CLONING AND EXPRESSION OF PERIPLASMIC (CLP P-LIKE) AND MEMBRANE-BOUND SERINE PROTEASE GENES OF *THERMOPLASMA VOLCANIUM* IN *ESCHERICHIA COLI*

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BURÇAK DEMİROK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BIOLOGY

JANUARY 2006

Approval of the Graduate School of Natural and Applied Sciences

_____

Prof. Dr. Canan ÖZGEN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science

_____

Prof. Dr. Semra KOCABIYIK
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science

_____

Prof. Dr. Semra KOCABIYIK
Supervisor

Examining Committee Members

Prof. Dr. Aykut KENCE (METU, Biology) _____

Prof. Dr. Semra KOCABIYIK (METU, Biology) _____

Prof. Dr. Orhan ADALI (METU, Biology) _____

Assoc. Prof. Dr. Ayşegül GÖZEN (METU, Biology) _____

Assist. Prof. Dr. Hakan ÖKTEM (Inst. of Applied Maths.) _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Burçak DEMİROK

ABSTRACT

CLONING AND EXPRESSION OF PERIPLASMIC (CLPP-LIKE) AND MEMBRANE-BOUND SERINE PROTEASE GENES OF *THERMOPLASMA VOLCANIUM* IN *ESCHERICHIA COLI*

DEMİROK, Burçak

M. Sc. , Department of Biological Sciences

Supervisor : Prof. Dr. Semra KOCABIYIK

January 2006, 149 pages

Serine proteases are a family of proteases that utilize an activated serine residue in the substrate-binding pocket to catalytically hydrolyze peptide bonds. Enzymes which belong to this family, with a diverse array of metabolic and regulatory functions, play critical roles in cell physiology and pathology. 'Clp's are a class of ATP dependent serine proteases which are composed of a protease (ClpP) and an ATPase (ClpA or ClpX) component. Their involvements in degrading proteins are especially implicated under stress conditions. In contrast to members of *Bacteria* and *Eukarya*, little is known about the energy-dependent proteolysis and there is no report on Clp family proteases in *Archaea*.

In this study, for the fist time, a periplasmic Clp P-like (PSP) and a membrane bound serine protease (MSP) genes from thermophilic archaeon *Thermoplasma volcanium* GSS1 were cloned and expressed in *E. coli*. PCR amplifications at 55 º C yielded unique fragments of 971 and 1521bp, for PSP and MSP genes, respectively, which were ligated to p-Drive cloning vectors and introduced into

E.coli TG1 competent cells. Recombinant clones were screened depending on blue/white colony selection. Putative recombinant plasmids were analyzed by restriction enzyme digestions. Serine protease activities of the three positive clones (*E. coli* TG-S1, *E. coli* TG-S4 and *E. coli* TG-M1) were determined spectrophotometrically by using chromogenic oligopeptide substrates. These results indicated that cloned PSP and MSP genes were successfully expressed in *E. coli* under the control of their own promoters. Heterologous expression of PSP gene was also attempted by adding 6xHis tag to the 5′ end of the PSP gene in pQE 30 expression vector. Competent *E.coli* TG1 cells were transformed by pQE expression constructs. Positive clones were detected on colony blots using Anti-His HRP conjugates and chromogenic DAB substrate. Plasmids of these colonies were analyzed by restriction digestions to select the true recombinants. Expression of the 6xHis-PSP fusion protein from the recombinant *E. coli* TG-pQE-S1.7 strain was confirmed by functional analysis and SDS-PAGE.

An NCBI domain search and multiple sequence alignment using Clustal W 1.82 program indicated homologies between PSP and MSP of *Tp. volcanium* and various bacterial ATP dependent ClpPs.

Signal peptide search using Signal P 3.0 server predicted a signal peptide sequence in MSP homologous to that of Gram (+) bacteria.

Key words: Serine protease, *Thermoplasma volcanium*, Cloning, Expression.

ÖZ

THERMOPLASMA VOLCANIUM'UN PERİPLAZMİK (CLP P-BENZERİ) ve MEMBRANA-BAĞLI SERİN PROTEAZ ENZİM GENLERİNİN *ESCHERICHIA COLI*'DE KLONLANMASI VE ANLATIMI

DEMİROK, Burçak
Yüksek Lisans, Biyolojik Bilimler Bölümü
Tez Yöneticisi: Prof. Dr. Semra KOCABIYIK

Ocak 2006, 149 sayfa

Serin proteazlar, peptid bağlarının katalitik hidrolizini gerçekleştirmek için substrat bağlanma bölgesindeki aktif serin grubunu kullanan bir proteaz ailesidir. Bu aileye üye, farklı metabolik ve düzenleyici fonksiyonlara sahip enzimler, hücre fizyolojisi ve patolojisinde kritik rol oynarlar. Bir proteaz (ClpP) ve bir ATPaz (ClpA veya ClpX) bileşeninden oluşan 'Clp'ler , ATP bağımlı serin proteazların bir sınıfını oluştururlar. Proteinlerin parçalanmasında rol oynamaları özellikle stres koşullarıyla ilişkilidir. Bakteri ve Ökaryotlar'ın aksine, Arkea'daki enerji-bağımlı proteoliz hakkında çok az şey bilinmektedir ve Clp ailesine ait proteazlar hakkında bir rapor bulunmamaktadır.

Bu çalışmada, ilk kez, termofilik bir arkea olan *Thermoplasma volcanium*'un Clp P benzeri periplazmik (PSP) ve membrane bağlı serin proteaz (MSP) enzim genleri *E. coli*'de klonlanmış ve anlatımı sağlanmıştır. 55ºC' de gerçekleştirilen PZR amplifikasyonları sonucu, PSP ve MSP genleri için sırasıyla 971 ve 1571

bç uzunluğunda fragmentler elde edilmiştir. Bu fragmentler pDrive klonlama vektörlerine bağlanmış ve kompetan *E. coli* hücrelerine aktarılmışlardır. Rekombinant klonlar mavi/beyaz koloni ayrımına göre seçilmişlerdir. Olası recombinant plazmitler restriksiyon enzimi kesimleriyle analiz edilmiştir. Üç pozitif klonun (*E. coli* TG-S1, *E. coli* TG-S4 ve *E. coli* TG-M1) serin proteaz aktiviteleri kromojenik oligopeptit substratlar kullanılarak spektrofotometrik yöntemle belirlenmiştir. Bu sonuçlar, klonlanan PSP ve MSP genlerinin *E. coli*'de , kendi promotorlarının kontrolünde, başarıyla anlatıldığını göstermiştir. Ayrıca, PSP geninin 5′ ucuna, pQE 30 anlatım vektöründe, 6xHis takısı eklenerek heterolog olarak anlatımı çalışılmıştır. Oluşturulan pQE anlatım yapıları, kompetan *E.coli* TG1 hücrelerine transformasyonla aktarılmışlardır. Pozitif klonlar, Anti-His HRP konjügatı ve DAB substratı kullanılarak koloni hibridizasyonu ile belirlenmiştir. Gerçek recombinantların seçilmesi için, bu kolonilerden izole edilen plazmitler restriksiyon enzimi kesimleriyle analiz edilmiştir. 6xHis-PSP füzyon proteinin rekombinant *E. coli* TG-pQE-S1.7 suşunda anlatıldığı işlevsel analizler ve SDS-PAGE yöntemiyle doğrulanmıştır.

NCBI domain araştırması ve Clustal W 1.82 programıyla yapılan çoklu dizi hizalamaları, *Tp. volcanium*'un PSP ve MSP genleri ile çeşitli bakteriyel ATP bağımlı ClpP'ler arasında homoloji olduğunu göstermiştir.

Signal P 3.0 servis sağlayıcısı ile yapılan sinyal peptit araştırması, MSP gen dizisi içinde Gram (+) bakterilerdeki diziler ile homolog bir sinyal peptitin var olabileceğini göstermiştir.

Anahtar Kelimeler: Serin proteaz, *Thermoplasma volcanium*, Klonlama, Anlatım.

To My Mother

With my only wish to be together in heaven forever…

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor Prof. Dr. Semra Kocabıyık for her guidance, advice, criticism, encouragements and insight throughout the research.

I would also like to thank my laboratory mate Res. Assist. Bilsev Koyuncu for her suggestions, comments and helps.

The spiritual power of my mother Fzt. Neşe Demirok gave me the resolve upon this study.

This study was supported by the State Planning Organization (DPT) Grant No: BAP-08-11-DPT-2002K125/510.

TABLE OF CONTENTS

3. RESULTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF ABBREVIATIONS

Ala	Alanine
Bis	Ammonium persulfate
BSA	Bovine serum albumin
CFE	Cell free extract
CBZ	Carbobenzoxyl
CDD	Conserved domain database
COG	Clusters of orthologous groups
dd	Double distilled
EDTA	Ethylenediaminetetraacetic acid
MSP	Membrane-bound serine protease
PAGE	Polyacrylamide gel electrophoresis
Phe	Phenylalanine
pNA	Para nitroanilide
Pro	Proline
PSP	Periplasmic (ClpP-like) serine protease
SDS	Sodium dodecylsulfate salt
Suc	Succinyl
TBS	Tris buffered saline

CHAPTER 1

INTRODUCTION

1.1. Proteases

Proteases are degradative enzymes which catalyze the hydrolysis of proteins and involved in the conversion of proteins into peptides or amino acids. Proteases execute a large variety of functions, extending from the cellular level to the organ and organism level, to produce cascade systems such as hemostasis and inflammation. They are responsible for the complex processes involved in the normal physiology of the cell such as cell growth and differentiation, protein catabolism, migration, tissue arrangement and morphogenesis in development. But, it has also been demonstrated that proteases conduct highly specific and selective modifications of proteins such as activation of zymogenic forms of enzymes by limited proteolysis, blood clotting and lysis of fibrin clots, and processing and transport of secretory proteins across the membranes. They, also, play a critical role in many pathological processes such as tumor growth and metastasis. Since they are physiologically essential molecules, proteases occur ubiquitously in a wide variety of organisms including viruses, prokaryotes, fungi, plants and animals (Gupta*, et al,* 2002; Rao, *et al*, 1998). These diverse physiological niches demand proteases with wildly varied specificities, ranging from digestive proteases that cleave after hydro-phobic or positively charged residues, to proteases that recognize a five-residue cleavage site or even a single protein (Hedstrom, 2002).

There is renewed interest in the study of proteolytic enzymes, since these

enzymes do not only play an important role in the cellular metabolic processes but have also gained considerable attention in the industrial community to exploit their biotechnological applications (Fox, *et al*, 1991; Poldermans, *et al*, 1990).

The amount of proteolytic enzymes produced worldwide on a commercial scale is larger than that of any of the other biotechnologically used enzymes (Niehaus, *et al*, 1999). Today, proteases account for approximately 40% of the total enzyme sales in various industrial market sectors. These enzymes have become widely used in the detergent industry, since their first introduction in 1914 as detergent additives (Gupta*, et al,* 2002). Serine alkaline proteases are used as additives to household detergents for laundering, where they have to resist denaturation by detergents and alkaline conditions. Therefore, until today, the largest share of this enzyme market has been held by serine alkaline proteases active and stable in the alkaline pH range (Niehaus, *et al*, 1999). Proteases, also, have a long history of application in the food industries. Their application in the leather industry for dehairing and bating of hides to substitute currently used toxic chemicals is a relatively new development and has conferred additional biotechnological importance. On the other hand, enzymatic peptide synthesis (e.g. dipeptides and tripeptides) that uses proteases offers several advantages over chemical methods (So, *et al*, 2000). In addition, their involvement in the life cycle of disease-causing organisms has led them to become a potential target for developing therapeutic agents against fatal diseases such as cancer and AIDS. The wide diversity and specifity of proteases are used as great advantage in developing effective therapeutic agents (Rao, *et al*, 1998; Kudrya, *et al*, 1994; Kim, *et al*, 1996). Recently, the use of alkaline proteases in the management of wastes from various food processing industries opened up a new era in the use of proteases in waste management (Gupta, *et al*, 2002). Alkaline proteases play a crucial role in the bioprocessing of used X-ray or photographic films for silver

recovery (Ishikawa, *et al*, 1993). One of the least explored areas for the use of proteases is the silk industry for the degumming of silk (Gupta, *et al*, 2002). Besides their industrial and medicinal applications, proteases play an important role in basic research. Their selective peptide bond cleavage is used in the elucidation of structure-function relationship and in the sequencing of proteins (Ladenstein, *et al*, 1998).

1.2. Protease Classification

According to the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology, proteases are classified in subgroup 4 of group 3 (hydrolases) (Enzyme Nomenclature, 1992). However, proteases do not comply easily with the general system of enzyme nomenclature due to their huge diversity of action and structure. Currently, proteases are classified on the basis of three major criteria: (i) type of reaction catalyzed, (ii) chemical nature of the catalytic site, and (iii) evolutionary relationship with reference to structure (Barret, *et al*, 1994).

Proteases are grossly subdivided into two major groups, exopeptidases and endopeptidases, depending on their site of action. Exopeptidases act only near the ends of polypeptide chains. Based on their site of action at the N or C terminus, they are classified as amino- and carboxypeptidases, respectively. Endopeptidases are characterized by their preferential action at the peptide bonds in the inner regions of the polypeptide chain away from the N and C termini.

Based on the functional group present at the active site, proteases are further classified into four prominent groups, i.e., serine proteases (EC.3.4.21), cysteine proteases (EC.3.4.22), aspartic proteases (EC.3.4.23), and metalloproteases (EC.3.4.24). There are a few miscellaneous proteases which do not precisely fit

into the standard classification, e.g., ATP-dependent proteases which require ATP for activity (Menon, *et al*, 1987). Based on their amino acid sequences, proteases are classified into different families and further subdivided into "clans" to accommodate sets of peptidases that have diverged from a common ancestor. To facilitate quick and unambiguous reference to a particular family of peptidases, Rawlings and Barrett have assigned a code letter denoting the catalytic type, i.e., S, C, A, M, or U followed by an artibrarily assigned number (Rawlings, *et al*, 1993).

1.3. Serine Proteases

Serine proteases (SP) are a family of enzymes that utilize a uniquely activated serine residue as a nucleophile in the substrate-binding pocket to catalytically hydrolyze peptide bonds (Schultz, *et al*, 1997). They are numerous and widespread among viruses, bacteria, and eukaryotes, suggesting that they are vital to the organisms.

Serine proteases are found in the exopeptidase, endopeptidase, oligopeptidase, and omega peptidase groups and constitute almost one-third of all proteases (Hedstrom, 2002). Based on their structural similarities, serine proteases have been grouped into 20 families, which have been further subdivided into about six clans with common ancestors (Barret, *et al*, 1994). The primary structures of the members of four clans, chymotrypsin (SA), subtilisin (SB), carboxypeptidase C (SC), and *Escherichia* D-Ala-D-Ala peptidase A (SE) are totally unrelated, suggesting that there are at least four separate evolutionary origins for serine proteases. Clans SA, SB, and SC have a common reaction mechanism consisting of a common catalytic triad (Figure 1.1) of the three amino acids, serine (nucleophile), aspartate (electrophile), and histidine (base). Although the geometric orientations of these residues are similar, the protein folds are quite different, forming a typical example of a convergent evolution. The catalytic mechanisms of clans SE and SF (repressor LexA) are distinctly different from those of clans SA, SB, and

SE, since they lack the classical Ser-His-Asp triad (Rao, *et al*, 1998) More recently, serine proteases with novel catalytic triads and dyads have been discovered, including Ser—His—Glu, Ser—Lys/His, His—Ser—His, and N-terminal Ser (Dodson, *et al*, 1998).

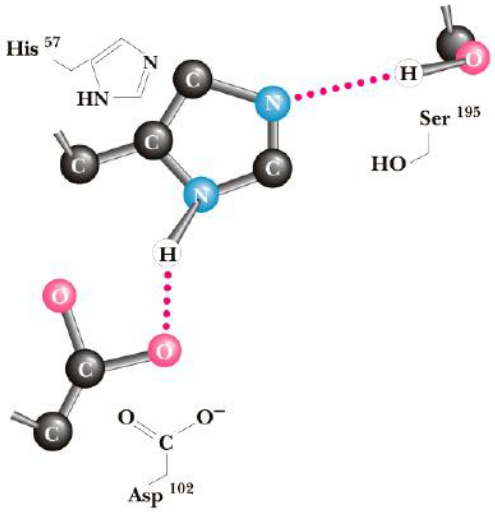

Figure 1.1 : Catalytic triad

Serine proteases are recognized by their irreversible inhibition by 3,4-dichloroisocoumarin (3,4-DCI), L-3-carboxytrans 2,3-epoxypropyl-leucylamido (4-guanidine) butane (E.64), di-isopropylfluorophosphate (DFP), phenylmethylsulfonyl fluoride (PMSF) and tosyl-L-lysine chloromethyl ketone (TLCK). Some of the serine proteases are inhibited by thiol reagents such asp-chloromercuribenzoate (PCMB) due to the presence of a cysteine residue near the active site (Rao, *et al*, 1998).

Serine proteases are generally active at neutral and alkaline pH, with an optimum between pH 7 and 11. They have broad substrate specificities including esterolytic and amidase activity. Their molecular masses range between 18 and 35 kDa,

although there are exceptions such as the serine protease from *Blakeslea trispora,* which has a molecular mass of 126 kDa. The isoelectric points of serine proteases are generally between pH 4 and 6. Serine alkaline proteases that are active at highly alkaline pH represent the largest subgroup of serine proteases (Rao, *et al*, 1998).

The carboxypeptidases (Clan SC) are unusual among the serine dependent enzymes in that they are maximally active at acidic pH. These enzymes are known to possess a Glu residue preceding the catalytic Ser, which is believed to be responsible for their acidic pH optimum.

Serine proteases exhibit preference for hydrolysis of peptide bonds adjacent to a particular class of amino acids. In the trypsin-like group, the protease cleaves peptide bonds following basic amino acids such as arginine or lysine, since it has an aspartate (or glutamate) in the substrate-binding pocket, which can form a strong electrostatic bond with these residues (George, *et al*, 2003) The recent studies suggest a putative general substrate binding scheme for proteases with specificity towards glutamic acid involving a histidine residue and a hydroxyl function (Rao, *et al*, 1998). The chymotrypsin-like proteases have a non-polar substrate-binding pocket, and thus, require an aromatic or bulky non-polar amino acid such as tryptophan, phenylalanine, tyrosine or leucine. The elastase-like enzymes, on the other hand, have bulky amino acids (valine or threo-nine) in their binding pockets, thus requiring small hydrophobic residues, such as alanine.

Serine protease specificity can usually be rationalized by the topology of the substrate binding sites adjacent to the catalytic triad (the "active site cleft"). The substrate recognition sites include the polypeptide binding site and the binding pockets for the side chains of the peptide substrate. Investigations of protease specificity have generally focused on the P1/S1 interaction, where PI—PI' denotes peptide residues on the acyl and leaving

group side of the scissile bond, respectively. The adjacent peptide residues are numbered outward and S1, SI', etc. denote the corresponding enzyme binding sites (Hedstrom, 2002). The Pl residue exclusively dictates the site of peptide bond cleavage. The primary specificity is affected only by the Pl residues; the residues at other positions affect the rate of cleavage.

1.3.1 Action Mechanism of Serine Proteases

All proteases must overcome three obstacles to hydrolyze a peptide bond: (a) amide bonds are very stable due to electron donation from the amide nitrogen to the carbonyl. Proteases usually activate an amide bond via the interaction of the carbonyl oxygen with a general acid, and may also distort the peptide bond to disrupt resonance stabilization; (b) water is a poor nucleophile; proteases always activate water, usually via a general base; and (c) amines are poor leaving groups; proteases protonate the amine prior to expulsion. Serine proteases perform these tasks very efficiently by following a two-step reaction for hydrolysis in which a covalently linked enzyme-peptide intermediate is formed with the loss of the amino acid or peptide fragment. This acylation step is followed by a deacylation process which occurs by a nucleophilic attack on the intermediate by water, resulting in hydrolysis of the peptide (Figure 1.2).

Once the substrate is bound, the hydroxyl group of serine 195 nucleophilically attacks the scissile peptide's carbonyl group. The hydroxyl group is made more nucleophilic by the neighboring histidine that extracts the hydroxyl hydrogen, a process that is facilitated by the polarizing effects of aspartate 102. A covalent bond forms between the serine and the substrate to yield a complex known as the tetrahedral intermediate. The tetrahedral intermediate, which resembles the reaction's transition state, is stabilized by two amide hydrogens that hydrogen bond to the anionic oxygen. This region is known as the "oxyanion hole"

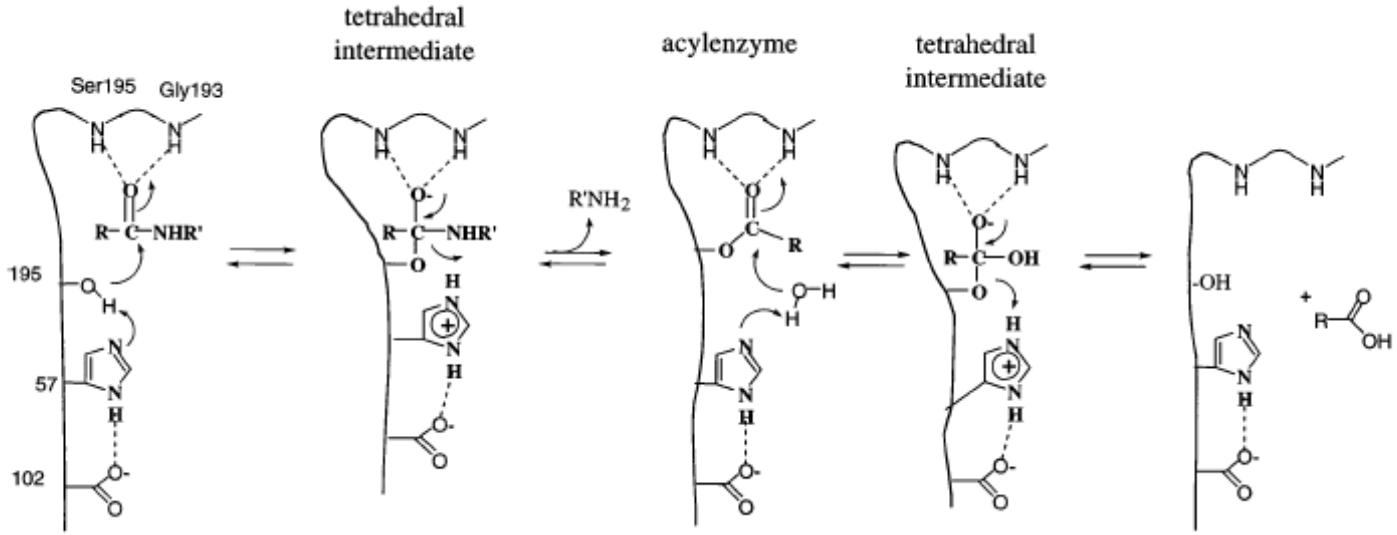

Figure 1.2 : Generally accepted mechanism for chymotrypsin-like serine proteases. In the acylation half of the reaction, Ser195 attacks the carbonyl of the peptide substrate, assisted by His57 acting as a general base, to yield a tetrahedral intermediate. The resulting His57-H^+ is stabilized by the hydrogen bond to Asp 102. The oxyanion of the tetrahedral intermediate is stabilized by interaction with the main chain NHs of the oxyanion hole. The tetrahedral intermediate collapses with expulsion of leaving group, assisted by His57-H^+ acting as a general acid, to yield the acylenzyme intermediate. The deacylation half of the reaction essentially repeats the above sequence: water attacks the acyl-enzyme, assisted by His57, yielding a second tetra-hedral intermediate. This intermediate collapses, expelling Ser195 and carboxylic acid product (Hedstrom, 2002).

because it is occupied by the intermediate's oxyanion group. The tetrahedral intermediate quickly decomposes back to a planar carbonyl group. This can happen in one of two ways. The bond to the serine can dissolve regenerating the original starting components (this is just the reverse of step 1), or the bond to the nitrogen can break, releasing the C-terminal portion of the substrate while forming an ester linkage to the N-terminal portion. This latter assembly is called the acyl-enzyme intermediate. The hydrogen temporarily held by the histidine has now been passed on to the leaving polypeptide fragment. Next, the ester bond of the acyl-enzyme must be broken. This is accomplished by another nucleophilic attack on the carbonyl group, this time by a water that has diffused into the active site. The water transfers one hydrogen to the histidine while forming a covalent bond to the carbonyl carbon. The result is another tetrahedral intermediate stabilized by the amide groups in the oxyanion hole.

In the last step, the tetrahedral intermediate decomposes by breaking the bond to the serine hydroxyl group. The hydrogen held by the histidine is transferred to the serine and the substrate is released with a carboxylic acid terminus. The enzyme is returned to its original state ready to bind the next protein.

1.3.2 Physiological Functions of Serine Proteases

SP carry out a diverse array of physiological functions, the best known being digestion, blood clotting, fibrinolysis, fertilization, and complement activation during immune responses in higher organisms. They have also been shown to be associated with many diseases including cancer, arthritis, and emphysema (George, *et al*, 2003).

Interactions between serine proteases are common, and substrates of serine proteases are sometimes other serine proteases that are activated from an inactive precursor (zymogen). The involvement of serine proteases in cascade pathways is

well documented. One important example is the blood coagulation cascade. Blood clots are formed by a series of zymogen activations. In this enzymatic cascade, the activated form of one factor catalyzes the activation of the next factor. Very small amounts of the initial factors are sufficient to trigger the cascade because of the catalytic nature of the process. These numerous steps yield a large amplification, thus ensuring a *rapid* and *amplified* response to trauma. A similar mechanism is involved in the dissolution of blood clots where activation of plasminogen activators leads to conversion of plasminogen to plasmin which is responsible for lysis of the fibrin clot. A third important example of the coordinated action of serine proteases is that of the intestinal digestive enzymes; coordinated control is achieved by the action of trypsin as the common activator of all pancreatic zymogens; trypsinogen, chymotrypsinogen, proelastase, and procarboxypeptidase. The apoptosis pathway is another important example of the coordinated action of proteases. More recently, a cascade mechanism is hypothesized for kallikrein involvement in cancer and inflammation (Yousef, *et al*, 2003, Bhoola, *et al*, 2001).

Owing to the expanding roles for serine proteases, there has been increasing interest in the identification, structural, and functional characterization of all members of the serine protease family of enzymes in humans and other organisms. The near completion of the 'Human Genome Project' provides a unique opportunity for such efforts. Structural characterization of all serine proteases and extensive analysis of their location are the first steps towards understanding the control of their gene expression and their involvement in various physiological and pathological conditions.

1.4 Energy Dependent Proteolysis and Protein Turnover

All living cells maintain a particular rate of protein turnover by continuous, albeit balanced, degradation and synthesis of proteins. Catabolism of

proteins provides a ready pool of amino acids as precursors of the synthesis of proteins.

Proteolysis plays a dynamic and vital role in the cell's response to environmental stimuli. It controls metabolic fluxes by regulating the levels of key rate-limiting enzymes, while also removing irreversibly damaged polypeptides that may interfere with these pathways. The appearance of such abnormal proteins occurs continuously through events such as spontaneous denaturation, biosynthetic errors or accumulative mutations. Proteolytic removal of these non-functional proteins is therefore crucial for cell homeostasis and optimal metabolic activities. This role for proteases is particularly vital during stresses that exacerbate the occurrence of damaged proteins. As a consequence, many proteases are induced during such adverse conditions to prevent the accumulation of abnormal and thus potentially toxic polypeptides (Porankiewicz, *et al*, 1999).

Apart from removing damaged and/or deleterious proteins for quality control purposes, proteases are involved in important regulatory mechanisms (Gottesman and Maurizi, 1992), among others the adaptation to stationary-phase (Weichart, *et a*l., 2003) and starvation (Schweder, *et al*., 1996), the heat-shock (Bukau, 1993) and general stress (Zhou, *et al*, 2001) responses, cell division (Schoemaker, *et al*, 1984), mutagenesis (Frank, *et al*, 1996), phage (Gayda and Markovitz, 1978) and plasmid replication (Maas, 2001) and capsule synthesis (Torres-Cabassa and Gottesman, 1987). The ATP-dependent serine proteases ClpP and Lon are critical for the removal of misfolded or otherwise damaged proteins and they are responsible for more than 70% of the proteolysis covering most of the regulatory functions listed in *E. coli* (Maurizi, 1992) above (Tomoyasu, *et al*, 2001; Rosen, *et al*, 2002). Furthermore, Lon and ClpP proteases participate in the physiological disintegration of bacterial inclusion bodies (Vera, *et al*, 2005).

1.4.1. Energy Dependent Proteolysis

The proteolysis can be divided into two parts based on the utilization of metabolic energy. The proximal ATP dependent steps are followed by ATP-independent events. Energy-dependent protein degradation in viable cells plays important roles in the rapid turnover of short-lived regulatory proteins and in the quality control of intracellular proteins by eliminating damaged or denatured proteins.

The basic model of protein unfolding followed by degradation during energy-dependent proteolysis is conserved in eubacteria, archaea and eukaryotes. In prokaryotic cells and in the organelles of higher eukaryotes, energy-dependent proteolysis is accomplished by oligomeric ATP-dependent proteases, such as Lon, FtsH, ClpAP, ClpXP, and HslVU (ClpYQ) (Gottesman, 1996), and, in a few cases, genuine 20S proteasomes whereas multicatalytic complexes (26 S proteasomes) perform that function in the cytosol of eukaryotic cells.

Proteins destined for degradation are recognized and unfolded by regulatory subcomplexes that invariably contain ATPase modules, before being translocated into another subcomplex, the proteolytic core, for degradation. These heterooligomeric ATP-dependent proteases, as well as 26 S proteasomes, display similar overall architecture despite a lack of similarity in sequence and three-dimensional folds of their proteolytic subunits. The AAA^+ modules are contained in separate subunits, which assemble into six- or seven-membered rings. The proteolytic subunits also assemble into six- or seven-membered rings that stack upon each other to form barrel-shaped complexes with a central cavity containing the proteolytic active sites accessible by narrow axial channels in the rings. The ATPase rings interact at both ends of the protease barrels (Zwickl, *et al*, 2000).

1.4.1.1 ATP-Dependent Proteases

The ATP-dependent proteases are composed of a chaperone component or domain that specifically recognizes protein targets and couples ATP hydrolysis to unfolding and translocation of the polypeptide chains into the interior chamber of an associated protease domain where processive proteolysis takes place. These enzymes are members of the extended AAA^+ family (**A**TPases **A**ssociated with a variety of cellular **A**ctivities), a group of proteins whose diverse activities often require the ATP-modulated assembly of oligomeric (often hexameric) rings. AAA^+ ATPases function as essential components of energy-dependent compartmental proteases in all biological kingdoms. Besides being involved in selective proteolysis, AAA^+ proteins participate in diverse cellular processes, including cell-cycle regulation, organelle biogenesis, vesicle-mediated protein transport, microtubule severing, membrane fusion, and DNA replication (Maurizi and Li, 2001; Ogura and Wilkinson, 2001).

ATPase functional or regulatory domains (also called AAA^+ modules) of the AAA^+ proteins consist of a single polypeptide chain containing 220–250 amino acids. Within the AAA^+ superfamily, these functional domains are present either once or as repeats (Neuwald, *et al*, 1999). Each AAA^+ module consists of two structural domains, the larger one at its N terminus (referred to as the α/β domain) and the smaller one at its C terminus (referred to as the α domain). The large domains are highly conserved and typically contain a Rossmann fold, common in many nucleotide-binding enzymes. The small domains typically contain three or four α-helices but vary in sizes and exhibit substantially greater structural variation (Hattendorf and Lindquist, 2002; Lupas and Martin, 2002). The α/β domains of AAA^+ proteins contain the Walker A and B nucleotide binding motifs (Walker, *et al*, 1982) shared by other P-loop type ATPases. In addition, the α/β domain contains the ‘‘sensor-1’’ motif, consisting of a polar residue, which forms a hydrogen bonding network crucial in positioning of a

nucleotide binding water molecule (Lenzen, *et al*, 1998; Yu, *et al*, 1998), as well as a highly conserved arginine within the Box VII motif that projects from one subunit into the active site of the neighboring subunit, acting as an ‘‘arginine finger’’ (Karata, *et al*, 1999; Neuwald, *et al*, 1999; Putnam, *et al*, 2001). Another motif, ‘‘sensor-2,’’ present in many but not all AAA^{+} proteins, is located near the start of helix 3 of the α domain. The sensor-2 residue, Arg or Lys, participates in binding and hydrolysis of the ATP, which binds in a crevice at the interface of the large and small domains (Hattendorf and Lindquist, 2002; Neuwald, *et al*, 1999).

The proteolytic components of ATP-dependent proteases contain several different active site types. ClpP has a classic serine protease triad (Wang, *et al*, 1997); HslV, like the proteasome, has a catalytic N-terminal threonine residue (Yoo, *et al*, 1997); and FtsH is a zinc-dependent metalloprotease (Tomoyasu, *et al*, 1995). The catalytic residues in Lon have remained uncertain until quite recently. Mutational studies suggested that Lon had a catalytic serine (Amerik, *et al*, 1991); however, other candidate catalytic residues could not be definitively identified.

1.4.1.1.1 Proteasomes

To date, the most fully characterized member of the family of cage-forming ATP dependent proteases is the proteasome. The proteasome is the central protease in non-lysosomal ubiquitin dependent protein degradation, and is involved in the clearance of misfolded proteins, breakdown of regulatory proteins, processing of proteins by specific partial proteolysis, cell cycle control as well as preparation of peptides for immune presentation (Voges, *et al*, 1999).

In eukaryotes, the 26S proteasome is the central protease in the energy-dependent degradation of proteins via the ubiquitin pathway. This quintessential molecular machine, with a cumulative mass of more than 2.5 MDa, comprises two subcomplexes — the 20S core particle and the 19S regulatory particle. The core particle performs the actual proteolysis, whereas the 19S particle is involved in the recognition and binding of substrate proteins, deubiquitination, unfolding and translocation into the core subcomplex. Translocation is thought to be coupled to the gating of the axial entry channel into the proteolytic cavity (Zwickl, *et al*, 2000).

The 20S proteasome is barrel-shaped, with a diameter of 11 nm, a height of 15 nm and a mass of 700 kDa. The barrel is a bipolar stack of four seven-membered rings, the two outer rings being formed by α subunits and the two inner rings by β subunits. As most prokaryotic proteasomes have single species of both subunits, they form homomeric rings, whereas eukaryotic proteasomes contain heteromeric rings formed by their seven distinct α- and β-type subunits. The central β rings enclose α cavity, approximately 5 nm in diameter that houses the active sites: this inner compartment is a key structural feature common to all self-compartmentalizing proteases so far characterized (Groll, *et al*, 2003).

The proteasome from the archaeon *Thermoplasma acidophilum* represents the prototype for the quaternary structure and topology of the core particle, 20S proteasome (Seemüller, *et al*, 1995). The core particle comprises 28 subunits, which are arranged in four seven membered rings that stack upon each other, yielding an α7β7β7α7 complex. The two inner rings are solely built of the β-subunits, which delimit the hydrolytic chamber. This central cavity is connected, via two narrow gates, to the two outer chambers – the antechambers – each consisting of one α- and one β-ring. A small axial pore in the center of the α-rings provides access to the antechambers, enabling substrate entry and

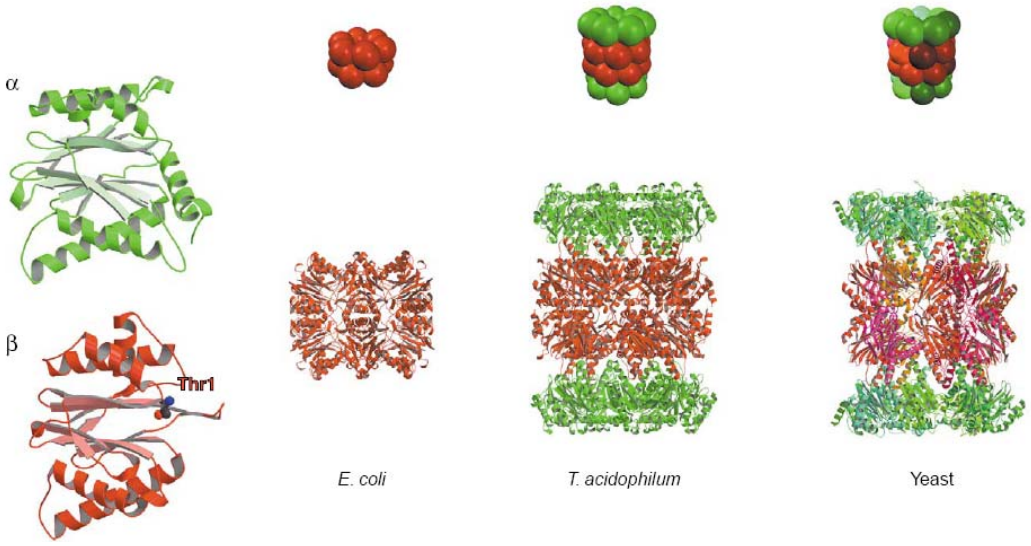

Figure 1.3 : Overall architecture of the 20S proteasome. The folds of the individual α- and β-subunits, and the subunit composition of eubacterial, archaeal and eukaryotic (yeast) proteasome complexes are illustrated. The catalytically inactive α-subunits are depicted in green and the active β-subunits in red. The position of the essential nucleophile, Thr1, is indicated. (Groll, *et al*, 2003)

product release. As anticipated from sequence comparison, the noncatalytic α-subunit and the catalytic β-subunit share the same fold, a four-layer structure in which two sheets of β-strands are sandwiched between two layers of α-helices. The α-subunits have an additional N-terminal extension, which is essential for sequestering the interior of the barrel from the outside world. Whereas the active sites of the α-subunits are buried within the protein, those of the β-subunits are freely accessible. They open up into the central cavity and define the proteolytic chamber. Mutagenesis studies and structural analysis of an enzyme–inhibitor complex of the *Thermoplasma* core particle identified the hydroxyl group of Thr1 of the β-subunits as the essential nucleophile that initiates peptide bond hydrolysis (Seemüller, *et al*, 1995). Thus, the proteasome

is a threonine protease belonging to the family of Ntn (N-terminal nucleophile) hydrolases.

The most elaborate version of the core particle is found in eukaryotes, as shown by the crystal structure of the yeast 20S proteasome (Groll, *et al*, 1997). Here, the α- and β-subunits have each diverged into seven different subunits. There are two copies of each subunit and they occupy precisely defined positions within the 20S complex. As in the archaebacterial proteasome, the α-subunits are inactive and contribute to the antechambers of the particle, whereas the β-subunits build up the inner hydrolytic chamber. Remarkably, four of the seven different eukaryotic β-subunits lack residues that are essential for peptide bond cleavage and are, therefore, proteolytically inactive. The remaining three subunits mature autoproteolytically to active threonine proteases, but contain different cleavage capabilities with a caspase-like (β1), trypsin-like (β2) and chymotrypsinlike (β5) activity (Groll, *et al*, 1997). The crystal structure of the bovine 20S proteasome has confirmed that yeast and mammalian core particles are highly homologous in their structural architecture, quaternary assembly and active site geometry.

Eubacteria, which lack ubiquitin, contain at least one proteasome-related system—a miniproteasome termed HslUV. The protease HslV (ClpQ), alone, has only minor peptidase activity and is regulated by the ATPase rings of HslU (ClpY), a member of the Clp100 subfamily. The architecture of HslV represents a simple form of the core particle, consisting of only two homo-oligomeric rings. Structural analyses of HslV from *Escherichia coli*, *Haemophilus influenzae* and *Thermotoga maritima* demonstrate its functional homology to the β-subunits of proteasome core particles (Groll, *et al*, 2003). The HslV protomers build up a complex that looks like a ‘double donut’, with a diameter of 100A ° and a length of 75A °. Although each core particle ring is formed by seven subunits, each bacterial donut consists of only six protomers. The change

from seven- to six-fold symmetry is the result of the absence of the sterically demanding H5 helix at the C-terminus of the core particle β subunit. Nevertheless, the hexameric double donut of HslV enfolds a large cavity that houses the active sites and, thus, conserves the proteolytic mechanism of core particles.

The activation mechanism of eubacterial HslV by HslU is fundamentally different from that of the eukaryotic proteasome. In the eubacterial HslUV complex, two hexameric ATPase rings of HslU bind intimately to opposite sides of the HslV protease, with the HslU 'intermediate domains' extending outward from the complex. In contrast to the eukaryotic core particle, alterations of the HslV entry/exit channel resulting from HslU binding are minimal. For the bacterial proteasome, a mechanism of allosteric activation, in which binding of the HslU chaperone triggers conformational changes in HslV, seems to be valid. The C-termini of HslU bind in pockets between subunits of HslV, thereby displacing its apical helices, which, in turn, transmit a conformational change to the active site region of the protease. This mechanism has been corroborated by the use of a peptide vinyl sulfone inhibitor that forms a covalent linkage with the active site threonine – a reaction that occurs only in the presence of both HslU and ATP. The inhibitor exploits the substrate-binding mode; it is bound to the protein as an antiparallel β-strand by several hydrogen bonds between protease and inhibitor backbones. Comparison of the structure of the HslUV–inhibitor complex with that of uncomplexed HslV reveals that binding of HslU generates a substrate-binding pocket in HslV, in which the substrate can be properly positioned for cleavage (Kwon, *et al*, 2003).

As the number of sequenced genomes grows, a progressively clearer picture of the species distribution of proteasomes is emerging. In general, it appears that proteasomes are ubiquitous and essential in eukaryotes; ubiquitous, but not

essential in archaea; and rare and nonessential in bacteria, in which other energy-dependent proteases abound.

The 20 S proteasome appears to be essential in eukarya. However, some bacteria contain neither a 20 S proteasome nor hslV. Those bacteria containing a 20 S proteasome or its ancestor hslV protease respond differently to its experimental deletion. While *Mycobacterium smegmatis* lacking the 20 S proteasome is phenotypically indistinguishable from the wild-type (Knipfer, *et al*, 1999), deletion of hslV in *Escherichia coli* affects growth at high temperature (Wu, *et al*, 1999) due to a reduced capacity to break down misfolded proteins. Other intracellular proteases such as ClpAP (Gottesman, 1996) and La, the product of the lon-gene, obviously compensate a missing hslV to a certain extent (Goldberg, *et al*, 1994; Gille, *et al*, 2003).

