligomeric ring formation is necessary for binding of nucleotides and ATP hydrolysis, which in turn, is necessary for the chaperone activity of ClpB (Barnett et al., 2000). However, besides the functionally relevant hexameric ring structure, under certain conditions ClpB is also able to form heptameric ring assemblies. Previous work from our lab indicates that in solution, ClpB shows a reversible equilibrium between monomers, heptamers, and intermediate-size oligomers. High protein concentration or low ionic-strength buffers stabilize heptameric ClpB, which switches to a hexamer upon binding of nucleotides (Akoev et al., 2004).

The functional diversity of AAA+ ATPases appears to rely not only on the common protein architecture, but also on similar principles of molecular mechanism. Thus, it can be proposed that the mechanism of other AAA+ ATPases might parallel that of ClpB and could involve insertion or enclosure of substrates in the central pore of the oligomeric ring followed by an ATP-dependent work resulting in a change of substrate conformation.

To understand that mechanism, it is essential to characterize the conformational effects that accompany formation of the AAA+ ring. In particular, it is important to understand the role of distinct protein domains within and outside of the AAA+ module in maintaining the ring structure. It is also important to identify the conformational

changes that occur in the various ClpB domains as a result of ATP binding and hydrolysis. We attempted to reach these goals, by investigating the thermodynamic stability of ClpB in the monomeric and oligomeric form and by following the local conformational rearrangements in ClpB upon ADP or ATP binding.

Materials and Methods

Mutagenesis and protein purification

Site-directed mutagenesis of ClpB was performed using the QuickChange method (Stratagene). Primers were ordered from IDT (Integrated DNA Technologies, INC). pET20b+ vector containing the wt ClpB sequence was used as the starting template for the PCR reactions. Wt ClpB contains two Trp residues, which were mutated to Phe to produce a Trp-less ClpB sequence that was used as template for the subsequent Trp mutations. Following digestion of the parental DNA the PCR products were transformed into XL1-Blue supercompetent cells (Stratagene). Plasmid DNA was purified from overnight cultures using QIAprep Spin Miniprep Kit (Qiagen) and sequenced at DNA Sequencing and Genotyping Facility, Department of Plant Pathology at KSU. Plasmid DNA containing the desired mutation was transformed into BL21 (DE3) cells (Novagen) using the provided protocol. Previously published procedures were used to produce and purify *Escherichia coli* ClpB and its mutants (Barnett et al., 2000).

Proteins were dialyzed against 50 mM Tris/HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, with 10% glycerol. Protein concentration of ClpB was determined spectrophotometrically. The extinction coefficients were recalculated for the Trp-mutants.

ClpB ATPase assay

ClpB samples were incubated for 15 min at 37 °C in 100 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 5 mM ATP, 1 mM EDTA, 1 mM DTT without or with 0.1 mg/ml κ-casein (Sigma). Inorganic phosphate production was determined using the malachite green method (Hess and Derr, 1975; Lanzetta et al., 1979). 2.5 µg ClpB was used for the basal ATPase measurements and 0.25 µg for the measurement of the κ-casein stimulated ATPase activity. Three separate measurements were performed for each ClpB variant.

Urea-induced unfolding of ClpB

Solutions containing urea at the concentration 0 - 8 M were prepared with buffer F (50 mM Tris/HCl pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT). To investigate monomeric ClpB, single-Trp ClpB variants were diluted to 0.1 mg/ml in the buffer F urea solutions containing 0.3 M KCl. To investigate heptameric ClpB, the ClpB variants were diluted to 0.5 mg/ml in the buffer F urea solutions. The urea concentration in case of each sample was corrected for the dilution by the protein solutions. ClpB samples were incubated overnight (~15 h) at room temperature before measuring fluorescence spectra.

Fluorescence spectroscopy

Fluorescence spectra of the single-Trp ClpB variants were measured with a Cary Eclipse spectrofluorimeter (Varian) using a 1-cm quartz fluorescence cuvette. The excitation wavelength was set at 300 nm and the emission spectrum was recorded from 280 nm to 450 nm. In case of the unfolding experiments with 0.1 mg/ml ClpB, the excitation slit was 5 nm and the emission slit was 10 nm. In experiments with 0.5 mg/ml

ClpB, both slits were 5 nm. The spectra of urea-containing buffers (without ClpB) were subtracted from the protein spectra.

For the detection of the nucleotide-induced conformational changes, proteins were diluted to 0.6 mg/ml in buffer F. The fluorescence emission spectra were recorded in case of each ClpB variant without nucleotides, than ADP or ATP was added from 0.2 M stock solutions. The final concentration of the nucleotides was 2 mM in order to saturate the nucleotide binding sites. Both excitation and emission slits were set to 5 nm.

Results and Discussion

Design and characterization of the Trp variants of ClpB

Stability of different structural domains, as well as nucleotide-induced conformational changes in ClpB, was monitored by following the changes in the intrinsic fluorescence of Trp residues. This required the production of several ClpB mutant proteins, each with a single Trp residue located in different domains. Aromatic residues located in proximity of the structurally or functionally important residues were chosen for mutation.

E. coli ClpB contains two Trp residues: Trp462 at the center of the coiled-coil domain (see Fig. 3.1) and Trp543 in the D1 small domain. First, we produced two single-Trp variants of ClpB, W543F (with the remaining Trp462) and W462F (with the remaining Trp543). Next, we produced a Trp-less ClpB with the double mutation W462F/W543F. Subsequently, five new variants of Trp-less ClpB were produced with a single Trp introduced in specific locations containing aromatic residues (see Fig. 3.1): F105W (within the N-terminal domain, close to the putative substrate binding site

(Barnett et al., 2005)), Y251W (at the D1 entrance to the ClpB channel), F276W (within Walker B motif in D1 large domain), F603W (within Walker A motif in D2 large domain), and Y812W (within sensor-2 motif in the D2 small domain). Altogether, we produced seven variants of ClpB, each with a single Trp located in each of the six protein domains (at positions 105, 251, 276, 462, 543, 603, and 812, see Fig. 3.1).

After protein production and purification the structural integrity and biochemical functionality of the seven single-Trp variants of ClpB was tested. Similar to wt ClpB, the seven single-Trp ClpB variants efficiently formed oligomers at high protein concentration, as determined by sedimentation velocity studies (data not shown), which indicates that the site mutations did not disrupt the structure of the ClpB variants. All seven ClpB variants were active ATPases (Fig. 3.2 A). The basal ATPase activity varied among the single-Trp variants with three proteins showing lower activity than the wt and three proteins showing higher activity. Interestingly, the introduction of a Trp in the vicinity of the Walker A motif in D2 (Trp603) significantly activated the ClpB ATPase. Casein, an unstructured protein, is often used as an activator of the ATPase activity of the wt ClpB. Fig. 3.2B shows that the ATPase of all seven variants of ClpB was activated by casein (compare the scale with Fig. 3.2A). The ATPase activity of the ClpB variants measured in the presence of casein paralleled their basal activity. We conclude that a replacement of the selected aromatic residues (either Phe or Tyr) by Trp may modulate the rate of ATP hydrolysis of ClpB. Interestingly, even mutations in the domains that do not contain nucleotide binding sites (W105 in the N-terminal domain, W462 in the coiled-coil domain) affect the ClpB ATPase, which suggests that concerted

conformational changes in all domains may contribute to the ATP-dependent functionality of ClpB.

Domain stability in ClpB

Thermal unfolding transitions of ClpB by differential scanning calorimetry revealed that ClpB contains multiple folding domains that likely correspond to the structural domains identified in Fig. 3.1B (Nagy et al., 2006). We used Trp fluorescence to monitor changes in conformation of each domain in ClpB at equilibrium with increasing concentrations of urea. It has been previously shown that ~0.1 mg/ml ClpB is monomeric in high ionic strength buffers (0.2-0.3 M KCl), whereas heptameric ClpB predominates at low ionic strength (Akoev et al., 2004; Zolkiewski et al., 1999). We compared the Trp fluorescence spectra of the ClpB variants under the conditions where either monomeric (see Fig. 3.3A) or oligomeric ClpB (see Fig. 3.3B) predominates. As shown in Fig. 3.3A, the Trp emission intensity and the wavelength position of the emission band strongly depend on the position of the fluorophore in monomeric ClpB. In particular, the fluorescence of Trp276 (near the D1 ATP binding site) shows the strongest intensity quenching, whereas that of Trp543 (in D1 small domain) shows the lowest quenching. The differences between emission spectra of different ClpB variants indicate that the solvent exposure and the local environment of each of the fluorophores are different for each ClpB variant. Importantly, the intensities and positions of Trp emission spectra for the oligomeric ClpB (Fig. 3.3B) correlate with those observed for monomeric ClpB (Fig. 3.3A). This result indicates that major conformational changes are not likely to occur in the ClpB domains during protein oligomerization.

To further characterize the effects of ClpB oligomerization on the properties of its domains, we investigated the stability of the ClpB variants and measured Trp fluorescence as the function of urea concentration. To prevent protein aggregation at high temperature, we performed chemical denaturation studies for ClpB at room temperature (25 °C). To ensure the thermodynamic equilibrium during protein denaturation studies, we incubated ClpB in urea-containing buffer overnight and determined that the fluorescence spectra of ClpB samples did not change upon longer incubation. We also investigated the urea-induced unfolding of ClpB in the presence of ATP and ADP. Interestingly, after prolonged incubation with nucleotides, ClpB solutions with low urea concentrations became cloudy, which indicated protein precipitation and made fluorescence measurements unfeasible. This result suggests that nucleotides induce a high-affinity protein-binding conformation in ClpB and promote its non-specific self-association. Indeed, ClpB binds substrates with the highest affinity in its ATP-bound state (Weibezahn et al., 2003).

Fig. 3.4 A-G shows urea-induced changes in the Trp fluorescence intensity of the seven ClpB variants under conditions stabilizing the ClpB monomer and the heptamer. For all chosen locations of the fluorescence probes in ClpB, except Trp251 and Trp276 (Fig. 3.4 B and C), the Trp emission intensities decrease as the concentration of urea increases and show cooperative urea-dependent transitions. The emission intensities of Trp251 and Trp276 are the lowest among all investigated ClpB variants (see Fig. 3.3), which indicates that both these locations of the fluorophores are significantly exposed to the solvent in the native ClpB. Trp251 is located within a flexible unstructured loop at the entrance to the ClpB channel. The Trp251 loop has not been resolved in the crystal

structure of the *T. thermophilus* ClpB (Lee et al., 2003) and is involved in binding aggregated substrates to ClpB (Barnett et al., 2005; Haslberger et al., 2007). Urea-induced changes in the fluorescence of Trp251 are small and do not show clear cooperative transitions, which is consistent with the lack of stable structure within the channel loop either for the monomeric or the oligomeric ClpB. Trp276 is located near the conserved acidic residues of the Walker B motif in the D1 large domain of ClpB, which are essential for the ATP hydrolysis (Weibezahn et al., 2003). The Walker B motif is located at the tip of a β-strand (Lee et al., 2003) and its side chains are exposed into the ATP binding site. Unlike for the remaining fluorophore positions, the fluorescence intensity of Trp276 increases with increasing concentration of urea (Fig. 3.4C), which suggests that urea decreases the emission quenching of the solvent-exposed Trp276. Characteristically, the changes in Trp276 fluorescence show similar patterns for the monomeric and the oligomeric ClpB. We conclude that Trp251 and Trp276 are highly exposed to the solvent in the ClpB structure and their emission intensities may not follow the urea-induced structural rearrangements in their corresponding ClpB domains.

The remaining five single-Trp ClpB variants (W105, W462, W543, W603, and W812) show cooperative urea-induced transitions, which are consistent with two-state unfolding of the corresponding structural domains of ClpB (see Fig. 3.1). The pre-transition increases in fluorescence intensity are found for W543 under monomeric conditions (Fig. 3.4E) and for W603 (Fig. 3.4F). We assume that pre-transition intensity increases are due to urea-induced changes in fluorescence quenching for the fluorophores that are partially solvent-exposed in the native structure of ClpB.

The urea-induced cooperative conformational transitions are often described with the following empirical relationship: $\Delta G([D]) = \Delta G_0 - m[D]$ (Pace, 1986), where ΔG_0 is the standard free energy of unfolding of a given domain under urea-free conditions, [D] is molar concentration of urea, and m is an empirical constant. The above equation predicts that the standard free energy of protein unfolding (ΔG) decreases with an increasing urea concentration, which increases the equilibrium population of the unfolded conformation and decreases that of the native state.

We used the above empirical relationship to describe the cooperative urea-induced transitions in the Trp fluorescence intensity (Fig. 3.4A and D-G). The parameters of the equation, ΔG_0 and m, are listed in Table 3.1 for five single-Trp ClpB variants. The D2 large domain that contains Trp603 is the most stable among the domains of monomeric ClpB (as shown by the values of ΔG_0). The coiled-coil domain that contains Trp462 is the least stable. Upon oligomerization of ClpB, no significant changes of domain stability are found for the N-terminal domain (W105), the small D1 domain (W543), and the large D2 domain (W603). However, the ClpB oligomerization significantly destabilizes the small D2 domain (W812) and, to a lower extent, the coiled-coil domain (W462). This result indicates that the two domains may undergo significant structural rearrangements during formation of the ClpB ring. In the oligomeric ClpB, the coiled-coil and D2 small domain may lose some inter-domain contacts that contribute to their stability in the monomeric form. Conversely, in the oligomeric ring, these domains may form new contacts with partner domains of the neighboring subunit. Indeed, our previous studies showed that the contacts mediated by the D2 small domain are necessary for the stability of the ClpB ring. Removal of the D2 small domain completely

inhibits the formation of ClpB oligomers in the absence or in the presence of nucleotides (Barnett et al., 2000). Deletions within the coiled-coil domain also destabilize the oligomers of ClpB (Kedzierska et al., 2003). In the high-resolution structure of the *T. thermophilus* ClpB oligomer (Lee et al., 2003), both the coiled-coil domain and the D2 small domain form the outside "layer" of the ring. The two domains form extensive contacts with the large D1 and D2 domains, respectively, of the neighboring subunit within the ring. The mechanism of stabilization of the ClpB ring by contacts maintained by the coiled-coil and the D2 small domain is consistent with the oligomerization-induced changes in those domains' stabilities that were detected in the present work.

The remaining ClpB domains: the mobile N-terminal domain, the D2 large domain, and the D1 small domain, which serves as a "hinge" for the coiled-coil (see Fig. 3.1) do not change their stability in the oligomeric ClpB (see Fig. 3.4). This result indicates that the core of the AAA+ module is a rigid structural unit that can be assembled into an oligomeric ring and "fastened" to its neighbors by contacts maintained by the coiled-coil and the D2 small domain. This work is the first protein stability study of a AAA+ ATPase.

Nucleotide-induced conformation changes in ClpB

The single-Trp variants produced for the stability studies were also used in another set of experiments, one designed to follow the nucleotide-induced conformational changes in ClpB. The change in the fluorescence emission spectra was followed after addition of saturating amounts of ADP or ATP. These experiments were performed under conditions which favored the formation of ClpB oligomers (low ionic strength buffer and 0.6 mg/ml protein concentration), thus the changes in the

fluorescence intensity were the consequences of local conformational changes which occurred after nucleotide binding to an already oligomeric ClpB variant (during the heptamer-hexamer switch). Several of the Trp variants of ClpB responded to nucleotide binding. The biggest change in the fluorescence emission intensity was observed in case of the channel probe (W251). In this ClpB variant, the fluorescence emission intensity decreased significantly after addition of saturating amounts of ADP (19.5 %) or ATP (17.5 %) (Fig. 3.5A). This decrease of intensity suggests that the Trp residue located on the mobile loop that lines the channel entrance of the ClpB oligomer, becomes more accessible to solvent quenching upon nucleotide binding. The middle domain probe (W462) showed a significant increase in the emission intensity in the presence of ADP (18.7 %) as well as ATP (13 %) (Fig. 3.5B). This decrease in quenching could signal a conformational change in the middle domain of ClpB upon nucleotide binding as a result of which Trp462 is buried and the emission quenching is decreased. Interestingly this naturally occurring Trp residue is located at the base of the propeller-like middle domain, and in the homology model of the *E. coli* ClpB, it is oriented towards the small D1 sub-domain. Indeed, recent data indicated that the middle domain contacts the first AAA+ domain of ClpB in a nucleotide dependent manner and could facilitate the cooperation between the different subunits of the ring (Haslberger et al., 2007; Lee et al., 2007).

The most interesting result was obtained for the Trp603 probe, the one located in the large domain of the second AAA+ motif near the conserved Walker B motif. For this variant, ADP binding caused a decrease (14 %), while ATP binding caused an increase in the fluorescence intensity of this Trp residue (10 %) (Fig. 3.5C). This result suggests

a significant change in the environment of Trp603 upon the hydrolysis of ATP molecules, possibly a conformational change which makes this residue more accessible to solvent quenching. It is interesting to note, that the domain stability studies indicated that the same domain does not undergo a significant conformational change during oligomerization (stability did not changed from monomer to heptamer). So, while the large domain of the second AAA+ module constitutes the relatively rigid core of the ClpB disaggregating machine, it is still able to undergo major conformational changes upon ATP hydrolysis. It is believed that these and similar conformational changes which occur in and around the central channel of the ClpB oligomer are responsible for the substrate translocation activity. Finally, in the case of the fluorescence probe located in the D2 small C-terminal domain, ADP binding caused a small (9%) increase in the emission intensity, while ATP binding had no effect on the intensity of the fluorescence (Fig. 3.5D). However, both ADP and ATP binding caused a shift toward longer wavelengths in the maximum of the emission spectrum of the Trp812 (from 348 nm to 352.5 nm and from 348 nm to 350.5 nm, respectively). This “red shift” towards longer wavelengths of the maximum of intensity suggests that the Trp812 becomes more exposed as a result of nucleotide binding (especially ADP binding) or may form different contacts with the rest of the protein. The domain stability studies suggested that the D2 small domain probably also undergoes conformational change upon oligomerization.

In case of the remaining fluorescence probes (W105, W276, W543), addition of saturating amounts of ADP or ATP did not result in significant change of the emission intensity. This is an indication that those locations do not undergo significant conformational changes upon ATP binding or hydrolysis. Interestingly, among these

probes Trp543 together with Trp105 showed the lowest quenching (highest intensity) among the probes and Trp276 probe showed the highest quenching (lowest intensity) (Fig. 3.3B).

References

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

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

Barnett,M.E., Zolkiewska,A., and Zolkiewski,M. (2000). Structure and activity of ClpB from *Escherichia coli*. Role of the amino- and -carboxyl-terminal domains. J. Biol. Chem. *275*, 37565-37571.

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

Haslberger,T., Weibezahn,J., Zahn,R., Lee,S., Tsai,F.T., Bukau,B., and Mogk,A. (2007). M domains couple the ClpB threading motor with the DnaK chaperone activity. Mol. Cell *25*, 247-260.

Hess,H.H. and Derr,J.E. (1975). Assay of inorganic and organic phosphorus in the 0.1-5 nanomole range. Anal. Biochem. *63*, 607-613.

Kedzierska,S., Akoev,V., Barnett,M.E., and Zolkiewski,M. (2003). Structure and function of the middle domain of ClpB from *Escherichia coli*. Biochemistry *42*, 14242-14248.

Lanzetta,P.A., Alvarez,L.J., Reinach,P.S., and Candia,O.A. (1979). An improved assay for nanomole amounts of inorganic phosphate. Anal. Biochem. *100*, 95-97.

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

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

Nagy,M., Akoev,V., and Zolkiewski,M. (2006). Domain stability in the AAA+ ATPase ClpB from *Escherichia coli*. Arch. Biochem. Biophys. *453*, 63-69.

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

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

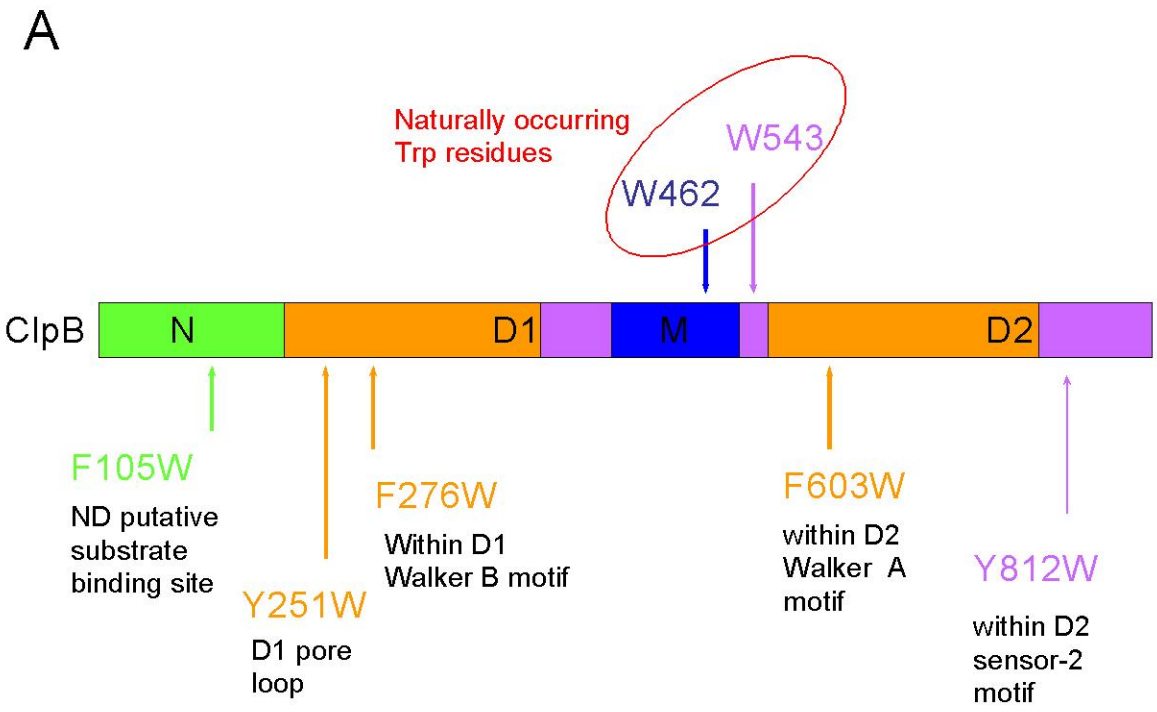
Pace,C.N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. Methods Enzymol. *131*, 266-280.

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

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

Figures

Figure 3.1 Location of Trp fluorescence probes in *E. coli* ClpB.
(A) Positions of single-Trp probes within ClpB. (B) Monomeric structure of the *E. coli* ClpB (as determined by homology modeling and presented in Chapter 2). The N-terminal domain (green) is on top of the figure, followed by the D1 AAA+ large (orange) and small domain (purple), the middle coiled-coil domain (blue), and the D2 AAA+ large (orange) and small (purple) domains. Positions of the aromatic residues mutated to Trp are shown in red spheres. The location of the entrance to-and exit from-the ClpB central channel is also indicated.



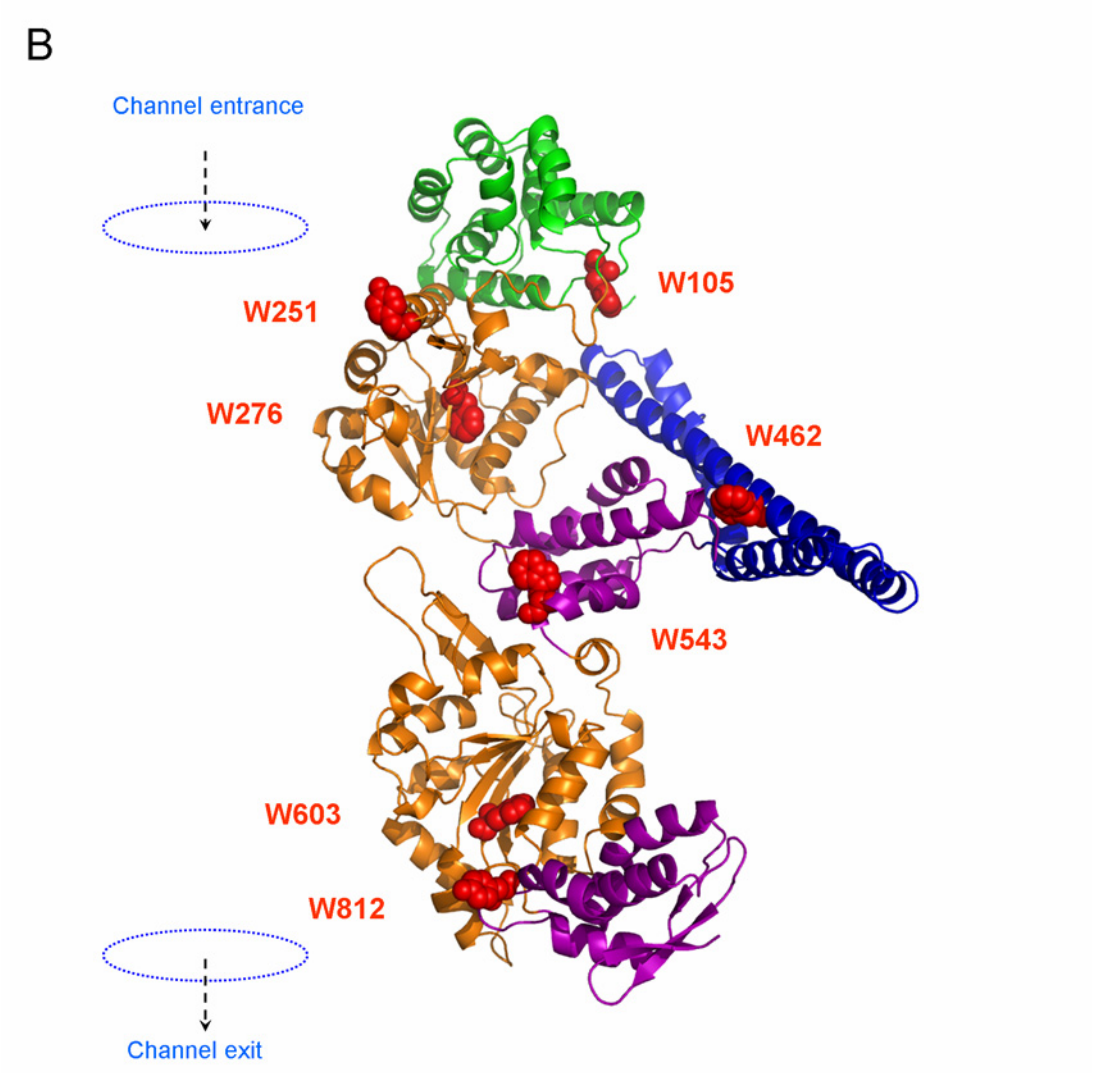

B
Channel entrance
W105
W251
W276
W462
W543
W603
W812
Channel exit

Figure 3.2 ATPase activity of wild type ClpB and its single-Trp variants.

The rate of ATP hydrolysis catalyzed by wt ClpB or its variants, each with a single Trp at the indicated position, was measured at 37 °C without (A) or with 0.1 mg/ml κ-casein (B). Averages from three measurements are shown with the standard deviations.

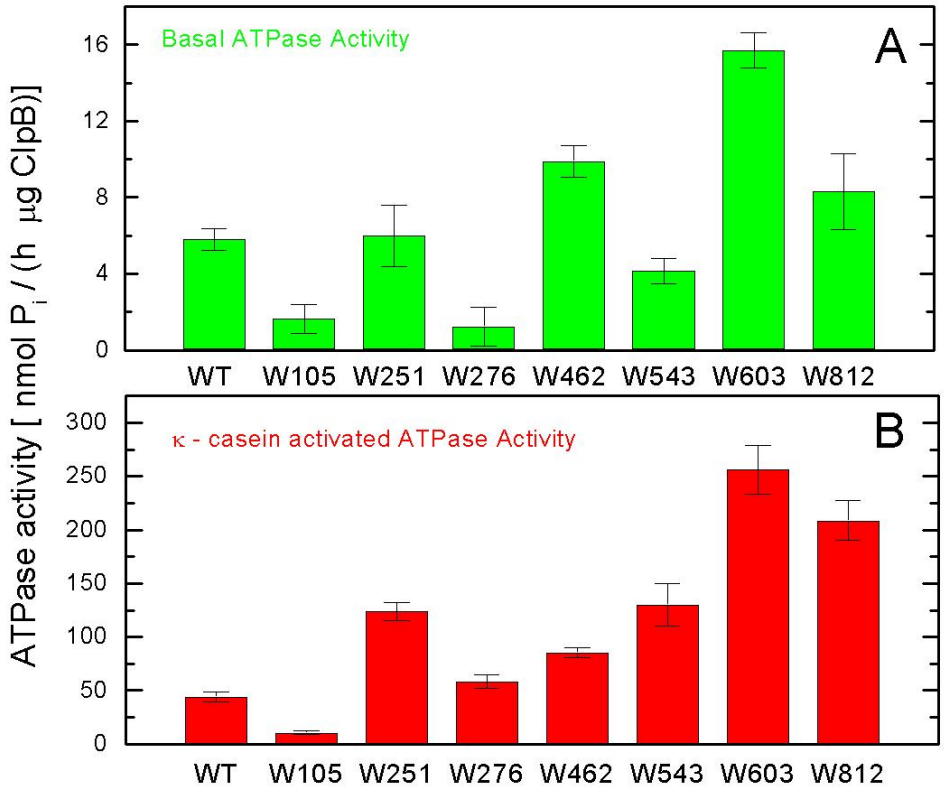


Figure 3.3 Fluorescence emission spectra of the single-Trp variants of ClpB. The excitation wavelength was 300 nm and the emission spectra were measured for each of the variants of ClpB indicated in the figure at 0.1 mg/ml in the buffer F (50 mM Tris/HCl pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT) with 0.3 M KCl (A) or at 0.5 mg/ml in buffer F without KCl (B).

