Sri Lankan J. Biol. 2025, 10 (2): 33-48

DOI: https://doi.org/10.4038/sljb.v10i2.124

Research Article Open Access

Isolation of ClpB & DnaK bi-chaperones from obligate pathogen *Anaplasma phagocytophilum* as future therapeutic targets

Chathuranga B Ranaweera[1,3,*], Swetha Madesh[2], Sunitha Shiva[1], Roman R Ganta[2] and Michal Zolkiewski[1]

[1]Department of Biochemistry & Molecular Biophysics, Kansas State University, Manhattan Kansas 66506, USA.

[2]Department of Diagnostic Medicine/Pathobiology, Kansas State University, Manhattan Kansas 66506, USA.

[3]Department of Medical Laboratory Sciences, Faculty of Allied Health Sciences, General Sir John Kotelawala Defence University, Sri Lanka.

Abstract

Anaplasma phagocytophilum is an obligate intracellular pathogen that causes Human Granulocytic Anaplasmosis (HGA). Genetic manipulation of obligate intracellular pathogens is extremely difficult, as these organisms possess small genomes made primarily of essential genes and rely significantly on the host cell for proliferation. To date, only a handful of genetic tools have been developed to study the pathobiology of *Anaplasma,* and doxycycline remains the only available treatment for HGA. ClpB and DnaK are molecular chaperones responsible for reactivating and resolubilizing protein aggregates within the cellular environment. Recently, ClpB has emerged as a promising target for the development of novel antimicrobials, as it is essential for the survival of many pathogens and lacks apparent orthologs in metazoans. Initial findings indicate that these chaperones have different substrate preferences, despite being almost identical across different species, providing a unique approach to develop novel therapeutics against specific species.

Received: 14 June 2024

Accepted: 27 April 2025

Keywords:
AAA+ ATPases;
ClpB; DnaK;
Molecular chaperones;
Anaplasma phagocytophilum;
Human Granulocytic Anaplasmosis.

*Corresponding author:
cbr2704@kdu.ac.lk

https://orcid.org/0000-0003-4173-0547

1. Introduction

Proteins are the most versatile and structurally complex macromolecules found in any biological system. They perform many essential functions and participate in the majority of biological reactions in a cell. As a consequence, protein synthesis, protein folding, protein refolding and protein degradation/recycling are vital components of cellular protein homeostasis (Morimoto and Cuervo, 2009). One of the most important biological functions of protein homeostasis within a living cell is to prevent protein aggregation and maintain cellular proteins in a soluble and active state. Many proteins, most prominently, molecular chaperones, support protein quality and protein homeostasis. According to current definition, a molecular chaperone is a protein that assists other proteins to reach their biologically active, native conformation. Chaperones interact transiently with their client proteins without forming stable complexes with their folded, active structure (Ranaweera, 2021; Hartl *et al*., 2011).

During biological evolution, certain cells have developed a specific type of chaperones with the unique ability to disassemble protein aggregates and convert them into soluble, unstructured polypeptides. These molecular chaperones are known as Caseinolytic Peptidase B (ClpB) in bacteria, Hsp104 in yeast, and Hsp101 in plants. As their monomeric molecular weight

is around 100 kDa (kilodaltons), they are also known as members of the Hsp100 family. Interestingly, while Hsp100 proteins are found in bacteria, yeast and plants, they are absent in animals and humans (Ranaweera, 2021; Zolkiewski *et al*., 2012).

Hsp100 proteins belong to the AAA+ (ATPase Associated with various cellular Activities) superfamily, and they utilize the energy from ATP hydrolysis to disassemble and re-solubilize large complexes of aggregated proteins (Zolkiewski *et al.*, 2012). However, there is limited understanding of how Hsp100 proteins recognize specific substrates, therefore this is an active research area. We recently demonstrated that the substrate recognition mechanism of *Escherichia coli* ClpB may rely on the global surface properties of aggregated proteins rather than local sequence motifs (Ranaweera *et al*., 2018).

Upon infecting a host, pathogens experience heat shock and oxidative stress, and their survival depends on molecular responses to these external conditions. The stress response in pathogens has therefore emerged as a critical mechanism for the development of novel antimicrobials. In particular, understanding the role of molecular chaperones in host-pathogen interactions has emerged as a novel direction in studying mechanisms of infection, with ClpB emerging as a promising target for the development of novel antimicrobials (Zhang *et al*., 2013; Glaza *et al*., 2021).

Anaplasma phagocytophilum is a Gram negative, obligate intracellular pathogen measuring 0.5 to 1.0 µm in size. It is considered one of the most important tick-borne bacteria, causing infections in both veterinary animals and humans (Dumler *et al*., 2005). This pathogen causes Human Granulocytic Anaplasmosis (HGA), the second most commonly reported tick-borne illness in the United States, resulting in significant morbidity and mortality (Zhang *et al*., 2021). Despite growing public health concerns regarding HGA, there is limited information about the physiology of the bacterium, host-pathogen interactions, virulence mechanisms and genetic manipulation strategies (Cheng *et al*., 2013; Atif, 2015). HGA can be fatal, particularly in immunocompromised individuals, the elderly, and children (Zhang *et al*., 2021). No vaccines are currently available, and doxycycline remains the only treatment. Due to the high clinical significance of *A. phagocytophilum* and the need for biochemical studies, we undertook the first cloning, expression, and purification of ClpB and DnaK chaperones from this bacterial pathogen. The protocol we describe has been optimized for cost-effectiveness (e.g., avoiding the use of ATP) and efficiency.

2. Materials and Methods

2.1 Cloning of Anaplasma phagocytophilum ClpB and DnaK

A. phagocytophilum strain HZ was used to infect HL-60 cells (American Type Culture Collection, CCL-240) for *Anaplasma* genomic DNA isolation. HL-60 cells were cultured in T25 flasks with complete RPMI 1640 medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum and 2 mM L-glutamine. Cells were maintained at 37ºC in an incubator with 5% CO_2. Infectivity of HL-60 host cells was tested daily by preparing cytospin slides from the culture and staining them with the Hema-3 kit (Sigma-Aldrich).

Once HL-60 cells were infected, about 80-90% of cells were collected from the T25 flasks and were harvested by centrifugation at 400 × g for 10 minutes at 4°C. The resulting pellet was resuspended in 5 mL of RPMI media, and the cells were lysed by sonication, using two series of 30 second pulses at a setting of 6.5 (Fisher Scientific Sonic Dismembrator Model 100) to release *Anaplasma* from HL-60 cells. The lysate was centrifuged at 100 × g for 10 minutes at 4°C, to pellet the host cell nuclei and other host cell debris. The supernatant was then passed through 2.0 µm sterile isopore membrane filters. The filtrate was then centrifuged at 20,000 × g for 20 minutes at 4°C to isolate a host cell free *Anaplasma* sample.

Lysozyme was added to the isolated *Anaplasma* sample to a final concentration of 50 µg/ml, followed by incubation for 10 minutes. Genomic DNA was then extracted from the *Anaplasma* sample using the Wizard Genomic Purification kit (Promega). To confirm the presence of *Anaplasma* genomic DNA, PCR was performed using the following *Anaplasma* 16s ribosomal RNA gene-specific primers (16S rRNA gene 0.48-kb, Forward: 5'CTCAGAACGAACGCTGG3', Reverse: 5'CATTTCTAGTGGCTATCCC3'). PCR reaction was performed as follows: pre-heating for 3 minutes at 95°C, followed by 45 cycles of 95°C for 15 seconds, 50°C for 30 seconds, and 60°C for 60 seconds. PCR products were resolved on a 1% agarose gel, stained with ethidium bromide and visualized under UV light.

The pET-28a(+) vector (Novagen) was digested with Nhe1 and Xho1 restriction enzymes, followed by purification with QIAquick PCR (Qiagen) and Gel Cleanup Kit (Qiagen). The following primers were designed to amplify *Anaplasma* ClpB and DnaK genes (restriction sites underlined):

Aph ClpB-Forward:
5'ACTGAGCTAGCATGGACTTGAATAAGTTTACT
GA3'

AphClpB-Reverse: 5’ TCAGTCTCGAGTTATGTAGCCTCTTTGATAACTA G3’

Aph DnaK-Forward: 5’ACTGAGCTAGCATGGGGGTAGTCATGG 3’

Aph DnaK-Reverse: 5’ TCAGTCTCGAGCTAAGTATTCTTCTTGTCCTCG3’

The protein-coding sequences of *A. phagocytophilum* ClpB and DnaK were amplified using the above primers, with the isolated genomic DNA as the template. Q5® High-Fidelity DNA Polymerase (New England Biolabs, NEB) was used for PCR reactions according to the manufacturer’s instructions. The amplified ClpB and DnaK PCR products were digested with NheI and XhoI (NEB) and purified using the QIAquick PCR Purification Kit to remove digested overhangs and other small fragments.

