Study on Caseinolytic protease chaperones and its cognate adaptors of *Leptospira interrogans*

A thesis submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy

by

Surbhi Kumari

Under the supervision of

Prof. Manish Kumar

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Guwahati, 781039, Assam, India

TH-3792_206106022

Indian Institute of Technology Guwahati

Department of Biosciences and Bioengineering

Guwahati-781039, Assam, India

Declaration

I do hereby declare that the matter embodied in this thesis entitled "**Study on Caseinolytic protease chaperones and its cognate adaptors of *Leptospira interrogans***" is the result of work carried out under the supervision of Prof. Manish Kumar, Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati. In keeping with the general practice of reporting scientific observations, due acknowledgement has been made wherever the work described is based on the findings of other investigators. The work illustrated in this thesis has not been submitted elsewhere for any other degree.

Surbhi Kumari

Roll no. 206106022

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Guwahati-781039, Assam, India

TH-3792_206106022

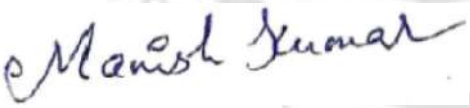
Indian Institute of Technology Guwahati

Department of Biosciences and Bioengineering

Guwahati-781039, Assam, India

Certificate

It is certified that the work described in this thesis entitled "**Study on Caseinolytic protease chaperones and its cognate adaptors of *Leptospira interrogans***" by **Mrs Surbhi Kumari (Roll no. 206106022)** for the award of the degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision in the Department of Biosciences and Bioengineering, IITG. The work embodied in this thesis has not been submitted elsewhere for a degree.



17 Nov 2025

Prof. Manish Kumar

(Thesis supervisor)

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Guwahati-781039, Assam, India

TH-3792_206106022

I dedicate this thesis to

my parents, in-laws, husband, sister and friends

TH-3792_206106022

Acknowledgment

First and foremost, I humbly acknowledge the grace of God, who has blessed me with opportunities and guided my instincts in pursuing my career path.

I am deeply grateful to Prof. Manish Kumar for giving me the opportunity to work in his lab and for his invaluable guidance during the formative years of my research. The invaluable scientific insights and life lessons I have gained will undoubtedly benefit my future pursuits.

I sincerely thank the esteemed members of my doctoral committee, Dr. Shirisha Nagotu, Dr. Priyadarshi Satpati and Prof. Debasis Manna for their invaluable insights and timely suggestions. Their guidance has significantly enhanced the quality of this study, and I deeply appreciate their support throughout the course of my research.

I am deeply grateful to Prof. Utpal Bora, Head of the Department, as well as former HODs, Prof. Latha Rangan and Prof. Rakhi Chaturvedi, for providing the departmental facilities that enabled me to conduct my research. I am indebted to the Department of Biosciences and Bioengineering and the Central Instrumentation Facility at IIT Guwahati for furnishing the essential research infrastructure crucial for achieving the objectives of my PhD thesis. My heartfelt gratitude goes out to all the staff members of department for their indispensable logistical support, which facilitated the progression of my research endeavors. Furthermore, I extend my appreciation to IIT Guwahati, the Ministry of Human Resources Development, India, and the Department of Biotechnology, Government of India, for their financial support. Department of Biotechnology (DBT) and Department of Science and Technology (DST) for providing a research grant to our laboratory.

I extend my deepest appreciation to all my past and present lab colleagues of MK lab: Dr. Anusua Dhara, Dr Md. Saddam Hussain, Dr Vineet Anand, Dr Saswat Hota, Pratibha Koundal, Prattusha Khan, Shibam Dey, Madhurima Bhowmik, Sawna Roy, Tania Sarkar, Kumari Anjali and Kiran Kumari for their constant support and assistance throughout my research journey. Specifically, I would like to acknowledge the help provided by my lab seniors, Dr Anusua Dhara (Generation of ClpP mutants), Mr Dipankar Dutta (Cloning of clpC, clpA, clpS1 and clpS2) and Dr. Md Saddam Hussain (polyclonal antibody generation in mice). Additionally, I am deeply thankful to my dear friends Swagatika Dash, Pratik Dasgupta,

TH-3792_206106022

Debolina Ghosh, Shubhangini Singh, and Sumona Koley, who stood by me through both the highs and lows and made this journey of five years enjoyable, filled with joys and sorrows.

Finally, I am ever grateful to my parents Mr. Ashok Kumar Singh and Mrs. Mandavi Devi, my elder sister, Mrs Priti kumari and my younger brother, Mr Amit Kumar, for their love and constant mental support during the hard times. This acknowledgement will be incomplete if I do not mention the role of my life partner, Mr. Manikant Singh, and my in-laws for strengthening my spirit and letting me complete my Ph.D. without any hindrances.

Additionally, I express my appreciation to all others whom I may have inadvertently omitted.

Surbhi Kumari

TH-3792_206106022

Table of contents

TH-3792_206106022

TH-3792_206106022

TH-3792_206106022

TH-3792_206106022

TH-3792_206106022

List of figures

TH-3792_206106022

TH-3792_206106022

Supplementary Figures

TH-3792_206106022

List of tables

TH-3792_206106022

ABSTRACT

Bacterial Caseinolytic protease (Clp) systems are essential for degrading misfolded and aggregated proteins, ensuring cellular protein homeostasis. These complexes comprise a proteolytic core, ClpP, formed by two stacked heptameric rings, and an ATP-dependent chaperone (ClpA, ClpC or ClpX), which assembles into a hexameric ring. The chaperone recognizes, unfolds, and translocates substrates into the ClpP chamber using energy from ATP hydrolysis. Substrate specificity is tightly regulated to prevent the degradation of essential proteins. While chaperones can directly identify specific substrates, many require adaptor proteins that facilitate substrate recognition and delivery to the protease complex, thereby enhancing selectivity and regulatory control in proteolysis. The spirochete *Leptospira interrogans* possesses a set of genes encoding core proteolytic components (*clpP1* and *clpP2*), ATPase chaperones (*clpA*, *clpC*, *clpX*), and adaptor proteins (*clpS1* and clpS2). The core proteolytic component of the Clp system in leptospires is composed of two different ClpP isoforms (LinClpP1 and LinClpP2), forming a hetero-tetradecameric assembly (ClpP1P2). The heterocomplex (LinClpP1P2) degrades small peptide substrates (2-4 amino acids) alone, and for degrading protein model substrates (casein), it requires association with the ATPase chaperone, LinClpX. This study aims to identify and study the role of key structural motifs of LinClpP1 and LinClpP2 isoforms. Further, we have explored the functionality of two ATPase chaperones (LinClpC and LinClpA) and their association with core proteolytic components (LinClpP) and adaptor proteins (LinClpS).

This study investigates the functional dynamics of five mutants (LinClpP1^{E170D}, LinClpP1^{N172D}, LinClpP2^{IG_del}, LinClpP2^{S40AK41N}, LinClpP2^{Y62A}) targeting key structural motifs. LinClpP2^{S40AK41N} and LinClpP2^{Y62A} showed a gain in peptidase activity independent of the LinClpP1 isoform and also exhibited greater hydrodynamic diameter than the LinClpP2 isoform. The natural antibiotic, ADEP1, binding to the heterocomplex (LinClpP1P2^{S40AK41N} and LinClpP1P2^{Y62A}) showed increased activity (1.7- and 1.5-fold) than LinClpP1P2. Further, the heterocomplex (LinClpP1P2^{S40AK41N} and LinClpP1P2^{Y62A}) exhibited a 3-fold surge in leptospiral trigger factor (LinTF) presence. In contrast, the heterocomplex (LinClpP1^{E170D}P2 and LinClpP1^{N172D}P2 displayed 0.5-fold lower peptidase activity than the LinClpP1P2, while the LinClpP1P2^{IG_del} showed loss of activity. These findings uncover critical residues governing LinClpP activation, revealing novel mechanisms of protease regulation in pathogenic leptospires. The LinClpC, a class-I chaperone with two ATPase domains,

TH-3792_206106022

undergoes nucleotide-induced oligomerization into a higher molecular weight complex. The LinClpC associates non-specifically with both LinClpP isoforms (ClpP1 or ClpP2) with equal affinity. The LinClpC associates with the LinClpP1P2 heterocomplex to form an active protease that degrades FITC-casein even without ATP. Adding ATP or ATPγS enhances protease activity by promoting LinClpC oligomerization. Additionally, ADEP1 is proposed as a more potent activator of LinClpP1P2 activity than LinClpC. Further, this study investigates the structural and functional characteristics of Clp regulatory components: LinClpA, LinClpS1, and LinClpS2. LinClpA is a 740-residue protein comprising an N-terminal domain and two AAA+ ATPase domains (D-I and D-II) with conserved ATP-binding and hydrolysis motifs. The LinClpS1 and LinClpS2, while structurally similar, differ in their binding pockets for the N-terminal domain of LinClpA and the N-degron-tagged proteins. Functional assays demonstrate that a variant of LinClpA lacking the N-terminal domain (LinClpA$^{\Delta N}$) still undergoes nucleotide-induced oligomerization, akin to the full-length protein, but shows increased ATPase activity. Interaction analyses revealed that LinClpA's ATPase activity is upregulated in the presence of LinClpP1 and LinClpP2, while inhibited when bound by LinClpS1 or LinClpS2. Interestingly, LinClpA$^{\Delta N}$ does not respond to these adaptor proteins, emphasizing the importance of the N-terminal domain of LinClpA in adaptor protein engagement. Furthermore, this study explores the functional significance of the SsrA tag (a C-terminal degradation signal; C-degron) in promoting substrate recognition and breakdown by the LinClpAP1P2 proteolytic machinery. The findings from our study shed light on the new insight into the structural reorganization and regulatory interactions of Clp components in *L. interrogans*, enhancing our understanding of bacterial proteostasis and offering potential targets for antimicrobial development.

TH-3792_206106022

Chapter 1

Overview of bacterial protein quality control systems (Caseinolytic proteases) and introduction to leptospirosis

TH-3792_206106022

1. Introduction

The survival of a cell relies on the coordinated interactions of proteins to perform various cellular processes. The newly synthesized proteins released as a linear chain of amino acids from ribosomes must fold into well-defined three-dimensional structures to attain cellular functionality (Ellis et al., 2006). Although some newly synthesized proteins can fold spontaneously into their native states, some exhibit lower folding efficiency and are susceptible to misfolding. Such proteins depend on several molecular chaperones and metabolic energy to efficiently attain their native structures (Frydman, 2001; Hartl & Hayer, 2002). The emerging polypeptides have their hydrophobic amino acid residues exposed in the cytosol, and thus, to prevent their aggregation and misfolding, the newly released chains are safeguarded by several chaperones (Dobson et al., 1998). Molecular chaperones were described as folding modulators that help other proteins reach their native conformation. Based on their action mechanisms, the chaperones are categorized into three families (holdases, foldases, and disaggregases) (Fatima et al., 2021). The holding chaperones (holdases), such as inclusion body-associated protein B (IbpB) and trigger factor (TF), stabilize the protein in their partially folded or misfolded state in an energy-independent manner (Veinger et al., 1998; Baneyx & Mujacic, 2004). The folding chaperones (foldases) such as growth essential large protein (GroEL), Caseinolytic peptidase A (ClpA), and DnaK drive the refolding or unfolding of poorly folded, damaged, or unwanted proteins in an energy-dependent manner (Baneyx & Mujacic, 2004). The disaggregases, such as Caseinolytic peptidase B (ClpB), facilitate the solubilization of protein aggregates (Mróz et al., 2020; Kim et al., 2013).

1.1. The Molecular Chaperone

The molecular chaperone aids in properly folding or assembling a protein without being a part of its folded functional structure. The synthesis of molecular chaperones is triggered under several environmental stress conditions (such as heat or oxidative stress); thus, they are categorized as stress proteins or Heat shock proteins (Hsps). The Hsp chaperones are named based on their molecular weight, such as Hsp40, Hsp60, Hsp70, Hsp90, Hsp100, and small Hsp (Kim et al., 2013). In prokaryotes, the coupling of translation with protein folding makes it a complex process. During translation, the exit tunnel of ribosomes restricts the folding of nascent polypeptide chains beyond small helices until a complete protein domain is synthesized (O'Brien et al., 2010; Wilson & Beckmann, 2011). The molecular chaperone TF binds at the exit tunnel of ribosomes and interacts with emerging nascent polypeptide chains of 100 amino

TH-3792_206106022

acid length (Ferbitz et al., 2004). The ribosome-associated TF is the first to encounter and bind to the hydrophobic segments of newly emerging polypeptide chains (Kaiser et al., 2006) (**Figure 1.1**). The TF assists in folding nascent chains and permits the handover of polypeptide chains to the downstream chaperones, such as DnaK, a Hsp70 protein family, in an ATP-independent manner (Kim et al., 2013). The TF inhibits the binding of DnaJ/K to nascent polypeptide chains of length up to 250 amino acids, thereby promoting the preferential interaction of DnaK with protein substrates exceeding 30 kDa *in vivo* (Roeselová et al., 2024). The Hsp70 holds a significant affinity for many newly synthesized proteins but does not associate with the ribosomes (Calloni et al., 2012). The Hsp70 chaperones (DnaK) work with a cochaperone of the Hsp40 family (DnaJ proteins) (Kampinga & Craig 2010). The DnaK-DnaJ complex binds to longer polypeptide chains and protects them from aberrant interaction, leading to misfolding or protein aggregation. They mediate the co- or post-translational folding of the nascent chain through the ATP binding and hydrolysis cycle and permit the transfer of partially folded proteins to the downstream chaperones (Kim et al., 2013).

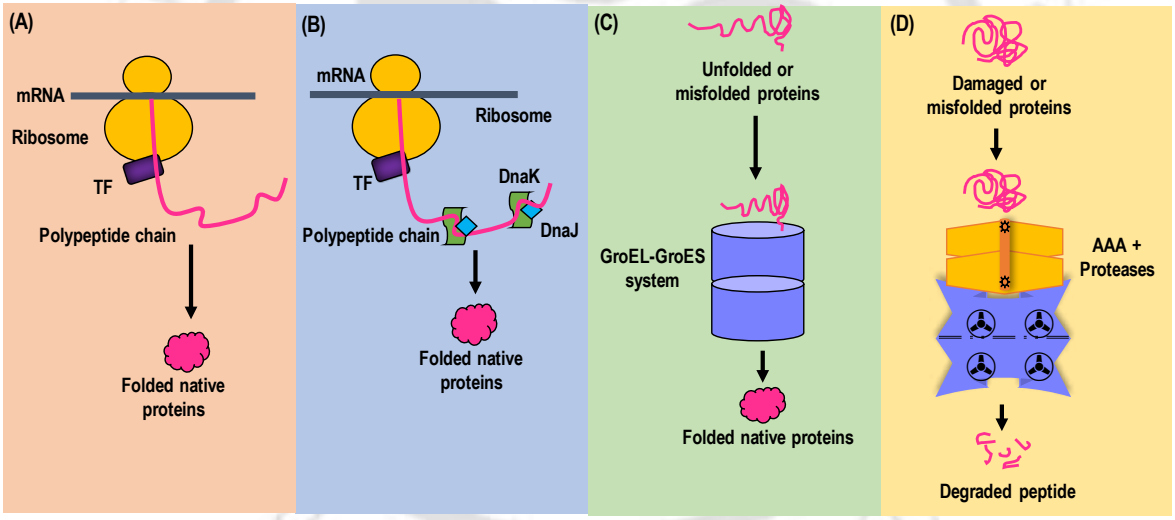


Figure 1.1: Schematic diagram showing chaperone-assisted folding of nascent polypeptides in the bacterial cytosol. (A) The ribosome-bound trigger factor (TF) assists folding of smaller polypeptide chains (65-80%) in an ATP-independent manner. **(B)** Longer polypeptide chains (10 to 20% of total) associate with chaperones (DnaK and DnaJ) for folding utilizing energy driven from ATP hydrolysis. **(C)** About 10 to 15% of larger polypeptides go to the chaperonin system—GroEL-GroES complex. The misfolded or multi-domain proteins are handed over to the GroEL-GroES complex from other chaperones like DnaK for proper folding. **(D)** The terminally damaged or unwanted proteins are removed from the cell using several AAA+ proteases to maintain proteostasis (adapted from Kim et al., 2013).

Proteins that cannot associate with Hsp70 are directed to the chaperonins or the Hsp90 system. The chaperonins are large molecular weight complexes (800-1000 kDa) comprising a double-ring complex with a central cavity. The central cavity of chaperonins provides an isolated

TH-3792_206106022

compartment for the proper folding of proteins and, thereby, bypasses the aggregation-prone cytosol (Bukau & Horwich, 1998). The chaperonins, such as GroEL, are associated with a cochaperone, GroES, which acts as a lid to protect protein substrates within the central cavity of GroEL. The complex multi-domain proteins are majorly the substrates of the GroEL-GroES system and comprise 10-15% of newly synthesised proteins (Fujiwara et al., 2010). Additionally, the Hsp70 system coordinates with AAA+ (ATPase associated with diverse cellular activities) chaperones such as Hsp100 and helps dissociate and refold aggregated proteins (Kim et al., 2013). The truncated or misfolded proteins beyond repair are delivered to the protein degradation machinery, such as the Caseinolytic peptidase P (ClpP) complex (Kim et al., 2013).

1.1.1. Trigger Factor (TF)

In bacteria, TF is an abundant protein that binds to the ribosomes and assists in the folding of newly synthesized proteins (Kaiser et al., 2006; Preissler & Deuerling, 2012). In *Escherichia coli*, TF is the primary chaperone that binds to the newly released proteins from the ribosomes (Craig et al., 2003). In *E. coli*, TF is one of the abundant proteins with 20,000 copies/cell (Crooke et al., 1988). The 432-amino acid long (48 kDa) *E. coli* TF (EcoTF) has three structural domains, namely the N-terminal domain (tail), the middle domain (head), and the C-terminal domain (arms) as shown in **Figure 1.2A**. The crystal structure of EcoTF reveals an elongated dragon-like structure upon folding (Ferbitz et al., 2004). The N-terminal domain (1-149 residues) contains a motif with the helix-loop-helix structural arrangement and consensus sequence (Gly-Phe-Arg-X-Gly-X-X-Pro), which is known as the TF-signature motif that docks at the large ribosomal subunit (Schlünzen et al., 2005; Kramer et al., 2002). The middle domain (150- 245 residues) possesses peptidyl-propyl cis/trans isomerase (PPIase activity) (Hesterkamp et al., 1997). It forms the head part of the structure and aligns away from the ribosome (Li et al., 2019). The CTD (246-432 residues) is the largest domain and possesses chaperone activity. The CTD forms the central part of the TF structure and has two protruding arms (arm1 and arm2), which facilitate the holding of the protein substrates. The linker segment (111-133 residues) between the NTD and PPIase domain folds so that the CTD aligns in the middle of the EcoTF three-dimensional structure (Ferbitz et al., 2004). The CTD holds great importance in EcoTF activity as its absence has impaired the refolding of denatured protein (Zeng et al., 2006).

TH-3792_206106022

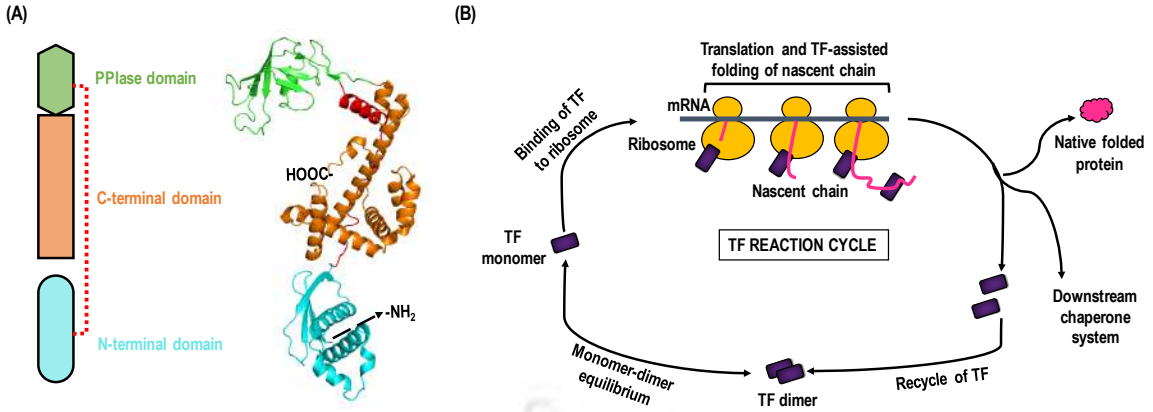


Figure 1.2. Structural and functional analysis of trigger factor (TF). (A) Crystal structure of *E. coli* TF (PDB: 1W26). The N-terminal domain (cyan), peptidylprolyl cis/trans isomerase (PPIase) domain (green), and the C-terminal domain (orange) are marked. The linker segment between the NTD and PPIase domain is marked in red colour. **(B)** Folding assistance by TF is shown in a cyclic reaction. The dimeric and monomeric states of TF are in equilibrium inside the cytosol. The monomeric form binds to the ribosome and interacts with nascent chains via their hydrophobic segments. Upon binding, TF assists in nascent chain folding, followed by its release from the ribosomes (adapted from Kim et al., 2013).

The monomeric EcoTF binds to the large ribosomal subunit with ribosomal proteins (L23 and L29), while the fraction of unbound EcoTF exists in a dimeric state (Ferbitz et al., 2004; Kramer et al., 2002). The L23 protein supports the newly synthesized polypeptide chain progression from the ribosomal exit channel and also binds to TF via hydrophobic interaction (Baram et al., 2005). The cellular environment in an *E. coli* cell maintains a 2-3 times excess molar concentration of TF than ribosomes (Crooke et al., 1988). The non-ribosome-bound TF are in continuous monomer-dimer equilibrium (Patzelt et al., 2002). The TF aids in the folding and proper assembly of nascent polypeptide chains upon a cyclic process of binding and release to the ribosome and the nascent chains in an ATP-independent TF reaction cycle (**Figure 1.2B**) (Kim et al., 2013). As the monomeric TF associates with the ribosome, it extends its CTD region to engage with the hydrophobic regions of the translating polypeptide chain. During the folding process, the nascent polypeptide arranges itself to bury the hydrophobic regions and drives the release of TF from the ribosome with a half-life of 10s (Kaiser et al., 2006). The released TF results in either folding of nascent polypeptide into its native state or handing over to other chaperones for further folding assistance (Kim et al., 2013).

1.1.2. The Hsp70 System

The Hsp70 proteins are abundant among the chaperone family and are found in bacteria, eukaryotic cytosol, organelles (ER, mitochondria, chloroplasts), and archaea (Powers & Balch,

TH-3792_206106022

2013). They assist in various cellular processes, including nascent protein folding, refolding stress-damaged proteins, protein transport, membrane translocation, and protein degradation. The Hsp70 binds to the exposed hydrophobic regions of unfolded or misfolded proteins to prevent their aggregation and acts on them through ATP binding and hydrolysis for proper folding. They work in association with their obligate co-chaperone partner, Hsp40 (DnaJ protein) (Kampinga & Craig, 2010).

The Hsp70 (DnaK protein) has two domains at the N-terminal and C-terminal ends, respectively: nucleotide-binding domain (NBD; 45 kDa) and substrate-binding domain (SBD; 25 kDa) (Kim et al., 2013). The NBD comprises two lobes (I and II), each with two subdomains (lobe I-IA and IB; lobe II- IIA and IIB) as shown in **Figure 1.3A**. The two lobes assemble to create a deep cleft for MgATP/MgADP binding (Bauer et al., 2015). The SBD is divided into two subdomains based on functionality: an α-helical domain (SBDα), which forms the lid, and a β sandwich domain (SBDβ) with a substrate binding site (Young, 2010; Melero et al., 2015). The SBDβ binds to a 5-7 amino acid-long stretch of hydrophobic residues flanked by basic amino acids on the nascent polypeptide chains. The interaction is mediated via hydrogen bonding between the nascent peptide backbone and SBDβ and by van der Waals contacts within hydrophobic moieties (Mayer, 2010). The NBD and SBD are linked via a hydrophobic inter-domain linker region (Swain et al., 2007). The binding of ATP in the NBD is coupled with the conformational changes within the SBD to regulate substrate binding. The ATP binding to the nucleotide-binding cleft triggers the attachment of the inter-domain linker and SBD α-helical lid to the NBD, displacing the lid to open up the substrate binding pocket (Kityk et al., 2012). The crystal structure of the DnaK protein in the open state is shown in **Figure 1.3 B**. Further, the hydrolysis of ATP to ADP induces conformational switching, which results in the detachment of the lid from NBD and the closing of the SBDβ binding pocket (Mapa et al., 2010).

The Hsp40 (DnaJ protein) contains a J-domain (70 residues long), which associates with the N-terminal ATPase domain and inter-domain linker region of the DnaK proteins (Ahmad et al., 2011). It has a C-terminal peptide-binding domain, which interacts with the conserved EEVD motif of DnaK SBD (Suzuki et al., 2010). The reaction cycle showing DnaJ-DnaK complex mediated protein folding through ATP hydrolysis is displayed in **Figure 1.3C**. The binding of Hsp40 proteins to Hsp70 strongly stimulates the ATPase activity of Hsp70 and modulates its substrate binding ability (Mayer, 2010). The Hsp40 co-chaperone also supports specialized functions of Hsp70, such as protein transport other than protein folding (Kampinga & Craig, 2010).

TH-3792_206106022

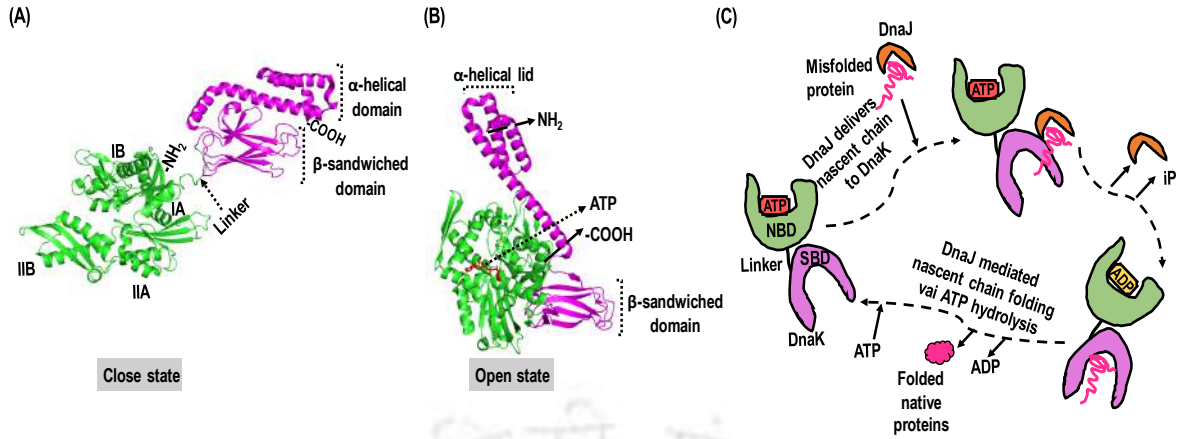


Figure 1.3. Structural and functional analysis of Hsp70 (DnaJ chaperone). Structure of different states of DnaJ **(A)** Close state (PDB: 2KHO) **(B)** Open state (PDB: 4B9Q). A linker region connects the nucleotide-binding domain (NBD; green) and the substrate-binding domain (SBD; magenta). The α-helical lid and β-sandwich domains of the SBD are marked. In the open state, the cavity in the NBD is the site for ATP binding. **(C)** Schematic representation of the DnaK function: DnaJ facilitates substrate recognition and delivery to the DnaK protein via interaction with its SBD region. ATP binding and hydrolysis mediated by NBD drive the folding of proteins. ATP hydrolysis triggers the transition of the open state to the closed form and, thereby, releases the folded substrate (adapted from Fatima et al., 2021).

1.1.3. The Hsp60 System

The Hsp60 system comprises oligomeric proteins (GroEL and GroES), which were discovered in the 1970s as an essential component of λ-phage assembly found in *E. coli* cytosol (Georgopoulos et al., 1973). The crystal structure of *E. coli* GroEL (547 residues) reveals a cylindrical structure with a central catalytic activity to encap the substrate protein (Braig et al., 1994). The monomeric subunit of GroEL (57 kDa) forms two heptameric rings arranged cylindrically (Braig et al., 1994). Each subunit contains three domains: an equatorial domain, a hinge domain, and an apical domain (**Figure 1.4A**). The equatorial domain (residues 6-133; E1 and 409-523; E2) is composed of a well-ordered helix, which forms the baseline of the GroEL ring that aids in the assembly of two heptameric rings and also exhibits nucleotide binding and hydrolysis ability (Hartl & Hayer, 2009). The apical domain (residues 233-267) has exposed hydrophobic residues that cover the apical surface of the GroEL ring cavity and interact with the protein substrate. The apical domain has a less ordered and flexible structure connected to the equatorial domain via an intermediate hinge domain (Sigler et al., 1998). The intermediate domain (I1; 138-192 and I2; 376-410) is responsible for the conformational switch of GroEL protein upon binding the nucleotide and protein substrate (Saibil et al., 2013).

The GroES (Hsp40) comprises seven identical subunits (10 kDa) to form a heptameric ring and caps the GroEL protein by binding to its apical domain (**Figure 1.4A**). The GroES subunit

TH-3792_206106022

contains nine beta-strands wherein a flexible hydrophobic loop of 22 amino acids between the second and third strands from each of the seven subunits protrudes outside to associate with GroEL (Lin & Rye, 2006). The substrate-bound GroEL heptameric rings bind to ATP molecules within their equatorial domain. Upon ATP binding, conformational changes occur within the intermediate and apical domains of GroEL, which allow the association of GroES to the apical domain of GroEL proteins (Sigler et al., 1998). After attachment, the GroES induces changes in the orientation of GroEL rings and, thereby, converts the hydrophobic central cavity of GroEL into a hydrophilic environment (Georgopoulos et al., 1973; Sigler et al., 1998). The entrapped misfolded protein substrate within this hydrophilic cavity can fold into its native state (Xu et al., 1997). During this encapsulation and folding stage, ATP molecules get hydrolyzed into ADP by the ATPase component of equatorial domains, facilitating the release of the substrate, GroES, and bound ADP (Hartl & Hayer, 2009).

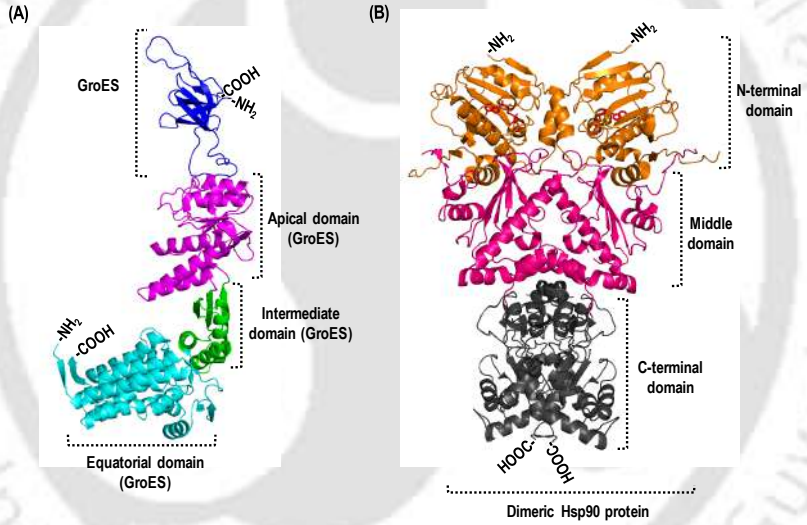


Figure 1.4. Structural analysis of Hsp chaperones. **(A)** The structure of the GroEL-GroES complex (PDB: 1PF9). The GroEL has three domains: the apical domain (magenta), which interacts with the substrate; the intermediate domain (green); and the equatorial domain (cyan), with nucleotide binding site. In the GroEL/ES complex, the GroEL subunits interact with each other through the equatorial domain and with GroES through its apical domain. **(B)** structure of ATP-bound closed state of Hsp90 chaperone in dimeric form (PDB: 2CG9). The N-terminal domain (orange), middle domain (pink), and C-terminal domain (grey) of Hsp90 are represented (adapted from Kim et al., 2013).

1.1.4. The Hsp90 System

The Hsp90 molecular chaperones were initially discovered to prevent misfolded protein aggregation and are proposed to be more substrate-specific than other Hsp chaperones (Picard, 2002). The Hsp90 chaperone acts upon two categories of protein substrates: transcription

TH-3792_206106022

factors (p53) and kinases (CDK4) (Li et al., 2012; Zuehlke et al., 2010). The structure of *E. coli* Hsp90 has three conserved domains: N-terminal, middle, and C-terminal domains (**Figure 1.4B**). The N-terminal domain comprises a two-layer α/β-sandwich structure with an ATP binding site. The middle domain consists of α-β-α motifs, which associate with protein substrates and regulate ATP hydrolysis activity. The C-terminal domain helps in the oligomerization of Hsp90 into its functional dimeric forms. The C-terminal end of Hsp90 proteins contains a conserved sequence (MEEVD), which binds to various cochaperones bearing tetratricopeptide (TPR) domains (Prodromou et al., 1997; Young et al., 1998). The dimer Hsp90 with a single inter-subunit contact between the CTD of monomers is considered an open state, which binds to the protein substrate, as observed in *E. coli* Hsp90 proteins (Shiau et al., 2006). The binding of ATP to the nucleotide-binding region of NTD induces the dimeric state to attain a closed, compact state upon interaction within the NTD of two subunits (shown in **Figure 1.4B**). The closed state provides an internal environment for protein substrate assembly and folding (Ali et al., 2006). Following this, a flexible loop of the middle domain induces ATP hydrolysis by the ATPase domain of the NTD. Once ATP gets converted into ADP after hydrolysis, the NTDs dissociate and transform into an open state (Dollins et al., 2007). The Hsp90 proteins do not clearly differentiate among their three domains for protein substrate binding sites. For example, the CDK4 kinase interacts with the middle domain, while some studies also provide evidence for substrate binding to the NTD or CTD of the Hsp90 (Vaughan et al., 2006; Scheibel et al., 1998; Fang et al., 2006).

1.1.5. The Hsp100 System

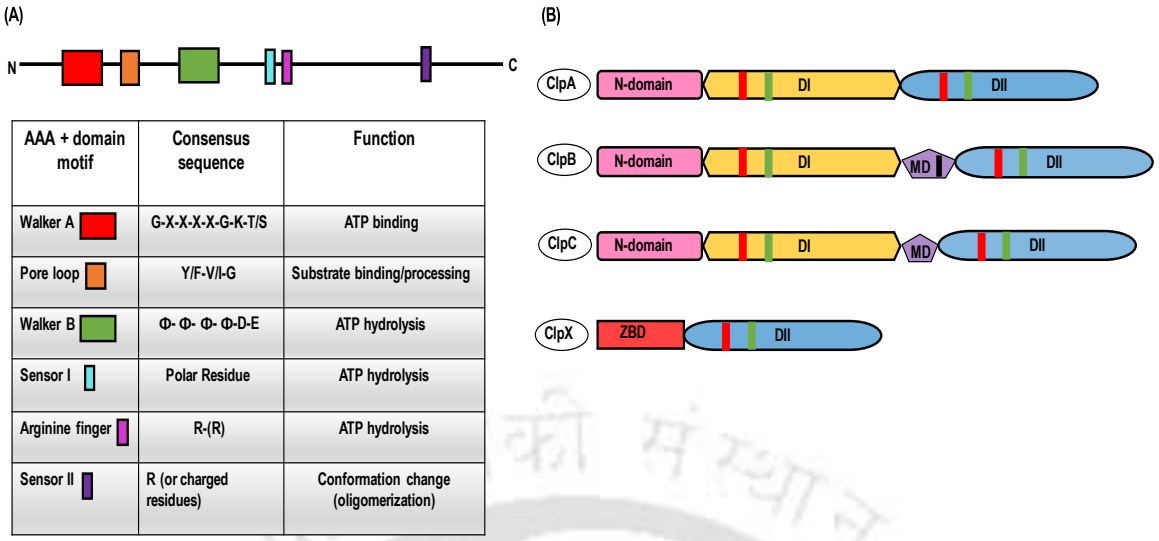
The Hsp100 chaperones are found across all domains of life and drive essential cellular processes such as protein degradation, cell cycle regulation, gene expression, and vesicular fusion. The Hsp100 proteins can act as an unfoldase as well as a disaggregase (Clare & Saibil, 2013). The unwanted, truncated, or misfolded proteins must be adequately cared for inside a cell to maintain cellular homeostasis. Several proteases perform energy-dependent proteolytic activities. The proteases are large, complex, multi-subunit structures that associate with Hsp100 molecular chaperones (Wickner et al., 1999). These molecular chaperones belong to the ATPase associated with diverse cellular activities (AAA+) superfamily of proteins (Baker & Sauer, 2006; Sauer & Baker, 2011). The AAA+ proteins have highly conserved ATPase domain modules of 200-250 amino acids, further dividing into two subdomains (large and small) (Hanson & Whiteheart, 2005). A large subdomain with αβα structural arrangement where parallel β-sheets sandwiched between two sets of α-helices possess a nucleotide-binding

TH-3792_206106022

site. The small subdomain consists of an α-helical arrangement (Ogura & Wilkinson, 2001). Within the ATPase module, several key structural motifs exist for the functionality of the Hsp100 chaperone and are displayed in **Figure 1.5A** (Hanson and Whiteheart, 2005). The large subdomain of AAA+ proteins possesses conserved Walker A and Walker B motifs. The Walker A motif (GX4GKT; X represents any amino acid) facilitates nucleotide binding and metal-ion (Mg^{2+}) coordination (Walker et al., 1982; Erzberger & Berger, 2006). The Walker A motif's conserved lysine (K) and threonine (T) residues interact with the β and γ-phosphate of ATP (Neuwald, 1999). The Walker B motif (RX4Φ4DE; Φ represents any hydrophobic aa) drives the hydrolysis of ATP through conserved aspartate (D) and glutamate (E) residues (Walker et al., 1982; Leipe et al., 2003). In addition, the AAA+ chaperone includes conserved motifs such as sensor I, sensor II, and Box sequences. The sensor motifs have conserved polar residues (asparagine, aspartate, threonine, or serine) that aid in the nucleotide binding and its hydrolysis by interacting with the γ-phosphate of ATP (Story & Steitz, 1992). Sensor II contains conserved arginine residues that interact with the γ-phosphate of bound-ATP (Neuwald, 1999). The Box sequences include two conserved motifs: Box II and Box VII. Box II recognizes the adenine moiety of ATP, and Box VII contains an arginine finger motif, which facilitates ATP hydrolysis and also aids in the association of AAA+ chaperone subunits (Davey et al., 2003).

A subset of AAA+ chaperones work in coordination with Caseinolytic protease P (Clp) and are called Clp ATPase chaperones. The Clp protease proteins were first found to degrade casein protein (found in milk), thus termed caseinolytic proteases (Katayama et al., 1988). The Clp proteins are found in all organisms, including bacteria and eukaryotic cells. The Clp ATPase chaperones are classified into two classes (Class I and II) based on the number of ATPase domains and diagrammatically represented in **Figure 1.5B**. The class I chaperones have two ATPase domains (DI and DII), including several Clp AAA+ proteins, ClpA, ClpB, ClpC, ClpD, ClpE, and ClpL. Class II chaperones with one ATPase domain similar to the DII of Class I chaperones include ClpX and ClpY AAA+ proteins (Schirmer et al., 1996). All mentioned Clp chaperones (except ClpB and ClpL) are reported to be associated with Clp proteases that form complex proteolytic machinery. The ClpB chaperone of the Hsp100 family acts as a disaggregase and possesses an additional coiled-coil insertion in its DI ATPase domain (Mogk et al., 2018).

TH-3792_206106022



AAA + domain motif	Consensus sequence	Function
Walker A	G-X-X-X-X-G-K-T/S	ATP binding
Pore loop	Y/F-V/I-G	Substrate binding/processing
Walker B	Φ- Φ- Φ- Φ-D-E	ATP hydrolysis
Sensor I	Polar Residue	ATP hydrolysis
Arginine finger	R-(R)	ATP hydrolysis
Sensor II	R (or charged residues)	Conformation change (oligomerization)

Figure 1.5. Domain organization of AAA+ chaperone. **(A)** The schematic representation of AAA+ protein domains and a table showing key structural motifs of AAA+ proteins with their consensus sequence and associated function. X is any amino acid, and Φ is a hydrophobic amino acid. **(B)** Schematic representation of the domain organization of Class I (ClpA, ClpB, and ClpC) and Class II (ClpX) Clp ATPases. The ATPase domains I and II are represented as DI and DII, respectively, and specific motifs are labelled with different coloured lines (Walker A: red and Walker B: green). The ZBD in ClpX stands for zinc-binding domain present in its N-terminal domain. The violet colour pentagon shape in ClpB and ClpC represents a middle domain with a black line showing insertion of the coiled coil domain within MD.

The ATPase chaperone assembles into functional hexamers with their nucleotide-binding sites at the interface between subunits. The hexameric assemblies have a central cavity where the protein substrate is unfolded, followed by translocation into the catalytic chamber of the protease component for degradation. The unfolding and translocation uses energy driven by ATP hydrolysis. The ATPases are coaxially aligned with the barrel-shaped peptidase component, which possesses an active catalytic site for protein degradation inside its cavity. In bacteria, there are multiple AAA+-associated proteases that regulate their proteome (Sauer & Baker, 2011). Among several bacterial species, the proteasomal complexes are extensively studied in *E. coli*, including Lon, FtsH, HslUV, ClpXP, and ClpAP (Sauer & Baker, 2011).

1.2. The AAA+ Proteases

The AAA+ proteases help to maintain protein homeostasis inside cells by degrading damaged or misfolded proteins. The process of protein degradation is irreversible; thus, the AAA+ protease system works with a specific recognition system mediated by adaptor proteins to prevent the uncontrolled destruction of functional proteins. The exposed region of a protein substrate interacts with the axial pore of the ATPase protein, which induces conformational

TH-3792_206106022

changes in the ATPase domain of AAA+ proteins. This drives the unfolding and translocation of unfolded protein substrate to the catalytic chamber of the peptidase component, where the protein is being degraded into small peptide fragments. In some cases, additional specificity factors known as adaptor proteins are involved in the substrate recognition step. The protein substrate with specific degradation tags (called degrons) at its C-termini or N-termini binds to the adaptor proteins, which will then be delivered to the AAA+ protease complex (Sauer & Baker, 2011). The basic working framework of AAA+ proteases is shown in **Figure 1.6**. The different AAA+ proteases are oligomeric proteins with similar structures but composed of variable polypeptides. The ATPase and protease domains are present in a single polypeptide chain for Lon and FtsH, while both domains are present separately in the case of ClpXP, ClpAP, and HslUV as represented in **Figure 1.7** (Sauer & Baker, 2011).

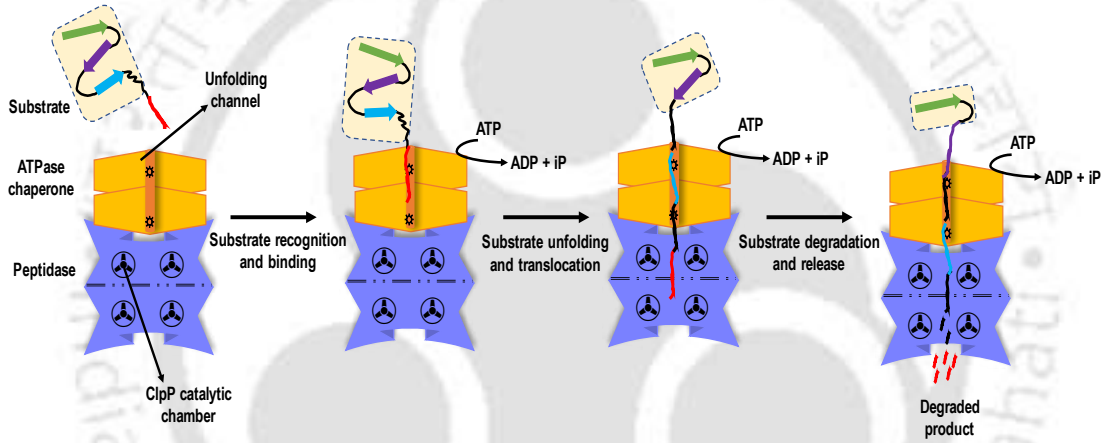


Figure 1.6. Basic working framework of AAA+ protease. A degradation tag (red coloured) in a substrate is initially recognized by a hexameric ATPase chaperone (green). ATP hydrolysis by the chaperone stimulates substrate unfolding and translocation into the catalytic chamber of the peptidase component. The degraded peptide products released from the exit channel of peptidase (adapted from Kim et al., 2013).

1.2.1. FtsH Protease

The FtsH is a membrane-bound ATP-dependent protease that targets abnormal integral membrane proteins (Ito & Akiyama, 2005). *E. coli* encodes a 647 amino-acid long FtsH protein (71 kDa), which exists in about 400 copies per cell under normal conditions (Tomoyasu et al., 1993). Out of five AAA+ proteases of *E. coli*, FtsH is only essential for its growth and survival (Ogura et al., 1991). The FtsH protease regulates the level of LpxC, which is important for lipopolysaccharide (LPS) biosynthesis and maintaining outer membrane integrity (Ogura et al., 1999). Thus, FtsH helps maintain an adequate LpxC enzyme level to maintain a proper LPS to other phospholipids ratio in the outer membrane (Führer et al., 2007).

TH-3792_206106022

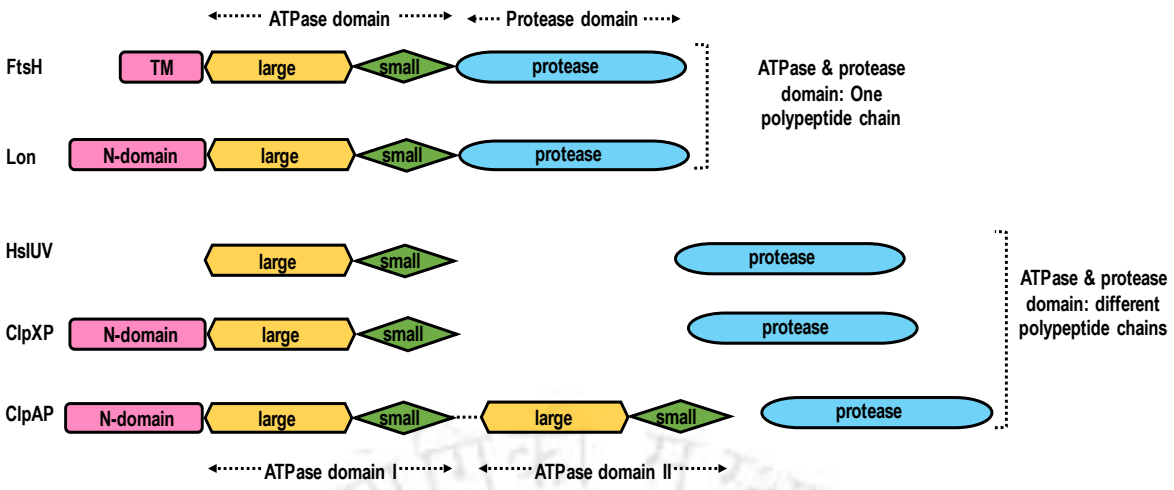


Figure 1.7. Domain organization of bacterial AAA+ proteases. The architecture of the five ATP-dependent proteases in *E. coli* is represented. The FtsH and Lon protease has ATPase and protease domains in a single polypeptide chain. The HslUV, ClpXP, and ClpAP proteases have different ATPase and protease domains. The N-terminal domain (pink), ATPase large subdomain (orange), small subdomain (green), and protease (blue) subunits are shown in different colours (adapted from Sauer & Baker, 2011).

The FtsH also helps sense several membrane stresses (Akiyama & Ito, 2000; Shimohata et al., 2002; Ito & Akiyama, 2005).

The FtsH is an inner membrane (cytoplasmic) embedded protein having a short N-terminal located towards the cytoplasmic region, followed by two transmembrane domains. The transmembrane domains are connected via a short periplasmic region (70 amino acids long), and the structure ends with a long cytoplasmic region (520 aa) (Tomoyasu et al., 1993) (**Figure 1.8A**). The residues 144-398 within the cytoplasmic region of FtsH possess characteristic sequence motifs of the AAA+ superfamily, which include the Walker A motif, Walker B motifs, and the Arginine finger (Krzywda et al., 2002; Neuwald, 1999). The remaining portion of the cytoplasmic regions functions as a protease domain with Zn^{+2} metalloprotease active site motif, a Zn^{+2} coordinating residue (Asp479), and a coiled-coil leucine zipper motif (Saikawa et al., 2002; Shotland[a] et al., 2000). The periplasmic, AAA+, and protease domain regions allow oligomerization of the ftsH subunit to form a functional homo-hexameric structure (Akiyama et al., 1998; Akiyama & Ito, 2000; Shimohata et al., 2002).

In *E. coli*, the FtsH is associated with a membrane protein complex (HflKC), forming a large holoenzyme bound to the cytoplasmic membrane (Kihara et al., 1996). The small periplasmic region of FtsH interacts with the HflKC complex, creating a large membrane protein with approximately 1000 kDa molecular weight (Kihara et al., 1996). The HflKC acts as an adaptor molecule mediating the recognition of membrane-bound substrates, followed by

TH-3792_206106022

its transfer to the ATPase domain of FtsH (Akiyama et al., 1998). The proteolytic degradation of protein substrates depends on a divalent metal cation (Zn^{+2}) and ATP hydrolysis (Tomoyasu et al., 1995). The FtsH shows processive degradation of protein substrate and releases degraded peptide products of approximately 20-30 amino acids long (Shotland[b] et al., 2000). The AAA+ protease FtsH is an endopeptidase that usually cleaves at the hydrophobic amino acid towards the N-terminal end of protein substrates (Asahara et al., 2000).

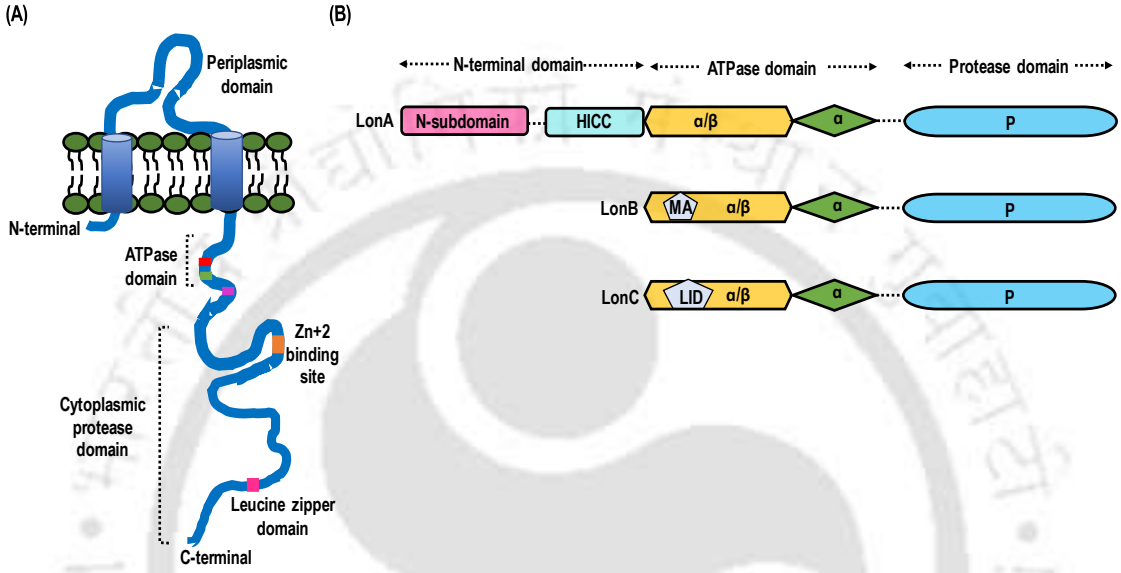


Figure 1.8. Domain representation of FtsH and Lon protease. **(A)** Cartoon representation of a FtsH protease subunit showing the short N-terminal followed by two transmembrane helices, a periplasmic domain, an ATPase domain, and a long cytoplasmic domain containing conserved regions such as Zn^{2+}-binding site and leucine zipper domain. **(B)** Block diagram comparing domains of LonA, LonB, and LonC. The N-terminal domain of LonA has two subdomains: The N-subdomain and an HICC (Helical Inserted with Coiled Coil) subdomain. The ATPase domains of all three Lon proteases have two subdomains, namely, α/β and α. The α/β subdomain of LonB has an insertion called the Membrane anchoring domain (MA), while LonC's insertion is called the Lon insertion domain (LID) (adapted from Nixon et al., 2005; Wlodawer et al., 2022).

1.2.2. Lon Protease

The Lon protease was identified as a key component of the protein quality control system about 30 years ago and was the first among various AAA+ proteases to be biochemically characterized (Charette et al., 1981; Chung & Goldberg, 1981). Lon is a universal protease system present in eubacteria, archaea, and eukaryotic organelles (Van Dyck et al., 1994). The Lon protease has been reported to control the degradation of approximately 50 % of unwanted proteins in *E. coli* cells (Chung & Goldberg, 1981). The Lon proteases are divided into subfamilies, including LonA, LonB, and LonC (Rotanova et al., 2004). The LonA subfamily is the most studied and present in several eubacteria and eukaryotic organelles. The LonB

TH-3792_206106022

subfamily belongs to the archaeal group, while LonC was recently discovered in the thermophilic bacteria, *Meiothermus taiwanensis* (Wlodawer et al., 2022).

The LonA comprises a long extended N-terminal region of size varying from 300-330 amino acids in bacteria to 420-570 aa in eukaryotes, which differentiates it from LonB and LonC proteases. The N-terminal region is further divided into the N-subdomain and HICC (Helical Inserted with Coiled-Coil fragment) domains (Rotanova et al., 2004; Rotanova & Melnikov, 2010). The catalytic site of the LonA subfamily is a duad of Ser and Lys residues. The catalytic site residues (S and K) represent a consensus sequence of KDGPSAG and (K/R)xKxF (x denotes any amino acid) (Wlodawer et al., 2022). The LonA functions as a hexamer, while in some organisms, such as *E. coli*, the EcoLonA exists in two states (hexamer and dodecamer) (Vieux et al., 2013). The LonB subfamily is characterized by the presence of a membrane anchoring region within the ATPase domain, which helps in its attachment to the cell membrane. LonB differs from LonA protease in its catalytic dyad consensus sequence and N-domain length. The consensus sequence represents UEGDSA(S/T) and (T/N)xKFE, where U denotes any hydrophobic amino acid (Wlodawer et al., 2022). The LonC protease also shows the absence of the N-terminal region. The LonC represents an insertion of helical segments called Helical Hairpin extension (HHE) or Lon Insertion domain (LID), similar to the membrane anchoring region of LonB (Li et al., 2013). Additionally, LonC shows mutations within its ATPase domain, thereby abolishing its nucleotide hydrolysis ability (Li et al., 2013; Wlodawer et al., 2022). The differences within domains of LonA, LonB, and LonC are displayed in **Figure 1.8B**. The structure of the complete LonA protease has not been solved to date by X-ray crystallography, but several fragments of the LonA polypeptide chain have been solved. The N-terminal domain of *E. coli* LonA (1-245) is divided into two subdomains; the first subdomain (1-116 residues) comprises seven β strands and two α helices. The first subdomain is connected to the second subdomain via an unstructured linker region (residues 117-123). The second subdomain has five α helices, where α 7 is a long, extended region at the carboxy end (Li et al., 2010).

1.2.3. The HslUV Protease

The HslUV (Heat shock locus UV), also known as ClpYQ, is a major AAA+ protease of *E. coli*, which degrades 70-80 % of protein substrates together with Lon and ClpAP protease (Jain & Chan, 2007; Maurizi, 1992). The Clp protease component (ClpY) exclusively interacts with its cognate ATPase partner, ClpY, while the other Clp protease (ClpP) can interact with multiple ATPase chaperones (ClpX, ClpA, ClpC, ClpE) (Kress et al., 2009; Hsieh et al., 2011).

TH-3792_206106022

The expression of genes *clpY* and *clpQ* is controlled by the heat shock transcription factor (σ32), which encodes a small protein (ClpQ; 19 kDa) and a large protein (ClpY; 49 kDa) (Missiakas et al., 1996; Yoo et al., 1996). The ClpY subunit oligomerizes into a hexamer, which has ATPase activity, while the peptidase subunit of ClpQ forms a double-ring dodecameric complex with a catalytic site located internally (Rohrwild et al., 1997). The ClpY crystal structure shows the presence of three domains: N, I, and C domains (Bochtler et al., 2000). The N-domain has ATP binding and hydrolysis activity (Wang et al., 2001). The domain I protrudes externally to capture the protein substrates while the C-domains are required for ClpQ association and activation (Ramachandran et al., 2002; Seong et al., 2002). Several natural proteins (SulA, RcsA, RpoH, TraJ, and σ32) have been reported to be degraded by ClpYQ protease of *E. coli* (Kanemori et al., 1999; Kuo et al., 2004; Wu et al., 1999; Khattar, 1997; Kanemori et al., 1997). There have been reports that ClpYQ helps in several SOS responses, including DNA damage, acidic stress, and unfolded or damaged protein accumulation (Missiakas et al., 1996; Kanemori et al., 1997; Kuo et al., 2004; Khattar, 1997). It is not an essential gene for the survival of *E. coli*, although its deletion has severely impacted the growth of *E. coli* above 43 °C (Missiakas et al., 1996).

The HslU component shows the presence of key structural motifs of AAA+ proteins, including Walker A, Walker B, sensor I, sensor II, and arginine finger motifs (Wang et al., 2001). The hexameric HslU shows asymmetric binding of nucleotides, where nucleotides bind to three or four subunits of the hexamer (Yakamavich et al., 2008). The structure of HslV shows that the catalytic site residues of HslV protease are properly aligned only when it is complexed with its cognate ATPase chaperone, HslV, making an HslUV complex (Sousa et al., 2000). Thus, the free HslV shows no activity on small substrates, but with the addition of HslU and a non-hydrolysable ATP analogue, it stimulates the degradation of peptide substrates (Rohrwild et al., 1996; Yoo et al., 1996). The C-terminal tail of HslU helps in coordinating with the activity of HslV peptidase subunit (Seong et al., 2002).

1.2.4. The ClpXP/ClpAP Protease

The Caseinolytic protease P (ClpP) is a self-compartmentalized complex having controlled proteolytic activity with a central catalytic chamber. The folded protein substrates cannot access the catalytic chamber of ClpP, and thereby, their degradation requires the assistance of Clp ATPase chaperones. The ATPase chaperone associates axially with the ClpP component, and with energy derived from ATP hydrolysis, it unfolds and transfers the unfolded protein substrates inside the catalytic chamber of ClpP. The *E. coli* ClpP component associates with

TH-3792_206106022

two Clp ATPase chaperones (ClpX and ClpA) and makes ClpXP/ClpAP proteolytic machinery (Liu et al., 2014). The ClpP is a highly conserved cytoplasmic protease in eubacteria and eukaryotic organelles (Yu & Houry, 2007). The ClpP subunits form a tetradecameric complex with a central catalytic chamber surrounded by an axial pore. The ClpP complex can degrade small peptide substrates (up to 5 amino acids) and requires association with the Clp ATPase component to degrade folded protein substrates (Thompson et al., 1994; Woo et al., 1989; Grimaud et al., 1998; Kessel et al., 1995). The activity of the ClpP is regulated by several ATPase chaperones, such as ClpA, ClpC, and ClpX, which bind directly to the ClpP ring, forming a larger complex (ClpAP/ClpCP/ClpXP) (Sauer & Baker, 2011). These ATPase chaperones are hexameric proteins containing one or two AAA^+ domains, which drive energy from ATP hydrolysis for protein unfolding. These chaperones can recognize protein substrates with specific degradation signals (called degrons) by themselves or by recruiting different adaptor proteins (Singh et al., 2001; Chien et al., 2007; Battesti & Gottesman, 2013; Dougan et al., 2003).

1.3. The Caseinolytic Protease P (ClpP)

The *E. coli* ClpP (EcoClpP) protein was the first caseinolytic protease to be purified and studied (Katayama et al., 1987; Katayama et al., 1988). The EcoClpP (207 amino acids) was first expressed as a proenzyme, which further undergoes post-translational modification to remove 14 amino acids from the N-terminal. The mature EcoClpP consists of 193 amino acids (Maurizi et al., 1990). The ClpP structure from several organisms revealed a similar structural architecture where the protein comprises 14 subunits arranged as two heptameric rings. The two heptameric rings associate to form a closed cylindrical structure with a central cavity possessing active sites (Yu & Houry 2007).

1.3.1. Overview of ClpP Structure

The crystal structure of EcoClpP shows a tetradecameric structure composed of two heptameric rings (Szyk & Maurizi 2006). A central cavity is formed upon two heptameric ring associations encompassing catalytic residues (**Figure 1.9A**). The 14-catalytic triad of EcoClpP (Ser, His, and Asp) was present inside the central cavity (Wang et al., 1997). The monomeric EcoClpP has three domains: The N-terminal, head, and handle domains (**Figure 1.9B**). The N-terminal domain is a highly flexible structure, forming an axial loop region. The axial loop restricts the opening of the ClpP central cavity on both ends and aids in its association with the cognate ATPase chaperone (Gribun et al., 2005; Bewley et al., 2006). The N-terminal axial

TH-3792_206106022

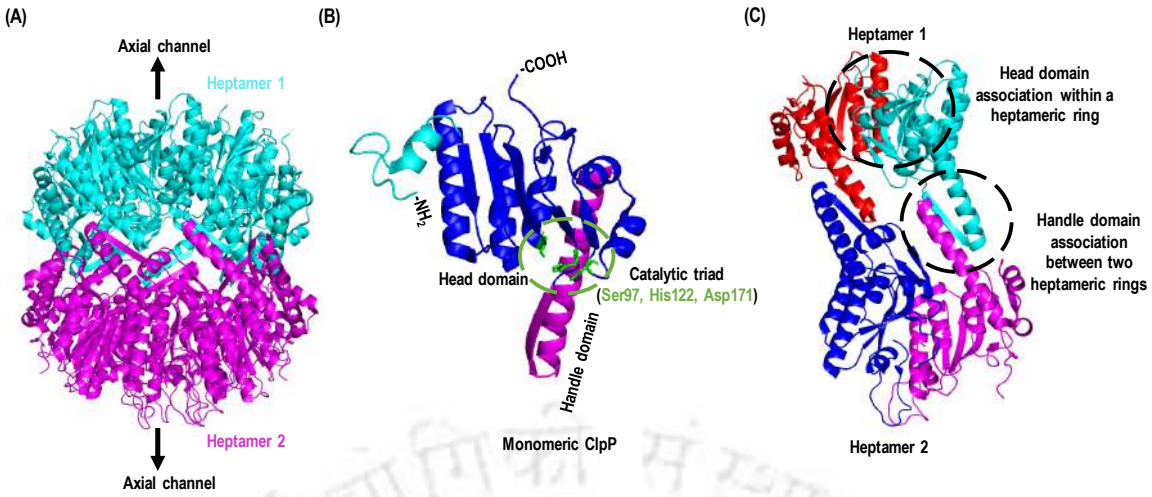


Figure 1.9. Overview of EcoClpP structure. (A) The side view of the EcoClpP crystal structure (PDB: 1YG6). The two heptamers (shown in cyan and magenta) form a tetradecamer complex (14-mer), and the N-terminal axial channel region is shown with black arrows in both heptameric rings. (B) The monomeric EcoClpP shows three characteristic domains: The N-terminal axial loop region (colored in cyan), the head domain (blue), and the handle domain (magenta). The catalytic triad residues (Ser97, His122, and Asp171) are shown as sticks and circled. (C) Four monomers (two from each heptameric ring) of the EcoClpP tetradecameric complex are shown. Two monomers from the same heptameric rings (red and cyan color) interact via their head domain. Two monomers from different heptameric rings (cyan and magenta) interact via their handle domain (adapted from Liu et al., 2014).

loop region has two parts: an axial pore lining (1-7 residues) composed of hydrophobic amino acids that line the ClpP central cavity. The second part, termed axial protrusion (8-16 residues), is composed of hydrophilic or charged amino acids that extend away from the axial pore of ClpP (Gribun et al., 2005; Bewley et al., 2006). The axial loop region is followed by a head domain (residues 28-120), which forms the body of the EcoClpP. The catalytic triad residues (Ser97, His1222, and Asp171) are within the head domain (Szyk & Maurizi 2006). The last part of the EcoClpP structure is a handle domain composed of a β-strand (β9) and an E helix (residues 125-157) (Gribun et al., 2005). The head domains mediate the interaction between monomers of a heptameric ring (inter-ring contact) via hydrophobic interactions and hydrogen bonding. The two monomers of different heptameric rings (intra-ring contact) are associated via their handle domains through hydrogen bonding (**Figure 1.9C**) (Wang et al., 1998).

The positioning of the N-terminal axial loops regulates the access to the catalytic central chamber of EcoClpP. The structure of axial loops is highly flexible and can adopt an up or down conformation across the axial entrance pore of ClpP. In the up conformation of EcoClpP, six out of seven axial loops of a heptameric ring were observed to protrude outwards from the axial surface of the EcoClpP cylinder. On the other hand, the axial loops are packed within the central proteolytic cavity in the down conformation. The entry of protein substrate inside the

TH-3792_206106022

catalytic chamber of ClpP is controlled by N-terminal axial loop positioning (Gribun et al., 2005; Szyk & Maurizi 2006). Thus, the up-and-down conformation of the axial loop is linked to the open and closed state of EcoClpP, respectively (Effantin et al., 2010; Bewley et al., 2006). In the open state, the axial loop acts as a gate and restricts the entry of folded proteins into the ClpP catalytic chamber. The axial loop deleted mutant of EcoClpP shows a much higher rate of peptide substrate degradation than wild-type EcoClpP, proving restriction by axial loops (Gribun et al., 2005). In the open state, the N-terminal axial loop residues form a specific beta-hairpin structure, which is essential for the efficient translocation of protein substrates inside the catalytic chamber of EcoClpP. A study shows that substituting Ile21 with Pro in EcoClpP prevents the formation of beta-hairpin and limits the proteolytic activity of EcoClpP (Alexopoulos et al., 2013; Gribun et al., 2005).

The N-terminal axial loops are also proposed to be significant for the association of EcoClpP with its cognate ATPase chaperones (Martin et al., 2007). The binding of the EcoClpA chaperone to the EcoClpP axial loops results in the conformational change of the EcoClpP axial loop into an up conformation, followed by the widening of the entrance pore of the catalytic chamber. This widening allows easy access to the substrate into the EcoClpP catalytic chamber (Jennings et al, 2008; Martin et al., 2007; Effantin et al., 2010).

