

T. R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

KINETIC CHARACTERIZATION OF
GEOBACILLUS KAUSTOPHILUS
ATP-DEPENDENT LON PROTEASE

LEMAN ÖZEL
A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE
DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

GEBZE
2021

T. R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

KINETIC CHARACTERIZATION OF
GEOBACILLUS KAUSTOPHILUS
ATP-DEPENDENT LON PROTEASE

LEMAN ÖZEL
A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE
DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

THESIS SUPERVISOR
ASSIST. PROF. DR. F. İNCİ ÖZDEMİR

GEBZE

2021

**T. C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**GEOBACILLUS KAUSTOPHILUS
ATP-BAĞIMLI LON PROTEAZIN
KİNETİK KARAKTERİZASYONU**

**LEMAN ÖZEL
YÜKSEK LİSANS TEZİ
MOLEKÜLER BİYOLOJİ VE GENETİK ANABİLİM DALI**

**DANIŞMANI
DR. ÖĞR. ÜYESİ F. İNCİ ÖZDEMİR**

**GEBZE
2021**

GEBZE TEKNİK ÜNİVERSİTESİ	YÜKSEK LİSANS JÜRİ ONAY FORMU
----------------------------------	--------------------------------------

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu’nun 11/11/2021 tarih ve/..... sayılı kararıyla oluşturulan jüri tarafından 19/11/2021 tarihinde tez savunma sınavı yapılan Leman ÖZEL’in tez çalışması Moleküler Biyoloji ve Genetik Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) : Dr. Öğr. Üyesi F. İnci ÖZDEMİR

ÜYE

: Prof. Dr. Melek ÖZKAN

ÜYE

: Dr. Öğr. Üyesi Tuğrul DORUK

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu’nun

...../...../..... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

ÖZET

Prokaryotik ve ökaryotik hücrelerde, hücre içi dengenin korunması için yanlış katlanmış ya da hasarlı proteinlerin yıkımı büyük önem teşkil etmektedir. Tüm organizmalarda bu rolü ATP-bağımlı proteazlar üstlenmektedir. Lon hemen hemen tüm canlı organizmalarda bulunan yegâne ATP-bağımlı proteazdır. Lon'un hücre içindeki birçok fonksiyonu açıklanmış olmasına rağmen, enzimin mekanizması ile ilgili hala cevaplanması gereken önemli sorular bulunmaktadır. Bu amaçla termofilik bir bakteri olan *Geobacillus kaustophilus*'tan elde edilen Lon (*GkLON*) ile yapılacak çalışmalar, enzimin mekanizmasını aydınlatmak açısından önemlidir.

Bu tez çalışmasında, laboratuvarımızda daha önce *E. coli* BL21 (DE3) hücrelerine klonlanan *GkLON* kısmi olarak biyokimyasal, kinetik ve yapısal özellikleri bakımından karakterize edilmiştir. *GkLON*'un bir kimotripsin benzeri serin proteaz substratı olan Glt-Ala-Ala-Phe-MNA'ya karşı kinetik parametreleri K_M ve V_{max} sırasıyla 0.6 mM ve 134 mM/min olarak belirlenmiştir. *GkLON*'un biyokimyasal karakterizasyonu, substrat olarak α -kazein kullanılarak gerçekleştirilmiştir. Enzimin proteaz aktivitesi ATP varlığında önemli ölçüde artmıştır. Bununla birlikte *GkLON*, α -caseini CTP ve GTP varlığında daha etkin şekilde yıkıma uğratmıştır. *GkLON*'un enzimatik aktivitesinin kesin olarak iki değerlikli katyon mevcudiyetine bağlı olduğu belirlenmiştir: en yüksek aktivite Mg^{2+} varlığında gerçekleşirken Mn^{2+} , Ca^{2+} ve Co^{2+} gibi katyonlar da protein yıkımını arttırmıştır. Enzim maksimum aktiviteyi 55°C'de ATP varlığında gerçekleştirmiştir. Bir saat inkübasyon sonunda α -casein yıkımı 55°C'da %62, 37°C'da ise %25 olarak belirlenmiştir. Ayrıca homoloji modellemesine göre *Bacillus subtilis* ile *Geobacillus kaustophilus* Lon proteazları arasında %78 oranında dizi benzerliği tespit edilmiştir.

Anahtar Kelimeler: *Geobacillus kaustophilus*, ATP-Bağımlı Lon Proteaz, Kinetik Nitelendirme, Enzim Kinetiği, Biyokimyasal Nitelendirme.

SUMMARY

In prokaryotic and eukaryotic cells, the degradation of misfolded or damaged proteins is of great importance for the maintenance of homeostasis. In all organisms, this role is accomplished by ATP-dependent proteases. Lon is the only ATP-dependent protease found in almost all living organisms. Although many functions of Lon in the cell have been explained, there are still important questions to be answered regarding the mechanism of the enzyme. For this purpose, studies with Lon (*GkLON*) obtained from a thermophilic bacterium *Geobacillus kaustophilus* are important.

In this thesis, in our laboratory previously cloned *GkLON* in *E. coli* BL21 (DE3) cells was partially characterized in terms of its biochemical, kinetic, and structural properties. The kinetic parameters K_M and V_{max} of *GkLON* against Glt-Ala-Ala-Phe-MNA, a chymotrypsin-like serine protease substrate, were determined as 0.6 mM and 134 mM/min, respectively. Biochemical characterization of *GkLON* was performed using α -casein as substrate. The protease activity of *GkLON* was significantly increased in the presence of ATP. *GkLON* degraded α -casein more efficiently in the presence of CTP and GTP. The enzymatic activity of *GkLON* was determined to be strictly dependent on the presence of the divalent cation: the highest activity observed in the presence of Mg^{2+} , while cations Mn^{2+} , Ca^{2+} and Co^{2+} increased protein degradation. The enzyme achieved its maximum activity at 55°C in the presence of ATP. After one hour of incubation, α -casein degradation was determined as 62% at 55°C and 25% at 37°C. In addition, according to homology modeling, 78% sequence similarity was detected between *Bacillus subtilis* and *Geobacillus kaustophilus* Lon.

Keywords: *Geobacillus kaustophilus*, ATP-Dependent Lon Protease, Kinetic Characterization, Enzyme Kinetics, Biochemical Characterization.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Assoc. Prof. Dr. F. İnci ÖZDEMİR who has led me with her vast knowledge, experience and foresight during my master's thesis and who has not withheld her help and support.

I would like to thank Prof. Dr. Nuri ÖZTÜRK and his laboratory team from Department of Molecular Biology and Genetics at Gebze Technical University, for their help and support in the Western blot experiment; to Dr. AİŞE ÜNLÜ who is a research assistant at the Gebze Technical University, Department of Chemistry, for her support in providing protein purification system; to Dr. Ahmet TÜLEK for Docking Demonstration and Homology Modeling; to Prof Dr. Sedef TUNCA GEDİK who we share the same laboratory and did not spare us her sincerity; finally, to my dear laboratory friends Nada ALSHAWI, Elif KALPAR, Nagihan GENEL, Zeynep GİRĞİN ERSOY, Halil YILMAZ, Aslı BAYRAKTAR KAVALCI, Burcu KARAASLAN, Eda CANSU, Ğarip DEMİR and Mouktar BARKAD for their help, support and friendship.

I also would like to express my sincere thanks to dear Muhammet KILIÇ, who entered my life in the last days of my master education and made my journey beautiful and did not spare me his moral support.

I am grateful to my precious family for their contribution and dedication to becoming the person I am today. I would like to express my endless thanks to my dear father Cezmi ÖZEL, my dear mother Aysel ÖZEL and my dear brother Cevdet ÖZEL, who are always there for me with their moral support.

TABLE OF CONTENTS

	<u>Page</u>
ÖZET	v
SUMMARY	vi
ACKNOWLEDGMENTS	vii
TABLE of CONTENTS	viii
LIST of ABBREVIATION and ACRONYMS	xi
LIST of FIGURES	xiv
1. INTRODUCTION	1
1.1. Aim and Contribution	2
2. GENERAL INFORMATIONs	3
2.1. Proteolysis and Proteases	3
2.2. ATP-Dependent Proteases	4
2.3. Lon Proteases	6
2.3.1. Physiological Roles of Lon Protease	11
2.4. Biochemical Features of Lon Protease	13
3. MATERIALS AND METHODS	17
3.1. Material	17
3.1.1. Microorganism Strain and Plasmid Vector	17
3.1.2. Enzymes and Antibiotics	17
3.1.3. Protein Marker	17
3.1.4. Nucleotides	18
3.1.5. Protein Purification Materials	18
3.1.6. Protein and Peptide	19
3.1.7. Western Blot Materials	19
3.1.8. Mediums	19
3.1.8.1. Luria-Bertani (LB) Broth	19
3.1.8.2. Luria Bertani (LB) Agar	19
3.1.9. Buffers and Solutions	20
3.1.9.1. Lysis Buffer for Sonication	20

3.1.9.2. Protein Purification Buffers	20
3.1.9.2.1. Binding Buffer	20
3.1.9.2.2. Elution Buffer	20
3.1.9.3. SDS-PAGE Buffers, Solutions and Gel Materials	20
3.1.9.3.1. Running Buffer (1X)	20
3.1.9.3.2. 1 M Tris-HCl Buffer Solution	20
3.1.9.3.3. Loading Buffer (4X)	21
3.1.9.3.4. Coomassie Blue Solution	21
3.1.9.3.5. Destaining Solution	21
3.1.9.3.6. 0.5 M Tris-HCl Solution	21
3.1.9.3.7. 1.5 M Tris-HCl Solution	21
3.1.9.3.8. Separation Gel	22
3.1.9.3.9. Stacking Gel	22
3.1.9.4. Western Blot Buffers and Solutions	22
3.1.9.4.1. Blocking solution (5%)	22
3.1.9.4.2. Transfer Buffer (1X)	22
3.1.9.4.3. TBS Buffer (10X)	22
3.1.9.4.4. TBS-T (1X)	23
3.1.9.5. Protein Quantification Solutions	23
3.1.9.5.1. Reagent A of BCA Assay	23
3.1.9.5.2. Reagent B of BCA Assay	23
3.1.9.5.3. BCA Solution	23
3.1.9.6. Protease Activity Assay Buffer	23
3.1.9.6.1. Tris-HCl Solution (50 mM)	24
3.1.9.6.2. MgCl ₂ Solution (7.5 mM)	24
3.1.9.6.3. DTT (1 mM)	24
3.1.10. Laboratory Equipment	24
3.1.11. Chemicals	25
3.2. Methods	25
3.2.1. Overexpression of <i>GkLON</i> Protease	25
3.2.2. SDS-PAGE Analysis	26
3.2.3. <i>GkLON</i> Protease Purification	27
3.2.4. Determination of Protein Concentration	27

3.2.5. Western Blot Analysis of Purified <i>GkLON</i>	28
3.2.6. Kinetic Analysis of <i>GkLON</i>	28
3.2.7. Degradation of α -casein by <i>GkLON</i>	29
3.2.8. Densitometric Analysis of SDS-PAGE gel by ImageJ Program	30
3.2.9. Homology Modelling and Structural Properties of <i>GkLON</i>	30
4. RESULTS	31
4.1. Overexpression of <i>GkLON</i> Protease	31
4.2. Purification of <i>GkLON</i> Protease	32
4.3. Western Blot Analysis of <i>GkLON</i> Protease	34
4.4. Kinetic Characterization of <i>GkLON</i> Protease	34
4.5. Biochemical Characterization of <i>GkLON</i> Protease	35
4.5.1. Effect of ATP Concentration on the <i>GkLON</i> Proteolytic Activity	36
4.5.2. Time-course Degradation of α -casein by <i>GkLON</i>	38
4.5.3. Nucleotide Specificity of <i>GkLON</i>	40
4.5.4. Effects of Cations on the <i>GkLON</i> Proteolytic Activity	42
4.6. Homology Modelling and Structural Analysis of <i>GkLON</i>	44
5. CONCLUSION	47
REFERENCES	52
BIOGRAPHY	62
APPENDICES	63

LIST OF ABBREVIATIONS AND ACRONYMS

<u>Abbreviations and Acronyms</u>	<u>Explanations</u>
~	: Approximately
ϕ	: Hydrophobic amino acid
α	: Alpha
β	: Beta
λ	: Lambda
AAA+	: ATPases Associated with a variety of cellular Activities
ADP	: Adenosine diphosphate
A ₁ /Lon	: Lon protease from <i>Archaeoglobus fulgidus</i>
Ala	: Alanine amino acid
AMP-PNP	: Adenylyl-imino-diphosphate (non-hydrolyzable analogue of ATP)
APS	: Ammonium persulfate
Arg	: Arginine amino acid
Asp	: Aspartic acid
ATP	: Adenosine triphosphate
ATP γ S	: Labeled non-hydrolyzable analogue of ATP
ATPase	: Enzyme subunit which catalyze ATP
Bp	: Base pair
BCA	: Bicinchoninic acid
Bt-Lon	: Lon protease from <i>Brevibacillus thermoruber</i>
C-	: Carboxy terminal of the enzyme
Ca ²⁺	: Calcium cation
Co ²⁺	: Cobalt cation
CTP	: Cytidine triphosphate
dH ₂ O	: Distilled water
ddH ₂ O	: Double distilled water
DNA	: Deoxyribonucleic acid
dsDNA	: Double stranded DNA

DTT	: Dithiothreitol
<i>Ec</i>	: <i>Escherichia coli</i>
<i>E. coli</i>	: <i>Escherichia coli</i>
EtOH	: Ethanol (Ethyl alcohol)
FtsH	: Filamentation temperature sensitive protein H
g	: Gram
<i>Gk</i>	: <i>Geobacillus kaustophilus</i>
<i>GkLON</i>	: Lon protease from <i>Geobacillus kaustophilus</i>
Glt	: Glutaryl
GTP	: Guanosine triphosphate
HCl	: Hydrochloric acid
His	: Histidine amino acid
HslVU	: Heat shock proteins HslV and HslU
kDa	: Kilo dalton
K_M	: Affinity of the substrate
LB	: Luria broth
Lon	: Lon protease
<i>Lon_{TK}</i>	: Lon protease from <i>Thermococcus kodakaraensis</i>
Lys	: Lysine amino acid
μ	: Micro
μ L	: Microliter
M	: Molar
mM	: Mini molar
Mg	: Milligram
Mg^{2+}	: Magnesium cation
mL	: Milliliter
Mn^{2+}	: Manganese cation
MNA	: Methoxy- β - naphthylamide
$MgCl_2$	: Magnesium chloride
N-	: Amino terminal of the enzyme
nm	: Nano meter
nmol	: Nano molar
NaOH	: Sodium hydroxide

NaCl	: Sodium chloride
Ni ²⁺	: Nickel cation
Ni-NTA	: Nickel- Nitrilotriacetic acid
PAGE	: Polyacrylamide gel electrophoresis
Phe	: Phenylalanine amino acid
Rpm	: Round per minute
s ⁻¹	: Frequency units per second
SDS	: Sodium dodecyl sulfate
Ser	: Serine amino acid
SSD	: Substrate Sensor and Discriminatory Domain
ssDNA	: Single stranded DNA
TEMED	: N,N,N',N'-Tetramethyl ethylenediamine
TaLon	: Lon protease from <i>Thermoplasma acidophilum</i>
Trp	: Tryptophan amino acid
TT-Lon	: Lon protease from <i>Thermus thermophilus</i>
UTP	: Uridine triphosphate
V _{max}	: Maximum velocity
Zn ²⁺	: Zinc cation

LIST OF FIGURES

<u>Figure No:</u>	<u>Page</u>
2. 1: Domain construction of ClpXP and Lon protease.	6
2. 2: Domain organization of LonA and LonB proteases for one subunit.	8
2. 3: Catalytic mechanism of serine proteases.	10
3. 1: PageRuler Plus Prestained Protein Ladder.	18
4. 1: Protein expression profile in <i>E. coli</i> BL21 (DE3) cells empty and carrying pET28a- <i>GkLON</i> plasmid.	31
4. 2: Protein expression profile of Lon protease.	32
4. 3: Purification profile of 6xHistidine labeled <i>GkLON</i> by Ni-NTA affinity chromatography.	33
4. 4: SDS-PAGE analysis of purified protein.	33
4. 5: Western blot analysis of <i>GkLON</i> with 6xHis tag.	34
4. 6: Kinetic parameters of <i>GkLON</i> .	35
4. 7: The effect of ATP concentration on α -casein hydrolysis by <i>GkLON</i> .	37
4. 8: Densitometric analysis of the effect of ATP concentration on α -casein hydrolysis by <i>GkLON</i> .	37
4. 9: Time and heat dependent α -casein hydrolysis by <i>GkLON</i> .	39
4. 10: Densitometric analysis of time and heat dependent α -casein degradation by <i>GkLON</i> .	39
4. 11: Nucleotide specific α -casein degradation of <i>GkLON</i> .	41
4. 12: Densitometric analysis of <i>GkLON</i> nucleotide specificity.	41
4. 13: Effect of cations on α -casein hydrolysis of <i>GkLON</i> .	43
4. 14: Densitometric analysis of effect of cations on <i>GkLON</i> α -casein hydrolysis.	43
4. 15: Homology modeling of <i>GkLON</i> and <i>BsLON</i> .	45
4. 16: Structural modelling of <i>GkLON</i> active site architecture.	46

1. INTRODUCTION

Lon protease, also known as La, is the first identified ATP-dependent serine protease from the *E. coli*. Lon proteases belong to the AAA+ protein superfamily (ATPase Associated with various cellular Activities) which also includes other ATP-dependent proteases such as FtsH, ClpAP, ClpXP, HslVU and 26S proteasome [1], [2], [3]. Lon differs from other ATP-dependent proteases since it ubiquitous from archaea to humans [4], [5], [6]. Lon protease is a soluble protein found in the cytoplasm in eubacteria and localized in the membrane in archaea but in Eukaryotes it is found in mitochondria, chloroplasts, and peroxisomes [7], [8]. Lon proteases are homo-oligomeric ring-shaped enzymes. The number of subunits in oligomers can vary between 4 and 8 depending on the organism. Members of the Lon proteases family share a high degree of sequence similarity [11].

Lon plays an important role in the protein quality control systems of cells, selectively eliminating unfolded, misfolded, and damaged and some short-lived regulatory proteins [9]. These responsibilities are performed in an ATP-dependent manner [10]. Lon protease hydrolyzes ATP, converting chemical energy into mechanical energy and using it for conformational changes of enzyme and translocation of substrate protein [11]. Since ATP is the physiological activator of peptide and protein degradation of Lon protease, it is important to understand the kinetic coordination between Lon's peptidase - protease activities and ATP requirement [12].

