Molecular Characterization of Thermostable Proteases from Thermophilic Bacterial Strains

Tanveer Majeed

2016

Department of Biotechnology
Pakistan Institute of Engineering and Applied Sciences
Nilore, Islamabad, Pakistan
Abstract

Thermophilic microorganisms have gained world-wide importance due to their tremendous potential to produce thermostable enzymes that have wide applications in pharmaceuticals and industries. Proteases are such enzymes which account for nearly 60% of the total world-wide enzyme sales. Thermostable proteases are of greater advantage in applications because they do not only denature at high temperatures, but they also remain active at such temperatures. This project explores the production and biochemical and molecular characterization of the thermostable proteases from thermophilic bacterial strains.

The growth conditions of *Bacillus subtilis* strains in submerged cultivation have been investigated to understand fermentation behavior of the microorganism and the production pattern of the thermostable enzyme. The optimum protease production (178 U/mL) was observed with 1% casein substrate after 48 h of cultivation at 60°C from *Bacillus subtilis* BSP in shake flasks study. The optimization of various factors affecting the protease production was statistically determined by Box Behnken design based on response surface methodology (RSM). The simultaneous effects of four test interacting factors on the protease production were monitored and conditions were optimized for maximum enzyme production. Optimized cultivation media formulation included: initial pH 8, casein concentration 2% (w/v) and inoculum density 1% when fermentation was carried out for 48 h. The enzyme was 9.8 fold purified in a 2-step procedure involving ammonium sulfate precipitation, DEAE cellulose and Sephadex G-200 gel permeation chromatography. The enzyme was shown to have a molecular weight of 36 kDa by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). It was most active at 50°C, pH 8, with casein as substrate. This enzyme was 100% stable at 50°C and retained its 82% activity at 60°C after 30 minutes of incubation. The enzyme became completely inactivated after incubation at 80°C for 3h. It was strongly activated by metal ions such as Ca^{2+} and Mn^{2+}. Enzyme activity was inhibited strongly by ethylene diamine tetra acetic acid (EDTA) confirming its identity as a metalloprotease. Protease was not only stable in the presence of organic solvents and surfactants studied but it also exhibited a higher activity than in the absence of some surfactants.
Consequently, the DNA encoding the thermostable protease of *Bacillus subtilis* BSP has been identified. The gene contained an open reading frame of 1638 bp and encoded for a mature peptide sequence of 318 amino acid residues. The protease gene was cloned and expressed in *E.coli* expression system and the recombinant protease was purified. Sequence analysis showed that this protease has close homology with thermolysin class of proteases. Primary structure analysis of thermostable protease showed 35% of its content to be alpha helix making it stable for three dimensional structure modeling. Homology model of thermolysin has been constructed using swiss model as the workspace. The model was validated by ProSA and RMSD. The results showed the final refined model is reliable. It has 0.06 Å as RMSD and has -2.19 as Z-score. For expression in *Bacillus subtilis* 1A751 two integration vectors (pDR111 and pSG1154) were compared in order to get the higher extracellular protease yield. 1A751/pDR-BSP-MprT transformant was found to be secreted higher amount of thermostable protease in the culture medium when compared to 1A751/pSG-BSP-MprT strain. The recombinant enzyme was undergone for further studies to check its stability under native enzyme conditions and it was found to be as stable as native protease BSP-MprT.

The thermostability of *Bacillus subtilis* BSP protease was improved by two rounds of error prone PCR mutagenesis. A random mutant library of *Bacillus subtilis* BSP protease was generated by ep-PCR. The generated mutants have been successfully expressed in *E. coli*. The mutant proteases obtained in this mutant library were screened for increased protease thermostability. The SDS-PAGE pattern of enzyme was exhibiting a well-defined band (36 kDa). In this study, the thermostability was enhanced from 60°C to 80°C by the single mutation Gly347Cys which has a stabilizing effect on the irreversible thermal inactivation. The BSP-mutant has exhibited increase in thermal stability and 2.7 folds half-life at 80°C when compared to wild type protease. Sequence analysis and comparison of both wt-BSP-MprT and Mutant-BSP-MprT showed that wt-BSP-MprT had only one cysteine residue at Cys518 position while another cysteine has substituted Gly347 in Mutant-BSP-MprT. Introduction of another cysteine resulted in the formation of disulfide bridge between the two cysteine residues which play important role in the stability of
protein tertiary structure. The inter subunit disulfide bond was estimated to be one of the factors for thermal stability.

Another putative alkaline serine thermostable protease gene (aprE) from the thermophilic bacterium *Coprothermobacter proteolyticus* was cloned and expressed in *Bacillus subtilis*. The enzyme was determined to be a serine protease based on inhibition by PMSF. Biochemical characterization demonstrated that the enzyme had optimal activity under alkaline conditions (pH 8–10). In addition, the enzyme had an elevated optimum temperature (60°C). The protease was also stable in the presence of many surfactants and oxidant. Thus, the *C. proteolyticus* protease has potential applications in industries such as the detergent market.
1 Introduction

1.1 Enzymes

Enzymes are well known biocatalysts that perform a large number of chemical reactions and are commercially exploited in the food, detergents, diagnostics and pharmaceutical fine chemical industries. More than 3000 different enzymes have been reported so far, the majority of which were isolated from mesophilic microorganisms. These enzymes mainly work in a narrow range of pH, temperature, and ionic strength. Furthermore, due to the technological application the currently known enzymes have become uncommendable under desired industrial condition. Therefore, the search for new microbial sources is a continual exercise, where one must respect biodiversity.

The role of enzymes in many processes has been recognized for a long time. Their existence was linked with the history of ancient Greece where they were using enzymes from microorganisms in alcohol production, baking, brewing, cheese making etc. The number of applications has been increased many folds with better knowledge and purification of enzymes and with the availability of engineered enzymes a number of new possibilities for industrial processes have emerged [1].

Enzymes derived from extremophilic microorganisms are able to withstand harsh conditions in industrial processes that were long thought to be destructive to proteins. Heat-stable and solvent-tolerant biocatalysts are valuable tools for processes in which for example hardly decomposable polymers need to be liquefied and degraded, while cold-active enzymes are of relevance for food and detergent industries. Extremophilic microorganisms are a rich source of naturally tailored enzymes, which are more superior over their mesophilic counterparts for applications at extreme conditions [2]. Moreover enzymes from extremophiles were revealed to have compact structure and a number of charge interactions compared to mesophilic counterparts, and such structural and functional analyses of extreme-enzymes have provided some insight into the design of enzymes with high stability [3].
1.2 Enzymes in Industrial Applications: Global Markets

Today enzymes enjoy a major market next to pharmaceuticals and industrial enzymes and are growing steadily. They are being used as eco-friendly and cost effective substitutes for chemical processing in several industries.

1.3 BCC Research Market Forecasting

Enzyme technology is being influenced almost every sector of industrial activity, extending from technical field to food, feed and healthcare activity. Enzymatic processes are rapidly becoming better alternatives to chemical processes both financially and ecologically due to enzymes biodegradable nature and cost effectiveness.

BCC Research is a market research and forecasting company that provides reports for current and future aspects of several business categories including biotech, chemicals, food and beverage, healthcare, pharmaceuticals, engineering, environment, battery technologies, information technology, nanotechnology, instrumentation and sensors, membrane and separation technology, safety and security, plastics, and semiconductor manufacturing. More than 250 market research and forecasting reports are being published by BCC annually.

According to the BCC market research the global market for industrial enzymes was calculated to be around US$4.5 billion in 2012 and almost worth US$4.8 billion in 2013. Registering a five-year CAGR (compound annual growth rate) of 8.2% from 2013 to 2018, this market is estimated to reach about $7.1 billion by the end of 2018 (Figure 1-1). The industrial enzyme market is dominated by Novozymes, DuPont and DSM. Maximum growth is estimated to be in the detergent enzyme market which was valued at nearly $1.1 billion in 2013 and is estimated to reach $1.8 billion by 2018 with a CAGR of 11.3%. Animal feed is the second largest segment with 10% CAGR during the forecast period. (http://www.bccresearch.com/report/enzymes-industrial-applications-bio030H.html).

Microorganisms produce a large variety of enzymes, most of which are made in only small amounts and are involved in cellular processes. Extracellular enzymes are usually capable of digesting insoluble nutrient materials such as cellulose, protein
and starch and the digested products are transported into the cell where they are used as nutrients for growth. Some extracellular enzymes are used in the food, dairy, pharmaceutical, and textile industries and are produced in large amounts by microbial synthesis [4]. Commercial enzymes are usually produced from strains of bacteria, molds, and yeasts [5]. Even over fifty years ago commercial enzymes were produced by fungi, bacteria and yeast [6].

![Figure 1-1 Global industrial enzymes market, 2012-2018 $ millions. Source: BCC research (BIO030H), June 2014](image)

The majority of currently used industrial enzymes are hydrolytic in action, being used for the degradation of various natural substances. Proteases remain the dominant enzyme type, because of their extensive use in the detergent and dairy industries [7].

### 1.4 Proteases

Proteases (EC 3.4.21-24 and 99; peptidyl-peptide hydrolases) are enzymes that hydrolyze proteins by the addition of water across peptide bonds and catalyze peptide synthesis in organic solvents and in solvents with low water content [8]. The hydrolysis of peptide bonds by proteases as shown in Figure 1-2 is termed as proteolysis; the products of proteolysis are protein and peptide fragments, and free amino acids.
Proteolytic enzymes are ubiquitous in occurrence and are found in all living organisms including prokaryotes, fungi, plants and animals and are essential for cell growth and differentiation. There is a changed interest in the study of proteolytic enzymes, mainly due to the recognition that these enzymes not only play an important role in the cellular metabolic processes but have also gained considerable attention in the industrial community [9]. They can be cultured in large quantities in comparatively short time by established fermentation methods and produce an abundant, regular supply of the desired product.

Proteases represent one of the three principal groups of industrial enzymes and have traditionally held the predominant share (Figure 1-3) of the industrial enzyme market accounting for about 60% of total worldwide sale of enzymes [10].

Figure 1-2 Protease catalysis of peptide bonds (Proteolysis)
Contribution of different enzymes to the total sale of enzymes.

The shaded portion indicates the total sale of proteases [11]

Proteases of commercial importance are produced from microbial, plant and animal sources. They constitute a very complex and large group of enzymes with different properties of substrate specificity, active site, catalytic mechanism, temperature and pH activity and stability profiles. Industrial proteases have application in a range of process taking advantage of the unique physical and catalytic properties of individual proteolytic enzyme types. In contrast to the specificity of their action this vast diversity of proteases has attracted worldwide attention in attempts to exploit their physiological and biotechnological applications [11].

Protease Enzymes Market worth $2,767 Million by 2019 (Figure 1-4) and it is growing at a CAGR of 5.3% from 2014 to 2019, out of which the detergent and cleaning industry accounts for the major share and it is growing at a CAGR of 5.6%. Few regions like North America and Western Europe have almost reached the maturity with regard to the protease enzymes detergent market due to which the trend in usage of the proteases is shifting towards pharmaceutical industries in these regions. On the other hand, the developing countries in Asia-Pacific and Rest of the
World (ROW) regions show a lot of untapped potential for the growth of protease enzymes in both detergent as well as pharmaceutical industries (Protein Hydrolysis Enzymes Market - Global Trends & Forecasts to 2019).

![Figure 1-4 Global protease enzymes market size, by geography. Source: BCC research report 2014](image)

### 1.5 Classification of Proteases

Proteases are classified in subgroup 4 of group hydrolases according to the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology [11].

Proteases can be classified according to three major criteria. Such as;

i) The reaction catalyzed,

ii) The chemical nature of the catalytic site,

iii) The evolutionary relationship which is revealed by the structure [11].

Proteases as shown in Table 1.1 are broadly classified as endo or exo enzymes on the basis of their site of action on protein substrates. They are further categorized as aspartic proteases, serine proteases, cysteine proteases, or metallo proteases depending on their catalytic mechanism. They are also classified into different families and clans according to their amino acid sequences and evolutionary
relationships. Depending on the pH of their optimal activity, they are referred to as acidic, neutral, or alkaline proteases [11].

### 1.5.1 Exoproteases

The exopeptidases act only close to the ends of polypeptide chains. They are classified as amino- and carboxypeptidases based on their site of action at the N or C terminus, respectively [11].

**Aminopeptidases**

Aminopeptidases act at free N terminus of the polypeptide chain and release a single amino acid residue, a dipeptide, or a tripeptide. They are known to eliminate the N terminal Met that may be found in heterologously expressed proteins but not in many naturally occurring mature proteins. Aminopeptidases occur in a variety of microbial species including bacteria and fungi [11].

**Carboxypeptidases**

These enzymes act at C terminals of the polypeptide chain and liberate a single amino acid or a dipeptide. Carboxypeptidases are divided into three major groups, serine carboxypeptidases, metallo carboxypeptidases, and cysteine carboxypeptidases, based on the nature of the amino acid residue at the active site of the enzymes [11].

### 1.5.2 Endopeptidases

Endopeptidases are characterized by their preferential action away from the N and C termini at the peptide bonds in the inner regions of the polypeptide chain. The endopeptidases are divided into four subgroups on the basis of their catalytic mechanism, (i) serine proteases, (ii) aspartic proteases, (iii) cysteine and (iv) metalloproteases [11].

**Serine Proteases**

Serine proteases are characterized due to the presence of a serine residue in the active site, termed as the catalytic triad, which contains asp (D), his (H), and ser (S) amino acids (Shi et al. 2009). They are abundant and widespread among viruses, bacteria and eukaryotes, suggesting that they are vital to the organisms [11].
Table 1.1 General Classification of proteases with their enzyme commission (EC) code, coupled with specific mechanism of action of each subgroup [12]

<table>
<thead>
<tr>
<th>Protease</th>
<th>EC code</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exopeptidases</td>
<td>3, 4, 11-19</td>
<td>Cleave the peptide bond proximal to the amino or carboxytermini of the substrate</td>
</tr>
<tr>
<td>Aminopeptidases</td>
<td>3, 4, 11</td>
<td>Those acting at a free N-terminus liberate a single amino acid residue</td>
</tr>
<tr>
<td>Dipeptidases</td>
<td>3, 4, 13</td>
<td>Exopeptidases specific for dipeptides</td>
</tr>
<tr>
<td>Dipeptidyl peptidase</td>
<td>3, 4, 14</td>
<td>Release of an N-terminal dipeptide from a polypeptide</td>
</tr>
<tr>
<td>Tripeptidyl peptidase</td>
<td>3, 4, 14</td>
<td>Release of an N-terminal tripeptide from a polypeptide</td>
</tr>
<tr>
<td>Peptidyl-dipeptidase</td>
<td>3, 4, 15</td>
<td>Release of free C-terminus liberate a dipeptide</td>
</tr>
<tr>
<td>Carboxypeptidase</td>
<td>3, 4, 16-18</td>
<td>Release of a single residue C-terminal from a polypeptide</td>
</tr>
<tr>
<td>Serine type protease</td>
<td>3, 4, 16</td>
<td>Carboxypeptidase have an active center serine involved in the catalytic process</td>
</tr>
<tr>
<td>Metalloprotease</td>
<td>3, 4, 17</td>
<td>Carboxypeptidase use a metal ion in the catalytic mechanism</td>
</tr>
<tr>
<td>Cysteine protease</td>
<td>3, 4, 18</td>
<td>Carboxypeptidase have a cysteine in the active center</td>
</tr>
<tr>
<td>Omega peptidases</td>
<td>3, 4, 19</td>
<td>Remove terminal residues that are linked by isopeptide bonds</td>
</tr>
<tr>
<td>Endopeptidases</td>
<td>3, 4, 21-24</td>
<td>Cleave internal bonds in polypeptide chains</td>
</tr>
<tr>
<td>Serine protease</td>
<td>3, 4, 21</td>
<td>Endopeptidases have an active centre serine involved in the catalytic proce</td>
</tr>
<tr>
<td>Cysteine protease</td>
<td>3, 4, 22</td>
<td>Possesses a cysteine in the active centre</td>
</tr>
<tr>
<td>Aspartic protease</td>
<td>3, 4, 23</td>
<td>An aspartic acid residue for their catalytic activity</td>
</tr>
<tr>
<td>Metalloprotease</td>
<td>3, 4, 24</td>
<td>Use a metal ion (often, but not always, Zn(^{2+})) in the catalytic mechanism</td>
</tr>
<tr>
<td>Endopeptidases of unknown catalytic mechanism</td>
<td>3, 4, 99</td>
<td>Acting on peptide bonds</td>
</tr>
</tbody>
</table>

Serine alkaline proteases are produced by several bacteria, yeast, molds, and fungi. They hydrolyse a peptide bond, which has phenylalanine, tyrosine, or leucine at the carboxyl site of the splitting bond. The optimal pH of alkaline proteases is about pH 10, and their isoelectric point is around pH 9. Their molecular weights are in the range at 15 and 30 kDa. Although serine alkaline proteases are produced by several bacteria such as *Arthrobacter*, *Streptomyces*, and *Flavobacterium* spp., subtilisins produced by *Bacillus* spp. are the best known ones [11].
Aspartic Proteases
Aspartic acid proteases are the endopeptidases that depend on aspartic acid residues for their catalytic activity. Acidic proteases have grouped into three families, namely, pepsin, retropepsin and enzymes from pararetroviruses. Most aspartic proteases show maximal activity at low pH and have isoelectric points in the range of pH 3 to 4.5. Their molecular weights are in the range of 30 to 45 kDa. The aspartic proteases are inhibited by pepstatin [11].

Cysteine / Thiol Proteases
Cysteine proteases occur in both prokaryotes and eukaryotes. This group is susceptible to oxidation and can react with a variety of reagents; heavy metals, iodoacetate, N-ethyl-maleimide etc. They are broadly divided into four groups based on their side-chain specificity: (i) papain-like, (ii) trypsin-like, (iii) specific to glutamic acid and (iv) others. Papain is the best-known cysteine proteases. Cysteine proteases have neutral pH optima.

1.5.2.1 Metalloproteases
Metallo proteases have characterized by the requirements for a divalent metal ion for their activity. According to the specificity of their action, metalloproteases can be divided into 4 groups, (i) neutral, (ii) alkaline, (iii) Myxobacter I and (iv) Myxobacter II.

1.6 Microbial Proteases
Microorganisms account for a two-third share of commercial protease production in the world [13]. Proteases of bacteria, fungi and viruses are extensively studied due to its importance and subsequent applications in industry and biotechnology as compared to proteases from plant and animals. Bacterial proteases possess almost all the characteristics desired for their biotechnological applications [14] and they can be cultured in large quantities in a relatively short time by established fermentation methods producing an abundant, regular supply of the desired product. In addition they can be genetically manipulated to generate new enzymes with altered properties that are desirable for their various applications [9,11].
Table 1.2 Commercial bacterial proteases, sources, applications and their industrial suppliers, Gupta et al. [9].