1.3.2. Dynamic Handle Domain of ClpP

The handle domain of ClpP is a dynamic region adopting different conformational states, which correlates with the catalytic activity of ClpP (Gersch et al., 2012; Ye et al., 2013; Zeiler et al., 2013; Zhang et al., 2011). Based on the various orientations of the handle domain, three other states of ClpP, extended, compact, and compressed, have been reported. The structure of ClpP from *E. coli*, *Francisella tularensis*, *Helicobacter pylori*, *Listeria monocytogenes* ClpP2 and human ClpP was observed in the extended state (Wang et al., 1997; Díaz-Sáez et al., 2017; Kim & Kim 2008; Zeiler et al., 2013; Kang et al., 2004). The compact state of ClpP was observed for *E. coli*, *Streptococcus pneumoniae*, *Mycobacterium tuberculosis*, *L. monocytogenes* (ClpP1), *Plasmodium falciparum,* and *Staphylococcus aureus* (Kimber et al., 2010; Gribun et al., 2005; Ingvarsson et al., 2007; Zeiler et al., 2013; El Bakkouri et al., 2010; Vedadi et al., 2007; Geiger et al., 2007). The compressed state of ClpP was reported in the case of *Bacillus subtilis* and *S. aureus* (Lee et al., 2011; Geiger et al., 2011; Ye et al., 2013; Zeiler et al., 2013; Zhang et al., 2011).

TH-3792_206106022

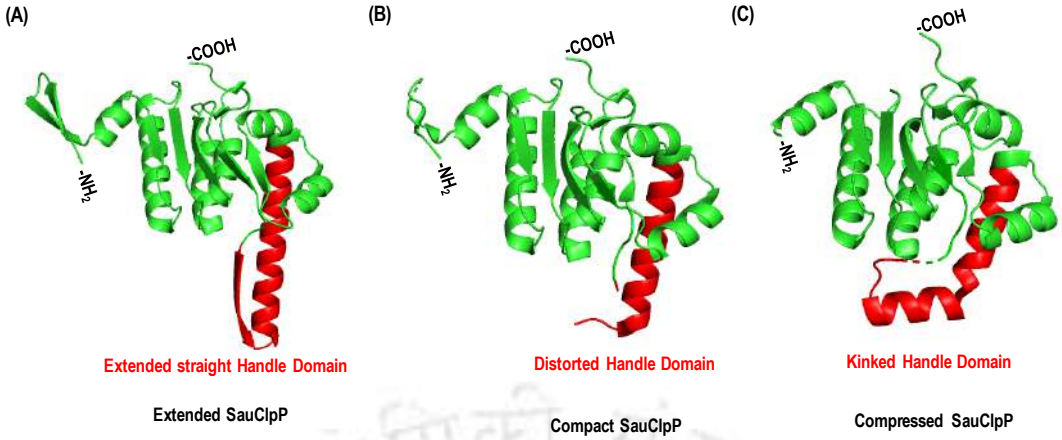


Figure 1.10. Three conformational states of SauClpP. The monomeric SauClpP in different conformational states **(A)** Extended (PDB: 3STA), **(B)** Compact (PDB: 4EMM), and **(C)** Compressed (PDB: 3ST9). The N-terminal and C-terminal ends are marked as -NH_2 and –COOH. The differentially aligned handle domain of each structure is represented in red color (adapted from Liu et al., 2014).

The crystal structure of *S. aureus* ClpP (SauClpP) has been solved in all states (Gersch et al., 2012; Zhang et al., 2011). In the extended state, the handle domain of the SauClpP monomer in the heptameric rings is well-ordered (**Figure 1.10A**). The oligomerization sensor residues of SauClpP (Arg171 and Asp170) stabilize two heptameric rings to form the tetradecameric structure. Also, the Asp170 and Arg171 residues form hydrogen bonds with Gln132 of the handle domain, thereby stabilizing the structure of the handle domain in an extended straight conformation. In the extended state, an extensive network of hydrogen bonding exists between residues Arg171-Asp170 and Gln132-Glu135 of monomers from opposite heptameric rings. This leads to the proper orientation of the SauClpP catalytic triad and allows the glycine-rich loop (Gly127, Gly128, and Gly131) of SauClpP to form an anti-parallel beta-sheet in the tetradecameric complex (Gersch et al., 2012; Zeiler et al., 2013). In the compact state of SauClpP, the hydrogen bond network between Arg171, Aps170, and Gln132 residues has been abolished, resulting in a distorted arrangement of the handle domain (**Figure 1.10B**). The compact state also revealed an improper orientation of the catalytic triad corresponding to an inactive state of SauClpP (Zhang et al., 2011).

The compressed structure of SauClpP represents an inactive state in which the cylindrical tetradecameric complex is shorter than that of extended or compact states. Within the handle domain of compressed SauClpP, the E helix is broken at residue Lys145, resulting in a kink that divides the E helix into two smaller helices (**Figure 1.10C**). The kink within the extended E helix brings the two heptameric rings closer, resulting in a shorter cylindrical structure (Zhang et al., 2011; Geiger et al., 2011). Additionally, all the hydrogen bond networking

TH-3792_206106022

between Arg171-Asp170-Gln132 residues is absent in the compressed state, which alters the proper alignment of the catalytic triad and thus makes the compressed state an inactive form (Zhang et al., 2011).

1.3.3. ClpP Modulators

A study by Brötz-Oesterhelt et al. characterized a cyclic compound, Acyldepsipeptides (ADEPs), derived from *Streptomyces hawaiiensis* and proposed it to be a modulator of ClpP activity (Brötz-Oesterhelt et al., 2005). The ADEPs are natural antibiotics that dysregulate ClpP's activity. ADEPs stimulate uncontrolled proteolysis in vivo and promote dissociation of the chaperones from the ClpP component, thus leading to cell death (Brötz-Oesterhelt et al., 2005). The structural changes associated with ADEPs binding have been reported for several ClpPs, including *B. subtilis* ClpP (BsuClpP), *E. coli* ClpP (EcoClpP), *M. tuberculosis* ClpP (MtuClpP1P2), and *N. meningitides* ClpP (NmeClpP) (Goodreid et al., 2016; Lee et al., 2010; Li et al., 2010; Schmitz et al., 2014). The structural analysis of ADEP-bound ClpP revealed that it binds to a hydrophobic cluster region (H-pocket) between two adjacent monomers on the apical surface of the ClpP ring. The binding of ADEP to the apical hydrophobic surface of ClpP results in the reorientation of the ClpP subunits, which ends up in the enlargement of the catalytic chamber pore (axial pore) (**Figure 1.11**). The opening of the entrance pore of ClpP leads to easy access to the protein substrate in the catalytic cavity of ClpP (Sass et al., 2011; Kirstein et al., 2009; Gersch et al., 2015). Recently, a study revealed that the axial surface of EcoClpP possesses a hydrophobic residue, Tyr63, which acts as a gatekeeper, keeping the entry of substrates restricted. Further, the binding of ADEP results in conformational twisting of the Tyr63 side chain that eliminates the restriction to the entry of substrates inside the EcoClpP catalytic chamber (Ni et al., 2016).

The discovery of ADEP has paved the way for several novel studies leading to the identification of new antibiotics to be used as an alternative therapeutic agent to target ClpP activity. A study by Leung et al. screened out several chemicals that structurally are not related to ADEP but functionally activate ClpP protease activity. These ClpP modulators are named activators of self-compartmentalizing proteases (ACPs) and show potent characteristics of activating ClpP proteolytic activity similar to ADEP. The molecular docking experiments revealed that ACPs have two binding pockets on the ClpP; one is the Hydrophobic cluster at the axial surface of ClpP (H-pocket), while the second pocket is named C-pocket, which comprises charged residues (Leung et al., 2011).

TH-3792_206106022

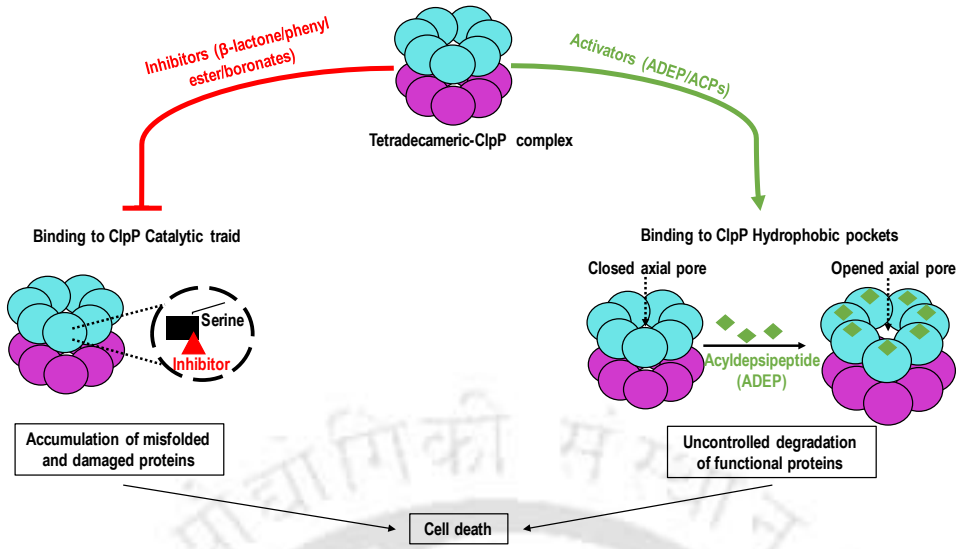


Figure 1.11. The modulators (activators or inhibitors) of ClpP activity. The catalytic activity of ClpP is inhibited by compounds that bind to the –OH group of the serine residue within the catalytic triad. The ClpP activator (ADEP) binds to the hydrophobic patches on the axial surface of ClpP tetradecamer and induces widening of the ClpP axial pore, resulting in uncontrolled proteolysis (adapted from Queraltó et al., 2023).

Another approach to modulate the activity of ClpP includes the development of drugs that inhibit the catalytic function of the protease by interfering with the serine residue within the catalytic triad (**Figure 1.11**). A group of compounds belonging to the β-lactones family were screened to irreversibly inhibit ClpP activity in *S. aureus*, *L. monocytogenes*, *P. falciparum* and *M. tuberculosis* (Compton et al., 2013; Böttcher & Sieber 2009; Gersch et al., 2013). Another potent drug to be identified as an activity modulator for ClpP was the phenyl ester. The phenyl esters covalently bind to the catalytic serine residue and trigger the dissociation of the tetradecameric proteolytic complex into inactive heptamers (Hackl et al., 2015). Another category of modulators includes β-sultams, which contain cyclic sulfonyl groups. The sulfonyl group attacks the serine residue and triggers the disassembly of the ClpP complex into inactive heptamers (Gersch et al., 2014). Another study has identified a boron-derived compound (Bortezomib) as a ClpP inhibitor in *M. tuberculosis* (Moreira et al., 2015). Another recent study reported Myricentin (a natural flavonoid compound) as a blocker of ClpP activity in *S. aureus* (Jing et al., 2021).

Furthermore, a recent study depicts allosteric modulation of *Staphylococcus epidermidis* ClpP by a peptidomimetic boronate, ixazomib. The ixazomib binds to the catalytic residue (Ser) and stimulates the uncontrolled proteolytic activity of SepClpP (França et al., 2024). Also, another natural β-lactone derivative, Cystargolides A, has been reported to inhibit the

TH-3792_206106022

proteolytic activity of *S. aureus* ClpP (SauClpP). The crystal structure solved for SauClpP-Cystargolides A complex revealed that Cystargolides A covalently binds to all 14 active site residues in the tetradecameric complex (Illigmann et al., 2023*)*.

1.4. The Clp Protease-ATPase Complex

The ATPase chaperone ClpX possesses a specific zinc-binding domain (ZBD) and one AAA+ domain. The ZBD of ClpX folds independently and can help dimerize ClpX independent of its AAA+ domain (Donaldson et al., 2003; Wojtyra et al., 2003; Sauer & Baker, 2011). The ZBD is a highly flexible region protruding from the AAA+ domain, which helps in substrate recognition and binding with adaptor proteins. The hexameric form of ClpX (trimer of dimers) is associated with heptameric rings of ClpP (Sauer & Baker, 2011). The ClpXP complex was first discovered to degrade a bacteriophage replication initiator protein (λO) (Gottesman et al., 1993). Other in-vivo protein substrates to be degraded by the ClpXP complex include MuA, Mu repressor, and SsrA-tagged proteins (Flynn et al., 2003).

The ClpA chaperone was the first ATPase to be discovered, which associates with the Clp protease system and was reported to degrade α-casein protein substrates as ClpAP complex (Hwang et al., 1987). Under in vitro conditions, the *E. coli* ClpA (EcoClpA) was reported to associate with the phage origin-binding protein (RepA) and promote the disassembly of inactive dimer into a monomeric state (Hoskins et al., 1998; Lopez et al., 2020). The ClpAP complex recognizes protein substrates with specific amino acid tags at the proteins' C- or N-terminus. For instance, ClpAP degrades RepA protein by identifying a specific N-terminal recognition tag. A fusion protein construct containing the first 25 residues of RepA fused with native GFP protein was reported to be degraded by the EcoClpAP complex (Lopez et al., 2020).

The Clp ATPases-protease complex formation involves the docking of a conserved Ile/Leu-Gly-Phe motif (I/LGF loops) of the Clp ATPase onto the axial surface (apical hydrophobic pocket) of the ClpP (Kessel et al., 1995; Kim et al., 2001). In the past few years, the structures of the Clp ATPase-Protease complex have been solved for very few organisms, including ClpXP from *N. meningitidis,* ClpC1P1P2 from *S. hawaiiensis*, ClpXP1P2 from *L. monocytogenes,* ClpXP and ClpAP from *E. coli* (Ripstein et al., 2020; Xu et al., 2024; Gatsogiannis et al., 2019; Fei et al., 2020; Lopez et al., 2020).

The *L. monocytogenes* comprises two isoforms of ClpP (ClpP1 and ClpP2), which assemble to form a hetero-tetradecamer complex (LmoClpP1P2) (Balogh et al., 2017). The catalytic residue mutant (Ser98Ala) of LmoClpP isoforms and nucleotide hydrolysis mutant of LmoClpX (Asp183Gln) were used for ClpXP1P2 complex formation. The obtained cryo-EM

TH-3792_206106022

structure shows a fraction of LmoClpXP1P2 dimers and LmoClpXP1P2 monomers. The dimerization of the LmoClpXP1P2 complex was due to the interaction between the flexible N-terminal ZBDs of adjacent LmoClpX hexamers. The structure of LmoClpXP1P2 complex represents an asymmetric organization where ClpP1 homoheptamers form the upper part while ClpP2 forms the base of the ClpP tetradecameric ring, which associates with the ATPase ClpX. The structure of the LmoClpP1P2 complex in the presence of LmoClpX was found to be identical to the crystal structure of the LmoClpP1P2 heterocomplex only. Further, the hexameric LmoClpX rings are not centrally aligned with the LmoClpP1P2 complex but were found to be tilted 11° toward LmoClpP2. The LmoClpX association with the LmoClpP1P2 heterocomplex changes the overall conformation of AAA+ domains of LmoClpX. The ATPase domain of LmoClpX was found to be more regularly arranged. All six subunits of hexameric LMoClpX are packed close to each other, and the small and large subdomains are in an open conformation with an accessible nucleotide-binding pocket (Gatsogiannis et al., 2019).

The interface of the LmoClpP2-ClpX complex shows a six-seven-fold symmetry mismatch. The flexible IGL loop from the large subdomain of each LmoClpX subunit is positioned below the hydrophobic cluster of the LmoClpP2 subunit. The hydrophobic cluster of LmoClpP2 occupies the interface of two subunits. This alignment allows the direct interaction between the IGL loop and the hydrophobic grooves below. Within the LmoClpP2-ClpX complex, five IGL loops display a symmetrical arrangement, while the sixth IGL loop is placed between the two hydrophobic grooves. The sixth IGL loop acquires an extended conformation and interacts with the hydrophobic groove on the left-hand side. However, the other hydrophobic groove remained empty. Thus, tilting the LmoClpX ring towards the LmoClpP2 heptamer and extending one of the IGL loops of LmoClpX is sufficient for adopting a C6-C7 LmoClpP2-ClpX complex (Gatsogiannis et al., 2019).

The *E. coli* possesses a single ClpP isoform forming a tetradecamer and its cognate ATPases (ClpA and ClpX) (Liu et al., 2014). The EcoClpX consists of an N-terminal domain and an ATPase domain, which contains specific conserved motifs for ATP binding and hydrolysis, ClpP binding, and substrate recognition. The EcoClpP structure reveals a flexible N-terminal region, which forms a specific beta-hairpin structure to create a central catalytic pore of the ClpP tetradecamer. The EcoClpX structure shows that the six subunits of the hexameric ring are arranged spirally with a central hollow cavity, which aligns with the EcoClpP catalytic central pore. The cryo-EM structure of the EcoClpXP complex shows the abundance of doubly capped complexes in a ClpXPX orientation (Fei et al., 2020). The EcoClpX has conserved IGF loops (Ile268-Gly269-Phe270), which dock at the axial hydrophobic pockets of EcoClpP.

TH-3792_206106022

Apart from hydrophobic clusters, the side chain of Arg192 of the EcoClpP rings was also observed to be involved in hydrogen bonding with the backbone of IGF loops. The structure of the EcoClpXP complex displays a diameter of ~30 A˚, similar to the axial diameter observed in the case of ADEP-bound EcoClpP. Thus, it concludes that EcoClpX binding induces the widening of the EcoClpP axial pore, as reported earlier for ADEP binding (Li et al., 2010; Fei et al., 2020). However, it contradicts recent findings from the *L. monocytogenes* ClpXP1P2 complex (Gatsogiannis et al., 2019; Fei et al., 2020).

The cryo-EM structure of *Neisseria meningitidis* ClpX-ClpP complex (NmeClpXP) shows that the NmeClpX is positioned with a slightly tilted orientation (15˚) relative to the NmeClpP (Ripstein et al., 2020) (**Figure 1.12A**). The association between hexameric ClpX and the double-heptameric rings of ClpP forms a continuous channel comprising a catalytic site for substrate unfolding, translocation, and degradation. The IGF loops of NmeClpX engage with six out of seven hydrophobic pockets of NmeClpP. The association of NmeClpX IGF loops allows the reorganization of N-terminal axial loops of NmeClpP into a well-ordered beta-hairpin structure, which resembles the structure observed for ADEP-bound NmeClpP (Goodreid et al., 2016; Ripstein et al., 2020). The NmeClpX binding induces the opening of the axial channels of NmeClpP, allowing easy access to the substrates in the catalytic chamber of NmeClpP. In all three reported cryo-EM structures of ClpXP complexes from *L. monocytogenes*, *E. coli*, and N. meningitidis, the general architecture of the complex remains conserved, where ClpX acquires a slightly tilted position related to the ClpP central axis. However, some dissimilarities exist between the structures obtained for LmoClpXP1P2 and NmeClpXP. In the LmoClpXP complex, dimerization of the LmoClpXP complex was observed with the involvement of the ZBD of LmoClpX. However, despite having a similar ZBD, the NmeClpX does not support the formation of such NmeClpXP dimers (Gatsogiannis et al., 2019; Ripstein et al., 2020). Additionally, the binding of ClpX to the apical surface of ClpP in the case of *E. coli* and *N. meningitidis* results in the widening of the axial pores of the ClpP catalytic chamber. However, the apical loops of LmoClpP1P2 do not have any change in their conformation upon LmoClpX binding (Gatsogiannis et al., 2019; Ripstein et al., 2020; Fei et al., 2020).

TH-3792_206106022

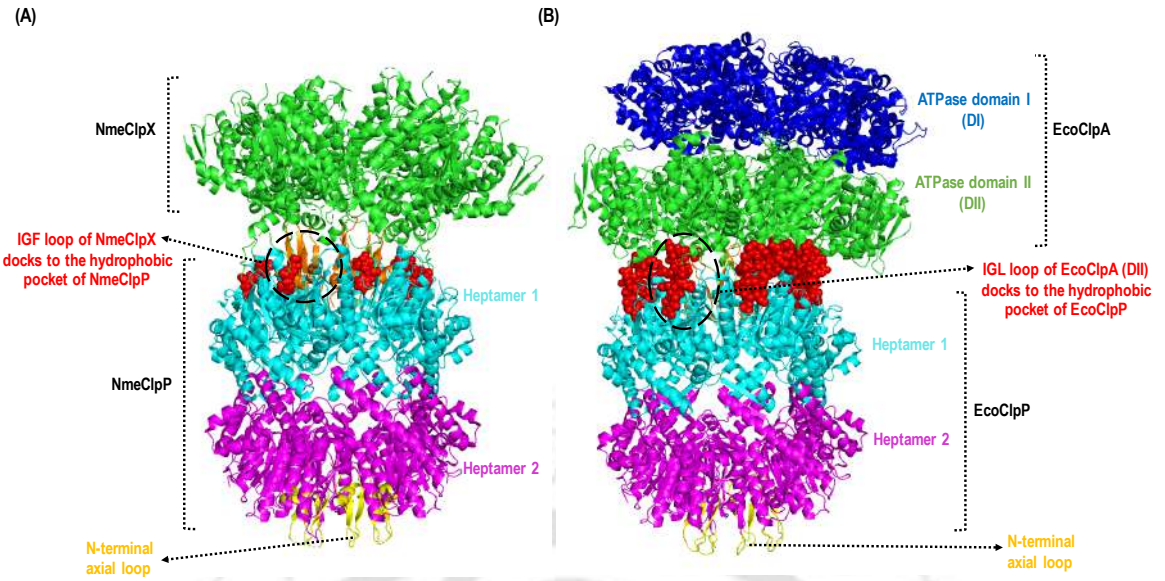


Figure 1.12. The cryo-EM structures of Clp ATPase-protease complex. **(A)** The ClpXP structure from *N. meningitides* (PDB*:* 6VFS). **(B)** The ClpAP structure from *E. coli* (PDB: 6UQE). The two heptameric rings of NmeClpP and EcoClpAP are shown in magenta and cyan colours. The NmeClpX hexameric ring is shown in green colour. The DI and DII ATPase domains of EcoClpA were shown in blue and green colour, respectively. The IGF loop of NmeClpX and the IGL loop of EcoClpA DII are marked as red spheres and protrude towards the hydrophobic axial region (highlighted in orange color) of NmeClpP and EcoClpP heptamer 1. The docking site of the IGF/L loop and the hydrophobic pocket is circled (Ripstein et al., 2020; Lopez et al., 2020).

The recently reported cryo-EM imaging of the *E. coli* ClpAP complex shows the EcoClpP tetradecameric complex to be double-capped with EcoClpA hexamers (Lopez et al., 2020). Out of two, one ClpA hexameric ring showed a well-resolved structure as a complex. In the EcoClpAP complex structure, the DII ring contacts the EcoClpP heptameric ring while the DI domain protrudes away from the EcoClpP ring (**Figure 1.12B)**. The IGL loop region (residues 611-623) of EcoClpA protrudes outwards from the DII large subdomain and extends into the hydrophobic pockets on the apical surface of EcoClpP rings. The six EcoClpA IGL loops occupy the six hydrophobic pockets, while the one remaining pocket on the apical surface of EcoClpP results in a symmetry mismatch. Further, the N-terminal region of EcoClpP (1-18 residues) forms the pore of the EcoClpP catalytic chamber, and the access to these pores is reported to be controlled by ADEP antibiotic or IGF loops of ClpX (Lopez et al., 2020). The EcoClpAP structure displays an open extended state conformation in line with the previous reports from ADEP-bound EcoClpP (Lopez et al., 2020; Lee et al., 2010). The hexameric EcoClpA translocation channel aligns with the catalytic chamber of EcoClpP.

The *Streptomycetes* strains encode for three to five ClpP homologs, and a recent study determined the cryo-EM structure of ClpP1P2 associated with ATPase chaperone ClpC1 from

TH-3792_206106022

S. hawaiiensis (Viala et al., 2000; Xu et al., 2024). The ClpP1 and ClpP2 isoforms of *S. hawaiiensis* form a heterocomplex (ShaClpP1P2), which exists in an extended conformation consisting of two homo-heptameric rings (Xu et al., 2024). The Cryo-EM structure of the ShaClpC1P1P2 complex reveals that the hexameric ShaClpC1 arranges along the central axis of the ShaClpP1P2 tetradeacmer ring with a tilting of 18° from the central ring. The ShaClpC1 subunit has two ATPase domains (DI and DII), each divided into a large and small subdomain. The large subdomain of the ShaClpC1 is associated with the ShaClpP2 subunits, while the small subdomain faces outside. In the ShaClpC1P2 complex, the LGF loop from the DII ATPase domains of ShaClpC1 is inserted into the hydrophobic pockets at the interface of two ClpP2 subunits. The six subunits of ShaClpX have their LGF loops inserted within ShaClpP2 grooves while one hydrophobic pocket remains empty.

1.5. Substrate Specificity by Clp Adaptor Proteins

The proteolysis event needs to be controlled and regulated to avoid the unnecessary degradation of proteins. To regulate this, proteolytic machinery usually recognizes specific signals called degrons associated with degradation substrates only. In bacteria, the signal degrons are usually a stretch of amino acids present at either the N-terminal or C-terminal of proteins. In some cases, certain internal regions or some tertiary conformations also serve as signal sequences for proteasomal degradation. These degrons in bacteria can be recognized directly by AAA+ chaperones or bind to adaptor proteins that further deliver them to the ATPase-Protease machinery. A similar way of controlled proteolysis exists in eukaryotes, which is more complex and advanced. The 26S proteasome of eukaryotes is an AAA+ cytoplasmic protease that targets proteins with a specific tag for degradation. The eukaryotic degron involves a polyubiquitin tag (4-6 ubiquitin monomers) attached to the target protein's lysine residue (Sauer & Baker, 2011). The Eukaryotes contain one predominant cytoplasmic AAA+ protease, the 26S proteasome. The proteins targeted to this degradation machine generally display a common degron, a polyubiquitin tag (at least four ubiquitin moieties) attached to a lysine residue. Recognizing the target substrate and adding a polyubiquitin tag is a primary function of E3 ubiquitin ligases. E3 ubiquitin ligases are a diverse class of enzymes with more than 600 subtypes found in humans and are responsible for the addition of polyubiquitin chains onto protein substrates (Varshavsky, 2011). The Eukaryotic cells rely on ubiquitin tagging to target proteins for proteasomal destruction, but prokaryotes employ several adaptor proteins to deliver specific substrates directly to the AAA+ proteases.

TH-3792_206106022

The gram-positive bacteria *Bacillus subtilis p*ossesses six different adaptor proteins, including CmpA and YjbH, controlling ClpXP activity, while McsB, YjbA, Gp53, and MecA control ClpCP activity (Wen et al., 2025). The Adaptor protein CmpA delivers the spore-inducing protein (SpoIVA) to the ClpXP complex, disrupting sporulation (Tan et al., 2015). The YjbH mediates the degradation of Spx by the ClpXP complex. The Spx is a transcription factor regulating the expression of genes associated with oxidative stress conditions (Awad et al., 2019). The YjbA is a recently reported adaptor protein mediating the degradation of sporulation-specific proteins in *B. subtilis* (Riley et al., 2025). The BsuClpC protein in complex with BsuMecA regulates the refolding of protein aggregates (Kirstein et al., 2009). The BsuMecA helps recognize specific protein substrates like ComK. The ComK is a transcription factor that regulates competence development in *B. subtilis*. The MecA binds to ComK and delivers it to the ClpCP proteolytic complex for degradation (Schlothauer et al., 2003). The McsB is another adaptor protein reported in *B. subtilis*, which regulates the function of BsuClpC. The McsB is an arginine kinase that requires an activator (McsA) for the phosphorylation role. The auto-phosphorylated MscB binds to the CtsR and delivers it to the BsuClpCP proteolytic complex (Kirstein et al., 2005; Krüger & Hecker, 1998). The CtsR is a transcription regulator that controls the transcription of several heat shock genes (Derré et al., 1999).

The gram-negative *E. coli* Clp system comprises two proteolytic machineries (ClpXP and ClpAP) with different substrate specificities controlled by several adaptor proteins. Three adaptor proteins (SspB, umuD, and rssB) have been reported to control EcoClpXP activity, while one adaptor protein (ClpS) alters EcoClpAP activity (Kirstein et al., 2009). The EcoSspB recognizes the SsrA-tagged substrates and mediates its efficient delivery to the EcoClpXP complex (Levchenko et al., 2000). The EcoUmuD is an adaptor protein that forms a heterodimer with the UmuD' (UmuD-UmuD' complex) and facilitates its delivery to the ClpXP proteolytic chamber. The UmuD' protein is a regulator of SOS responses in *E. coli* (Gonzalez et al., 2000; Neher et al., 2003). The EcoRssB mediates the degradation of RpoS (a regulator of stress response) by the ClpXP complex (Muffler et al., 1996; Zhou et al., 2001). Lastly, the EcoClpS possesses the unique ability to negatively or positively regulate the activity of EcoClpAP (Erbse et al., 2006). EcoClpS inhibits the EcoClpAP-mediated degradation of SsrA-tagged substrates. Alternatively, the EcoClpS stimulates the delivery and degradation of N-end rule substrates to the EcoClpAP complex (Dougan et al., 2002; Schmidt et al., 2009).

TH-3792_206106022

1.5.1. The Bacterial C-Degron

Identifying specific protein substrates for their degradation whenever they are not required, damaged, or truncated is critical to maintaining cell proteome integrity. Most commonly, the substrates for proteolysis have signal sequences called recognition or degradation tags or degrons required for their recognition by adaptors and AAA+ proteases. The bacterial degrons help in substrate recognition in several ways: (i) direct binding with the pore of AAA+ hexameric rings, (ii) binding with an auxiliary site of core protease complex, (iii) interacting with adaptor proteins, which further delivers the substrate to protease complex (iv) mediates covalent modifications in the substrate to allow recognition by AAA+ proteases (Sauer & Baker, 2011). In a widespread phenomenon, the *E. coli* ClpXP recognizes and mediates the degradation of substrates with a C-terminal sequence AANDENYALAA-COOH, which is known as the SsrA-tag. The C-terminal dipeptide, AA of the SsrA-tag, is recognized by the AAA+ ClpX ring (Flynn et al., 2001; Levchenko et al., 2005). These ssrA-tags are one of the most extensively characterized C-terminal degrons in several bacterial species.

A general mechanism in bacteria to rescue stalled ribosomes involves a small stable RNA encoded by the *ssrA* gene. A complete description of SsrA-mediated substrate degradation is given in **Figure 1.13**. The small stable RNA A (ssRNA) was discovered in *E. coli* when a 10S RNA fraction was resolved into two species, the 10Sa or SsrA (small stable RNA A) molecule and the 10Sb RNA, which is the catalytic subunit of ribonuclease P (Subbarao & Apirion, 1989). Sequence similarities between parts of SsrA and tRNAs were initially recognized based on the *ssrA* gene sequence from *Mycobacterium tuberculosis* (Tyagi & Kinger, 1992). The 10Sa RNA encodes a proteolytic peptide tag (SsrA tag), which is co-translationally added to truncated nascent polypeptides, thereby targeting them for rapid proteolysis to prevent the accumulation of aberrant, unfinished proteins. Thus, SsrA RNA shows both tRNA and mRNA-like domains and is called transfer-messenger RNA (tmRNA). The tmRNA has a dual function of liberating stalled ribosomes and tagging the incomplete nascent proteins for degradation. The tmRNA first recognizes the stalled ribosomes with its tRNA-like domain and occupies the A-site of the stalled ribosomes. Further, the mRNA-like domain provides a template to continue the translation process, adding a short peptide stretch (SsrA tag) to the growing polypeptide chains (Moore & Sauer, 2007). The SsrA tag, usually 10-11 amino acids long, causes degradation by providing a recognition site for AAA+ proteases. In *E. coli*, the sequence ANDENYALAA is added to the C-terminus of the nascent polypeptide (Tu et al., 1995).

TH-3792_206106022

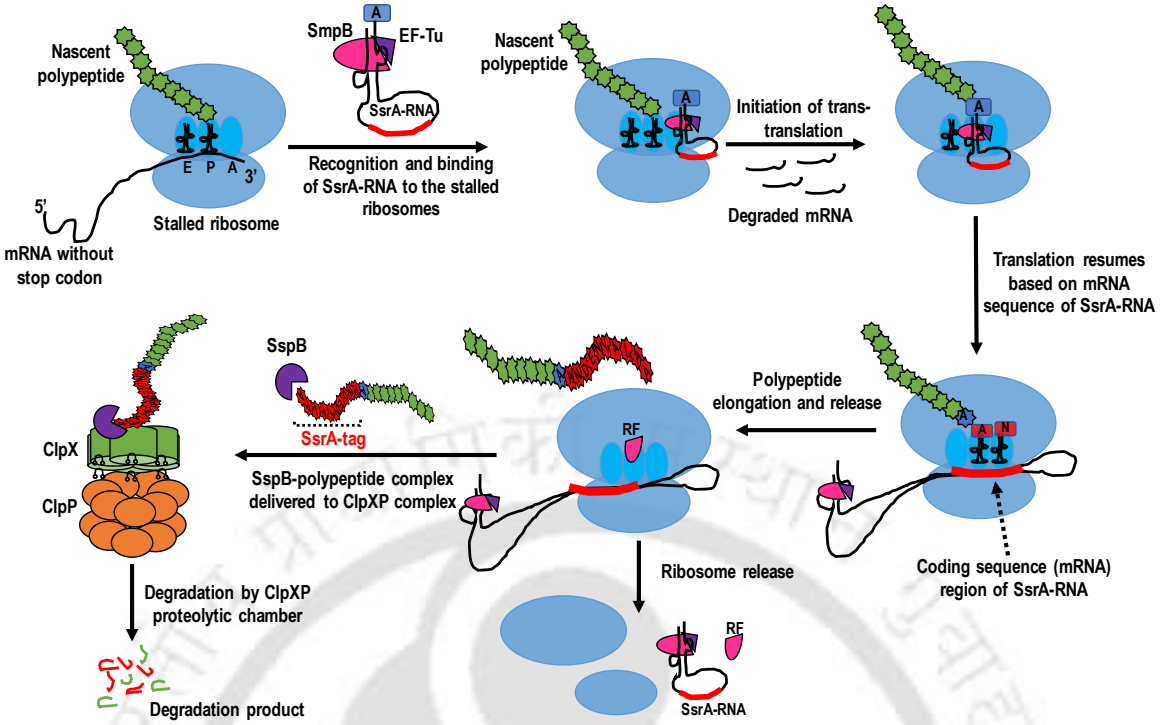


Figure 1.13. Schematic representation of trans-translation (SsrA-RNA mediated ribosome rescue). The ribosome stalls when it encounters a damaged mRNA (without a stop codon at the 3' end) recognized by the SmpB-SsrA-RNA^{Ala} complex. The ternary complex (SmpB + SsrA-RNA^{Ala} + EF-tu elongation factor) binds to the A-site of the ribosome, and the translation process resumes using the mRNA domain of SsrA-RNA as a template. The nascent polypeptide chain tagged with SsrA sequence (AANDENYALAA) was recognized by adaptor protein (SspB), which delivers the SsrA-tagged polypeptide to the ClpXP proteolytic chamber (adapted from Fritze et al., 2020).

The LAA sequence at the C-terminal of the SsrA sequence is conserved among various species, while the rest of the sequence may or may not vary. The significance of the conserved LAA sequence was demonstrated by comparing the degradation of model proteins with the tags AANDENYALAA and AANDENYALDD. The model proteins with LAA at their C-terminal were rapidly degraded, while those containing LDD were not degraded (Keiler et al., 1996). At least five proteases are involved in the degradation of SsrA-tagged protein in *E. coli* (Lies & Maurizi, 2008). Amongst them, two intracellular AAA+ proteases, ClpXP and ClpAP, degrade most of the SsrA-tagged proteins (Gottesman et al., 1998). The hexameric form of the AAA+ component of the ClpXP or ClpAP complex recognizes and unfolds the SsrA-tagged proteins using ATP hydrolysis. The ClpXP complex targets the majority of the fraction of SsrA-tagged substrates, but some studies suggest the significance of ClpAP in the degradation of SsrA-tagged proteins. This uncertainty is mainly because of the involvement of two adaptor proteins, ClpS and SspB, in modulating the proteolysis of the SsrA-tagged

TH-3792_206106022

substrates. The SspB is a ribosome-associated protein that is responsible for the degradation of SsrA-tagged substrates by EcoClpXP (Levchenko et al., 2005). The SspB is a dimeric protein that interacts with the N-terminal region of EcoClpX through its flexible C-terminal tail (Wah et al., 2003). Mutagenesis experiments also revealed that *E. coli* SspB makes key interactions with residues 1-4 (AAND) and 7 (Y) of the *E. coli* SsrA tag (Flynn et al., 2001). The EcoSspB adaptors act as a molecular mediator that binds to EcoClpX with one hand and SsrA-tagged substrates with another hand and finally delivers the substrate to the EcoClpXP proteolytic chamber (Wah et al., 2003). The presence of SspB increases the degradation rate for SsrA tag substrates, and it was also reported that SspB decreases the Km (increases affinity) for SsrA-tagged substrate degradation (Levchenko et al., 2000). Also, SspB competitively inhibits SsrA-tagged substrate degradation by ClpAP (Flynn 2001). Thus, for SsrA-tagged substrates, EcoSspB acts as an enhancer of EcoClpXP recognition and an inhibitor of EcoClpAP recognition.

1.5.2. The Bacterial N-Degron

A study by Alexander Varshavsky's group discovered that the stability of proteins depends on their N-terminal residue and proposed the N-end rule (Bachmair et al., 1986). Substituting the N-terminal residues of beta-galactosidase with specific residues has stimulated its degradation, while placing other amino acids did not have the same impact. Thus, the residues that promote rapid protein degradation are termed destabilizing residues, and the ones that prevent a protein from degradation are termed stabilizing residues. The N-end rule pathway has been studied in several organisms, including *E. coli*, plants, and mammals (Tobias et al., 1991; Bachmair et al., 1993; Graciet et al., 2009; Gonda et al.,1989). The destabilizing residues are classified into various sub-categories (primary, secondary, or tertiary) (Varshavsky, 1996). The prokaryotic N-end rule pathways harbour a set of enzymes for recruiting primary and secondary destabilizing residues, while the eukaryotic systems have an additional tertiary destabilizing residue. The enzymes responsible for modifying the N-terminal residue of proteins are different for eukaryotes and prokaryotes (Dougan et al., 2010). For example, in eukaryotes, the secondary destabilizing residues are recruited with the help of enzyme N-terminal amidase (NTAN1 or NTAQ1), while the primary destabilizing residue is attached by arginyl-tRNA protein transferases (ATE1) (Grigoryevet al., 1996; Wang et al., 2009; Balzi et al., 1990). However, in prokaryotes, the primary destabilizing residues are attached with the help of the enzyme Leucine/phenylalanine-tRNA protein transferase (LFTR) (**Figure 1.14**). The LFTR was first discovered in *E. coli* as a protein that mediates the transfer of Leu or Phe

TH-3792_206106022

from tRNA to a protein without the involvement of ribosomes (Leibowitz & Soffer, 1969). Later, the direct role of LFTR in the *in-vivo* stability of a protein was reported (Tobias et al., 1991).

Several studies using model substrates have identified two secondary destabilizing residues, Arg and Lys, in *E. coli*. However, the first identified natural substrate of EcoLFTR is putrescine aminotransferase (PATase), in which the primary destabilizing residue (Leu or Phe) is added upon identifying the secondary destabilizing residues (Met) (Ninnis et al., 2009). Thus, apart from Arg and Lys, the amino acid Met can act as a secondary destabilizing residue for EcoLFTR. Apart from *E. coli* LFTR, two other aminoacyl-transferases have also been reported with diverse specificity (**Figure 1.14**). The *Plasmodium falciparum* ATEL1 mediates the transfer of the primary destabilizing residue (Arg) to the acidic secondary destabilizing residues (Graciet et al., 2006). The other enzyme discovered is bacterial protein transferase (Bpt) in *Vibrio vulnificus*, which mediates transfers of Leu to the acidic N-termini of proteins (Graciet et al., 2006).

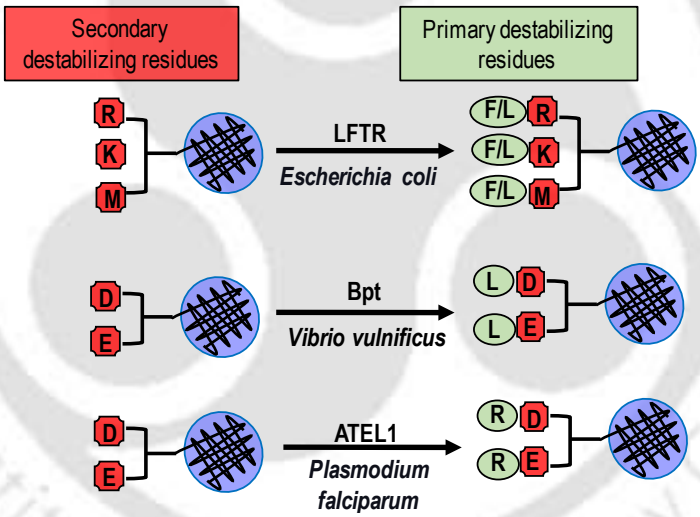


Figure. 1.14. The N-end rule pathway of bacteria. In *E. coli*, the primary destabilizing residues, F/L (green), are attached to the secondary stabilizing residues, R, K, and M (red) by Leu/Phe tRNA transferase (LFTR). In *V. vulnificus*, the primary destabilizing residue, L (green), is attached to the secondary destabilizing residues, D and E (red), by the bacterial protein transferase enzyme (Bpt). In *P. falciparum*, the primary destabilizing residue, R (green), is attached to the secondary destabilizing residues, D and E (red) by the ATEL1 enzyme (Dougan et al., 2010).

The recognition of a protein substrate tagged via the N-end rule pathway was mediated by a class of proteins classified as N-recognins. In eukaryotes, four different recognins have been discovered (UBR1, UBR2, UBR4, and UBR5), while bacterial species show the presence of a single recognin, ClpS, for recognizing the primary destabilizing residues of protein substrates

TH-3792_206106022

(Tasaki et al., 2005; Erbse et al., 2006; Tasaki et al., 2009). The *clpS* gene is present upstream of the *clpA* gene in *E. coli*, encoding a protein of 12 kDa. Initial studies of ClpS in *E. coli* found that it interacts with EcoClpAP, inhibiting ClpA auto-degradation and directing EcoClpAP towards artificial protein aggregates (Dougan et al., 2002). Also, EcoClpS was reported for recognizing N-end rule substrates and directing them to the EcoClpAP complex for degradation (Erbse et al., 2006). The structures of *C. crescentus* and *E. coli* ClpS have been solved, revealing how ClpS binds to the N-end rule substrates and interacts with ClpA (Guo et al., 2002; Román-Hernández et al., 2009). Structural studies on ClpS reveal it contains a well-folded core domain (ClpS core) comprising three α-helices and three β-sheets. In addition, the N terminus of ClpS is disordered in most crystal structures. The first 25 N-terminal residues of EcoClpS are termed as N-terminal extension (NTE) and shown in **Figure 1.15A**. Structures, along with mutagenic studies, show that a patch of charged residues on the α3 helix of EcoClpS interacts with the flexible N-terminal domain of EcoClpA (Guo et al., 2002; Zeth et al., 2002; Wang et al., 2008). In particular, the residues Glu79, Glu82, and Lys84 of EcoClpS interact via two salt bridges and two hydrogen bonds with the charged residues (Thr81, Glu23, Glu28, and Arg86) on EcoClpA N-domain (Schuenemann et al., 2009). Further, the EcoClpS has a conserved hydrophobic pocket which provides binding sites for the hydrophobic side chains of N-degrons (Phe, Tyr, Trp, and Leu). Also, three residues in the EcoClpS pocket (Asn34, Asp36, and His66) form a conserved hydrogen bond network with the α-amino group of the N-degron (Román-Hernández et al., 2009). The importance of these H-bonding has been reported by several mutational studies where mutations of Asn34, Asp36, or His66 to Ala impede the degradation of model N-end rule substrates (Wang et al., 2008). Apart from the N-terminal amino acids of N-degrons, the second and third residues from the N-terminal of substrates are also reported to have a role in interaction with EcoClpS (Wang et al., 2008). Also, it was observed that EcoClpS shows a higher preference towards Leu than Met residue, although both Leu and Met have similar side-chain organization. A peptide with N-terminal Met binds to EcoClpS with 1000 times less affinity than a peptide with N-terminal Leu (Román-Hernández et al., 2009). The recognition of N-degron by EcoClpS must be followed by its handover from EcoClpS binding pocket to the EcoClpAP catalytic chamber for degradation (Wang et al., 2007) (**Figure 1.15B**). When EcoClpA, EcoClpS, and N-end rule peptides interact, they form a high-affinity ternary complex (Román-Hernández et al., 2011), and the disruption of this complex is essential for substrate delivery to the EcoClpP complex. This efficient delivery process is mediated by the N-terminal extension (NTE) of EcoClpS in the ternary complex. The chaperone, EcoClpA, interacts with EcoClpS NTE, which allows a

TH-3792_206106022

structural rearrangement in EcoClpS-bound N-degron and thus permits the translocation of the N-end rule substrate inside the core catalytic chamber of EcoClpAP (Dougan et al., 2012; Rivera-Rivera et al., 2014).

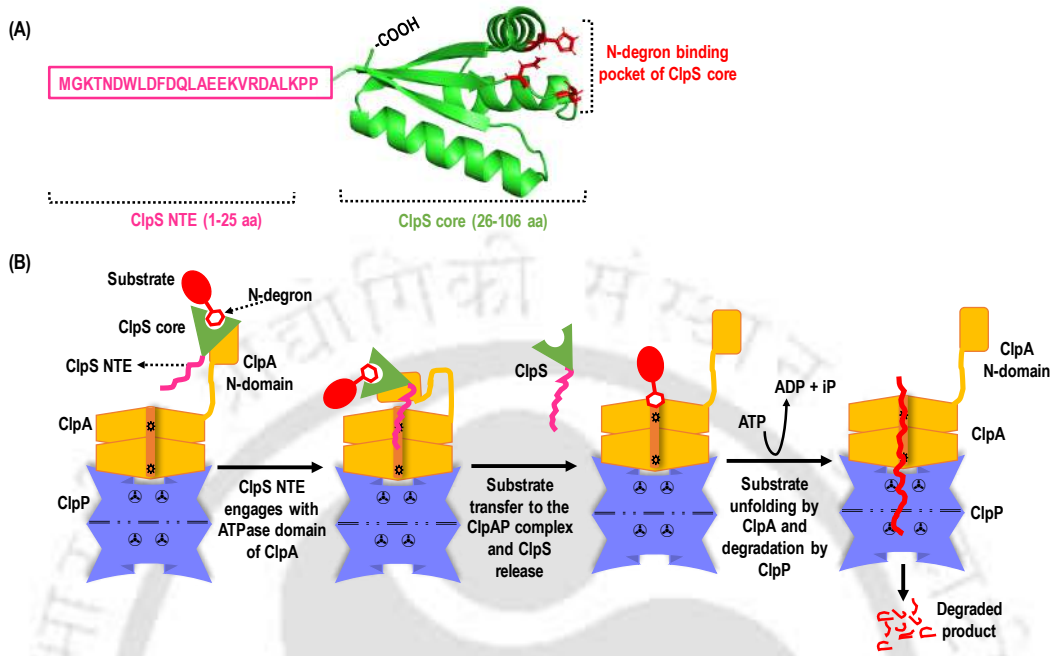


Figure 1.15. Model for ClpS-mediated delivery of N-degron substrate to ClpAP complex. **(A)** The bacterial N-recognin (ClpS) has a flexible N-terminal extension (NTE; 1-25 aa) and a folded core domain (26-106 aa). The crystal structure of ClpS core (PDB: 3O2O) has three α-helices and three β-sheets with N-degron peptide binding pocket residues marked in red color (Asn34, Asp36, and His66). **(B)** Complex formation between ClpA, ClpS, and N-degron substrate. The N-degron-tagged substrate (red) binds to the ClpS core region (green), followed by ClpS core binding to the ClpA N-domain (orange). Next, the ClpS NTE (pink) engages with the ATPase domain of ClpA, followed by substrate delivery to the ClpAP complex and release of ClpS adaptor. The N-degron substrate undergoes proteolysis by ClpAP complex in an energy-dependent manner (adapted from Rivera-Rivera et al., 2014).

1.6. Leptospirosis

A global zoonosis originated from the contamination of pathogenic *Leptospira* species has been re-emerging and is affecting almost all mammals (Adler & de, 2010). These zoonotic infections are known as leptospirosis, leading to about half a million severe cases yearly worldwide, with a mortality rate of 10% (Karpagam & Ganesh, 2020). The entry of *Leptospira* into the host cell is facilitated by damaged skin cells or wounds, followed by a wide range of mild symptoms like fever, chills, and headache. Severe clinical manifestations may result in pulmonary haemorrhage syndrome and Weil's syndrome. Leptospirosis also causes many reproductive disorders in farm animals (e.g., horses, cows, pigs, and sheep), leading to economic loss (Bharti et al., 2003). Cases of leptospirosis have been reported in rural and urban areas in tropical, subtropical, and temperate regions. Most human cases are associated with

TH-3792_206106022

factors such as sanitation, housing, income level, occupation, and travel exposure (Hartskeerl et al., 2011). The *Leptospira* was first discovered around 100 years ago by Dr Weil (Professor of Medicine at Heidelberg). He specified that a disease called Weil syndrome originated from *Leptospira interrogans* serovar Copenhageni (Inada et al., 1916). After discovery, the in-vitro culturing of leptospires had not been observed until 1907, but in 1916 (first in Japan), it was isolated and cultivated (Kumbhare et al., 2019). Leptospirosis has been declared a waterborne infection based on several outbreaks observed during the rainy season. During monsoon, the disease has been declared an epidemic in areas with high rainfall, which signifies a direct correlation between rainfall and the frequency of leptospirosis. Leptospires grow at an optimum temperature of 28°C to 32°C and survive for several weeks under suitable environmental conditions (Kumbhare et al., 2019). These spirochetes can be transmitted directly or via an indirect mode among animals. The transmission is mainly facilitated by direct contact through injured skin, exposure to contaminated urine, or ingestion of infected tissues. *Leptospira* can be sporadically detected in the urine of recovered dogs for months to years following infection. In humans, leptospirosis was connected with pulmonary haemorrhage syndrome (Bharti et al., 2003; Antima & Banerjee, 2023).

1.6.1. Impact of Leptospirosis

It has been estimated that worldwide, there are 1.3 million cases and 58,900 deaths due to leptospirosis annually (Costa et al., 2015 Karpagam & Ganesh, 2020). The estimated disease burden occurs mainly in tropical regions and the world's poorest countries. Adult males are the leading risk group for leptospirosis. It is estimated that morbidity and mortality are high in regions like South and Southeast Asia, where leptospirosis is an under-recognized public health problem. Leptospirosis predominantly affects males (80% of the total burden) and young adults aged 20-49 (52% of the total burden) (Torgerson et al., 2015). Numerous outbreaks have been documented globally following natural disasters in urban areas, including in Guyana, India, Italy, Indonesia, Malaysia, and the Philippines (Dechet et al., 2012; Sehgal et al., 2003; Al-shere et al., 2012). Endemic transmission has also been reported in regions such as Thailand, Italy, New Caledonia, Vietnam, Brazil, Laos, Mexico, Germany, Argentina, Sri Lanka, Korea, and California (Chadsuthi et al., 2021; Goarant et al., 2009; Galan et al., 2021; Vado-Solis et al., 2002; Schmidt et al., 2021; Warnasekara et al., 2019; Kim, 2019; Meites et al., 2004).

Leptospirosis is most prevalent in coastal areas. Most outbreaks of Leptospirosis in India are reported in the coastal regions of Gujarat, Maharashtra, West Bengal, Orissa, Kerala, Tamil Nadu, Karnataka, and the Andaman Islands. A study in India found that leptospirosis accounted

TH-3792_206106022

for about 12.7% of hospital acute febrile illness cases (Sehgal et al., 2003). Surveillance data from 1988 to 2005 in the Andaman and Nicobar Islands reported 1,126 cases, while Gujarat reported 3,121 cases with 383 deaths from 1997 to 2005. Chennai, Tamil Nadu, reported 5,452 cases between 2004 and 2006, and Maharashtra reported 4,484 cases with 394 deaths from 1998 to 2005 (Antima & Banerjee, 2023). Other states, including Punjab, Haryana, Himachal Pradesh, and Karnataka, have also reported cases (Antima & Banerjee, 2023). The predominant *Leptospira* serovars causing leptospirosis in India include Copenhageni, Autumnalis, Canicola, Pyrogenes, Grippotyphosa, Australis, Javanica, Sejroe, Louisiana, Valbuzzi, and Pomona (Krishnan et al., 2024).

1.6.2. *LEPTOSPIRA*, The Causative Agent

Leptospires belong to the family Leptospiraceae, order Spirochaetale, and are thin, obligate aerobes, fine spiral-shaped organisms of about 0.1 mm diameter and 6–20 mm long (Faine, 1994). They have periplasmic flagella at both ends, which provide specific corkscrew motion to the bacteria and differentiate them from other spirochetes. These bacteria are visualized under a dark-field microscope without staining (Miguel et al., 2020). These bacteria are obligate aerobes that utilize long-chain fatty acids as carbon sources. The *Leptospira* requires a specific enrichment medium derived from rabbit or cattle serum for optimal growth under in vitro conditions (Murray et al., 2015; Adler & de, 2010). Unlike other spirochetes such as *Treponema* and *Borrelia*, *Leptospira* possesses surface-expressed lipopolysaccharides (LPS), a key determinant of their pathogenic potential. Additional virulence factors include outer membrane proteins, hemolysins, surface proteins, and adhesins, contributing to host interactions and disease progression (Adler & de, 2010; Karpagam & Ganesh, 2020).

Recent phylogenetic analysis using whole-genome sequences (previously based on 16S rRNA gene sequencing) has reclassified *Leptospira* into two major clades: saprophytic and pathogenic (Vincent et al., 2019). The saprophytes include species that inhabit natural environments and do not cause infections. In contrast, the pathogenic *Leptospira* species causes infections in humans and animals. To date, 64 *Leptospira* species have been identified. These species are further divided into 20 serogroups and over 300 serovars based on variations in their lipopolysaccharide (LPS) layer (Picardeau, 2017; Vincent et al., 2019; Picardeau, 2020). The differentiation of several serovars is based on their surface-expressed molecules, which are constituents of its outer LPS layer. The antigenicity of each serovar depends on the composition of polysaccharides within the LPS (Miguel et al., 2020). The Leptospires are also characterized by the presence of periplasmic flagella, which originate from the periplasmic

TH-3792_206106022

space. FlaA and FlaB are the major constituents of periplasmic flagella, which provide motility to the spirochetes (Picardeau et al., 2001).

1.6.3. Transmission of Leptospirosis

The transmission cycle of leptospirosis involves the asymptomatic hosts (reservoir), the environment, and the susceptible host (human beings and domestic animals) (**Figure 1.16**). Almost every known species of rodents and mammals can be carriers and excretors of leptospires (Faine, 1994). Leptospires are excreted into the environment from the renal tubes of the excretor animals. Transmission can be direct or indirect. Direct transmission occurs when leptospires from tissues, body fluids, or urine of acutely infected or asymptomatic carrier animals enter the body of the new host and initiate infection. Direct transmission among animals can be transplacental, haematogenous, by sexual contact, or by suckling milk from infected mothers (Faine, 1994). The rodents are the main reservoir of leptospires, which rarely develop the disease (leptospirosis). The renal system of rodents has alkaline pH conditions, which prevent bacterial survival and allow the disposal of leptospires into the environment (soil, water) through urine (Miguel et al., 2020). Upon entry into a host, the *Leptospira* invades the tissue and vascular system and begins to damage the endothelium layer of blood vessels. This causes bacteria to release into the bloodstream within the first few days of bacterial infection (Haake & Levett, 2015). In severe cases, leptospires colonize within the renal proximal tubule, causing a chronic tubulo-interstitial inflammation.

1.6.4. Treatment

The conventional antibiotic treatment approach for leptospirosis in case of early disease diagnosis should be administered within 7-10 days of infection (Watt et al., 1988). Leptospires are sensitive to most antibiotics, as tested in the laboratory, but can resist antibiotics like rifampicin and chloramphenicol. The most extensively administered antibiotic in high doses is penicillin, and in patients with hypersensitivity towards penicillin, erythromycin has been recommended. Tetracycline can also be used, but it has a limitation as it cannot be provided to young children and pregnant women (Griffith et al., 2006). No antibiotic can overturn the chronic state of infection when leptospires invade the tissues and organs. Still, penicillin was found to have valuable effects, reducing mortality and the extent of illness in severe leptospirosis when given intravenously. The antibiotic benzylpenicillin is administered by injection in doses of 5 million

TH-3792_206106022

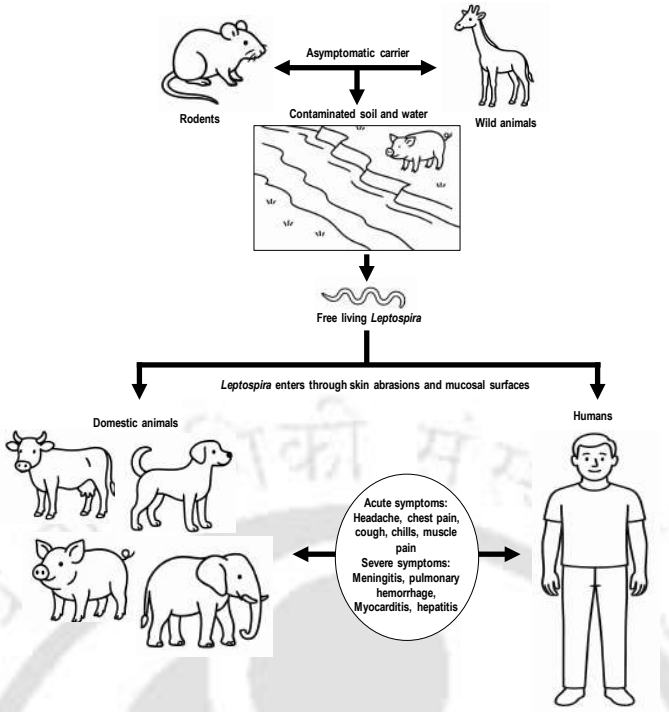


Figure 1.16. Transmission cycle of leptospirosis. Rodents and other wild animals are asymptomatic carriers or reservoirs of leptospires. The carrier host releases the leptospires into water and soil, and from there the free leptospire enters the human or other domestic animals' body via skin abrasions or mucous surfaces, causing the initiation of acute symptoms (adapted from Miguel et al., 2020).

units per day for five days. Erythromycin 245 mg is administered four times daily for five days to patients hypersensitive to penicillin. Doxycycline 100 mg twice daily for ten days is also recommended for patients with mild leptospirosis (Griffith et al., 2006). Leptospirosis, under severe conditions, may result in kidney failure. These critical conditions require dialysis or the use of an artificial kidney, as the administration of antibiotics is not sufficient (Levett, 2015). Most of these antibiotics target cellular processes like DNA replication, translation, and cell wall formation, which are vital for actively dividing cells. To survive the host defence mechanism, many bacterial species have adapted themselves to undergo a dormant, non-replicative state (Chan et al., 2015). These non-growing persistent cells are thus critical for antibiotic treatment (Balaban et al., 2019). Also, continuous and unmanaged antibiotic usage may drive resistance towards conventional antibiotics (Kåhrström, 2013). Therefore, recent searches should focus on finding alternative approaches to combat leptospires even in their persistent state.

1.6.5. Alternative Approaches of Treating Leptospirosis

Vaccination is the most viable strategy for controlling infectious diseases, including leptospirosis. However, even after more than 30 years of discovery, designing

TH-3792_206106022

vaccines for leptospires serves as a significant challenge. Due to the presence of more than 300 serovars of *Leptospira* species with different virulence factors and pathogenicity, it is difficult to use a single universal vaccine targeting all (Teixeira et al., 2019). Recent studies have shown the development of new vaccines wherein the LPS layer and outer membrane proteins (OMPs) have emerged as potential vaccine candidates. Several categories of vaccines, including recombinant, subunit, and killed, have provided partial immunity to the disease but are ineffective in conferring long-term immunity. In recent years, alternatives have been investigated, including novel antibiotics, antibodies, probiotics, bacteriophages, and antimicrobial peptides (Ghosh et al., 2019). Anti-virulence approaches, phage therapy, and therapeutic antibodies are recent methodologies that may gain advantages over conventional treatment strategies. In this context, Caseinolytic proteases P (ClpP) have garnered considerable attention as targets for antibacterial action after their direct relationship with bacterial virulence was proven in gram-positive and gram-negative bacteria (Culp et al., 2017; Bhandari et al., 2018; Gaillot et al., 2000). The bacterial ClpPs are essential in every aspect of life as cells, tissues, and organelles are full of proteins and their controlled regulation is a primary concern for maintaining a healthy cellular condition. Bacterial ClpPs are an important class of proteases required for cellular homeostasis and protein turnover to maintain vital cellular functions, particularly under stress conditions. The ClpP plays a significant role in controlling several cellular processes like cell motility, cell sporulation, cell division, and differentiation, and is conserved among diverse bacterial species, including human pathogens, which has led to its consideration as a potential antibacterial target (Thomy et al., 2019; Conlon et al., 2013; Famulla et al., 2016).

1.7. The Caseinolytic Proteases of *Leptospira*

The *Leptospira interogans* serovar Copenhageni's genome sequence analysis predicts an array of genes related to the caseinolytic protease (Clp) system (Lourdault et al., 2011). **Figure 1.17** shows a schematic arrangement of genes associated with the leptospiral Clp system. This includes the core proteolytic component genes (*clpP1*, *clpP2*, and *clpQ*), the ATPase chaperones (*clpA*, *clpC*, *clpX*, and *clpY*), and the adaptor proteins (*clpS1*, *clpS2*, and *mcsB*).

TH-3792_206106022

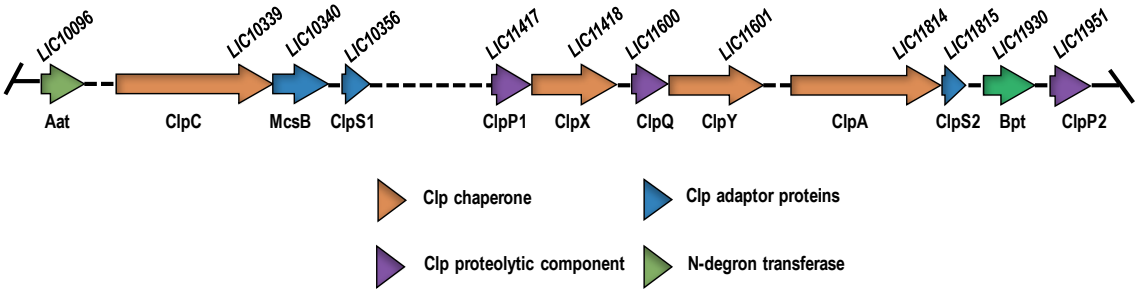


Figure 1.17. A schematic arrangement of genes related to the caseinolytic protease (clp) system in the *L. interrogans* genome. The various Clp system-associated genes are represented in the arrow diagram (drawn approximately to scale). The core Clp protease component (*clpP1*, *clpP2,* and *clpQ*; violet), Clp ATPase (*clpC*, *clpX*, *clpY,* and *clpA*; orange), clp adaptor proteins (*mcsB*, *clpS1,* and *clpS2*; blue), and N-degron transferase (*lftr* and *bpt*; green) are shown (adapted from Dhara et al., 2019).

Further, genes associated with post-translational transfer of N-degrons (*lftr* and *bpt*) were also predicted to be present in the leptospiral genome. The ClpC chaperones are usually absent in most gram-negative bacteria but present in spirochetes like *Leptospira interrogans*, *Treponema pallidum,* and *Borrelia burgdorferi*. The Clp ATPase proteins ClpA, ClpC, and ClpX possess the ClpP recognition binding motif composed of 3 amino acids (L/I/V)G(F/L), which may allow their association with the ClpP peptidase, resulting in a Clp proteolytic complex. In *Leptospira* spp., the genes for chaperone (*clpA*) and adaptor protein (*clpS2*) are adjacent on the chromosome, while another *clpS* ortholog (*clpS1*) was found elsewhere. Also, the genes associated with the chaperone *clpC* and its predicted adaptor protein (*mcsB*) are present together. However, no genes for adaptor proteins that were reported to associate with *E. coli* ClpX (SspB, RssB, and UmuD) have been identified in the *Leptospira* genome. Further, the genome analysis of *L. interrogans* serovar Copenhageni reveals that the *tig* gene encoding a ribosome-associated chaperone, trigger factor (TF), is located within the Clp array. The *tig* gene is located upstream of the *clpP1* gene, with a four-base pair overlap between the coding regions.

The pure isoforms of leptospiral ClpPs (LinClpP1 and LinClpP2) are enzymatically inactive. The heterocomplex LinClpP1P2 shows optimum peptidase activity on small model peptides (2-4 amino acids long) when mixed in equimolar concentration. The LinClpP1P2 heterocomplex shows protease activity against model protein substrate (casein) in the presence of LinClpX in an energy-dependent manner (Dhara et al., 2019). The LinClpP1 isoform was reported to exist in a higher oligomeric form containing 14-21 monomers, while LinClpP2 exists in a tetradecamer. The equimolar mixture of LinClpP1 and LinClpP2 associates with a 1:1 stoichiometry and forms a functional tetradecameric complex containing two heptameric

TH-3792_206106022

rings stacked together (Dhara et al., 2019). Additionally, the impact of antibiotic ADEP1 on the leptospiral ClpP system has been explored. The ADEP1 impaired the leptospire's growth and stimulated the protease activity of LinClpP1P2 heterocomplex without the involvement of cognate ATPase chaperone (Dhara et al., 2021). As the genes encoding leptospiral *clpP1* and trigger factor (*tig*) are located near each other, their functional association has been explored. The LinTF has been speculated to promote functional oligomerization of LinClpP isoforms, thereby enhancing the peptidase activity of LinClpP1P2 heterocomplex (Choudhary et al., 2021).

1.7.1. Research Gap

Leptospirosis is a serious economic problem in developing tropical and subtropical countries. Despite the severity of leptospirosis and its global burden, little is known about *Leptospira* pathogenesis and its protein quality control system. Thus, understanding the functionality of the leptospiral Clp system (a key regulator of the cell proteome) might help us to reduce the indiscriminate use of antibiotics for the treatment of leptospirosis by substituting with new therapeutic drugs directly affecting the functions of Clp proteins. This study focuses on the biochemical characterization of the Clp system-associated proteins, which can provide hints for designing novel antibiotics to intervene or accelerate the protein degradation process, leading to the self-destruction of bacteria. The current study aims to characterize the roles of the key structural motifs of the core proteolytic proteins (LinClpP1 and LinClpP2) and the individual domains of the ATP-independent chaperone, LinTF. Also, the role of ClpP regulatory proteins (LinClpA or LinClpC) has not been studied in *Leptospira*. Further, the nature and mechanism of substrate specificity of Clp chaperones and adaptor proteins in *L. interrogans* remained unexplored. The findings from this study provide a better understanding of the physiological role of Clp chaperone and its cognate adaptor proteins in modulating the ClpP protease activity of *Leptospira*. Thus, targeting the leptospiral Clp chaperone (LinClpC or LinClpA) or adaptor proteins (LinClpS1 and LinClpS2), which intervene in the substrate specificity of LinClpPs, could be a novel strategy for leptospirosis treatment. Also, studying the functional mechanism of these protein quality control enzymes might help us understand this bacterium's pathogenesis.