1.4.1.1.2 Lon Proteases

Energy-dependent proteases of the Lon family (MEROPS:1 clan SJ, ID S16) are peptide hydrolases with a catalytic Ser-Lys dyad that participate in the rapid turnover of short-lived regulatory proteins and control protein quality through eliminating mutant and abnormal proteins. Lon proteases all include AAA^+-type ATPase domains (AAA^+ modules) that precede their proteolytic domains, and are found in all three domains of life, *Bacteria*, *Archaea* and *Eucarya*. (Iyer, *et al*, 2004 ; Botos, *et al*, 2005).

The enzymatic properties of Lon protease and its mode of activity have been the subject of extensive studies (Goldberg, *et al*, 1994; Gottesman, *et al*, 1997; Melnikov, *et al*, 2000; Rotanova, 1999).

Escherichia coli Lon protease (Lon*Ec*), an 87-kDa protein, was the first ATP-dependent protease described in detail (Swamy and Goldberg, 1981), with

homologs later identified in other prokaryotes and eukaryotes. This enzyme is active as a homooligomer consisting of four to eight copies (Goldberg, *et al*, 1994) of a single 784 amino acid polypeptide chain (Amerik, *et al*, 1990). Three functional domains with different activities can be identified within each subunit of Lon. The N-terminal domain (N domain, also referred to as LAN) is putatively involved in the recognition and binding of target proteins (Lupas and Martin, 2002). The central part of the chain (A domain or AAA^+ module) is the ATPase domain, while the proteolytic domain (P domain) is located at its C-terminus (Amerik, *et al*, 1991). The crucial residues of the latter domain that are directly involved in the proteolytic activity of Lon were identified by site directed mutagenesis as Ser679 and Lys722 (Amerik, *et al*, 1991; Rotanova, *et al*, 2003). Sequence analysis and crystal structure all independently support the presence of a Ser–Lys catalytic dyad in the active site of Lon protease (Rotanova, *et al*, 2004). The tertiary structure of the Lon proteolytic domain also represented a unique, previously unreported protein fold. Based on these observations, the *E. coli* Lon protease became the founding member of a newly introduced clan SJ in the MEROPS classification of proteolytic enzymes (Barrett, *et al*, 2001).

Division of the Lon family into two subfamilies, based primarily on the characteristics of their catalytic sites, is in agreement with the differences in the respective consensus sequences of their AAA^+ modules. In the LonA subfamily, the Walker A and B motifs are located in the conserved fragments GPPGVGKTS and $P\Phi_4DEIDK$, whereas in the LonB subfamily these motifs are represented by the sequences GXPGXGKSΦ and $G\Phi_4DEIXX$, respectively. The sequences in the vicinity of the conserved sensor-1, arginine finger, and sensor-2 residues (Asn473, Arg484, and Arg542 in *E. coli* LonA protease) are also notably different in LonA and LonB proteases. The other very important differences between the two subfamilies of Lon proteases are the absence of N-terminal domain and the presence of transmembrane fragment in LonB

proteases. According to the evolutionary classification of the AAA^+ ATPases (Iyer, *et al*, 2004, Lupas and Martin, 2002), Lon family belongs to the HslU/ClpX/Lon/ ClpAB-C clade and consists of two distinct branches, bacterial and archaeal Lon, on the basis of the differences in their AAA^+ modules.

The LonA subfamily consists mainly of bacterial and eukaryotic enzymes (MEROPS, clan SJ, ID: S16.001–16.004, S16.006 and partially S16.00X), accounting for >80% of the presently known Lon proteases. The LonA subfamily members mimic the 'classical' Lon protease from *E. coli* and they all contain the N and P domains that flank the AAA^+ module. The overall length of LonA proteases ranges from 772 (*Oceanobacillus iheyensis*) to 1133 (*Saccharomyces cerevisiae*) amino acid residues. The N domains are found to be the most variable, both in their length (220–510 amino acids) and in their amino acid sequences. The P domains of LonA proteases have similar lengths (188–224 amino acids) and are highly homologous. LonA AAA^+ modules show very high homology for their nucleotide binding α/β domains, whereas their α-helical domains vary significantly due to C-terminal insertions or extensions.

ATP-dependent enzymes from the LonB subfamily (< 20% of known Lon proteases) are found only in archaebacteria (MEROPS, ID: S16.005). LonB-like proteins with homologous proteolytic domains but no clearly defined AAA^+ domains are also found in other bacteria (ID: S16.00X, partially). The subunit architecture of archaeal LonB proteases is significantly different from that of LonA proteases. LonB enzymes (621–1127 amino acids) consist of AAA^+ modules and proteolytic domains (205–232 amino acids), but lack the N (LAN) domains (Iyer, *et al*, 2004). These proteins are membrane bound via one or two potential transmembrane (TM) segments that may be part of additional TM domains. The putative TM domains are inserted within the nucleotide-binding domains (α/β), between the Walker A and B motifs. Thus, the architecture of the LonB AAA^+ module is similar to the HslU subunit of HslUV

protease with an insertion domain (I domain) between its Walker motifs (Dougan, *et al*, 2002) . Rotanova, *et al,* (2004) have noticed that some lonB genes (e.g. from *Pyrococcus sp.*) contain selfsplicing elements that encode polypeptides (inteins, 333–474 amino acids), also located between the Walker A and B motifs and following the TM domains. The α domain of archaeal LonB proteases typically consists of 118 residues, except for *Methanocaldococcus jannaschii* LonB, which has 139 residues in its α domain. The recently published crystal structure of the proteolytic domain of LonB protease from *Methanococcus jannaschii* (*Mj*LonB) showed that, whereas the overall fold of the catalytic domain of *Mj*LonB resembled that of *Ec*LonA, significant differences were found in the active site (Im, *et al*, 2004). This observation led to suggestion of a different catalytic mechanism for LonB than that of LonA, although only very limited biochemical characterization of *Mj*LonB was presented.

Site-directed mutagenesis of a number of amino acid residues located in the vicinity of the active site was carried out by Botos, *et al,* (2005) in order to study the role of these residues in the enzymatic mechanism and to explain the structural results obtained for the isolated P-domain. They cloned the full-length cDNA encoding LonB gene of *Archaeoglobus fulgidus* (*Af*LonB) by a PCR approach into expression plasmid pET24 (+). The full-length proteins were expressed in the *E. coli* strain Rosetta (DE3)pLysS using plasmid constructs pET24a(+)-lonAf carrying genes encoding either wild-type enzyme or its mutants. The cell suspensions were disrupted by sonication and the protein fractions were derived after several chromatographic methods. For cloning, expression and purification of the P domain, two recombinant versions of the P-domains of *Af*LonB were amplified from the full-length wild-type gene by PCR. The amplicons from the second PCR were inserted by recombinational cloning into the entry vector pDONR201. The open reading frame encoding the P-domains and containing a recognition site for tobacco etch virus (TEV)

protease at their N terminus was moved by recombinational cloning into the destination vector pDEST-HisMBP to produce plasmids. These plasmids direct the expression of the two different (recombinant) P-domains of *Af*LonB as N-terminal fusions with *E. coli* maltose binding protein (MBP) and with an intervening TEV protease recognition site. The MBP contained an N-terminal hexahistidine tag for affinity purification by immobilized metal affinity chromatography. The fusion proteins were expressed in the *lon*-deficient *E. coli* strain BL21(DE3)pRIL. Fractions containing the recombinant fusion proteins were collected by Ni-NTA Superflow column and incubated with hexahistidine-tagged TEV protease. Fractions containing the Lon P-domains were pooled and concentrated following second Ni-NTA Superflow column. Proteolytic activities of the full-length, wild-type *Af*LonB and of its mutants towards protein substrate (β-casein) were analyzed by SDS-PAGE. The atomic-resolution crystal structure of the proteolytic domain (P-domain, residues 415–621) of *Archaeoglobus fulgidus* B-type Lon protease (wt*Af*LonB) and the structures of several mutants have revealed significant differences in the conformation of the active-site residues when compared to other known Lon P-domains, despite the conservation of the overall fold.

The enzymatic properties of two LonB proteases, isolated from *Thermococcus kodakaraensis* (*Tk*LonB) and *Thermoplasma acidophilum* (*Ta*LonB), were described recently, and they were found to be noticeably different from those of *Ec*LonA.

The first membrane-bound LonB protease to be purified was isolated from *Thermococcus kodakarensis* (*Tk*Lon) and it was shown to be located in the membrane, but was purified from the soluble fraction when expressed in *E. coli* (Fukui, *et al*, 2002).

Sequence analysis predicted two transmembrane regions for the *Thermoplasma* Lon homologue (*Ta*Lon), while all bacterial and eukaryotic Lon polypeptides are devoid of transmembrane regions and are isolated as soluble proteases. The *T. acidophilum* Lon (*Ta*Lon) protease encompasses an N-terminal AAA^+ domain and a C-terminal protease domain, but lacks the N-terminal α-helical domain inherent in most bacterial and eukaryotic Lon homologues. However, the AAA^+ domain of *Ta*Lon contains an insert between the Walker A and Walker B ATPase signatures, which is not present in any bacterial or eukaryotic Lon sequence. This insert is found in all archaeal Lon homologues and is predicted to contain two consecutive transmembrane helices, suggesting that archaeal Lon proteases are membrane associated (Ruepp, *et al*, 2000).

Furthermore, Besche, *et al*, (2004) showed that the Lon protease of *Thermoplasma acidophilum* (*Ta*Lon) is membrane- associated in *Thermoplasma* cells and when expressed in *E. coli*. The recombinant *Ta*Lon protease was purified from the membrane-fraction as a defined complex with ATPase activity. The proteolytic activity was processive, as described for related ATP-dependent enzymes. A systematic mutational analysis of conserved AAA^+ residues yielded *Ta*Lon mutants impaired to different extents in ATPase and protease activity. Comparative analysis of wt and mutant TaLon revealed insights into the individual role of conserved AAA^+ residues in the intramolecular regulation of this ATPdependent protease (Besche, *et al*, 2004)

LonB proteases are expected to bear the functions of the only bacterial membrane-bound ATP-dependent protease, FtsH (MEROPS, ID: M41.001), because the latter enzymes are not present in Archaea (Ward, *et al*, 2002). However, one should not postulate that Archaea contain solely LonB proteases, because the *Methanosarcinacae* genomes are known to encode both LonA and LonB proteases. The metabolic role and biochemical specificity of these bacterial LonB-like proteases are still obscure (Rotanova, *et al*, 2004).

1.4.1.1.3 Clp Proteases

Clps (also called Ti proteases), are barrel-shaped, high molecular mass (700,000 Da) ATP-dependent proteases composed of two components: one catalytic and the other regulatory. Two distinct Clp proteolytic subunits of 23 kDa are known: ClpP and ClpQ (HslV). ClpP is a serine protease, whereas ClpQ belongs to the N-terminal hydrolase family of enzymes with Thr-1 as the active site residue, similar to the β subunit of 20S proteasomes, as mentioned in section 1.7.1.2.1. Clp regulators belong to the Clp/Hsp100 family of ATPases, ClpA, B, X and Y (HslU), and are structurally and functionally distinct (Porankiewicz, *et al*, 1999; Chandu, *et al*, 2004).

Clp found in the cytoplasm of *E. coli* consists of ClpP (the protease component) and ClpA (the ATPase component). The DNA sequence of the *clpP* gene in *E. coli* predicts a polypeptide chain of 207 amino acid residues, including a 14 residue leader peptide, which is rapidly cleaved *in vivo* to yield the 193 residue protein of molecular mass 21,567 Da. Ser111 and His136 of ClpP were found to be essential for the protease activity by site-directed mutagenesis. The amino acid sequence around these active-site residues suggests that ClpP represents a unique class of serine proteases (Maurizi, *et al*, 1990).

Self-compartmentalization is the principal mechanism whereby proteins that are not specifically targeted for degradation by ClpP are shielded from their active sites. As with the 20S proteasome, the proteolytic compartment of ClpP is accessible only through a channel that is so narrow (~10 Å) that it may be negotiated only by unfolded polypeptides. Consistent with this scenario, ClpP on its own is capable of degrading peptides with up to about 30 amino acids. The degradation of larger folded protein substrates requires the collusion of an ATPase. Most large peptides and proteins are degraded at multiple sites of the ClpAP complex without release of high molecular mass intermediates. ClpP

and ClpA associate in the presence of ATP to form an active proteolytic complex, ClpAP, which is composed of a tetradecameric ClpP and a hexameric ClpA. The bacterial protein ClpA, a member of the Hsp100 chaperone family, forms hexameric rings that bind to the free ends of the double-ring serine protease ClpP.

The crystal structure of ClpP revealed a subunit fold quite distinct from that of the 20S proteasome. In essence, ClpP may be considered to be an analog of the 20S proteasome, but with a different subunit fold and without the outer rings of

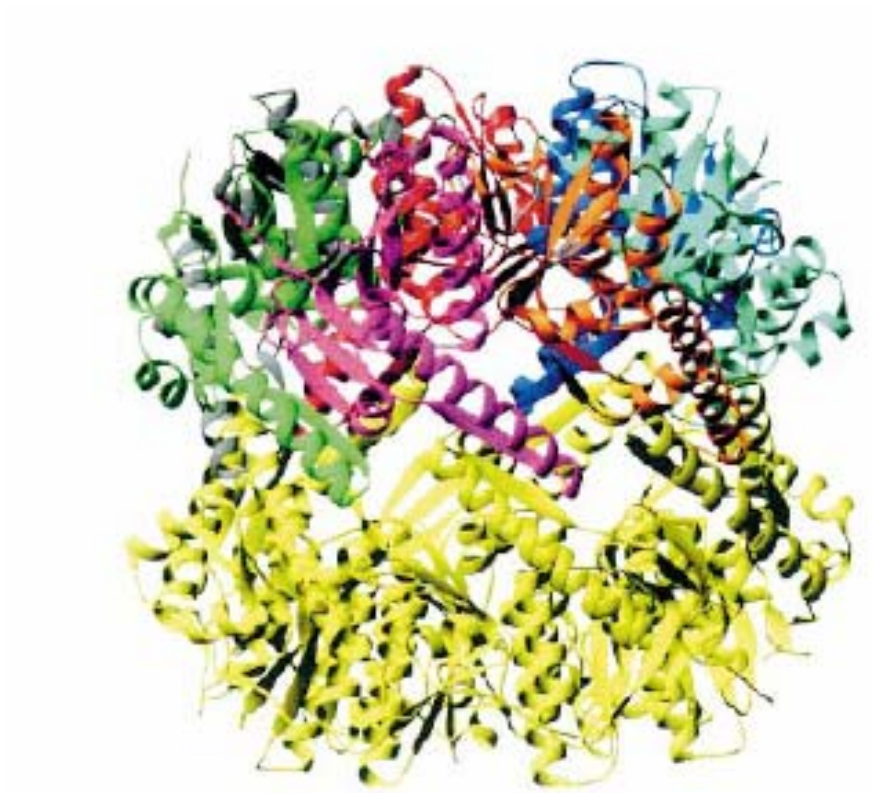

Figure 1.4 : A ribbon representation of the *E. coli* ClpP oligomer showing the interdigitated stacking of two heptameric rings. The bottom ring is shown in yellow, while the top ring is coloured by subunits (Porankiewicz, *et al*, 1999).

α subunits. Although its active site, as a serine protease, also differs from that of the proteasome, a threonine protease, it nevertheless contrives to generate peptides of a similar length to the 20S proteasome in processive degradation (Zwickl, *et al*, 2000).

The ATP-dependent proteolytic activity and the structural organization of the ClpAP complex suggest that it is functionally related to the eukaryotic proteasome (Figure 1.5). The tetradecameric subunit arrangement of ClpP is similar to that of the inner (β-type) subunits of the eukaryotic and archaeal 20 S proteasomes in both the ClpAP complex and the eukaryotic 26 S proteasome, while the oligomeric ATPase, ClpA or 19 S particle, is attached to one or both layers of the protease catalytic core, ClpP or 20 S proteasome. ClpA directs the ATPdependent degradation of substrate proteins bearing specific sequences, much as the 19S ATPase `cap' of eukaryotic proteasomes functions in the degradation of ubiquitinated proteins. This may reflect a possible common functional mechanism between the two protease complexes in cellular protein degradation. (Fukui, *et al*, 2002; Chandu, *et al*, 2004).

1.4.1.1.3.1 The Hsp100/Clp ATPases

Clp/Hsp100 proteins have been implicated in two distinct pathways: protein (un)folding/disaggregation on one hand, and protein degradation and quality control on the other (Dougan, *et al*, 2002; Gottesman, *et al*, 1997; Horwich, *et al*, 1999; Wickner, *et al*, 1999). While ClpB/Hsp104 promotes the former, the latter is mediated by other Clp/Hsp100 proteins, such as ClpA and ClpX, which associate with the ClpP protease and direct the degradation of substrate proteins bearing specific sequences (Horwich, *et al*, 1999; Kim, *et al*, 2001; Wickner, *et al*, 1999; Maurizi, *et al*, 2004).

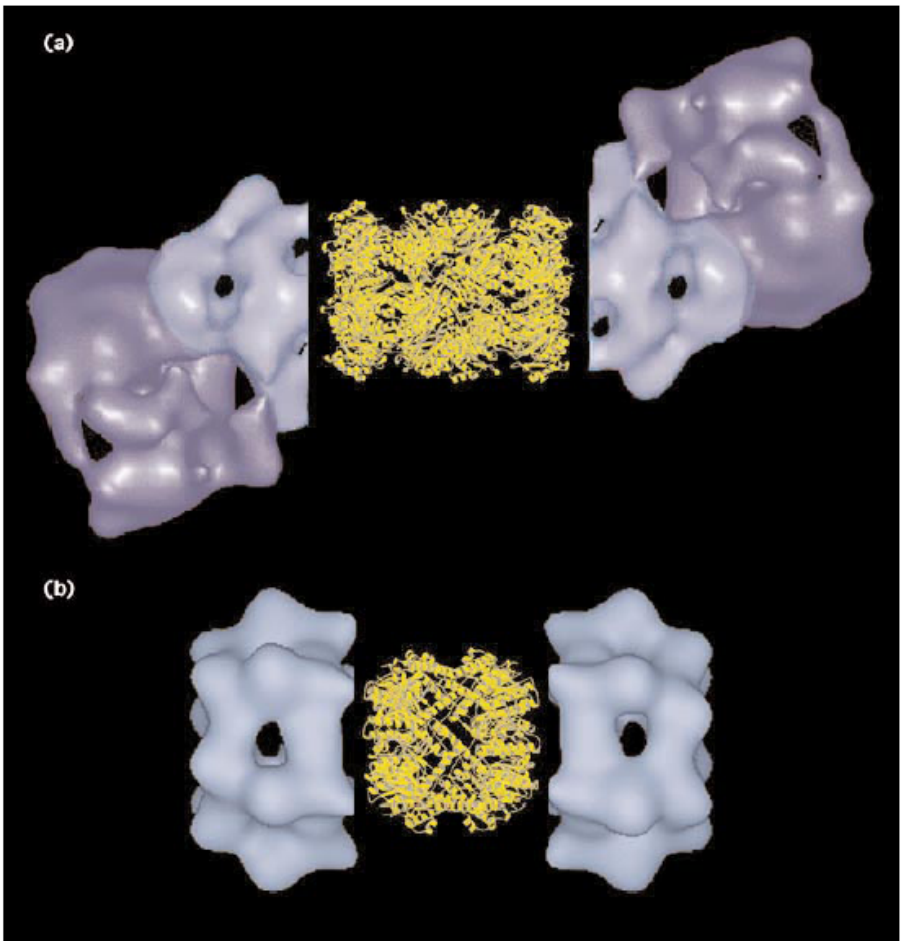

Figure 1.5 : A comparison of the structures of the 26S proteasome and the ClpAP protease (Zwickl, *et al*, 2000).

(a) Model of the 26S proteasome obtained after combining the 3D reconstruction of the19S regulator (blue) — the lid (distal) and base (proximal) subcomplexes are indicated by different shades of blue — from *Drosophila* with the crystal structure of the 20S core (yellow) from *Thermoplasma* .

(b) A model of the ClpAP protease from *E. coli* derived from the combination of the 3D reconstruction of the ClpAP protease (blue) with the crystal structure of the ClpP protease (yellow).

The Clp ATPase family belongs to the AAA^+ superfamily that is characterized by a conserved AAA domain, which contains several conserved motifs including those necessary for ATP binding and hydrolysis, the Walker A and Walker B motifs respectively (Neuwald, *et al*,1999), as mentioned in section 1.7.1.2.. Members of the Clp ATPase family are classified based on the presence of either one or two ATP binding domains (Clarke, 1996; Schirmer, *et al*, 1996). The class 1 Clp proteins posses two ATP binding sites, ATP-1 and ATP-2 and are relatively large (ranging from around 70 to 110 kDa). The variable length of the spacer region separating ATP-1 and ATP-2 as well as the occurrence of specific signature sequences is the basis for subdividing into ClpA, ClpB, ClpC, ClpD, ClpE and ClpL families of Clp 1 ATPases. The smaller, class 2 Clp proteins, such as ClpX and ClpY (HslU), contain a single ATP binding site with most similarity to ATP-2 (Schirmer, *et al*, 1996). Figure 1.6 shows the structurel comparison of ClpA, ClpB, ClpC and ClpX proteins.

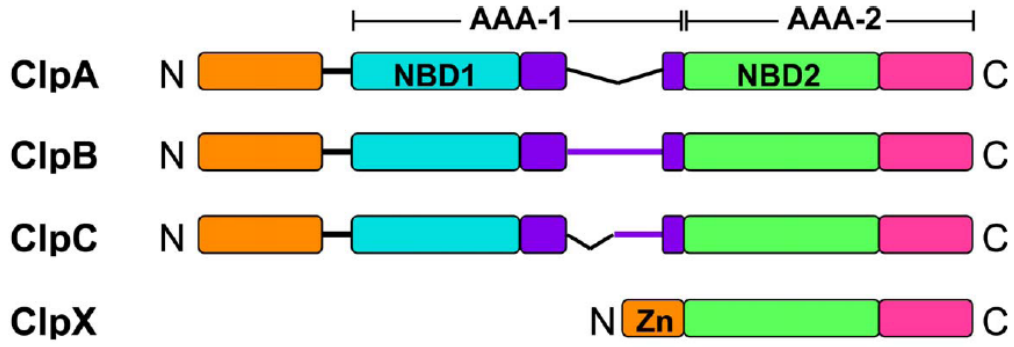

Figure 1.6 : Comparison of Clp/Hsp100 proteins. The N-terminal domain is colored in gold, the NBD1 or large D1 domain in cyan, the α-helical or small D1 domain in purple, the NBD2 or large D2 domain in green, and the C-terminal α-helical or small D2 domain in pink. The N and C indicate the amino- and carboxy termini, and Zn indicates a zinc binding site in ClpX (Lee, *et al*, 2004).

Chaperones and ATP-dependent proteases play a major role for bacterial survival under conditions of stress where proteins tend to unfold and aggregate. In *Escherichia coli*, the expression of ClpA is unaffected by stress, whereas expression of both ClpB and ClpX is induced by heat shock. However, only mutants lacking ClpB have a heat stress-sensitive phenotype (Squires, *et al*, 1991; Kim, *et al*, 1998). This phenotype was later linked to the indispensable role of ClpB for solubilization of larger protein aggregates in cooperation with the DnaK chaperone system (Goloubinoff, *et al*, 1999; Mogk, *et al*, 1999). Similarly, Hsp104 and HSP70, the yeast ClpB and DnaK orthologues, respectively, can solubilize protein aggregates during heat shock (Parsell, *et al*., 1994; Glover and Lindquist, 1998), and this ability has been suggested to be a unique property of the Clp/Hsp100 family (Schirmer, *et al*, 1996). Irreversibly damaged proteins that cannot be refolded to the native state by the sequential action of chaperones are degraded by stress-inducible ATP-dependent proteases.

While non-native polypeptides are mainly disposed by the Lon protease in *E. coli* (Maurizi, *et al*, 1985; Tomoyasu, *et al*, 2001), novel data support that this role is mediated by ClpP proteolytic complexes in Gram-positive bacteria (Frees and Ingmer, 1999; Gaillot, *et al*, 2000; Krüger, *et al*, 2000). In accordance with this assumption, inactivation of *clpP* in *Bacillus subtilis*, *Lactococcus lactis*, *Listeria monocytogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* and *Staphylococcus aureus* resulted in mutants that were significantly more sensitive to heat and other stress conditions than the corresponding wild-type strains (Msadek, *et al*, 1998; Frees and Ingmer, 1999; Gaillot, *et al*, 2000; Lemos and Burne, 2002; Robertson, *et al*, 2002; Frees, *et al*, 2003). In *B. subtilis*, ClpC and to a lesser extent ClpX seem to target heat-denatured proteins for degradation by ClpP (Krüger, *et al*, 2000; Wiegert and Schumann, 2001). The requirement for Clp proteolytic complexes is paralleled by increased synthesis of the constituents under stress conditions.

In an attempt to predict the regulatory elements controlling *clp* gene expression, Frees, *et al,* (2004) mapped the transcriptional start site of *clpP*, *clpX*, *clpL*, *clpC* and *clpB* genes of *Staphylococcus aureus* by primer extension analysis using RNA prepared from cells prior or subsequently to a heat shock treatment. The cells devoid of ClpP could not grow under heat stress conditions supporting the notion that ClpP is essential for degrading heat-denatured proteins in staphylococci (Frees, *et al*, 2003). In contrast, inactivation of ClpX led to improved growth at high temperature, demonstrating that ClpP and ClpX are not cooperating in degrading heat damaged proteins. In order to identify the Clp ATPase(s) that target(s) non-native proteins for degradation by ClpP and to examine the significance of other Clp ATPases for growth under heat stress, they compared growth of wildtype and mutant strains at elevated temperatures by spotting dilutions of cells in early exponential growth phase onto TSA plates and incubating overnight at 37°C, 45°C or 45.5°C. They demonstrated that the action of a ClpP proteolytic complex(es) is essential for staphylococci to grow under heat-shock conditions, and that ClpX does not act as the specificity factor of ClpP under these conditions (Frees, *et al*, 2004).

ClpA**,** alone and as a component of the ClpAP complex, functions like the ATP-dependent chaperones DnaK and DnaJ. In the presence of ATP, ClpA mediates the unfolding of proteins bound after recognition of specific terminal amino-acid sequences. In association with ClpP protease, these destabilized structures are committed to translocation from the cavity of ClpA into the hydrolytic cavity of the protease, where they are degraded. In the absence of associated ClpP, substrate proteins are released from ClpA into solution, where they may spontaneously refold or become bound, by other components such as chaperonins (Weber-Ban, *et al*, 1999). The ClpA subunit has two ATPase domains, which, in the hexamer, form two tiers with a cavity between them.

ClpX ATPase is a member of the large family of closely related ATPases found in both prokaryotic and eukaryotic cells (Schirmer, *et al*, 1996). It can, also, function as a molecular chaperone, such as ClpA, independent of ClpP (Wawrzynow, *et al*, 1995; Wickner, *et al*, 1994). ClpX has a single ATPase domain and forms a single-tiered ring.

When bound to ClpP, both ClpX and ClpA play a role as protein specificity factors by presenting different polypeptide substrates in a form competent for proteolysis by ClpP (Fukui, *et al*, 2002).

To probe functional communication between the ClpX and ClpP enzymes of *Escherichia coli*, Joshi, *et al,* (2004) used changes in ATPase activity to monitor and quantify the strength of the ClpX-ClpP interaction. They found that ClpP binds most tightly to ClpX when the ATPase is denaturing protein substrates, less tightly during translocation of substrates and least tightly in the absence of substrates. Their results suggested that interactions between subunits in ClpX have an important role in the unfolding of protein substrates by preventing quaternary distortions in ClpX that block substrate denaturation and weaken ClpP binding (Joshi, *et al,* 2004).

Bacterial ClpB (also known as heat-shock protein F84.1) and its eukaryotic orthologs, yeast Hsp104 and plant Hsp101, are heat-shock induced and are essential proteins of the stress response (Queitsch, *et al*, 2000; Sanchez and Lindquist, 1990; Squires, *et al*, 1991). ClpB/Hsp104 belongs to the Clp/Hsp100 family of AAA^+ ATPases (Dougan, *et al*,; Neuwald, *et al*, 1999; Schirmer, *et al*, 1996) In contrast to ClpA and ClpX, ClpB/Hsp104 does not associate with a cellular protease (Parsell, *et al*, 1994; Woo, *et al*, 1992) but, instead, functions exclusively as a molecular chaperone. However, unlike other chaperones, ClpB/Hsp104 neither promotes the ‘‘forward’’ folding nor prevents the aggregation of proteins. Rather, ClpB/Hsp104 has the remarkable ability to

rescue proteins from an aggregated state by mediating the disaggregation of stress-damaged proteins (Mogk, *et al*, 1999; Parsell, *et al*, 1994; Sanchez and Lindquist, 1990; Sanchez, *et al*, 1992). The full recovery of these proteins requires the concerted effort of the cognate DnaK/Hsp70 chaperone system (Glover and Lindquist, 1998; Goloubinoff, *et al*, 1999; Mogk, *et al*, 1999; Motohashi, *et al*, 1999; Zolkiewski, 1999). However, the mechanism of this bi-chaperone system remains elusive, in part because a stable complex between ClpB and the DnaK chaperone system has not been observed to date. It is possible that ClpB and the DnaK system interact only transiently (Cashikar, *et al*, 2002; Glover and Lindquist, 1998) or may recognize substrate proteins via a sequential binding mechanism (Goloubinoff, *et al*, 1999).

The *clpB* gene of *S. aureus* encodes a polypeptide with a predicted molecular mass of 98 kDa and has high similarity to ClpB proteins from other bacteria (63% and 52% identity to ClpB from *L. monocytogenes* and *E. coli* respectively) exclusively acts as a chaperone (Frees, *et al*, 2004).

ClpYQ (HslUV) was identified as a heat shock operon with similarity to Clp ATPases and 20S proteasomal β subunits. ClpY and ClpQ can exist separately or they can specifically interact to form the ClpYQ protease in which ATP is required for oligomerization of the ClpYQ complex and protease activity. The crystal structures of ClpYQ, also, revealed a close resemblance with 20S proteasomes with a similar subunit structure and intersubunit interactions. However, 20S proteasomes are composed of heptameric rings, whereas ClpYQ are composed of hexameric rings (Chandu, *et al*, 2004).

Even though the ClpP and HslV peptidases have unrelated structures and active site architectures, ClpXP and HslUV both have mechanisms that allow communication between the ATPase active sites and the peptidase active sites. This functional conservation, even in the absence of structural conservation,

emphasizes the importance of communication between the processing and protease compartments of these energy-dependent proteases during protein degradation (Joshi, *et al*, 2004).

HslVU (ClpQX), itself a heat shock protein, affects the heat shock response by degradation of the heat shock sigma factor and affects the SOS response through the degradation of the cell-division inhibitor SulA. HslVU had eluded all attempts to isolate the ATP-dependent proteolytic activities from the bacterial cytosol, probably because of the lability of the complex (Groll, *et al*, 2005).

1.4.1.1.4 Clp Like Proteases in *Archaeabacteria*

In contrast to members of the domains *Bacteria* and *Eukarya*, little is known about the energy-dependent proteolysis system in members of the third domain, *Archaea*. Complete genome sequences have revealed that archaeal genomes also encode putative orthologs of 20S proteasome and proteasome-activating nucleotidases, as well as open reading frames that share significant identity with Lon proteases from bacterial cytosol and eucaryal mitochondria. *clp* and other ATP-dependent protease genes are not found in archaeal genomes, except for the thermophilic methanogen *Methanobacterium thermautotrophicum*, which encodes a Clp-like ATPase (MTH284) but not a Clp-like protease in the genome (Fukui, *et al*, 2002).

Based on available genome sequence data, hyperthermophilic archaea lack the Clp and HflB (FtsH) family of proteins and have a different version of the Lon protease. Hyperthermophilic archaea, which typically have proteasomes, lack the eukaryotic ubiquitination pathway for selective protein degradation by the proteasome and, therefore, seem to modulate proteolysis at the protease level. Another interesting feature of hyperthermophilic archaeal heat shock response

is the induced formation of unique compatible solutes that have been proposed to stabilize intracellular proteins against thermal denaturation. Whether compatible solutes reduce the need for protein turnover mechanisms is not known. The relative contributions to the collective response of chaperones, chaperonins, proteases, and compatible solutes during heat shock in hyperthermophilic archaea have yet to be examined (Shockley, *et al*, 2003).

1.5 Scope and Aim of the Study

Proteases are no longer viewed as non-specific degradative enzymes associated solely with protein catabolism, but are known to be intimately involved in controlling biological processes that span life to death. They maintain this exquisite control by catalyzing the precisely-timed and rapid turnover of key regulatory proteins. They also interplay with chaperones, in energy dependent proteolysis, to ensure protein quality and to readjust the composition of the proteome following stress.

In contrast to members of the domains *Bacteria* and *Eucarya*, little is known about the energy-dependent proteolysis system in members of the third domain, *Archaea*. Complete genome sequences have revealed that archaeal genomes also encode putative orthologs of 20S proteasome and proteasome-activating nucleotidases, as well as open reading frames that share significant identity with *clp, lon* and other ATP-dependent protease genes. The enzymes isolated from thermophilic microorganisms have been recognized to be of high interest for clarifying the molecular mechanisms at the base of their adaptation and activity at high temperatures, as well as their responses to heat-shock.

In our reaserch we focused on ATP dependent protease inventory of thermoacidophilic *archaeon* (*Euryarchaeota* family) *Thermoplasma volcanium* to investigate their possible roles in regulated proteolysis as well as in heat-

shock response. Analysis of the *T.volcanium* genome revealed the presence of 28 genes encoding peptidases and proteases, among which two are putative Clp-like proteases (ClpP family). ClpP is the functional serine protease unit of ATP dependent Clp protease (ClpXP). To decipher the roles of these potential ClpP orthologs in *Thermoplasma volcanium* in this study, we aimed at cloning and expression of two putative ClpP like serine proteases (Periplasmic ClpP like and membrane-bound) in *E. coli*. Enzymatic activities associated with the heterologous proteins were studied in the cell free extracts. An extensive domain and sequence homology searches were also undertaken to further characterize the proteases based on their orthologs.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

2.1.1. Chemicals, Enzymes and Kits

Agarose (low melting point gel), ammonium sulfate [$(NH_4)_2SO_4$], ampicillin, calcium chloride dihydrate ($CaCl_2$. $2H_2O$), chromogenic substrates: N-Suc-Alanine-Alanine-Phenilalanine-pNA (N-Suc-Ala-Ala-Phe-pNA), N-Suc-Alanine-Alanine-Proline-Phenilalanine-pNA (N-Suc-Ala-Ala-Pro-Phe-pNA), N-Carbobenzoxyl-Arginine-pNA (N-CBZ-Arg-pNA), ethidium bromide (Et-Br), glycerol, phenol, phytagel, sodium perclorate ($NaClO_4$), Triton X-100 and Tween 20 were purchased from Sigma Chemical Co., St. Louis, Missouri, USA.

Acetic acid, α-D-glucose, chloroform, disodium hydrogen phosphate (Na_2HPO_4), ethylenediaminetetraacetic acid (EDTA), isoamyl alcohol, magnesium chloride ($MgCl_2$), N,N´-methylene-bisacrylamide, potassium dihydrogen phosphate (KH_2PO_4), sodium acetate anhydrous, sodium chloride (NaCl), sodium dodecyl sulfate (SDS), sodium hydroxide (NaOH), sucrose, Tris-base, tryptone were from Merck; Darmstadt, Germany.

Agar was purchased from Acumedia, Baltimore, USA. Agarose was from Applichem, Darmstadt, Germany. Ethanol was from Reidel de Häen, N,N dimethyl formamide was from Fluka, Garantie. Yeast extract was from Difco,

Detroit, USA and Oxoid Basingstoke, England.

RNAase was purchased from Sigma Chemical Co., St. Louis, Missouri, USA. Taq DNA Polymerase, T4 DNA ligase and T4 DNA ligase buffer, restriction endonucleases *BamHI*, *EcoRI*, *HindIII*, *PstI*, *SalI*, and their buffers Tango™ Yellow, Buffer R, Buffer G were from MBI Fermentas AB, Vilnius, Lithuania.

QIAGEN PCR Cloning Kit was from QIAGEN Inc. Valencia, USA. Wizard® Plus SV Minipreps DNA Purification System Kit was purchased from Promega Corporation, Madison, WI, USA. DNA Extraction Kit was from BIO 101. QIAExpressionist Expression Kit was from QIAGEN Inc. Valencia, USA.

BSA and DAB substrate were from Roche Diagnostics GmbH, Mannheim, Germany.

All of the primers were synthesized by Applichem, Darmstadt, Germany.

2.1.2. Buffers and Solutions

Compositions of buffers and solutions used in the experiments are given in Appendix A.

2.1.3. Plasmid vectors, Molecular Size Markers and Genomic DNA Sequence Data

The maps of cloning vectors (pDrive and pQE30-32, QIAGEN) are illustrated in the Appendix B. DNA molecular size marker (Lambda DNA/*EcoRI*+HindIII Marker, Fermentas) is given in Appendix C. Nucleotide sequences of periplasmic

(ClpP-like) (PSP) and membrane-bound (MSP) serine proteases of *Tp. volcanium* are presented in Appendix D.

2.2. Strain and Medium

2.2.1 Archaeal and Bacterial Strains

Thermophilic archaea *Thermoplasma volcanium* GSS1 (strain type 4299) was used as the source organism in this study. The culture was purchased from DSM (Deutsche Summlung von Microorganismen und Zellkulturen, GmbH, Braunschweig, Germany).

Escherichia coli TGI strain from our laboratory collection was used as the recipient in transformations.

2.2.2. Growth and Maintenance of Microorganisms

Tp. volcanium was grown in Volcanium Medium (pH 2.7) (Robb, 1995) supplemented with glucose (Merck) and yeast extract (Oxoid) to final concentrations of 0,5 % and 0,1 %, respectively, at 60 °C. Cultures were subcultured with 30 days intervals onto fresh Darland's agar plates (Robb, 1995).

E. coli TGI cells were grown on LB agar plates at 37 °C. The cultures were subcultured with 30 days intervals**.**

Recombinant *E.coli* TGI strains were cultured on LB agar plates including ampicillin at 37 °C. The cultures were subcultured with 30 days intervals**.**

2.3. Gene Manipulation Methods

2.3.1. Genomic DNA Isolation from *Thermoplasma volcanium*

Genomic DNA of *Thermoplasma volcanium* was isolated by using the modified method of Sutherland, *et al,* (1990). 200 ml overnight culture was centrifuged (Sigma 3K30 Centrifuge, Germany) at 8.000 rpm for 10 min at +4ºC. After washing steps in saline-EDTA (pH:8.0) pellet was dissolved in lysosyme/saline EDTA solution (0.4 mg/ml lysosyme in saline-EDTA). After cell lysis, $NaClO_4$ was added to obtain a final concentration of 1 M. Then, one volume chloroform:isoamylalcohol (24:1, v/v) was added and the mixture was stirred in ice-water bath for 30 minutes until it was homogenous. After centrifugation at 12 000 rpm for 10 min at 4°C, the aqueous phase was transferred into a new tube and two volumes of ice cold absolute ethanol was added. The precipitated DNA, following incubation at -20°C for 1 hour, was collected by centrifugation at 11 000 rpm for 45 min., at 4°C. The supernatant was discarded and the pellet was washed with 70 % ethanol. After centrifugation the supernatant was poured and the remaining ethanol was removed by air drying. The DNA pellet was then dissolved in 1 ml of sterile ddH_2O, stored at - 20 °C, until use.

2.3.2. Purification of Chromosomal DNA of *Thermoplasma volcanium*

Isolated chromosomal DNA was purified as described by Sambrook, *et al,*. (2001). After removal of RNA by RNAase (Sigma Chemical Co., S. Louis, Missouri, USA) treatment, two successive (1/1, v/v) phenol:chloroform extractions were performed. The contaminating phenol was removed by ether extraction. The purified chromosomal DNA was precipitated with ethanol and washed with 70 % (v/v) ethanol. After centrifuge at 13 000 rpm (Biofuge 15 Centrifuge, Heraeus Sepatech, Germany) for 45 min at

+4°C, resulting DNA pellet was air dried and dissolved in minimal volume of sterile double distilled water (ddH_2O).

2.3.3. PCR Amplification of Gene Fragments of Periplasmic (ClpP-like) (PSP) and Membrane-bound (MSP) Serine Proteases

2.3.3.1. Design of PCR Amplification Primers

For amplification of PSP gene, forward and reverse amplification primers (FP1 and RP1) (Table 2.1) were designed so that 176 bp upstream and 102 bp downstream sequences to be included into amplification fragment. For amplification of MSP gene, forward and reverse amplification primers (FP2 and RP2) were so designed that 156 bp upstream, 45 bp downstream sequences to be included into amplification product. These gene fragments of PSP and MSP genes were expected to be amplified from *Tp. volcanium* genomic DNA as 971 bp and 1521 bp amplicons, respectively.

2.3.3.2 Amplification of PSP and MSP Genes

Tp. volcanium genomic DNA was used as a template for PCR amplifications. Each PCR reaction mixture with a total volume of 100µl included: lx PCR buffer [750 mM Tris-HCl pH 8.8 at 25°C, 200 mM $(NH4)_2SO_4$, 0.1 % Tween 20], 50-500 ng of the template DNA, 200 µM of each deoxyribonucleoside triphosphate (dNTP), 1.5 mM $MgCl_2$, 100 pmoles each of primers and 2.5 U of Taq DNA polymerase (MBI Fermentas AB, Vilnius, Lithuania). Total volumes of the mixtures were completed to 100 µl by using sterile ddH_2O.

Table 2.1 : Primers designed for periplasmic (ClpP-like) and MSP genes of *Thermoplasma volcanium*.

Primer	Sequence of the Primer
FP1 Forward Primer for Periplasmic (ClpP-like) Serine Protease Gene	**5' gtg atc cgt aca cag ctg ttc 3'**
RP1 Reverse Primer for Periplasmic (ClpP-like) Serine Protease Gene	**5' c gta gcc gat ata gag aaa tcg 3'**
FP2 Forward Primer for Membrane-Bound Serine Protease Gene	**5' gcc gca tcg ata tta ttt cac tg 3'**
RP2 Reverse Primer for Membrane-Bound Serine Protease Gene	**5' gt atg ttt aga tgc atg tct gac 3'**

The samples were pre-incubated at 95°C for 5 minutes, before the addition of Taq DNA polymerase. The reactions were carried out with 30 cycles of amplifications in a thermal cycler (Techgene, Techne Inc. NJ. USA). The programme for each cycle was as follows: Denaturation at 94°C for 1min, annealing at 55°C for 2 min, extension at 72°C for 3 min and final extension at 72°C for 10 min.