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Product name</th>
<th>trade name</th>
<th>Microbial sources</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genencor International, USA</td>
<td>Purafact</td>
<td>Primatan</td>
<td>B. lentus</td>
<td>Detergent</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bacterial source</td>
<td>Leather</td>
</tr>
<tr>
<td>Novo Nordisk, Denmark</td>
<td>Novoyme 243</td>
<td>Nue</td>
<td>B. licheniformis</td>
<td>Denture cleaners</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Durazyme</td>
<td>Bacillus sp.</td>
<td>Leather</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alcalase</td>
<td>Bacillus sp.</td>
<td>hydrolysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Savinase</td>
<td>B. licheniformis</td>
<td>Feed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bacillus sp</td>
<td>Detergent, silk degumming</td>
</tr>
<tr>
<td>Solvay Enzymes, Germany</td>
<td>Optimase</td>
<td>Protein engineered variant of Bacillus sp</td>
<td>Detergent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maxapem</td>
<td></td>
<td>B. subtilis</td>
<td>Detergent</td>
</tr>
<tr>
<td></td>
<td>Opticlean</td>
<td></td>
<td>B. licheniformis</td>
<td>Alcohol, baking, leather, photographic waste, Food, waste</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Detergent, food</td>
</tr>
<tr>
<td>Godo Shusei, Japan</td>
<td>Godo-Bap</td>
<td></td>
<td>B. licheniformis</td>
<td>Detergent</td>
</tr>
<tr>
<td>Gist-Brocades, The Netherlands</td>
<td>Maxatase</td>
<td>Bacillus sp.</td>
<td>Detergent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maxacal</td>
<td>Bacillus sp.</td>
<td>Detergent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtilisin</td>
<td>B. alcalophilus</td>
<td>Detergent</td>
<td></td>
</tr>
<tr>
<td>Nagase Biochemicals, Japan</td>
<td>Bioprase SP-10</td>
<td>B. subtilis</td>
<td>Food</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bioprase</td>
<td>B. subtilis</td>
<td>Detergent, cleaning</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cryst. Protease</td>
<td>B. subtilis (Bioteus)</td>
<td>Research</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ps. Elastase</td>
<td>Pseudomonas aeruginosa</td>
<td>Research</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ps. Protease</td>
<td>Pseudomonas aeruginosa</td>
<td>Research</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bioprase concentrate</td>
<td>B. subtilis</td>
<td>Cosmetic, pharmaceuticals</td>
<td></td>
</tr>
<tr>
<td>Rohm, Germany Enzyme Development, USA</td>
<td>Corolase 7089</td>
<td>B. subtilis</td>
<td>Food</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enzeco high alkaline protease</td>
<td>Bacillus sp.</td>
<td>Industrial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enzeco alkaline protease</td>
<td>B. licheniformis</td>
<td>Industrial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wuxi</td>
<td>Bacillus sp.</td>
<td>Detergent</td>
<td></td>
</tr>
<tr>
<td>Wuxi Synder Bioproducts, China</td>
<td>Protosol</td>
<td>Bacillus sp.</td>
<td>Detergent</td>
<td></td>
</tr>
<tr>
<td>Advance Biochemicals, India</td>
<td>Proleather</td>
<td>Bacillus sp.</td>
<td>Food</td>
<td></td>
</tr>
<tr>
<td>Amano Pharmaceica, Japan</td>
<td>Collagenase</td>
<td>Clostridium sp</td>
<td>Technical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Amano protease</td>
<td>Bacillus sp.</td>
<td>Food</td>
<td></td>
</tr>
</tbody>
</table>
Among various microorganisms *Bacillus* strains are one of the most important producers of commercially applicable proteases [15]. Commonly microbial proteases are extracellular in nature and are directly secreted into the fermentation broth by the producer, thus simplifying downstream processing of the enzyme as compared to proteases obtained from plants and animals [9].

### 1.7 Thermostable Proteases

The industrial use of proteases, in detergents and in leather processing, requires that the enzymes be stable at higher temperatures. Thermostable proteases like other enzymes from thermophilic and hyperthermophilic microorganisms (often called thermozymes) are of particular interest in some applications because they are stable and active at temperatures above 60-70°C [16]. Thermostable proteases are advantageous in some applications because of the higher processing temperatures which can be employed, resulting in much faster reaction rates, increasing the solubility of non-gaseous reactants and products, and reducing the incidence of microbial contamination by mesophilic organisms [17]. Industrially important thermostable proteases are usually produced using thermophilic strains belonging to the genus *Bacillus* [18].

The use of proteases at higher temperatures is advantageous, because unfolded form of proteinaceous substrate is better susceptible for the still active thermostable protease. It results in higher specific activities for proteases sourced from thermophiles and optimization of some industrial processes, particularly with the enzymes that are active at temperatures near 100°C. Stability in organic solvents makes the thermostable proteases useful for synthesis of high molecular weight peptides, carried out in reaction media with low water content [16, 19]. Performing enzyme reactions at elevated temperatures permit for higher substrate concentrations, lower viscosity, reduction of microbial contamination risk and high reaction rates [20]. With the industrially desirable attributes of very high stability at elevated temperature, in wide range of pH and resistance to detergents, chelators, organic solvents along with increased susceptibility of protein substrates at high temperature, proteases from thermophilic organism are of considerable biotechnological interest and may open unexplored avenues of biocatalysis which were otherwise limited due
to use of mesophilic proteases. Given the potential uses of the thermostable proteases and their high demand, there exists a need for the screening strains of thermophilic bacteria that produce proteases with industrially desirable characteristics and the development of industrial fermentation processes for the same.

1.8 Optimization of Protease Production

Process optimizations play a major role in industrial production processes. With particular regard to biotechnological production processes, in which even small improvements can be decisive for commercial success, process optimization is presently an undisputed component of the agenda of any commercial concern.

The overall cost of enzyme production and downstream processing is the major obstacle against the successful application of any technology in the enzyme industry. Researchers and process engineers have used several methods to increase the yields of alkaline proteases with respect to their industrial requirements. Recent approaches for increasing protease yield include screening for hyper-producing strains, cloning and over-expression, fed-batch, chemostat fermentations, and optimization of the fermentation medium through a statistical approach, such as response surface methodology. Extracellular protease production in microorganisms is also strongly influenced by media components, e.g. variation in C/N ratio, presence of some easily metabolizable sugars, such as glucose [21] and metal ions [22]. Besides these, several other physical factors, such as aeration, agitation rate, inoculum density, pH, temperature and incubation, also affect the amount of protease produced [22-24]. It is also important to produce the enzyme in inexpensive and optimized media on large scale for the process to be commercially viable [24]. In order to scale up protease production from microorganisms at the industrial level, biochemical and process engineers use several strategies to obtain high yields of protease in a fermentor. The industrial production of enzymes is mainly performed under submerged fermentation (SmF). Use of SmF is advantageous because of ease of sterilization and easier process control during fermentation [25]. Approximately 90% of all industrial enzymes are being produced by SmF, by using specifically optimized and genetically manipulated microorganisms [26]. In fermentation technology, improvements in the productivity of the microbial metabolite are achieved, in general,
via the manipulation of nutritional and physical parameters and by strain improvements as the result of mutation selection [27]. These measures can alter the product yield significantly.

### 1.8.1 Statistical Approach for Optimization of Protease Production

Traditional methods of optimization involve changing one independent variable while keeping the others fixed at a certain level. This single dimensional approach was laborious, time consuming, expensive and incapable of reaching the optimum due to the interactions among variables. Statistical methods viz. Box Behnken design and response surface methodology are an important strategy for seeking the significant variables and their optimal conditions. It has been successfully employed for optimizing medium ingredients and operating conditions in protease production [29-30].

Application of properly designed approaches with multi-factor models allows process and biochemical engineers to design scale-up strategies for increasing enzyme production.

#### Table 1.3 Statistical methods used to improve protease production from microorganisms

<table>
<thead>
<tr>
<th>Microorganism</th>
<th>Design</th>
<th>Software</th>
<th>Yield improvement</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacillus subtilis</em> NCIM 2724</td>
<td>Box-Behnken design</td>
<td>‘STATISTICA’ software</td>
<td>1.02 fold</td>
<td>[30]</td>
</tr>
<tr>
<td><em>Bacillus sp.</em> HTS102</td>
<td>Central composite design</td>
<td>Design-Expert ver. 8.0.3</td>
<td>&gt;2.0 fold</td>
<td>[31]</td>
</tr>
<tr>
<td><em>Bacillus amyloliquefaciens</em> B7</td>
<td>Plackett-Burman design</td>
<td>Design Expert® 7.1.6 (Stat-Ease)</td>
<td>3.92 fold</td>
<td>[32]</td>
</tr>
<tr>
<td><em>Bacillus licheniformis</em> DSM</td>
<td>Plackett-Burman design</td>
<td>Design-Expert version 7.1.6 (Stat-Ease)</td>
<td>n.s</td>
<td>[33]</td>
</tr>
<tr>
<td><em>Penicillium citrinum</em> YL-1</td>
<td>Plackett-Burman and Box-Behnken design</td>
<td>Minitab 16.0 Design-Expert 7.0 (Stat Ease)</td>
<td>n.s</td>
<td>[34]</td>
</tr>
</tbody>
</table>

n.s. Not specified


1.9 Purification of Proteases

The development of techniques for protein purification has been an essential prerequisite for many of the advancements made in biotechnology. Protein purification varies from simple one-step precipitation procedures to large scale validated production processes.

Occasionally when a sample is readily available, purity can be achieved by simply adding or repeating steps. However, experience shows that even for the most challenging applications, high purity and yield can be achieved efficiently in fewer than four well-chosen and optimized purification steps. To minimize reduction of the overall yield, high-resolution separations such as chromatography should be utilized as early as possible. Most purification schemes involve some form of chromatography. Different chromatographic techniques with different selectivity can form powerful combinations for the purification of any biomolecule. Protease can be purified by a combination of chromatographic procedures such as affinity chromatography [35], ion-exchange chromatography [36], hydrophobic interactions and gel filtration [37-38]. For thermostable enzymes, purification processes involving one step or combination of all the steps has been reported [39-41].

1.10 General Properties Thermostable Proteases

The enzymatic and physico-chemical properties of thermostable proteases have been studied extensively and are given below:

1.10.1 Optimum Temperature and Stability

There are mainly two factors which affects the heat stability of enzymes alone or in combination. First is primary structure of enzyme. More the number of hydrophobic groups, more condensed will be the structure and the enzyme will not denature easily. Disulphide bridges impart stability against heat inactivation and chemical denaturation. Secondly, specific molecules such as polysaccharides and divalent ions confer stability to enzyme [40]. Generally thermostable proteases are identified to be active over a wide range of temperature. The optimum temperatures of
thermostable proteases range from 60°C to 100°C. Some exceptions like the protease from an extreme thermophile \textit{Pyrococcus furiosus} showed a very high optimum temperature of 115°C \cite{41}. The optimum temperature and thermal stability of some thermostable proteases from various thermophiles is summarized in the Table 4.1.

\begin{table}[h]
\centering
\caption{Properties of various alkaline proteases produced from thermophilic microbial Sp.}
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Thermophilic Strain} & \textbf{Mol Wt (kDa)} & \textbf{Temp (°C)} & \textbf{pH} & \textbf{Reference} \\
\hline
\textit{Aeropyrum} Pernix & 34 & 90 & 8 & [53] \\
\hline
\textit{Bacillus} sp & 33 & 55 & 10-11 & [54] \\
\hline
\textit{Aquifex aelocic} & 54 & 80 & 8.5 & [55] \\
\hline
\textit{Aquifex pyrophilus} & 43 & 85 & 9 & [56] \\
\hline
\textit{Bacillus pumilus CBS} & 34.6 & 65 & 10.6 & [57] \\
\hline
\textit{Bacillus subtilis} & 65 & 8 & 8.0 & [58] \\
\hline
\textit{Fervidobacterium islandicum} AW-1 & >200 & 100 & 9.0 & [59] \\
\hline
\textit{Brevibacillus thermoruber} LII & 36 & 60 & 8.0 & [60] \\
\hline
\textit{Bacillus} alcalophilus TCCC11004 & 26 & 50 & 11 & [61] \\
\hline
\textit{Bacillus strain HS08} & 30.9 & 65 & 7.5 & [62] \\
\hline
\textit{Bacillus strain HUTBS71} & 49 & 65 & 7.8 & [63] \\
\hline
\textit{Bacillus strain HUTBS62} & 49 & 80 & 6.8 & [48] \\
\hline
\textit{Bacillus stearothermophilus} TLS33 & 71 & 85 & 8.5 & [64] \\
\hline
\textit{Bacillus subtilis} & 45 & 60 & 9.1 & [65] \\
\hline
\textit{Bacillus licheniformis} RSP-09-37 & 55 & 50 & 10 & [66] \\
\hline
\textit{Bacillus cereus} MCM B-326 & 36-40 & 55 & 9.0 & [67] \\
\hline
\textit{Bacillus subtilis} BP-36 & 40 & 60 & 9.0 & [68] \\
\hline
\end{tabular}
\end{table}
1.10.2 Optimum pH

Enzymes are protein molecules containing side chains of weak basic and acidic amino acids i.e. they are amphoteric in nature and hence the surface charge of enzyme molecules changes with change in the pH of the environment. These effects are especially important in the zone of the active sites, which affects the activity, structural stability and solubility of the enzyme (Chaplin, 1990). Thermostable proteases are generally alkaliphilic in nature active in the range of pH 9-11 [42-44] with a few exceptions like *pyrococcus furiosus* active at pH 6 [44] and *sulfolobus acidocaldarius* active at pH 2 [45].

1.10.3 Molecular Weight

The proteases produced by thermophiles have molecular weight in range of 25 kDa [46] to 66 kDa [47] however there are exceptions like *Pyrobaculum aerophilum* with very high molecular weight of 401 kDa.

1.10.4 Metal Ion Requirement of Thermostable Proteases

Thermostable proteases require a divalent cation like Mg$^{2+}$, Ca$^{2+}$, and Mn$^{2+}$ or a combination of these cations, for maximum activity. These cations were found to enhance the thermal stability of a *Bacillus* alkaline protease. It is thought that these cations protect the enzyme against thermal denaturation and play a vital role in maintaining the active conformation of the enzyme at high temperatures [48].

1.10.5 Effect of Surfactant and Detergents

The thermostable proteases are most commercial viable enzyme and widely applied for the tannery and detergent industries, due to its unique properties. More and more proteases were explored from varied categories of microbes and characterized from this point of view. The protease to be used in the detergent preparations must be stable and compatible with all detergent constituents. The most important detergent components are the ionic surfactants like sodium dodecyl sulfate (SDS) and sodium laurylsulfate (SLS), nonionic surfactants such as t-octyl phenoxy polyethoxy ethanol.
(Triton X-100), polyoxyethylene (20) sorbitan monolaurate (Tween 20), and polyoxyethylene (20) sorbitan monoleate (Tween-80), oxidizing agents such as hydrogen peroxide (H₂O₂) and sodium perborate (NaBO₃), and bleaching agents like sodium hypochlorite (NaClO) [49]. Some alkaline proteases were stable with sodium dodecyl sulphate (SDS) and sodium linear alkyl benzene sulphonate (Joshi et al. 2008). Alkaline protease from Bacillus clausii was highly stable with 5% SDS and 10% H₂O₂ [50]. The thermostable proteases from the thermophilic Bacillus sp. were stable in the presence of SDS, Triton X-100 and Tween-80 [18, 39, 53-54].

### 1.11 Molecular Approaches of Proteases

Many worldwide corporations have recognized the bio-based technologies as one of the key drivers of sustainable growth. However, the biological process is often considered only when the chemical arsenal has failed to achieve synthesis of the target molecule. This is primarily because the unavailability of the desired enzyme to catalyze the reaction in an efficient manner. The exploitation of new types of enzymes, improvements of enzyme properties and of the production processes are overall goals of innovation in the enzyme manufacturing industry. Accordingly, systematic methods in the field of enzyme and reaction engineering have allowed access to means to achieve the ends, i) screening for novel enzymes from natural samples with improved characteristics as a good starting point, ii) engineering the existing enzymes using genetic engineering approaches, iii) fining the enzyme processes in the enzyme manipulation to overcome catalyst limitation, e.g. downstream processing in enzyme manufacturing, formulation of enzyme preparations and enzyme immobilization, etc. [69].

### 1.11.1 Cloning and Overexpression of Microbial Proteases

Advancement in biotechnology, especially in the area of genetics and protein engineering, have opened a new era of enzyme application in many industrial processes and experiencing major R&D initiatives, resulting not only in the development of a number of new products but improvement of in the process and performance of several existing processes also [70]. Gene cloning is a rapidly progressing technology that has been helpful in improving our understanding of the
structure-function relationship of genetic systems. It provides an excellent method for the control and manipulation of genes. More than 50% of the industrially important enzymes are produced from genetically engineered microorganisms now. Several reports have been published on the isolation and manipulation of microbial protease genes [72-76] with the aim of (i) enzyme overproduction by the gene dosage effect, (ii) studying the primary structure of the protein and its role in the pathogenicity of the secreting microorganism, and (iii) protein engineering to locate the active-site residues and/or to alter the enzyme properties to suit its commercial applications. Many protease genes from fungi, bacteria, and viruses have been cloned and sequenced.

The objectives of cloning bacterial protease genes have been mainly the over production of enzymes for various commercial applications in the food, detergent and pharmaceutical industries [77]. The virulence of many bacteria is related to the secretion of several extracellular proteases. In these microbes gene cloning was studied to understand the basis of their pathogenicity and to develop therapeutics against them. Proteases play a significant role in cell physiology, and protease gene cloning, has been attempted to study the regulatory aspects of proteases especially in *E. coli*. [43, 78-79].

Plasmid transformation is a crucial and efficient biotechnological tool that enables the enhancement of many important microbial characters that would be beneficial in a lot of industrial, agricultural and environmental applications (Sabir & El-Bestawy, 2009).

Many *Bacillus* proteases have been characterized and their encoding genes have been cloned and sequenced. Genetic transformation (plasmid transfer) is the process through which exogenous DNA is introduced into a receiving cell to permanently alter its heredity such as overproduction of cellular *Bacillus* protease and to increase the corresponding gene copy number [80].
1.11.2 Protein Engineering: Improving Stability and Catalytic Behavior

The tailoring of the microbial enzymes has become a trend in the field of protein engineering to overcome the limitations of natural biocatalysts and to develop process-specific enzymes. Scientists have been attempting to generate enzymes which can withstand harsh and unfavourable conditions prevailing in industrial processes. The tolerance to high or low temperatures, exhibiting activity in the alkaline or acidic environments, high performance in non-aqueous media and others are a few of the requisite properties. The ultimate goal is to redesign the proteins in such a way that the industrial processes can be carried out in a more economic and greener way [13].

The greatest contribution of proteases occupies in the application in detergent industry, so that efforts are being made for novel proteases in this field. Moreover, proteases are being used in various cleaning and washing process as an additive for the removal of protein pollution related to environmental issues. Even though, the opportunities for protease use are numerous, identification of enzyme with desirable properties remains one of the key limiting steps in commercial application of protease [81]. For this purpose, the stability and functionality of proteases under unfavorable industrial conditions is of great interest. The term “stability” refers to the resistance of proteins to adverse conditions such as heat or denaturants, that is, to the persistence of its molecular integrity or biological function in the presence of high temperatures or other deleterious influences.

Naturally occurring enzymes are usually not optimally suitable for industrial applications due to their reduced performance under industrial process conditions. The stability of an enzyme is affected by many factors, such as pH, temperature, solvent properties, oxidative stress, the binding of metal ions or co-factors, and the presence of surfactants. Among these variables, the effect of temperature is the best studied. Mainly, there are three biotechnological approaches to generate stable and improved variants of a given protein: i) isolating enzymes from organisms living in suitable extreme environments [82]; ii) rationale-based mutagenesis [83] and iii) directed evolution [84]. Very interestingly, proteases were among the first enzymes which were targeted for directed evolution study, mainly because of the need for
stable and functional variants in detergent applications. In one of the early studies on family shuffling of 26 subtilisins from *Bacillus* species for multiple properties: activity at 23°C and at three different pH values, thermostability and solvent stability (35% DMF) [85]. It was reported that the best variants had up to three times more residual activity after heat treatment or up to 50% greater residual activity in DMF. This indicated that new, valuable combinations of properties had been achieved, including considerable increases in stability. The directed evolution application in protease stability is mainly on subtilisins, such as subtilisin E and its homologs. For instance, subtilisin BPN’, an engineered industrial protease (with Y217L mutation) derived from *Bacillus amyloliquefaciens* has made entry to the market and is sold under the brand name Purafect Prime®. Purafect Prime® exhibits alkali stability because of the replacement of Y217 (positioned close to the substrate binding region of the enzyme and prone to ionization in alkaline environment) with a leucine (L) [13].