TH-3792_206106022

1.7.2. Objectives of the study

1. Cloning, over-expression and purification of *L. interrogans* caseinolytic adaptor proteins; ClpS1 (LIC11356), ClpS2 (LIC11815) and its cognate ATPase chaperones; ClpA (LIC11814), ClpC (LIC10339).
2. Biochemical characterization of purified Clp chaperones (ClpA and ClpC) and their association with leptospiral ClpP isoforms.
3. Understanding the role of adaptor proteins (ClpS1 and ClpS2) in N-end rule degradation pathways of spirochetes and their association with caseinolytic chaperones and protease complex.
4. Understanding the association of *Leptospira* trigger factor TF with the caseinolytic chaperones, ClpA and ClpC.

TH-3792_206106022

Chapter 2

Biochemical characterization of key structural motif mutants of leptospiral ClpP isoforms revealed unprecedented gain-of-function

This chapter is a modified version of the manuscript published as:

Kumari, S., Dhara, A., & Kumar, M. (2024). *Leptospira* ClpP mutant variants in association with the ClpX, acyldepsipeptide, and the trigger factor displays unprecedented gain-of-function. ***International Journal of Biological Macromolecules***, *254*, 127753.

TH-3792_206106022

2.1. Abstract

The tetradecameric active caseinolytic protease (LinClpP) of *Leptospira interrogans* is composed of two different isoforms (LinClpP1 and LinClpP2). In this study, five mutants of LinClpP (LinClpP1^{E170D}, LinClpP1^{N172D}, LinClpP2^{IG_del}, LinClpP2^{S40AK41N}, LinClpP2^{Y62A}) targeting its critical hotspot residues were generated. The functional activity of pure LinClpP mutant variants or its heterocomplex and its effect when associated with a chaperone (LinClpX)/antibiotic acyldepsipeptide (ADEP1)/trigger factor (LinTF) was examined. The two mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}) displayed gain-of-function (GOF) in peptidase activity. The ADEP1-bound heterocomplex (LinClpP1P2^{S40AK41N} and LinClpP1P2^{Y62A}) measured 1.7 and 1.5-fold higher protease activity than ADEP-bound LinClpP1P2. The dynamic light scattering analysis of ADEP1-bound GOF mutants displayed increased hydrodynamic diameter. In the presence of LinTF, the heterocomplex (LinClpP1P2^{S40AK41N} and LinClpP1P2^{Y62A}) exhibited a 3-fold surge in peptidase activity. The deletion mutant (LinClpP2^{IG_del}) or its heterocomplex (LinClpP1P2^{IG_del}) displayed no activity. Similarly, the pure LinClpP1^{E170D} and LinClpP1^{N172D} could not cleave a model dipeptide. However, its heterocomplex (LinClpP1^{E170D}P2 and LinClpP1^{N172D}P2) showed 0.5-fold lower peptidase activity than the LinClpP1P2. Collectively, two mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}) have GOF and can degrade the model dipeptide substrate without the aid of the LinClpP1 isoform and thus provide new insights into unprecedented LinClpP activation.

TH-3792_206106022

2.2. Introduction

The degradation of proteins is a critical process mediated by several regulatory proteins and proteases to control and maintain protein homeostasis (Sauer et al., 2004). The caseinolytic serine protease (ClpP) system in bacteria is generally composed of a core protease component (ClpP) along with its cognate chaperones and adaptor proteins (Sauer et al., 2004; Martin et al., 2008). The monomeric ClpP orthologs consist of three folded domains, namely the N-terminal region, the globular head region, and the handle domain, wherein the head region forms the ClpP body (Bewley et al., 2006; Alexopoulos et al., 2012). Multiple monomeric subunits of ClpP self-assemble into two heptameric rings. These heptameric rings then develop a functional homo or hetero tetradecamer depending on the number of ClpP isoforms (Alexopoulos et al., 2012). Microorganisms like *Escherichia coli* (EcoClpP), *Staphylococcus aureus* (SauClpP), *Bacillus subtilis* (BsuClpP) possess a single ClpP isoform and tend to self-assemble and form homo tetradecamer ClpP (Maurizi et al., 1994; Zhang et al., 2011; Krüger et al., 2000). While few microorganisms like *Leptospira interrogans* (LinClpP1P2), *Listeria monocytogenes* (LmoClpP1P2), *Pseudomonas aeruginosa* (PaeClpP1P2), *Mycobacterium tuberculosis* (MtuClpP1P2), *Chlamydia trachomatis* (CtrClpP1P2), and *Clostridium difficile* (CdiClpP1P2) has two ClpP isoforms and thus, can form hetero tetradecamer ClpPs (Dhara et al., 2019; Gaillot et al., 2000; Hall et al., 2017; Raju et al., 2012; Pan et al., 2019; Lavey et al., 2018). The basic framework of homo- or hetero ClpP tetradecamer consists of a catalytic triad (Ser, His, and Asp) essential for proteolysis (Wang et al., 1997), a central E-helix which is highly flexible (Kang et al., 2004; Gribun et al., 2005), a Gly-rich loop required for inter-ring communication (Gersch et al., 2012) and an outward projecting N-terminal region for interaction with its cognate Clp-ATPase chaperone (Gribun et al., 2005). In addition, a conserved aspartate/arginine sensor (Asp170/Arg171) of SauClpP is reportedly involved in tetradecamer assembly and the proper alignment of the catalytic triad (Gersch et al., 2012; Ye et al., 2013).

The caseinolytic protease (ClpP) complex is an energy-dependent system, and thus, it associates with the cognate hexameric ATPase chaperones such as ClpA, ClpX, or ClpC to become functional (Olivares et al., 2018). The EcoClpP homo tetradecamer without its ATPase chaperone has restricted axial pore size and only allows entry of small peptide fragments (up to 5 amino acids) for degradation (Woo et al., 1989). To gain protease activity, the EcoClpP homo tetradecamer associates with its chaperone (ClpX or ClpA) and forms an active EcoClpXP or EcoClpAP complex (Woo et al., 1989; Hoskins et al., 1998). The binding of

TH-3792_206106022

ATPase chaperones to the apical surface of EcoClpP as a hexameric assembly induces the widening of the axial pore and allows entry of unfolded polypeptides of considerable size (Gribun et al., 2005; Hoskins et al., 1989; Singh et al., 1998; Martin et al., 2007). In the homo-tetradecameric EcoClpP, two ATPases hexamer docks, one at each side of ClpP heptamers (Beuron et al.,1998; Grimaud et al., 1998). While the ClpPs from bacteria expressing single isoforms have been studied in detail, the systems with more than one ClpP isoform are less explored. To date, the peptidase activity of ClpP on model dipeptide (Suc-LY-AMC) from multi-isoforms expressing bacteria revealed a functional discrepancy. For LinClpP, MtuClpP, and CtrClpP, the two isoforms (ClpP1 and ClpP2) assemble into a hetero-tetradecamer complex to gain functionality (Dhara et al., 2019; Pan et al., 2019; Schmitz et al., 2014). While ClpP1 isoform alone for CdiClpP and PaeClpP and ClpP2 isoform alone for LmoClpP were reportedly active (Gaillot et al., 2000; Hall et al., 2017; Lavey et al., 2018).

Among hetero tetradecamers, the disparity in regulation by their interacting cognate ATPase chaperones is also observed. For MtuClpP1P2, LmoClpP1P2, and CtrClpP1P2, the cognate ATPase chaperone interacts exclusively with the ClpP2 apical hydrophobic pockets (Pan et al., 2019; Schmitz et al., 2014; Leodolter et al., 2015; Gatsogiannis et al., 2019). While for the PaeClpP1P2, the ATPase chaperone interacts with the ClpP1 isoform (Hall et al., 2017). In addition, acyldepsipeptide (ADEP), a natural antibiotic obtained from *Streptomyces hawaiiensis,* has also been proven to regulate the ClpP proteolytic activity (Brötz-Oesterhelt et al., 2005; Kirstein et al., 2009). The ADEP interaction has reprogrammed the EcoClpP and BsuClpP homo tetradecamer into an unregulated active complex that can degrade unfolded and folded native proteins without the aid of any cognate ATPase chaperones (Brötz-Oesterhelt et al., 2005; Kirstein et al., 2009). In BsuClpP homo tetradecamer, the ADEP binding disrupts the hydrophobic interactions within the BsuClpP body, resulting in a conformational twist of the N-terminal regions. Such a twist of the N-terminal part in the BsuClpP tetradecamer enforces axial pore widening in the BsuClpP machinery and allows uncontrolled substrate entry (Lee et al., 2010). In the case of ADEP-bound heterocomplex, the ADEP exclusively docks at the apical surface of the MtuClpP2 ring and allows the existence of both heptameric rings (MtuClpP1 and MtuClpP2) in an open axial pore state (Schmitz et al., 2014). Previous studies on LinClpP1P2 and MtuClpP1P2 suggest that ADEP can also reprogram the hetero tetradecameric ClpP1P2 into an unregulated complex (Schmitz et al., 2014; Dhara et al., 2021). In the *Leptospira genome,* the ClpP complex consists of a core LinClpP1P2 heterocomplex (LinClpP1 heptameric ring toppled on LinClpP2 heptameric ring), hexameric cognate ATPase chaperone (ClpA/ClpC/ClpX), and adaptor proteins (Dhara et al., 2019). Various biochemical

TH-3792_206106022

studies suggest that the sole LinClpP isoforms are inactive on model peptide substrates, but not its heterocomplex LinClpP1P2. Moreover, for the protease activity, the heterocomplex LinClpP1P2 requires the association of ATPase chaperone LinClpX or the antibiotic ADEP1 (Dhara et al., 2019; 2021). Recently, the functional link of LinClpP1P2 and EcoClpP with trigger factor (TF) has been reported, which shows that TF can enhance the ClpP protease activity (Choudhury et al., 2021; Rizzolo et al., 2021). Nevertheless, the functional involvement of crucial structural motifs in the core LinClpP protease has not yet been explored. Thus, a comprehensive analysis of LinClpP structural motifs was executed. Based on the multiple sequence alignment with the available well-studied ClpP orthologs, the LinClpP isoforms exhibit sequence disparity within the N-terminal hydrophobic cluster, glycine-rich loop, catalytic triad, Tyr activation trigger, and oligomerization sensor domain, referred here as critical hotspots residues. The five mutants of LinClpP (LinClpP1^{E170D}, LinClpP1^{N172D}, LinClpP2^{IG_del}, LinClpP2^{S40AK41N}, LinClpP2^{Y62A}) targeting the critical hotspot residues in the two LinClpP isoforms were generated. Then, the effect on their biochemical activity in the presence of a chaperone, ADEP1, and trigger factor was explored.

2.3. Materials and Methods

2.3.1. Site-directed mutagenesis of LinClpP based on multiple sequence alignment with its orthologs

Amino acid sequences of ClpP from various pathogenic bacteria were retrieved from the UniProtKB database (Consortium et al., 2014). Multiple sequence alignment (MSA) using Clustal Omega software was conducted (Sievers et al., 2011). For better clarity, MSA was refined using the online tool ESPript (Easy Sequencing in PostScript) (Robert et al., 2014). The secondary structures of ClpP orthologs were fetched from the protein data bank (PDB) (Berman et al., 2000).

Within the conserved motifs of LinClpP1, residues Glu170 of the oligomerization sensor domain and Asn172 of the catalytic triad were selected for mutagenesis. In contrast, for LinClpP2, the residues Ile126Gly127 of the glycine-rich loop, Ser40Lys41 of the hydrophobic cluster, and Tyr62 of the tyrosine trigger motif were selected for mutagenesis. A Q5 site-directed mutagenesis kit (NEB, catalog no. E0554S) was used for performing mutation in LinClpP isoforms (LinClpP1; *LIC11417* and LinClpP2; *LIC11951*). The primers were designed using an NEBaseChanger tool to generate the LinClpP mutant variants (**Table 2.1**). The generated plasmid variants were sequenced by outsourcing to a local company (Eurofins, India) and the specific mutations were confirmed.

TH-3792_206106022

Table 2.1. Primers used in this study

Primer Name	Primer sequence (5′-3′)
ClpP1(N172D) F	ACAGAAAGAgatTTTTACATGACAGCAG
ClpP1(N172D) R	ATCTTTTTGAATTTGTTCCAC
ClpP1(E170D) F	AAGATACAgatAGAAATTTTTACATGAC
ClpP1(E170D) R	TTTGAATTTGTTCCACAGTTTTAC
ClpP2(IG_del) F	GGACAGATTGTAGCACCTG
ClpP2(IG_del) R	ACTAGGTTGATGAATCATCAC
ClpP2(S40AK41N) F	AGACGAATCTgctaatGATTTAGTTGGTAAAC
ClpP2(S40AK41N) R	GTAACAGGACCCCAGAGA
ClpP2(Y62A) F	AATCACTTTTgctATCAATAGTCCCG
ClpP2(Y62A) R	TTTTTACCTGGATCTTTCATTT

Nucleotide sequences in small letters denote the site of mutation.

2.3.2. Generation of LinClpP model structure

The structures of LinClpP1 and LinClpP2 were retrieved from the AlphaFold (AF) protein structure database (Jumper et al., 2021). The model structures of mutant variants were generated via SWISS-MODEL using wild-type LinClpP model structures as a template (Waterhouse et al., 2018). The program PyMOL was used for structure visualization and representation (DeLano et al., 2002).

2.3.3. Overexpression and purification of recombinant proteins

The LinClpP isoforms or its mutant variants were overexpressed and purified from *E. coli* BL21 (DE3) cells, as described previously in our laboratory (Dhara et al., 2019). Briefly, the *E. coli* cultures were grown till OD_{600} = 0.6 and then chemically induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37 °C to overexpress all recombinant proteins. For LinClpP2^{Y62A}, induction was exclusively given with a lower IPTG concentration (0.5 mM) and was incubated for 16 h at 18 °C. The recombinant proteins were purified by affinity column chromatography using nickel-nitrilotriacetic acid (Ni-NTA) resins (Invitrogen). The concentration of recombinant proteins was estimated by Bradford assay, as described previously (Dhara et al., 2019).

2.3.4. LinClpP self-assembly condition

Equimolar concentrations of purified LinClpP1 and LinClpP2 were mixed (1:1) and pre-incubated for 24 hours at 4 °C. The resulting product constitutes the catalytically active heterocomplex LinClpP1P2. A similar step was performed for other mutant variants of

TH-3792_206106022

LinClpP. All the downstream biochemical assays were conducted only after the pre-incubation step.

2.3.5. Peptide degradation assay

The fluorogenic model dipeptide substrate Suc-LY-AMC (N-succinyl-Leu-Tyr-7-amino-4-methyl coumarin; Sigma) was solubilized in dimethyl sulfoxide (DMSO). The cleavage of the peptide substrate (100 µM) by LinClpP or its mutant variants (1 µM) was carried out in a total of 50 µL reaction buffer (50 mM phosphate buffer, pH 7.6, 100 mM KCl, and 5% (v/v) glycerol). Hydrolysis of the peptide was monitored in black, flat-bottom 96-well microplates (Invitrogen) at 37 °C for 240 min. The release of AMC during hydrolysis was estimated in a spectrofluorometer (iTECAN Infinite M PLEX) at a specific excitation (380 nm) and emission wavelength (460 nm). All the experiments were performed thrice independently and in duplicates.

2.3.6. Protein degradation assay

Degradation of the fluorogenic model protein, FITC-casein (fluorescein isothiocyanate-casein; Sigma), was carried out in a total of 20 µL reaction buffer (50 mM Tris-Cl pH 7.8, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$) as described previously (Dhara et al., 2021). Briefly, 4 µL of the substrate (1.5 µg μL^{-1}) was pre-incubated with LinClpP or its mutant variant (1 µM) and LinClpX (1 µM) for 10 min before the addition of ATP (2 mM). Similar pre-incubation of LinClpP or its mutant variants with ADEP1 (15 µM) and LinTF (0.4 µM) was done before adding the substrate. The protease reactions were performed in the dark for 120 min at 37 °C, followed by the termination of the reaction with trichloroacetic acid (0.6 N). The fluorescence of released FITC during proteolysis was monitored in black, flat-bottom 96-well microplates (Invitrogen) at 37 °C in a spectrofluorometer (iTECAN infinite M PLEX). The excitation and emission wavelengths used were 490 nm and 525 nm, respectively. All experiments were performed thrice independently, each in duplicate.

2.3.7. Protein analyses on Native-PAGE

Native PAGE analysis of LinClpPs or its mutant variants was performed as described previously (Dhara et al., 2019). Briefly, recombinant proteins (1 µM) were incubated for 24 h at 4 °C in a buffer (50 mM Tris-Cl, pH 8.0, 100 mM NaCl, and 10% (v/v) glycerol). The pre-incubated samples were mixed with 3× native sample buffer (240 mM Tris-HCl pH 6.8, 30% (v/v) glycerol, 0.03% (w/v) bromophenol blue). The samples were loaded on a 7.5% native gel,

TH-3792_206106022

followed by protein separation by electrophoresis at 120 V for 2 h. The resolved proteins were stained with Coomassie Blue dye for visualization and compared with standard protein markers (Invitrogen, catalog no. 928387) of known molecular weight.

2.3.8. Size exclusion chromatography (SEC)

A HiLoad™ 16/600 Superdex™ 200 pg column (GE Healthcare #28-9893-35) was used for SEC analysis on a next-generation chromatography system (Bio-Rad) as described earlier for wild-type LinClpP isoforms (Dhara et al., 2019). Briefly, the column was calibrated with standard proteins of known sizes, including β-amylase (200 kDa), alcohol dehydrogenase (158 kDa), albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome C (12.4 kDa) (Sigma, catalog no. MWGF-200). The column was then buffer exchanged using an equilibration buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, and 10% glycerol). The LinClpP or its mutant variants (0.2-0.5 mg mL^{-1}) were pre-incubated for 24 hours at 4 °C. The pre-incubated proteins (300 µg) were loaded onto the column, and gel-filtration experiments were carried out at room temperature with a flow rate of 1 mL min^{-1}.

2.3.9. Dynamic Light Scattering (DLS)

DLS experiments were performed on a LITESIZER 500 (Anton Paar) at 37°C. The LinClpP2 (5 µM) or its mutants (LinClpP2^{S40AK41N} or LinClpP2^{Y62A}) were incubated in an equilibration buffer to self-assemble for 24 hours at 4 °C. Then, the self-assembled LinClpP2 or its mutant variants were added to the polystyrene cuvettes to record the light scattering. Next, to inspect the impact of ADEP1 on the hydrodynamic diameter, the LinClpP2 or its mutant variants were incubated with ADEP1 (1 h at 37 °C), followed by DLS analysis. A total of 15 autocorrelation functions (technical replicates) were recorded for each protein sample, and the hydrodynamic diameters were determined as described previously (Dhara et al., 2019). The hydrodynamic diameter is represented as the average of three replicates performed independently.

2.4. Results

2.4.1. *In-silico* analyses of LinClpP sequences and identification of its conserved motifs

The primary sequences of LinClpP1 and LinClpP2 from *Leptospira* were aligned to well-characterized ClpP orthologs from various pathogenic microorganisms, including *Clostridium difficile* (CdiClpP), *Mycobacterium tuberculosis* (MtuClpP), *Listeria monocytogenes* (LmoClpP), *Chlamydia trachomatis* (CtrClpP), *Staphylococcus aureus* (SauClpP),

TH-3792_206106022

Escherichia coli (EcoClpP), *Bacillus subtilis* (BsuClpP), *Helicobacter pylor*i (HpyClpP), and *Neisseria meningitides* (NmeClpP). The LinClpP1 shared a high degree (>60%) of sequence identity with LmoClpP2 (62%), EcoClpP (61%), SauClpP (62%), and HpyClpP of *Helicobacter pylori* (64%). While the LinClpP2 shared sequence identity (>45%) with MtuClpP2 (46%), EcoClpP (48%), SauClpP (49%), and HpyClpP (45%) (Dhara et al., 2019). The LinClpP isoforms were assessed for various conserved domains recognized to date, including the catalytic triad, activation trigger, oligomerization sensor, and the Gly-rich heptamer dimerization domain. To comprehend the role of conserved hotspots in the LinClpP isoforms, specific residues in LinClpP1 (Glu170, oligomerization sensor domain; Asn172, catalytic triad) and LinClpP2 (Ile126Gly127, glycine-rich loop; Ser40Lys41, hydrophobic cluster; Tyr62, activation trigger) were chosen for site-directed mutagenesis **(Figure 2.1)**.

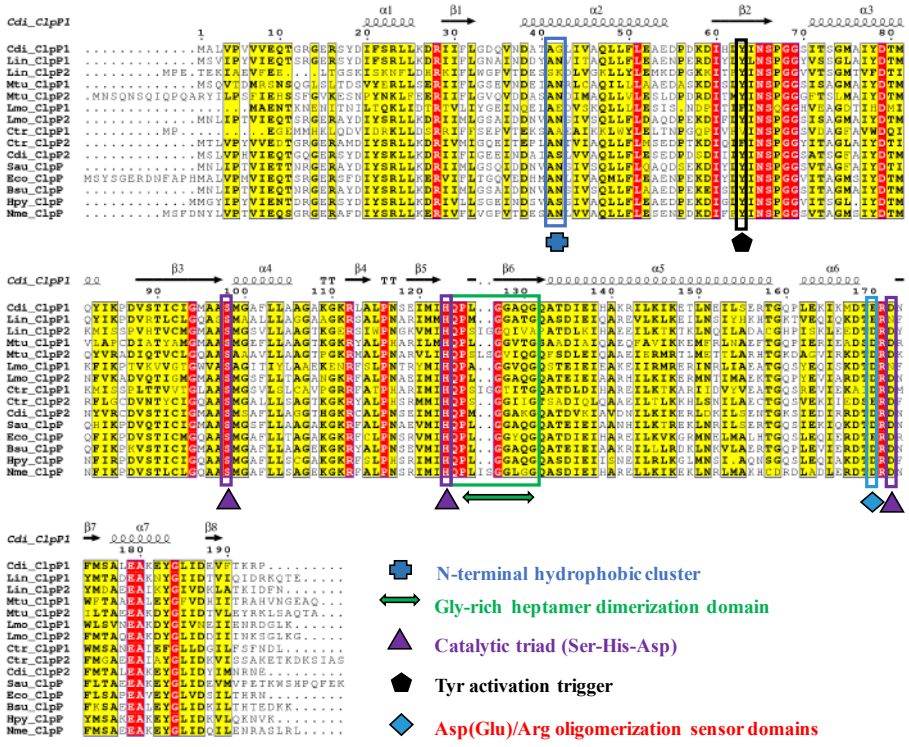


Figure 2.1. **Multiple sequence alignment of ClpP proteases**. Multiple sequence alignment of ClpP orthologs from various pathogenic bacteria was performed using Clustal Omega software. The ClpP orthologs used for amino acid alignment are CdiClpP1; (Q180F0), LinClpP1; (Q72SG6), LinClpP2; (Q72R01), MtuClpP1; (P9WPC5), MtuClpP2; (P9WPC3), LmoClpP1; (Q8Y7Y1), LmoClpP2; (Q9RQI6), CtrClpP1; (P38002), CtrClpP2; (084712), CdiClpP2; (Q180J6), SauClpP; (P63786), EcoClpP; (P0A6G7), BsuClpP; (P80244), HpyClpP; (P56156), NmeClpP; (Q9JZ38). Secondary structural elements in the CdiClpP1 (PDB ID: 6MX2) structure are shown on top of the sequence alignment. Identical and semi-identical residues in ClpP proteases are highlighted in red and yellow, respectively, with residue numbers at the top of the alignment after CdiClpP1. The conserved key structural motifs in the LinClpP isoforms, including the Tyr activation trigger, catalytic triad (Ser-His-Asp), Asp(Glu)/Arg oligomerization sensor domains, hydrophobic cluster, and the Gly-rich heptamer dimerization domain, are outlined in different colored boxes and marked with specific symbols at the bottom of the box.

TH-3792_206106022

The ClpP orthologs have a conserved catalytic triad (Ser-His-Asp), an essential motif for the charge-relay in the serine peptidase family. The deputed catalytic triad of LinClpP1 (Ser98-His123-Asn172) and LinClpP2 (Ser97-His122-Asp173), after alignment with its orthologs, appeared to vary at one residue (Asn172 versus Asp173). However, among multiple ClpP1 orthologs, the Asn172 residue exhibited conservation within LinClpP1 and LmoClpP1. Interestingly, in LmoClpP1, the substitution of Asn172 with the more conserved Asp172 residue (LmoClpP1^{N172D}) resulted in a gain-of-function (GOF) (Zeiler et al., 2013). Enlightened by the GOF recorded in LmoClpP^{N172D}, a mutant LinClpP1^{N172D} was generated to estimate the GOF, if any.

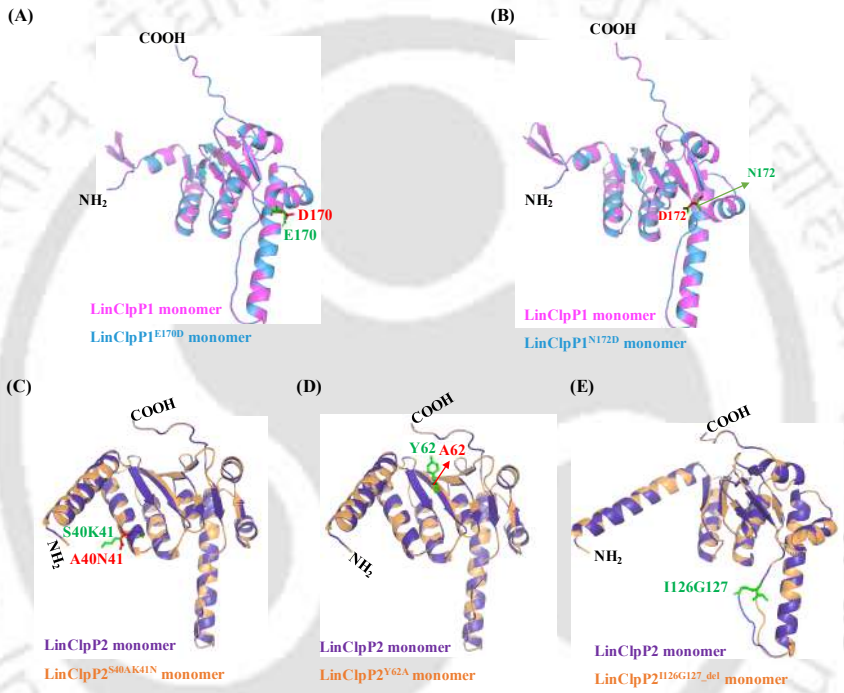


Figure 2.2 Modeled structure of LinClpP and its mutant variants. The modeled structure of the LinClpP1 monomer (pink) superimposed with **(A)** LinClpP1^{E170D} (blue) **and (B)** LinClpP1^{N172D} (blue). The residues (E170 and N172) of wild-type LinClpP1 and mutants (D170 and D172) are represented in green and red colored sticks, respectively. The modeled LinClpP2 structure (violet) superimposed with **(C)** LinClpP2^{S40AK41N}, (D) LinClpP2^{Y62A}, **(E)** LinClpP2^{I126G127_del}. The LinClpP2 mutant structures are represented in orange color. The residues (S40K41, Y62, and I126G127) of wild-type LinClpP2 and mutants (A40N41 and A62) are represented in green and red colored sticks, respectively.

Similarly, a semi-conserved residue Glu170 in place of conserved Asp within the oligomerization sensor domain of LinClpP1 and CdiClpP1 was also observed. Natural substitution of Asp residue within the oligomerization domain was not only limited to the ClpP1 isoform but also observed in the case of CtrClpP2 (Glu170) and EcoClpP (Glu183).

TH-3792_206106022

Thus, another mutant LinClpP1^{E170D} was generated to determine the effect of the conserved Asp in the oligomerization of LinClpP1.

In a recent study, the hydrophobic side chain of Tyr63 residue was illustrated to regulate the entrance pore of SauClpP tetradecamer, wherein the interaction of antibiotic, ADEP1, with the hydrophobic cluster of SauClpP induces an open-gate activation (Ni et al., 2016). Substitution of Tyr63 with Ala63 residue that contains a small aliphatic side chain recorded a GOF in SauClpP (Ni et al., 2016). Similarly, in an earlier study on LinClpP2, the residue Tyr62 was speculated to interact with the ADEP1 hydrophobic cluster (Dhara et al., 2021). Thus, in this study, the substitution of Tyr62 with Ala62 residue was executed (LinClpP2^{Y62A}) to understand its implication on biological activity. Among other LinClpP2 hotspots, the conserved Gly residues within the glycine-rich motif were also investigated. The multiple sequence alignment (MSA) illustrates a conserved glycine-rich loop (Gly127, Gly128, and Gly131) in LinClpP1; however, LinClpP2 has only Gly128, and the other equivalent conserved Gly residue was replaced with Gln129 and Ala132, respectively (**Figure 2.1**). In addition, within its Gly-rich loop, the LinClpP2 holds two extra residues (Ile126 and Gly127) insertions similar to those of CtrClpP1. Thus, to comprehend the role of these extra residues (Ile126 and Gly127) in LinClpP2, a deletion mutant (LinClpP2^{IG_del}) was created. Among other disparities, Ser40 and Lys41 have replaced conserved Ala and Asn in LinClpP2 at the α2 helix of the N-terminal. In contrast, LinClpP1 retained the conserved aliphatic residues (Ala and Asn) within the N-terminal region as observed in other ClpP orthologs (**Figure 2.1**). Ser40 and Lys41 residues reside in the hydrophobic cluster near the N-terminal segment of LinClpP2. The significance of these polar and charged residues within the hydrophobic cluster of LinClpP2 is unclear. Thus, the double mutant, LinClpP2^{S40AK41N}, was generated. For clarity, the modeled structure of LinClpP1 (**Figure 2.2A – 2.2B)** and LinClpP2 (**Figure 2.2C – 2.2E**) subunits was aligned to various mutant forms to demarcate the critical hotspot residues selected in this study.

2.4.2. LinClpP mutant variants exhibit a functional disparity

The site-directed mutagenesis was conducted individually in pET23a-clpP1 and pET23a-clpP2 expression vectors. The sequencing result suggested that all the plasmids were substituted at the desired site through site-directed mutagenesis. The purification of all recombinant mutant variants of LinClpP1 and LinClpP2 was conducted using the Ni-NTA affinity column chromatography (**Figure 2.3A**). The final yield for the four mutants (LinClpP1^{E170D}, LinClpP1^{N172D}, LinClpP2^{S40AK41N}, and LinClpP2^{IG_del}) was approximately 1 mg L^{-1}. At the

TH-3792_206106022

same time, as most of the overexpressed LinClpP2^{Y62A} was present in the insoluble fraction, the recombinant LinClpP2^{Y62A} was purified with a maximum yield of 0.3 mg L^{-1}.

The fluorogenic model peptide (Suc-LY-AMC) degradation is often used to determine the functionality of several ClpP peptidase complexes under *in-vitro* conditions (Gaillot et al., 2000; Woo et al., 1989; Schmitz et al., 2014). Therefore, the peptidase activity of pure LinClpP1 mutants (LinClpP1^{E170D} and LinClpP1^{N172D}) and LinClpP2 mutants (LinClpP2^{IG_del}, LinClpP2^{S40AK41N}, and LinClpP2^{Y62A}), along with their heterocomplex, was determined using the fluorogenic dipeptide model substrate (S1; Suc-LY-AMC).

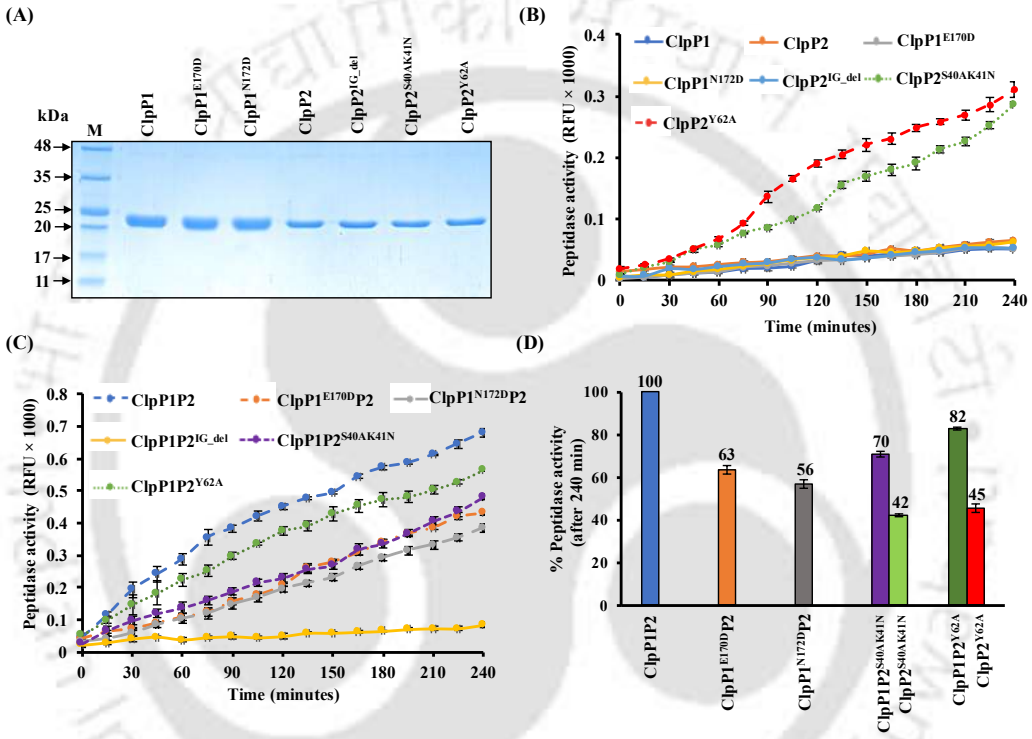


Figure 2.3. Biochemical characterization of recombinant mutant variants of LinClpP. (A) Purified recombinant LinClpP1 (23 kDa), LinClpP2 (22 kDa), and their mutants (LinClpP1^{E170D}, LinClpP1^{N172D}, LinClpP2^{IG_del}, LinClpP2^{S40AK41N}, LinClpP2Y62A) resolved on a 12% SDS-polyacrylamide gel and stained with Coomassie blue. **(B)** Peptidase activity of pure LinClpP isoforms (ClpP1 and ClpP2) or its mutants on the peptide substrate (S1; Suc-LY-AMC) with time. Peptide degradation of S1 at 30 min intervals for 240 min was measured as a relative fluorescent unit (RFU × 1000) at a specific excitation (380 nm) and emission (460 nm) wavelength. **(C)** Comparison of peptidase activity of heterocomplex variants (LinClpP1^{E170D}P2, LinClpP1^{N172D}P2, LinClpP1P2^{S40AK41N}, LinClpP1P2^{Y62A}, and LinClpP1P2^{IG_del}) with LinClpP1P2 on substrate S1 at 30 min intervals for 240 min. **(D)** Peptidase activity (at 240 minutes) in percentage for pure LinClpP mutants or its heterocomplex variants and compared with LinClpP1P2 (100%). The data represent the Mean ± Standard Error Mean (SEM) from the three independent experiments performed in duplicates.

Till 30 min of the peptidase reaction, none of the pure LinClpP mutant variants indicated any activity on substrate S1. Therefore, the peptidase reaction on S1 was extended and measured at every 30-minute interval up to 240 minutes. With increasing time, two pure LinClpP2

TH-3792_206106022

mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}) demonstrated a gain in peptidase activity on model substrate S1 in comparison to the inactive pure LinClpP isoforms (**Figure 2.3B**). However, no such gain in peptidase activity was observed for the pure LinClpP1^{N172D} mutant as reported earlier for LmoClpP1^{N172D} upon redesigning the catalytic triad (Zeiler et al., 2013). Next, we were curious to address whether the heterocomplex formed from such a GOF mutant can exhibit higher peptidase activity than the wild-type LinClpP1P2.

Thus, various heterocomplexes were constituted where one isoform was of wild-type LinClpP and the other was a mutant variant at an equivalent concentration. The peptidase activity on substrate S1 of reconstituted heterocomplexes was estimated (**Figure 2.3C**). The measured peptidase activity of wild-type LinClpP1P2 heterocomplex at 240 min was considered 100 %, and the percent peptidase activity of pure LinClpP isoforms and their mutant variants or their heterocomplex after 240 min was compared (**Table 2.2** and **Figure 2.3D**). Interestingly, the relative peptidase activity of two variants of heterocomplex LinClpP1^{E170D}P2 (63%) and LinClpP1^{N172D}P2 (56%) containing mutations in LinClpP1 isoform was measured lower than that of the wild-type (**Figure 2.3D**). Thus, the mixing of wild-type LinClpP2 might have partially restored the functionality of the mutant variants of LinClpP1 as a heterocomplex. Likewise, the relative peptidase activity of heterocomplex variants LinClpP1P2^{S40AK41N} (70%) and LinClpP1P2^{Y62A} (82%) with the mutation in LinClpP2 isoform measured lower than the wild-type heterocomplex (**Figure 2.3D**). Interestingly, the pure LinClpP2^{S40AK41N} and LinClpP2^{Y62A} exhibited only 42% and 45% of LinClpP1P2 activity. In contrast, the heterocomplex variant, LinClpP1P2^{IG_del} was enzymatically inactive and establishes the importance of Ile126 and Gly127 residues within the Gly-rich loop. A similar abolishment in the enzymatic activity of SauClpP was also reported upon substituting Glycine residues of the Gly-rich loop with alanine (Gersch et al., 2012).

2.4.3. LinClpP mutant variants' protease activity in the presence of LinClpX

A previous study on the *Leptospira* ClpP system illustrated that LinClpX forms an active complex with LinClpP isoforms and stimulates its protease activity on fluorogenic model protein substrate, FITC-casein, exclusively in the presence of ATP (Dhara et al., 2019). In the same context, in this study, the protease activity of LinClpP mutant variants was compared with wild-type LinClpP isoforms in the presence of a chaperone, LinClpX. The protease activity of wild-type LinClpP isoforms and their mutant variants on the FITC-casein model substrate was evaluated in the presence and absence of LinClpX. The degradation of FITC-casein after 120 min was represented as a relative fluorescent unit (RFU) (**Figure 2.4A**).

TH-3792_206106022

Further, to determine the percent (%) protease activity after 120 min (**Figure 2.4B**), the background RFU of inactive LinClpP isoforms (LinClpP1 and LinClpP2) was subtracted from all RFU values and then % protease activity was calculated with respect to LinClpXP1P2 (100%) (**Table 2.2**). In the case of pure LinClpP mutants (LinClpP1^{E170D}, LinClpP1^{N172D}, and LinClpP2^{S40AK41N}), the addition of the LinClpX could not stimulate any protease activity.

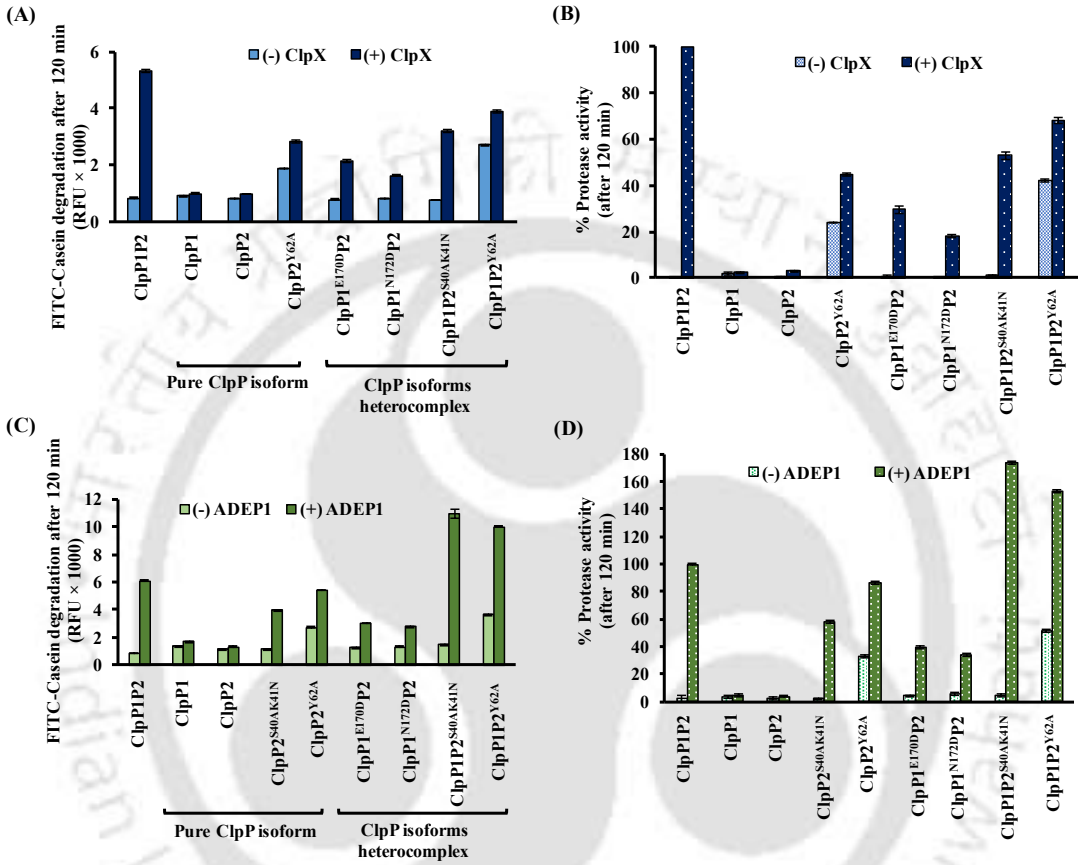


Figure 2.4. The effect of LinClpX or ADEP1 on the protease activity of LinClpP and its mutant. (A) The effect of adding LinClpX (+) to the protease activity of pure LinClpP or its mutant (LinClpP2^{Y62A}) or the heterocomplex variants (LinClpP1P2, LinClpP1^{E170D}P2, LinClpP1^{N172D}P2, LinClpP1P2^{S40AK41N}, LinClpP1P2^{Y62A}) on substrate S2. After 120 min of protease activity, the degradation of S2 was measured fluorometrically (RFU × 1000) at specific excitation (490 nm) and emission (525 nm) wavelengths and compared with an equivalent reaction in the absence (-) of LinClpX. **(B)** The effect of the presence (+) or absence (-) of LinClpX on the protease activity of pure LinClpP or its mutant heterocomplex variants at 120 min is shown in percentage. The background RFU for inactive LinClpP isoforms was subtracted from all RFU values before calculating % protease activity. **(C)** Effect of presence (+) and absence (-) of ADEP1 on the protease activity of pure LinClpP or its mutants heterocomplex on substrate S2 at 120 min. **(D)** Effect of the presence and absence of ADEP1 on the protease activity of pure LinClpP or its mutants heterocomplex after 120 min in percentage. The background RFU for inactive LinClpP isoforms was subtracted from all RFU values before determining % protease activity. For comparison, the protease activity of ADEP1-bound LinClpP1P2 was considered 100%. The data here represents Mean ± Standard Error Mean (SEM) from three independent experiments performed in duplicates.

TH-3792_206106022

However, the protease activity (%) of LinClpP heterocomplex variants (LinClpP1^{E170D}P2; 30%, LinClpP1^{N172D}P2; 18%, and LinClpP1P2^{S40AK41N}; 53%) was partly restored in the presence of LinClpX to that of the wild-type LinClpXP1P2 (100%) (**Figure 2.4B**). Consistently, in the deletion mutant LinClpP2^{IG_del}, it was also inactive in the presence of LinClpP1 isoform and LinClpX (data not shown). Interestingly, the protease activity of one of the pure LinClpP2^{Y62A} mutants had increased compared to inactive pure wild-type LinClpP isoforms and measured 23% activity compared to that of LinClpXP1P2 (100%). A similar GOF in protease activity without ATPase chaperone was reported elsewhere for homo-tetradecameric SauClpP^{Y63A}, EcoClpP^{Y67A,} and BsuClpP^{Y63A} mutants (Ni et al., 2016). Adding ATPase chaperone LinClpX enhanced LinClpXP2^{Y62A} protease activity to 45%. Also, when wild-type LinClpP1 is added to constitute the heterocomplex, LinClpP1P2^{Y62A}, the protease activity increases from 23% to 42% of that of LinClpXP1P2 (100%). Although in the pure LinClpP2^{Y62A}, a GOF was observed, the ternary complex LinClpXP1P2^{Y62A} demonstrated only 68% protease activity compared to that of LinClpXP1P2 (**Figure 2.4B**).

2.4.4. Protease activity of LinClpP mutants under the influence of ADEP1

In *Leptospira*, the antibiotic acyldepsipeptide (ADEP1) has been responsible for the functional gain of protease activity in the LinClpP1P2 isoforms even in the absence of the chaperone, LinClpX (Dhara et al., 2021). Further, to get an insight into the effect of ADEP1 on the protease activity of LinClpP mutant variants generated in this study, a protease reaction was conducted (**Figure 2.4C** and **2.4D**). To determine the percentage (%) protease activity after 120 minutes (**Figure 2.4D**), the background RFU for inactive LinClpP isoforms (LinClpP1 and LinClpP2) was subtracted from all RFU values, and then, % protease activity was calculated considering ADEP1-bound LinClpP1P2 as 100% (**Table 2.2**). The pure mutant LinClpP1^{E170D}, having a mutation at the oligomerization sensor domain, and the LinClpP1^{N172D} with a mutation at the active site, did not demonstrate GOF in the presence of ADEP1. However, the mutant heterocomplex variants, LinClpP1^{E170D}P2 and LinClpP1^{N172D}P2, showed protease activity upon ADEP1 supplementation. The measured protease activity (%) of mutant heterocomplex variants (LinClpP1^{E170D}P2; 40% and LinClpP1^{N172D}P2; 34%) in the presence of ADEP1 was relatively lower than the ADEP1-bound LinClpP1P2 (100%). In contrast, adding ADEP1 to the pure double mutant LinClpP2^{S40AK41N}, having a mutation at the N-terminal hydrophobic cluster, measured GOF in protease activity versus the wild-type LinClpP2 (no activity). Nevertheless, the relative protease activity of ADEP1-bound double mutant LinClpP2^{S40AK41N} reached 58% of that of ADEP1-bound LinClpP1P2 heterocomplex.

TH-3792_206106022

Table 2.2. Comparison of the peptidase/protease activity of LinClpP isoforms or their mutant variants under the influence of chaperone LinClpX and antibiotic ADEP1

LinClpP isoforms or their mutant variants	Peptidase activity in percentage	Protease activity in percentage		
		Only LinClpP isoform(s)	+ Chaperone LinClpX	+ Antibiotic ADEP1
LinClpP1/LinClpP2	NA	NA	NA	NA
LinClpP1P2	100	NA	100	100
LinClpP1^{E170D}	NA	NA	NA	NA
LinClpP1^{E170D}P2	63 ± 1.1	NA	30 ± 1.3	40 ± 0.8
LinClpP1^{N172D}	NA	NA	NA	NA
LinClpP1^{N172D}P2	56 ± 2	NA	18 ± 0.6	34 ± 0.8
LinClpP2^{IG_del}	NA	NA	NA	NA
LinClpP1P2^{IG_del}	NA	NA	NA	NA
LinClpP2^{S40AK41N}	42 ± 0.5	NA	NA	58 ± 0.9
LinClpP1P2^{S40AK41N}	70 ± 1.3	NA	53 ± 1.4	170 ± 1.3
LinClpP2^{Y62A}	45 ± 1.9	23 ± 0.2	45 ± 0.6	85 ± 1.1
LinClpP1P2^{Y62A}	82 ± 0.6	42 ± 0.7	68 ± 1.1	150 ± 1.2

NA indicates NOT ACTIVE LinClpP isoforms or their mutant variants. The LinClpP protease activity in the presence (+) of either chaperone LinClpX or antibiotic ADEP1 is indicated.

It is presumed that the binding of ADEP1 stimulates structural changes within the hydrophobic region of LinClpPs. Thus, the ADEP1-bound double mutant LinClpP2^{S40AK41N} gained protease activity more than its wild-type counterpart. Moreover, unlike other mutant heterocomplex variants, LinClpP1P2^{S40AK41N}, under the influence of ADEP1, measured 1.7-fold higher protease activity than its counter wild-type (**Figure 2.4D**). Similarly, in the LinClpP2^{Y62A} mutant, which was alone active in degrading protein substrate, the relative protease activity on ADEP1 supplementation was further enhanced and achieved 85% of that of LinClpP1P2. Moreover, as expected, in the mutant heterocomplex variant LinClpP1P2^{Y62A}, the protease activity showed GOF by 1.5 times that of LinClpP1P2 on ADEP1 supplementation (**Figure 2.4D**). On the contrary, ADEP1 showed no influence on the protease activity of the pure deletion mutant LinClpP2^{IG_del} or its heterocomplex LinClpP1P2^{IG_del}, suggesting the importance of the glycine-rich loop.

TH-3792_206106022

2.4.5. LinTF promotes the peptidase/protease activity of the ClpP variants on model substrates

In *Leptospira*, the trigger factor (LinTF), an ATP-independent chaperone, has been previously demonstrated to have a functional association with the caseinolytic protease system (Choudhury et al., 2021). The aforementioned study provided insight into the structural stabilization of heterocomplex LinClpP1P2 by LinTF and was measured in terms of gain in peptidase activity. However, LinTF showed no improvement in peptidase activity in pure LinClpP1 and LinClpP2 (Choudhury et al., 2021). In this study, we demonstrated that the generation of the mutant variants of LinClpP1 did not gain any peptidase or protease activity, while selective LinClpP2 mutant variants showed a GOF. Thus, we tested if LinTF could further promote the functionality in GOF mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}). Supplementation of LinTF increased the peptidase activity of the double mutant LinClpP2^{S40AK41N} from 12% to 24% compared to LinClpP1P2 activity (100 %) under similar conditions (**Figure 2.5A**). Similarly, with the addition of LinTF, the LinClpP2^{Y62A} shows a boost in peptidase activity from 15% to 30%. For the mutants heterocomplex (LinClpP1P2^{S40AK41N} and LinClpP1P2^{Y62A}), the stimulation in peptidase activity in the presence of LinTF was more pronounced than that of pure LinClpP2 mutant variants. The addition of LinTF boosted (22% to 64%) the peptidase activity of the mutant heterocomplex LinClpP1P2^{S40AK41N}. Similarly, the peptidase activity of mutant heterocomplex LinClpP1P2^{Y62A} was also boosted (25% to 82%) of LinClpP1P2 in the presence of LinTF (100%).

Among the various LinClpP mutant variants in this study, only LinClpP2^{Y62A} showed GOF for protease activity. Thus, we also investigated the influence of LinTF on the protease activity of the pure LinClpP2^{Y62A} mutant on the model protein substrate (FITC-Casein; S2) (**Figure 2.5B**). The supplementation of LinTF to the pure LinClpP2^{Y62A} mutant or its heterocomplex (LinClpP1P2^{Y62A}) stimulated the protease activity to 2.3-fold and 2.7-fold, respectively (**Figure 2.5B**). However, the pure LinClpP2 or heterocomplex LinClpP1P2 showed no change in its protease activity in the presence of LinTF and thus remained inactive. The LinClpP1P2-TF complex is protease inactive, and therefore, it was not considered 100% for comparison.

Similarly, the synergistic influence of LinTF and ADEP1 on the protease activity of LinClpP2 mutant variants was studied and compared to LinClpP1P2 activity (100%) under similar conditions (**Figure 2.5C**). The protease activity of pure ADEP1-bound LinClpP2^{S40AK41N} mutant in the presence of LinTF was stimulated from 15% to 43%, while under similar conditions, its heterocomplex LinClpP1P2^{S40AK41N} was stimulated from 115% to 150%.

TH-3792_206106022

Similarly, the addition of LinTF to ADEP1-bound LinClpP2^{Y62A} showed a change in protease activity from 44% to 80%, while its heterocomplex ADEP1-LinClpP1P2^{Y62A} protease activity increased from 87% to 130%. Collectively, LinTF can also associate with GOF LinClpP2 mutant variants as reported for the LinClpP1P2 heterocomplex and thus assists in its functional assembly.

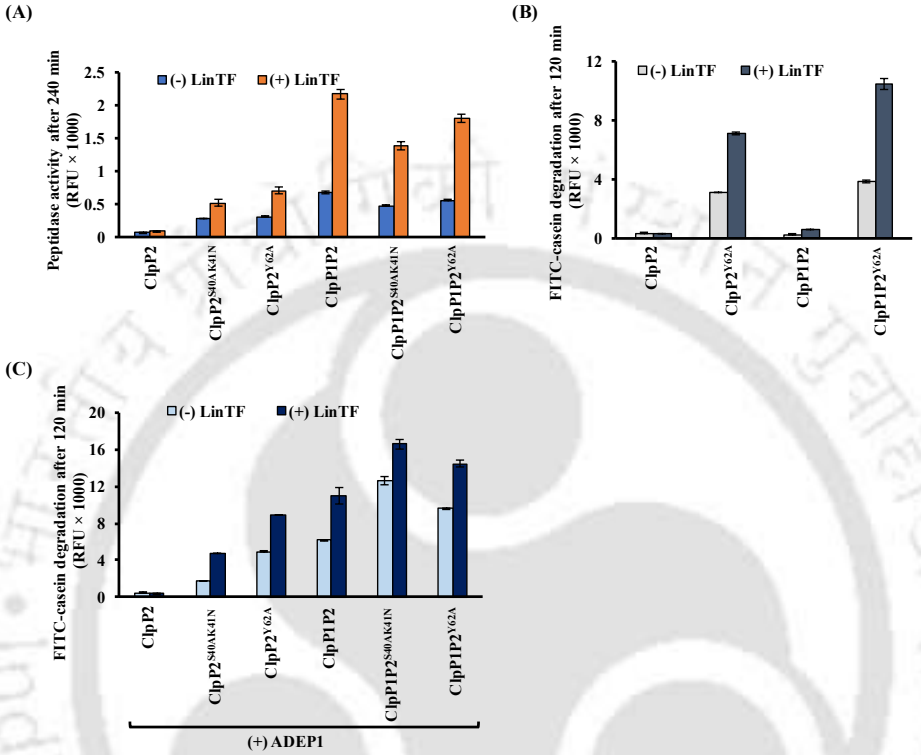


Figure 2.5. Effect of LinTF on the peptidase and protease activity of LinClpP mutant variants. (A) Effect of presence (+) or absence (-) of LinTF on the peptidase activity of pure LinClpP2, GOF mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}), and the heterocomplex variants (LinClpP1P2, LinClpP1P2^{S40AK41N}, LinClpP1P2^{Y62A}) on substrate S1 at 240 min. **(B)** Protease activity of the pure GOF mutant (LinClpP2^{Y62A}) or its heterocomplex on S2 after 120 min in the presence and absence of LinTF. **(C)** Impact of the presence (+) and absence (-) of LinTF on the protease activity of ADEP1-bound GOF mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}) on S2 after 120 min. The data represent the Mean± Standard error mean (SEM) from the three independent experiments performed in duplicates.

2.4.6. Oligomerization of LinClpP and its mutants

The structural organization of LinClpP mutant variants was evaluated using size exclusion chromatography (SEC) and native polyacrylamide gel electrophoresis (native-PAGE), as reported before (Dhara et al., 2019; Akopian et al., 2012). The oligomerization abilities of the LinClpP mutant variant were analyzed by SEC and compared with the wild-type LinClpP isoforms. On SEC, the elution profile of the mutant LinClpP1^{N172D} indicated its existence as a heptameric species similar to wild-type LinClpP1 (**Figure 2.6A**). The one having a mutation at the oligomerization domain (LinClpP1^{E170D}) showed an elution profile at monomeric and dimeric forms, and thus varied with the wild-type LinClpP1. The existence of LinClpP1^{E170D}

TH-3792_206106022

in a lower oligomeric form signified the importance of the evolutionary replacement of Asp170 with the Glu170 residue within the oligomerization domain.

In the pure wild-type LinClpP2 isoform, the predominant SEC peak corresponds to the tetradecameric species (308 kDa). In comparison, its selective mutant variants demonstrated a significantly different elution profile ranging from 154-308 kDa (**Figure 2.6B**). The double mutant LinClpP2^{S40AK41N} predominantly eluted at a heptameric size (154 kDa), while the deletion mutant LinClpP2^{IG_del} was eluted as a tetradecameric species (308 kDa) (**Figure 2.6B**). The inactive deletion mutant LinClpP2^{IG_del} showing tetradecameric assembly further corroborates that only tetradecameric assembly is insufficient for activity. In fact, proper alignment of the ClpP catalytic triad is also required. A similar loss in activity, even after tetradecameric assembly, was observed in the case of the Gly-rich loop mutant of SauClpP (Gersch et al., 2012). For the mutant LinClpP2^{Y62A}, despite multiple attempts, the protein failed to elute out through SEC.

The oligomerization result of the LinClpP mutant variant was also bolstered by an independent analysis using native polyacrylamide gel electrophoresis (native-PAGE). The mutants LinClpP1^{E170D} and LinClpP1^{N172D} resolved at a higher oligomeric species (14-21 mer) similar to the wild-type LinClpP1 in the native-PAGE (**Figure 2.6C**). As previously reported for LinClpP1 (Dhara et al., 2019), the migration of proteins on native-PAGE depends not only on its size but also on the shape and hydrodynamic properties of proteins. Thus, a similar discrepancy in size was observed from native-PAGE and SEC analysis of LinClpP1^{E170D} and LinClpP1^{N172D} mutants. However, for the LinClpP2 mutants, the oligomerization results analyzed through SEC agreed with those of native-PAGE. The native-PAGE analysis suggests that the predominant oligomeric species of LinClpP2^{S40AK41N} and LinClpP2^{IG_del} were heptamers and tetradecamer, respectively (**Figure 2.6D**).

2.4.7. The LinClpP mutant variants with a gain-of-function display disparity in conformations

In one elegant review, the structure of the ClpP monomer was reported to exist in three different conformational states, namely, extended, compact, and compressed forms (Malik et al., 2017). At the same time, the binding of ADEP to the axial surface of ClpP allows the enlargement of the entrance pore and converts the inactive ClpP complex into an open-gate active form (Lee et al., 2010; Li et al., 2010; Gersch et al., 2015).

TH-3792_206106022

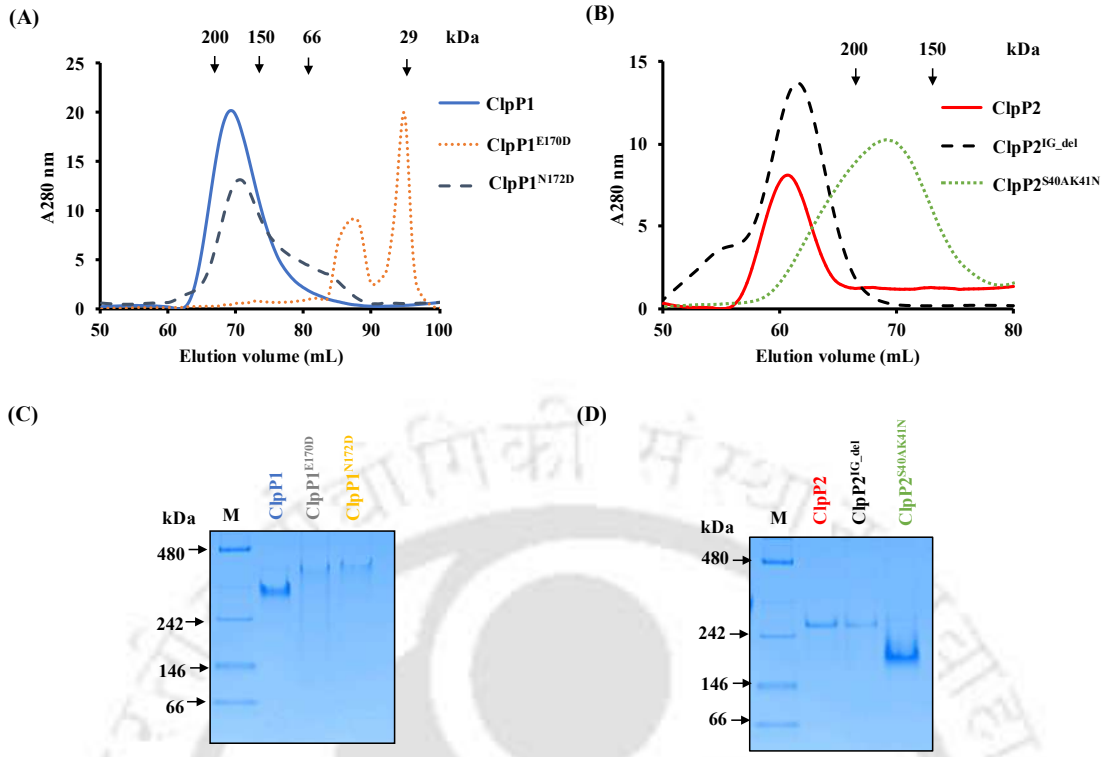


Figure 2.6. Oligomerization property of LinClpP isoforms and their mutants. The oligomeric states of LinClpP1 and LinClpP2 mutants were monitored by size exclusion chromatography (SEC). The elution profile for LinClpP isoforms and their mutants was represented as Absorbance 280 nm (A280 nm) versus elution volume (mL). The elution positions of standard molecular weight proteins in kilodaltons (kDa) are marked by arrows pointing downward and include β-amylase (200), alcohol dehydrogenase (158), albumin (66), and carbonic anhydrase (29). **(A)** Elution profile of LinClpP1 and its mutants (LinClpP1^{E170D} and LinClpP1^{N72D}). **(B)** Elution profile of LinClpP2 and its mutants (LinClpP2^{IG_del} and LinClpP2^{S40AK41N}). **(C)** The LinClpP1 or its mutants (LinClpP1^{E170D} and LinClpP1^{N72D}), **(D)** LinClpP2 or its mutants (LinClpP2^{IG_del} and LinClpP2^{S40AK41N}), resolved on the native-polyacrylamide gel (native-PAGE). The pure LinClpP or its mutant variants (1 µM) were resolved on a native-PAGE (7.5%) and stained with Coomassie brilliant blue. The standard molecular weight marker is resolved in lane 1 (M).

Similarly, a conformational switch of the LinClpP1P2 complex from an inactive compressed state to an active extended state was proposed upon ADEP1 binding (Dhara et al., 2021). Since the mutants, LinClpP2^{Y62A} and LinClpP2^{S40AK41N}, with GOF, overpoweringly degrade model protein substrate, we wondered if these mutants also undergo similar conformational changes. The change in conformation in ClpP has been previously explored by measuring the hydrodynamic diameter (D_h) of protein molecules in a solution using the dynamic light scattering technique for LinClpP1P2, TthClpP of *Thermus theromophilus*, and SauClpP (Dhara et al., 2019; Gersch et al., 2015; Felix et al., 2019). In this study, the diameter (D_h) of pure LinClpP2 (8.3 nm), LinClpP2^{S40AK41N} (8.2 nm), and LinClpP2^{Y62A} (11.44 nm) was measured and analyzed (**Figure 2.7A**).

TH-3792_206106022

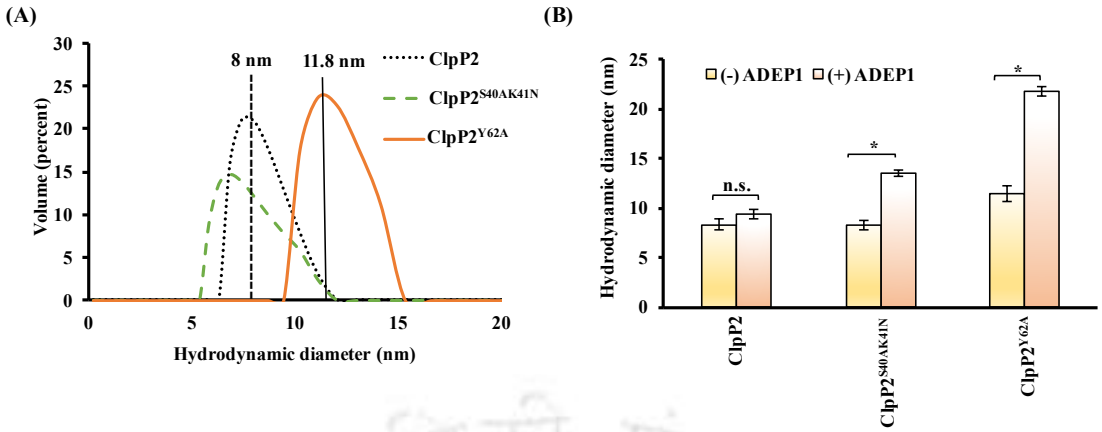


Figure 2.7. The oligomeric size distribution of LinClpP2 and its mutants by dynamic light scattering (DLS). (A) The oligomeric size distribution profile of LinClpP2 or its mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}) was examined using DLS. The volume (percent) versus particle size in nanometers (hydrodynamic diameter, D_h) is plotted. **(B)** Effect of presence (+) and absence (-) of ADEP1 on the hydrodynamic diameter (D_h) of the LinClpP2 and mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}). The data represent the Mean ± SEM from three independent experiments. Statistical analysis was performed by the Student's *t*-test for comparing the measured D_h obtained for LinClpP2 and its mutants in the presence and absence of ADEP1 (* denotes p-value < 0.05 and n.s. denotes p-value > 0.05).

The mutant LinClpP2^{Y62A} exhibited a 45 % increase in measured diameter (D_h) relative to LinClpP2 (**Figure 2.7A**). Increased diameter (D_h) of LinClpP2^{Y62A} suggests the existence of an open gate conformational state allowing unrestricted protein substrate entry into its catalytic chamber. The lesser diameter (D_h) of LinClpP2 and one of its double mutant LinClpP2^{S40AK41N} suggests their restricted protein substrate entry and justifies their incompetence.

After that, the diameter (D_h) measurement was also done for the ADEP1-bound LinClpP2^{S40AK41N} and LinClpP2^{Y62A} (**Figure 2.7B**). The LinClpP2 and its mutant variants (5 µM) were incubated with ADEP1 (15 µM) for 1 h at 37 °C before measuring the diameter (D_h). The measured diameters (D_h) of double mutant LinClpP2^{S40AK41N} in the presence of ADEP1 increased from 8.2 nm to 13.6 nm. Similarly, in LinClpP2^{Y62A}, on ADEP1 supplementation, the measured diameter (D_h) increased from 11.44 nm to 21.8 nm. Interestingly, adding ADEP1 to LinClpP2 did not significantly change its diameter (D_h). The influence of ADEP1 on the measured diameter (D_h) of LinClpP2 mutants (LinClpP2^{S40AK41N} and LinClpP2^{Y62A}) suggests conformational transitions from a compressed to an extended state and thus justifies their GOF in terms of their activity.