To optimize hybridization temperature for PCR amplification, all the parameters were kept constant and PCRs were performed at four different annealing temperatures (40-50-55 and 60°C).

PCR products (15 µl samples) were analyzed by agarose gel electrophoresis (1% agarose gel). Lambda DNA/*EcoRI+HindIII* Marker (Fermentas) was used for length calibration.

2.3.4 Agarose Gel Electrophoresis

PCR amplicons, DNA fragments produced by restriction digestions and DNA samples following any manipulation were analyzed by agarose gel elecrophoresis on 1% (w/v) agarose gel (Applichem, Darmstadt, Germany) in 1X TAE buffer. Submarine agarose gel apparatus (Mini Sub™ DNA Cell, Bio Rad, Richmond, CA, U.S.A) was used in these experiments. 1/10 vol of 4x tracking dye was added into each 10-20 µl DNA sample and this mixture was loaded to gel which was supplemented by ethidium bromide at final concentration of 1.25 µg/ml. Electrophoresis was carried out at 70 volts (Fotodyne Power Supply, Foto/Force™ 300, USA).

After electrophoresis, the bands were visualized with a UV transilluminator (Vilber Lourmat TFP-M/WL, Marne La Vallee Cedex 1, France) and gel photographs were taken using a gel imaging and documentation system (Vilber Lourmat Gel Imaging and Analysis System, Marne La Vallee Cedex 1, France). The molecular sizes of DNA fragments were determined by referring to calibration curve (Appendix F), which was obtained by plotting the log molecular weights of known size marker fragment against migration distance on the gel.

2.3.5 Cloning of PCR Amplified PSP and MSP Gene Fragments

PCR amplified PSP and MSP gene fragments were cloned using the QIAGEN PCR Cloning Kit (QIAGEN Inc., Valencia, USA) following the manufacturer's instructions. The PCR fragment was ligated to the pDrive Cloning Vector, that was provided with the Kit. The ligation reaction was carried out as described in the manufacturer’s manual.

2.3.5.1 Introduction of Recombinant Plasmids into Competent Cells

2.3.5.1.1 Preparation of Competent *E. coli* cells

Competent *E.coli* TGI cells were prepared according to the modified method of Chung, *et al,* (1989). The *E.coli* TGI cells were grown in 20 ml of LB medium through vigorous shaking (Heidolph Unimax 1010 Shaking Incubator, Heidolph Instruments GmbH, Kelheim, Germany). The growth was followed by measuring the optical absorbance of cell culture at 600 nm using the Hitachi U-2800 double beam spectrophotometer (Hitachi High Technologies Corporation, Tokyo, Japan). When the cells reached to early log phase (about 10^8 cell/ml), they were collected by centrifugation at 4000 rpm for 10 min. After discarding the supernatant, the pellet was dissolved in 1/10 TSS solution and distributed into eppendorf tubes as aliquots (100 ml). Competent cells were stored at - 80 °C until use (Thermo Forma -86ºC ULT Freezer, Thermo Electron Corp., USA).

2.3.6. Transformation

Competent *E. coli* TGI cells were taken from -80°C deep freezer and thawed on ice to be used in transformation. 1/10 (v/v) volume of ligation mixture was added and mixed by gentle tabbing. After incubation on ice for 30 min, cells were transferred to glass tubes containing 0.16 mM glucose in LB. The cells were incubated at 37° for 1 hr with vigorous shaking at 225 rpm (Heidolph Unimax 1010 Shaking Incubator, Heidolph Instruments GmbH, Kelheim, Germany). Then, appropriate dilutions of competent cells were spread onto selective LB agar plates containing Ampicillin, IPTG and X-Gal and grown at 37°C for 24 hours.

2.3.7 Isolation of Plasmid DNA from Recombinant Colonies

Putative recombinant colonies were collected depending on blue/white colony selection. Cells from the white colonies were inoculated into 10 ml LB culture containing ampicillin and incubated at 37 °C overnight. Plasmid isolation was achieved by using Wizard® Plus SV Minipreps DNA Purification System (Promega Corporation, Madison, WI, USA) according to the procedure provided by the manufacturer. The bacterial cells were derived by centrifugation at 4000 rpm for 15 min (IEC Clinical Centrifuge, Damon/IEC Division, U.S.A). The pellet was thoroughly re-suspended in 250 µl Cell Resuspension Solution. The cells were lysed by adding 250 µl Cell Lysis Solution and the tubes were inverted for 4-5 times to mix the ingredients. After 4 min of incubation at room temperature, proteins were digested with 10 µl Alkaline Protease Solution. Tubes were inverted for 4-5 times to mix the ingredients and incubated for 5 min at room temperature. Then, neutralization solution (350 µl) was added; tubes were inverted for 6-8 times to mix the ingredients and incubated on ice for 1 min. The cell debris was pelleted by centrifugation at 13 000 rpm for 10 minutes in microfuge (Biofuge 15 Centrifuge, Heraeus Sepatech, Germany). The supernatant was transferred into Spin Column which was provided with kit and centrifuged at 13 000 rpm for 1 min to bind the DNA to the column. The bound DNA was washed twice with the Washing Solution with the volumes of 750 µl and 250 µl, respectively and eluted with 100 µl Nuclease Free Water. After centrifugation at 13 000 rpm for 1 min, 100 µl of plasmid sample, at a concentration of 1.5 µg/ml, was obtained from the starting material. Approximately 15 µl of the sample was run on the agarose gel to check the efficiency of purification and the remaining sample was stored at -20°C for further use.

2.3.8 Restriction Enzyme Digestion Analysis of Putative Recombinant Plasmids

Plasmid DNAs isolated from the transformed colonies were cut in several single and double digestion reactions by using *EcoRI*, *BamHI*, *HindIII*, *SalI* and *PstI* restriction endonucleases (MBI Fermentas Co, Lithuania). Digestions were carried out in reaction mixtures with total volumes of 20 µl including sample plasmid DNA, restriction endonuclease and convenient buffer following the instructions of manufacturer.

2.4 Biochemical Methods

2.4.1 Preparation of Culture Supernatants

Recombinant *E.coli* TGI cells were grown in LB broth including ampicillin. Non-recombinant cells, which were used as control groups in enzyme assays, were cultured in LB broth without antibiotics. Cells were removed by centrifugation at 8000 rpm for 15 min by using Sigma Centrifuge (Sigma 3K30 Centrifuge, Germany) at +4°C. The supernatants were stored at -20 °C until they were used in the extracellular serine protease activity assays.

2.4.2 Preparation of Cell Free Extracts

Overnight cultures (250 ml) of recombinant and non-recombinant *E.coli* TGI cells were incubated at 37ºC by vigorous shaking at 160 rpm (Heidolph Unimax 1010, Germany) and then, cells were harvested by centrifugation at 8000 rpm (Sigma 3K30 Centrifuge, Germany) for 15 minute at +4°C. The cells were washed in serine protease assay buffer. The cell pellet was then resuspended in 1/5 volume of serine protease assay buffer. The cells were disrupted by

sonication (Sonicator VC 100, Sonics and Materials, CT, U.S.A) for 1 minute with 30 second intervals, until the solution became transparent. The cell free extract (CFE) was obtained by centrifugation (Sigma 3K30 Centrifuge, Germany) at 12 000 rpm for 2 hours at +4°C. A 1.5 ml of each CFE was heat treated at 60ºC for 15 min in waterbath and then centrifuged (Micromax 230 RF Centrifuge, Thermo IEC, USA) at 12 000 rpm for 1,5 h min at +4 ºC. CFEs were stored at -20°C until use.

2.4.3 Determination of Extracellular and Intracellular Serine Protease Activities by Using Chromogenic Substrates

The activities of recombinant PSP and MSP enzymes were determined by measuring their abilities to hydrolyze synthetic peptide substrates which are N-terrninally blocked by succinyl (Suc) and that have aromatic amines such as para-nitroanilide (pNA) at their C-termini. These activities were measured based on the p-nitroanilide liberation from N-Succinyl-Alanine-Alanine-Phenylalanine-para-Nitroanilide (N-Suc-Ala-Ala-Phe-pNA) and N-Succinyl-Alanine-Alanine-Proline-Phenylalanine-para-Nitroanilide (N-Suc-Ala-Ala-Pro-Phe-pNA). The continuous spectrophotometric assays were carried out in a total volume of 800µl which includes serine protease assay buffer (pH: 7.0), susbstrate and CFE or supernatants derived from recombinant and non-recombinant (for control) cultures, at 55 ºC. After pre-incubation of the CFE or supernatants at 55ºC for 5 min, reactions were started by adding substrate solution to give final concentration of 0,625 mM. The absorbance was measured at 410 nm using UV/visible spectrophotometer (Shimadzu l601 UV/Visible Spectrophotometer, Shimadzu Analytical Co., Kyoto, Japan).

2.4.4 Heat Shock Response of *Tp. volcanium* Cells

Tp. volcanium was grown on Volcanium Medium (pH: 2.7) supplemented with 0.5 % glucose and 0.1 % yeast extract as carbon and energy sources. The cells were grown until mid-exponential phase (OD_{640}=194) at 60°C and then, in two different experiments, the temperature was shifted to 72°C for 2 h and 80°C for 2 h to perform heat-shock. Control cultures were allowed to continue growing at 60°C. The cell growth was monitored through measurements of OD at 640 nm, in a time dependent manner.

Control and heat-shocked culture samples were removed at regular time intervals and supernatants and cell free extracts were prepared to be used in enzyme assays. Enzymatic activities were measured based on the p-nitroanilide liberation from N-Succinyl-Ala-Ala-Pro-Phe-pNA, N-Succinyl-Ala-Ala-Phe-pNA and N-CBZ-Arg as described in 2.4.3 to find out possible effect of heat shock on proteolytic activities of *Tp. volcanium*.

2.5 Heterologous Expression of PSP Gene of *Thermoplasma volcanium* in *Escherichia coli*

Expression of PSP gene was carried out by using QIAExpress® Type IV Kit (QIAGEN Inc., Valencia, USA), following the instructions of the manufacturer. Expression product comprised vectors (pQE30-32) that could be used for expression of 6xHis tagged recombinant colonies. Expression of the gene required the construction of expression clones, as described below.

PSP gene fragments were excised from the recombinant plasmids isolated from recombinant *E. coli* TG-S1 colonies by double-digestion with *SalI* and *PstI* restriction endonucleases (MBI Fermentas, Lithuania). Also, the expression vectors pQE30 and pQE32 were digested with the same enzymes. Gene and

vector fragments were run on 1% low melting agarose gel (Sigma-Ag414-105, Sigma Chemical Co., St. Louis, Missouri, USA) by electrophoresis (Mini Sub™ DNA Cell, Bio Rad, Richmond, CA, U.S.A) in 1X TAE (Tris-acetate-EDTA) running buffer. Low melting agarose gel was supplemented with ethidium bromide at a final concentration of 1,25 µg/ml before it solidified. Serine protease gene fragments were mixed with 1/10 volume of 4X tracking dye before loading the completely solidified gel. Electrophoresis was carried out at 50 volts (Fotodyne Power Supply, Foto/Force™ 300, U.S.A).

2.5.1 Isolation of DNA Fragments from Agarose Gel

BIO 101 GeneClean® Kit (QBioGene, Carlsbad, CA, USA) was used for the isolation of DNA fragments from the gel. Isolation was carried out according to the recommendations of the manufacturer. The gel slices containing the fragments were excised, weighed and solubilized in NaI with a value of x3 weight of gel slices. After incubation at 55°C for 5 min, glass milk was added into each mix (10µl glass milk for the volumes between 500-1000 µl). Mixtures were incubated for 10 min at room temperature by inverting the tubes in every minute to ensure the complete gel dissolution. Desired DNA-silica resins were obtained by centrifugation at 13 000 rpm for 45 secs (Eppendorf Mini Spin Centrifuge, Hamburg, Germany). Pellets were washed two times with BIO101 wash solution and then, the silica resin was air dried. Pellets were dissolved in minimal volumes sddH_2O and incubated at 55°C for 5 min**.** After centrifugation at 13 000 rpm for 60 secs at room temperature, supernatants were transferred into sterile eppendorf tubes and kept at -20°C, until use.

2.5.2. Ligation of the PSP Gene and pQE Vector Fragments

The gene encoding the PSP was cloned into the pQE30 and pQE32 vectors in the same reading frame as the 6xHis affinity tag. To be able to provide convenient binding, the gene fragment was excised from pDrive vector, where the fragment was located in 3´-5´ direction, with respect to *lacZ* promoter by using *SalI* and PstI restriction endonucleases. The ligation of the insert to the expression vectors cut with the same restriction enzymes was carried out using T4 DNA ligase under standard conditions (Sambrook, *et al*, 2001). Ligation mixtures with a total volume of 20µl included variable volumes (10-14 µl, depending on fragment concentration) of gene fragments and pQE vectors (1-2 µl, depending on vector fragment concentration), 1/10 (v/v) of T4 DNA ligase (5 weiss unit/ µl) (MBI Fermentas, Lithuania), 1/10 (v/v) of 10X T4 DNA ligase buffer (MBI Fermentas, Lithuania) and variable amounts of sterile double distilled water in order to complete the final volume to 20 µl. Ligation mixtures were incubated at +4ºC for 22 h.

2.5.3. Transformation of the *E.coli* TGI Cells with Expression Constructs

E. coli TGI competent cells were used as hosts in the transformation. They were taken from -80°C deep freezer and thawed on ice, and mixed with a 1/5 (v/v) volume of ligation mixture and incubated for 30 min on ice. Samples were transferred into tubes containing 0.16 mM glucose in LB broth and incubated at 37°C at 230 rpm for 1h 30 by vigorous shaking (Heidolph Unimax 1010 Shaking Incubator, Heidolph Instruments GmbH, Kelheim, Germany). A series of Dilutions (10^{-1}, 10^{-2} and 10^{-3} fold) were prepared and 0.1 ml from each dilution was spread onto LB agar containing 50µg ampicillin. Plates were incubated for 16 h at 30ºC.

2.5.4 Colony Blotting and Immunoscreening

Transformants were screened for correct insertion in frame of the coding fragment by colony-blotting procedure, since it also allowed simultaneous screening of expression levels and in-frame translation of the 6xHis tag.

Detection of the expressed proteins, so the recombinant colonies, was carried out by immobilization of the expressed protein to membranes directly from bacterial colonies. Colonies were transferred to nitrocellulose membrane that was laid on the surface of the agar plate. Protein expression was induced, and after incubation, the cells were lysed with alkali. The proteins that are released bound to the membrane in the position of the colony, since the colony blotting procedure distinguished the clones that express 6xHis tagged proteins.

Freshly transformed cells were grown on LB agar plates containing ampicillin. Nitrocellulose membranes (Millipore HATF08250, Billerica, MA, USA) were placed on the agar surfaces in contact with the colonies. Membranes were transferred to fresh LB agar plates containing 20μg ampicillin and 250μM IPTG and incubated at 37ºC for 4 h. Membranes were transferred onto 3MM paper (Whatmann 3MM Chromatography paper, Maldstone, England) sheets soaked in a number of solutions and incubated for several minutes on each sheet (Incubation with SDS solution for 10 min, incubation with denaturing solution for 5 min, incubation with neutralization solution for 5 min, incubation with neutralization solution for 5 min, incubation with 2X SSC buffer for 15 min) for binding of protein to the membranes. Then, the membrane blots were washed for 10 min in TBS buffer and then incubated for 1 h in blocking buffer (3% (w/v) BSA in TBS) at room temperature. They were washed in TBS-Tween/Triton buffer and again in TBS buffer, for 10 min. Following the incubation in Anti-His HRP conjugate solution (1/1000 dilution of conjugate solution in blocking buffer, washing steps in TBS-Tween/Triton and TBS

buffers were repeated. Staining was performed by using DAB as chromogenic substrate for the HRP reaction. Color reaction was stopped by rinsing the membranes in double distilled water. The formation of pink color (in the form of spots) on the membrane blots might indicate the expression of the 6xHis tagged fusion protein. Colony blotting and immunohybridization procedures are summarized in Figure 2.1.

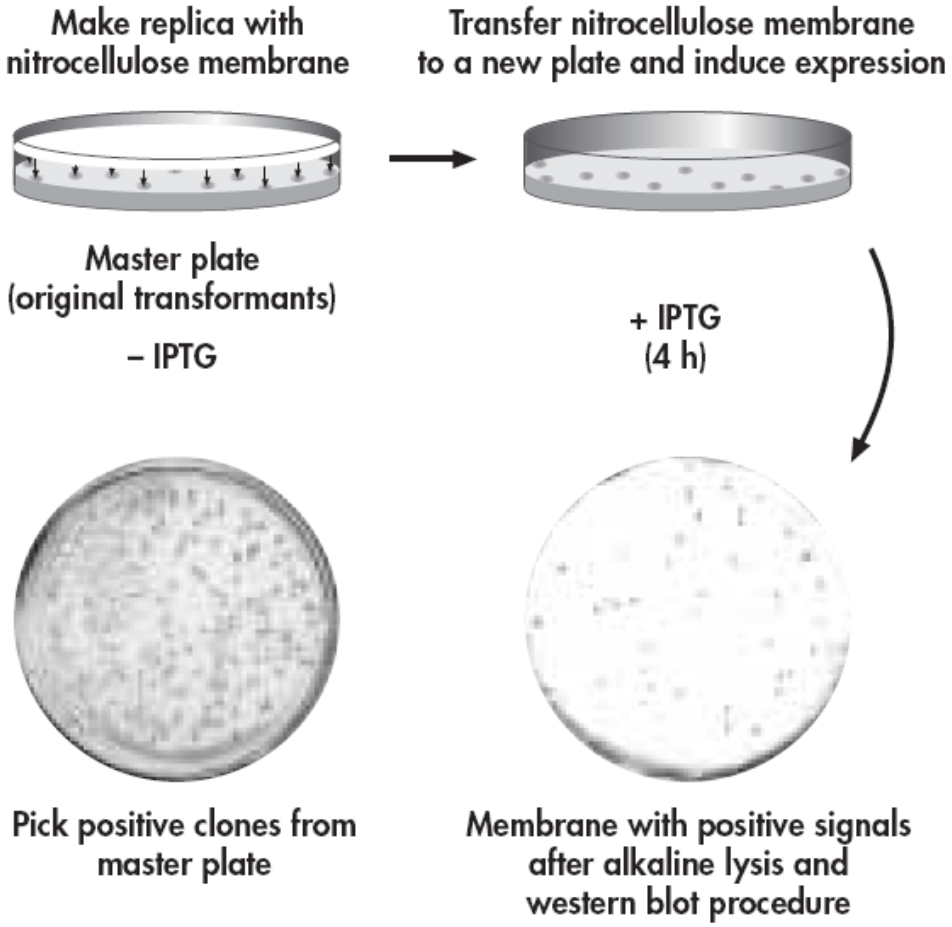

Figure 2.1: Detection of positive expression clones by colony blotting (QIAExpressionist Kit Handbook, Fifth Edition, QIAGEN)

2.5.5 Characterization of Expression Constructs

Transformants were screened for presence of the coding fragment by restriction analysis of the recombinant pQE plasmid DNA using restriction enzymes *EcoRI, BamHI*, *HindIII*, *SalI* and *PstI*. Recombinant pQE plasmids were isolated and restriction enzyme digestions were carried out as described in section 2.3.7 and 2.3.8., respectively.

2.5.6 SDS-Polyacrylamide Gel Electrophoresis

Polyacrylamide slab gel electrophoresis in the presence of 0.5 % anionic detergent SDS was performed on the 5 % stacking gel and 12.5 % separating gel according to Laemmli (1970). SDS Molecular Weight Marker MW-ND 500 kit (Sigma Chem. Corp., St. Louis, MO, USA) was used as molecular weight standard.

Vertical slab gel electrophoresis was carried out using the Biometra Vertical Minigel System (Biometra Biomedzinische Analytic Gmbh, Göttingen, Germany). The assembly of the glass plate cassettes and the process of the gel casting were done according to the instruction manual provided with the apparatus.

10µl aliquots of enzyme samples were mixed with 2X sample buffer containing 5% SDS and incubated in boiling water bath for 5 minutes to denaturate the proteins. Afterwards, protein sample and molecular weight standart were loaded onto gel. Running process was carried out at 100 V at +4ºC. After the electrophoresis, polyacrylamide gel was stained in Coomassie Brillant Blue R250 and then, destained by placing in the destaining solution that was changed several times.

2.6 Gene and Protein Structure Analyses

Total genome of *Thermoplasma volcanium* was derived from NCBI (National Center for Biotechnology Information, US).

Nucleotide and protein sequences of PSP and MSPs were derived from the online-database of KEGG.

Restriction Mapper Version 3 was used to map the sites for restriction endonucleases in PSP and MSP nucleotide sequences.

To search for the highly conserved motifs in the proteins, COG (Clusters of orthologous group) and CDD (Conserved Domain Database) analyses were carried out by using the online-database of NCBI.

Multiple alignments of protein sequences across their entire lengths with Archaeal, Gram (+) and Gram (-) bacterial serine proteases were achieved by using ClustalW 1.82 program (European Bioinformatics Institute server). Also, phylograms based on ClustalW (1.82) multiple sequence alignment of PSP and MSPs were derived to estimate their phylogenies.

Signal peptide prediction results derived by SignalP 3.0 server due to Neural Networks and Hidden Markov Models.

CHAPTER 3

RESULTS

3.1. Gene Manipulation

3.1.1 Isolation of Genomic DNA from *Thermoplasma volcanium*

Genomic DNA of *Thermoplasma volcanium* was isolated and purified by RNAase application as described in Materials and Methods, section 2.3.2. Isolated and purified DNA samples were visualized by agarose gel electrophoresis. The convenience of the purified DNA samples for gene manipulation experiments was controlled by restriction enzyme digestion with *EcoRI* endonuclease (Figure 3.1).

In each purification, isolated DNA samples were analyzed spectrophotometrically (Sambrook, *et al*, 2001). The absorbance values at 260 nm and 280 nm were obtained to calculate the quality and purity of DNA samples. The concentrations of the samples varied over the range of 1.0-1.6 mg/ml.

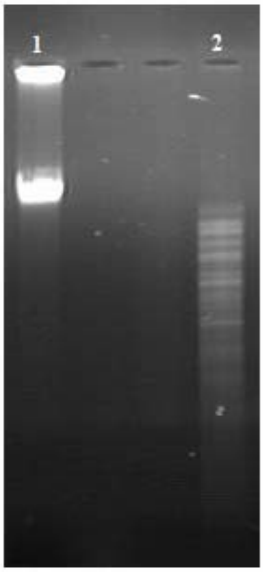

Figure 3.1: *Tp. volcanium* genomic DNA isolation and digestion. Lane 1: Isolated and purified chromosomal DNA of *Tp. volcanium*, Lane 2: Chromosomal DNA after EcoRI digestion.

3.1.2 PCR Amplifications of Periplasmic (ClpP-like) and Membrane-bound Serine Protease Gene Fragments

PCRs were carried out by using the FP1/RP1 and FP2/RP2 primer sets which were given in Table 2.1 to amplify the gene fragments of PSP and MSPs from genomic DNA of *Tp. volcanium*. PSP and MSP genes were amplified as 971 bp and 1521 bp amplicons, respectively, as visualized by agarose gel electrophoresis.

When PCRs were performed at 40ºC, some nonspecific amplification products were observed in addition to the target fragment i.e., PSP gene in Figure 3.2.

In order to get rid of the nonspecific hybridizations and to obtain specific amplification products, annealing temperature was increased gradually to 50ºC,

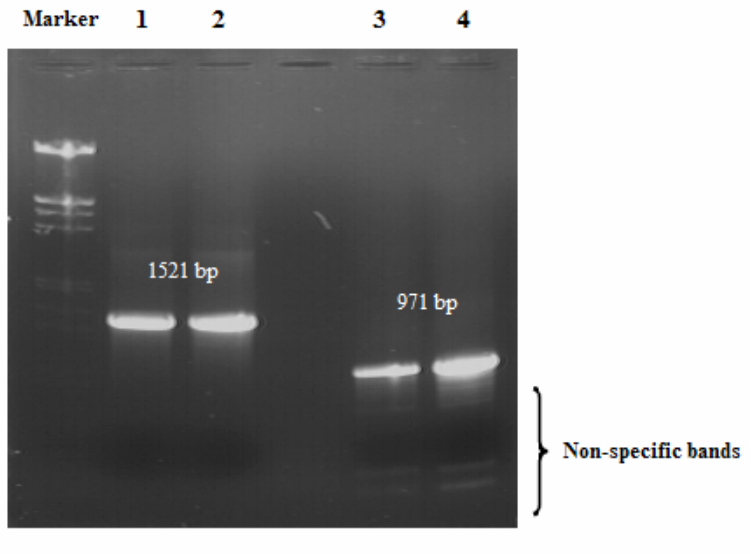

Figure 3.2 : PCR amplifications of MSP (Lane 1 and 2) and PSP (Lane 3 and 4) gene fragments at annealing temperature 40ºC. Marker : *EcoRI/HindIII* cut Lambda DNA (MBI Fermentas, Lithuania)

55ºC and 60ºC (Figure 3.3). At the annealing temperatures $\geq$ 50ºC, single PCR bands were observed for both periplasmic (ClpP-like) and MSP gene fragments with the estimated lengths of approximately 971 bp and 1531 bp, respectively.

Although the amplifications at 50ºC, 55ºC and 60ºC produced the specific bands, concentrations of the products were variable. At 50ºC, the product concentrations were almost the same for both gene fragments, but at 60ºC concentration of PSP gene fragment was higher than the concentration of MSP gene fragment. When the purity and concentrations of the amplicons were considered , we realized that amplification at an annealing temperature of 55ºC was the most efficient one (Figure3.4).

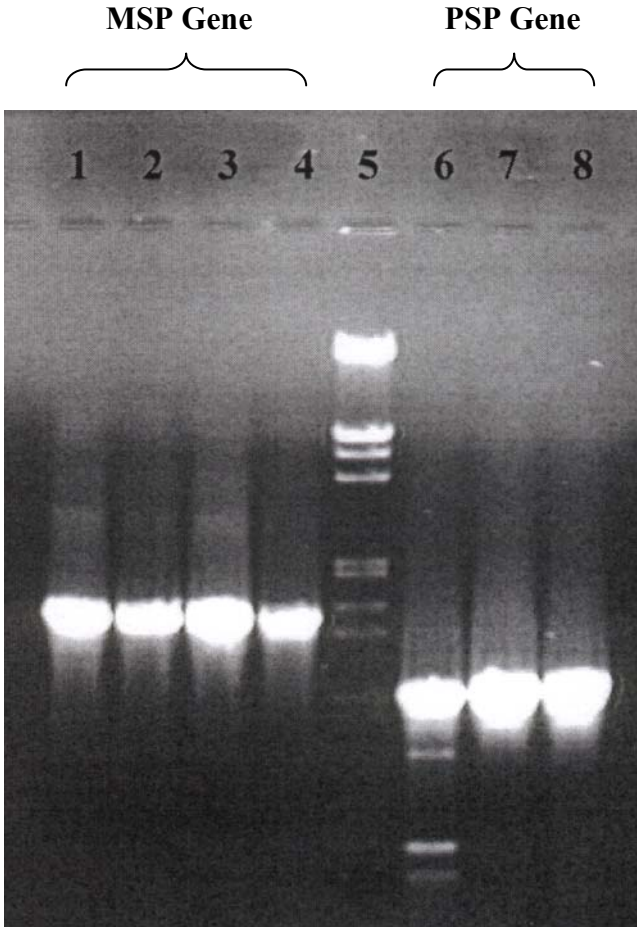

Figure 3.3 : PCR amplifications of MSP and PSP genes at different annealing temperatures: Lane 1 at 40ºC, Lane 2 at 50ºC, Lane 3 at 55ºC and Lane 4 at 60ºC amplifications of MPS gene. Lane 5 represents the Lambda DNA *EcoRI/HindIII* cut Marker (MBI Fermentas, Lithuania). Lane 6 at 40ºC, Lane 7 at 55ºC and Lane 8 at 60ºC amplifications of PSP gene.

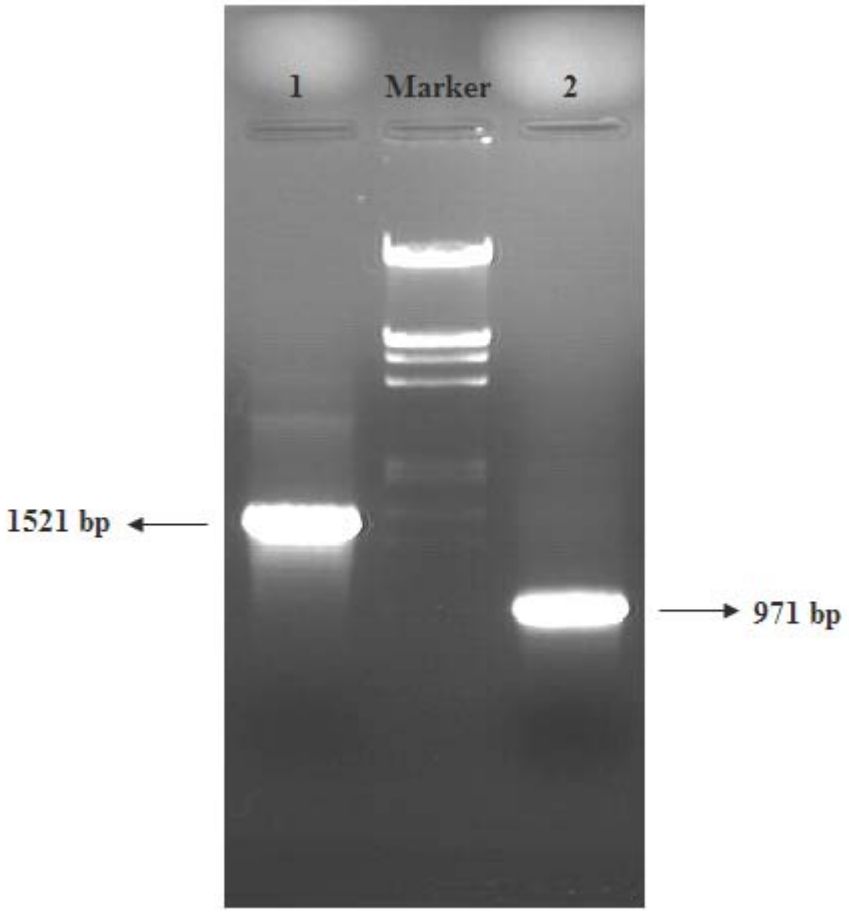

Figure 3.4 : PCR amplification products of MSP (Lane1) and PSP (Lane2) gene fragments produced at an annealing temperature of 55ºC. Marker: *EcoRI/HindIII* cut Lambda DNA (MBI Fermentas, Lithuania).

3.1.3. Cloning of PCR Amplified Gene Fragments

PCR amplicons obtained at 55ºC and 60ºC annealing temperatures by using *Taq* DNA polymerase were used for cloning of PSP and MSP genes, respectively. These amplicons were ligated into pDrive cloning vector which is supplied in a linear form with the QIAGEN PCR Cloning Kit (QIAGEN Inc., Valencia, USA). This kit takes the advantage of the single A overhang at each end of PCR products generated using *Taq* and other non-proofreading DNA polymerases. The pDrive cloning vector, supplied in a linear form with a U overhang at each end, hybridizes with high specificity to such PCR products with an A overhang. The pDrive cloning vector provides not only the high performance through UA-based ligation, but allows blue/white colony screening as well. The vector also carries several unique restriction endonuclease recognition sites around the cloning site, allowing easy restriction analysis of recombinant plasmids. Furthermore, it contains a T7 and SP6 promoter on either side of the cloning site which allows *in-vitro* transcription and facilitates the analyses of cloned PCR products.

The cloning schemes for the PCR amplified PSP and MSP gene fragments using pDrive cloning vector are shown in Figure 3.5 and Figure 3.6, respectively.

E. coli TG1 competent cells which were capable of efficient transformation (Transformation frequency: $8x10^6$ transformant/µg DNA) were used as recipients in transformation experiments. These cells were plated immediately onto agar plates, including IPTG, X-Gal and Ampicillin, following transformation. Recombinant colonies were selected depending on blue/white colony screening.

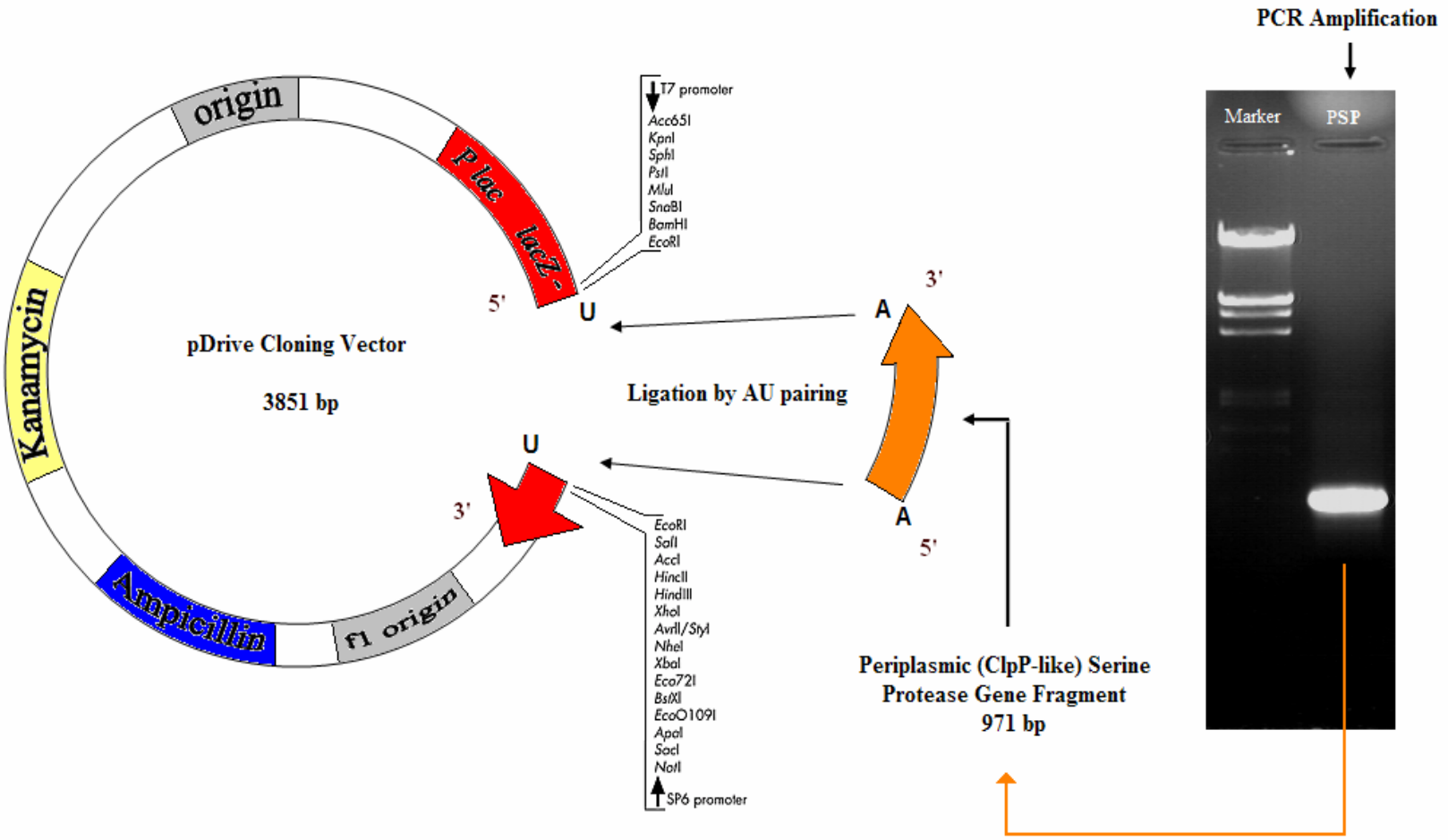

Figure 3.5 : Schematic representation of cloning of PCR amplified 971 bp PSP gene fragment by using pDrive cloning vector (QIAGEN Inc. Valencia, USA). Marker : Lamda DNA *EcoRI/HindIII* Cut Molecular Weight Marker (MBI Fermentas, AB, Vilnuis)

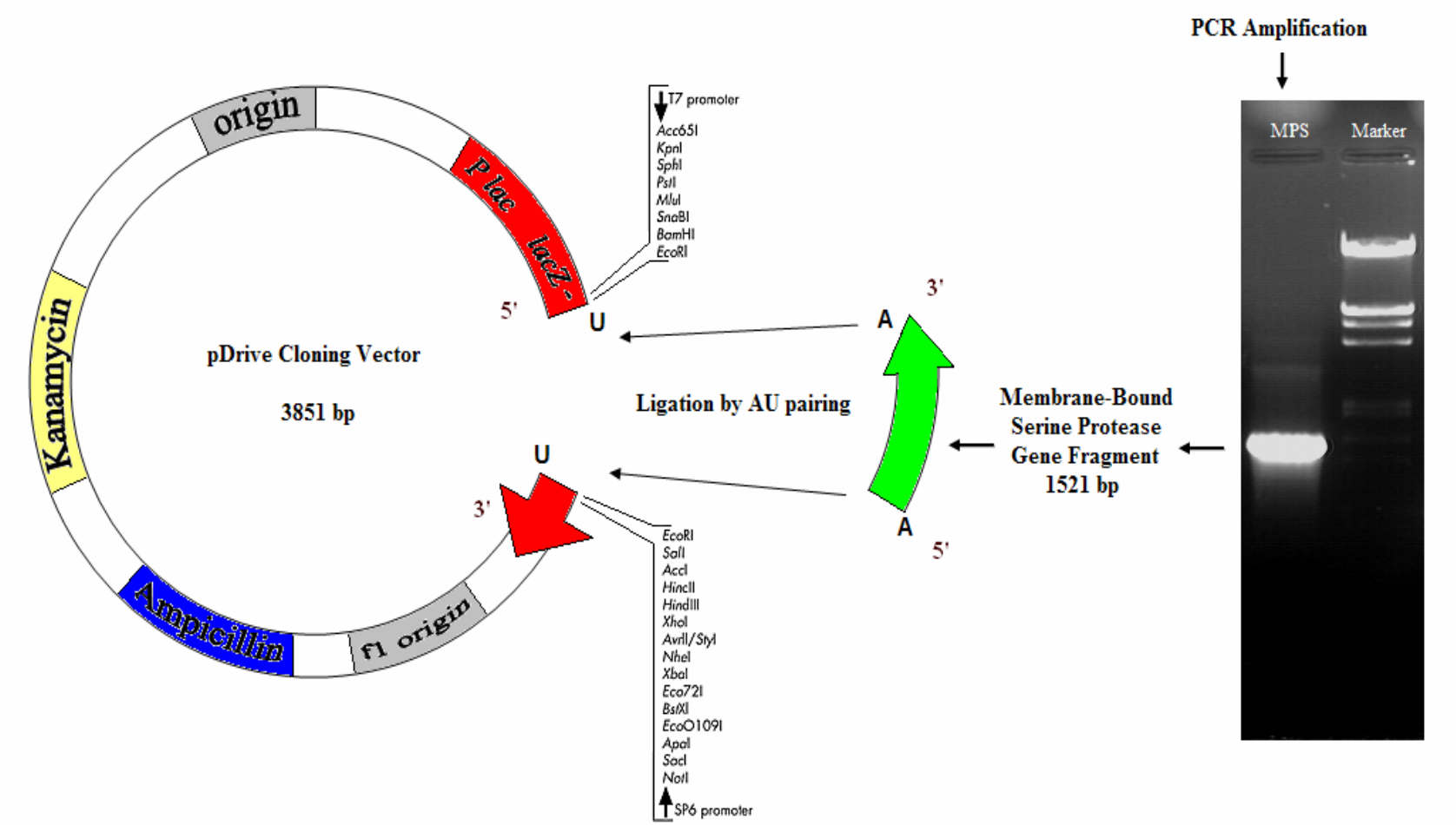

Figure 3.6 : Schematic representation of cloning of PCR amplified 1521 bp MSP gene fragment by using pDrive cloning vector (QIAGEN Inc. Valencia, USA). Marker : Lamda DNA *EcoRI/HindIII* Cut Molecular Weight Marker (MBI Fermentas, AB, Vilnuis)

Transformation efficiency was $3x10^4$ transformant/µg DNA and recombinant frequency was $3x10^6$ recombinant/ng DNA for PSP.

Transformation efficiency was $7x10^4$ transformant/µg DNA and recombinant frequency was $7x10^6$ recombinant/ng DNA for MSP.

A total of 12 and 34 putative recombinant (white-colored) colonies were obtained from PSP and MSP cloning experiments, respectively. Plasmid DNAs were isolated from 5 clones from each group (Figure 3.7).

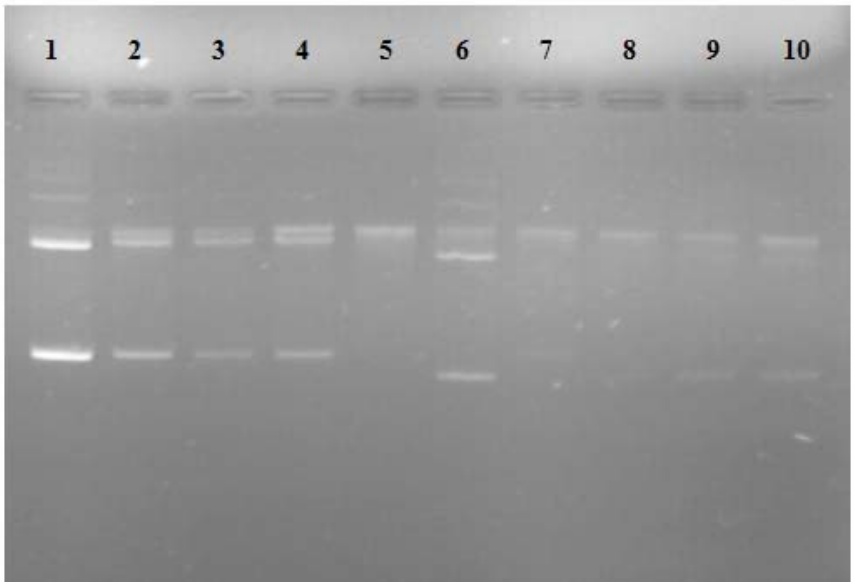

Figure 3.7 : Plasmids isolated from putative recombinant colonies. Lane 1, colony M1; Lane 2, colony M2, Lane 3, colony M3; Lane 4, colony M4; Lane 5, colony M5 that were expected to be recombinants that include MSP gene. Lane 6, colony S1; Lane 7, colony S2, Lane 8, colony S3, Lane 9, colony S4 and Lane 10, colony S5 that were expected to be recombinants with cloned PSP gene.