There are two different hypotheses about the ATP dependency of Lon protease. Some studies have concluded that ATP hydrolysis occurs just before the protein degradation and ATP is indispensable for protein degradation [6]; The findings of other studies reveal that protein degradation triggers ATP hydrolysis and ATP facilitates protein translocation and degradation [5]. In addition, a group of studies have shown that the proteolytic activity of Lon proteases depends on other nucleotides such as CTP, GTP and UTP as well as ATP [6], [9]. It is also known that Lon proteases require Mg^{2+} in addition to ATP for maximum proteolytic activity. Cofactors such as Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} and Zn^{2+} can also affect the proteolytic activity of the enzyme [9]. The affinity of the enzyme for different nucleotides and cofactors differs between organisms.

1.1. Aim and Contribution

The elucidation of the mechanism of Lon proteases has become even more important with the discovery that Lon proteases play key roles not only in the protein quality control mechanism in the cell, but also in cell physiology. The structure, functions and physiological roles of many Lon proteases belonging to various organisms have been investigated and clarified to date. However, there are still important questions to be clarified regarding the mechanism of Lon proteases belonging to the *Bacillus* species. Therefore, this study carried out with Lon protease isolated from a thermophilic bacterium *Geobacillus kaustophilus* is important to revealing the biochemistry and structure of the enzyme.

Based on this idea, it was aimed to perform biochemical, kinetic and structural analyzes of *Geobacillus kaustophilus* ATP-dependent Lon protease in this thesis. *GkLON*, was cloned into the pET (pET28a) vector and transformed into *E. coli* BL21 (DE3) cells in previous study, was partially purified by Ni-NTA method as 6xHis-tagged for use in the characterization of the enzyme in this thesis. A fluorogenic chymotrypsin-like serine protease substrate Glt-Ala-Ala-Phe-MNA peptide was used for the kinetic characterization of the enzyme. Kinetic parameters, K_M and V_{max} of *GkLON* were determined in the presence of ATP and Mg^{+2} . Biochemical characterization of the enzyme was performed using α -casein, which is a serine protease substrate. The proteolytic activity of the enzyme was examined in the presence of various divalent cations such as Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} and Zn^{2+} and various nucleotides such as ATP, ADP, CTP, GTP and UTP. In addition, homology modeling of the *GkLON* protease, whose gene sequence was determined previously in our laboratory, was performed by comparing the sequence similarity with *Bacillus subtilis* Lon protease. The structure of the active site of the *GkLON* protease was determined by docking analysis with ADP.

2. GENERAL INFORMATIONS

2.1. Proteolysis and Proteases

Proteolysis is a key mechanism for the maintenance of cellular homeostasis and the regular functioning of metabolic processes. Damaged, dysfunctional or unwanted proteins are broken down into smaller peptides or amino acids by breaking peptide bonds via proteolysis. A class of enzymes called proteases or peptidases are responsible for proteolysis [13]. Broadly, proteases are categorized as exo- (exopeptidases) and endo- (endopeptidases) peptidase.

Endopeptidases catalyze the cleavage of internal peptide bonds of proteins and release peptide chains at the end of the reaction. On the other hand, exopeptidases cleave peptide bonds at the N- and C- terminals of proteins and individual amino acids are released as the reaction product [14]. Proteases are divided into 6 groups according to MEROPS, an extensive database of more than 1000 proteases; serine proteases, cysteine proteases, metallo proteases, aspartic proteases, glutamic and threonine proteases [15]. Aspartic, glutamic, and metalloproteases use the activated water molecule for a nucleophilic attack on their substrate (act as electrophile), whereas cysteine, serine, and threonine proteases act as nucleophile and target an amino acid in the substrate [14].

Conventional proteases do not show an energy requirement for protein degradation, since hydrolysis of peptide bonds is an exogenous process according to the laws of thermodynamics. However, a group of proteases use ATP during protein breakdown and this energy utilization only serves some kinetic and regulatory functions, protein substrate is not phosphorylated or adenylated [16]. This class of proteases is called ATP-dependent proteases. This type of proteases is present in all living organisms and play major roles in intracellular proteolysis and cell homeostasis. They rid the cells of old and misfolded proteins and are also play role in the control of the stability of most regulatory proteins [17].

2.2. ATP-Dependent Proteases

Protein synthesis and degradation reactions play important roles in all living organisms in maintaining homeostasis and responding to environmental changes or developmental needs. In this process, cells have developed protein quality control mechanisms to protect functional proteins from this destruction while degrading unnecessary or damaged proteins [18]. A class of proteolytic enzymes has known as ATP-dependent proteases which belonging to AAA+ protein superfamily is responsible for protein quality control in all organisms from archaea to eukaryotes [2], [19], [20], [21], [22]. ATP-dependent proteases are responsible for the elimination of proteins that are non-functional, misfolded, mistranscribed or translated, and damaged by environmental conditions such as heat or UV exposure in both prokaryotic and eukaryotic cells [23], [24]. In addition, energy-dependent proteases play role in the degradation of short-lived regulatory proteins, regulation of the activity of transcription factors, antigen recognition, bacteriophage λ lytic cycle, and apoptosis [25].

In eukaryotic cells, protein degradation is predominantly eliminated by the multimeric 26S proteasome after being covalently modified with ubiquitin, however, ATP-dependent proteases belonging to the same protein family; Lon, ClpAP, ClpXP, HslVU and FtsH play important roles in selective protein degradation in both prokaryotic and eukaryotic cells [19], [20], [21]. All ATP-dependent proteases show a common oligomeric barrel-like structure composed of ATPase and protease domains. In ClpXP, ClpAP, and HslVU proteinases, ATPase and protease domain are located on an independent polypeptide, however in Lon and FtsH, these domains are located on a single polypeptide [26].

The oligomeric structure of the holoenzyme can be varying among organism. Generally, 6 or 7 identical subunits come together to form a hexameric or heptameric enzyme in bacteria [11], [27]. The proteolytic sites are located inside the barrel-shaped enzymes and form a narrow space. Protein substrates have to be recognized by the enzyme before they accepted to proteolytic cavity [28]. Generally, recognition of the substrate by the protease is mediated by an adapter protein and a recognition motif called degron. Degrons are located in the inner regions or terminal ends of the protein substrates. Adapter proteins recognize the substrate via degrons but

sometimes the protease can directly recognize the substrate recognition sequence without the adapter protein [18], [26], [28], [29]. After the recognition step, the protein substrates are unfolded so that they can pass through the proteolytic cavity.

The process of directing protein substrates to the proteolytic site is called translocation. Unfolding and translocation of substrate requires repetitive cycles of ATP binding, hydrolysis, ADP and free phosphate release [26], [28]. ATP cycles take place in the AAA+ ATPase ring of the proteases. Generally, cycles of ATP binding and cleavage trigger a series of conformational changes in the enzyme for substrate acceptance into the proteolytic compartment of the enzyme and allows the protein substrate to pass through the degradation pore [18], [26], [28].

The proteolytic chambers of ATP-dependent proteases contain catalytic residues that provide proteolytic activity. The number and type of catalytic actives can vary among proteases. For example, ClpP harbors the His-Asp-Ser catalytic triad in the proteolytic ring for cleavage of peptide bonds; the Lon protease Lys-Ser dyad and FtsH contain only Zn^{2+} [28], [30], [31], [32]. Catalytic residues allow the formation of temporary covalent bonds between enzyme and substrate. The substrate taken up by the enzyme is completely cleaved into smaller peptide fragments. Degradation cycle ends until another abnormal protein degradation begin [18], [26].

Protease domain and the ATPase domain have to be communicating for efficient protein degradation. All ATP-dependent proteases have linkers enable to interaction between the both domains. Linker is well conserved among FtsH proteases while highly diverse among Lon proteases [26].

The ATPase domains of ATP-dependent proteases are highly conserved. A significant degree of sequence homology is shared between the ATPase domains of different proteases. The ATP binding and hydrolysis functions of these enzymes are also predicted to be similar. The ATPase domain of all proteases contains at least one nucleotide binding pocket and an arginine residue containing the Walker A (GXXXXGKT/S) and Walker B (ϕ ₄DE) modules (X and ϕ represent any amino acid and hydrophobic amino acid, respectively) [11].

The Walker A module also known as the P-loop, is responsible for the ATPase subunit oligomerization, and the Walker B module is involved in ATP binding [26]. The AAA+ motif of all ATP-dependent proteinases consists of a large and a small unit called the α/β Rossman loop (α/β -domain) and the α -helix (α -domain) [33]. α -

domain is also known as Substrate Sensor and Discriminatory Domain (SSD) [11]. The discriminator domain helps distinguish protease-specific substrates. Because the substrates are unique to protease, the SSD region is not conserved among proteases and is highly diverse [34]. α/β - and α -domain are connected by a connector called the N-linker. The N-linker is a flexible arm that controls binding and hydrolysis of ATP to aid structural changes of the enzyme. Additionally, some proteinases contain Sensor-1 and Sensor-2 portions within the nucleotide binding cavity. Sensor-1 assists the Walker B motif, and Sensor-2 acts like an arginine finger [26].

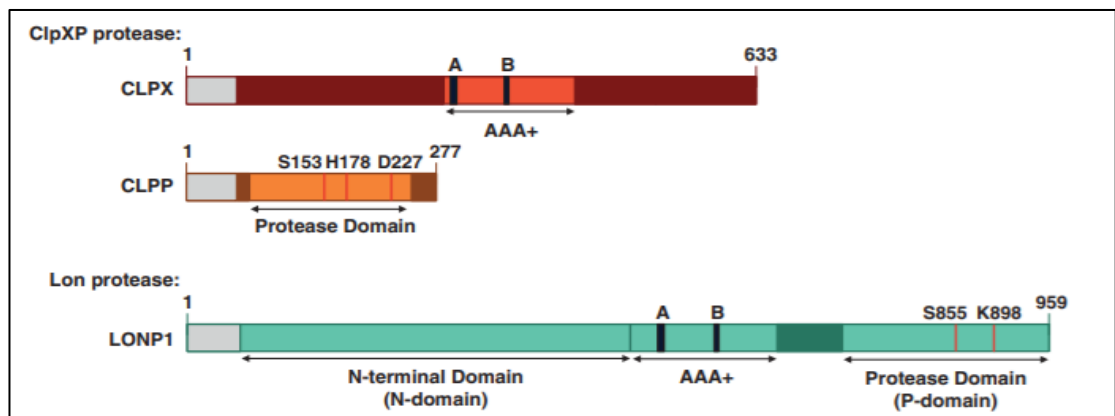


Figure 2. 1: Domain construction of ClpXP and Lon protease.

CLPX and CLPP: ATPase and protease domain of ClpXP, respectively, LONP1: Human Lon, A and B: Walker A and Walker B motif, respectively [35].

2.3. Lon Proteases

Lon protease was first discovered from *Escherichia coli* as an oligomer with 87 kDa identical subunits in 1981 and named as ‘‘La’’ [36]. Following this discovery, it was found in 1988 that Lon protease is encoded by the Lon gene that produces heat shock response [37]. After the initial discovery in *E. coli*, Lon protease has been identified in many organisms, from archaea to humans. Some of the most studied are *Bacillus brevis* [38], *Thermus thermophilus* [39], *Mycobacterium smegmatis* [40], *Brevibacillus thermoruber* [41], *Thermoplasma acidophilum* [42], and *Bacillus subtilis* [43].

The first discovery of Lon protease in eukaryotic organism was realized in 1992 with the finding of a Lon-like ATP-dependent protease in mitochondrial matrices of rat lungs [23]. Immediately after this discovery, human Lon protease was founded by cloning the human Lon gene in 1993 [44]. The finding of Lon in yeast occurred with the detection of Lon proteins belonging to *Myxococcus xanthus* in 1993 [45] and *Saccharomyces cerevisiae* in 1994 [46].

Lon protease is found in all organisms as a homo-oligomeric ring shape complex. The number of subunits in oligomers varies between 4 and 8 depending on the organism. Lon protease gene of *E. coli* encodes a 783 amino acid long tetrameric Lon protein with monomers each with a molecular weight of 94 kDa [36]. The Lon protease from *Thermus thermophilus* is a homo-octamer with each subunit 2385 bp in length and 89 kDa [39]. *Brevibacillus thermoruber* Lon protease is a 2749 bp length protein contains 779 amino acids and it's hexameric molecular weight of 549 kDa [41]. The *Bacillus subtilis* hexameric Lon protein has a total length of 774 amino acids and each monomer has a molecular weight of 87.4 kDa [43]. The Lon protease gene of *Myxococcus xanthus* encodes the hexameric Lon protein with a length of 827 amino acids, a monomeric molecular weight of 90 kDa [45]. Human Lon protease is expressed by the Lon gene located on 19th chromosome as a hexameric protein with a length of 963 amino acids, a monomer weight of 106 kDa [44], [47].

Lon protease is evolutionarily highly conserved among organisms. The sequence homologies determined in some studies are as follows. *E.coli* Lon shows 56%, 50%, 60% and 58% sequence identity with *Thermus thermophilus* [39], *Brevibacillus thermoruber* [41], *Myxococcus xanthus* [45] and human Lon proteases, respectively [44], [47]. The N-terminus of *Bacillus subtilis* and *E. coli* Lon proteases show 49.14% homology similarity [43].

Lon proteases are structurally divided into two subfamilies, LonA and LonB [48], [49], [50]. LonA is found in bacteria and humans in hexameric form, and in yeast in heptameric form. LonB proteases are generally found in all archaea and are in hexameric form [48], [49], [51].

Domain organization of these two types of Lon proteases shows differences for one subunit (Figure 2.2). Lon has a C-terminal proteolytic domain in both subfamilies. LonA enzymes carry a large N-terminal domain. LonB proteases are

membrane-bound proteases that carry a large hydrophobic transmembrane region located between the Walker A and B motifs within the AAA+ module instead of the N-terminal domain [48], [50], [52]. The transmembrane domain anchors the protease to the inner surface of the cell membrane [50], [53] and also this region is not conserved between different organisms; shows diversity [53].

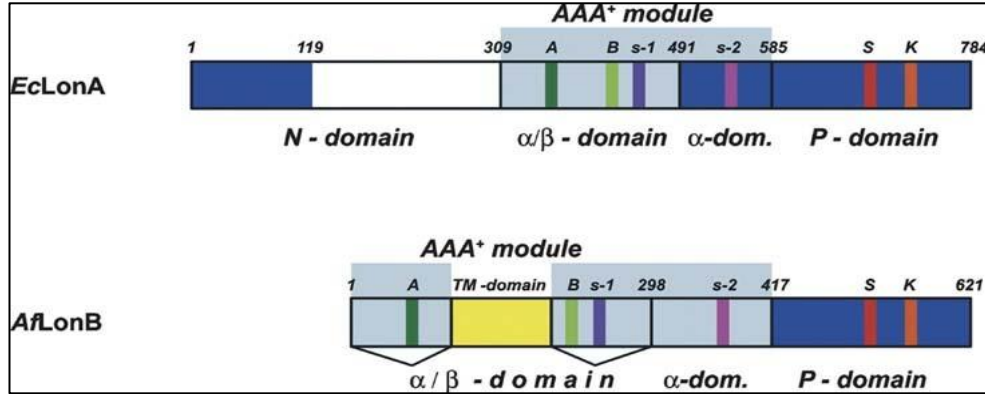


Figure 2. 2: Domain organization of LonA and LonB proteases for one subunit.

EcLonA and *AflonB*: LonA and LonB proteases from *E. coli* and *Archaeoglobus fulgidus*, respectively, TM-domain: Hydrophobic transmembrane domain, A and B: Walker A and Walker B motifs, respectively, S-1 and S-2: Sensor-1 and Sensor-2, respectively, S and K: Serine and lysine residues, respectively [50].

The N-domain of LonA proteases contains α/β - and α -domains within the AAA+ module. The α/β -domain carries Walker A, Walker B, and Sensor-1 (Figure 2.2). In LonA proteases, Walker A and B modules are found in the conserved GPPGVGKTS and P ϕ ₄DEIDK sequences, respectively (ϕ : hydrophobic amino acid, X: any an amino acid) [53], [54].

LonB proteases contain α/β - and α -domains in the AAA+ module. Walker A and Walker B motifs within the α/β -domain are linked by the transmembrane domain and Sensor-1 is located next to the Walker B motif (Figure 2.2). Walker A and B modules of LonB enzymes are represented by the GXPGXGKS ϕ and G ϕ ₄DEIXX fragments, respectively [53], [54].

In both LonA and LonB, Walker A motif consists phosphate binding loop (P-loop) required for nucleotide triphosphate binding. The P-loop consists of a conserved lysine and serine or threonine residue following the glycine-rich sequence

[54]. The α -domain of both enzymes carries Sensor-2 which contains the Arg or Lys residue. In addition, both LonA and LonB proteases have a P-domain at the C-terminal of the enzyme. The P-domain carry serine and lysine residues that provide proteolytic activity (Figure 2.2).

α/β -domain controls nucleotide binding and hydrolysis with Sensor-1, Walker A and Walker B motifs. The arginine finger, within this domain activates nucleotide hydrolysis [1], [11], [27], [53]. The helical α -domain (SSD) is responsible for recognizing protein substrates and directing them into the proteolytic cavity. SSD distinguishes forms of enzymes bound to ATP and ADP [11]. Sensor-2 plays role in ATP hydrolysis and protein substrate remodeling with the help of Arg or Lys residue [1]. The movement of α/β - and α -domains together during the nucleotide binding and hydrolysis cycles generates mechanical force acting on the associated substrates [2].

In the ATPase domain, repeated cycles of ATP binding and hydrolysis take place. This energy cycle induces a series of conformational changes within the proteolytic site of the enzyme to aiding substrate acceptance [50], [55], [56]. During conformational changes in the enzyme, the serine residue is activated. Thus, catalytic active site can recognize and binds to the substrate [50], [57].

Serine proteases have the Serine, Histidine, Aspartic acid amino acids (catalytic triad) in the active site of the enzyme. Catalytic triad is conserved in all super families of serine proteases [58], [59]. Serine proteases hydrolyze the peptide bond in two steps. Firstly, acylation reaction takes place and serine makes the nucleophilic attack to the scissile bond of the substrate. As a result of this attack, a tetrahedral intermediate is formed and then the C-terminus of the substrate is released to form a covalent acyl-enzyme. Deacylation takes place in the second step of the reaction. A water molecule enters the reaction and attacks the acyl-enzyme. A second tetrahedral intermediate is formed and then the N-terminus of the substrate is released (Figure 2.3). It is hypothesized that the histidine residue acts as a proton acceptor to activate the serine and then donates a proton to the N-terminal of the cleaved peptide [118].