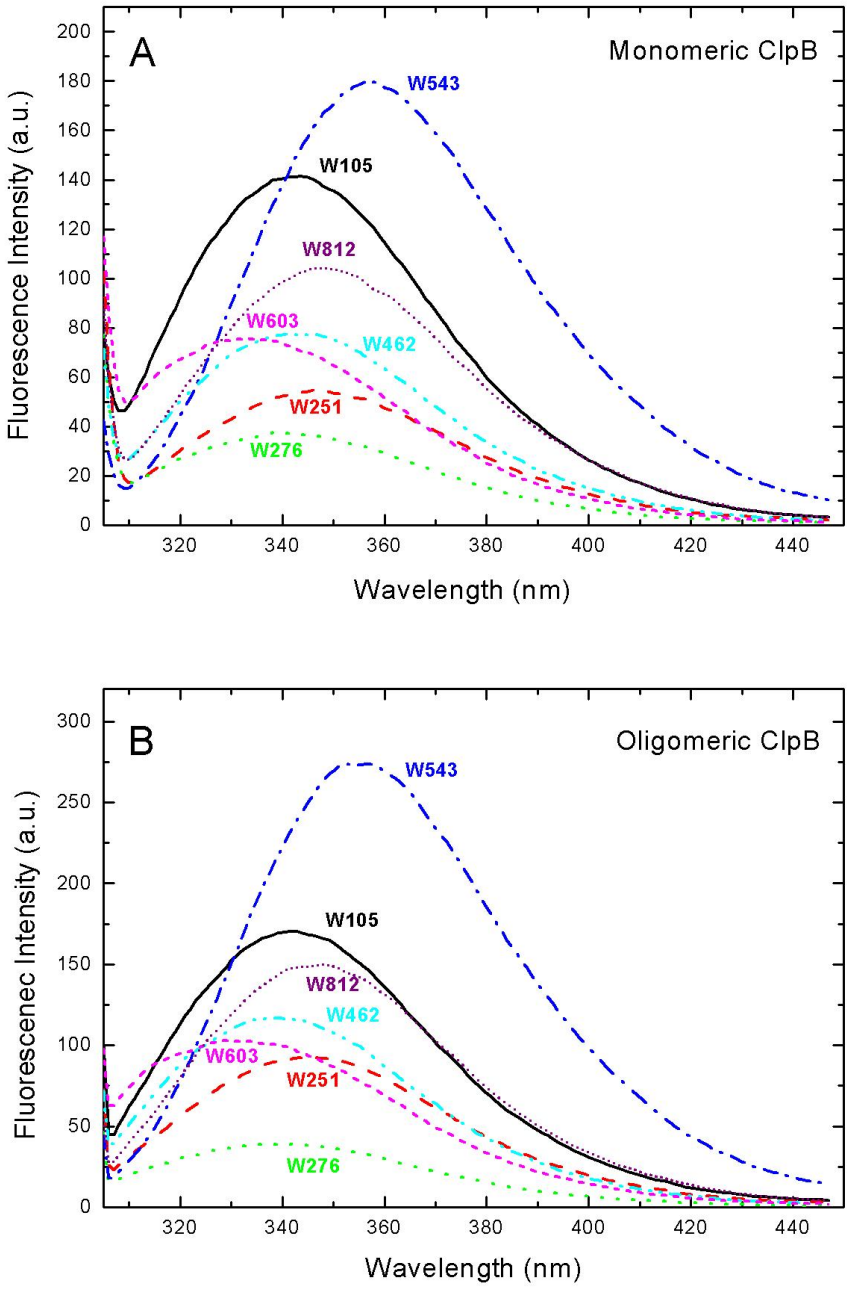
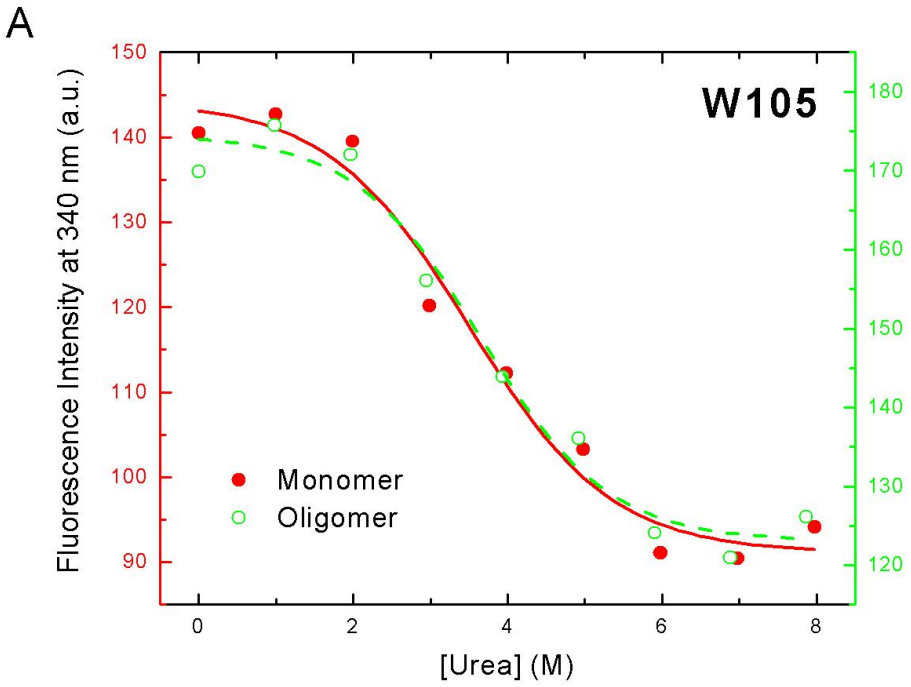
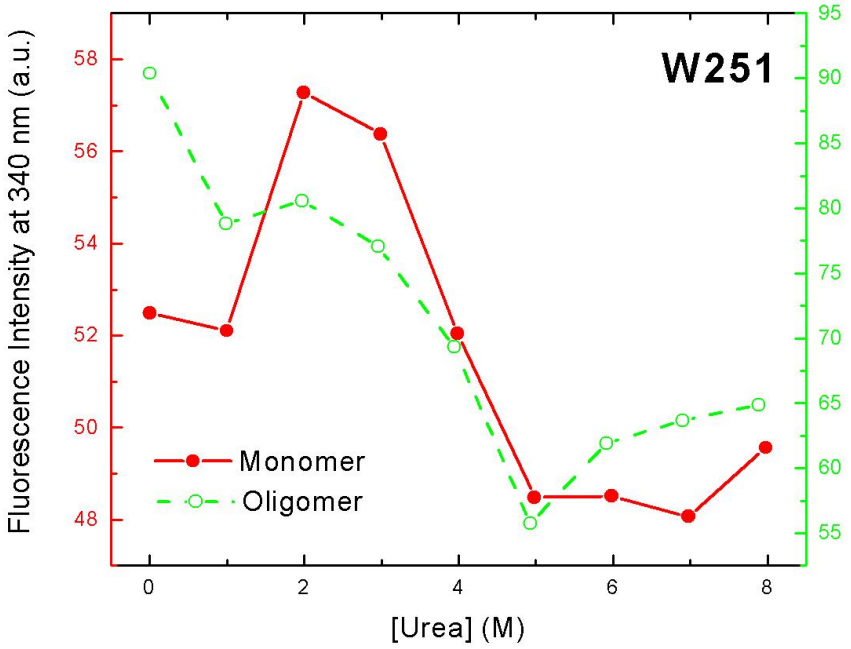


Figure 3.4 Urea-induced changes of fluorescence intensity in single-Trp ClpB variants.

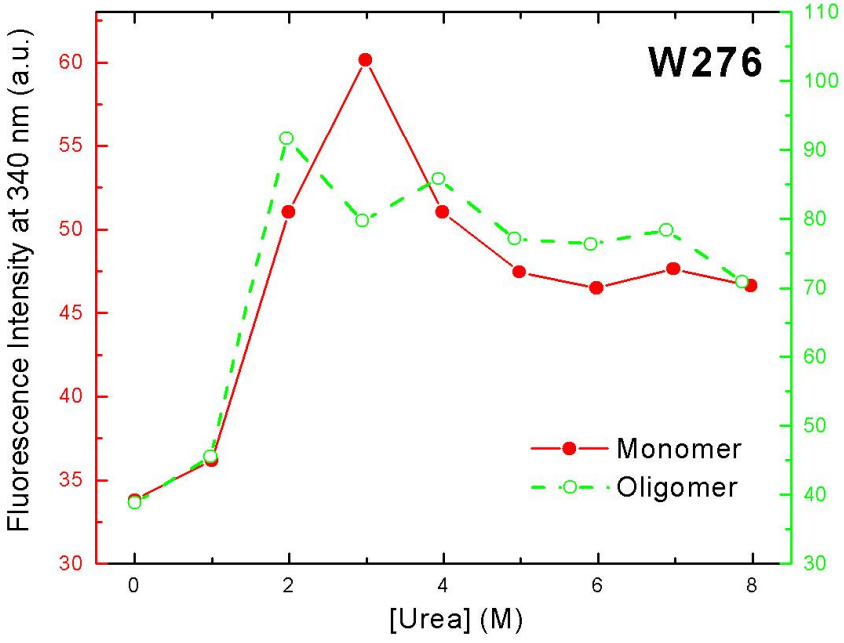
Shown are the data for monomeric ClpB (0.1 mg/ml in 50 mM Tris/HCl pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT with 0.3 M KCl and indicated urea concentrations, red filled symbols) and oligomeric ClpB (0.5 mg/ml in 50 mM Tris/HCl pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT with indicated urea concentrations, green open symbols). The emission intensity was measured at equilibrium for the ClpB variants containing Trp at the following positions: W105 (A), W251 (B), W276 (C), W462 (D), W543 (E), W603 (F), and W812 (G). In panels A and D through G, shown are the nonlinear least-squares fits of the data using the two-state unfolding model (see text) for monomeric ClpB (red solid line) and oligomeric ClpB (green dashed line).



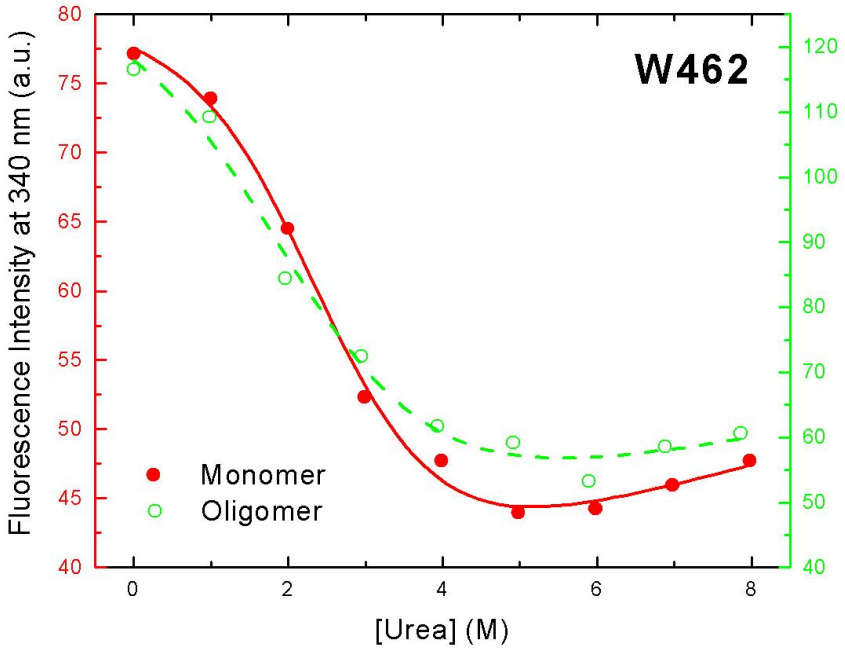
B


W251
Fluorescence Intensity at 340 nm (a.u.)
[Urea] (M)
Monomer
Oligomer
58
56
54
52
50
48
95
90
85
80
75
70
65
60
55
0
2
4
6
8

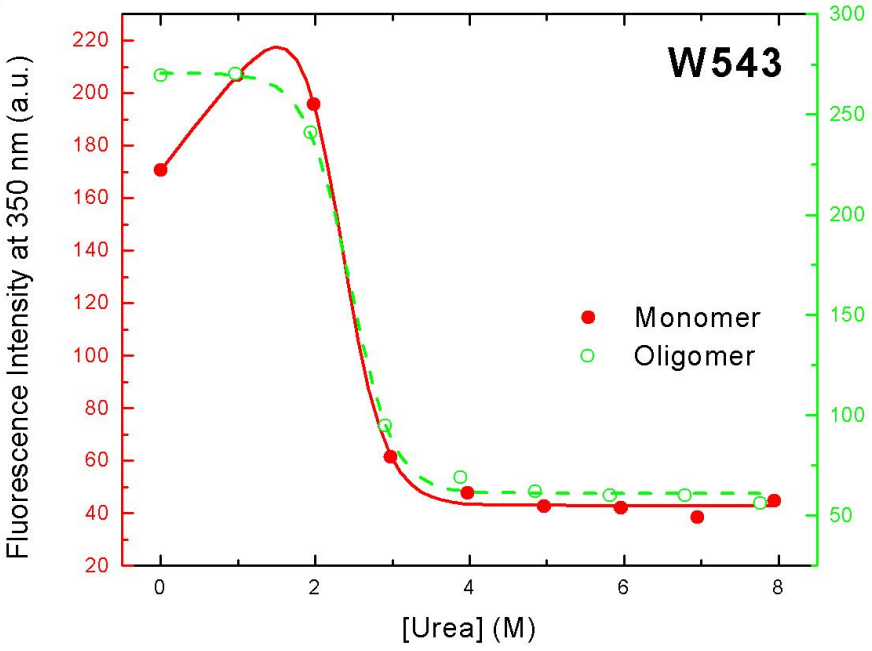
C


W276
Fluorescence Intensity at 340 nm (a.u.)
[Urea] (M)
Monomer
Oligomer
60
55
50
45
40
35
30
110
100
90
80
70
60
50
40
30
0
2
4
6
8

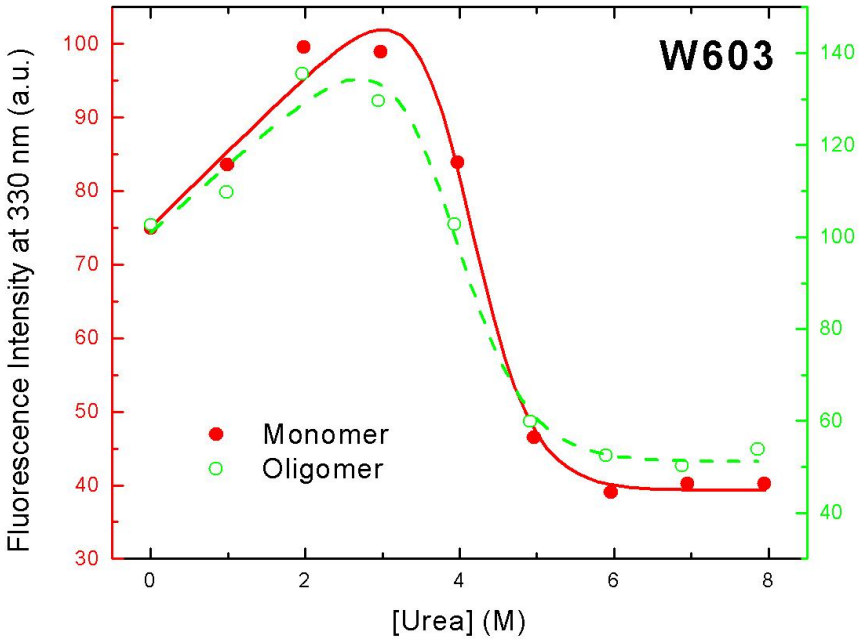
D


W462
Fluorescence Intensity at 340 nm (a.u.)
[Urea] (M)
Monomer
Oligomer
80
75
70
65
60
55
50
45
40
120
110
100
90
80
70
60
50
40
0
2
4
6
8

E


W543
Fluorescence Intensity at 350 nm (a.u.)
[Urea] (M)
Monomer
Oligomer
220
200
180
160
140
120
100
80
60
40
20
300
250
200
150
100
50
0
2
4
6
8

F


W603
Fluorescence Intensity at 330 nm (a.u.)
[Urea] (M)
Monomer
Oligomer
100
90
80
70
60
50
40
30
140
120
100
80
60
40
0
2
4
6
8

G

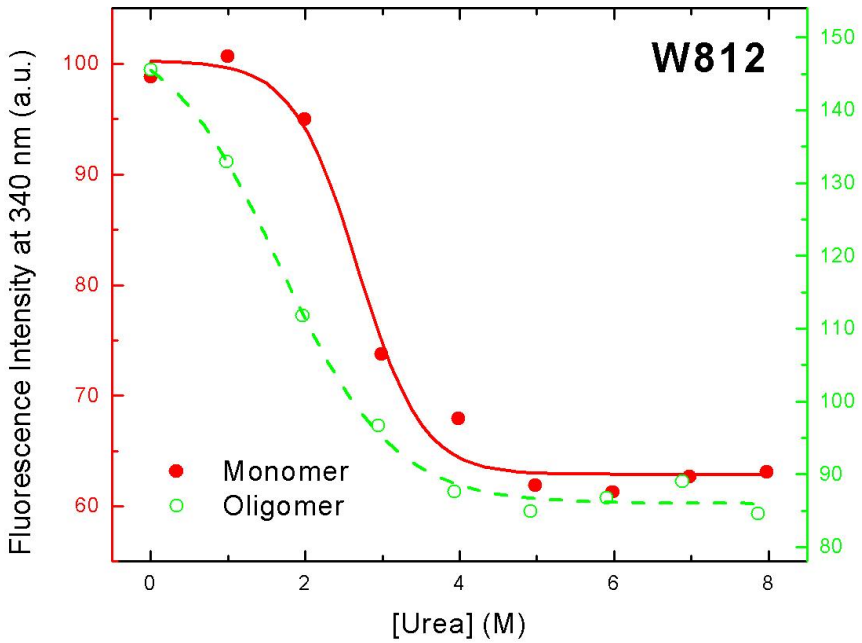

W812
Fluorescence Intensity at 340 nm (a.u.)
[Urea] (M)
Monomer
Oligomer
100
90
80
70
60
150
140
130
120
110
100
90
80
0
2
4
6
8

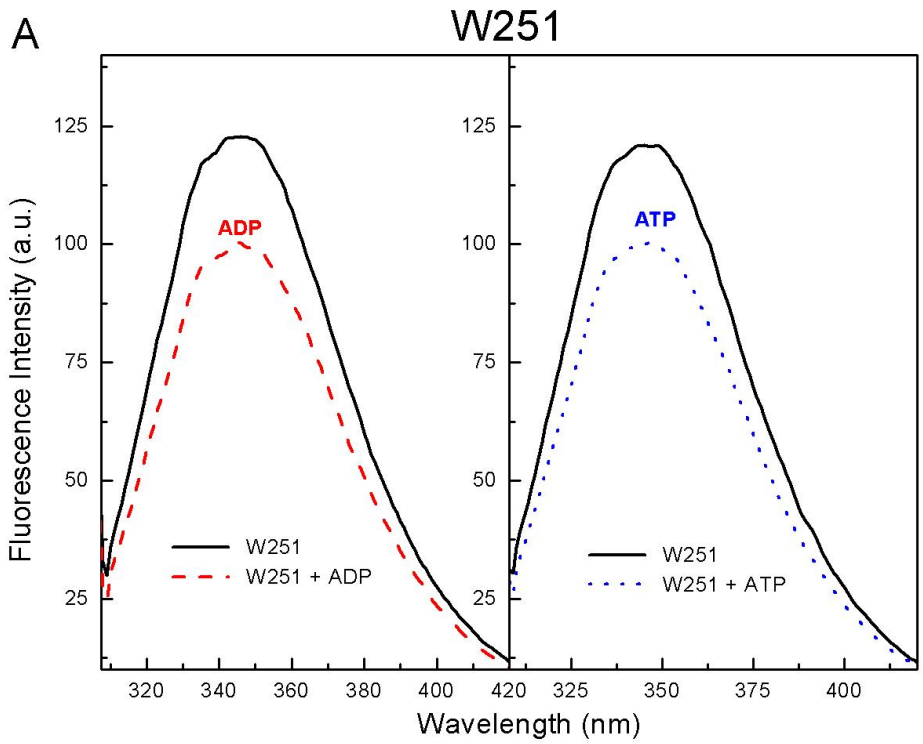
Table 3.1 Parameters of the urea-induced two-state unfolding transitions in single-Trp variants of ClpB.

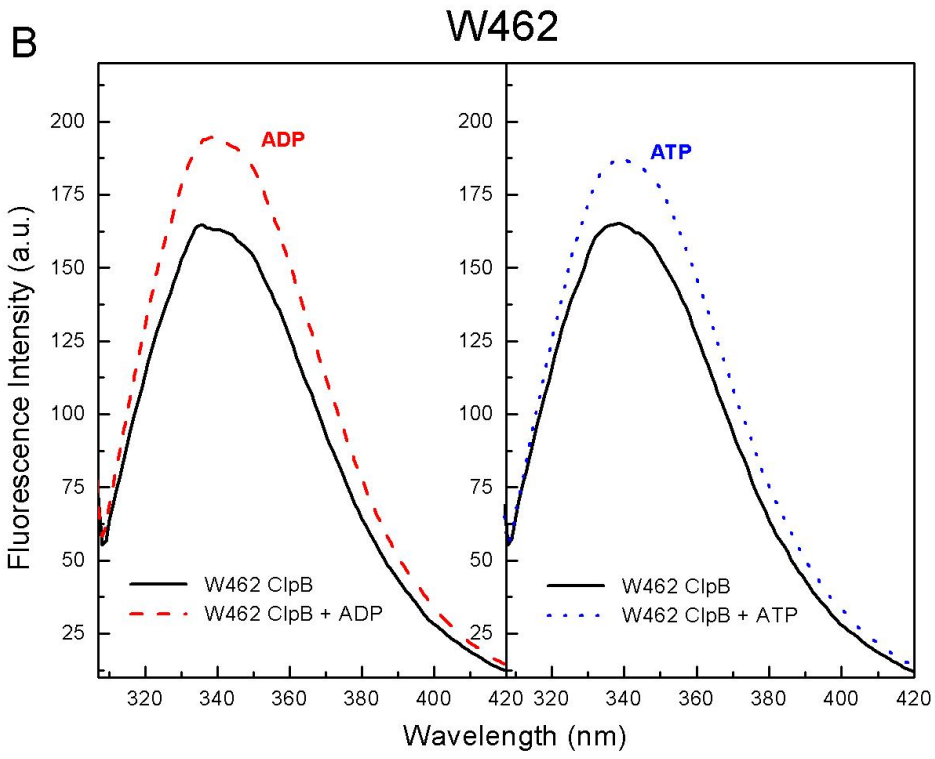
The standard energy of unfolding under urea-free conditions (ΔG_0) and the empirical constant (m) is shown for the single-Trp variants (that showed cooperative urea-induced transitions) for the monomeric and the oligomeric condition (see text and ref.(Pace, 1986)).

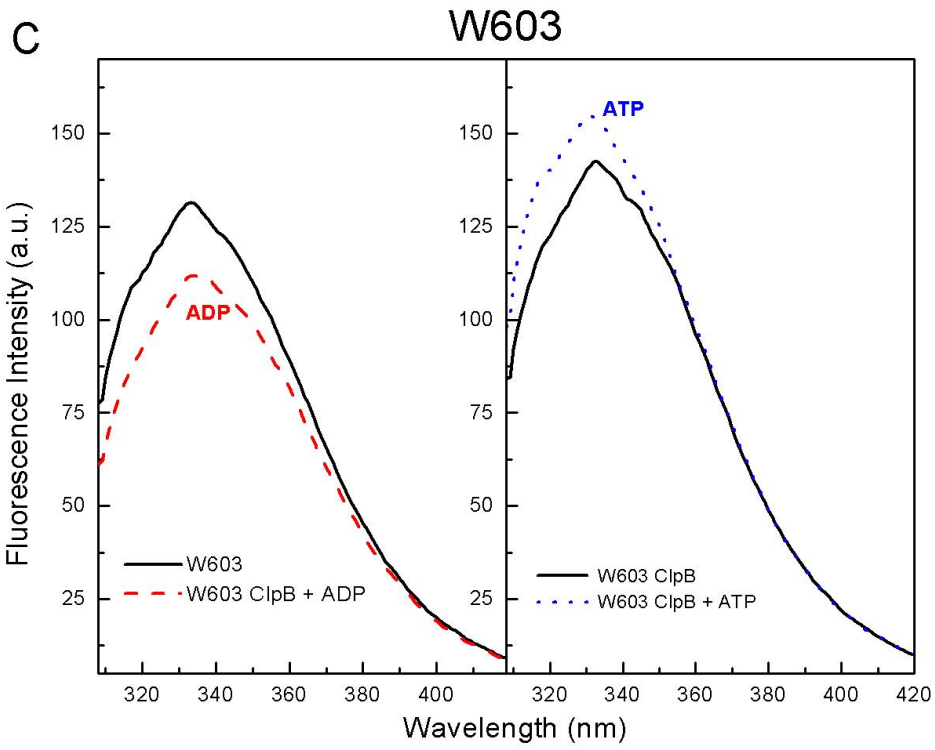
ClpB variant	Monomer		Oligomer	
	ΔG_0 [kJ/mol]	m [kJ/(mol M)]	ΔG_0 [kJ/mol]	m [kJ/(mol M)]
W105	9.6 ± 3.2	2.7 ± 0.9	11.1 ± 3.7	3.0 ± 1.0
W462	9.3 ± 2.1	4.2 ± 0.8	6.0 ± 2.1	3.2 ± 0.8
W543	20.8 ± 1.8	9.0 ± 0.8	21.8 ± 1.8	8.9 ± 0.7
W603	26.1 ± 6.8	6.4 ± 1.6	21.3 ± 7.0	5.5 ± 1.7
W812	16.2 ± 4.1	6.1 ± 1.4	5.7 ± 1.2	3.4 ± 0.5

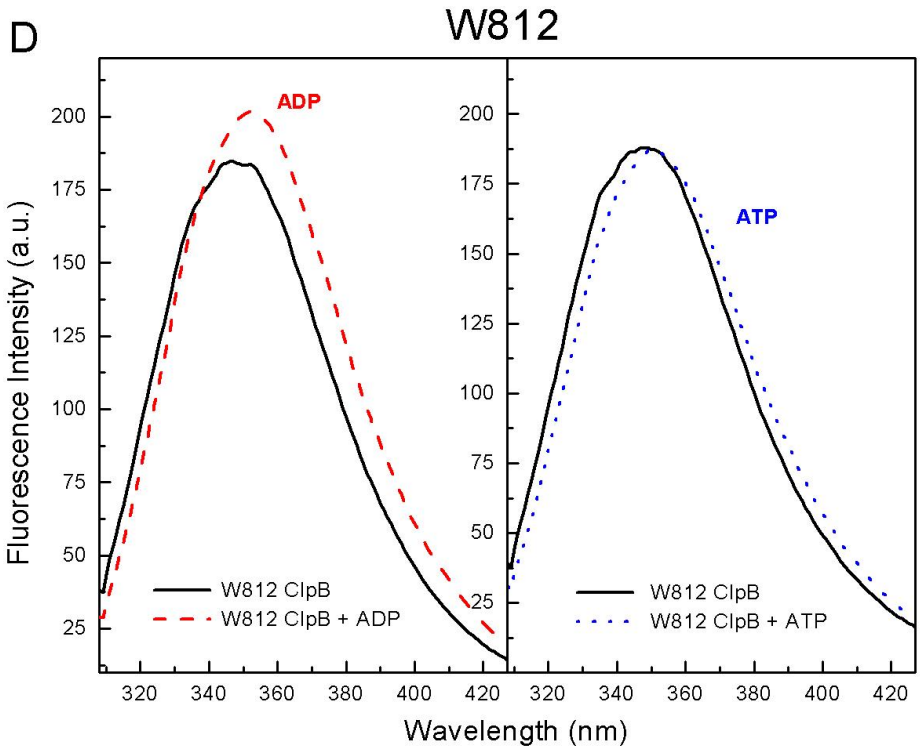
Figure 3.5 Nucleotide-induced changes in fluorescence intensity of single-Trp ClpB variants.

Fluorescence emission spectra of the four single-Trp variants of ClpB that responded to the nucleotide binding: W251 (A), W462 (B), W603 (C), W812 (D). Spectra were measured in low-salt buffer (50 mM Tris, pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA, and 1mM DTT) without nucleotides (left and right panels, black solid line), with 2 mM ADP (left panels, red dashed line) or with 2 mM ATP (right panels, blue dotted line). The concentration of the ClpB variants was 0.6 mg/ ml and the excitation wavelength was set at 300 nm.




B
W462
ADP
ATP
Fluorescence Intensity (a.u.)
Wavelength (nm)
W462 ClpB
W462 ClpB + ADP
W462 ClpB
W462 ClpB + ATP


C
W603
ADP
ATP
Fluorescence Intensity (a.u.)
Wavelength (nm)
W603
W603 ClpB + ADP
W603 ClpB
W603 ClpB + ATP


D
W812
ADP
ATP
Fluorescence Intensity (a.u.)
Wavelength (nm)
W812 ClpB
W812 ClpB + ADP
W812 ClpB
W812 ClpB + ATP
200
175
150
125
100
75
50
25
320
340
360
380
400
420

CHAPTER 4 - Development of a novel assay for studying ClpB interaction with strongly aggregated proteins

Introduction

Understanding the mechanism of protein disaggregation mediated by ClpB involves characterization of the initial step of recognition and binding of aggregated substrates. However, isolation of complexes of ClpB with aggregated substrates is a challenging task partly due to the poorly defined nature of the aggregated proteins and partly due to the fact that ClpB is a highly dynamic protein. While DnaK has a high affinity for substrates in the ADP-bound form (Mayer and Bukau, 2005), stabilization of ClpB interaction with substrate proteins requires the presence of ATP. In early reports, even as aggregated proteins were detected in complex with the DnaK system, both *in vivo* and in cell extracts, complexes between ClpB and proteins aggregates were not identified (Mogk et al., 1999). This observation was attributed to the transient nature of the ClpB-substrate interaction, as ClpB quickly hydrolyzes the bound ATP.

The first evidence for a direct interaction between ClpB and substrate proteins was obtained with the help of a ClpB variant with mutations of the conserved glutamates in the Walker B motifs in both AAA+ domains (E279A/E678A) (Weibezahn et al., 2003). This double mutant, referred to as ClpBTrap, is able to bind ATP, but is unable the hydrolyze it, thus “traps” ClpB in ATP-bound conformation that has high affinity for substrates. Although ClpBTrap is able to stabilize interaction with substrates, it

represents a dead-end chaperone as it is unable to perform work on the bound substrates. An analogous double Walker B mutant was also produced previously in our laboratory (E279Q/E678Q) and it was successfully used to detect the ATP-dependent interaction of ClpB with aggregated substrates using gel filtration (GF) chromatography (Barnett et al., 2005). The interaction of ClpBTrap with aggregated proteins is detected in presence of ATP by co-elution of ClpB with the large aggregates. Although GF chromatography is a straightforward and undoubtedly useful approach for the study of ClpB-aggregate interaction, it has several limitations. The borderline solubility of the strongly aggregated proteins, the long separation time, and large dilution of the proteins on the column are the most obvious drawbacks. The goal of this project was to develop a new assay that can be used quickly and efficiently to detect interaction between large aggregates and ClpB.

Characteristics of the aggregated substrates

Enzymes are often used as aggregated substrates for chaperones since their refolding/ reactivation can be easily followed by the regaining of their enzymatic activity. Previously, we developed a modified version of the procedure used by Goloubinoff and coworkers (Diamant et al., 2000) for the production of aggregated glucose-6-phosphate dehydrogenase (G6PDH) (Barnett et al., 2005). G6PDH aggregates can be produced by first unfolding the native G6PDH in 5 M urea at 47˚C followed by rapid dilution into a refolding buffer and further incubation at 47˚C. The size of the aggregates can be controlled by the initial concentration of G6PDH and the incubation time in the refolding buffer. The aggregation process can be arrested by a rapid cooling of the samples and the resulting aggregates remain stable for several hours. It has been shown that DnaK

system alone can reactivate small aggregates, but as the size of the aggregates increases the reactivation process requires the chaperone activity of ClpB (Diamant et al., 2000). Thus the aggregates used as substrates for the ClpB-DnaK bichaperone system are often very large and have borderline solubility.