Some trends in amino acid composition have been observed, e.g. non-polar residues tend to substitute uncharged polar residues (Ser, Asn, Thr and Gln) that favors increase in hydrophobic interactions which correlate with elevated thermostability in thermophiles [85-88]. An increased residue volume [85-86, 89-90], higher number of intramolecular hydrogen bonds [71-74] and larger number of α-helices [76-77] have been suggested as contributors for the stabilization of proteins from thermophiles.

The thermodynamic thermal stability of a protein can be estimated by the value of its \( T_m \), the midpoint of the thermal transition. \( T_m \) is the temperature at which half of all protein molecules are in native state and the remaining half are in the denatured state [93]. The denatured state is an ensemble of many conformations and possesses high conformational entropy [94]. Higher values of \( T_m \) result in an increase in thermal stability. Molecules in the denatured state are more likely to be affected by proteolysis, aggregation, adsorption and precipitation. These processes are usually irreversible and result in continuous decrease in the concentration of the protein in the native state.
Disulfide bridging between cysteine residues is an important tertiary structural element that is paramount in determining the overall structure of a protein. Organisms within all domains of life have adapted their own systems of keeping proper bridging in check, as some are favorable and some completely inactivate enzymes. Within thermostable enzymes these structural elements are significant, since they have been shown to increase stability within thermophilic proteins and play a role in preventing alteration of quaternary structure [95-96]. Yang et al. [97] attempted to enhance the stability by the formation of a disulfide bridge between two molecules of Ser236Cys subtilisin E and found that Ser236Cys subtilisin E has remarkably enhanced thermostability at 60 °C. Meizhi and Xiongwei [98] reported the improved thermostability of three nattokinase mutants (G61C/S98C, T22C/S87C and S24C/S87C) by disulfide bond formation in protein engineering.

1.11.3 Protease Engineering by Directed Evolution

Most common molecular level approaches in protein engineering to generate and isolate an enzyme having tailored characteristics are directed evolution and rational design.

During the past decades directed evolution has been a promising and successful approach for tailoring protein properties to meet industrial demands and advancing our understanding of structure-function relationships of biocatalysts. The concept of directed evolution was introduced in 1967 [99]. Directed evolution is a method developed to overcome the limitations of natural enzymes by tailoring their properties to the needs of a particular application or working conditions and simultaneously identify and investigate structure–function relationships [100]. Besides improving an activity that already exists, directed evolution can be used for combining properties not necessarily found together in natural enzymes [101]. It has been successfully applied in the field of biotransformations, biosensors, bioremediation, vaccines, therapeutic proteins and others [100].

The directed evolution mainly consists of three steps. i) Diversity generation ii) Screening of improved variants iii) Isolating the gene encoding the improved
protein variant. Repetitive cycles are usually required until the desired property has been obtained or improved significantly.

During the whole process, the first two steps are crucial for a successful directed evolution experiments. However, it is still a challenge to generate an unbiased mutant library and a challenge in setting up the screening systems.

1.12 Objectives of the Study

- To explore the studies of thermostable proteases from thermophilic bacterial strains for their potential application in detergent formulation as well dehairing agent in leather processing to uplift the socio-economic status of country
- Production of thermostable protease from thermophilic Bacillus subtilis BSP
- Statistical optimization of culture conditions for the selected thermophilic Bacillus subtilis BSP in shake flasks using response surface methodology
- Scale up study in stirred tank bioreactor of thermostable protease
- Purification and characterization of thermostable protease
- Kinetic studies of the purified thermostable proteases
- Cloning and expression of thermostable proteases into E. coli and Bacillus subtilis expression systems
- Directed evolution of the thermostable protease to improve the enzyme properties
- Identification of DNA encoding thermostable alkaline serine of Coprothermobacter proteolyticus DSM 5265 and subsequently cloning and expression in B. subtilis. Optimization of parameters affecting the expression level of the and the study of main technological properties of the recombinant crude enzyme preparation
2 Material and Methods

In order to avoid unnecessary repetition, techniques that are used in more than of the subsequent results chapters are described below.

2.1 Chemicals and Growth Media

Caseins were procured from Lab M Limited (UK) and Sigma Aldrich (USA), Tryptone, agar-agar, yeast extract (YE) and pepton were purchased from Acumedia (Michigan, USA). All chemicals were of analytical-reagent grade or higher quality and were purchased from Sigma Aldrich (St. Louis, MO, USA) unless otherwise specified. All enzymes were purchased from Fermentas (St. Leon-Rot, Germany) and New England Biolabs (Ipswich, MA, USA). Thermal cycler (Perkin Elmer GeneAmp PCR System 2400 Thermal Cycler, CT, USA) and thin-wall PCR tubes (Multi-ultra tubes, 0.2 mL, Carl Roth, Karlsruhe, Germany) were used in all PCRs. The PCR volume was always 50 µL. The amount of DNA in cloning experiments was quantified by using a NanoDrop photometer (ND-1000, NanoDrop Technologies, Wilmington, DE, USA).

Water used in experiments was purified to 18.2 MΩ resistances by passing through a MilliQ Purification system (Millipore) which was maintained at room temperature. Reagents used in molecular techniques such as restriction enzymes, PCR and cloning reagents were purchased from Fermentas and New England Biolab (NEB).

Growth media used in the study were TBAB, Luria Bertani (LB), Basal salt, LB casein and LB starch. Compositions of all growth media are given in appendices. Inoculum was prepared by transferring a loop full of cells from LB agar slants to 250 mL Erlenmeyer flask containing 50 mL LB (Luria Bertani) medium and incubated at desired temperature in orbital incubator shaker (Gallenkamp, UK) set at 180 rpm for 24 h. Inoculum at a concentration of 1% (v/v) was used to seed the production media.
2.2 Microorganisms

Protease producing thermophilic Bacillus subtilis BSP strain obtained from fermentation technology group of NIBGE (previously isolated and identified by Dr. Romana Tabassum, DCS IBD) was employed in the present study. Escherichia coli JM109 was used as a host for plasmid propagation and Escherichia coli BL21, Bacillus subtilis 1A751 and Bacillus subtilis PY79 were used as an expression host. The pET29b(+), pSG1154 and pDR111 vectors were used for cloning and expression.

2.3 Maintenance and Preservation of the Bacterial Strains

Proteases producing bacterial strains were maintained in LB (see appendix 2) supplemented with respective antibiotic (if necessary). Glycerol stocks (50% v/v) were prepared by aseptically mixing bacterial cultures grown in LB and the 50% glycerol. The stocks were preserved at -80°C for months.

2.4 Detection of Protease Activity

Proteolytic activity in liquid was determined using the EnzChek protease assay kit (Molecular Probes, Eugene, Oreg.) essentially according to the manufacturer's instructions and read using a spectrophotometer with an excitation/emission maxima at 590/645 nm. Values from controls without added enzyme was subtracted from sample values.

Protease sample corresponding to 0.5 µL initial protease reaction was transferred to the wells of a 96-well black plate. The EnzChek Protease Assay kit (Life Technologies, Carlsbad, CA, USA) consists of casein protein that is labeled with green-fluorescent BODIPY FL dye. The conjugation density is such that the fluorophore is quenched until smaller fragments are released into solution upon cleavage by proteases; this approach eliminates the need for a TCA precipitation step. Protease was added to 95 µL reconstituted BODIPY FL-casein (10 µg/mL) and 100 µL 1X digestion buffer (10 mM Tris-HCl, pH 7.8) in wells of a 96-well black plate and incubated in the dark at required
temperature. Fluorescence was measured after 20 min incubation to obtain maximal sensitivity as recommended by the manufacturer.

2.5 Assay for Protease Activity

The proteolytic activity was determined using a modified version of the method described by Anbu [88] using casein as the substrate. In general, 1 mL of enzyme and 1 mL of 1% casein were pre incubated separately at the desired temperature for 30 min. In all cases, the pH of the enzyme and casein substrate was measured and adjusted prior to pre incubation. After pre incubation, the enzyme and casein were mixed and incubated for 20 min. The reactions were stopped by the addition of 3 mL of 5% trichloro acetic acid (TCA). After 10 min, the reactions were centrifuged for 5 min at 11 K×g, and the amount of released tyrosine was measured at 280 nm.

2.5.1 Principle

The absorbance of UV light by tyrosine and tryptophan residues is high at 280 nm. The peptides mainly tyrosine which are liberated in proteolytic digestion are measured in a filtrate of Trichloroacetic acid (TCA), by means of their absorbance at 280 nm. The activity of protease is estimated in “Units”. One unit is defined as the liberation of 1 mg of tyrosine in one min equivalent of substrate per mL of enzyme under standard assay conditions.

2.5.2 Standard Curve of Tyrosine

One milligram of tyrosine was dissolved in a 1 mL of distilled water and the final volume was adjusted to 100 mL. Each mL of the solution thus contained 1 µg of tyrosine. Nine standard solutions were prepared from this stock solution by serial dilutions ranging from 10.0 - 100 µg/mL. The absorbance of the samples was measured at 280 nm with the help of spectrophotometer and calibration curve was plotted for each set of assay.
2.6 Protein Assay

The protein concentration was determined by the Bradford method [72]. The amount of protein was measured in terms of the absorbance at 280 nm. The specific activity was expressed as the enzyme activity per mg of protein. Protein concentrations were also obtained using NanoDrop 1000 Spectrophotometer (Thermo Scientific, USA) according to manufacturer’s instructions.

2.7 Optimization of Physico Chemical Parameters

The effect of various physico-chemical and nutritional parameters (substrate, time, inoculum concentration, initial pH and temperature) on protease production in 250 mL flasks was studied.

2.8 Sodium Deodecyl Sulphate-Polyacrylamide Gel Elecrophoresis (SDS-PAGE) Gel Elecrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the Laemmli method [103]. Resolving and stacking gels were prepared (see appendix 6) and allowed to polymerize at room temperature. The compositions of resolving and stacking gels were given in appendices.
2.8.1 Resolving Gel

Two glass plates of 8cm x 10 cm dimensions and 1.5 mm spacers were assembled to set up the gel cast. Resolving gel mix was poured till ¾ th of the size of the gel and butanol was over layered to remove air bubbles.

2.8.2 Stacking Gel

Once the gel polymerized, butanol was decanted and comb was inserted in the head space. Following this, stacking gel mix was overlaid and allowed to polymerize.

After polymerization, the comb was removed and the wells were washed with electrode buffer. The protein sample was mixed with sample buffer and heated at 100°C for 5 min in a boiling water bath. The samples were loaded in to respective wells and using 1X electrode buffer the gel was electrophoresed at 100 V for 1 h. After electrophoresis, the gel was washed for 40 min in 500 mL 2.5% Triton X-100 buffer at room temperature for the removal of SDS. Gel was the incubated in 500 mL incubation buffer (100 mM Tris-HCl pH 7.8) for 20 h at 37°C. After incubation the part of the gel containing different molecular weight markers was stained with Comassie R-250 for 1-2 h. The gel was destained with a solution of methanol, acetic acid and water with a ratio of 9:2:9 respectively.

2.9 NuPAGE® Pre-Cast Gel Electrophoresis

For gel electrophoresis using NuPAGE® Bis-Tris Precast Gels, protein samples were prepared by heating for 10 min in NuPAGE® LDS Sample Buffer (4X). It is used to prepare protein samples for denaturing gel electrophoresis with NuPAGE® Novex® gels. Optimal protein sample preparation for polyacrylamide gel electrophoresis (PAGE) requires denaturing and reducing protein disulfide bonds. It contains lithium dodecyl sulfate, pH 8.4, which allows for maximum activity of the reducing agent. Samples were denatured by heating for 10 min at 70°C. Protein separation was done in NuPage 4-12% Novex® Bis Tris gels (Invitrogen) in MOPS buffer (composition) Reducing agent was
added to the upper buffer chamber to ensure that reduced samples did not reoxidise during electrophoresis. Gels were run at 200 V constant voltages for one h.

2.10 Zymography for Activity Staining

Substrate gel electrophoresis was carried out using a modified version of the method described by [104] using 10% polyacrylamide running gel polymerized in the presence of 0.1% gelatin and 4% stacking gel without gelatin were used. Following electrophoresis, the SDS was removed by incubation with 2.5% Triton X-100 for two hours at 25 °C, and the gel was then incubated in neutral buffer (0.1 M Tris-HCl, pH 7.8 containing 1 mM CaCl$_2$) at 37 °C for 24 hours. The gel was stained with Coomassie blue, and protease bands were identified as clear areas in the gel after being destained with destaining solution (40% methanol and 10% acetic acid in H$_2$O).

2.11 Molecular Characterization

2.11.1 DNA Isolation

Genomic DNA was extracted from bacterial isolates by CTAB (Cetyl Tri-methyl Ammonium Bromide) method. Bacterial cells were grown in LB medium at 60°C overnight and harvested by centrifugation at 7500×g. Cells were re-suspended in 5 mL T.E buffer & 20 mg lysozyme was added. Suspended cells were incubated at 37°C for 5 minutes. Then 500 µL of 10% SDS, 25 µL proteinase K (25 mg/mL) and 3 µL RNase were added to the incubated suspension. The contents were mixed thoroughly and incubated at 37°C for 10 minutes. After incubation, 0.9 mL 5M NaCl & 0.75 mL NaCl/CTAB were added, mixed thoroughly and incubated at 65°C for 20 minutes. The protein contents were extracted (twice) with an equal volume of Phenol-chloroform-isoamyl alcohol (25:24:1) and centrifuged for 10 minutes at 7500×g and 4°C to separate the organic and aqueous phase (containing DNA). Aqueous supernatant was separated in fresh tube and 0.6 volume of isopropanol was added and mixed thoroughly until a white DNA thread precipitated out of solution and condensed into a tight mass. The DNA was
pelleted by centrifugation at 7500 g for 5 minutes and supernatant was discarded. Finally the pellet was washed with 70% ethanol (twice) and dried at room temperature. Dried pellet was re-suspended in 100 µL T.E for further use.

2.11.2 Genomic DNA Isolation by FastDNA kit from Bio101

Genomic DNA was also extracted from bacterial isolates using FastDNA kit. From bacterial cells grown in LB medium at 60°C overnight, took 1 mL culture in lysing tube, spun down and trashed supe. Then 1 mL of CLS-TC lysis buffer was added and homogenized in FastPrep instrument for 10 seconds at a 5 setting. The mixture was spun for 15 min at maximum speed in centrifuge. In 750 uL of supernatant, 600 uL binding matrix solution was added and rotated for 5 min at room temperature then spun for 5 seconds and trashed supernatant. Matrix was resuspended in 500 uL SEW-M wash buffer then added the other 400 uL SEW-M. Transferred the matrix to spin filter and spun down for 1 min in centrifuge. It was made sure that wash buffer came out clean by repetitive addition of 500 uL SEW-M wash buffer to the top and spin (Did not resuspend the pellet). A final spin of spin column was done to dry it out. Filter unit was transferred to new catch tube and resuspended the pellet in 100 uL DES (H2O). Kept at room temperature for 3 min and then spun for 1 min and DNA was collected.

2.11.3 Protein and DNA Concentration Measurement

Protein and DNA concentrations of various samples were obtained using NanoDrop 1000 Spectrophotometer (Thermo Scientific, USA) according to manufacturer's instructions.

2.11.4 Preparation of Heat Shock Competent Cells

Competent E. coli were prepared by the method described by Cohen et al. [104]. Using a sterile wire loop, a single colony from freshly grown plate of E. coli (Top 10, JM109) was transferred into a 50mL flask containing 20mL LB medium. The flask was incubated at 37°C overnight with vigorous shaking at 220rpm (Pamico technologies Pakistan,
Model GLSC-OSI-HC-196-10). The next morning 1 mL of the culture was transferred aseptically into a flask containing 100mL Lauria bertani (LB) media (1% tryptone, 0.5% yeast extract and 1% NaCl) and incubated in shaker at 220 rpm at 37°C until an OD at 600 nm (OD600) of 0.5-1 was attained. The flask was placed on ice for 30 minutes and then 50mL of bacterial culture was transferred aseptically to pre-chilled sterile propylene tubes and centrifuged (Eppendorf, 5810R) at 3220xg for 8-10 minutes at 4°C to pellet the cells. The supernatant was discarded and the cell pellet was resuspended in 20 mL of 0.1M MgCl₂ and centrifuged again. The cell pellet was resuspended in 20mL of 0.1M CaCl₂, incubated on ice for 30 minutes and then centrifuged as above. Finally the cell pellet was resuspended in 3-4 mL of 0.1 M CaCl₂. Filter sterilized cold glycerol (200 µL per 1 mL of CaCl₂) was added. Cells were stored in aliquots of 200 µL at -70°C. All the steps were performed under aseptic conditions.

2.11.5 Primer Designing

Primers were designed and synthesized based on the protease gene sequences of thermophilic bacterial strains from Genbank to amplify the potential protease genes from genomic DNA isolated from thermophilic bacterial strains. Primer description, concentrations and annealing temperatures for each primer set are provided in each respective chapter. Oligonucleotide primers were synthesized by Elim Biopharmaceuticals Inc CA, USA.

2.11.6 Amplification of Protease Gene

PCR reactions were performed in 50 µl reaction volumes containing

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward Primer</td>
<td>5.0 µL</td>
<td>(2.5 µM)</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>5.0 µL</td>
<td>(2.5 µM)</td>
</tr>
<tr>
<td>Genomic DNA</td>
<td>1.0 µL</td>
<td>(50 ng/µL)</td>
</tr>
<tr>
<td>10XPCR reaction buffer</td>
<td>5.0 µL</td>
<td></td>
</tr>
</tbody>
</table>
dNTP \hspace{1cm} 1.0 \mu L (10 \text{ mM})

pfu DNA polymerase \hspace{1cm} 0.5 \mu L (2.5 \text{ U/\mu l})

Water \hspace{1cm} 31.5 \mu L

All the ingredients were added to PCR tube and the following PCR conditions were set.

**PCR conditions for amplification of protease gene**

<table>
<thead>
<tr>
<th>Steps</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step-1</td>
<td>Initial denaturation 94°C, 4 min</td>
</tr>
<tr>
<td></td>
<td>Denaturation 94°C, 1 min</td>
</tr>
<tr>
<td>Step-2</td>
<td>Annealing 60-65°C (Depending on T_m of primer pair), 30 sec</td>
</tr>
<tr>
<td></td>
<td>Extension 72°C, 2 min</td>
</tr>
<tr>
<td></td>
<td>Repeat step 2 for 30 Cycles</td>
</tr>
<tr>
<td>Step-3</td>
<td>Final extension 72 °C, 10 min</td>
</tr>
<tr>
<td></td>
<td>Hold at 4°C</td>
</tr>
</tbody>
</table>

All the process for the preparation of reaction mixture was carried out on ice. After the completion of PCR the amplified product was run on 0.9% agarose gel along with NEB 1Kb ladder to estimate size of amplicon.