Chymotrypsin-like proteases hydrolyze the peptide bonds of substrate from bulky aromatic side chains such as tyrosine, phenylalanine, and tryptophan (Figure 2.3). Lon is a chymotrypsin-like serine protease and it cleaves substrate bond from partially nonpolar adjacent amino acids, particularly phenylalanine [24], [60].

Hydrophobic, ionized side chains, such as glutamic acid, cannot attach to the proteolytic cavity [57].

Although their catalytic mechanism is common among serine proteases, there are also exceptions. One of the examples is the *E. coli* Lon protease. Botos and colleagues determined that the Ser-Lys catalytic diad is present instead of the Ser-His-Asp triad in the proteolytic region of the *E. coli* Lon protease [30]. Lon cleaves the peptide bonds of the substrate protein with the Ser-Lys catalytic dyad [30], [53], [61]. It is known that the Ser-Lys catalytic dyad of Lon proteases is highly conserved among different organisms [30], [53], [62]. In various studies, it has been revealed that mutations in the catalytic region only affect the proteolytic function, but do not prevent ATPase activity [54], [62], [63].

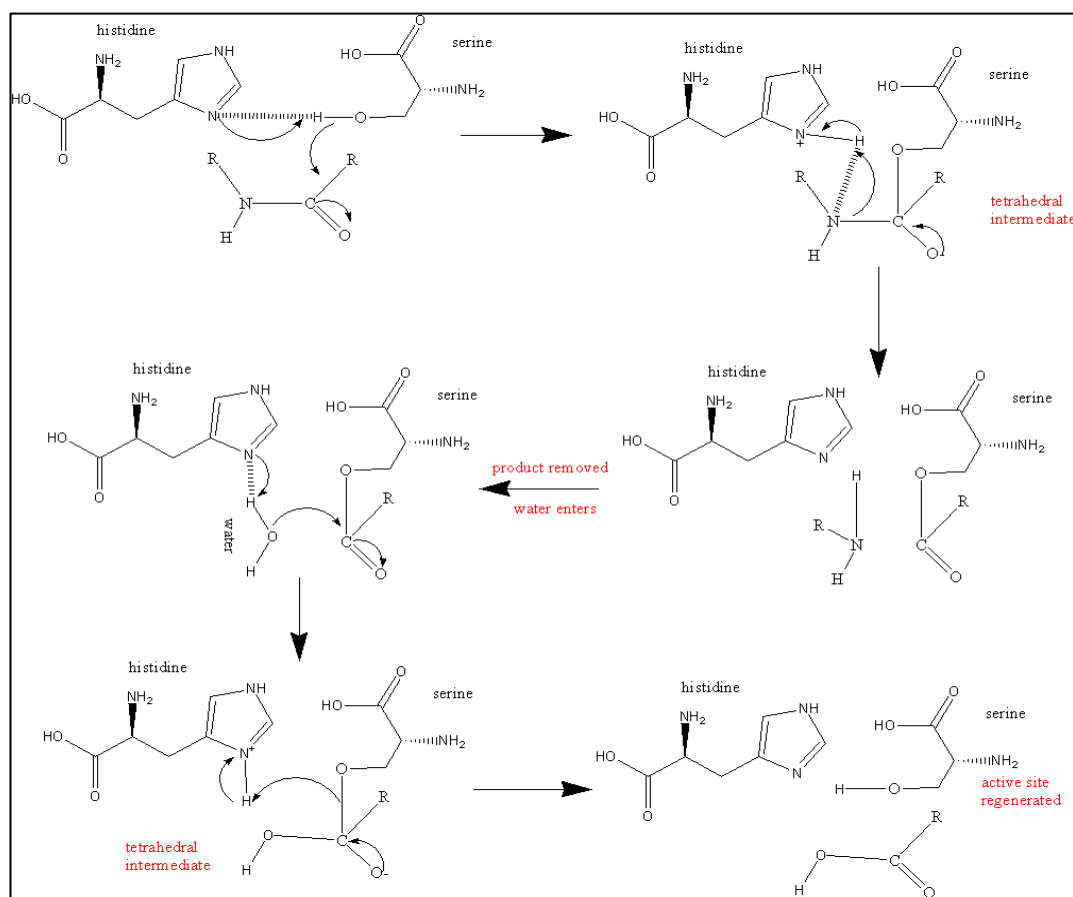


Figure 2. 3: Catalytic mechanism of serine proteases.

2.3.1. Physiological Roles of Lon Protease

Lon proteases play important roles in many biological processes in all organisms. Some of the studies investigating the roles of the enzyme in cell physiology are as follows:

Lon is required for the normal progression of cell division in many organisms. For example, *Caulobacter crescentus* needs to chromosome methylation by CcrM DNA methyltransferase to complete cell division. Regulation of the CcrM level depends on the balance between transcription of the CcrM gene and CcrM degradation by the Lon protease. If Lon protease is mutant in *C. crescentus* cell, cell division is disrupted and also the initiation of DNA replication in the cell cannot be controlled [64].

In the study, in which the name Lon was given for the first time, it was shown that many strains of *E. coli* became resistant to UV or X-rays in the presence of Lon. When Lon was mutated, the cells become sensitive to radioactive lights. For this reason, cell growth slows down and cell division errors occur [65]. Because cell growth takes longer than the wild type, this enzyme was named Lon protease, which originates from the adjective "Longer" [66].

Lon protease has the ability to bind to DNA and plays a role in the protection of DNA metabolism via this ability. For the destruction of unwanted proteins that interact with DNA, Lon forms a nucleoprotein complex [67]. Also Lon stabilizes the genome by regulating DNA replication in stress conditions such as heat shock, nutritional deficiency, osmotic, and oxidative stress [68], [69], [70].

Lon is involved in the regulation of virulence genes in various pathogenic bacteria. For example, Lon protects the intestinal epithelium of host by degrading HilC and HilD proteins. These proteins are activators of the SPI-1 virulence gene, which allows *Salmonella typhimurium* to invade the host's intestinal epithelial cells. Conversely, Lon activates the SPI-2 virulence gene, allowing *S. typhimurium* to survive within the host's macrophages [71], [72].

In some organisms, Lon can indirectly control the motility of organism. For instance, activation and inhibition of flagella biosynthesis of *Bacillus subtilis* are mediated by the proteins SwrA and SmiA, respectively. Lon protease directs the motility of *Bacillus subtilis* by acting to these two proteins [73]. Besides, *Bacillus*

subtilis Lon protease is thought to regulate motility by acting on LytC and LytD proteins, which are involved in the remodeling of peptidoglycan [74].

Another responsibility of Lon is to control mitochondrial metabolism. Lon dysfunction can trigger mitochondrial disorders [47]. Altered expression of Lon, leads to reprogramming of cell metabolism [48]. Studies shows that upregulation of Lon promote tumorigenesis by inhibiting apoptosis and promote cell proliferation [75].

Mitochondrial Lon protease plays role proliferation and apoptosis of cancer cells by acting on the Hsp60 and mtHsp70 proteins. In the upregulation of Lon, apoptosis is inhibited and Hsp60 and mtHsp70 proteins are overexpressed. Cancer cells proliferate rapidly. Loss of Lon triggers apoptosis and expression of Hsp60 and mtHsp70 proteins decrease [76].

Lon proteases also have role in the regulation of viral lytic cycle. Lambda (λ) N protein is the transcriptional antitermination factor that allows begins the viral lytic phase. λ N protein makes resistant the RNA polymerase of the host to termination signals. Expression of delayed viral early genes is triggered by Lambda N [77], [78]. Gottesman and colleagues determined that Lambda N protein is a natural physiological substrate of Lon protease. In this study, it was concluded that when the Lon protease is mutant in the *E. coli* cell, the phage enters the lytic cycle because Lambda N protein cannot be degraded [79].

Lon protease also mediates antibiotic resistance in some organisms. Lon related antibiotic resistance is regulated by the marRAB operon and the AcrAB-TolC pump in *E.coli*. MarA is natural substrates of Lon protease and *acrA*, *acrB* and *tolC* are target genes of MarA. These genes are responsible antibiotic resistance. Bhaskarla and colleagues conducted a study investigating the role of Lon protease in cell growth and antibiotic resistance in the presence of sodium salicylate (NaSal). NaSal is a compound that affects bacterial cell growth and multiple antibiotic resistance. According to this study, when Lon was mutant and high amount of NaSal was applied, the expression of MarA, *acrA*, *acrB* and *tolC* increased, and cell growth was very slow. High level of MarA was caused growth in cell filamentation [80].

Frank and coworkers investigated how the SOS response is regulated by Lon mediated proteolysis. It is known that the UmuC and UmuD proteins are associated with the SOS response. Mutations resulting from DNA damage in *E. coli* are

controlled by the activity of the UmuC and UmuD proteins. Transcription and translation level of these two proteins is controlled by Lon in the *E. coli* cell. In this study, it was shown that Lon was responsible for the degradation of UmuC and UmuD. Lon is thought to rapidly degrade UmuC and UmuD just before they mutate [81].

SulA is a physiological substrate of Lon and is also a member of SOS response regulation. LexA is the negative regulator of the SOS response. Breidenstein and colleagues observed that mutations of LexA did not affect Lon in *Pseudomonas aeruginosa*. However, when SulA was overexpressed, cell motility was restricted and cell division defects were observed in Lon mutants [82].

2.4. Biochemical Features of Lon Protease

In order to better understand the mechanisms of Lon proteases, the enzyme belonging to various organisms have been biochemically characterized in many studies. It has been demonstrated that Lon protease degrades various fluorogenic substrates such as Glt-Ala-Ala-Phe-MNA and Suc-Ala-Ala-Phe-MNA in the presence of ATP and Mg^{2+} [60], [83]. On the other hand, the degradation of peptides and proteins (like caseins) without secondary structures can be carried out without ATP.

Some of the studies performed to date to elucidate the biochemistry of Lon proteases are described below.

Watanabe and colleagues molecularly cloned the *Thermus thermophilus* Lon protease (TT-Lon) gene in 1999 and characterized the gene product. They tested the ATP-dependent peptidase activity of TT-Lon by using two different fluorogenic substrates. TT-Lon degraded the Suc-Phe-Leu-Phe-MNA substrate more effectively than the Suc-Ala-Ala-Phe-MNA fluorogenic substrate. When Suc-Phe-Leu-Phe-MNA was supplemented with α -casein, they observed 4.5-fold and 2.5-fold increase in fluorescent signal in the absence and presence of ATP, respectively. The temperature required for optimum enzymatic activity was determined as 70°C [39].

In 2000, Rudyak and coworkers investigated the peptidase activity of Lon protease isolated from *Mycobacterium smegmatis*, using α -casein and various peptides specified as rCAA's as substrate. According to their findings, peptidase

activity of Lon increased when rcAA peptides were supplemented with α -casein in the reaction. However, the usage of α -casein at very high concentrations was inhibited peptidase activity. In addition, the kinetic parameter K_M of *Mycobacterium smegmatis* Lon protease against Glt-Ala-Ala-Phe-MNA fluorogenic peptide was found as >1 mM in this study [84].

In 2002, Fukui and colleagues cloned and characterized the Lon protease from a hyperthermophilic archaea, *Thermococcus kodakaraensis* (Lon_{TK}). In this study, optimum ATPase and peptidase activity of *Thermococcus kodakaraensis* Lon (Lon_{TK}) was determined at pH 9.0 and 95 and 70°C, respectively. Lon_{TK} showed neither ATPase nor peptidase activity in the absence of Mg²⁺. It has been determined that the activity of the enzyme is strictly dependent on divalent cations such as Mg²⁺, Ni²⁺, Ca²⁺, Mn²⁺ and Co²⁺. The Lon protease of *T. kodakaraensis* does not require ATP for α -casein degradation; on the other hand hydrolysis of substrates with secondary structures such as hemoglobin A and Rubisco is completely ATP-dependent [19].

According to the study by Lee and colleagues in 2004, *Brevibacillus thermoruber* Lon (Bt-Lon) protease was able to cleave Glt-Ala-Ala-Phe-MNA fluorogenic peptide 2-3 times more efficient than Suc-Phe-Leu-Phe-MNA. The optimum temperature for the maximum peptidase and ATPase activity of Bt-Lon was determined as 50°C and 70°C, respectively [41].

Lon protease belonging *Thermoplasma acidophilum* (TaLon) was characterized in a study by Besche and colleagues in 2004. The optimum temperature, pH and MgCl₂ concentration for the highest ATPase activity were determined as 75°C, 5.75 and 10 mM, respectively. 18 different fluorogenic peptide substrates were tested to determine peptidase activity of TaLon. Increase in the fluorescent signal was detected only for Suc-Leu-Leu-Val-Tyr-AMC. The peptidase activity of TaLon was at basal level in the absence of nucleotides; however, it was decreased in the presence of ATP or non-hydrolyzable analogues AMP-PNP and ATP γ S. The peptidase activity was inhibited in the presence of ADP [61].

Makhovskaya and colleagues showed in 2007 that when the transmembrane domain (TM-domain) of the *Archaeoglobus fulgidus* LonB protease (AfLon) is deleted, the activation of the proteolytic domain of the enzyme cannot be controlled. In this study, the optimum temperature for the enzymatic activity of AfLon was determined as 70°C [85].

In 2014, Ambro and colleagues examined ATPase, peptidase, and protease activities by mutating the Gly- residue in the catalytic active site of human Lon protease. They used Glt-Ala-Ala-Phe-MNA and β -casein as substrates to determine peptidase and protease activities, respectively. Gly893 and Gly894 residues are located in the proteolytic active compartment and direct the conformational changes of the enzyme when the substrate is bound. Mutations in these residues were expected to impair peptidase and protease activity of the enzyme. Surprisingly, as a result of the mutation in the Gly894 residue, ATPase activity increased. However no change was observed in peptidase activity. It was concluded that presence of functional Gly893 and Gly894 residues is not essential for degradation of small peptides, but is absolutely necessary for β -casein degradation [86].

Jihua Pei and colleagues identified and characterized Lon-like protease (Lon-like-Ms) from *Methanobrevibacter smithii* in 2016. This Lon protease, like its eukaryotic and bacterial counterparts, contains ATPase and protease domains, but shows low sequence similarity to both LonA and LonB. For this reason, it has been named Lon-like protease. ATPase, peptidase and protease activity of Lon-like-Ms was observed toward fluorogenic peptides and casein, respectively in this study. They showed that peptidase activity of Lon-like-Ms strictly depend on divalent cations and nucleotides. It was determined that the most effective cofactor for maximum peptidase and protease activity was Mg^{2+} , but Mn^{2+} , Ca^{2+} , Co^{2+} could support enzymatic activity, respectively. Ni^{2+} and Zn^{2+} did not cause a significant change in both peptidase and protease activities. The order of efficiency of nucleotides in both peptide and casein degradation was determined as CTP>ATP>GTP>UTP. ADP, which is known to be an effective inhibitor of Lon proteases, was observed to inhibit the activity of the enzyme in this study as well. Also, the molecular weight of Lon-like-Ms was determined as 77.4 kDa. The optimum temperature for maximum enzymatic activity was found to be 37°C [9].

In 2009 and 2011, the LON gene obtained from a thermophilic bacterium, *Geobacillus kaustophilus*, was molecularly cloned and Lon protease (GkLON) biochemically characterized by the Camcı and Asal in our laboratory. GkLON has been observed to effectively hydrolyze the fluorogenic peptide Glt-Ala-Ala-Phe-MNA. According to this study, GkLON was able to perform peptide hydrolysis even in the absence of ATP, but it required Mg^{2+} and ATP together for maximum

peptidase activity. It has been observed that ADP inhibits the peptidase activity of the enzyme. The optimum temperature and pH for *GkLON* peptidase activity was determined as 55°C and 8.5, respectively [87], [88].

In this thesis, *GkLON* was biochemically and kinetically characterized in more detail. The kinetic parameters of *GkLON* against the peptide Glt-Ala-Ala-Phe-MNA were determined. Biochemical conditions such as the temperature, incubation time, the effects of nucleotides and divalent cations required for the degradation of α -casein were investigated. In addition, structural properties of *GkLON* were also examined.

3. MATERIALS AND METHODS

3.1. Material

3.1.1. Microorganism Strain and Plasmid Vector

In this study, pET28a-*GkLON* plasmid in *E. coli* BL21 (DE3) strain was used for the heterologous Lon protein expression. pET28a-*GkLON* was constructed in our laboratory previously.

3.1.2. Enzymes and Antibiotics

Kanamycin (50 µg/mL) (Caisson labs Kanamycin Sulfate) was used as selective antibiotic in the growth medium of the *E. coli* BL21 (DE3) strain bearing pET28a-*GkLON* plasmid.

Lysozyme (1 mg/mL) was applied for cell lysis before the sonication. Lysozyme was supplied from Thermo ScientificTM (Waltham, Massachusetts).

The enzymes and antibiotics used were filter-sterilized.

3.1.3. Protein Marker

PageRulerTM Plus Prestained Protein Ladder (Fermentas) (10 to 250 kDa) was used as protein marker in SDS-PAGE visualization (Figure 3.1).

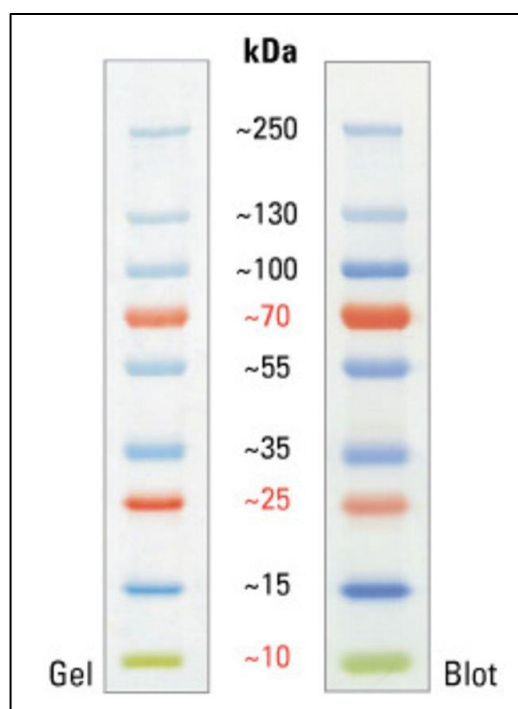


Figure 3. 1: PageRuler™ Plus Prestained Protein Ladder.

Image is belonging to band profile of PageRuler™ Plus Prestained Protein Ladder from SDS-PAGE (4-20% Tris-glycine gel) and western blot (transfer to membrane).

3.1.4. Nucleotides

The nucleotides; Adenosine-5'-triphosphate (ATP), Adenosine-5'-diphosphate (ADP), Cytidine -5'-Triphosphate (CTP), Uridine-5'-triphosphate (UTP), and Guanosine-5'-triphosphate (GTP) were purchased from (Sigma-Aldrich, USA).