Indeed, GF chromatography confirmed that large G6PDH aggregates elute at the void volume fraction of a Superose 6 column corresponding to a molecular weight larger than 2,000 kDa (Barnett et al., 2005). Imaging with TEM showed that the aggregated G6PDH has a globular shape with an average diameter of approximately 50 nm (results not shown). The aggregation process is very sensitive to the initial concentration of the protein and the environmental conditions. Thus, there is a fine line between producing soluble non-turbid aggregates or turbid aggregates with low solubility. The HPLC system and the GF column can be seriously damaged by large particulate aggregates, thus extra care must be taken to produce aggregates that are large but still applicable to the HPLC system.

Filter assay procedure

The new filter assay is based on the large difference between the size of the hexameric ClpB and aggregated proteins. It is designed especially to detect interaction of ClpB with large, even turbid, aggregated substrates. With this newly developed filter assay, the separation of the aggregate-bound fraction of ClpB from the free ClpB is achieved simply by using centrifugal filter devices. We chose filter units with Durapore membrane (it has the lowest protein binding capacity among the microporous membranes) with 0.1 μm pore size. These units have a maximum capacity 0.5 ml and a hold-up volume of only 5 μl (Amicon-Ultrafree MC, Millipore).

To study ClpB binding to aggregated substrates, the aggregated proteins were mixed with a solution containing ClpB and excess of nucleotides, than incubated for a short time while shaking at a slow speed. A small aliquot of the aggregate-ClpB mixture was applied to the filter unit and centrifuged. Small proteins passed through, while large aggregates were retained by the filter. An additional washing step with a buffer containing the appropriate nucleotide was included to decrease the amount of nonspecifically or weakly bound proteins. Proteins retained on the membrane were eluted with SDS buffer after short incubation at high temperature. The resulting filtrates (small proteins not retained by the filter) and eluates (fractions eluted from the filter with SDS buffer) were analyzed by SDS-PAGE followed by Coomassie blue staining (see flowchart below, Fig. 4.1).

Results and Discussion

An example of a filter assay used for the detection of ClpB bound to aggregated G6PDH is illustrated in Fig. 4.2. Native G6PDH as well as the majority of ClpB are found in the filtrate fraction (left panel), while aggregated G6PDH was retained on the filter and found in the eluant fraction (right panel). Trace amounts of ClpB were also retained on the filter in control samples of native G6PDH, as well as in aggregate containing samples in absence of nucleotides or with ADP and ATP. The interaction between wild type ClpB and G6PDH aggregates was detected only in presence of the poorly-hydrolyzable ATP analog, as significant amounts of ClpB were found in the aggregate fraction only in presence of ATPγS. This is in agreement with the previously observed nucleotide-dependence of the ClpB-aggregate interaction (Diamant et al., 2000; Barnett et al., 2005), namely with the observation that the ClpB-aggregated

substrate interaction is stabilized in the presence of ATP when the ClpBTrap variant is used. Similarly, our results show that binding of the poorly-hydrolysable ATPγS to wild type ClpB reproduces the frozen ATP-bound conformation of ClpBTrap, and thus stimulates the interaction of ClpB with the aggregated G6PDH. In the next chapter, examples will be provided for detection of the interaction of ClpBTrap with the aggregated substrates in the presence of ATP and the DnaK cochaperone system.

Additionally, the filter assay was successfully used to detect binding-deficiencies of a ClpB double Walker A mutant protein ClpB (T213N/T612N). For this experiment a second enzyme was also used as aggregated substrate for ClpB. Similar to G6PDH aggregates, it has been also shown that malate dehydrogenase (MDH) can be reactivated by the combined action of ClpB and the DnaK chaperone system (Goloubinoff et al., 1999). The filter assay was performed as described above and the fractions retained on the filter and subsequently solubilized with an SDS buffer (eluates) were analyzed by SDS-PAGE (Fig. 4.3). The control samples in lanes 1-4 confirm that the native G6PDH (Fig. 4.3A) or MDH (Fig. 4.3B) is not retained on the filter, while both aggregated G6PDH and aggregated MDH are retained on the filter (lanes 5-12). Only background amounts of wt ClpB and ClpB (T213N/T612N) were retained on filters in the absence of aggregates without nucleotides (lanes 1-2) or in presence of ATPγS (lanes 3-4), as well as in presence of aggregates in nucleotide free condition (lanes 5-6) or in presence of ATP (lanes 7-8) or ADP (lanes 11-12). As expected, binding of wt ClpB to G6PDH aggregates or to MDH aggregates was stimulated by ATPγS (lane 9). However, these experiments revealed the aggregate-binding deficiencies of ClpB (T213N/T612N), as even in presence of ATPγS, the double Walker A mutant showed

only background binding to aggregated G6PDH (lane 10, panel A) or aggregated MDH (lane10, panel B).

As shown above, the filter assay is a simple procedure that can be used to study the interaction between ClpB, or different mutated variants of ClpB, with aggregated substrates. This assay has several advantages over the previous assay used for the detection of ClpB-aggregates interaction.

Detection of ClpB-aggregate interaction with GF chromatography requires ten times more concentrated solutions of ClpB and aggregates, larger volumes and longer running times for each experiment. When more than one condition has to be checked (like binding in presence of different nucleotides) a separate run has to be performed for each condition and the GF column has to be equilibrated accordingly between the different runs. Extra care has to be taken to prevent accumulation of aggregated proteins that can damage the GF column and the HPLC system. In addition, fresh sample of aggregated protein has to be prepared immediately prior to each run. Also, in order to detect wt ClpB interaction with aggregates in the presence of ATP, the collected fractions have to be concentrated approximately 20 times to allow detection with SDS-PAGE. For the most part the ClpBTrap version has to be used to stabilize the interaction in the presence of ATP (Barnett et al., 2005). Due to the high cost of ATPγS, equilibration of the columns and running of a large number of experiments can become a major hassle.

In contrast the filter assay does not require highly concentrated solutions, works well with small volumes and large, even turbid aggregates. In addition, the separation of the aggregate-bound and free ClpB is a very quick process and allows for the

simultaneous processing of a large number of samples. Moreover, the same preparation of aggregated proteins can be divided between the different samples, which in turn makes the comparison among various samples and conditions more reliable. Due to the small volume required for the assay, ATPγS can be successfully used to stabilize the ClpB-aggregate interaction and the production of the ClpBTrap version is not required. This is especially important when several other ClpB mutants have to be screened, and the introduction of two more mutations is a time consuming process.

Most importantly, this simple assay could reproduce the previously published results regarding ClpB-aggregate interaction (Fig. 4.2) and additionally it has been used to characterize aggregate-binding capabilities of new ClpB mutants (Fig. 4.3). The disadvantage of the filter assay is that it can be only used to detect ClpB binding to large aggregates (like the examples presented above) and it is not possible to follow the ClpB-aggregate interaction in an aggregate-size dependent manner. For this purpose the GF chromatography-based method still remains the experiment of choice.

References

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

Diamant,S., Ben Zvi,A.P., Bukau,B., and Goloubinoff,P. (2000). Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. J. Biol. Chem. *275*, 21107-21113.

Goloubinoff,P., Mogk,A., Zvi,A.P., Tomoyasu,T., and Bukau,B. (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.

Mayer,M.P. and Bukau,B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol. Life Sci. *62*, 670-684.

Mogk,A., Tomoyasu,T., Goloubinoff,P., Rudiger,S., Roder,D., Langen,H., and Bukau,B. (1999). Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. EMBO J. *18*, 6934-6949.

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

Figures

Figure 4.1 Filter Assay Flow Chart.

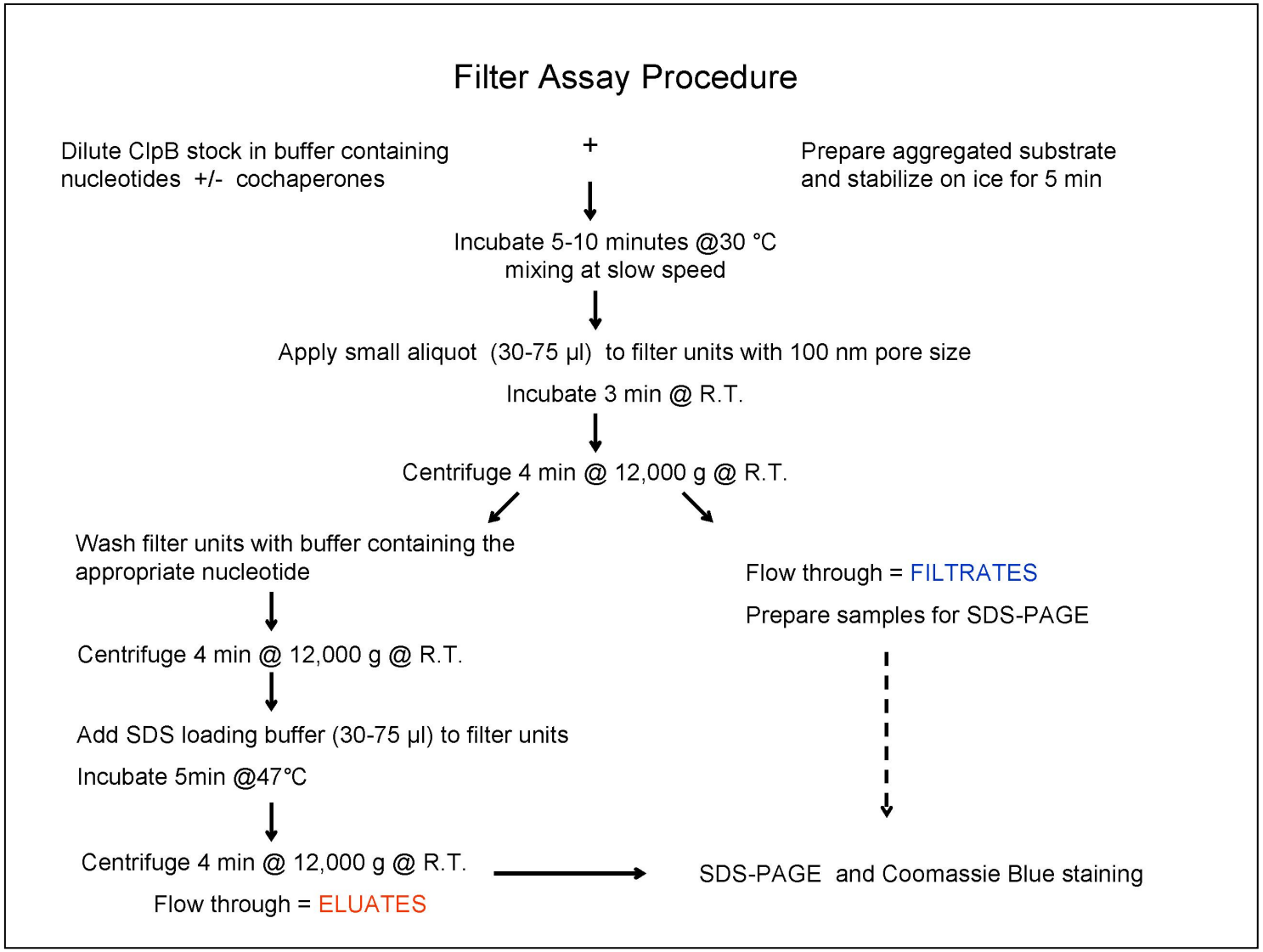


Figure 4.2 ClpB binds to aggregated G6PDH in presence of ATPγS.
1.5 µM ClpB was incubated with 2.7 µM native (N) or aggregated G6PDH without nucleotides or in the presence of 5 mM ATP, ATPγS or ADP, as indicated. 75 µl aliquots were applied to filter units and filtrates and eluates were separated by centrifugation. The results of the SDS-PAGE are shown with the filtrate and eluate fractions on the left and right side of the figure, respectively.

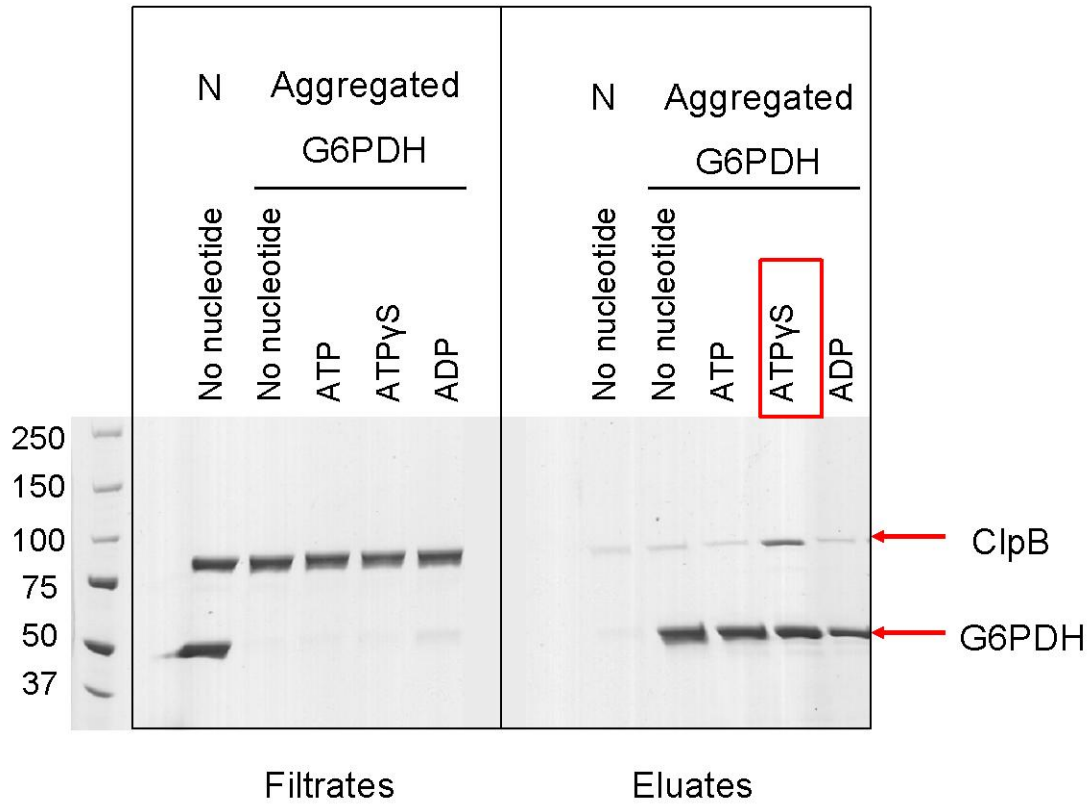
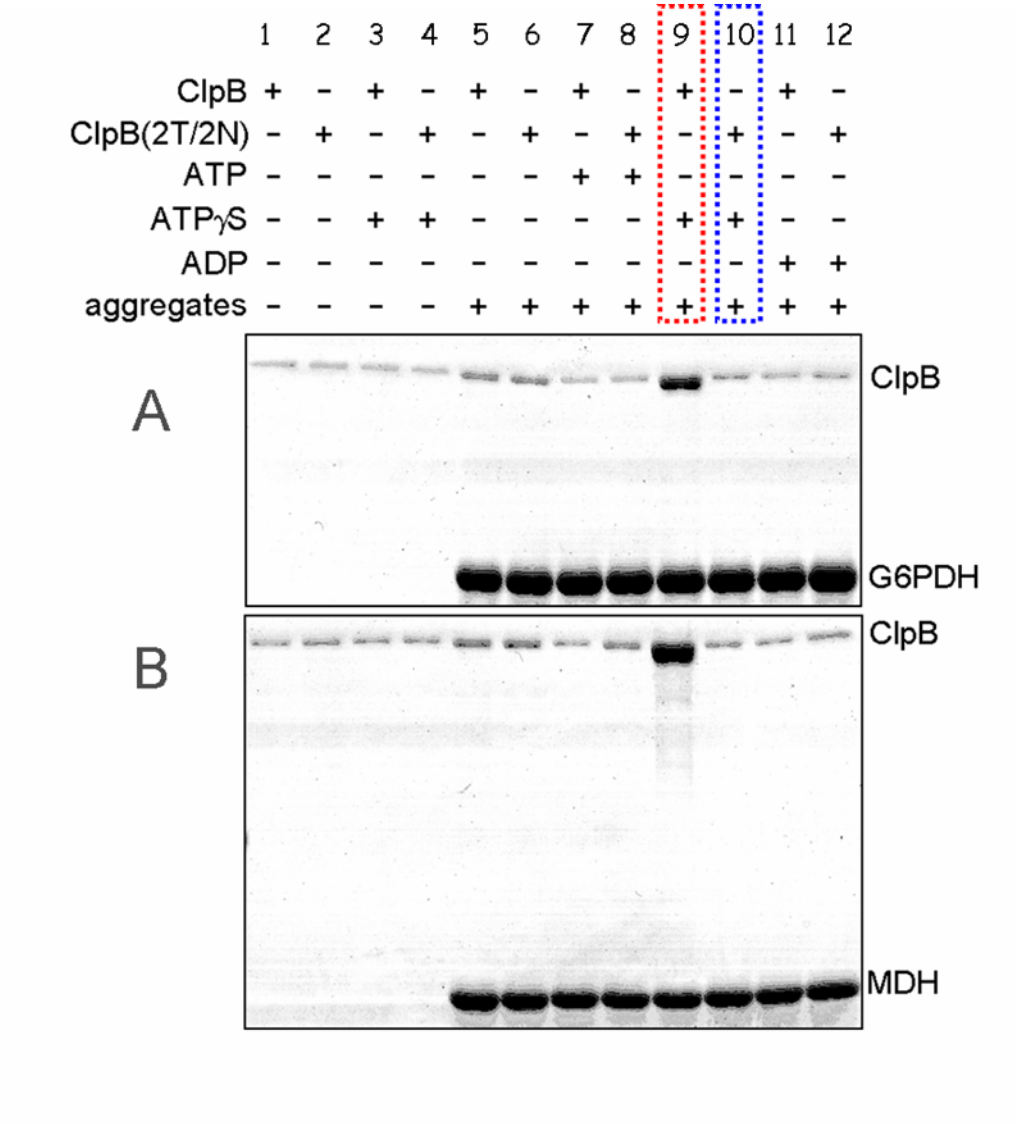


Figure 4.3 Interaction of ClpB and ClpB (T213N/T612N) with aggregated substrates.

1.5 μM wt ClpB, or ClpB(T213N/T612N) were incubated with 2.7 μM native or aggregated G6PDH (A) or 3 μM native or aggregated MDH (B) without nucleotides or with 5 mM ATP, ATPγS, or ADP. The result of SDS-PAGE is shown only for the fractions retained on the filter and subsequently solubilized with an SDS buffer (eluates). Lanes 1-4 are the control lanes with native proteins, while lanes 5-12 contain aggregated substrates.



CHAPTER 5 - Studies on functionality of the ClpB95/ClpB80 hetero-oligomers

Introduction

The ClpB transcript contains an alternative translation initiation site at the N-terminus of the first AAA+ module. Consequently, two isoforms of ClpB are produced *in vivo*: the full-length 95-kDa ClpB (ClpB95) and a truncated 80-kDa ClpB (ClpB80) that lacks the whole N-terminal domain of the full-length ClpB (Fig. 5.1) (Woo et al., 1992). Both isoforms of ClpB are localized in the bacterial cytoplasm and the truncated ClpB80 accounts approximately for one third of the total amount of ClpB (Chow and Baneyx, 2005).

Within the family of AAA+ ATPases, production of different-size isoforms is not restricted to ClpB. Truncated isoforms arising from alternative translation initiation sites were found for the close bacterial Clp ATPase ClpA, (Seol et al., 1994) and for spastin, a mammalian AAA+ ATPase involved in microtubule severing (Claudiani et al., 2005). The alternative N-termini in isoforms of AAA+ ATPases are located between the AAA+ modules and the “attachment” domains, which suggests that such domains may be functionally dispensable. The reasons for production of multiple isoforms of AAA+ ATPases are, however, poorly understood.

It is relevant to investigate the biological significance of production of two ClpB isoforms and the role of the truncated ClpB isoform in the ClpB-mediated

disaggregation process, especially in the light of the controversy surrounding the role of the N-terminal domain of ClpB (as described in Chapter 1).

Recent published data indicated that the optimal survival of *E. coli* during heat-shock is achieved when both ClpB95 and ClpB80 are produced *in vivo* (Chow and Baneyx, 2005). This result suggests that the two isoforms of ClpB may cooperate in producing a highly efficient chaperone system.

Materials and Methods

Proteins

ClpB95, ClpB80, and their double-Walker B mutant variants ClpB95Trap and ClpB80Trap, as well as the glycine insertion variant GI144ClpB were produced as described in Chapter 3. Previously published procedures were used to purify DnaK (Zmijewski et al., 2004), and DnaJ (Zylicz et al., 1985). GrpE was obtained from StressGen Biotechnologies (Victoria, Canada). Glucose-6-phosphate dehydrogenase (G6PDH) from *Leuconostoc mesenteroides* was obtained from Sigma (St. Louis, Missouri) in a lyophilized form. Porcine heart mitochondrial malate dehydrogenase (MDH) used in the reactivation assays and in the filter assays were obtained from MP Biomedicals (Irvine, California) and from Roche Pharmaceuticals, respectively. Protein concentrations were determined spectrophotometrically and are given in monomer units.

Production of protein aggregates

To prepare aggregated G6PDH, 440 μM stock solution of G6PDH was diluted two times with the heated denaturation buffer (10 M urea, 16 % glycerol, and 40 mM DTT). The mixture was incubated at 47 °C for 5 min, after which the G6PDH sample

was quickly diluted ten times with buffer A (50 mM triethanolamine/Cl, pH 7.5, 20 mM $Mg(OAc)_2$, 30 mM KCl, 1 mM β-mercaptoethanol, and 1 mM EDTA) and was mixed vigorously and incubated at 47 °C for 1 min or 15 min. The solution was then mixed briefly and incubated further on ice for 2 min followed by centrifugation at 4 °C for 10 min at 13,000 g. The samples prepared for the HPLC runs contained an additional 2 mM ATP in buffer A.

To prepare aggregated MDH for the reactivation assays, 250 μM stock solution was diluted 20-fold in buffer B (100 mM Tris pH 7.4, 120 mM KCl, 15 mM $Mg(OAc)_2$, 8 mM DTT) and incubated at 47 °C for 60 min. The solution was then mixed and incubated on ice for 2 min. MDH aggregates for the filter assays were prepared the same way as aggregated G6PDH.

Protein reactivation assay

Aggregates of G6PDH (22 μM) were diluted 8-fold in buffer A with 6 mM ATP containing either no chaperones, 1 μM DnaK, 1 μM DnaJ and 0.5 μM GrpE (KJE) with 1.5 μM ClpB95, KJE with 1.5 μM ClpB80, or KJE with the mixtures of ClpB95 and ClpB80 (total 1.5 μM ClpB). Aggregates of MDH (12.5 μM) were diluted to 3 μM in buffer B with 6 mM ATP containing either no chaperones, KJE with 1.5 μM ClpB95, KJE with 1.5 μM ClpB80, or KJE with the mixtures of ClpB95 and ClpB80 (total 1.5 μM ClpB). After incubation at 30 °C, aliquots were withdrawn and the G6PDH or MDH activity was measured.

G6PDH and MDH activity assay

G6PDH samples were incubated at 30 °C in 50 mM Tris-HCl pH 7.8, 5 mM $MgCl_2$ with 2 mM glucose-6-phosphate and 1 mM $NADP^+$. The absorption at 340 nm was measured after 5 min.

MDH samples were incubated at 25 °C in 50 mM Tris-HCl pH 7.4, 2 mM DTT, 0.3 mM oxaloacetate, and 0.15 mM NADH. The absorption at 340 nm was measured after 20 min.

Analytical ultracentrifugation

A Beckman Optima XL-I analytical ultracentrifuge was used in sedimentation studies. Sedimentation velocity experiments were performed at 50,000 rpm at 20 °C using two-channel aluminum analytical cells. ClpB solutions were prepared in 50 mM Tris-HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$ 1mM EDTA and 2 mM β-mercaptoethanol. The data were analyzed using the time-derivative method (Stafford, III, 1992) and software distributed with the instrument.

Gel filtration chromatography

Protein samples were injected onto a 25-cm Pharmacia HR10/30 column packed with Superose 6 gel filtration media (Amersham Biosciences) equilibrated with buffer A with 2 mM ATP. Elution at 0.5 ml/min was performed at room temperature using a Shimadzu HPLC LC10ATvp equipped with a SPD-M10Avp diode-array detector. One-minute fractions (0.5 ml) were collected from the column and analyzed by SDS-PAGE followed by Coomassie staining. Band densitometry was performed using Scion Image software (www.scioncorp.com).

ClpB ATPase assay

ClpB samples were incubated for 15 min at 37 °C in 100 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 5 mM ATP, 1 mM EDTA, 1 mM DTT with 0.2 mg/ml α-casein (Sigma) or 5 µM peptide B1. Inorganic phosphate production was determined using the malachite green method (Hess and Derr, 1975; Lanzetta et al., 1979).

Peptide B1 (AHAWQHQGKTLFISRKTYRIC) was synthesized at KSU Biotech Core Facility and further purified on reverse phase-HPLC (RP-HPLC).

Normal mode analysis

The PDB coordinates of the iMolTalk hexameric model of full-length *T.thermophilus* ClpB were downloaded from http://i.moltalk.org/ . The ElNemo web server (http://www.igs.cnrs-mrs.fr/elnemo/) was used to compute a b-factor or temperature factor for each residue based on the first 100 normal modes that characterizes the mobility of the residues. One set of calculations was performed on the hexamer composed of six full-length chains, while in the second case, the hexamer was composed of three long chains (A, C, E) and three short ones (B, D, F) with the N-terminal domain (residues 4-139) removed.

Results

Chaperone activity of ClpB80 and ClpB95

Research regarding the role of the N-terminal domain of ClpB was initiated by former members of our lab. It has been shown that the N-terminal domain of ClpB95, which is missing in ClpB80, contributes significantly to the aggregate-binding affinity. Most importantly, the N-terminal domain becomes essential for binding large protein aggregates of G6PDH (Barnett et al., 2005).

In order to compare the chaperone activity of the two isoforms we followed the reactivation of G6PDH in presence of ClpB and the DnaK/DnaJ/GrpE system of cochaperones as a function of the size of the aggregates. Small or large aggregates of G6PDH were produced (as described in Chapter 4) by arresting the aggregation process after 1min or 15 min refolding time, respectively. Gel filtration chromatography confirmed that aggregates produced after 1 min of refolding time, showed a broad distribution of particles with a maximum at ~1,000 kDa, while the aggregates produced after 15 min of refolding time eluted in the void-volume fraction corresponding to >~2,000 kDa with a minor fraction of smaller aggregates (Fig.1 from ref. (Barnett et al., 2005)).

We compared the rates of G6PDH reactivation by different chaperones for the aggregate population prepared with 1 min and 15 min refolding time. As shown in Fig. 5.2A (open circles), the 1-min aggregate population showed measurable G6PDH activity that did not increase over time, which indicates that the aggregate population contains a fraction of refolded native G6PDH. The fact that the amount of refolded G6PDH did not increase spontaneously over time is consistent with the irreversible nature of protein aggregation. Either DnaK/DnaJ/GrpE (KJE) alone or in cooperation with ClpB reactivated aggregated G6PDH, but the reactivation rate was ~4-fold higher when ClpB95 was present. The rate of G6PDH reactivation was ~2-fold lower when the full-length ClpB was replaced with ClpB80.

In contrast to the 1-min aggregate sample, the 15-min aggregate population had no measurable G6PDH activity (Fig. 5.2B, open circles), which indicates that essentially all G6PDH aggregated during 15 min of refolding. The DnaK/DnaJ/GrpE system alone

did not reactivate G6PDH, which is consistent with previous observations (Diamant et al., 2000).

Either ClpB95 or ClpB80 reactivated G6PDH, but the reactivation rate was in this case ~5-fold lower with ClpB80 than with the full-length ClpB. Collectively, the results in Fig. 5.2 indicate that the chaperone activity of ClpB80 is weaker than that of ClpB95. The ability of ClpB80 to reactivate aggregated G6PDH decreases as the size of the aggregates increases (comp. Figs. 5.2A and 5.2B). This result is consistent with the weaker binding of ClpB80 to the large aggregates of G6PDH. The fact that ClpB80 is less efficient than ClpB95 in reactivation of strongly aggregated G6PDH questions the role of ClpB80 isoform in the cellular protein disaggregation. Also, if the truncated version is less efficient then why has the alternative translation initiation site in the ClpB gene not been eliminated during evolution?

Combined chaperone activity of ClpB80 and ClpB95

Since DnaK system alone could not reactivate the large aggregates of G6PDH (produced with 15 min refolding time), and also the requirement for the N-terminal domain of ClpB95 was essential for the reactivation of the aggregates, we used large aggregates of G6PDH to follow the combined chaperone activity of ClpB95 and ClpB80. The reactivation assays were performed in the presence of the DnaK system and different molar fractions of ClpB80 (f_{ClpB80}) while keeping the total concentration of ClpB constant among the various samples.

Since ClpB80 was unable to reactivate efficiently the G6PDH aggreagates, it was expected that the “dilution” of the active ClpB95 chaperone with ClpB80 would lower the chaperone activity of the mixture as compared to the same concentration of ClpB95.

Unexpectedly, we observed a strong activation of the chaperone activity (over 100% increase in the G6PDH reactivation rate) when both ClpB95 and ClpB80 were used in the reactivation assay (Fig. 5.3). The highest chaperone activity was observed for 1:1 molar ratio of ClpB80 and ClpB95 ($f_{ClpB80} = 0.5$) (Fig. 5.3, inset). We wanted to see if this enhancement of chaperone activity, observed in presence of the mixtures of ClpB80 and ClpB95, is specific to the aggregates of G6PDH, or it is a more general property of ClpB. Consequently, we followed the reactivation rates of thermally-aggregated MDH, whose reactivation does not appear to depend on the presence of the N-terminal domain of ClpB to the extent observed for aggregated G6PDH (Mogk et al., 2003). Indeed, these reactivation assays showed that, in contrast to the result obtained in case of G6PDH aggregates, the MDH aggregates were reactivated with similar rate by ClpB80 or ClpB95 in the presence of the DnaK system (note the lack of difference between the traces for $f_{ClpB80} = 0$ and $f_{ClpB80} = 1$) (Fig. 5.4). But most importantly, just like in case of the G6PDH, the reactivation rate of aggregated MDH significantly increased in the presence of both ClpB95 and ClpB80 with the maximum at 1:1 molar ratio of the ClpB isoforms (Fig. 5.4, inset). These results demonstrate that the mixtures of ClpB95 and ClpB80 work more efficiently as aggregate reactivating chaperones than either ClpB95 or ClpB80 alone.