3.1.4. Characterization of Cloned DNA Fragments

3.1.4.1. Characterisation of Putative Recombinant Plasmids with PSP and MSP Gene Fragments

Digestions with *EcoRI* restriction endonuclease were performed to confirm the presence of PSP and MSP inserts in putative recombinant plasmids which were isolated from white colonies. Restriction maps of our genes of interest were constructed, based on their sequences which had already been released to genomic databases, by using the software analysis system as mentioned in Materials and Methods, section 2.6 (Table 3.1). There are two *EcoRI* restriction sites in the MCS of the vector flanking the insertion site which facilitated the intact excision of the insert. Since there is no other restriction site for *EcoRI* in the PSP insert, the putative recombinant plasmids which were digested with *EcoRI* restriction enzyme yielded two fragments of 3851 bp and 971 bp in length corresponding to that of linear pDrive cloning vector and insert, respectively. There is no other restriction site for *EcoRI* in the MSP insert, either. So, digestion with *EcoRI* restriction enzyme yielded two fragments with the lenghts of 3851 bp and 1521 bp corresponding to that of linear pDrive cloning vector and MSP insert, respectively.

Three (S1, S4 and S5) of the five plasmids analysed by *EcoRI* digestion were recognized as true recombinants including the PSP insert. Among these three recombinant plasmids , S1 was named as pDrive-S1 and selected to be used for further caharacterisation by restriction analyses (Figure 3.8).

Table 3.1 : Cut positions of several restriction enzymes that were determined by Restriction Mapper software analysis system.

RESTRICTION MAPS				
PSP Gene Fragment				
Non-Cutter Restriction Enzymes	*EcoRI, EcoRV, KpnI, PstI, SacI, SalI, SmaI, SphI*			
Cutter Restriction Enzymes	**Name**	**Sequence**	**Overhang**	**Cut Position**
	BamHI	GGATCC	Five_prime	466
	HindIII	AAGCTT	Five_prime	94
	PvuII	CAGCTG	Five_prime	16
	SspI	AATATT	Five_prime	232
	ApoI	RAATTY	Five_prime	823
	SmlI	CTYRAG	Five_prime	811
	TaqI	TCGA	Five_prime	484, 950
MSP Gene Fragment				
Non-Cutter Restriction Enzymes	*EcoRI, KpnI, PstI, SacI, SalI, SmaI, SphI*			
Cutter Restriction Enzymes	**Name**	**Sequence**	**Overhang**	**Cut Position**
	BamHI	GGATCC	Five_prime	492
	HindIII	AAGCTT	Five_prime	830
	PvuII	CAGCTG	Five_prime	349, 661
	SspI	AATATT	Five_prime	860
	EcoRV	GATATC	Five_prime	974
	SmlI	CTYRAG	Five_prime	583,825
	XhoII	RGATCY	Five_prime	490, 1458

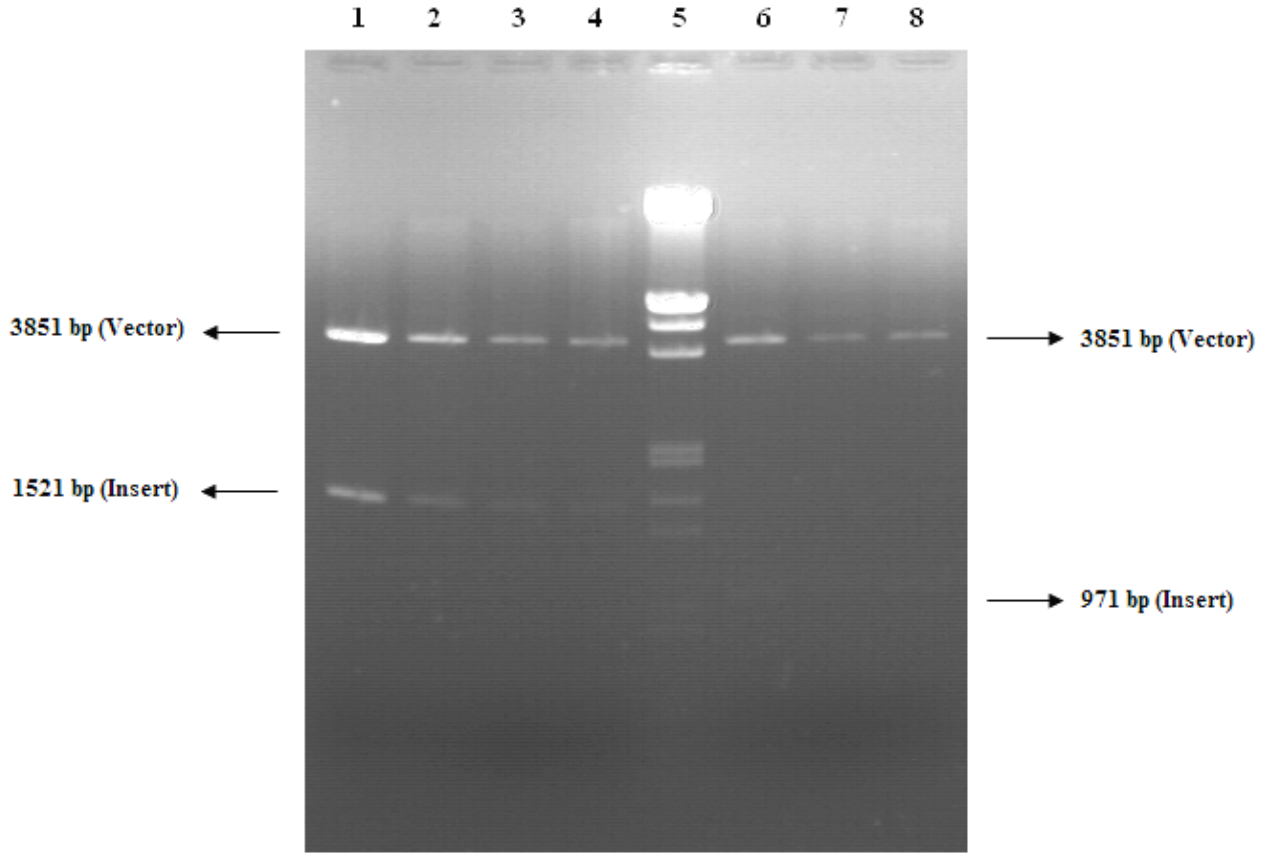

Figure 3.8 : Restriction enzyme analyses of putative recombinant plasmids isolated from white colonies with *EcoRI* digestion. Lane 1 (M1), Lane 2 (M2), Lane 3 (M3) and Lane 4 (M4) include MSP gene fragment and Lane 6 (S1), Lane 7 (S4) and Lane 8 (S5) include PSP gene fragment. Lane 1 is the *EcoRI/HindIII* cut Lamda DNA molecular weight marker (MBI Fermentas, AB, Vilnius).

Four (M1, M2, M3 and M4) of the five plasmids analysed by *EcoRI* digestion were found as true recombinants possessing the MSP insert. Among these four recombinant plasmids , M1 was named as pDrive-M1 and selected to be used for further caharacterisation by restriction analyses (Figure 3.8).

3.1.4.2. Further Characterization of Recombinant Plasmids Containing PSP Gene

Plasmids isolated from recombinant pDrive-S1 cells (Figure 3.10.A) were digested with *EcoRI*, *BamHI* and *HindIII* endonucleases for further characterization and restriction mapping. *EcoRI* digestions (Figure 3.9, Figure3.13) were carried out again , in order to confirm the results from initial screening of recombinant plasmids isolated from putative recombinant colonies, which was mentioned in section 3.2.4.1.

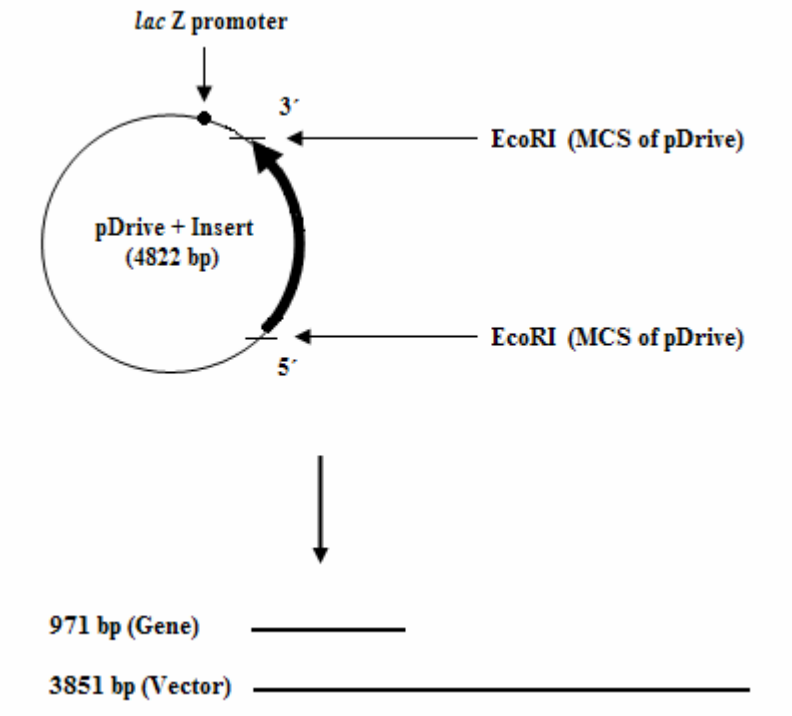

Figure 3.9 : *EcoRI* digestion profile of recombinant plasmid construct with PSP gene.

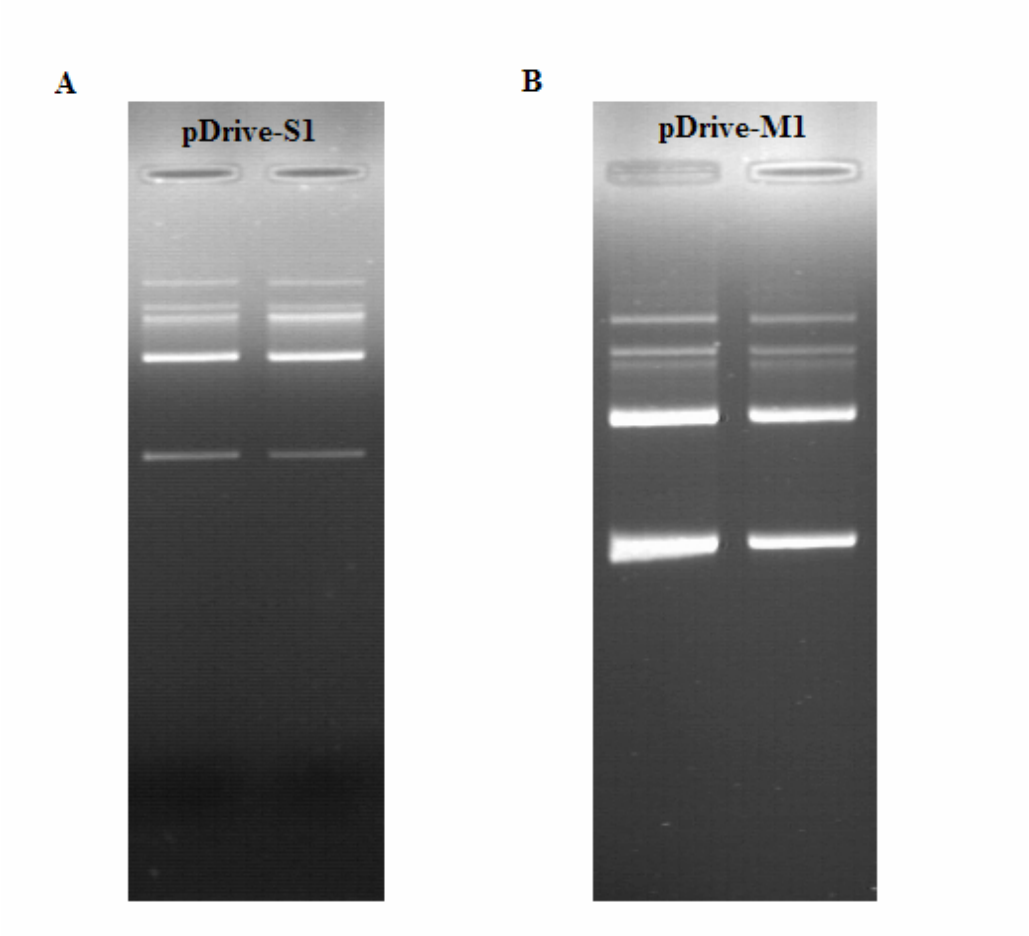

Figure 3.10 : Isolated recombinant plasmids. **A)** Recombinant plasmids (pDrive-S1) that include PSP gene fragment. **B)** Recombinant plasmids (pDrive-M1) that carry MSP gene fragment.

Recombinant plasmid has two restriction sites for *BamHI* endonuclease. One of them is within the MCS of the vector and the other one is located in the gene fragment at position 492 bp. Therefore, *BamHI* single digestion yielded two digestion products with the lenghts of 4317 bp and 405 bp (Figure 3.11, Figure 3.13).

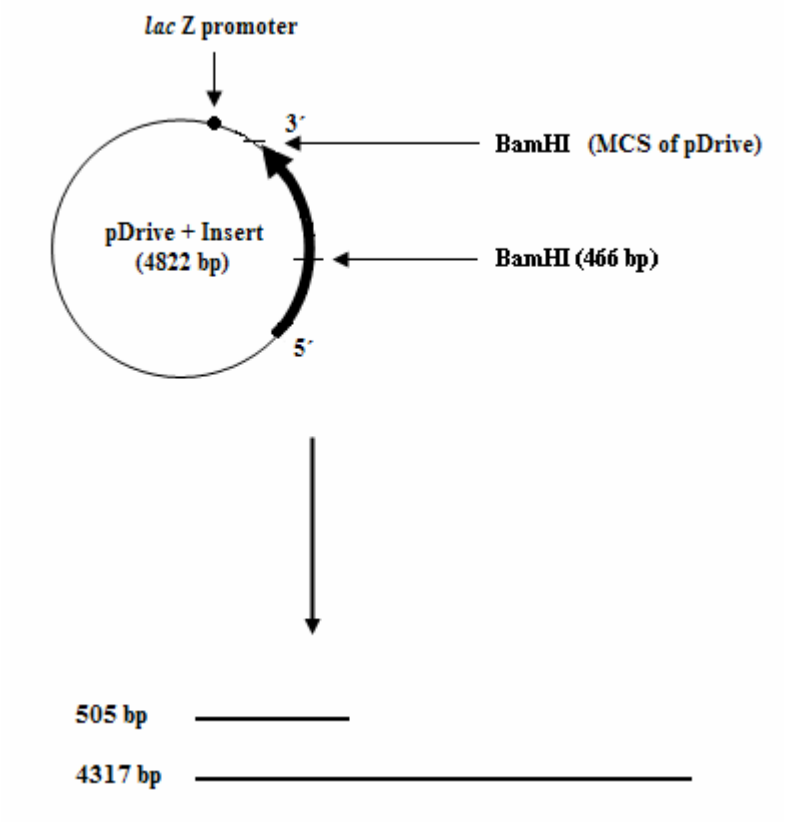

Figure 3.11 : *BamHI* digestion profile of the recombinant plasmid construct with PSP gene.

There are two restriciton sites for *HindIII* endonuclease in recombinant vectors pDrive-S1. This enzyme cuts the vector at 333rd bp of MCS and cuts the gene fragment at 94th bp. Therefore, single digestion of the recombinant plasmids gave rise to two digestion products with the lengths of 4728 bp and 94 bp (Figure 3.12, Figure 3.13).

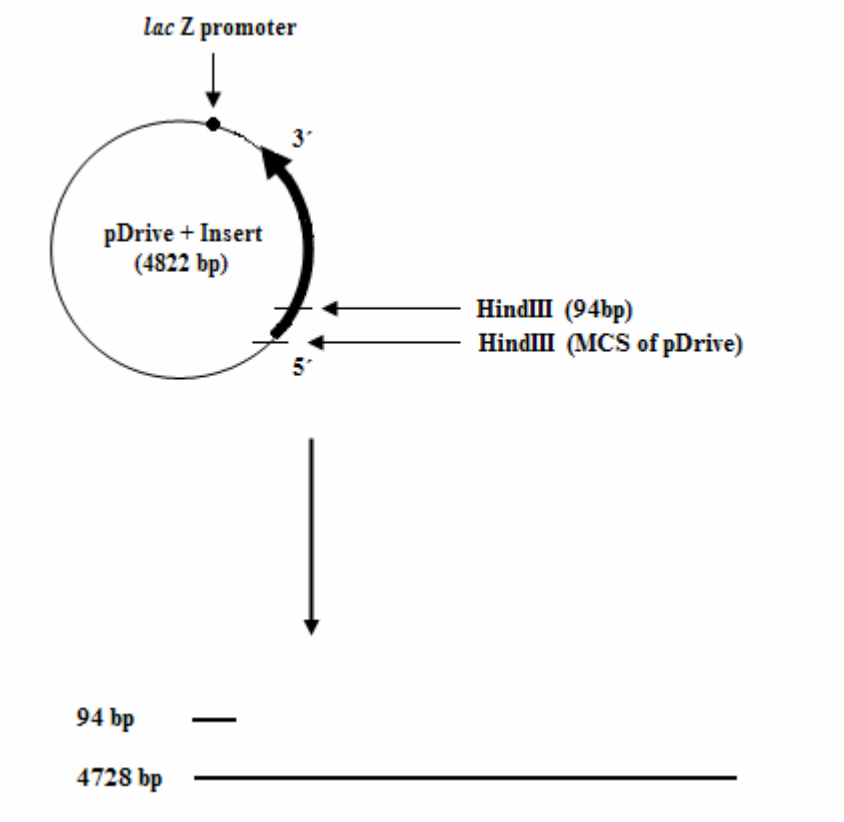

Figure 3.12 : *HindIII* digestion of cloning construct with PSP gene.

The restriction map which was derived from our experimental data (Figure 3. 14) was in good agreement with the restriction map derived by using Restriction Mapper Version 3 program from the known sequence of PSP gene. This result confirmed that we have successfully cloned PSP gene of *Tp. volcanium* in pDrive cloning vector. Furthermore, based on the restriction enzyme analyses of the recombinant vector, we conceived that the gene fragment was located in reverse direction relative to the *lacZ* promoter of the pDrive vector (Figure 3.14).

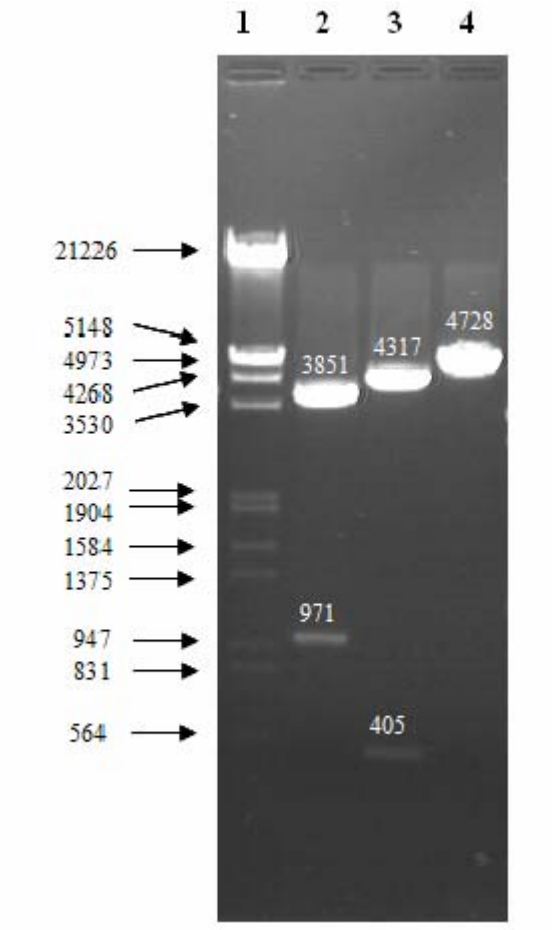

Figure 3.13 : Restriction enzyme digestions of recombinant pDrive-S1 plasmids including PSP gene fragment. *EcoRI* excised the gene fragment (971 bp) from the vector (3851 bp) (Lane 2). *BamHI* digestion yielded two fragments with the lenghts of 4317 bp and 405 bp (Lane 3). *HindIII* digestion gave rise to two fragments that were 4728 bp and 94 bp in length (Lane 4). Since the 94 bp is a very small fragment, it cannot be seen on agarose gel. Lane 1 is the *EcoRI*/*HindIII* cut Lamda DNA molecular weight marker (MBI Fermentas, AB, Vilnius).

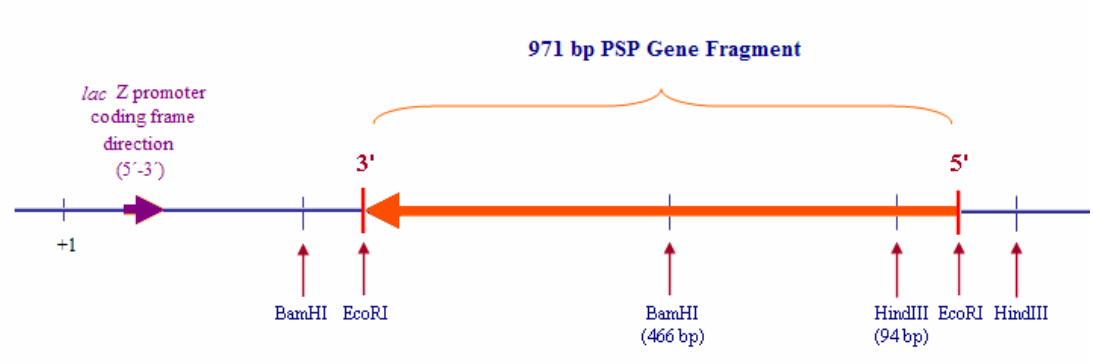

Figure 3.14 : Schematic representation of the restriction map of recombinant plasmid which include PSP gene.

3.1.4.3 Further Characterization of Recombinant Plasmids Containing MSP Gene

Plasmids isolated from recombinant pDrive-M1 cells (Figure 3.10.B) were digested with *EcoRI*, *BamHI* and *HindIII* endonucleases for further caharacterization and restriction mapping (Figure 3.18).

There are two restriction sites for *EcoRI* on MCS of pDrive vector which are located on the flanking sequences. Restriction map which was derived by software analysis showed that *EcoRI* is a non-cutter for MSP gene fragment. Therefore, this enzyme excised the gene fragment from the vector and gave rise to two products with the sizes of 1521bp and 3851 bp corresponding to the lengths of insert and vector, respectively (Figure 3.15; Figure 3.18).

BamHI restriction endonuclease cuts the vector in MCS and cuts the gene fragment at position 492 bp. For that reason, digestion with *BamHI* yielded two fragments that were 1029 bp and 4343 bp in lengths (Figure 3.16; Figure 3.18).

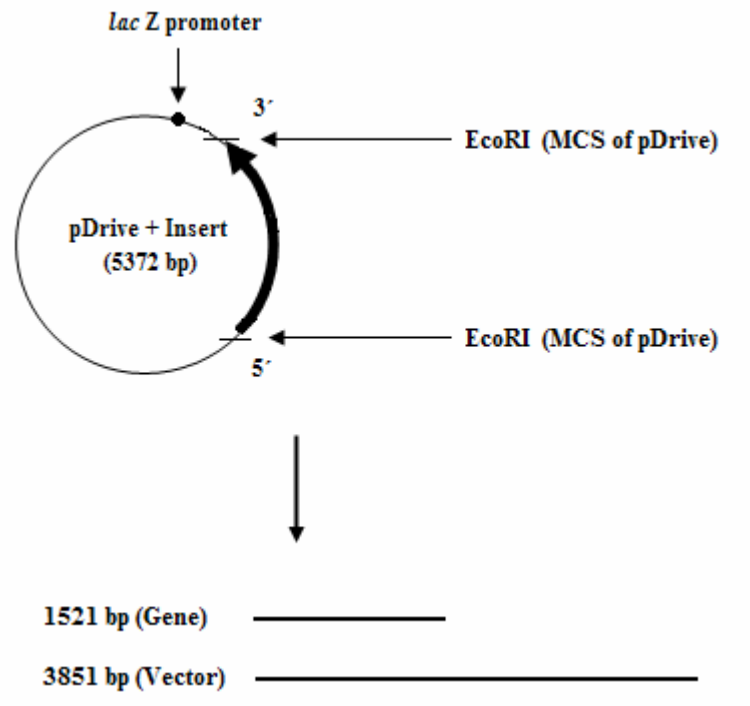

Figure 3.15 : *EcoRI* digestion of cloning construct with MSP gene.

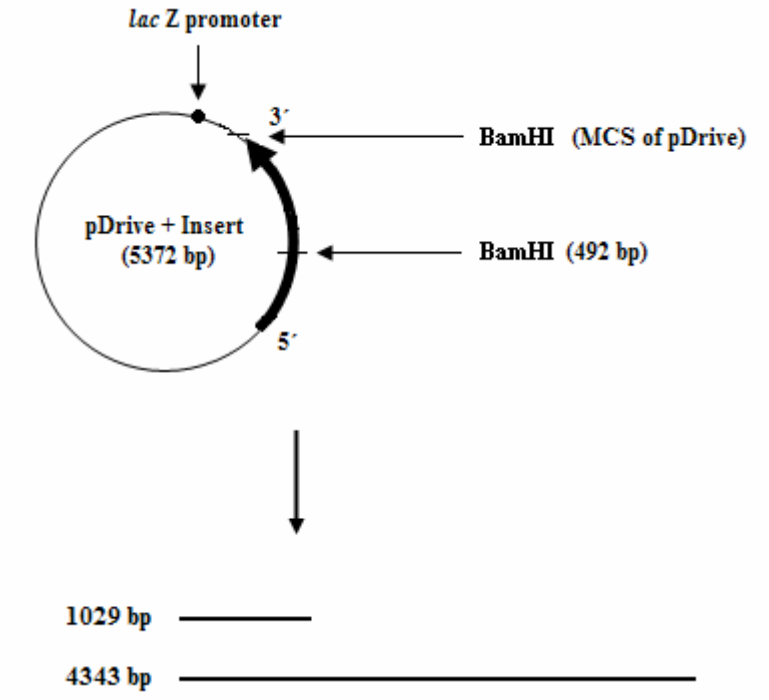

Figure 3.16 : *BamHI* digestion of cloning construct with MSP gene.

Recombinant plasmids with MSP gene fragment has two restriction sites for *HindIII* endonuclease. One of them is in MCS of the vector, while the other one is in the gene fragment at position 830 bp. Therefore, the products of this digestion were fragments with the lenghts of 830 bp and 4542bp (Figure 3.17; Figure 3.18).

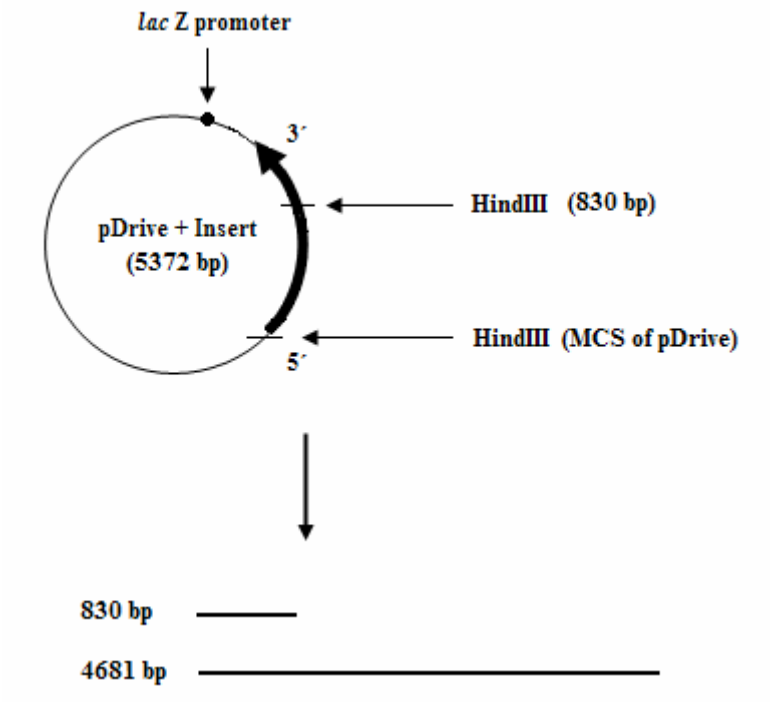

Figure 3.17 : *HindIII* digestion of cloning construct with PSP gene.

The restriction map which was derived from our experimental data (Figure 3.19) was in good agreement with the restriction map derived from the known sequence of MSP gene by using Restriction Mapper Version 3 program. This result showed that we cloned the MSP gene of *Tp. volcanium* successfully.

This result also conceived us that the gene fragment was located in reverse direction respect to the *lacZ* promoter of the pDrive vector.

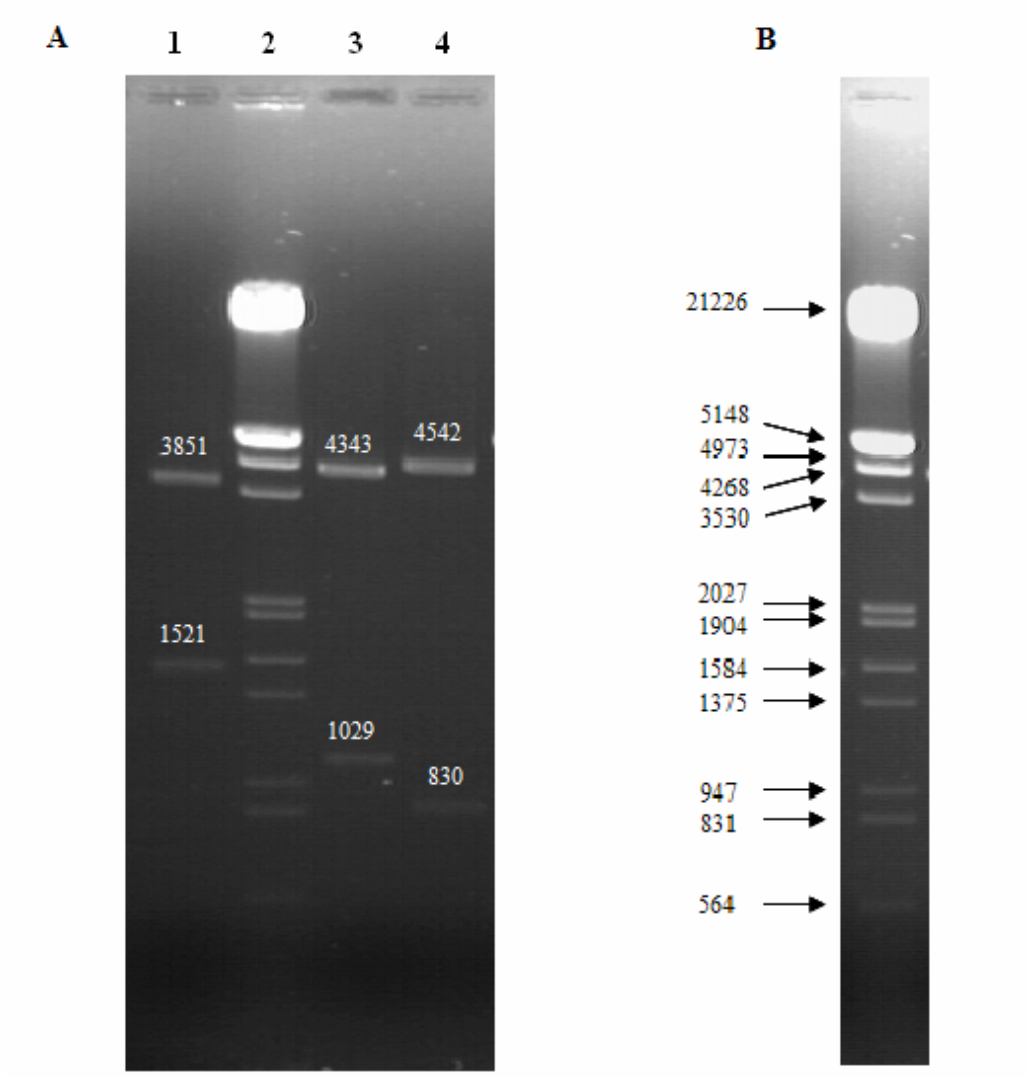

Figure 3.18 : A) Restriction enzyme digestions of recombinant pDrive-M1 plasmids including MSP gene fragment. *EcoRI* digestion excised the gene fragment (1521 bp) from the vector (1521 bp) (Lane 1). *BamHI* digestion yielded two fragments with the lenghts of 4343 bp and 1029 bp (Lane 3). *HindIII* gave rise to two fragments that were 4542 bp and 830 bp in length (Lane4). Lane 2 and **(B)** *EcoRI*/*HindIII* cut Lamda DNA show molecular weight marker (MBI Fermentas, AB, Vilnius).

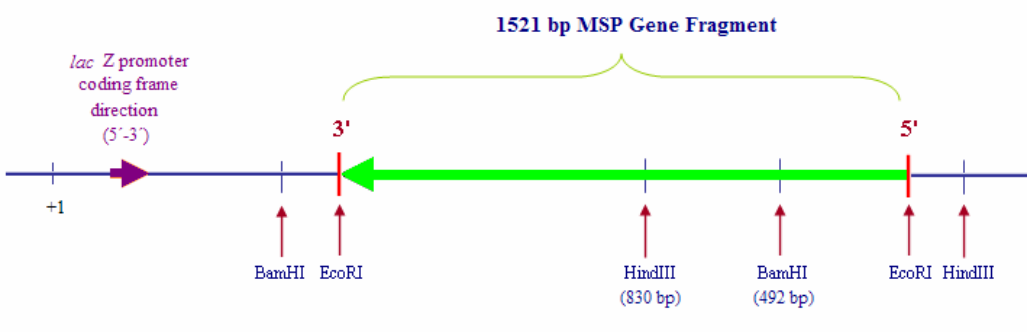

Figure 3.19 : Schematic representation of the restriction map of recombinant plasmid which include MSP gene.

3.2 Heterologous Expression of PSP gene of *Thermoplasma volcanium* in *Escherichia coli*

3.2.1 Preparation of the insert and the expression vectors for ligation

QIAExpress® Type IV Kit (QIAGEN Inc., Valencia, USA) provided a complete system for the expression and detection of 6xHis-tagged proteins. Since the 6x His tag is small, poorly immunogenic and at pH 8.0 it is uncharged, and it does not generally affect secretion, compartmentalization, or folding of the fusion protein within the cell. Also, it does not interfere with the structure or function of the purified protein. Furthermore, it allows the immobilization of the protein on metal chelating surfaces (Ni-NTA) and therefore simplifies protein interaction and purification studies. In addition, Anti-His Antibodies can be used for detection. High-level expression of 6xHis-tagged proteins in *E. coli* using pQE vectors is based on the T5 promoter/*lac* operator transcription–translation system. pQE vectors have optimized promoter–operator element consisting of phage T5 promoter and two *lac* operator sequences which increase *lac* repressor binding and ensure efficient repression of the powerful T5 promoter. Also they have multiple cloning sites and translational stop codons in

all reading frames for convenient preparation of expression constructs. Furthermore, two strong transcriptional terminators: *t0* from phage lambda, and T1 from the *rrnB* operon of *E. coli*, prevent read-through transcription and ensure stability of the expression construct.

Recombinant constructs based on the pQE vectors were produced by placing the 6xHis tag at the 5' end of the PSP gene (Figure 3.20). Therefore, recombinant pDrive vectors with the PSP gene fragments in reverse direction with respect to the *lacZ* promoter and the expression vectors, pQE 30 and pQE 32 vectors were cut by *SalI* and PstI restriction endonucleases to be able to ligate the 5' end of the ORF to the 6xHis tag. Following digestion, fragments were run in low melting point agarose gel and then purified from gel prior to ligation (Figure 3.21).

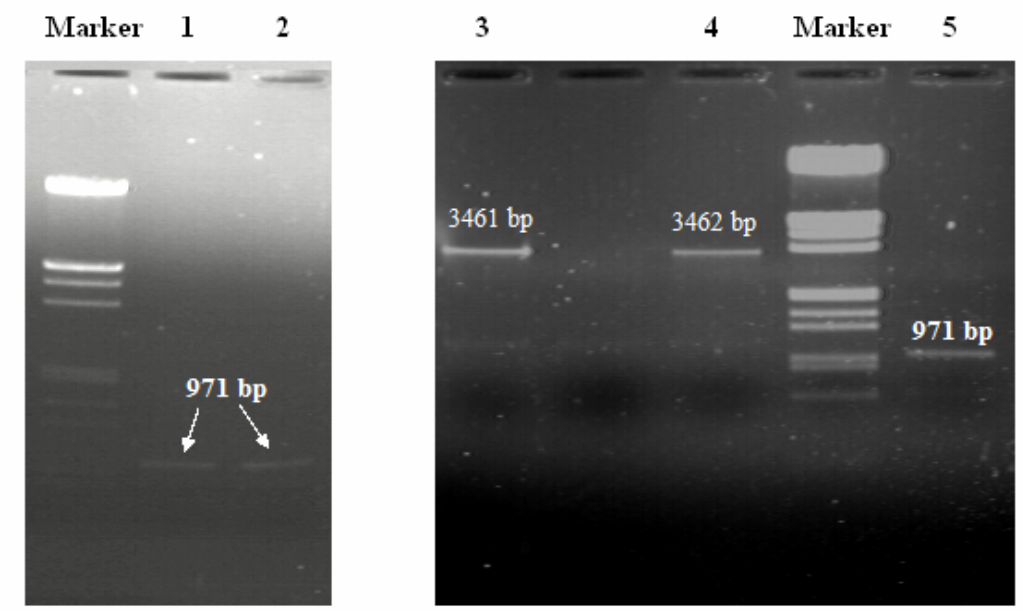

Figure 3.21 : Samples in lanes 1, 2 and 5 are PSP gene fragments (971 bp) isolated from low melting agarose gel. Gel-isolated linear pQE30 and pQE32 vectors with the lengths of 3461 and 3462 bp are in lane 3 and 4, respectively. Marker : *EcoRI*/*HindIII* cut Lamda DNA (MBI Fermentas, AB, Vilnius).

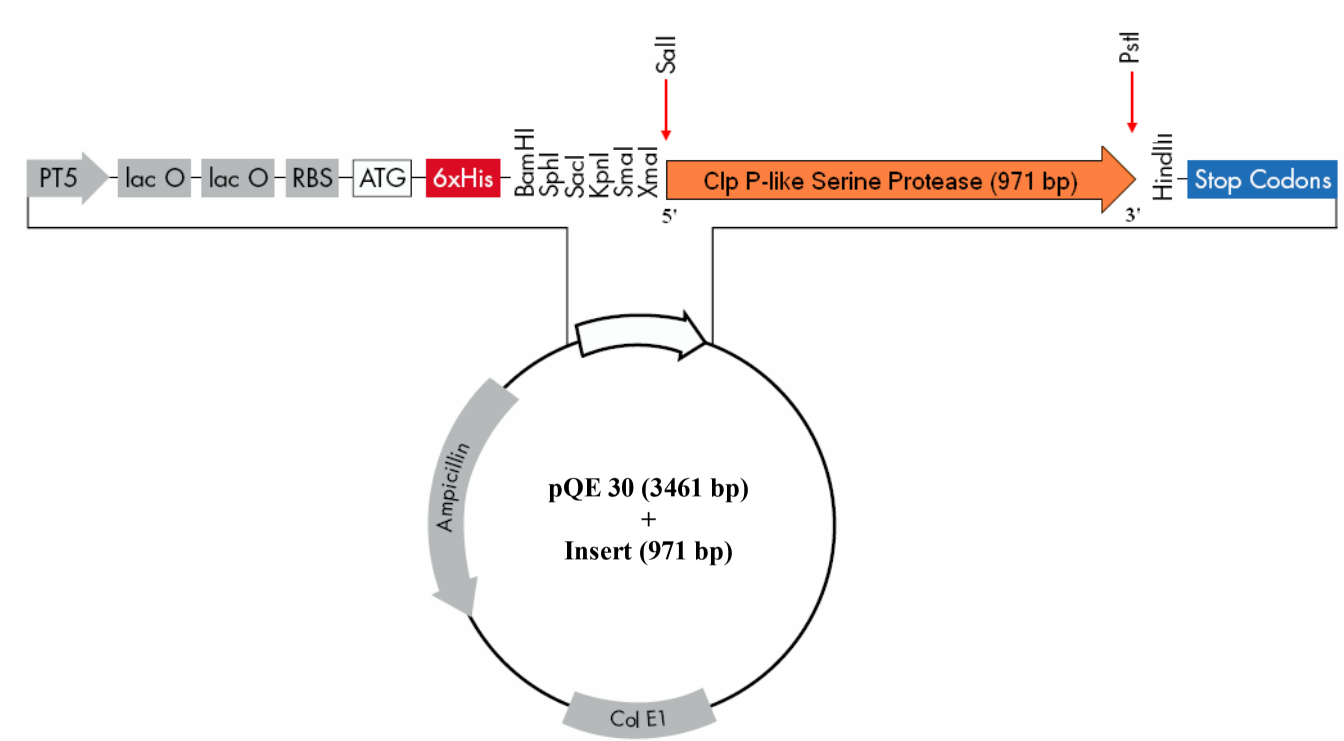

Figure 3.20 : Schematic representation of the pQE expression construct.

The inserts and the pQE 30, pQE 32 vectors which were isolated from gel were ligated by using T4 DNA ligase under standard conditions as mentioned in Materials and Methods, section 2.5.2.

3.2.2 Transformation

The pQE expression constructs produced by ligation were transformed into the competent *E.coli* TGI and XLI Blue strains. Transformation frequencies were $7x10^4$ transformants/µg DNA and 10^3 transformants/µg DNA, respectively.