3.1.5. Protein Purification Materials

Protein purification column (HisTrap™ FF 5 mL) (GE Healthcare Biosciences) and ultrafiltration centrifugal filter unit (Amicon® Ultra-3 30K) was supplied from Sigma Aldrich and MERCK, respectively.

3.1.6. Protein and Peptide

Glt-Ala-Ala-Phe-MNA chymotrypsin-like serine protease fluorogenic peptide and serine protease protein substrate α -casein (from bovine milk) was supplied from Sigma Aldrich.

3.1.7. Western Blot Materials

Primary antibody; His-Tag (27E8) Mouse mAb, used in 1:1000 dilution, Secondary antibody; Horse Anti-mouse IgG HRP-linked Antibody, used in 1:3000 dilution and Signal Fire TM ECL Reagent (Cat: 6883S), includes ECL Reagent A and B, were obtained from (Cell Signaling Technology, USA). WesternBright Sirius HRP substrate (Cat: K-12043-D10), includes Western Bright Peroxide (Component 1) and Western Bright Sirius (Component 2), was obtained from (Advansta Inc, USA).

3.1.8. Mediums

3.1.8.1. Luria-Bertani (LB) Broth

Yeast extract (0.5%) (5 g) (Biolife), tryptone (1%) (10 g) (Biolife) and NaCl (1%) (10 g) (Emsue[®]) were dissolved in double distilled water and the final volume was completed to 1 L. pH was adjusted to 7.5 and medium was sterilized via autoclave at 121°C for 20 minutes.

3.1.8.2. Luria Bertani (LB) Agar

Agar (1.5%) (BactoTM Agar, Becton Dickinson Company) was added within the LB medium (pH 7.5). LB-Agar media was sterilized by autoclaving at 121°C for 20 minutes.

3.1.9. Buffers and Solutions

3.1.9.1. Lysis Buffer for Sonication

NaH_2PO_4 (50 mM) (5.99 g), NaCl (300 mM) (17.53 g), and Imidazole (5 mM) (0.34 g) were dissolved in 1 L double distilled water and the pH was adjusted to 8.0.

3.1.9.2. Protein Purification Buffers

3.1.9.2.1. Binding Buffer

NaH_2PO_4 (20 mM) (2.39 g), NaCl (0.5 M) (29.22 g), and Imidazole (20 mM) (1.36 g) were dissolved in 1 L double distilled water and the pH was adjusted to 7.4.

3.1.9.2.2. Elution Buffer

NaH_2PO_4 (20 mM) (2.39 g), NaCl (0.5 M) (29.22 g), Imidazole (500 mM) (34.08 g) were dissolved in 1 L double distilled water and the pH was adjusted to 7.4.

3.1.9.3. SDS-PAGE Buffers, Solutions and Gel Materials

3.1.9.3.1. Running Buffer (1X)

Tris base (Tris hydroxymethyl aminomethane) (25 mM) (3.03 g), glycine (Glycine Electrophoresis Grade) (191 mM) (14.44 g), SDS (3 mM) (1 g) were dissolved in 1 L double distilled water.

3.1.9.3.2. 1 M Tris-HCl Buffer Solution

$\text{C}_4\text{H}_{11}\text{NO}_3$ (1 M) (12.11 g) was dissolved within double distilled water and pH was adjusted to 6.8 with 1% HCl. The final volume was completed to 500 mL with double distilled water.

3.1.9.3.3. Loading Buffer (4X)

Tris-HCl (1 M) (2 mL) (pH 6.8), SDS (0.8 g), 100% glycerol (4 mL), β -mercaptoethanol (14.7 M) (0.4 mL), EDTA (0.5 M) (1 mL), and bromophenol blue (1.194 M) (8 mg) were dissolved in 10 mL double distilled water.

3.1.9.3.4. Coomassie Blue Solution

Coomassie Brilliant Blue R-250 (Instant Blue[®] Coomassie Protein Stain) (1 g) was dissolved 450 mL Methanol and 100 mL Glacial acetic acid. Final volume was completed 1 L with double distilled water. Solution was filter-sterilized.

3.1.9.3.5. Destaining Solution

Methanol (50%) (v/v) and glacial acetic acid (10%) (v/v) were dissolved in double distilled water. This solution used for removing excess gel dye.

3.1.9.3.6. 0.5 M Tris-HCl Solution

$\text{C}_4\text{H}_{11}\text{NO}_3$ (0.5 M) (121.14 g) was dissolved within double distilled water and pH was adjusted 6.8 by HCl. The final volume was completed to 1 L with double distilled water.

3.1.9.3.7. 1.5 M Tris-HCl Solution

$\text{C}_4\text{H}_{11}\text{NO}_3$ (1.5 M) (121.14 g) was dissolved within double distilled water and pH was adjusted 8.8 by HCl. The final volume was completed to 1 L with double distilled water.

3.1.9.3.8. Separation Gel

Acrylamide (12%) (4 mL), (30%/0.8%/w/v), Tris-HCl (1.5 M) (2.6 mL) (pH 8.8), and SDS (10%) (100 μ L) (w/v) were dissolved within 3.2 mL double distilled water. APS (10%) (100 μ L) and TEMED (10 μ L) were added to mixture and mixture was poured between the glasses rapidly.

3.1.9.3.9. Stacking Gel

Acrylamide (12%) (0.67 mL), (30%/0.8%/w/v), Tris-HCl (0.5 M) (0.625 mL) (pH 6.8), and SDS (10%) (50 μ L) were dissolved within 2.975 mL double distilled water. APS (10%) (50 μ L) and TEMED (5 μ L) were added to mixture and mixture was poured rapidly to top of the separation gel after it was polymerized.

3.1.9.4. Western Blot Buffers and Solutions

3.1.9.4.1. Blocking solution (5%)

Non-fat dry milk (5 g) was dissolved within the 100 mL with 1XTBS-T.

3.1.9.4.2. Transfer Buffer (1X)

SDS Running buffer (10X) (100 mL), SDS (10%) (10 mL), and methanol (200 mL) were completed to 1 L with double distilled water.

3.1.9.4.3. TBS Buffer (10X)

Tris base ($C_4H_{11}NO_3$) (96.8 g) and NaCl (350.64 g) were dissolved within 4 L double distilled water and pH was adjusted to 7.6.

3.1.9.4.4. TBS-T (1X)

TBS (10X) (500 mL) and Tween 20 (5 mL) were dissolved within 5 L double distilled water.

3.1.9.5. Protein Quantification Solutions

PierceTM BCA Protein Assay Kit was used for protein quantification obtained from Thermo ScientificTM (Waltham, Massachusetts).

3.1.9.5.1. Reagent A of BCA Assay

Sodium bicinchoninic acid (BCA) (1 g), sodium carbonate (2 g) for 100 mL solution, sodium tartrate (0.16 g), NaOH (0.4 g), and sodium bicarbonate (0.95 g) were dissolved within double distilled water and pH was adjusted to 11.25.

3.1.9.5.2. Reagent B of BCA Assay

Copper sulfate (0.4 g) was dissolved within 10 mL 5H₂O dH₂O.

3.1.9.5.3. BCA Solution

50 volume BCA reagent A and 1 volume BCA reagent B were mixed with 50:1 (v/v) ratio.

3.1.9.6. Protease Activity Assay Buffer

Tris-HCl (50 mM) (pH 8.0) and DTT (1 mM) were mixed and final volume was completed to 30 mL with double distilled water. According to assay condition, MgCl₂ (7.5 mM) was added to assay buffer when it was necessary.

3.1.9.6.1. Tris-HCl Solution (50 mM)

$\text{C}_4\text{H}_{11}\text{NO}_3$ (50 mM) (0.168 g) was dissolved within double distilled water and pH was adjusted 8.0 by HCl. The final volume was completed to 30 mL with double distilled water.

3.1.9.6.2. MgCl_2 Solution (7.5 mM)

MgCl_2 (7.5 mM) was prepared from MgCl_2 (1 M) stock solution (Sigma Aldrich, USA).

3.1.9.6.3. DTT (1 mM)

DTT (1 mM) (4.62 mg) was dissolved within the 30 mL double distilled water.

3.1.10. Laboratory Equipment

+4°C and -20 °C refrigerator (Arçelik), -80 °C deep-freeze (Thermo, Forma 86°C ULT Freezer), sterile cabinet (Heraeus, HP48), incubator shaker (New Brunswick Scientific Innova[®] 40 Incubator Shaker Series), magnetic stirrer (Heidolph MR Hei-Standard), ultracentrifuge (Beckman Coulter[®] Optima[™] Max-XP Ultracentrifuge), sonicator (Branson 450 Digital Sonifier[®] Cell Disruptor), spectrometer (SHIMADZU UV-1280 UV-VIS Spectrometer), power supply (BIO-RAD, PowerPac Basic), autoclave (HIRAYAMA, HICLAHVE-50), vertical electrophoresis system (Bio-Rad Mini-Slab Cell CT), UV imaging device (Azure Biosystems, Gel Imaging System c600), ChemiDoc[™] XRS+ Imaging System (BioRad), pH meter (Hana Instruments, pH 211 Microprocessor pH Meter), water bath (TECHNE, TE-10A Tempette), ice machine (BAR-LINE BF85), fluorescence spectrophotometer (Varian Eclipse), fluorescent spectrometer (FLUOstar Omega, BMG Labtech), fluorescent spectrometer (Varioskan FLASH, Thermo Scientific[™] Waltham, Massachusetts), microwave oven (Samsung), heating block (WiseTherm). All devices used were supplied from Gebze Technical University. ÄKTA Prime Plus

Purification system (Cytiva Life Science, GE Healthcare) was supplied from Department of Chemistry at Gebze Technical University.

3.1.11. Chemicals

Tris base (Tris hydroxymethyl aminomethane), SDS, β -mercaptoethanol, APS, NaCl, methanol and acetic acid supplied from Merck; IPTG from neofroxx GmbH; glycine (Glycine Electrophoresis Grade) supplied from MP; Coomassie Brilliant Blue-R250 (Instant Blue[®] Coomassie Protein Stain) supplied from Expedeon; TEMED, EDTA and Bis-Acrylamide supplied from AppliChem; glycerol and sodium phosphate (NaH_2PO_4) supplied from Riedel-de-Haën; Imidazole, bromophenolblue, acrylamide, DL-Dithiothreitol 99% (DTT), and magnesium chloride solution (1 M) supplied from Sigma Aldrich.

3.2. Methods

GkLON gene was cloned into pET28a vector for the first time and transformed into *E. coli* BL21 (DE3) cells in the thesis study conducted by Zişan Asal, a graduate student, in our laboratory. In this thesis study, in partial biochemical characterization studies using the fluorogenic substrate Glt-Ala-Ala-Phe-MNA, it was determined that the optimum temperature of *GkLON* was 55°C and its optimum pH was 8.0. In addition, it was determined that *GkLON*, which is an ATP-dependent protease, needs both ATP and Mg^{2+} (7.5 mM) for maximum enzyme activity [88]. In the current thesis study, enzyme activity experiments were carried out under these optimum conditions.

3.2.1. Overexpression of *GkLON* Protease

Escherichia coli BL21 (DE3) cells harboring pET28a-*GkLON* plasmids was taken from fresh LB-Agar culture plates and inoculated into sterilized LB broth (10 mL) containing 50 $\mu\text{g/mL}$ kanamycin and incubated at 200 rpm at 37°C. Cell cultures grown overnight were transferred to 1 L of fresh LB medium with 1:100 dilution and incubated again at 37°C with shaking at 200 rpm. The growth of the

cultures was monitored by continuously measuring their optical density (OD) at 600 nm. Cell culture with OD 0.6 was induced with 0.5 mM IPTG to initiate gene expression and cells were incubated for an additional 4 hours. Bacterial cells were then harvested by centrifugation at 4000 rpm and +4°C for 30 min. Supernatant was discarded and cell pellets were dissolved in lysis buffer (pH 8.0) containing 5 mM imidazole and 1 mg/mL lysozyme. Samples were incubated on ice for 20 minutes. After lysis step, the cells were disrupted by sonication on ice for 2.5 minutes with amplitude of 10% with the pulse being on for 10 seconds and off for 10 seconds. Homogenate then clarified by centrifugation at +4°C, 12000 rpm for 1 hour and 30 minutes. After centrifugation, 15 mL protein supernatant was transferred to clean glass tubes stored at -20°C for purification until use. Additionally, cell extracts containing 6xHisGkLON proteins were visualized on 12% SDS-PAGE. (A similar set up was prepared for the control culture (empty *E. coli* BL21 (DE3)).

3.2.2. SDS-PAGE Analysis

SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) is a method used for separating proteins according to their molecular size. SDS-PAGE in the presence of 0.1% anionic detergent (SDS) was performed on 12% gel slabs according to the method developed by Laemmli [89]. Separation and stacking gels were prepared according to the Table 3.3. Vertical slab gel electrophoresis was carried out using Bio-Rad Mini-Slab Cell CT system (Bio-Rad Laboratories INC). 20 µL aliquots of protein samples were mixed with 5 µL 4X sample buffer containing 10% SDS and incubated in 95°C bath for 3 minutes to denature the proteins. Denaturized samples were loaded onto gel. PageRuler™ Plus Prestained Protein Ladder was used as protein marker. Gel electrophoresis was performed at 95 V within SDS-PAGE running buffer (1X) (25 mM Tris Base, 191 mM Glycine, 3 mM SDS within 1 L ddH₂O.). After the electrophoresis, proteins in the gel were visualized by Coomassie Brilliant Blue-R250 staining for an hour at room temperature. Gels were then transferred to destaining solution and rinsed several times with fresh destaining solution until excess dye was removed. Gel images were taken with a UV imaging device (Azure Biosystems, Gel Imaging System c600) and the molecular weights of the obtained protein bands were analyzed with a known standard protein molecular

marker (PageRulerTM Plus Prestained Protein Molecular Ladder (SM26619)) and image analyzer (ImageJ software).

3.2.3. *GkLON* Protease Purification

The purification of 6xHis*GkLON* proteins was performed with nickel affinity chromatography in a single step. N-terminally histidine tagged proteins from the cell free extract was purified by using HisTrapFFTM (5 mL) column with ÄKTA Prime Plus Purification system. The system was initially washed with 5 column volume (CV) 20% ethanol followed by 5 CV ultrapure distilled water. The column was equilibrated with 5 CV of binding buffer containing 20 mM imidazole at a flow rate of 1 mL/min. After the protein sample was loaded onto column, unbound proteins were washed with 5 CV of binding buffer. Specifically bound His₆-tagged *GkLON* protease was eluted with 500 mM imidazole containing elution buffer (pH 7.4). Elutes were collected in fractions of 1 mL each in glass tubes.

The elution buffer containing the phosphate is adversely affects the enzymatic activity of the *GkLON* protease. In order to change the buffer, Amicon Ultra-3 30K filter was used. The enzyme solution was washed 2.5 fold with 50 mM Tris-HCl pH 8.0 and concentrated up to 2 mL. Concentrated purified enzyme was stored at +4°C. *Lon* protease purification and concentration quality was observed by electrophoretic analysis with 12% SDS-PAGE.

3.2.4. Determination of Protein Concentration

The total protein concentrations of purified enzymes were determined by the Bicinchoninic Acid Assay (BCA) protein assay. Bovine Serum Albumin (BSA) was used as a standard in varying concentration from 20 to 2000 µg/mL. BCA protein assay was performed in a 96 well microplate. Briefly, 10 µL of each protein sample, standard and blank were mixed with 200 µL of BCA working reagent (50:1 BCA Reagent A:BCA Reagent B). Then incubated at 37°C for 30 min and after cooling down the absorbance was read in a microplate reader (FLUOstar Omega, BMG Labtech) at 562 nm. The protein concentration in each sample was calculated using

BSA standard curve. The concentration of purified and concentrated *GkLON* protease was determined by BCA assay 0.7 mg/mL and 0.9 mg/mL, respectively.

3.2.5. Western Blot Analysis of Purified *GkLON*

Following the purified protein samples running on a SDS-PAGE gel, the gel was transferred to a nitrocellulose membrane sandwich within 1X transfer buffer (with ingredients 25 mM Tris, 190 mM Glycine, 0.1% of SDS and 20% Methanol) at stable 12 V voltage for 90 min using a transblot transfer system (BIO-RAD Transblot Turbo Transfer System). After transfer, membrane was blot with 5% skimmed milk extract in 1X TBS-T solution (20 mM Tris, 150 mM NaCl, 0.05% Tween 20 content) for 1 hour at room temperature. The membrane was later exposed to primary antibody (His-Tag (27E8) Mouse mAb) in 1X TBS-T, 0.05% sodium azide overnight at 4°C. The membrane was then washed with 1X TBS-T 3 times for 5 min, 1 time for 15 min on a shaker (5-15-5-5 minutes, respectively). Then membrane was incubated with secondary antibody (Horse Anti-mouse IgG HRP-linked Antibody) in 1X TBS-T for 1 hour at room temperature. The blot was then washed with 1X TBS-T 3 times for 5 min, 1 time 15 min on a shaker (5-15-5-5 mins, respectively). After washing the membrane, detection was made in an imager analyzer (BIORAD Molecular Imager ChemiDoc™ XRS+with Image Lab™ Software) using enzyme linked chemiluminescence (ECL) detection reagents (Cell Signaling Technology). Western blot study was carried out in the laboratory of Molecular Biology and Genetics in our department which is working under the coordination of Prof. Dr. Nuri ÖZTÜRK.

3.2.6. Kinetic Analysis of *GkLON*

All studies were performed 6xHis-tagged enzyme without cleavage of the tag and the enzyme concentration was calculated on the basis of monomeric form of protein. The peptidase activity was assayed using fluorogenic peptide Glt-Ala-Ala-Phe-MNA. The purified 6xHis-*GkLON* (22 ng) was mixed with 7.5 mM MgCl₂ in a buffer of 50 mM Tris-HCl, pH 8.0, with or without 1 mM ATP. After 10 min preincubation at room temperature, the reaction was started by adding fluorogenic

peptide. The reaction was terminated on ice after 30 min incubation at 55°C. The fluorescence of the reaction mixture was measured with a fluorescence spectrophotometer (Varian Eclipse) with excitation and emission at 335 and 415 nm, respectively. One unit of peptidase activity was defined as the amount of the enzyme which released 1 nmol of MNA per min under the assay conditions. The amount of MNA released during peptidase assays was calibrated using the free compound (Bachem).

In order to determine the Michaelis-Menten constant (K_M) and maximum velocity (V_{max}) values of the enzyme, the enzymatic reactions were performed using different concentration of substrate (0-5 mM) under standard experimental conditions. The K_M and V_{max} constants of the enzyme were determined using the Hill plot obtained from growth/sigmoidal parameters with OriginLab software 2019b.