The digested vector and the digested PCR products were then ligated using T4 DNA Ligase (NEB), following the manufacturer’s instructions. The resulting recombinant *Anaplasma* ClpB (pCBR_AphClpB) and DnaK (pCBR_AphDnaK) plasmids were transformed into NEB® 5-alpha *E. coli* cells. Following transformation, plasmids were propagated in *E. coli* and then purified using a plasmid purification kit (Qiagen). To identify correct clones of *Anaplasma* ClpB (pCBR_AphClpB) and DnaK (pCBR_AphDnaK), plasmid DNA was digested with the following restriction enzymes (Nhe1, Xho1, Cla1, Sph1 and HindIII) and was then analyzed by agarose gel electrophoresis. The correct insertion of the *Anaplasma* ClpB and DnaK genes was further verified by DNA sequencing using the following primers: T7 promoter primer, T7 terminator primer,

AphClpB Internal 1 Forward 5’ CTCGCTAGGAGATTTCAACCA G 3’,

AphClpB Internal 1 Reverse 5’GCTGCATTACCTGACGATCAATAC3’,

AphClpB Internal 2 Forward 5’ ACAAAAAACAATCCTGCCCTAATTG3’,

AphClpB Internal 2 Reverse 5’CTAATA CCGTGGTGCAGCTC 3’,

AphDnaK Internal 1 Forward 5’ GCTGGTACGATTGCTGGCCTA G 3’,

AphDnaK Internal 1 Reverse 5’ ATCACTCAA AGCCTTCTTACATGGCTCTATAGT 3’,

AphDnaK Internal 2 Forward 5’ GCGGGTATTAAGGAT AACAGTAAGGTCGAC3’,

AphDnaK Internal 2 Reverse 5’ GCCTTATCCTTTGCTGAAACG TGCA3’.

2.2 Purification of Anaplasma phagocytophilum ClpB and DnaK

Anaplasma ClpB (pCBR_AphClpB) and DnaK (pCBR_AphDnaK) plasmids were transformed into *E. coli* Bl21(DE3) cells (Novagen). The cells (50–100 µl) were spread on agar plates supplemented with kanamycin (50 µg/ml) and incubated overnight at 37°C. A single colony of each recombinant plasmid was used to inoculate fresh LB media supplemented with kanamycin, followed by overnight incubation at 37°C. The final culture was used to prepare glycerol stocks for further use.

To produce recombinant *Anaplasma* ClpB and DnaK, a previously prepared glycerol stock sample was used to inoculate the LB medium supplemented with kanamycin (50 µg/ml). After overnight incubation at 30°C, the culture was diluted 50-fold in 1 liter of LB containing kanamycin and incubated again at 30°C. When the absorbance at 600 nm reached approximately 0.4, IPTG (Isopropyl ß-D-1-thiogalactopyranoside) was added to a final concentration of 1 mM, and the culture was incubated overnight at 20°C with shaking at 250 RPM. Cells were collected by centrifugation (Beckman JA-14 rotor, 3000 × g, 15 minutes, 4°C). The cell pellet was then resuspended in 20 ml of cold Ni-NTA lysis buffer (40 mM Tris-HCl, pH 8.0, 100 mM KCl, 10 mM imidazole) by vortexing. Phenylmethylsulfonyl fluoride (PMSF) was added to a final concentration of 1 mM together with 1 tablet of cOmpleteTM Protease Inhibitor (Promega) per 50ml of sample. The cells were disrupted by sonication on ice using pre-chilled tubes. The soluble fraction was collected by centrifugation (Beckman JA-20 rotor, 11,500 RPM, 60 min, 4°C). The supernatant was loaded onto a column containing 5 ml of nickel nitrilotriacetic acid (Ni-NTA) resin (Qiagen) equilibrated with Ni-NTA lysis buffer. The column was washed with 50 ml of Ni-NTA wash buffer (40 mM Tris-HCl, pH 8.0, 100 mM KCl, 20 mM imidazole), and proteins were eluted in 5-ml fractions using Ni-NTA elution buffer (40 mM Tris-HCl, pH 8.0, 100- mM KCl, 250 mM imidazole).

The fractions containing either *Anaplasma* ClpB or DnaK were pooled and concentrated using a centrifugal filter device (10,000 Da molecular weight cut-off, 3500 RPM, 4°C). The concentrated *Anaplasma* ClpB or DnaK were further purified by gel filtration chromatography on a Superdex 200 column (Bio-Rad) equilibrated with gel filtration buffer (50 mM Tris-HCl, pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 10% glycerol, 1 mM EDTA, 1 mM DTT). Gel filtration chromatography was performed at 4°C with 0.3 ml/min flow rate.

Fractions containing proteins with >95% purity were pooled and subjected to thrombin cleavage (Sigma-Aldrich) to remove the N-terminal 6-His tag. Purity of the thrombin-cleaved samples was assessed by SDS-PAGE (Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis), and the final protein concentration was determined using the Bradford method. Aliquots of purified recombinant *Anaplasma* ClpB and DnaK were stored at −20°C in gel filtration buffer.

2.3 Activity assay of purified recombinant proteins

The ATPase activity of purified recombinant *Anaplasma* ClpB and DnaK was determined as previously described by Glaza and colleagues (Glaza *et al*., 2021). *Ehrlichia chaffeensis* ClpB has been cloned earlier (Zhang *et al*., 2013) and *E. chaffeensis* recombinant ClpB proteins were used for comparison of ATPase activity assays.

3. Results

3.1 Expression of recombinant Anaplasma ClpB and DnaK in E. coli

To produce functional proteins, *Anaplasma* ClpB and DnaK genes were amplified by PCR from genomic DNA of *Anaplasma* using gene specific primers and were cloned into the bacterial expression vector pET-28a(+), placing the genes under the control of the T7 promoter. The size of the newly constructed AphClpB was 7876 bp. As expected, a single band of approximately 8 kb was observed (Fig:1, top left). Cloning of the inserts was confirmed by restriction enzyme digestion of the recombinant plasmids (pCBR_AphClpB and pCBR_AphDnaK) (Fig. 1). The pCBR_AphClpB plasmid was digested with Nhe1 and Xho1 (2586 bp and 5296 bp expected fragments), Cla1 and Sph1 (1501 bp, 2856 bp and 3525 bp expected fragments). The restriction-enzyme digestion results confirmed the correct assembly of the pCBR_AphClpB construct (Fig: 1, bottom left).

For the pCBR_AphDnaK, a single band of 1950 bp was observed as expected on 1% Agarose gel following ethidium bromide staining (band for full length pCBR_AphDnaK is not shown). Next, pCBR_AphDnaK plasmid was further digested with restriction enzyme combinations, Nhe1 and Xho1 (2586 bp & 5296 bp) and with Cla1 and HindIII (4556 bp & 2696 bp) for identification of predicted band sizes. Clones 5,6 and 7 were initially selected based on NheI and XhoI digestion of pCBR_AphDnaK. However, an extra band was observed in clone 5 after Cla1 and HindIII digestion (Fig: 1, bottom right). Restriction digestion results confirmed that clones 6 and 7 contained the correct clone of pCBR_AphDnaK (Fig: 1, bottom right).

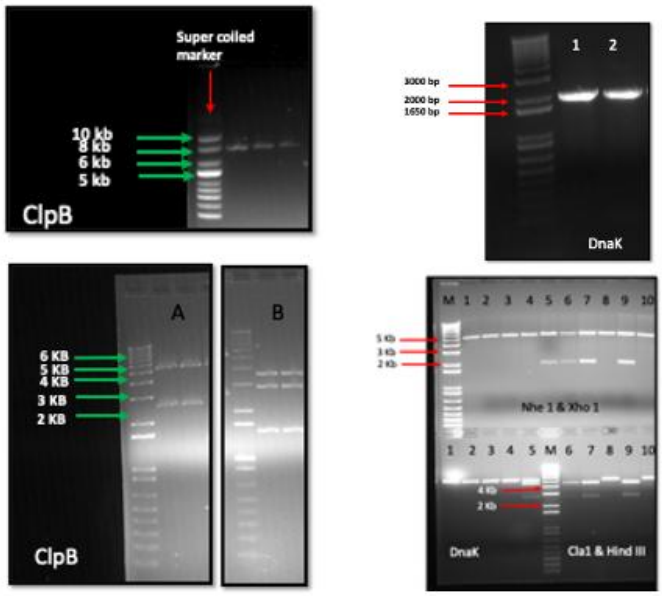


Figure 1: Cloning results for *Anaplasma* ClpB and DnaK. Each lane represents an individual colony.