Self-association of ClpB80 and ClpB95

The synergistic activation of the ClpB chaperone system by a concerted action of ClpB95 and ClpB80 might be linked to the formation of hetero-oligomers containing both ClpB isoforms. Indeed, Kim et al. detected interactions between ClpB80 and His-tagged ClpB95 on a nickel column in presence of ATP (Kim et al., 2000). It has been shown by

our lab previously that the self-association affinity of ClpB80 is higher than that of ClpB95, which correlates with a higher basal ATPase activity of ClpB80 (Barnett et al., 2000) (see also Fig. 5.10, below). We asked, whether ClpB80 could stimulate oligomerization of ClpB under conditions that do not favor the oligomerization of ClpB95, i.e. at low protein concentration and in the absence of nucleotides. Sedimentation velocity experiments showed that 5 μM ClpB95 was monomeric in the absence of nucleotides (sedimentation coefficient $s_{20,w}$ ~ 4 S), as also was 2.5 μM ClpB80 (Fig. 5.5A). However, when 5 μM ClpB95 was combined with 2.5 μM ClpB80, a significant population of oligomers ($s_{20,w}$ ~ 13 S) was observed (Fig. 5.5B). Since neither of the two isoforms formed homo-oligomers at such protein concentration, the oligomers found in Fig. 5.5B represent hetero-oligomers of ClpB95 and ClpB80. This result indicates that the formation of hetero-oligomers is preferred over the formation of homo-oligomers at low concentration of ClpB95 and ClpB80.

In the experiment shown in Fig. 5.5B, the total ClpB concentration was 7.5 μM. ClpB95 at similar concentration remained monomeric (Fig. 5.5C), but ClpB80 formed homo-oligomers with the sedimentation coefficient similar to that of the hetero-oligomers shown in Fig. 5.5B. This result indicates, first, that the self-association affinity of ClpB80 is indeed higher than that of ClpB95 and, second, that the size of the hetero-oligomers is similar to that of ClpB80 homo-oligomers (i.e. hexameric). Since the difference in molecular weight between ClpB95 and ClpB80 is only ~15%, the resolution of sedimentation studies is not sufficient to determine the exact composition of hetero-oligomers, nor it is possible to determine whether their population is homo-or heterogeneous.

Binding of the ClpB hetero-oligomers to aggregated substrates

Sedimentation velocity confirmed that the two isoforms can form hetero-oligomers. In an attempt to explain the enhanced chaperone activity of the hetero or mixed-oligomers, we investigated interactions of the aggregates with substrate-trapping variants of ClpB95 and ClpB80 (Weibezahn et al., 2003). As described in Chapter 4 of this work, interaction between ClpBTrap (E279Q/E678Q) and protein aggregates can be detected by gel filtration chromatography in presence of ATP (Barnett et al., 2005). Large aggregates of G6PDH eluted from the gel filtration column equilibrated with ATP at 11-13 min (see Fig. 5.6 and ref. (Barnett et al., 2005)). ClpB95Trap and ClpB80Trap formed hexamers in the presence of ATP and eluted at 21-23 min. In the presence of aggregated G6PDH, a significant fraction of ClpB95, but only a trace amount of ClpB80, eluted at 11-13 min (Fig. 5.6AB) This result indicates that the N-terminal domain of ClpB, which is missing in ClpB80, is essential for binding to the aggregates, consistent with our previous results (Barnett et al., 2005). When ClpB95 and ClpB80 (1:1 molar ratio) were incubated with aggregated G6PDH, both isoforms of ClpB were found in the aggregate containing fractions (Fig.5.6C). We performed band densitometry of the aggregate-containing gel filtration fractions eluting at 11, 12, and 13 min (Fig. 5.6D). We found that the total amount of ClpB (ClpB95+ClpB80) bound to the aggregates in Fig. 5.6C was not higher than the amount of ClpB95 bound to the aggregates in Fig. 5.6A. The ratio of ClpB80/ClpB95 was 0.40 ± 0.07 in the three aggregate fractions (see Fig. 5.6D, filled circles), which suggested that the hetero-hexamers of the Walker B mutant ClpB that formed in the presence of ATP and bound to G6PDH aggregates contained an average number of 2 ClpB80 and 4 ClpB95 subunits.

Binding of ClpB95Trap and ClpB80Trap to aggregated G6PDH was also investigated with the filter assay. Aggregates were incubated with the chaperones in presence of ATP prior to filtration and the fractions retained on the filters were analyzed by SDS-PAGE (as described in Chapter 4). The results obtained in Fig. 5.7A are consistent with the results of the gel filtration, and confirm that the hetero-oligomers do not possess a higher affinity for aggregated G6PDH as compared to the homo-oligomers of ClpB95 (compare lane 3 and lane 7, Fig. 5.7A). The weaker binding of ClpB80 to G6PDH aggregates was also observed with this assay. Also in case of aggregated MDH the total amount of ClpB (ClpB95+ClpB80) bound to the MDH aggregates (lane 7, Fig. 5.7B) is not higher than the amount of ClpB95 (lane 3, Fig. 5.7B). The N-terminally truncated isoform interacted stronger with the aggregates of MDH than with the G6PDH aggregates (lane 5, Fig. 5.7 AB). Indeed, during the reactivation of MDH aggregates, ClpB80 was just as effective as ClpB95 (Fig. 5.4).

ClpB cooperation with the DnaK system

In addition to ClpB, the G6PDH and MDH reactivation assays contained also the DnaK/DnaJ/GrpE system of cochaperones. Existing data indicate an essential role of the DnaK system during the early steps of the disaggregation process, before the substrates are threaded through the central channel of the ClpB oligomers (Weibezahn et al., 2005; Zietkiewicz et al., 2006). Thus, the DnaK system could influence the ClpB-aggregate interaction. When DnaK/DnaJ/GrpE was included in the binding assays, we did not detect with the filter assay an increase in the amount of ClpB95, ClpB80 or the total amount of ClpB (ClpB95+ClpB80) bound to the aggregated G6PDH or aggregated MDH (lanes 4, 6, and 8, compared to lanes 3, 5, and 7, respectively, Fig. 5.7 A and B).

Instead, we could detect a slight decrease in the aggregate-bound ClpB that was clearly visible in case of ClpB95 (lane 3 compared to lane 4, Fig. 5.7A and B). This suggests that ClpB and DnaK compete for the same binding sites on the aggregates. Thus, the hetero-oligomers do not bind more aggregated G6PDH or aggregated MDH than the homo-oligomers of ClpB95, not even in the presence of the DnaK system.

It has been shown that preincubation of certain protein aggregates (GFP and luciferase) with DnaK/DnaJ results in a faster initiation of the disaggregation and subsequent reactivation by the ClpB-DnaK system of cochaperones (Zietkiewicz et al., 2004). Recently the same group proposed that large aggregates are fragmentated into smaller ones upon preincubation with DnaK (Zietkiewicz et al., 2006).

We tested the role of the DnaK system on the aggregate-binding affinity of ClpB and on the protein reactivation mediated by ClpB in order-of-addition experiments. To compare binding affinities, G6PDH aggregates were preincubated 10 minutes either with the DnaK/DnaJ/GrpE or with ClpB95Trap in presence of ATP, after which the missing chaperone was added and the mixture was incubated for another 10 minutes. No change in the amount of ClpB95 bound to the aggregated G6PDH was observed between the samples that were preincubated with DnaK and the samples that were preincubated with ClpB (lane 13 and 14, Fig. 5.8). Also, just like before (Fig. 5.7A) slightly more ClpB was found in the aggregate fraction in absence of DnaK system (lanes 9, 10 versus lanes 11, 12; Fig. 5.8) and this amount was similar to the amount of ClpB associated with aggregates in the preincubated samples (lanes 9, 10 and lanes 13, 14; Fig. 5.8). Overall the preincubation with DnaK does not seem to increase the amount of ClpB bound to the large aggregates of G6PDH. However, the reactivation of

G6PDH was affected by the preincubation with the DnaK system as compared to the preincubation with ClpB (Fig. 5.9). G6PDH reactivation was more efficient when the samples were preincubated with the DnaK system prior addition of ClpB95. This is in agreement with the proposed function of the DnaK upstream from the ClpB-translocation, as it could "prepare" the aggregates prior threading through the ClpB channel.

ATPase activity of the ClpB hetero-oligomers

Next, we investigated the ATPase activity of ClpB as a function of the ClpB80 fraction. In addition to the basal ATPase activity, the ATP hydrolysis was also measured in presence of pseudo-substrates that are known activators of the ClpB ATPase: α-casein and a positively-charged peptide, "peptide B1" (AHAWQHQGKTLFISRKTYRIC) that has been shown to interact with ClpB (Schlieker et al., 2004). In agreement with previous observations, the basal ATPase activity of ClpB80 was higher, while the casein-activated ATPase was lower, than that of ClpB95 (Fig. 5.10.A and C) (Barnett et al., 2000). However, the ATPase of ClpB80 was higher than that of the ClpB95 in presence of peptide B1 (Fig. 5.10 A). Additionally, we found that the basal ATPase activity of mixtures ClpB95/ClpB80 (Fig. 5.10A), as well as the activated ATPase in the presence of α-casein or peptide B1 (Fig. 5.10B, C) was approximately proportional to the ClpB80 fraction. Thus, unlike their chaperone activity (see insets in Fig. 5.3 and 5.4), the ATPase of ClpB95/ClpB80 mixtures did not show any "excess" activation above the capabilities of both isolated components. It has been shown previously that α-casein and positively-charged peptides bind to ClpB and are translocated through the channel in the ClpB hexamer even in the absence of the cochaperones (Weibezahn et

al., 2004). We assume, therefore, that substrate threading occurred under the conditions of experiments shown in Figs. 5.10B and C. Importantly, the rate of ATP hydrolysis by AAA+ “polypeptide-threading machines” is linked to the speed of substrate translocation (Kenniston et al., 2003). Thus, taking into account the results of Figs. 5.10B and C, we conclude that the enhanced chaperone activity of hetero-hexamers of ClpB95 with ClpB80 is not associated with an enhanced rate of substrate threading through the ClpB channel.

We also attempted to measure the ATP consumption during the time-course of the reactivation of aggregated G6PDH by ClpB95/ClpB80 in presence of the DnaK system under conditions of Fig. 5.3. Just like in case of basal and pseudo-substrate stimulated ATPase activity, we did not detect an enhanced ATPase in case of the hetero-hexamers as compared to homo-oligomers (Fig 5.10D) during the actual disaggregation process. However, it has to be mentioned that under the conditions of the reactivation assay, ClpB is not saturated by the aggregates, and also we are measuring the combined ATPase of ClpB and DnaK.

Normal mode analysis

The results of the sedimentation velocity experiments shown in Fig. 5.5 imply that the formation of a full ring of the N-terminal domains in hexameric ClpB is thermodynamically unfavorable, as compared to rings without some or all N-terminal domains. The “penalty” for forming a ring of the N-terminal domains may be due to an unfavorable enthalpy (steric repulsion) and/or unfavorable entropy (restriction of mobility). Indeed, a high mobility of the N-terminal domain in a non-physiological incomplete ClpB oligomer has been shown by the crystal structure as illustrated in Fig.

2.3 (Chapter 2) (Lee et al., 2003). We used a structural model of the hexameric ClpB from *T. thermophilus* (referred to as the "iMolTalk" hexameric model in Chapter 2) (Diemand and Lupas, 2006) to compute and compare the mobility of the N-terminal domain in a homo-hexamer of ClpB95 and in a hetero-hexamer with alternating ClpB95 and ClpB80 subunits. Fig. 5.11 shows the results of the elastic normal mode analysis (Suhre and Sanejouand, 2004) of oscillatory motions in ClpB hexamers. In the homo-hexamer of ClpB95 (Fig. 5.11A, C), the normal mode analysis detects high mobility of two domains located at the outside of the ClpB ring: the tip of the coiled-coil middle domain and the C-terminal small sub-domain of D2 AAA+ module.

The crystal structure of the non-physiological ClpB trimer also revealed that both domains are mobile, and among them the middle domain was found to possess the higher mobility. Also, earlier studies indicated that the mobility of the middle domain supports chaperone activity and cooperation with the DnaK system (Lee et al., 2003; Haslberger et al., 2007). It has been shown also that the D2 small sub-domain maintains the stability of the ring (Barnett et al., 2000). Domain stability studies (Chapter 3) (Nagy et al., 2006) revealed that only the middle domains and the D2 small sub-domains are destabilized upon oligomerization. These earlier observations validate the results of the normal mode analysis. Aditionally, the AAA+ modules showed no mobility as it was expected from domains that form the core of the ring-like hexamers of ClpB.

However, in the ClpB95/ClpB80 hetero-hexamer, a third mobile structural region was detected by normal mode analysis, which corresponds to the N-terminal domain (Fig. 5.11B, C). The mobility of the N-terminal domain was suppressed in the ClpB95 homo-hexamer as compared to the hetero-hexamer. This result indicates that the

excess thermodynamic stability of the ClpB hexamers with an incomplete N-terminal ring may arise from elimination of the entropic penalty due to restriction of mobility of the N-terminal domains. It was also apparent that the mobility of the small D2 sub-domain is increased slightly in the hetero-hexamers compared to the mobility in the homo-hexamers (Fig. 5.11C).

ClpB variant with a modified N-domain linker region

The crystal structure of the *T. thermophilus* ClpB revealed that the highly mobile N-terminal domain is connected to the first AAA+ module via a 14-residue long flexible linker (residues 145-158 in *E. coli* ClpB, Fig. 5.12A). At the beginning of this linker region we identified two Gly residues followed shortly by a Ser (-GGES-). Due to their flexibility, glycine residues are often found in linker regions, but alanines and serines are also common. In order to gain additional insights about the importance of the N-domain mobility in ClpB-mediated protein disaggregation we produced a variant of ClpB with a modified N-domain linker region. We proposed to modify the length, and possibly the flexibility of the ClpB N-domain linker region by introducing four additional glycines adjacent to the existing glycines (Fig. 5.12B). The resulting ClpB variant, GI144ClpB (glycine insertion after residue # 144), was tested for structural and functional integrity. Sedimentation velocity studies confirmed that the self-association ability of ClpB was not disrupted by the Gly insertions, since GI144ClpB formed oligomers at high protein concentration similar to ClpB95 (results not shown). Moreover, the glycine insertion mutant was able to hydrolyze ATP at the same rate as ClpB95 (0.46 v.s. 0.48 [nmol ATP hydrolyzed / pmol ClpB hr], Fig. 5.13), which is another indication that the structural integrity of the mutant was retained.

Additionally, the ATPase activity of GI144ClpB was stimulated by poly-lysine, a potent activator of the ClpB ATPase activity, to similar extent as that of the ClpB95. Also, while casein increased the ATPase activity of ClpB95 over 12 times, only a 2-fold increase of the ATP hydrolysis rate was detected in case of the glycine insertion mutant (Fig. 5.13).

Next, we tested the aggregate-binding capability and the chaperone activity of GI144ClpB. Interestingly, the ClpB variant with a modified N-domain linker region lost its ability to interact with aggregates of G6PDH in the presence of ATPγS (Fig. 5.14). This variant of ClpB was also inactive as a protein disaggregase, as shown by the G6PDH reactivation assay (Fig. 5.15). The chaperone activity could not be stimulated by additional presence of ClpB95, or ClpB80 in the reactivation assays.

Thus, even slight variations in the length and composition of the linker region that connects the N-terminal domain to the rest of the molecule leads to defect in aggregate binding and processing. The fact that the oligomerization and the basal ATPase activity was not affected at all by the insertion of the additional glycines to the N-domain linker suggests that the AAA+ module is functional. It has been shown that casein and poly-lysine affect the ATPase activity of ClpB by different mechanisms. The naturally unstructured casein acts more like a bona fide ClpB substrate and stimulates the ATPase of both AAA+ domains, while poly-lysine influences ATP hydrolysis by the second AAA+ domain (Strub et al., 2003).

Discussion

Biochemical properties of the two isolated ClpB isoforms were compared previously by several research groups (Park et al., 1993; Barnett et al., 2000; Chow et

al., 2005; Beinker et al., 2002). What has been possibly overlooked so far is the fact that both isoforms are produced and co-exist in bacterial cytoplasm. In this work, we discovered that ClpB achieves its full potential as an aggregate-reactivating chaperone through utilization of both isoforms and formation of hetero-oligomers.

We showed that the N-terminally truncated isoform, ClpB80 becomes a less effective chaperone as the size of the aggregates increases (Fig.5.2). This is in agreement with previous results that showed that the N-terminal domain contributes significantly to the aggregate binding affinity of ClpB and becomes essential for mediating the interaction with large protein aggregates (Barnett et al., 2005). It has been also reported that *in vivo* under conditions that lead to increased protein misfolding (background of defective DnaK, that could suppress misfolding and aggregation) ClpB95 is a more effective chaperone than ClpB80 (Chow et al., 2005).

Unexpectedly, we found that the chaperone system built from two ClpB isoforms works much better than the sum of its components (Figs. 5.3 and 5.4). We asked the question: where does the “excess” chaperone capability of ClpB95/ClpB80 come from, which is absent in either ClpB95 or ClpB80? As we have shown, the ClpB95/ClpB80 hetero-oligomers do not have a significantly stronger aggregate binding capability than ClpB95 alone (Fig. 5.6 A, B and Fig. 5.7 A, B). Moreover, the additional presence of the DnaK system did not increase the aggregate-binding affinity of ClpB homo-or hetero-oligomers (Fig. 5.7 and Fig. 5.8). Also, in agreement with published data (Haslberger et al., 2007; Zietkiewicz et al., 2004), we found that the DnaK system has an important role upstream of the translocation, during the initial steps of the disaggregation process, as we could detect a slightly enhanced reactivation of G6PDH aggregates upon

preincubation with DnaK system, as compared to the preincubation with ClpB95 (Fig. 5.9).

We also observed that the hetero-oligomers do not hydrolyze ATP faster than the sum of their components (Fig. 5.10). Importantly, we can assume that the threading of substrates through the ClpB channel, which is an essential part of aggregate reactivation (Weibezahn et al., 2004) and is tightly coupled to the ATP consumption, does not occur faster with the ClpB95/ClpB80 hetero-oligomers than with the homo-oligomers (Fig. 5.10B, C). This result suggests that the polypeptide threading is not the rate-limiting step of disaggregation and that another process, in which the hetero-oligomers are superior to homo-oligomers, determines the overall rate of aggregate reactivation, as measured in Figs. 5.3 and 5.4.

Sedimentation velocity experiments demonstrated that the self-association affinity of ClpB95 is lower than that of ClpB80 (see Fig. 5.5 C). This indicates that the formation of a full ring of the N-terminal domains is thermodynamically unfavorable.

Normal mode analysis revealed that the N-terminal domains in the homo-hexamer formed by the full-length ClpB95 are significantly inhibited in their mobility as compared to the hetero-hexamer of ClpB95 and ClpB80 (Fig. 5.11C). This explains the results of the sedimentation velocity experiments, that the formation of hetero-oligomers of ClpB95 with ClpB80 is thermodynamically favorable over the formation of ClpB95 homo-oligomers (see Fig. 5.5A, B) since the formation of an incomplete ring of the N-terminal domains removes the “penalty” and restores the intrinsic mobility of the N-terminal domain.

We have recently shown that conserved acidic residues at the surface of the N-terminal domain of ClpB are involved in binding to aggregated proteins (Barnett et al., 2005). We speculated that the mobility of the N-terminal domain might help ClpB seek and bind the exposed recognition sites on the surface of large protein aggregates. We have found now that even an incomplete ring of the N-terminal domains provides sufficient molecular contacts for interactions with an aggregated substrate and the binding efficiency of hetero-oligomers is similar to that of ClpB95 homo-oligomers. Thus, the hetero-oligomers do not possess excess aggregate binding capability (see Fig. 5.6 and 5.7). Also, since there is no difference between ClpB95 and ClpB80 within the AAA+ modules, the efficiency of substrate threading through the central channel of the ClpB, does not increase in hetero-oligomers, as shown by our ATPase assays (Fig. 5.10 B, C).

We propose that there is an intermediate step between substrate binding and threading that is rate limiting for the disaggregation process and this step is supported by the mobility of the N-terminal domain of ClpB. In conclusion, the enhancement of the chaperone activity of ClpB95/ClpB80 hetero-oligomers is linked to an enhancement of mobility of the N-terminal domains which could support efficient substrate extraction from an aggregate and initiation of substrate translocation through the ClpB channel. Interestingly, in addition to the increased mobility of the N-terminal domains in the hetero-oligomers of ClpB, the normal mode analysis also detected a small stimulation of the C-terminal D2 small sub-domain mobility (Fig. 5.11C). This could suggest a linkage between the movements of the two domains with a possible role in coordinating the entrance and the exit of substrates into and from the central channel of ClpB. Thus the

N-domain mobility could influence the activity of ClpB also in an indirect way. At this point, we have no evidence for such mechanism.

Based on the results obtained with a ClpB variant that contained a longer N-domain linker region than ClpB95, it can be concluded that the mobility of the N-domain is optimized for the chaperone activity. The linker region (Fig. 5.12 A, B) that connects the N-domain of ClpB to the first AAA+ domain allows the movements of the N-terminal domain. We have found that even slight variation in the length and flexibility of the linker leads to a total loss of aggregate binding (Fig. 5.14) and solubilizing ability of ClpB (Fig. 5.15). Also, the alteration of the linker length decreased the ability of α- casein to stimulate the ATPase activity of ClpB95, and not the basal or poly-lysine activated ATPase (Fig 5.13). Thus, it can be hypothesized that perhaps the linker insertions allow less controlled movements of the N-terminal domains, which may disable the substrate binding sites located in the N-terminal domain and at the entrance to the ClpB channel.

Interestingly, similar results were obtained in case of another bacterial Clp ATPase, ClpA that also contains a mobile N-terminal domain and an internal translation initiation site at the C-terminal end of the N domain. ClpA also contains a 25 amino acid-long linker region that connects the N-terminal domain to the rest of the protein and controls its mobility (Fig. 5.12B). It was found that the optimal function of ClpA in protein degradation requires an unaltered linker region, as shortening or elongation of the linker caused defects in ClpA activity (Cranz-Mileva et al., 2008). Recently, it has been found that the N-terminal domain of another Clp ATPase, ClpX is likewise highly mobile and that the motions of the N-terminal domain are linked to the feeding of substrates into the ClpX channel (Thibault et al., 2006).

In conclusion, this study explains why the alternative translation initiation site in the ClpB gene has not been eliminated: *E. coli* produces two isoforms of ClpB to optimize activity of this chaperone through formation of hetero-oligomers. We showed that ClpB80 plays an important role in the ClpB function in spite of its low intrinsic substrate-binding capability. Since oligomer formation is an essential element of the mechanism of AAA+ ATPases, our results raise a possibility that the functionality of other AAA+ ATPases may be regulated by production of multiple protein isoforms and formation of hetero-oligomers. Moreover, since substrate insertion into the channel in the oligomeric ring is emerging as an underlying paradigm of the mechanism of AAA+ ATPases (Zolkiewski, 2006), our results suggest a common function for the N-terminal domains that are attached to the AAA+ modules, which is linked to substrate recognition and its delivery to the channel entry. The high mobility of the N-terminal domains may be essential for achieving efficient rates of substrate delivery. In the case of ClpB, formation of the hetero-oligomers of its two isoforms optimizes the functionality of the disaggregating chaperone by increasing the mobility of the N-terminal domains.

Acknowledgements

I thank Dr. Vladimir Akoev and Dr. Micheal Barnett for their participation in experiments shown in Figs. 5.5 and 5.6.

References

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

Barnett,M.E., Zolkiewska,A., and Zolkiewski,M. (2000). Structure and activity of ClpB from *Escherichia coli*. Role of the amino- and -carboxyl-terminal domains. J. Biol. Chem. *275*, 37565-37571.

Beinker,P., Schlee,S., Groemping,Y., Seidel,R., and Reinstein,J. (2002). The N terminus of ClpB from *Thermus thermophilus* is not essential for the chaperone activity. J. Biol. Chem. *277*, 47160-47166.

Chow,I.T. and Baneyx,F. (2005). Coordinated synthesis of the two ClpB isoforms improves the ability of *Escherichia coli* to survive thermal stress. FEBS Lett. *579*, 4235-4241.

Chow,I.T., Barnett,M.E., Zolkiewski,M., and Baneyx,F. (2005). The N-terminal domain of *Escherichia coli* ClpB enhances chaperone function. FEBS Lett. *579*, 4242-4248.

Claudiani,P., Riano,E., Errico,A., Andolfi,G., and Rugarli,E.I. (2005). Spastin subcellular localization is regulated through usage of different translation start sites and active export from the nucleus. Exp. Cell Res. *309*, 358-369.

Cranz-Mileva,S., Imkamp,F., Kolygo,K., Maglica,Z., Kress,W., and Weber-Ban,E. (2008). The flexible attachment of the N-domains to the ClpA ring body allows their use on demand. J. Mol. Biol. *378*, 412-424.

Diamant,S., Ben Zvi,A.P., Bukau,B., and Goloubinoff,P. (2000). Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. J. Biol. Chem. *275*, 21107-21113.

Diemand,A.V. and Lupas,A.N. (2006). Modeling AAA+ ring complexes from monomeric structures. J. Struct. Biol. *156*, 230-243.

Haslberger,T., Weibezahn,J., Zahn,R., Lee,S., Tsai,F.T., Bukau,B., and Mogk,A. (2007). M domains couple the ClpB threading motor with the DnaK chaperone activity. Mol. Cell *25*, 247-260.

Hess,H.H. and Derr,J.E. (1975). Assay of inorganic and organic phosphorus in the 0.1-5 nanomole range. Anal. Biochem. *63*, 607-613.

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

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

Lanzetta,P.A., Alvarez,L.J., Reinach,P.S., and Candia,O.A. (1979). An improved assay for nanomole amounts of inorganic phosphate. Anal. Biochem. *100*, 95-97.

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

Mogk,A., Schlieker,C., Strub,C., Rist,W., Weibezahn,J., and Bukau,B. (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.

Nagy,M., Akoev,V., and Zolkiewski,M. (2006). Domain stability in the AAA+ ATPase ClpB from *Escherichia coli*. Arch. Biochem. Biophys. *453*, 63-69.

Park,S.K., Kim,K.I., Woo,K.M., Seol,J.H., Tanaka,K., Ichihara,A., Ha,D.B., and Chung,C.H. (1993). Site-directed mutagenesis of the dual translational initiation sites of the clpB gene of *Escherichia coli* and characterization of its gene products. J. Biol. Chem. *268*, 20170-20174.

Schlieker,C., Weibezahn,J., Patzelt,H., Tessarz,P., Strub,C., Zeth,K., Erbse,A., Schneider-Mergener,J., Chin,J.W., Schultz,P.G., Bukau,B., and Mogk,A. (2004). Substrate recognition by the AAA+ chaperone ClpB. Nat. Struct. Mol. Biol. *11*, 607-615.

Seol,J.H., Yoo,S.J., Kim,K.I., Kang,M.S., Ha,D.B., and Chung,C.H. (1994). The 65-kDa protein derived from the internal translational initiation site of the clpA gene inhibits the ATP-dependent protease Ti in Escherichia coli. J. Biol. Chem. *269*, 29468-29473.

Stafford,W.F., III (1992). Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. Anal. Biochem. *203*, 295-301.

Strub,C., Schlieker,C., Bukau,B., and Mogk,A. (2003). Poly-L-lysine enhances the protein disaggregation activity of ClpB. FEBS Lett. *553*, 125-130.

Suhre,K. and Sanejouand,Y.H. (2004). ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic Acids Res. *32*, W610-W614.

Thibault,G., Tsitrin,Y., Davidson,T., Gribun,A., and Houry,W.A. (2006). Large nucleotide-dependent movement of the N-terminal domain of the ClpX chaperone. EMBO J. *25*, 3367-3376.

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

Weibezahn,J., Schlieker,C., Tessarz,P., Mogk,A., and Bukau,B. (2005). Novel insights into the mechanism of chaperone-assisted protein disaggregation. Biol. Chem. *386*, 739-744.

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

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

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

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

Zmijewski,M.A., Macario,A.J., and Lipinska,B. (2004). Functional similarities and differences of an archaeal Hsp70(DnaK) stress protein compared with its homologue from the bacterium *Escherichia coli*. J. Mol. Biol. *336*, 539-549.

Zolkiewski,M. (2006). A camel passes through the eye of a needle: protein unfolding activity of Clp ATPases. Mol. Microbiol. *61*, 1094-1100.

Zylicz,M., Yamamoto,T., McKittrick,N., Sell,S., and Georgopoulos,C. (1985). Purification and properties of the dnaJ replication protein of *Escherichia coli*. J. Biol. Chem. *260*, 7591-7598.

Figures

Figure 5.1 Diagram of the two naturally-occurring isoforms of ClpB.
Besides the full-length 95 kDa protein (ClpB95) an N-terminally truncated isoform (ClpB80) is also produced in *E. coli* cells. The truncated isoform is missing the 148 residue-long N-terminal domain of the full-length ClpB. N-terminal domain (N) is colored in green, the two AAA+ modules, AAA-1 and AAA-2 (D1 and D2) are colored in purple and gray, respectively; the coiled-coil middle domain (M) is blue. The *in vivo* ratio of ClpB95/ClpB80 is ~ 2.

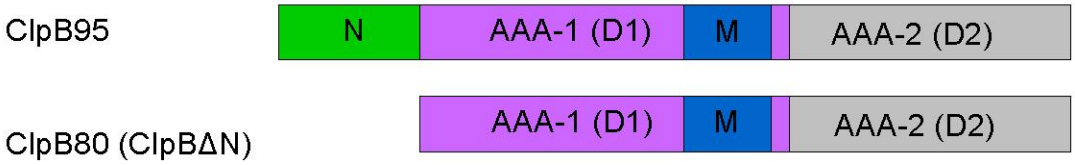


Figure 5.2 Reactivation of aggregated G6PDH in the presence of the ClpB/DnaK bichaperone system.
Intermediate-size (A) or large (B) aggregates of G6PDH were prepared by incubating 22 µM denatured G6PDH in buffer A for 1 min or 15 min, respectively, at 47 °C. After cooling on ice for 2 min, aggregates were diluted 8-fold into buffer A with 5 mM ATP containing either no chaperones (open circles), 5 µM DnaK, 1 µM DnaJ, 0.5µM GrpE (KJE, filled squares), KJE with 1.5 µM ClpB95 (full circles), or KJE with 1.5 µM ClpB80 (full triangles). After incubation at 30 °C for the indicated periods of time, aliquots were withdrawn and the G6PDH activity was measured as described in Materials and Methods.