Transformants were grown on plates containing both ampicillin and IPTG. However, since the transformation efficiency was very low (10^3 transformants/µg DNA) when we used XL1 Blue strain, we could not use any of the XL1 Blue plates as a master plate for colony blotting.

Six plates that were prepared with 10^{-1} dilutions of TG1 competent cells were selected as the master plates for colony blotting. Four of these plates contained the colonies with expression constructs of pQE30 vector and two of them contained the colonies with expression constructs of pQE32 vector (Table 3.2).

Recombination frequencies of pQE 30 and pQE 32 vectors were $2x10^6$ recombinants/ng DNA and 10^6 recombinants/ng DNA, respectively.

3.2.3 Colony Blotting

The colony-blot hybridization procedure was used as described in the Kit Manual for the identification of clones expressing a 6xHis-tagged protein. Positive expression clones were detected directly on colony blots using Anti-His HRP Conjugates and DAB substrate. Transformed colonies were transferred to nitrocellulose membranes that were laid on the surface of the six

Table 3.2 : Numbers of colonies that were selected from the colonies which gave (+) signals on nitrocellulose membranes.

RESULTS OF COLONY BLOTTING		
Master Plate	**Vector**	**Number of Colonies that Gave Significant Signals**
Plate 1	PQE 32	1
Plate 2	PQE 32	6
Plate 3	PQE 30	2
Plate 5	PQE 30	2
Plate 6	PQE 30	3
Plate 7	PQE 30	4

master plates. Protein expression was induced, and after incubation, the cells were lysed with alkali as mentioned in section 2.5.4. The proteins that were released bound to the membranes in the positions of the colonies.

Anti-His HRP Conjugates (QIAGEN) that are consisted of mouse monoclonal IgG1 Anti-His Antibodies coupled to horseradish peroxidase were used for direct detection of proteins with accessible 6xHis tags by chromogenic method. After transfer of proteins (including the 6xHis-tagged protein), the remaining protein-free sites on the membrane were blocked so as to prevent high background due to binding of Anti-His HRP conjugate directly to the membrane. BSA 3% in TBS buffer was used as the blocking solution following the manufacturer's instructions. Generally, BSA gives consistently good, reproducibly high signal-to-noise ratios and is the reagent of choice for using chromogenic substrates.

Chromogenic DAB substrate gave rise to insoluble, colored products that formed a precipitate on the membrane at the location where the enzyme was situated. The results were obtained quickly and easily, but the membrane had to be photographed to provide a permanent record, since the DAB substrate is light sensitive. The reaction monitored and stopped at the optimal time to prevent overdevelopment and high background.

Out of 18 colonies that produced positive signals, 12 putative recombinant colonies (From master plates #2, #3, #6 and #7) were selected for restriction enzyme digestion characterizations. Plasmids isolated from these 12 colonies were numbered from 1 to 12.

3.2.4 Characterizations of Expression Constructs

Plasmids were isolated from putative recombinant colonies (Figure 3.22) that gave positive signals on the membrane blots and initially analyzed by restriction digestion with *SalI* with expectation of linearization of the recombinant vectors, since there is one cut site for *SalI* endonuclease enzyme in MCS of the pQE cloning vectors. This enzyme is non-cutter for PSP gene fragment. Therefore, digestion with this enzyme gives rise to linearization of the expression construct. Digestion of six plasmids (from the colonies #2, #3, #7, #8, #10 and #11) with *SalI* yielded ~ 4432 bp fragment as expected, and these plasmids were selected for further caharacterizations (Figure 3.23).

SalI-PstI double digestion results in the excision of the gene fragment since the restriction sites for these enzymes are located at each one of the flanking sequences. As a result, two fragments with the lengths of 971 bp and 3461 bp are released corresponding to the lengths of insert and vector, respectively, as consistant with the digestion profiles of the plasmid samples from the colonies #3, #7 and #10 (Figure 3.24; Figure 3.25; Figure 3.29).

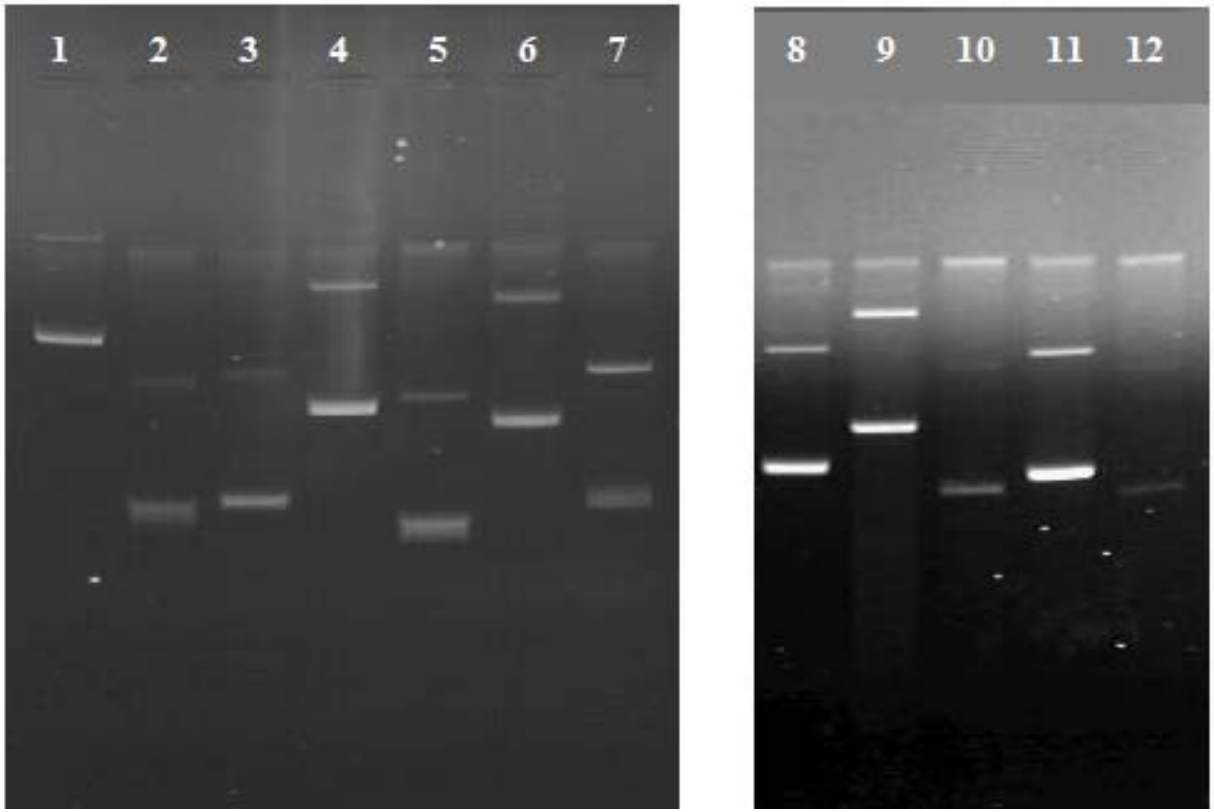

Figure 3.22 : Putative recombinant plasmids isolated from colonies that gave positive signals. Lane **1**: Colony #1 from master plate 2, Lane **2**: Colony #2 from master plate 2, Lane **3**: Colony #3 from master plate 2, Lane **4**: Colony #4 from master plate 2, Lane **5**: Colony #6 from master plate 2, Lane **6**: Colony #2 from master plate 3, Lane **7:** Colony #1 from master plate 6, Lane **8:** Colony #1 from master plate 7, Lane **9:** Colony #2 from master plate 7, Lane **10:** Colony #3 from master plate 7, Lane **11:** Colony #4 from master plate 7, Lane **12:** Colony #1 from master plate 3)

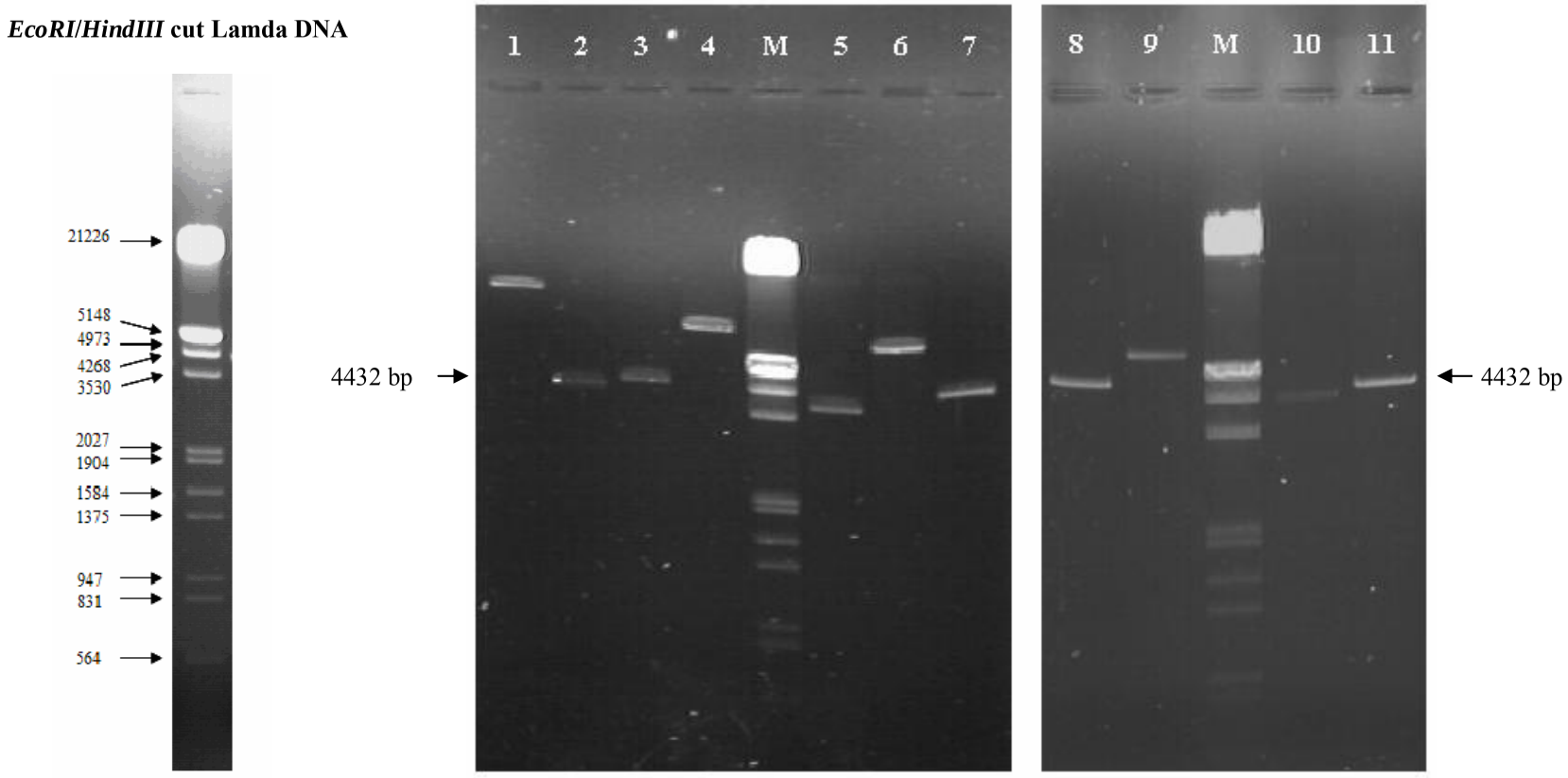

Figure 3.23 : Single digestion of the putative recombinant plasmids with SalI restriction endonuclease which linearizes the expression construct and gives rise to a fragment of 4432 bp (3461 bp vector + 971 bp gene). Samples represented as 2, 3, 7, 8, 10 and 11 were selected for further characterization since the lengths of the fragments were between 4268 bp and 4973 bp

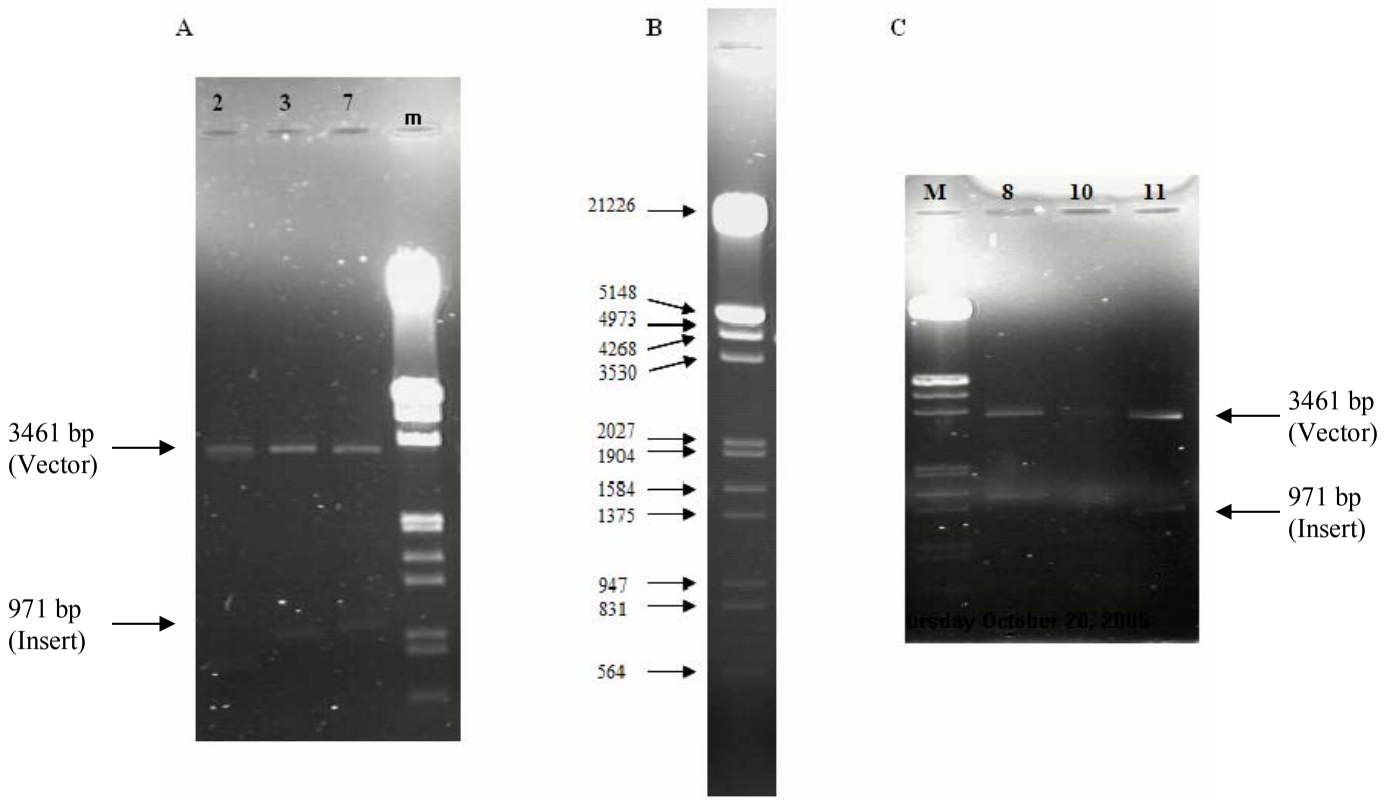

Figure 3.24 : *SalI-PstI* double digestions of plasmid samples from the colonies #2, #3, #7, #8, #10 and #11. Digestion profiles implied that plasmids #3, #7 and #10 could be true recombinants. These samples were selected for further caharacterization.

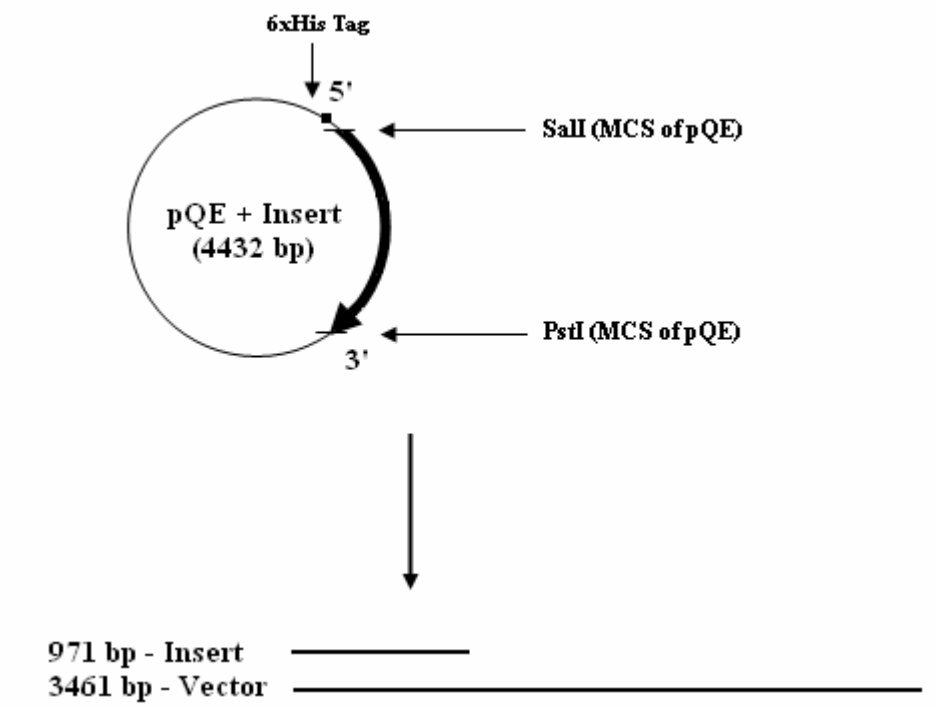

Figure 3.25: *SalI-PstI* double digestion of expression constructs.

BamHI cuts the PQE vectors within the MCS. There are two restriction sites for *BamHI* in the insert sequence. One of them comes from the pDrive vector when the fragment is excised with the *SalI-PstI* double digestion in order to insert it into pQE vectors. The other one is at the 446th position of the PSP gene. Therefore, Bam*HI* single digestion gives rise to three fragments with the sizes of 466 bp, 505 bp and 3461 bp (Figure 3.26; Figure 3.27; Figure 3.29).

HindIII endonuclease cuts the gene fragment at position 94 bp. It also cuts the pQE vectors at MCS. Single digestion with *HindIII* yields two fragments of 877 bp and 3555 bp in lengths (Figure 3.28; Figure 3.29).

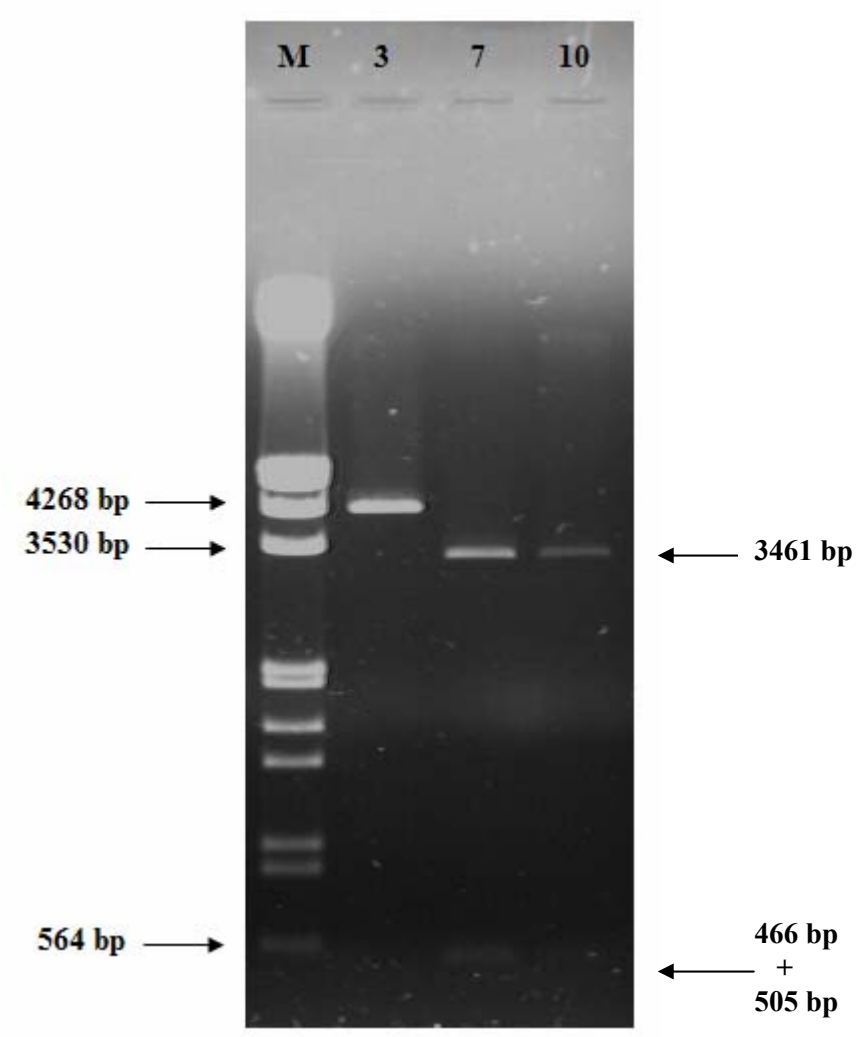

Figure 3.26 : Restriction enzyme digestions of samples 3, 7 and 10 with *BamHI* endonuclease.

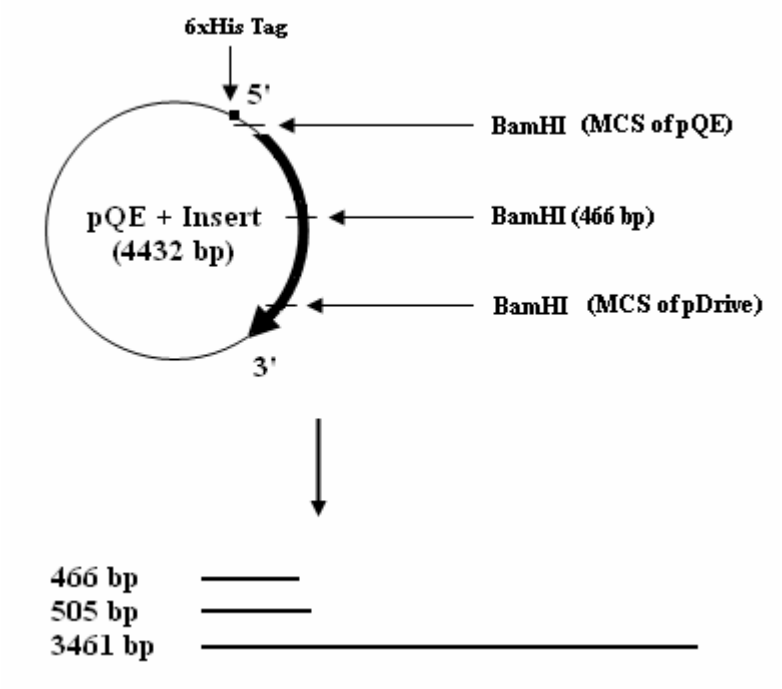

Figure 3.27 : Schematic representations of *BamHI* single digestions of expression constructs.

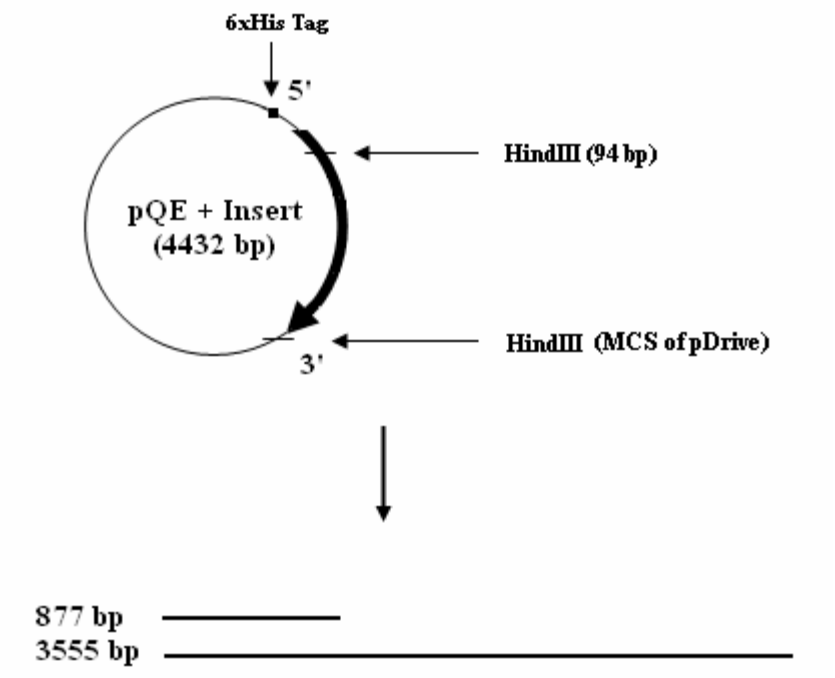

Figure 3.28 : Schematic representations of and Hind*III* single digestions of expression constructs.

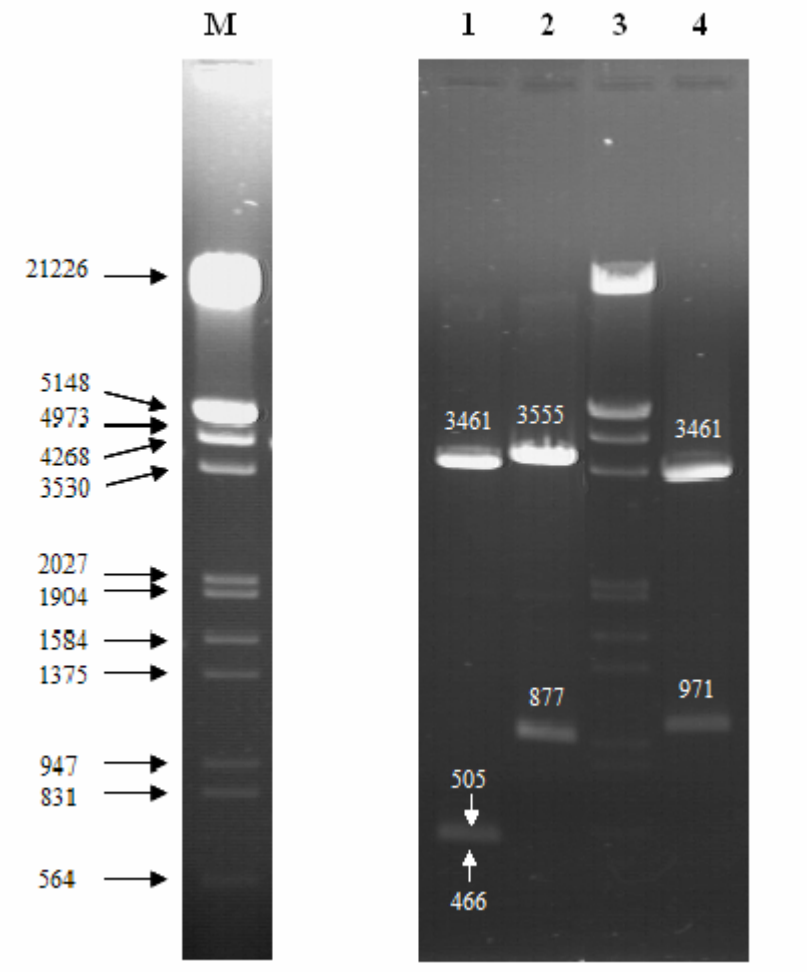

Figure 3.29 : Restriction enzyme digestion profile of the recombinant plasmid from colony #7 which was originally from master plate #6. Lane1: *BamHI* digestion; Lane 2: *HindIII* digestion; Lane 3: *EcoRI*/*HindIII* cut Lambda DNA Marker; Lane 4: *SalI-PstI* double digestion.

Restriction enzyme digestions with *BamHI, HindIII* and *SalI-PstI* confirmed that the plasmid sample from colony #7 carries the cloned PSP gene fragment in correct position (Figure 3.29). This plasmid is isolated from the clone #1 that was originally on master plate #6 (Figure 3.30; Figure 3.31). Schematic representation of the restriction digestion profile of the recombinant vector and its insert are shown in Figure 3.32. Colonies which carry these recombinant plasmids were named as *E. coli* TG- pQE30-S1.7.

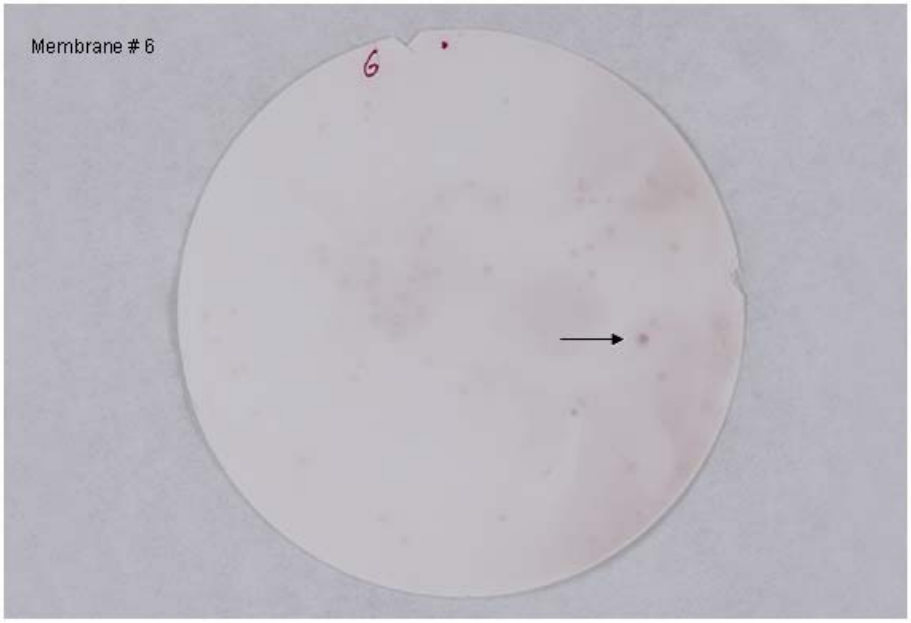

Figure 3.30 : Membrane blot of the master plate # 6, which contained protein blots of transformed cells with pQE 30 expression constructs. Colony # 7, indicated by the arrow was found to be a true recombinant depending on restriction enzyme profile and named as *E. coli* TG-pQE30-S1.7.

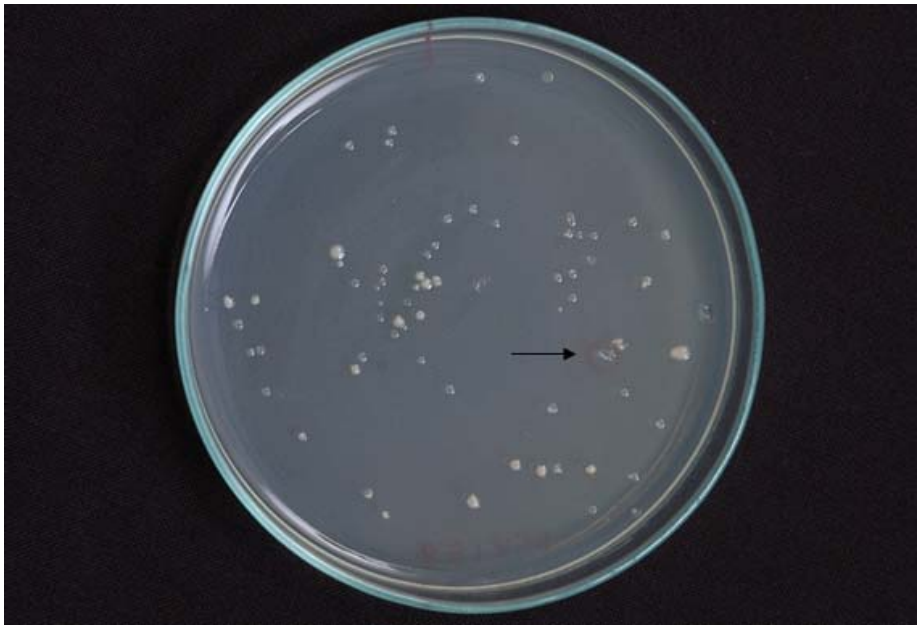

Figure 3.31 : Recombinant clone *E. coli* TG-pQE30-S1.7 (indicated by the arrow) on master plate # 6.

3.3. Enzymatic Studies

3.3.1. Determination of Extracellular and Intracellular Serine Protease Activities of the Recombinant Cells by Using Chromogenic Substrates

3.3.1.1 Enzymatic Activities of *E. coli* TG-S1, *E. coli* TG-S4 and *E. coli* TG-M1 Recombinant Colonies

Supernatants and cell free extracts of the recombinant clones *E.coli* TG-M1 (with MSP gene), *E. coli* TG-S1 and *E. coli* TG-S4 (with PSP gene) which were grown in LB medium were used for the enzyme assays to detect expression of serine-like protease activities by using N- Suc- Ala- Ala- Pro- Phe- pNA and N-Suc-Ala-Ala-Phe-pNA as substrates.

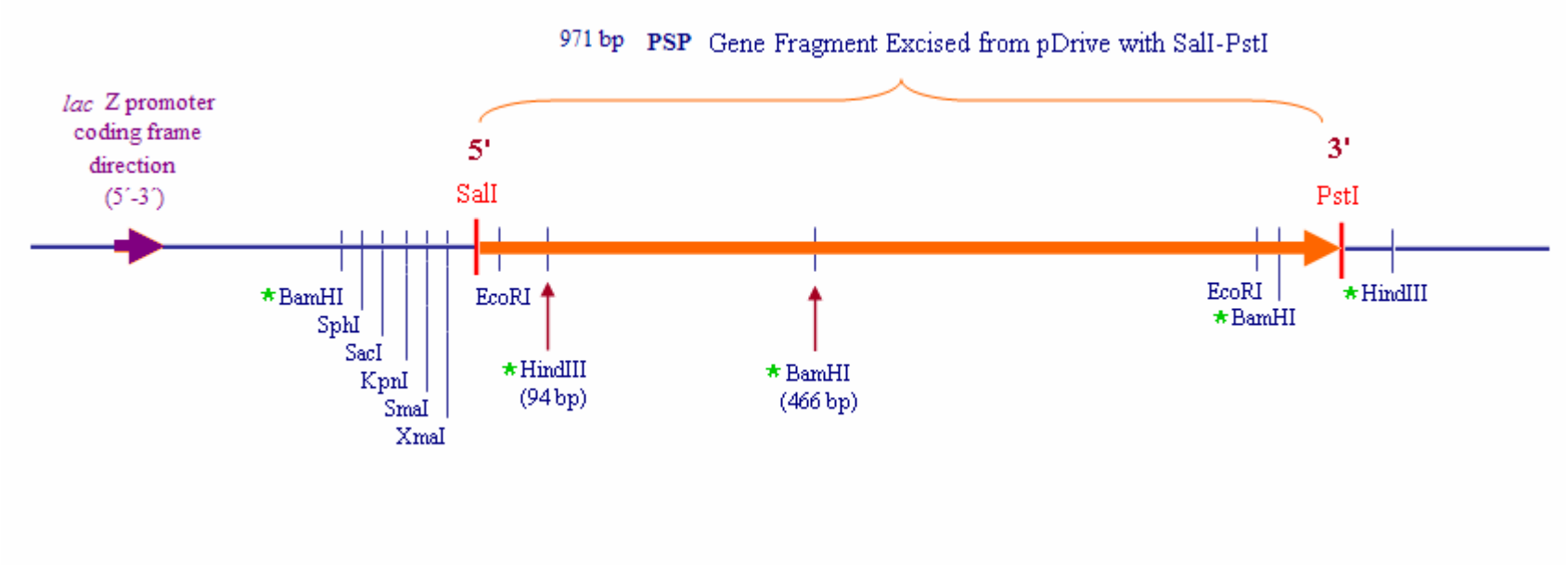

Figure 3.32 : Schematic representation of the restriction map of insertion site of the gene fragment which was excised from the pDrive cloning vector and subcloned into the pQE expression vector.

Activities were monitored spectrophotometrically based on measurement of the rate of p-nitroaniline release as described in section 2.4.3. All of the samples showed activity against both of the substrates.

Intracellular activities ($\Delta Abs_{410} x min^{-1}$) of *E.coli* TG-M1, *E. coli* TG-S1 and *E. coli* TG-S4 cell free extracts were measured as 137 x 10^{-3}, 110 x 10^{-3} and 84 x 10^{-3}, respectively, for N- Suc- Ala- Ala- Pro- Phe- pNA being the substrate. Cell free extract of *E.coli* TG1 was used as control and the activity of the control was $\Delta Abs_{410} x min^{-1}$=6,2 x 10^{-3} (Figure 3.33; Figure 3.34; Figure 3.35).

Activities of the cell free extracts derived from recombinant clones were significantly higher than the cell free extract of *E. coli* TG1 (control) strain. Expression levels of *E.coli* TG-M1, *E. coli* TG-S1 and *E. coli* TG-S4 strains were 23-folds, 18-folds and 14-folds higher than the wild-type *E. coli* TG1 (control) strain. These results showed that the recombinant protease genes were successfully overexpressed under the control of their own promoters.

Experimental data, also, showed that recombinant enzyme activities of cell free extracts were significantly higher than the supernatants of each clone. Activities of the supernatants were found to be $0.7x10^{-3}$, $1.7x10^{-3}$ and $1.4x10^{-3}$ $\Delta Abs_{410} x min^{-1}$ for *E. coli* TG-M1, *E. coli* TG-S1 and *E. coli* TG1 (control), respectively. This indicates that there is no difference between the recombinants and wild-type strain with respect to extracellular serine protease activities.

Furthermore, after heat treatment at 60°C for 15 min, serine protease-like activities of the *E.coli* TG-M1 ($\Delta Abs_{410} x min^{-1}$ = $17x10^{-3}$), *E. coli* TG-S1($\Delta Abs_{410} x min^{-1}$ = $21x10^{-3}$), and *E. coli* TG-S4 ($\Delta Abs_{410} x min^{-1}$ = $1.8x10^{-3}$) strains decreased approximately 6-folds. This result indicates that recombinant PSP and MPS retained about 15% of their original activities at 60°C after 15 min.

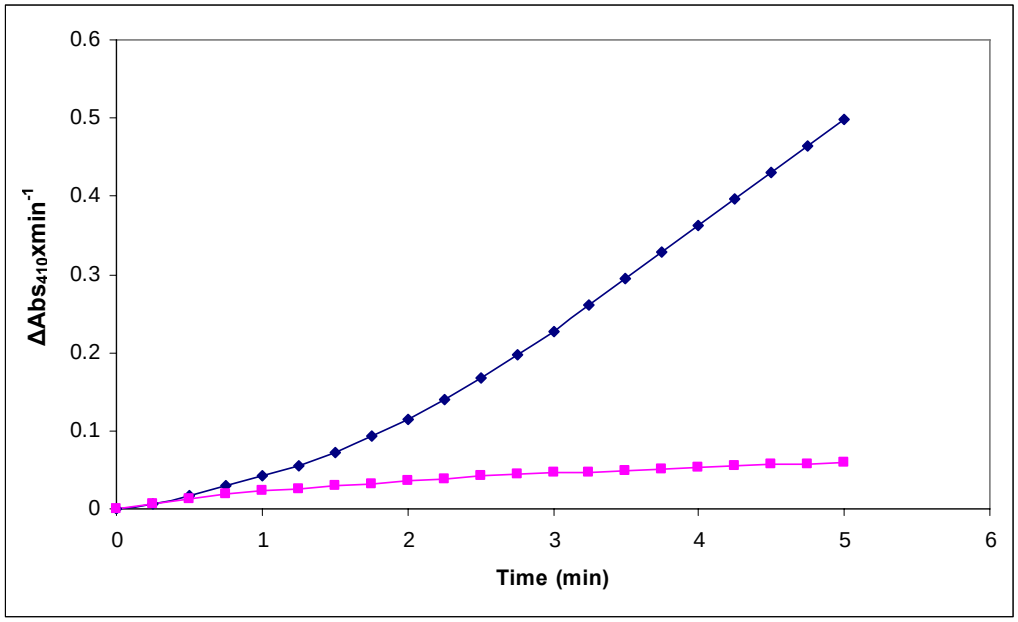

Figure 3.33 : Serine protease activity of *E. coli* TG-M1 cell free extract using N- Suc-Ala- Ala- Pro- Phe- pNA as substrate. Activity of *E. coli* TG1 cell free extract was used as control. *E. coli* TG-M1 *E. coli* TG1 (control)

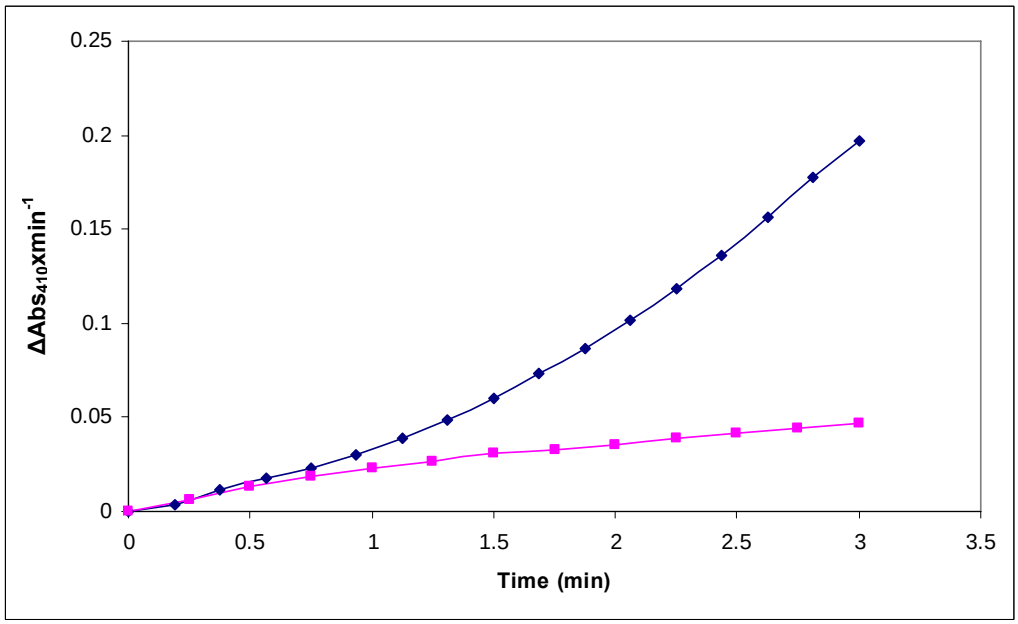

Figure 3.34 : Serine protease activity of *E. coli* TG-S1 cell free extract using N- Suc-Ala- Ala- Pro- Phe- pNA as substrate. Activity of *E. coli* TG1 cell free extract was used as control. *E. coli* TG-S1 *E. coli* TG1 (control)

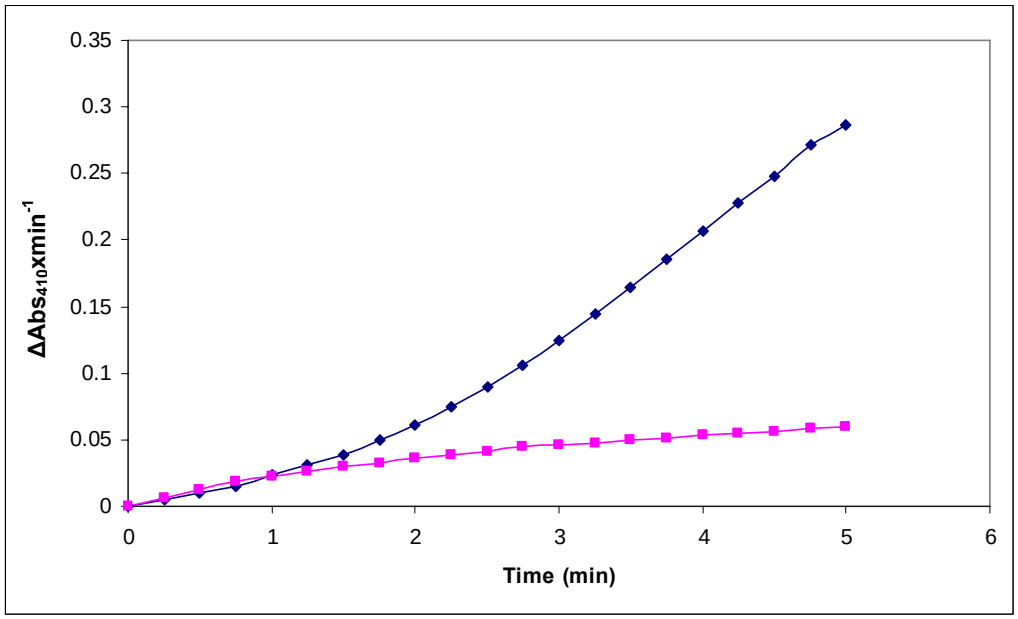

Figure 3.35 : Serine protease activity of *E. coli* TG-S4 cell free extract using N- Suc- Ala- Ala- Pro- Phe- pNA as substrate. Activity of *E. coli* TG1 cell free extract was used as control. *E. coli* TG-S4 *E. coli* TG1 (control)

3.3.1.2 Serine Protease Activity of 6xHis Tagged Recombinant Enzyme

N- Suc- Ala- Ala- Pro- Phe- pNA and N-Suc-Ala-Ala-Phe-pNA synthetic substrates were used to detect the activity of 6xHis tagged recombinant enzyme derived from *E. coli* TG-pQE30-S1.7 colonies which possess the PSP gene in pQE 30 expression vectors.