Top left: 1% agarose gel of recombinant *Anaplasma* pCBR_AphClpB plasmid (7882 bp). Bottom left: A- *Anaplasma* ClpB recombinant plasmid digested with Nhe1 and Xho1 enzymes, B- *Anaplasma* ClpB recombinant plasmid digested with Cla1 and Sph1 enzymes. Top right: 1% agarose gel for PCR amplified new recombinant *Anaplasma* DnaK fragments (1950 bp). Full length pCBR_AphDnaK plasmid not shown. Bottom right: *Anaplasma* DnaK recombinant plasmid digested with Nhe1 and Xho1, Cla1 and HindIII enzymes.

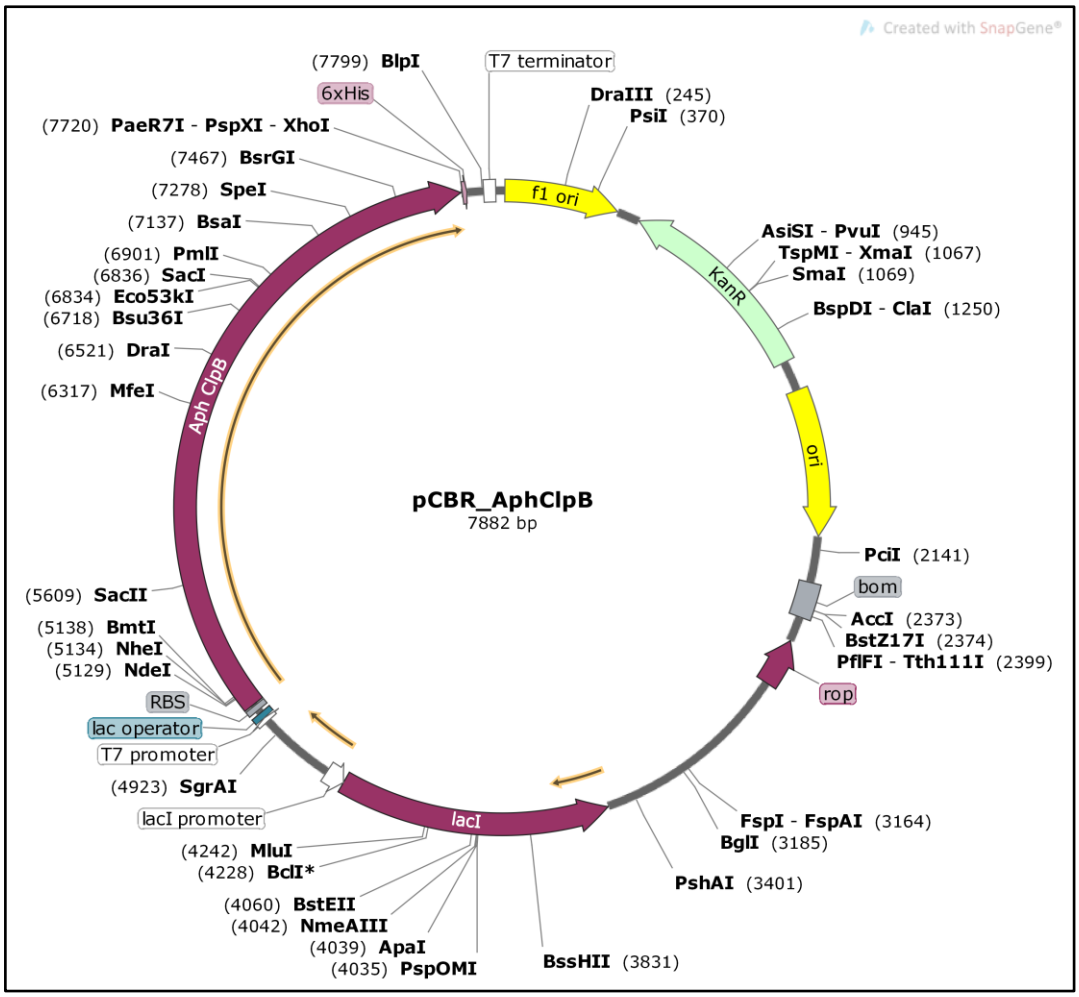


Figure 2. *A. phagocytophilum* ClpB plasmid map.
Anaplasma ClpB gene (2649bp) was ligated into pET-28a(+) vector backbone after restriction digestion with Nhe1 and Xho1 restriction enzymes. This newly constructed plasmid has Kanamycin resistance. When *Anaplasma* ClpB gene gets translated from this new plasmid, six Histidine tags will be attached to the N-terminus of *Anaplasma* ClpB. A Thrombin cleavage site is located between six Histidine tags and the first Methionine of *Anaplasma* ClpB gene. The six Histidine tags and Thrombin cleavage site are not shown in the figure. A stop codon is located in front of the six Histidine tags at the C-terminus of *Anaplasma* ClpB plasmid construct. Hence, no C-terminus Histidine tag gets attached with translated *Anaplasma* ClpB. Different restriction enzyme locations are also shown in the plasmid map. Plasmid map was created using SnapGene software.

Both pCBR_AphClpB and pCBR_AphDnaK plasmids were further confirmed by DNA sequencing using overlapping gene specific primers (primers are shown in the method section; sequencing results are not shown). Correct recombinant plasmids (Figures 2 & 3) were transformed into *E. coli* BL21(DE3) cells, and over-production was performed as described in materials and methods (Fig. 3).

To examine the expression of the recombinant proteins, the genes of interest were cloned under the control of the T7 promoter.

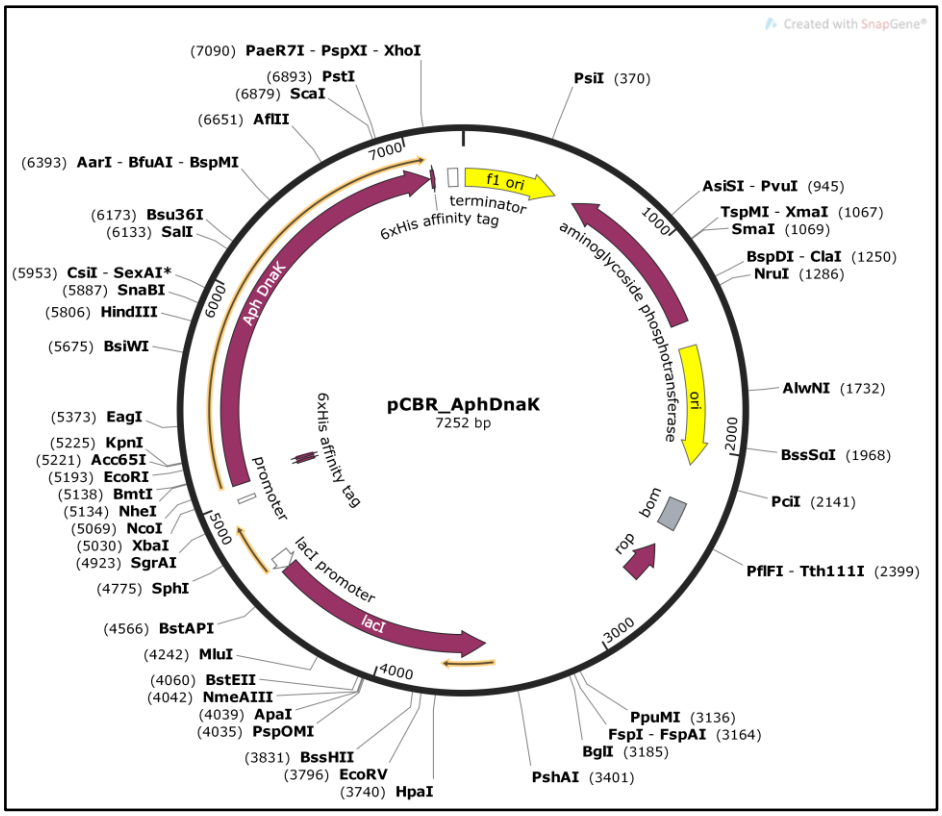


Figure 3. *A. Phagocytophilum* DnaK plasmid map.
Anaplasma DnaK gene (2019bp) was ligated into pET-28a (+) vector backbone after restriction digestion with Nhe1 and Xho1 restriction enzymes. This newly constructed plasmid has Kanamycin resistance and contains the same features described in figure 2.

After IPTG induction, cells were harvested, and SDS-PAGE analysis of the induced culture pellet revealed major bands around 95 kDa (for *Anaplasma* ClpB) and 70 kDa (for *Anaplasma* DnaK) in *E. coli* BL21 (DE3) cells (Fig. 4).