TH-3792_206106022

2.5. Discussion

The structure determination of different ClpP orthologs from various microorganisms, including *Escherichia coli* (Wang et al., 1997), *Bacillus subtilis* (Lee et al., 2011), and *Plasmodium falciparum* (Bakkouri et al., 2010) reveals the existence of three specific regions: (a) an N-terminal flexible loop that aligns at the axial pore, (b) a head domain with a catalytic center, and (c) a dynamic handle domain. The structural analysis of SauClpP suggests that the handle domain exists in different conformations, which allows the switching of ClpPs into three forms: extended, compact, and compressed (Ye et al., 2013). The tetradecameric SauClpP, undergoing a conformational switch from a compressed or compact state to an extended form, is closely associated with its functionality (Ye et al., 2013). The active SauClpP in an extended state requires two types of inter-ring communication for the proper alignment of the catalytic triad. Firstly, the handle domains of two SauClpP monomers from different heptameric rings form an anti-parallel β-sheet arrangement. The second communication involves hydrogen bond networking of the oligomerization sensor domain (Ye et al., 2013) In SauClpP, a hydrogen bridge network involving Gln132, Asp170, and Arg171 connects two heptameric rings to form an active tetradecamer assembly. In agreement, the SauClpP dissociated into inactive oligomeric complexes when Asp170 and Arg171 residues were substituted with Ala (Gersch et al., 2012). Multiple sequence alignments of LinClpP isoforms revealed the existence of conserved Asp and Arg within the oligomerization sensor domain of LinClpP2; however, in LinClpP1, the conserved Asp170 is naturally replaced with Glu170. Thus, to understand the evolutionary role of Glu170 residue in the LinClpP1 oligomerization sensor domain, a mutant LinClpP1^{E170D} was generated by substituting it with an evolutionarily conserved Asp residue found in other orthologs. The LinClpP1^{E170D} was inactive on peptide substrate, like its wild-type counterpart, LinClpP1. Similarly, the peptidase activity of the mutant heterocomplex LinClpP1^{E170D}P2 was compromised to 63% of that of LinClpP1P2. The peptidase assay suggested that substituting Glu170 within the oligomerization sensor domain with a more conserved Asp residue limits the peptidase activity of the heterocomplex LinClpP1^{E170D}P2. Also, the oligomeric size for LinClpP1^{E170D} obtained from SEC analysis differs from its heptameric LinClpP1. The monomeric and dimeric state of LinClpP1^{E170D} signifies the importance of the Glu170 residue and thus confirm the role of the oligomerization sensor domain in limiting the catalytic role of LinClpP1^{E170D}.

The catalytic triad arrangement in ClpP serine proteases reveals the existence of a network of hydrogen bonding. The active extended state of SauClpP shows the first interconnection between the hydroxyl group (-OH) of the active site, Ser98, and the amine group (-NH) of

TH-3792_206106022

His123 and the second interconnection between the carbonyl group (-CO) of Asp172 with the amine group (-NH) of His123 (Gersch et al., 2012). Hydrogen bonding within the catalytic triad of SauClpP is responsible for the adequate alignment of active site residues (Ser98, His123, and Asp172) (Gersch et al., 2012). The analyses of catalytic triad arrangement in ClpP orthologs from various pathogenic bacteria reveal a disparity at the 172nd residues of LinClpP1 and LmoClpP1 (*Listeria monocytogenes*). In LinClpP1 and LmoClpP1, unlike its ClpP2 isoforms, instead of the conserved Asp residue of the catalytic triad, Asn172 is present. In the case of *L. monocytogenes*, the model dipeptide substrate (Suc-LY-AMC) is hydrolyzed by LmoClpP2 alone but not with LmoClpP1 or its mutant LmoClpP1^{N172D} (Zeiler et al., 2013). Interestingly, each LmoClpPs isoform on another tripeptide model substrate (Leu-ACC) showed activity, while the mutant LmoClpP1^{N172D} activity enhanced 20-fold (Zeiler et al., 2013). Leu-ACC is a tripeptide model substrate (Ala-Arg-Leu) with 7-Amino-4-carbamoyl methyl coumarin (ACC) as a fluorophore. Since the LmoClpP1 catalytic triad has Asn172 in place of conserved Asp172, it is argued to have reduced strength of hydrogen bonding and, thus, low activity. Influenced by the LmoClpP1 findings, in this study, the Asn172 residue of LinClpP1 was also substituted with the Asp residue. The pure LinClpP1^{N172D} was enzymatically inactive even after extending the peptidase reaction on S1 (Suc-LY-AMC) for 240 min. Also, the mutant heterocomplex variant LinClpP1^{N172D}P2 displayed reduced (56%) peptidase activity than LinClpP1P2. Adding an ATPase chaperone, LinClpX or antibiotic ADEP1, could not stimulate the protease activity in the LinClpP1^{N172D}. Moreover, the protease activity of mutant heterocomplex, LinClpP1^{N172D}P2, was reduced (18% and 34%) in the presence of LinClpX and ADEP1, respectively. Retention of protease activity by mutant heterocomplex LinClpP1^{N172D}P2, albeit of lower efficiency, signifies that substituting the Asn172 residue with Asp still allows alignment of the catalytic triad in the tetradecameric complex.

Based on the conformation of the handle domain in SauClpP, an intermediate compact state was reported during the transitions between compressed and extended states (Ye et al., 2013). The extended form has a straight and ordered alignment of the handle domain, which is required for inter-ring interaction and forms an active tetradecameric assembly (Ye et al., 2013). The E-helix and glycine-rich motif from the handle domain forms an antiparallel β-sheet upon the interaction between two heptameric rings. The formation of an antiparallel β-sheet allows the straight alignment of the E-helix. Also, it induces His123 of the catalytic triad to undergo structural rearrangement for proper alignment of the active catalytic triad (Gersch et al., 2012). In SauClpP, three conserved glycines (Gly127, Gly128, and Gly131) form a part of the inter-

TH-3792_206106022

ring contact. Substitution of these three conserved glycines (Gly127, Gly128, and Gly131) with an Ala residue not only completely abolished the SauClpP protease activity but also disassociated 80% of the tetradecameric complex into heptamers (Gersch et al., 2012). Similarly, in LinClpP1, MSA analyses demonstrated a conserved glycine-rich loop (G127, G128, and G131). In contrast, the LinClpP2 showed a disparity in the glycine-rich loop sequence at Gln129 and Ala132. In addition, within the Gly-rich loop, the LinClpP2 has two extra residues (Ile126 and Gly127). A deletion mutant LinClpP2^{IG_del} in its pure state was enzymatically inactive despite retaining its tetradecameric form. Adding LinClpP1 to constitute mutant heterocomplex LinClpP1P2^{IG_del} did not revert its enzymatic activity. The loss of function in the mutant LinClpP2^{IG_del} signifies the importance of Ile126 and Gly127 evolutionary insertion in the Gly-rich loop of the handle domain. Deleting these residues may have compromised the proper formation of the anti-parallel β-sheet between two heptameric rings and thus impaired the adequate alignment of the catalytic triad.

In BsuClpP, the interaction between ADEP1-Clp proteases is mainly mediated by the bulky core of ADEP1 and hydrophobic residues at the axial pore of the BsuClpP complex (Lee et al., 2010). The examination of the N-terminal segment of the BsuClpP tertiary structure revealed the presence of a hydrophobic cluster formed between two different monomers, which is critical for the activation of BsuClpP (Lee et al., 2010). It has been hypothesized that the binding of ADEP1 to BsuClpP tetradecamer causes a structural transition in the hydrophobic cluster and thus allows an open axial pore conformation (Li et al., 2010; Alexopoulos et al., 2013). Moreover, ADEP1 has also been involved in hydrogen bonding with a specific tyrosine residue within the N-terminal segment of ClpP protease. For BsuClpP, MtuClpP1P2, and SauClpP, it has been reported that ADEP1 binding induces a rotation in the Tyr aromatic ring within the hydrophobic cluster (Schmitz et al., 2014; Ni et al., 2016; Li et al., 2010). In the extended or compressed state of *Sau*ClpP, due to the energy barrier, the side chains of Tyr63 have always been in a staggered conformation and do not allow the ClpP to achieve its active open conformation (Ni et al., 2016). However, the binding of ADEP1 to the hydrophobic cluster of SauClpP enables the alignment of Tyr63 in the eclipsed form and surpasses the energy barrier (Martin et al., 2007; Li et al., 2010). The Tyr63 has been substituted with Ala63 in SauClpP to overcome the energy barrier of Tyr63 side chain rotation without ADEP1 or ATPase chaperones. Tyr63 substitution with Ala63 residue converts SauClp into an ATPase-independent and uncontrollable protease (Ni et al., 2016).

In some pathogenic bacteria where the ClpP protein is a heterocomplex composed of two heptameric rings of ClpP1 and ClpP2, where ClpP2 acts as a chaperone interaction platform

TH-3792_206106022

(Schmitz et al., 2014; Leodolter et al., 2015; Gatsogiannis et al., 2019). Similarly, in the heterocomplex LinClpP1P2, computational modeling has predicted ADEP1 to interact with the hydrophobic cluster of LinClpP2 (Dhara et al., 2021). Moreover, the hydrophobic cluster in the N-terminal segment of LinClpP2 shows the existence of a conserved Tyr62 trigger residue. In agreement, the mutant LinClpP2^{Y62A} gained peptidase activity compared to the inactive LinClpP2. The substitution of a tyrosine aromatic ring with a smaller residue like alanine (Y62A) reduces the energy barrier for substrate entry at the N-terminal axial pore of the LinClpP2^{Y62A} mutant. It thus allows LinClpP2^{Y62A} to achieve an active open conformation. Additionally, the mutant LinClpP2^{Y62A} alone had gained protease activity compared to the LinClpP1P2, and the addition of chaperone LinClpX or antibiotic ADEP1 enhanced its protease activity. Similarly, the heterocomplex mutant LinClpP1P2^{Y62A} also demonstrated enhancement in protease activity.

In BsuClpP, secondary chains (β1, α1, and α2) of adjacent ClpP monomers are involved in inter-subunit interaction, and the significance of the N-terminal hydrophobic cluster has been demonstrated (Lee et al., 2010). In the MSAs of ClpP isoforms from the various pathogenic organisms, the existence of conserved aliphatic residues, Ala and Asn, in the α2 helix close to the N-terminal β-hairpin loop was observed in the majority of organisms. However, in LinClpP2, we were surprised to observe that the residues Ser40 and Lys41 evolutionarily replaced the conserved Ala and Asn, but not in LinClpP1. In LinClpP2, the residues Ser40 and Lys41 lie within the hydrophobic cluster of the N-terminal domain. The significance of these polar and charged residues within the hydrophobic cluster of LinClpP2 is unknown. The results in this study suggest that the double mutant LinClpP2^{S40AK41N} had gained peptidase activity (42%) compared to the inactive LinClpP2. However, adding the LinClpX could not stimulate the protease activity in the pure mutant LinClpP2^{S40AK41N}. In contrast, the mutant heterocomplex LinClpP1P2^{S40AK41N} gained protease activity (52%) with the addition of LinClpX.

Further, the effect of ADEP1 on the protease activity of the double mutant LinClpP2^{S40AK41N} was evaluated to understand the role of Ser40 and Lys41 residues within the N-terminal hydrophobic cluster. The ADEP bound to pure LinClpP2^{S40AK41N} double mutant or its heterocomplex LinClpP1P2^{S40AK41N} showed 58% and 173% protease activity, respectively. We believe the gain in protease activity upon ADEP1 binding to ClpP could be due to the reorganization of electrostatic interactions at the axial pore of ClpP isoforms. Recent studies concerning ADEP1-mediated ClpP activation have demonstrated the role of specific

TH-3792_206106022

electrostatic bonds near the N-terminal region in stabilizing the interconnection between two subunits (Schmitz et al.,2014; Lee et al., 2010; Mabanglo et al., 2019). In NmeClpP of *Neisseria meningitidis* and EcoClpP, the loss of electrostatic interaction mediated by inter-subunit contact has been associated with a gain in proteolytic activity (Mabanglo et al., 2019). In the case of LinClpP2, it is speculated that the Ser40 and Lys41 residues within the α2 helix of the N-terminal region might be electrostatically involved in forming the inter-subunit contact. The substitution of Ser40 with Ala and Lys41 with Asn abolishes inter-subunit communication. It thus allows structural changes within the N-terminal region, which may facilitate the gain of peptidase activity in the pure LinClpP2^{S40AK41N} double mutant. The ADEP1 binding to the hydrophobic cluster of NmeClpP and EcoClpP eliminates inter-subunit contact and thus allows lateral displacement of adjacent subunits to expand the apical surface (Mabanglo et al., 2019). Similarly, the binding of ADEP1 to the double mutant LinClpP2^{S40AK41N} with weakened electrostatic inter-subunit contact causes the enlargement of the axial pore and thus provides an unregulated gain in protease activity.

Thus, the present study led us to conclude the unprecedented GOF by LinClpP having mutations within critical structural motifs (hydrophobic cluster and tyrosine trigger motif). The generation of the double mutant, LinClpP2^{S40AK41N}, results in structural reorganization within the N-terminal hydrophobic cluster and thereby imparts peptidase activity. Similarly, substituting Tyr62 with Ala reduces the energy barrier along the axial pore of mutant LinClpP2^{Y62A} and consequently allows LinClpP2^{Y62A} to achieve an open-gate active conformation. Taken together, the characterization of the GOF mutants of LinClpP furnished novel insights into the activation mechanism of caseinolytic serine proteases.

TH-3792_206106022

Chapter 3

Functional insight into the association of ATPase chaperone, ClpC, with the caseinolytic protease of pathogenic *Leptospira*

This chapter is a modified version of the manuscript published as:

Kumari, S., Ali, A., & Kumar, M. (2024). Nucleotide-induced ClpC oligomerization and its non-preferential association with ClpP isoforms of pathogenic *Leptospira*. ***International Journal of Biological Macromolecules*****,** *266*, 131371.

TH-3792_206106022

3.1. Abstract

The protein homeostasis in bacteria is controlled by Caseinolytic protease-chaperone complexes (Clp complex). The spirochete *Leptospira interrogans* possesses a set of Clp-chaperones (ClpX, ClpA, and ClpC), which may associate functionally with two different isoforms of LinClpP (ClpP1 and ClpP2). The *L. interrogans* ClpC (LinClpC) belongs to class-I chaperone with two ATPase domains separated by a middle domain. Using the size exclusion chromatography, ANS dye binding, and dynamic light scattering analysis, the LinClpC is suggested to undergo nucleotide-induced oligomerization into a higher molecular weight complex. LinClpC associates with either pure LinClpP1 or LinClpP2 isoforms non-preferentially and with equal affinity. Regardless, pure LinClpP isoforms cannot constitute an active protease complex with LinClpC. Interestingly, the heterocomplex LinClpP1P2 in association with LinClpC forms a functional proteolytic machinery and degrades β-casein or FITC-casein in an energy-independent manner. Adding either ATP or $ATP^{\gamma S}$ further fosters the LinClpCP1P2 complex protease activity by nurturing the functional oligomerization of LinClpC. The antibiotic, acyldepsipeptides (ADEP1), displays a higher activatory role on LinClpP1P2 protease activity than LinClpC. Altogether, this work illustrates an in-depth study of hetero-tetradecamer LinClpP1P2 association with its cognate ATPase and unveils a new insight into the structural reorganization of LinClpP1P2 in the presence of chaperone, LinClpC, to gain protease activity.

TH-3792_206106022

3.2. Introduction

In prokaryotic organisms, proteins exist in a complex cellular environment, making up approximately 17 % of the total cell volume (Vendeville et al., 2011). In such an intricate environment, any damaged, misfolded, or unfolded proteins become an inherent menace, as they can facilitate the formation of protein aggregates and impair the function of essential proteins. In *Escherichia coli* and *Salmonella typhimurium*, abnormal or aggregated protein is induced at a higher rate under temperature or oxidative stress conditions (Farr et al., 1991). To counteract such stress conditions, bacteria have evolved pathways for protein quality control that aid in proper protein folding, repair protein damage, and remove terminally damaged, misfolded, or unfolded proteins within the cell (Mogk et al., 2011; Kim et al., 2013). The quality control enzyme expression is often upregulated in response to stress conditions compromising proteostasis (Mogk et al., 2011). Several classes of molecular enzymes (foldases, unfoldases, and disaggregases), also known as chaperones, work in coordination and maintain protein homeostasis (Mogk et al., 2011; Kim et al., 2013). Foldases such as trigger factor, DnaK, and GroE system facilitate the optimal folding of newly synthesized proteins (Mogk et al., 2011). The unfoldases play an essential role when a protein is busted beyond the repair or dispensable and thus deemed to be eliminated (Sauer et al., 2011). For elimination, the unfoldases from the HSP100 (Heat Shock Protein 100) family of proteins can unfold and deliver the damaged protein into the proteolysis compartments (Sauer et al., 2011). Contrary to this, the enzyme disaggregase from the HSP100 family of proteins acts on aggregated polypeptides to untangle them (Mogk et al., 2018).

The HSP100 family of molecular chaperones, comprising Caseinolytic protein (Clp), belongs to the ATPase associated with diverse cellular activities (AAA+) superfamily of proteins. The HSP100/Clp proteins are categorized into two major classes: class I and class II. The ClpA-D belongs to the class I chaperone and contains two nucleotide-binding domains (NBDs) separated by a middle domain of variable length. The class II chaperones, ClpX-Y, are shorter in size and have single NBD (Schirmer et al., 1996) The chaperone ClpC is present in most gram-positive bacteria, including *Bacillus subtilis*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, *Streptococcus mutans*, *Listeria monocytogenes,* and some gram-negative bacteria, including *Synechococcus elongatus*, *and Chlamydia trachomatis* (Kruger et al., 1994; Chatterjee et al., 2005; Kar et al., 2008; Lemos et al., 2002; Rouquette et al., 1998; Andersson et al., 2006; Pan et al., 2023). Although ClpCs are encoded in the genome of diverse microorganisms, their specific functions and association with caseinolytic proteases lack clarity. The ClpCs orthologs from various microbes, including *B. subtilis* (BsuClpC), *S. aureus*

TH-3792_206106022

(SauClpC), *S. mutans* (SmuClpC), and *L. monocytogenes* (LmoClpC), have been documented to provide heat stress tolerance and play an essential role in virulence (Kruger et al., 1994; Chatterjee et al., 2005; Lemos et al., 2002; Rouquette et al., 1998).

Additionally, a subset of either of the Clp ATPases (ClpA, ClpC, ClpX) from diverse microorganisms is coaxially stacked with ring-like protease complexes, ClpP (Casinolytic Protease P), and functions to regulate the access of folded substrates to the ClpP proteolysis chamber (Clare et al., 2013). The *Escherichia coli* ClpP (EcoClpP) belongs to the family of serine proteases where proteolytic active sites are confined within an internal chamber formed by the association of two EcoClpP heptameric rings (Wang et al., 1997). Access to the catalytic chamber of EcoClpP homotetradecamer is regulated by a narrow axial pore located at the center of each EcoClpP heptameric ring, allowing entry of only small peptides (up to 30 residues) (Thompson et al., 1994). The N-terminal region of the EcoClpP homotetradecamer has been divided into two parts. The first part with 6-7 hydrophobic residues circumventing the axial pore is termed 'axial pore lining'. While the second part, called ‘axial protrusion’, has 8-9 charged or hydrophilic residues projecting outwards from the axial surface (Gribun et al., 2005; Bewley et al., 2006). The hydrophobic pockets on the axial surface of EcoClpP provide a binding site for the conserved IGF/L loop of EcoClp chaperones (Gribun et al., 2005). Among other functions, the *E. coli* chaperones ClpA (EcoClpA) and ClpX (EcoClpX) exhibit nucleotide-induced hexameric assembly and bind to the heptameric rings present in both outer faces of EcoClpP (Grimaud et al., 1998). The axial channels of hexameric EcoClpX align with the axial pore of EcoClpP and permit the translocation of unfolded substrates into the EcoClpP catalytic chamber (Grimaud et al., 1998).

Recently, an antibiotic, acyldepsipeptides (ADEPs), produced by *Streptomyces hawaiiensis* has been reported to regulate the functionality of bacterial ClpP (Brötz-Oesterhelt et al., 2005). The ADEP docks at the hydrophobic cluster of SauClpP, thereby rendering a broadening of its axial pore and permitting protein substrate entry (Brötz-Oesterhelt et al., 2005). In *E. coli* and *B. subitilis*, various biochemical studies suggested that ADEPs reprogram ClpP into unregulated protease to degrade newly synthesized bacterial proteins, leading to proteome imbalance and cell death (Kirstein et al., 2009). The ADEP binding to BsuClpP prevents its interaction with BsuClpC and additionally can dissociate the preformed BsuClpCP complex (Kirstein et al., 2009).

In most genera, including *Bacillus, Streptococcus, Neisseria, Helicobacter, Escherichia, Salmonella*, and *Plasmodium*, caseinolytic proteases are composed of a single ClpP isoform, creating a homo-tetradecamer (Yu et al., 2007). However, certain genera of microorganisms

TH-3792_206106022

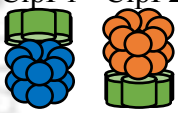
like *Mycobacterium, Chlamydia, Pseudomonas*, and *Leptospira* exhibit two different ClpP isoforms forming a hetero-tetradecamer (Yu et al., 2007; Dhara et al., 2019). The well-conserved and characterized homo-tetradecameric ClpPs reveal a consistency in their structure and functionality (Queralto et al., 2023). However, in the case of organisms with multiple ClpP isoforms, a diversity among functional and structural organization of ClpP was observed (**Table 3.1**). In the multi-ClpP containing microorganisms, including *L. monocytogenes*, *Pseudomonas aeruginosa,* and *Clostridium difficile,* the active peptidase is composed of either homo-tetradecamer or hetero-tetradecamer of ClpP (Zeiler et al., 2013; Hall et al., 2017; Lavey et al., 2018). However, in the case of ClpP of *M. tuberculosis*, *C. trachomatis,* and *L. interrogans*, the active peptidase is exclusively constituted of a heptameric assembly of ClpP1 and ClpP2 stacked together to form a hetero-tetradecamer (Li et al., 2016; Pan et al., 2019; Dhara et al., 2019).

Additionally, among different pathogenic bacteria, significant functional divergence in the Clp chaperones (ClpX, ClpA and ClpC) has been on record (Grimaud et al., 1998; Wang et al., 2011). The two well-characterized hexameric chaperones (EcoClpA and EcoClpX) tend to associate with homo-tetradecamer EcoClpP and impart protease activity (Grimaud et al., 1998). While, the chaperone BsuClpC is dependent upon an adaptor protein (MecA) for activation of protease activity in BsuClpP homo-tetradecamer (Wang et al., 2011). Among other Clp chaperone functional divergences, the MtuClpC and CtrClpC exclusively are reported to interact with ClpP2 isoforms and impart a gain in protease activity of hetero-tetradecamer (ClpP1P2) (Schmitz et al., 2014; Pan et al., 2019).

In this study, to further have a better understanding of Clp chaperone functional divergence, the gene encoding for LinClpC of *L. interrogans* was cloned, over-expressed, and functionally characterized. In this study, LinClpC possesses an intrinsic ATPase activity that remains unaltered in the presence of the model substrate, β-casein. The LinClpC interacts unconditionally and with equal affinity to each LinClpP isoform (LinClpP1 and LinClpP2); however, it imparts protease activity in LinClpP when associated with a single self-assembled isoform (LinClpP1 or LinClpP2). Unprecedently, the association of LinClpC with the LinClpP1P2 heterocomplex transforms the complex machinery into an energy-independent active LinClpP protease on unstructured model protein substrates. Additionally, supplementation of either hydrolysable or non-hydrolysable ATP promotes the association of LinClpC with LinClpP1P2 and thus augmented LinClpCP1P2 protease activity. Altogether, this study provides insight into the functionality of LinClpC and its association with LinClpP isoforms.

TH-3792_206106022

Table 3.1. Overview of Clp protease systems from bacterial species that encode two ClpP isoforms

	Mycobacterium tuberculosis	*Listeria monocytogenes*	*Chlamydia trachomatis*	*Pseudomonas aeruginosa*	*Clostridium difficile*
ClpP1	Tetradecamer  (inactive)	Heptamer  (inactive)	Heptamer  (inactive)	Tetradecamer  (active)	Tetradecamer  (active)
ClpP2	Tetradecamer  (inactive)	Tetradecamer  (active)	Heptamer  (inactive)	Heptamer  (inactive)	Tetradecamer  (inactive)
ClpP1P2	Tetradecamer  (active)	Tetradecamer  (active)	Tetradecamer  (active)	Tetradecamer  (active)	Tetradecamer  (active)
Cognate chaperone association with ClpP isoform	 ClpP2	 ClpP2	 ClpP2	ClpP1 	ClpP1 ClpP2 
References	Schmitz et al., 2014; Leodolter et al., 2015	Zeiler et al., 2013; Gatsogiannis et al., 2019	Pan et al., 2019	Hall et al., 2017	Lavey et al., 2018

3.3. Materials and Methods

3.3.1. Bacterial strains, media, and growth conditions

Leptospira interrogans serovar Copenhageni strain Fiocruz L1-130 was grown in Ellinghausen-McCullough-Johnson-Harris (EMJH) medium (Difco, India) at 29 °C supplemented with enrichment media (Difco, India) along with 5-fluorouracil (100 µg/mL). The bacteria were sub-cultured for maintenance after 6-7 days of growth. The *E. coli* strains DH5α and BL21 (Novagen) were employed to maintain and overexpress recombinant proteins, respectively. Both *E. coli* strains were grown in Luria Bertani (LB) broth (Himedia) or LB-agar (Himedia) medium supplemented with appropriate antibiotic (100 µg/mL) at 37 °C.

TH-3792_206106022

3.3.2. Cloning, over-expression, and purification of recombinant proteins

The genomic sequence of *L. interrogans* serovar Copenhageni strain Fiocruz L1–130, available at the National Centre for Biotechnology Information (NCBI), was used for primer designing and cloning. The full-length *clpC* (*LIC10339*; 2541 bp) was PCR amplified from the genomic DNA of *L. interrogans* using sequence-specific primers (*clpC*_F: CTAGCTAGCATGCTGGAATTTACAAAAAGAGCT, *clpC*_R: CCGCTCGAGAACTAA GGCAATTTCTTCCGA. The full-length ClpC (LinClpC) with a C-terminal 6×His tag was cloned in the pET-23a plasmid and over-expressed in *E. coli* BL21 cells. Bacterial cultures were induced with isopropyl β-D-1-thiogalactopyranoside (IPTG; 1 mM) for 16 h at 18 °C. The over-expressed LinClpC was purified by affinity column chromatography as described previously for LinClpP isoforms (Dhara et al., 2019). The purified protein was visualized on a 12% sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel by Coomassie staining. Protein concentrations were estimated by the Bradford method with bovine serum albumin as the standard.

3.3.3. Generation of polyclonal antibodies and immunoblot analysis

Polyclonal antibodies specific to purified LinClpC of *L. interrogans* were generated in mice as described elsewhere (Dhara et al., 2019). Briefly, BALB/c mice 4–6 weeks old (n = 5) were immunized with purified LinClpC. The purified protein (15–20 µg in phosphate buffer saline) emulsified in Freund's complete adjuvant (FCA; Santa Cruz Biotechnology #SC-3727) was injected subcutaneously for primary immunization. Immunized mice were given two subsequent booster injections of LinClpC protein (15–20 µg) emulsified in Freund's incomplete adjuvant (FIA; Santa Cruz Biotechnology #SC-3726) at 14 and 21 days of primary immunization. After seven days of the second booster, blood was collected from each mouse by retro-orbital bleeding, and then the mouse was sacrificed by the atlantooccipital dislocation method. Sera obtained were pooled, and the antibody titer was analyzed by enzyme-linked immunosorbent assay (ELISA). The immunization experiments were performed at the Department of Microbiology, College of Veterinary Science, Assam Agricultural University, Guwahati, India, after approval by the Institutional Animal Ethics Committee (approval no.770/ac/CPCSEA/FVSC/AAU/IAEC/13-14).

To detect LinClpC expressions in *Leptospira*, whole-cell lysates of 3 $\times 10^9$ spirochetes were re-suspended in sodium dodecyl sulfate (SDS) loading dye. The resulting lysates of *Leptospira* were resolved in 12% SDS-PAGE and transferred to a nitrocellulose membrane (Bio-Rad). The membranes were blocked with 5% non-fat dried milk prepared in Tris-buffered

TH-3792_206106022

saline (TBS; pH 8.0) containing 0.1% Tween 20 (TBS-T) and probed with mouse anti-LinClpC (1:1000) antibodies for 2 h at room temperature. After being washed, the membranes were incubated with horse radish peroxidase (HRP)-conjugated goat anti-mouse IgG (1:5000; Sigma) for 1 h, and immunoblots were developed by adding the chemiluminescence substrate (Thermo Scientific, catalog no. 32209). All dilutions of antibodies were prepared using 2% non-fat dried milk in 0.1% TBS-T.

3.3.4. LinClpC ATPase activity assay

The LinClpC ATPase activity (μM min^{-1}) was measured by determining the release of free phosphate using a malachite green phosphate assay kit (Sigma; MAK307). For a standard assay, LinClpC (1 μM) was incubated with ATP (5 mM) in 40 μL assay buffer (50 mM Tris-Cl, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$; pH 7.8) for 60 minutes at 37 °C. The free phosphate released during the ATPase reaction was determined from the standard phosphate graph as per the manufacturer's protocol. The reaction kinetics study with increasing ATP concentration (0-8 mM) was performed to monitor ATPase activity. The result obtained was analyzed on the Hill plot in Origin 9.0 software described previously for LinClpX (Dhara et al., 2019). In addition, the ATPase activity of LinClpC was measured in the presence of model substrate, β-casein, and ClpP isoforms (LinClpP1 and LinClpP2). In every ATPase assay, the reaction mixture containing LinClpC and the model substrate was prepared in the assay buffer and pre-incubated for 10 minutes at 37 °C before the addition of ATP. The absorbance measurements at 620 nm were performed in a spectrofluorometer (iTECAN infinite M PLEX). All the experiments were performed thrice independently and in duplicates.

3.3.5. Size exclusion chromatography (SEC)

The purified LinClpC (~5 μM) was loaded onto a HiLoad™ 16/600 Superdex™ 200 pg column (GE Healthcare #28-9893-35) after equilibration in a buffer (50 mM Tris-HCl; pH 8, 300 mM NaCl, 10% glycerol). In another assay, the LinClpC was incubated with ATP (2 mM) for 20 minutes at 4 °C in assay buffer (50 mM Tris-Cl; pH 7.8, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$). The pre-incubated LinClpC complex was loaded onto the column equilibrated with the buffer (50 mM Tris-HCl; pH 8, 300 mM NaCl, 10% glycerol) and run with a 1 mL/min flow rate at room temperature. The molecular weight standards (Sigma, catalog no MWGF-200) used for the determination of oligomeric size are β-amylase (200 kDa), alcohol

TH-3792_206106022

dehydrogenase (158 kDa), albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome C (12.4 kDa).

Similarly, the elution profile of pure LinClpP1 and LinClpP2 isoforms was monitored. Next, the chaperone-protease complex was allowed to be formed by pre-incubating equimolar LinClpC with LinClpP1 or LinClpP2 for 10 minutes at 37 °C in assay buffer. The assembled LinClpCP1 or LinClpCP2 complex was then loaded onto the column, and the elution profile for each complex was examined. Aliquots were collected at different elution volumes (mL), and the proteins were resolved on 12% SDS-PAGE. The respective LinClpC and LinClpP isoform bands were visualized by Coomassie staining.

3.3.6. ANS binding assay

A hydrophobic fluorophore, 8-anilino-1-naphthalenesulfonic acid (ANS), was used for the protein oligomerization study (Semisotnov et al., 1991; Bernot et al., 2005). Briefly, the LinClpC (1 µM) was pre-incubated with or without ATP (2 mM) for 20 minutes at 4 °C in an assay buffer (50 mM Tris-Cl; pH 7.8, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$). Next, ANS (10 µM) was mixed with pre-incubated LinClpC and incubated for 30 minutes at room temperature in the dark. A control reaction having ANS (10 µM) in assay buffer was also kept under similar conditions. Fluorescence spectra were recorded in a black, flat-bottom 96-well microplate (Invitrogen) using a spectrofluorometer (iTECAN infinite M PLEX). All samples were excited at 350 nm, and emission spectra were recorded from 400 to 750 nm with 2-nm increments.

3.3.7. Dynamic light scattering

DLS experiments were performed on a LITESIZER 500 (Anton Paar) at 25 °C as described previously (Kumari et al., 2024). Briefly, the LinClpC (3 µM) was pre-incubated with or without ATP (2 mM) for 20 minutes at 4 °C in assay buffer (50 mM Tris-Cl; pH 7.8, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$). The pre-incubated samples were added to the polystyrene cuvettes, followed by the DLS analysis. Each investigation was performed thrice, and the hydrodynamic diameter was measured as the average of these replicates.

3.3.8. Fluorescence Resonance Energy Transfer (FRET) analysis

The FRET analysis between a fluorophore donor molecule, Cy3, and an acceptor molecule, Cy5, was performed as described previously (Li et al., 2010). The pure LinClpC and LinClpP isoforms were labeled with fluorophores, Cy3 and Cy5, respectively, using BioVision's

TH-3792_206106022

EZLabel™ kit (Cat. No. K837-5 and K839-5). The labeled proteins (1 µM) were pre-incubated for 10 minutes at 37 °C in 50 µL assay buffer (50 mM Tris-Cl; pH 7.8, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$) to allow the formation of complexes (LinClpC-Cy3 + LinClpP1-Cy5 and LinClpC-Cy3 + LinClpP2-Cy5). Additionally, FRET analysis was performed using ATP-bound LinClpC-Cy3 with LinClpP-Cy5. The control reactions contained pure labeled proteins (LinClpC-Cy3, LinClpP1-Cy5, and LinClpP2-Cy5) in an assay buffer. The spectra were acquired at a fixed excitation wavelength of 488 nm, and emission was acquired at a range of 540 - 750 nm.

3.3.9. Immunoassay

The interaction between LinClpC and LinClpP isoforms was examined by ELISA as described previously (Dixit et al., 2021). Briefly, 96-well microtiter plates were coated with pure LinClpP1/ LinClpP2/ BSA (1 µM/well) in a total volume of 50 µL overnight at 4 °C. Thereafter, the unbound proteins were discarded from the wells, and each well was blocked with 3% BSA solution for 2 h at 37 °C and then was overlaid using different concentrations of free or ATP bound LinClpC (0-3 µM) for 2 h at room temperature. After three washes with phosphate buffer saline + 0.05% Tween 20 (PBS-T) buffer, mouse anti-LinClpC (1:1000) was used. The interaction was detected using HRP-conjugated anti-mouse secondary antibodies (1:5000) with TMB (Trimethylbenzidine, Thermo Scientific) as substrate. The absorbance was measured at a wavelength of 450 nm after terminating the reaction using 1 M H_2SO_4 as per the manufacturer's instructions. All the experiments were performed thrice independently and in duplicates.

3.3.10. Protease activity of LinClpP isoforms

Degradation of protein substrates, β-Casein (S1) and FITC-casein (S2), was performed in assay buffer (50 mM Tris-Cl; pH 7.8, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$). For each 50 µL reaction, the complex (LinClpCP1, LinClpCP2, and LinClpCP1P2) was constituted by pre-incubating equimolar (3 µM) of LinClpC with pure LinClpP1, LinClpP2, or heterocomplex LinClpP1P2 for 10 minutes at 37 °C, followed by the addition of model substrate, β-casein (5 µM). The reaction was conducted with or without ATP (2 mM). The protease assay was performed for 120 minutes at 37 °C, where 15 µL of the reaction mixture was aliquoted at 0, 60, and 120-minute intervals. The aliquots were mixed with SDS loading dye and heated for 10 minutes at 95 °C to terminate the degradation reaction. The time-dependent degradation of β-casein was visualized by resolving the reaction aliquots on 12% SDS-PAGE, followed by

TH-3792_206106022

Coomassie staining. To further validate the protease activity in LinClpCP1P2 heterocomplex, a more sensitive fluorometric analysis was performed using FITC-casein (S2). Briefly, LinClpCP1P2 (1 µM) complex in assay buffer was mixed with 5 µL of FITC-casein (1.5 µg/µL) in a total reaction volume of 25 µL. The reaction was allowed with or without the addition of ATP or $ATP_{\gamma}S$ (2 mM) for 120 minutes at 37 °C, followed by termination using trichloroacetic acid (0.6 N). A similar protease assay on S2 was performed using equimolar concentrations of LinClpC, ADEP1, or both. The fluorescence of released FITC during proteolysis was monitored in black, flat-bottom 96-well microplates (Invitrogen) every 15-minute interval for 120 minutes at 37 °C in a spectrofluorometer (iTECAN infinite M PLEX). The excitation and emission wavelengths used were 490 nm and 525 nm, respectively. All experiments were performed thrice independently, each in duplicate.

The proteolytic activity of LinClpCP1P2 machinery (1 µM) on eGFP model substrate (2 µM) was checked in the presence and absence of ATP (2 mM) for 15 minutes at 37 °C. The degradation was checked by measuring the loss of fluorescence of eGFP at an excitation wavelength of 485 nm and an emission wavelength of 525 nm.

Table 3.2. Percent identity of ClpC orthologs in selective pathogenic bacteria with LinClpC

Organism (protein name)	Uniprot ID	Percent identity with LinClpC sequence (Percent query coverage)
Chlamydia trachomatis (CtrClpC)	O84288	47.1 (93%)
Staphylococcus aureus (SauClpC)	Q2G0P5	53.2 (95%)
Bacillus subtilis (BsuClpC)	P37571	52.6 (96%)
Mycobacterium tuberculosis (MtuClpC1)	P9WPC9	49.8 (93%)
Synechococcus elongatus (SelClpC)	A0A3G6WWS0	52.2 (95%)
Streptococcus mutans (SmuClpC)	Q8DS14	41.4 (95%)
Listeria monocytogenes (LmoClpC)	Q8YAB6	52.1 (95%)

3.4. Results

3.4.1. *Leptospira* ClpC (LinClpC) sequence analysis and domain characterization

The primary amino acid sequence of LinClpC (846-amino-acid protein) was aligned with well-studied ClpC orthologs from various pathogenic microorganisms to identify the critical functional motifs. The LinClpC shows a high degree of sequence identity (> 50%) with S. aureus ClpC (SauClpC; 53.2%), B. subtilis ClpC (BsuClpC; 52.6%), S. elongatus ClpC (SelClpC; 52.2%), and *L. monocytogenes* ClpC (LmoClpC; 52.1%) **(Table 3.2)**. The LinClpC shows the existence of conserved functional domains (N-terminal domain, nucleotide-binding

TH-3792_206106022

domain, middle domain, and C-terminal domain) which are critical features of the ATPase family of chaperones **(Figure 3.1)**. In addition, the functional and the conserved primary sequences of LinClpC components with their known orthologs (CtrClpC, SauClpC, BsuClpC, MtuClpC1, SelClpC, SmuClpC, and LmoClpC) were aligned **(Figure 3.1)**. For clarity, the full-length MSA of LinClpC with its orthologs is also represented in the supplementary (**Figure 3S1**). The LinClpC has four functional domains and includes the N-terminal domain (NTD), nucleotide-binding domain I (NBD-I), middle domain (MD), and nucleotide-binding domain II (NBD-II). The NBD-II has a long carboxy-terminal sub-domain (CTD). One of the conserved features of the ATPase associated with the diverse cellular activity (AAA+) family of proteins is the occurrence of Walker A and Walker B motifs in the nucleotide-binding domain (NBD). In SmuClpL of *Streptococcus mutans*, the lysine residues (K128; K459) within Walker A motifs of NBD-I and NBD-II are reported to be important for ATP binding (Jana et al., 2020).

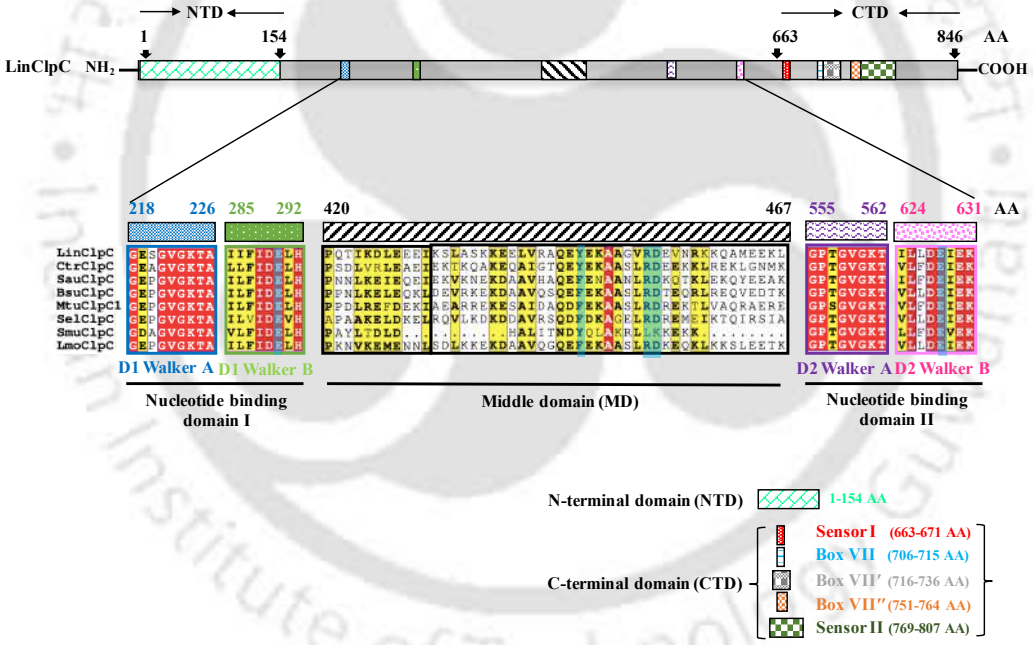


Figure 3.1. Conserved functional domains of LinClpC and its comparison to its orthologs. Multiple sequence alignment of ClpC orthologs from various pathogenic bacteria was performed using Clustal Omega software. The ClpP orthologs used for amino acid alignment are LinClpC (Q72VF8), CtrClpC (O84288), SauClpC (Q2G0P5), BsuClpC (P37571), MtuClpC1 (P9WPC9), SelClpC (A0A3G6WWS0), SmuClpC (Q8DS14), and LmoClpC (Q8YAB6). Identical and semi-identical residues in ATPases are highlighted in red and yellow colors, respectively. The conserved key structural motifs in the LinClpC subunits are indicated in different colored boxes, and the consensus sequence of each domain is mentioned in the lower panel. In LinClpC, the N-terminal domain (154 amino acids) is followed by 2 nucleotide-binding domains (NBD-I and II) having conserved Walker A and Walker B motifs. The key residues for ATP hydrolysis within domain I (E290) and domain II (E628) of Walker B motifs were highlighted in blue color. The two NBDs were separated by a middle domain (MD) from residue 420 to 467. The non-conserved Y446 and conserved R454D455 within MD were highlighted with green color. The residues beyond

TH-3792_206106022

663 to 846 were considered as the C-terminal domain having different sensor motifs (sensor I, Box VII, Box VII', VII", and Box II) denoted with different colored boxes.

The Walker B motifs within NBD-I and NBD-II of CtrClpC of *Chlamydia trachomatis* contain conserved Glu residues (E306; E644), which are essential for hydrolysis of ATP (Pan et al., 2023). Likewise, the Walker A motifs of LinClpC revealed the existence of conserved Lys residues (K224; K561) and are proposed to be important in binding with ATP. Similarly, the critical Glu residue (E290; E628) involved in ATP hydrolysis was conserved within the Walker B motifs of LinClpC. Next, the middle domain of LinClpC, involved in oligomerization, shows a sequence disparity at Tyr446 (replaced with conserved Phe) with well-studied ClpC orthologs (SauClpC, BsuClpC, and MtuClpC). There is a long C-terminal domain (CTD) in LinClpC and its orthologs, which contains five conserved motifs: Sensor I, Box VII, Box VII', Box VII", and Sensor II. In *M. tuberculosis*, these CTD motifs were reported to be involved in nucleotide-induced oligomerization of the MtuClpC1 (Bajaj et al., 2012). Although the ClpC of several pathogenic bacteria shows sequence conservation, one remarkable difference was the presence of variable lengths of the CTD. A long C-terminal extension beyond the conserved sensor II was observed in LinClpC and agrees with that of MtuClpC1 and CtrClpC (Kar et al., 2008; Pan et al., 2023).

Table 3.3. Comparative ATPase activity in K_M of different Clp chaperones

ATPase chaperone	K_M (mM)	Reference
MtuClpC1 of *Mycobacterium tuberculosis*	5.6 ± 1.4	Kar et al., 2008
AthClpB3 of *Arabidopsis thaliana*	1.07 ± 0.08	Rosano et al., 2011
AthClpC2 of *Arabidopsis thaliana*	1.42	Rosano et al., 2011
CcrClpB of *Corynebacterium crenatum*	0.29	Huang[a] et al., 2020
CcrClpX of *Corynebacterium crenatum*	0.39	Huang[b] et al., 2020
SelClpC of *Synechococcus elongatus*	2.05	Andersson et al., 2006
LinClpX of *Leptospira interrogans*	0.78 ± 0.07	Dhara et al., 2019

TH-3792_206106022

3.4.2. LinClpC displays intrinsic ATPase activity

The gene encoding leptospiral LinClpC was cloned, over-expressed, and purified by Ni-NTA affinity chromatography to understand its function (**Figure 3S2a**). The anti-mouse LinClpC generated using purified LinClpC could recognize the native LinClpC (~ 96 kDa) expression on the immunoblot of *Leptospira* lysate, suggesting the protein to be immunogenic and expressed by *Leptospira* under *in vitro* growth conditions (**Figure 3S2b**). Next, a malachite green assay was performed to measure the LinClpC intrinsic ATP hydrolysis activity. The malachite green assay measures the free phosphate released after ATP hydrolysis. The LinClpC exhibited dose-dependent ATPase activity, which further allows us to determine the saturation level of ATP for the chaperone. The LinClpC exhibited a surge in ATP hydrolysis rate with increasing ATP concentration (0.05 – 8 mM) and attained saturation at a specific concentration (>6 mM) of ATP (**Figure 3.2A**). The LinClpC presented Michelis-Menten kinetics with V_{max} (13.8 ± 1.2 µM) and K_M (2.5 ± 0.5 mM) at different reaction variables. The calculated K_M (ATP affinity) for LinClpC was similar to well-studied AAA+ chaperones of various organisms (**Table 3.3**).

Further, the presence of β-casein (an unstructured-model substrate) as an activator on LinClpC ATPase activity was evaluated. Interestingly, the LinClpC did not exhibit any surge in ATPase activity in the presence of β-casein substrate (**Figure 3.2B**), and was in contrast to various other reported AAA+ proteins, including EcoClpA, LinClpB, and LinClpX (Dhara et al., 2019; Seol et al., 1994; Krajewska et al., 2017). Similar to LinClpC, the ATPase activator (β-casein) has not shown any boost on the ATPase activity of EchClpB (*Ehrlichia chaffeensis*) and BsuClpC (Zhang et al., 2013; Trentini et al., 2016). However, the BsuClpC ATPase activity was stimulated in the presence of activator phosphorylated casein substrate ($casein^{pArg}$) and adaptor proteins (BsuMcsB and BsuMecA) (Trentini et al., 2016; Kirstein et al., 2006; Kirstein et al., 2007). The BsuMcsB is an arginine kinase that was proposed to assist the oligomerization of BsuClpC and assembly into a functional BsuClpCP complex to degrade protein substrate (Kirstein et al., 2007).

The BsuMcsB role as activator of ATPase activity in BsuClpC was analogous to the BsuMecA activity and, thus, acts as an adaptor protein for BsuClpC (Kress et al., 2007). Preliminary analysis of the *Leptospira* genome suggests the existence of the *Linmcs*B gene (*LIC10340; 786 bp*) with an intergenic distance of 149 bp, downstream to *Linclp*C (*LIC10339*) (**Figure 3S3)**. However, the genes encoding for the adaptor protein, mecA, and the activator protein *mcsA* were absent in *the Leptospira* genome.

TH-3792_206106022

For comparative analysis, the schematics of *mcsB* and *clpC* arrangements in *S. aureus*, *B. subtilis,* and *L. monocytogenes* genome is also shown (**Figure 3S3)**. The presence of arginine kinase (LinMcsB) adjacent to LinClpC in the *Leptospira* genome hints that LinMcsB may act as an adaptor protein, similar to *B. subtilis*, and warrants further study.

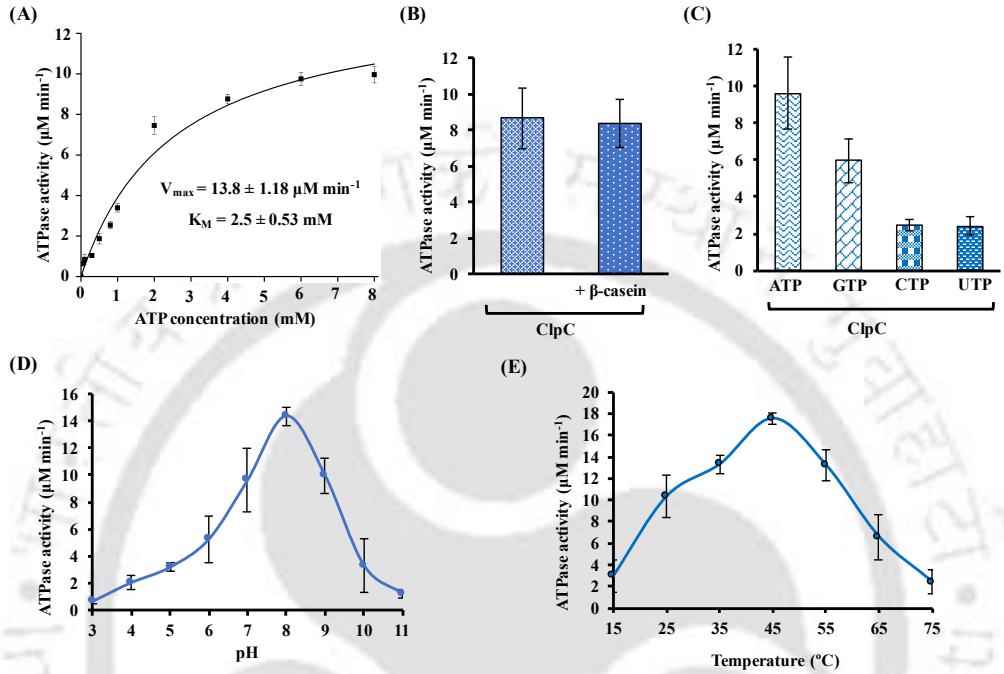


Figure 3.2. Nucleotide hydrolysis property of LinClpC under diverse parameters. (A) Effect of ATP at increasing concentration on the LinClpC ATPase activity. The kinetics of the LinClpC enzymatic reaction on ATP was determined by non-linear curve fitting in the Origin software. The V_{max} (13.8 ± 1.2 µM min^{-1}) and K_M (2.5 ± 0.5 mM) of the reaction were estimated. **(B)** ATPase activity analysis in the absence (-) and presence (+) of model substrate, β-casein. **(C)** The LinClpC nucleotide hydrolysis rate on various nucleotides (GTP, CTP, UTP) and its comparison with the rate of hydrolysis of ATP. **(D)** The ATPase activity of LinClpC over a broad range of pH 3-11. **(E)** The ATPase activity of LinClpC in the temperature range of 25 - 75 °C. The data represent the Mean ± Standard error mean (SEM) from the three independent experiments performed in duplicates.

Next, LinClpC's ability to hydrolyse other nucleotides (GTP, CTP, and UTP) was examined. LinClpC displayed hydrolysis of GTP, CTP, and UTP substrates; however, the hydrolytic efficiency on GTP (60%), CTP (25%), and UTP (25%) was lower than that of ATP (100%) (**Figure 3.2C**). The influence of environmental conditions (pH and temperature) on the LinClpC ATPase activity was also examined. The ATPase activity of LinClpC increased gradually from pH 3 - 8 and was optimum at pH 8. An increment of pH beyond 8 resulted in reduced ATPase activity (**Figure 3.2D**). The ATPase activity of LinClpC, influenced by the change of temperature (15 - 75 °C), displayed a bell-shaped curve. LinClpC retained its activity

TH-3792_206106022

over a broad range of temperatures (25 - 65 °C), and the optimal activity was measured at 45 °C (**Figure 3.2E**).

3.4.3. The LinClpC undergoes nucleotide-induced oligomerization

Several chaperones, including EcoClpA (Kress et al., 2007), MtuClpC1 (Kar et al., 2008), SmuClpL (Jana et al., 2020), and LinClpB (Krajewska et al., 2017) have been documented for undergoing nucleotide-induced oligomerization into higher-order complexes. Thus, the effect of ATP on the oligomerization of purified recombinant LinClpC was determined using size exclusion chromatography (**Figure 3.3A**). The LinClpC elution profile indicates its existence as a dimer (192 kDa) in the solution in the absence of ATP. However, the elution profile shifted towards a higher molecular weight complex when LinClpC was incubated with ATP (2 mM). To bolster our findings, the nucleotide-induced oligomerization of LinClpC was evaluated by two additional independent assays. The binding of 8-anilino-1-naphthalenesulfonic acid (ANS), a hydrophobic fluorophore, to the exposed hydrophobic regions of proteins increases its fluorescence and shifts the fluorescence spectra towards a shorter wavelength (Semisotnov et al., 1991; Bernot et al., 2005).

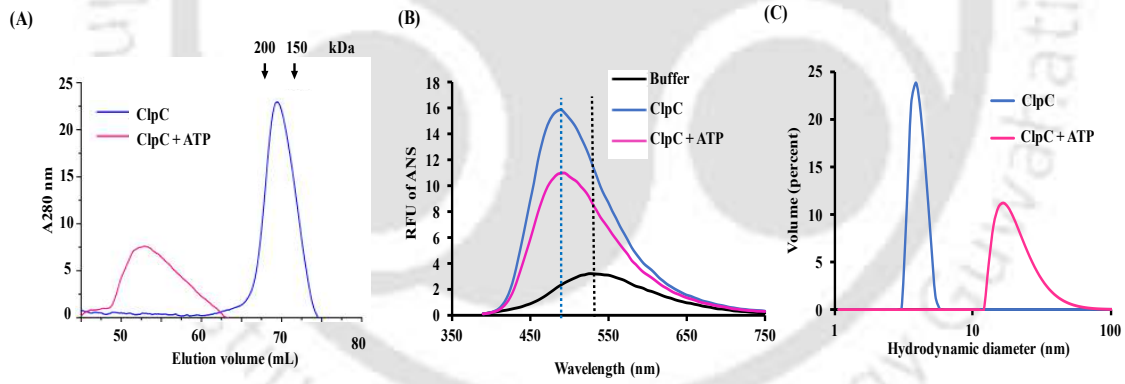


Figure 3.3. Nucleotide-induced oligomerization of LinClpC. **(A)** SEC analysis of LinClpC in the absence (-) and presence (+) of ATP. The elution profile of LinClpC shifts towards a higher molecular weight complex in the (+) of ATP (blue line). **(B)** ANS binding assay to study the oligomerization of LinClpC in the absence (-) or presence (+) of ATP. The emission spectra of hydrophobic fluorophore (ANS) were recorded in the range of 400-750 nm with an excitation wavelength of 350 nm. **(C)** The size distribution profile of LinClpC in the absence (-) and presence (+) of ATP was checked using DLS. The distribution of LinClpC species in the (-) and (+) of ATP was represented in terms of volume (percent). Each experiment to study the nucleotide-induced oligomerization of LinClpC (SEC analysis, ANS binding assay, and DLS analysis) was performed thrice, and one representative data point is shown.

However, the folding or oligomerization of proteins masks the hydrophobic regions, and so does the magnitude of ANS binding (Bernot et al., 2005). Therefore, using the same principle

TH-3792_206106022

of ANS dye binding, the oligomerization of LinClpC from dimer to higher molecular weight complexes was tested. The pure LinClpC or the one pre-incubated with ATP was allowed to bind with ANS, and the fluorescence spectra were analyzed at an excitation wavelength of 350 nm and the emission wavelength range of 400-750 nm (**Figure 3.3B**). The control reaction (only ANS) exhibited maximum emission at a wavelength of 525 nm, and the ANS bound with LinClpC exhibited a maximum emission at a relatively shorter wavelength (490 nm) and with higher fluorescence intensity. However, as expected, the ANS fluorescence was reduced when binding to LinClpC supplemented with ATP was allowed. We believe that nucleotide-induced oligomerization of LinClpC may mask hydrophobic residues within each subunit, thereby reducing ANS binding.

Additionally, macromolecules' oligomerization states can be validated using dynamic light scattering (DLS) analysis (Shamir et al., 2006). Previous studies on DLS analysis of globular proteins suggest that the hydrodynamic diameter (D_h) is linearly correlated with protein concentration (Lorber et al., 2012; Takeuchi et al., 2014). Thus, we have performed the DLS analysis of different concentrations of LinClpC (1 - 5 µM) and observed a concentration-dependent increase in the D_h of LinClpC, which attains saturation at 3 µM of LinClpC (**Figure 3S4**). The D_h of pure LinClpC (3 µM) measured 3.63 ± 0.15 nm with a sharp peak (**Figure 3.3C**). While the LinClpC (3 µM) supplemented with ATP (2 mM) showed a broad peak with an average diameter (D_h) of 18.04 ± 1.79 nm, indicating its higher oligomeric assembly. Altogether, using the SEC, ANS binding, and DLS analysis, it is suggested that the LinClpC undergoes nucleotide-induced oligomerization into higher complexes.

3.4.4. Non-preferential association of LinClpC to LinClpP isoforms

To date, whether the LinClpP isoforms show asymmetry towards their cognate Clp chaperone's binding is not explicit. Previously, in one of the organisms with two ClpP isoforms, it was reported that CtrClpC exclusively interacts with its CtrClpP2 isoform but not CtrClpP1 (Pan et al., 2023). Thus, various protein-protein interaction studies were conducted to comprehend the interaction between LinClpC and the LinClpP isoforms. Based on previous reports on the association of Clp chaperones and proteases of *M. tuberculosis*, *C. trachomatis*, *P. aeruginosa,* and *C. difficile* (Pan et al., 2023; Hall et al., 2017; Lavey et al., 2018; Schmitz et al., 2014), it is hypothesized that LinClpC may interact with either LinClpP1 or LinClpP2 or both non-preferentially. In this study, using three independent techniques, namely, SEC, fluorescence resonance energy transfer (FRET), and enzyme-linked immunosorbent assay (ELISA), the interaction of LinClpC with its LinClpP isoforms was performed. Using SEC elution profile

TH-3792_206106022

analysis, the pure recombinant LinClpC (dimer; ~192 kDa), LinClpP1 (heptamer; ~161 kDa), and LinClpP2 (tetradecamer; ~308 kDa) demonstrated their existence in an oligomeric state (**Figure 3.4A** and **3.4B**). Additionally, the self-assembled protein complex (LinClpCP1) or (LinClpCP2) was independently loaded into the SEC column, and elution profiles were monitored by measuring the absorbance at 280 nm. One representative elute collected from different peaks of SEC was then analyzed on SDS-PAGE (12%). The SEC elution profile of LinClpCP1 complex displayed 3 major peaks (Peak 1, Peak 2, and Peak 3) (**Figure 3.4A**). On resolving specific elutes of these peaks on SDS-PAGE, bands corresponding to LinClpC (96 kDa) from Peak 1 (60 mL) and bands of LinClpP1 (23 kDa) from Peak 3 (88 mL) was visualized after Coomassie Blue staining (**Figure 3.4A**). While from the elutes of Peak 2 (75 mL) (LinClpCP1 complex), bands of both LinClpC and LinClpP1 could be resolved by SDS-PAGE on the same lane (96 and 23 kDa) (**Figure 3.4A**).

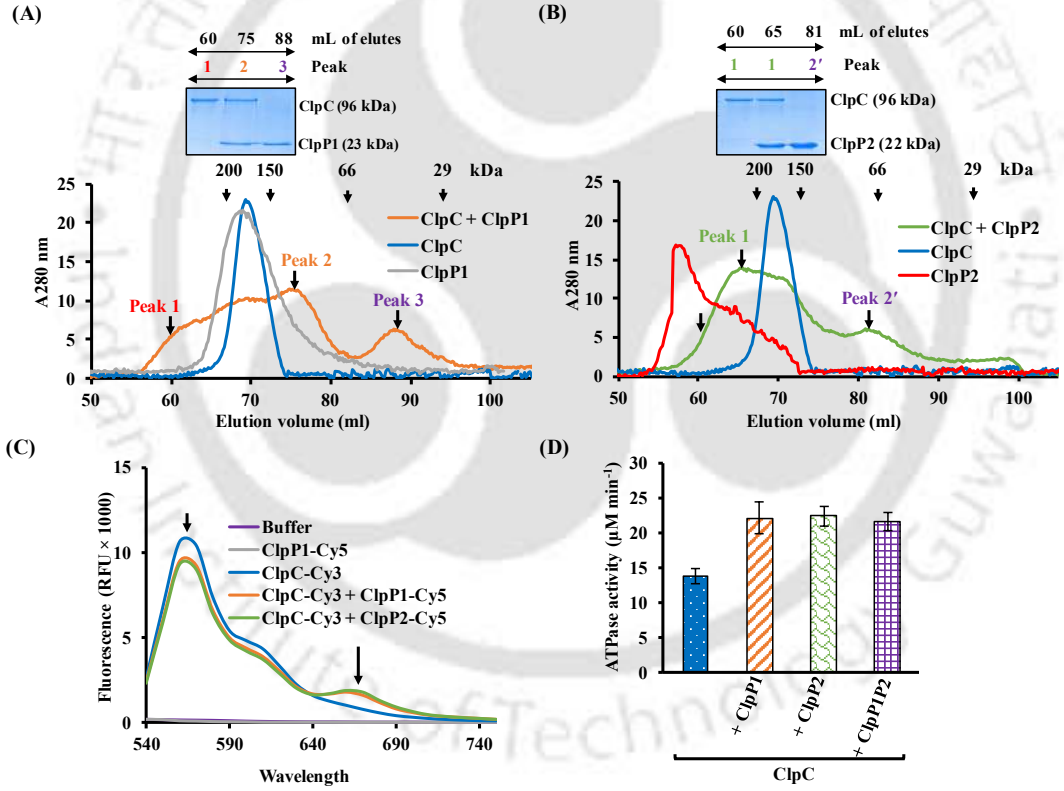


Figure 3.4. LinClpP isoforms associate with the ATPase chaperone, LinClpC. (A) SEC of pure LinClpP1, LinClpC, and complex LinClpCP1 was performed using 16/600 Superdex™ 200 pg column (GE Healthcare). The elution profile of the complex LinClpCP1 shows 3 different peaks and one elute of specific volume (indicated with black arrow mark) from each peak was analysed on 12% SDS PAGE (shown in upper panel). **(B)** SEC of pure LinClpP2, LinClpC, and complex LinClpCP2. The elution profile of LinClpCP2 complex shows a broad peak (Peak 1) and a shoulder peak (Peak 2′) and one elute of specific volume (indicated with black arrow mark) from each peak was analysed on 12% SDS PAGE (shown in upper panel). **(C)** FRET study between LinClpC-Cy3 donor and LinClpP1-Cy5 or LinClpP2-Cy5 acceptor. The fluorescence emission spectra was recorded at a fixed excitation wavelength of 488 nm and emission wavelength from 540-740 nm. The fluorescence emission spectra of LinClpC-Cy3 (blue line), LinClpC-Cy3 + LinClpP1-Cy5 (orange line), LinClpC-Cy3 + LinClpP2-Cy5 (green line), LinClpP1-

TH-3792_206106022

Cy5 (grey line), and buffer (violet line) are represented with their fluorescence intensity as RFU × 1000. Emission peaks for Cy3 (570 nm) and Cy5 (670 nm) are indicated by small and large arrows, respectively. **(D)** ATPase activity analysis of LinClpC in the presence (+) of equimolar concentration of pure LinClpP isoforms or heterocomplex. The data represent the Mean ± Standard error mean (SEM) from the three independent experiments performed in duplicates.

Similarly, the SEC elution profile for the LinClpCP2 complex displayed a broad peak (Peak 1) and another shoulder peak (Peak 2′) (**Figure 3.4B)**. On SDS-PAGE, the broad peak (Peak 1) of the LinClpCP2 complex indicated bands of pure LinClpC (96 kDa) (60 mL), followed by LinClpC-LinClpP2 together (96 kDa and 22 kDa) (65 mL) on a single lane. While the shoulder peak (Peak 2′; 81 mL) of the LinClpCP2 complex demonstrates the presence of only pure LinClpP2 (22 kDa). In our study, it was observed that when LinClpC and LinClpP isoforms are incubated together, they undergo a dynamic state of interaction where heptamers of LinClpP1 or tetradecamers of LinClpP2 tend to dissociate into monomeric or dimeric states, followed by their interaction with a monomeric or dimeric form of LinClpC. The co-elution of LinClpP1 or LinClpP2 with LinClpC in SEC reveals that the chaperone, LinClpC, can interact with each LinClpP isoform.

Next, a more sensitive fluorescence-based technique (FRET) was employed to study the LinClpC interaction with LinClpP isoforms. The emission wavelength scans (540–750 nm) of pure LinClpC-Cy3, or LinClpP1-Cy5, as well as co-incubated fluorescent partners, were measured at an excitation wavelength of 488 nm (**Figure 3.4C**). The excitation of pure LinClpP1-Cy5 or LinClpP2-Cy5 at 488 nm resulted in insignificant fluorescence emission across 540–750 nm, and was comparable to the control fluorescence. The excitation of LinClpC-Cy3 at 488 nm resulted in an anticipated emission peak at 570 nm (**Figure 3.4C**). Excitation of complexes (LinClpC-Cy3 + LinClpP1-Cy5 or LinClpC-Cy3 + LinClpP2-Cy5) resulted in emission peaks of Cy3 (570 nm) and Cy5 (670 nm) (**Figure 3.4C)**. It can be ascertained that the second peak at 670 nm is the outcome of energy transfer between the donor (LinClpC-Cy3) and acceptors (LinClpP1-Cy5 or LinClpP2-Cy5) fluorophores and thus, unequivocally justifies the LinClpC interaction with each LinClpP isoform. Additionally, to study the effect of ATP on the complexes (LinClpCP1 or LinClpCP2), FRET analysis was performed using ATP-bound LinClpC-Cy3 and LinClpP-Cy5 (**Figure 3S5a**). A similar level of energy transfer between free or ATP-bound LinClpC-Cy3 (donor) and LinClpP-Cy5 (acceptor) was observed.

Among other techniques, ELISA has also been widely used to measure protein-protein interaction (Murray et al., 2007; Dixit et al., 2021). In this study, a 96-well microtiter plate was

TH-3792_206106022

coated with LinClpP1 or LinClpP2 or BSA (1 μM) and was overlaid with increasing concentration (0-3 μM) of free or ATP-bound LinClpC (**Figure 3S5b** and **3S5c**). A linear increase in absorbance at 450 nm was observed for both complexes (LinClpCP1 or LinClpCP2) with an increase in LinClpC concentration that achieved a saturation point (1 μM). Immunoassay results suggest that LinClpC can interact to LinClpP1 or LinClpP2 with similar affinity. Additionally, the ATP-bound LinClpC shows a similar extent of interaction with LinClpP isoforms as compared to free LinClpC, which corroborates that the nucleotide-induced oligomerization has a non-significant impact on LinClpCP1 or LinClpCP2 complex formation.