Serine protease activity of the cell free extract of *E. coli* TG-S1.7 was $46.2x10^{-3}$ $\Delta Abs_{410}xmin^{-1}$ and 50.01 $\Delta Abs_{410}xmin^{-1}$ for N- Suc- Ala- Ala- Pro- Phe- pNA and N-Suc-Ala-Ala-Phe-pNA substrates (Represented in Figure 3.36 and Figure 3.37), respectively. Cell free extract of *E.coli* TG1 was used as control and the activity of the control was 8,1 x 10^{-3} $\Delta Abs_{410}xmin^{-1}$. These results indicate the successful expression of PSP gene in *E. coli* as 6xHis tag fusion protein.

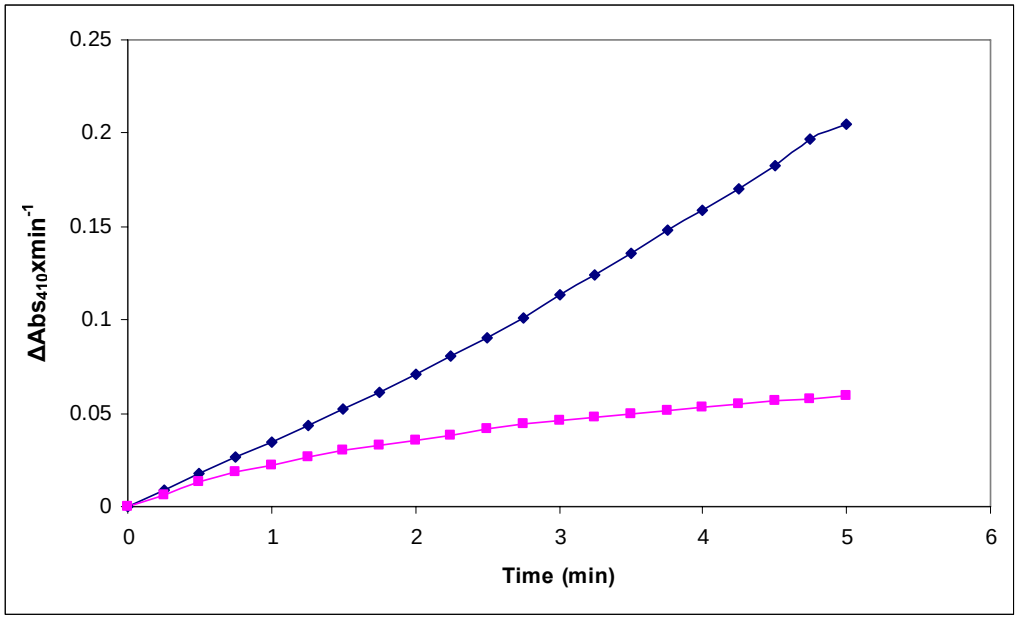

Figure 3.36 : Intracellular serine protease activity of *E. coli* TG-pQE30-S1.7 cells with the substrate N- Suc- Ala- Ala-Pro-Phe- pNA. Activity of *E. coli* TG1 cell free extract was used as control. *E. coli* TG-pQE30-S1.7 *E. coli* TG1 (control).

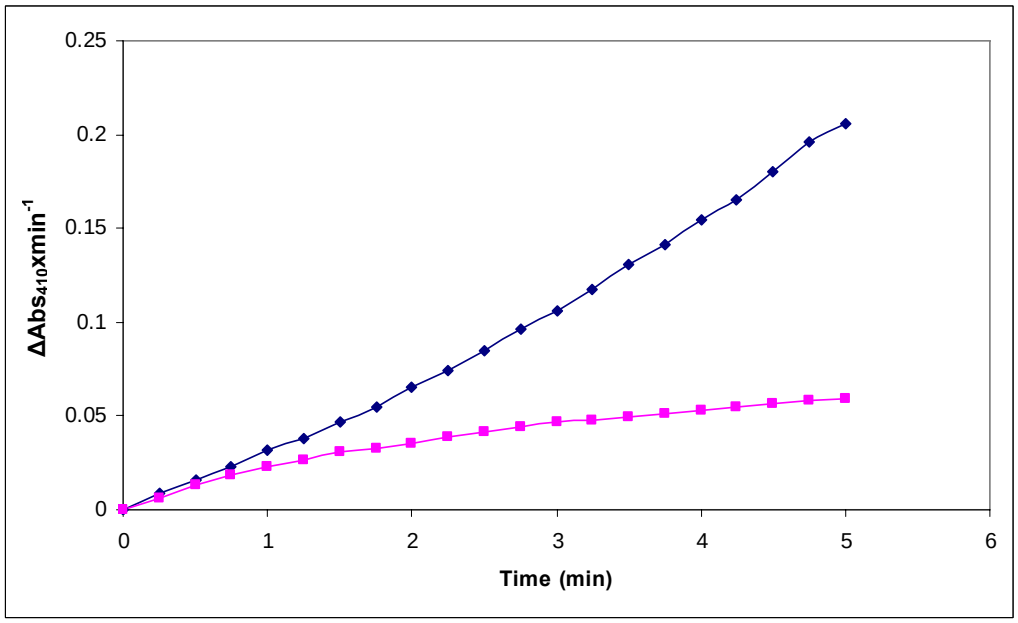

Figure 3.37 : Intracellular serine protease activity of *E. coli* TG-pQE30-S1.7 cells with the substrate N- Suc- Ala- Ala- Phe- pNA. Activity of *E. coli* TG1 cell free extract was used as control. *E. coli* TG-pQE30-S1.7 *E. coli* TG1 (control).

3.3.2 Heat Shock Response of *Thermoplasma volcanium* Cells

Tp. volcanium cells were grown until mid-exponential phase at 60°C and then, the temperature was shifted to 72°C for 2 h and 80°C for 2 h, in two different experiments. Control cultures were allowed to grow at 60°C. The cell growth was monitored through spectrophotometric measurements of the cell density at 610 nm, in a time dependent manner as mentioned in section 2.4.4.

Heat shock at 72°C for 2 h resulted in an increase in the growth rate of the cells which were allowed to grow at 60°C. The growth rate [$\Delta Abs_{640} x (10h)^{-1}$] which was $0.95 x 10^{-2}$ before heat shock increased to 1.62×10^{-2} after heat shock while the growth rate of the control cells was $0.95 x 10^{-2}$ $\Delta Abs_{640} x (10h)^{-1}$ throughout the log phase. This result shows that the heat shock at 72°C for 2 h induced the cell growth about 2-folds as compared to the control which was not exposed to the heat shock.

Following heat shock at 80°C for 2 h, when the cells were allowed to grow at 60°C, the growth ceased soon after as revealed by the plateau at 145^{th} h in the growth curve (Figure 3.39).

Enzymatic activities were measured based on the p-nitroanilide liberation from N-Succinyl-Ala-Ala-Pro-Phe-pNA, N-Succinyl-Ala-Ala-Phe-pNA and N-CBZ-Arg oligopeptide substrates. There was no activity when N-CBZ-Arg was used. Serine protease activities measured using N-Succinyl-Ala-Ala-Pro-Phe-pNA in the cell free extract and supernatant were $0.4 x 10^{-3}$ and $1.8 x 10^{-3}$ $\Delta Abs_{410} x min^{-1}$, respectively, before heat shock. After heat shock at 72ºC for 2h, the activities were $2.6 x 10^{-3}$ and $2.4 x 10^{-3}$ $\Delta Abs_{410} x min^{-1}$ for cell free extract and supernatant, respectively, for the same substrate. These resuls show that the activities of cell free extract and supernatant increased approximately 6.5 and 1.3 folds, respectively, as shown in Table 3.3.

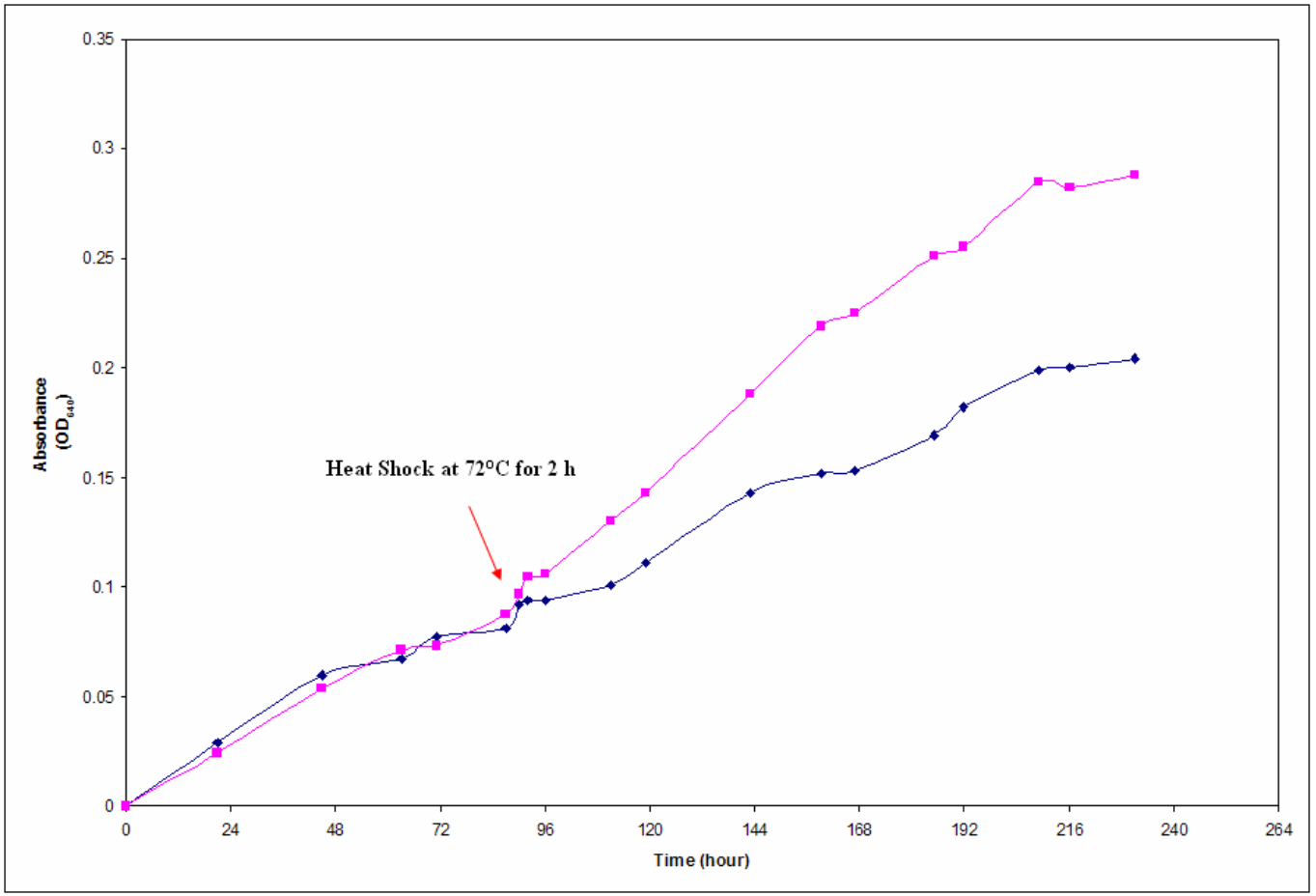

Figure 3.38 : Effect of heat shock at 72°C for 2 h on the growth rate of *Thermoplasma volcanium*.
Tp. volcanium culture that exposed to heat shock Control *Tp. volcanium* culture.

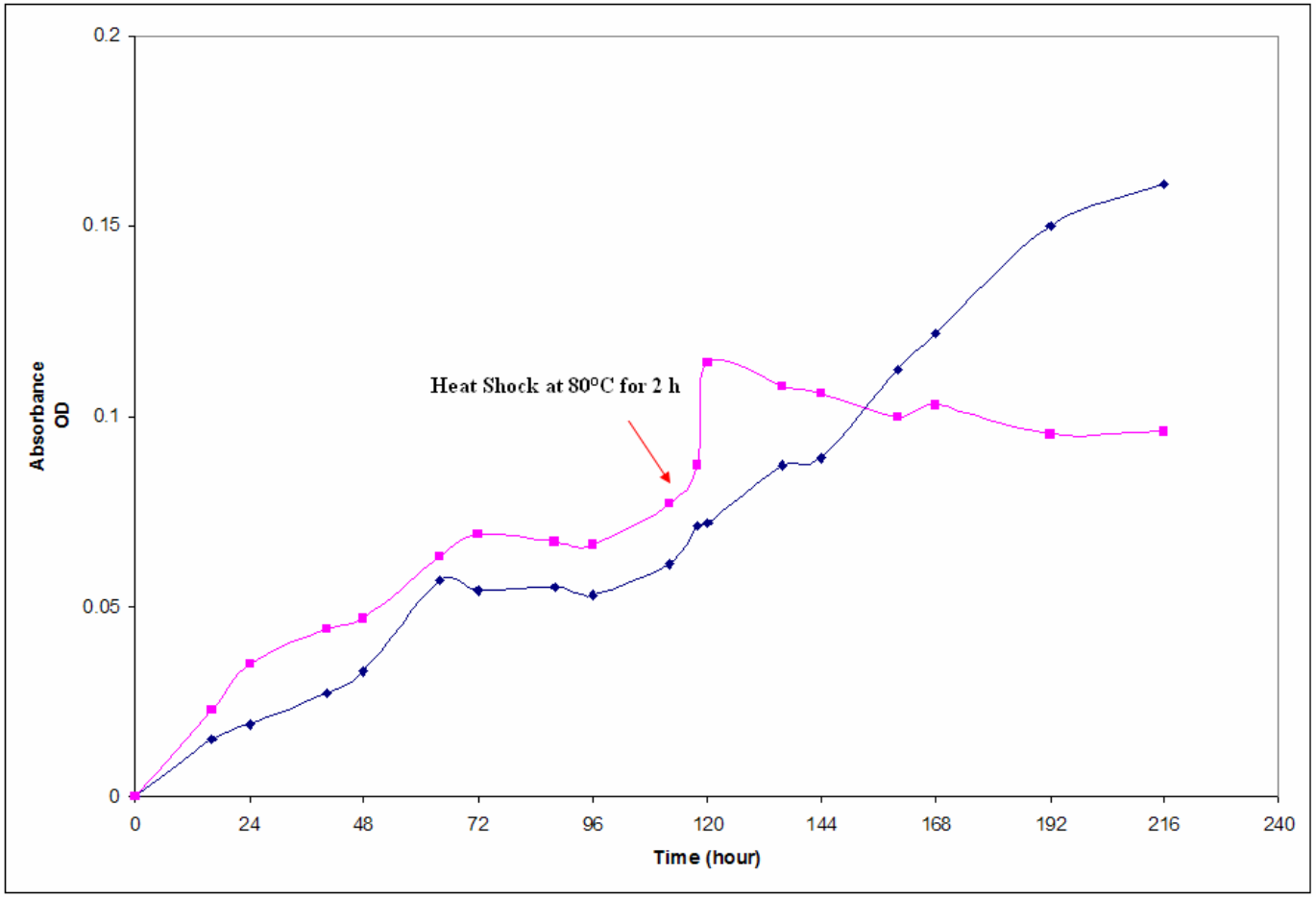

Figure 3.39 : Effect of heat shock at 80°C for 2 h on the growth rate of *Thermoplasma volcanium*.

Tp. volcanium culture that exposed to heat shock Control *Tp. volcanium* culture.

Table 3.3 : Serine protease activities of *Tp. volcanium* before and after heat shock determined by using N- Suc- Ala- Ala- Pro- Phe- pNA.

	Serine Protease Activity Δ **Ab$_{410}$ x min^{-1}**	
	Supernatant	**Cell Free Extract**
Before Heat Shock (76 h culture)	1.8×10^{-3}	0.4×10^{-3}
After Heat Shock at 72ºC for 2h	2.4×10^{-3}	2.6×10^{-3}
After Heat Shock (96 h subculture)	1.4×10^{-3}	1.08×10^{-3}

3.4 SDS-PAGE Analysis of *E. coli* TG-pQE30-S1.7 Proteins

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out according to Laemmli (1970) using 12.5% polyacrylamide gel. The molecular mass of the recombinant protein was estimated by SDS-PAGE by comparison with migration of standard protein marker. Protein profile of the cell free extract of *E. coli* TG-pQE30-S1.7 strain revealed a strong protein band with a molecular mass of approximately 25 000 Dalton (Figure 3.40). Estimated molecular weight of the enzyme encoded by PSP gene is 25 585 Dalton according to NCBI and KEGG databases. This result might implay that 6xHis-PSP fusion protein was expressed in *E. coli*.

3.5 Software Analyses of Periplasmic (ClpP-like) and Membrane-Bound Serine Proteases

3.5.1 Basic Properties of the Genes

Our genes of interest were selected after a search through the complete genome

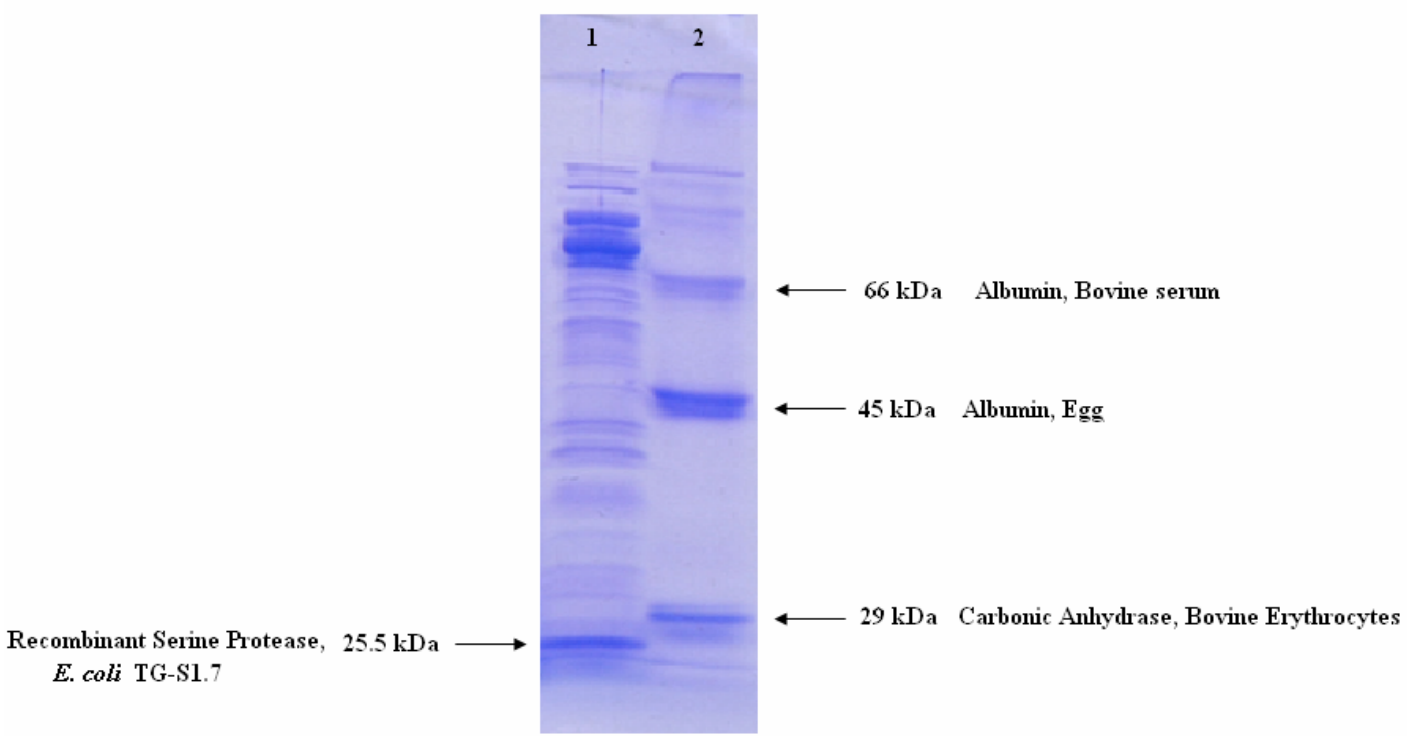

Figure 3.40 : SDS-PAGE result of *E.coli* TG-S1.7 proteins (Lane 1). Lane 2 is SDS Molecular Weight Marker MWV-ND-500 Kit, Sigma Chem. Corp., St. Louis, MO, USA.

of *Tp. volcanium* (1,584, 799 bp, G-C content 38%) that has been sequenced and released to genomic databases (Kawashima, *et al*., 2000). Analysis of the *T.volcanium* genome revealed the presence of 28 genes encoding peptidases and proteases, among which two are putative Clp-like proteases (ClpP family).

PSP was named as 'Secreted Serine Protease' on, 'Signal Peptide Peptidase' or 'Proteinase IV' in different databases (NCBI, KEGG) depending on the motifs that are found in its protein structure. Generally, the putative identification for the product is ' ClpP class serine protease'. Database searches showed that the gene (693 bp) coordinates between nucleotides 29176 and 29868. Locus of the gene is named as TVN0029. The GC content of the gene is 41.45%. It encodes a 230 aa protein with a molecular weight of 25 585.83 Dalton. pI of the protein is at pH 9.8040.

MSP (Membrane-bound serine protease) gene is located between the nucleotides 475673 and 476992 in total genome of *Tp. volcanium*. Gene length is 1320 bp with a GC content of 41.76 %. Locus of the gene is named as TVN0492. Calculated molecular weight of the 439 aa protein encoded by this gene is 46 832.22 Dalton. pI of the protein is at pH 4.8559.

3.5.2 Clusters of Orthologous Groups and Conserved Domain Database Analyses

Clusters of Orthologous Groups of proteins (COGs) were delineated by comparing protein sequences encoded in complete genomes, representing major phylogenetic lineages (phylogenetic classifications of proteins from complete genomes). They were also defined as recurring units in molecular evolution, whose extents can be determined by sequence and structure analysis. Our enzymes were classified in COG0616 = Periplasmic serine proteases (ClpP class) and COG1030 = Membrane bound serine proteases (ClpP class) which

are serine proteases of the Clp family Moreover, *Tpv* periplasmic serine protease was classified in COG0740 = Protease subunit (ClpP) of ATP-dependent Clp proteases. Members of COG0616, COG0740 and COG1030 all belong to the Peptidase family S49 and they are associated with posttranslational modification, protein turnover and chaperones.

In consistant with this classification, NCBI domain search indicated that *Tp. volcanium* PSP has peptidase S49, Clp protease, Signal peptide peptidase A (SppA) and ClpP protease family domains in its structure (Figure 3.41).

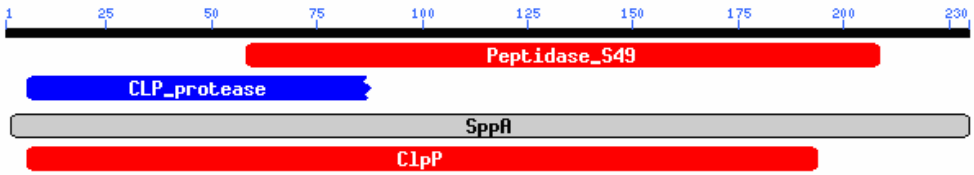

Figure 3.41 : Domains found in *Tp. volcanium* PSP structure.

ClpP domain (COG0740) of *Tp. volcanium* PSP is found to be related to a number of bacterial ATP-dependent ClpP Protease proteolytic subunits which are shown in the phylogram (Figure 3.42).

Furthermore, the homology is quite high between *Tp. volcanium* PSP and members of MEROPS peptidase family S49 (protease IV family, clan S). Protease of the same class from *Tp. acidophilum* was grouped together with the *Tp. volcanium* protease in the phylogram represented in Figure 3.44.

Multiple sequence alignment of PSP gene is given in Figure 3.43. The homology scores derived depending on the multiple sequence alignment are represented in Table 3.4.

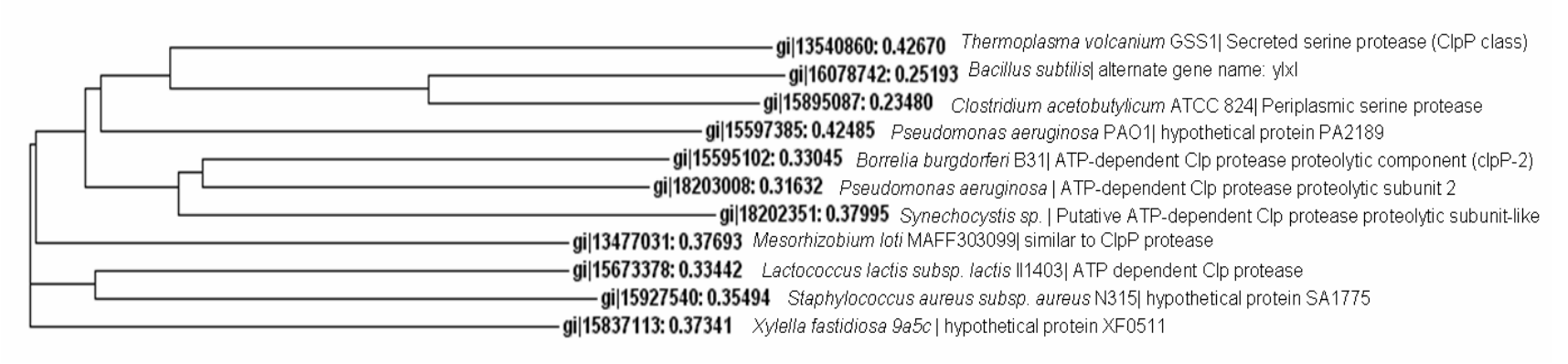

Figure 3.42 : Phylogram of *Tp. volcanium* periplasmic protease depending on its ClpP domain (COG0740) which shows close relationship with a number of bacterial ATP-dependent ClpP Protease proteolytic subunits. (Phylogram was derived by using Clustal W -1.82)

Figure 3.43 Continued

```
gi|15673378      NPPKEEPIIENKQADLRSRKLAILLEK-------------------------------- 235
gi|15927540      IDAIANKVIEKINMKEKESEIDVADSKVSANGFSRFLF---------------------- 245
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      CSPQQHALITHIEDPTMKHHLNVILAQAGTTPPPSLTPPALSPQTGHTPSPSQGNTPPQS 305
gi|13477031      SGQRDATRPAPRDAGFDPLAVQRLIQTIKS------------------------------ 281
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------

gi|15673378      ------------------------------------------------------------
gi|15927540      ------------------------------------------------------------
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      PAPGSAANDDPLTALEARNARIRDVFAAFADVPGVHDLEASCLANPRLSVEHAQAQLLQR 365
gi|13477031      ------------------------------------------------------------
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------

gi|15673378      ------------------------------------------------------------
gi|15927540      ------------------------------------------------------------
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      LPGGAGPLAATPRHGIHHLHLVHDEHTTRRQRVADGILARAGILTGPEADAARQDNPAAH 425
gi|13477031      ------------------------------------------------------------
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------

gi|15673378      ------------------------------------------------------------
gi|15927540      ------------------------------------------------------------
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      EALWVLAEQSLKATGVDTRGLDREKVSKTALAQSTGDFPVILENVMHKMLLTAYRLQSYT 485
gi|13477031      ------------------------------------------------------------
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------

gi|15673378      ------------------------------------------------------------
gi|15927540      ------------------------------------------------------------
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      WMRFCATGSLSDYRPHHRYHMGGFSDLKKVNENGEYENGVLSDAEKETIQAVRKGRILQI 545
gi|13477031      ------------------------------------------------------------
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------
```

Figure 3.43 Continued

```
gi|15673378      ------------------------------------------------------------
gi|15927540      ------------------------------------------------------------
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      TPEVLVNDDLGAFSRPTSRLAQAAARTIEKDVYALLALNAGSGPRMSDGKPLFHADHRNI 605
gi|13477031      ------------------------------------------------------------
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------

gi|15673378      ------------------------------------------------------------
gi|15927540      ------------------------------------------------------------
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      PVAAALSVESIDAARQLMAQQMDVGGNDFLDIVPALWLGPLSLGSKARELNAQEYNDEAG 665
gi|13477031      ------------------------------------------------------------
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------

gi|15673378      ------------------------------------------------------------
gi|15927540      ------------------------------------------------------------
gi|15595102      ------------------------------------------------------------
gi|18203008      ------------------------------------------------------------
gi|16078742      ------------------------------------------------------------
gi|15895087      ------------------------------------------------------------
gi|18202351      ------------------------------------------------------------
gi|15837113      KQQRKPNVVRGLFSDVVDSPRLKENAWYTFADPTLEPVIEVAFLNGVHTPTLEQDTNFRT 725
gi|13477031      ------------------------------------------------------------
gi|15597385      ------------------------------------------------------------
gi|13540860      ------------------------------------------------------------

gi|15673378      ---------------------------
gi|15927540      ---------------------------
gi|15595102      ---------------------------
gi|18203008      ---------------------------
gi|16078742      ---------------------------
gi|15895087      ---------------------------
gi|18202351      ---------------------------
gi|15837113      DGLSWKVVHRYGVAAVGWRGATRTSGT 752
gi|13477031      ---------------------------
gi|15597385      ---------------------------
gi|13540860      ---------------------------
```

Figure 3.43 : Multiple sequence alignement of ClpP protease subunit of PSP.

[gi|13540860|*Thermoplasma volcanium* GSS1 Secreted serine protease (ClpP class)(PSP); gi|13477031|*Mesorhizobium loti* MAFF303099 similar to ClpP protease; gi|15595102|*Borrelia burgdorferi* B31 ATP-dependent Clp protease proteolytic component (clpP-2); gi|16078742|*Bacillus subtilis* alternate gene name: ylxI; gi|15895087|*Clostridium acetobutylicum* ATCC 824-Periplasmic serine protease; gi|15673378|*Lactococcus lactis subsp. lactis* Il1403 ATP dependent Clp protease; gi|15597385|*Pseudomonas aeruginosa* PAO1 hypothetical protein PA2189; gi|18203008|*Pseudomonas aeruginosa* CLPP2_PSEAE ATPdependent Clp protease proteolytic subunit 2; gi|15927540|*Staphylococcus aureus subsp. aureus* N315 hypothetical protein SA1775; gi|15837113|*Xylella fastidiosa 9a5c* hypothetical protein XF0511; gi|18202351|*Synechocystis sp.* PCC 6803 Putative ATP-dependent Clp protease proteolytic subunit-like protein.]

Table 3.4 : Scores of multiple sequence alignments of PSP gene

	1	2	3	4	5	6	7	8	9	10	11
1		10	12	15	14	11	10	14	7	6	5
2	10		17	7	9	25	17	16	20	24	13
3	12	17		5	17	18	10	35	16	15	29
4	15	7	5		51	8	9	13	11	9	7
5	14	9	17	51		8	8	13	11	7	12
6	11	25	18	8	8		9	16	31	25	13
7	10	17	10	9	8	9		14	8	19	11
8	14	16	35	13	13	16	14		21	18	26
9	7	20	16	11	11	31	8	21		21	12
10	6	24	15	9	7	25	19	18	21		14
11	5	13	29	7	12	13	11	26	12	14	

1. **gi|13540860|***Thermoplasma volcanium* GSS1| Secreted serine protease (ClpP class)
2. **gi|13477031|***Mesorhizobium loti* MAFF303099| similar to ClpP protease
3. **gi|15595102|***Borrelia burgdorferi* B31| ATP-dependent Clp protease proteolytic component (clpP-2)
4. **gi|16078742|***Bacillus subtilis*| alternate gene name: ylxI
5. **gi|15895087|***Clostridium acetobutylicum* ATCC 824| Periplasmic serine protease, YMFB *B.subtilis* ortholog
6. **gi|15673378|***Lactococcus lactis subsp. lactis* Il1403| ATP dependent Clp protease
7. **gi|15597385|***Pseudomonas aeruginosa* PAO1| hypothetical protein PA2189
8. **gi|18203008|***Pseudomonas aeruginosa* |CLPP2_PSEAE ATP-dependent Clp protease proteolytic subunit 2 (Endopeptidase Clp 2)
9. **gi|15927540|***Staphylococcus aureus subsp. aureus* N315| hypothetical protein SA1775
10. **gi|15837113|***Xylella fastidiosa 9a5c* | hypothetical protein XF0511
11. **gi|18202351|***Synechocystis sp.* PCC 6803| Putative ATP-dependent Clp protease proteolytic subunit-like (Endopeptidase Clp-like)

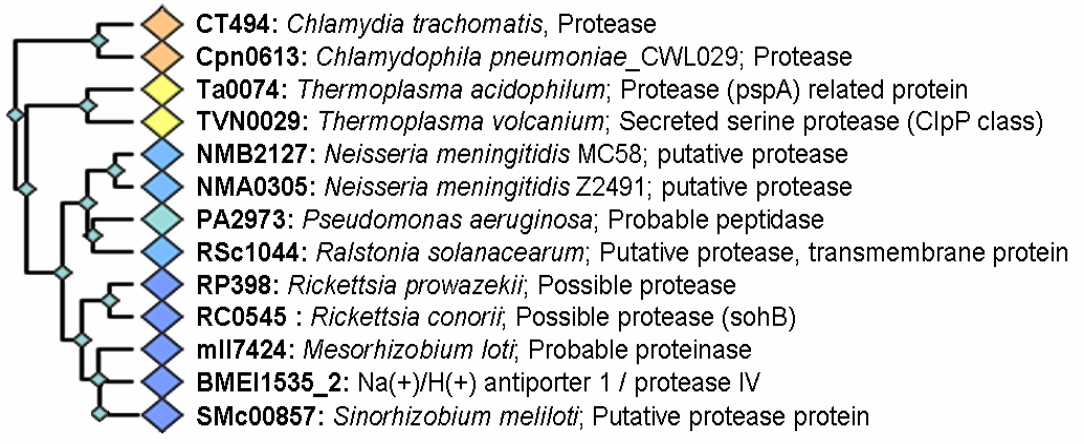

Figure 3.44 : Phylogram of periplasmic serine protease with other relative members of COG0616 (Periplasmic serine proteases – ClpP class)

The data derived from NCBI domain search for *Tp. volcanium* MSP showed that, this protease has ClpP protease, Signal peptide peptidase A (SppA) and Nodulation efficiency protein D (NfeD) domains (Figure 3.45).

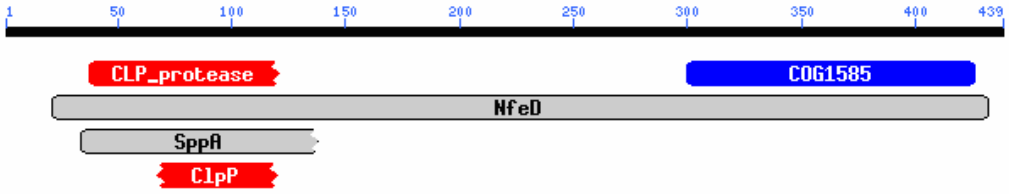

Figure 3.45: Domains of *Tp. volcanium* MSP.

This search indicated that MSP produces significant alignments with NfeD proteins, which are involved in the nodulation efficiency and competitiveness of the *Sinorhizobium meliloti* strain GRY on alfalfa roots. The specific function

of this family is unknown although it is unlikely that the NfeD is specifically involved in nodulation as the family contains several different archaeal and bacterial species most of which are not symbionts. The phylogram based on NfeD domain is given in Figure 3.46.

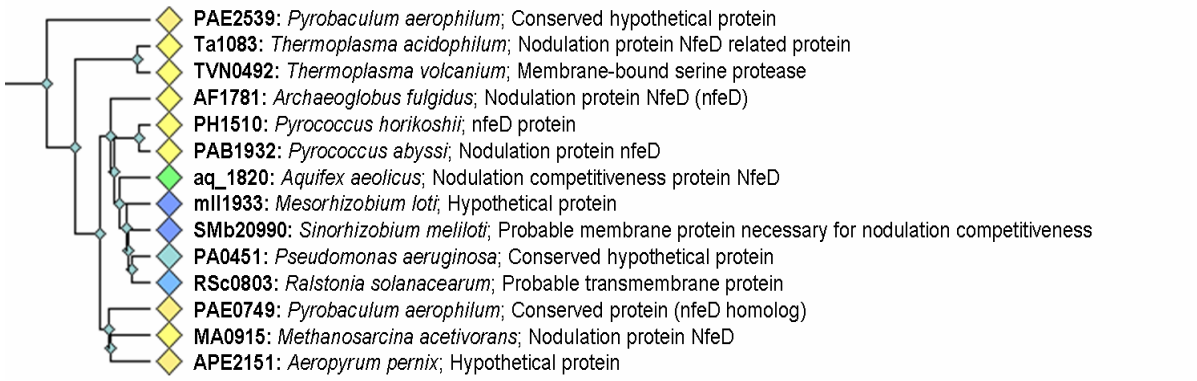

Figure 3.46 : Phylogram for *Tp. volcanium* MSP depending on its NfeD domain.

Clp protease domain, which was found in N-terminal part of the *Tp. volcanium* membrane bound protease, was found to be related to endopeptidase Clps of various G(+) bacteria (*B. halodurans, Listeria innocua, S. aureus, C. acetobutylicum*, *Synechococcus elongatus*) and *Streptomyces coelicolor* as revealed by multiple sequence alignment (Figure 3.47; Figure 4.48).

Although homology between ATP dependent ClpPs of various Gram (+) bacteria is very high (41% to 78%) (Table 3.5), *Tp. volcanium* MSP was recognized as a distinct lineage.