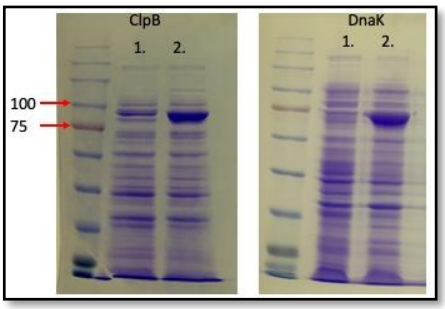


Figure 4. SDS-PAGE gel analysis results after IPTG induction.

The image of 10% SDS-PAGE gel is shown here. Arrows indicate apparent molecular weights of bands in protein weight marker. 1= uninduced sample. 2= induced sample. An equal number of cells were used for protein isolation and gel loading. Cells were suspended in 160 µl of 2x Laemmli sample buffer (Laemmli 1970). Sixty microliters of each sample were resolved in duplicates in 10% SDS-PAGE gel. The gel was stained in Coomassie R-250.

3.3 Ni-NTA purification of recombinant proteins

Ni-NTA purification was performed as mentioned in the methodology, and results are shown in Figures 5 and 6. The majority of the Histidine-tagged recombinant *Anaplasma* ClpB was eluted within the first three eluted fractions, while recombinant *Anaplasma* DnaK got eluted in the first two elution fractions.

3.4 Gel filtration results of recombinant proteins

Gel filtration was performed after Ni-NTA purification as mentioned in the methodology, and results are shown in Figures 7 and 8. The majority of His-tagged recombinant *Anaplasma* ClpB was collected between the 15th and 30th fractions, while the majority of His-tagged recombinant *Anaplasma* DnaK was collected between the 2nd and 35th fractions.

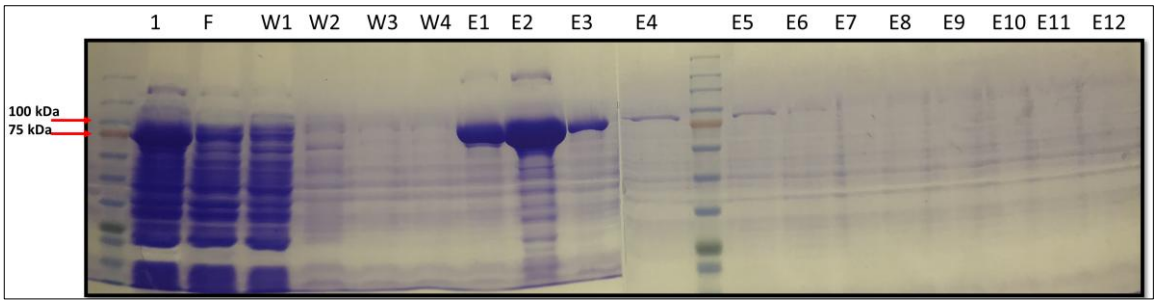


Figure 5. Ni-NTA purification for *Anaplasma* ClpB.
SDS-PAGE gel for *Anaplasma* ClpB (~95 kDa) after Ni-NTA purification is shown here.
1= input, F= Flow fraction, W1 to W3= Wash 1 to 3 fractions, E1 to E12= Elution1 to 12 fractions.

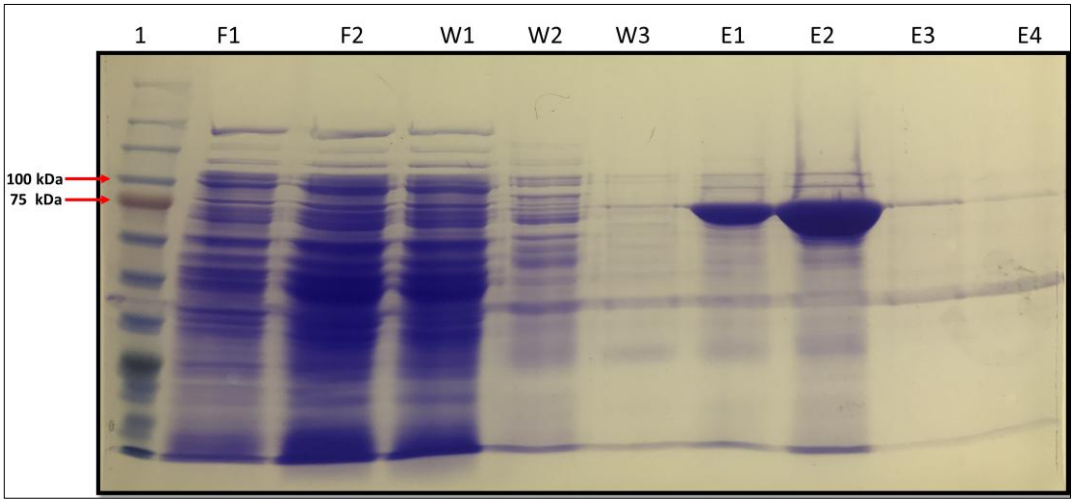


Figure 6. Ni-NTA purification for *Anaplasma* DnaK.
SDS-PAGE gel for *Anaplasma* DnaK (~70 kDa) after Ni-NTA purification is shown here.
1= Protein ladder, F1 to F2 = Flow fractions 1 and 2, W1 to W3= Wash 1 to 3 fractions, E1 to E4= Elution1 to 4 fractions.

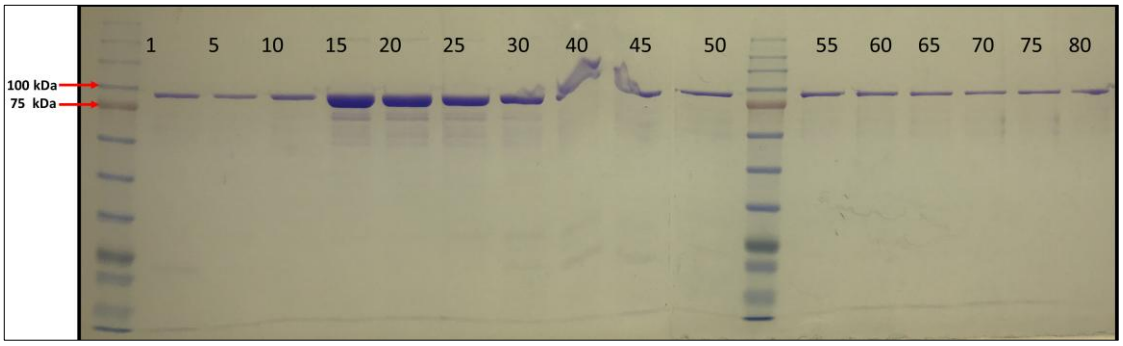


Figure 7. Gel filtration results for *Anaplasma* ClpB.
SDS-PAGE gel for *Anaplasma* ClpB (~95 kDa) after gel filtration using Superdex 200 resin is shown here.
Numbers correspond to fraction numbers.

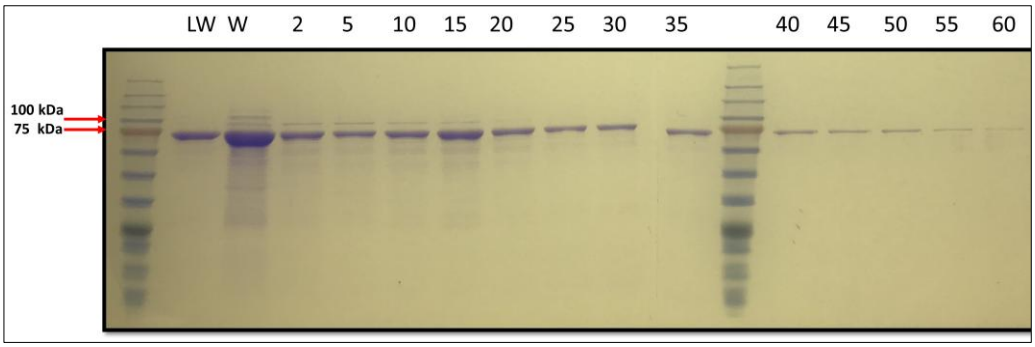


Figure 8. Gel filtration results for *Anaplasma* DnaK.
SDS-PAGE gel for *Anaplasma* DnaK (~70 kDa) after gel filtration using Superdex 200 resin is shown here. LW= Loading sample, W= Waste, Numbers correspond to fraction numbers.

3.5 ATPase assays (activity assays) of recombinant proteins

ATPase assays were performed as previously described by Glaza and colleagues (Glaza et al., 2021). It is well-documented that *E. coli* ClpB prefers casein as a model substrate, and its ATPase activity gets elevated in the presence of casein (Glaza et al., 2021; Barnett et al., 2000; Doyle et al., 2007; Deville et al., 2019). Similarly, bacterial DnaK also contains an ATPase domain that regulates its ATPase activity (McCarty and Walker 1991; Lee et al., 2018; Yu et al., 2018). Consistent with these findings, *Anaplasma* ClpB and DnaK ATPase activity also changed based on the availability of casein (Fig: 9).