As LinClpC associates with each LinClpP isoform (LinClpP1 or LinClpP2) with equal affinity, it was interesting to address whether each LinClpP isoform can modify the ATPase activity in LinClpC. Previous studies have shown that the interaction of EcoClpA with EcoClpP activates its ATPase activity in a dose-dependent manner (Bewley et al., 2009). Thus, in this study, the rate of ATP hydrolysis by LinClpC in the presence of either pure LinClpP1 or LinClpP2 and its heterocomplex was measured (**Figure 3.4D**). The ATPase activity of LinClpC in the presence of pure LinClpP isoforms or its heterocomplex was boosted 2-fold and agrees with EcoClpA activity (Bewley et al., 2009). However, in MtuClpC, only the heterocomplex MtuClpP1P2 bound with activator-dipeptide (Bz-LL) stimulated the ATPase activity by 2-fold (Li et al., 2016).

3.4.5. Association of LinClpC to LinClpP1P2 heterocomplex stimulates protein degradation

The LinClpC interaction with each LinClpP isoform encouraged us to probe the activity of complexes (LinClpCP1, LinClpCP2, or LinClpCP1P2) on the model protein substrate. Using the model-protein substrates, β-casein (S1) and FITC-casein (S2), LinClpPs protease activity in the presence of LinClpC was assessed. In protease activity, each self-assembled complex (3 μM) (LinClpCP1, LinClpCP2 and LinClpCP1P2) was tested on S1 (5 μM), with or without ATP supplementation. The proteolytic capacity of various complexes (LinClpCP1, LinClpCP2, LinClpCP1P2) was analysed by a time-dependent degradation of S1 on 12% SDS-PAGE (**Figure 3.5A**). Analysis of the result suggests that LinClpP1P2 exclusively in the presence of LinClpC can degrade S1 in a time-dependent manner. Neither LinClpP1 nor LinClpP2 showed proteolytic activity in the presence of LinClpC. Interestingly, the LinClpCP1P2 machinery could degrade S1 without the aid of ATP. This ATP-independent activity of the LinClpCP1P2 machinery on S1 did not agree with previously reported EcoClpP

TH-3792_206106022

activity (Thompson et al., 1994). Thus, another protease assay was performed using a fluorophore-tagged model substrate (FITC-casein; S2) to validate the unprecedented ATP-independent activity of the LinClpCP1P2 machinery in this study. The LinClpCP1P2 protease activity on S2 was measured in the presence and absence of ATP. Additionally, a non-hydrolysable ATP analogue, ATP$_\gamma$S, was added to comprehend the essentiality of energy or nucleotide. The degradation of S2 was monitored at different time points (15-minute intervals for 120 minutes) by measuring the fluorescence of the released FITC (**Figure 3S6a**).

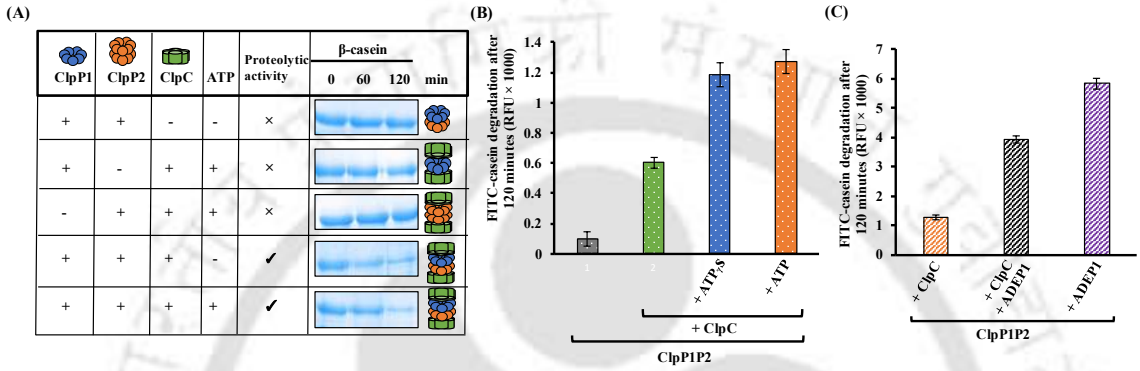


Figure 3.5. The association of LinClpC with heterocomplex LinClpP1P2 to form functional protease machinery. (A) Protease activity of pure LinClpP isoforms or heterocomplex in association with LinClpC was checked on model substrate (S1; β-casein) in the absence (-) or presence (+) of ATP (2 mM). The degradation of β-casein was examined on 12 % SDS-PAGE within a time interval of 0, 60, and 120 minutes. **(B)** Degradation of model substrate (S2; FITC-Casein) by LinClpP1P2 heterocomplex in the absence (-) and presence (+) of hydrolysable or non-hydrolysable ATP was examined by measuring the fluorescence intensity of released FITC at an excitation wavelength of 490 nm and emission wavelength of 525 nm after 120 minutes of proteolytic activity. **(C)** The degradation of FITC-casein by LinClpP1P2 heterocomplex in the presence (+) of either LinClpC or ADEP1 or both (LinClpC + ADEP1) was measured after 120 minutes. The data represent the Mean ± Standard error mean (SEM) from the three independent experiments performed in duplicates.

It was observed that the LinClpC-associated LinClpP1P2 heterocomplex shows protease activity, while the LinClpP1P2 heterocomplex alone remained inactive. The supplementation of either ATP or ATP$_\gamma$S provides an additional 0.5-fold boost in the LinClpCP1P2 proteolytic activity after 120 minutes of degradation (**Figure 3.5B**). This infers that hydrolysis of ATP by LinClpC is dispensable to stimulate the protease activity of the LinClpCP1P2 machinery on an unstructured model substrate (casein). Additionally, degradation of folded model substrate (eGFP) by LinClpCP1P2 machinery in the presence and absence of ATP was checked (**Figure 3S6b**). The LinClpCP1P2 machinery does not show ATP-independent or dependent proteolytic activity on eGFP substrate. Thus, our study signifies that ATP-independent degradation by LinClpCP1P2 complex was restricted to the unstructured model substrate, casein only. We

TH-3792_206106022

believe that the binding of nucleotide (ATP or ATP$_\gamma$S) to LinClpC nucleotide-binding domains promotes its oligomerization and thereby fosters LinClpC association with the heptameric rings of LinClpP isoforms. The stabilization of the LinClpCP1P2 machinery in the presence of nucleotide (ATP or ATP$_\gamma$S) might provide further stimulation in its protease activity.

3.4.6. Differential association of LinClpP isoforms with LinClpC and antibiotic ADEP1

Previous studies on the effect of antibiotic ADEP1 on LinClpP isoforms proposed an additional allosteric mode of regulation of LinClpP1P2 protease activity (Dhara et al., 2021). Apart from ADEP1 binding to the conserved hydrophobic cluster of LinClpP heterocomplex, it also interacts allosterically along the catalytic Ser98 residue of LinClpP1 subunits. The ADEP1, upon interaction with Ser98 of LinClpP1 subunits, modulated the required alignment of catalytic triads within the mutant heterocomplex LinClpP1P2^{S97A} but not the LinClpP1^{S98A}P2 (Dhara et al., 2021). However, LinClpX, an ATPase chaperone, did not display an allosteric mode of activation on LinClpP1P2^{S97A} (Dhara et al., 2019). To further understand the ADEP1-mediated ClpP activation, a study on the LinClpP heterocomplex in the presence of ADEP1 and LinClpC was conducted. The model protein substrate (FITC-casein) degradation was measured fluorometrically in the presence of equimolar concentration (1 µM) of LinClpC or ADEP1 or both at different time points for an interval of 120 minutes (**Figure 3S6c**). The ADEP1-bound LinClpP heterocomplex displayed approximately 4-fold higher protease activity compared to LinClpCP1P2 machinery after 120 minutes of degradation (**Figure 3.5C**). Supplementation of ADEP1 to the LinClpCP1P2 machinery, resulted in a 2.8-fold increment in its protease activity. (**Figure 3.5C**). Thus, from this study, it can be ascertained that among the two activators (ADEP1 and LinClpC) of LinClpP1P2 protease activity, ADEP1 outcompetes LinClpC.

3.4.7. Effect of LinTF and its variants on LinClpC and ADEP1-bound LinClpP1P2 activity

Chaudhary et al., 2021 delineated the domain architecture of *Leptospira interrogans* Trigger Factor (LinTF), categorizing the N-terminal domain within residues 1–144, the central PPIase domain between residues 145–219, and the C-terminal domain spanning residues 220–451. The C-terminal portion was further subdivided into two structural arms: arm-1 (residues 250–350) and arm-2 (residues 351–451). Following earlier evidence from our laboratory demonstrating the stimulatory effect of recombinant LinTF (LinTF) on the proteolytic activity of both LinClpP1P2 and LinClpXP1P2 complexes, we proceeded to analyze the contributions

TH-3792_206106022

of LinTF individual domains. Guided by its defined domain structure, we designed three deletion constructs: LinTF$^{\Delta pCTD}$, which lacks residues 351–451 (corresponding to the C-terminal arm-2); LinTF$^{\Delta CTD}$, omitting the full C-terminal domain (residues 220–451); and LinTF$^{\Delta NTD}$, lacking the N-terminal region (residues 1–144).

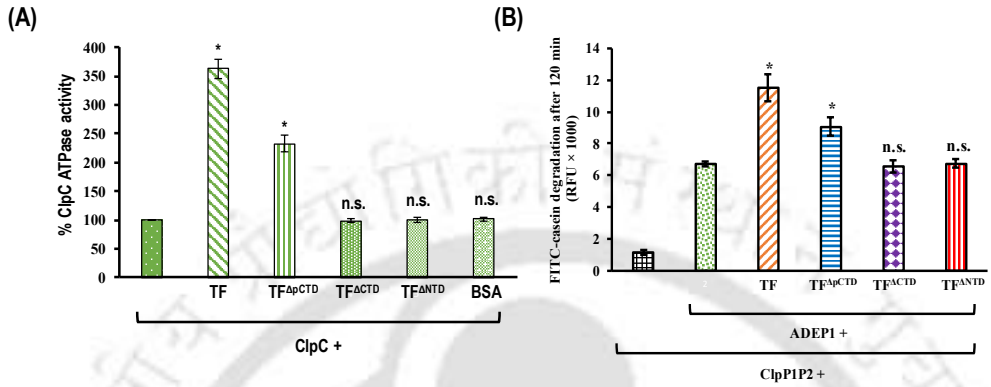


Figure 3.6. Evaluation of role of LinTF and its variants on LinClp activity (A) ATPase activity of LinClpC was checked in the presence ofLinTF and variants (LinTF$^{\Delta pCTD}$, LinTF$^{\Delta CTD}$ and LinTF$^{\Delta NTD}$). The LinClpC activity was considered 100% and BSA was used a control. **(B)** Degradation of FITC-Casein by ADEP1-bound LinClpP1P2 heterocomplex in the presence of LinTF and variants was examined by measuring the fluorescence intensity of released FITC at an excitation wavelength of 490 nm and emission wavelength of 525 nm after 120 minutes of proteolytic activity.

Prior findings revealed that truncation of the C-terminal arm-2 in LinTF$^{\Delta pCTD}$ led to diminished enhancement of ClpP1P2 activity, suggesting this region plays a key role in protease complex stabilization or assembly.

The current study further investigates the functional significance of the LinTF N-terminal and complete C-terminal domains. The two additional truncated forms, LinTF$^{\Delta CTD}$ and rLinTF$^{\Delta NTD}$, were amplified via PCR using *L. interrogans* serovar Copenhageni genomic DNA as the template, cloned into the pET-23a expression system, heterologously expressed in *E. coli* BL21 (DE3), and purified through Ni-NTA affinity chromatography for downstream functional assays. To further dissect the functional roles of individual LinTF domains, we assessed their impact on the ATPase activity of LinClpC and the proteolytic activity of the ADEP1-activated LinClpP1P2 complex. The ATPase activity assays were conducted in the presence or absence of full-length LinTF and its deletion variants (LinTF$^{\Delta pCTD}$, LinTF$^{\Delta CTD}$, and LinTF$^{\Delta NTD}$), using the intrinsic ATPase activity of LinClpC (considered 100%). Addition of full-length LinTF resulted in a three-fold increase in LinClpC ATPase activity, while LinTF$^{\Delta pCTD}$ caused a moderate two-fold stimulation. In contrast, neither LinTF$^{\Delta CTD}$ nor rLinTF$^{\Delta NTD}$ elicited any measurable enhancement, indicating that both the N-terminal and full C-terminal regions are

TH-3792_206106022

essential for LinTF-mediated stimulation of LinClpC ATPase activity. In parallel, we evaluated the influence of these LinTF variants on the proteolytic function of the ADEP1-bound LinClpP1P2 heterocomplex, using FITC-casein as a model substrate. After a 2-hour incubation, the protease activity of the ADEP1-activated complex alone was set to 100% as a reference. The addition of full-length LinTF led to a 1.7-fold increase in substrate degradation, while LinTF$^{\Delta pCTD}$ showed a modest enhancement of 1.2-fold. However, the presence of either LinTF$^{\Delta CTD}$ or LinTF$^{\Delta NTD}$ did not significantly alter the proteolytic activity, reinforcing the critical role of both terminal domains in modulating ClpP1P2 function. These results collectively highlight the domain-dependent contribution of LinTF in regulating the Clp protease machinery of *L. interrogans*. The failure of these variants in stimulating LinClpC ATPase activity or ADEP1-bound LinClpP1P2 heterocomplex protease activity suggests that both the N-terminal domain, likely involved in ribosome binding, and the C-terminal domain, associated with chaperone function, are critical for LinTF-mediated activation of the protease complex. Their deletion appears to impair the formation of the substrate-binding hydrophobic pocket, thereby compromising the structural integrity and functional capacity of LinTF.

3.5. Discussion

Classic approaches in antimicrobial therapy are based on targeting cellular processes like DNA replication, translation, and cell wall formation, which are vital for actively dividing cells (Culp et al., 2017). Although conventional antibiotics have saved millions of lives, excessive overuse and inappropriate prescription have led to the emergence of antibiotic-resistant bacteria (Culp et al., 2017; Kahrstrom et al., 2013). In addition, some pathogenic bacteria become persistent in the host and enter an antibiotic-tolerant life cycle. These non-replicative and dormant persistent cells in the host are thus challenging for antibiotic treatment (Balaban et al., 2019). Thus, to combat the rising emergence of antibiotic-resistant and persistent bacteria, there is a need for new alternative targets in antimicrobial therapy. In this context, caseinolytic proteases and chaperones have garnered considerable attention as targets for antibacterial action after their direct relationship with bacterial viability and virulence was proven in Gram-positive and Gram-negative bacteria (Culp et al., 2017; Mei et al., 2003; Gaillot et al., 2000; Sassetti et al., 2003; Bhandari et al., 2018). Recently, some antibiotic compounds were reported to target the protein quality control machinery either by activating or inhibiting the ClpP to reduce infection caused by *S. aureus*, *M. tuberculosis*, and *S. hawaiiensis* (Conlon et al., 2013; Famulla et al., 2016; Thomy et al., 2019). Additionally, the cyclic peptides ecumicin and lassomycin were discovered to increase the activity of caseinolytic chaperone, ClpC1, in *M. tuberculosis* and

TH-3792_206106022

thereby uncouple it from the ClpP complex, potentially resulting in the accumulation of unwanted proteins (Gao et al., 2015; Gavrish et al., 2014). Similarly, CyclomarinA also targets MtuClpC1, causing it to become more active (Schmitt et al., 2011). These Clp chaperone-targeting peptide derivatives were reported to be bactericidal against replicative multi-drug resistant and dormant *M. tuberculosis* (Schmitt et al., 2011). However, the bactericidal effects of such peptide derivatives have not been reported to date for bacterial species other than actinobacteria.

The pathogenic *L. interrogans* is the causative agent of leptospirosis, an emerging disease with more than 1 million severe cases worldwide and 60,000 deaths yearly (Costa et al., 2015). The complete genome sequence analysis of *L. interrogans* shows the existence of an array of genes belonging to the Clp system (Laurdault et al., 2011). The core proteolytic component of the Clp system in leptospires is composed of two different ClpP isoforms (ClpP1 and ClpP2), forming an active hetero-tetradecameric assembly (LinClpP1P2) (Dhara et al., 2019). Previously, by targeting LinClpPs, the impact of antibiotic ADEP1 on the growth impairment of leptospires was reported (Dhara et al., 2021). Based on *in-vitro* biochemical studies and *in-vivo* growth inhibition assay, ADEP1 was proposed as an alternative to conventional antibiotics (Dhara et al., 2021). Taken together, these pieces of evidences from previous studies indicate that targeting the leptospiral Clp system could be a novel strategy for leptospirosis treatment. Thus, a better and elaborate understanding of individual Clp components of *Leptospira* will pave the way for several aspects of new antibacterial targets.

The *L. interrogans* possesses an ATPase chaperone, LinClpC, which is often present in gram-positive bacteria. The LinClpC showed high sequence similarity with various ClpC orthologs (SauClpC, BsuClpC, SelClpC, and LmoClpC) within the ATPase binding and hydrolysis domains. In this study, LinClpC showed a dose-dependent intrinsic ATPase and followed Michaelis-Menten kinetics. The calculated half-maximal ATP concentration (K_M) for LinClpC ATPase activity was in a range (0.5 - 6 mM) similar to that of MtuClpC1 (Kar et al., 2008), CcrClpB *of Corynebacterium crenatum* (Huang et al., 2020[a]), CcrClpX (Huang et al., 2020[b], SelClpC (Andersson et al., 2006) and LinClpX (Dhara et al., 2019). Interestingly, in this study, unlike previously reported AAA+ ATPase activity of EcoClpA (Seol et al., 1994), LinClpB (Krajewska et al., 2017), and LinClpX (Dhara et al., 2019), the β-casein substrate could not stimulate the LinClpC ATPase activity. Similar to LinClpC, the ATPase activity of selective Clp chaperones (BsuClpC and SauClpC) in the presence of casein substrate also was undeterred (Trentini et al., 2016; Carroni et al., 2017). In a recent interesting study, a phosphoarginine (pArg) degradation tag was reported to be essential for BsuClpC substrate (casein-pArg)

TH-3792_206106022

recognition and to permit BsuClpP protease activity. Also, the ATPase activity of BsuClpC was significantly increased in the presence of casein-pArg but not by natural casein (Trentini et al., 2016). A similar degradation of casein-pArg substrate was reported for CtrClpCP1P2 protease (Pan et al., 2023). Likewise, a phosphorylation tag in the β-casein substrate may stimulate LinClpC ATPase activity and is warranted for future study.

Further, the LinClpC nucleotide phosphorylase (NTPase) activity was not limited to conventional ATP (100%), but also holds the potential to hydrolyse other nucleotides, including GTP (60%), UTP (25%), and CTP (25%). On the same line, EcoClpB (Woo et al., 1992) and *Streptococcus pneumoniae* SpnClpL (Park et al., 2015) also hold the ability to hydrolyse different types of nucleotides (GTP and CTP). However, the reason for NTPase activity on diverse nucleotides in specific bacterial ClpP chaperones is not clear.

In this study, the optimum temperature for LinClpC ATPase activity was above 35 °C, which led us to speculate its role in thermo-tolerance or resisting heat shock. Similarly, CcrClpB (Huang[a] et al., 2020) and MtuClpC1 (Kar et al., 2008) optimum ATPase activity was reported beyond 37 °C and thus suggested these chaperone's role in heat stress conditions. In agreement, BsuClpC (Kruger et al., 1994), SauClpC (Chatterjee et al., 2005), and CcrClpB (Huang[a] et al., 2020) are reported to play an important role in developing thermo-tolerance. In another study on *Leptospira*, a 3-fold increase in the transcription level of the *LinclpB* was reported upon exposure to 42 °C (heat shock) (Laurdault et al.,2011). However, a more comprehensive *in-vivo* analysis of the *clpC* gene in *L. interrogans* may reveal a correlation with heat stress tolerance.

Next, the assembly of LinClpC in the presence of nucleotide was assessed by multiple independent techniques, including size exclusion chromatography (SEC), ANS binding assay, and dynamic light scattering. SEC analysis of LinClpC suggested its existence in a dimeric form (192 kDa) in solution, which eventually can self-assemble into higher molecular weight complexes in the presence of ATP. A similar ATP-dependent oligomerization via SEC was reported for MtuClpC1 (Bajaj et al., 2012) and SmuClpL (Jana et al., 2020). Further, we have also demonstrated the nucleotide-induced oligomerization of LinClpC by ANS binding assay and DLS analysis. However, on SEC, the SauClpC shows a broad elution profile with large, non-hexameric assembly in the presence or absence of ATP (Carroni et al., 2017). Interestingly, the addition of adaptor protein, SauMecA and ATP shifts the elution profile of the SauClpC towards hexameric form (Carroni et al., 2017). The existence of large assemblies of SauClpC in the absence of ATP depends on the Phe436 residue of the middle domain (MD) as substitution with Ala ($SauClpC^{F436A}$) shifts the higher oligomeric form towards dimeric size

TH-3792_206106022

(Carroni et al., 2017). However, in this study, a significant sequence disparity within the MD of LinClpC was observed compared to its well-studied orthologs (MtuClpC, BsuClpC, and SauClpC). The Tyr446 residue has been naturally evolved in the MD of LinClpC (F446Y) as well as CtrClpC (F462Y) in place of the conventional Phe residue in its orthologs. Thus, such an evolutionary substitution of Phe with Tyr within the MD of LinClpC might restrict its existence in larger, higher-order complexes above 576 kDa.

The AAA+ chaperone interaction with Clp protease (ClpP) involves an association of hexameric chaperones with heptameric ClpP rings, which is conserved among diverse bacterial groups (Olivares et al., 2016). The EcoClpX hexamer and EcoClpP heptamer contacts are primarily mediated by the distal IGF/L loops of EcoClpX and the apical hydrophobic clusters of EcoClpP (Amor et al., 2019). The MSA affirmed the existence of semi-conserved IGF/L loops in the C-terminal domain of different ClpC orthologs, including LinClpC. The LinClpC possesses an LGF loop at residues 681-883 which might be involved in docking at the hydrophobic cluster of either one or both LinClpP isoforms. We thus examined the docking extent of the LGF loop in LinClpC with the hydrophobic cluster present in LinClpP isoforms (LinClpP1 and LinClpP2) using SEC, FRET analysis, and ELISA. Our SEC results signified that the LinClpC was co-eluted with each LinClpP isoform. However, for both complexes (LinClpCP1 and LinClpCP2), the elution profile demonstrated more than one peak suggesting that all the LinClpC molecules are not being able to dock its LGF loop with LinClpP isoforms. Also, in the presence of LinClpP, the elution profile of LinClpC switched to high molecular weight complex rather than its dimeric size (192 kDa). Thus, it was speculated that each LinClpP isoform promotes the LinClpC oligomerization to acquire its functional form. Likewise, the FRET analysis and ELISA experiments in this study proclaimed that each LinClpP isoforms have the same affinity with the chaperone, LinClpC. Additionally, the interaction study via FRET and ELISA using free or ATP-bound LinClpC demonstrates that the dimeric or higher oligomeric form of LinClpC interacts with LinClpP isoforms in a similar manner. Our study signifies that oligomerization of LinClpC in the presence of ATP has no influence on the formation of complexes (LinClpCP1 or LinClpP2). However, among bacteria harboring two ClpP isoforms, the interacting ClpP partner for cognate ATPase varies significantly (Leodolter et al., 2015; Gatsogiannis et al., 2019; Pan et al., 2019; Hall et al., 2017). The ClpP2 isoform exclusively serves as the docking site for the semi-conserved IGF/L loops of ATPase chaperone in the case of hetero-tetradecameric MtuClpP1P2 (Leodolter et al., 2015), LmoClpP1P2 (Gatsogiannis et al., 2019), and CtrClpP1P2 (Pan et al., 2019). Meanwhile, PauClpP1 exclusively interacts with its cognate ATPase IGF/L loops in the case

TH-3792_206106022

of PauClpP1P2 (Hall et al., 2017). Therefore, our results, combined with previous reports on Clp chaperone-protease interaction (Leodolter et al., 2015; Gatsogiannis et al., 2019; Pan et al., 2019; Hall et al., 2017), suggested that the bacteria with multiple ClpP isoforms evinced disparity in association with the cognate ATPase chaperones.

Next, we addressed whether the complexes (LinClpCP1 and LinClpCP2) are functionally active in degrading the model protein substrate (β-casein). Neither LinClpCP1 nor LinClpCP2 complex showed any functional proteolytic activity despite having the ability to interact with its chaperone. To gain protease activity, an extended active state conformation of LinClpP1P2 is mandatory (Dhara et al., 2021). Thus, LinClpC failed to alter the conformational state of each LinClpP isoform. In contrast, LinClpC interaction with the LinClpP1P2 heterocomplex achieved protease activity. A similar functional interdependence of both ClpP isoforms (ClpP1 and ClpP2) in association with AAA+ chaperone to form an active protease hetero-tetradecamer was reported earlier for CtrClpP1P2 and MtuClpP1P2 (Pan et al., 2019; Schmitz et al., 2014).

For a well-studied bacterial protease, the ATPase component is essential for the processive degradation of protein substrates by ClpP (Thompson et al., 1994). However, in this study, the proteolytic activity of the LinClpCP1P2 machinery was ATP-independent on the casein substrate. Supplementation of ATP or its non-hydrolysable analogue stimulated LinClpCP1P2 machinery for the protease activity. The nucleotide-induced oligomerization of LinClpC might stimulate the association of functional LinClpC with LinClpP isoforms and thereby increase the activity. The three major categories of substrates for EcoClpP are short peptides (2-5 amino acids long), longer peptides (10-30 amino acids long), and larger protein substrates (Thompson et al., 1994). Short peptides can be cleaved by EcoClpP alone, while to cleave longer peptide substrates, the EcoClpP requires association with the ATPase chaperone in the presence of a non-hydrolysable nucleotide (Thompson et al., 1994). The EcoClpP can degrade protein substrates exclusively in the presence of chaperone EcoClpA in an energy-dependent manner. It has been argued that the nucleotide promotes the interaction of EcoClpA with EcoClpP and induces conformational changes within the catalytic triad of EcoClpP and thus boosting its catalytic efficiency (Thompson et al., 1994). Later, in another independent study, processive degradation of casein was reported by EcoClpP alone, albeit at a much lower rate (Jennings et al., 2008). A similar protease activity without the aid of any ATPase chaperone on the FITC-casein substrate was observed for CtrClpP1P2 (Pan et al., 2019). However, a truncated smaller N-terminal loop of CtrClpP1 is proposed to provide a wider entrance pore for substrates within the CtrClpP1P2 catalytic chamber (Pan et al., 2019). The EcoClpAP cryo-EM structure

TH-3792_206106022

suggests that EcoClpA binds at a peripheral site on the EcoClpP apical surface through its IGF/L loops and elicits a conformational change within the axial region of the EcoClpP ring (Effantin et al., 2010). The EcoClpP, upon binding with IGF/L loops of EcoClpA, its N-terminal flexible loops are displaced from the axial surface and thereby its axial pore gets opened (Effantin et al., 2010). The EcoClpP axial region appeared to be solidly blocked at the free end, while the EcoClpA bound end represents an axial channel of 12 Å (Effantin et al., 2010). Additionally, the crystal structure of ADEP-bound EcoClpP suggested that ADEP binds to the hydrophobic pockets present near the outer edge of EcoClpP rings (Li et al., 2010). These EcoClpP hydrophobic clusters coincide with the EcoClpA binding pockets of IGF/L loops. In the ADEP-bound EcoClpP structure, the N-terminal regions are locked in an extended state from the axial surface of EcoClpP, creating an axial pore of 20 Å diameter. Such locked conformation of the EcoClpP N-terminal region with wider axial pore diameter allows easy access to large substrates and, thus, higher protease activity (Li et al., 2010). Thus, in the case of *E. coli*, the cryo-EM structure of ClpP bound to ClpA and the X-ray structure of ADEP-bound ClpP revealed that the diameter of the axial channel of the ADEP-bound ClpP structure was wider (20 Å) compared to the ClpA-bound ClpP (12 Å) (Alexopoulos et al., 2012). Similarly, in this study, the ADEP-bound LinClpP1P2 heterocomplex displayed approximately 4-fold higher protease activity than the LinClpCP1P2 complex. Thus, from this study, it can be conceived that among the two activators (ADEP1 and LinClpC) of LinClpP isoforms, ADEP1 activates the LinClpP1P2 protease machinery more efficiently than LinClpC. We also speculate that, like EcoClpAP, the association of LinClpC with LinClpP1P2 heterocomplex might allow conformational switching of LinClpP N-terminal axial loops and form an open axial pore state to allow entry of protein substrates for degradation.

In summary, this study advanced the understanding of the LinClpP system in a pathogenic spirochete, *Leptospira interrogans*. The LinClpC demonstrated nucleotide-induced oligomerization and also exhibited intrinsic ATPase activity. The LinClpC can associate with both the pure LinClpP isoforms; however, such association (LinClpCP1 and LinClpCP2) failed to impart functional activity. In contrast, the LinClpP1P2 heterocomplex, in association with LinClpC, exhibited protease activity.

TH-3792_206106022

Chapter 4

Modulation of ClpA chaperone activity by Clp adaptor proteins and fate of SsrA-tagged substrates of *Leptospira interrogans*

This chapter is a modified version of the manuscript published as:

Kumari, S., & Kumar, M. (2025). The ClpA chaperone and the two adaptor proteins modulate the fate of the model substrate tagged with a SsrA-degron of Leptospira. ***Biochemical Journal***, 482(17), 1253-1275.

TH-3792_206106022

4.1. Abstract

Caseinolytic protease (Clp) complexes play a vital role in bacterial protein quality control by targeting misfolded or aggregated proteins for degradation. In the pathogenic spirochete *Leptospira interrogans*, Clp system includes several adaptor proteins (ClpS1 and ClpS2) and chaperones (ClpX, ClpA, and ClpC). These components are thought to function alongside two ClpP isoforms, ClpP1 and ClpP2, to mediate targeted proteolysis. This study investigates the molecular and biochemical characteristics of three *L. interrogans*-derived Clp proteins: LinClpA, LinClpS1, and LinClpS2. LinClpA, a 740-amino acid chaperone, contains a regulatory N-terminal domain and two conserved AAA+ ATPase modules (D-I and D-II), essential for nucleotide binding and hydrolysis. The LinClpS1 and LinClpS2, while structurally similar, differ in their binding pockets for N-degron-tagged proteins. Functional assays demonstrate that a variant of LinClpA lacking the N-terminal domain (LinClpA$^{\Delta N}$) still undergoes nucleotide-induced oligomerization, akin to the full-length protein, but shows increased ATPase activity. Interaction analyses revealed that LinClpA's ATPase activity is upregulated in the presence of LinClpP1 and LinClpP2, while inhibited when bound by LinClpS1 or LinClpS2. Interestingly, LinClpA$^{\Delta N}$ does not respond to these adaptor proteins, emphasizing the importance of the N-terminal domain of LinClpA in adaptor proteins engagement. Furthermore, this study explores the functional significance of the SsrA tag (a C-terminal degradation signal; C-degron) in promoting substrate recognition and breakdown by the LinClpAP1P2 proteolytic machinery.

TH-3792_206106022

4.2. Introduction

Targeted proteolysis is vital in regulating cellular processes, maintaining proteome balance, and ensuring protein quality control across all cells (Gottesman, 1996). In bacteria, these functions are facilitated by energy-dependent ATPase chaperone-protease complexes, which utilize energy from ATP hydrolysis to unfold and transport protein substrates into the proteolytic chamber of proteases (Hanson & whiteheart, 2005; Gottesman, 2003; Sauer et al., 2004). Several proteases, including members of the Clp protease family (ClpAP, ClpXP, and ClpYQ), as well as FtsH and Lon, are essential for sustaining protein homeostasis (Mahmoud & Chien, 2018; Olivares et al., 2016). The Lon and FtsH proteases comprise a single polypeptide chain encompassing both ATPase and protease domains. In contrast, the Clp proteases consist of distinct polypeptide chains for their ATPase components (ClpA, ClpX, or ClpY) and protease components (ClpP and ClpQ) (Lee & Suzuki, 2008; Langklotz et al., 2012; Baker & Sauer, 2012; Rohrwild 1996; Maurizi et al., 1998).

The tetra-decameric complex EcoClpP, a core proteolytic component of *Escherichia coli*, demonstrates a unique capability to interact with two ATPase chaperone partners, namely EcoClpX and EcoClpA (Fujimura et al., 1987). Upon the binding of nucleotides, EcoClpA assembles into an active hexameric chaperone, thereby facilitating the translocation and unfolding of protein substrates through its translocation channel into the catalytic chamber of EcoClpP (Kress et al., 2007; Baker & Sauer, 2006). This process is driven by the energy released during ATP hydrolysis by EcoClpA's two ATPase domains. The degradation of proteins necessitates a highly specific recognition mechanism to safeguard essential and properly folded proteins within the bacteria (Baker & Sauer, 2006). Certain protein substrates are directly identified by the regulatory components ClpA and ClpX associated with the ATP-dependent EcoClpAP and EcoClpXP proteases (Baker & Sauer, 2006). In contrast, other substrates are initially recognized by specialized adaptor proteins that subsequently facilitate their delivery to the respective proteases. Proteins designated for degradation typically possess specific motifs known as degrons, which are located at either the carboxy- or amino-terminal ends (Gottesman et al., 1998). The carboxy-terminal tagging of substrate proteins represents the most thoroughly studied C-degrons in *E. coli* (Gottesman et al., 1998). During protein synthesis, premature mRNA, which lacks a stop codon, results in stalled ribosomes. Trans-translation is required to rescue the stalled ribosomes and tag the nascent polypeptide chain with a specific C-degron for its degradation. In *E. coli*, a *s*mall *s*table *R*NA *A* (SsrA), also known as *t*ransfer *m*essenger RNA (tmRNA), works with the EcoSmpB protein to rescue stalled ribosomes. This process restores translation, followed by the release of the ribosome

TH-3792_206106022

and the production of proteins with a C-terminal SsrA tag (Gottesman et al., 1998). The EcoClpAP and EcoClpXP complexes and two distinct adaptor proteins exhibit varying affinities for SsrA-tagged substrates. The adaptor protein EcoSspB enhances the degradation of SsrA-tagged substrates by facilitating their delivery to EcoClpXP (Dougan et al., 2003). Conversely, the adaptor protein EcoClpS mitigates the binding and subsequent degradation of SsrA-tagged proteins by the EcoClpAP degradation complex (Dougan et al., 2002). EcoClpS interacts with the N-domain of EcoClpA, thereby steering the specificity of the EcoClpAP complex toward N-degron-tagged substrates and away from those categorized as C-degron substrates (Dougan et al., 2002; Erbse et al., 2006). In *E. coli*, destabilizing N-terminal residues, referred to as N-degrons, include phenylalanine (F), tyrosine (Y), tryptophan (W), and leucine (L) (Erbse et al., 2006). The primary destabilizing residues (F/L) are conjugated to a target protein substrate upon recognition of secondary destabilizing residues (Lysine; K or Arginine; R) by a specific F/L transferase (Erbse et al., 2006).

The EcoClpA belongs to the ATPase associated with diverse cellular activities (AAA+) superfamily of proteins, and its subunit comprises an N-terminal domain (N-domain) and two ATPase domains, D1 and D2. The N-domain of EcoClpA consists entirely of α-helices, exhibiting a 2-fold symmetry between helices H1-H4 and H5-H8. The D1 and D2 domains each include a large sub-domain with Walker A and B motifs alongside a smaller domain (Guo et al., 2002). Within the EcoClpA hexamer, the D1 and D2 domains are arranged as homomeric rings stacked vertically. The D2 ring of EcoClpA connects to the EcoClpP component, while the D1 domain is linked to the N-domain through a flexible linker (Guo et al., 2002). The N-domain can adopt multiple conformations above the D1 ring, functioning as a mobile element (Ishikawa et al., 2004). When the substrate-bound EcoClpS interacts with the highly mobile N-domain of EcoClpA, it facilitates the transfer of the substrate to the ATPase domain (Hernandez et al., 2011). EcoClpS is a small protein characterized by an unstructured N-terminal extension (NTE; amino acids 1-25) and a folded core domain (amino acids 26-106) (Zeth et al., 2002). The monomeric form of EcoClpS contains a single binding site for peptides that include phenylalanine (F), tyrosine (Y), tryptophan (W), or leucine (L) at the N-terminus. The hydrophobic side chains of the N-degron peptide bind to the hydrophobic pockets within the core region of EcoClpS (Schuenemann et al., 2009). The interface between the core region of EcoClpS and the N-domain of EcoClpA is primarily characterized by salt bridges and hydrogen bonds (Zeth et al., 2002). Upon engagement with the N-domain of EcoClpA, the exposed NTE of EcoClpS interacts with the pore loops present within the ATPase domains of EcoClpA, thereby facilitating the delivery of substrates to the protease complex (Hernandez et al., 2011).

TH-3792_206106022

The Clp family of proteins comprises several ATPase chaperones and their associated serine protease, ClpP. Based on the number of ATPase domains, these chaperones are divided into class I (ClpA and ClpC) with two ATPase domains and class II (ClpX) with one ATPase domain (Schirmer et al., 1996). The proteolytic complex, ClpXP, is present in most bacteria and has been studied in detail in gram-positive and gram-negative bacteria (Wojkowiak et al., 1993; Li et al., 2016; Griffith & Grossman, 2008; Kirsch et al., 2020; Pan et al., 2019; Lavey et al., 2018). The other chaperone, ClpA, is present in Gram-negative bacteria, while ClpC is found in Gram-positive bacteria and cyanobacteria (Queralto et al., 2023).

The ClpCP-mediated protein degradation has been studied in various bacteria, including *Bacillus subtilis*, *Streptomyces hawaiiensis*, *Chlamydia trachomatis*, *Mycobacterium tuberculosis*, and *L. interrogans* (Mei et al., 2009; Xu et al., 2024; Pan et al., 2023; Leodolter et al., 2015; Kumari et al., 2024). However, a detailed *in-vitro* biochemical characterization of ClpA and its association with Clp protease and adaptor proteins has only been performed in *E. coli* (Maurizi et al., 1998; Dougan et al., 2002; Erbse et al., 2006; Ishikawa et al., 2004; Hernandez et al., 2011). Several *in-vivo* studies have been reported regarding ClpA chaperone's functionality in gram-negative bacteria (Loughlin et al., 2009; Ekaza et al., 2000; Liu et al., 2016; Lo et al., 2022). For instance, in *Helicobacter pylori*, HpyClpA plays a role in providing resistance to oxidative and antibiotic stress, while in *Brucella suis*, BruClpA was reported to be involved in regulating thermal stress (Loughlin et al., 2009; Ekaza et al., 2000). Another study on *Caulobacter crescentus* highlights the importance of CcrClpA in controlling chromosome number and content (Liu et al., 2016). In *Xanthomonas campestris*, XcamClpA plays a role in pathogenicity and thermal stress tolerance (Lo et al., 2022). In this study, ClpA of *Leptospira*, a spirochete that causes leptospirosis disease in humans and animals (Mendu et al., 2025) has been explored. The *Leptospira* genome is predicted to have ten genes that encode for Clp family proteins (Lourdault et al., 2011). These genes include two adaptor protein-encoding genes (*clpS1; LIC11356* and *clpS2; LIC11815*), five genes coding for AAA+ ATPase (*clpA; LIC11814, clpB; LIC12017, clpC; LIC10339, clpX; LIC11418,* and *clpY; LIC11601*), and three genes coding for Clp protease (*clpP1; LIC11417, clpP2; LIC11951,* and *clpQ; LIC11600*) (Lourdault et al., 2011). In a prior study, the *clpB* (*LIC12017)* was essential for the survival of *Leptospira* under thermal and oxidative stress conditions (Lourdault et al., 2011). Biochemical analysis of LinClpB (LIC12017) has shown that nucleotide induces the formation of LinClpB hexamers with ATPase and disaggregase activity (Krajewska et al., 2017). Additionally, the core catalytic components of the Clp system in *Leptospira* (LinClpP1 and LinClpP2) have been well studied and characterized, revealing that the pure LinClpP1 and

TH-3792_206106022

LinClpP2 isoforms are catalytically inactive, while the heterocomplex, LinClpP1P2, shows peptidase activity on the di-peptide model substrate (Dhara et al., 2019). The modulating role of the natural antibiotic acyldepsipeptide (ADEP1) on the LinClpP1P2 protease activity has also been recently studied (Dhara et al., 2021; Kumari et al., 2024). Furthermore, a comprehensive study on LinClpC, a Class-I chaperone, has revealed that it has two ATPase domains separated by a middle domain and undergoes nucleotide-induced oligomerization, and the LinClpC oligomers interact with both LinClpP isoforms (LinClpP1 and LinClpP2) with similar affinity (Kumari et al., 2024).

This study examined the Clp chaperone, LinClpA, and its associated adaptor proteins (LinClpS1 and LinClpS2) in *L. interrogans*. Our findings indicate that the N-domain of LinClpA does not play a crucial role in its ATPase activity and interaction with LinClpP isoforms. Furthermore, we show that LinClpA undergoes time-dependent auto-degradation when in a complex with LinClpAP1P2, and the signal for this degradation is not located within the N-domain. Our study also provides evidence of the functionality of the C-degron (SsrA-tag) of *Leptospira*. The enhanced green fluorescent protein (eGFP), which is usually stable, became a model substrate for LinClpAP1P2 machinery when tagged with the predicted LinSsrA (eGFP-SsrA). Additionally, the inhibitory effect of binding with two different adaptor proteins (LinClpS1 and LinClpS2) on the ATPase activity of LinClpA and protease activity of LinClpAP1P2 machinery has been explored.

4.3. Materials and Methods

4.3.1. Bioinformatics analysis

The amino acid sequences of leptospiral ClpA, ClpS1, and ClpS2 proteins were compared using the protein Blast (Blastp) online server (Altschul et al., 1990). The sequences of ClpA (*Leptospira interrogans* ClpA*;* Q72RD2, *Escherichia coli* ClpA; P0ABH9, *Xanthomonas campestris* ClpA; Q8P998, *Borrelia burgdorferi* ClpA; O51342, *Helicobacter pylori* ClpA; A0AB33XJV6 and *Pseudomonas aeruginosa* ClpA; Q9I0L8*)* and ClpS (*Leptospira interrogans* ClpS1: Q72SM3, LinClpS2; Q72RD1, *Escherichia coli* ClpS; P0A8Q6, *Caulobacter crescentus* ClpS; B8GZM8, *Mycobacterium tuberculosis* ClpS; P9WPC0, *Agrobacterium tumefaciens* ClpS1; Q8UFN4, AtuClpS2; Q8UD95, *Synechococcus elongatus* ClpS1; Q31QE7 and SelClpS2; Q31R11) from various pathogenic bacteria were obtained from the UniProtKB database (UniProt consortium, 2015). A multiple sequence alignment (MSA) was conducted using Clustal Omega software and analyzed with the online tool ESPript (Easy Sequencing in PostScript) (Sievers et al., 2011; Robert et al., 2014). The tertiary structures of

TH-3792_206106022

LinClpA, LinClpS1, and LinClpS2 were retrieved from the AlphaFold (AF) protein structure database (Jumper et al., 2021). The available structures of EcoClpA (1KSF) and EcoClpS (3O2O) were obtained from the protein data bank (PDB) (Berman et al., 2000). Structural visualizations and superimpositions were performed using PyMol (DeLano, 2002).

Table 4.1. Primers used in this study

Primer Name	Primer sequence (5′-3′)
ClpA_F	CTAGCTAGCATGGAACGTACTTTAAGAAAGGCTT
ClpA_R	CCGCTCGAGATTCTTTTTTCCGGAAGAGAAT
ClpS1_F	CTAGCTAGCATGGCGAGTACACAAACTCCAG
ClpS1_R	CCGCTCGAGATCTTTCCAATGTAGCACTCA
ClpS2_F	CTAGCTAGCATGAGTGATATCTTTCGATTCGA
ClpS2_R	CCGCTCGAGTGACTCCTCCTCTCCTTCC
ClpA$^{\Delta N}$_F	CGGCTAGCAAAGATTCTAAAAAAAATCCAGGC
ClpA$^{\Delta N}$_R	CCGCTCGAGATTCTTTTTTCCGGAAGAGAAT
eGFP-SsrA_F	gccctggccgcctaaGAATTCGAGCTCCGTCGA
eGFP-SsrA_R	cagctcgttgttggcCTTGTACAGCTCGTCCATG
LIC13341-SsrA_F	gccctggccgcctaaTGACATCATCATCATCATCAC
LIC13341-SsrA_R	cagctcgttgttggcCTCGAGTTCTTGCTTGGAAAC

Nucleotide sequences in lowercase letters denote the inserted nucleotides.

4.3.2. Cloning, overexpression, and purification of recombinant proteins

The full-length genes (*clpA*; *LIC11814*; 2220 bp, *clpS1*; *LIC11356*; 318 bp, *clpS2*; *LIC11815*; 333 bp), and partial open reading frame of *clpA* (*clpA*$^{\Delta N}$; 1806 bp), lacking the sequence encoding N-domain of LinClpA of *L. interrogans* serovar Copenhageni strain Fiocruz L1-130, were amplified using PCR with its genomic DNA as a template. The oligonucleotides used for PCR amplification were designed based on the genomic sequence of *L. interrogan*s available on the National Center for Biotechnology Information (NCBI) and are listed in **Table 4.1**. Each gene was individually cloned into the pET-23a vector at the *Nhe*I and *Xho*I restriction sites to overexpress recombinant proteins with a C-terminal 6×His tag. The transformed recombinant vectors were introduced into *E. coli* BL21 (DE3) strain for induction, overexpression, and purification of the protein of interest. The transformed bacterial cells were cultured at 37°C in Luria Bertani medium, and protein overexpression was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG; 1 mM) in the pilot experiments to get the best yield during purification. For LinClpA and LinClpA$^{\Delta N}$, the overexpression was induced for 16 hours at 18°C, while for LinClpS1 and LinClpS2, induction was performed for 4 hours at 37°C. The recombinant proteins were purified by nickel-nitrilotriacetic acid (Ni-NTA) affinity column chromatography under native conditions, as previously described (Dhara et al., 2019). The

TH-3792_206106022

purified proteins were visualized on a 12% sodium dodecyl sulfate-polyacrylamide gel by Coomassie staining.

4.3.3. Generation of Anti-LinClpA antibodies and immunoblot analysis

Polyclonal antibodies specific to purified LinClpA of *L. interrogans* were generated in mice as described previously (Kumari et al., 2024). Briefly, BALB/c mice (4–6 weeks old, n = 5) were immunized with purified LinClpA (15–20 µg) emulsified in Freund's complete adjuvant (FCA; Santa Cruz Biotechnology #SC-3727). The primary immunization was administered subcutaneously, followed by two booster injections of LinClpA emulsified in Freund's incomplete adjuvant FIA; Santa Cruz Biotechnology #SC-3726 at days 14 and 21 of primary immunization. Blood was collected by retro-orbital bleeding seven days after the second booster, and the sera were pooled. The LinClpA antibody titer was analyzed by enzyme-linked immunosorbent assay (ELISA). The immunization experiments were performed at the Department of Microbiology, College of Veterinary Science, Assam Agricultural University, Guwahati, India, after approval by the Institutional Animal Ethics Committee (approval no.770/ac/CPCSEA/FVSC/AAU/IAEC/13-14).

To detect native LinClpA expression in *L. interrogans* serovar Copenhageni, whole-cell lysates of 3 $\times 10^9$ spirochetes were re-suspended in sodium dodecyl sulfate (SDS) loading dye. The pure recombinant LinClpA protein and lysate of *Leptospira* were resolved in 12% SDS-PAGE and transferred to a nitrocellulose membrane (Bio-Rad). The membranes were blocked with 5% non-fat dried milk prepared in Tris-buffered saline (TBS; pH 8.0) containing 0.1% Tween 20 (TBS-T) and probed with mouse anti-LinClpA (1:1000) antibodies for 2 h at room temperature. After being washed, the membranes were incubated with horse radish peroxidase (HRP)-conjugated goat anti-mice IgG (1:5000; Sigma) for 1 h, and immunoblots were developed by adding the chemiluminescence substrate (Thermo Scientific, catalog no. 32209). All dilutions of antibodies were prepared using 2% non-fat dried milk in 0.1% TBS-T.

4.3.4. Leptospiral SsrA-tag sequence prediction and generation of model substrates

The amino acid sequences of the SsrA-tag of *L. interrogans* and several pathogenic bacteria were retrieved from the transfer-messenger RNA database (tmRDB) (Knudsen et al., 2001). The last nine amino-acid sequences of predicted SsrA-tag from several bacteria were analyzed by the WebLogo webserver (Crooks et al., 2004). Site-directed insertion mutagenesis was performed to generate model protein substrates with a LinSsrA-tag at the C-terminal end. A Q5 site-directed mutagenesis kit (NEB, catalog no. E0554S) was used to insert the 30-

TH-3792_206106022

nucleotide sequence (GCCAACAACGAGCTGGCCCTGGCCGCCTAA) encoding a 9-amino acid long SsrA-tag (ANNELALAA) of *Leptospira* at the downstream of *eGFP* and *LIC13341*genes in the pET-21d and pET-28a vectors, respectively. The recombinant proteins (eGFP-SsrA; 31 kDa and LIC13341-SsrA; 43 kDa) were purified as previously described (Ghosh et al., 2018).

4.3.5. ANS binding assay

The nucleotide-induced oligomerization tendency of LinClpA was assessed using a hydrophobic fluorophore, 8-anilino-1-naphthalenesulfonic acid (ANS), as outlined in a previous study (Kumari et al., 2024). LinClpA (3 µM) was pre-incubated with or without different nucleotide analogs (ATP, ADP, PolyP, AMP; 2 mM) for 20 minutes at 4 °C in an assay buffer (50 mM Tris-Cl; pH 7.8, 50 mM KCl, 1 mM DTT, and 8 mM $MgCl_2$). The ATP-bound or pure LinClpA was then mixed with ANS (10 µM) and incubated for 30 minutes at room temperature in the dark. A control reaction with ANS (10 µM) in assay buffer was also kept under similar conditions. Fluorescence spectra were recorded in a black, flat-bottom 96-well microplate (Invitrogen) using a spectrofluorometer (iTECAN infinite M PLEX). All samples were excited at 350 nm, and emission spectra were recorded from 400 to 750 nm with 5 nm increments.

4.3.6. ATPase activity analysis

The ATPase activity was measured by determining the rate of release of free phosphate (µM min^{-1}) upon ATP hydrolysis using a malachite green phosphate assay kit (Sigma; MAK307), as previously described (Kumari et al., 2024). Each 40 µL ATPase assay reaction contained the enzymes (LinClpA or LinClpA$^{\Delta N}$; 1 µM) and the required amount of ATP (1-8 mM) in assay buffer and was incubated for 1 hour at 37 °C. To study the effect of the presence of Clp protease and adaptor proteins on ATPase activity, the reactions containing LinClpA or LinClpA$^{\Delta N}$ were pre-incubated with increasing concentrations (1-10 µM) of LinClpP (ClpP1 and ClpP2) and LinClpS (ClpS1 and ClpS2) proteins for 10 minutes at 37 °C, followed by ATP addition. The absorbance measurements at 620 nm were performed using a spectrofluorometer (iTECAN infinite M PLEX). All experiments were performed independently three times and in duplicate.

4.3.7. Immunoassay for interaction analysis

An indirect enzyme-linked immunosorbent assay (ELISA) was conducted to examine the interaction between chaperone-protease (LinClpAP complex) and chaperone-adaptor proteins

TH-3792_206106022

(LinClpAS complex). Test proteins (LinClpP1/LinClpP2/LinClpS1/LinClpS2; 1 µM/well) and a control protein (bovine serum albumin; BSA; 1 µM/well) were coated on a 96-well microtiter plate overnight at 4 °C. Following this, each well was blocked with a 3% BSA solution in phosphate-buffered saline (PBS) for 2 hours at 37 °C and then overlaid with increasing concentrations of ATP-bound LinClpA or LinClpA$^{\Delta N}$ (0-3 µM) for 2 hours at 37 °C. After three washes with PBS containing 0.05% Tween 20 (PBS-T) buffer, mouse anti-LinClpA (1:1000) was added. The interaction was detected using HRP-conjugated anti-mouse secondary antibodies (1:5000) with TMB (Trimethylbenzidine, Thermo Scientific) as the substrate. The absorbance was measured at a wavelength of 450 nm after terminating the reaction using 1 M H_2SO_4 as per the manufacturer's instructions. All experiments were independently performed three times and in duplicate.

4.3.8. Auto-degradation assay of LinClpA and its variant

The LinClpA or its N-domain deleted variant (LinClpA$^{\Delta N}$), at a specific concentration (1.5 µM), was pre-incubated with either pure LinClpP isoforms (3 µM) or the LinClpP1P2 heterocomplex in assay buffer for 10 minutes at 37 °C before the addition of ATP (5 mM). When necessary, individual adaptor proteins (LinClpS1 and LinClpS2, 5 µM) were added to the reaction mixture. The reactions were carried out for 180 minutes at 37 °C, with 25 µL of the reaction mixture aliquoted at 0, 60, 120, and 180-minute intervals. These aliquots were mixed with 5× SDS sample buffer and heated for 10 minutes at 95 °C to halt the degradation reaction. Following separation by SDS-PAGE and Coomassie staining, the amount of LinClpA and LinClpA$^{\Delta N}$ remaining was quantified using densitometry analysis with Image Lab software from Bio-Rad.

4.3.9. Degradation of SsrA-tagged substrates by LinClpAP1P2 protease machinery

For the protease assay, equimolar amounts of LinClpP1 and LinClpP2 were allowed to incubate for 10 minutes at 37°C to form the LinClpP1P2 heteromplex. Following this, LinClpA was added to the self-assembled LinClpP1P2 heterocomplex and allowed to incubate for 10 minutes at 37°C to create the LinClpAP1P2 machinery. Whenever necessary, the individual adaptor proteins (LinClpS1 and LinClpS2) were added at a concentration of 5 µM. The proteolytic activity of the LinClpAP1P2 machinery (3 µM) was tested on the eGFP-SsrA model substrate (2 µM) in an assay buffer containing ATP (5 mM). The degradation of eGFP-SsrA was monitored in black, flat-bottom 96-well microplates (Invitrogen) at 37°C for 60 minutes by measuring the loss of fluorescence at an excitation wavelength of 485 nm and emission

TH-3792_206106022

wavelength of 525 nm. The change in relative fluorescence units (ΔRFU) for eGFP-SsrA (RFU_{Final} - $RFU_{Initial}$) was determined and plotted against time.

Similarly, the LinClpAP1P2 machinery was allowed to degrade the LIC13341-SsrA substrate (2 µM) in assay buffer over 180 minutes at 37°C. At 0, 60, 120, and 180-minute intervals, 25 µL of the reaction mixture was sampled, mixed with SDS sample buffer, and then visualized by running the aliquots on a 12% SDS-PAGE gel, followed by Coomassie staining.

4.4. Results

4.4.1. Structural and sequence analysis of the leptospiral Clp ATPase (LinClpA)

The full-length LinClpA protein consists of 740 amino acids and includes an N-domain followed by two ATPase domains- I and II. Each ATPase domain (I &II) is further subdivided into large and small subdomains (**Figure 4.1A**). The N-domain is composed of 138 residues and is connected to a large subdomain by a flexible linker region of 22 residues (from amino acids 139 to 160). The D-I domain spans 267 residues, while the second domain (D-II) comprises 313 residues. A sequence comparison of LinClpA with other bacterial orthologs reveals a high sequence identity with *Pseudomonas aeruginosa* (PaeClpA; 52%), *E. coli* (EcoClpA; 51%), and *Xanthomonas campestris* (XcaClpA; 50%). In contrast, LinClpA shows lower identity with *Helicobacter pylori* (HpyClpA; 38%) and *Borrelia burgdorferi* (BbuClpA; 28%) (**Figure 4S1**). The pairwise sequence alignment of the N-domain, D-I, and D-II of LinClpA and EcoClpA indicates that each ATPase domains share more than 50% sequence identity, while LinClpA's N-domain shows 40% identity with EcoClpA's N-domain. Key elements critical for ATPase function, such as Walker A, Walker B, the Arginine finger, Sensor I, and Sensor II, are highly conserved in both the D-I and D-II domains of LinClpA (**Figure 4.1A**). The Walker A motif of LinClpA is located in the D-I domain at residues 206-213 (GEAGVGKT) and in the D-II domain at residues 485-492 (GPTGVGKT); this motif is essential for ATP binding and oligomerization. Similarly, the Walker B motif, in the D-I domain of LinClpA at residues 273-280 (VLFIDEIH) and in the D-II domain at residues 551-557 (LLLDEIE), is crucial for ATP hydrolysis. The Sensor I motif and Arginine finger responsible for ATP hydrolysis are located at residues 310-315 (CIGTTT; D1 Sensor I) and 591-596 (LVMTTN; D2 Sensor I), as well as at Arg331, Arg332, and Arg633, respectively. Additionally, the Sensor II motifs of the D-I and D-II domains of LinClpA, which are involved in conformational changes during ATP binding and hydrolysis, are situated at residues 383-392 (DRKLPDKAID) and 689-694 (FGARPV), respectively. Furthermore, the tertiary structure of the LinClpA monomer was modeled using AlphaFold and showed similarity to the

TH-3792_206106022

structure of EcoClpA (PDB: 1KSF) (**Figure 4.1B – 4.1C**). The modeled structure of the LinClpA N-domain consists of eight helices (α1–α8) with approximately 60% helical content, similar to the EcoClpA N-domain (Guo et al., 2002). The LinClpA N-domain contains two conserved repeats, each approximately 60 residues long, showing a pseudo-two-fold symmetry. Repeat 1 (α1-α4) can be superimposed onto repeat 2 (α5-α8) of LinClpA with a root mean square deviation (r.m.s.d.) of 1.2 (**Figure 4S2a**). The large subdomains of D-I and D-II in LinClpA possess a core α/β folding motif characterized by five β-sheets flanked by α-helices on either side, which is analogous to the reported structures of EcoClpA's D-I and D-II domains (**Figure 4S2b – 4S2c)** (Guo et al., 2002). The small subdomain of D-I in LinClpA contains four α-helices, while the small subdomain of D-II has three α-helices and three β-sheets.

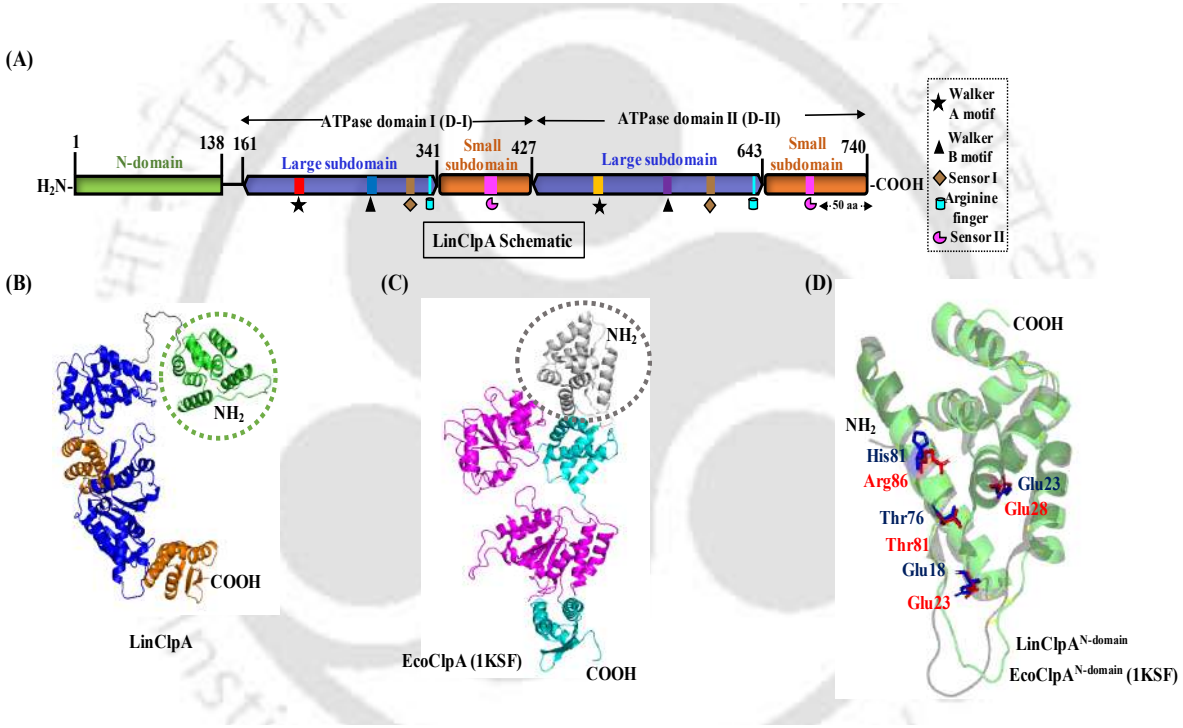


Figure 4.1. Prediction for LinClpA domain organization and the three-dimensional structure. (A) The diagram illustrates the predicted N-domain, ATPase domain I (D-I), and ATPase domain II (D-II), presented roughly to scale. The large and small subdomains of the ATPase domains are depicted as hexagonal and square shapes, respectively. The essential components of the ATPase domains are color-coded as follows: Walker A (red and orange), Walker B (blue and violet), Sensor I (brown), Arginine finger (cyan), and Sensor II (magenta). **(B)** The modeled tertiary structure of LinClpA **(C)** tertiary structure of *E. coli* ClpA (EcoClpA; PDB- 1KSF) shown, with the N- and C-terminal ends indicated as NH_2 and COOH, respectively. The N-domain, ATPase large subdomain, and ATPase small subdomain of LinClpA (green, blue, and orange) and EcoClpA (grey, magenta, and cyan), respectively. The N-domains of LinClpA and EcoClpA are highlighted by green and grey circles and zoomed in for structural superimposition. **(D)** The modeled LinClpA N-domain (green) and EcoClpA N-domain (1KSF; grey) were structurally superimposed with an r.m.s.d of 0.9. Essential residues for interaction with adaptor proteins are marked as red sticks (EcoClpA N-domain) and blue sticks (LinClpA N-domain).

TH-3792_206106022

The N-domain of EcoClpA has multiple binding sites for EcoClpS, which are crucial for EcoClpS docking, followed by substrate recognition and degradation (Zeth et al., 2002). The structural superimposition of the modeled LinClpA N-domain with EcoClpA N-domain (PDB: 1KSF) indicates that residues facilitating interactions between EcoClpA's N-domain and EcoClpS are highly conserved in LinClpA's N-domain (**Figure 4.1D**). The predicted contact site with adaptor proteins in LinClpA contains one aliphatic residue (Thr76) and three charged residues (Glu18, Glu23, and His81), which align with EcoClpA N-domain residues (Thr81, Glu23, Glu28, and Arg86) (Zeth et al., 2002).

4.4.2. Structural and sequence analysis of the leptospiral Clp adaptor proteins (LinClpS1 and LinClpS2)

The complete genome sequence of *L. interrogans* shows the existence of two *clpS* genes (*LIC11356*; *clpS1* and *LIC11815*; *clpS2*). Interestingly, the genes *clpS2* (333 bp) and *clpA* (2220 bp) in *Leptospira* are situated next to each other on the chromosome with a 14-nucleotide gap, while the *clpS1* (318 bp) is located elsewhere in the genome. LinClpS1 (12 kDa) and LinClpS2 (13 kDa) are quite close in terms of size and structure, where both have an N-domain (22 vs 25 residues) followed by a C-domain (84 vs 86 residues) (**Figure 4.2A – 4.2B**). A comparison of the amino acid sequences of LinClpS1 and LinClpS2 with well-studied ClpS orthologs was carried out using ClustalW and represented as a heat map (**Figure 4S3a**). The length and sequence conservation of the N-domain within various ClpS orthologs vary widely, while the C-domain is moderately conserved. LinClpS2 shares 30-40% sequence identity with EcoClpS, CcrClpS (*Caulobacter crescentus* ClpS), AtuClpS1 (*Agrobacterium tumefaciens* ClpS1), AtuClpS2, and SelClpS1 (*Synechococcus elongatus* ClpS1). In contrast, LinClpS1 shares a lower sequence identity (15-25%) with EcoClpS, CcrClpS, AtuClpS1, AtuClpS2, SelClpS1, SelClpS2, and MtuClpS (*M. tuberculosis* ClpS). The co-crystal structure of CcrClpS with a decapeptide having a Tyr residue at the N-terminal demonstrates that the α-amino group of the peptide interacts with specific residues (Asn47, Asp49, and His79) of CcrClpS (Wang et al., 2008). The residues corresponding to the peptide interaction site in CcrClpS are highlighted for LinClpS1 and LinClpS2 (**Figure 4.2A – 4.2B**). The LinClpS1 protein retains conserved Asp30 and Asn32 residues, while the conserved His residue has been evolutionarily replaced with Asp62 within the binding pocket. The binding pocket of LinClpS2 maintained the conserved residues (Asn33, Asp35, and His65) alike its orthologs. In the CcrClpS-peptide co-crystal, the tyrosine ring of the peptide fits into a deep hydrophobic pocket on the surface of CcrClpS. This hydrophobic pocket for CcrClpS, known as the specificity pocket, is formed

TH-3792_206106022

by a list of residues (Ile45, Leu46, Asn47, Asp48, Thr51, Met53, Val56, Met75, Val78, His79, and Leu112) (Wang et al., 2008). Accordingly, the residues corresponding to the specificity pocket of LinClpS1 (Leu28, Trp29, Asp 31, His34, Tyr36, Val39, Ala58, Val61, Met100) and LinClpS2 (Ile31, Leu32, Asp34, Thr37, Met39, Val42, Met61, Ala64, Leu98) were assessed and highlighted (**Figure 4.2A - 4.2B**). In EcoClpS of *E. coli*, the residue Met40 acts as a specific gatekeeper, accommodating the hydrophobic side chains of a few residues (Phe, Trp, Tyr, and Leu) while excluding β-branched amino acid-containing peptides (Wang et al., 2008). The residue Met40 of the specificity pocket recorded in EcoClpS is also conserved in LinClpS2; however, LinClpS1 has a non-conserved residue Tyr36. The natural replacement of Met with Tyr residue in the LinClpS1 peptide binding pocket might indicate a different substrate specificity compared to EcoClpS, CcrClpS, and LinClpS2.