**2.11.7 Agarose Gel Electrophoresis**

Agarose gel electrophoresis is a method used to separate and identify DNA on the basis of their size and rate of movement through a gel under the influence of an electric field. Since the nucleic acids are negatively charged due to the negatively charged phosphate group, they will migrate to the positively charged electrode (anode). Distinct bands, based on the molecule size, will form on the gel. Shorter DNA molecules will move faster than longer, since they are able to slip through the matrix more easily. The band can be visualized by staining the DNA with ethidium bromide, which causes the DNA to
fluoresce in ultra violet light. Ethidium bromide is a ring-formed structure that intercalates between the base pairs in the DNA double helix. It is a mutagen and was handled with care.

The amplified PCR products were analyzed on 1% agarose gel in 0.5X TAE (20 mM Tris-acetate and 0.5 mM EDTA (pH 8.0) buffer. The agarose was dissolved in the buffer by heated for 2 minutes in microwave oven. After cooling the clear solution to 55°C, ethidium bromide (EtBr; 0.5 μg/mL) or SYBR® Safe DNA Gel Stain (Life Technologies, USA) was added and poured into the casting tray with a suitable comb. The gel was allowed to solidify at room temperature. The comb was removed carefully after placing the gel tray in the gel tank and submerging in 0.5X TAE buffer. Loading dye (Fermentas or NEB) was mixed with the DNA samples and loaded into the wells. DNA was resolved on the gel by applying 100 volts for 1 h followed by the examination of gel under UV transilluminator for analyzing the PCR products.

2.11.8 PCR Purification

PCR product was purified using Wizard®SV Gel and PCR Clean-Up System (Promega). Briefly, an equal volume of Membrane Binding Solution was added to the PCR amplification. Applied this dissolved mixture in SV minicolumn into collection tube and centrifuged at 16,000 × g for 1 min after incubation at room temperature for 1 min. Flow through was discarded and washed the column with 700 μL membrane wash solution (ethanol added). Centrifuged at 16,000 × g for 1 min and discarded flowthrough. This step was repeated with 500 μL Membrane Wash Solution. Then centrifuged the column for 5 min at 16,000 × g. Recentrifuged the column assembly for 1 min with the microcentrifuge lid open to allow evaporation of any residual ethanol. Minicolumn was carefully transfered to a clean 1.5 mL microcentrifuge tube and added 50 μL of nuclease-free water. Incubated at room temperature for 1 min and centrifuged at 16,000 × g for 1 min. Minicolumn was discarded and stored DNA at –20°C.
2.11.9 Restriction Digestion of PCR Product and Vector

Purified plasmids and PCR products were digested using restriction enzymes in their respective buffers as recommended by the supplier (New England Biolab). A total volume of 20 μL (containing 1 -2μg DNA, 10 units restriction enzyme, buffer and SDW) was used to digest the plasmid for cloning and sub-cloning, while to screen the plasmid for presence of clone, volume of digestion mixture was reduced to 10 μL (500 ng/ μL DNA, 04 units restriction enzyme, buffer and SDW). Reaction mixtures were kept at the optimum temperature (usually 37˚C) for 1 h. Sizes of DNA fragments of digested product were determined on ethidium bromide or cyber safe stained agarose gels by comparison to an appropriate co-electrophoresed DNA marker.

2.11.10 Vector Dephosphorylation

Dephosphorylation is a common step in traditional cloning workflows to ensure that the vector does not re-circularize during ligation. If a vector is linearized by a single restriction enzyme, or has been cut with two enzymes with compatible ends, use of a phosphatase, to remove the 5' phosphate reduces the occurrence of vector re-closure by intramolecular ligation. Decreased re-circularization reduces the background during subsequent transformation. If the vector is dephosphorylated, it is essential to ensure that the insert contain a 5' phosphate to allow ligation to proceed. Each double-strand break requires that one intact phosphodiester bond be created before transformation.

Double digested (pET29b(+), pSG1154, pDR111) vector was dephosphorylated using NEB dephosphorylation protocol. Briefly, 1/10 volume of 10X Antarctic Phosphatase Reaction Buffer and 1 μL of Antarctic Phosphatase (5 units) was added to 1-5 μg of double digested DNA and incubated for 15 min at 37˚C for 5´ extensions or blunt-ends. Then heat inactivated for 5 min at 65˚C.
2.11.11 Purification of DNA Fragments from Agarose Gel

Double digested PCR product and vector (dephosphorylated) was run on 0.9% agarose gel along with broad range DNA marker (New England Biolabs) to analyze the restriction patterns. Desired fragments (amplicon and vector) were excised from the agarose gel with a clean, sharp scalpel. DNA fragments were eluted from agarose gel using Wizard®SV Gel and PCR Clean-Up System as described above. Concentrations were measured using nanodrop and proceeded with ligation.

2.11.12 Quantification of DNA

The concentration of DNA samples isolated from bacterial strains, purified PCR and of purified plasmids was measured by using a NanoDrop 1000 Spectrophotometer (Thermo Scientific, USA). The absorbance was measured at 280nm.

2.11.13 Ligation

The ligation, below described, was next performed by incubating the treated plasmid and the purified gene overnight at 16°C.

- 2 µL ligation buffer (10X)
- 25 ng Vector DNA
- 75 ng Insert DNA (purified digestion product)
- 1 µL of T4 DNA ligase
- H2O to a total of 20 µL

Control was used without insert and 3 to 5 µL of ligated mixture was used for transformation into competent bacteria.

2.11.14 Transformation

Transformation of Heat-Shock Competent E. coli Cells

Competent E. coli cells were transformed according to previously described methods.
Cometent cells were taken out of the -70°C freezer and immediately placed on ice to thaw. Then the ligation mixture (2 μL) was added to the competent cells (100 μL) and after gentle mixing was incubated on ice for half an hour. The cells were then given a heat shock by incubating at 42°C in a heating block for 2 min. After incubating on ice for 2 min 1 mL of liquid LB media was added and cells were allowed to grow at 37°C for 1 h. Finally the cells were pelleted, most of the supernatant was discarded, and cells were resuspended in remaining medium and spread on solid LB media petri plates (containing suitable antibiotics) and kept inverted at 37°C for 16 h in an incubator.

**Transformation of E. coli Competent Cells provided by Promega**

Promega E. coli competent cells were transformed using quick transformation protocol of Promega (Promega, USA) according to manufacturer’s guidelines. Briefly competent bacteria (Promega E.coli JM 109) were allowed to thaw on ice for 5-10 min. For a single transformation 1 aliquot (50 μl) of bacterial suspension was mixed with 1 μL of ligation product in chilled sterile 17×100 mm polypropylene culture tubes, equaling approximately 1-50 ng of circular plasmid DNA. Flicked the tube several times and immediately after heat-shock (42°C for 45-50 s) the tube was placed on ice for 2 min. 450μl of cold (4°C) SOC (Media composition in appendices) medium was added to each transformation reaction and this suspension was incubated for 60 min at 37°C with shaking. For each transformation reaction, diluted the cells 1:10 and 1:100 and spread 100μl of the undiluted, 1:10 and 1:100 dilutions onto LB agar plates containing antibiotics (kanamycin 30 μg/mL, ampicillin 100 μg/mL) and incubated at 37°C for overnight.

**2.11.15 Verification of Recombinants**

**Colony PCR**

The presence of the recombinant plasmid containing the insert of interest in the colonies growing on Luria Broth agar containing appropriate antibiotic (kanamycin 30 μg/mL, ampicillin 100 μg/mL) plates was verified by PCR using the cells directly as a DNA template.
Colonies were marked and picked for direct utilization in the PCR tubes and mixed into the PCR reaction. The PCR reaction mix was set up as:

- H₂O 6.5 µL
- 10x PCR buffer 1 µL
- Primer 1 (2.5 uM) 1 µL
- Primer 2 (2.5 uM) 1 µL
- dNTP (10 mM) 0.2 µL
- Herculase (Stratagene, #600312-51) 0.2 µL
- Total Volume 9.9 µL

Clone culture From plates, small part of colony

The total volume was adjusted to 10µl without a DNA template. In each screen, the primers specific for the pET29b vector were used i.e. T7-rev-long ((TTATGCTAGTTATTGCTCAGCGGTGGCA)) and p29-708 (ATTCGATGGTGTCGGGGATCTCG).

The program for the PCR as:

**Cycle Settings**

1 cycle 95°C for 3 minutes
22 cycles 95°C for 1 minute
55°C for 1 minute

72°C for (1 min/kb, size of gene including flanking region of oligos)

The amplified products were separated on 0.9% agar gel containing TAE buffer (see appendix 5) and visualized as described in agarose gel electrophoresis protocol. Positive colonies showed a fragment of the size combining the insert flanking regions and
the insert size.

**Extraction of Plasmids from Bacterial Cells**

To screen the colonies for correct insert a selection of colonies from the transformation procedure was picked and transferred into tubes of LB medium containing the appropriate antibiotic (Kanamycin, 30 μg/mL) before the tubes were incubated at 37°C with shaking overnight. Plasmid DNA was isolated using QIAprep spin Miniprep Kit according to manufacturer’s guidelines. Briefly, 1.5 mL overnight culture of *E. coli* in LB medium was transferred to a microcentrifuge tube and centrifuged 30 seconds at 13000 rpm in a Microcentrifuge, Mini Spin (ependorf). The supernatant was discarded. The pelleted bacterial cells were resuspended in 250 μL Buffer P1 by vortexing, then 250 μL Buffer P2 (lysis buffer) was added and mixed carefully. The solution became viscous and slightly clear. For 5 min the solution was incubated at room temperature. 350 μL Buffer N3 (neutralization buffer) was added and mixed immediately, but thoroughly by inverting the tube. The solution was centrifuged for 10 min at 13000 rpm. A compact white pellet was formed. The supernatants from above step were applied to a QIAprep spin column by decanting and pipetting before it was centrifuged for 60 seconds. After discarding the flow-through, the QIAprep spin column was washed by adding 0.5 mL Buffer PB and centrifuged for 60 seconds. The flow-through was discarded. Further the QIAprep spin column was washed by adding 0.75 mL Buffer PE and centrifuged for 60. The flow-through was discarded, and centrifuged for an additional minute to remove residual wash buffer. The QIAprep spin column was placed in a clean 1.5 mL microcentrifuge tube and the DNA was eluted from the QIAprep column with 50 μL Buffer EB by adding it to the center of the membrane. The column was left for 1 min, before it was centrifuged for 1 min. The DNA yield of the elute was then determined using nanodrop.

**Cloning Confirmation by Restriction Analysis**

In order to verify the recombination, digestion of plasmids was performed using appropriate restriction enzymes and digestion products were analyzed in 0.9% agarose gel along with broad range DNA marker to analyze the restriction patterns.
Sequencing of Protease Genes

Plasmid clones for sequencing were purified using QIAprep spin Miniprep Kit (QIAGEN Inc CA, USA) and sequenced by the dideoxy chain termination method [105] using gene specific primers. The protease gene sequences were compared to the already known nucleotide sequences using BLAST search algorithm (http://www.ncbi.nlm.nih.gov/BLAST) and molecular evolutionary relationships were accomplished using MEGA 6.0 version with the Kimura two-parameter model and the neighbor-joining algorithm [106]. The sequences of the isolates obtained were submitted in GenBank.

2.11.16 Transformation of Recombinant Plasmid into E. coli Expression System

Plasmid isolated from the putative recombinants was transformed in E.coli BL-21 cells for expression of the protease under of isopropyl-β-D-thiogalactopyranoside (IPTG) induction (1 mM). Complete genes for thermophilic proteases were attempted to be expressed. The expressed proteins were activated into mature active protease by its autocatalytic cleavage at 60-70°C by heat treatment for 3-4 h.

2.11.17 Expression Analysis of Protease Enzyme

A liquid culture of the transformed colony was grown in Luria-Bertani (LB) broth by shaking at 37°C. Recombinant protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM when the optical cell density at 600 nm reached 1.5. After 12 h of additional growth, the bacteria were harvested by centrifugation and lysed with CellLytic™ B Cell Lysis Reagent (Sigma-aldrich St. Louis, MO, USA) containing lysozyme (Novagen) according to the manufacturer’s protocol as: Cells from 1 mL of culture were collected by centrifugation at 4,000×g for 20 min at 4 °C and then suspended in in 0.2 mL of CellLytic B 2X containing 0.2 mg/mL lysozyme. The solution was briefly vortexed to resuspend the cell pellet and mix for 20 minutes to ensure full extraction of the soluble proteins. Cell lysate
was centrifuged at full speed for 5 minutes to pellet any insoluble material. The supernatant was used for the detection of protease activity.

2.11.18 Subcloning of Protease Gene in *Bacillus-E.coli* Shuttle Vector

*Bacillus-E.coli* shuttle vector (pDR111 and pSG1154) were linearized by restriction digestion with appropriate restriction enzymes. Linearize vector DNA was undergone phosphorylation using calf intestinal alkaline phosphatase (NEB) as described above to prevent self ligation of the vector. Protease gene was amplified from cloned pET29b+ plasmids (in each case) using primers specifically designed for cloning in *Bacillus* strain. PCR product was purified using SV Wizard as described above and undergone restriction digestion by using the same restriction enzymes which were used to linearize *Bacillus-E. coli* shuttle vector. Restricted PCR product and vectors were run on 0.9% agarose gel and fragment were excised and purified as discussed above. DNA concentration was estimated and ligation reactions were set as described above. *E. coli* JM109 competent were transformed with ligation mixture according to the manufacturer guideline (Promega, USA) and spread on LB agar plate containing Ampicillin (100 µg/mL) for the selection of recombinants as described above. Plates were incubated at 37°C for overnight. Successful cloning was verified by PCR using the cells directly as a DNA template. Colonies were marked and picked for direct utilization in the PCR tubes and mixed into the PCR reaction. The PCR reaction mix was set up as described above. Colonies were picked and processed for plasmid DNA isolation using QIAprep Spin Miniprep Kit. Cloning was further confirmed by digesting the isolated plasmid was with appropriate restriction enzymes (specific for each primer pair) and sequencing thereafter.

**Sequencing of the Protease Gene**

Positive clones were used for nucleotide sequencing of insert to verify the correct gene frame. The plasmid with accurate gene sequence was transformed in *Bacillus subtilis* strain 1A751 for protein expression analysis.
Transformation of Positive Clone in Bacillus Subtilis 1A751 and PY79

The plasmids containing protease apr were transformed into Bacillus subtilis and integrated into the bacterial chromosome using standard protocols found on the BGSC website (http://www.bgsc.org/catalogs/Catpart4.pdf). Compositions of media used in transformation protocol are given in appendices.

Transformation Protocol

Recipient strain (1A751) was streaked on one-half of a Tryptose Blood Agar Base plate and incubated overnight (18 h) at 37°C. A few colonies were inoculated into 4.5 mL of Medium A in a 16×125 mm test tube that lacked visible scratches. The contents of the tube were mixed thoroughly. Read its optical density at 650 nm in the spectrophotometer and adjusted the OD$_{650}$ to be 0.1-0.2, maintaining the volume at 4.5 mL. The culture was incubated at 37°C with vigorous aeration. Read the OD$_{650}$ every 20 min, plotting OD$_{650}$ against time on semi-log paper. After a brief lag, the OD increased logarithmically. The point at which the culture left log growth was noted the graph points fell below the straight line. In B. subtilis genetics, this point is known as t0. It took 60-90 min of incubation and occurred at OD$_{650}$=0.4-0.6.

Incubation was continued for 90 min after the cessation of log growth (t90). 0.05 mL of this culture was transferred into 0.45 mL of pre-warmed Medium B in a 16×125 mm test tube. Set up one tube for each transformation, plus an extra for a DNA-less control.

Diluted cultures were kept incubating at 37°C with vigorous aeration for 90 min. At this point, the cultures were highly competent. Then 1 µg of DNA (plasmid) was added to the competent cells and incubated at 37°C with aeration for 30 minutes. Transformed cells were plated onto tryptose blood agar base (TBAB) containing spectinomycin (100 µg/mL) and incubated at 37°C for overnight.
2.11.19 Confirmation of Cloning

Plate Assay
Colonies from plates were picked and spotted on LB agar plates containing 1% starch for the selection of starch non hydrolysers, indicating insertional inactivation of amylases gene. Such colonies were picked and spotted and LB agar plate containing 1% casein and 1 mM isopropylβ-D-1-thiogalactopyranoside (IPTG) which formed casein hydrolysis zones confirming the presence of protease gene in there.

Confirmation of Clones by PCR
Cloning of protease gene in Bacillus subtilis was further confirmed by PCR amplification of protease gene from 1A751-MprT strain. For this purpose, genomic DNA from 1A751/pDR-BSP-MprT was isolated using FastDNA® KIT according to manufacturer’s instructions. Then PCR was performed using the same primer set which was previously used for the amplification of MprT gene and same conditions for PCR set up. 16sRNA gene amplification was used as control using primer set (27F/1525R) for 16sRNA gene.

2.12 Characterization of Wild and Recombinant Protease
The proteases were characterized for its biochemical properties to gauge its suitability for industrial application. Protease activity was measured at different pH (3.0-11.0) to find the optimum pH and stability of enzymes at various pH. The details of buffers and other conditions are discussed in respective section. Likewise optimum temperature and thermostability of protease were determined at various temperatures (40-90C). Moreover, effect of metal ions, inhibitors and reducing agents, stability of proteases in organic solvents, detergents, oxidizing agents and surfactants were also studied. These experiments are discussed in details in respective sections.
3 Optimization of Extracellular Thermostable Protease Production and its Characterization from Thermophilic Bacillus subtilis BSP Strain

3.1 Introduction

Proteases (EC 3:4, 11-19, 20-24, 99) (synonymous as peptidase or proteinase) constitute a very large and complex group of enzymes, widely utilized in a host of industries. Microorganisms are the preferred source for obtaining proteases because of their fast growth rate, easy to manipulate for getting highly stable enzymes through genetic engineering and require shorter time for production and purification steps [11]. Proteases execute a large variety of functions and have numerous applications in detergent, food, pharmaceutical and leather industries. Although there are many microbial sources available for protease production, only a few are considered as commercial producers. Of these, species of Bacillus dominate in the industry [9]. Due to enhancement of such demand of proteases for specific properties, scientists are looking for newer sources of proteases. About 30-40% of the cost of industrial enzymes depends on the cost of the growth medium [107]. In addition, extracellular protease production from Bacillus species is significantly influenced by medium composition and some physical factors, such as fermentation period, aeration, inoculum density, incubation temperature and pH of growth medium [108].

The growth of thermophilic bacteria at high temperature and production of extracellular thermostable enzymes was attributed to the possibility of increasing their enzyme application. Therefore, these microorganisms were the first candidates for massive enzyme production for industrial applications [109]. Bacillus species are major sources of industrial and commercial valuable enzymes however some of thermophilic Bacillus species have been reported to produce useful enzymes [11]. Few thermophilic strains such as Bacillus stearothermophilus [64], Bacillus licheniformis [46] and
Bacillus pumilus [110] were explored for protease production. However, the earlier studies did not deal with the stability and the application evaluation of proteases, which offers new possibility and potential for their biotechnological applications.

In order to scale up protease production from microorganisms at the industrial level, biochemical and process engineers use several strategies to obtain high yields of protease in a fermenter. The submerged, solid state and fed-batch systems are reported to produce the enzymes from bacteria and fungi. Although fed-batch and solid state fermentation techniques are reported to be more efficient than SmF, approximately 90% of the industrial enzymes are produced by SmF techniques with enzyme titers in the range of g/L [26]. Moreover, enzyme production processes involve aerobic fermentation. Efficient production always depends upon a critical bacterial cell biomass, which in turn, depends on the availability/proper distribution of oxygen to the growing cells. In shake-flask batch fermentation, only regulation of agitation speed is possible, while the aeration remains uncontrolled. This may lead to reduced availability of dissolved oxygen (DO) to the growing bacterial biomass, ultimately leading to decreased protease production [111].