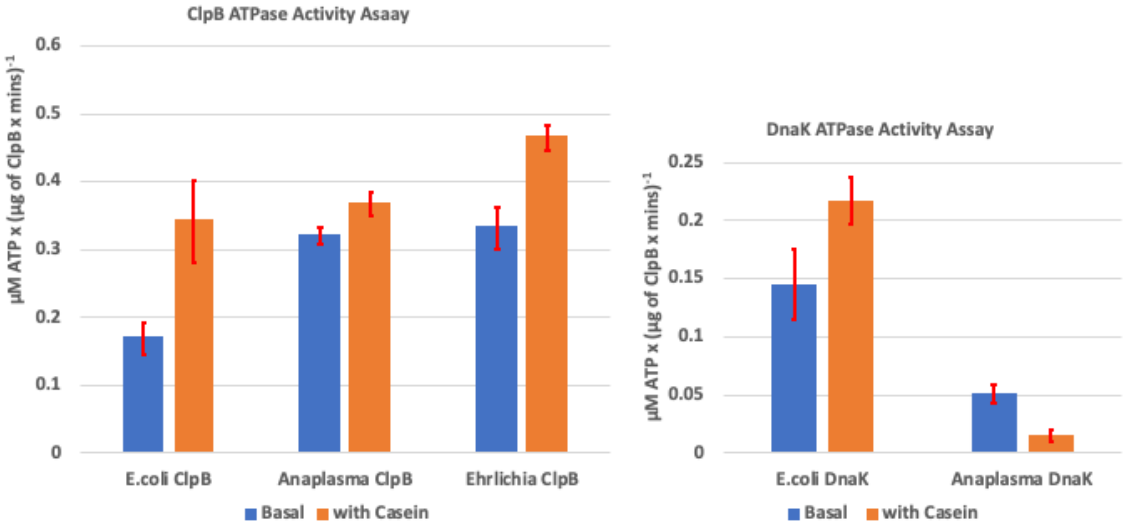


Figure 9: ATPase activity of purified recombinant *Anaplasma* ClpB and DnaK

.

Basal ATPase activity was measured without casein. *Ehrlichia chaffeensis* is an obligate intra cellular pathogen.

3.6 Analysis of the amino-acid sequence of ClpB and DnaK from Anaplasma phagocytophilum, Escherichia coli and Ehrlichia chaffeensis

Since there was a drastic drop in DnaK ATPase activity in the presence of kappa casein, we investigated the amino acid sequences of ClpB and DnaK to gain insight into the structure-function relationship of these two chaperones. AAA+ ATPases contain one or two ATP-binding AAA+ modules, known as Walker-A/B and sensor-1/2, as well as less conserved extra domains that mediate interactions with substrates or adaptors (Zhang *et al*., 2013). Sequence alignment of *Anaplasma* ClpB and DnaK with two other AAA+ proteins from two different organisms (supplementary figures 1 and 2) revealed that the *Anaplasma* ClpB and DnaK structure is similar to those of other AAA+ proteins and that all characteristic AAA+ motifs are present in the sequence of *Anaplasma* ClpB and DnaK. Hence, the significant change in DnaK ATPase activity may be linked to *Anaplasma*'s need to survive within host cells, which is often achieved by evading immune defenses and manipulating cellular processes of the host. Consequently, *Anaplasma* DnaK may have evolved additional or altered functions, such as interacting with host proteins, modulating immune responses, or even assisting protein folding under very different

environmental conditions than *E. coli*. While *E. coli* DnaK is optimized for general stress responses (mainly heat shock), *Anaplasma* DnaK might be less responsive to thermal stress but more attuned to oxidative or immune-related stressors within host cells.

4. Discussion

Ticks are now recognized as the second most prevalent vector of human infectious illnesses, trailing behind only mosquitoes. These obligate hematophagous arthropods are thought to parasitize nearly every class of vertebrates on the planet (Parola and Raoult, 2001). In recent years, stress and stress responses in host-pathogen interactions have opened new avenues for studying infectious disease development processes, the development of new antimicrobials, and discovering new antibiotic targets. Keeping this approach in sight, we initiated the first detailed biochemical characterization of two essential molecular chaperones from *A. phagocytophilum*.

Oligomer formation of Hsp100 proteins is linked to their chaperone activity, and oligomer formation is dictated by availability of ATP or ADP (Zolkiewski *et al*., 2012). Therefore, traditional purification of ClpB is carried out in the presence of a significant excess of ATP. This purification method is extremely expensive, more time-consuming, and inefficient. In this study, we demonstrate a process that can be completed without using ATP while retaining chaperone activity, making this the most cost-effective and efficient method described to date. Contamination with host genomic DNA has historically posed significant difficulty for researchers attempting to isolate the genomic DNA of obligate intracellular pathogens. This study presents an optimized protocol that enhances the isolation and separation of intracellular pathogen DNA from host cells.

We also investigated the ATPase activities of *Anaplasma* ClpB and *Anaplasma* DnaK. Both *Anaplasma* ClpB and *Ehrlichia* ClpB exhibited similar basal ATPase activities (Fig: 9); however, *E. coli* ClpB showed a two-fold lower basal ATPase activity. In the presence of kappa casein, all three ClpB variants showed enhanced ATPase activity. However, *E. coli* ClpB showed the greatest ATPase activity with kappa casein (2.1 times higher than basal activity), while *Anaplasma* ClpB showed only a marginal increase. Interestingly, our findings revealed that kappa casein significantly inhibited the ATPase activity of *Anaplasma* DnaK (3.4 times lower than basal activity), in contrast to the increasing effect observed for *E. coli* DnaK (1.5 times higher).

Our ATPase activity results for *Anaplasma* ClpB and DnaK demonstrate that these two chaperones do not respond to a pseudo-substrate in the same manner. As ClpB and DnaK had two opposite responses with kappa casein, we conducted an amino acid sequence analysis to gain more insight into the structure-function relationship of these chaperones. However, our sequence analysis showed that both ClpB and DnaK proteins of *A. phagocytophilum* exhibit highly conserved amino acid sequences in all signature motifs, which is characteristic of the AAA+ superfamily (Hanson and Whiteheart, 2005).

What could then be the reason/s for the above-mentioned differences in the substrate selectivity of *Anaplasma* ClpB and DnaK in different species? It has been firmly established that Hsp70/DnaK proteins act as the central hub of the protein homeostasis network in both bacteria and eukaryotes and are believed to be among the most important and abundant cellular proteins (Balchin *et al*., 2016). One could speculate that the substrate specificities of *Anaplasma* chaperones have evolved, possibly because of different stresses experienced by these bacteria in their natural habitat. Perhaps an obligate intracellular organism may need chaperones with distinct substrate recognition motifs compared to an extracellular organism like *E. coli*. Hence, *Anaplasma* chaperones might have evolved to meet specific demands of the pathogen's survival under host-induced stress.

5. Conclusions

The bi-chaperone system from an obligatory intracellular bacterium has not been previously reported. Studies on infection mechanisms have recently taken a novel direction toward better understanding the function of molecular chaperones in pathogen-host interactions. An effective method of limiting pathogen survival and infectivity may involve interfering with chaperone activity. Among the known molecular chaperones, Hsp100 proteins represent a particularly interesting target for antibiotics, as they are not produced in animal cells.

In this study, we report for the first time the successful cloning of ClpB and DnaK from *A. phagocytophilum*, as well as the purification of these recombinant chaperones under ATP-free conditions. This method enables a more cost-effective and efficient purification of biologically active chaperones. The ATPase assay results indicate that *Anaplasma* chaperones may have evolved to meet specific demands of the pathogen's survival under host-induced stress conditions.

Declaration of Competing Interest

The authors declare that they have no competing financial interests.

Acknowledgements

This work was performed as part of the Ph.D. thesis of C.B.R. titled "Bacterial AAA+ Disaggregase ClpB: Mechanism and inhibition" at Kansas State University,

USA.

We thank Dr. Przemyslaw Glaza, Dr. Yiva Wang, Dr. Arathy Nair, Dr. Chandramouli Kondethimmanahalli, Dr. Huitao Liu, Dr. Li Chen, and Claudia Clausi for their valuable suggestions and support throughout this project.

This study was supported by a grant from the National Institute of Health, USA AI121366 and AI141586 (to M.Z.) and the Kansas Agricultural Experiment Station, USA (contribution 18–623-J).

References

Atif, F. A. 2015. 'Anaplasma marginale and Anaplasma phagocytophilum: Rickettsiales Pathogens of Veterinary and Public Health Significance'. Parasitology Research 114 (11): 3941–57. https://doi.org/10.1007/s00436-015-4698-2.