3.2.7. Degradation of α -casein by *GkLON*

The protease activity of purified *GkLON* was determined using α -casein as a substrate. α -casein is a loosely folded 30 kDa protein. 9 μ g of purified enzyme were incubated in the presence of ATP within the protease assay buffer (50 mM Tris-HCl, pH 8.0, 7.5 mM $MgCl_2$ 1 mM DTT) in a final volume of 25 μ L. The reactions were initiated by the addition of 3 μ g of α -casein and incubated at 55°C for 60 min. At the end of the incubation time period, the reaction mixture was terminated by adding with 5 μ L of 4X SDS sample buffer. Samples were subjected to electrophoresis on 12% polyacrylamide SDS gels. At the end of the electrophoresis, the gels were stained with Coomassie Brilliant Blue R-250 solution. Later on the densities of the protein bands of α -casein were analyzed with an image analyzer and quantified with ImageJ software. The densitometry results were shown relative to the initial density of the α -casein band.

3.2.8. Densitometric Analysis of SDS-PAGE gel by ImageJ Program

In this study, ImageJ software was used to quantify the density of each protein band separated with SDS-PAGE. ImageJ is a Java-based image processing program [111]. It facilitates to solve image process and analysis problems in many areas such as three dimensional cell images [112], radiological images [113], and hematology systems [114]. Area and density value of selected images can be calculated.

According to this program, the gel image was originally captured in RGB (red green blue) mode and was converted into 8-bit mode. The color intensity of each band, corresponding to the control protein bands, was quantified following subtraction of the background density. Rectangular selection tool was used to select protein bands and plot lane option was used to generate the profile plot for each band. The straight line selection was used to draw base lines of the peak of interest to create completely enclosed area. Wand tool was used to calculate the area of each selected peak. Obtained data were normalized and displayed with densitogram to determine significance of the results.

3.2.9. Homology Modelling and Structural Properties of *GkLON*

To analyse 3D homology model of *GkLON*, the structural model obtained from i-TASSER [90] was evaluated by its T_m and C-score. The C-score, provided as an estimate of the accuracy of the models generated, typically ranges between -5 and +2, with a higher value indicating higher confidence, and vice versa. The published structure of *Bacillus subtilis* (PDB ID: 3M6A) [43] was used as a template in homology modelling. The ligand, ADP (CID 6022) obtained from PubChem database was docked with the optimized model of *GkLON* using AutoDock Vina [91]. All of the generated docking poses were analyzed through Schrödinger Maestro (USA, New York) [92]. Amino acid sequences of *GkLON* and the structurally characterized Lon protease from the Research Collaboratory for Structural Bioinformatics Protein Database (RCSB PDB) were aligned with Clustal Omega (*Bacillus subtilis*, PDB ID: 3M6A) [93]. The secondary structure of the protein was predicted with the ESPript [94].

4. RESULTS

4.1. Overexpression of *GkLON* Protease

Empty and carrying the pET28a-*GkLON* plasmid *E. coli* BL21 (DE3) cells were grown in LB broth containing kanamycin (50 µg/mL) and the protein expression was stimulated with IPTG when their optical density reached 0.6. The expression level of His₆-tagged *GkLON* was checked by removing samples before and after IPTG induction. Expression did not observed in cells carrying empty vector (Figure 4.1). As seen in Fig. 4.2., *LON* gene was successfully expressed and the apparent subunit molecular weight of His₆-tagged *GkLON* about 100 kDa was determined by SDS-PAGE.

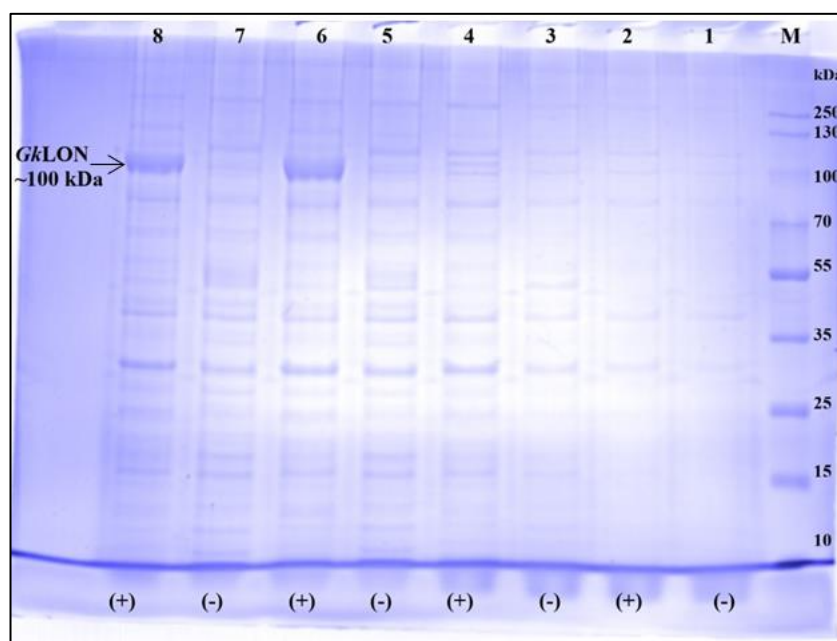


Figure 4. 1: Protein expression profile in *E. coli* BL21 (DE3) cells empty and carrying pET28a-*GkLON* plasmid.

(-) and (+): *E. coli* BL21 (DE3) cells empty and carrying pET28a-*GkLON* plasmid, respectively. 1-2: Before IPTG induction, 3-4: IPTG induction after 2h, 5-6: IPTG induction after 3h, 7-8: IPTG induction after 4h, M: Molecular Protein Marker.

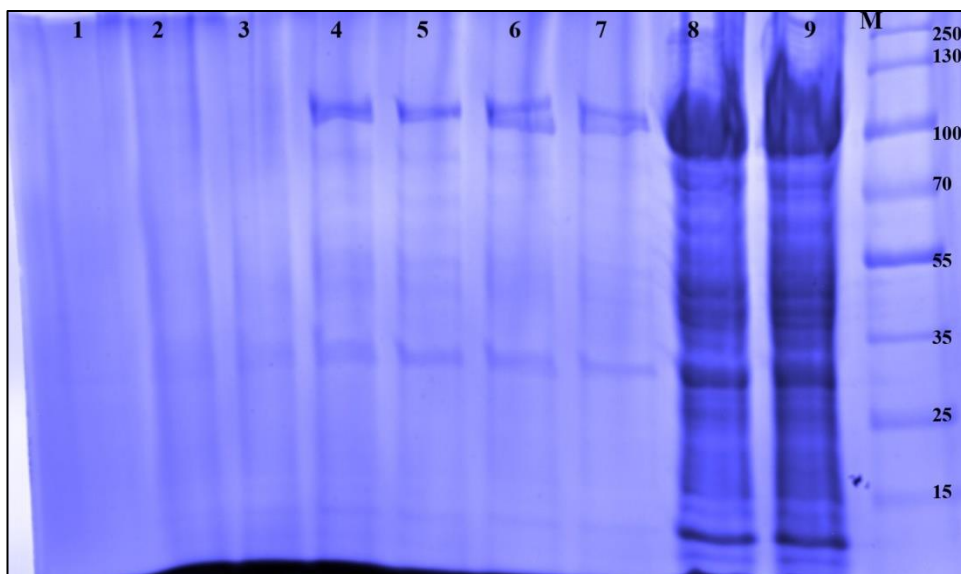


Figure 4. 2: Protein expression profile of Lon protease.

1-2: Before IPTG induction, 3: IPTG induction after 1h, 4-5: IPTG induction after 3h, 6-7: IPTG induction after 4h, 8-9: Supernatant after sonication, M: Molecular Protein Marker.

4.2. Purification of *GkLON* Protease

After confirming the expression of the recombinant protein, large scale culture growth for protein expression was carried out. Following overnight expression and subsequent lysis by sonication, obtained cell extract was loaded onto the HisTrapFFTM 5 mL column and eluted with the help of 500 mM Imidazole containing elution buffer (Figure 4.3). Fractions containing *GkLON* with a 6xHistidine labeled at the N-terminal end were collected and the protein content was measured at 280 nm during purification with the aid of a spectrophotometer integrated in the purification system. After that the buffer solution was changed from phosphate buffer solution (pH 7.4) to Tris-HCl buffer solution (pH 8.0) and concentrated with the 30000 MWCO Amicon[®] concentrator falcon. As seen in Fig. 4.4, purified *GkLON* was observed approximately 100 kDa. The purified and concentrated protein amounts were determined by BCA method as 0.7 and 0.9 mg/mL, respectively.

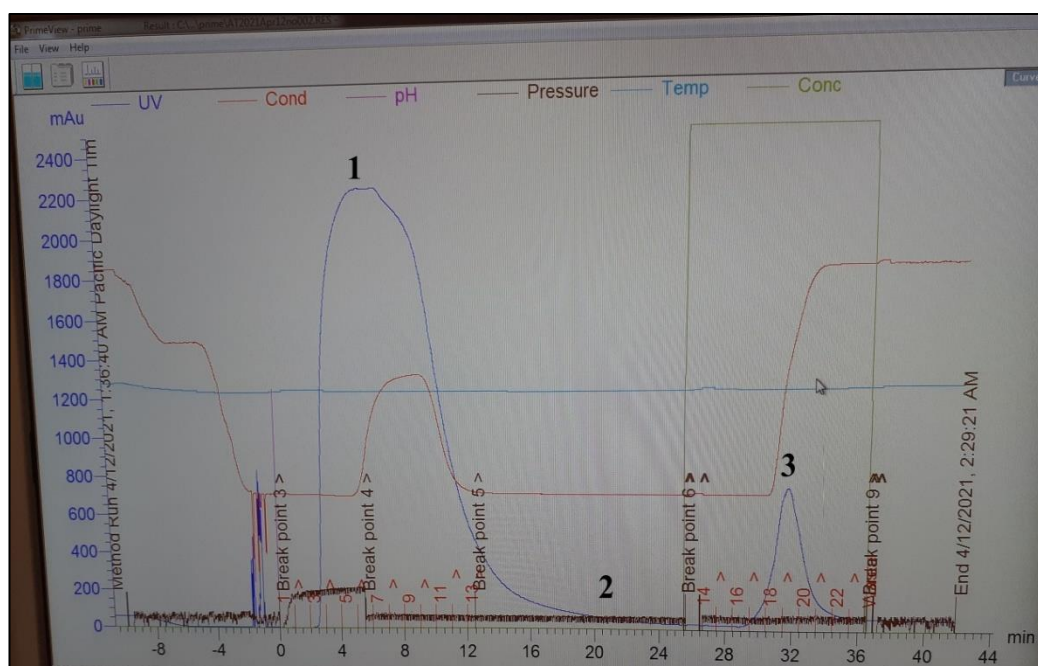


Figure 4. 3: Purification profile of 6xHistidine labeled *GkLON* by Ni-NTA affinity chromatography.

1: Unbound protein fractions to Ni-NTA resin, 2: Washing of unbound proteins with binding buffer, 3: 6xHis*GkLON* collection with elution buffer.

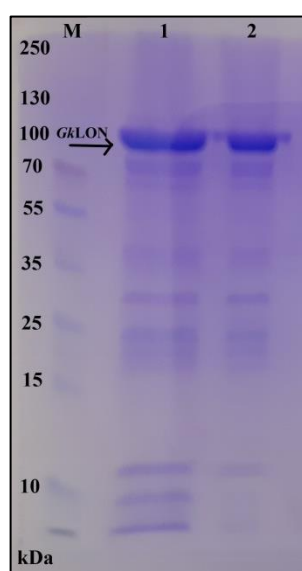


Figure 4. 4: SDS-PAGE analysis of purified protein.

M: Molecular Protein Marker, 1: Purified *GkLON*, 2: Concentrated *GkLON*.

4.3. Western Blot Analysis of *GkLON* Protease

GkLON fused with 6xHis-tagged at the N-terminal was also confirmed by Western blot analysis using anti-His antibody. The monomeric *GkLON* was determined to be ~100 kDa (Figure 4.5).

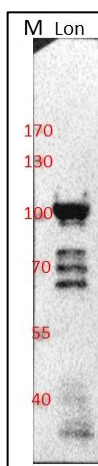


Figure 4. 5: Western blot analysis of *GkLON* with 6xHis tag.

M: Molecular Protein Marker, 1: *GkLON* (~100 kDa).

4.4. Kinetic Characterization of *GkLON* Protease

The kinetic properties of *GkLON* with histidine tag were determined using fluorogenic peptide, Glt-Ala-Ala-Phe-MNA which is the excellent substrates for mammalian chymotrypsin and related serine proteases. Glt-Ala-Ala-Phe-MNA degradation by Lon exhibited sigmoidal kinetics, indicative of cooperative allosteric behavior, and fit best to the Hill equation. K_M and V_{max} were determined as 0.6 mM and 134 mM/min respectively in the presence of 1 mM ATP (Figure 4. 6).

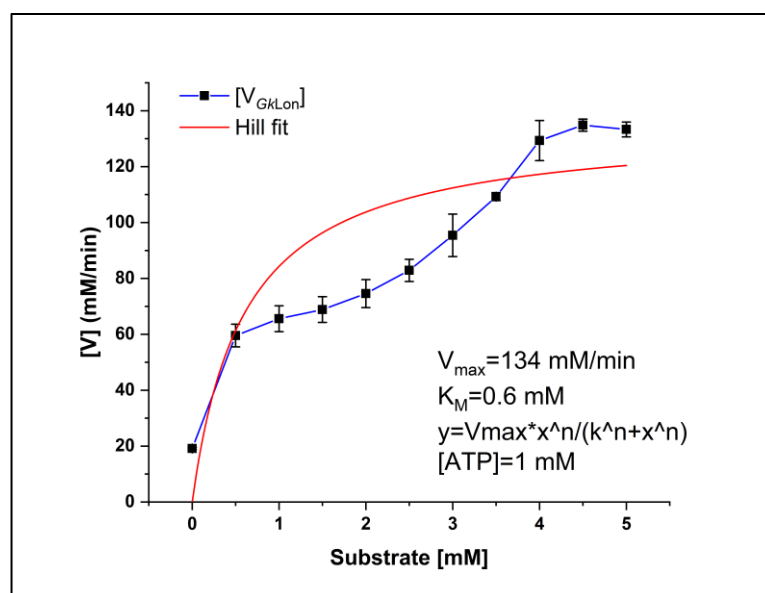


Figure 4. 6: Kinetic parameters of *GkLON*.

4.5. Biochemical Characterization of *GkLON* Protease

The protease activity of purified *GkLON* was determined using α -casein as a substrate. As mentioned in Material Method section 3.2.7 and 3.2.8, *GkLON* was incubated with α -casein in the presence (or absence) of nucleotides in the protease assay buffer at 55°C for 60 min unless mentioned. The enzyme activity of *GkLON* was determined by analyzing the degradation of α -casein on SDS slab gels. The densities of α -casein bands were measured with ImageJ analyzer software. The reaction without enzyme was used as control. The enzyme activity was expressed as the percentage of the activity resulting from the reaction without enzyme.

4.5.1. Effect of ATP Concentration on the *GkLON* Proteolytic Activity

Although Lon protease belongs to the family of ATP-dependent proteases, it is known that it can perform protein hydrolysis at the basal level without ATP. In order to determine the effect of ATP concentration on Lon protease activity, the enzyme was incubated with different concentrations of ATP at 55°C. After 1 h enzymatic reaction with α -casein, protease activity of *GkLON* was demonstrated by SDS-PAGE analysis (Figure 4.7). As the result of densitometric analysis, a relatively increased rate of substrate degradation was observed due to increasing ATP concentration (at 1 mM 32%, at 2 mM 40%, at 4 mM 42%, at 6 mM 54% and at 8 mM 57%) (Figure 4.8). However, even absence of ATP, 31% α -casein degradation was observed.

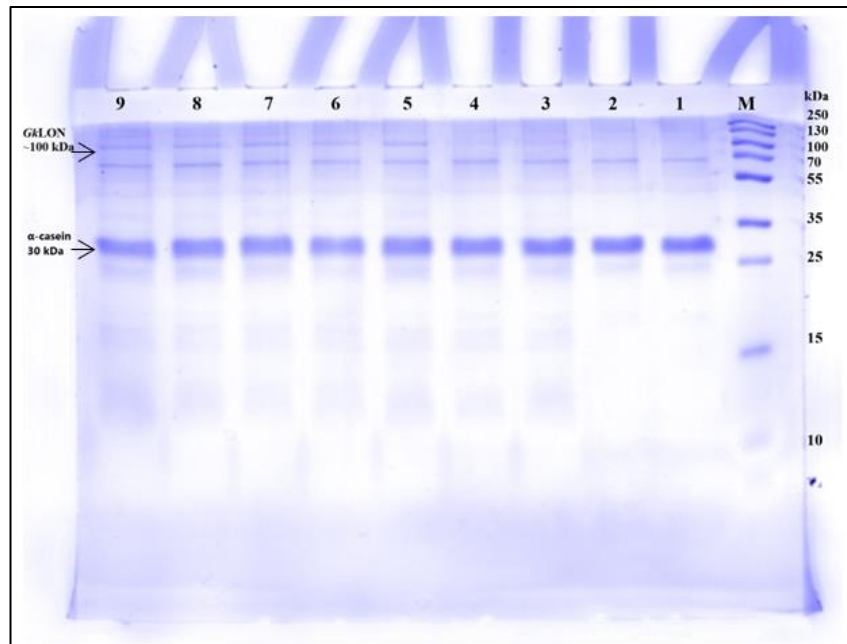


Figure 4. 7: The effect of ATP concentration on α -casein hydrolysis by *GkLON*.

M: Molecular Protein Marker, 1: only α -casein, 2: 1 mM ATP without enzyme, 3: without ATP and MgCl_2 , 4: without ATP, 5-6-7-8-9: 1, 2, 4, 6 and 8 mM ATP, respectively.