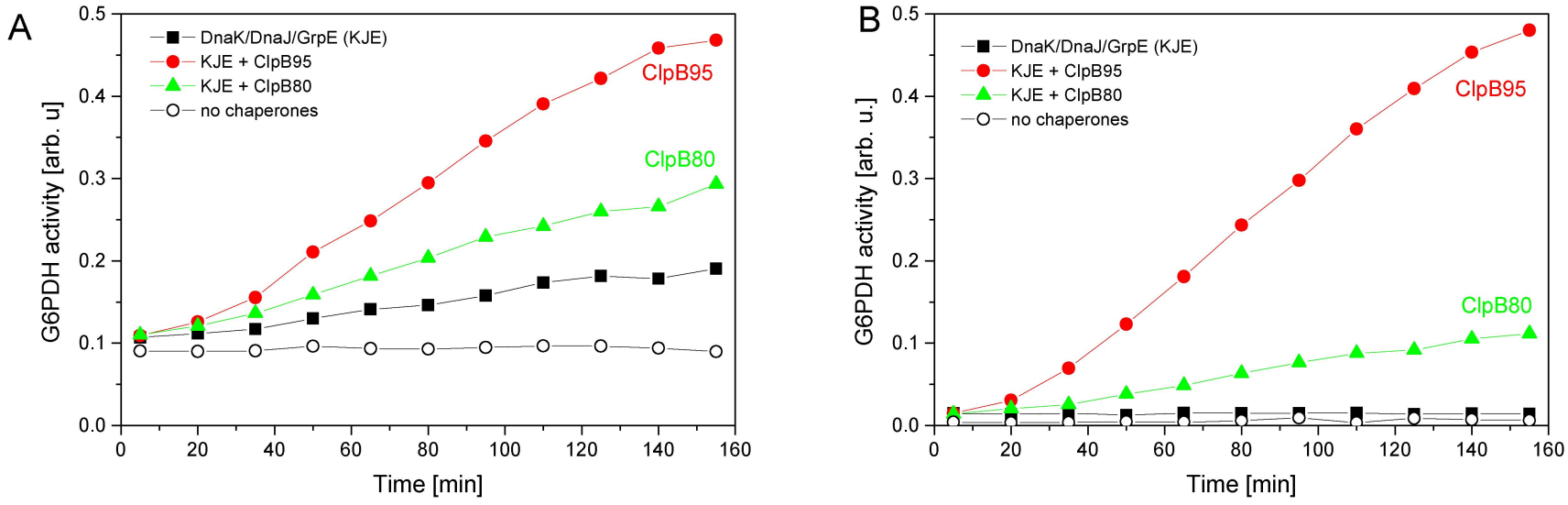


Figure 5.3 Reactivation of aggregated G6PDH by the combined action of ClpB95/ClpB 80 and the DnaK chaperone system.

Aggregated G6PDH was prepared as described in Methods. The aggregates were diluted into buffer A with 6 mM ATP containing no chaperones (black crosses), 1 μM DnaK, 1 μM DnaJ, 0.5 μM GrpE (KJE) with 1.5 μM ClpB95 (red trace), KJE with 1.5 μM ClpB80 (green trace), KJE with 1 μM ClpB95 and 0.5 μM ClpB80 (cyan trace), KJE with 0.75 μM ClpB95 and 0.75 μM ClpB80 (magenta), or KJE with 0.5 μM ClpB95 and 1 μM ClpB80 (blue trace). After incubation at 30 °C for the indicated periods of time, aliquots were withdrawn and the G6PDH activity was measured as described in Materials and Methods. Insets show the G6PDH activity after 80 min vs. the molar fraction of ClpB80, f_{ClpB80}=[ClpB80]/([ClpB95]+[ClpB80]).

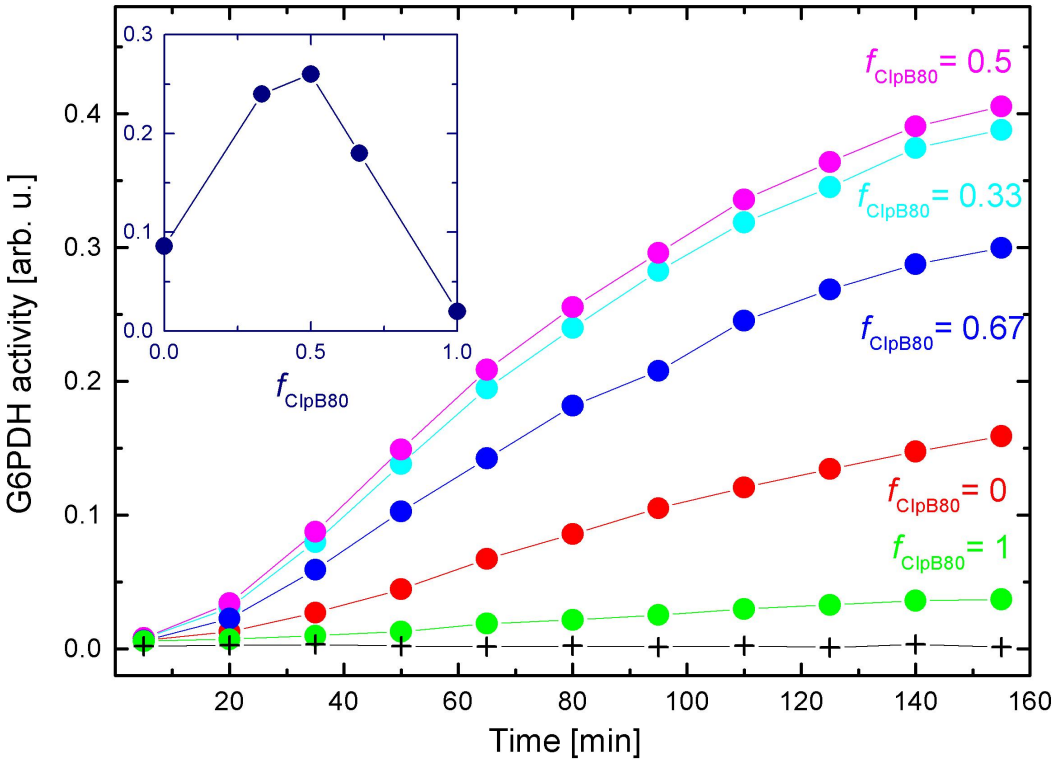


Figure 5.4 Reactivation of aggregated MDH by the combined action of ClpB95/ ClpB 80 and the DnaK chaperone system.

Aggregated MDH was prepared as described in Methods. The aggregates were diluted into buffer B with 6 mM ATP containing either no chaperones (black crosses), 1 μM DnaK, 1 μM DnaJ, 0.5 μM GrpE (KJE) with 1.5 μM ClpB95 (red trace), KJE with 1.5 μM ClpB80 (green trace), KJE with 1 μM ClpB95 and 0.5 μM ClpB80 (cyan trace), KJE with 0.75 μM ClpB95 and 0.75 μM ClpB80 (magenta), or KJE with 0.5 μM ClpB95 and 1 μM ClpB80 (blue trace). After incubation at 30 °C for the indicated periods of time, aliquots were withdrawn and the MDH activity was measured as described in Materials and Methods. Insets show the MDH activity after 120 min vs. the molar fraction of ClpB80 f_{ClpB80}=[ClpB80]/([ClpB95]+[ClpB80]).

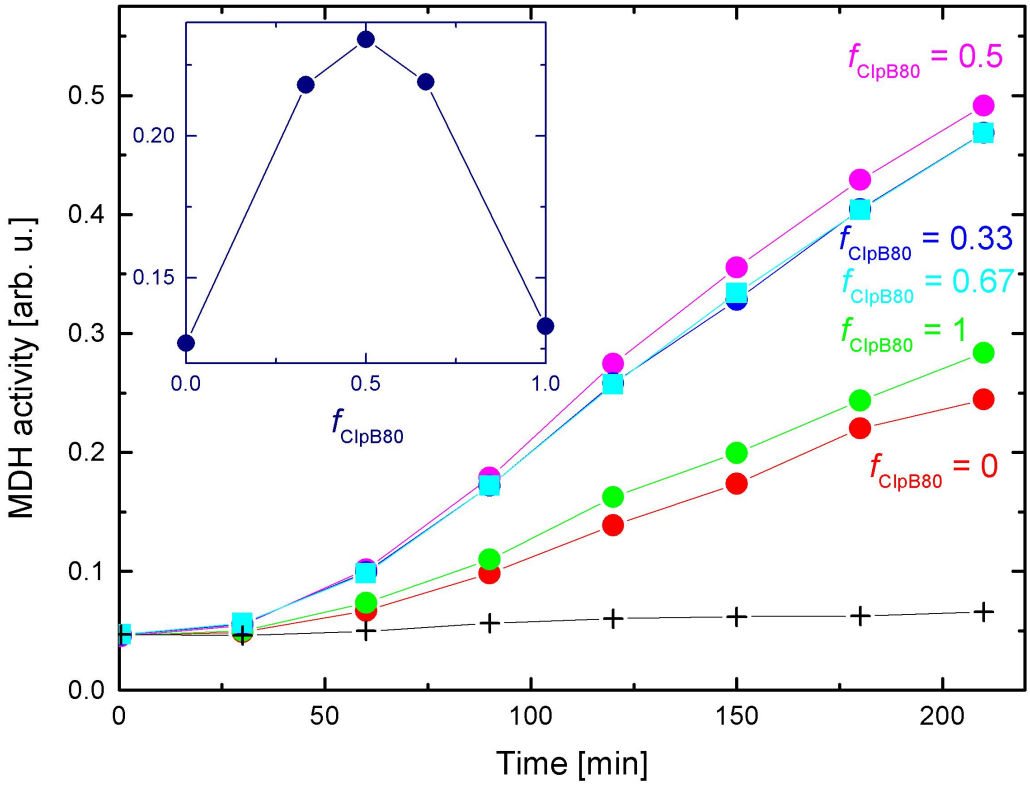


Figure 5.5 Sedimentation velocity analysis of ClpB95 and ClpB80.
Ultracentrifugation was performed at 50,000 rpm and 20 °C in 50 mM Tris-HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1 mM EDTA, and 2 mM β-mercaptoethanol. Apparent sedimentation coefficient distributions $g(s^*)$ vs. $s^*_{20,w}$ in Svedbergs (S) were calculated from protein concentration profiles measured with absorption at 238 nm and are shown for 5 μM ClpB95 (A, red line), 2.5 μM ClpB80 (A, green line), 5 μM ClpB95 with 2.5 μM ClpB80 (B), 8 μM ClpB95 (C, red line), and 8 μM ClpB80 (C, green line).

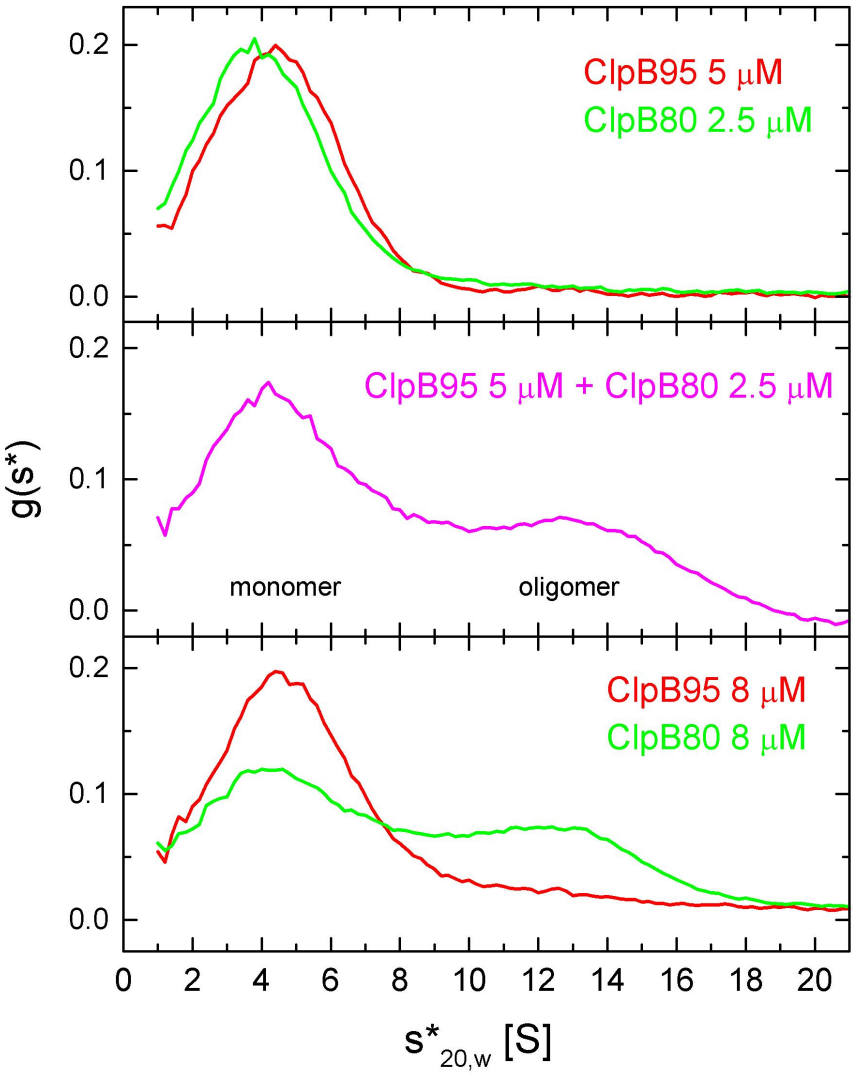
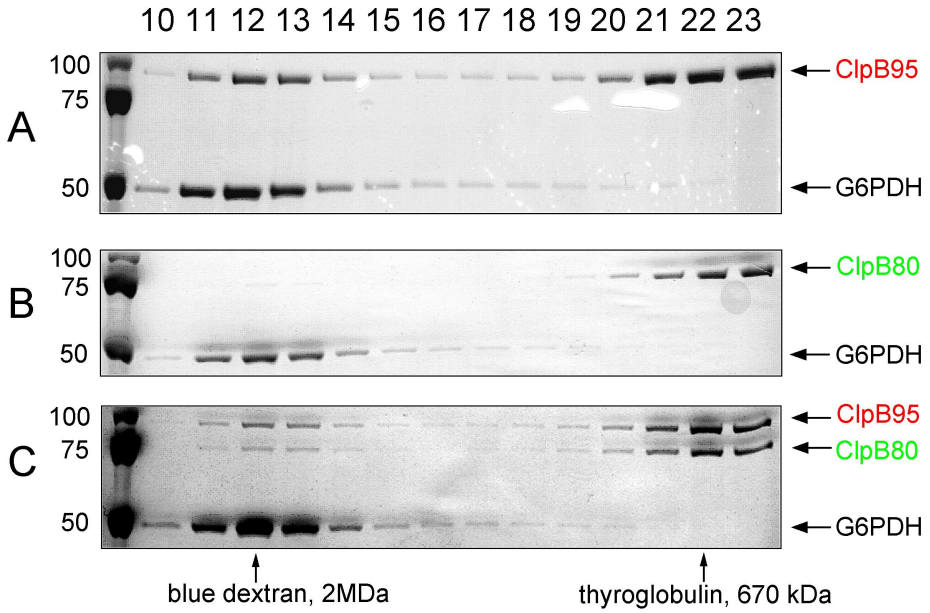


Figure 5.6 Interaction of ClpB95 and ClpB80 substrate-trapping Walker B mutants with strongly aggregated G6PDH detected by gel filtration chromatography.

Large aggregates of G6PDH were incubated with 15 μM ClpB95Trap (A), 15 μM ClpB80Trap (B), or 7.5 μM ClpB95Trap and 7.5 μM ClpB80Trap (C) for 5 min at 25 °C in buffer A with 2 mM ATP and injected onto the gel filtration column. 1-min elution fractions were analyzed by SDS-PAGE followed by Coomassie staining. Elution time (min) is indicated at the top of the figure. The arrows indicate the elution time of blue dextran and thyroglobulin; the position of the aggregates is marked with a red box. (D) The amounts of ClpB95 and ClpB80 in gel filtration fractions eluting at 11, 12, and 13 min (see panels A, B, C) were analyzed by band densitometry and corrected for the amount of G6PDH in each fraction. Shown are the amounts of ClpB95 (from panel A, red bars) and total amounts of ClpB95+ClpB80 (from panel C, green bars). The relative amounts of ClpB80 and ClpB95 bound to G6PDH in panel C are also shown (circles, right axis).


10 11 12 13 14 15 16 17 18 19 20 21 22 23
A
100
75
50
ClpB95
G6PDH
B
100
75
50
ClpB80
G6PDH
C
100
75
50
ClpB95
ClpB80
G6PDH
blue dextran, 2MDa
thyroglobulin, 670 kDa

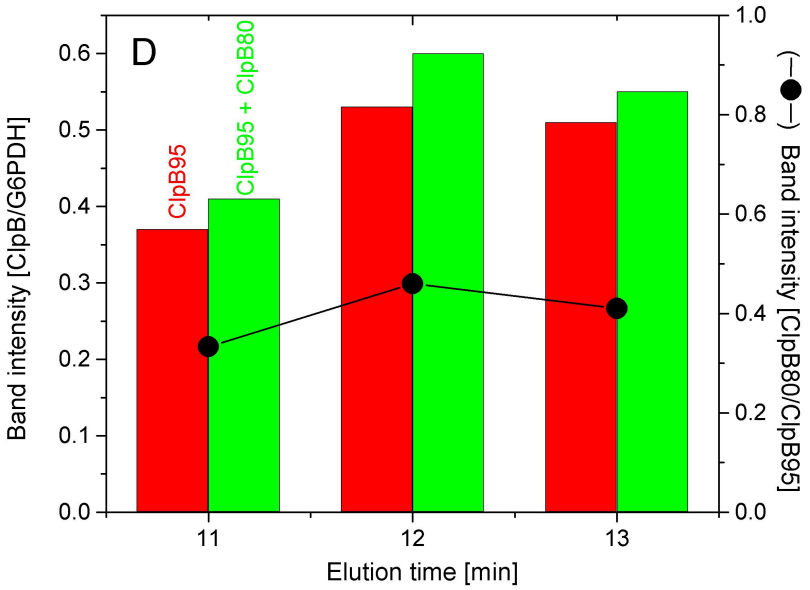
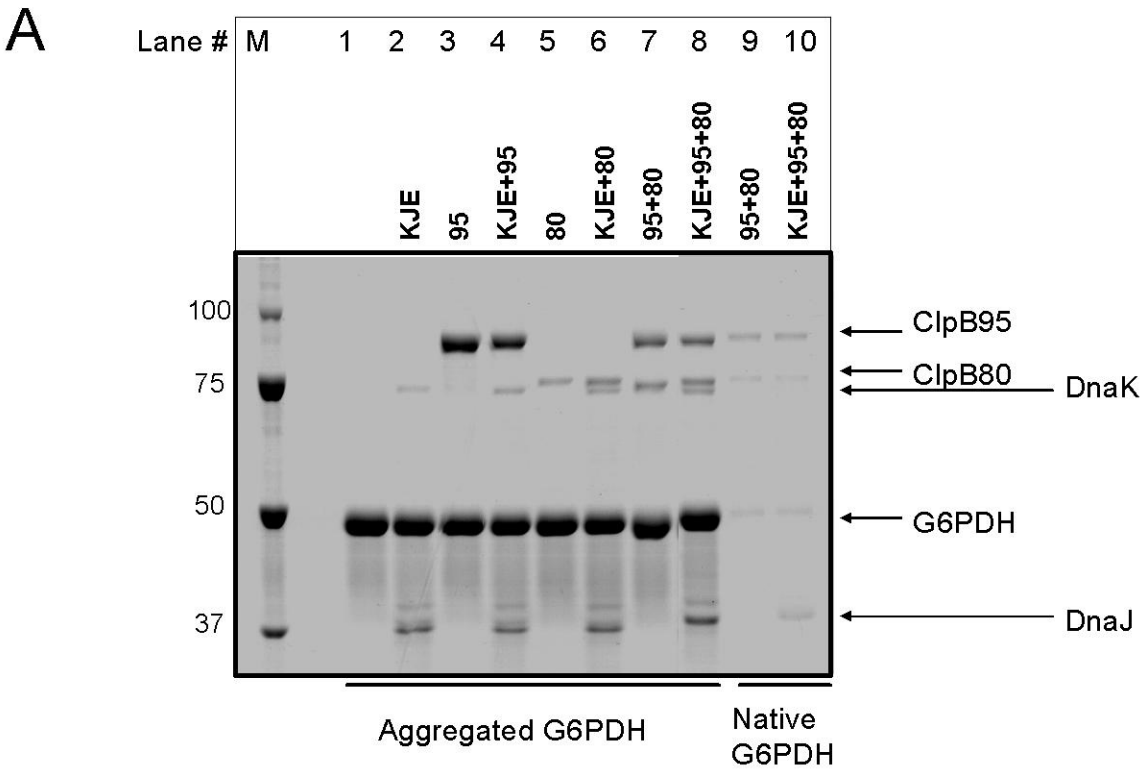

D
ClpB95
ClpB95 + ClpB80
Band intensity [ClpB/G6PDH]
(—●—) Band intensity [ClpB80/ClpB95]
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.0
0.2
0.4
0.6
0.8
1.0
11
12
13
Elution time [min]

Figure 5.7 Interaction of ClpB95 and ClpB80 substrate-trapping Walker B mutants with aggregated substrates as detected by filter assay.

2.7 μM aggregated G6PDH (A) or 3 μM MDH (B) were incubated in buffer A in presence of 5 mM ATP with 1.5 μM ClpB95Trap (95, lanes 3, 4), 1.5 μM ClpB80Trap (80, lanes 5,6) or 0.75 μM ClpB95Trap and 0.75 μM ClpB80Trap (95+80, lanes 7,8), without (lanes 3, 5, 7)or with 1μM DnaK, 1 μM DnaJ, and 0.5 μM GrpE (KJE, lanes 4,6,8) for 10 min at 30 °C with mixing at slow speed. Control samples containing 2.7 μM aggregated G6PDH (A) or 3 μM aggregated MDH (B) were also incubated under similar conditions in presence of ATP without (lane 1) and with the DnaK system (lane 2) in addition to the controls with 2.7 μM native G6PDH (A) or 3 μM native MDH (B) in presence of 0.75 μM ClpB95Trap and 0.75 μM ClpB80Trap without (lane 9) or with the DnaK system (lane 10). The solutions (35μl) were passed through a 0.1 μm filter and the fraction retained on the filter was eluted with SDS buffer and analyzed by SDS-PAGE, followed by Coomassie blue staining.


A
Lane #
M
1
2
3
4
5
6
7
8
9
10
KJE
95
KJE+95
80
KJE+80
95+80
KJE+95+80
95+80
KJE+95+80
100
75
50
37
ClpB95
ClpB80
DnaK
G6PDH
DnaJ
Aggregated G6PDH
Native G6PDH

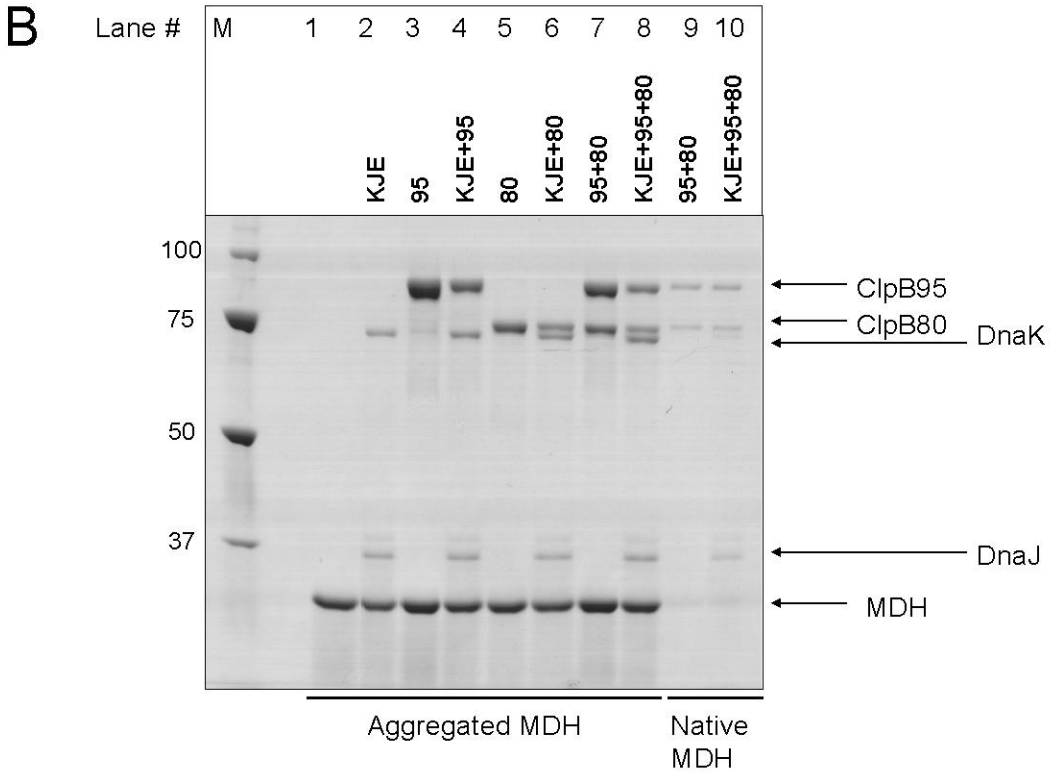

B
Lane #
M
1
2
3
4
5
6
7
8
9
10
KJE
95
KJE+95
80
KJE+80
95+80
KJE+95+80
95+80
KJE+95+80
100
75
50
37
ClpB95
ClpB80
DnaK
DnaJ
MDH
Aggregated MDH
Native MDH

Figure 5.8 Interaction of ClpB95 with aggregated G6PDH upon preincubation with the DnaK chaperone system.

G6PDH aggregates were incubated in buffer A in presence of 5 mM ATP with DnaK/DnaJ/GrpE (KJE, lane 13) or with ClpB95 (lane 14) for 10 min at 30 ˚C after which the missing chaperones (ClpBTrap or KJE, respectively) were added and incubated for an additional 10 min. Aggregated G6PDH was also incubated in presence of ATP (lanes 5, 6) with KJE (lanes 7, 8), ClpB95Trap (lanes 9, 10) or mixtures of KJE + ClpB95Trap (lanes 11, 12) for 10 min (lanes 5, 7, 9, 11) or 20 min (lanes 6, 8, 10, 12). In case of the control samples native G6PDH was similarly incubated in presence of ATP (lane1) with KJE (lane 2), ClpB95Trap (lane 3) or mixtures of KJE + ClpB95Trap (lane 4) for 20 minutes. The concentrations used were: 4 µM native or aggregated G6PDH, 1.5 µM ClpB95Trap, 1µM DnaK, 1 µM DnaJ, 0.5 µM GrpE. After incubation 35 µl was applied to the filter units and the aggregate containing fractions were separated as described in Materials and Methods. The results of the SDS-PAGE are shown.

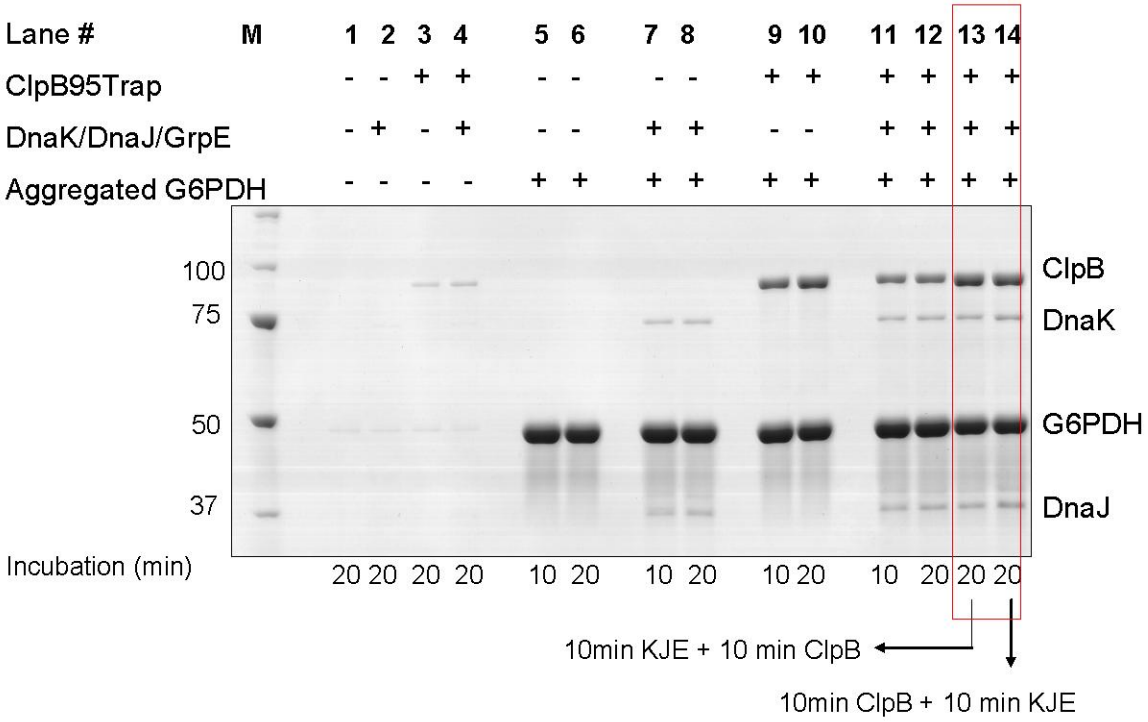


Figure 5.9 Reactivation of aggregated G6PDH upon preincubation with DnaK chaperone system.

The aggregates were diluted into buffer A with 6 mM ATP containing 1 μM DnaK, 1 μM DnaJ, 0.5 μM GrpE (KJE) (blue squares) and 1.5 μM ClpB95 (red circles) and after 10 min incubation at 30 °C the missing chaperones were added at time 0 of the reactivation assay. At the indicated periods of time, aliquots were withdrawn and the G6PDH activity was measured as described in Materials and Methods. Averages of two measurements are shown for the preincubated samples. Controls samples were also included that contained aggregated G6PDH without chaperones (magenta open triangles) or with 1 μM DnaK, 1 μM DnaJ, 0.5 μM GrpE (KJE) (black X-es).