Barnett, M. E., Zolkiewska, A., Zolkiewski, M. 2000. 'Structure and Activity of ClpB from Escherichia coli. Role of the Amino-and -Carboxyl-Terminal Domains'. The Journal of Biological Chemistry 275 (48):37565–71. https://doi.org/10.1074/jbc.M005211200.

Balchin, D., M. Hayer-Hartl and F. U. Hartl (2016). "In vivo aspects of protein folding and quality control." Science 353(6294): aac4354. DOI: 10.1126/science.aac4354.

Cheng, Chuanmin, Arathy D. S. Nair, Vijaya V. Indukuri, Shanzhong Gong, Roderick F. Felsheim, Deborah Jaworski, Ulrike G. Munderloh, and Roman R. Ganta. 2013. 'Targeted and Random Mutagenesis of Ehrlichia Chaffeensis for the Identification of Genes Required for In Vivo Infection'. PLoS Pathogens 9 (2): e1003171. https://doi.org/10.1371/journal.ppat.1003171.

Deville, Célia, Kamila Franke, Axel Mogk, Bernd Bukau, and Helen R. Saibil. 2019. 'Two-Step Activation Mechanism of the ClpB Disaggregase for Sequential Substrate Threading by the Main ATPase Motor'. Cell Reports 27 (12): 3433-3446.e4. https://doi.org/10.1016/j.celrep.2019.05.075.

Doyle, Shannon M., James Shorter, Michal Zolkiewski, Joel R. Hoskins, Susan Lindquist, and Sue Wickner. 2007. 'Asymmetric Deceleration of ClpB or Hsp104 ATPase Activity Unleashes Protein-Remodeling Activity'. Nature Structural & Molecular Biology 14 (2): 114–22. https://doi.org/10.1038/nsmb1198.

*Dumler, J. Stephen, Kyoung-Seong Choi, Jose Carlos Garcia-Garcia, Nicole S. Barat, Diana G. Scorpio, Justin W. Garyu, Dennis J. Grab, and Johan S. Bakken. 2005. 'Human Granulocytic Anaplasmosis and Anaplasma phagocytophilum'. Emerging Infectious Diseases 11 (12): 1828–34. https://doi.org/10.3201/eid1112.050898.

Glaza, Przemyslaw, Chathurange B. Ranaweera, Sunitha Shiva, Anuradha Roy, Brian V. Geisbrecht, Frank J. Schoenen, and Michal Zolkiewski. 2021. 'Repurposing P97 Inhibitors for Chemical Modulation of the Bacterial ClpB–DnaK Bichaperone System'. Journal of Biological Chemistry 296 (January). https://doi.org/10.1074/jbc.RA120.015413.

Hanson, P. I. and S. W. Whiteheart (2005). "AAA+ proteins: have engine, will work." Nature Reviews Molecular Cell Biology 6(7): 519-529. https://doi.org/10.1038/nrm1684.

Hartl, F. Ulrich, Andreas Bracher, and Manajit Hayer-Hartl. 2011. 'Molecular Chaperones in Protein Folding and Proteostasis'. Nature 475 (7356): 324–32. https://doi.org/10.1038/nature10317.

Laemmli, U. K. 1970. 'Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4'. Nature 227 (5259): 680–85. https://doi.org/10.1038/227680a0.

Lee, Changhan, Kamila B. Franke, Shady Mansour Kamal, Hyunhee Kim, Heinrich Lünsdorf, Jasmin Jäger, Manfred Nimtz, et al. 2018. 'Stand-Alone ClpG Disaggregase Confers Superior Heat Tolerance to Bacteria'. Proceedings of the National Academy of Sciences 115 (2): E273. https://doi.org/10.1073/pnas.1712051115.

McCarty, J. S., and G. C. Walker. 1991. 'DnaK as a Thermometer: Threonine-199 Is Site of Autophosphorylation and is Critical for ATPase Activity.' Proceedings of the National Academy of Sciences 88 (21): 9513–17. https://doi.org/10.1073/pnas.88.21.9513.

Morimoto, Richard I., and Ana M. Cuervo. 2009. 'Protein Homeostasis and Aging: Taking Care of Proteins From the Cradle to the Grave'. The Journals of Gerontology: Series A 64A (2): 167–70. https://doi.org/10.1093/gerona/gln071.

Parola, P., and D. Raoult. 2001. 'Ticks and Tickborne Bacterial Diseases in Humans: An Emerging Infectious Threat'. Clinical Infectious Diseases: 32 (6): 897–928. https://doi.org/10.1086/319347.

Ranaweera, Chathurange B., Przemyslaw Glaza, Taihao Yang, and Michal Zolkiewski. 2018. 'Interaction of Substrate-Mimicking Peptides with the AAA+ ATPase ClpB from Escherichia coli'. Archives of Biochemistry and Biophysics 655 (October): 12–17. https://doi.org/10.1016/j.abb.2018.08.002.

Ranaweera, Chathurange Bharathee. 2021. 'Bacterial AAA+ Disaggregase ClpB: Mechanism and Inhibition', Diss. 2021. https://krex.k-state.edu/dspace/handle/2097/41435

Yu, Hongjun, Tania J. Lupoli, Amanda Kovach, Xing Meng, Gongpu Zhao, Carl F. Nathan, and Huilin Li. 2018. 'ATP Hydrolysis-Coupled Peptide Translocation Mechanism of Mycobacterium tuberculosis ClpB'. Proceedings of the National Academy of Sciences 115 (41): E9560. https://doi.org/10.1073/pnas.1810648115.

Zhang, Ting, Sabina Kedzierska-Mieszkowska, Huitao Liu, Chuanmin Cheng, Roman R. Ganta, and Michal Zolkiewski. 2013. 'Aggregate-Reactivation Activity of the Molecular Chaperone ClpB from Ehrlichia

chaffeensis'. PLoS ONE 8 (5): e62454. https://doi.org/10.1371/journal.pone.0062454.

Zhang, Yuntao, Li Chen, Chandramouli Kondethimmanahalli, Huitao Liu, and Roman R. Ganta. 2021. 'Protein and DNA Biosynthesis Demonstrated in Host Cell-Free Phagosomes Containing Anaplasma phagocytophilum or Ehrlichia chaffeensis in Axenic Media'. Infection and Immunity, January. https://doi.org/10.1128/IAI.00638-20.

Zolkiewski, Michal, Ting Zhang, and Maria Nagy. 2012. 'Aggregate Reactivation Mediated by the Hsp100 Chaperones'. Archives of Biochemistry and Biophysics 520 (1): 1–6. https://doi.org/10.1016/j.abb.2012.01.012.