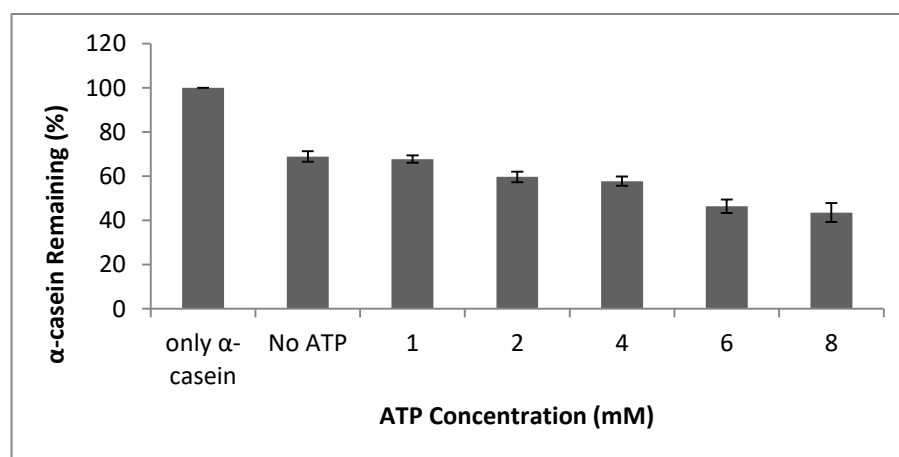


Figure 4. 8: Densitometric analysis of the effect of ATP concentration on α -casein hydrolysis by *GkLON*.

4.5.2. Time-course Degradation of α -casein by *GkLON*

Time-dependent proteolysis of α -casein was investigated with 9 μ g of purified *GkLON* incubated in the presence and absence of 4 mM ATP at 37 and 55°C with 0, 10, 30 and 60 min time intervals (Figure 4.9). According to densitometric gel analysis, *Geobacillus kaustophilus* Lon protease degraded α -casein most efficiently at 55°C. α -casein was degraded even in the absence of ATP at 55°C at the end of the 10 minutes reaction. However, in the following time courses, *GkLON* efficiently degraded α -casein in the presence of ATP in a time-dependent manner, while the enzyme activity remained stable in the absence of ATP. At the end of the 60 minutes incubation period at 55°C, degradation of α -casein by *GkLON* was determined as 62% in the presence of ATP and 38% in the absence of ATP. Densitometric analyzes performed for the experiment conducted at 37°C, there was no time-dependent change in the degradation of α -casein in the absence of ATP, besides time-dependent destruction was observed in the presence of ATP. Maximum substrate degradation was determined as 25% after 1 hour (Figure 4.10).

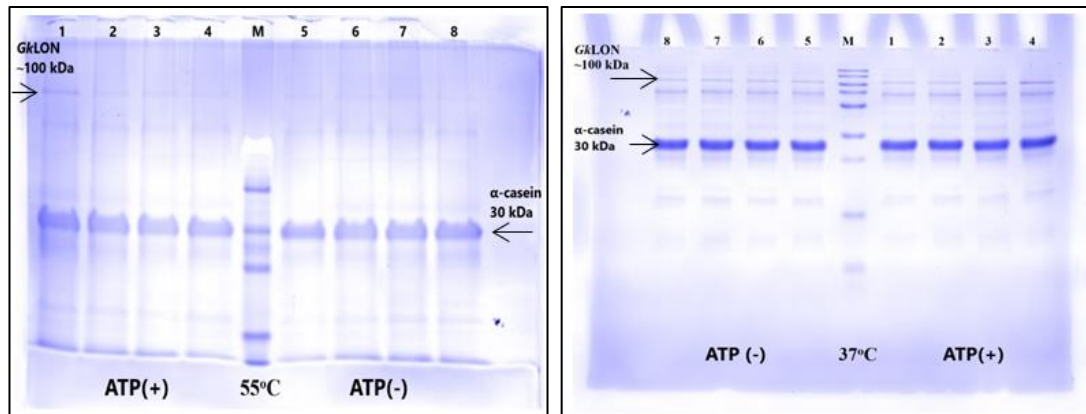


Figure 4. 9: Time and heat dependent α -casein hydrolysis by *GkLON*.

α -casein degradation at 55°C and 37°C. 1-4: in the presence of 4 mM ATP at 0, 10, 30, 60 minutes, respectively, M: Molecular Protein Marker, 5-8: in the absence of ATP at 0, 10, 30, 60 minutes, respectively.

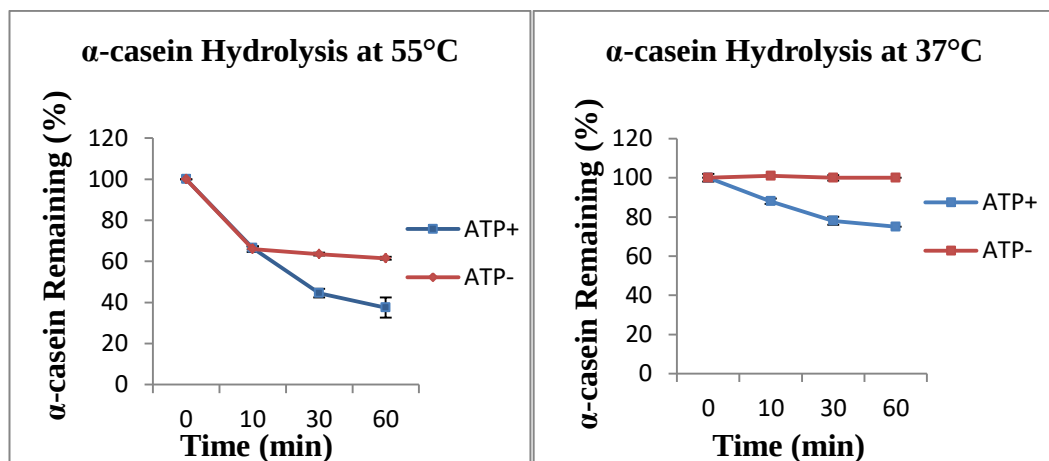


Figure 4. 10: Densitometric analysis of time and heat dependent α -casein degradation by *GkLON*.

4.5.3. Nucleotide Specificity of *GkLON*

Recombinant *GkLON* protease has been shown to utilize ATP as an energy source. To further investigate the nucleotide specificity of the *GkLON*, the protease activity of the enzyme was tested in the presence of various nucleoside triphosphates (ATP, ADP, CTP, UTP, and GTP) were tested. Considering the above study (Section 4.5.1) the nucleotide concentration was chosen to be 4 mM. The experiment was carried out as describes in Material Method section 3.2.7. The effect of different nucleotides was observed by SDS-PAGE analysis shown in Fig. 4.11. According to densitometric analysis, *GkLON* was able to degrade α -casein even in the absence of NTP (32.5%) (Figure 4.12). Interestingly, *GkLON* destroyed α -casein in the presence of GTP and CTP (63.5%) more effectively than in the presence of ATP (56.5%). However, a lower rate of substrate hydrolysis was observed in the use of UTP (50.5%). Very low *GkLON* proteolytic activity was observed in the presence of ADP (40%). When the same amount of ADP as ATP was added to the reaction, it was determined that ADP inhibited the ATP-dependent proteolytic activity (41%).

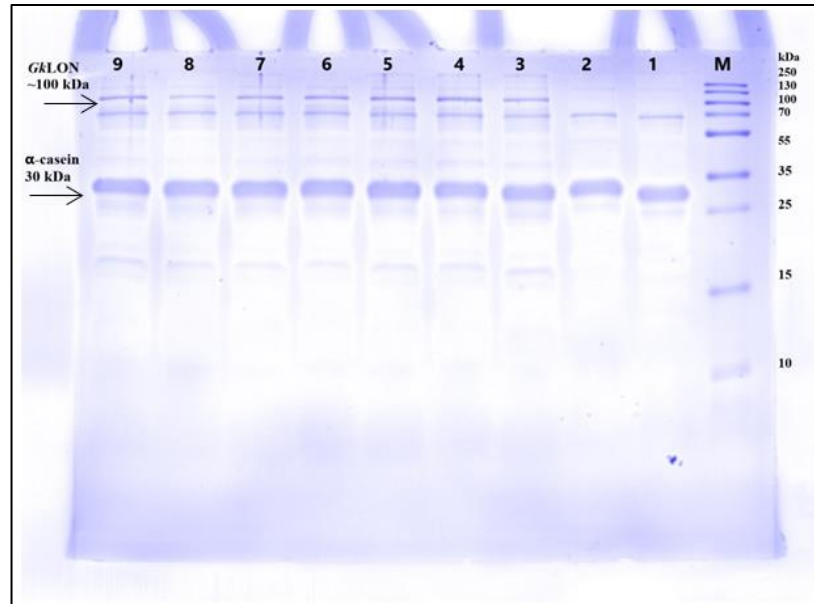


Figure 4. 11: Nucleotide specific α -casein degradation of *GkLON*.

M: Molecular Protein Marker, 1: Only α -casein, 2: 4 mM ATP without enzyme, 3: without nucleotides, 4-5-6-7-8-9: ATP, ADP, ATP+ADP, UTP, GTP and CTP, respectively.

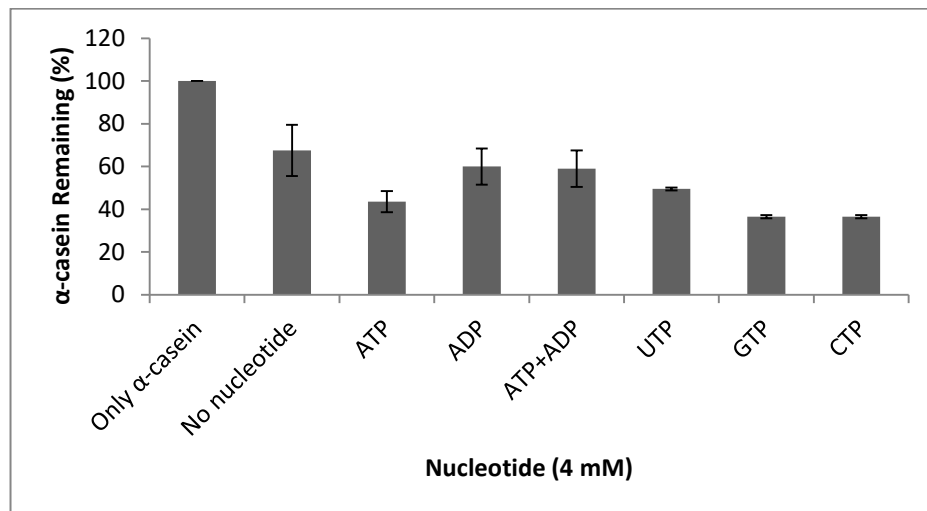


Figure 4. 12: Densitometric analysis of *GkLON* nucleotide specificity.

4.5.4. Effects of Cations on the *GkLON* Proteolytic Activity

The effects of some divalent cations on proteolytic activity of *GkLON* was investigated in the presence 7.5 mM cations and 4 mM ATP (Figure 4.13). Densitometric analysis showed that the most efficient α -casein degradation was observed in the presence of Mg^{2+} (56.4%) (Figure 4.14). Mn^{2+} , Ca^{2+} and Co^{2+} also stimulated proteolytic activity, although slightly to a lesser extent, 46%, 44.6% and 41.3%, respectively. Moreover, Ni^{2+} and Zn^{2+} had no effect on *GkLON* protease activity.

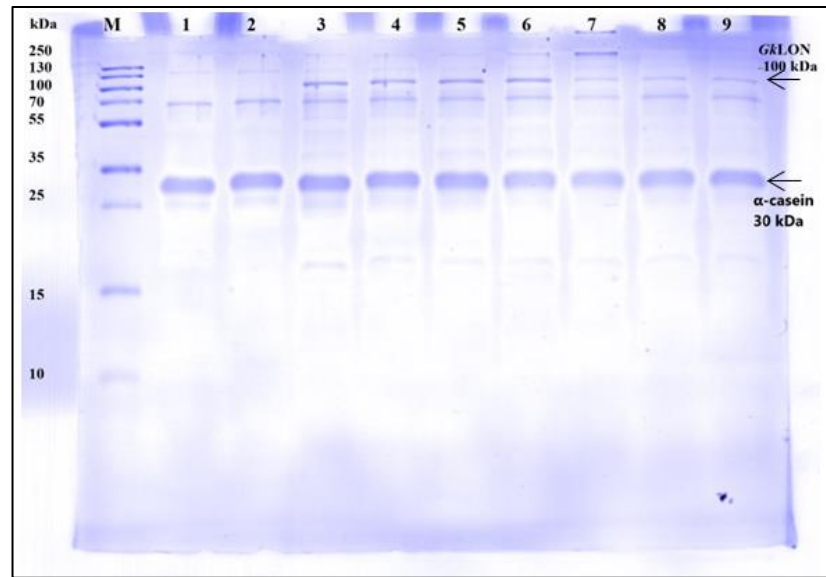


Figure 4. 13: Effect of cations on α -casein hydrolysis of *GkLON*.

M: Molecular Protein Marker, 1: Only α -casein, 2: Mg^{+2} without enzyme, 3: without divalent cations, 4-5-6-7-8-9: Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} and Zn^{2+} , respectively.

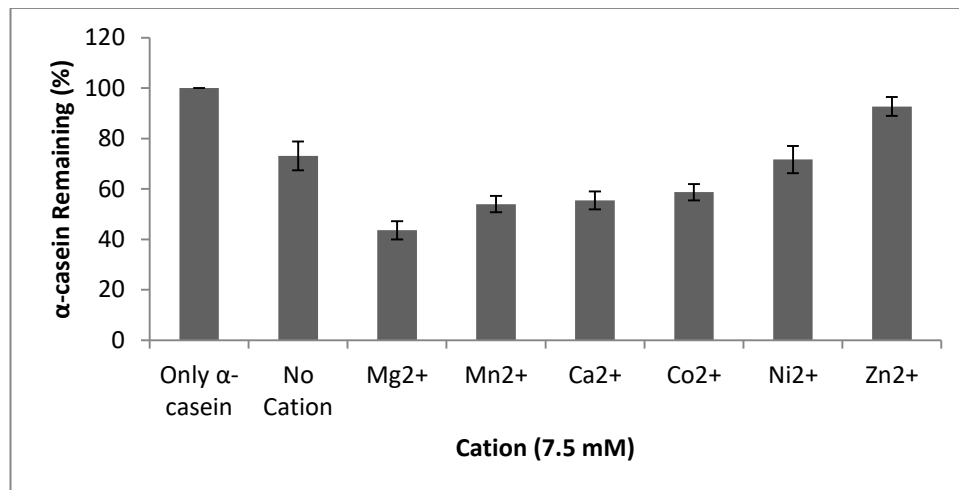


Figure 4. 14: Densitometric analysis of effect of cations on *GkLON* α -casein hydrolysis.

4.6. Homology Modelling and Structural Analysis of *GkLON*

GkLON structural model was produced through i-TASSER with T_m score of 0.91±0.06 and C-score of 1.42. C-score values between [-5, 2] show confidence of the model. A high C-score value for *GkLON* suggests a highly likelihood of the model correctness. The predicted model showed highest homology with Lon protease from *Bacillus subtilis*. *GkLON* exhibits sequence similarities with Lon protease from *Bacillus subtilis* (78%) (Figure 4.15).



Figure 4.15: Homology modeling of *GkLON* and *BsLON*.

GkLON is active as a homo-hexameric and many residues from both subunits are involved in substrate recognition and catalysis. Each subunit of *GkLON* consists of an N-terminal and a C-terminal domain. The active site architecture is shown in Fig. 4.16. The docking analysis results confirm that the substrate-binding pocket of ADP is stabilized by strong hydrogen bonds ($<3 \text{ \AA}$) to Glu305 and Glu82. The binding energy was recorded to be -5.2 kcal/mol . Furthermore, substrate-binding pocket of ADP is stabilized by Pi-Pi stacking and Pi-cation to His83 and His265.

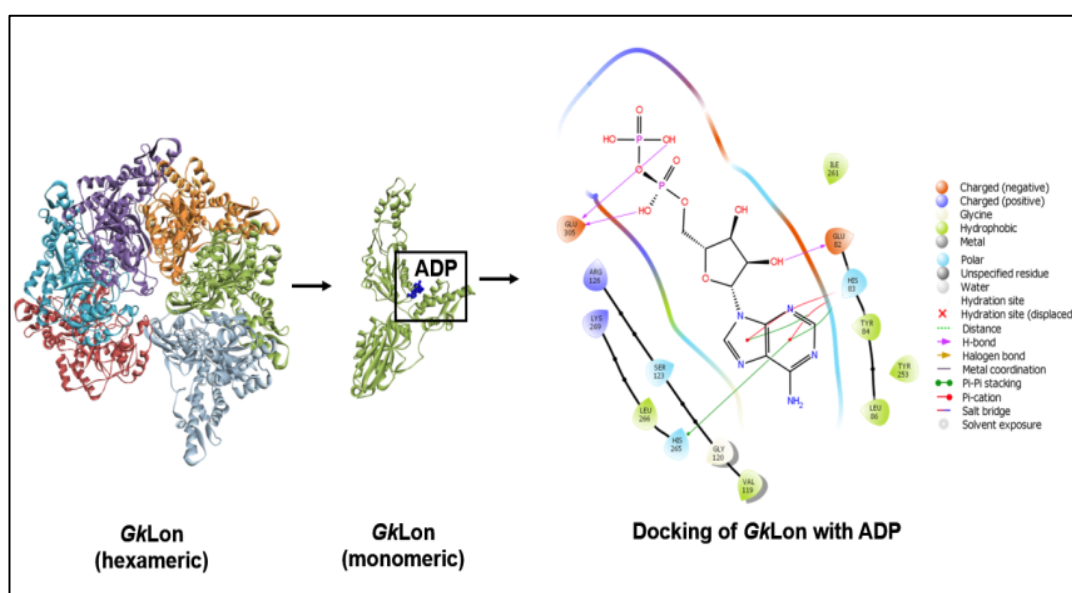


Figure 4. 16: Structural modelling of *GkLON* active site architecture.

5. CONCLUSION

Lon proteases belong to the AAA+ protein as well as other ATP-dependent proteases (FtsH, ClpAP, ClpXP, HslVU and 26S proteasome) [1], [2], [3]. ATP-dependent Lon is the only proteases found in all organisms, from archaea to human [4], [5], [6]. Lon protease is a soluble cytoplasmic protein in eubacteria, and it is essential for cellular homeostasis, for survival from DNA damage and developmental changes in response to numerous stresses, through the control of protein machinery and the regulation of metabolism [95]. Although many functions of Lon protease, which plays a key role in the cell, have been explained, there are still a number of important questions to be answered regarding the mechanism of the enzyme. For this purpose, studies with Lon obtained from a thermophilic bacterium *Geobacillus kaustophilus* are important to elucidate the mechanism of this protease.

In this thesis study, the *Geobacillus kaustophilus* Lon protease gene, which was previously cloned into the pET vector (pET28a-*GkLON*) and transformed into *E. coli* BL21 (DE3) cells in our laboratory, was expressed and purified for kinetic and further biochemical characterization. In order to obtain purified *GkLON*, BL21 (DE3) cells containing pET28a/*GkLON* vector incubated in LB medium (kanamycin 50 µg/mL) at 37°C were stimulated with IPTG to ensure high intracellular expression of the enzyme. After obtaining the cell free extract, N-terminally histidine-labeled proteins were purified in a single step using Ni-NTA affinity chromatography. For activity studies, the buffer solution of the purified protein was changed and concentrated in Tris-HCl buffer solution (pH 8.0).