In recent years, there has been a great amount of research and development effort focusing on the use of statistical approach methods, using different statistical software packages during process optimization studies, with the aim of obtaining high yields of protease in the fermentation medium. The application of properly designed approaches with multi-factor models allows process and biochemical engineers to design scale-up strategies for increasing enzyme production.

Though Bacillus licheniformis is the model system for the production of thermostable alkaline protease, thermotolerant and thermophilic strains are very less reported. Moreover commercial production of protease for detergent and leather industries are in the budding stage in the developing country like Pakistan. Having based on the ideal properties of thermostable protease, the aim of the study was protease production from proteolytic thermophilic Bacillus sp through optimization of nutritional/physical parameters by conventional (one-variable-at-a-time) and statistical approach (RSM) at
shake flask level. The enzymatic properties of this thermostable protease are also reported.

3.2 Materials and Methods

3.2.1 Microorganism and Inoculum Preparation

The bacterial strain used in this study was a thermophilic Bacillus subtilis BSP strain previously isolated from a soil sample. Production of the protease was carried out in Basal salt medium (see appendix 1). The pH was adjusted to 6.9-7.0 with 1.0 M NaOH and this basal medium was sterilized. The production medium (50 mL in 250 mL Erlenmeyer flasks) was inoculated with 1 mL of an overnight culture and incubated at 60°C under shaking (180 rpm) in an orbital shaker (Gallenkamp U.K). The fermenting broth (1.0 mL) was periodically drawn aseptically, and growth was assessed by turbidity measurement at 660 nm. Each sample was then centrifuged at 7500 g (4°C) for 15 min, and cell-free supernatant was used for protease assay.

3.2.2 Protease Assay

Protease activity assay was performed by some modifications of Lowry method [99] using casein at alkaline (pH-10) as substrate. The protein hydrolysate was estimated by using tyrosine as standard as described in general material and methods section 2.5.

3.2.3 Optimization of Culture Conditions

Initial screening of significant parameters was carried out by one-variable at-a-time approach in submerged fermentation system by keeping all factors at a constant level in the basal medium, except the one under study. The media (50 mL) were prepared in 250 mL Erlenmeyer flasks and autoclaved at 121°C for 20 min. Bacillus subtilis BSP was inoculated and incubated on a rotary shaker (180 rpm) for 72 h. An aliquot of 2 mL was collected in every 12 h interval, centrifuged at 10,000×g for 15 min to remove bacterial cells and the supernatant was used to check the protease activity. The growth
of the microorganism was determined by measuring absorption at 600 nm. All experiments were carried out in triplicate and average values were reported. Initial pH of the medium (3-10.0), temperature (40-90°C), substrate (casein) concentration (0.5-3.0%) and time of incubation (0-72 h) were the physical parameters studied for their effect on bacterial growth and protease production.

### 3.2.4 Statistical Optimization of Factors Affecting Protease Production by Response Surface Methodology (RSM)

The conventional method of optimization involves variation of one parameter at a time while keeping the others constant. This is a very time consuming and expensive method when a large number of variables are need to be taken and it does not often bring up the effect of interaction of various parameters as compared to factorial design [113]. A statistical experimental design using response surface methodology (RSM) was employed to eliminate the limitations of single factor optimization process. This study involves the optimization of the fermentation medium for maximum alkaline protease production from *Bacillus subtilis* BSP as a result of the interaction between four variables (pH, time of incubation, substrate (casein) concentrations, and inoculum size) by employing Box Behnken design (design expert 9.0.3.1 Stat Ease, Inc., Statistical made Easy, Minneapolis, MN) to identify the optimal level of four important variables, viz., casein concentration (A), incubation time (B), pH (C) and inoculum density (D) for protease production. The effect of each variable on enzyme production was studied at three different levels (1, 0 and 1) with minimum, central, and maximum values, and twenty nine (29) experimental setups were obtained (Table 3.1). The design experiments were carried out in Erlenmeyer flasks containing Basal salt broth medium and inoculated with *Bacillus subtilis* BSP and incubated at 180 rpm at 60°C in shaking conditions. After incubation, the amount of protease was calculated by taking samples. All experiments were carried out in triplicate and average values were reported. A second order polynomial equation was used for the analysis of protease production, and the data obtained were fitted to the equation by multiple regression.
Table 3.1 Experimental designs used in RSM studies for four independent variables with six center points showing observed and predicted values of protease production

<table>
<thead>
<tr>
<th>Run order</th>
<th>Factor A (Incubation Time hr)</th>
<th>Factor B (pH)</th>
<th>Factor C (substrate conc. %)</th>
<th>Factor D (Inoculum Density %)</th>
<th>Observed response (U/mL)</th>
<th>Predicted response (U/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>7</td>
<td>2</td>
<td>1.5</td>
<td>163.00</td>
<td>165.87</td>
</tr>
<tr>
<td>2</td>
<td>72</td>
<td>8</td>
<td>3</td>
<td>1</td>
<td>228.63</td>
<td>227.60</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>8</td>
<td>3</td>
<td>1.5</td>
<td>232.38</td>
<td>235.23</td>
</tr>
<tr>
<td>4</td>
<td>72</td>
<td>8</td>
<td>2</td>
<td>1.5</td>
<td>259.88</td>
<td>256.47</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>8</td>
<td>2</td>
<td>0.5</td>
<td>228.19</td>
<td>230.70</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>8</td>
<td>3</td>
<td>1</td>
<td>224.56</td>
<td>225.27</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>163.88</td>
<td>161.69</td>
</tr>
<tr>
<td>8</td>
<td>72</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>262.00</td>
<td>262.56</td>
</tr>
<tr>
<td>9</td>
<td>24</td>
<td>8</td>
<td>2</td>
<td>1.5</td>
<td>195.44</td>
<td>193.97</td>
</tr>
<tr>
<td>10</td>
<td>48</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>209.06</td>
<td>210.97</td>
</tr>
<tr>
<td>11</td>
<td>24</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>208.76</td>
<td>211.06</td>
</tr>
<tr>
<td>12</td>
<td>48</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>279.38</td>
<td>280.87</td>
</tr>
<tr>
<td>13</td>
<td>48</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>278.00</td>
<td>280.87</td>
</tr>
<tr>
<td>14</td>
<td>72</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>199.25</td>
<td>201.05</td>
</tr>
<tr>
<td>15</td>
<td>72</td>
<td>8</td>
<td>2</td>
<td>0.5</td>
<td>221.44</td>
<td>222.03</td>
</tr>
<tr>
<td>16</td>
<td>24</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>177.44</td>
<td>175.57</td>
</tr>
<tr>
<td>17</td>
<td>72</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>188.56</td>
<td>190.05</td>
</tr>
<tr>
<td>18</td>
<td>48</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>169.25</td>
<td>169.40</td>
</tr>
<tr>
<td>19</td>
<td>48</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>172.50</td>
<td>171.46</td>
</tr>
<tr>
<td>20</td>
<td>48</td>
<td>8</td>
<td>1</td>
<td>1.5</td>
<td>222.69</td>
<td>221.02</td>
</tr>
<tr>
<td>21</td>
<td>48</td>
<td>8</td>
<td>1</td>
<td>0.5</td>
<td>250.00</td>
<td>246.76</td>
</tr>
<tr>
<td>22</td>
<td>48</td>
<td>9</td>
<td>2</td>
<td>0.5</td>
<td>181.06</td>
<td>179.46</td>
</tr>
<tr>
<td>23</td>
<td>48</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>295.00</td>
<td>280.87</td>
</tr>
<tr>
<td>24</td>
<td>48</td>
<td>8</td>
<td>3</td>
<td>0.5</td>
<td>210.50</td>
<td>211.79</td>
</tr>
<tr>
<td>25</td>
<td>48</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>190.94</td>
<td>188.15</td>
</tr>
<tr>
<td>26</td>
<td>48</td>
<td>9</td>
<td>2</td>
<td>1.5</td>
<td>190.50</td>
<td>191.31</td>
</tr>
<tr>
<td>27</td>
<td>48</td>
<td>7</td>
<td>2</td>
<td>0.5</td>
<td>179.56</td>
<td>180.02</td>
</tr>
<tr>
<td>28</td>
<td>48</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>280.00</td>
<td>280.87</td>
</tr>
<tr>
<td>29</td>
<td>48</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>272.00</td>
<td>280.87</td>
</tr>
</tbody>
</table>

This resulted in an empirical model. The model equation for analysis was:

\[ Y = \beta_0 + \sum \beta_n X_n + \sum \beta_{nm} X_n^2 + \sum \beta_{nm} X_n X_m \]  

(1)

Where, \( y \) is the predicted response, \( \beta_0 \) offset term, \( \beta_n \) linear coefficient, \( \beta_{nm} \) squared coefficient, \( \beta_{nm} \) interaction coefficient, \( X_n \) nth independent variable, \( X_n^2 \) squared effect, and \( X_n X_m \) interaction effects. For four variable systems, the model equation was as follows:
\[ Y = \beta_o + \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{11} A^2 + \beta_{22} B^2 + \beta_{33} C^2 + \beta_{44} D^2 \]
\[ + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD \]  

**3.2.5 Scale-Up Studies in Fermenter**

Cultivation of *Bacillus subtilis* BSP for protease production was also carried out in a 20 L fermenter (B. BraunBiostat®C, Illinois, USA) with a working volume of 10 L. The Basal medium containing 1% casein was sterilized in situ at 121°C for 20 min and was inoculated with 1% of the seed inoculum (OD660nm ≈0.600). The bioreactor was equipped with monitors to measure and control the variables such as foam, temperature, pH, stirring rate and dissolved oxygen. Silicone oil (Sigma) was used to control foaming during fermentation. Fermentation was maintained at 60°C and pH 8 at 150 rpm. The present study investigated the influence of incubation time on enzyme. The fermentation parameters, such as pH, enzyme activity and airflow rate were continuously monitored. Dissolved oxygen was maintained at 0.8 vvm.

**3.2.6 Purification of Extracellular Protease Produced by Thermophilic *Bacillus subtilis* BSP strain**

Cell-free supernatant was collected by centrifugation. Ammonium sulfate was slowly added to the supernatant to 70% saturation. The precipitated protein was obtained by centrifugation for 20 min at 7500 rpm at 4 °C. The obtained pellet was re-suspended in a 10 mL of ice-cold 0.1 mM phosphate buffer pH, 7.5 and then subjected to a process of dialysis against the same buffer to get rid of the excess of ammonium sulfate. The dialyzed enzyme was applied to DEAE-cellulose column (Sigma-Aldrich) previously equilibrated with phosphate buffer 0.1 M (pH 7.5). The adsorbed protein was eluted with 0.5 M NaCl prepared in 0.1 M phosphate buffer (pH 8) at a flow rate of 0.8 mL min⁻¹. 2 mL fractions were collected. The protein content was measured according to the method as described in chapter 2 section 2.6. The peaks with the highest protease activity pooled. Active fractions were loaded into Sephadex G-200 column (Sigma-Aldrich) previously
equilibrated with 0.1M phosphate buffer (pH 7.5) and developed at a flow rate of 0.5 mL min\(^{-1}\). 2 mL fractions were eluted with the same buffer. The protease activity and protein concentration were measured and the specific activity was calculated at each step.

### 3.2.7 SDS-Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and zymography for activity staining was carried out using 12% cross linked polyacrylamide gel as described in chapter 2 section 2.8. Crude, partially purified and purified fractions were monitored on the SDS-PAGE. The status of purity and molecular weight was determined with reference to molecular weight marker (Broad Range protein ladder, Thermo Scientific).

### 3.2.8 Stability and Activity Profiles of Protease

**Determination of Optimum pH and Temperature**

The activity of purified proteases were measured at different pH values (4.0-11.0) by using 1% (w/v) solution of casein as a substrate dissolved in different buffers (0.05 M): Sodium acetate buffer (4-5), potassium phosphate buffer (6-9) and glycine-NaOH (pH 10-11). Reaction mixtures were incubated at 50°C for 20 min and then the relative activities of the enzyme were measured. Optimum temperature for protease activity was determined by incubating the enzyme in casein substrate solution at different temperatures (40, 50, 60, 70, 80 and 90°C) and the level of proteolysis was assayed as described above.

**Determination of Thermal Stability**

The thermostability of protease was examined by incubating the enzyme at 50-90°C. The enzyme samples were taken after 1 hour of incubation. The reaction proceeded at 50°C and assayed as described in chapter 2 section 2.5.
Chapter 3

Production & Characterization of Protease

**Stability of Protease in Presence of Organic Solvents and Surfactants**

The effect of some organic solvents (ethanol, methanol, hexane, benzene, dimethyl sulfoxide (DMSO)) and various surfactants (sodium dodecyl sulfate (SDS), Triton X-100, Tween-20, Tween-80, H₂O₂) on enzyme activity were investigated. 10% (v/v) of the above mentioned organic solvents and surfactants in 1 mL protease enzyme solution were incubated at 50°C. Incubation was done for 10 min for organic solvent and 30 min for surfactants. All incubations occurred at pH 8. After the incubation period the substrate (1 mL of 1% casein) was added and the reaction was initiated. Then, the activity of the enzyme was measured under the standard assay conditions and the residual activity was obtained.

**Effect of Metal Ions and Inhibitors on Enzyme Activity**

The effect of the metal ions (ions (Ca²⁺, Mg²⁺, Na⁺¹, Co²⁺, Sr²⁺, Fe²⁺, Mn²⁺ and K⁺¹) and inhibitors (Dithiothritol (DTT), Phenyl Methane Salfonayl Fluoride (PMSF), Thiourea, EDTA (Ethylene Diamine Tetra Acetate) and β-mercaptoethanol) on protease activity was studied. The protease enzyme solution and metal ion were incubated for 30 minutes at 50°C. Inhibitors and enzyme were incubated at 37°C for 30 min. All cation solutions and inhibitors were used at a final concentration of 5mM. All incubations occurred at pH 8. After the incubation, the reaction was initiated by adding 1 mL of 1% casein substrate in the reaction mixture. Then, protease activity was determined under the standard assay conditions. While preparing the control, the metal ion or inhibitor was not added into the enzyme solution and its activity was regarded as 100%. The tested cations included the following corresponding salts: NaCl, KCl, CaCl₂, FeSO₄, MgCl₂, CoCl₂, SrCl₂ and MnSO₄.

**Determination of Kinetic Parameters**

The difference between the energy of the reactants and that of the highest energy transition state is known as the activation energy of the reactants. Activation energy of the protease was determined by using the data for optimum temperature. For determination of Vₘₐₓ and Kₘ, protease was assayed in 50mM phosphate buffer of pH 8,
with variable amounts (0.2%-4%) of soluble casein solution. The reciprocal of reaction velocity was plotted against the reciprocal of the corresponding substrate concentration, giving the Lineweaver-Burk plot for the purified protease with the substrate casein. The Michaelis-Menten constant (Km) and maximum reaction velocity (Vmax) of the purified enzyme were determined from the Lineweaver-Burk plot [115].

3.3 Results

3.3.1 Screening for Proteolytic Activity

For the screening of casein utilizing bacterial strains for thermophilic alkaline protease production the *Bacillus subtilis* BSP was checked on casein (1%) agar plates for zone of clearance. The figure 3-1 showed clear zones of hydrolysis on casein agar plate. The strain was selected for further studies of enzyme production, characterization and scale up studies. The isolate was a gram positive, motile, rod shaped bacterium and aerobic.

For the growth and extracellular enzyme production the *Bacillus subtilis* BSP was grown on basal salt medium containing 1% casein at 60°C and pH 7.0. Growth of the organism and enzyme production by the culture was determined periodically by taking out samples after 8 hours interval up to 72 hours. Determination of cell biomass indicated that lag phase was very short and the log phase was up to 24 h followed by the stationary phase till 48th h, after that decline phase was observed. Protease production was observed during the stationary phase. Enzyme production was maximum in 48 h (15.2 g/L, 178 U/mL) after which the production started to decline, precisely in the decline growth phase (Figure 3-2).
Figure 3-1 Thermophilic *Bacillus subtilis* BSP strain indicating a zone of clearance on casein agar plate

Figure 3-2 Growth curve of *Bacillus subtilis* BSP strain in Basal salt medium
3.3.2 Optimization of Nutritional and Physical Parameters

Effect of pH and Temperature on Enzyme Production

The effect of initial pH (3.0-11.0) of the growth medium on the enzyme production during the course of a four days fermentation period is shown in Figure 3-3. It was found that *Bacillus subtilis* BSP was capable of producing protease over a wide range of initial culture pH (5-9) but the enzyme production was highest (184 U/mL) at pH 8.0 after 48 h. The enzyme production was drastically reduced at pH 10.0 (Figure 3-3).

*Bacillus subtilis* BSP exhibited protease production at all temperatures studied (40°C to 90°C). Optimum temperature for protease production (178.64 U/mL) was found to be 60°C (Figure 3-4). Above the optimum temperature the increase in temperature was accompanied by a decrease in the protease activity.

Effect of Incubation Time

The monitoring of enzyme production at different time interval indicated the optimum production of enzyme 179.37 U/mL after 48 h of incubation period. The enzyme activity decreased gradually after 60-72 h (Figure 3-5).

Effect of Substrate (Casein) Concentration on Protease Production

Optimum production of the enzyme (165 U/mL) was observed by 1 % (w/v) of casein. Lower concentration (0.5%) of casein dramatically decreases the activity and high concentration of substrate 3% inhibited the enzyme production (Figure 3-6). However, the present investigation indicated that the enzyme production has relation with different substrate concentrations which suggest the applicability of this protease in substrate rich systems.
Chapter 3

Production & Characterization of Protease

Figure 3-3 Effect of initial pH on protease production for 4 days fermentation period at 60°C

Figure 3-4 Effect of temperature on protease production after 48 h fermentation period at pH 8

53
Figure 3-5 Effect of incubation time on protease production at 60°C and pH 8

Figure 3-6 Effect of casein concentration on protease production at 60°C and pH 8 after 48 h of incubation
3.3.3 Scale-Up Fermentation Studies

The batch fermentation experiment in 20 L fermenter was studied for protease production with 1% substrate (casein) concentration. The enzyme production and pH was monitored after every 8h. It was observed that on the first day there was a drop in the pH from 8.0 to 7.6 and thereafter the pH of the fermentation medium was almost the same with some fluctuations till the end of fermentation. Maximum enzyme production of 248 U/mL was obtained on the 48th h of incubation (Figure 3-7). These results were beneficial for further studies, as it cut down the necessity to adjust the pH till the end of fermentation. A significant increase in protease production was observed in a bioreactor as compared to enzyme production at shake flask level, when all the parameters optimized at flask level were used in the bioreactor studies. This could be due to continuous supply of oxygen to bacterial culture and also uniform distribution of medium components by the agitation.

![Graph](image)

**Figure 3-7 Growth of Bacillus subtilis BSP and protease production in bench scale fermenter at 60°C and pH 8**
3.3.4 Statistical Optimization of Factors Affecting Protease Production by RSM

In order to search for the optimum formulation of the medium, Box-Behnken design with four factors, which indicated that 29 experiments were required for this procedure, was employed to fit a second order polynomial model. The experimental design and corresponding alkaline protease yields are shown in Table 3.2.