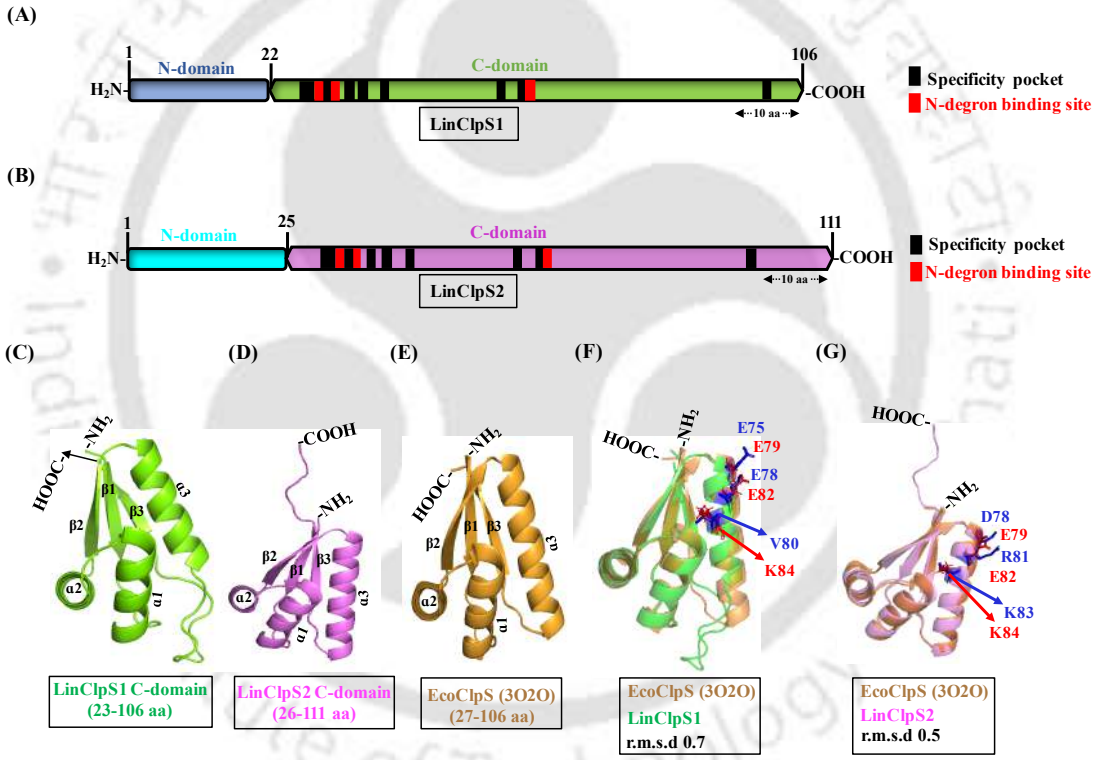


Figure 4.2. Prediction of the domain organization and three-dimensional structure of LinClpS1 and LinClpS2. The diagram illustrates the predicted domains of **(A)** LinClpS1 and **(B)** LinClpS2 to scale. In the diagram, the N-domains of LinClpS1 and LinClpS2 are represented as square shapes, while the C-domains are depicted as hexagonal shapes. The residues within the C-domain of LinClpS1 and LinClpS2 that are predicted to bind to the LinClpA N-domain are marked with red boxes. Additionally, black boxes in the C-domain indicate the residues likely important for binding with the N-degron peptide (specificity pocket). The modeled tertiary structure of the C-domain of **(C)** LinClpS1 is shown in green, while **(D)** LinClpS2 is represented in magenta. In **(E)**, the tertiary structure of *E. coli* ClpS (EcoClpS; 3O2O) is depicted in orange. The structural superimposition in **(F)** demonstrates modeled LinClpS1 (green) and EcoClpS (3O2O; orange) with an r.m.s.d of 0.7, while **(G)** shows modeled LinClpS2 (magenta) superimposed with EcoClpS (3O2O; orange) with an r.m.s.d of 0.5. The residues corresponding to the chaperone binding pocket are highlighted as red sticks for EcoClpS and blue sticks for LinClpS1 and LinClpS2.

TH-3792_206106022

The three-dimensional structure of LinClpS1 and LinClpS2 was modeled using AlphaFold and compared to the available crystal structure of EcoClpS (PDB: 3O2O) (**Figure 4.2C – 4.2E**). EcoClpS has a disordered N-domain, referred to as the N-terminal extension (NTE), and a folded core domain (26-106 residues) (Zeth et al., 2002). The disordered region in the N-domain of LinClpS1 and LinClpS2 was assessed using the PrDOS online server and compared with EcoClpS (Ishida et al., 2007). The server predicts the disordered regions of a protein from its amino acid sequence and provides a disorder probability for each residue. The average disorder probability of the N-domain of LinClpS1, LinClpS2, and EcoClpS was estimated to be 0.6-0.7 (**Figure 4S3b**). This suggests that LinClpS1 and LinClpS2 also possess an unstructured N-terminal extension region similar to EcoClpS (Zeth et al., 2002).

The modeled structures of LinClpS1 (23-106 residues) and LinClpS2 (26-111 residues) are composed of three α-helices (α1–α3) linked to three β-strands (β1–β3) in a βααβαβ arrangement (**Figure 4.2C – 4.2D**). In EcoClpS, the helix α3 contains highly conserved residues (Glu82 and Lys84), which form salt bridges with the EcoClpA residues, Agr86 and Glu23, respectively (Zeth et al., 2002). Additionally, the Glu79 of EcoClpS forms hydrogen bonds with Glu28 and Thr81 of EcoClpA (Zeth et al., 2002). The modeled structures of LinClpS1 and LinClpS2 were superimposed with EcoClpS (PDB: 3O2O), with r.m.s.d values of 0.7 and 0.5, respectively (**Figure 4.2F – 4.2G**). The structure comparison revealed that in LinClpS1, the conserved charged residue Lys84 present in EcoClpS has been evolutionarily replaced with the aliphatic Val80, while LinClpS2 retains the conserved Lys83 residue. These variations within the chaperone binding pocket of LinClpS1 and LinClpS2 may indicate their different affinities for the N-domain of LinClpA. To comprehensively study Clp ATPase and adaptor proteins of *Leptospira*, the genes of interest (*clpA*, *clpS1*, and *clpS2*) were cloned, and recombinant proteins were over-expressed and purified from *E. coli* BL21 cells. A partial *clpA* gene ($clpA^{\Delta N}$) encoding LinClpA with a deleted N-domain (138 residues) was also cloned, over-expressed, and purified. The quality of the recombinant proteins (LinClpA, $LinClpA^{\Delta N}$, LinClpS1, and LinClpS2) purified using Ni-NTA affinity chromatography was analyzed on polyacrylamide gel and stained with Coomassie Blue (**Figure 4S3c**). Polyclonal antibodies were generated against recombinant LinClpA protein in BALB/c mice, and the resulting sera were assessed for adequate antibody titer (1:1000). Subsequently, immunoblot analysis was performed using whole-cell lysates of *Leptospira interrogans* and the corresponding recombinant protein. The anti-LinClpA antibodies detected the native LinClpA protein (~82 kDa) expression in *L. interrogans* serovar Copenhageni. Additionally, the immunoblot results confirmed the specificity of the anti-LinClpA antibodies, as no cross-reactivity was observed

TH-3792_206106022

with other related proteins such as LinClpP isoforms or the LinClpC chaperone present in the whole-cell lysate of serovar Copenhageni (**Figure 4S3d**).

4.4.3. Analysis of nucleotide-induced oligomerization of LinClpA

The AAA+ proteins typically assemble into higher oligomeric structures upon binding to nucleotides (Maurizi, 1991). The influence of nucleotide (ATP; 2 mM) on the oligomeric state of purified recombinant LinClpA was analysed using size-exclusion chromatography (**Figure 4.3A**). In the absence of ATP, LinClpA predominantly eluted in its monomeric form (~82 kDa). However, upon incubation with ATP, the elution profile shifted towards a higher molecular weight species (size ~492 kDa), suggesting ATP-induced oligomerization of LinClpA.

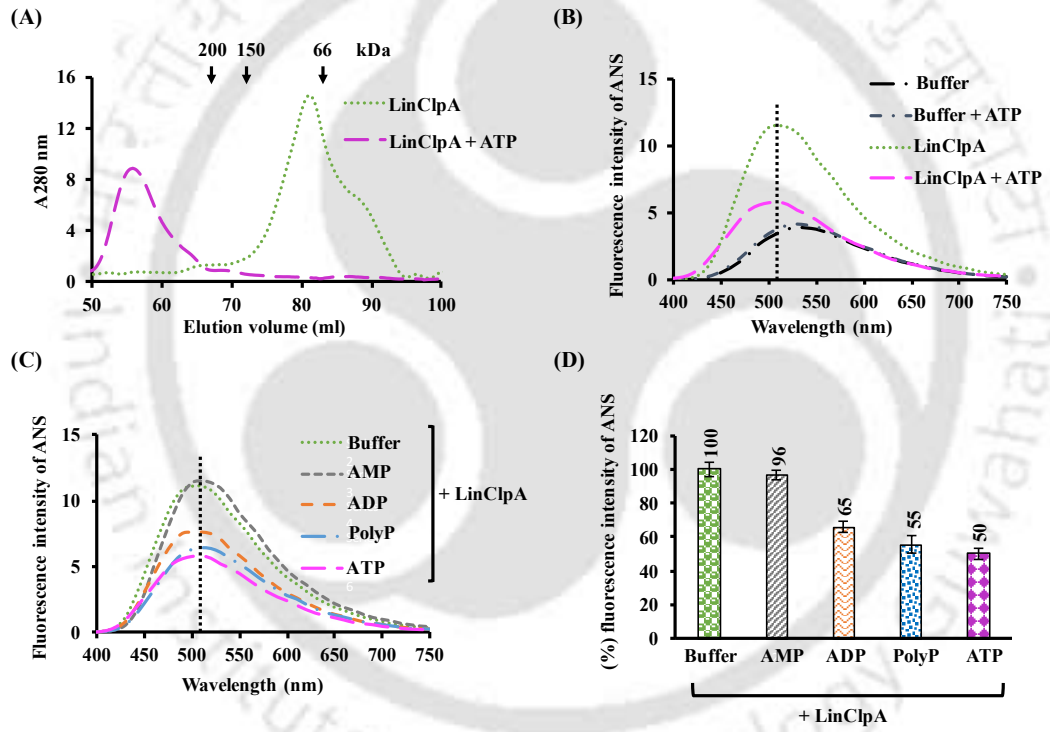


Figure 4.3. Oligomerization property of LinClpA. **(A)** Size exclusion chromatography analysis of LinClpA in the absence (-) and presence (+) of ATP. An ANS binding assay was conducted to study the oligomerization of LinClpA. This assay measured the emission spectra of a hydrophobic fluorophore, ANS, in the 400-750 nm range, using an excitation wavelength of 350 nm. The fluorescence intensity of ANS was recorded when bound to LinClpA **(A)**, both in the absence (-) and presence (+) of ATP **(B)**, alongside other molecules such as AMP, ADP, and PolyP. **(C)** The percentage change in fluorescence intensity of ANS was calculated when LinClpA was bound to AMP, ADP, PolyP, and ATP against a sample of pure LinClpA, which served as the baseline (considered as 100%). Each ANS binding assay experiment was performed three times, and one representative spectrum was displayed.

Furthermore, previous s research has explored the oligomerization propensity of a leptospiral ATPase chaperone LinClpC using a hydrophobic fluorophore 8-anilino-1-naphthalenesulfonic

TH-3792_206106022

acid (ANS) dye-binding assay (Kumari et al., 2024). To examine the nucleotide-induced oligomerization of LinClpA, a binding assay utilizing the ANS dye was also conducted (**Figure 4.3B**). ANS is inclined to bind to the exposed hydrophobic regions of proteins, resulting in increased fluorescence than the unbound form. Oligomerization of proteins leads to the masking of hydrophobic regions, causing reduced binding with ANS (Semisotnov et al., 1991). Pure LinClpA or LinClpA pre-incubated with ATP was mixed with ANS in assay buffer, and fluorescence spectra were examined at a fixed excitation wavelength of 350 nm and emission wavelength ranging from 400 - 750 nm (**Figure 4.3B**). Pre-incubation of LinClpA with ATP resulted in the reorganization of its subunits into higher molecular weight complexes, leading to reduced binding with ANS and a consequent decrease in fluorescence intensity, indicating high-order oligomerization of LinClpA in the presence of ATP. Previous reports have indicated that EcoClpA forms hexamers upon binding with nucleotide triphosphates (ATP or ATPγS), while the addition of nucleotide diphosphate (ADP) promotes its dissociation into dimers (Maurizi, 1991; Wickner et al., 1994). Another study on EcoClpA oligomerization suggests that the oligomerization of EcoClpA is induced by nucleotide triphosphate analogs (ATPγS, AMP-PNP, and AMP-PCP) as well as nucleotide diphosphates (ADP·BeF and ADP) (Veronese et al., 2011). These significant differences in the oligomerization behaviour of EcoClpA prompted an investigation into the oligomerization tendency of LinClpA in the presence of various nucleotide analogs. The fluorescence spectra of LinClpA incubated with either ATP, ADP, AMP, or polyphosphates (PolyP) were analyzed and compared with pure LinClpA (**Figure 4.3C**). The spectral analysis revealed that incubation of LinClpA with ATP, ADP, or PolyP promoted its oligomerization into higher molecular weight complexes and resulted in a reduction of ANS fluorescence. Conversely, pre-incubation of LinClpA with AMP resulted in no spectral changes to those of pure LinClpA. Furthermore, the percentage change in the fluorescence intensity of ANS for each reaction was determined using the fluorescence intensity of ANS bound to pure LinClpA as the reference (**Figure 4.3D**). The findings showed that pre-incubation of LinClpA with either ADP, PolyP, or ATP led to reductions in ANS fluorescence intensity from 100% to 65%, 55% and 50% respectively. Therefore, it can be concluded that the presence of two or more phosphate moieties (ATP, ADP, or PolyP) promotes the oligomerization of LinClpA, while AMP with only one phosphate does not have the same effect.

TH-3792_206106022

4.4.4. LinClpA intrinsic ATPase activity

A malachite green assay was utilized to quantify the intrinsic ATP hydrolysis activity of LinClpA, which detects the liberated phosphate subsequent to ATP hydrolysis into ADP.

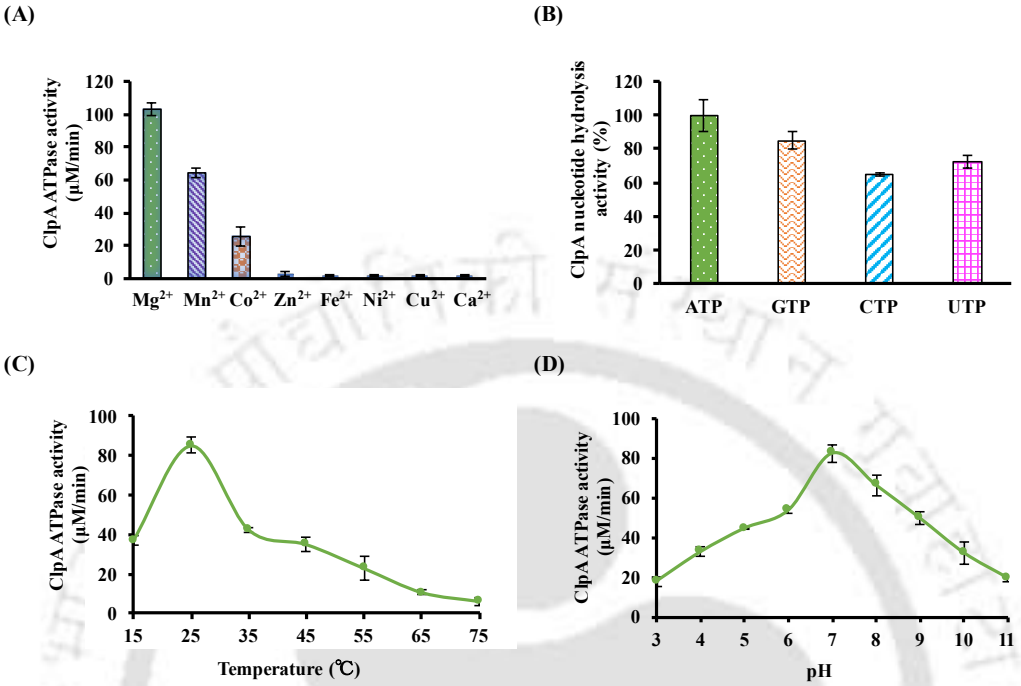


Figure 4.4. Nucleotide hydrolysis property of LinClpA under diverse parameters. (A) LinClpA ATPase activity was measured in the presence of various divalent metal cations (Mg^{2+}, Mn^{2+}, Co^{2+}, Zn^{2+}, Fe^{2+}, Ni^{2+}, Cu^{2+}, and Ca^{2+}). **(B)** The nucleotide hydrolysis activity of LinClpA on various nucleotides (GTP, CTP, UTP) was compared with the rate of hydrolysis of ATP (considering 100%). **(C)** LinClpA ATPase activity in the temperature range 25 - 75 °C **(D)** LinClpA ATPase activity over a broad pH 3-11 range. The data represent the Mean ± Standard error mean (SEM) from the three independent experiments.

Firstly, the influence of different divalent metal cations on LinClpA's (1 µM) ATP hydrolysis activity was investigated for 1 hour at 37 °C by introducing various metal ions (8 mM) to the buffer (50 mM Tris-Cl, 50 mM KCl, 1 mM DTT; pH 7.8) (**Figure 4.4A**). The highest observed LinClpA ATPase activity occurred in the presence of Mg^{2+} (8 mM). Under similar conditions, when equimolar (8 mM) amounts of either Mn2+ or Co^{2+} were utilized, the ATPase activity of LinClpA was only 64% and 25%, respectively, compared to that of Mg^{2+} (100%). Other tested cations (Zn^{2+}, Fe^{2+}, Ni^{2+}, Cu^{2+}, and Ca^{2+}) did not support LinClpA ATPase activity (**Figure 4.4A**).

Subsequently, the investigation into the release of phosphate during the hydrolysis of various nucleotides, including GTP, CTP, and UTP, was aimed at evaluating LinClpA's ability to hydrolyze these nucleotides compared to ATP. Although LinClpA primarily utilizes ATP as its substrate, it was determined that it could also hydrolyze other triphosphates: GTP, CTP, and

TH-3792_206106022

UTP (**Figure 4.4B**). Specifically, LinClpA demonstrated phosphate hydrolysis rates of 84% for GTP, 64% for CTP, and 72% for UTP, with ATP serving as the standard at 100%. Furthermore, the ATPase activity of LinClpA was examined to understand how environmental factors, such as pH and temperature, influence its functionality. Optimal ATPase activity was observed within a temperature range of 20-30°C and at a pH of 7 (**Figure 4.4C and 4.4D**).

A comparison of nucleotide-induced oligomerization between LinClpA and LinClpA$^{\Delta N}$ revealed no significant differences, suggesting that the N-domain is not critical for the protein's oligomerization (**Figure 4S4a**). Additionally, both LinClpA and LinClpA$^{\Delta N}$ demonstrated dose-dependent ATPase activity, reaching saturation at specific ATP concentrations (**Figure 4S4b**). The maximum activity for LinClpA was recorded at 94.6 ± 4 μM/min, while for LinClpA$^{\Delta N}$, it was 102 ± 4.5 μM/min, indicating a 15% increase in V_{max} for LinClpA$^{\Delta N}$.

4.4.5. Differential association of LinClpA and LinClpA$^{\Delta N}$ with Clp protease and adaptor proteins

A study was conducted to assess the capacity of LinClpA to form complexes with the LinClpP isoforms (LinClpP1 and LinClpP2), and this was compared to the association abilities of LinClpA with the N-domain deleted variant (LinClpA$^{\Delta N}$). An enzyme-linked immunosorbent assay (ELISA) utilizing antibodies against LinClpA (anti-LinClpA) was employed to monitor the interactions of LinClpA or LinClpA$^{\Delta N}$ with the LinClpP isoforms. The results demonstrated a linear increase in absorbance at 450 nm for the complexes (LinClpAP1, LinClpAP2, LinClpA$^{\Delta N}$P1, and LinClpA$^{\Delta N}$P2) as the concentrations of LinClpA and LinClpA$^{\Delta N}$ increased (**Figure 4.5A and 4.5B**). The dissociation constant (K_d) for these complexes was determined through a Hill plot analysis, and the obtained K_d values for the complexes are displayed in **Figure 4.5C**. Our immunoassay results indicate that LinClpA and LinClpA$^{\Delta N}$ exhibit similar affinities for the pure LinClpP1 and LinClpP2 isoforms, as reflected by a K_d value of 0.28 ± 0.02 μM.

Additionally, an immunoassay was performed to explore the association between LinClpS adaptor proteins (LinClpS1 and LinClpS2) and LinClpA or LinClpA$^{\Delta N}$ (**Figure 4.5A and 4.5B**). The results reveal that LinClpA exhibits differential interactions with LinClpS1 and LinClpS2. Notably, a lower dissociation constant was observed for the LinClpAS2 complex (0.32 ± 0.03 μM) relative to the LinClpAS1 complex (0.44 ± 0.025 μM), indicating that the LinClpAS2 complex is more stable than its LinClpAS1 counterpart (**Figure 4.5C**). Furthermore, a significant change in the dissociation constants for the complexes LinClpA$^{\Delta N}$S1

TH-3792_206106022

and LinClpA$^{\Delta N}$S2 was noted when compared to LinClpAS1 and LinClpAS2, respectively (**Figure 4.5C**).

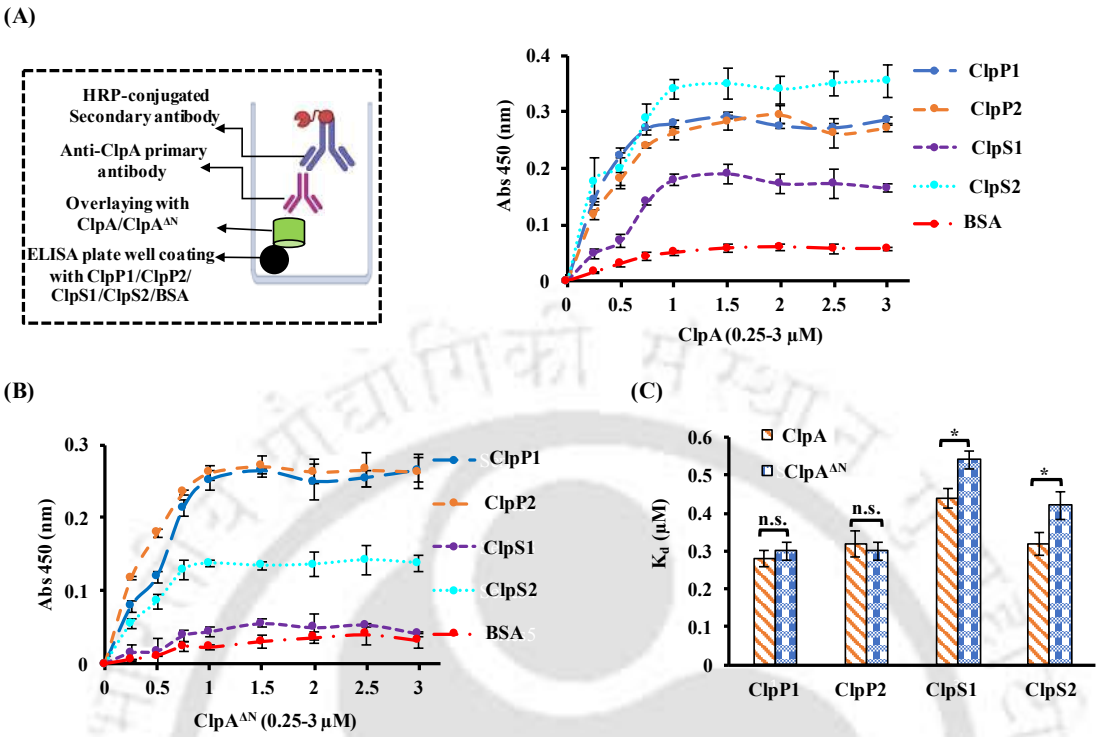


Figure 4.5. Association of LinClpA or its variant LinClpA$^{\Delta N}$ with LinClp protease and adaptor proteins. The interaction study used an enzyme-linked immunosorbent assay (ELISA) under *in vitro* conditions. The interaction between **(A)** LinClpA and LinClpP1/LinClpP2 or LinClpS1/LinClpS2 or BSA, **(B)** LinClpA$^{\Delta N}$ and LinClpP/LinClpP2 or LinClpS1/LinClpS2 or BSA. **(C)** The dissociation constant (K_d) of the complexes (LinClpAP1, LinClpAP2, LinClpAS1, LinClpAS2 and LinClpA$^{\Delta N}$P1, LinClpA$^{\Delta N}$P2, LinClpA$^{\Delta N}$S1, LinClpA$^{\Delta N}$S2) was determined via Hill plot analysis and plotted. The data represent the Mean ± Standard Error of the Mean (SEM) from three independent experiments.

These findings suggest that the presence of the N-domain in LinClpA is essential for stable complex formation with LinClpS adaptor proteins, while it does not affect the formation of complexes with LinClpP isoforms. Furthermore, we confirmed the impaired interaction between LinClpA$^{\Delta N}$ and LinClpS adaptor proteins through ATPase activity analysis. We assessed the rate of ATP hydrolysis by LinClpA and LinClpA$^{\Delta N}$ in the presence of various LinClpP isoforms or LinClpS adaptor proteins (**Figure 4.6A – 4.6B**). Our findings indicate that as the concentration of pure LinClpP isoforms increased, both LinClpA and LinClpA$^{\Delta N}$ exhibited a linear rise in ATPase activity. In contrast, LinClpA's ATPase activity was inhibited in a dose-dependent manner by either LinClpS1 or LinClpS2. Adding either LinClpS1 or LinClpS2 (10 µM) resulted in a 57 % and 90 % reduction in the LinClpA ATPase activity, respectively. Thus, the LinClpAS2 demonstrated a more stable complex than LinClpAS1 (**Figure 4.6A**). Nevertheless, the presence of LinClpS1 or LinClpS2 did not impact the ATPase

TH-3792_206106022

activity of LinClpA$^{\Delta N}$, which supports the need for the adaptor protein to bind to the N-domain of the ATPase chaperone (**Figure 4.6B**).

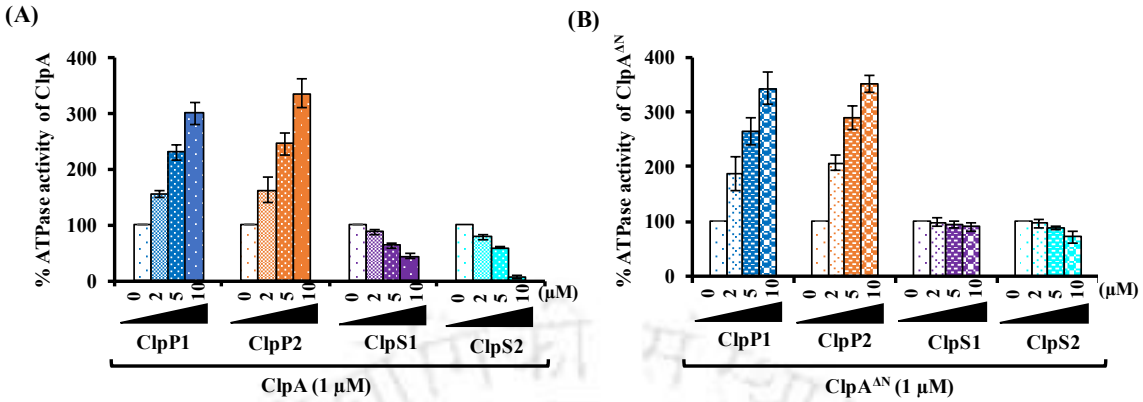


Figure 4.6. Influence of LinClpP and LinClpS on the ATPase activity of LinClpA and its variant. The ATPase activity of **(A)** LinClpA and **(B)** LinClpA$^{\Delta N}$ was measured in the presence of increasing concentrations of either LinClpP1, LinClpP2, LinClpS1, or LinClpS2. The percentage change in ATPase activity of LinClpA in the presence of an activator (LinClpP1 and LinClpP2) or inhibitor (LinClpS1 and LinClpS2) was calculated considering the activity of pure LinClpA to be 100%. The data represent the Mean ± Standard error mean (SEM) from the three independent experiments.

Next, we aimed to investigate whether the LinClpS adaptor protein, which inhibits the ATPase activity of LinClpA, also interferes with the activity of other class I chaperone in *Leptospira interrogans* (LinClpC). For this purpose, we assessed the ATPase activity of LinClpC in the presence of adaptor proteins. Our results showed that LinClpS1 and LinClpS2 also inhibit the ATPase activity of LinClpC (**Figure 4S5**). Interestingly, although neither *clpS1* nor *clpS2* genes are located in close proximity to the *clpC* gene within the *clp* gene cluster, but adaptor protein (LinClpS) is still capable of interfering with LinClpC activity. A similar regulatory mechanism has been reported in *Mycobacterium*, where MtuClpC activity is modulated by its ClpS adaptor proteins (Marsee et al., 2018). These findings suggest that LinClpS may also play a regulatory role in modulating LinClpC chaperone and warrants detailed investigation in future studies.

4.4.6. Auto-degradation of ClpA is retained in the absence of its N-domain

In *E. coli*, the chaperone EcoClpA has been recognized as a substrate for degradation by EcoClpAP under *in vivo* and *in vitro* conditions (Cranz-Mileva et al., 2008). Once the natural substrate is depleted, EcoClpA becomes the target and undergoes *in vitro* degradation in a time- and ATP-dependent manner, by EcoClpAP (Cranz-Mileva et al.,2008). Therefore, there was a strong interest in investigating whether LinClpA exhibits auto-degradation in the absence of substrates when interacting with the LinClpP1P2 complex.

TH-3792_206106022

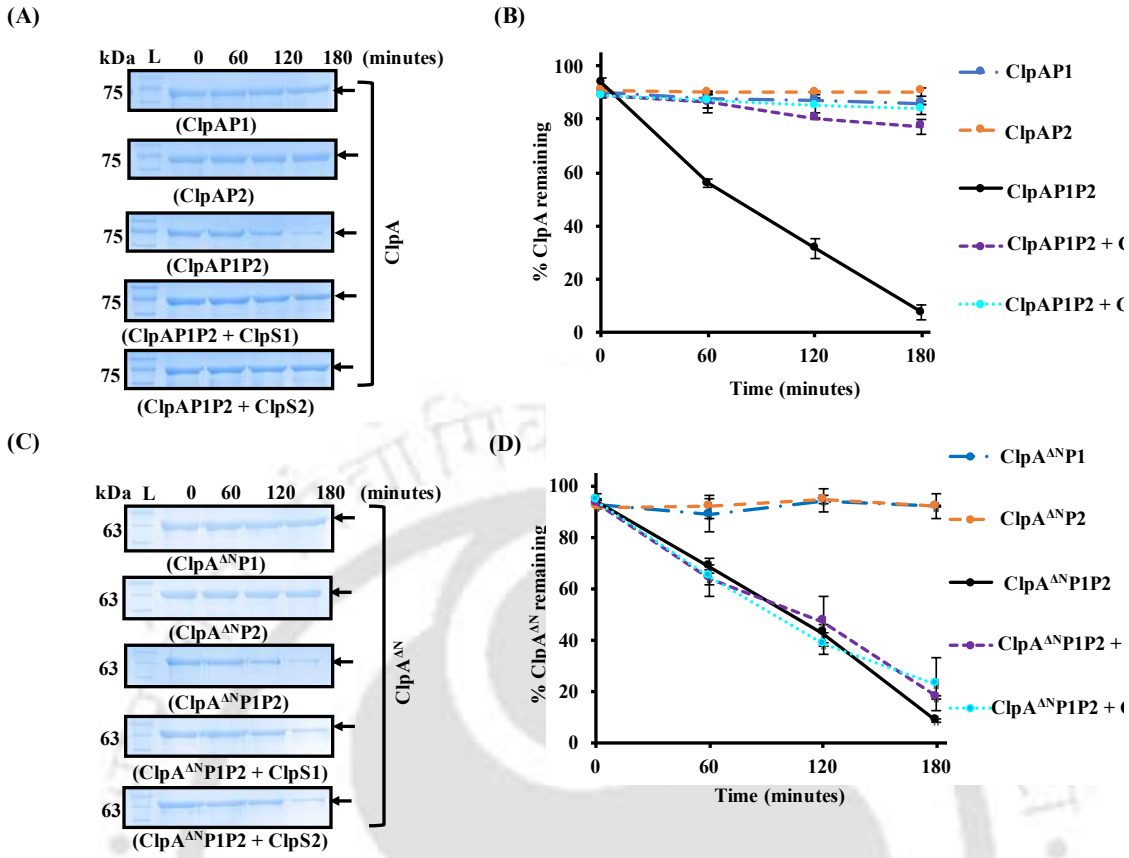


Figure 4.7. LinClpS inhibits LinClpP1P2-mediated auto-degradation exclusively to LinClpA, not its variant. **(A)** The LinClpA auto-degradation was checked using different self-assembled complexes (LinClpAP1, LinClpAP2, or LinClpAP1P2) in the presence of ATP for 180 minutes. The reaction was supplemented with LinClpS1 or LinClpS2, wherever required. **(B)** The remnant LinClpA was estimated by densitometry analysis using ImageLab software of the Coomassie-stained LinClpA bands in the SDS-PAGE gel. The percentage of LinClpA remaining at different time points of reaction was plotted. **(C)** The LinClpA$^{\Delta N}$ auto-degradation was checked using different self-assembled complexes (LinClpA$^{\Delta N}$P1, LinClpA$^{\Delta N}$P2, or LinClpA$^{\Delta N}$P1P2) in the presence of ATP for 180 min. The reaction was supplemented with LinClpS1 or LinClpS2, wherever required. **(D)** The remnant LinClpA$^{\Delta N}$ in the reaction was determined by densitometry analysis using ImageLab software of the Coomassie-stained LinClpA$^{\Delta N}$ bands.

The leptospiral Clp system contains two ClpP isoforms, LinClpP1 and LinClpP2, which necessitated the initial examination of LinClpA's auto-degradation using these pure LinClpP isoforms in their respective complexes (**Figure 4.7A**). As expected, the association of LinClpA with the pure LinClpP isoforms did not confer any activity to the complex. Subsequently, we assessed the time-dependent auto-degradation of LinClpA by the LinClpAP1P2 machinery. This reaction involved mixing the self-assembled LinClpP1P2 heterocomplex with ATP-bound LinClpA and incubating it for 180 minutes at 37 °C. At intervals of 0, 60, 120, and 180 minutes, 25 µL of the reaction mixture was resolved on polyacrylamide gel and stained with Coomassie blue. The LinClpAP1P2 complex demonstrated a time-dependent auto-degradation of LinClpA

TH-3792_206106022

on the polyacrylamide gel. Densitometry analysis of the LinClpA protein band, depicted in **Figure 4.7A**, was conducted at various time points and represented as a percentage of LinClpA remaining over time (**Figure 4.7B**). A 90% reduction in the band intensity of LinClpA was observed after 180 minutes. These findings prompted further investigation into the effects of adaptor proteins (LinClpS1 and LinClpS2) on LinClpAP1P2-mediated auto-degradation of LinClpA. Consistent with previous studies on EcoClpA, both LinClpS1 and LinClpS2 inhibited the auto-degradation of LinClpA mediated by the LinClpAP1P2 complex (**Figure 4.7A**). It was noted that the adaptor proteins remained stable throughout the experiment, suggesting that they do not directly compete with the ATPase chaperone to inhibit LinClpA's auto-degradation.

Interestingly, LinClpA$^{\Delta N}$ also displayed time-dependent auto-degradation by the LinClpA$^{\Delta N}$P1P2 complex (**Figure 4.7C – 4.7D**). This indicates that the N-domain of the ATPase chaperone has an insignificant role in modulating the protease activity of the core ClpP protease. However, the auto-degradation of LinClpA$^{\Delta N}$ was not inhibited in the presence of LinClpS by the LinClpA$^{\Delta N}$P1P2 complex. Therefore, it can be concluded that the binding of LinClpS to the N-domain of LinClpA is necessary to prevent auto-degradation.

4.4.7. The *ssrA* gene of *Leptospira,* encoding a c-degron, mediates the rapid degradation of eGFP-SsrA substrate

The substrate specificity of the leptospiral Clp system and its associated mechanisms remains unexplored. The ClpAP system of *E. coli*, which has been extensively studied, is known to be involved in degrading substrates with specific degradation tags located at the amino-terminal (N-degron tag) or carboxy-terminal (SsrA-tag) (Tobias et al., 1992; Keiler et al., 1996). Proteins starting with primary destabilizing residues (N-degrons; Phe, Leu, Trp, or Tyr) are recognized by the binding pocket on the core region of the EcoClpS adaptor. This adaptor delivers the proteins to EcoClpAP for degradation (Tobias et al., 1990). In *E. coli*, proteins starting with N-terminal Arg or Lys residues are recognized by leucyl/phenylalanyl-tRNA-protein transferase (LFTR transferase), which conjugates the protein with N-terminal Leu or Phe residue for recognition by EcoClpS (Ninnis et al., 2009). Another primary destabilizing residue transferase, Aspartate/glutamate leucyltransferase (Bpt), has been reported in *Vibrio vulnificus*. This Bpt transferase conjugates Leu residue to proteins beginning with Asp or Glu (Graciet et al., 2006). A preliminary *Leptospira* genome analysis suggested genes encoding for both LFTR (*LIC10096*; 660 bp) and Bpt (*LIC11930*; 774 bp) transferases are present. The

TH-3792_206106022

functionality of these genes encoding transferases for protein homeostasis in spirochetes is warranted to be explored in future studies.

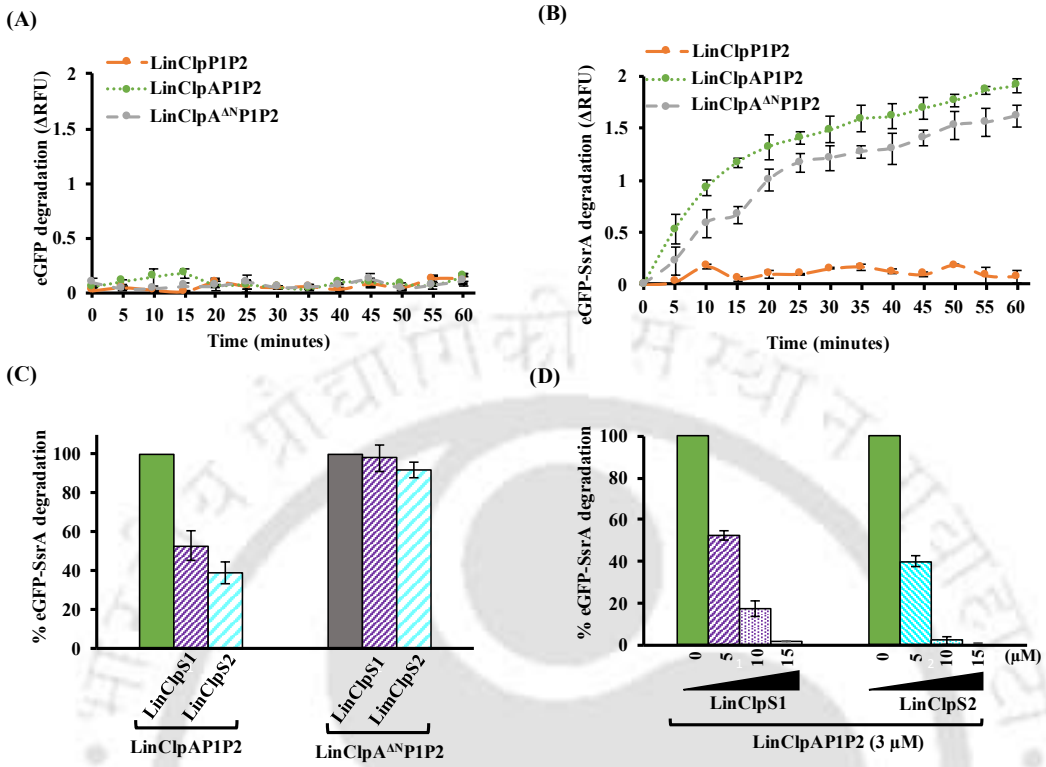


Figure 4.8. Proteolysis of the model substrate (eGFP) tagged with predicted leptospiral SsrA-tag. The degradation of **(A)** eGFP and **(B)** eGFP-SsrA was monitored by measuring fluorescence intensity at excitation and emission wavelengths of 485 nm and 510 nm, respectively, in the presence of various self-assembled complexes (LinClpP1P2, LinClpAP1P2, and LinClpA$^{\Delta N}$P1P2). The eGFP or eGFP-SsrA fluorescence intensity change (ΔRFU) was calculated using $RFU_{INITIAL} - RFU_{FINAL}$ and plotted with time. **(C)** The LinClpAP1P2 or LinClpA$^{\Delta N}$P1P2-mediated degradation of eGFP-SsrA was measured after 60 min in the presence of either LinClpS1 or LinClpS2. **(D)** The LinClpAP1P2 proteolytic activity on eGFP-SsrA substrate in the presence of increasing concentrations (5, 10, 15 μM) of LinClpS proteins was measured after 60 min. The change in the degradation rate of eGFP-SsrA with LinClpS1 or LinClpS2 addition was calculated considering LinClpAP1P2 and LinClpA$^{\Delta N}$P1P2 machinery activity as 100%.

The SsrA-tag is utilized to rescue stalled ribosomes during protein synthesis in prokaryotes, particularly in the presence of premature mRNA (Moore & Sauer, 2007). The stalled ribosome is recognized by a hybrid RNA called tmRNA (transfer-messenger RNA) encoded by the *ssrA* gene in association with the SmpB protein. This recognition facilitates the completion of translation and subsequent ribosome release (Moore & Sauer, 2007). The resulting polypeptide, marked with a C-terminal SsrA tag, is targeted for degradation by ClpAP and ClpXP in *E. coli* or ClpXP in *B. subtilis* (Gottesman et al., 1998; Weigert et al., 2001). The genome of *L. interrogans* serovar Copenhageni is predicted to contain the *ssrA* gene (348 bp) and its associated *smpB* gene (*LIC12418*; 480 bp). While the sequences of C-terminal proteolytic tags

TH-3792_206106022

encoded by the *ssrA* gene from various bacterial species have been listed in the tmRNA database, their potential in targeting polypeptides for proteolysis remains to be evaluated (Gottesman et al., 1990). The functionality of the SsrA-tag has been extensively studied and established in selected bacteria such as *E. coli*, *B. subtilis*, *S. aureus*, S. *pneumoniae*, and *M. tuberculosis* (Keiler et al., 1991; Wiegert et al., 2001; Donegan et al., 2014; Ahlawat et al., 2009; Marsee et al., 2018). This study compared the SsrA-tag sequences of several bacterial species with well-characterized ClpP systems (**Table 4.2**). The length of the SsrA-tag varies significantly among organisms, ranging from 9 residues (*L. interrogans*, *M. tuberculosis*, and *C. difficile*) to 22 residues (*C. trachomatis*) (Knudsen et al., 2001). WebLogo representation was used to compare the sequence of the last 9 residues of the SsrA-tag from the listed bacteria (**Figure 4S6a**) (Crooks et al., 2004). Notably, the last two Ala residues are conserved across all sequences and are crucial for docking onto the EcoClpX ATPase chaperone (Flynn et al.,2001). Additionally, the critical residues for recognition by EcoClpA in the 10-amino acid-long SsrA-tag of *E. coli* were reported to be Ala-8, Leu-9, and Ala-10 (Flynn et al., 2001). The last four residues of the leptospiral SsrA sequence (ANNELALAA), which resemble those in *E. coli* (ANDENYALAA), may be involved in the association of SsrA-tagged substrates with LinClpA. To study the role of the leptospiral ClpAP1P2 machinery in degrading proteins tagged with SsrA, we created an eGFP-SsrA model substrate. This was achieved by inserting the DNA sequence encoding the predicted amino acid sequence (ANNELALAA) of SsrA from *Leptospira* into the pET21d-eGFP plasmid. This modification allows for the expression of the eGFP protein, which is tagged with SsrA at its C-terminus. Subsequently, the recombinant eGFP and eGFP-SsrA with N-terminal 6×His-tags were purified by Ni-NTA affinity chromatography.

The *in-vitro* degradation of the eGFP and eGFP-SsrA substrates was assessed by LinClpP1P2 heterocomplex alone or in association with ATP-bound LinClpA or LinClpA$^{\Delta N}$ (LinClpAP1P2 and LinClpA$^{\Delta N}$P1P2 machinery) (**Figure 4.8A - 4.8B**). The incubation of the eGFP-SsrA substrate with the LinClpAP1P2 machinery resulted in a noticeable decrease in fluorescence over time, while the eGFP retained stable fluorescence intensity. In *E. coli*, in-vitro studies have demonstrated that the auto-degradation of EcoClpA by the EcoClpAP complex is inhibited in the presence of SsrA-tagged model substrate (Dougan et al., 2002). In a similar approach, we conducted a degradation assay using the LinClpAP1P2 complex with eGFP-SsrA as a substrate, and analyzed the reaction products by SDS-PAGE (**Figure 4S6b**). A time-dependent reduction in the intensity of the eGFP-SsrA band was observed, confirming the proteolytic activity of the LinClpAP1P2 machinery. Notably, LinClpA auto-degradation was

TH-3792_206106022

suppressed in the presence of the LinSsrA-tagged substrate, indicating substrate-dependent protection from self-degradation.

Table 4.2. Comparison of the SsrA-tag of *Leptospira* and other pathogenic bacteria

Organism	SsrA-tag (total number of residues)
A. tumefaciens	ANDNNAKEYALAA (13)
B. subtilis	GKTNSFNQNVALAA (14)
C. difficile	ADDNFAIAA (9)
C. trachomatis	AEPKAECEIISFADLEDLRVAA (22)
E. coli	ANDENYALAA (10)
H. influenza	ANDEQYALAA (10)
H. pylori	VNNTDYAPAYAKAA (14)
L. interrogans	ANNELALAA (9)
L. monocytogenes	GKEKQNLAFAA (11)
M. tuberculosis	ADSHQRLAA (9)
N. meningitidis	ANDETYALAA (10)
P. aeruginosa	ANDDNYALAA (10)
S. aureus	GKSNNNFAVAA (11)
S. mutans	AKNTNSYAVAA (11)
S. pneumoniae	AKNNTSYALAA (11)
S. pyogenes	AKNTNSYALAA (11)
S. typhimurium	ANDETYALAA (10)
X. campestris	ANDDNYGSDFAIAA (14)

Furthermore, the association of ATP-bound LinClpA$^{\Delta N}$ with the LinClpP1P2 heterocomplex supported the degradation of eGFP-SsrA, albeit with 15 % lower efficiency than the LinClpAP1P2 machinery after 60 minutes (**Figure 4.8B**). In *E. coli*, the EcoClpS has been reported to redirect the specificity of the EcoClpAP protease complex away from SsrA-tagged substrates by significantly inhibiting their degradation (Dougan et al., 2002). Therefore, we conducted a comparison of the degradation of the eGFP-SsrA substrate by the LinClpAP1P2 machinery (3 µM) in the absence and presence of adaptor proteins (LinClpS1 and LinClpS2; 5 µM) (**Figure 4.8C**). As expected, the presence of LinClpS1 and LinClpS2 reduced the degradation of the eGFP-SsrA substrate to 53% and 39%, respectively compared to LinClpAP1P2 machinery (100%). To achieve complete inhibition of LinClpAP1P2 activity, increasing concentrations of LinClpS1 and LinClpS2 (5, 10, and 15 µM) were added to the reaction mixture and the degradation rate of eGFP-SsrA was subsequently monitored (**Figure 4.8D**). Notably, the addition of LinClpS1 (15 µM) and LinClpS2 (10 µM) led to a complete suppression of the proteolytic activity of the LinClpAP1P2 complex against the SsrA-tagged model substrate. To investigate how the LinClpA$^{\Delta N}$ would influence this inhibitory effect of adaptor proteins, the degradation of eGFP-SsrA by LinClpA$^{\Delta N}$P1P2 in the absence and

TH-3792_206106022

presence of adaptor proteins was measured. In contrast to LinClpA, the LinClpA$^{\Delta N}$, in association with the LinClpP1P2 heterocomplex, showed full activity on eGFP-SsrA measured after 60 minutes of degradation at 37°C in the presence of LinClpS1 and LinClpS2 (**Figure 4.8C**).

To further validate the functionality of the determined SsrA-tag of *Leptospira*, another model protein substrate was generated by adding the 9-residue (ANNELALAA) SsrA-tag at the C-terminal of an outer membrane leptospiral protein (LIC13314) previously characterized in our laboratory (Gosh et al., 2018). The recombinant LIC13341-SsrA substrate was purified, and its time-dependent degradation was analyzed using ATP-bound LinClpAP1P2 machinery. Interestingly, the LIC13341-SsrA tagged substrate remained stable over 180 min when visualized on SDS-PAGE by Coomassie staining (**Figure 4S6c**). When the model substrate (LIC13341-SsrA) was not recognized by the LinClpAP1P2 machinery, it led to the auto-degradation of LinClpA in a time-dependent manner (**Figure 4S6c**). This study indicates that simply tagging any protein with the cognate SsrA tag is not enough for it to be recognized by the Clp chaperone-protease complex. Additional specificity factors within the substrates are likely necessary and should be thoroughly validated in future research.

4.4.8. Investigation of LinClpAP1P2-mediated degradation of peptides bearing the N-degron (Y) tag

To assess the functionality of the N-degron (Y) tag, we designed a synthetic peptide consisting of 10 amino acids with the sequence YLFVQHHHHC, which was labeled at the C-terminal end with a fluorescent tag, AMC. We monitored the degradation of this peptide by measuring the release of free AMC, which fluoresces upon cleavage. In our experimental setup, we incubated the LinClpAP1P2 proteolytic machinery with the YLFVQHHHHC-AMC peptide and tracked its degradation over time. Each 100 µL reaction mixture contained the LinClpAP1P2 complex (3 µM), supplemented with ATP (5 mM) and 100 µM of the Y-Pep substrate (YLFVQHHHHC-AMC). The mixture was incubated for 1 hour at 37 °C in an assay buffer consisting of 50 mM Tris-Cl, 50 mM KCl, 1 mM DTT, and 8 mM MgCl2 at pH 7.8. Surprisingly, we observed no increase in AMC fluorescence, indicating that the YLFVQHHHHC-AMC peptide was not degraded by the LinClpP1P2 or LinClpAP1P2 complex under the tested conditions (**Figure 4.9**). To explore this further, we added the adaptor protein LinClpS2 (5 µM), which features a more conserved N-degron recognition site. However, even with the addition of LinClpS2, no degradation was detected.

TH-3792_206106022

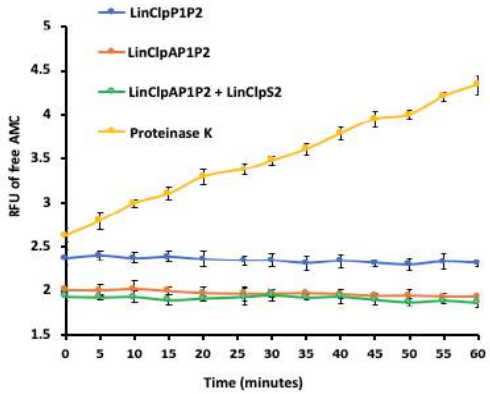


Figure 4.9. N-degron peptide degradation assays using LinClpP1P2 complex. Peptidase activity of LinClpP1P2 alone or in association with LinClpA chaperone and LinClpS2 adaptor protein on the N-degron tagged fluorogenic peptide (YLFVQHHHHC-AMC). Peptide degradation was measured fluorometrically as a relative fluorescent unit (RFU × 1000) using the fluorogenic substrates for 1 h. The serine protease, proteinase K, used as a positive control.

To verify that the AMC fluorophore and peptide were both intact and functional, we treated the peptide with proteinase K, a broad-spectrum serine protease commonly used in molecular biology. This enzyme cleaves peptide bonds adjacent to aliphatic, hydrophobic, and aromatic amino acids. As anticipated, treatment with proteinase K (3 μM) led to a time-dependent increase in AMC fluorescence, confirming the structural and functional integrity of the peptide and fluorophore. Our initial study, which utilized the YLFVQHHHHC-AMC peptide cleavage assay, suggested that the LinClpAP1P2 machinery might cleave the peptide, but this cleavage may not result in the release of free AMC fluorophore, explaining the absence of any observable increase in fluorescence. In contrast, when proteinase K cleaves the peptide at the C-terminal end, a distinct fluorescence signal from the released AMC is detected. This indicates that, in our current setup, we cannot reliably assess the functionality of the N-degron using the YLFVQHHHHC-AMC peptide and the LinClpAP1P2 system since the position of the cleavage hinders fluorophore release. To address this limitation, we can design model substrates where the positioning of the cleavage site does not interfere with detection. Specifically, we can generate GFP-based substrates tagged with various peptide sequences, such as YLFVQ, FLFVQ, WLFVQ, and LLFVQ, at the N-terminal end of the GFP protein (Tobias et al., 1990). These N-degron-tagged GFP constructs can then be utilized to monitor degradation by the LinClpAP1P2 machinery. Employing this strategy, future studies will aim to explore the functionality of N-degrons in *Leptospira* more effectively.

TH-3792_206106022

4.5. Discussion

The Clp family of proteins is widely present across different life forms and plays a crucial role in the general stress response and the virulence of many pathogenic bacteria (Lourdault et al., 2011). In the genome of *L. interrogans*, several typical Clp family proteins have been predicted, including ClpS, ClpA, ClpX, ClpP, ClpQ, and ClpY. The *Leptospira* species possess a diderm structure, including an outer membrane containing lipopolysaccharide (LPS), linking them to Gram-negative bacteria (Lourdault et al., 2011). While most Gram-negative bacteria have the class I ATPase chaperone ClpA but lack ClpC chaperone, the *Leptospira* species contain both ClpA and ClpC ATPase chaperones (Lourdault et al., 2011). The functionality of ClpC chaperone and its association with Clp protease have been well-described in organisms such as *S. aureus*, *B. subtilis*, *M. tuberculosis*, and *L. interrogans* (Donegan et al., 2014; Taylor et al., 2022; Mei et al., 2009; Kumari et al., 2024). However, the detailed biochemical characterization of ClpA in complex with Clp protease and adaptor proteins has only been limited to *E. coli* (Maurizi et al., 1998).

The assembly of the Clp ATPase chaperone into hexamers is essential for its functionality (Maurizi et al., 1991). Multiple studies have investigated nucleotide-induced oligomerization of EcoClpA, yielding inconsistent findings. According to a study by Maurizi and colleagues, the hexameric assembly of EcoClpA is contingent solely upon the binding of adenosine triphosphate (ATP) (Maurizi et al., 1991). Conversely, another study suggested that EcoClpA oligomerization is not exclusively dependent on the presence of the γ phosphate of ATP. The presence of ATPγS, AMP-PNP, AMP-PCP, ADP·BeF, and ADP all facilitates the formation of a hexameric EcoClpA (Veronese et al., 2011). Similarly, the oligomerization of LinClpA was not solely restricted to the presence of nucleotide triphosphates. Adenosine diphosphate (ADP) also promotes LinClpA oligomerization, albeit to a lesser extent than ATP. Intriguingly, this study recorded a shift in the oligomerization propensity of LinClpA in the presence of polyphosphate (PolyP). In bacterial cells, the accumulation of inorganic polyP occurs during oxidative stress or amino acid starvation (Kuroda et al., 2001). In certain bacteria, such as *C. crescentus* and *P. aeruginosa*, polyP regulates various cell cycle events during stress conditions (Ropelewska et al., 2020). In *E. coli*, it has been reported that polyP interacts with the Lon protease via its ATPase domain, thereby regulating its activity (Kuroda et al., 2001). However, the effect of polyP on other bacterial cell proteases has not been explored. While this study provided indications of the binding of polyP with the ATPase chaperone LinClpA and its promotion of oligomerization, a more comprehensive analysis is required to grasp the significance of polyP in leptospiral cell survival under stress conditions.

TH-3792_206106022

Biochemical analysis of LinClpA displayed optimal ATPase activity at pH 7 and a temperature of 25 ºC in the presence of specific divalent metals (Mg^{2+}, Mn^{2+}, and Co^{2+}), with the highest activity observed with Mg^{2+}. In contrast, both EcoClpA and EcoClpB exhibit ATPase activity in the presence of Mg^{2+}, Mn^{2+}, and Ca^{2+}, while SmuClpL and SpnClpL demonstrate ATPase activity with Mg^{2+} and Mn^{2+} only (Hwang et al., 1988; Woo et al., 1992; Jana et al., 2020; Park et al., 2015). Furthermore, like EcoClpB, LinClpC, and SpnClpL, LinClpA hydrolyzes various nucleotides (ATP, GTP, CTP, and UTP) (Woo et al., 1992; Park et al., 2015; Kumari et al., 2024). However, the phosphate hydrolysis activity of EcoClpA is limited to ATP only (Hwang et al., 1988). The *in-silico* analysis of LinClpA indicates the conservation of crucial elements, including the Walker A and Walker B motifs, Arginine finger, Sensor I, and Sensor II motifs, underscoring their significance in Clp ATPase activity. A comparison of the amino-acid sequence of the N-domain of LinClpA with EcoClpA reveals an approximate 40% sequence identity. Additionally, the tertiary structure analysis of the modeled LinClpA N-domain using AlphaFold demonstrates a high degree of structural similarity with the crystal structure of the EcoClpA N-domain.

In the context of *E. coli*, there exists substantial variation in studies characterizing different N-domain deleted EcoClpA variants. The EcoClpA N-domain comprises 143 residues with an attached linker region of 24 residues, followed by two ATPase domains (Guo et al., 2002). One study focuses on an EcoClpA variant ($EcoClpA^{\Delta 168}$) lacking 1-168 residues from the N-terminal, which exhibited a 20-fold reduction in ATPase activity. However, the complex formed with EcoClpP ($EcoClpA^{\Delta 168}P$) was inactive in degrading the casein model substrate (Seol et al., 1995). Similarly, another variant of EcoClpA ($EcoClpA^{\Delta 161}$) recorded a 50% reduction in ATPase activity; nevertheless, the $EcoClpA^{\Delta 161}P$ complex degraded casein and SsrA-tagged substrates at a 20-30% reduced rate (Lo et al., 2001). Conversely, another $EcoClpA^{\Delta 153}$ variant exhibited higher ATPase activity and a higher rate of degradation of casein and GFP-SsrA substrates mediated by the $EcoClpA^{\Delta 153}P$ complex (Xia et al., 2004). Further, the N-terminal domain deleted variant $EcoClpA^{\Delta 143}$ retained ATPase activity as EcoClpA but compromised the $EcoClpA^{\Delta 143}P$ complex protease activity. The $EcoClpA^{\Delta 143}P$ complex exhibited a slower rate of degradation of GFP-SsrA than EcoClpAP (Hinnerwisch et al., 2005).

A recent study showed that the deletion of the N-domain and complete linker region ($EcoClpA^{\Delta N+\Delta L}$) did not alter its ATPase activity but increased the catalytic efficiency of the complex $EcoClpA^{\Delta N+\Delta L}P$ in degrading the cognate GFP-SsrA substrate (Cranz-Mileva et al., 2008). On the other hand, in another variant, $EcoClpA^{\Delta L}$, where only the linker region was

TH-3792_206106022

deleted, the ATPase activity of the chaperone and the protease degradation rate of GFP-SsrA mediated by the complex EcoClpA$^{\Delta L}$P was retained (Cranz-Mileva et al., 2008).

In this study, a deletion variant of LinClpA has been generated where the predicted N-domain (1-138 residues) has been deleted, and thus, the LinClpA$^{\Delta N}$ comprises a linker region (22 residues) and the 2 ATPase domains. Similar to earlier reported variants of EcoClpA (ClpA$^{\Delta N+\Delta L}$ and ClpA$^{\Delta 143}$), the LinClpA$^{\Delta N}$ efficiently assembled into higher-order structures in the presence of ATP (Cranz-Mileva et al., 2008; Hinnerwisch et al., 2005). This indicates that the N-domain has no role in ATP-induced oligomerization of ATPase proteins. Moreover, the LinClpA$^{\Delta N}$ exhibits a higher ATPase activity, which agrees with the EcoClpA$^{\Delta 153}$ variant study (Xia et al., 2004). However, other reported EcoClpA N-domain variants exhibit either similar (ClpA$^{\Delta N+\Delta L}$ and ClpA$^{\Delta L}$) or reduced (ClpA$^{\Delta 168}$ and ClpA$^{\Delta 161}$) ATPase activity to wild-type EcoClpA (Cranz-Mileva et al., 2008; Seol et al., 1994; Lo et al., 2001).

Next, the studies concerning the association of EcoClpP with N-domain variants of EcoClpA show discrepancies. The association between EcoClpA$^{\Delta 143}$ and EcoClpP is reported to be less stable than the EcoClpAP association, as the addition of EcoClpP to EcoClpA$^{\Delta 143}$ failed to stimulate its ATPase activity (Hinnerwisch et al., 2005). However, another study shows that the EcoClpA$^{\Delta N+\Delta L}$, EcoClpA$^{\Delta L}$, and EcoClpA show a similar ability to associate with EcoClpP (Cranz-Mileva et al., 2008). To our dismay, the LinClpA or its deletion variant (LinClpA$^{\Delta N}$) shows a similar affinity for interaction with LinClpP isoforms (ClpP1 and ClpP2) in this study. Thus, it can be ascertained that the deletion of the LinClpA N-domain does not influence the ability of Clp Chaperone-Protease complex formation.

The structural superimposition of modeled LinClpA with the crystal structure of EcoClpA (PDB: 1KSF) shows that the N-domain residues which are reported to be important for binding with the adaptor proteins in EcoClpA N-domain (Thr81, Glu23, Glu28, and Arg86) are conserved in LinClpAN-domain (Thr76, Glu18, Glu23, and His81) (Zeth et al., 2002). The adaptor proteins serve as substrate recognition components and deliver the recognized substrate to their cognate ATPase chaperone (Dougan et al., 2002). The EcoClpS adaptor exhibits dual functionality. EcoClpS acts as either an efficient stimulator for the degradation of N-degron substrates or as an inhibitor for the degradation of C-degron (SsrA-tagged) substrates by EcoClpAP (Dougan et al., 2002). The core region of EcoClpS binds to the N-domain of EcoClpA via a salt bridge and hydrogen bonds, while the N-terminal extension (NTE) region of EcoClpS engages with the EcoClpA ATPase domain. The EcoClpS NTE engagement with the EcoClpA translocation channel suppresses its ATPase activity (Torres-Delgado et al., 2020). Without EcoClpS, the EcoClpAP efficiently degrades SsrA-tagged substrates more than

TH-3792_206106022

N-degron substrates. When EcoClpS is present, its core region has hydrophobic binding pockets for N-degrons, facilitating enhanced delivery of N-degron substrates to the EcoClpAP machinery. Also, EcoClpS decreases the affinity of SsrA-tagged to the EcoClpA by a non-competitive binding mechanism (Torres-Delgado et al., 2020). Thus, the EcoClpS binding to the EcoClpA switches the substrate specificity of the complex EcoClpAP from SsrA-tagged substrates to N-degron substrates (Torres-Delgado et al., 2020).

The Clp system of *Leptospira* possesses two adaptor proteins (LinClpS1 and LinClpS2), which share 17 % sequence identity among themselves while having 23% and 42% sequence identity with EcoClpS, respectively. The LinClpS residues predicted to be important for the association with LinClpA N-domain are more conserved in LinClpS2 than in LinClpS1. These variations within the chaperone binding pocket of LinClpS1 and LinClpS2 might hint towards their different affinity for the LinClpA N-domain. The determined K_d for the LinClpAS2 complex (0.32 µM) was similar to the reported (Zeth et al., 2002) K_d for the EcoClpAS complex (0.33 µM), while a higher K_d (0.44 µM) was observed for the LinClpAS1 complex. Further, a dose-dependent inhibition of ATPase activity in LinClpA was observed in the presence of LinClpS1 and LinClpS2. Although LinClpS1 forms a less stable complex with LinClpA, its presence (10 µM) inhibits 57 % ATPase activity in LinClpA. The stable complex formation between LinClpA and LinClpS2 results in more than 90% ATPase activity inhibition in LinClpA. Moreover, no inhibitory effect of either LinClpS1 or LinClpS2 on the ATPase activity of variant LinClpA$^{\Delta N}$ was observed, implying that LinClpS can only interact with LinClpA with an intact N-domain.

Apart from mediating the degradation of SsrA-tagged or N-degron protein substrates, the EcoClpA serves as a substrate and undergoes proteolytic cleavage mediated by EcoClpAP. The EcoClpA was reported to undergo such auto-degradation to regulate the levels of EcoClpA protein *in-vivo* (Gottesman et al., 1990). The EcoClpA undergoes EcoClpAP-mediated auto-degradation *in-vitro* with a half-life of 15 minutes, and the binding of EcoClpS prevents the auto-degradation process (Dougan et al., 2002). In line with earlier findings in *E. coli*, the LinClpA also undergoes auto-degradation by LinClpAP1P2 machinery under *in-vitro* conditions with a half-life of approximately 60 minutes. As contemplated, in the presence of adaptor proteins (LinClpS1 and LinClpS2), the auto-degradation of LinClpA by the protease complex is inhibited.

Additionally, the functionality of the predicted SsrA-tag sequence (ANNELALAA) in *Leptospira* was explored in this study. The LinClpAP1P2 machinery degraded the cognate eGFP-SsrA protein substrate under *in-vitro* conditions. The model eGFP protein lacking

TH-3792_206106022

LinSsrA-tag remained stable for 60 minutes of reaction time under similar conditions with LinClpAP1P2 machinery and agrees with earlier reported studies on EcoClpAP and GFP (Weber-Ban et al., 1999). Also, LinClpS binding to the N-domain of LinClpA inhibits SsrA-tagged protein degradation. In contrast, adding LinClpS to LinClpA$^{\Delta N}$P1P2 machinery did not affect the degradation rate of cognate eGFP-SsrA. Thus, the N-domain of LinClpA is considered indispensable for adaptor protein-mediated regulation of LinClpAP1P2 activity. A recent study reports that GFP protein contains nine unstructured amino acids (THGMDELYK) at its C-terminal, and the insertion of *E. coli* SsrA tag (AANDENYALAA; 11 aa) extends the unstructured region to a total of 20 residues at the C-terminus and thus renders it as a model substrate for EcoClpAP (Shih et al., 2024). However, a variant of GFP (lacking the nine unstructured C-terminal aa) tagged with EcoSsrA remained resistant to degradation by EcoClpAP. This resistance is likely due to the inability of the 11-residue-long EcoSsrA tag to access the binding sites within the axial channel of EcoClpA (Shih et al., 2024). In our study, the C-terminal LinSsrA tagging of the protein LIC13341 does not support its degradation by LinClpAP1P2 complex. The inability of the LIC13341-SsrA protein to be degraded by the LinClpAP1P2 complex might be due to limited accessibility of the 10-aa long LinSsrA tag to the key binding sites within the axial channel of LinClpA and it requires an additional unstructured region at the C-terminal end of protein substrates. In summary, this study reports the functional characterization of the molecular chaperone, LinClpA, and modulation of its activity upon binding with adaptor proteins (LinClpS1 and LinClpS2) and protease isoforms (LinClpP1 and LinClpP2).