Lon is homo-oligomeric (consisting of more than one subunit) serine protease that is activated by ATP [96]. Each subunit of the enzyme contains 3 covalently linked domains on a single polypeptide chain: An N-terminal domain involved in substrate recognition, a central ATPase domain that helps to substrate for translocation by supplying energy and a C-terminal protease domain [97]. By using ATP as a chemical energy, Lon protease converts it into the mechanical energy that required for protein conformational change and translocation in proteolysis [11].

It is known that multimeric ATP-dependent proteases such as 26S proteasome and ClpXP show positive cooperativity in substrate binding and peptide bond hydrolysis [115]. These proteases exhibit allosteric behavior towards their specific

substrates and unfolded proteins such as casein. Both ATPase and protease activity of the enzyme are activated with the stimulation of the allosteric mechanism [11]. Many studies have shown that *E. coli* and *Yersinia pestis* Lon proteases exhibit sigmoidal kinetics, indicating positive cooperativity during the hydrolysis of various substrates [115], [116], [117].

In our study, kinetic analysis of *GkLON* was determined using Glt-Ala-Ala-Phe-MNA as a substrate. This fluorogenic peptide substrate is readily soluble in aqueous buffers and is a widely used chymotrypsin like substrate for serine proteases [60]. The degradation of Glt-Ala-Ala-Phe-MNA by *GkLON* in the presence of 1 mM ATP showed sigmoidal kinetics indicative of cooperative allosteric behavior, and fit best to the Hill equation. The K_M and V_{max} parameters of *GkLON* were found to be 0.6 mM and 134 mM/min, respectively. The kinetic parameters of Lon protease can be varying between organisms.

Since the thesis study coincided with Covid-2019, we had problems in the supply of fluorogenic substrate Glt-Ala-Ala-Phe-MNA, so we continued the next studies with α -casein, which is a serine protease substrate. For the enzyme activity studies with α -casein, *GkLON* was incubated with the substrate at 55°C for 1 hour in the presence and absence of ATP, and then the reaction mixture was loaded on SDS-PAGE to observe the change in the intact substrate. The proteolytic activity was determined by measuring the density of the α -casein band on the gel with ImageJ analysis program.

The previous activity studies in our laboratory with Glt-Ala-Ala-Phe-MNA and α -casein, the optimum temperature of the enzyme was determined as 55°C [88]. α -casein is a loosely folded 30 kDa protein and it has no secondary structure at temperatures between 37°C and 70°C [60].

In the study performed with ATP-dependent FtsH in our laboratory, an increase in α -casein hydrolysis was observed due to an increased ATP concentration up to 6 mM [98]. In the study of Larimore and colleagues in 1881, maximum casein hydrolysis by *E. coli* Lon was observed in the presence of 5 mM ATP in a study with ATP concentrations ranging from 0.1 to 5 mM [10]. Similarly, in our study in the presence of ATP at concentrations ranging from 0-8 mM, the highest α -casein degradation by *GkLON* was detected in the presence of 8 mM ATP. However, 4 mM ATP was used in all experiments.

To determine the effect of incubation time on *GkLON* activity, the enzyme was incubated with α -casein in the presence and absence of 4 mM ATP at 37 and 55°C for different time intervals up to one hour. According to results, after the first 10 min incubation at 55°C, the substrate with or without ATP were degraded by *GkLON* at a rate of 34%. However, during the 1 h incubation after 10 minutes, an increase in the presence of ATP was observed in the time-dependent substrate degradation, and no significant change was observed in the absence of ATP. These results showed that at the end of the first 10 minutes, the substrate was degraded by the effect of high temperature, but protein degradation was effective in the presence of ATP in the following incubation times. The data we obtained are similar to other ATP-dependent proteases in the literature and ATP-dependent FtsH [98] and HslVU [99] studied in our laboratory. In our laboratory, studies with FtsH, an ATP-dependent protease, have also shown that spontaneous degradation is observed in α -casein at high temperature, but degradation is effectively increased in the presence of ATP [98]. When the same experiment with *GkLON* was conducted at 37°C, no change was observed in protein substrate degradation in the absence of ATP, in addition the maximum protein degradation was 25% in the presence of ATP. From these results, it was concluded that *GkLON* is a thermophilic enzyme and shows maximum activity at high temperatures and requires ATP for protein degradation.

The polypeptide hydrolysis reaction of Lon protease with the aid of ATP ends with the release of small peptide fragments of 5-10 amino acids and ADP [12]. ADP is known to bind to Lon with a higher affinity and thus it is a potent inhibitor of *E. coli* [11], [83], [100], [101], [102]. Also, as observed in the *GkFtsH* studied in our laboratory, in the presence of ADP, the proteolytic activity of *GkLON* decreased by approximately 16.5% compared to ATP. When the same amounts of ATP and ADP were added to the enzyme reaction, ADP inhibited the *GkLON* activity and decreased the enzyme activity by 15.5%.

It has been stated that Lon proteases from different organisms show different nucleotide dependence. For example, *E. coli* Lon protease showed the highest affinity for ATP against peptide substrate when compared with other nucleotides, followed by CTP with the least enzyme activity in the presence of UTP [55]. Similarly, in another study with *Thermococcus kodakaraensis* Lon protease, the highest enzyme activity in peptide substrate degradation was observed in the

presence of CTP, ATP, UTP and GTP, respectively [19]. In the α -casein degradation experiment that we conducted with FtsH in our laboratory, the highest enzyme activity was shown to be ATP and CTP (similar ratio), GTP, UTP, respectively. In the activity experiment we conducted with GkLON in this thesis study, α -casein was most effectively destroyed in the presence of CTP and GTP (similar ratio), ATP and UTP, respectively [98]. These results are similar to the data obtained with Lon-like *Methanobrevibacter Smithii* in 2016 [9].

Lon require cofactor for enzymatic activity. Various studies have been conducted on the cofactor dependence of this enzyme. In 1981, Goldberg and colleagues showed that the enzyme required ATP and Mg^{2+} together for protein substrate (α - and β -casein) hydrolysis in their study with *E. coli* Lon [5]. In 1985, Goldberg, in his study with Waxman, stated that proteins definitely require ATP and Mg^{2+} or Ca^{2+} for degradation by Lon, but peptide hydrolysis by Lon only depends on the presence of divalent cations. According to this study, it was shown that Ca^{2+} and Mn^{2+} had a comparable effect with Mg^{2+} in promoting peptide hydrolysis, nevertheless protein degradation occurred most in the presence of Mg^{2+} [60]. Similarly, we concluded that GkLON enzymatic activity is strictly dependent on the presence of divalent cation, Mg^{2+} . Only in the presence of ATP, GkLON destroyed α -casein by 27%, however, a 2-fold increase was observed in protein degradation when Mg^{2+} was added. Although divalent cations such as Mn^{2+} , Ca^{2+} , Co^{2+} increased the protein degradation of GkLON, they did not show significant an effect as Mg^{2+} .

Lon protease, one of the most conserved ATP-dependent proteases, is known to vary in sequence similarity between different organisms [11], [27]. According to the findings in the literature, 15%, 55%, 88%, 51%, 41%, 47%, and 67% sequence similarity has been seen between *Brevibacillus thermoruber* Lon protease and Lon proteases from *Thermococcus kodakaraensis* [19], *E. coli* [37], *Brevibacillus brevis* [38], *Thermus thermophilus* [39], *Mycobacterium smegmatis* [40], *Myxococcus xanthus* [45], and *Bacillus subtilis* [103], respectively [41]. Similarly, in our study, 78% sequence similarity was found between *Geobacillus kaustophilus* and *Bacillus subtilis* Lon proteases. This finding reveals that high homology has been seen between the two organisms.

Active site of bacterial Lon protease, which has a barrel-like hexameric ring structure, shows structural differences between ADP-bound and nucleotide free

forms [110]. Various studies to reveal the molecular infrastructure of ATP-dependent proteases have demonstrated that nucleotide free subunits of the proteases are more flexible than nucleotide-bound subunits [104], [105], [106], [107], [108]. Conformational changes are dictated especially by ADP binding and release since ADP binds to Lon protease with higher affinity than other nucleotides [109]. Therefore, examining the ADP-bound or ADP-free conformational structure of the *GkLON* protease is important to identify the structural features that affect Lon activity in both conformations. For this purpose, the docking analysis performed in this thesis study was similar to the findings in the literature was shown that the active site of the *GkLON* protease is in the form of a hexameric ring. Also it was determined that the ADP-bound form of *GkLON* was stabilized by strong hydrogen bonds to Glu305 and Glu82 and Pi-Pi stacking and Pi cation to His83 and His265. The binding energy was found to be -5.2 kcal/mol.

In conclusion of this thesis, it has been contributed to the literature that *GkLON* is a nucleotide and cofactor dependent, thermophilic enzyme that shows maximum activity at high temperatures. In addition, *GkLON* has been demonstrated to be a hexameric Lon protease with a ring-shaped proteolytic active site. A high degree of homology similarity was detected between *Geobacillus kaustophilus* and *Bacillus subtilis* Lon proteases.

REFERENCES

- [1] Neuwald A. F., Aravind L., Spouge J. L., Koonin E. V., (1999), ‘‘AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes’’, *Genome Res.*, 9, 27–43.
- [2] Ogura T., Wilkinson A. J., (2001), ‘‘AAA+ superfamily ATPases: common structure-diverse function’’, *Genes Cells.*, 6, 575–597.
- [3] Lupas A. N., Martin J., (2002), ‘‘AAA proteins’’, *Curr. Opin. Struct. Biol.*, 12, 746-53.
- [4] Swamy K. H., Goldberg A. L., (1981), ‘‘E. coli contains eight soluble proteolytic activities, one being ATP dependent’’, *Nature*, 292, 652-4.
- [5] Goldberg A. L., Moerschell R. P., Chung C. H., Maurizi M. R., (1981), ‘‘ATP-dependent protease La (Lon) from *Escherichia coli*’’, *Methods Enzymol.*, 244, 350-75.
- [6] Chiarette M. F., Henderson G. W., Markovitz A., (1981), ‘‘ATP hydrolysis-dependent protease activity of the Lon (capR) protein of *Escherichia coli* K-12’’, *Proc. Natl. Acad. Sci.*, 78, 4728-32.
- [7] Dyck L. V., Langer T., (1999), ‘‘ATP-dependent proteases controlling mitochondrial function in the yeast *Saccharomyces cerevisiae*’’, *Cell Mol. Life Sci.*, 56, 825–842.
- [8] Rep M., Grivell L. A., (1996), ‘‘The role of protein degradation in mitochondrial function and biogenesis’’, *Curr. Genet.*, 30, 367–380.
- [9] Pei J., Yan J., Jiang Y., (2016), ‘‘Characterization of the ATP-Dependent Lon-like protease in *Methanobrevibacter smithii*’’, *Archaea*, 2016: 1-7.
- [10] Larimore F. S., Waxman L., Goldberg A. L., (1982), ‘‘Studies of the ATP-dependent proteolytic enzyme, protease La, from *Escherichia coli*’’, 257, 4187–4195.
- [11] Lee I., Suzuki C. K., (2008), ‘‘Functional mechanics of the ATP-dependent Lon protease lessons from endogenous protein and synthetic peptide substrates’’, *Biochim. Biophys. Acta.*, 1784, 727–735.
- [12] Thomas-Wohlever J., Lee I., (2002), ‘‘Kinetic characterization of the peptidase activity of *Escherichia coli* Lon reveals the mechanistic similarities in ATP-dependent hydrolysis of peptide and protein substrate’’, *Biochem.*, 41, 9418-9425.

- [13] Ryan B. J., Hennehan G. T., (2013), "Overview of approaches to preventing and avoiding proteolysis during expression and purification of proteins", *Curr. Protoc. Protein Sci.*, 5.25.1-5.25.7.
- [14] Rao M. B., Tanksale A. M., Ghatge M. S., Deshpande V. V., (1998), "Molecular and biotechnological aspects of microbial proteases", *MMBR.*, 62, 597–635.
- [15] Bond J. S., (2019), "Proteases: history, discovery, and roles in health and disease", *J. Biol. Chem.*, 294, 1643–1651.
- [16] Menon A. S., Waxman L., Goldberg A. L., (1987), "The energy utilized in protein breakdown by the ATP-dependent protease (La) from *Escherichia coli*", *J. Biol. Chem.*, 262, 722-726.
- [17] Kebe N. M., Samanta K., Singh P., Lai-Kee-Him J., Apicella V., Payrot N., Lauraire N., Legrand B., Lisowski V., Mbang-Benet D. E., Pages M., Bastien P., Kajava A. V., Bron P., Hernandez J.F., Coux O., (2019), "The HslV protease from *Leishmania major* and its activation by C-terminal HslU peptides", *Int. J. Mol. Sci.*, 20, 1021.
- [18] Olivares A. O., Baker T. A., Sauer R. T., (2018), "Mechanical protein unfolding and degradation", *Annu. Rev. Physiol.*, 80, 413–29.
- [19] Fukui T., Eguchi T., Atomi H., Imanaka T., (2002), "A membrane-bound archaeal Lon protease displays ATP-independent proteolytic activity towards unfolded proteins and ATP-dependent activity for folded protein", *J. Bacteriol.*, 184, 3689–3698.
- [20] Gottesman S., (1999), "Regulation by proteolysis: developmental switches", *Curr. Opin. Microbiol.*, 2, 142–147.
- [21] Gottesman S., Maurizi M. R., (1992), "Regulation by proteolysis: energy dependent proteases and their targets", *Microbiol. Rev.*, 56, 592–621.
- [22] Hanson P. I., Whiteheart S. W., (2005), "AAA+ proteins: have engine, will work", *Nat. Rev. Mol. Cell. Biol.*, 6, 519–529.
- [23] Goldberg A. L., (1992), "The mechanism and functions of ATP-dependent proteases in bacterial and animal cells", *Eur. J. Biochem.*, 203, 9–23.
- [24] Van Melder L., Thi M. H., Lecchi P., Gottesman S., Couturier M., Maurizi M. R., (1996), "ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions", *J. Biol. Chem.*, 271, 27730-8.
- [25] Starkova N. N., Koroleva E. P., Rotanova T. V., (2000), "Intracellular proteolysis: signals of selective protein degradation", *Bioorg. Khim.*, 26, 83-96.

- [26] Bittner L. M., Arends J., Narberhaus F., (2016), "Mini review ATP-dependent proteases in bacteria", *Biopolymers.*, 105, 505-517.
- [27] Venkatesh S., Lee J., Singh K., Lee I., Suzuki C. K., (2012), "Multitasking in the mitochondrion by the ATP-dependent Lon protease", *Biochim. Biophys. Acta.*, 1823, 56-66.
- [28] Sauer R. T., Baker T. A., (2011), "AAA+ proteases: ATP-fueled machines of protein destruction", *Annu. Rev. Biochem.*, 80, 587-612.
- [29] Koodathingal P., Jaffe N. E., Kraut D. A., Prakash S., Fishbain S., Herman C., Matouschek A., (2009), "ATP-dependent proteases differ substantially in their ability to unfold globular proteins", *J. Biol. Chem.*, 284, 18674-18684.
- [30] Botos I., Melnikov E. E., Cherry S., Tropea J. E., Khalatova A. G., Rasulova F., Dauter Z., Maurizi M. R., Rotanova T. V., Wlodawer A., Gustchina A., (2004), "The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site", *J. Biol. Chem.*, 279, 8140-8148.
- [31] Bochtler M., Hartmann C., Song H. K., Bourenkov G. P., Bartunik H. D., Huber R., (2000), "The structures of HsIU and the ATP-dependent protease HsIU-HsIV", *Nature*, 403, 800-5.
- [32] Bieniossek C., Schalch T., Bumann M., Meister M., Meier R., Baumann U., (2006), "The molecular architecture of the metalloprotease FtsH", *Proc. Natl. Acad. Sci.*, 103, 3066-71.
- [33] Lee A. Y. L., Hsu C. H., Wu S. H., (2004), "Functional domains of Brevibacillus thermoruber Lon protease for oligomerization and DNA binding" *J. Biol. Chem.*, 279, 34903-34912.
- [34] Smith C. K., Baker T. A., Sauer R. T., (1999), "Lon and Clp family proteases and chaperones share homologous substrate-recognition domains", *Proc. Natl. Acad. Sci.*, 96, 6678-6682.
- [35] Goard C. A., Schimmer A. D., (2014), "Mitochondrial matrix proteases as novel therapeutic targets in malignancy", *Oncogene*, 33, 2690-2699.
- [36] Chung C. H., Goldberg A. L., (1981), "The product of the Lon (capR) gene in Escherichia coli is the ATP-dependent protease, protease La", *Proc. Natl. Acad. Sci.*, 78, 4931-4935.
- [37] Chin D. T., Goff S. A., Webster T., Smith T., Goldberg A. L., (1988), "Sequence of the Lon gene in Escherichia coli. A heat-shock gene which encodes the ATP-dependent protease La", *J. Biol. Chem.*, 263, 11718-11728.
- [38] Ito K., Udaka S., Yamagata H., (1992), "Cloning, characterization, and inactivation of the Bacillus brevis Lon gene", *J. Bacteriol.*, 174, 2281-2287.

- [39] Watanabe S., Muramatsu T., Ao H., Hirayama Y., Takahashi K., Tanokura M., Kuchino Y., (1999), "Molecular cloning of the Lon protease gene from *Thermus thermophilus* HB8 and characterization of its gene product" *Eur. J. Biochem.*, 266, 811–819.
- [40] Roudiak S. G., Seth A., Knipfer N., Shrader T. E., (1998), "The Lon protease from *Mycobacterium smegmatis*: molecular cloning, sequence analysis, functional expression, and enzymatic characterization", *Biochem.*, 37, 377–386.
- [41] Lee A. Y., Tsay S. S., Chen M. Y., Wu S. H., (2004), "Identification of a gene encoding Lon protease from *Brevibacillus thermoruber* WR-249 and biochemical characterization of its thermostable recombinant enzyme", *Eur. J. Biochem.*, 271, 834–844.
- [42] Besche H., Zwickl P., (2004), "The *Thermoplasma acidophilum* Lon protease has a Ser-Lys dyad active site", *Eur. J. Biochem.*, 271, 4361–4365.
- [43] Duman R. E., Löwe J., (2010), "Crystal structures of *Bacillus subtilis* Lon protease", *J. Mol. Biol.*, 401, 653–670.
- [44] Wang N., Gottesman S., Willingham M. C., Gottesman M., Maurizi M. R., (1993), "A human mitochondrial ATP-dependent protease that is highly homologous to bacterial Lon protease", *Proc. Natl. Acad. Sci.*, 90, 11247–11251.
- [45] Tojo N., Inouye S., Komano T., (1993), "The LonD gene is homologous to the Lon gene encoding an ATP-dependent protease and is essential for the development of *Myxococcus xanthus*", *J. Bacteriol.*, 175, 4545–4549.
- [46] Suzuki C. K., Suda K., Wang N., Schatz G., (1994), "Requirement for the yeast gene Lon in intramitochondrial proteolysis and maintenance of respiration", *Science*, 264, 273–276.
- [47] Bota D. A., Davies K. J. A., (2016), "Mitochondrial Lon protease in human disease and aging: including an etiologic classification of Lon-related diseases and disorders", *Free Radic. Biol. Med.*, 100, 188–198.
- [48] Pintia M., Gibellini L., Nasi M., De Biasi S., Bortolotti C. A., Iannone A., Cossarizza A., (2016), "Emerging role of Lon protease as a master regulator of mitochondrial functions", *Biochim. Biophys.*, 1857, 1300–1306.
- [49] He L., Luo D., Yang F., Li C., Zhang X., Deng H., Zhang J.R., (2018), "Multiple domains of bacterial and human Lon proteases define substrate selectivity", *Emerg. Microbes. Infect.*, 7, 149.