**Model Fitting and Analysis of Variance (ANOVA)**

The combined interactions of factors, viz., time of incubation, pH, substrate concentration and inoculum density on protease production by *Bacillus subtilis* BSP were examined by RSM employing a Box–Behnken design. Table 3.2 presents the predicted responses of the design based on a polynomial equation. ANOVA yielded the following regression equation in terms of protease levels produced (Y) as a function of incubation time (A), pH (B), substrate concentration (C), and inoculum density (D):

\[
Y=280.88+13.46A+6.22B-5.19C-0.57D-0.72AB-
12.29AC+17.80AD- 14.56BC+ 6.50BD+12.30CD-
26.08A^2-72.71B^2-23.17C^2-29.00D^2
\]

(3)

The coefficient of determination (R²) is defined as the ratio of the explained variation to the total variation, and is a measure of the degree of fit [116]. It is suggested that a good model fit should yield an R² of at least 0.8. This means that the response model evaluated in this study can explain the reaction very well, with an R² of 0.9913 and an Adj-R² of 0.9827 at a confidence level of 95%. The "Pred R-Squared" of 0.9779 is in reasonable agreement with the "Adj R-Squared" of 0.9827; i.e. the difference is less than 0.2.
Table 3.2 ANOVA for response surface quadratic model

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F-Value</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>43212.23</td>
<td>14</td>
<td>3086.59</td>
<td>114.27</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>A-Time</td>
<td>2173.22</td>
<td>1</td>
<td>2173.22</td>
<td>80.46</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>B-pH</td>
<td>464.07</td>
<td>1</td>
<td>464.07</td>
<td>17.18</td>
<td>0.0010</td>
</tr>
<tr>
<td>C-Substrate</td>
<td>323.03</td>
<td>1</td>
<td>323.03</td>
<td>11.96</td>
<td>0.0038</td>
</tr>
<tr>
<td>D-Inoculum density</td>
<td>3.94</td>
<td>1</td>
<td>3.94</td>
<td>0.15</td>
<td>0.7082</td>
</tr>
<tr>
<td>AB</td>
<td>2.07</td>
<td>1</td>
<td>2.07</td>
<td>0.077</td>
<td>0.7861</td>
</tr>
<tr>
<td>AC</td>
<td>604.62</td>
<td>1</td>
<td>604.62</td>
<td>22.38</td>
<td>0.0003</td>
</tr>
<tr>
<td>AD</td>
<td>1266.90</td>
<td>1</td>
<td>1266.90</td>
<td>46.90</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>BC</td>
<td>848.29</td>
<td>1</td>
<td>848.29</td>
<td>31.41</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>BD</td>
<td>169.00</td>
<td>1</td>
<td>169.00</td>
<td>6.26</td>
<td>0.0254</td>
</tr>
<tr>
<td>CD</td>
<td>604.84</td>
<td>1</td>
<td>604.84</td>
<td>22.39</td>
<td>0.0003</td>
</tr>
<tr>
<td>A²</td>
<td>4411.81</td>
<td>1</td>
<td>4411.81</td>
<td>163.34</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>B²</td>
<td>34288.62</td>
<td>1</td>
<td>34288.62</td>
<td>1269.47</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>C²</td>
<td>3483.35</td>
<td>1</td>
<td>3483.35</td>
<td>128.96</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>D²</td>
<td>5456.22</td>
<td>1</td>
<td>5456.22</td>
<td>202.01</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>378.14</td>
<td>14</td>
<td>27.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of Fit</td>
<td>88.58</td>
<td>10</td>
<td>8.86</td>
<td>0.12</td>
<td>0.9409</td>
</tr>
<tr>
<td>Pure Error</td>
<td>289.56</td>
<td>4</td>
<td>72.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cor Total</td>
<td>43590.37</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard deviation 5.20  R-Squared 0.9913
Mean 218.41  Adj R- Squared 0.9827
Coefficient of variation (C.V. %) 2.38  Pred R-Squared 0.9779
PRESS 962.67  Adeq Precision 31.886
"Adeq Precision" measures the signal to noise ratio. A ratio greater than 4 is desirable. Our ratio of 31.886 indicates an adequate signal. This model can be used to navigate the design space. The Model F-value of 114.27 implies the model is significant. There is only a 0.01% chance that an F-value this large could occur due to noise. Values of "Prob> F" less than 0.0500 indicate model terms are significant. In this case model term linear incubation time ($P < 0.0001$), interaction of incubation time and inoculum size ($P < 0.0001$), interaction of pH and substrate concentration ($P < 0.0001$), incubation time $A_2$ ($P < 0.0001$), pH $B_2$ ($P < 0.0001$), Substrate concentration $C_2$ ($P < 0.0001$) and inoculum size $B_2$ ($P < 0.0001$) are significant model terms. Values greater than 0.1000 indicate the model terms are not significant.

Table 3.2 showed the ANOVA results for the acquired model. ANOVA is an analytical technique that is used to identify the importance of the model and its parameters, using Fisher’s F-test and Student’s t-test [117]. In general, larger F-values and smaller p-values indicate more significant coefficient terms.

**Response Analysis**

The three dimensional response surface plots and two dimensional contour plots were used to understand the interaction effects of factors under study, and the optimum concentration of each factor required for maximum protease yield. Response surface curves for variation in alkaline protease production were constructed, and are depicted in Figure 3-8. In each set, two parameters were varied within the tested experimental range, and the other two variables were kept at their central level.

Figure 3-8a depicts protease production with respect to pH versus incubation time in the fermentation broth. Protease production increased with increasing pH up to 8.0 at the central level (48 h) of time of incubation. The protease yield decreased when the pH deviated from 8.0 and incubation time exceeded beyond 51 h. Figure 3-8 B depicts 3D response surface plot for the interaction between casein concentration and incubation time. The effect contributed by between casein concentration and incubation time towards protease production was not significant when pH and inoculum size were kept
Figure 3-8a Response-surface curve of alkaline protease production *Bacillus subtilis* BSP showing mutual interactions between (A) time of incubation and pH (B) casein concentration and time (C) inoculum density and time.
Figure 3-8b Response-surface curves of alkaline protease production *Bacillus subtilis* BSP showing mutual interactions between (D) casein concentration and pH (E) pH and inoculum density (F) casein concentration and inoculum density.
constant at 7 and 0.5 % respectively since an increase in the concentration of these two affect the protease production. As shown in Figure 3-8 C, the maximum production of alkaline protease was found to occur when the inoculum density range from 0.9% to 1.0% and the incubation time ranged from 48 to 51 h. Figure 3-8 B shows that protease production increased with increased level of substrate concentration up to 2% and pH 8.0. Figure 3-8 E reveals the interaction between pH and inoculum density on protease production, which increased with an increasing inoculum density up to 1% and then decreased at higher value and at pH (8.0). A further increase in pH gave a lower enzyme yield. Figure 3.8 D illustrates the interaction effect of inoculum density and casein concentration (1-3%) on alkaline protease production. Maximum enzyme yield was observed 1.8% casein and inoculum density of 1 %. Any deviation in casein concentration or inoculum density from their central optimal levels decreased the enzyme yield.

The optimized values of variables for maximum protease activity were determined by solving the quadratic mode equation. The corresponding optimum values of each variable were determined to be pH 8.0, inoculum density 1%, incubation time of 48h and 2% casein concentration at which conditions a maximum protease activity response of 295 U/mL was predicted. The results were validated by conducting a shake-flask culture under optimized conditions. This result shows that the regression model developed in this study resulted in good agreement between the actual and predicted responses.

### 3.3.5 Purification of Protease

The bacterial culture supernatant containing the active protease enzyme having a specific activity of 39 U mg⁻¹ was first purified through protein precipitation using ammonium sulfate. The partially purified enzyme was further purified by DEAE-cellulose ion-exchange chromatography and then applied into a gel filtration column Sephadex G-200 column. The purified fraction was 9.8 folds with specific activity of 387 U mg⁻¹ of protein (Table 3.3).
Table 3.3 Purification of alkaline protease from *Bacillus subtilis* BSP

<table>
<thead>
<tr>
<th>Purification Steps</th>
<th>Volume (ml)</th>
<th>Total Activity (U)</th>
<th>Total Protein (mg)</th>
<th>Specific Activity (U/ml)</th>
<th>Fold Purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude Enzyme (supernatant)</td>
<td>1000</td>
<td>100800</td>
<td>2560</td>
<td>39.3</td>
<td>1</td>
</tr>
<tr>
<td>70% Ammonium Sulfate precipitation</td>
<td>200</td>
<td>12120</td>
<td>208</td>
<td>58.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Dialysis</td>
<td>50</td>
<td>7667</td>
<td>107</td>
<td>71.6</td>
<td>1.8</td>
</tr>
<tr>
<td>DEAE Cellulose Chromatography</td>
<td>10</td>
<td>3997</td>
<td>36</td>
<td>111</td>
<td>2.8</td>
</tr>
<tr>
<td>Gel filtration chromatography on Sephadex G-200</td>
<td>15</td>
<td>602</td>
<td>3.34</td>
<td>180.2</td>
<td>4.5</td>
</tr>
</tbody>
</table>

![SDS-PAGE](image)

**Figure 3-9** SDS–PAGE of the thermostable proteases from *Bacillus subtilis* BSP (A) Activity staining of purified protease B) Lane 1, standard markers; lane 2, culture supernatant; lane 3, Freeze dried enzyme; lane 3, dialyzed protease after ammonium sulfate ppt; lane 6, DEAE cellulose column purified protease and lane 6, Sephadex G-200 purified thermostable protease.
3.3.6 Electrophoretic Analysis

The extracellular protease activity was subject to electrophoresis on SDS/polyacrylamide gels with and without 0.1% gelatin. A clear zone of proteolytic activity was observed, the results being the same on native gel. The purified enzyme showed single homogeneous band of 36 kDa on in SDS–PAGE (Figure 3-9). The molecular mass of this protein was compared with molecular mass standard, and was estimated as 36kDa.

3.3.7 Characterization Profile of Thermostable Protease

Effect of pH on Protease Activity, Optimum pH

The activity of the protease was determined over a range of pH values in 50 mM acetate buffer (3-5), potassium phosphate buffer (6-9) and glycine-NaOH (pH 10-11). The maximum % relative activity was obtained at pH 8.0 in 50 mM potassium phosphate buffers. Thermostable protease was active over a very broad pH range 6-9 and about 83% and 51% relative activity proteolytic activity was obtained at the pH extremes respectively (Figure 3-10).

Effect of Temperature on Protease Activity and Stability

The protease activities were assayed at different temperatures ranging from 40ºC-90ºC at pH 8 (Figure 3-11). Enzyme activity increased with temperature within the range of 40ºC to 50ºC. A reduction in enzyme activity was observed at values above 60ºC. The optimum temperature of this protease was 50ºC. The thermostability of the protease was examined by incubating the enzyme at different temperatures (40-90ºC) without substrate for 60 min and the relative activity was measured at 50ºC. Maximal enzyme activity observed was set as 100% relative activity. Thermostability profile indicated that the enzyme was stable at 40ºC and 50ºC for 30 min while the enzyme retained 82% of the original activity after 30 min heat treatment at 60ºC. At 70ºC, 73% of the original activity was retained after heat treatment. At 80ºC 84% of the original activities was lost. The
Figure 3-10 Effects of pH on alkaline protease activity

Maximal enzyme activity observed was set as 100% relative activity

Figure 3-11 Effects of temperature on activity and stability of protease
decrease in enzyme activity at higher temperatures at 80 °C might be due to the destruction of an enzyme at certain temperatures.

**Effect of Organic Solvents, Surfactants, Metal Ions and Inhibitors on Protease Activity**

The organic solvents namely ethanol, methanol, hexane, benzene and dimethyl sulfoxide (DMSO), various surfactants (sodium dodecyl sulfate (SDS), Triton X-100, Tween-20, Tween-80, H₂O₂), inhibitors (dithiothritol (DTT), Phenyl Methane Salfonyl Fluoride (PMSF), Thiourea, EDTA (Ethylene Diamine Tetra Acetate) and β-mercaptoethanol) and different metal ions (Ca²⁺, Mg²⁺, Na⁺, Co²⁺, Sr²⁺, Fe²⁺, Mn²⁺ and K⁺) were checked to determine the stability of alkaline protease against these reagents. Protease activity was activated by divalent cations in decreasing order Ca²⁺ > Mn²⁺ > Mg²⁺ > Fe²⁺ for thermostable protease enzyme (Figure 3-12). The protease enzyme was metalloprotease and requires metal Ca²⁺ and Mn²⁺ ions to enhance the enzyme activity. These results suggest that these metal ions apparently protected the enzyme against thermal denaturation and played a vital role in maintaining the active conformation of the enzyme at higher temperatures.

The enzyme activity was studied against different organic solvents and incubated at 50°C, pH 8 for 20 min. The relative activity of enzyme was more than 90% stable against 10% benzene, ethanol and methanol. On the other hand, Hexane caused much effect on the activity among the other organic solvents and the protease enzyme dropped to nearly 75% of its activity in the presence of hexane. Moreover in the presence of 10% DMSO the enzyme showed 83% relative activity. Enzyme was found highly stable in presence of organic solvents (Table 3.4). The performance of enzyme in organic solvents can be attributed to many factors. The protein-solvent interactions favor exclusion of the solvent molecules from the enzyme hydration layer resulting in compaction of the enzyme because of that unfavorable surface energy rise which stabilizes the enzyme [19]. Protease was stable in 10% (v/v) Triton X-100 and showed nearly 97% of its activity in the presence of Triton X-100. On the other hand, SDS and H₂O₂ enhanced the activity among the other surfactants and the protease enzyme showed 112% and 105% of its activity respectively. In the Table 3.5 the results indicated that thermophilic protease
Chapter 3  
Production & Characterization of Protease

Figure 3-12 Effect of metal ions on protease activity

Table 3.4 Effect of organic solvents on protease activity (at 50 °C, pH 8, for 20 min.)

<table>
<thead>
<tr>
<th>Organic Solvent</th>
<th>Concentrations (%)</th>
<th>Relative Activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>-</td>
<td>100</td>
</tr>
<tr>
<td>DMSO</td>
<td>10</td>
<td>83.2</td>
</tr>
<tr>
<td>Benzene</td>
<td>10</td>
<td>95.8</td>
</tr>
<tr>
<td>Ethanol</td>
<td>10</td>
<td>95.77</td>
</tr>
<tr>
<td>Methanol</td>
<td>10</td>
<td>92.18</td>
</tr>
<tr>
<td>Hexane</td>
<td>10</td>
<td>75.6</td>
</tr>
</tbody>
</table>
Table 3.5 Effect of surfactants on protease activity (at 50 °C, pH 8, for 20 min.)

<table>
<thead>
<tr>
<th>Surfactants</th>
<th>Concentration (%)</th>
<th>Relative Activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>-</td>
<td>100</td>
</tr>
<tr>
<td>SDS</td>
<td>10</td>
<td>112</td>
</tr>
<tr>
<td>Triton X-100</td>
<td>10</td>
<td>97.3</td>
</tr>
<tr>
<td>H₂O₂</td>
<td>10</td>
<td>105.2</td>
</tr>
<tr>
<td>Tween-80</td>
<td>10</td>
<td>87.6</td>
</tr>
<tr>
<td>Tween-20</td>
<td>10</td>
<td>95.4</td>
</tr>
</tbody>
</table>

Figure 3-13 Effect of inhibitors on protease activity
enzyme was stable in the presence of 10% (v/v) of these surfactants. When the effects of the surfactants on protease activity were compared in the literature, similar and different results were observed. The protease showed extreme stability towards non-ionic (10% Tween 20 and Triton X-100) and anionic (SDS) surfactant. The anionic surfactant, sodium dodecyl sulfate (SDS) enhanced the activity. The enzyme was stable in the presence of H₂O₂.

All the inhibitors tested at 5mM inhibited the protease (Figure 3-13). The enzyme lost 90.7% of its original activity in the presence of EDTA which infers that EDTA chelates metal ions at the active site of the purified protease which possibly serve as cofactors for the protease confirming the enzyme is a metalloprotease. β-ME inhibited protease activity by 45%. DTT and PMSF both induced moderate inhibition while IAA had no effect.

**Determination of Kinetic Parameters**

Michaelis-Menten constant (Kₘ) and maximum velocity (Vₘₐₓ) was calculated with the help of Line-weaver-Burk plot of purified protease produced from indigenous *Bacillus subtilis* BSP strain. The value of Kₘ was 0.01mg/mL/min while Vₘₐₓ showed a value of 79μmol/min (Figure 3-14). It was observed that the reaction was dependent on the amount of casein present in the reaction mixture. Activation energy of protease for the formation of ES*-complex was determined by Arrhenius plot. The activation energy of protease at optimum temperature was 63.7 KJmol⁻¹ and deactivation was 66.8 KJmol⁻¹ (Figure 3.14). The Arrhenius plot of protease activity exhibited activation energy of 63.7 kJmol⁻¹ and in the temperature range 40-70°C for thermostable protease.

The reaction was dependent on the amount of casein present in the reaction mixture. A Lineweaver Burk Plot of the data revealed a Kₘ and Vₘₐₓ of 0.01 mg casein mL⁻¹ min⁻¹ and 79 μ mol min⁻¹ respectively for BSP protease.
Figure 3-14 Lineweaver-Burk plot to estimate Km and Vmax of protease produced from wild strain of thermophilic *Bacillus subtilis* BSP

Vmax = 79 micro mol/min
Km = 0.01 mg/mol/min

Figure 3-15 Arrhenius plot to calculate activation energy of BSP protease produced from thermophilic *Bacillus subtilis* BSP

Ea = 63.7 KJ/mol
Ea (n) = 66.8 KJ/mol
3.4 Discussion

Production of proteases of commercial significances by employing a suitable organism and economical growth medium has been a worth praising achievement in the field of fermentation biotechnology. Although proteases are wide spread in nature, microbes have been identified as preferred source of these enzymes because of rapid growth, limited space required for cultivation and ease of product separation from fermented broth [11]. Protease production is an inherent capacity of all microorganisms; and large numbers of bacterial species are known to produce proteases [9]. Among various bacteria, the *Bacillus* species are most significant and specific producers of proteases [9, 118]. Considering the richness of microbial diversity, there is always a chance of searching new organisms producing enzyme with better properties and suitability for commercial exploitation.

Understanding the characteristics of a candidate enzyme for industrial applications is a serious concern for industrial scientists, as it would help to establish the usefulness, or otherwise, of the enzyme in such processes. We report herein the production and characterization of protease from the thermophilic *Bacillus subtilis* strain BSP. One of the main concerns of this study was to identify thermophilic *Bacillus* sp. having a vital tendency to secrete extra-cellular thermostable proteolytic enzyme. *Bacillus subtilis* BSP strain, used in this study, is a thermophilic, gram positive, spore forming and motile exhibiting an optimum growth at 60ºC.

The strain of *Bacillus* presented a typical kinetic growth profile, showing a short lag phase, a marked exponential and stationary phases. The decline phase was observed after 56 h that produced a significant decrease in the absorbance values. Our experimental results showed that the maximum amount of enzyme was produced by the bacterium in its stationary growth phase and a reduction in enzyme expression was noted beyond this period, precisely in the decline growth phase. Similar studies on *Bacillus sp.* [119-120] showed maximum enzyme production in stationary phase.

Protease production is highly depending on the medium pH as it strongly influences the transport of components across the cell membrane and enzymatic
processes which in turn affects the cell growth and product formation [121]. Summarizing, the initial pH of the cultivation media is a parameter impacting on both maximum enzyme production levels and on the properties of the crude extract. The optimal pH for the microorganism growth and optimal pH for the enzyme production was found at pH 8.0. This result is accordance with several earlier reports showing pH optima for *Bacillus cereus* sp. protease expression being alkaline [122-123].