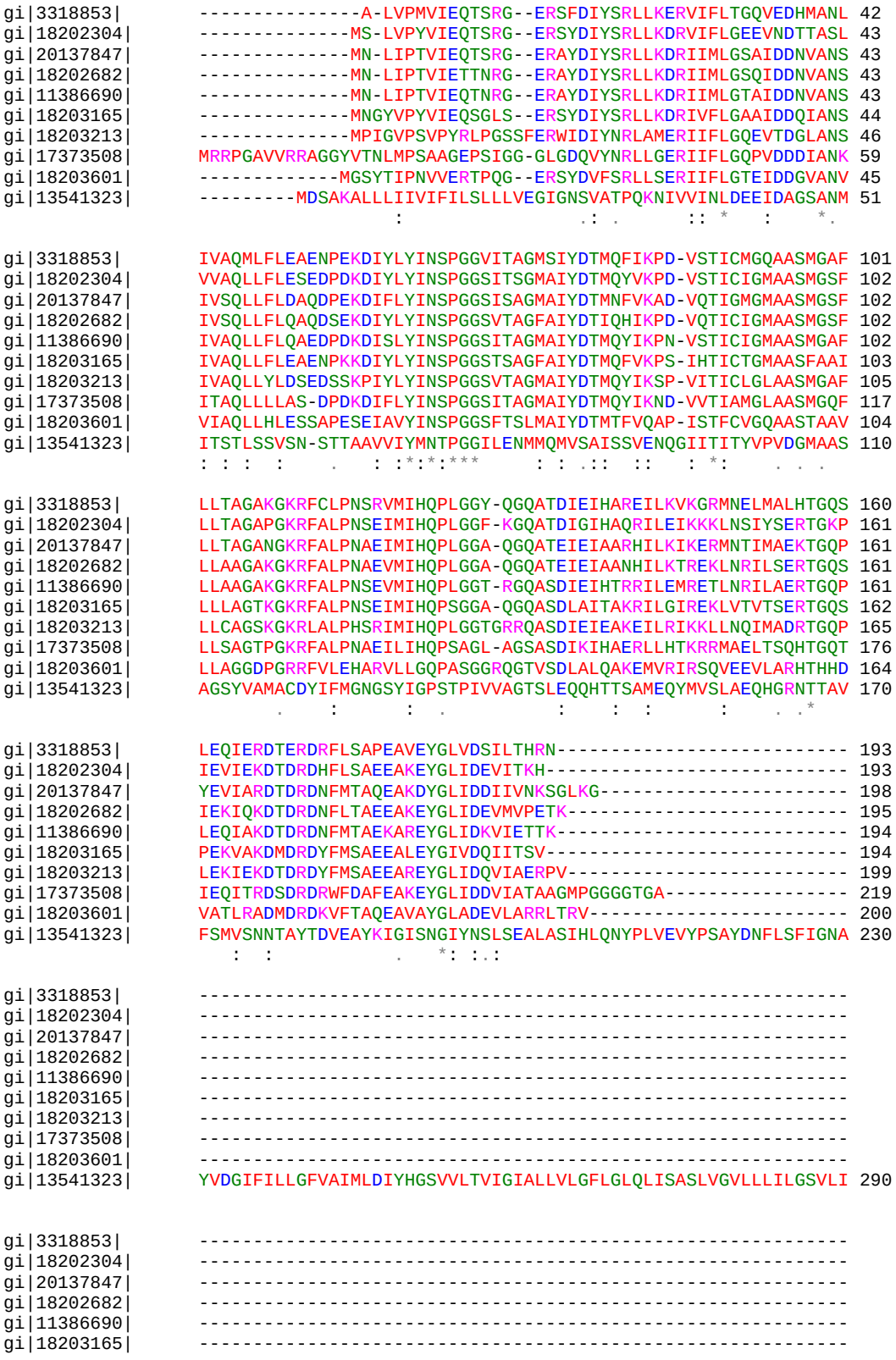

Figure 3.47 Continued

```
gi|18203213|      ------------------------------------------------------------
gi|17373508|      ------------------------------------------------------------
gi|18203601|      ------------------------------------------------------------
gi|13541323|      LLEAKMGHGFALLSGVVIGLVGTFMLASPYYSSNPGYSPSPFTTFDLLTSILIVIVAGFL 350

gi|3318853|       ------------------------------------------------------------
gi|18202304|      ------------------------------------------------------------
gi|20137847|      ------------------------------------------------------------
gi|18202682|      ------------------------------------------------------------
gi|11386690|      ------------------------------------------------------------
gi|18203165|      ------------------------------------------------------------
gi|18203213|      ------------------------------------------------------------
gi|17373508|      ------------------------------------------------------------
gi|18203601|      ------------------------------------------------------------
gi|13541323|      AFYIRRIVRTIKFRGRWTGAESLIGRSAKAVSNINSHGWVSVDGIEWQAKSNDGKPINKG 410

gi|3318853|       ----------------------------
gi|18202304|      ----------------------------
gi|20137847|      ----------------------------
gi|18202682|      ----------------------------
gi|11386690|      ----------------------------
gi|18203165|      ----------------------------
gi|18203213|      ----------------------------
gi|17373508|      ----------------------------
gi|18203601|      ----------------------------
gi|13541323|      EPVKIVDRSGLVLIVERTNTEEKQDLKLK 439
```

Figure 3.47: Multiple sequence alignement of Clp protease domain of MPS gene derived by Clustal W (1.82) program.

>gi|13541323| *Thermoplasma volcanium* GSS1| Membrane-bound serine protease
>gi|3318853| *Escherichia coli* | Chain A, The Structure of ClpP
>gi|17373508| *Streptomyces coelicolor* | ATP-dependent Clp protease proteolytic subunit 1 (Endopeptidase Clp 1)
>gi|18203165| *Bacillus halodurans* | ATP-dependent Clp protease proteolytic subunit 2 (Endopeptidase Clp 2)
>gi|20137847| *Listeria innocua* | ATP-dependent Clp protease proteolytic subunit (Endopeptidase Clp)
>gi|11386690| *Bacillus halodurans* | ATP-dependent Clp protease proteolytic subunit 1 (Endopeptidase Clp 1)
>gi|18202682| *Staphylococcus aureus subsp. aureus* Mu50| ATP-dependent Clp protease proteolytic subunit (Endopeptidase Clp)
>gi|18202304| *Clostridium acetobutylicum* | ATP-dependent Clp protease proteolytic subunit (Endopeptidase Clp)
>gi|18203213| *Synechococcus elongatus* PCC 7942| ATP-dependent Clp protease proteolytic subunit 3 (Endopeptidase Clp 3)
>gi|18203601| *Streptomyces coelicolor* | Putative ATP-dependent Clp protease proteolytic subunit-like (Endopeptidase Clp-like)

Table 3.5 : Scores of multiple sequence alignments of MSP gene.

	1	2	3	4	5	6	7	8	9	10
1		9	12	9	12	10	9	15	15	13
2	9		55	59	66	66	63	68	59	45
3	12	55		50	56	56	53	57	52	40
4	9	59	50		62	65	63	64	56	46
5	12	66	56	62		74	77	68	55	43
6	10	66	56	65	74		78	73	64	45
7	9	63	53	63	77	78		70	58	43
8	15	68	57	64	68	73	70		62	45
9	15	59	52	56	55	64	58	62		41
10	13	45	40	46	43	45	43	45	41	

1. **gi|13541323|** *Thermoplasma volcanium* GSS1| Membrane-bound serine protease
2. **gi|3318853|** *Escherichia coli* | Chain A, The Structure of ClpP
3. **gi|17373508|** *Streptomyces coelicolor* | ATP-dependent Clp protease proteolytic subunit 1 (Endopeptidase Clp 1)
4. **gi|18203165|** *Bacillus halodurans* | ATP-dependent Clp protease proteolytic subunit 2 (Endopeptidase Clp 2)
5. **gi|20137847|** *Listeria innocua* | ATP-dependent Clp protease proteolytic subunit (Endopeptidase Clp)
6. **gi|11386690|** *Bacillus halodurans* | ATP-dependent Clp protease proteolytic subunit 1 (Endopeptidase Clp 1)
7. **gi|18202682|** *Staphylococcus aureus subsp. aureus* Mu50| ATP-dependent Clp protease proteolytic subunit (Endopeptidase Clp)
8. **gi|18202304|** *Clostridium acetobutylicum* | ATP-dependent Clp protease proteolytic subunit (Endopeptidase Clp)
9. **gi|18203213|** *Synechococcus elongatus* PCC 7942| ATP-dependent Clp protease proteolytic subunit 3 (Endopeptidase Clp 3)
10. **gi|18203601|** *Streptomyces coelicolor* | Putative ATP-dependent Clp protease proteolytic subunit-like (Endopeptidase Clp-like)

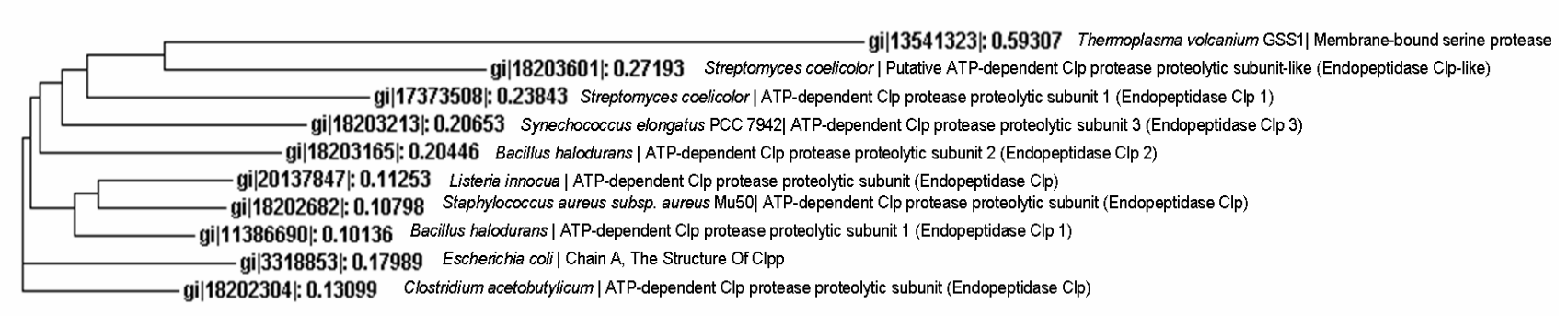

Figure 3.48 : Phylogram of *Tp. volcanium* membrane-bound serine protease depending on its Clp domain which shows relationship with a number of bacterial Clp Proteases. (Phylogram was derived by using Clustal W -1.82)

3.5.3 Signal Peptide Search

Protein translocation across the bacterial plasma membrane and the eukaryal endoplasmic reticulum membrane relies on cleavable N-terminal signal peptides. Although the signal peptides of secreted proteins in *Bacteria* and *Eukarya* have been extensively studied at the sequence, structure, and functional levels, little is known of the nature of archaeal signal peptides. Accordingly, analysis of archaeal signal peptides would fail to detect the presence of uniquely archaeal putative signal peptides, that is, those that do not resemble either their eukaryal or bacterial counterparts. Because of the uncertainty concerning the actual cleavage sites used in *Archaea*, non secreted proteins would be erroneously included in the lists of secreted proteins, or vice versa.

Although classified as periplasmic serine protease, signal peptide search using Signal P 3.0 server based on Gram (+) (Figure 3.49), Gram (-) (Figure 3.50) and eukaryotic networks (Figure 3.51) (Hidden Markov Model and Neural Networks) predicted this protease as a non-secretory protein due to absence of signal sequence.

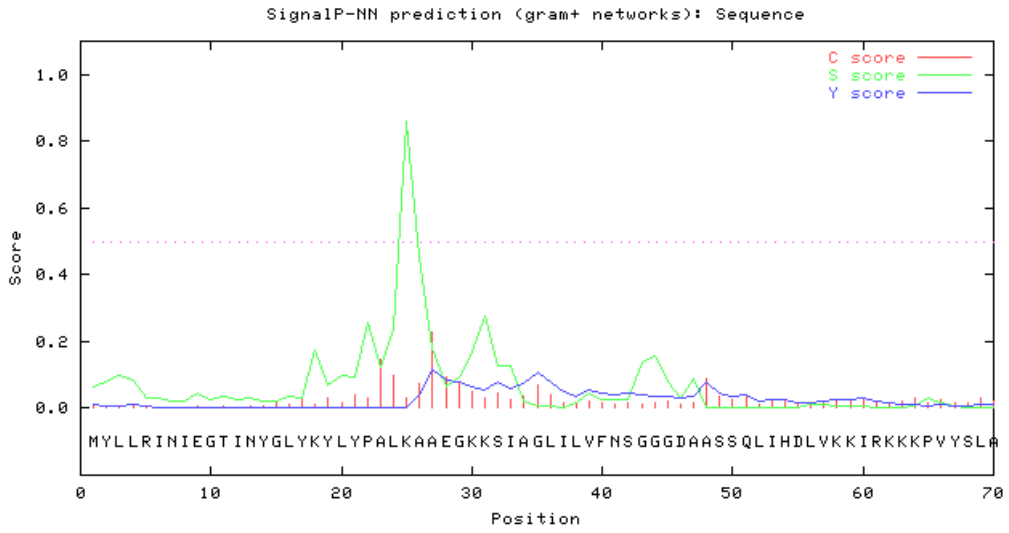

Figure 3.49 : Signal peptide prediction results for PSP gene derived by Gram (+) bacterial networks.

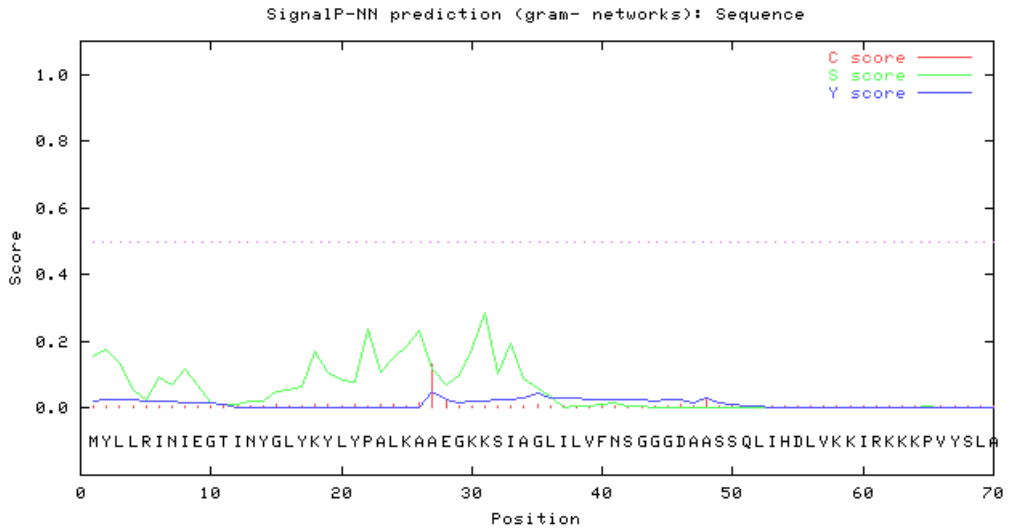

Figure 3.50 : Signal peptide prediction results for PSP gene derived by Gram (-) bacterial networks.

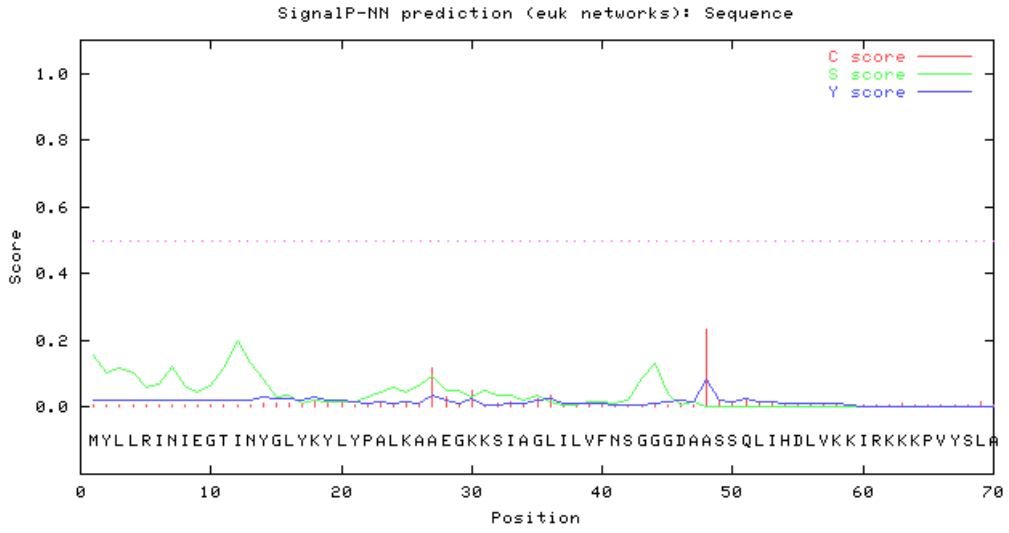

Figure 3.51 : Signal peptide prediction results for PSP gene derived by Eukaryotic networks.

The search for MSP revealed the presence of signal peptide sequence homologous to that of Gram (+) bacteria with 95.9% probability and a cleavage site between position 27 and 28 (Figure 3.52). Signal peptide search based on eukaryotic and Gram (-) networks also implicated the existence of the homologous sequences with a probability of 0.878 (Cleavage site between position 23 and 24) (Figure 3.53) and 0.576 (Cleavage site between positions 29 and 30) (Figure 3.54), respectively.

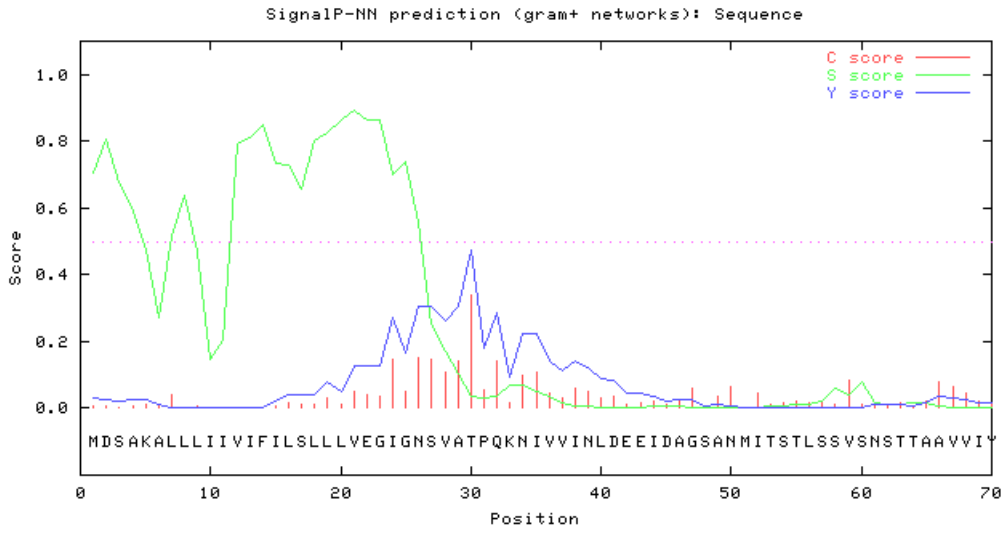

Figure 3.52 : Signal peptide prediction results for MSP gene derived by Gram (+) bacterial networks.

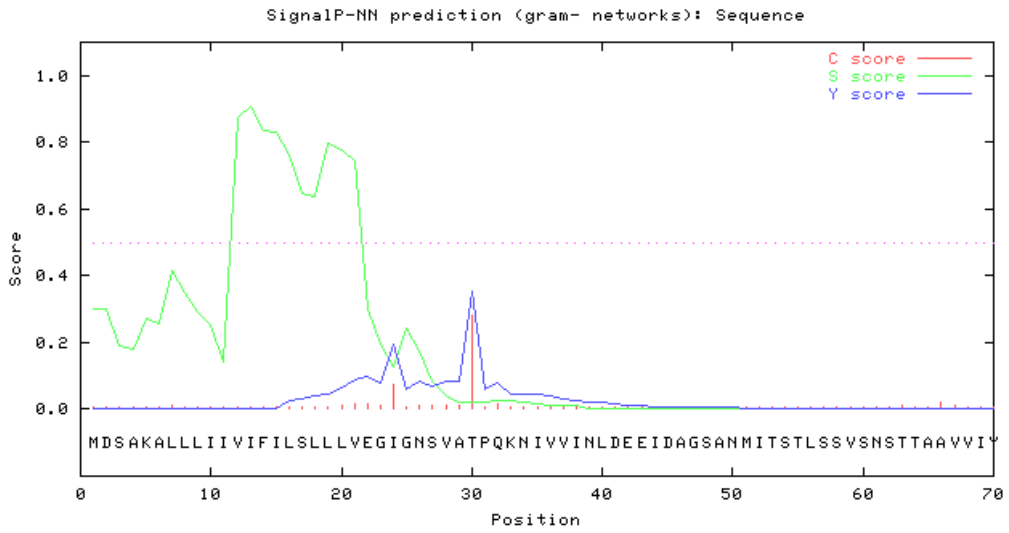

Figure 3.53 : Signal peptide prediction results for MSP gene derived by Gram (-) bacterial networks.

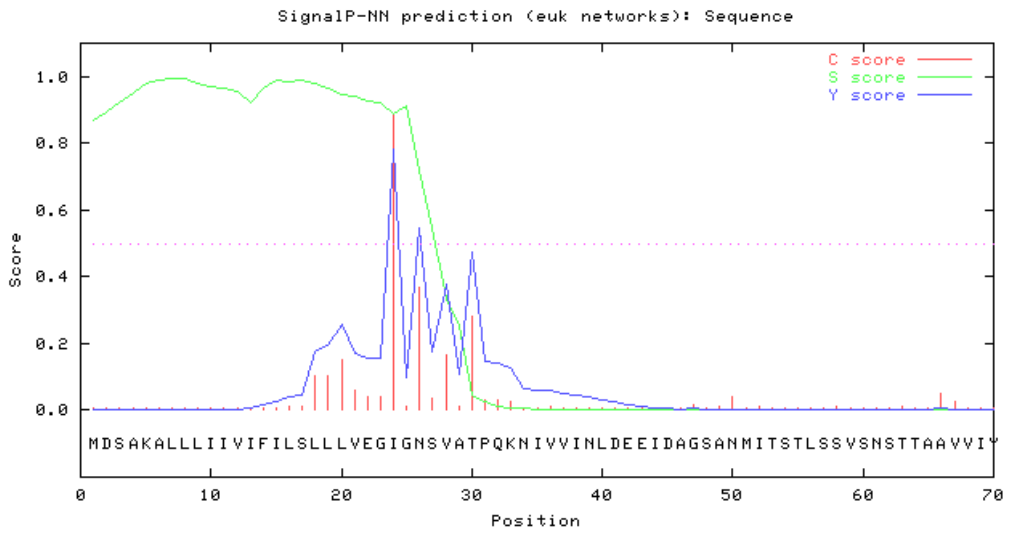

Figure 3.54 : Signal peptide prediction results for MSP gene derived by Eukaryotic networks.

CHAPTER 4

DISCUSSION

Recent years have witnessed a number of significant advances in the structural biology of proteases, including aspects of intracellular protein and peptide degradation by self-compartmentalizing proteases, activation of proteases in proteolytic cascades of regulatory pathways, and mechanisms of microbial proteases in pathogenicity and stress conditions. Proteases are no longer viewed as non-specific degradative enzymes associated solely with protein catabolism, but are known to be intimately involved in controlling biological processes that span life to death. They maintain this exquisite control by catalyzing the precisely-timed and rapid turnover of key regulatory proteins. Proteases also interplay with chaperones to ensure protein quality and to readjust the composition of the proteome following stress. Although much less is known about proteolysis in the context of an archaeal cell than that of bacteria and eukaryotes, recent advances in genomics and biochemistry provide unprecedented insight into the complete repertoire of archaeal proteases that might regulate cell function (Maupin-Furlow, *et al*, 2005; Rother, *et al*, 2005).

Rapid proteolysis plays a major role in post-translational cellular control by the targeted degradation of short-lived regulatory proteins and also serves an important function in protein quality control by eliminating defective and potentially damaging proteins from the cell. In all cells, protein degradation is predominantly carried out by ATP-dependent proteases, which are complex

enzymes containing both ATPase and proteolytic activities expressed as separate domains within a single polypeptide chain or as individual subunits in complex assemblies. Five ATP-dependent proteases, Lon, FtsH, ClpAP, ClpXP, and HslUV, have been discovered in *Escherichia coli*, and homologous proteases have been found in all eubacteria and in many eukaryotes (Botos, *et al*, 2005).

Energy-dependent proteolysis plays a major role during heat shock in bacteria in which genes encoding ATP-dependent proteases, such as *lon*, *clp*, and *hfl*, are linked to heat shock promoters. However, based on available genome sequence data, hyperthermophilic archaea lack the Clp and HflB (FtsH) family of proteins and have a different version of the Lon protease. Hyperthermophilic archaea, which typically have proteasomes, lack the eukaryotic ubiquitination pathway for selective protein degradation by the proteasome and, therefore, seem to modulate proteolysis at the protease level (Shockley, *et al*, 2003).

The enzymes isolated from thermophilic microorganisms, especially produced in heterologous expression systems such as *E. coli*, have been recognized to be of high interest both for clarifying the molecular mechanisms at the base of their stability and activity at high temperature, and for their peculiar properties, such as resistance to organic solvents, detergents, etc., that are of interest for potential biotechnological applications. For instance, Catara, *et al*, reported the purification to homogeneity and characterization of a novel extracellular protease isolated from *Aeropyrum pernix* K1 which is a strictly aerobic hyperthermophile, growing optimally above 90 ºC. The protease, designated as Pernisine, is an endopeptidase belonging to the family of subtilisin like serine proteases, stable at high temperature, and active in a broad range of pH values. Its gene, identified from the sequenced genome of *A. pernix* K1 (Kawarabayasi, *et al*, 1999), has been expressed in *Escherichia coli*, and the recombinant enzyme has been purified and characterized (Catara, *et al*, 2003).

In our study, we have used a PCR based strategy for cloning both periplasmic (ClpP-like) (PSP) and membrane-bound serine protease (MSP) genes of *Thermoplasma volcanium* GSS1. Forward and reverse degenerate primers were designed considering the codon usage of *Thermoplasma volcanium* GSS1 to be able to amplify the desired genes. Forward primers fit the highly conserved regions in the N-terminals of the genes and reverse primers fit the chromosomal DNA of *Thermoplasma volcanium* GSS1, which has been sequenced and released to genomic data bases (Kawashima, *et al*, 2000). The amplification of the PSP gene resulted in a fragment with the length of 971 bp while the amplification of the MSP gene resulted in a fragment of 1521 bp. Obtained PCR products were cloned into pDrive cloning vector by AU ligation. Putative *E. coli* TG-S1, *E. coli* TG-S4 and *E. coli* TG-M1 recombinant clones were selected and analyzed by restriction digestions and it was confirmed that both of the genes were located in reverse direction relative to the *lacZ* promoter of the pDrive vector. This data helped us to select the restriction enzymes in order to excise the gene fragment from the recombinant pDrive vector (pDrive-S1) and ligate it into the pQE expression vector in correct positions to be able to locate the 6xHis tag at N-terminus of the gene during expression experiments.

Intracellular and extracellular serine protease activities of the recombinant enzymes were determined spectrophotometrically by using synthetic chromogenic substrates. Cell free extracts derived from the recombinant colonies which include PSP and MSP genes showed higher activities (Intracellular activities: *E.coli* TG-M1 = 137 x $10^{-3}\Delta Abs_{410}xmin^{-1}$; *E. coli* TG-S1 = 110 x 10^{-3} $\Delta Abs_{410}xmin^{-1}$ and *E. coli* TG-S4 = 84 x 10^{-3} $\Delta Abs_{410}xmin^{-1}$) as compared to the wild-type *E. coli* TG1 ($\Delta Abs_{410}xmin^{-1}$=6,2 x 10^{-3}) strain. This result indicated expression of the cloned genes under control of their own promoters, since the genes had been determined to be localized in reverse direction relative to the *lacZ* promoter.

Also, it has been observed that the activities of cell free extracts were higher than that of the supernatants for both enzymes. This comparison represents the fact that both of the enzymes are intracellular such as other ClpP like proteases.

Currently, the most attractive method available for cloning purifying recombinant thermophile proteins at the lab scale involves the use of protein/polypeptide affinity tags which are cloned along with the protein of interest in the form of a genetic fusion (e.g., with a GST tag, or a 6x His tag, fused at the N-terminus, or C-terminus, of the protein of interest). Importantly, these tags can serve a dual purpose. Besides their obvious utility in affinity purification strategies, these tags [especially the larger ones, such as the GST tag and also the maltose-binding protein (MBP) tag] are found to increase the probability of synthesis and folding of the protein of interest in the *E. coli* cytoplasm. Thus, even notwithstanding their utility in regard to purification strategies, affinity fusion tags facilitate protein synthesis and folding (Mukherjee, *et al*, 2005).

Recombinant PSP enzyme gene was selected for expression using His-tag strategy. This gene was excised from the pDrive vectors by using convenient restriction enzymes in order to locate the 6xHis tag at the 5´ end of the cloning gene and ligated into pQE vectors. Generally, if the insert is to be cloned into a single restriction site, it is especially important to dephosphorylate the vector ends after digestion (Sambrook, *et al*, 1995). In vectors cut with two enzymes, when the sites are close together, dephosphorylation decreases the nonrecombinant background caused by incomplete digestion with one of the enzymes. Since our restriction sites for *SalI* and *PstI* enzymes were next to each other, we did not need to dephosphorilate the vector ends to avoid from this undesired background.

Most systems use *E. coli* as an expression host due to the convenience and

economy of working with bacterial cultures. Generally, QIA*express* pQE vectors and constructs can be maintained in any *E. coli* strain that is ampicillin-sensitive or harbors the *lacI*q gene on the F-factor episome (The QIA*expressionist*™ handbook for high-level expression and purification of 6xHis-tagged proteins). XL1-Blue series contain an episomal copy of *lacI*q. *lacI*q is a mutation of *lacI* that produces very high levels of the *lac* repressor. Therefore, initial cloning and propagation using XL1-Blue is recommended by the manufacturer. In spite of this, our cloning trials were not successful when we used XL1-Blue as a host, since the transformation efficiency was very low (10^3 transformants/µg DNA).

Expression vectors such as the pQE plasmids do not allow color selection of clones that contain plasmids with inserts. Therefore, positive expression clones were detected directly on colony blots using Anti-His HRP conjugate and DAB substrate. This strategy depends on the affinity of 6xHis tagged recombinant proteins to specific antibodies. 6His-tags are a popular choice due to their small size and resultant lack of effect on the physical and biological properties of the expressed protein (Kuznetsova, *et al*, 2005). Several antibodies and chelator/enzyme conjugates have been reported to detect histidine-tagged proteins (Field, *et al*, 1988). However these are restricted by requiring additional amino acids adjacent to the histidine sequence (requiring specific expression vectors) or recognize histidine-tagged proteins with only moderate affinity. Anti-His HRP conjugate is a monoclonal antibody directly conjugated to horseradish peroxidase which allows specific and sensitive detection of histidine-tagged proteins irrespective of the expression system used. Therefore, we used colony blotting procedure which was recommended by the manufacturer to select the recombinant clones which include 6xHis tagged PSP gene.

Colonies which gave putative positive signals in colony blotting experiments

were further analyzed by restriction enzyme digestions. Only one colony which was named as *E. coli* TG-S1.7 was confirmed to be a true recombinant. This colony included the recombinant pQE 30 plasmid. Other samples which included the pQE32 vectors were identified as non-recombinants. This result was most probably due to frame shifting.

Enzymatic activity analyses of the cell free extracts derived from recombinant *E. coli* TG-pQE30-S1.7 cells showed that this enzyme has a significant serine protease activity which was 5.7-fold higher than the control samples derived from *E. coli* TG1 cells. Higher activity with the substrate N-Suc-Ala-Ala-Pro-Phe-pNA, typically for serine protease, suggested that the enzyme preferably cleaves hydrophobic residues at the P1 site (nomenclature according to Schechter and Berger 1967), similarly to other subtilisin-like serine protease enzymes. The enzyme also hydrolyzed N-Suc-Ala-Ala-Pro-Phe-pNA which is also a substrate specific for serine proteases.

Also, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out according to Laemmli (1970) using 12.5% polyacrylamide gel. Cell free extracts of this recombinant strain revealed a strong protein band with a molecular mass of approximately 25 kDa which is equal to the estimated molecular weight of the enzyme encoded by PSP gene (25 585 Dalton) revealed by NCBI and KEGG databases search. This result might indicate the expression of the PSP fusion protein in *E. coli* in consistant with the activity data.

Energy-dependent proteolysis plays a major role during heat shock in bacteria in which genes encoding ATP-dependent proteases, such as *lon*, *clp*, and *hfl*, are linked to heat shock promoters. The ClpP subunit of the Clp protease complex was recently identified in *L. lactis* and found to be required for survival at high temperatures and its transcription was induced by both heat and acid shock (Frees and Ingmer, 1999). The *clpB*, *clpC*, and *clpE* ATPase genes

have also been identified in *L. lactis* (Ingmer, *et al*, 1999). Ingmer, *et al*, (1999) showed that *L. lactis clpB* and *clpE* transcription was induced at least 10-fold by heat shock.

Skinner, *et al*, cloned and sequenced the *L. lactis clpX* gene and surrounding 5′ and 3′ sequences by using a PCR based technique. PCR amplification from *L. lactis* MG1363 chromosomal DNA resulted in a 600-bp product with high homology to other *clpX* genes. The entire *clpX* gene and surrounding sequence from *L. lactis* was then cloned. The open reading frame (ORF) for ClpX starts with an ATG codon and encodes a protein consisting of 411 AA residues and has a calculated molecular mass of 45.8 kDa with a predicted pI of 4.86.

Thermal stress was apparent by the effects on growth and the induction of known and putative stress genes present in *P. furiosus* which is a hyperthermophilic archaeon (Shockley, *et al*, 2003). The given results were based on the combined effects of thermal stress and reduced cell growth that are collectively described by the term heat shock. Based on available genome sequence data, hyperthermophilic archaea lack the Clp and HflB (FtsH) family of proteins and have a different version of the Lon protease (Ward, *et al*, 2002). The genes encoding the major Hsp60-like chaperonin (thermosome) in *P. furiosus* and the Hsp20-like small heat shock protein were found to be strongly induced, as were two other molecular chaperones (VAT) belonging to the CDC48/p97 branch of the AAA^+ family.

We focused on heat-shock response of *Thermoplasma volcanium* to decipher the roles of our proteins of interest during heat stress conditions, since they include putative ortologs of ClpP proteases, which functions in response to the heat shock. After heat-shock, there was 6.5 fold and 1.3 fold increase in the serine protease-like activities of cell free extract and supernatant, respectively. This might be a result of induction of the heat-shock related serine proteases.

The research to understand the roles of PSP and MSP in heat shock response of *Tp. volcanium* is underway in our laboratory.

Several sequence-based algorithms have been developed to identify homologous proteins in databases (Gille, *et al*, 2003). The deduced amino acid sequences were compared to sequences deposited in a database. The results of the alignment of the Clp protease domain, which was found in N-terminal part of the *Tp. volcanium* membrane bound protease revealed that, it is related to endopeptidase Clps of various G(+) bacteria (*B. halodurans, Listeria innocua, S. aureus, C. acetobutylicum*, *Synechococcus elongatus* and *Streptomyces coelicolor* .) This result may be in consistant with the given support to a phylogenetic classification that places the archaea closer to gram-positive bacteria than to eukaryotes (Gupta, 1998), although the evolutionary and functional significance of the sequence gap in archaea and gram-positive bacteria has not been elucidated.

The homology is quite high between *Tp. volcanium* periplasmic serine protease and members of MEROPS peptidase family S49 (protease IV family, clan S). On the other hand, Clp domain (COG0740) of this protein is found to be closely related to a number of bacterial ATP-dependent ClpP Protease proteolytic subunits.

The presequence is a signal peptide that leads the secretion of proteins across the cytoplasmic membrane, while the prosequence is believed to act as an intramolecular chaperone that guides the correct folding of the protein, and is then removed by autoproteolysis from the active enzymatic form (Catara, *et al*, 2003). A search using Signal P 3.0 world wide server revealed the presence of signal peptide sequence homologous to that of Gram (+) bacteria with 95.9% probability and a cleavage site between 27 and 28^{th} position of *Tp. volcanium* MSP. Signal peptide search based on eukaryotic and Gram (-) networks also

implicated the existence of homologous sequences with a probability of 0.878 (Cleavage site between position 23 and 24) and 0.576 (Cleavage site between positions 29 and 30), respectively. On the other hand, our intracellular and extracellular enzymatic activity analyzes showed that MSP is an intracellular enzyme.

Although classified as secreted serine protease in NCBI database, signal peptide search based on Gram (+), Gram (-) and eukaryotic networks (Hidden Markov and Neural Networks) predicted PSP protease as a non-secretory protein due to absence of signal sequence. This data is in consistant with the enzymatic activity results which showed that PSP is intracellular.

Analysis of archaeal signal peptides would fail to detect the presence or absence of uniquely archaeal putative signal peptides since the sequence, structure, and functional levels of signal peptides of secreted proteins in *Archaea* have been extensively studied. Therefore, non secreted proteins would be erroneously included in the lists of secreted proteins, or vice versa.

CHAPTER 5

CONCLUSION

1. *Thermoplasma volcanium* genome contains the gene sequences encoding periplasmic (ClpP like) (PSP) (TVN0029) and membrane-bound serine protease (MSP) (TVN0492) enzymes which include ClpP domains.

2. A unique 971 bp PCR fragment containing the PSP gene of *Tp. volcanium* was amplified with the degenerate primer set (FP1 and RP1) at an optimum annealing temperature of 55ºC and 1.5 mM $MgCl_2$ concentration. This fragment was cloned in *E. coli* by using a pDrive PCR cloning kit. The cloning was confirmed by restriction enzyme digestions. The recombinant PSP gene was successfully expressed, under control of its own promoter, in *E. coli*, and associated serine protease activity in the cell free extract was 110 x 10^{-3} $\Delta Abs_{410} x min^{-1}$.

3. A unique 1521 bp PCR fragment containing the MSP gene of *Tp. volcanium* was amplified with the degenerate primer set (FP2 and RP2) at an optimum annealing temperature of 55ºC and 1.5 mM $MgCl_2$ concentration. This fragment was cloned in *E. coli* by using a PCR based strategy. PCR cloning was confirmed by restriction enzyme digestions of plasmids isolated from recombinant colonies and serine protease activity assays. Expression of the recombinant MSP gene in pDrive cloning vector was achieved in *E. coli* under the control of its own promoter and serine protease activity measured in the cell free extract was 137 x 10^{-3} $\Delta Abs_{410} x min^{-1}$.

4. Expression of the cloned PSP gene as 6xHis tag fusion was achieved using QIAExpress® Type IV Kit. PSP gene fragment was excised and ligated to pQE cloning vectors. Transformed colonies were selected by colony blotting and hybridization. The recombinant *E. coli* TG-pQE30-S1.7 strain carried PSP gene as a 6xHis tag fusion, in pQE30 expression vector. Recombinant colonies were further analyzed by restriction enzyme digestions. The serine protease activity related to this fusion protein was spectrophotometrically determined as $46.2x10^{-3}$ $\Delta Abs_{410}xmin^{-1}$.

5. The SDS-PAGE profile of the heat treated cell free extract of the *E. coli* TG-pQE30-S1.7 revealed the presence of a protein band with an estimated MW of 25kDa, which is consistent with the molecular mass of the ClpP's reported so far. Serine protease activity assays of crude cell free extracts showed the expression of the 6xHis-PSP fusion enzyme.

Purification of the expressed recombinant enzymes by column chromatography including Ni-NTA affinity matrices is still in progress in our laboratory.

6. To determine the heat shock response of *Tp. volcanium*, two different experiments at different temperatures were carried out. Exposure to the heat shock at 72ºC for 2 h resulted in an increase in growth rate of *Tp. volcanium*, while the heat shock at 80ºC for 2 h significantly decreased the growth rate. Studies on possible involvements of PSP, MSP and other heat-shock related proteins in heat shock response of *Tp. volcanium* is also underway in our laboratory.

7. Clusters of orthologous groups and conserved domain database searches indicated the relationships between PSP and MPS genes and other ClpPs from various bacteria.

REFERENCES

1. Amerik A.Y., Antonov V.K., Gorbalenya A.E., Kotova S.A., Rotanova T. V., and Shimbarevich E. V.,1991, FEBS Lett., 287, 211–214.

2. Amerik A.Y., Antonov V.K., Ostroumova N.I., Rotanova T.V., and Chistyakova L.G., 1990, Cloning, structure and expression of the full-size lon gene in *Escherichia coli* coding for ATPdependent La-proteinase. Bioorg. Khim., 16, 869–880.

3. Barett A.J. 1994. Proteolytic enzymes: serine and cysteine peptidases, Methods Enzymol., 244, 1-15.

4. Barrett A.J., Rawlings N.D., O'Brien E.A., 2001, The MEROPS database as a protease information system, J. Struct. Biol., 134, 95–102.

5. Besche H., Tamura N., Tamura T., Zwickl P., 2004, Mutational analysis of conserved AAA^{+} residues in the archaeal Lon protease from *Thermoplasma acidophilum*, FEBS Letters 574,161–166.

6. Bhoola K., Ramsaroop R., Plendl J., Cassim B., Dlamini Z., Naicker S., 2001, Kallikrein and kinin receptor expression in inflammation and cancer, Biol. Chem., 382, 77-89.

7. Botos I, Melnikov E.E., Cherry S., Kozlov S., Makhovskaya O.V., Tropea J.E., Gustchina A., Rotanova T.V., and Wlodawer A, 2005, Atomic-resolution Crystal Structure of the Proteolytic Domain of *Archaeoglobus fulgidus* Lon Reveals the Conformational Variability in the Active Sites of Lon Proteases, J. Mol. Biol., 351, 144–157.

8. Bukau B., 1993. Regulation of the *Escherichia coli* heat-shock response, Mol. Microbiol., 9, 4, 671–680.

9. Cashikar A.G., Schirmer E.C. Hattendorf D.A., Glover J.R., Ramakrishnan M.S., Ware D.M., Lindquist S.L., 2002, Defining a pathway of communication from the C-terminal peptide binding domain to the N-terminal ATPase domain in a AAA protein, Mol. Cell, 9, 751–760.

10. Chandu D., Nandi D., 2004, Research in Microbiology, Comparative genomics and functional roles of the ATP-dependent proteases Lon and Clp during cytosolic protein degradation, 155, 710–719.

11. Cheng S.W. , Hu M.N. , Shen S.W., Takagi H., Asano M., Tsai Y.C., 1995, Production and characterization of keratinase of a feather-degrading *Bacillus licheniformis* PWD-1. Biosci. Biotech. Biochem., 59, 2239-2243.

12. Chiba S., Akiyama Y. and Ito K., 2002, Journal of Bacteriology, 184, 4775–4782.

13. Chung C.T., Niemela S.L., and Miller R.H., 1989, One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution, Proc. Nat. Acad. Sci. U.S.A., 86, 2172.

14. Clarke A.K., 1996, Variations on a theme: combined molecular chaperone and proteolysis functions in Clp/HSP100 proteins, J. Biosci., 21: 161–177.

15. Coux O., Tanaka K., Goldberg A.L., 1996, Structure and functions of the 20S and 26S proteasomes, Annu. Rev. Biochem., 65, 801-847.

16. De Mot R., Nagy I., Walz J., and Baumeister W., 1999, Trends in Microbiology, 7, 88–92.

17. Dodson G., Wlodawr A., 1998, Catalytic triads and their relatives, Trends Biochem. Sci., 23, 347-352

18. Dong G., Vieille C., Savchenko A., Zeikus J. G., 1997, Cloning, sequencing, and expression of the gene encoding extracellular a-amylase from *Pyrococcus furiosus* and biochemical characterization of the recombinant enzyme, Appl. Environ. Microbiol., 63, 3569–3576.