TH-3792_206106022

Chapter 5

Conclusions and future prospects

TH-3792_206106022

5.1. Conclusions

This comprehensive study provides a critical, detailed analysis of the structural and regulatory mechanisms governing the Clp protease system in the pathogen *Leptospira interrogans*. Bacterial caseinolytic protease (Clp) chaperone-protease complexes are vital for the degradation of misfolded and aggregated protein substrates, thereby ensuring protein homeostasis. Typically, these complexes consist of a core tetradecameric peptidase made up of two heptameric rings stacked together, along with a hexameric ring-shaped ATPase chaperone. The protein degradation process relies on a precise recognition mechanism to safeguard essential and correctly folded proteins within the bacteria. The spirochete *L. interrogans* has a set of genes that encode adaptor proteins (*clpS1* and *clpS2*), as well as genes for Clp ATPase (*clpA*, *clpC*, *clpX*) and Clp protease (*clpP1* and *clpP2*). The core proteolytic component of the Clp system in leptospires comprises two distinct ClpP isoforms (LinClpP1 and LinClpP2), which together form a functional hetero-tetradecameric assembly known as LinClpP1P2. By investigating key LinClpP structural motifs and characterizing the associated chaperone LinClpC and LinClpA and adaptor proteins LinClpS1/S2, the research establishes a deeper understanding of this complex protein quality control pathway. Notably, the unique presence and function of the Gram-positive like ClpC chaperone and the regulatory role of the LinClpS adaptors shed new light on how leptospires manage C-degron substrate degradation. These findings are crucial for future therapeutic development, as they pinpoint specific Clp components as potential drug targets to disrupt bacterial survival and reduce reliance on indiscriminate antibiotics for treating leptospirosis.

Key Research Highlights

LinClpP Isoform Structure and Function

- Mutant Generation: Five mutants targeting critical hotspot residues were created across LinClpP1 and LinClpP2 isoforms to study structure-function relationships.
- Gain-of-Function Mutants: The LinClpP2 mutants S40AK41N and Y62A displayed increased protease activity (1.7-fold and 1.5-fold, respectively) when their heterocomplexes were bound to the antibiotic ADEP1.
- Loss-of-Function Mutant: The LinClpP2 deletion mutant (IG_del) maintained its tetradecameric structure but exhibited a significant loss of activity.

TH-3792_206106022

- Assembly Changes: The LinClpP1 oligomerization sensor domain mutant E170D and the catalytic triad mutant N172D both led to decreased activity in the LinClpP1P2 heterocomplex.

LinClpC Chaperone Characterization

- *L. interrogans* Anomaly: The bacterium contains the Gram-positive-like ClpC ATPase chaperone within its Clp array.
- Activity: LinClpC exhibits optimal ATPase activity at 45°C and pH 8 and undergoes nucleotide-induced oligomerization.
- Novel Activation: When associated with the LinClpP1P2 heterocomplex, LinClpC imparts protease activity on a casein substrate without any energy input.
- Enhancement: LinClpC's functional oligomerization and the resulting protease activity are further enhanced by the addition of ATP and the antibiotic ADEP1.

LinClpA and LinClpS Adaptor Function

- Adaptor Discovery: Two adaptor proteins, ClpS1 and ClpS2, were identified in the *Leptospira* genome.
- LinClpA$^{\Delta N}$ Activity: A mutant of the ClpA chaperone lacking the N-terminal domain (LinClpA$^{\Delta N}$) showed increased ATPase activity compared to the full-length protein.
- Regulatory Role of LinClpS: The LinClpS adaptor proteins suppress the time-dependent degradation of SsrA-tagged eGFP by the LinClpAP1P2 complex, indicating a regulatory role in C-degron substrate processing.
- Adaptor Interaction: The LinClpA$^{\Delta N}$ mutant, despite maintaining oligomerization capacity, showed a loss of significant interaction with LinClpS1 and LinClpS2.

The findings from this study advance our understanding of the physiological role of the Clp chaperone and its associated adaptor proteins in regulating the activity of the ClpP protease in *Leptospira*. Biochemical characterization of these Clp system-associated proteins may offer valuable insights for designing drugs that can modify or enhance the protein degradation process, potentially leading to the self-destruction of bacteria. By comprehending these functions, we could reduce the indiscriminate use of antibiotics for treating leptospirosis and develop new therapeutic agents that specifically target these Clp proteins. Thus, exploring the functional mechanisms of these protein quality control enzymes could deepen our

TH-3792_206106022

understanding of this bacterium's pathogenesis and illuminate pathways for potential drug development.

5.2. Future prospects

The findings of this study regarding various leptospiral Clp proteins, including ClpP isoforms, ClpA/ClpC chaperones, and ClpS adaptor proteins, provide a robust foundation for future research aimed at exploring their therapeutic applications. Recent reports have identified Clp chaperone-targeting peptide derivatives such as ecumicin, lassomycin, and Cyclomarin A as bactericidal agents effective against both replicative multidrug-resistant and dormant *Mycobacterium tuberculosis* (Gao et al., 2015; Gavrish et al., 2014; Schmitt et al., 2011). Lassomycin, an antimicrobial compound derived from the *Lentzea kentuckyensis* species, specifically targets the ClpC1 ATPase, a crucial enzyme in *Mycobacterium tuberculosis*. It binds to a highly acidic N-terminal region of the ClpC1 ATPase complex, significantly enhancing its ATPase activity (Gavrish et al., 2014).

The multiple sequence alignment conducted in this study demonstrates that leptospiral ClpC shares 50% sequence identity with mycobacterial ClpC1, suggesting that lassomycin might disrupt the activity of the LinClpC ATPase. Consequently, lassomycin presents a promising target for inhibiting the caseinolytic protease of *Leptospira* by interfering with the ClpC chaperone's function, which warrants further investigation in future studies.

Preliminary analysis of the *Leptospira* genome indicates the presence of a gene encoding an adaptor protein, designated mcsB (*LIC10340*; 786 bp), situated 149 base pairs downstream of the ATPase chaperone gene, LinclpC (*LIC10339*). The McsB protein functions as an arginine kinase that supports the oligomerization of *Bacillus subtilis* ClpC (BsuClpP) (Trentini et al., 2016). The proximity of the arginine kinase (LinMcsB) to LinClpC within the *Leptospira* genome suggests that LinMcsB may function as an adaptor protein analogous to its *B. subtilis* counterpart, warranting further investigation.

In the current study, it was observed that the model substrate casein does not activate LinClpC ATPase activity, which is consistent with earlier research regarding BsuClpC (Trentini et al., 2016). The substrate casein tagged with a phosphoarginine (pArg) generates casein-pArg, which stimulates BsuClpC activity (Trentini et al., 2016). The significance of the phosphorylation tag in the adaptor protein's role concerning substrate specificity within the leptospiral ClpP system can be explored to enhance our understanding of diverse regulatory proteolysis events.

TH-3792_206106022

Proteins designated for degradation typically contain specific motifs referred to as degrons, located at either the carboxy-terminal (C-degron) or amino-terminal (N-degron). The C-degron tagging system incorporates SsrA-tagged proteins created during the rescue of stalled ribosomes. Conversely, N-degrons involve tagging and destabilizing specific N-terminal residues, including phenylalanine (F), tyrosine (Y), tryptophan (W), and leucine (L). The present study provides evidence supporting the functionality of the leptospiral C-degron (SsrA tag) and demonstrates time-dependent degradation of SsrA-tagged model substrate (eGFP-SsrA) by the LinClpAP1P2 machinery. However, the N-end rule pathway in *Leptospira* remains unexamined. N-terminal destabilizing residues (F or L) are conjugated to proteins by the LFTR transferase enzyme present in *Escherichia coli* (Ninnis et al., 2009), while the Bpt transferase enzyme from *Vibrio vulnificus* conjugates the Leu residue to proteins (Graciet et al., 2006). Preliminary bioinformatics insights into the *Leptospira* genome indicate the existence of genes encoding both LFTR (LIC10096; 660 bp) and Bpt (LIC11930; 774 bp) transferases, which are implicated in the N-end rule pathway. Future investigations can assess the functionality of these transferase genes within the protein quality control system of *L. interrogans*. This may involve generating various eGFP protein variants by replacing the N-terminal residue with destabilizing residues (F, Y, W, and L) and evaluating their time-dependent degradation by the leptospiral ClpAP1P2 machinery.

In addition to the previously mentioned findings, two Class I ATPase chaperones, ClpA and ClpC, have been identified in *Leptospira*. ClpC chaperones are generally longer in amino acid sequence due to the presence of a middle domain (MD) that is inserted between the DI and DII ATPase domains. This middle domain is known to play a crucial role in the oligomerization of ClpC proteins. In our study, we identified a similar middle domain in LinClpC, and it would be valuable to investigate its functional significance through deletion analysis or site-directed mutagenesis of key residues within this domain in the future. Furthermore, our analyses suggest that LinClpC can associate with the LinClpS adaptor protein to modulate its ATPase activity, similar to the regulatory mechanism observed in LinClpA. Future studies could explore whether LinClpC is involved in canonical C-degron or N-degron mediated protein degradation pathways and whether this process is modulated by LinClpS.

In conclusion, this study provides novel insights into the interaction between Clp proteases and chaperones. In the context of the leptospiral hetero-tetradecameric ClpP1P2 complex, the ATPase chaperone was found to interact with both ClpP isoforms. In *L. interrogans*, both class I ATPase chaperones (LinClpA and LinClpC) can assemble with the LinClpP1P2

TH-3792_206106022

heterocomplex, thereby facilitating protease activity. It is anticipated that the LinClpAP1P2 and LinClpCP1P2 machineries may exhibit variations in substrate specificity, necessitating further exploration to achieve a comprehensive understanding of the protein quality control system in *Leptospira.*

TH-3792_206106022

Appendix

Supplementary data to Chapter 3

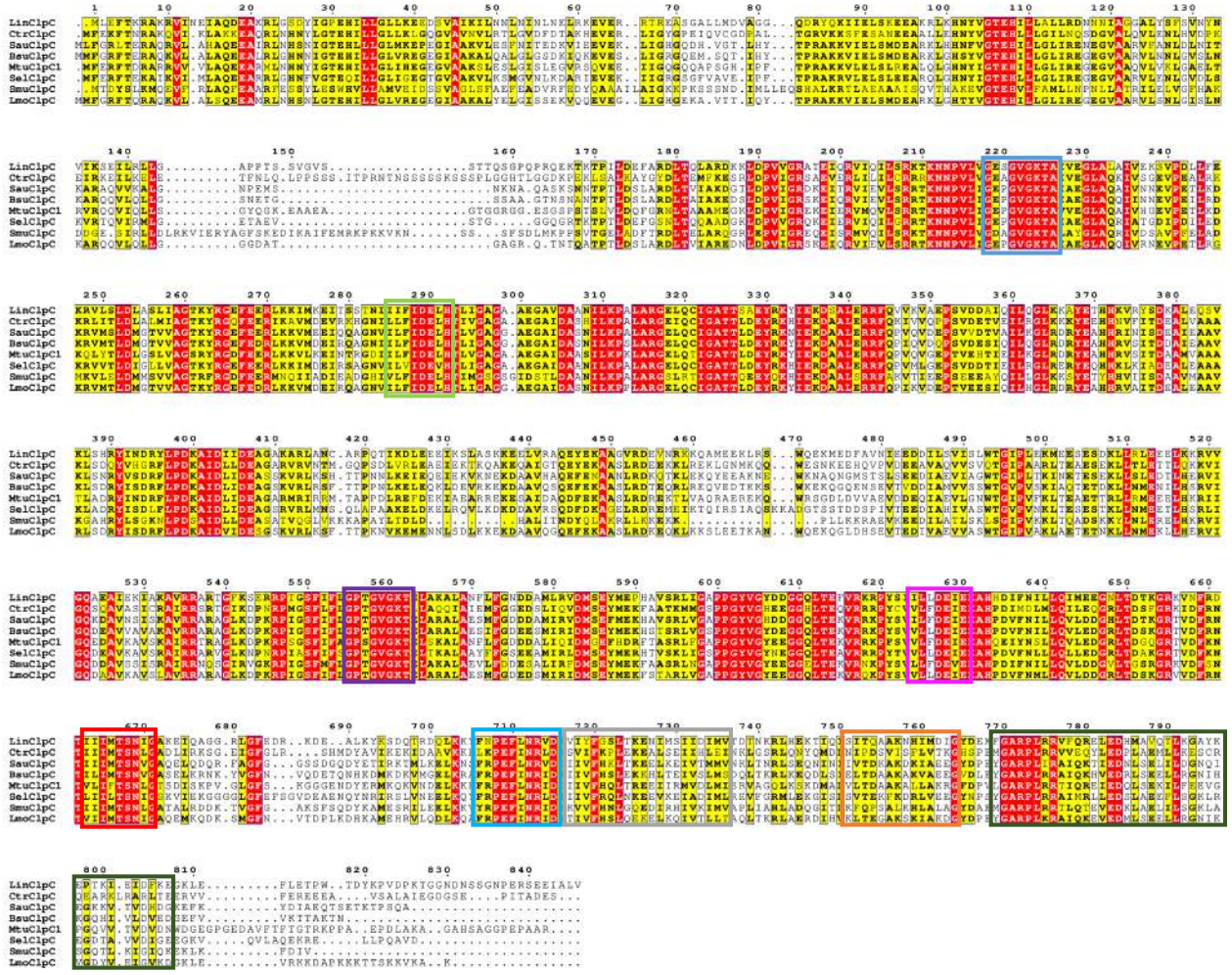


Figure 3S1. The conserved domains of *Leptospira* ClpC compared to bacterial orthologs. Multiple sequence alignment of ClpC protein sequence from *Leptospira interrogans*, *Chlamydia trachomatis*, *Staphylococcus aureus*, *Bacillus subtilis*, *Mycobacterium tuberculosis*, *Synechococcus elongatus*, *Streptococcus mutans*, and *Listeria monocytogenes*. The conserved functional domains of ClpC proteins are indicated in colored boxes: nucleotide-binding domain I (Walker A- Blue and Walker B- Green), nucleotide-binding domain II (Walker A-Violet and Walker B- Pink), C-terminal domain (Sensor I- Red, Box VII- Cyan, Box VII'- Grey, Box VII"- Orange and Sensor II- Dark green).

TH-3792_206106022

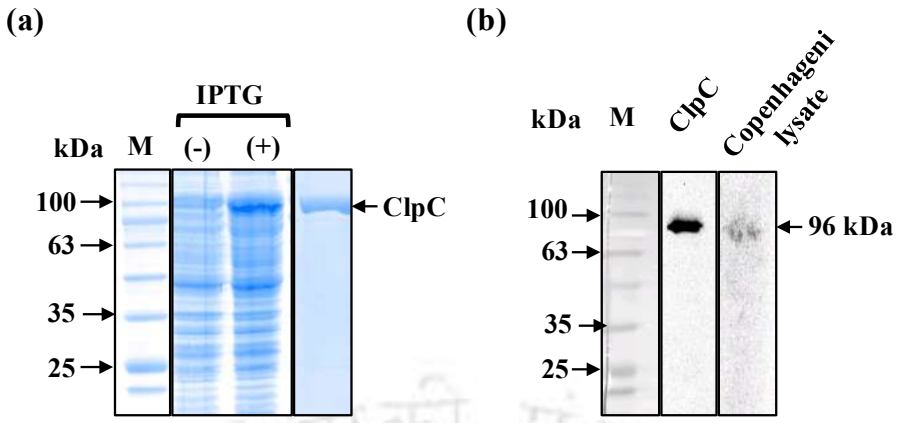


Figure 3S2. Molecular characterization of *L. interrogans* ClpC (LinClpC). (a) Over-expression and purification of LinClpC in *E. coli* BL 21 cells. The cell lysates after induction without (-) or (+) with IPTG (1 mM) and the Ni-NTA affinity purified recombinant LinClpC resolved on 12 % SDS-PAGE stained with Coomassie blue. M denotes the standard protein molecular size marker (kDa). **(b)** Immunoblot analysis of *L. interrogans* serovar Copenhageni lysate. The expression of native ClpC was detected in *Leptospira* using mouse anti-LinClpC (1:1000 dilution). The recombinant LinClpC and native ClpC of *Leptospira* were detected at a similar molecular size (~ 96 kDa).

TH-3792_206106022

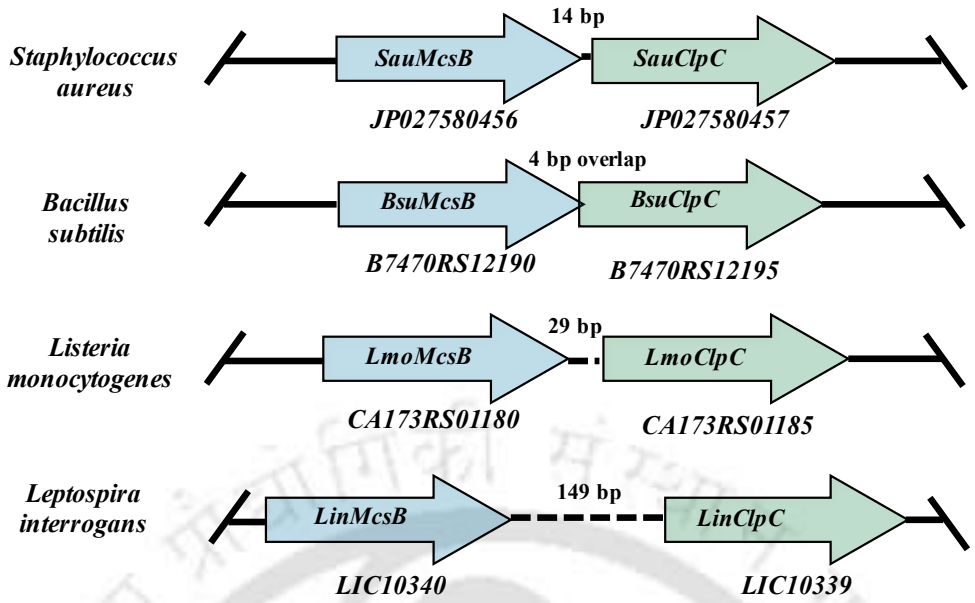


Figure 3S3. Schematic arrangement of the genes encoding arginine kinase, *mcsB,* and caseinolytic chaperone, *clpC*. The organization of the *mcsB* and *clpC* genes in the genome of *L. interrogans* and other pathogenic bacteria was represented. Arrow diagrams based on bioinformatics analysis represent the arrangement of the genes and their annotations, and intergenic regions are depicted as interrupted lines.

TH-3792_206106022

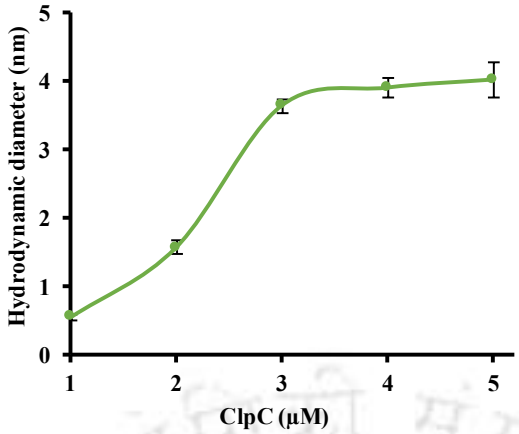


Figure 3S4. The size distribution profile of LinClpC via DLS analysis. The hydrodynamic diameter of LinClpC (1-5 µM) was determined using Dynamic light scattering (DLS). A concentration-dependent increase in D_h was observed for LinClpC.

TH-3792_206106022

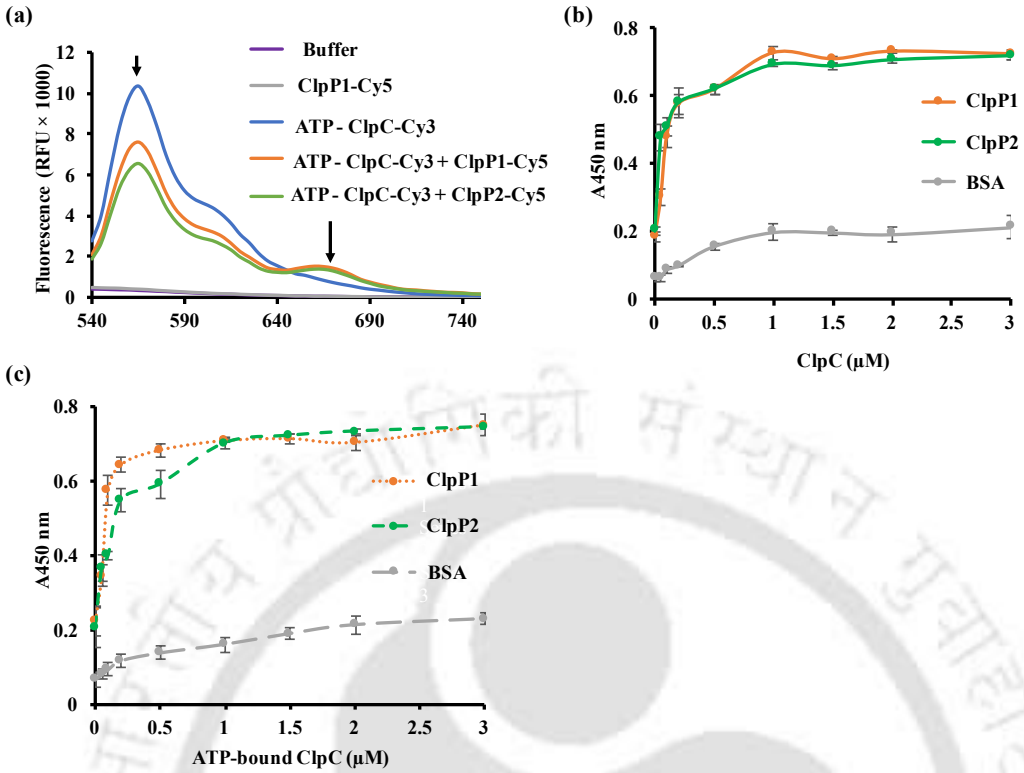


Figure 3S5. Study of LinClpC association with LinClpP isoforms (a) Fluorescence resonance energy transfer (FRET) study between ATP-bound LinClpC-Cy3 (donor) and LinClpP1-Cy5 or LinClpP2-Cy5 (acceptor). The fluorescence emission spectra were recorded at a fixed excitation wavelength of 488 nm and a 540-740 nm emission wavelength. The Emission peaks for Cy3 (570 nm) and Cy5 (670 nm) are indicated by small and large arrows, respectively. The ATP-bound LinClpC-Cy3 displays a single peak at 570 nm. A second peak at 670 nm was observed for ATP-bound LinClpC-Cy3 + LinClpP1-Cy5 and ATP-bound LinClpC-Cy3 + LinClpP2-Cy5 complex, which indicates the interaction of chaperone, LinClpC, with both isoforms of LinClpP (ClpP1 and ClpP2). **(b)** Immunoassay for the interaction of LinClpC with LinClpP isoforms under *in vitro* conditions. The LinClpC-LinClpP isoform interaction study used an enzyme-linked immunosorbent assay (ELISA). **(c)** Immunoassay for the interaction of ATP-bound LinClpC with LinClpP isoforms under *in vitro* conditions was done using indirect ELISA. The LinClpC showed a non-preferential association with each ClpP isoform or its heterocomplex. BSA was used as a negative control.

TH-3792_206106022

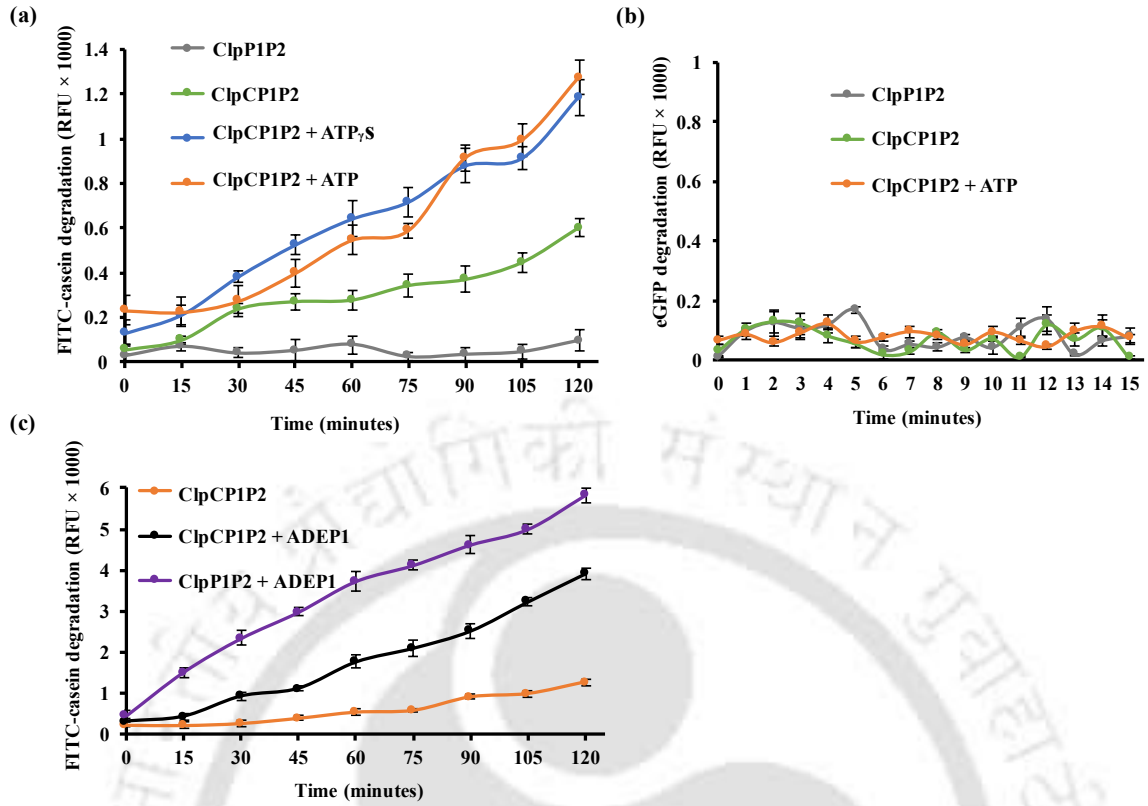


Figure 3S6. Proteolytic activity of LinClpP1P2 heterocomplex in association with LinClpC (a) Degradation of the model substrate (FITC-Casein) was examined by measuring the fluorescence intensity of released FITC at an excitation wavelength of 490 nm and emission wavelength of 525 nm after every 15 minutes for 120 minutes. The protease activity of LinClpC complexed with LinClpP1P2 was checked in the absence (-) of ATP, presence (+) of ATP, and non-hydrolysable ATP analogue, $ATP_{\gamma}S$. **(b)** The degradation of the eGFP model substrate by the LinClpCP1P2 machinery in the presence and absence of ATP was measured for 15 minutes. No degradation of eGFP was observed by ATP-bound LinClpCP1P2 machinery. **(c)** The protease activity of the LinClpP1P2 heterocomplex was checked in the presence of either LinClpC or ADEP1 or both. The degradation was measured every 15 minutes for 120 minutes.

TH-3792_206106022

Supplementary data to Chapter 4

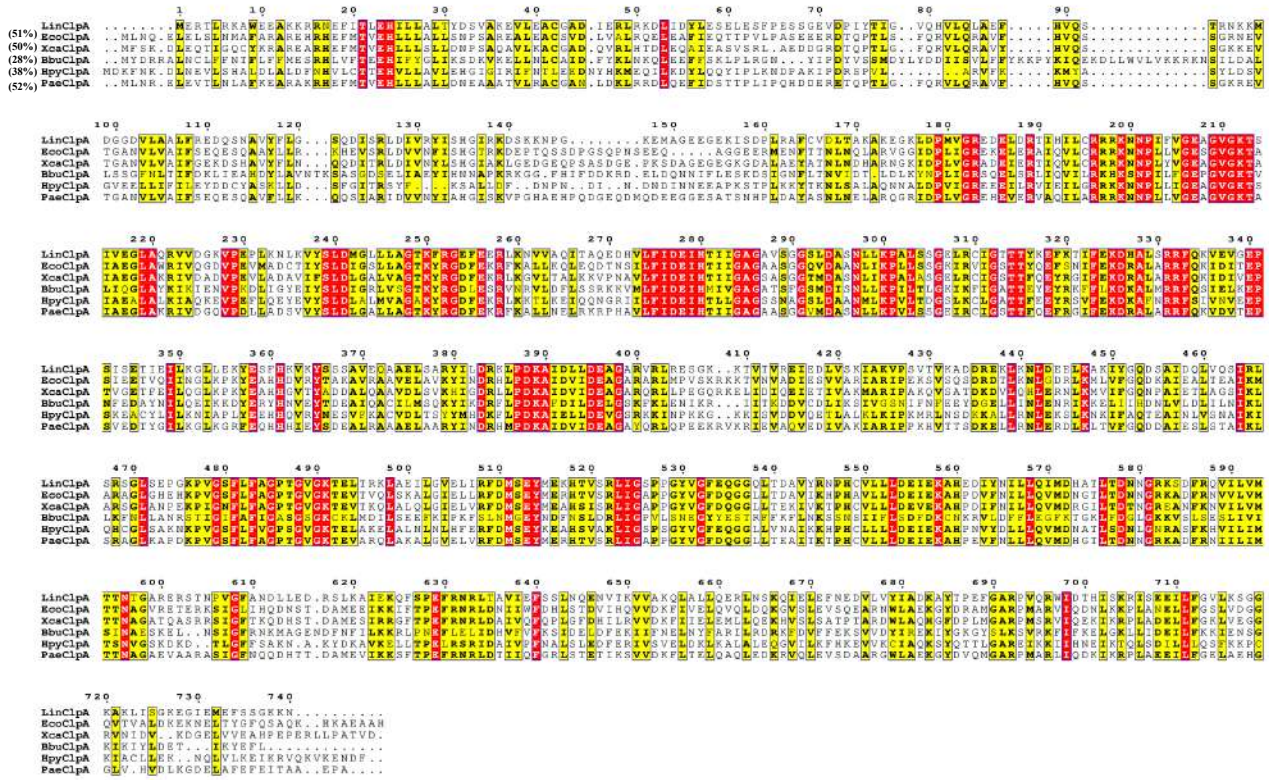


Figure 4S1. The sequence comparison of *Leptospira* ClpA with bacterial orthologs. Multiple sequence alignment of ClpA protein sequences from *Leptospira interrogans* (LinClpA; Q72RD2), *Escherichia coli* (EcoClpA; P0ABH9), *Xanthomonas campestris* (XcaClpA; Q8P998), *Borrelia burgdorferi* (BbuClpA; O51342), *Helicobacter pylori* (HpyClpA; A0AB33XJV6), and *Pseudomonas aeruginosa* PaeClpA; Q9I0L8). The conserved and semi-conserved regions are highlighted in red and yellow, respectively.

TH-3792_206106022

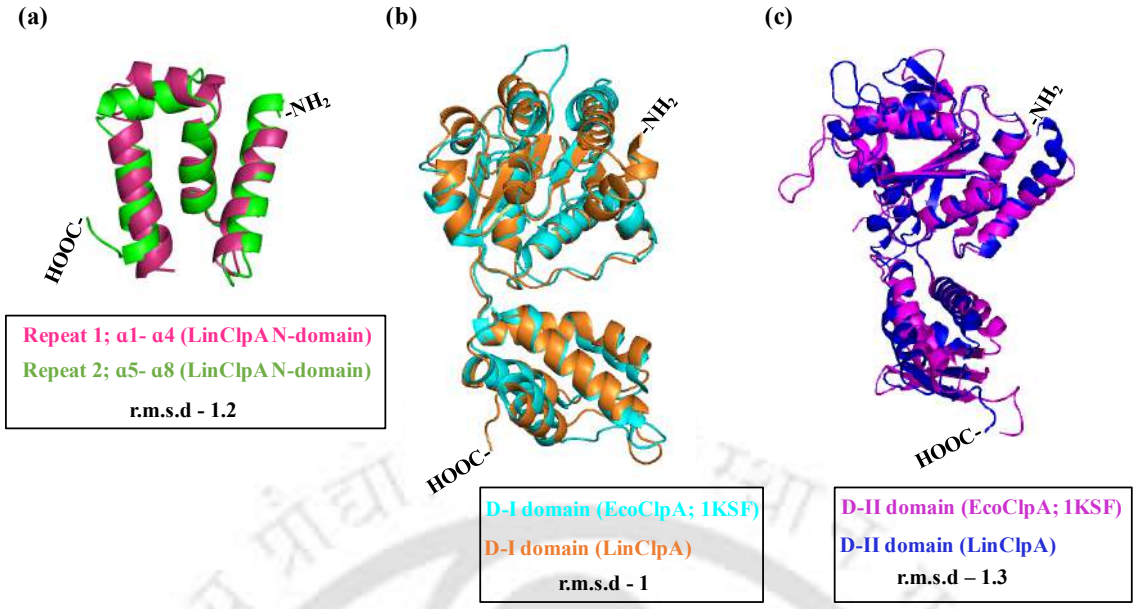


Figure 4S2. Tertiary structural comparison of N-domain, ATPase domain I and II of LinClpA (a) The structural superimposition of repeat 1 and 2 of the modeled LinClpA N-domain. The helices α1- α4 (repeat 1; pink) and α5- α8 (repeat 2; green) represents a pseudo-symmetry with r.m.s.d of 1.2. **(b)** The structural superimposition of modeled LinClpA ATPase domain I (orange) and EcoClpA ATPase domain I (Cyan; PDB: 1KSF) with r.m.s.d of 1. **(c)** The structural superimposition of modeled ATPase domain II of LinClpA (blue) and EcoClpA ATPase domain II (Magenta; PDB: 1KSF) with r.m.s.d of 1.3.

TH-3792_206106022

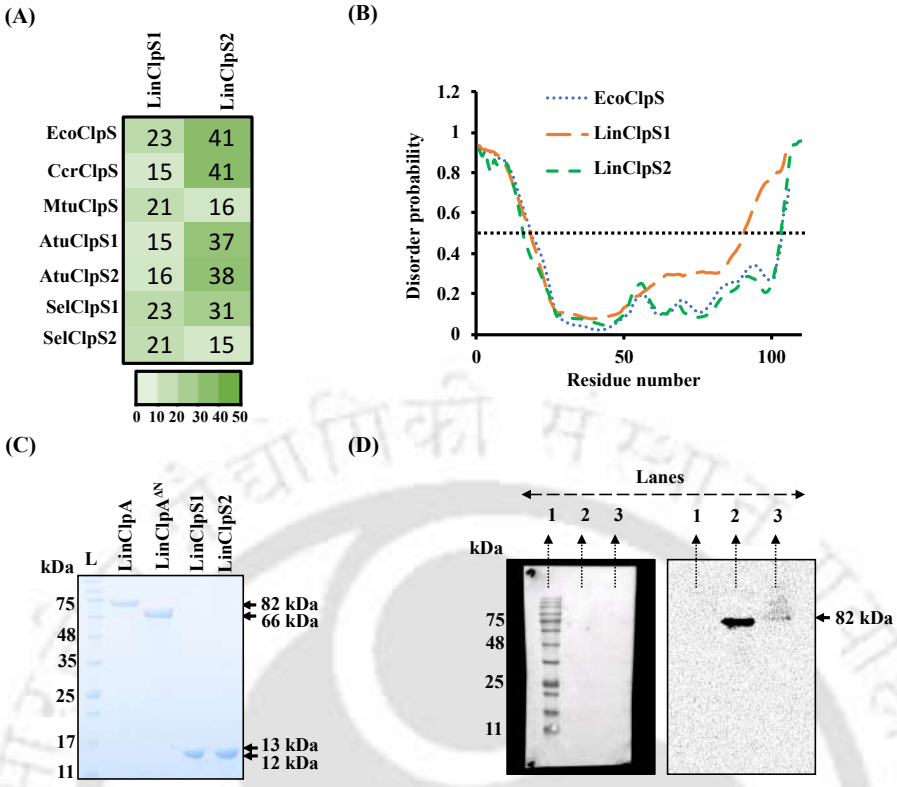


Figure 4S3. Sequence comparison of ClpS orthologs and molecular characterization of LinClp ATPase and adaptor proteins. (a) Heat map showing comparison of amino-acids sequence identity of LinClpS1; Q72SM3 and LinClpS2; Q72RD1 with ClpS proteins from various microorganisms, including *Escherichia coli* (EcoClpS; P0A8Q6), *Caulobacter crescentus* (CcrClpS; B8GZM8), *Mycobacterium tuberculosis* (MtuClpS; P9WPC0), *Agrobacterium tumefaciens* (AtuClpS1; Q8UFN4, AtuClpS2; Q8UD95), *Synechococcus elongatus* (SelClpS1; Q31QE7, SelClpS2; Q31R11). **(b)** The disorder of amino-acid sequences of LinClpS1 (106 aa), LinClpS2 (111 aa), and EcoClpS (106 aa) was determined using PrDOS software, and the disordered probability was plotted with the respective residue number. The cut-off for the disordered region (0.5) is indicated with a dotted line. **(c)** Purification of LinClpA, LinClpA$^{\Delta N}$, LinClpS1, and LinClpS2 via Ni-NTA affinity chromatography. The purified recombinant LinClpA (82 kDa), LinClpA$^{\Delta N}$ (66 kDa), LinClpS1 (12 kDa) and LinClpS2 (13 kDa) proteins resolved on 12 % SDS-PAGE stained with Coomassie blue. **(d)** Immunoblot analysis of *L. interrogans* serovar Copenhageni lysate. Lanes 1, 2 and 3 represent ladder, recombinant LinClpA and leptospiral whole cell lysate, respectively. The expression of native ClpA was detected in *Leptospira* lysate (lane 3) using mouse anti-LinClpA (1:1000 dilution). The recombinant LinClpA and native ClpA of *Leptospira* were detected at a similar molecular size (~ 82 kDa).

TH-3792_206106022

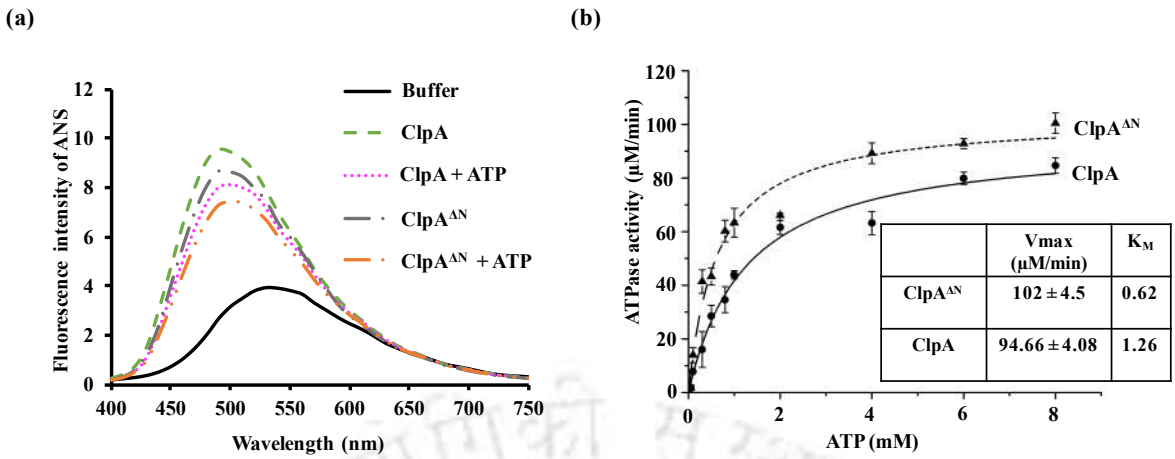


Figure 4S4. Comparison of biochemical activity of LinClpA and LinClpA$^{\Delta N}$ (a) ANS binding assay to study the oligomerization of LinClpA and LinClpA$^{\Delta N}$ in the absence (-) or presence (+) of ATP. The fluorescence intensity of ANS was measured at an excitation wavelength of 350 nm and emission spectra in the range of 400-750 nm. **(b)** Effect of ATP at increasing concentration on the LinClpA and LinClpA$^{\Delta N}$ ATPase activity. The kinetics of the LinClpA and LinClpA$^{\Delta N}$ enzymatic reaction on ATP were determined by non-linear curve fitting in the Origin software, and the Vmax and KM values of the reaction were estimated.

TH-3792_206106022

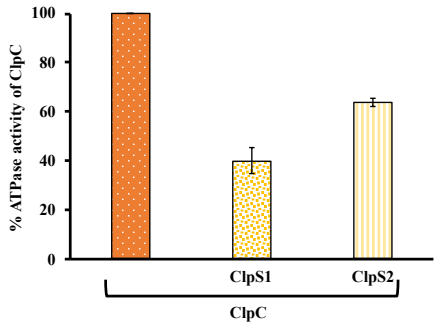


Figure 4S5. Influence of LinClpS on the ATPase activity of LinClpC. The ATPase activity of LinClpC (1 μM) was measured in the presence of LinClpS1 or LinClpS2 proteins (5 μM). The percentage change in ATPase activity of LinClpC in the presence of LinClpS proteins was calculated considering the activity of pure LinClpC to be 100%.

TH-3792_206106022

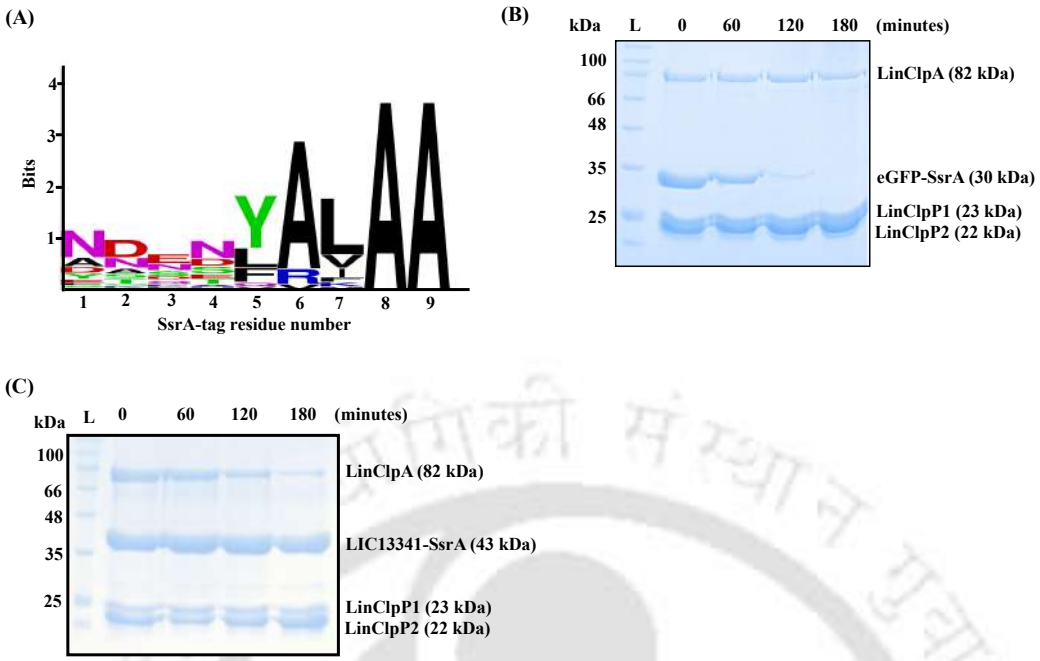


Figure 4S6. The comparison of the SsrA-tag from various pathogenic bacteria and validation of the leptospiral SsrA-tag. (a) WebLogo representation of SsrA-tag sequences from *Leptospira* and other pathogenic bacteria with a well-studied ClpP system. The sequences of SsrA-tag for various bacteria was retrieved from tmRNA database and the last 9 aa sequences were selected for WebLogo representation. The degradation of the **(b)** eGFP-SsrA model substrate **(c)** LIC13341-SsrA substrate by the LinClpAP1P2 machinery. The time-dependent degradation of LinSsrA-tagged substrates was examined on 12 % SDS-PAGE within a time interval of 0, 60, 120 and 180 minutes.

TH-3792_206106022

References

TH-3792_206106022

Adler, B., & de la Peña Moctezuma, A. (2010). *Leptospira* and leptospirosis. *Veterinary microbiology*, *140*(3-4), 287-296.

Ahlawat, S., & Morrison, D. A. (2009). ClpXP degrades SsrA-tagged proteins in Streptococcus pneumoniae. *Journal of bacteriology*, *191*(8), 2894-2898.

Ahmad, A., Bhattacharya, A., McDonald, R. A., Cordes, M., Ellington, B., Bertelsen, E. B., & Zuiderweg, E. R. (2011). Heat shock protein 70 kDa chaperone/DnaJ cochaperone complex employs an unusual dynamic interface. *Proceedings of the National Academy of Sciences*, *108*(47), 18966-18971.

Akiyama, Y., & Ito, K. (2000). Roles of multimerization and membrane association in the proteolytic functions of FtsH (HflB). *The EMBO Journal*.

Akiyama, Y., Kihara, A., Mori, H., Ogura, T., & Ito, K. (1998). Roles of the Periplasmic Domain of Escherichia coliFtsH (HflB) in Protein Interactions and Activity Modulation. *Journal of Biological Chemistry*, *273*(35), 22326-22333.

Akopian, T., Kandror, O., Raju, R. M., UnniKrishnan, M., Rubin, E. J., & Goldberg, A. L. (2012). The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *The EMBO journal*, *31*(6), 1529-1541.

Al-shere, T. A., Ujiie, M., Suzuki, M., Salva, E., Belo, M. C. P., Koizumi, N., ... & Ariyoshi, K. (2012). Outbreak of leptospirosis after flood, the Philippines, 2009. *Emerging infectious diseases*, *18*(1), 91.

Alexopoulos, J. A., Guarné, A., & Ortega, J. (2012). ClpP: a structurally dynamic protease regulated by AAA+ proteins. *Journal of structural biology*, *179*(2), 202-210.

Alexopoulos, J. A., Guarné, A., & Ortega, J. (2012). ClpP: a structurally dynamic protease regulated by AAA+ proteins. *Journal of structural biology*, *179*(2), 202-210.

Alexopoulos, J., Ahsan, B., Homchaudhuri, L., Husain, N., Cheng, Y. Q., & Ortega, J. (2013). Structural determinants stabilizing the axial channel of ClpP for substrate translocation. *Molecular microbiology*, *90*(1), 167-180.

Ali, M. M., Roe, S. M., Vaughan, C. K., Meyer, P., Panaretou, B., Piper, P. W., ... & Pearl, L. H. (2006). Crystal structure of an Hsp90–nucleotide–p23/Sba1 closed chaperone complex. *Nature*, *440*(7087), 1013-1017.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410.

Alves França, B., Falke, S., Rohde, H., & Betzel, C. (2024). Molecular insights into the dynamic modulation of bacterial ClpP function and oligomerization by peptidomimetic boronate compounds. *Scientific reports*, *14*(1), 2572.

Amor, A. J., Schmitz, K. R., Baker, T. A., & Sauer, R. T. (2019). Roles of the ClpX IGF loops in ClpP association, dissociation, and protein degradation. *Protein Science*, *28*(4), 756-765.

Andersson, F. I., Blakytny, R., Kirstein, J., Turgay, K., Bukau, B., Mogk, A., & Clarke, A. K. (2006). Cyanobacterial ClpC/HSP100 protein displays intrinsic chaperone activity. *Journal of Biological Chemistry*, *281*(9), 5468-5475.

Antima, & Banerjee, S. (2023). Modeling the dynamics of leptospirosis in India. *Scientific Reports*, *13*(1), 19791.

TH-3792_206106022

Asahara, Y., Atsuta, K., Motohashi, K., Taguchi, H., Yohda, M., & Yoshida, M. (2000). FtsH recognizes proteins with unfolded structure and hydrolyzes the carboxyl side of hydrophobic residues. *The journal of biochemistry*, *127*(5), 931-937.

Awad, W., Al-Eryani, Y., Ekström, S., Logan, D. T., & von Wachenfeldt, C. (2019). Structural basis for YjbH adaptor-mediated recognition of transcription factor Spx. *Structure*, *27*(6), 923-936.

Bachmair, A., Finley, D., & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *science*, *234*(4773), 179-186.

Bajaj, D., & Batra, J. K. (2012). The C-terminus of ClpC1 of Mycobacterium tuberculosis is crucial for its oligomerization and function. *PloS one*, *7*(12), e51261.

Baker, T. A., & Sauer, R. T. (2006). ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends in biochemical sciences*, *31*(12), 647-653.

Baker, T. A., & Sauer, R. T. (2012). ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1823*(1), 15-28.

Balaban, N. Q., Helaine, S., Lewis, K., Ackermann, M., Aldridge, B., Andersson, D. I., ... & Zinkernagel, A. (2019). Definitions and guidelines for research on antibiotic persistence. *Nature Reviews Microbiology*, *17*(7), 441-448.

Balogh, D., Dahmen, M., Stahl, M., Poreba, M., Gersch, M., Drag, M., & Sieber, S. A. (2017). Insights into ClpXP proteolysis: heterooligomerization and partial deactivation enhance chaperone affinity and substrate turnover in Listeria monocytogenes. *Chemical Science*, *8*(2), 1592-1600.

Balzi, E., Choder, M., Chen, W. N., Varshavsky, A., & Goffeau, A. (1990). Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *265*(13), 7464-7471.

Baneyx, F., & Mujacic, M. (2004). Recombinant protein folding and misfolding in Escherichia coli. *Nature biotechnology*, *22*(11), 1399-1408.

Baram, D., Pyetan, E., Sittner, A., Auerbach-Nevo, T., Bashan, A., & Yonath, A. (2005). Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action. *Proceedings of the National Academy of Sciences*, *102*(34), 12017-12022.

Battesti, A., & Gottesman, S. (2013). Roles of adaptor proteins in regulation of bacterial proteolysis. *Current opinion in microbiology*, *16*(2), 140-147.

Bauer, D., Merz, D. R., Pelz, B., Theisen, K. E., Yacyshyn, G., Mokranjac, D., ... & Žoldák, G. (2015). Nucleotides regulate the mechanical hierarchy between subdomains of the nucleotide binding domain of the Hsp70 chaperone DnaK. *Proceedings of the National Academy of Sciences*, *112*(33), 10389-10394.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... & Bourne, P. E. (2000). The protein data bank. *Nucleic acids research*, *28*(1), 235-242.

Bernot, K. M., Lee, C. H., & Coulombe, P. A. (2005). A small surface hydrophobic stripe in the coiled-coil domain of type I keratins mediates tetramer stability. *The Journal of cell biology*, *168*(6), 965-974.

TH-3792_206106022

Beuron, F., Maurizi, M. R., Belnap, D. M., Kocsis, E., Booy, F. P., Kessel, M., & Steven, A. C. (1998). At sixes and sevens: characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. *Journal of structural biology*, *123*(3), 248-259.

Bewley, M. C., Graziano, V., Griffin, K., & Flanagan, J. M. (2006). The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *Journal of structural biology*, *153*(2), 113-128.

Bewley, M. C., Graziano, V., Griffin, K., & Flanagan, J. M. (2009). Turned on for degradation: ATPase-independent degradation by ClpP. *Journal of structural biology*, *165*(2), 118-125.

Bhandari, V., Wong, K. S., Zhou, J. L., Mabanglo, M. F., Batey, R. A., & Houry, W. A. (2018). The role of ClpP protease in bacterial pathogenesis and human diseases. *ACS chemical biology*, *13*(6), 1413-1425.

Bharti, A. R., Nally, J. E., Ricaldi, J. N., Matthias, M. A., Diaz, M. M., Lovett, M. A., ... & Vinetz, J. M. (2003). Leptospirosis: a zoonotic disease of global importance. *The Lancet infectious diseases*, *3*(12), 757-771.

Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D., & Huber, R. (2000). The structures of HslU and the ATP-dependent protease HslU–HslV. *Nature*, *403*(6771), 800-805.

Böttcher, T., & Sieber, S. A. (2008). β-lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of Staphylococcus aureus. *Journal of the American Chemical Society*, *130*(44), 14400-14401.

Böttcher, T., & Sieber, S. A. (2009). Beta-lactones decrease the intracellular virulence of Listeria monocytogenes in macrophages.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., & Sigler, P. B. (1994). The crystal structure of the bacterial chaperonln GroEL at 2.8 Å. *Nature*, *371*(6498), 578-586.

Brötz-Oesterhelt, H., Beyer, D., Kroll, H. P., Endermann, R., Ladel, C., Schroeder, W., ... & Labischinski, H. (2005). Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nature medicine*, *11*(10), 1082-1087.

Bukau, B., & Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell*, *92*(3), 351-366.

Calloni, G., Chen, T., Schermann, S. M., Chang, H. C., Genevaux, P., Agostini, F., ... & Hartl, F. U. (2012). DnaK functions as a central hub in the E. coli chaperone network. *Cell reports*, *1*(3), 251-264.

Carroni, M., Franke, K. B., Maurer, M., Jäger, J., Hantke, I., Gloge, F., ... & Mogk, A. (2017). Regulatory coiled-coil domains promote head-to-head assemblies of AAA+ chaperones essential for tunable activity control. *Elife*, *6*, e30120.

Chadsuthi, S., Chalvet-Monfray, K., Wiratsudakul, A., & Modchang, C. (2021). The effects of flooding and weather conditions on leptospirosis transmission in Thailand. *Scientific Reports*, *11*(1), 1486.

Chan, J. F., Lau, S. K., To, K. K., Cheng, V. C., Woo, P. C., & Yuen, K. Y. (2015). Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. *Clinical microbiology reviews*, *28*(2), 465-522.

TH-3792_206106022

Charette, M. F., Henderson, G. W., & Markovitz, A. (1981). ATP hydrolysis-dependent protease activity of the lon (capR) protein of Escherichia coli K-12. *Proceedings of the National Academy of Sciences*, *78*(8), 4728-4732.

Chatterjee, I., Becker, P., Grundmeier, M., Bischoff, M., Somerville, G. A., Peters, G., ... & Herrmann, M. (2005). Staphylococcus aureus ClpC is required for stress resistance, aconitase activity, growth recovery, and death. *Journal of bacteriology*, *187*(13), 4488-4496.

Chien, P., Perchuk, B. S., Laub, M. T., Sauer, R. T., & Baker, T. A. (2007). Direct and adaptor-mediated substrate recognition by an essential AAA+ protease. *Proceedings of the National Academy of Sciences*, *104*(16), 6590-6595.

Choudhury, M., Dhara, A., & Kumar, M. (2021). Trigger Factor in Association with the ClpP1P2 Heterocomplex of Leptospira Promotes Protease/Peptidase Activity. *ACS omega*, *6*(2), 1400-1409.

Chung, C. H., & Goldberg, A. L. (1981). The product of the lon (capR) gene in Escherichia coli is the ATP-dependent protease, protease La. *Proceedings of the National Academy of Sciences*, *78*(8), 4931-4935.

Clare, D. K., & Saibil, H. R. (2013). ATP-driven molecular chaperone machines. *Biopolymers*, *99*(11), 846-859.

Compton, C. L., Schmitz, K. R., Sauer, R. T., & Sello, J. K. (2013). Antibacterial activity of and resistance to small molecule inhibitors of the ClpP peptidase. *ACS chemical biology*, *8*(12), 2669-2677.

Conlon, B. P., Nakayasu, E. S., Fleck, L. E., LaFleur, M. D., Isabella, V. M., Coleman, K., ... & Lewis, K. (2013). Activated ClpP kills persisters and eradicates a chronic biofilm infection. *Nature*, *503*(7476), 365-370.

Costa, F., Hagan, J. E., Calcagno, J., Kane, M., Torgerson, P., Martinez-Silveira, M. S., ... & Ko, A. I. (2015). Global morbidity and mortality of leptospirosis: a systematic review. *PLoS neglected tropical diseases*, *9*(9), e0003898.

Craig, E. A., Eisenman, H. C., & Hundley, H. A. (2003). Ribosome-tethered molecular chaperones: the first line of defense against protein misfolding?. *Current opinion in microbiology*, *6*(2), 157-162.

Cranz-Mileva, S., Imkamp, F., Kolygo, K., Maglica, Ž., Kress, W., & Weber-Ban, E. (2008). The flexible attachment of the N-domains to the ClpA ring body allows their use on demand. *Journal of molecular biology*, *378*(2), 412-424.

Crooke, E., Guthrie, B., Lecker, S., Lill, R., & Wickner, W. (1988). ProOmpA is stabilized for membrane translocation by either purified E. coli trigger factor or canine signal recognition particle. *Cell*, *54*(7), 1003-1011.

Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome research*, *14*(6), 1188-1190.

Culp, E., & Wright, G. D. (2017). Bacterial proteases, untapped antimicrobial drug targets. *The Journal of antibiotics*, *70*(4), 366-377.

Davey, M. J., Indiani, C., & O'Donnell, M. (2003). Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture. *Journal of Biological Chemistry*, *278*(7), 4491-4499.

Dechet, A. M., Parsons, M., Rambaran, M., Mohamed-Rambaran, P., Florendo-Cumbermack, A., Persaud, S., ... & Mintz, E. D. (2012). Leptospirosis outbreak following severe flooding: a

TH-3792_206106022

rapid assessment and mass prophylaxis campaign; Guyana, January–February 2005. *PloS one*, *7*(7), e39672.

DeLano, W. L. (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsl. protein crystallogr*, *40*(1), 82-92.

Derré, I., Rapoport, G., & Msadek, T. (1999). CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in Gram-positive bacteria. *Molecular microbiology*, *31*(1), 117-131.

Dhara, A., Hussain, M. S., Datta, D., & Kumar, M. (2019). Insights to the assembly of a functionally active leptospiral ClpP1P2 protease complex along with its ATPase chaperone ClpX. *ACS omega*, *4*(7), 12880-12895.

Dhara, A., Hussain, M. S., Kanaujia, S. P., & Kumar, M. (2021). Acyldepsipeptide activated ClpP1P2 macromolecule of Leptospira, an ideal Achilles' heel to hamper the cell survival and deregulate ClpP proteolytic activity. *Research in Microbiology*, *172*(2), 103797.

Díaz-Sáez, L., Pankov, G., & Hunter, W. N. (2017). Open and compressed conformations of Francisella tularensis ClpP. *Proteins: Structure, Function, and Bioinformatics*, *85*(1), 188-194.

Dixit, B., Prakash, A., Kumar, P., Gogoi, P., & Kumar, M. (2021). The core Cas1 protein of CRISPR-Cas IB in Leptospira shows metal-tunable nuclease activity. *Current Research in Microbial Sciences*, *2*, 100059.

Dobson CM, Sali A, Karplus M. 1998. Protein folding: a perspective from theory and experiment. Angew. Chem. Int. Ed.37:868–93

Dollins, D. E., Warren, J. J., Immormino, R. M., & Gewirth, D. T. (2007). Structures of GRP94-nucleotide complexes reveal mechanistic differences between the hsp90 chaperones. *Molecular cell*, *28*(1), 41-56.

Donaldson, L. W., Wojtyra, U., & Houry, W. A. (2003). Solution structure of the dimeric zinc binding domain of the chaperone ClpX. *Journal of Biological Chemistry*, *278*(49), 48991-48996.

Donegan, N. P., Marvin, J. S., & Cheung, A. L. (2014). Role of adaptor TrfA and ClpPC in controlling levels of SsrA-tagged proteins and antitoxins in Staphylococcus aureus. *Journal of bacteriology*, *196*(23), 4140-4151.

Dougan, D. A., Micevski, D., & Truscott, K. N. (2012). The N-end rule pathway: from recognition by N-recognins, to destruction by AAA+ proteases. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1823*(1), 83-91.

Dougan, D. A., Reid, B. G., Horwich, A. L., & Bukau, B. (2002). ClpS, a substrate modulator of the ClpAP machine. *Molecular cell*, *9*(3), 673-683.

Dougan, D. A., Truscott, K. N., & Zeth, K. (2010). The bacterial N-end rule pathway: expect the unexpected. *Molecular microbiology*, *76*(3), 545-558.

Dougan, D. A., Weber-Ban, E., & Bukau, B. (2003). Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. *Molecular cell*, *12*(2), 373-380.

Effantin, G., Maurizi, M. R., & Steven, A. C. (2010). Binding of the ClpA unfoldase opens the axial gate of ClpP peptidase. *Journal of Biological Chemistry*, *285*(19), 14834-14840.

TH-3792_206106022

Ekaza, E., Guilloteau, L., Teyssier, J., Liautard, J. P., & Kohler, S. (2000). Functional analysis of the ClpATPase ClpA of Brucella suis, and persistence of a knockout mutant in BALB/c mice. *Microbiology*, *146*(7), 1605-1616.

El Bakkouri, M., Pow, A., Mulichak, A., Cheung, K. L., Artz, J. D., Amani, M., ... & Houry, W. A. (2010). The Clp chaperones and proteases of the human malaria parasite Plasmodium falciparum. *Journal of molecular biology*, *404*(3), 456-477.

Ellis, R. John, and Allen P. Minton. "Protein aggregation in crowded environments." (2006): 485-497.

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., ... & Bukau, B. (2006). ClpS is an essential component of the N-end rule pathway in Escherichia coli. *Nature*, *439*(7077), 753-756.

Erzberger, J. P., & Berger, J. M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.*, *35*(1), 93-114.

Faine, S. (1994). *Leptospira and leptospirosis* (pp. 353-pp).

Famulla, K., Sass, P., Malik, I., Akopian, T., Kandror, O., Alber, M., ... & Brötz-Oesterhelt, H. (2016). Acyldepsipeptide antibiotics kill mycobacteria by preventing the physiological functions of the ClpP1P2 protease. *Molecular microbiology*, *101*(2), 194-209.

Fang, L., Ricketson, D., Getubig, L., & Darimont, B. (2006). Unliganded and hormone-bound glucocorticoid receptors interact with distinct hydrophobic sites in the Hsp90 C-terminal domain. *Proceedings of the National Academy of Sciences*, *103*(49), 18487-18492.

Farr, S. B., & Kogoma, T. O. K. I. O. (1991). Oxidative stress responses in Escherichia coli and Salmonella typhimurium. *Microbiological reviews*, *55*(4), 561-585.

Fatima, K., Naqvi, F., & Younas, H. (2021). A review: Molecular chaperone-mediated folding, unfolding and disaggregation of expressed recombinant proteins. *Cell Biochemistry and Biophysics*, *79*(2), 153-174.

Fei, X., Bell, T. A., Jenni, S., Stinson, B. M., Baker, T. A., Harrison, S. C., & Sauer, R. T. (2020). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *Elife*, *9*, e52774.

Felix, J., Weinhäupl, K., Chipot, C., Dehez, F., Hessel, A., Gauto, D. F., ... & Fraga, H. (2019). Mechanism of the allosteric activation of the ClpP protease machinery by substrates and active-site inhibitors. *Science advances*, *5*(9), eaaw3818.

Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., & Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature*, *431*(7008), 590-596.

Flynn, J. M., Levchenko, I., Seidel, M., Wickner, S. H., Sauer, R. T., & Baker, T. A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proceedings of the National Academy of Sciences*, *98*(19), 10584-10589.

Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., & Baker, T. A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular cell*, *11*(3), 671-683.

Fritze, J., Zhang, M., Luo, Q., & Lu, X. (2020). An overview of the bacterial SsrA system modulating intracellular protein levels and activities. *Applied microbiology and biotechnology*, *104*, 5229-5241.

TH-3792_206106022

Frydman J. 2001. Folding of newly translated proteins in vivo: the role of molecular chaperones. Annu. Rev. Biochem. 70:603–47

Führer, F., Müller, A., Baumann, H., Langklotz, S., Kutscher, B., & Narberhaus, F. (2007). Sequence and length recognition of the C-terminal turnover element of LpxC, a soluble substrate of the membrane-bound FtsH protease. *Journal of molecular biology*, *372*(2), 485-496.

Fujiwara, K., Ishihama, Y., Nakahigashi, K., Soga, T., & Taguchi, H. (2010). A systematic survey of in vivo obligate chaperonin-dependent substrates. *The EMBO journal*, *29*(9), 1552-1564.

Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S., & Berche, P. (2000). The ClpP serine protease is essential for the intracellular parasitism and virulence of Listeria monocytogenes. *Molecular microbiology*, *35*(6), 1286-1294.

Galan, D. I., Roess, A. A., Pereira, S. V. C., & Schneider, M. C. (2021). Epidemiology of human leptospirosis in urban and rural areas of Brazil, 2000–2015. *PloS one*, *16*(3), e0247763.

Gao, W., Kim, J. Y., Anderson, J. R., Akopian, T., Hong, S., Jin, Y. Y., ... & Cho, S. (2015). The cyclic peptide ecumicin targeting ClpC1 is active against Mycobacterium tuberculosis in vivo. *Antimicrobial agents and chemotherapy*, *59*(2), 880-889.

Gatsogiannis, C., Balogh, D., Merino, F., Sieber, S. A., & Raunser, S. (2019). Cryo-EM structure of the ClpXP protein degradation machinery. *Nature structural & molecular biology*, *26*(10), 946-954.

Gavrish, E., Sit, C. S., Cao, S., Kandror, O., Spoering, A., Peoples, A., ... & Lewis, K. (2014). Lassomycin, a ribosomally synthesized cyclic peptide, kills Mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chemistry & biology*, *21*(4), 509-518.

Geiger, S. R., Böttcher, T., Sieber, S. A., & Cramer, P. (2011). A conformational switch underlies ClpP protease function. *Angewandte Chemie International Edition*, *50*(25), 5749-5752.

Georgopoulos, C. P., Hendrix, R. W., Casjens, S. R., & Kaiser, A. D. (1973). Host participation in bacteriophage lambda head assembly. *Journal of molecular biology*, *76*(1), 45-60.

Gersch, M., Famulla, K., Dahmen, M., Göbl, C., Malik, I., Richter, K., ... & Sieber, S. A. (2015). AAA+ chaperones and acyldepsipeptides activate the ClpP protease via conformational control. *Nature communications*, *6*(1), 6320.

Gersch, M., Gut, F., Korotkov, V. S., Lehmann, J., Böttcher, T., Rusch, M., ... & Sieber, S. A. (2013). The mechanism of caseinolytic protease (ClpP) inhibition. *Angewandte Chemie International Edition*, *52*(10), 3009-3014.

Gersch, M., Kolb, R., Alte, F., Groll, M., & Sieber, S. A. (2014). Disruption of oligomerization and dehydroalanine formation as mechanisms for ClpP protease inhibition. *Journal of the American Chemical Society*, *136*(4), 1360-1366.

Gersch, M., List, A., Groll, M., & Sieber, S. A. (2012). Insights into structural network responsible for oligomerization and activity of bacterial virulence regulator caseinolytic protease P (ClpP) protein. *Journal of Biological Chemistry*, *287*(12), 9484-9494.

Ghosh, C., Sarkar, P., Issa, R., & Haldar, J. (2019). Alternatives to conventional antibiotics in the era of antimicrobial resistance. *Trends in microbiology*, *27*(4), 323-338.

TH-3792_206106022

Ghosh, K. K., Prakash, A., Shrivastav, P., Balamurugan, V., & Kumar, M. (2018). Evaluation of a novel outer membrane surface-exposed protein, LIC13341 of Leptospira, as an adhesin and serodiagnostic candidate marker for leptospirosis. *Microbiology*, *164*(8), 1023-1037.