The time course data revealed that maximum level of thermostableprotease was produced after 48 h of cultivation period. Seifzadef et al. [123] and Sareket et al. [124] reported the same incubation period for maximum production of protease from *Bacillus* sp. GUS1 and *Bacillus* sp P003 respectively. Different fermentation time has also been reported: 36 h in *Bacillus* sp. to produce subtilisin [125] under submerged fermentation and 72 hr [21] in *Bacillus licheniformis* (MTC NO. 7053) for protease production.

Incubation temperature also effect on the culturing of the organism. The effect of temperature on protease production at varying temperatures depicted that most suitable temperature for the production of alkaline protease was 60 °C. These findings were similar to that of Serkar et al. [124]. Zeikus [126] reported that majority of the thermophilic *Bacilli* are found to grow at pH and temperature range of 5.8-8.0 and 50-65°C, respectively.

Nascimento and Martins [14] reported an optimum temperature of 60°C for *Bacillus* sp.SMIA-2 while Prabhavathy et al. [127] and Rani et al. [128] reported the production of thermostable proteases at 45°C and 55°C respectively . Cultivation temperature affects protein synthesis by influencing rate of biochemical reactions within the cell and consequently inducing or repressing enzyme production.

Study on the substrate concentration indicated that increasing concentration of casein enhanced the enzyme production. This result correlates with the findings of Lakshmi [129]. In view of other investigators, 0.6% (w/v) casein was the best substrate concentration for thermostable protease production by *Bacillus licheniformis* U1 [109]. Ferro et al. [46] observed that casein was the best source for production of alkaline protease.
Role of oxygen in metabolism and the growth of aerobic microbes is well known. The enzyme production in aerobic microorganisms is also depended on the availability of oxygen for their growth in the fermentation medium. Therefore, effect of different incubation period on enzyme yield and growth of microorganism was investigated in lab scale fermenter ((B. Braun Biostat®C, Illinois, USA). A significant increase in protease production (248 U/mL) was observed in a bioreactor as compared to enzyme production at shake flask level, when all the parameters optimized at flask level were used in the bioreactor studies. This could be due to continuous supply of oxygen to bacterial culture and also uniform distribution of medium components by the agitation.

The bacterium used in this study was shown to constitutively secrete extracellular proteolytic enzyme. The optimization of culture medium was carried out by a combination of non-statistical and statistical based experimental designs. The selection of medium compositions was carried out through the one-factor-at-a-time experiments, and 184 U/mL of the protease activity was observed. The Box Behnken experimental design indicated that incubation time and pH were the most significant factors.

The three-dimensional response surface presentations were then plotted to study the interaction among various physicochemical factors used and to find out the optimum concentration of each factor for maximum thermostable protease production from Bacillus subtilis BSP. The corresponding optimum values of each variable were determined to be pH 8.0, inoculum density 1%, incubation time of 48h and 2% casein concentration at which conditions a maximum protease activity response was predicted. The results were validated by conducting a shake-flask culture under optimized conditions. This result shows that the regression model developed in this study resulted in good agreement between the actual and predicted responses. Meruvu and Vangalapati [31] and Queiroga et al.[32] reported 1.2 and 2.0 folds increase in protease production by using Box Behnken experimental design and central composite design of response surface methodology respectively.

The purification of proteases is essential from the better understanding of the functioning of the enzyme. Precipitation is the most commonly used method for the recovery of proteins from crude biological mixtures [130]. It also performs both
purification and concentration steps. Our data showed that the 70% ammonium sulfate saturation fractions correlated with high proteolytic and specific activity. Abdulkhair and Rashad [132] found that protease from *Rumexve sicarius L* was precipitated by 40-50% saturation of ammonium sulfate. Moreover, protease from *Bacillus circulans* precipitated by 80% saturation of ammonium sulfate [133].

In the present study the molecular weights of the protease was estimated to be 36 kDa, based on the zymogram gel stained by Coomassei Brilliant Blue G-250. A varieties of molecular masses for proteases from other *Bacillus* species had been reported: 49 kDa thermostable protease from hyper thermophilic *Bacillus* strain HUTBS71 [63]; 36 kDa thermostable proteases from *Bacillus stearothermophilus* TLS33 [64]; 40kDa *Bacillus licheniformis* VSG1 [134]; 28 kDa alkaline protease from *Bacillus pumilus* MK6-5 [110]; 36.0 kDa metalo serine protease from *Brevibacillus thermoruber LII* [60] and 37-40 kDa Alkaline serine proteases from thermophilic *Bacillus* sp. GUS1[123]

### 3.4.1 Characterization of Extracellular Protease Produced by Thermophilic *Bacillus subtilis* BSP Strain

The effect of temperature on the activity of enzyme was studied. Maximum activity was shown at 50°C for protease enzyme [109]. The decrease in enzyme activity at higher temperatures at 80 °C might be due to the destruction of an enzyme at certain temperatures. Zilda et al. [60] reported the optimum temperature for protease activity was 60°C.

Sookkheo et al. [64] reported to three proteases, S, N and B from thermophilic *Bacillus stearothermophilus* TLS33, optimum pH values of 8.5, 7.5, and 7.0, respectively. The protease S was active over a very broad pH range, and about 60% of proteolytic activity was still detectable at pH 6 and 10 in the presence of 5mM CaCl₂. In contrast, proteases N and B retained relatively little activity above pH 9.0.

Protease activity was activated by divalent cations in decreasing order Ca²⁺ > Mn²⁺ > Mg²⁺ > Fe²⁺ for this thermostable protease enzyme. The protease enzyme was
metalloprotease and requires metal Ca$^{2+}$ and Mn$^{2+}$ ions to enhance the enzyme activity. Most previous studies showed that the effect of calcium and manganese was to increase protease activity, which was similar to our results. Oberoi et al. [135] reported two fold enhancements in protease activity in the presence of Mn$^{2+}$. Akel et al. [63] and Rao et al. [136] also reported that protease produced by *Bacillus* sp. was enhanced by Ca$^{2+}$ and Mn$^{2+}$. These results suggest that these metal ions apparently protected the enzyme against thermal denaturation and played a vital role in maintaining the active conformation of the enzyme at higher temperatures.

Among the organic solvents used in the study benzene and ethanol were found to be best for the stability of protease. Enzymes, in general, get denatured or give very low rates of reaction in solvent media because of the unfolding, structural disfunctioning and stripping of the essential water layer from the enzyme molecule. The presence of organic solvents alters the catalytic process of enzyme by disruption of hydrogen bonds and hydrophobic interactions as well as changes in the dynamics and conformation of the protein [137]. The performance of enzyme in organic solvents can be attributed to many factors. The protein-solvent interactions favor exclusion of the solvent molecules from the enzyme hydration layer resulting in compaction of the enzyme because of that unfavorable surface energy rise which stabilizes the enzyme. Previous studies have reported some organic solvents tolerant proteases [138] [10].

When the effects of the surfactants on protease activity were compared in the literature, similar and different results were observed. The protease showed extreme stability towards non-ionic (5 mM Tween 20 and 5 mM Triton X-100) and anionic (SDS) surfactants [139]. The anionic surfactant, sodium dodecyl sulfate (SDS) enhanced the activity [54, 140]. The enzyme was stable in the presence of H$_2$O$_2$. The activity was not affected by Tween 80 [54], SDS [123] and Triton X-100 [54].

The Arrhenius plot of protease activity exhibited activation energy of 63.7 kJmol$^{-1}$ and in the temperature range 40-70°C for thermostable protease. Similar type of break point was observed by Rao et al. [136] for serine protease from *Bacillus circulans*. The activation energy of alkaline protease for casein hydrolysis was calculated from the
slope of an Arrhenius linear plot as 44.30 kJ mol\(^{-1}\). Margesin et al. [141] also reported the high Ea values of 59.09 KJ/mol, 62.19 KJ/mol and 64.14 KJ/mol for two commercial proteases and an alkaline serine protease from psychrotrophic *Bacillus* sp. strain 2/2 respectively.

The reaction was dependent on the amount of casein present in the reaction mixture. The kinetic constants \(K_\text{m}\) (0.01 mg casein mL\(^{-1}\) min\(^{-1}\)) and \(V_{\text{max}}\) (79 μ mol min\(^{-1}\)) were comparable to the \(K_\text{m}\) (0.8 mg mL\(^{-1}\) min\(^{-1}\)) and \(V_{\text{max}}\) (85.0 μ mol min\(^{-1}\)) reported by Devi et al. [142]. A higher \(K_\text{m}\) value of 50 mg/mL has been reported for the enzyme from *Lactobacillus brevis* indicating higher affinity of the enzyme towards casein [143].

In conclusion, thermophilic *Bacillus subtilis* BSP was found to be an excellent and efficient producer of thermostable metalloprotease which can be used for wide range of applications. Further studies on enzyme kinetics and fermenter scale data will enable to translate the production of protease enzyme for commercial avenues with much ease.
4 Cloning and Expression of Protease Gene from Thermophilic *Bacillus subtilis* BSP Strain into *E.coli* and *Bacillus* Expression Systems

4.1 Introduction

The microorganisms capable of living at high temperatures, widely known as thermophiles, hold a particular fascination for microbiologists and biochemists. They survive and adapt to ecological niches defined as ‘extreme’ for humans, and they offer stable thermophilic enzymes which tolerate high temperatures and exhibit a long half-life. Thermophilic bacteria are considered to be a good model for better understanding the mechanism and molecular structure of thermostability [144].

Production of proteases by thermophilic bacteria at higher temperature has several advantages over mesophiles like easy handling, increased solubility, increased reaction rate and reduced incidence of microbial contamination from mesophilic organisms benefits [145]. Therefore, attention has been focused on thermophilic bacteria and attempts are being made to find and produce new thermostable enzymes [64]. Many thermophiles, such as *Bacillus stearothermophilus*, *Thermus aquaticus*, *Bacillus licheniformis*, *Bacillus pumilus* and *Thermoanaerobacter yonseiensis* can produce a variety of thermostable extracellular proteases, some of which are industrially important. Since thermophilic enzymes are also stable under a wide range of acidic and alkaline conditions, their production can be achieved either from thermophiles, grown under optimum fermentation conditions or by cloning their genes encoding for thermophilic enzymes. Various protease-encoding genes have been cloned and characterized from mesophilic bacteria, such as *Bacillus subtilis* [136, 146-148]. Although there have been many examples of protease genes from mesophiles that have been successfully cloned and expressed in *Escherichia coli*, little information is available about thermophilic proteases. It is necessary and interesting to investigate, whether the folding and functioning of recombinant protein is identical to normal protein. Recombinant protein
production is an enormous field. There seems no sign that the expansion of this field will abate anytime soon. *Escherichia coli* expression system continues to dominate the bacterial expression systems and remain to be the first choice for laboratory investigations and initial development in commercial activities or as a useful benchmark for comparison among various expression platforms. *E. coli* system is also the basis for efforts in protein engineering and high-throughput structural analysis [149]. Extracellular production of proteins is highly desirable as it could greatly reduce the complexity of a downstream processing and improve product quality. In this regard, the weak ability of *E. coli* in secretion of protein products to growth medium has long been considered as a major drawback of the system.

After *Escherichia coli*, gram-positive *Bacilli* strains have proved to be the most attractive hosts for heterologous protein production because *Bacillus subtilis* is considered as a GRAS organism (generally recognized as safe) in contrast to *E. coli* [150] and they have a naturally high secretion capacity, and they export proteins directly into the extracellular medium [151].

*Bacillus subtilis* BSP characterized by significant thermostable alkaline protease, was biochemically characterized (chapter 3). The activity of the protease reached 178 U/mL after cultivating for 48 h at 60°C, and the molecular mass of the expressed enzyme was about 36 kDa. The optimal temperature and pH of the protease production were 60°C and 8, respectively. However, it has a growth temperature too high for large-scale cultivation. For this reason, we realized that there was a need to develop a system that can operate at a lower temperature. Meanwhile, *B. subtilis* is an efficient system for the extracellular expression of foreign gene. Its greatest advantage is the ability to efficiently express the foreign genes with biological activity and secrete proteins directly into the culture medium. Furthermore, *Bacillus subtilis* is not a human pathogen and can be considered biologically safe.

In this study, we have therefore cloned the gene coding thermophilic protease into *Bacillus subtilis* under the hyper spank promoter which was followed by the signal peptide sequence of the thermophilic neutral protease gene. We also described the purification and biochemical characterization of this protease.


4.2 Materials and Methods

4.2.1 Strains, Plasmids and Growth Conditions

All strains used in this study are listed in Table 4.2. *Escherichia coli* strain BL21(DE3)pLysE was obtained from Novagen (Madison, WI, USA). *Bacillus subtilis* 1A751, which was deficient in two proteases from the *Bacillus* Genetic Stock Center (BGSC; OH, USA) was used as expression host. *Escherichia coli* strain JM109 (endA1, recA1, gyrA96, thi, hsdR17 (rK-, mK+), relA1, supE44, λ–, Δ (lac-proAB), [F’, traD36, proAB, lacIqZΔM15] was also used in all DNA manipulations. The plasmid used for *E. coli* and *Bacillus subtilis* expression were pET29b (Novagen, Madison, WI, USA), pSG1154 (BGSC) and pDR111 (a gift from David Rudner, Harvard Medical School Boston, MA, USA). All the constructs were transformed in *E. coli* strain JM109 by using the standard protocol described by the manufacturer and transformants were screened on LB agar ampicillin (100 μg/mL) and LB agar kanamycin (30μg/mL) plates. Some colonies were selected from plates for plasmid isolation and cloning was confirmed by restriction digestion of plasmids and then by sequencing. Recombinant plasmids (pET-BSP-MprT, pSG1154-BSP-MprT and pDR111-BSP-MprT) were then used to transform *E. coli* strain BL21 (DE3) pLysE (used for the high-level expression of genes cloned into vectors for expression of sequence downstream from the T7 promoter, provided that the cloned sequence contains a ribosomal binding) and *Bacillus subtilis* 1A751 as described below.

Bacteria were grown on LB and TBAB (see appendix 4) medium at 37°C. The recombinant *E. coli* was screened and cultured on LB medium containing kanamycin (30μg/mL) or ampicillin (100 μg/mL), and the recombinant *B. subtilis* on LB medium containing spectinomycin (100 μg/mL).

4.2.2 DNA Manipulation

Restriction endonucleases, T4 DNA ligase, and Taq DNA polymerase were purchased from New England Biolab (USA). Standard techniques were used for DNA isolation, enzyme digestion, agarose gel electrophoresis, DNA fragment purification, ligation, and transformation of *E. coli* as described in chapter 2 section 2.11.
4.2.3 Cloning of the Thermostable Protease Gene

According to the whole sequence of the thermostable metalloprotease gene, one pair of primers was designed for cloning the thermostable protease (MprT) gene, including the signal peptide, pro-peptide and mature peptide sequences.

**Table 4.1 List of primers used in this study**

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5’→3’)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSP forward</td>
<td>CGC<strong>CATATGAACAAACGGCGATG</strong></td>
<td>Forward primer for BSP-Mpr gene amplification, <em>Ndel</em> recognition sequence is underlined</td>
</tr>
<tr>
<td>BSP reverse</td>
<td>CGA<strong>CTCGAGTTATACACTCCAA</strong> CGC</td>
<td>Reverse primer of BSP-Mpr gene amplification, <em>XhoI</em> recognition sequence is underlined</td>
</tr>
<tr>
<td>TM-BSP-5H</td>
<td>GCGAAGCTTACATAAAGGAGGAAC<strong>TACTATGAACAAACGGCGATGC</strong> TT</td>
<td>Forward primer for Mpr gene cloning in pDR111, <em>HindIII</em> recognition sequence is underlined</td>
</tr>
<tr>
<td>P29-26-sp</td>
<td>GCGGCATG<strong>TCAAAAAAACCCCTCA</strong> AGACCCGTTTAG</td>
<td>Reverse primer for Mpr gene cloning in pDR111, <em>SphI</em> recognition sequence is underlined</td>
</tr>
<tr>
<td>TM-BSP-5Av</td>
<td>GCG<strong>CCT AAG ATG</strong> AAC AAA CGG GCG ATG CTT</td>
<td>Forward primer for Mpr gene cloning in pSG1154, <em>AvrII</em> recognition sequence is underlined</td>
</tr>
<tr>
<td>P29-26-Hi</td>
<td>GCG<strong>AAG CTTCAA AAA ACC CCT CA AGA CCC GTT</strong> TAG</td>
<td>Reverse primer for Mpr gene cloning in pSG1154, <em>HindIII</em> recognition sequence is underlined</td>
</tr>
</tbody>
</table>

Using genomic DNA of *Bacillus* strain BSP as a template, the nucleotide sequence of the Mpr gene was amplified by the BSP forward primer and the reverse primer (table 4.1). The PCR reaction was subjected to the following thermal cycles: one cycle at 94°C for 4 min, 30 cycles (94°C, 1 min; 62°C, 30 sec; and 72°C, 2 min), and a final
extension at 72°C for 10 min. According to the sequence of the MprT gene obtained, two sets of primers for MprT gene were designed for expression in *Bacillus subtilis* 1A751. The nucleotide sequence of the MprT gene was amplified using the forward primer (TM-BSP-5H) and the reverse primer (P29-26-sp) for cloning in pDR111 integration vector. Similarly Mpr gene was amplified using forward and reverse primers (TM-BSP-5Av and P29-26-Hi) for cloning in pSG1154 integration vector.

**Table 4.2 List of strains used in this study**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E.coli</em> JM109</td>
<td>endA1, recA1, gyrA96, thi, hsdR17 (r−, m+)</td>
<td>Promega (Madison, WI, USA)</td>
</tr>
<tr>
<td></td>
<td>relA1, supE44, A(lac-proAB), [F′, traD36, proAB, laqP]</td>
<td></td>
</tr>
<tr>
<td>BL21(DE3)pLysE</td>
<td>F−, ompT, hsdS_{B}(r_{B}−m_{B}−), gal, dcm, (DE3).</td>
<td>Novagen (Madison, WI, USA)</td>
</tr>
<tr>
<td><em>B. subtilis</em> 1A751</td>
<td>eglS^102 bglT/bglS^EV npr Mpr his</td>
<td>BGSC</td>
</tr>
<tr>
<td>1A751- pSG1154-BSP-Mpr</td>
<td>1A751 integrated with plasmid pSG1154 containing Mpr gene of BSP strain</td>
<td>This study</td>
</tr>
<tr>
<td>1A751- pSG1154</td>
<td>1A751 integrated with plasmid pDR111 alone</td>
<td>This study</td>
</tr>
<tr>
<td>1A751- pDR111-BSP-Mpr</td>
<td>1A751 integrated with plasmid pDR111 containing Mpr gene of BSP strain</td>
<td>This study</td>
</tr>
<tr>
<td>1A751- pDR111</td>
<td>1A751 integrated with plasmid pDR111 alone</td>
<td>This study</td>
</tr>
</tbody>
</table>

The final products were purified from the agarose gel using Wizard®SV Gel and PCR Clean-Up System according to the manufacturer’s guidelines as described in chapter 2 section 2.11.8 and all constructs were verified through DNA sequencing.
4.2.4 Construction of the Inducible Expression Vectors

For cloning the MprT gene into the vector pET29b, the MprT gene was digested by NdeI and XhoI restriction enzymes, and the resulting 1638 bp DNA fragment was purified and inserted into the same endonuclease sites of vector pET29b. In this way, the plasmid pET29b-BSP-MprT which harbors the MprT gene under the control of the T7 promoter and the native signal peptide sequence was gained.