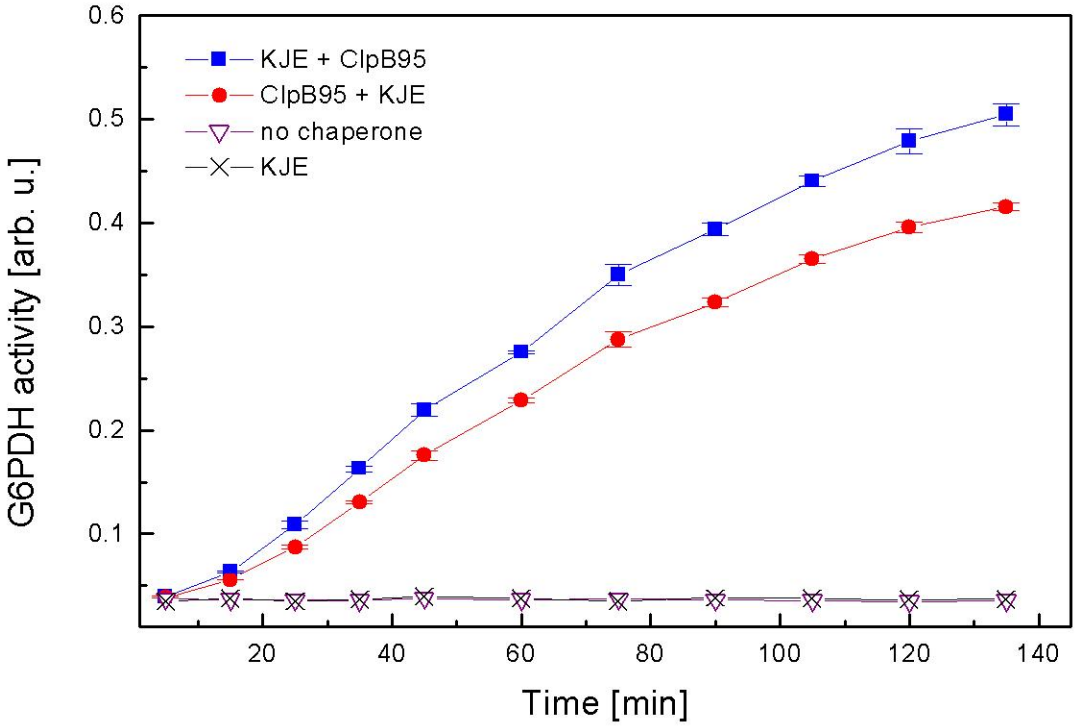
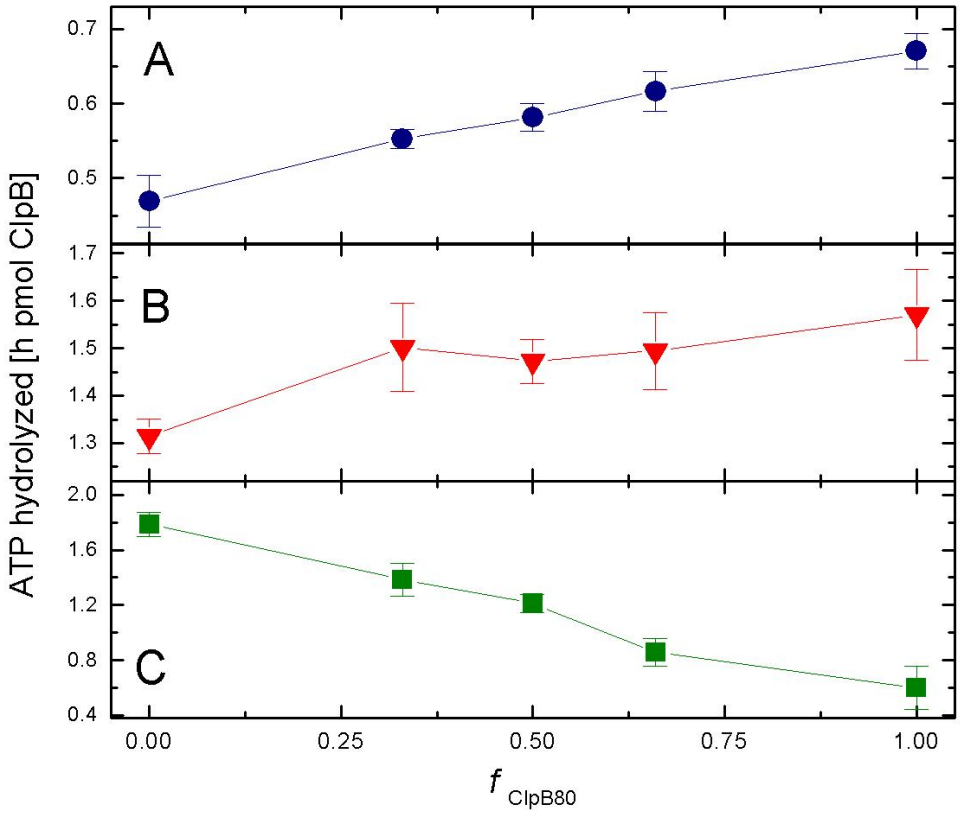


Figure 5.10 ATPase activity of ClpB95 and ClpB80.

The rate of ATP hydrolysis catalyzed by ClpB95 and ClpB80 at 37 °C was measured as the function of the molar fraction of ClpB80 (f_{ClpB80}) without activators (A), with 5 μM peptide B1 (B), or with 0.2 mg/ml α-casein (C). Inorganic phosphate production from ATP was determined in the presence of 30 pmol ClpB (A), 15 pmol ClpB (B), or 6 pmol ClpB (C) after 15 minute. (D) Inorganic phosphate production was also measured during the reactivation of aggregated G6PDH (2.7 μM) with 1 μM DnaK, 1 μM DnaJ, 0.5 μM GrpE (KJE) and 1.5 μM ClpB95 (red circles), KJE with 1.5 μM ClpB80 (green circles), KJE with 1 μM ClpB95 and 0.5 μM ClpB80 (cyan circles), KJE with 0.75 μM ClpB95 and 0.75 μM ClpB80 (magenta circles), or KJE with 0.5 μM ClpB95 and 1 μM ClpB80 (blue circles).


A
B
C
ATP hydrolyzed [h pmol ClpB]
f_{ClpB80}
0.00
0.25
0.50
0.75
1.00

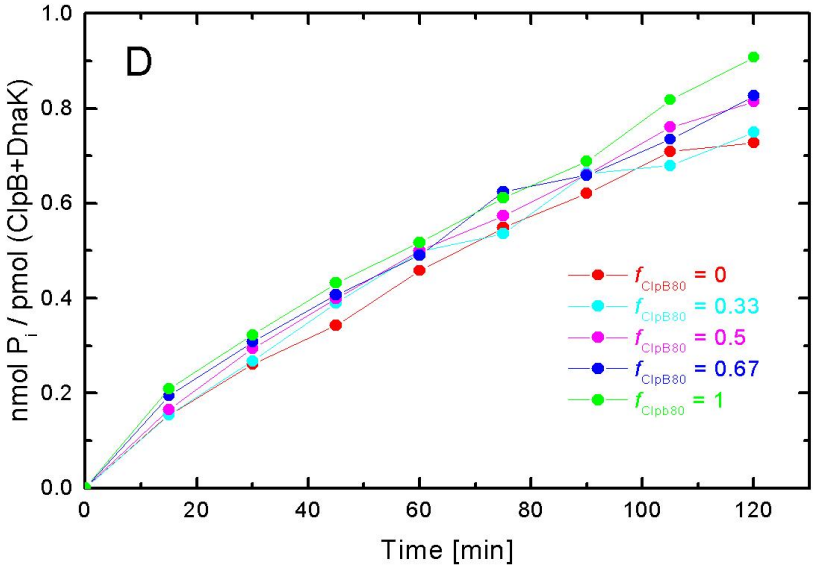
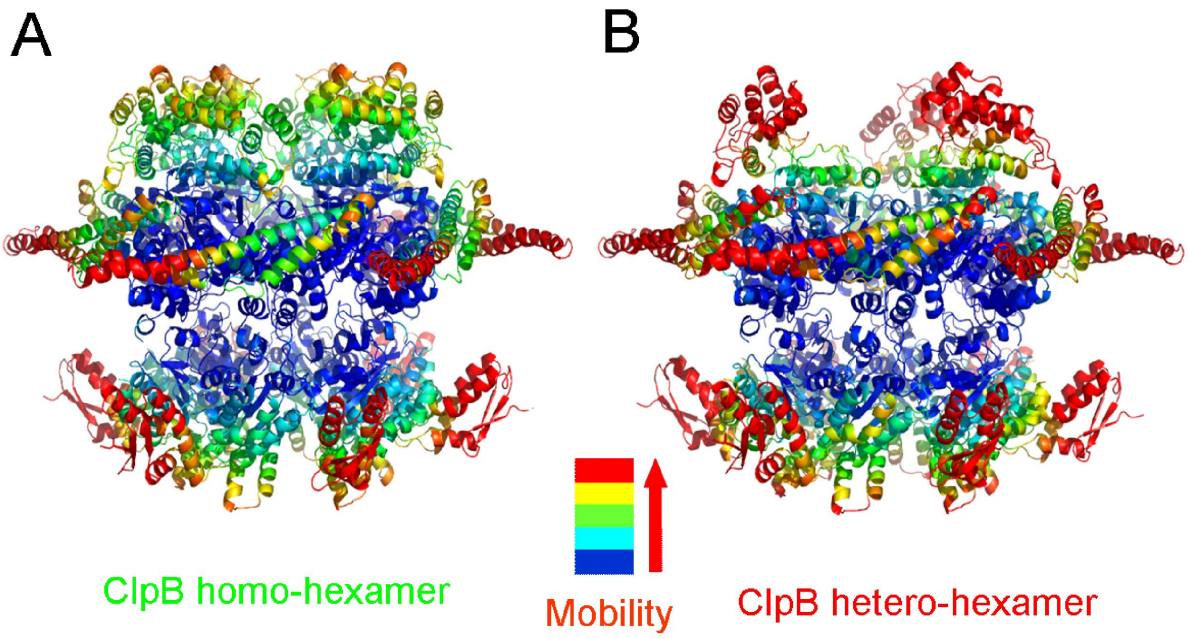

D
nmol P_i / pmol (ClpB+DnaK)
Time [min]
f_{ClpB80} = 0
f_{ClpB80} = 0.33
f_{ClpB80} = 0.5
f_{ClpB80} = 0.67
f_{Clpb80} = 1

Figure 5.11 Domain mobility in the homo-hexamer of ClpB95 and hetero-hexamer of ClpB95/ClpB80.

Elastic normal mode analysis was performed on the iMolTalk hexameric model of *T. thermophilus* ClpB (Diemand and Lupas, 2006) using the ElNemo server (Suhre and Sanejouand, 2004). The α-Carbon B-factors computed from the first 100 normal modes of a ClpB95 subunit located in a homo-hexamer or hetero-hexamer containing 3 ClpB95 and 3 ClpB80 subunits. The iMolTalk hexameric model of ClpB homo-hexamer (A) or hetero-hexamer (B) are colored by B-factors (values increase from dark blue to red). (C) Plot of the normalized α-Carbon B-factors of the homo-hexamer (green) and the hetero-hexamer (red).


A
B
ClpB homo-hexamer
Mobility
ClpB hetero-hexamer

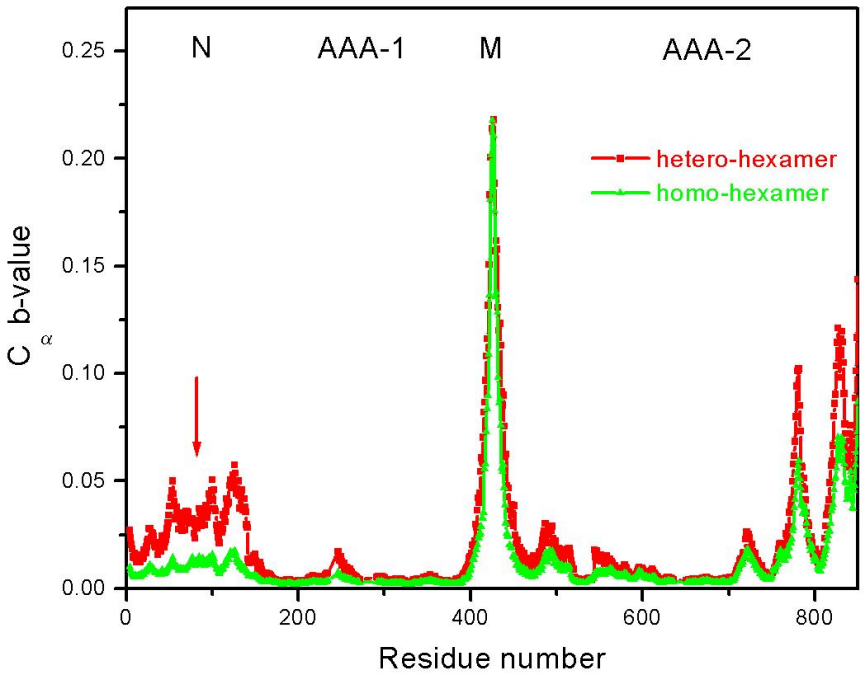
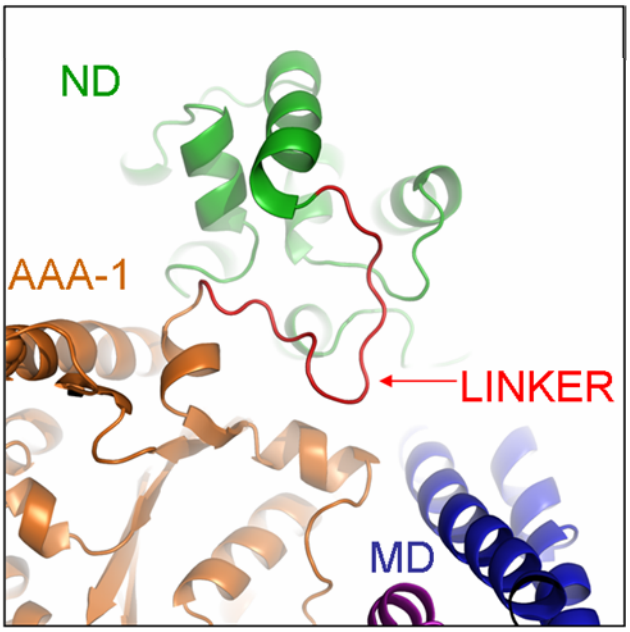

N
AAA-1
M
AAA-2
hetero-hexamer
homo-hexamer
C α b-value
0.25
0.20
0.15
0.10
0.05
0.00
0
200
400
600
800
Residue number

Figure 5.12 Localization and sequence of the N-terminal flexible linker.
Ribbon representation of the *E. coli* ClpB95 monomer model (A). The 14 residue-long linker that connects the N-terminal domain of ClpB to the first AAA+ module is colored in red. Parts of the N-terminal domain (green), the first AAA+ module (orange) and the tip of the middle domain (blue) are also visible. (B) Sequence of the *E. coli* ClpB95 N-domain linker and that of GI144ClpB compared to the sequence of the *E. coli* ClpA linker. The red arrow indicates the place where four Gly residues were inserted; the first residue of the AAA-1 module is colored in blue.

A



B

```
                          ↓
EC_ ClpB          143 MR----G--GESVNDQGAEDQR---------QALKK
EC_GI144ClpB      143 MRGGGGG--GESVNDQGAEDQR---------QALKK
EC_ClpA           140 HG----TRKDEPTQSSDPGSQPNSEEQAGGEERMEN
```

Figure 5.13 ATPase activity of GI144ClpB.

The rate of ATP hydrolysis was measured for ClpB95 and GI144ClpB glycine insertion mutant, in the absence of other polypeptides (red bars), in the presence of 0.1 mg/ml κ-casein (green bars), or with 0.04 mg/ml poly-L-lysine (yellow bars) after a 15 min incubation at 37 ˚C in presence of 2.5 µg ClpB in the absence of stimulators and 0.25 µg ClpB in presence of κ-casein or poly-L-lysine.

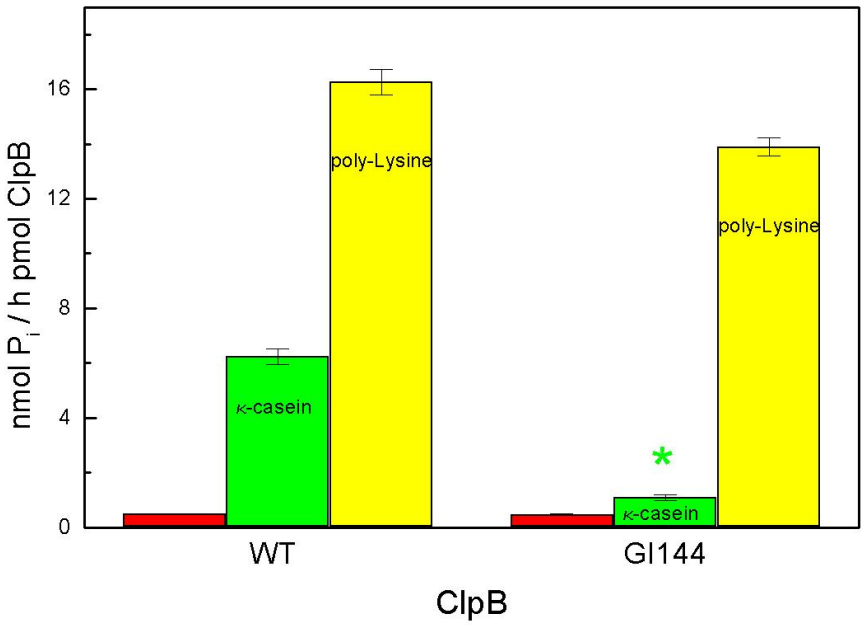


Figure 5.14 Interaction of GI144ClpB with aggregated G6PDH.

1.5 µM ClpB95 and GI144ClpB were incubated with native (lanes 1-4) or with aggregated G6PDH (3 µM) (lanes 5-8) without nucleotides (lanes 1-2, 5-6) or with 5 mM ATPγS (lanes 3-4, 7-8). The solutions (35 µl) were passed through a 0.1-µm filter.The result of the SDS-PAGE is shown for the fractions retained on the filter and subsequently solubilized with an SDS buffer.

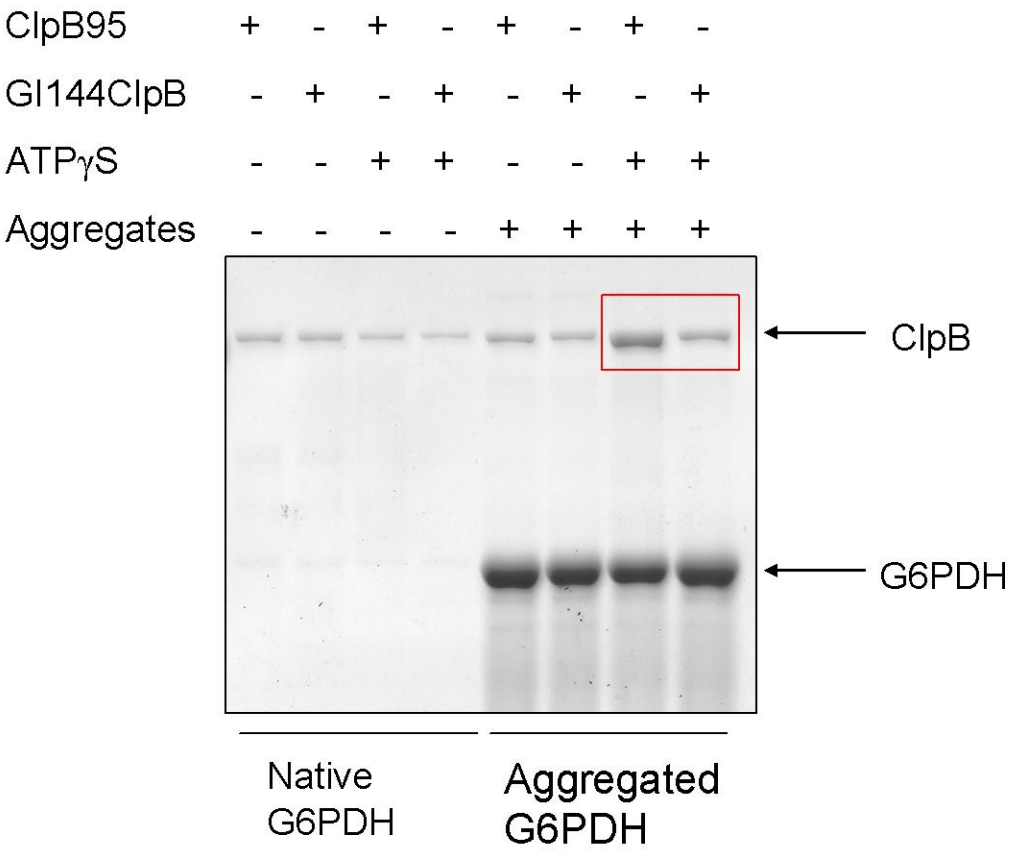
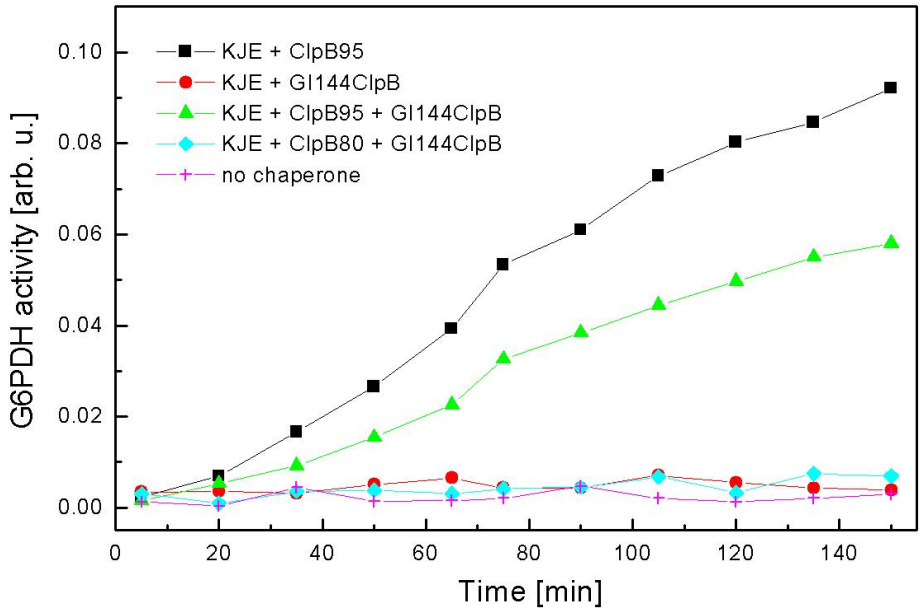


Figure 5.15 Reactivation of aggregated G6PDH in the presence of GI144ClpB and the DnaK chaperone system.

Aggregated 3 µM G6PDH was incubated in buffer A in presence of 5 mM ATP without chaperones (magenta cross) or with 1 µM DnaK, 1 µM DnaJ, 0.5 µM GrpE (KJE) and 1.5 µM ClpB95 (black squares), KJE with 1.5 µM GI144ClpB (red circles), KJE with 0.75 µM ClpB95 + 0. 75 µM GI144ClpB (green triangles), and KJE with 0.75 µM ClpB80 + 0. 75 µM GI144ClpB (cyan diamonds). After incubation at 30 °C for the indicated time, the G6PDH activity was measured as described in Methods.



CHAPTER 6 - Design and production of new fluorescent tools for mapping the ClpB-substrate interaction sites

Introduction

As the studies on the functionality of the ClpB95/ClpB80 hetero-oligomers exemplify, the events preceding the translocation of polypeptides through the central channel of the ClpB oligomer have a major, but so far not well described, role in the overall protein disaggregation mediated by ClpB. Based on our results obtained with the hetero-oligomers of ClpB, we postulated that the extraction of polypeptides from the aggregate and their insertion into the ClpB channel represent the rate-limiting intermediate step that occurs between the initial substrate recognition/binding and the translocation step. Currently, it is not known how ClpB is able to selectively recognize aggregated substrates over the correctly folded native proteins.

It has been proposed that molecular chaperones specifically recognize aggregation-prone hydrophobic sequences that are capped by positively charged gatekeeper residues (Chapter 1) (Rousseau et al., 2006). Indeed, peptide-binding studies revealed that ClpB preferentially binds to peptides that are enriched in positively charged and aromatic residues (Schlieker et al., 2004). Conserved residues, a tyrosine and two acidic residues, located on the flexible loops that line the entrance to the ClpB channel, were shown to be involved in substrate binding (Schlieker et al., 2004). We have shown that a pair of acidic residues located in the N-terminal domain of ClpB could

also support binding of aggregated proteins (Barnett et al., 2005). Moreover, we propose that the conserved residues located on the flexible loops that line the entrance to the ClpB channel are involved in the initiation of substrate translocation through the ClpB channel rather than in the initial event of aggregate recognition (as described in Chapter 1). Existing data suggests that the N-terminal domain (Chapter 5) and the coiled-coil middle domain of ClpB, as well as the DnaK system (Weibezahn et al., 2004; Haslberger et al., 2007), all play important roles in extraction of substrates from aggregates and insertion into the ClpB channel for translocation.

Thus, in order to dissect the overall mechanism of ClpB-mediated aggregate reactivation, it is crucial to characterize the initial interaction between ClpB and aggregated substrates and to investigate the mechanisms that control the substrate insertion into the ClpB channel. We intend to use fluorescently labeled ClpB and aggregated substrate to study both the initial interaction event and the subsequent extraction/insertion step of the disaggregation process by fluorescence resonance energy transfer (FRET). FRET is based on the transfer of the excited state energy from a donor (D) fluorophore, to an acceptor (A) molecule without the emission of a photon. This “radiationless” transfer of energy occurs only if the absorption spectrum of the acceptor molecule overlaps with the emission spectrum of the donor. The efficiency of FRET however, is influenced by the distance between the donor and the acceptor among other factors, and thus it is often used to detect protein-protein interactions.

This part of the dissertation describes ClpB variants designed and produced for labeling with florescent dyes. It also presents preliminary data, which confirms that

FRET can be successfully used to detect interaction between ClpB and model substrates.

Materials and Methods

Mutagenesis and protein purification

Cys-free ClpB (Cys-less ClpB) as well as the single-Cys variants of ClpB and their substrate trapping versions were produced by site-directed mutagenesis as described in Chapter 3. After the purification the single-Cys variants N104C ClpB, N104C ClpBTrap and V246C ClpBTrap were dialyzed against a buffer that contained TCEP (Tris(2-carboxyethyl)phosphine hydrochloride; Invitrogen) as a reducing agent instead of the commonly used DTT (50 mM Tris/HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1mM EDTA, 5 mM TCEP, 10% glycerol). This was necessary to allow for efficient labeling of cysteines with the thiol-reactive probe, because TCEP, unlike DTT and β-mercaptoethanol, does not contains thiols and does not have to be removed prior to labeling.

Model substrates and labeling reagents

FITC-casein (fluorescein isothiocyanate-casein) and κ-casein were obtained from Sigma-Aldrich (St. Louis, MO). Fluorescent probes used for labeling: IAEDANS (5-((((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid) and F5M (fluorescein-5-maleimide) were purchased from Invitrogen (formerly Molecular Probes).

Peptide B1 (2.54 kDa) was synthesized at the Kansas State University Biotech Core Facility. The peptide was solubilized in 5 % acetic acid and purified by reverse phase HPLC (RP-HPLC) on a 15 cm C12 column (Synergi Max-RP) equilibrated with 10% acetonitrile (ACN). The peptide was eluted at 0.8 ml/min when the organic gradient

reached ~23 % of ACN. The fractions containing peptide B1 (0.5 ml) from several runs were combined and lyophilized. The lyophilized sample containing the purified peptide B1 was dissolved in buffer containing 50 mM Tris/HCl pH 7.5 and 5 mM TCEP. The concentration was determined by measuring the absorbance at 280 nm and using the calculated extinction coefficient ($\varepsilon_{(280\ nm)}$ = 7115 M^{-1} cm^{-1}).

Protein and peptide labeling

N104C ClpBTrap, V246C ClpBTrap, and Cys-less ClpB were labeled separately with IAEDANS, while peptide B1 was labeled with F5M (fluorescein) by using previously published methods (Rye, 2001) and the protocol provided by Invitrogen. ClpB solutions were diluted to 100 µM with labeling buffer (50 mM Tris/HCl pH 7.5, 0.2 M KCl, and 5 mM TCEP) to a final volume of 2.5 ml. IAEDANS stock solution (50 mM) was prepared immediately prior to use by dissolving the dry powder in DMF (N, N-dimethylformamide). Approximately a 10-fold molar excess of the label was added to the protein solution (50 µl IAEDANS from the 50 mM stock solution) drop wise while stirring the mixture in the dark. The reaction was allowed to proceed for 4 hours in the dark at 25˚ C (with constant stirring), after which the reaction was quenched by addition of 10 mM β-mercaptoethanol. Following the labeling procedure the proteins were separated from the unreacted dye on a Sephadex G-25 gel filtration column (PD-10 desalting column; GE Healthcare) equilibrated with 50 mM Tris/HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1 mM EDTA, and 1 mM DTT. The proteins eluted from the gel filtration column were dialyzed extensively against a buffer containing 50 mM Tris/HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, and 10 % glycerol (dialysis buffer).

The concentration of the proteins was determined by measuring the absorbance at 280 nm.

For the labeling of the purified peptide B1, 40 mM stock solution of fluorescein-5-maleimide was prepared freshly by dissolving the dry powder in DMF. The labeling procedure was similar to the procedure used to label ClpB with IAEDANS except peptide B1 was diluted to 90 µM with the labeling buffer without KCl (50 mM Tris/HCl pH 7.5, 5 mM TCEP) and a 20-fold excess of F5M over the peptide was used during labeling (for 2 mM F5M in the 800 µl final reaction volume). Instead of using gel filtration for the removal of the unreacted dye, the reaction mixture was extensively dialyzed against the dialysis buffer without glycerol and the labeled peptide was further purified on RP-HPLC, lyophilized and dissolved in buffer containing 50 mM Tris/HCl pH 7.5, 20 mM $MgCl_2$, 1 mM EDTA, and 1 mM DTT.

ATP-dependent self association

Proteins (wt ClpB, Cys-less ClpB, N104C ClpB, N104C ClpBTrap and V246C ClpBTrap, as well as the IAEDANS labeled versions of Cys-less ClpB, N104C ClpBTrap and V246C ClpBTrap were diluted to 2 mg/ml in running buffer (50 mM Tris/HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT). For each run 20 µl was injected onto a 25 cm (2.4 ml) Superose 6 PC 3.2/30 column (GE Healthcare, formerly Amersham Biosciences) equilibrated with running buffer (monomeric condition) or with running buffer supplemented with 2 mM ATP (oligomeric condition). The runs were performed at 0.05 ml/min at room temperature with a Shimadzu HPLC LC10ATvp equipped with M10Avp diode-array detector. The absorption was recorded at 290 nm.

Thyroglobulin (670 kDa), apoferritin (443 kDa), alcohol dehydrogenase (150 kDa), BSA (66 kDa) and cytochrome C (12.4 kDa) were used as protein standards (from Sigma).

ATPase activity

The rate of ATP hydrolysis was measured after 15 min incubation at 37˚C in 100 mM Tris/HCl pH 8, 10 mM $MgCl_2$, 5 mM ATP, 1 mM EDTA, 1 mM DTT by using the malachite green method (Hess and Derr, 1975; Lanzetta et al., 1979). Assay solutions contained 2.5 µg ClpB in case of the basal ATPase measurements and 0.25 µg ClpB in case of the casein activated ATPase measurements (0.1 mg/ml κ-casein). Peptide B1 stimulated ATPase activity was measured for 2.5 µg wt ClpB in absence or in presence of 10 µM purified peptide B1.

Fluorescence spectroscopy

Fluorescence spectra were measured with a Cary Eclipse spectrofluorimeter (Varian) using a 1-cm quartz fluorescence cuvette. The excitation wavelength was set at 336 nm and the emission spectrum was recorded from 350 nm to 650 nm. Both emission and excitation slits were set to 5 nm and scanning was performed at low speed. All experiments were conducted at room temperature in buffer containing 50 mM Tris/HCl pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT. Protein solutions were mixed with labeled or non-labeled model substrates without nucleotides or with 3 mM ATP; the spectrum was recorded after 5 minute incubation at room temperature.