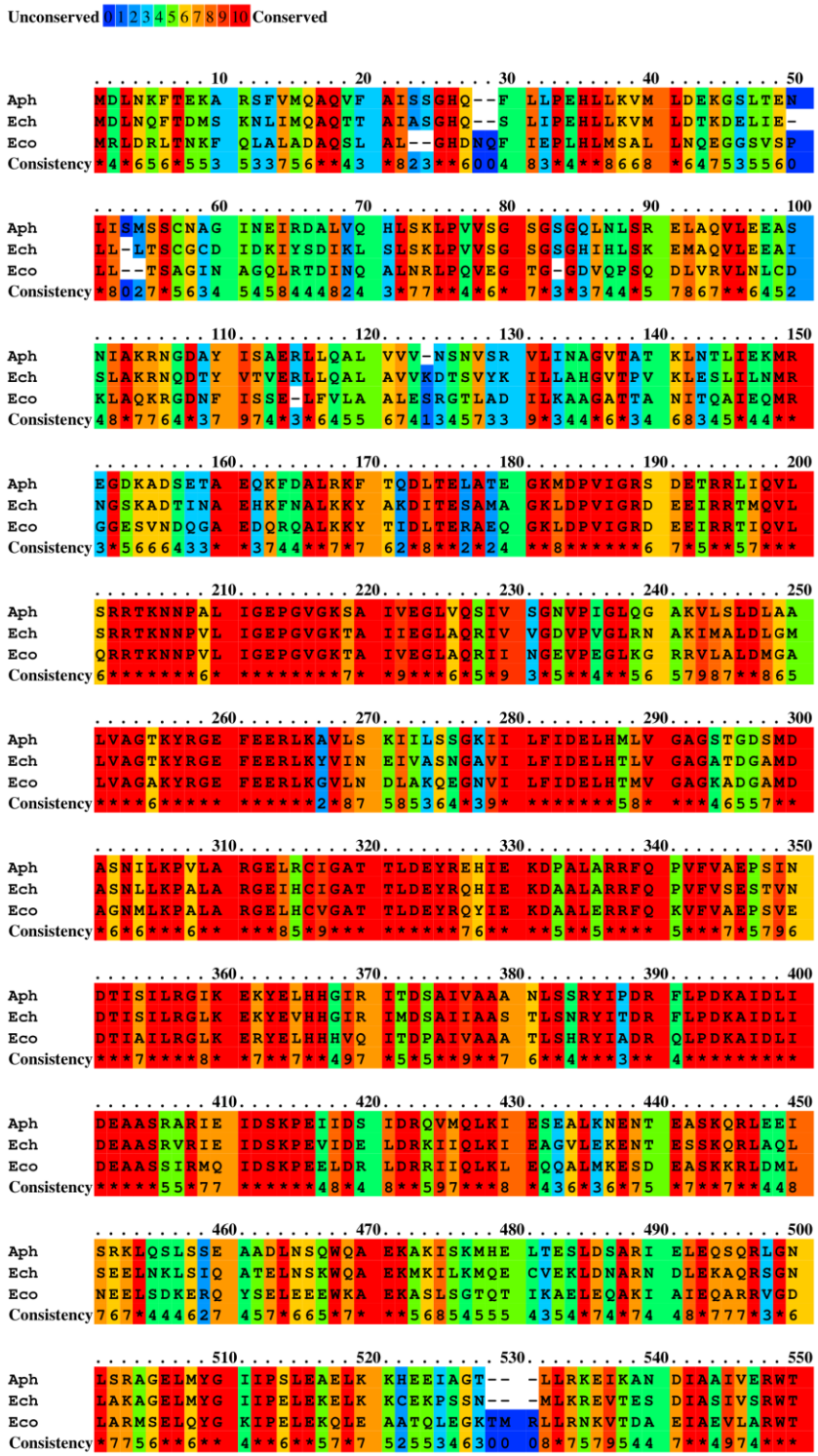
Supplementary Figure 1

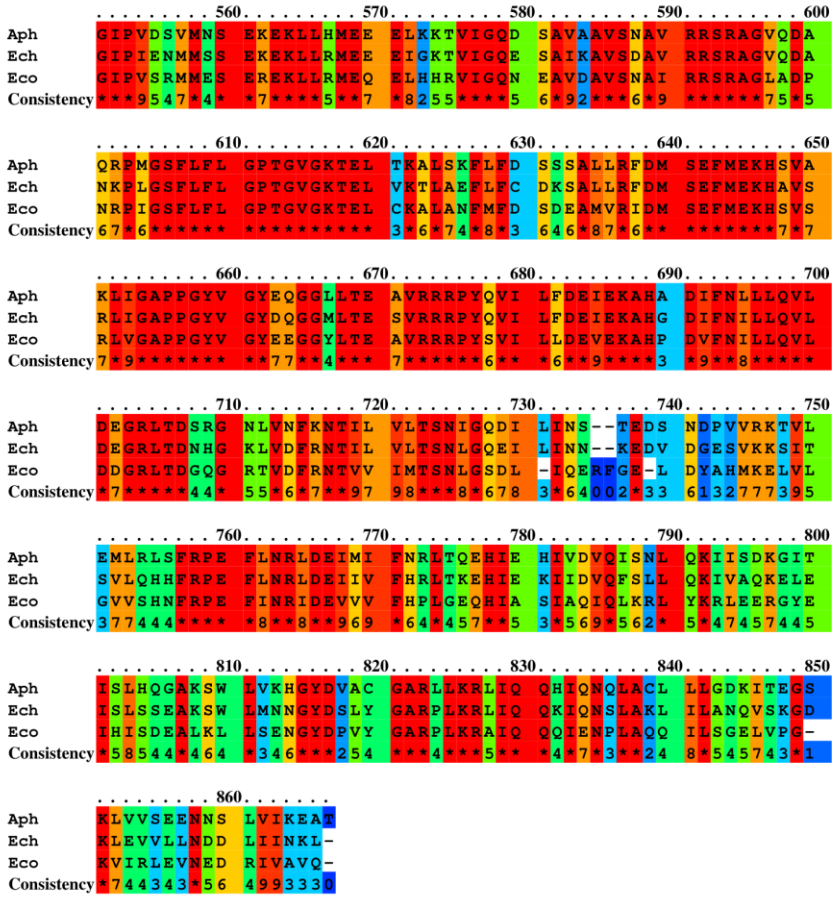
Amino acid sequence analysis results for ClpB with colour coded for amino acid conservation

The current colour scheme of the alignment is for amino acid conservation. The conservation scoring is performed by the PRALINE multiple sequence alignment program. The scoring scheme works from 0 for the least conserved alignment position, up to 10 for the most conserved alignment position.

Aph= *Anaplasma phagocytophilum*. Ech= *Ehrlichia chaffeensis*. Eco= *Escherichia coli*.

The colour assignments are:



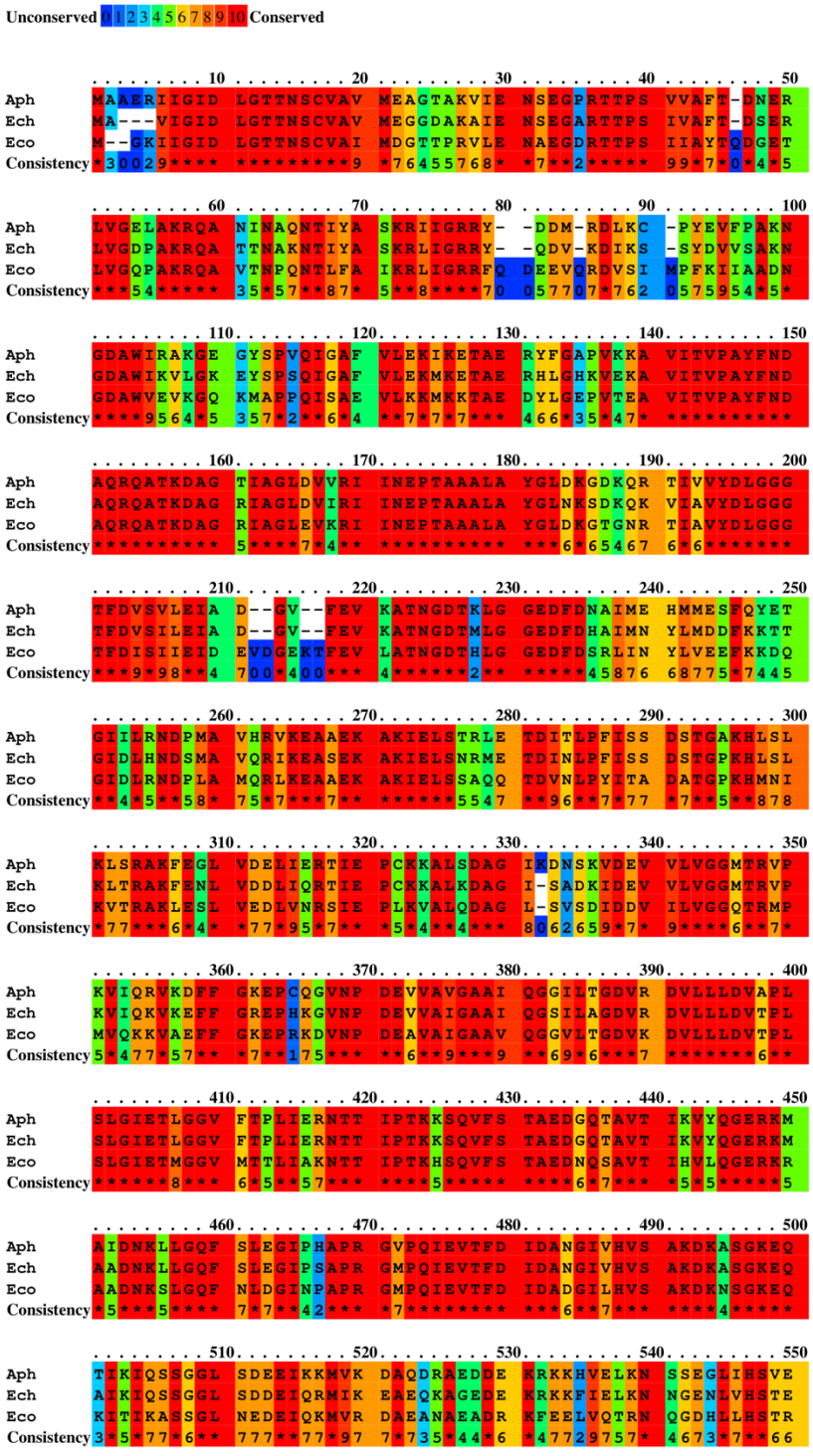

560 570 580 590 600
Aph GIPVDSVMNS EKEKLLHMEE ELKKTVIGQD SAVAAVSNAV RRSRAGVQDA
Ech GIPIENMMSS EKEKLLRMEE EIGKTVIGQE SAIKAVSDAV RRSRAGVQDA
Eco GIPVSRMMES EREKLLRMEQ ELHHRVIGQN EAVDAVSNAI RRSRAGLADP
Consistency ***954*4* *7****5**7 *8255****5 6*92***6*9 ******75*5
610 620 630 640 650
Aph QRPMGSFLFL GPTGVGKTEL TKALSKFLFD SSSALLRFDM SEFMEKHSVA
Ech NKPLGSFLFL GPTGVGKTEL VKTLAEFLFC DKSALLRFDM SEFMEKHAVS
Eco NRPIGSFLFL GPTGVGKTEL CKALANFMFD SDEAMVRIDM SEFMEKHSVS
Consistency 67*6****** ********** 3*6*74*8*3 646*87*6** *******7*7
660 670 680 690 700
Aph KLIGAPPGYV GYEQGGLLTE AVRRRPYQVI LFDEIEKAHA DIFNLLLQVL
Ech RLIGAPPGYV GYDQGGMLTE SVRRRPYQVI LFDEIEKAHG DIFNILLQVL
Eco RLVGAPPGYV GYEEGGYLTE AVRRRPYSVI LLDEVEKAHP DVFNILLQVL
Consistency 7*9******* **77**4*** 7******6** *6**9****3 *9**8*****
710 720 730 740 750
Aph DEGRLTDSRG NLVNFKNTIL VLTSNIGQDI LINS--TEDS NDPVVRKTVL
Ech DEGRLTDNHG KLVDFRNTIL VLTSNLGQEI LINN--KEDV DGESVKKSIT
Eco DDGRLTDGQG RTVDFRNTVV IMTSNLGSDL -IQERFGE-L DYAHMKELVL
Consistency *7*****44* 55*6*7**97 98***8*678 3*64002*33 6132777395
760 770 780 790 800
Aph EMLRLSFRPE FLNRLDEIMI FNRLTQEHIE HIVDVQISNL QKIISDKGIT
Ech SVLQHHFRPE FLNRLDEIIV FHRLTKEHIE KIIDVQFSLL QKIVAQKELE
Eco GVVSHNFRPE FINRIDEVVV FHPLGEQHIA SIAQIQLKRL YKRLEERGYE
Consistency 377444**** *8**8**969 *64*457**5 3*5695*562* 5*474574 45
810 820 830 840 850
Aph ISLHQGAKSW LVKHGYDVAC GARLLKRLIQ QHIQNQLACL LLGDKITEGS
Ech ISLSSEAKSW LMNNGYDSLY GARPLKRLIQ QKIQNSLAKL ILANQVSKGD
Eco IHISDEALKL LSENGYDPVY GARPLKRAIQ QQIENPLAQQ ILSGELVPG-
Consistency *5854*4644 *346***254 ***4***5** *4*7*3**24 8*545743*1
860
Aph KLVVSEENNS LVIKEAT
Ech KLEVVLLNDD LIINKL-
Eco KVIRLEVNED RIVAVQ-
Consistency *744343*56 4993330

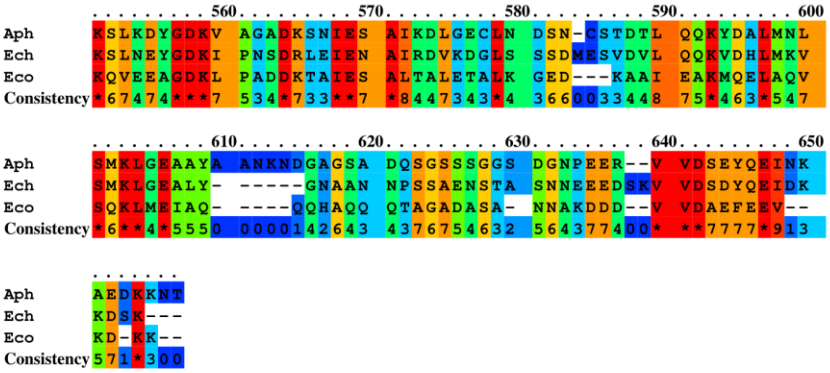
Amino acid sequence analysis results for DnaK with colour coded for amino acid conservation

The current colour scheme of the alignment is for amino acid conservation. The conservation scoring is performed by the PRALINE multiple sequence alignment program. The scoring scheme works from 0 for the least conserved alignment position, up to 10 for the most conserved alignment position.

Aph= *Anaplasma phagocytophilum*. Ech= *Ehrlichia chaffeensis*. Eco= *Escherichia coli*.

The colour assignments are:

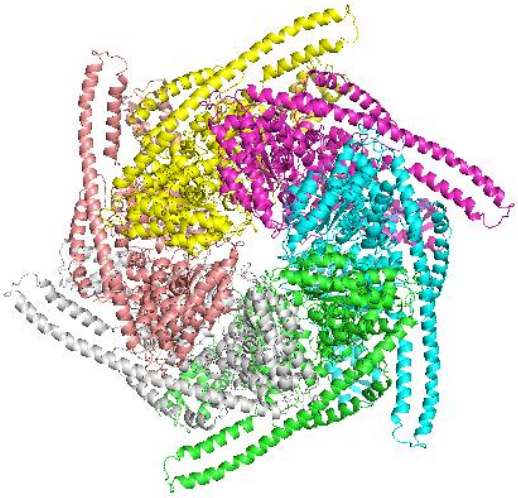



560 570 580 590 600
Aph KSLKDYGDKV AGADKSNIES AIKDLGECLN DSN-CSTDTL QQKYDALMNL
Ech KSLNEYGDKI PNSDRLEIEN AIRDVKDGLS SSDMESVDVL QQKVDHLMKV
Eco KQVEEAGDKL PADDKTAIES ALTALETALK GED---KAAI EAKMQELAQV
Consistency *67474***7 534*733**7 *8447343*4 366003448 75*463*547
610 620 630 640 650
Aph SMKLGEAAYA ANKNDGAGSA DQSGSSSGGS DGNPEER--V VDSEYQEINK
Ech SMKLGEALY- -----GNAAN NPSSAENSTA SNNEEEDSKV VDSDYQEIDK
Eco SQKLMEIAQ- ----QQHAQQ QTAGADASA- NNAKDDD--V VDAEFEEV--
Consistency *6**4*5550 0000142643 4376754632 564377400* **7777*913
Aph AEDKKNT
Ech KDSK---
Eco KD-KK--
Consistency 571*300

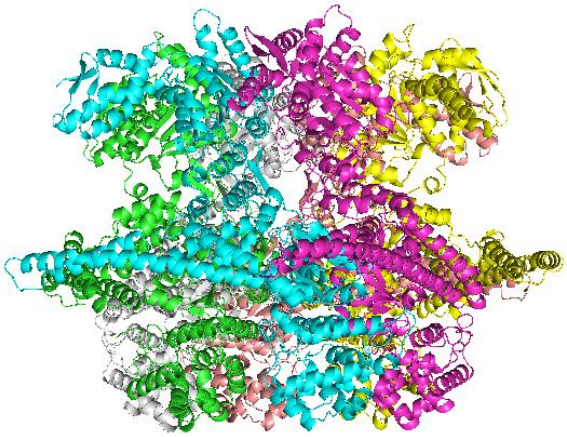
Supplementary Figure 2

***E.coli* ClpB and DnaK structures**

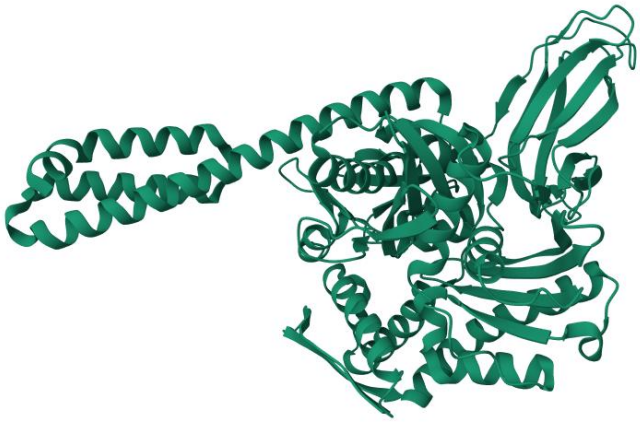
The top (a) and side view (b) of the active hexameric ClpB structure are shown here. Each ClpB monomer is coloured individually for representation.
Side view of monomeric DnaK (c) is shown in light green colour. The hexamer structure was obtained by assembling six monomers into a ring and performing a 500-step energy minimization with CHARMM and Generalized Born solvent model (Zolkiewski, Zhang, and Nagy 2012).



(a) Top view of ClpB hexamer



(b) Side view of ClpB hexamer



(c) Side view of DnaK. Generated from PDB: 4B9Q.

Kityk, R., Kopp, J., Sinning, I. and Mayer, Matthias P. (2012). Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. *Molecular Cell*, 48(6), pp.863–874. doi:https://doi.org/10.1016/j.molcel.2012.09.023.