Goarant, C., Laumond-Barny, S., Perez, J., Vernel-Pauillac, F., Chanteau, S., & Guigon, A. (2009). Outbreak of leptospirosis in New Caledonia: diagnosis issues and burden of disease. *Tropical Medicine & International Health*, *14*(8), 926-929.

Gonda, D. K., Bachmair, A., Wünning, I., Tobias, J. W., Lane, W. S., & Varshavsky, A. (1989). Universality and structure of the N-end rule. *Journal of Biological Chemistry*, *264*(28), 16700-16712.

Gonzalez, M., Rasulova, F., Maurizi, M. R., & Woodgate, R. (2000). Subunit-specific degradation of the UmuD/D′ heterodimer by the ClpXP protease: the role of trans recognition in UmuD′ stability. *The EMBO journal*.

Goodreid, J. D., Janetzko, J., Santa Maria Jr, J. P., Wong, K. S., Leung, E., Eger, B. T., ... & Batey, R. A. (2016). Development and characterization of potent cyclic acyldepsipeptide analogues with increased antimicrobial activity. *Journal of medicinal chemistry*, *59*(2), 624-646.

Goodreid, J. D., Janetzko, J., Santa Maria Jr, J. P., Wong, K. S., Leung, E., Eger, B. T., ... & Batey, R. A. (2016). Development and characterization of potent cyclic acyldepsipeptide analogues with increased antimicrobial activity. *Journal of medicinal chemistry*, *59*(2), 624-646.

Gottesman, S. (1996). Proteases and their targets in Escherichia coli. *Annual review of genetics*, *30*(1), 465-506.

Gottesman, S. (2003). Proteolysis in bacterial regulatory circuits. *Annual review of cell and developmental biology*, *19*(1), 565-587.

Gottesman, S., Clark, W. P., de Crecy-Lagard, V., & Maurizi, M. R. (1993). ClpX, an alternative subunit for the ATP-dependent Clp protease of Escherichia coli. Sequence and in vivo activities. *Journal of Biological Chemistry*, *268*(30), 22618-22626.

Gottesman, S., Roche, E., Zhou, Y., & Sauer, R. T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes & development*, *12*(9), 1338-1347.

Graciet, E., Hu, R.G., Piatkov, K., Rhee, J.H., Schwarz, E.M.,and Varshavsky, A. (2006) Aminoacyl-transferases and theN-end rule pathway of prokaryotic/eukaryotic specificity ina human pathogen. Proc Natl Acad Sci USA 103: 3078–3083

Graciet, E., Walter, F., Maoiléidigh, D. Ó., Pollmann, S., Meyerowitz, E. M., Varshavsky, A., & Wellmer, F. (2009). The N-end rule pathway controls multiple functions during Arabidopsis shoot and leaf development. *Proceedings of the National Academy of Sciences*, *106*(32), 13618-13623.

Gribun, A., Kimber, M. S., Ching, R., Sprangers, R., Fiebig, K. M., & Houry, W. A. (2005). The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation. *Journal of Biological Chemistry*, *280*(16), 16185-16196.

Griffith, K. L., & Grossman, A. D. (2008). Inducible protein degradation in Bacillus subtilis using heterologous peptide tags and adaptor proteins to target substrates to the protease ClpXP. *Molecular microbiology*, *70*(4), 1012-1025.

Griffith, M. E., Hospenthal, D. R., & Murray, C. K. (2006). Antimicrobial therapy of leptospirosis. *Current opinion in infectious diseases*, *19*(6), 533-537.

Grigoryev, S., Stewart, A. E., Kwon, Y. T., Arfin, S. M., Bradshaw, R. A., Jenkins, N. A., ... & Varshavsky, A. (1996). A mouse amidase specific for N-terminal asparagine: the gene, the enzyme, and their function in the N-end rule pathway. *Journal of Biological Chemistry*, *271*(45), 28521-28532.

Grimaud, R., Kessel, M., Beuron, F., Steven, A. C., & Maurizi, M. R. (1998). Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. *Journal of Biological Chemistry*, *273*(20), 12476-12481.

Guo, F., Esser, L., Singh, S. K., Maurizi, M. R., & Xia, D. (2002). Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. *Journal of Biological Chemistry*, *277*(48), 46753-46762.

Haake, D. A., & Levett, P. N. (2015). Leptospirosis in humans. *Leptospira and leptospirosis*, 65-97.

Hackl, M. W., Lakemeyer, M., Dahmen, M., Glaser, M., Pahl, A., Lorenz-Baath, K., ... & Sieber, S. A. (2015). Phenyl esters are potent inhibitors of caseinolytic protease P and reveal a stereogenic switch for deoligomerization. *Journal of the American Chemical Society*, *137*(26), 8475-8483.

Hall, B. M., Breidenstein, E. B., de La Fuente-Núñez, C., Reffuveille, F., Mawla, G. D., Hancock, R. E., & Baker, T. A. (2017). Two isoforms of Clp peptidase in Pseudomonas aeruginosa control distinct aspects of cellular physiology. *Journal of bacteriology*, *199*(3), 10-1128.

Hanson, P. I., & Whiteheart, S. W. (2005). AAA+ proteins: have engine, will work. *Nature reviews Molecular cell biology*, *6*(7), 519-529.

Hartl FU, Hayer-Hartl M. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295:1852–58

Hartl, F. U., & Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nature structural & molecular biology*, *16*(6), 574-581.

Hartskeerl, R. A., Collares-Pereira, M., & Ellis, W. A. (2011). Emergence, control and re-emerging leptospirosis: dynamics of infection in the changing world. *Clinical microbiology and infection*, *17*(4), 494-501.

Hesterkamp, T., Deuerling, E., & Bukau, B. (1997). The amino-terminal 118 amino acids of Escherichia coli trigger factor constitute a domain that is necessary and sufficient for binding to ribosomes. *Journal of Biological Chemistry*, *272*(35), 21865-21871.

Hinnerwisch, J., Reid, B. G., Fenton, W. A., & Horwich, A. L. (2005). Roles of the N-domains of the ClpA unfoldase in binding substrate proteins and in stable complex formation with the ClpP protease. *Journal of Biological Chemistry*, *280*(49), 40838-40844.

Hoskins, J. R., Pak, M., Maurizi, M. R., & Wickner, S. (1998). The role of the ClpA chaperone in proteolysis by ClpAP. *Proceedings of the National Academy of Sciences*, *95*(21), 12135-12140.

Hsieh, F. C., Chen, C. T., Weng, Y. T., Peng, S. S., Chen, Y. C., Huang, L. Y., ... & Wu, W. F. (2011). Stepwise activity of ClpY (HslU) mutants in the processive degradation of Escherichia coli ClpYQ (HslUV) protease substrates. *Journal of bacteriology*, *193*(19), 5465-5476.

TH-3792_206106022

Huang, M., Zhao, Y., Feng, L., Zhu, L., Zhan, L., & Chen, X. (2020[a]). Role of ClpB from Corynebacterium crenatum in thermal stress and arginine fermentation. *Frontiers in Microbiology*, *11*, 1660.

Huang, M., Zhao, Y., Feng, L., Zhu, L., Zhan, L., & Chen, X. (2020[b]). Role of the ClpX from Corynebacterium crenatum involved in stress responses and energy metabolism. *Applied Microbiology and Biotechnology*, *104*(12), 5505-5517.

Hwang, B. J., Park, W. J., Chung, C. H., & Goldberg, A. L. (1987). Escherichia coli contains a soluble ATP-dependent protease (Ti) distinct from protease La. *Proceedings of the National Academy of Sciences*, *84*(16), 5550-5554.

Hwang, B. J., Woo, K. M., Goldberg, A. L., & Chung, C. H. (1988). Protease Ti, a new ATP-dependent protease in Escherichia coli, contains protein-activated ATPase and proteolytic functions in distinct subunits. *Journal of Biological Chemistry*, *263*(18), 8727-8734.

Illigmann, A., Vielberg, M. T., Lakemeyer, M., Wolf, F., Dema, T., Stange, P., ... & Brötz-Oesterhelt, H. (2024). Structure of Staphylococcus aureus ClpP Bound to the Covalent Active-Site Inhibitor Cystargolide A. *Angewandte Chemie International Edition*, *63*(3), e202314028.

Inada, R., Ido, Y., Hoki, R., Kaneko, R., & Ito, H. (1916). The etiology, mode of infection, and specific therapy of Weil's disease (spirochaetosis icterohaemorrhagica). *The Journal of experimental medicine*, *23*(3), 377.

Ingvarsson, H., Maté, M. J., Högbom, M., Portnoï, D., Benaroudj, N., Alzari, P. M., ... & Unge, T. (2007). Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of Mycobacterium tuberculosis ClpP1. *Biological Crystallography*, *63*(2), 249-259.

Ishida, T., & Kinoshita, K. (2007). PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic acids research*, *35*(suppl_2), W460-W464.

Ishikawa, T., Maurizi, M. R., & Steven, A. C. (2004). The N-terminal substrate-binding domain of ClpA unfoldase is highly mobile and extends axially from the distal surface of ClpAP protease. *Journal of structural biology*, *146*(1-2), 180-188.

Ito, K., & Akiyama, Y. (2005). Cellular functions, mechanism of action, and regulation of FtsH protease. *Annu. Rev. Microbiol.*, *59*(1), 211-231.

Jain, R., & Chan, M. K. (2007). Support for a potential role of E. coli oligopeptidase A in protein degradation. *Biochemical and biophysical research communications*, *359*(3), 486-490.

Jana, B., & Biswas, I. (2020). Significance of individual domains of ClpL: A novel chaperone from Streptococcus mutans. *Biochemistry*, *59*(36), 3368-3379.

Jennings, L. D., Bohon, J., Chance, M. R., & Licht, S. (2008). The ClpP N-terminus coordinates substrate access with protease active site reactivity. *Biochemistry*, *47*(42), 11031-11040.

Jing, S., Wang, L., Wang, T., Fan, L., Chen, L., Xiang, H., ... & Wang, D. (2021). Myricetin protects mice against MRSA-related lethal pneumonia by targeting ClpP. *Biochemical Pharmacology*, *192*, 114753.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *nature*, *596*(7873), 583-589.

Kåhrström, C. T. (2013). Entering a post-antibiotic era?. *Nature Reviews Microbiology*, *11*(3), 146-146.

TH-3792_206106022

Kåhrström, C. T. (2013). Entering a post-antibiotic era?. *Nature Reviews Microbiology*, *11*(3), 146-146.

Kaiser, C. M., Chang, H. C., Agashe, V. R., Lakshmipathy, S. K., Etchells, S. A., Hayer-Hartl, M., ... & Barral, J. M. (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature*, *444*(7118), 455-460.

Kampinga, H. H., & Craig, E. A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature reviews Molecular cell biology*, *11*(8), 579-592.

Kanemori, M. A. S. A. A. K. I., Nishihara, K. A. Z. U. Y. O., Yanagi, H. I. D. E. K. I., & Yura, T. A. K. A. S. H. I. (1997). Synergistic roles of HslVU and other ATP-dependent proteases in controlling in vivo turnover of sigma32 and abnormal proteins in Escherichia coli. *Journal of bacteriology*, *179*(23), 7219-7225.

Kanemori, M., Yanagi, H., & Yura, T. (1999). The ATP-dependent HslVU/ClpQY protease participates in turnover of cell division inhibitor SulA in Escherichia coli. *Journal of bacteriology*, *181*(12), 3674-3680.

Kang, S. G., Maurizi, M. R., Thompson, M., Mueser, T., & Ahvazi, B. (2004). Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *Journal of structural biology*, *148*(3), 338-352.

Kar, N. P., Sikriwal, D., Rath, P., Choudhary, R. K., & Batra, J. K. (2008). Mycobacterium tuberculosis ClpC1: Characterization and role of the N-terminal domain in its function. *The FEBS journal*, *275*(24), 6149-6158.

Karpagam, K. B., & Ganesh, B. (2020). Leptospirosis: a neglected tropical zoonotic infection of public health importance—an updated review. *European Journal of Clinical Microbiology & Infectious Diseases*, *39*(5), 835-846.

Katayama-Fujimura, Y., Gottesman, S., & Maurizi, M. R. (1987). A multiple-component, ATP-dependent protease from Escherichia coli. *Journal of Biological Chemistry*, *262*(10), 4477-4485.

Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W. P., & Maurizi, M. R. (1988). The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component. *Journal of Biological Chemistry*, *263*(29), 15226-15236.

Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W. P., & Maurizi, M. R. (1988). The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component. *Journal of Biological Chemistry*, *263*(29), 15226-15236.

Keiler, K. C., Waller, P. R., & Sauer, R. T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, *271*(5251), 990-993.

Kessel, M., Maurizi, M.R., Kim, B., Kocsis, E., Trus, B.L., et al., 1995. Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26S proteasome. J. Mol. Biol. 250, 587–594.

Khattar, M. M. (1997). Overexpression of the hslVU operon suppresses SOS-mediated inhibition of cell division in Escherichia coli. *FEBS letters*, *414*(2), 402-404.

TH-3792_206106022

Kihara, A., Akiyama, Y., & Ito, K. (1996). A protease complex in the Escherichia coli plasma membrane: HflKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. *The EMBO journal*, *15*(22), 6122-6131.

Kim, D. Y., & Kim, K. K. (2008). The structural basis for the activation and peptide recognition of bacterial ClpP. *Journal of molecular biology*, *379*(4), 760-771.

Kim, M. J. (2019). Historical review of leptospirosis in the Korea (1945–2015). *Infection & Chemotherapy*, *51*(3), 315-329.

Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M., & Ulrich Hartl, F. (2013). Molecular chaperone functions in protein folding and proteostasis. *Annual review of biochemistry*, *82*(1), 323-355.

Kim, Y. I., Levchenko, I., Fraczkowska, K., Woodruff, R. V., Sauer, R. T., & Baker, T. A. (2001). Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nature structural biology*, *8*(3), 230-233.

Kimber, M. S., Yu, A. Y. H., Borg, M., Leung, E., Chan, H. S., & Houry, W. A. (2010). Structural and theoretical studies indicate that the cylindrical protease ClpP samples extended and compact conformations. *Structure*, *18*(7), 798-808.

Kirsch, V. C., Fetzer, C., & Sieber, S. A. (2020). Global inventory of ClpP-and ClpX-regulated proteins in Staphylococcus aureus. *Journal of Proteome Research*, *20*(1), 867-879.

Kirstein, J., Dougan, D. A., Gerth, U., Hecker, M., & Turgay, K. (2007). The tyrosine kinase McsB is a regulated adaptor protein for ClpCP. *The EMBO Journal*, *26*(8), 2061-2070.

Kirstein, J., Hoffmann, A., Lilie, H., Schmidt, R., Rübsamen-Waigmann, H., Brötz-Oesterhelt, H., ... & Turgay, K. (2009). The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO molecular medicine*, *1*(1), 37-49.

Kirstein, J., Molière, N., Dougan, D. A., & Turgay, K. (2009). Adapting the machine: adaptor proteins for Hsp100/Clp and AAA+ proteases. *Nature Reviews Microbiology*, *7*(8), 589-599.

Kirstein, J., Schlothauer, T., Dougan, D. A., Lilie, H., Tischendorf, G., Mogk, A., ... & Turgay, K. (2006). Adaptor protein controlled oligomerization activates the AAA+ protein ClpC. *The EMBO journal*, *25*(7), 1481-1491.

Kirstein, J., Zühlke, D., Gerth, U., Turgay, K., & Hecker, M. (2005). A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in B. subtilis. *The EMBO journal*, *24*(19), 3435-3445.

Kityk, R., Kopp, J., Sinning, I., & Mayer, M. P. (2012). Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. *Molecular cell*, *48*(6), 863-874.

Knudsen, B., Wower, J., Zwieb, C., & Gorodkin, J. (2001). tmRDB (tmRNA database). *Nucleic Acids Research*, *29*(1), 171-172.

Krajewska, J., Modrak-Wójcik, A., Arent, Z. J., Więckowski, D., Zolkiewski, M., Bzowska, A., & Kędzierska-Mieszkowska, S. (2017). Characterization of the molecular chaperone ClpB from the pathogenic spirochaete Leptospira interrogans. *PLoS One*, *12*(7), e0181118.

Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., ... & Bukau, B. (2002). L23 protein functions as a chaperone docking site on the ribosome. *Nature*, *419*(6903), 171-174.

Kress, W., Maglica, Ž., & Weber-Ban, E. (2009). Clp chaperone–proteases: structure and function. *Research in microbiology*, *160*(9), 618-628.

TH-3792_206106022

Kress, W., Mutschler, H., & Weber-Ban, E. (2007). Assembly pathway of an AAA+ protein: tracking ClpA and ClpAP complex formation in real time. *Biochemistry*, *46*(21), 6183-6193.

Krishnan, B. K., Balasubramanian, G., & Kumar, P. P. (2024). Leptospirosis in India: insights on circulating serovars, research lacunae and proposed strategies to control through one health approach. *One Health Outlook*, *6*(1), 11.

Krüger, E., & Hecker, M. (1998). The first gene of the Bacillus subtilis clpC operon, ctsR, encodes a negative regulator of its own operon and other class III heat shock genes. *Journal of bacteriology*, *180*(24), 6681-6688.

Krüger, E., Völker, U., & Hecker, M. (1994). Stress induction of clpC in Bacillus subtilis and its involvement in stress tolerance. *Journal of bacteriology*, *176*(11), 3360-3367.

Krüger, E., Witt, E., Ohlmeier, S., Hanschke, R., & Hecker, M. (2000). The clp proteases of Bacillus subtilis are directly involved in degradation of misfolded proteins. *Journal of Bacteriology*, *182*(11), 3259-3265.

Krzywda, S., Brzozowski, A. M., Verma, C., Karata, K., Ogura, T., & Wilkinson, A. J. (2002). The crystal structure of the AAA domain of the ATP-dependent protease FtsH of Escherichia coli at 1.5 Å resolution. *Structure*, *10*(8), 1073-1083.

Kumari, S., Ali, A., & Kumar, M. (2024). Nucleotide-induced ClpC oligomerization and its non-preferential association with ClpP isoforms of pathogenic Leptospira. *International Journal of Biological Macromolecules*, *266*, 131371.

Kumbhare, M. R., Surana, A. R., Arote, R. A., & Borse, G. D. (2019). Current status of Leptospirosis: A zoonotic tropical disease. *Int J Microbiol Curr Res*, *1*(1), 14-19.

Kuo, M. S., Chen, K. P., & Wu, W. F. (2004). Regulation of RcsA by the ClpYQ (HslUV) protease in Escherichia coli. *Microbiology*, *150*(2), 437-446.

Kuroda, A., Nomura, K., Ohtomo, R., Kato, J., Ikeda, T., Takiguchi, N., ... & Kornberg, A. (2001). Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in E. coli. *Science*, *293*(5530), 705-708.

Langklotz, S., Baumann, U., & Narberhaus, F. (2012). Structure and function of the bacterial AAA protease FtsH. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1823*(1), 40-48.

Lavey, N. P., Shadid, T., Ballard, J. D., & Duerfeldt, A. S. (2018). Clostridium difficile ClpP homologues are capable of uncoupled activity and exhibit different levels of susceptibility to acyldepsipeptide modulation. *ACS infectious diseases*, *5*(1), 79-89.

Lee, B. G., Kim, M. K., & Song, H. K. (2011). Structural insights into the conformational diversity of ClpP from Bacillus subtilis. *Molecules and cells*, *32*, 589-595.

Lee, B. G., Park, E. Y., Lee, K. E., Jeon, H., Sung, K. H., Paulsen, H., ... & Song, H. K. (2010). Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nature structural & molecular biology*, *17*(4), 471-478.

Lee, I., & Suzuki, C. K. (2008). Functional mechanics of the ATP-dependent Lon protease-lessons from endogenous protein and synthetic peptide substrates. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1784*(5), 727-735.

Leibowitz, M. J., & Soffer, R. L. (1969). A soluble enzyme from escherichiacoli which catalyzes the transfer of leucine and phenylalanine from tRNA to acceptor proteins. *Biochemical and biophysical research communications*, *36*(1), 47-53.

TH-3792_206106022

Leipe, D. D., Koonin, E. V., & Aravind, L. J. J. O. M. B. (2003). Evolution and classification of P-loop kinases and related proteins. *Journal of molecular biology*, *333*(4), 781-815.

Lemos, J. A., & Burne, R. A. (2002). Regulation and physiological significance of ClpC and ClpP in Streptococcus mutans.

Leodolter, J., Warweg, J., & Weber-Ban, E. (2015). The Mycobacterium tuberculosis ClpP1P2 protease interacts asymmetrically with its ATPase partners ClpX and ClpC1. *PloS one*, *10*(5), e0125345.

Leung, E., Datti, A., Cossette, M., Goodreid, J., McCaw, S. E., Mah, M., ... & Houry, W. A. (2011). Activators of cylindrical proteases as antimicrobials: identification and development of small molecule activators of ClpP protease. *Chemistry & biology*, *18*(9), 1167-1178.

Levchenko, I., Grant, R. A., Flynn, J. M., Sauer, R. T., & Baker, T. A. (2005). Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. *Nature structural & molecular biology*, *12*(6), 520-525.

Levchenko, I., Seidel, M., Sauer, R. T., & Baker, T. A. (2000). A specificity-enhancing factor for the ClpXP degradation machine. *Science*, *289*(5488), 2354-2356.

Levett, P. N. (2015). Leptospira. *Manual of clinical microbiology*, 1028-1036.

Li, C., Reddy, T. R., Fischer, P. M., & Dekker, L. V. (2010). A Cy5-labeled S100A10 tracer used to identify inhibitors of the protein interaction with annexin A2. *ASSAY and Drug Development Technologies*, *8*(1), 85-95.

Li, D. H. S., Chung, Y. S., Gloyd, M., Joseph, E., Ghirlando, R., Wright, G. D., ... & Ortega, J. (2010). Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chemistry & biology*, *17*(9), 959-969.

Li, D. H. S., Chung, Y. S., Gloyd, M., Joseph, E., Ghirlando, R., Wright, G. D., ... & Ortega, J. (2010). Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chemistry & biology*, *17*(9), 959-969.

Li, J. K., Liao, J. H., Li, H., Kuo, C. I., Huang, K. F., Yang, L. W., ... & Chang, C. I. (2013). The N-terminal substrate-recognition domain of a LonC protease exhibits structural and functional similarity to cytosolic chaperones. *Biological Crystallography*, *69*(9), 1789-1797.

Li, J., Soroka, J., & Buchner, J. (2012). The Hsp90 chaperone machinery: conformational dynamics and regulation by co-chaperones. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1823*(3), 624-635.

Li, M., Gustchina, A., Rasulova, F. S., Melnikov, E. E., Maurizi, M. R., Rotanova, T. V., ... & Wlodawer, A. (2010). Structure of the N-terminal fragment of Escherichia coli Lon protease. *Biological Crystallography*, *66*(8), 865-873.

Li, M., Kandror, O., Akopian, T., Dharkar, P., Wlodawer, A., Maurizi, M. R., & Goldberg, A. L. (2016). Structure and functional properties of the active form of the proteolytic complex, ClpP1P2, from Mycobacterium tuberculosis. *Journal of Biological Chemistry*, *291*(14), 7465-7476.

Li, Z., Wu, D., Zhan, B., Hu, X., Gan, J., Ji, C., & Li, J. (2019). Structural insights into the complex of trigger factor chaperone and ribosomal protein S7 from Mycobacterium tuberculosis. *Biochemical and Biophysical Research Communications*, *512*(4), 838-844.

Lies, M., & Maurizi, M. R. (2008). Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli. *Journal of Biological Chemistry*, *283*(34), 22918-22929.

Lin, Z., & Rye, H. S. (2006). GroEL-mediated protein folding: making the impossible, possible. *Critical reviews in biochemistry and molecular biology*, *41*(4), 211-239.

Liu, J., Francis, L. I., Jonas, K., Laub, M. T., & Chien, P. (2016). ClpAP is an auxiliary protease for DnaA degradation in Caulobacter crescentus. *Molecular microbiology*, *102*(6), 1075-1085.

Liu, K., Ologbenla, A., & Houry, W. A. (2014). Dynamics of the ClpP serine protease: a model for self-compartmentalized proteases. *Critical reviews in biochemistry and molecular biology*, *49*(5), 400-412.

Lo, H. H., Chang, H. C., Liao, C. T., & Hsiao, Y. M. (2022). Expression and function of clpS and clpA in Xanthomonas campestri s pv. campestris. *Antonie Van Leeuwenhoek*, *115*(5), 589-607.

Lo, J. H., Baker, T. A., & Sauer, R. T. (2001). Characterization of the N-terminal repeat domain of Escherichia coli ClpA—A class I Clp/HSP100 ATPase. *Protein Science*, *10*(3), 551-559.

Lopez, K. E., Rizo, A. N., Tse, E., Lin, J., Scull, N. W., Thwin, A. C., ... & Southworth, D. R. (2020). Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. *Nature structural & molecular biology*, *27*(5), 406-416.

Lorber, B., Fischer, F., Bailly, M., Roy, H., & Kern, D. (2012). Protein analysis by dynamic light scattering: Methods and techniques for students. *Biochemistry and molecular biology education*, *40*(6), 372-382.

Loughlin, M. F., Arandhara, V., Okolie, C., Aldsworth, T. G., & Jenks, P. J. (2009). Helicobacter pylori mutants defective in the clpP ATP-dependant protease and the chaperone clpA display reduced macrophage and murine survival. *Microbial pathogenesis*, *46*(1), 53-57.

Lourdault, K., Cerqueira, G. M., Wunder Jr, E. A., & Picardeau, M. (2011). Inactivation of clpB in the pathogen Leptospira interrogans reduces virulence and resistance to stress conditions. *Infection and immunity*, *79*(9), 3711-3717.

Lourdault, K., Cerqueira, G. M., Wunder Jr, E. A., & Picardeau, M. (2011). Inactivation of clpB in the pathogen Leptospira interrogans reduces virulence and resistance to stress conditions. *Infection and immunity*, *79*(9), 3711-3717.

Mabanglo, M. F., Leung, E., Vahidi, S., Seraphim, T. V., Eger, B. T., Bryson, S., ... & Houry, W. A. (2019). ClpP protease activation results from the reorganization of the electrostatic interaction networks at the entrance pores. *Communications biology*, *2*(1), 410.

Mahmoud, S. A., & Chien, P. (2018). Regulated proteolysis in bacteria. *Annual review of biochemistry*, *87*(1), 677-696.

Malik, I. T., & Brötz-Oesterhelt, H. (2017). Conformational control of the bacterial Clp protease by natural product antibiotics. *Natural product reports*, *34*(7), 815-831.

Mapa, K., Sikor, M., Kudryavtsev, V., Waegemann, K., Kalinin, S., Seidel, C. A., ... & Mokranjac, D. (2010). The conformational dynamics of the mitochondrial Hsp70 chaperone. *Molecular cell*, *38*(1), 89-100.

Marsee, J. D., Ridings, A., Yu, T., & Miller, J. M. (2018). Mycobacterium tuberculosis ClpC1 N-terminal domain is dispensable for adaptor protein-dependent allosteric regulation. *International journal of molecular sciences*, *19*(11), 3651.

Martin, A., Baker, T. A., & Sauer, R. T. (2007). Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. *Molecular cell*, *27*(1), 41-52.

TH-3792_206106022

Martin, A., Baker, T. A., & Sauer, R. T. (2008). Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Molecular cell*, *29*(4), 441-450.

Maurizi, M. (1992). Proteases and protein degradation in Escherichia coli. *Experientia*, *48*(2), 178-201.

Maurizi, M. R. (1991). ATP-promoted interaction between Clp A and Clp P in activation of Clp protease from Escherichia coli.

Maurizi, M. R., Clark, W. P., Katayama, Y., Rudikoff, S., Pumphrey, J., Bowers, B., & Gottesman, S. (1990). Sequence and structure of Clp P, the proteolytic component of the ATP-dependent Clp protease of Escherichia coli. *Journal of Biological Chemistry*, *265*(21), 12536-12545.

Maurizi, M. R., Singh, S. K., Thompson, M. W., Kessel, M., & Ginsburg, A. (1998). Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. *Biochemistry*, *37*(21), 7778-7786.

Maurizi, M. R., Thompson, M. W., Singh, S. K., & Kim, S. H. (1994). Endopeptidase Clp: ATP-dependent Clp protease from Escherichia coli. In *Methods in enzymology* (Vol. 244, pp. 314-331). Academic Press.

Mayer, M. P. (2010). Gymnastics of molecular chaperones. *Molecular cell*, *39*(3), 321-331.

Mei, J. M., Nourbakhsh, F., Ford, C. W., & Holden, D. W. (1997). Identification of Staphylococcus aureus virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Molecular microbiology*, *26*(2), 399-407.

Mei, Z., Wang, F., Qi, Y., Zhou, Z., Hu, Q., Li, H., ... & Shi, Y. (2009). Molecular determinants of MecA as a degradation tag for the ClpCP protease. *Journal of Biological Chemistry*, *284*(49), 34366-34375.

Meites, E., Jay, M. T., Deresinski, S., Shieh, W. J., Zaki, S. R., Tompkins, L., & Smith, D. S. (2004). Reemerging leptospirosis, California. *Emerging Infectious Diseases*, *10*(3), 406.

Melero, R., Moro, F., Pérez-Calvo, M. Á., Perales-Calvo, J., Quintana-Gallardo, L., Llorca, O., ... & Valpuesta, J. M. (2015). Modulation of the chaperone DnaK allosterism by the nucleotide exchange factor GrpE. *Journal of Biological Chemistry*, *290*(16), 10083-10092.

Miguel, P. S. B., de Carvalho Meireles, M. A., Feitosa, R. B., da Motta, O. J. R., de Oliveira Pereira, S., Junior, A. N. R., ... & Santana, L. A. (2020). Leptospirosis, a clinical update regarding a neglected infectious disease. *Revista de Patologia Tropical/Journal of Tropical Pathology*, *49*(4), 229-242.

Missiakas, D., Schwager, F., Betton, J. M., Georgopoulos, C., & Raina, S. (1996). Identification and characterization of HsIV HsIU (ClpQ ClpY) proteins involved in overall proteolysis of misfolded proteins in Escherichia coli. *The EMBO journal*, *15*(24), 6899-6909.

Mogk, A., Bukau, B., & Kampinga, H. H. (2018). Cellular handling of protein aggregates by disaggregation machines. *Molecular cell*, *69*(2), 214-226.

Mogk, A., Bukau, B., & Kampinga, H. H. (2018). Cellular handling of protein aggregates by disaggregation machines. *Molecular cell*, *69*(2), 214-226.

Mogk, A., Huber, D., & Bukau, B. (2011). Integrating protein homeostasis strategies in prokaryotes. *Cold Spring Harbor perspectives in biology*, *3*(4), a004366.

TH-3792_206106022

Moore, S. D., & Sauer, R. T. (2007). The tmRNA system for translational surveillance and ribosome rescue. *Annu. Rev. Biochem.*, *76*(1), 101-124.

Moreira, W., Ngan, G. J., Low, J. L., Poulsen, A., Chia, B. C., Ang, M. J., ... & Dick, T. (2015). Target mechanism-based whole-cell screening identifies bortezomib as an inhibitor of caseinolytic protease in mycobacteria. *MBio*, *6*(3), 10-1128.

Mróz, D., Wyszkowski, H., Szablewski, T., Zawieracz, K., Dutkiewicz, R., Bury, K., ... & Ziętkiewicz, S. (2020). CLPB (caseinolytic peptidase B homolog), the first mitochondrial protein refoldase associated with human disease. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1864*(4), 129512.

Muffler, A., Fischer, D., Altuvia, S., Storz, G., & Hengge-Aronis, R. (1996). The response regulator RssB controls stability of the sigma (S) subunit of RNA polymerase in Escherichia coli. *The EMBO journal*, *15*(6), 1333-1339.

Murray, J. K., & Gellman, S. H. (2007). Targeting protein–protein interactions: lessons from p53/MDM2. *Peptide Science: Original Research on Biomolecules*, *88*(5), 657-686.

Murray, P. R., Rosenthal, K. S., & Pfaller, M. A. (2015). *Medical Microbiology E-Book: Medical Microbiology E-Book*. Elsevier Health Sciences.

Neher, S. B., Sauer, R. T., & Baker, T. A. (2003). Distinct peptide signals in the UmuD and UmuD′ subunits of UmuD/D′ mediate tethering and substrate processing by the ClpXP protease. *Proceedings of the National Academy of Sciences*, *100*(23), 13219-13224.

Neuwald AF. (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res*, *9*, 27-43.

Ni, T., Ye, F., Liu, X., Zhang, J., Liu, H., Li, J., ... & Yang, C. G. (2016). Characterization of gain-of-function mutant provides new insights into ClpP structure. *ACS chemical biology*, *11*(7), 1964-1972.

Ninnis, R. L., Spall, S. K., Talbo, G. H., Truscott, K. N., & Dougan, D. A. (2009). Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in Escherichia coli. *The EMBO journal*, *28*(12), 1732-1744.

Nixon, P. J., Barker, M., Boehm, M., de Vries, R., & Komenda, J. (2005). FtsH-mediated repair of the photosystem II complex in response to light stress. *Journal of experimental botany*, *56*(411), 357-363.

O'Brien, E. P., Hsu, S. T. D., Christodoulou, J., Vendruscolo, M., & Dobson, C. M. (2010). Transient tertiary structure formation within the ribosome exit port. *Journal of the American Chemical Society*, *132*(47), 16928-16937.

Ogura, T., & Wilkinson, A. J. (2001). AAA+ superfamily ATPases: common structure–diverse function. *Genes to Cells*, *6*(7), 575-597.

Ogura, T., Tomoyasu, T., Yuki, T., Morimura, S., Begg, K. J., Donachie, W. D., ... & Hiraga, S. (1991). Structure and function of the ftsH gene in Escherichia coli. *Research in microbiology*, *142*(2-3), 279-282.

Olivares, A. O., Baker, T. A., & Sauer, R. T. (2016). Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nature Reviews Microbiology*, *14*(1), 33-44.

Olivares, A. O., Baker, T. A., & Sauer, R. T. (2018). Mechanical protein unfolding and degradation. *Annual review of physiology*, *80*, 413-429.

TH-3792_206106022

Pan, S., Jensen, A. A., Wood, N. A., Henrichfreise, B., Brötz-Oesterhelt, H., Fisher, D. J., ... & Ouellette, S. P. (2023). Molecular Characterization of the ClpC AAA+ ATPase in the Biology of Chlamydia trachomatis. *MBio*, *14*(2), e00075-23.

Pan, S., Malik, I. T., Thomy, D., Henrichfreise, B., & Sass, P. (2019). The functional ClpXP protease of Chlamydia trachomatis requires distinct clpP genes from separate genetic loci. *Scientific reports*, *9*(1), 14129.

Park, S. S., Kwon, H. Y., Tran, T. D. H., Choi, M. H., Jung, S. H., Lee, S., ... & Rhee, D. K. (2015). ClpL is a chaperone without auxiliary factors. *The FEBS journal*, *282*(8), 1352-1367.

Patzelt, H., Kramer, G., Rauch, T., Schönfeld, H. J., Bukau, B., & Deuerling, E. (2002). Three-state equilibrium of Escherichia coli trigger factor.

Picard, D. (2002). Heat-shock protein 90, a chaperone for folding and regulation. *Cellular and Molecular Life Sciences CMLS*, *59*, 1640-1648.

Picardeau, M. (2017). Virulence of the zoonotic agent of leptospirosis: still terra incognita? *Nature Reviews Microbiology*, *15*(5), 297-307.

Picardeau, M. (2020). Leptospira and leptospirosis. *Leptospira spp. Methods and Protocols*, 271-275.

Picardeau, M., Brenot, A., & Saint Girons, I. (2001). First evidence for gene replacement in Leptospira spp. Inactivation of L. biflexa flaB results in non-motile mutants deficient in endoflagella. *Molecular microbiology*, *40*(1), 189-199.

Powers, E. T., & Balch, W. E. (2013). Diversity in the origins of proteostasis networks—a driver for protein function in evolution. *Nature reviews Molecular cell biology*, *14*(4), 237-248.

Preissler, S., & Deuerling, E. (2012). Ribosome-associated chaperones as key players in proteostasis. *Trends in biochemical sciences*, *37*(7), 274-283.

Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., & Pearl, L. H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell*, *90*(1), 65-75.

Queraltó, C., Álvarez, R., Ortega, C., Díaz-Yáñez, F., Paredes-Sabja, D., & Gil, F. (2023). Role and regulation of Clp proteases: A target against gram-positive bacteria. *Bacteria*, *2*(1), 21-36.

Queraltó, C., Álvarez, R., Ortega, C., Díaz-Yáñez, F., Paredes-Sabja, D., & Gil, F. (2023). Role and regulation of Clp proteases: A target against gram-positive bacteria. *Bacteria*, *2*(1), 21-36.

Raju, R. M., Unnikrishnan, M., Rubin, D. H., Krishnamoorthy, V., Kandror, O., Akopian, T. N., ... & Rubin, E. J. (2012). Mycobacterium tuberculosis ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS pathogens*, *8*(2), e1002511.

Ramachandran, R., Hartmann, C., Song, H. K., Huber, R., & Bochtler, M. (2002). Functional interactions of HslV (ClpQ) with the ATPase HslU (ClpY). *Proceedings of the National Academy of Sciences*, *99*(11), 7396-7401.

Reference

Riley, E. P., Lyda, J. A., Reyes-Matte, O., Sugie, J., Kasu, I. R., Enustun, E., ... & Pogliano, K. (2025). Developmentally regulated proteolysis by MdfA and ClpCP mediates metabolic differentiation during Bacillus subtilis sporulation. *Genes & development*, *39*(7-8), 524-537.

TH-3792_206106022

Ripstein, Z. A., Vahidi, S., Houry, W. A., Rubinstein, J. L., & Kay, L. E. (2020). A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. *Elife*, *9*, e52158.

Rivera-Rivera, I., Román-Hernández, G., Sauer, R. T., & Baker, T. A. (2014). Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. *Proceedings of the National Academy of Sciences*, *111*(37), E3853-E3859.

Rizzolo, K., Yu, A. Y. H., Ologbenla, A., Kim, S. R., Zhu, H., Ishimori, K., ... & Houry, W. A. (2021). Functional cooperativity between the trigger factor chaperone and the ClpXP proteolytic complex. *Nature communications*, *12*(1), 281.

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research*, *42*(W1), W320-W324.

Roeselová, A., Maslen, S. L., Shivakumaraswamy, S., Pellowe, G. A., Howell, S., Joshi, D., ... & Balchin, D. (2024). Mechanism of chaperone coordination during cotranslational protein folding in bacteria. *Molecular Cell*, *84*(13), 2455-2471.

Rohrwild, M., Coux, O., Huang, H. C., Moerschell, R. P., Yoo, S. J., Seol, J. H., ... & Goldberg, A. L. (1996). HslV-HslU: A novel ATP-dependent protease complex in Escherichia coli related to the eukaryotic proteasome. *Proceedings of the National Academy of Sciences*, *93*(12), 5808-5813.

Rohrwild, M., Pfeifer, G., Santarius, U., Müller, S. A., Huang, H. C., Engel, A., ... & Goldberg, A. L. (1997). The ATP-dependent HslVU protease from Escherichia coli is a four-ring structure resembling the proteasome. *Nature structural biology*, *4*(2), 133-139.

Román-Hernández, G., Grant, R. A., Sauer, R. T., & Baker, T. A. (2009). Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. *Proceedings of the National Academy of Sciences*, *106*(22), 8888-8893.

Román-Hernández, G., Hou, J. Y., Grant, R. A., Sauer, R. T., & Baker, T. A. (2011). The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. *Molecular cell*, *43*(2), 217-228.

Ropelewska, M., Gross, M. H., & Konieczny, I. (2020). DNA and polyphosphate in directed proteolysis for DNA replication control. *Frontiers in Microbiology*, *11*, 585717.

Rosano, GL., Bruch, EM., and Ceccarelli, EA., Insights into the Clp/HSP100 chaperone system from chloroplasts of Arabidopsis thaliana, J Biol Chem. 286 (2011) 29671-29680, 10.1074/jbc.M110.211946

Rotanova, T. V., & Melnikov, E. E. (2010). A novel view on the architecture of the non-catalytic N-terminal region of ATP-dependent LonA proteases. *Biochemistry (Moscow) Supplement Series B: Biomedical Chemistry*, *4*, 404-408.

Rotanova, T. V., Melnikov, E. E., Khalatova, A. G., Makhovskaya, O. V., Botos, I., Wlodawer, A., & Gustchina, A. (2004). Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains. *European journal of biochemistry*, *271*(23-24), 4865-4871.

Rouquette, C., De Chastellier, C., Nair, S., & Berche, P. (1998). The ClpC ATPase of Listeria monocytogenes is a general stress protein required for virulence and promoting early bacterial escape from the phagosome of macrophages. *Molecular microbiology*, *27*(6), 1235-1245.

Saibil, H. R., Fenton, W. A., Clare, D. K., & Horwich, A. L. (2013). Structure and allostery of the chaperonin GroEL. *Journal of molecular biology*, *425*(9), 1476-1487.

TH-3792_206106022

Saikawa, N., Ito, K., & Akiyama, Y. (2002). Identification of glutamic acid 479 as the gluzincin coordinator of zinc in FtsH (HflB). *Biochemistry*, *41*(6), 1861-1868.

Sass, P., Josten, M., Famulla, K., Schiffer, G., Sahl, H. G., Hamoen, L., & Brötz-Oesterhelt, H. (2011). Antibiotic acyldepsipeptides activate ClpP peptidase to degrade the cell division protein FtsZ. *Proceedings of the National Academy of Sciences*, *108*(42), 17474-17479.

Sassetti, C. M., Boyd, D. H., & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology*, *48*(1), 77-84.

Sauer, R. T., & Baker, T. A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. *Annual review of biochemistry*, *80*(1), 587-612.

Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., ... & Baker, T. A. (2004). Sculpting the proteome with AAA+ proteases and disassembly machines. *Cell*, *119*(1), 9-18.

Scheibel, T., Weikl, T., & Buchner, J. (1998). Two chaperone sites in Hsp90 differing in substrate specificity and ATP dependence. *Proceedings of the National Academy of Sciences*, *95*(4), 1495-1499.

Schirmer, E. C., Glover, J. R., Singer, M. A., & Lindquist, S. (1996). HSP100/Clp proteins: a common mechanism explains diverse functions. *Trends in biochemical sciences*, *21*(8), 289-296.

Schlothauer, T., Mogk, A., Dougan, D. A., Bukau, B., & Turgay, K. (2003). MecA, an adaptor protein necessary for ClpC chaperone activity. *Proceedings of the National Academy of Sciences*, *100*(5), 2306-2311.

Schlünzen, F., Wilson, D. N., Tian, P., Harms, J. M., McInnes, S. J., Hansen, H. A., ... & Fucini, P. (2005). The binding mode of the trigger factor on the ribosome: implications for protein folding and SRP interaction. *Structure*, *13*(11), 1685-1694.

Schmidt, E., Obiegala, A., Imholt, C., Drewes, S., Saathoff, M., Freise, J., ... & Pfeffer, M. (2021). Influence of season, population and individual characteristics on the prevalence of Leptospira spp. in bank voles in North-West Germany. *Biology*, *10*(9), 933.

Schmidt, R., Zahn, R., Bukau, B., & Mogk, A. (2009). ClpS is the recognition component for Escherichia coli substrates of the N-end rule degradation pathway. *Molecular microbiology*, *72*(2), 506-517.

Schmitt, E. K., Riwanto, M., Sambandamurthy, V., Roggo, S., Miault, C., Zwingelstein, C., ... & Camacho, L. R. (2011). The natural product cyclomarin kills Mycobacterium tuberculosis by targeting the ClpC1 subunit of the caseinolytic protease. *Angewandte Chemie (International ed. in English)*, *50*(26), 5889-5891.

Schmitz, K. R., Carney, D. W., Sello, J. K., & Sauer, R. T. (2014). Crystal structure of Mycobacterium tuberculosis ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. *Proceedings of the National Academy of Sciences*, *111*(43), E4587-E4595.

Schuenemann, V. J., Kralik, S. M., Albrecht, R., Spall, S. K., Truscott, K. N., Dougan, D. A., & Zeth, K. (2009). Structural basis of N-end rule substrate recognition in Escherichia coli by the ClpAP adaptor protein ClpS. *EMBO reports*, *10*(5), 508-514.

Sehgal, S. C., Sugunan, A. P., & Vijayachari, P. (2003). Leptospirosis disease burden estimation and surveillance networking in India. *Southeast Asian journal of tropical medicine and public health*, *34*, 170-177.

TH-3792_206106022

Sehgal, S. C., Sugunan, A. P., & Vijayachari, P. (2003). Leptospirosis disease burden estimation and surveillance networking in India. *Southeast Asian journal of tropical medicine and public health*, *34*, 170-177.

Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas', A. F., & Gilmanshin, R. I. (1991). Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers: Original Research on Biomolecules*, *31*(1), 119-128.

Seol, J. H., Yoo, S. J., Kim, K. I., Kang, M. S., Ha, D. B., & 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. *Journal of Biological Chemistry*, *269*(47), 29468-29473.

Seong, I. S., Kang, M. S., Choi, M. K., Lee, J. W., Koh, O. J., Wang, J., ... & Chung, C. H. (2002). The C-terminal tails of HslU ATPase act as a molecular switch for activation of HslV peptidase. *Journal of Biological Chemistry*, *277*(29), 25976-25982.

Seong, I. S., Kang, M. S., Choi, M. K., Lee, J. W., Koh, O. J., Wang, J., ... & Chung, C. H. (2002). The C-terminal tails of HslU ATPase act as a molecular switch for activation of HslV peptidase. *Journal of Biological Chemistry*, *277*(29), 25976-25982.

Shamir, M., Amartely, H., Lebendiker, M., & Friedler, A. (2006). Characterization of Protein Oligomers by Multi-angle Light Scattering. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*, 1-17.

Shiau, A. K., Harris, S. F., Southworth, D. R., & Agard, D. A. (2006). Structural analysis of E. coli hsp90 reveals dramatic nucleotide-dependent conformational rearrangements. *Cell*, *127*(2), 329-340.

Shih, T. T., Sauer, R. T., & Baker, T. A. (2024). How the double-ring ClpAP protease motor grips the substrate to unfold and degrade stable proteins. *Journal of Biological Chemistry*, *300*(11).

Shimohata, N., Chiba, S., Saikawa, N., Ito, K., & Akiyama, Y. (2002). The Cpx stress response system of Escherichia coli senses plasma membrane proteins and controls HtpX, a membrane protease with a cytosolic active site. *Genes to Cells*, *7*(7), 653-662.

Shotland[a], Y., Teff, D., Koby, S., Kobiler, O., & Oppenheim, A. B. (2000). Characterization of a conserved α-helical, coiled-coil motif at the C-terminal domain of the ATP-dependent FtsH (HflB) protease of Escherichia coli. *Journal of molecular biology*, *299*(4), 953-964.

Shotland[b], Y., Shifrin, A., Ziv, T., Teff, D., Koby, S., Kobiler, O., & Oppenheim, A. B. (2000). Proteolysis of bacteriophage λ CII by Escherichia coli FtsH (HflB). *Journal of Bacteriology*, *182*(11), 3111-3116.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, *7*(1), 539.

Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., & Horwich, A. L. (1998). Structure and function in GroEL-mediated protein folding. *Annual review of biochemistry*, *67*(1), 581-608.

Singh, S. K., Grimaud, R., Hoskins, J. R., Wickner, S., & Maurizi, M. R. (2000). Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. *Proceedings of the National Academy of Sciences*, *97*(16), 8898-8903.

TH-3792_206106022

Singh, S. K., Rozycki, J., Ortega, J., Ishikawa, T., Lo, J., Steven, A. C., & Maurizi, M. R. (2001). Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis. *Journal of biological chemistry*, *276*(31), 29420-29429.

Sousa, M. C., Trame, C. B., Tsuruta, H., Wilbanks, S. M., Reddy, V. S., & McKay, D. B. (2000). Crystal and solution structures of an HslUV protease–chaperone complex. *Cell*, *103*(4), 633-643.

Story, R. M., & Steitz, T. A. (1992). Structure of the recA protein–ADP complex. *Nature*, *355*(6358), 374-376.

Subbarao, M. N., & Apirion, D. (1989). A precursor for a small stable RNA (10Sa RNA) of Escherichia coli. *Molecular and General Genetics MGG*, *217*, 499-504.

Suzuki, H., Noguchi, S., Arakawa, H., Tokida, T., Hashimoto, M., & Satow, Y. (2010). Peptide-binding sites as revealed by the crystal structures of the human Hsp40 Hdj1 C-terminal domain in complex with the octapeptide from human Hsp70. *Biochemistry*, *49*(39), 8577-8584.

Swain, J. F., Dinler, G., Sivendran, R., Montgomery, D. L., Stotz, M., & Gierasch, L. M. (2007). Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Molecular cell*, *26*(1), 27-39.

Szyk, A., & Maurizi, M. R. (2006). Crystal structure at 1.9 Å of E. coli ClpP with a peptide covalently bound at the active site. *Journal of structural biology*, *156*(1), 165-174.

Takeuchi, K., Nakatani, Y., & Hisatomi, O. (2014). Accuracy of protein size estimates based on light scattering measurements. *Open Journal of Biophysics*, *2014*.

Tan, I. S., Weiss, C. A., Popham, D. L., & Ramamurthi, K. S. (2015). A quality-control mechanism removes unfit cells from a population of sporulating bacteria. *Developmental cell*, *34*(6), 682-693.

Tasaki, T., Mulder, L. C., Iwamatsu, A., Lee, M. J., Davydov, I. V., Varshavsky, A., ... & Kwon, Y. T. (2005). A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Molecular and cellular biology*.

Tasaki, T., Zakrzewska, A., Dudgeon, D. D., Jiang, Y., Lazo, J. S., & Kwon, Y. T. (2009). The substrate recognition domains of the N-end rule pathway. *Journal of Biological Chemistry*, *284*(3), 1884-1895.

Taylor, G., Frommherz, Y., Katikaridis, P., Layer, D., Sinning, I., Carroni, M., ... & Mogk, A. (2022). Antibacterial peptide CyclomarinA creates toxicity by deregulating the Mycobacterium tuberculosis ClpC1–ClpP1P2 protease. *Journal of Biological Chemistry*, *298*(8).

Teixeira, A. F., Fernandes, L. G., Cavenague, M. F., Takahashi, M. B., Santos, J. C., Passalia, F. J., ... & Nascimento, A. L. (2019). Adjuvanted leptospiral vaccines: Challenges and future development of new leptospirosis vaccines. *Vaccine*, *37*(30), 3961-3973.

Thompson, M. W., Singh, S. K., & Maurizi, M. R. (1994). Processive degradation of proteins by the ATP-dependent Clp protease from Escherichia coli. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis. *Journal of Biological Chemistry*, *269*(27), 18209-18215.

Thomy, D., Culp, E., Adamek, M., Cheng, E. Y., Ziemert, N., Wright, G. D., ... & Brötz-Oesterhelt, H. (2019). The ADEP biosynthetic gene cluster in Streptomyces hawaiiensis NRRL 15010 reveals an accessory clpP gene as a novel antibiotic resistance factor. *Applied and environmental microbiology*, *85*(20), e01292-19.

TH-3792_206106022

Tobias, J. W., Shrader, T. E., Rocap, G., & Varshavsky, A. (1991). The N-end rule in bacteria. *Science*, *254*(5036), 1374-1377.

Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A. J., ... & Niki, H. (1995). Escherichia coli FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32. *The EMBO journal*, *14*(11), 2551-2560.

Tomoyasu, T., Yamanaka, K. U. N. I. T. O. S. H. I., Murata, K. A. Z. U. Y. O. S. H. I., Suzaki, T., Bouloc, P., Kato, A., ... & Ogura, T. (1993). Topology and subcellular localization of FtsH protein in Escherichia coli. *Journal of bacteriology*, *175*(5), 1352-1357.

Torgerson, P. R., Hagan, J. E., Costa, F., Calcagno, J., Kane, M., Martinez-Silveira, M. S., ... & Abela-Ridder, B. (2015). Global burden of leptospirosis: estimated in terms of disability adjusted life years. *PLoS neglected tropical diseases*, *9*(10), e0004122.

Torres-Delgado, A., Kotamarthi, H. C., Sauer, R. T., & Baker, T. A. (2020). The intrinsically disordered N-terminal extension of the ClpS adaptor reprograms its partner AAA+ ClpAP protease. *Journal of molecular biology*, *432*(17), 4908-4921.

Trentini, D. B., Suskiewicz, M. J., Heuck, A., Kurzbauer, R., Deszcz, L., Mechtler, K., & Clausen, T. (2016). Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature*, *539*(7627), 48-53.

Tu, G. F., Reid, G. E., Zhang, J. G., Moritz, R. L., & Simpson, R. J. (1995). C-terminal extension of truncated recombinant proteins in Escherichia coli with a 10Sa RNA decapeptide. *Journal of Biological Chemistry*, *270*(16), 9322-9326.

Tyagi, J. S., & Kinger, A. K. (1992). Identification of the 10Sa RNA structural gene of Mycobacterium tuberculosis. *Nucleic acids research*, *20*(1), 138.

UniProt Consortium. (2015). UniProt: a hub for protein information. *Nucleic acids research*, *43*(D1), D204-D212.

Vado-Solis, I., Cardenas-Marrufo, M. F., Jimenez-Delgadillo, B., Alzina-López, A., Laviada-Molina, H., Suarez-Solis, V., & Zavala-Velazquez, J. E. (2002). Clinical-epidemiological study of leptospirosis in humans and reservoirs in Yucatán, México. *Revista do Instituto de Medicina Tropical de São Paulo*, *44*, 335-340.

Van Dyck, L., Pearce, D. A., & Sherman, F. (1994). PIM1 encodes a mitochondrial ATP-dependent protease that is required for mitochondrial function in the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *269*(1), 238-242.

Varshavsky, A. (1996). The N-end rule: functions, mysteries, uses. *Proceedings of the National Academy of Sciences*, *93*(22), 12142-12149.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. *Protein science*, *20*(8), 1298-1345.

Vaughan, C. K., Gohlke, U., Sobott, F., Good, V. M., Ali, M. M., Prodromou, C., ... & Pearl, L. H. (2006). Structure of an hsp90-cdc37-cdk4 complex. *Molecular cell*, *23*(5), 697-707.

Vedadi, M., Lew, J., Artz, J., Amani, M., Zhao, Y., Dong, A., ... & Hui, R. (2007). Genome-scale protein expression and structural biology of Plasmodium falciparum and related Apicomplexan organisms. *Molecular and biochemical parasitology*, *151*(1), 100-110.

Veinger, L., Diamant, S., Buchner, J., & Goloubinoff, P. (1998). The small heat-shock protein IbpB from Escherichia coli stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *Journal of Biological Chemistry*, *273*(18), 11032-11037.

TH-3792_206106022

Vendeville, A., Larivière, D., & Fourmentin, E. (2011). An inventory of the bacterial macromolecular components and their spatial organization. *FEMS microbiology reviews*, *35*(2), 395-414.

Veronese, P. K., Rajendar, B., & Lucius, A. L. (2011). Activity of E. coli ClpA bound by nucleoside diphosphates and triphosphates. *Journal of molecular biology*, *409*(3), 333-347.

Viala, J., Rapoport, G., & Mazodier, P. (2000). The clpP multigenic family in Streptomyces lividans: conditional expression of the clpP3 clpP4 operon is controlled by PopR, a novel transcriptional activator. *Molecular Microbiology*, *38*(3), 602-612.

Vieux, E. F., Wohlever, M. L., Chen, J. Z., Sauer, R. T., & Baker, T. A. (2013). Distinct quaternary structures of the AAA+ Lon protease control substrate degradation. *Proceedings of the National Academy of Sciences*, *110*(22), E2002-E2008.

Vincent, A. T., Schiettekatte, O., Goarant, C., Neela, V. K., Bernet, E., Thibeaux, R., ... & Picardeau, M. (2019). Revisiting the taxonomy and evolution of pathogenicity of the genus Leptospira through the prism of genomics. *PLoS neglected tropical diseases*, *13*(5), e0007270.

Wah, D. A., Levchenko, I., Rieckhof, G. E., Bolon, D. N., Baker, T. A., & Sauer, R. T. (2003). Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease. *Molecular cell*, *12*(2), 355-363.

Walker, J. E., Saraste, M., Runswick, M. J., & Gay, N. J. (1982). Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO journal*, *1*(8), 945-951.

Wang, F., Mei, Z., Qi, Y., Yan, C., Hu, Q., Wang, J., & Shi, Y. (2011). Structure and mechanism of the hexameric MecA–ClpC molecular machine. *Nature*, *471*(7338), 331-335.

Wang, H., Piatkov, K. I., Brower, C. S., & Varshavsky, A. (2009). Glutamine-specific N-terminal amidase, a component of the N-end rule pathway. *Molecular cell*, *34*(6), 686-695.

Wang, J., Hartling, J. A., & Flanagan, J. M. (1997). The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. *Cell*, *91*(4), 447-456.

Wang, J., Hartling, J. A., & Flanagan, J. M. (1998). Crystal Structure Determination of Escherichia coli ClpP Starting from an EM-Derived Mask. *Journal of structural biology*, *124*(2-3), 151-163.

Wang, J., Song, J. J., Franklin, M. C., Kamtekar, S., Im, Y. J., Rho, S. H., ... & Eom, S. H. (2001). Crystal structures of the HslVU peptidase–ATPase complex reveal an ATP-dependent proteolysis mechanism. *Structure*, *9*(2), 177-184.

Wang, K. H., Román-Hernández, G., Grant, R. A., Sauer, R. T., & Baker, T. A. (2008). The molecular basis of N-end rule recognition. *Molecular cell*, *32*(3), 406-414.

Wang, K. H., Sauer, R. T., & Baker, T. A. (2007). ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes & development*, *21*(4), 403-408.

Warnasekara, J., Koralegedara, I., & Agampodi, S. (2019). Estimating the burden of leptospirosis in Sri Lanka; a systematic review. *BMC Infectious Diseases*, *19*, 1-12.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., ... & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, *46*(W1), W296-W303.

TH-3792_206106022

Watt, G., Tuazon, M. L., Santiago, E., Padre, L., Calubaquib, C., Ranoa, C., & Laughlin, L. (1988). Placebo-controlled trial of intravenous penicillin for severe and late leptospirosis. *The Lancet*, *331*(8583), 433-435.

Weber-Ban, E. U., Reid, B. G., Miranker, A. D., & Horwich, A. L. (1999). Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. *Nature*, *401*(6748), 90-93.

Wen, Z. T., Ellepola, K., & Wu, H. (2025). MecA: A Multifunctional ClpP-Dependent and Independent Regulator in Gram-Positive Bacteria. *Molecular Microbiology*.

Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., McKenney, K., & Maurizi, M. R. (1994). A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proceedings of the National Academy of Sciences*, *91*(25), 12218-12222.

Wickner, S., Maurizi, M. R., & Gottesman, S. (1999). Posttranslational quality control: folding, refolding, and degrading proteins. *Science*, *286*(5446), 1888-1893.

Wiegert, T., & Schumann, W. (2001). SsrA-mediated tagging in Bacillus subtilis. *Journal of bacteriology*, *183*(13), 3885-3889.

Wilson, D. N., & Beckmann, R. (2011). The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Current opinion in structural biology*, *21*(2), 274-282.

Wlodawer, A., Sekula, B., Gustchina, A., & Rotanova, T. V. (2022). Structure and the mode of activity of Lon proteases from diverse organisms. *Journal of molecular biology*, *434*(7), 167504.

Wojtkowiak, D., Georgopoulos, C., & Zylicz, M. (1993). Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of Escherichia coli. *Journal of Biological Chemistry*, *268*(30), 22609-22617.

Wojtyra, U. A., Thibault, G., Tuite, A., & Houry, W. A. (2003). The N-terminal zinc binding domain of ClpX is a dimerization domain that modulates the chaperone function. *Journal of biological chemistry*, *278*(49), 48981-48990.

Woo, K. M., Chung, W. J., Ha, D. B., Goldberg, A. L., & Chung, C. H. (1989). Protease Ti from Escherichia coli requires ATP hydrolysis for protein breakdown but not for hydrolysis of small peptides. *Journal of Biological Chemistry*, *264*(4), 2088-2091.

Woo, K. M., Kim, K. I., Goldberg, A. L., Ha, D. B., & Chung, C. H. (1992). The heat-shock protein ClpB in Escherichia coli is a protein-activated ATPase. *Journal of Biological Chemistry*, *267*(28), 20429-20434.

Wu, W. F., Zhou, Y., & Gottesman, S. (1999). Redundant in vivo proteolytic activities of Escherichia coli Lon and the ClpYQ (HslUV) protease. *Journal of bacteriology*, *181*(12), 3681-3687.

Xia, D., Esser, L., Singh, S. K., Guo, F., & Maurizi, M. R. (2004). Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone. *Journal of structural biology*, *146*(1-2), 166-179.

Xu, X., Wang, Y., Huang, W., Li, D., Deng, Z., & Long, F. (2024). Structural insights into the Clp protein degradation machinery. *Mbio*, *15*(4), e00031-24.

Xu, Z., Horwich, A. L., & Sigler, P. B. (1997). The crystal structure of the asymmetric GroEL–GroES–(ADP) 7 chaperonin complex. *Nature*, *388*(6644), 741-750.

TH-3792_206106022

Yakamavich, J. A., Baker, T. A., & Sauer, R. T. (2008). Asymmetric nucleotide transactions of the HslUV protease. *Journal of molecular biology*, *380*(5), 946-957.

Ye, F., Zhang, J., Liu, H., Hilgenfeld, R., Zhang, R., Kong, X., ... & Luo, C. (2013). Helix unfolding/refolding characterizes the functional dynamics of Staphylococcus aureus Clp protease. *Journal of Biological Chemistry*, *288*(24), 17643-17653.

Yoo, S. J., Seol, J. H., Shin, D. H., Rohrwild, M., Kang, M. S., Tanaka, K., ... & Chung, C. H. (1996). Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in Escherichia coli. *Journal of biological chemistry*, *271*(24), 14035-14040.

Young, J. C. (2010). Mechanisms of the Hsp70 chaperone system. *Biochemistry and Cell Biology*, *88*(2), 291-300.

Young, J. C., Obermann, W. M., & Hartl, F. U. (1998). Specific binding of tetratricopeptide repeat proteins to the C-terminal 12-kDa domain of hsp90. *Journal of Biological Chemistry*, *273*(29), 18007-18010.

Yu, A. Y. H., & Houry, W. A. (2007). ClpP: a distinctive family of cylindrical energy-dependent serine proteases. *FEBS letters*, *581*(19), 3749-3757.

Zeiler, E., List, A., Alte, F., Gersch, M., Wachtel, R., Poreba, M., ... & Sieber, S. A. (2013). Structural and functional insights into caseinolytic proteases reveal an unprecedented regulation principle of their catalytic triad. *Proceedings of the National Academy of Sciences*, *110*(28), 11302-11307.

Zeng, L. L., Yu, L., Li, Z. Y., Perrett, S., & Zhou, J. M. (2006). Effect of C-terminal truncation on the molecular chaperone function and dimerization of Escherichia coli trigger factor. *Biochimie*, *88*(6), 613-619.

Zeth, K., Ravelli, R. B., Paal, K., Cusack, S., Bukau, B., & Dougan, D. A. (2002). Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nature structural biology*, *9*(12), 906-911.

Zhang, J., Ye, F., Lan, L., Jiang, H., Luo, C., & Yang, C. G. (2011). Structural switching of Staphylococcus aureus Clp protease: a key to understanding protease dynamics. *Journal of Biological chemistry*, *286*(43), 37590-37601.

Zhang, T., Kedzierska-Mieszkowska, S., Liu, H., Cheng, C., Ganta, R. R., & Zolkiewski, M. (2013). Aggregate-reactivation activity of the molecular chaperone ClpB from Ehrlichia chaffeensis. *PLoS One*, *8*(5), e62454.

Zhou, Y., Gottesman, S., Hoskins, J. R., Maurizi, M. R., & Wickner, S. (2001). The RssB response regulator directly targets ςS for degradation by ClpXP. *Genes & development*, *15*(5), 627-637.

Zuehlke, A., & Johnson, J. L. (2010). Hsp90 and co-chaperones twist the functions of diverse client proteins. *Biopolymers: Original Research on Biomolecules*, *93*(3), 211-217.

TH-3792_206106022

Research output

Publication from thesis work

1. **Kumari, S**., Dhara, A., & Kumar, M. (2024). *Leptospira* ClpP mutant variants in association with the ClpX, acyldepsipeptide, and the trigger factor displays unprecedented gain-of-function. *International Journal of Biological Macromolecules*, *254*, 127753.
2. **Kumari, S**., Ali, A., & Kumar, M. (2024). Nucleotide-induced ClpC oligomerization and its non-preferential association with ClpP isoforms of pathogenic *Leptospira*. *International Journal of Biological Macromolecules*, *266*, 131371.
3. **Kumari, S**., & Kumar, M (2025). The ClpA chaperone and the two adaptor proteins modulate the fate of the model substrate tagged with a SsrA-degron of *Leptospira*. *Biochemical Journal, 482*(17), 1253-1275.

TH-3792_206106022

Conference presentations and workshop attended

1. **Surbhi Kumari**, Manish Kumar "Functional insight into association of ATPase chaperones with the caseinolytic protease of pathogenic Leptospira", Serine Proteases in Pericellular Proteolysis and Signaling. Oct 29, ASBMB, Rockville, Maryland (Virtual). (2021)
2. **Surbhi Kumari**, Manish Kumar "Functional insight into leptospiral caseinolytic protease and their association with ATPase chaperones", International Symposium on "Zoonotic and Transboundary Diseases: Breaking the Chain through Multidisciplinary Approach and IAVPHS, 1-2 December 2022, ICAR Research complex, Umiam, Meghalaya. (2022)
3. **Surbhi Kumari**, Manish Kumar "Site-directed ClpP mutation of Leptospira demonstrate unprecedented gain-of-function", VII Annual Convention of SVBBI and International Symposium on "Multiomics to One Health: Challenges and Way Forward in Biomedical Research", 14-15 December 2023, Division of Biochemistry, IVRI, Izatnagar, Bareilly-243122 (U.P.). (2023)
4. **Surbhi Kumari**, Manish Kumar "Molecular Characterization of the ClpC ATPase Chaperone and its Interaction with ClpP isoforms of Pathogenic Leptospira", International Conference on Frontier Areas of Science & Technology-2024 (ICFAST2024), 6 -7 September 2024. (2024)
5. Manish Kumar, **Surbhi Kumari** "Insight into the association of ClpC chaperone with the caseinoytic protese of pathogenic Leptospira", International Leptospirosis Society, 2024, Brussels, Belgium, 2-4 Sept. (2024) (Presented by thesis supervisor)
6. ICAR, NIVEDI, Bengaluru, Karnataka, India, 1-5 February 2022: participated in the National Centre for Disease Control (NCDC) sponsored Online Training Programme on "Laboratory Diagnosis of Leptospirosis".

TH-3792_206106022