Results and Discussion

Design and characterization of the Cys variants of ClpB

As described above the overall goal of the FRET studies is to characterize the initial interaction event between ClpB and aggregated proteins by following the interaction-induced responses of small fluorescent probes positioned at selected locations within the putative substrate binding sites in ClpB and at selected sites of an aggregated substrate. As a first step towards this goal we designed mutated versions of ClpB that allow for the coupling of exogenous fluorescent dyes to specific sites in ClpB. The thiol side-chain of the cysteines are commonly used as a site for the labeling with thiol-reactive fluorescent dyes. *E. coli* ClpB contains three Cys residues: Cys94 located in the N-terminal domain of ClpB, Cys310 located in the large sub-domain of the first AAA+ module (D1 large) and Cys615 within the large sub-domain of the second AAA+ domain (D2 large) (see Fig. 6.1). None of these endogenous Cys residues are located in the proximity of the putative substrate binding sites of ClpB and thus are not useful in FRET based experiments.

In order to specifically label selected sites in ClpB we produced a Cys-free version of ClpB (Cys-less ClpB) by mutating all three endogenous Cys residues to Ala. Based on the available crystal structures we chose two residues, located near the substrate binding sites of ClpB, for the subsequent introduction of single Cys residues. The first site selected for replacement with Cys, Asn104, is located on the surface of the N-terminal domain in the vicinity of the three residues: Thr7, Asp103 and Glu109, which were shown to support chaperone activity of ClpB (Liu et al., 2002) and thus were proposed to define a substrate binding site for aggregates (Barnett et al., 2005) (Fig. 6.1

and Fig. 6.2). The second site chosen for the replacement with Cys residue; Val246, is located in the first AAA+ module, in proximity of the flexible loop that defines the entrance to the ClpB channel (Fig.6.1). As mentioned previously, the conserved Tyr251 located on the flexible channel loop, as well as two neighboring acidic residues (Glu254 and Glu257) were also proposed to contribute to the substrate binding site of ClpB (Schlieker et al., 2004). As Fig. 6.3 illustrates, the Val246 selected for replacement with Cys and for the subsequent labeling with the fluorescent dye is positioned below the flexible central loop of the D1 large sub-domain and this way has a lower chance to disturb substrate binding. Aside from the N104C or the V246C mutations introduced in the Cys-less ClpB (C94A/C310A/C615A) we further mutated the conserved Walker B residues in both AAA+ modules (E279Q/E678Q) to produce the trap versions of ClpB (N104C ClpBTrap and V246C ClpBTrap) that form stable interaction with substrates in presence of ATP (as described in Chapter 4).

Consequently, we purified a total of four ClpB variants: Cys-less ClpB, N104C ClpB, N104C ClpBTrap (C104ClpB*) and V246C ClpBTrap (C246ClpB*) and followed the nucleotide-induced self association by gel filtration chromatography. We have found that despite of a large number of mutations introduced in wt ClpB sequence (a total of six mutations for the Trap versions of the single-Cys variants) the corresponding proteins were able to form wild type ClpB-like oligomers in presence of ATP (Fig. 6.4). Since the trap versions of ClpB are unable to hydrolyze ATP (Weibezahn et al., 2003) we only measured the ATPase activity of N104C ClpB and found that this ClpB variant retained both the basal and the casein stimulated ATPase activity of the wt ClpB after the quadruple cysteine mutation (Fig. 6.5).

Labeling of ClpB and its pseudo-substrate with fluorescent dyes

For the FRET experiments we decided to use the IAEDANS (5-((((2-iodoacetyl) amino)ethyl)amino)naphthalene-1-sulfonic acid) as energy donor and fluorescein as energy acceptor. The IAEDANS/fluorescein is a commonly used FRET pair since the emission spectrum of IAEDANS (emission maximum at 490 nm) overlaps with the absorption spectrum of fluorescein. The Forster radius (the distance at which FRET is 50% efficient) for IAEDANS/fluorescein pair is approximately 46 Å.

The trap versions of the single-Cys ClpB variants C104ClpB* and C246ClpB* were labeled with IAEDANS as described in Materials and Methods. In order to establish that the fluorescent probes were conjugated to the intended sites and not to nonspecific sites, we also attempted to label the Cys-less ClpB with IAEDANS as a negative control. As Fig. 6.6 illustrates, both C104ClpB* and C246ClpB* were successfully labeled with IAEDANS as they show the characteristic fluorescence emission spectra with a maximum intensity at ~490 nm when excited at 336 nm. But most importantly, only a very low fluorescence corresponding to IAEDANS was observed in case of the Cys-less ClpB after labeling. This result confirms that IAEDANS label was selectively targeted to the single engineered Cys residues located either in the N-terminal domain or at the entrance to the ClpB channel. By comparing the protein absorbance at 280nm to the IAEDANS absorbance at 336 nm we estimated that the labeling yield for C104ClpB* was ~30 %, as for C246ClpB* was ~50 %. Gel filtration confirmed that the labeled proteins were not affected in their ability to form oligomers in presence of ATP (Fig. 6.7).

We have also used a 21 amino acid-long positively charged “peptide B1” as pseudo-substrate for ClpB. Peptide B1 (AHAWQHQGKTLFISRKTYRIC) was originally

identified by Bukau and coworkers through peptide scans and it was shown to bind to the central pore of the hexameric ClpB (Schlieker et al., 2004). Peptide B1 was synthesized and the crude product was purified on a reversed-phase column. As expected, the purified peptide B1 was able to stimulate the ATPase activity of ClpB (Fig. 6.8). The single Cys residue located at the C-terminal end of peptide B1 was labeled with fluorescein-5-maleimide by using published procedures (Rye, 2001) as described in Materials and Methods. The labeled peptide was further purified by reversed-phase HPLC and used in FRET experiments.

Detection of ClpB-model substrate interaction by FRET

The ATP-dependent interaction between donor labeled ClpB and acceptor labeled model substrate can be detected by FRET as a decrease in the emission intensity of the donor (IAEDANS) and a concomitant increase in the emission intensity of the acceptor (fluorescein). As mentioned above the IAEDANS emission band has maximum intensity at 490 nm when excited at 336 nm, while the maximum intensity of fluorescein occurs at ~515 nm. As shown in Fig. 6.9 for the individual and combined spectras of the IAEDANS-labeled ClpB variants (Fig 6.9A and Fig. 6.9B for IAEDANS-C104ClpB* and IAEDANS-C246ClpB*, respectively) and the fluorescein-labeled peptide B1, besides the overlap between the emission spectrum of IAEDANS (400 nm-650 nm) and the absorption spectrum of fluorescein, the emission spectrum of IAEDANS overlaps significantly also with the emission spectrum of fluorescein (480 nm-650 nm). Due to this overlap, the FRET-induced decrease in the IAEDANS emission may hinder the observation of the increase in the emission intensity of fluorescein. However, fluorescein shows less interference with the emission spectrum of IAEDANS around the

wavelength corresponding to the maximum IAEDANS emission intensity (450 nm-480 nm). Based on these observations we focus on the decrease in the donor intensity as an indication of energy transfer between ClpB and the bound substrate.

As the combined fluorescence emission spectrum of IAEDANS-labeled C104ClpB* and fluorescein-labeled peptide B1 shown in Fig. 6.10 (left panel) illustrates, we observed a decrease in the intensity of the IAEDANS/donor peak ~475 nm and an increase in the intensity of the fluorescein/acceptor peak at 515 nm upon addition of ATP. This is an indication that energy transfer occurred between the IAEDANS label conjugated to Cys104 and the fluorescein conjugated to the C-terminal end of the model peptide as a result of their ATP-dependent interaction. Importantly, this decrease in the IAEDANS fluorescence was not detected when the non-labeled version of the peptide B1was used in presence of ATP (Fig. 6.10, right panel). FRET was also detected between the IAEDANS label conjugated to Cys246 and fluorescein-peptide B1 as shown in Fig. 6.11 (left panel), although the FRET effect was observed only as a decrease of the donor peak around 475 nm. The ratio of ClpB/peptide was the same as in case of the N-terminally conjugated label, but due to the higher labeling efficiency of C246ClpB* (i.e. higher fluorescence) and the above described effect of the decrease in the IAEDANS fluorescence on the fluorescence peak of the acceptor, the increase in the acceptor fluorescence intensity is not observed. Similarly to C104ClpB*, the addition of ATP to the C246ClpB* and non-labeled peptide B1 did not cause a decrease in the IAEDANS intensity (Fig. 6.11, right panel), indicating that the effect observed with the labeled ClpB and substrate is a consequence of the energy transfer that occurs as a result of their interaction.

Finally, we have also followed binding of fluorescein isothiocyanate-casein (FITC-casein) to the two single-Cys variants of ClpB labeled with IAEDANS. The ATP-induced interaction with IAEDANS-labeled C104ClpB* (Fig. 6.12, left panel) or C246ClpB* (Fig. 6.13, left panel), and the corresponding FRET signal was also observed in case of this unstructured model substrate. In order to detect the increase in the acceptor fluorescence with the IAEDANS label conjugated to Cys246, we used higher FITC-casein/ClpB ratio (Fig. 6.13) than for the IAEDANS conjugated to Cys104 (Fig. 6.12). Again, there was no decrease in the donor fluorescence when the non-labeled casein was incubated with the IAEDANS-labeled C104ClpB* (Fig. 6.12, right panel) or IAEDANS-labeled C246ClpB* (Fig. 6.13, right panel) in presence of ATP.

In conclusion, we have shown that we were able to engineer single-Cys variants of ClpB that retained wild type ClpB-like oligomerization properties (Fig. 6.7) and remained active as ATPases (as seen for the non-trap version of N104C ClpB, Fig. 6.5). We have also successfully labeled the trap version of ClpB with IAEDANS at two separate locations in the vicinity of putative substrate binding sites of the N-terminal domain and of the central pore region of the hexameric ClpB (C104ClpB* and C246ClpB*, respectively). Moreover, we have demonstrated that FRET can be observed between ClpB and model pseudo-substrates in the presence of ATP (Figs. 6.10 through 6.13). Importantly, the relative efficiency of FRET observed throughout our experiments is comparable to the FRET efficiency detected in case of the GroEL-GroES interaction (Rye, 2001). The preliminary experiments presented in this chapter show that it is feasible to use FRET as a proximity sensor to detect ClpB-substrate interaction.

It will be worth experimenting with other FRET donor-acceptor pairs that have a better separation between the emission spectrum of the donor and the acceptor. For example the existing single-Cys variants of ClpB as well as substrate proteins could be labeled with the highly superior thiol-reactive Alexa Fluor dyes like Alexa Fluor 350 (donor) and Alexa Fluor 488 (acceptor). This FRET pair has also a Forster radius ~50 Å, but there is only a minimal overlap between the fluorescence emission spectra of the two dyes. Alexa Fluor dyes can be also used in conjunction with the nonfluorescent acceptors such as dabcyl and the QSY dyes (http://probes.invitrogen.com/handbook/).

Although so far we used only soluble pseudo-substrates in the FRET experiments, future studies will be focused on labeling the G6PDH aggregates (or aggregating labeled G6PDH) at selected locations with FRET acceptor dyes in an attempt to identify the sites on aggregates that are involved in the initial binding to ClpB, before the substrate insertion and the subsequent translocation step occurs. Additionally, the substrate trapping ClpB variant with IAEDANS donor conjugated to the single Cys residue located at the entrance to the ClpB channel (IAEDANS-C246ClpB*) could be used to investigate the mechanism that controls the access of fluorescein labeled aggregated substrates to the ClpB pore site by following the kinetics of FRET development with stopped flow-fluorescence experiments.

References

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

Haslberger,T., Weibezahn,J., Zahn,R., Lee,S., Tsai,F.T., Bukau,B., and Mogk,A. (2007). M domains couple the ClpB threading motor with the DnaK chaperone activity. Mol. Cell *25*, 247-260.

Hess,H.H. and Derr,J.E. (1975). Assay of inorganic and organic phosphorus in the 0.1-5 nanomole range. Anal. Biochem. *63*, 607-613.

Lanzetta,P.A., Alvarez,L.J., Reinach,P.S., and Candia,O.A. (1979). An improved assay for nanomole amounts of inorganic phosphate. Anal. Biochem. *100*, 95-97.

Liu,Z., Tek,V., Akoev,V., and Zolkiewski,M. (2002). Conserved amino acid residues within the amino-terminal domain of ClpB are essential for the chaperone activity. J. Mol. Biol. *321*, 111-120.

Rousseau,F., Serrano,L., and Schymkowitz,J.W. (2006). How evolutionary pressure against protein aggregation shaped chaperone specificity. J. Mol. Biol. *355*, 1037-1047.

Rye,H.S. (2001). Application of fluorescence resonance energy transfer to the GroEL-GroES chaperonin reaction. Methods *24*, 278-288.

Schlieker,C., Weibezahn,J., Patzelt,H., Tessarz,P., Strub,C., Zeth,K., Erbse,A., Schneider-Mergener,J., Chin,J.W., Schultz,P.G., Bukau,B., and Mogk,A. (2004). Substrate recognition by the AAA+ chaperone ClpB. Nat. Struct. Mol. Biol. *11*, 607-615.

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

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

Figures

Figure 6.1 Positions of amino acid mutations produced for this study.
The diagram shows the domain organization of *E. coli* ClpB. Three endogenous Cys residues located in the N-terminal domain (N, green), in the large sub-domain of the first AAA+ module (D1, orange), and in the large sub-domain of the second AAA+ module (D2, orange), were mutated to Ala in order to produce a Cys-less ClpB (C94A/C310A/C615A). Double Walker B mutations were further introduced in the Cys-less ClpB (E279Q/E678Q). Single-Cys variants used for fluorescent labeling contained Cys at position Asn104 in the N-terminal domain of ClpB or at place of Val246 in the D1 large sub-domain. The middle domain is colored in blue (M) while the small sub-domains of D1 and D2 are colored in magenta.

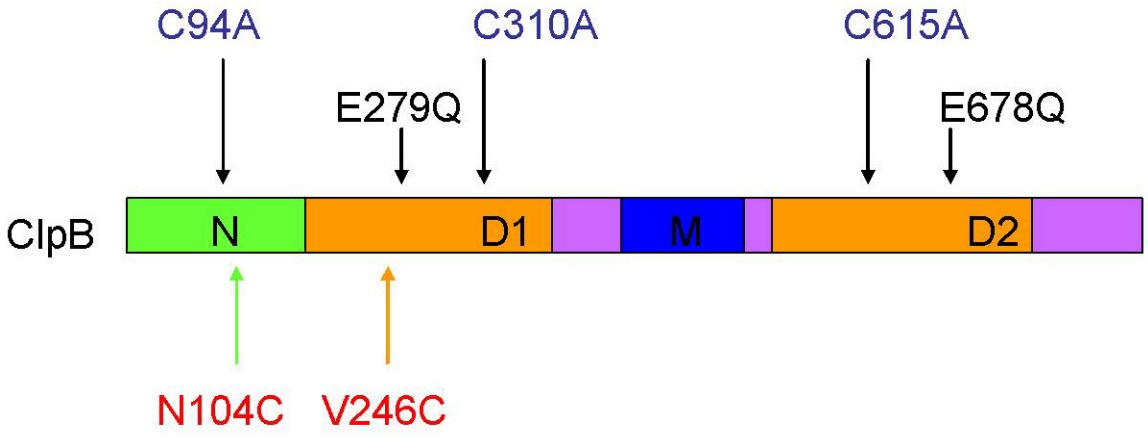


Figure 6.2 Location of the labeling site in the N-terminal domain.

The ribbon structure of the isolated N-terminal domain of *E. coli* ClpB (1khy) (Li and Sha, 2003) with indicated locations of the putative substrate binding site defined by Thr7 (magenta), Asp103 (red) and Glu109 (yellow) (Liu et al., 2002). Asn104 located in proximity of the substrate binding residues and selected for replacement with Cys is shown in green.

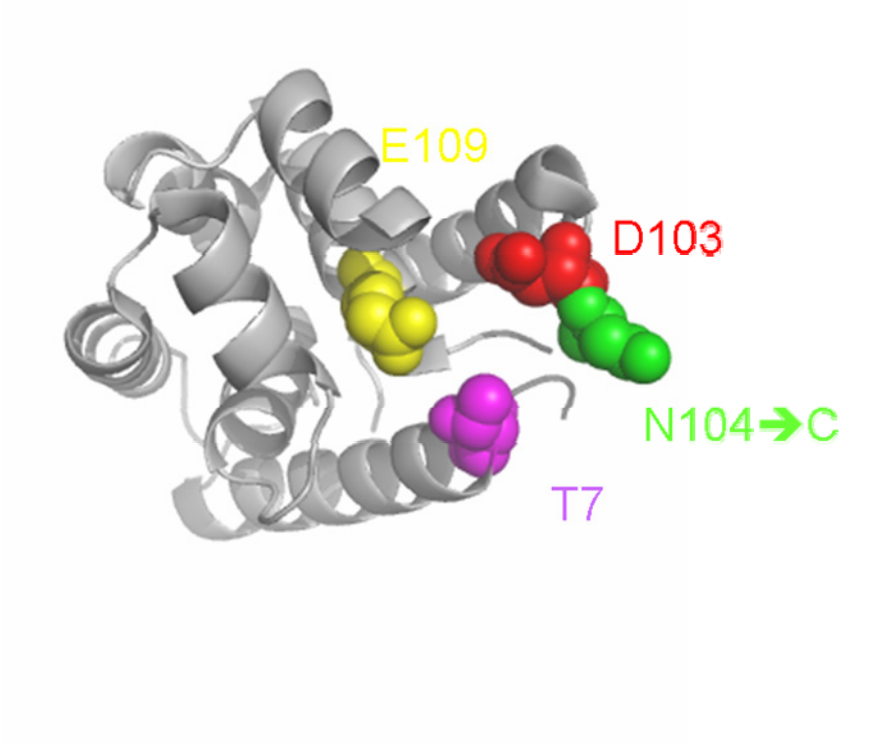


Figure 6.3 Location of the labeling site at the entrance to the ClpB channel. Ribbon representation of the isolated D1 large sub-domain of *E.coli* ClpB (1jbk) (Li and Sha, 2002) showing the conserved substrate-binding Tyr251 residue (red) of the flexible pore loop that lines the entrance to the ClpB translocating channel (Schlieker et al., 2004b; Barnett et al., 2005). Val246 selected for replacement with Cys is located below the conserved pore loop is shown in green. The dashed line indicates the position of the ClpB channel axis.

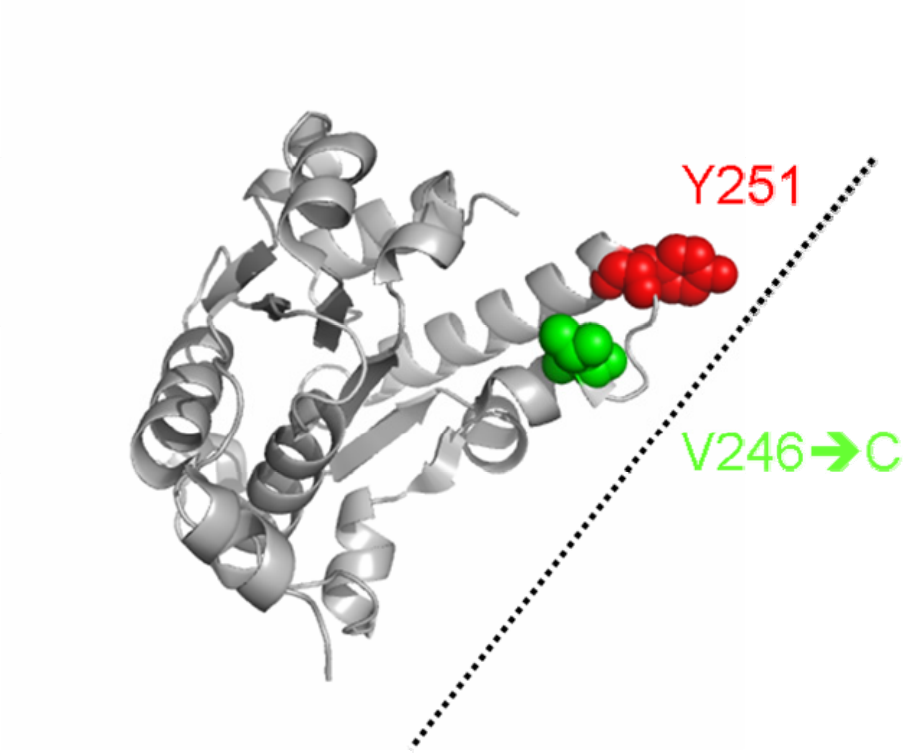


Figure 6.4 Oligomerization of wt ClpB and the various Cys mutants.

Wt ClpB, Cys-less ClpB, N104C ClpB as well as the substrate trap versions of the single-Cys variants of ClpB C104ClpB* (N104C ClpBTrap) and C246ClpB* (V246C ClpBTrap) were run on a Superose 6 gel filtration column at 0.05 ml/min without nucleotides (solid black lines) or in presence of 2 mM ATP (broken red lines). The elution times of four molecular weight standards (thyroglobulin (670 kDa), apoferritin (443 kDa), alcohol dehydrogenase (150 kDa), BSA (66 kDa)) are indicated with black circles in the upper panel.

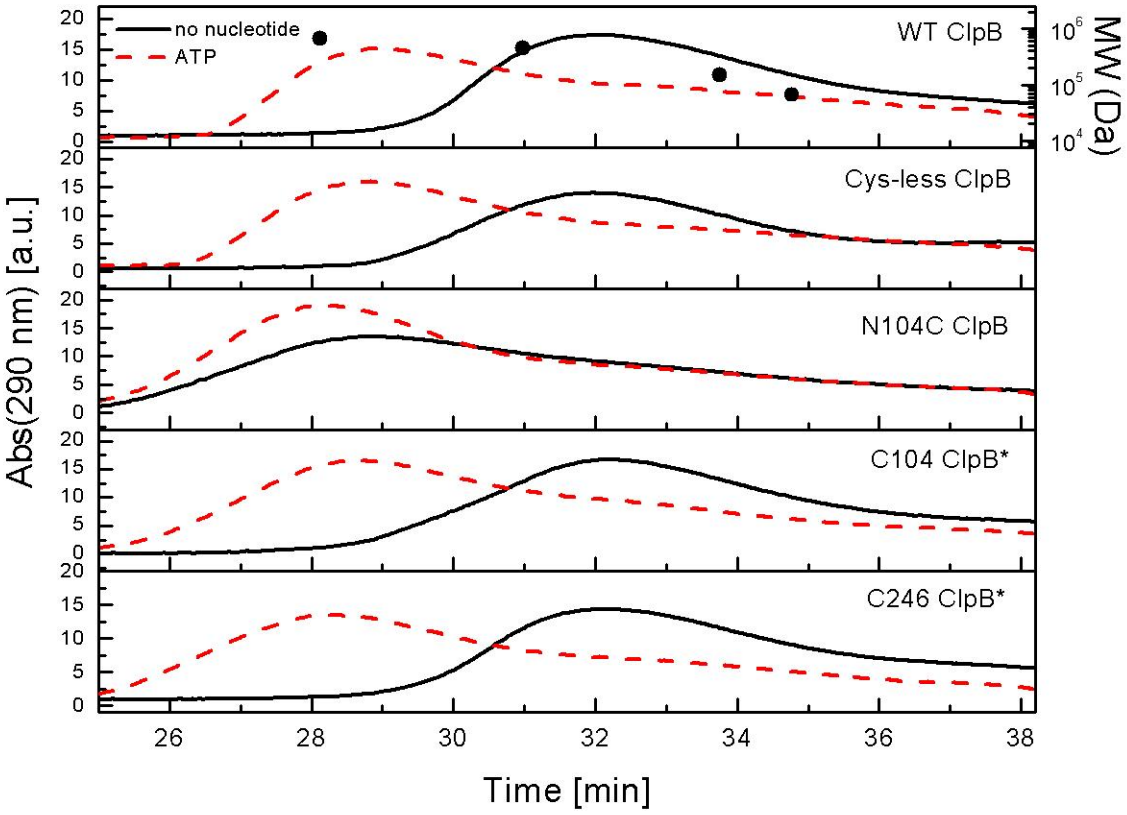


Figure 6.5 ATPase activity of N104C ClpB.

The rate of ATP hydrolysis was measures at 37 ˚C for wt ClpB and N104C ClpB (C94A/C310A/C615A/N104C) without an activator (basal ATPase, red bars) and in presence of 0.1 mg/ml κ-casein (hatched bars). The averages of three measurements are shown with the standard deviation.

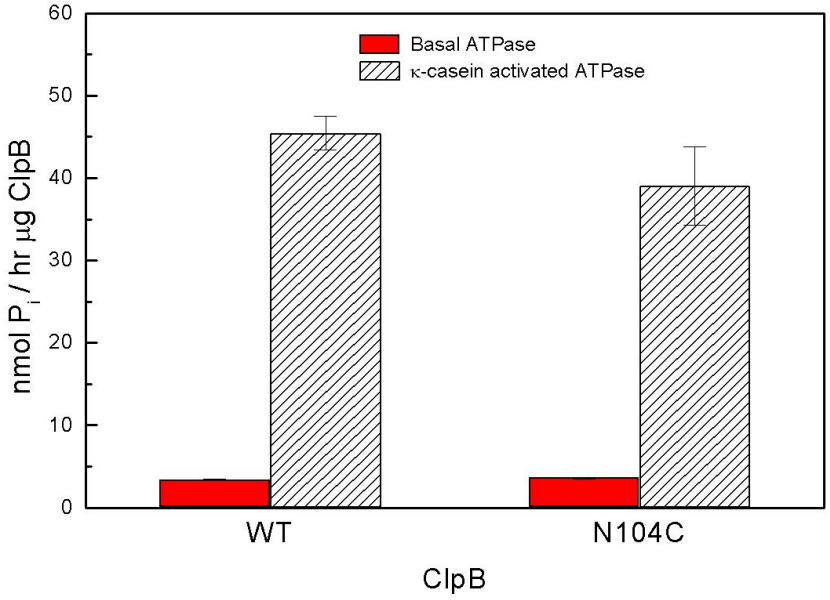


Figure 6.6 Fluorescence emission spectra of the single-Cys variants of ClpB and the Cys-less ClpB after labeling with IAEDANS.

Left panel shows Trp fluorescence spectrum of the IAEDANS labeled: C104ClpB* (blue thin line), C246ClpB* (red thick line), and Cys-less ClpB (black dotted line) obtained with excitation at 290 nm. The right panel shows the IAEDANS fluorescence spectrum obtained with excitation at 336 nm. The total concentration of the various proteins was 5 μM.

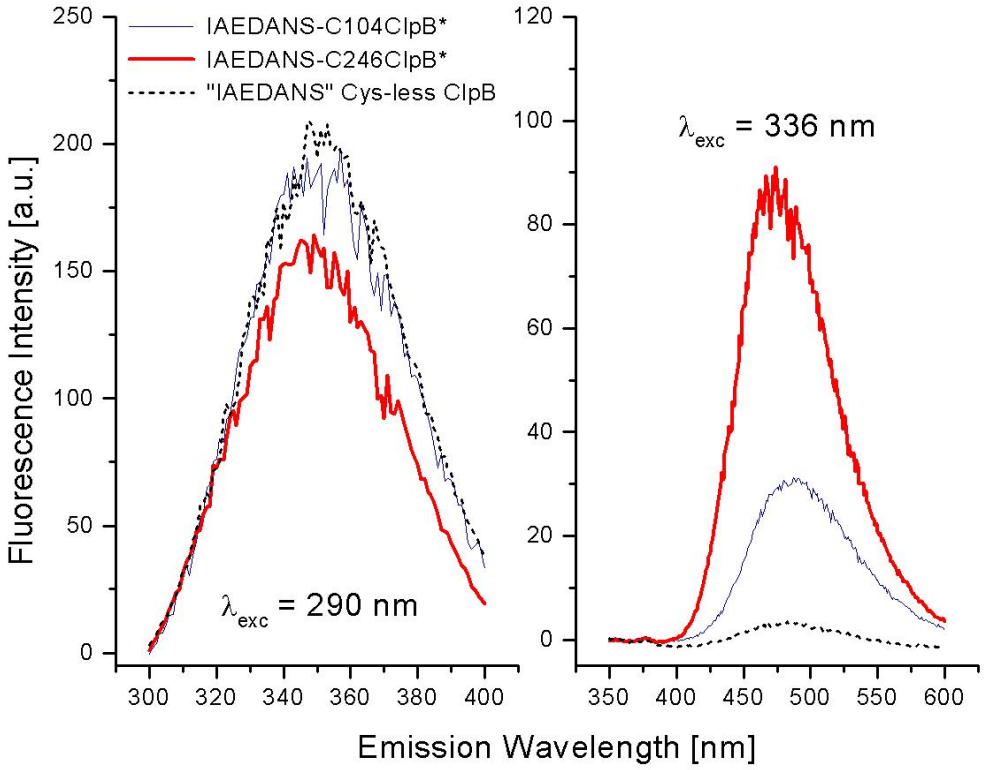


Figure 6.7 Oligomerization of wt ClpB and the various Cys mutants after IAEDANS labeling.

Wt ClpB and the IAEDANS labeled Cys-less ClpB, as well as the substrate trap versions of the single-Cys variants of ClpB C104ClpB* (N104C ClpBTrap) and C246ClpB* (V246C ClpBTrap) were run on a Superose 6 gel filtration column at 0.05 ml/min without nucleotides (solid black lines) or in presence of 2 mM ATP (broken red lines). The elution times of four molecular weight standards (thyroglobulin (670 kDa), apoferritin (443 kDa), alcohol dehydrogenase (150 kDa), BSA (66 kDa)) are indicated with black circles in the upper panel.

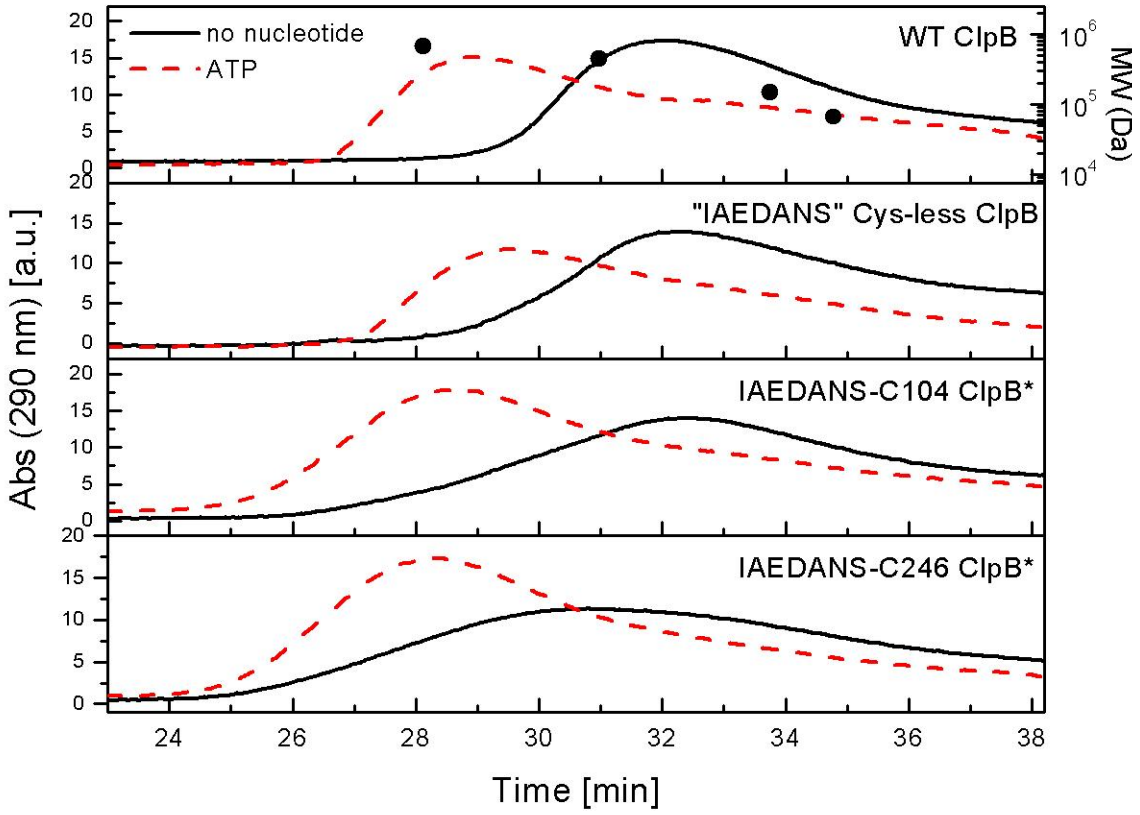


Figure 6.8 ATPase activity of wt ClpB in presence of peptide B1.