- [50] Rotanova T. V., Botos I., Melnikov E. E., Rasulova F., Gustchina A., Maurizi M. R., Wlodawer A., (2006), "Slicing a protease: structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains", *Protein Sci.*, 15, 1815–1828.
- [51] Cha S. S., An Y. J., Lee C. R., Lee H. S., Kim Y. G., Kim S. J., Kwon K. K., De Donatis G. M., Lee J. H., Maurizi M. R., Kang S. G., (2010), "Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber", *EMBO Jour.*, 29, 3520–3530.
- [52] Park S. C., Jia B., Yang J. K., Van D. L., Shao Y. G., Han S. W., Jeon Y. J., Chung C. H., Cheong G. W., (2006), "Oligomeric structure of the ATP-dependent protease La (Lon) of *Escherichia coli*", *Mol. Cell.*, 21, 129–134.
- [53] Rotanova¹ 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", *Eur. J. Biochem.*, 271, 4865–4871.
- [54] Fischer H., Glockshuber R., (1994), "A point mutation within the ATP-binding site inactivates both catalytic functions of the ATP-dependent protease La (Lon) from *Escherichia coli*", *FEBS Lett.*, 356, 101–103.
- [55] Patterson J., Vineyard D., Wohlever J. T., Behshad R., Burke M., Lee I., (2004), "Correlation of an adenine-specific conformational change with the ATP-dependent peptidase activity of *Escherichia coli* Lon", *Biochem.*, 43, 7432–7442.
- [56] Vasilyeva O. V., Kolygo K. B., Leonova Y. F., Potapenko N. A., Ovchinnikova T. V., (2002), "Domain structure and ATP-induced conformational changes in *Escherichia coli* protease Lon revealed by limited proteolysis and autolysis", *FEBS Letters.*, 526, 66–70.
- [57] Ahern K., Rajagopal I., Tan T., (2021), "Mechanisms of catalysis", *Biochem.*, 380–394.
- [58] Web 1, (2020), http://www2.csudh.edu/nsturm/CHE450/11_Enz.%20Mech.-Ser%20Protea.htm, (Access Date: 05/11/2021).
- [59] Hedstrom L., (2002), "Serine protease mechanism and specificity", *Chem. Rev.*, 102, 4501–24.
- [60] Waxman L., Goldberg A. L., (1985), "Protease La, the Lon gene product, cleaves specific fluorogenic peptides in an ATP-dependent reaction", *J. Biol. Chem.*, 260, 12022–12028.
- [61] Besche H., Tamura N., Tamura T., Zwickl P., (2004), "Mutational analysis of conserved AAA+ residues in the archaeal Lon protease from *Thermoplasma acidophilum*", *FEBS Lett.*, 574, 161–6.

- [62] Rotanova T. V., Melnikov E. E., Khalatova A. G., Makhovskaya O. V., Botos I., Wlodawer A., Gustchina A., (2004), ‘‘Atomic-resolution crystal structure of the proteolytic domain of *Archaeoglobus fulgidus* Lon reveals the conformational variability in the active sites of Lon proteases’’, *Eur. J. Biochem.*, 271, 4865.
- [63] Starkova N. N., Koroleva E. P., Rumsh L. D., Ginodman L. M., Rotanova T. V., (1998), ‘‘Mutations in the proteolytic domain of *Escherichia coli* protease Lon impair the ATPase activity of the enzyme’’, *FEBS Letters.*, 422, 218–220.
- [64] Wright R., Stephens C., Zweiger G., Shapiro L., Alley M. R., (1996), ‘‘*Caulobacter* Lon protease has a critical role in cell-cycle control of DNA methylation’’, *Genes Dev.*, 10, 1532–1542.
- [65] Howard-Flanders P., Simson E., Theriot L., (1964), ‘‘A locus that controls filament formation and sensitivity to radiation in *Escherichia coli* K-12’’, *Genetics.*, 49, 237–246.
- [66] Lee I., Berdis A. J., Suzuki C. K., (2006), ‘‘Recent developments in the mechanistic enzymology of the ATP-dependent Lon protease from *Escherichia coli*: highlights from kinetic studies’’, *Mol. Biosyst.*, 2, 477–483.
- [67] Sonezaki S., Okita K., Oba T., Ishii Y., Kondo A., Kato Y., (1995), ‘‘Protein substrates and heat shock reduce the DNA-binding ability of *Escherichia coli* Lon protease’’, *Appl. Microbiol. Biotechnol.*, 44, 484–488.
- [68] Jonas K., Liu J., Chien P., Laub M. T., (2013), ‘‘Proteotoxic stress induces a cell-cycle arrest by stimulating Lon to degrade the replication initiator DnaA’’*Cell.*, 154, 623–636.
- [69] Leslie D. J., Heinen C., Schramm F. D., Thüring M., Aakre C. D., Murray S. M., (2015), ‘‘Nutritional control of DNA replication initiation through the proteolysis and regulated translation of DnaA’’, *PLoS Genet.*, 11, e10053432.
- [70] Gross M. H., Konieczny I., (2020), ‘‘Polyphosphate induces the proteolysis of ADP-bound fraction of initiator to inhibit DNA replication initiation upon stress in *Escherichia coli*’’, *Nuc. Acids. Res.*, 48, 5457–5466.
- [71] Takaya A., Suzuki M., Matsui H., (2003), ‘‘Lon, a stress-induced ATP-dependent protease, is critically important for systemic *Salmonella enterica* serovar typhimurium infection of mice’’, *Infect. Immun.*, 71, 690–696.
- [72] Takaya A., Kubota Y., Isogai E., Yamamoto T., (2005), ‘‘Degradation of the HilC and HilD regulator proteins by ATP-dependent Lon protease leads to downregulation of *Salmonella* pathogenicity island 1 gene expression’’, *Mol. Microbiol.*, 55, 839–852.
- [73] Mukherjee S., Bree A. C., Liu J., Patrick J. E., Chie P., Kearns D. B., (2015), ‘‘Adaptor mediated Lon proteolysis restricts *Bacillus subtilis* hyperflagellation’’, *Proc. Natl. Acad. Sci.*, 112, 250–5.

- [74] Mukherjee S., Bree A. C., Liu J., Patrick J. E., Chie P., Kearns D. B., (2014), "The structure and regulation of flagella in *Bacillus subtilis*", *Annu. Rev. Genet.*, 48, 319–40.
- [75] De Matteis S., Carloni S., Napolitano R., Zanoni M., Gibellini L., Lucchesi A., Cossarizza A., Martinelli G., (2014), "Targeting Lon protease activity promotes growth inhibition and apoptosis in acute myeloid leukemia cell lines", *Blood*, 124, 5130.
- [76] Kao T. Y., Chiu Y. C., Fang W. C., Cheng C. W., Kuo C. Y., Juan H. F., Wu S. H., Lee A. L., (2015), "Mitochondrial Lon regulates apoptosis through the association with Hsp60–mtHsp70 complex", *CDD.*, 6, e1642.
- [77] Krebs J. E., Goldstein E. S., Lewin B., Kilpatrick S.T., (2010), "Antitermination may be a regulated event", *Lew. essent. genes.*, 2nd ed, 287–291.
- [78] Weisberg R. A., Gottesman M. E., (1999), "Processive antitermination", *J. Bacteriol.*, 181, 359–367.
- [79] Gottesman M., Gottesman S., (1981), "Protein degradation in *E. coli*: The Lon mutation and bacteriophage lambda N and cII protein stability", *Cell*, 24, 225–233.
- [80] Bhaskarla C., Das M., Verma T., Kumar A., Mahadevan S., Dipankar N., (2016), "Roles of Lon protease and its substrate MarA during sodium salicylate-mediated growth reduction and antibiotic resistance in *Escherichia coli*", *Microbiol.*, 162, 764–776.
- [81] Frank E. G., Ennis D. G., Gonzalez M., Levine A. S., Woodgate R., (1996), "Regulation of SOS mutagenesis by proteolysis", *Proc. Natl. Acad. Sci.*, 93, 10291–10296.
- [82] Breidenstein E. B. M., Bains M., Hancock R. E. W., (2012), "Involvement of the Lon protease in the SOS response triggered by ciprofloxacin in *Pseudomonas aeruginosa* PAO1", *Antimic. Agent. Chem.*, 56, 2879 –2887.
- [83] Goldberg A. L. Waxman L., (1985), "The role of ATP hydrolysis in the breakdown of proteins and peptides by protease La from *Escherichia coli*", *J. Biol. Chem.*, 260, 12029–12034.
- [84] Rudyak S. G., Shrader T. E., (2000), "Polypeptide stimulators of the Ms-Lon protease" *Prot. Sci.*, 9, 1810–1817.

- [85] Makhovskayaa O. V., Kozlovb S., Botosc I., Stepnova A. A., Andrianovaa A. G., Gushchinac A. E., Wlodawerc A., Mel'nikova E. E., Rotanovaa T. V., (2007), "Forms of LonB protease from *Archaeoglobus fulgidus* devoid of the transmembrane domain: the contribution of the quaternary structure to the regulation of enzyme proteolytic activity", *Russ. J. Bioorg. Chem.*, 33, 610–613.
- [86] Ambro L., Pevala V., Ondrovicova G., Bellova J., Kunova N., Kutejova E., Bauer J., (2014), "Mutations to a glycine loop in the catalytic site of human Lon changes its protease, peptidase and ATPase activities", *FEBS J.* 281, 1784–1797.
- [87] Camcı H., (2009), "Cloning, sequencing and expression of *Geobacillus kaustophilus* ATP dependent Lon protease gene", M.Sc. Thesis, Gebze Technical University.
- [88] Asal H. Z., (2011), "Purification and characterization of recombinant *Geobacillus kaustophilus* protein Lon from *E. coli*", M.Sc. Thesis, Gebze Technical University.
- [89] Laemmli U. K., (1970), "Cleavage of structural proteins during the assembly of the head of bacteriophage T4", *Nature*, 227, 680-685.
- [90] Yang J., Zhang Y., (2015), "I-TASSER server: New development for protein structure and function prediction", *Nuc. Acids. Res.*, 43 W174–W181.
- [91] Trott O., Olson A. J., (2009), "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading", *J. Comput. Chem.*, 31, NA-NA.
- [92] Madhavi Sastry G., Adzhigirey M., Day T., Annabhimoju R., Sherman W., (2013), "Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments", *J. Comput. Aided. Mol. Des.*, 27, 221–234.
- [93] Sievers F., Wilm A., Dineen D., Gibson T.J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D., Higgins D. G., (2011), "Fast scalable generation of high-quality protein multiple sequence alignments using Clustal Omega", *Mol. Syst. Biol.*, 7, 539.
- [94] Gouet P., Courcelle E., Stuart D. I., Métoz F. (1999), "ESPrpt: Analysis of multiple sequence alignments in PostScript", *Bioinformatics*, 15, 305–308.
- [95] Jonas K., Liu J., Chien P., Laub M. T., (2013), "Proteotoxic stress induces a cell-cycle arrest by stimulating Lon to degrade the replication initiator DnaA", *Cell*, 154, 623–636.
- [96] Patterson-Ward J., Huang J., Lee I., (2007), "Detection and characterization of two ATP-dependent conformational changes in proteolytically inactive *Escherichia coli* Lon mutants by stopped flow kinetic technique", *Biochem.*, 46, 13593–13605.

- [97] Van Melder L., Aertsen A., (2009), ‘‘Regulation and quality control by Lon-dependent proteolysis’’ *Res. Microbiol.*, 160, 645–651.
- [98] Tülek A., Özdemir F. İ., Ramadhan S. S., (2020), ‘‘Cloning, expression and characterization of membrane bound FtsH protease of *Geobacillus kaustophilus*’’, *App. Biochem. Microbiol.*, 56, 678–684.
- [99] Altın G., (2009), ‘‘Purification and characterization of recombinant *Geobacillus kaustophilus* protein HslV and HslU from *E. coli*’’, M.Sc. Thesis, Gebze Technical University.
- [100] Fishovitz J., Sha Z., Chilakala S., Cheng I., Xu Y., Lee I., (2017), ‘‘Utilization of mechanistic enzymology to evaluate the significance of ADP binding to human Lon protease’’, *Front. Mol. Biosci.*, 4, 47.
- [101] Thomas-Wohlever J., Lee I., (2002), ‘‘Kinetic characterization of the peptidase activity of *Escherichia coli* Lon reveals the mechanistic similarities in ATPdependent hydrolysis of peptide and protein substrates’’, *Biochem.*, 41, 9418-25.
- [102] Lee I., Berdis A. J., (2001), ‘‘Adenosine triphosphate-dependent degradation of a fluorescent lambda N substrate mimic by Lon protease’’, *Anal. Biochem.*, 291, 74-83.
- [103] Riethdorf S., Volker U., Gerth U., Winkler A., Engelmann S., Hecker M., (1994), ‘‘Cloning, nucleotide sequence, and expression of the *Bacillus subtilis* Lon gene’’, *J. Bacteriol.*, 176, 6518–6527.
- [104] De la Peña A. H., Goodall E. A., Gates S. N., Lander G. C., Martin A., (2018), ‘‘Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation’’, *Sci.*, 362, eaav0725.
- [105] Puchades C., Rampello A. J., Shin M., Giuliano C. J., Wiseman R. L., Glynn S. E., Lander G. C., (2017), ‘‘Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing’’, *Sci.*, 358, eaao0464.
- [106] Rizo A. N., Lin J., Gates S. N., Tse E., Bart S. M., Castellano L. M., DiMaio F., Shorter J., Southworth D.R., (2019), ‘‘Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase’’, *Nat. Commun.*, 10, 2393.
- [107] White K. I., Zhao M., Choi U. B., Pfuetzner R. A., Brunger A. T., (2018), ‘‘Structural principles of SNARE complex recognition by the AAA+ protein NSF’’, *eLife.*, 7, e38888.
- [108] Yu H., Lupoli T. J., Kovach A., Meng X., Zhao G., Nathan C. F., Li H., (2018), ‘‘ATP hydrolysis-coupled peptide translocation mechanism of *Mycobacterium tuberculosis* ClpB’’, *Proc. Natl. Acad. Sci.*, 115, E9560–E9569.

- [109] Shin M., Puchades C., Asmita A., Puri N., Adjei E., Wiseman R. L., Karzai A. W., Lander G. C., (2020), ‘‘Structural basis for distinct operational modes and protease activation in AAA+ protease Lon’’, *Sci. Adv.*, 6, eaba8404.
- [110] Su S. C., Lin C. C., Tai H. C., Chang Y. C., Su S. C., Ho M. R., Babu C. S., Liao J. H., Wu S. H., Lim C., (2016), ‘‘Chang C.I. Structural basis for the magnesium-dependent activation and hexamerization of the Lon AAA+ protease’’, *Structure*, 24, 1–11.
- [111] Schneider C. A., Rasband W. S., Eliceiri K. W., (2012), ‘‘NIH image to ImageJ: 25 years of image analysis’’, *Nat. Methods.*, 9, 671–675.
- [112] Eliceiri K., Rueden C., (2005), ‘‘Tools for visualizing multidimensional images from living specimens’’, *Photochem. Photobiol.*, 81, 1116–22.
- [113] Barboriak D., Padua A., York G., Macfall J., (2005), ‘‘Creation of DICOM aware applications using ImageJ’’, *J. Digit. Imaging.*, 18, 91–9.
- [114] Gering E., Atkinson C., (2004), ‘‘A rapid method for counting nucleated erythrocytes on stained blood smears by digital image analysis’’, *J. Parasitol.*, 90, 879–81.
- [115] Granot Z., Kobiler O., Melamed-Book N., Eimerl S., Bahat A., Lu B., Braun S., Maurizi M. R., Suzuki C. K., Oppenheim A. B., Orly J., (2007), ‘‘Turnover of mitochondrial steroidogenic acute regulatory (StAR) protein by Lon protease: the unexpected effect of proteasome inhibitors’’, *Mol. Endoc.*, 21, 2164–2177.
- [116] Puri N., Karzai A. W., (2017), ‘‘HspQ functions as a unique specificity enhancing factor for the AAA+ Lon protease’’, *Mol. Cell.*, 66, 672–683.e4.
- [117] Gur E., Sauer R. T., (2009), ‘‘Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine’’, *PNAS.*, 106, 18503–18508.
- [118] Radisky E. S., Lee J. M., Lu C. K., Koshland D. E., (2006), ‘‘Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates’’, *PNAS.*, 103, 6835– 6840.

BIOGRAPHY

Leman Özel completed her undergraduate education between 2012 and 2016 in the Molecular Biology and Genetics department at the Gebze Technical University. Between 2017 and 2020, she worked as a HLA tissue typing biologist with Prof. Dr. Zafer GÜLBAŞ (the Director of Hematology Oncology Department) and his team at Anadolu Medical Center Hospital in affiliation with John's Hopkins Medicine. During her HLA tissue typing career, she performed tissue type matching between leukemia patients and potential donor candidates by using PCR and Sanger Sequencing methods. She took an active role in the related and unrelated donor screening process in domestic and international tissue banks. She started her master's degree in the Molecular Biology and Genetics Department at Gebze Technical University in 2018.

APPENDICES

Appendix A: Publication related with the thesis

Özel L., Özdemir İ. *Kinetic Characterization of Geobacillus kaustophilus ATP Dependent Lon Protease*. 5th GTU Graduate Studies Symposium. Pg. 38. 2021. 10-11 June.