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

20. Feng Y.Y., Yang W.B. , Ong S.L. , Hu J.Y. , Ng W.J. , 2001, Fermentation of starch for enhanced alkaline protease production by constructing an alkalophilic *Bacillus pumilis* strain, Appl. Microboil. Biotechnol., 57:153-160.

21. Fox J.W., Shannon J.D., and Bjarnason J. B., 1991, Proteinases and their inhibitors in biotechnology, Enzymes in biomass conversion. ACS Symp., 460, 62-79.

22. Frank E.G., Ennis D.G., Gonzalez M., Levine A.S.,Woodgate R., 1996, Regulation of SOS mutagenesis by proteolysis, Proc. Natl. Acad. Sci. USA 93, 19, 10291–10296.

23. Frees D., and Ingmer H., 1999, ClpP participates in the degradation of misfolded protein in *Lactococcus lactis*, Mol. Microbiol., 31, 79–87.

24. Frees D., Qazi S., Hill P., Ingmer H., 2003, Alternative roles of ClpX and ClpP in *Staphylococcus aureus* stress tolerance and virulence, Mol. Microbiol., 48, 1565–1578.

25. Frees D., Chastanet A., Qazi S., Sørensen K., Hill P., Msadek T., and Ingmer H., 2004, Clp ATPases are required for stress tolerance, intracellular replication and biofilm formation in *Staphylococcus aureus,* Molecular Microbiology, 54, 5, 1445–1462.

26. Fukui T., Eguchi T., Atomi H., Imanaka,T., A Membrane-Bound Archaeal Lon Protease Displays ATP-Independent Proteolytic Activity towards Unfolded Proteins and ATP-Dependent Activity for Folded Proteins, Jurnal of Bacteriology, July 2002, p. 3689–3698

27. Gaillot O., Pellegrini E., Bregenholt S., Nair S., Berche P., 2000, The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*, Mol. Microbiol., 35, 1286–1294.

28. Gayda R.C., Markovitz A., 1978, Altered bacteriophage lambda expression in cell division mutants capR(lon) of *Escherichia coli* K-12. Mol. Gen. Genet., 159, 1, 1–11.

29. George M.Y., Kopolovic A.D., Elliott M.B., Diamandis E.P., 2003, Genomic overview of serine proteases, Biochemical and Biophysical Research Communications, 305, 1, 28-36.

30. Gille C., Goede A., Schlötelburg C., Preißner R., Kloetzel P.,Göbel U.B., and Frömmel C., 2003, A Comprehensive View on Proteasomal Sequences: Implications for the Evolution of the Proteasome, J. Mol. Biol., 326, 1437–1448.

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

32. Goldberg A. L., Moerschell R. P., Chung C. H., and Maurizi M. R., 1994, ATP-dependent protease La (lon) from *Escherichia coli*, Methods Enzymol., 244, 350–375.

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

34. Gottesman S., 1996, Proteases and their targets in *Escherichia coli*. Annu. Rev. Genet., 30, 465–506.

35. Gottesman S., Maurizi M.R., 1992, Regulation by proteolysis: energy-dependent proteases and their targets, Microbiol. Rev., 56, 4, 592–621.

36. Gottesman S., Wickner S., Maurizi M.R., 1997, Protein quality control: triage by chaperones and proteases, Genes Dev., 11, 815– 823.

37. Groll M. and Clausen T., 2003, Molecular shredders: how proteasomes fulfill their role, Current Opinion in Structural Biology, 13, 665–673.

38. Groll M., Ditzel L., Löwe J., Stock D., Bochtler M., Bartunik H.D., Huber R., 1997, Structure of 20S proteasome from yeast at 2.4 Aº resolution, Nature, 386:463-471.

39. Groll M., Bochtler M., Brandstetter H., Clausen T., and Huber R., 2005, Molecular Machines for Protein Degradation, ChemBioChem, 6, 222 – 256

40. Gupta R., Beg Q.K., Lorenz P., 2002, Bacterial alkaline proteases: molecular approaches and industrial applications, Appl. Microbiol. Biotechnol., 59, 1, 15-32.

41. Gupta R.S., 1998, What are archaebacteria: life's third domain or monoderm prokaryotes related to Gram-positive bacteria? A new proposal for the classification of prokaryotic organisms, Mol. Microbiol. 29:695–707.

42. Hattendorf D.A., and Lindquist S.L., 2002, Cooperative kinetics of both Hsp 104 ATPase domains and interdomain communication revealed by AAA sensor-1 mutants, EMBO J., 21, 12-21.

43. Hedstrom L., Serine Protease Mechanism and Specifity, Chem. Rev. 2002, 102, 4501-4532.

44. Hengge R., and Bukau B., 2003, Proteolysis in prokaryotes: protein quality control and regulatory principles, Molecular Microbiology, 49, 6, 1451–1462.

45. Holden J.F., Baross J.A., 1993, Enhanced thermotolerance and temperature-induced changes in protein composition in hyperthermophilic archaea, J. Bacteriol., 175, 2839-2843.

46. Horl W.H., Proteinases: potential role in health and disease, in: M. Sandler, H.J. Smith (Eds.), Design of Enzyme Inhibitors as Drugs, Oxford University Press, Oxford, 1989, pp. 573–581.

47. Horwich A.L., Weber-Ban E.U., Finle, D., 1999. Chaperone rings in protein folding and degradation, Proc. Natl. Acad. Sci. USA 96, 11033–11040.

48. Huang Q., Peng Y., Li X., Wang H., Zhang Y., 2003, Purification and Characterization of an Extracellular Alkaline Serine Protease with Dehairing Function from *Bacillus pumilus,*Current Microbiology, 46 , 169-173

49. Im Y. J., Na Y., Kang G. B., Rho S. H., Kim M. K., Lee J. H., 2004, The active site of a Lon protease from *Methanococcus jannaschii* distinctly differs from the canonical catalytic dyad of Lon proteases, J. Biol. Chem. 279, 53451–53457.

50. Ingmer H., Vogensen F. K., Hammer K., Kilstrup M., 1999, Disruption and analysis of the *clpB*, *clpC*, and *clpE* genes in *Lactococcus lactis*: ClpE, a new Clp family in gram-positive bacteria, J. Bacteriol., 181, 2075–2083.

51. International Union of Biochemistry, 1992, Enzyme nomenclature, Academic Press, Inc., Orlando, Fla.

52. Ishikawa H., Ishimi K., Sugiura M., Sowa A., Fujiwara N., 1993, Kinetics and mechanism of enzymatic hydrolysis of gelatin layers of X-ray film and release of silver particles, J Ferment Bioeng., 76, 300-305

53. Iyer L.M., Leipe D.D., Koonin E.V., Aravind L., 2004, Evolutionary history and higher order classification of AAA^{+} ATPases. J. Struct. Biol., 146, 11–31.

54. Joshi S.A., Hersch G.L., Baker T.A., and Sauer R.T., 2004, Communication between ClpX and ClpP during substrate processing and degradation, Nature Structural&Molecular Biology, Vol. 11, 5, 404-411.

55. Kanemori M., Nishihara K., Yanagi H., and Yura T., 1997, Synergistic roles of HslVU and other ATPdependent proteases in controlling in vivo turnover of sigma32 and abnormal proteins in *Escherichia coli*., J. Bacteriol., 179, 7219–7225.

56. Karata K., Inagawa T., Wilkinson A.J., Tatsuta T., Ogura T., 1999, Dissecting the role of a conserved motif (the second region of homology) in the AAA family of ATPases, Site-directed mutagenesis of the ATP-dependent protease FtsH., J. Biol. Chem., 274, 26225–26232.

57. Kawashima T., Amano N., Koike H., Makino S., Higuchi S., Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T., Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M., 2000, Archaeal adaptation to higher temperatures revealed by genomic sequence of *Thermoplasma volcanium*, Proc. Natl. Acad. Sci. U.S.A., 97, 14257-14262.

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

59. Kim K.I., Woo K.M., Seong I.S., Lee Z.W., Baek S.H., Chung C.H., 1998, Mutational analysis of the two ATPbinding sites in ClpB, a heat shock protein with proteinactivated ATPase activity in *Escherichia coli*, Biochem J., 333, 671–676.

60. Kim W. , Choi K. , Kim Y. , Park H. , Chol J. , Lee Y. , Oh H. , Kwon I. , Lee S., 1996, Purification and characterization of a fibrinolytic enzyme produced from *Bacillus* sp. strain CK 11-4 screened from chungkook-jangs. Appl Environ Microbiol., 62, 2482-2488

61. Kim Y.I., Levchenko I., Fraczkowska K., Woodruff R.V., Sauer R.T., Baker T.A., 2001, Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. Nat. Struct. Biol., 8, 230–233.

62. Knipfer N., Seth A., Roudiak S. G., and Shrader T. E., 1999, Species variation in ATP-dependent protein degradation: protease profiles differ between mycobacteria and protease functions differ between *Mycobacterium smegmatis* and *Escherichia coli*. Gene, 231, 95–104.

63. Krüger E., Witt E., Ohlmeier S., Hanschke R., and Hecker M., 2000, The *clp* proteases of *Bacillus subtilis* are directly involved in degradation of misfolded proteins, J Bacteriol., 182, 3259–3265.

64. Kudrya V.A., Simonenko I.A. , 1994, Alkaline serine proteinase and lectin isolation from the culture fluid of *Bacillus subtilis,* Appl Microbiol Biotechnol., 41, 505-509.

65. Kwon A., Kessler B., Overkleeft H., McKay D., 2003, Structure and reactivity of an asymmetric complex between HslV and Idomain deleted HslU, a prokaryotic homolog of the eukaryotic proteasome, J Mol Biol., 330, 185-195.

66. Ladenstein R., Antranikian G., 1998, Proteins from hyperthermophiles: stability and enzymatic catalysis close to the boiling point of water, Adv Biochem Eng Biotechnol., 61, 37-85.

67. Laemmli U.K., 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227, 259, 680–685.

68. Lee S., Sowa E.M., Choi J., and Tsai F.T.F, 2004, The ClpB/Hsp104 molecular chaperone—a protein disaggregating machine, Journal of Structural Biology, 146, 99–105.

69. Lemos J.A., Burne R.A., 2002, Regulation and physiological significance of ClpC and ClpP in *Streptococcus* mutans, J. Bacteriol., 184, 6357–6366.

70. Lenzen C.U., Steinmann D., Whitehaert S.W., and Weiss W.I., 1998, Crystal structure of the hexamerization domain of N-ethylmaleimide sensitive fusion protein, Cell, 94, 525-536.

71. Lupas A.N., and Martin J., 2002, AAA proteins, Curr. Opin. Struct. Biol., 12, 746–753.

72. Maas R., 2001, Change of plasmid DNA structure, hypermethylation, and Lon-proteolysis as steps in a replicative cascade, Cell 105, 7, 945–955.

73. Maurizi M.R., Clark W. P., Kim S. H., Gottesman S., 1990, ClpP represents a unique family of serine proteases, The Journal of Biological Chemistry, 265, 21, 12546-12552.

74. Maurizi M.R., 1992, Proteases and protein degradation in *Escherichia coli,* Experientia, 48, 178–201.

75. Maurizi, M.R. & Li, C.C.H., 2001, AAA proteins: in search of a common molecular basis, EMBO Report 2, 980–985.

76. Maurizi M.R., Trisler P., Gottesman S., 1985, Insertional mutagenesis of the *lon* gene in *Escherichia coli*: *lon* is dispensible, J. Bacteriol., 164: 1124–1135.

77. Maurizi M.R., and Xia D., 2004, Protein Binding and Disruption by Clp/Hsp100 Chaperones, Structure, Vol. 12, 175–183.

78. Melnikov E.E., Tsirulnikov K.B., and Rotanova T.V., 2000, Coupling of proteolysis with ATP hydrolysis by *Escherichia coli* Lon proteinase. I. Kinetic aspects of ATP hydrolysis. Bioorgan. Khim. 26, 530–538.

79. Menon A.S., Goldberg A.L., 1987, Protein substrates activate the ATP-dependent protease La by promoting nucleotide binding and release of bound ADP, J Biol Chem., Nov 5, 262, 31, 14929-34.

80. Mogk A., Tomoyasu T., Goloubinoff P., Rüdiger S., Röder D., Langen H., Bukau B., 1999, Identification of thermolabile *E. coli* proteins: prevention and reversion of aggregation by DnaK and ClpB, EMBO J.,18, 6934–6949.

81. Motohashi K., Watanabe Y., Yohda M., Yoshida M., 1999, Heat inactivated proteins are rescued by the DnaK.J-GrpE set and ClpB chaperones, Proc. Natl. Acad. Sci. USA 96, 7184–7189.

82. Msadek T., Dartois V., Kunst F., Herbaud M.L., Denizot F., Rapoport G., 1998, ClpP of *Bacillus subtilis* is required for competence development, motility, degradative enzyme synthesis, growth at high temperature and sporulation, Mol. Microbiol., 27, 899–914.

83. Neidhardt F.C., Ingraham J.L., Schaechter M., 1990, Physiology of bacterial cell: a molecular approach, Sinauer Associates, Inc, Sunderland, Massachusetts.

84. Neklyudov A. D. , Ivankin A. N. , Berdutina A. V., 2000, Properties and uses of protein hydrolysates., Appl Biochem Microbiol., 36,452-159.

85. Neuwald A.F., Aravind L., Spounge J.L., and Koonin E.V., 1999, AAA^{+}: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes, Genome Res, 9, 27–43.

86. Niehaus F., Bertoldo C., Kahler M.,Antranikian G., 1999, Extremophiles as a source of novel enzymes for industrial application, Appl Microbiol Biotechnol, 51, 711-729.

87. Ogura T., and Wilkinson A.J., 2001, AAA^+ superfamily ATPases: common structure—diverse function. Genes Cells, **6**, 575–597.

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

89. Poldermans B., 1990, Proteolytic enzymes, 108-123. *In* W. Gerhartz (ed.), Proteolytic enzymes in industry: production and applications. Publishers, Weinheim, Germany.

90. Porankiewicz J., Wang J., Clarke A. K., 1999, New insights into the ATP-dependent Clp protease: *Escherichia coli* and beyond, Molecular Microbiology, 32, 3, 449–458.

91. Puri S., 2001, An alkaline protease from *a Bacillus* sp.: Production and potential applications in detergent formulation and degumming of silk. MSc thesis, University of Delhi, New Delhi.

92. Putnam C.D., 2001, Structure and mechanism of the RuvB Holliday junction branch migration motor, J. Mol. Biol., 311, 297–310.

93. Queitsch C., Hong S.W., Vierling E., Lindquist S., 2000, Heat shock protein 101 plays a crucial role in thermotolerance in Arabidopsis, Plant Cell, 12, 479–492.

94. Rao M.B., Tanksale A.M., Ghatge M.S., Deshpande V.V., 1998, Molecular and biotechnological aspects of microbial proteases, Microbiol Mol Biol Rev., 62, 3, 597-635.

95. Rawlings N.D., and Barrett A.J., 1993. Evolutionary families of peptidases, Biochem. J., 290, 205-218.

96. Robb F.T., and Place A.R. 1995, Archaea: a laboratory manual, Cold Spring Harbor Laboratory Press.

97. Robertson G.T., Ng W.L., Foley J., Gilmour R., Winkler M.E., 2002, Global transcriptional analysis of *clpP* mutations of type 2 *Streptococcus pneumonia*e and their effects on physiology and virulence, J Bacteriol., 184, 3508–3520.

98. Rosen R., Biran D., Gur E., Becher D., Hecker M., Ron E.Z., 2002, Protein aggregation in *Escherichia coli*: role of proteases, FEMS Microbiol. Lett., 207, 1, 9–12.

99. Rotanova T. V., Melnikov E. E., Khalatova A. G., Makhovskaya O. V., Botos I., Wlodawer A., Gustchina, A., 2004, Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains, Eur. J. Biochem. 271, 4865–4871.

100. Rotanova T.V., Melnikov E.E., and Tsirulnikov K.B., 2003, A catalytic Ser – Lys dyad in the active site of the ATP-dependent Lon protease from *Escherichia coli*. Bioorgan. Khim., 29, 97–99.

101. Sambrook J.,and Russel D.W., 2001, Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

102. Sanchez Y., Taulien J., Borkovich K.A., Lindquist S., 1992, Hsp104 is required for tolerance to many forms of stress, EMBO J., 11, 2357– 2364.

103. Sanchez Y., and Lindquist S., 1990, Hsp104 required for induced thermotolerance, Science 248, 1112–1115.

104. Schirmer E. C., Glover J. R, Singer M. A, and Lindquist S., 1996, HSP100/Clp proteins: a common mechanism explains diverse functions, Trends Biochem. Sci., 21, 289–296.

105. Schoemaker J.M., Gayda R.C., Markovitz A., 1984, Regulation of cell division in *Escherichia coli*: SOS induction and cellular location of the sulA protein, a key to lon-associated filamentation and death, J. Bacteriol., 158, 2, 551–561.

106. Schultz R.M., Liebman M.N., Structure–function relationship in protein families, in: T.M. Devlin (Ed.), Textbook of Biochemistry with Clinical Correlations, fourth ed., Wiley-Liss, New York, 1997, pp. 1–2,116.

107. Schweder T., Lee K.H., Lomovskaya O., Matin A., 1996, Regulation of *Escherichia coli* starvation sigma factor (sigma s) by ClpXP protease, J. Bacteriol., 178, 2, 470–476.

108. Seemüller E., Lupas A., Stock D., Löve J., Huber R., Baumeister W., 1995, Proteasome from *Thermoplasma acidophilum*: a threonine protease, Science, 268, 579-582.

109. Shockley K. R., Ward D. E., Chhabra S. R., Conners S. B., Montero C. I., and Kelly R. M., 2003, Heat Shock Response by the

Hyperthermophilic Archaeon *Pyrococcus furiosus*, Applied and Environmental Microbiology, Apr., 2365–2371

110. Siezen R.J., Leunissen J.A.M., 1997, Subtilases: the superfamily of subtilisin-like serine proteases., Protein Sci, 6,501-523.

111. Skinner M. M., and Trempy J. E., 2001, Expression of *clpX*, an ATPase Subunit of the Clp Protease, Is Heat and Cold Shock Inducible in *Lactococcus lactis,* J. Dairy Sci., 84, 1783–1785.

112. So J.E., Shin J.S., Kim B.G., 2000, Protease-catalyzed tripeptide (RGD) synthesis, Enzyme Microb Technol., 26, 108-114.

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

114. Sutherland K.J., Henneke C.M., Towner P., hough D.W., and Danson M.J., 1990, Citrate synthase from the thermophilic *ArchaebacteriaThermoplasma acidophilum*: Cloning and sequencing of the gene, Europ. Jour. Biochem., 194, 839-844.

115. Swamy K. H., and Goldberg A. L., 1981, *E. coli* contains eight soluble proteolytic activities, one being ATP dependent, Nature, 292, 652–654.

116. Tomoyasu T., Gamer J., Bukau B., Kanemori M., Mori H., Rutman A. J.,Oppenheim A. B., Yura T., Yamanaka K., Niki H., 1995, *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32, EMBO J. 14, 2551–2560.

117. Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P., Bukau, B., 2001. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol, Mol. Microbiol., 40, 2, 397–413.

118. Torres-Cabassa, A.S., Gottesman, S., 1987, Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis, J. Bacteriol., 169, 3, 981–989.

119. Vera A., Arís A., Carrió M., González-Montalbán N., Villaverde A., 2005, Lon and ClpP proteases participate in the physiological disintegration of bacterial inclusion bodies, Journal of Biotechnology, 119, 163–171.

120. Vieille C., Zeikus G. J., 2001, Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability, Microbiology and Molecular Biology Rewievs, 65, 1, 1-43.

121. Voges D., Zwickl P., Baumeister W., 1999, The 26S proteasome: a molecular machine designed for controlled proteolysis, Annu. Rev. Biochem., 68, 1015-1068.

122. Wang J., Hartling J. A., and Flanagan J. M., 1997, Cell, 91, 447–456

123. Walker J.E., *et al*., 1982, Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATPrequiring enzymes and a common nucleotide binding fold, EMBO J., 1, 945–951.

124. Ward D.E., Shockley K.R., Chang L.S., Levy R.D., Michel J. K., Conners S.B., Kelly R.M., 2002, Proteolysis in hyperthermophilic microorganisms, Archaea, 1, 63–74.

125. Wawrzynow A., Wojtkowiak D., Marszalek J., Banecki B., Jonsen M., Graves B., Georgopoulos C., and Zylicz M., 1995, The ClpX heatshock protein of *Escherichia coli*, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone, EMBO J., 14, 1867–1877.

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

127. Weichart D., Querfurth N., Dreger M., Hengge-Aronis R., 2003, Global role for ClpP-containing proteases in stationary-phase adaptation of *Escherichia coli*, J. Bacteriol., 185, 1, 115–125.

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

129. Wickner S., Maurizi M.R., Gottesman S., 1999, Posttranslational quality control: folding, refolding, and degrading proteins, Science,286, 1888–1893.

130. Wiegert T., and Schumann W., 2001, SsrA-mediated tagging in *Bacillus subtilis* , J. Bacteriol., 183, 3885–3889.

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

132. Wu W. F., Zhou Y., and Gottesman S., 1999, Redundant in vivo proteolytic activities of Escherichia coli Lon and the ClpYQ (HslUV) protease, J. Bacteriol., 181, 3681–3687.

133. Yoo S. J., Shim Y. K., Seong I. S., Seol J. H., Kang M. S., and Chung C. H., 1997, FEBS Lett. 412, 57–60

134. Yousef G.M., Kopolovic A.D., Elliott M.B., and Diamandis E.P., 2003, Genomic overview of serine proteases, Biochemical and Biophysical Research Communications, 305, 28-36.

135. Yu R.C., Hanson P.I., Jahn R., and Brunger A.T., 1998, Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP, Nat. Struct. Bio., 9, 906-911.

136. Zeikus J. G. , Vieille C. , Savchenko A., 1998, Thermozymes: biotechnology and structure-function relationships., Extremophiles 2, 179-183

137. Zhou Y., Gottesman S., Hoskins J.R., Maurizi M.R., and Wickner S., 2001, The RssB response regulator directly targets sigma(S) for degredadtion by ClpXP, Genes Dev., 15, 627-637.

138. Zolkiewski M., 1999, ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. J. Biol. Chem. 274, 28083–28086.

139. Zwickl P., Baumeister W., Steven A., Dis-assembly lines: the proteasome and related ATPase-assisted proteases Current Opinion in Structural Biology 2000, 10:242–250.

APPENDIX A

MEDIA, BUFFERS AND SOLUTIONS

- **Serine Protease Assay Buffer pH: 7.0**
  - Tris ..... 50 mM
  - $CaCl_2$ ..... 5 mM

- **TE Buffer pH: 8.0**
  - TrisHCl ..... 10 mM
  - EDTA ..... 1 mM

- **TAE Buffer pH: 8.0**
  - Tris acetate ..... 0.04 M
  - EDTA ..... 0.001 M

- **LB Medium (1 L) Ph:7.4**
  - Tryptone ..... 10 g
  - Yeast extract ..... 5 g
  - NaOH .....10 g

- **SDS Solution**
  - 10% (w/v) sodium dodecyl sulfate

- **Denaturing Solution**
  - NaOH ……..... 0.5 M
  - NaCl …… 1.5 M

- **Neutralization Solution pH 7.4**
  - NaCl …… 1.5 M
  - Tris.Cl ……... 0.5 M

- **20X SSC Buffer (500ml)**
  - NaCl …….. 87.65 g
  - Trisodium citrate.$2H_2O$ ……. 50.25 g

- **TBS Buffer pH: 7.5**
  - Tris.Cl ……. 10 mM
  - NaCl ..... 150 mM

- **TBS-Tween Buffer pH: 7.5**
  - Tris.Cl ……. 20 mM
  - NaCl …….. 500 mM
  - Tween 20 …… 0.05 % (v/v)

- **TBS-Tween/Triton Buffer pH: 7.5**
  - Tris.Cl …….. 20 mM
  - NaCl ……500 mM
  - Tween 20 ……. 0.05 % (v/v)
  - TritonX-100 ……... 0.2 % (v/v)

- **Blocking Buffer**
  - 3 % (w/v) BSA in TBS buffer

- **DAB Working Solution (100ml)**
  - Chromogenic DAB Substrate ……11 ml
  - Double distilled H_2O …….. 89 ml

APPENDIX B

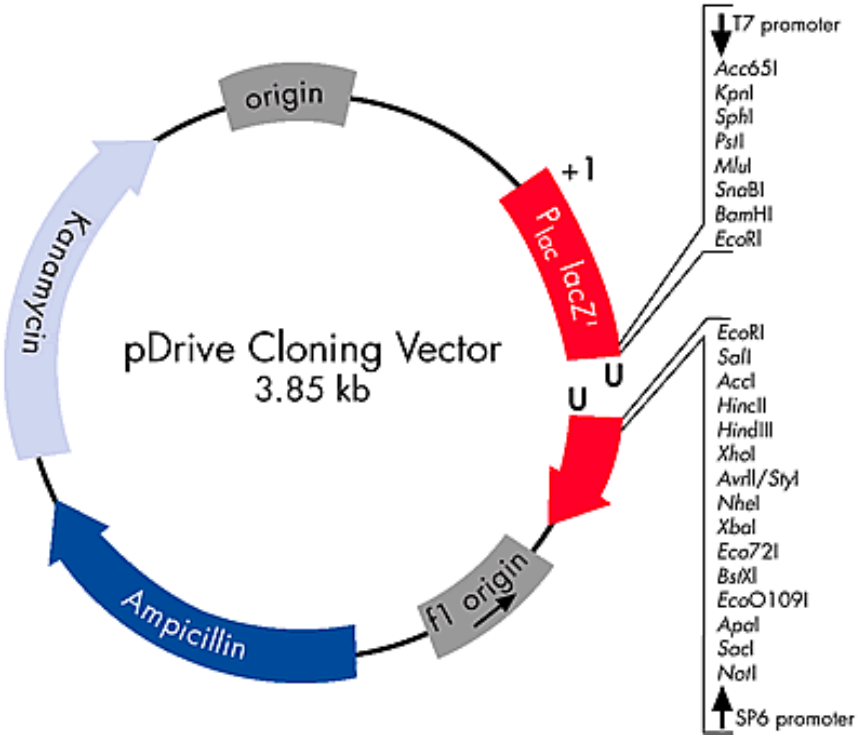

Figure B.1: pDrive Cloning Vector (QIAGEN)

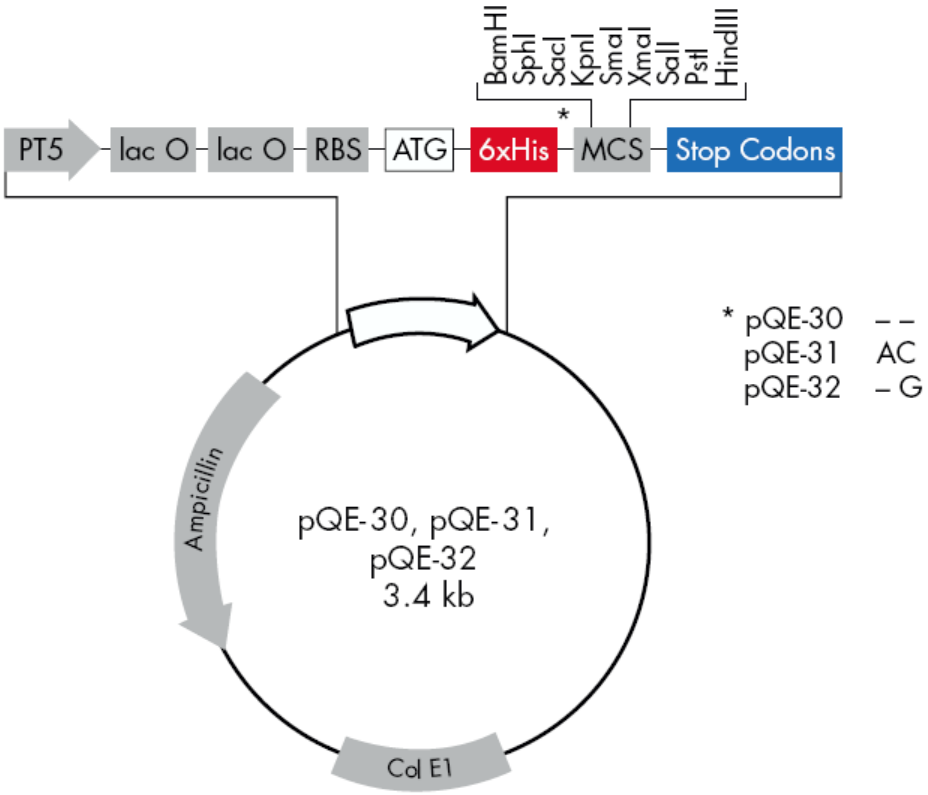

Figure B.2: pQE 30, pQE 31, pQE 32 Cloning Vector (QIAGEN)

APPENDIX C

MOLECULAR SIZE MARKER

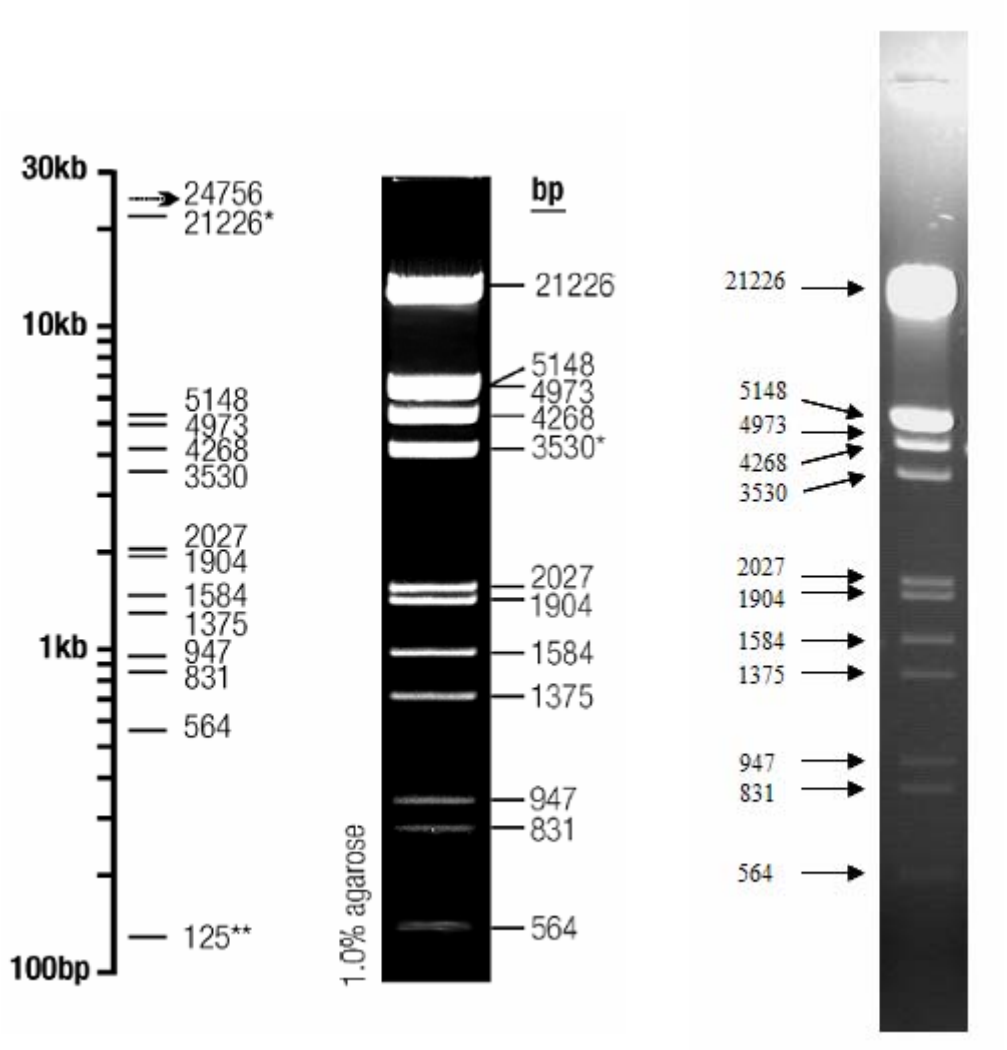

Figure C.1: Lambda DNA *EcoRI/HindIII* Molecular Marker (MBI Fermentas, AB, Vilnuis, Lithuania)

APPENDIX D

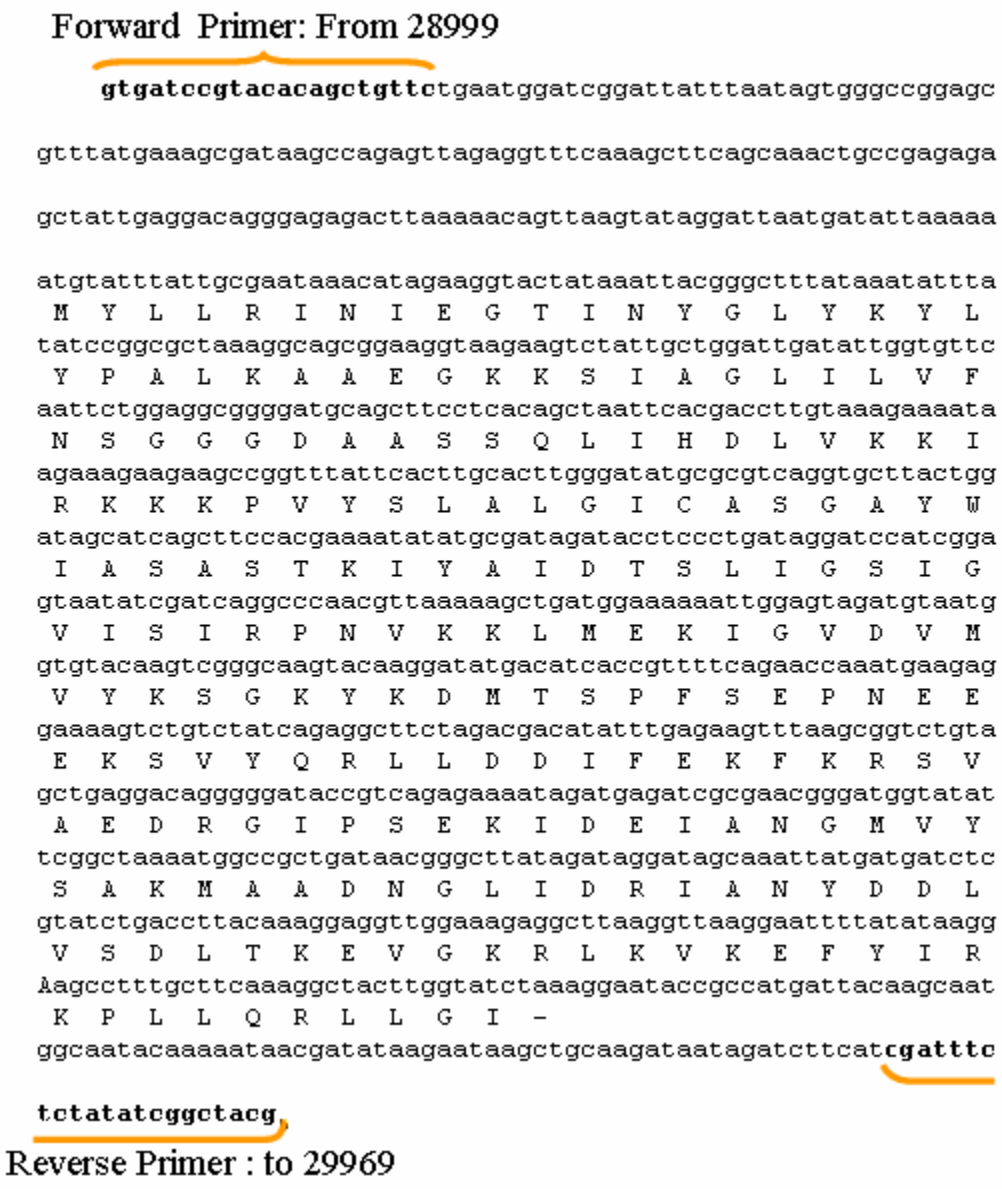

Figure D.1: Nucleotide and Aminoacid Sequence of Periplasmic (ClpP-like) Serine Protease Insert

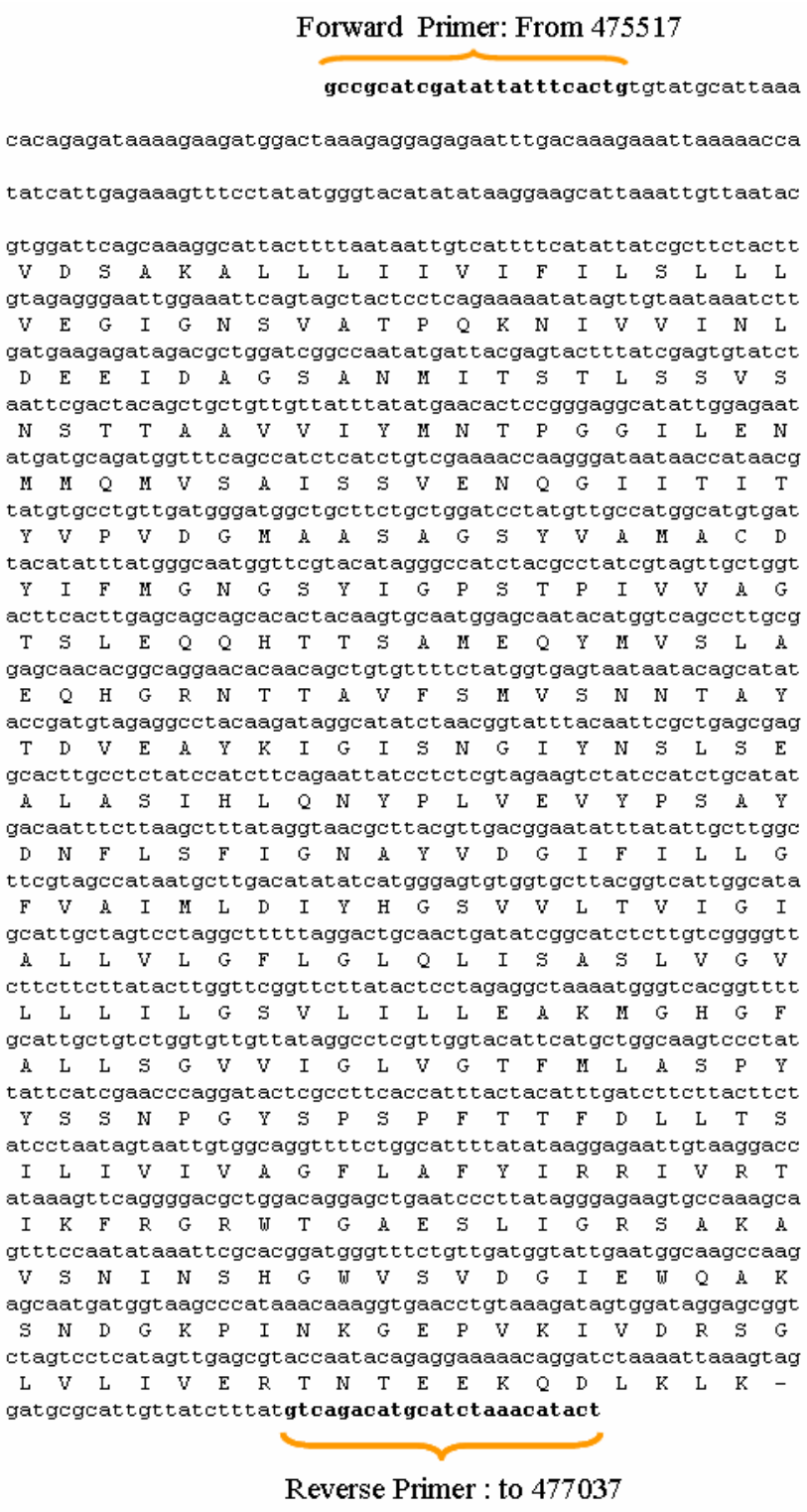

Figure D.2: Nucleotide and Aminoacid Sequence of Membrane-Bound Serine Protease Insert

APPENDIX E

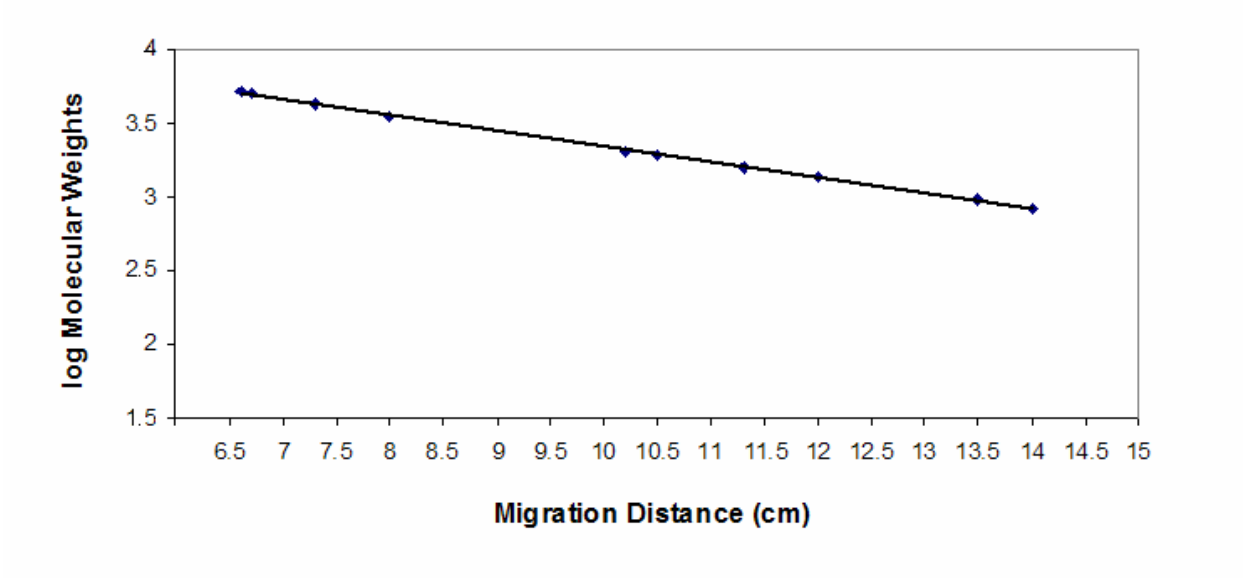

Figure E.1: Calibration Curve for Lambda DNA *EcoRI/HindIII* Cut Molecular Size Marker