The rate of ATP hydrolysis was measured at 37 ˚C for wt ClpB (0.5 μM monomeric) in absence or in presence of the purified peptide B1 (10 μM). The averages of three measurements are shown with the standard deviation.

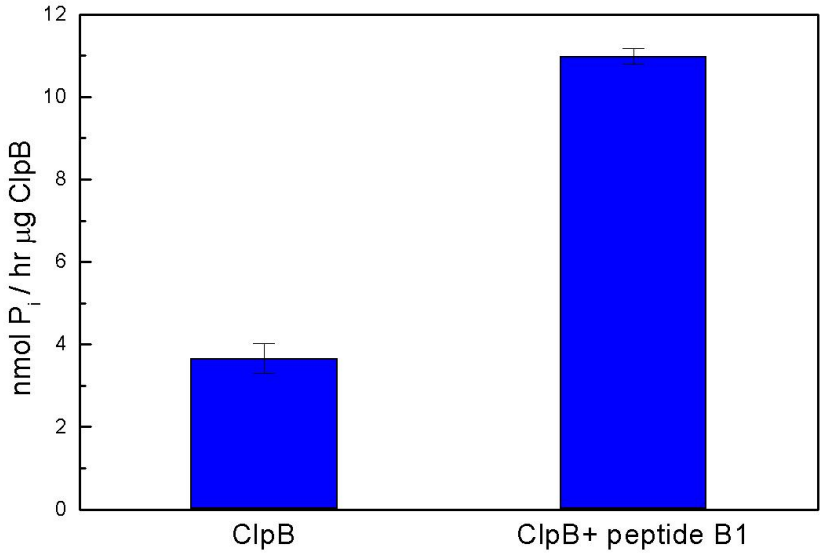


Figure 6.9 Emission spectra of the donor labeled ClpB variants and acceptor labeled model substrate used for FRET.

Fluorescence spectra of 5 μM (0.8 μM hexameric) IAEDANS labeled ClpB (IAEDANS-C104ClpB* shown in panel A or IAEDANS-C246ClpB* shown in panel B) and 0.8 μM fluorescein-peptideB1 individually and mixed together. Excitation wavelength was set at 336 nm in all cases.

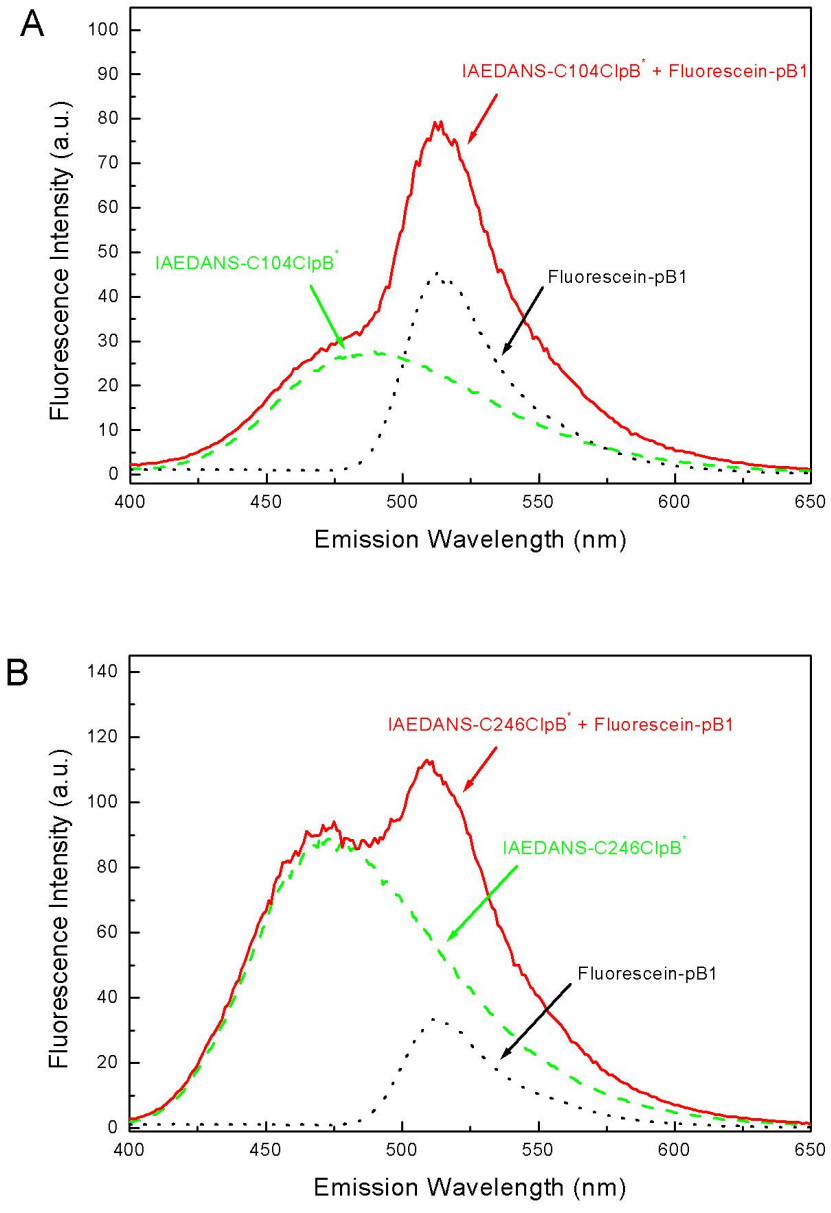


Figure 6.10 FRET between IAEDANS labeled C104ClpB* and fluorescein labeled peptide B1.

Left Panel: fluorescence emission spectra of 0.8 μM (hexameric) IAEDANS-C104ClpB* (D; donor) and 0.8 μM fluorescein-peptide B1 (A; acceptor) without nucleotides (D + A) or with 3 mM ATP (D + A + ATP). Right panel: fluorescence emission spectra of 0.8 μM (hexameric) IAEDANS-C104ClpB* (D; donor) and 0.8 μM non-labeled peptide B1 (pB1) without nucleotides (D + pB1) or with 3 mM ATP (D + pB1 + ATP). Excitation wavelength was set at 336 nm.

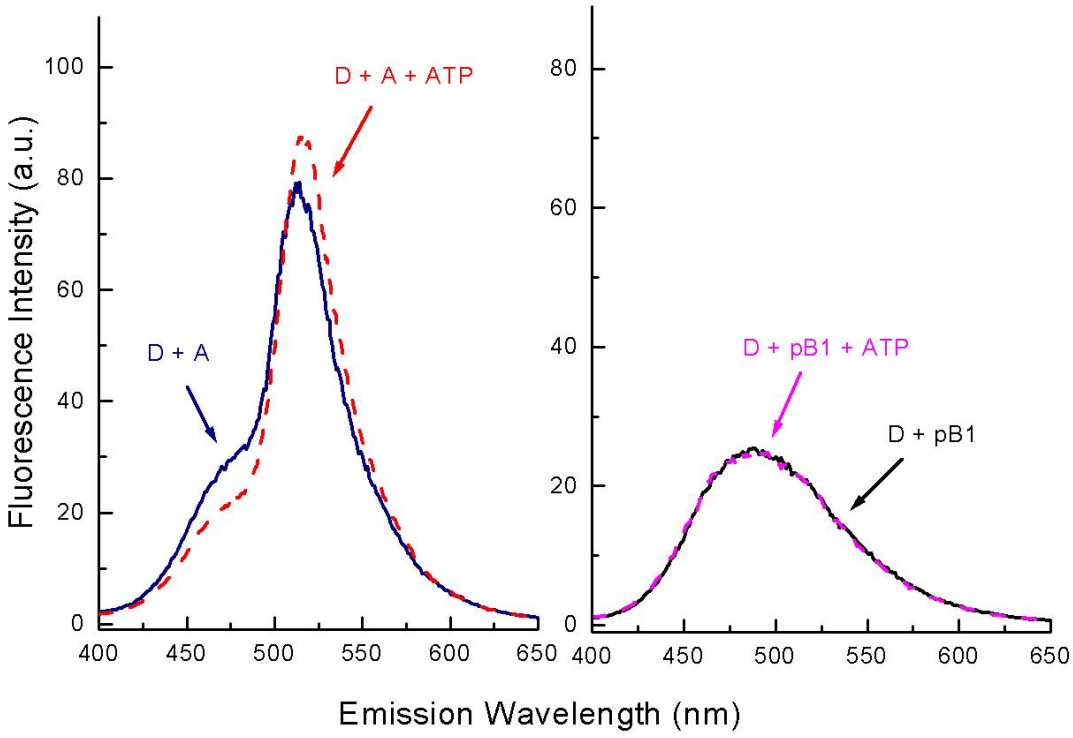


Figure 6.11 FRET between IAEDANS labeled C246ClpB* and fluorescein labeled peptide B1.

Left Panel: fluorescence emission spectra of 0.8 μM (hexameric) IAEDANS-C246ClpB* (D; donor) and 0.8 μM fluorescein-peptide B1 (A; acceptor) without nucleotides (D + A) or with 3 mM ATP (D + A + ATP). Right panel: fluorescence emission spectra of 0.8 μM (hexameric) IAEDANS-C246ClpB* (D; donor) and 0.8 μM non-labeled peptide B1 (pB1) without nucleotides (D + pB1) or with 3 mM ATP (D + pB1 + ATP). Excitation wavelength was set at 336 nm.

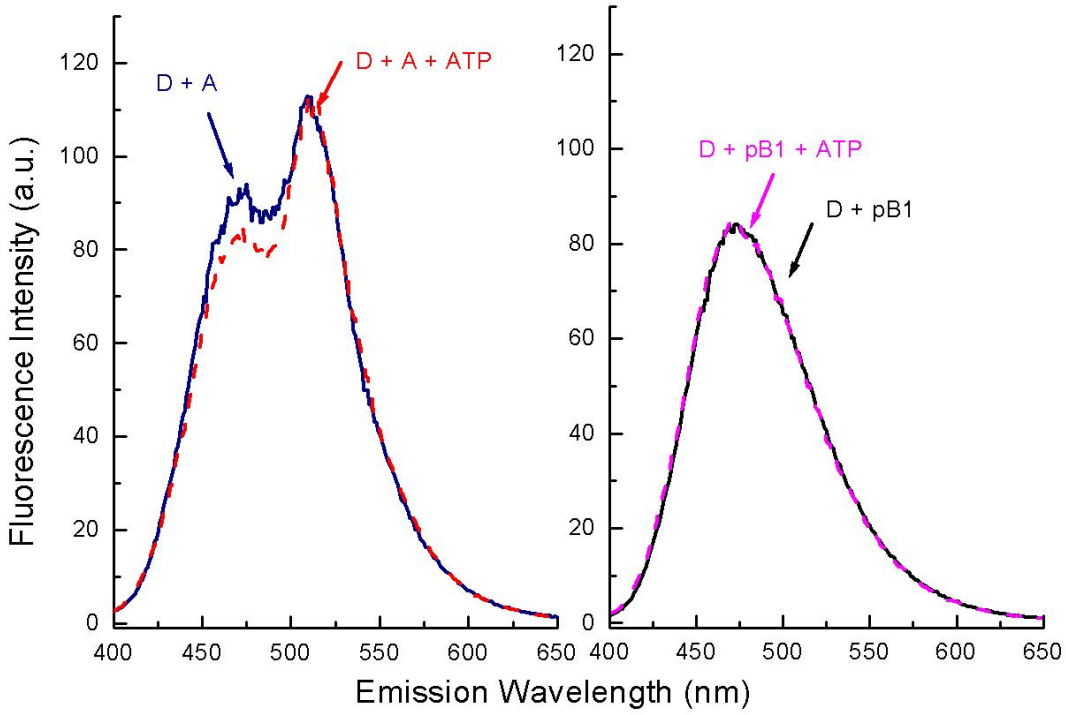


Figure 6.12 FRET between IAEDANS labeled C104ClpB* and FITC-casein.

Left Panel: fluorescence emission spectra of 0.8 μM (hexameric) IAEDANS-C104ClpB* (D; donor) and 1.5 μM FITC-casein (A; acceptor) without nucleotides (D + A) or with 3 mM ATP (D + A + ATP). Right panel: fluorescence emission spectra of 0.8 μM (hexameric) IAEDANS-C104ClpB* (D; donor) and 1.5 μM non-labeled κ-casein (casein) without nucleotides (D + casein) or with 3 mM ATP (D + casein + ATP). Excitation wavelength was set at 336 nm.

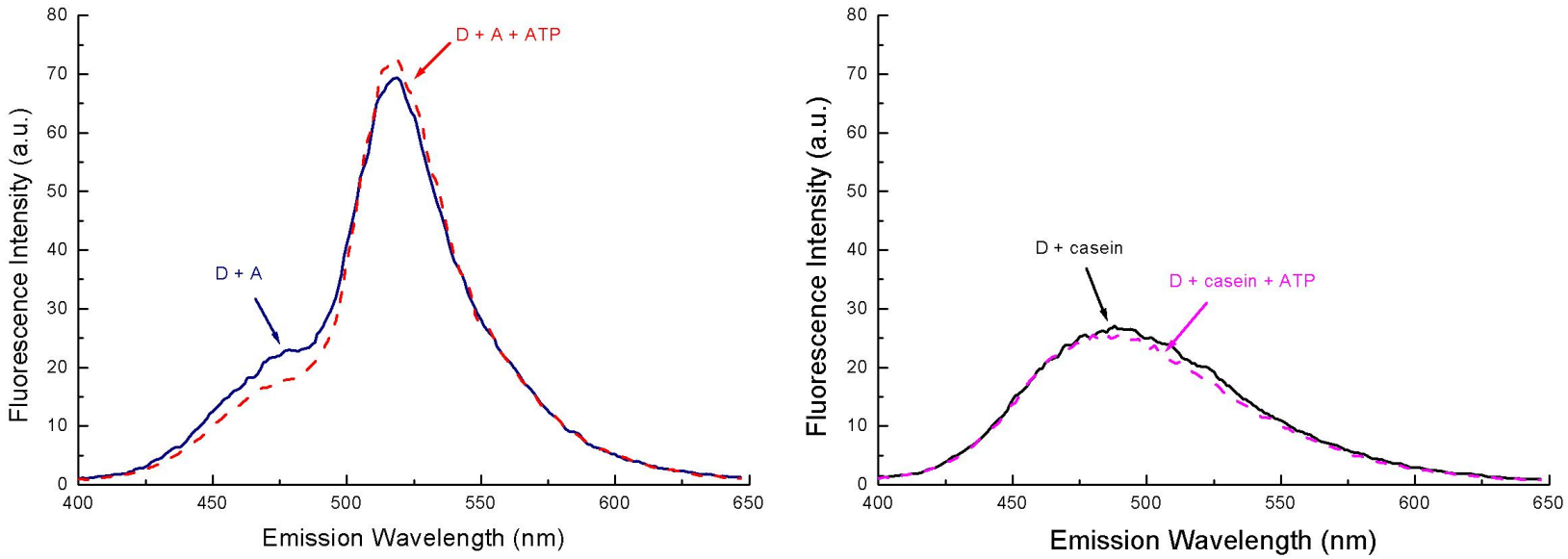
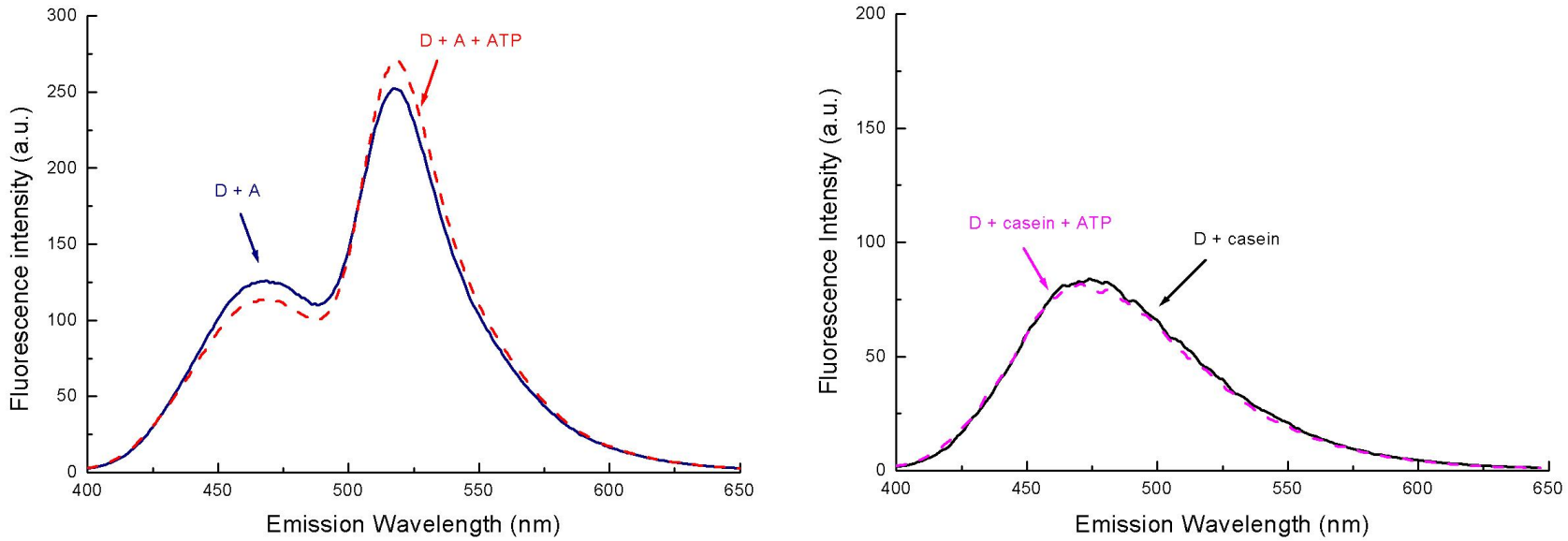


Figure 6.13 FRET between IAEDANS labeled C246ClpB* and FITC-casein.

Left Panel: fluorescence emission spectra of 1.6 μM (hexameric) IAEDANS-C246ClpB* (D; donor) and 6 μM FITC-casein (A; acceptor) without nucleotides (D + A) or with 3 mM ATP (D + A + ATP). Right panel: fluorescence emission spectra of 0.8 μM (hexameric) IAEDANS-C246ClpB* (D; donor) and 1.5 μM non-labeled κ-casein (casein) without nucleotides (D + casein) or with 3 mM ATP (D + casein + ATP). Excitation wavelength was set at 336 nm.



CHAPTER 7 - Summary and Final Discussion

The first part of this dissertation has been dedicated to the description of research oriented towards a better characterization of the ClpB architecture that supports the unique aggregation reversing capability of ClpB. Formation of ring-shaped oligomers is required for nucleotide binding and for the ATP-dependent protein remodeling activity of ClpB (Barnett et al., 2000).

As high-resolution 3D structure of the oligomeric ClpB is lacking, a homology model of the *E. coli* ClpB hexamer was built based on an existing oligomeric model of the *T. thermophilus* ClpB (Diemand and Lupas, 2006) (Chapter 2). Although, such model is not a substitute for high resolution structure, it is a useful tool in exploration of the ClpB structure-function relationship.

Additionally, in order to describe the conformational changes that occur in the various ClpB domains upon the oligomerization of the ClpB ring, we designed seven variants of ClpB each with a single Trp located in each of the ClpB domains. The thermodynamic stability of the distinct ClpB domains in monomeric and oligomeric forms was investigated by following the changes in Trp fluorescence during the equilibrium urea-induced unfolding. We found that the large sub-domain is the most stable, while the N-terminal domain and the middle domain are the least stable among the ClpB domains. Additionally, only two domains of ClpB were destabilized in the oligomeric form: the coiled-coil middle domain and the small sub-domain of the second AAA module. We conclude that only these two domains undergo significant conformational

changes during the formation of the oligomeric rings. The nucleotide-induced conformational changes of the Trp probes were also investigated, and several locations were found to respond to nucleotide-binding or hydrolysis (Chapter 3).

The second part of this work includes the description of a novel filter assay, developed to study ClpB interactions with aggregated proteins. This assay was designed for the quick detection of the interaction between ClpB and strongly aggregated proteins with marginal solubility. Examples were provided where the filter assay was instrumental in detection of binding deficiencies of ClpB variants (Chapter 4).

Chapter 5 describes the studies on the functionality of the two naturally-occurring ClpB isoforms. G6PDH reactivation assays demonstrated that the N-terminal domain of ClpB becomes essential for substrate binding as the size of the aggregates increase (Barnett et al., 2005). However, the major contribution of this work to the ClpB field is the discovery that the optimum chaperone activity of ClpB requires the cooperation of its two isoforms. In an attempt to decipher the mechanism of this synergistic cooperation, we investigated the structural and functional characteristics of the two naturally occurring isoforms in isolation and when combined at different ratios. We found that the hetero-oligomers of ClpB95/ClpB80 posses a higher self-association affinity and an enhanced N-terminal domain mobility when compared to the homo-oligomers of Clp95. We also investigated the three major steps during the reactivation of aggregated substrates: binding, rate of substrate translocation through the ClpB oligomer (which is coupled to the ATPase activity), and the rate of substrate reactivation. We found that hetero-oligomers bind to aggregated substrates with a similar efficiency as homo-oligomers of ClpB95, do not show enhanced ATPase activity over that of the homo-

oligomers, but display a strongly stimulated chaperone activity during the reactivation of strongly aggregated proteins. We proposed that extraction of single polypeptides from aggregates and their delivery to the ClpB channel for translocation is the rate-limiting step in aggregate reactivation and this step is supported by the mobility of the N-terminal domain of ClpB. We also suggest that the mobility of the N-domains is optimized for ClpB function.

The last chapter of this work (Chapter 6) describes the design and production of single Cys variants of ClpB and their subsequent use in FRET experiments. Single Cys residues were introduced in proximity of two putative substrates binding sites (located in the N-terminal domain and at the entrance to the ClpB channel) and were successfully labeled with a fluorescent donor dye. We have demonstrated that FRET can be used to detect ATP-dependent interaction between donor labeled ClpB variants and acceptor labeled pseudo-substrates of ClpB. These new fluorescent tools will be instrumental in future research for the mapping of ClpB-aggregated substrate interaction sites, as well as for the investigation of the mechanism that controls the substrate extraction from the aggregated conglomerates and their insertion into the ClpB channel.

Recently, major advances have been made that enhanced our understanding of the ClpB/DnaK system-mediated protein disaggregation. While several mechanisms were proposed, direct evidence only exist for the threading mechanism (as described in Chapter 1) (Weibezahn et al., 2004). The model presented in Fig. 7.1 summarizes our current understanding of the overall disaggregation process. Based on the accumulated experimental evidence, we propose to distinguish between three major steps during the

disaggregation: substrate recognition, insertion of the extracted polypeptide segment into the ClpB channel, and the subsequent threading through the ClpB channel.

The overall disagregation process is dependent on the additional presence of the DnaK system and cooperation between ClpB and DnaK, but the precise role of this co-chaperone system during the above described steps is still poorly understood. Direct interaction between ClpB and DnaK was not detected, but it was suggested by the observation that cooperation between the two chaperone systems is detected only between proteins of the same species. (Glover and Lindquist, 1998; Schlee et al., 2004)

Besides functioning as a co-chaperone during ClpB-mediated protein disaggregation, DnaK system is able to disaggregate certain small-sized aggregates in absence of ClpB also (top pathway in Fig. 7.1) (Diamant et al., 2000). Recent evidence also suggests that the conformational characteristics of the aggregates determine whether the additional presence of ClpB is needed for the disaggregation and the size of the aggregates is not the sole determinant. Interestingly disordered aggregates of a model substrate were disaggregated by DnaK system, while the reactivation of the aggregates with high β-structure content required the presence of ClpB (Lewandowska et al., 2007).

In case of the aggregated substrates that require the cooperation between both chaperone systems, the order of interaction of the co-chaperones with the aggregates has not been elucidated. It has been shown that the DnaK system has an essential role during the initial steps of the disaggregation process, upstream from substrate translocation (Weibezahn et al., 2004). Kinetic studies revealed that preincubation of

aggregates with DnaK system results is a faster reactivation (Zietkiewicz et al., 2004) (also Chapter 5 of this work). This could suggest that productive ClpB-aggregate complex formation requires the DnaK system. ClpB interacts with substrates in an ATP-bound conformation, but the detection of this complex is a challenging task due to the short life-time of the ATP-bond conformation of ClpB. Experiments with the trap version of ClpB that represents a frozen ATP-bond state of ClpB demonstrate that ClpB can recognize aggregated substrates in absence of the DnaK system (Weibezahn et al., 2003) (also Chapter 5 of this work).

Interestingly, it was shown that ClpB (and also the yeast homolog: Hsp104) can disaggregate and reactivate small aggregates independently of the DnaK system, when mixtures of ATP and ATPγS are used in the reactivation assays. This DnaK independent activity of ClpB was also observed with ClpB variants that contained mutations in the ATP-binding sites in the first or the second AAA+ domain, in presence of ATP (Doyle et al., 2007). In can be suggested that when the ATPase activity of ClpB is decelerated, ClpB-aggregate interaction becomes more stable and ClpB can unfold the bound substrates. Based on the above observations we suggest that ClpBtrap stabilizes a functionally relevant state of ClpB that probably represents an intermediate state to the reactivation process (bottom pathway in Fig. 7.1). The pre-incubation experiments suggest that under physiological conditions (that is in absence of the trap variant of ClpB and ATPγS) DnaK system can prepare the aggregates for efficient interaction with ClpB by remodeling the surface of the aggregates and thus providing “stronger” recognition signals that allow for a stable initial interaction between aggregates and ClpB, which in turn is a prerequisite for the subsequent threading

(recognition step in Fig. 7.1). However, this initial recognition and stable interaction is not sufficient for initialization of the threading.

Experiments with the two naturally occurring isoforms of ClpB prompted as to suggest that there is an intermediate step between the initial binding and the translocation that is rate limiting for the overall disaggregation. We proposed that this rate-limiting step (red arrow in Fig. 7.1) involves extraction of single polypeptides from the already bound aggregated substrates and their insertion into the ClpB channel for translocation. This is in agreement with previous experimental results that showed that binding of aggregated substrates to the central pore of the ClpB channel, as well as the degradation by the engineered BAP-ClpP, requires cooperation with the DnaK system (Weibezahn et al., 2004; Haslberger et al., 2007) (also described in Chapter 1). It was proposed that nucleotide-dependent motions of middle domain of ClpB are necessary for the cooperation with the DnaK system (Haslberger et al., 2007).

In conclusion we proposed that the N-terminal domains not only support the initial recognition and binding of aggregates, but their mobility possibly supports the rate limiting step of the disaggregation. Additionally, this rate-limiting step is also supported by the DnaK system of co-chaperones and possibly by the mobility of the middle domains.

The ClpB-DnaK mediated protein solubilization and reactivation of aggregated proteins is of major importance in living cells. Bukau and co-workers have shown that removal of aggregated proteins by degradation is not sufficient. In order to ensure proper cell function, refolding and thus reactivation of the major cellular proteins is essential (Weibezahn et al., 2004).

The consequences of accumulation of aggregated proteins are deleterious for the living cells, as exemplified by the large number of severe neurorodegenarative diseases. Quite intriguingly, ClpB is only conserved in bacteria, plants, and fungi, but no orthologs were identified so far in higher eukaryotes. It has been emphasized that species that contain ClpB lack large-scale mobility and are unable to escape sudden changes in environment (Weibezahn et al., 2005).

Recently, it has been proposed that the yeast Hsp104 may be useful as a therapeutic agent in elimination of the disease associated amyloids (Shorter, 2008). It remains to be seen if such attempts could contribute to the development of effective treatments for the various neurodegenerative diseases. Meanwhile, future research regarding the details of substrate recognition, as well as the role of the ClpB-DnaK co-operation , will be essential in understanding the intricate mechanism by which ClpB achieves its remarkable aggregation reversing activity.

References

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

Barnett,M.E., Zolkiewska,A., and Zolkiewski,M. (2000). Structure and activity of ClpB from *Escherichia coli*. Role of the amino- and -carboxyl-terminal domains. J. Biol. Chem. *275*, 37565-37571.

Diamant,S., Ben Zvi,A.P., Bukau,B., and Goloubinoff,P. (2000). Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. J. Biol. Chem. *275*, 21107-21113.

Diemand,A.V. and Lupas,A.N. (2006). Modeling AAA+ ring complexes from monomeric structures. J. Struct. Biol. *156*, 230-243.

Doyle,S.M., Shorter,J., Zolkiewski,M., Hoskins,J.R., Lindquist,S., and Wickner,S. (2007). Asymmetric deceleration of ClpB or Hsp104 ATPase activity unleashes protein-remodeling activity. Nat. Struct. Mol. Biol. *14*, 114-122.

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

Haslberger,T., Weibezahn,J., Zahn,R., Lee,S., Tsai,F.T., Bukau,B., and Mogk,A. (2007). M domains couple the ClpB threading motor with the DnaK chaperone activity. Mol. Cell *25*, 247-260.

Lewandowska,A., Matuszewska,M., and Liberek,K. (2007). Conformational properties of aggregated polypeptides determine ClpB-dependence in the disaggregation process. J. Mol. Biol. *371*, 800-811.

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

Shorter,J. (2008). Hsp104: a weapon to combat diverse neurodegenerative disorders. Neurosignals. *16*, 63-74.

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

Weibezahn,J., Schlieker,C., Tessarz,P., Mogk,A., and Bukau,B. (2005). Novel insights into the mechanism of chaperone-assisted protein disaggregation. Biol. Chem. *386*, 739-744.

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

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

Figures

Figure 7.1 Model of ClpB-mediated protein disaggregation.