An integration cassette was created, and integrated in the Bacillus subtilis genome by homologous recombination. The cassette was designed to include regions flanking the gene of interest. These regions were homologous to the Bacillus subtilis genome. Thereby, the cassette was integrated at a specific site on the genome. The cassette was created by joining DNA fragment with the flanking regions homologous to the B. subtilis genome.

For cloning the MprT gene into pdr111-amyE-hyper-SPANK, the MprT gene was digested using HindIII and SphI restriction enzymes, and the resulting 1638 bp DNA fragment was inserted into the same endonuclease sites of vector pdr111-amyE-hyperSPANK placing the genes under the control of the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter, thereby gaining the expression plasmid pDR-BSP-MprT which harbors the MprT gene under the control of the hyper spank promoter and the signal peptide sequence of the thermophilic neutral protease gene. Similarly, for cloning the MprT gene into pSG1154, the MprT gene was digested using AvrII and HindIII restriction enzymes, and the resulting 1638 bp DNA fragment was inserted into the same endonuclease sites of vector pSG1154 placing the genes under the control of the xylose-inducible promoter, thereby gaining the expression plasmid pSG-BSP-MprT which harbors the MprT gene under the control of the Pxyl and the signal peptide sequence of the thermophilic alkaline protease gene.

4.2.5 Modeling of 3-D Structures

With a view to studying the protein for its structural attributes, the theoretical 3D structure of metalloprotease was modeled using the Swiss Model Workspace (Expasy)
Figure 4-1 Vectors used in this study. (A) pET29b+ (Novagen, Madison, WI, USA). (B) pSG1154 integration vector having Pxyl (xylose inducible promoter) from BGSC. (C) pDR111-amyE-hyperSPANK (a gift from David Rudner, Harvard Medical School Boston, MA, USA) having IPTG inducible promoter.
server for protein 3D structure prediction. The output suggested one or more predicted 3D models from the repository, and a detailed information about the target protein and model building process, functional annotation, template selection log, target template alignment, and summary of building model with model quality assessment. The overall report generated a 3D template, the reliability of which was estimated by Q-mean 4 Global scores (Z-scores) for model quality estimation. In order to facilitate interpretation of the Z-score of the specified protein, its particular value is displayed in a plot that contains the Z-scores of all the experimentally determined protein chains in current PDB. The structure groups from different sources (X-ray, NMR) are distinguished by different colors (NMR dark blue and X-ray light blue). The plot can be used to examine whether the Z-score of the protein in question is within the range of scores usually found for proteins of similar size belonging to one of these groups. The Z-scores for high resolution X-ray structures are zero.

4.2.6 Transformation and Expression in *Escherichia coli*

The pET29-BSP-MprT plasmid was transformed into BL21(DE3) pLysE host cells. A liquid culture of the transformed colony was grown in Luria-Bertani (LB) broth by shaking at 37°C. Recombinant protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM when the optical cell density at 600 nm reached 1.5. After 12 h of additional growth, the bacteria were harvested by centrifugation and lysed with CelLytic™ B Cell Lysis Reagent (Sigma-aldrich St. Louis, MO, USA) containing lysozyme (Novagen) according to the manufacturer’s protocol as described in chapter 2 section 2.11.17.

4.2.7 Purification of Protease

The cell lysate was collected by centrifugation at 5000x g for 15 min and diluted up to 5 mL with 25 mM HEPES and filtered the supernatant using a 0.45 um (micrometer) pore membrane into syringe with Aktaprise adapter. The AktaPrime system and all other equipment and chromatography resins were from GE Healthcare Life Sciences (Piscataway, NJ). Buffer A was 25mL 1M HEPES (25mM final) and Buffer B was
Control programs were developed to complete cation exchange purifications without user intervention.

Cell lysate was sonicated for 30s on and 1 min off on ice (3 cycles) and all subsequent purification steps were conducted at 4 °C. The sonicated cell suspension was centrifuged for 30 min at 95,000 × g and the soluble fraction was retained. The soluble fraction was loaded into a 5 mL loading loop and then loaded onto purification system at 0.1 mL/min. This purifier system had 1 mL HiTrap Q HP columns equilibrated with buffer A. The columns were washed with eight volumes of a mixture of 100% buffer A and 0% buffer B. During the wash, the flow rate was increased to 1.5 mL/min. The bound protease was eluted from purification system 1 by a step-wise change to 100% buffer B. At the start of the elution step, the flow rate of buffer B was decreased to 0.3 mL/min. The fractions were used directly for activity assays.

4.2.8 SDS-PAGE

The fractions of purified protein on chromatogram were picked for precast 4-12% SDS-PAGE gel (Invitrogen) as described in chapter 2 section 2.9. The resolved protein visualized by silver staining (GelCode® SilverSNAP™ Stain Kit, Pierce IL, USA) according to the manufacture’s guidelines (see appendix 7).

4.2.9 Transformation of B. subtilis Strain 1A751

The resulting plasmids and vector controls (pSG1154 and pDR111) were transformed into B. subtilis 1A751 (see appendix 8 for media used), gaining the recombinants 1A751/pDR-BSP-MprT and 1A751/pSG-BSP-MprT. To check the development of competency in the pattern-producing strains and the ability of their DNA to transform B. subtilis, freshly prepared competent cells were prepared according to the two-step shift down procedure, as described by Anagnostopoulos and Spizizen [142] and Rudner et al. [143].
4.2.10 Screening Recombinant Clones for Alkaline Protease Activity

All Bacillus recombinant clones were spread on TBAB plates containing spectinomycin (100 ug/mL) overnight to screen recombinant colonies. After that the colonies were picked from TBAB plates and spotted on LB+starch (no antibiotic), LB+casein and LB+casein+IPTG for 1A751/pDR-BSP-MprT and on LB+starch, LB+casein and LB+casein+xylose (0.8%) for 1A751/pSG-BSP-MprT. All the plates were kept at 37°C overnight. Next day plates were checked and LB casein plates (with and without IPTG and xylose) were transferred to 60°C for overnight again. The reason for transferring casein plates to higher temperature is that BSP-MprT is a thermophilic enzyme and can only expressed at higher temperature.

4.2.11 Growth Conditions

Escherichia coli JM109 was used as the intermediary cloning host. E. coli strains and B. subtilis 1A751 were grown aerobically at 37°C in LB medium. B. subtilis growth media were supplemented with spectinomycin (100 mg/mL) when required. E. coli growth media were supplemented with ampicillin (100 mg/mL) when required.

4.2.12 Induction and Expression of the Thermostable Metalloprotease Gene

Expression of the pSG-BSP-MprT gene cluster (in strain 1A751) was induced by addition of xylose to the growth medium to a final concentration of 0.8% (w/v) and expression of the pDR-BSP-MprT gene cluster was induced by addition of 1 mM IPTG (in strain 1A751). B. subtilis strains were induced when OD_{600}=0.9. E. coli strains were grown up to OD_{600}=0.6 and then induced.

After induction recombinant Bacillus subtilis strains were incubated at 37°C with shaking (180 rpm) for 24 h. Samples were centrifuged at 7500xg for 20 min. The supernatant was used as crude enzyme to measure protease activity. Meanwhile, SDS-PAGE at 12% (w/v) was carried out to examine the expression of the
thermophilic alkaline protease gene; the proteins were visualized with Coomassie Brilliant Blue R-250.

**4.2.13 Detection of Protease Activity**

Protease activity on solid media was detected by visible clearings around microorganisms growing on LB agar plates containing 1% casein substrate. Proteolytic activity in liquid was determined using the EnzChek protease assay kit (Molecular Probes, Eugene, Oreg.) essentially according to the manufacturer's instructions as described in chapter 2 section 2.4.

**4.2.14 Protease Assay-Solid Phase**

*Escherichia coli* strain BL21 transformed with pET29-BSP-MprT and *B. Subtilis* colonies transformed with either pDR111 and pSG1154 (vector control) or pDR-BSP-MprT and pSG-BSP-MprT expression plasmid were spotted onto LB agar plates containing 1% casein, 1 mM IPTG or 0.8% xylose, and kanamycin (30μg/mL) or spectinomycin (100 μg/mL) antibiotic. The plates were incubated overnight at 37°C and then transferred to a higher temperature (60°C) at which the protease was active.

**4.2.15 Protease Assay-Liquid**

The proteolytic activity was determined using the protocol described in chapter 2 section 2.5.

**4.2.16 Effect of pH and Temperature on the Activity and Stability of Protease**

The effect of temperature on the activity of protease was determined by carrying out the enzyme assay at different temperatures viz. 40°C to 90°C. The thermostability of enzyme was determined by pre-incubating the enzyme at different temperatures ranging from 40°C to 90°C for 30 min (pH 8) and then determining the residual activity.
For determining the effect of pH on protease activity, different buffers (50 mM) used in enzyme assay mixture were: 50 mM sodium succinate buffer (pH 4-6) and sodium phosphate buffer (pH 7 - 11). For examining the pH stability of the protease the enzyme was pre-incubated with the buffer of appropriate pH for 30 min and then residual activity was assayed. Relative activity was determined by considering maximum activity as the standard reference.

4.2.17 Determination of Additives Effect

Various metal cations (Ca^{2+}, calcium chloride; Co^{2+}, cobalt chloride; Fe^{2+}, ferrous chloride; Mg^{2+}, magnesium chloride; Mn^{2+}, manganese sulfate; Ni^{2+}, nickel chloride; and Zn^{2+}, zinc sulfate) were pre-incubated with the protease for 30 min at 60°C. Various protease inhibitors (β-mercaptoethanol (β-ME; dithiothreitol, (DTT; ethylenediaminetetraacetic acid, EDTA; iodoacetate, IAA; and phenylmethylsulfonyl fluoride, PMSF) were added to the protease for 30 min at 37°C and then pre-incubated at 50°C for 30 min. All incubations occurred at pH 9. After pre-incubations, casein was added to all the enzymes, the reactions proceeded at 50°C, and residual protease activities were assayed as described above.

4.3 Results

The BSP-MprT gene was amplified by PCR, the nucleotide sequence of the MprT gene consists of the signal peptide sequence (1—75 bp), pro-peptide sequence (76—684 bp) and mature peptide sequence (685—1638 bp) (Figure 4.3). The MprT gene was cloned into an E.coli expression vector (pET29b+; Novagen, WI, USA). This plasmid was transformed into an E. coli expression strain (BL21(DE3) pLysE (Novagen).
Figure 4-2 Agarose gel illustrating the cloning strategy of BSP-MprT in pET29b(+) 

(A) Amplification by using BSP-F and BSP-R primer. Lane 1 represents 62°C PCR for BSP-Mpr. (B) Restricted fragments of pET29b and PSP-MprT PCR product. (C) Colony PCR of the transformants. Lane 1-30 represent colonies selected for PCR. (D) Restriction digestion result selected 8 colonies. Lane 1-8 represent the corresponding colonies. M: 1kb DNA ladder (NEB, USA).
Figure 4-3 Nucleotide and amino acid sequence of the BSP-MprT gene with a signal peptide (boxed), a propeptide (underlined) three residues (HExxH) at the active site (boldface) and the stop codon (indicated with an asterisk). The maturation site is indicated by the filled arrow.
Figure 4-4 Neighbor joining tree showing the phylogenetic relationship of protease gene from *Bacillus subtilis* BSP strain with the related species based on the protease gene sequences. Bootstrap values that are expressed as the percentages of 1000 replications are shown at the nodes of the branches.

### 4.3.1 Phylogenetic Analysis

Metalloprotease gene sequence of strain BSP was analyzed to determine its phylogenetic position. The protease gene sequence data of strain BSP was deposited in the GenBank (accession number: KP792450). A 1638-bp protease gene sequence was compared with the protease gene sequences of other related bacteria. The phylogenetic tree is shown in Figure 4-4. The BSP-MprT gene sequence similarity demonstrated that this protease was a member of the thermolysin family of metalloproteases. A 100% gene sequence similarity value with *Alicyclobacillus acidocaldarius* was observed.
4.3.2 Thermolysin Structure Analysis

The secondary structure analysis showed that the thermostable thermolysin protease seems stable. The target protein is good for 3D modeling as the alpha helix content is 35% which will make the protein stable (Figure 4-5). The predicted 3D structure of the BSP-Mpr protease was elucidated (Fig 4-6). The hypothetical protein model created is stored as PDB output file. The RMSD between the predicted model and templates was 0.060 Å. The lower values of RMSD suggest a higher homology between the predicted model and templates from the database. The Q-mean 4 Global Z-score value for the BSP-Mpr protease was -2.19 (Figure 4-7). A lower value of the Z-score indicates higher similarity in the structures. In general, the positive values correspond to problematic or erroneous parts of a model.

4.3.3 Expression Study of PSP-MprT in Escherichia coli

The ligated plasmid was used to transform E. coli BL21 for the expression study of protease gene. It was cultured in LB media in the presence of 30 µg/mL kanamycin which was induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 6 h. One culture of E. coli was kept without induction as control. The samples collected were analyzed for the protease expression both extracellular and intracellular. There was no extracellular expression of protease or any other protein as can be seen by Lane 1-3 (12 h, 20 h 30 h). The supernatant was precipitated by TCA and checked for protease expression but no expression was observed as can be seen by Lane 4-6 (12 h, 20 h 30 h) on SDS-PAGE (Figure 4-8). Intracellular expression of protease was observed which was good at even 12 h (Lane 7) and it was equally good at 20 h and 30 h (Lane 8 and 9). The molecular weight of protease was approximately 36 kDa as indicated by the marker used.
Figure 4-5 Secondary structure analysis of the metallprotease using SOPMA Expasy where (Hh) Alpha helix 35.16%; (Ee) Extended strand 25.82%; (Tt) Beta turn 11.36%; (Cc) Random coil 27.66%

Figure 4-6 Predicted 3D structure of BSp-Mpr protease enzyme by Swiss model workspace structure prediction tool
Figure 4-7 Z-Score Plot shows Z-score value of protein (dark blue region represents protein determined by NMR and light blue represents protein determined by X-ray). The black dot represents Z-score of the model protein.

Figure 4-8 SDS-PAGE of BSP-MprT expressed in E.coli. Lane M shows BenchMark™ Protein Ladder. Lane 1-3 culture supernatant of BL21-BSP-MprT, Lane 4-6 TCA precipitated supernatant, Lane 7-9 cell lysate.
4.3.4 Purification of BSP-MprT Expressed in *E. coli* BL21 (DE3)

At the end of the expression period samples were taken for the preparation of protein extracts. Intracellular expressed protease was collected and undergone purification through Hi trap Q column using AKTA prime (GE Healthcare Life Sciences).

After Q-column, peak fractions scene on chromatogram (Figure 4-9) are selected for SDS-PAGE analysis. Gel was silver stained and the results showed that the size of BSP-MprT expressed was as expected (36 kDa), as can be seen from the gel.

4.3.5 Expression of BSP-MprT in *Bacillus subtilis* Expression System

To study the expression of the thermophilic neutral protease gene in *B. subtilis* 1A751, integration vectors pDR111 and pSG1154 were tested. One plasmid pDR111-BSP-MprT was created by inserting the BSP-MprT gene into the plasmid (Figure 4-11). Approximately 30—40 transformants were obtained after the recombinants were screened on LB/agar plates containing spectinomycin. The insertions were then confirmed by selection on LB/starch/agar and LB/casein/agar plates. Transformants forming halos on LB/casein/agar but not on LB/starch/agar confirmed the insertion of protease gene.

Similarly for the expression of BSP-MprT gene using pSG1154 integration vector, plasmid pSG-BSP-MprT was created by inserting the BSP-MprT gene into the pSG1154 (Figure 4-12). Transformants obtained after the transformation were screened on LB/agar plates containing spectinomycin. The recombinants were then confirmed by selection on LB/starch/agar and LB/casein/agar plates. Transformants forming halos on LB/casein/agar but not on LB/starch/agar confirmed the insertion of protease gene.

In order to ensure that the nucleotide sequence of the thermophilic neutral protease gene in integration vectors was correct, sequencing was done and whole sequence was also measured. The result indicated that BSP-MprT sequence was right
Figure 4-9 Elution profile of anion exchange chromatography of protease (BSP-MprT) on HiTrap Q

Figure 4-10 Silver-stained SDS-polyacrylamide gel showing the expression of BSP-Mpr in eluted fractions
Figure 4-11 Agarose gel illustrating the cloning strategy of BSP-MprT in pDR111. (A) Amplification by using TM-BSP-5H and P29-26-sp primers. Lane 1-2 represent PCR for BSP-MprT. (B) Restricted fragments of pDR111 and PSP-MprT PCR product. (C) Restriction digestion result of 6 selected colonies. Lane 1-6 represent the corresponding colonies. M: 1kb DNA ladder (NEB, USA).
Figure 4.12 Agarose gel illustrating the cloning strategy of BSP-MprT in pSG1154. (A) Amplification by using TM-BSP-5Av and P29-26-Hi primer. Lane 1 represent PCR for BSP-MprT. (B) Restricted fragments of pSG1154 and PSP-MprT PCR product. (C) Restriction digestion of plasmid isolated from selected 5 colonies. Lane 1-6 represent the corresponding colonies. M: 1kb DNA ladder (NEB, USA).
and that, in fact, the cloned thermophilic neutral protease gene shared a 99% similarity with the reported sequence.

### 4.3.6 Expression of the Thermostable Protease Gene in *B. subtilis* 1A751

The levels of expressions and solubility of recombinant proteins are directly related to the expression vector used and a variety of known and unknown protein characteristics. Consequently, the solubility levels of expressed proteins directly contribute to the final yield of purified proteins.

The recombinant plasmids pDR-BSP-MprT and pSG-BSP-MprT were transformed into *B. subtilis* 1A751, while the expression of the inserted thermophilic alkaline protease gene in the recombinant strain was induced by adding IPTG or xylose respectively into the medium.

To determine the relative amounts of pDR-MprT and pSG-MprT proteins produced by the expression systems being tested, protease assays were performed (Figure 4-14). Comparison of the expression levels of pDR-MprT and pSG-MprT obtained in the xyl system and the hyper-spank systems showed slightly higher

**Figure 4-13** Assay of the *B. subtilis*-BSP-Mpr on a casein plate
Colonies transformed with either pDR11-BSP-MprT (right) or pSG-BSP-MprT (left) were spotted onto LB agar plates containing 1% casein and 1 mM IPTG or 0.8 % xylose
amounts of protein generated after induction in the hyper spank system than in xyl system (Figure 4-14). So, we selected pDR-BSP-MprT for further characterization of protease.

![Protease activities measured in B. subtilis expression systems under the P_\text{xyl} and P_\text{hyper-spank} promoters. Negative controls: 1, wild type strain 1A751; 2, 1A751-pSG; and 3, 1A751-pDR. Lane 4 shows protease activity measured in the xylose-inducible system; lane 5 shows protease activity in the IPTG-inducible expression. All values represent an average of at least three separate protease activity measurements.](image)

**Figure 4-14**

**4.3.7 SDS-PAGE Analysis of BSP-MprT Expressed in B. subtilis**

The SDS-PAGE result of the supernatant indicated that MprT was found in the supernatant of the culture medium, which in turn, revealed that MprT was secreted from *B. subtilis* cells successfully (Figure 4-15).
Figure 4-15 SDS-PAGE of BSP-MprT expressed in *B. subtilis* 1A751

Lane M shows BenchMark™ Protein Ladder (Invitrogen Cat. No.10747-012). Lane 1 is the negative control (*Bacillus subtilis* 1A751), Lane 2 culture supernatant of 1A751/pDR-BSP-MprT, Lane 3 culture supernatant of 1A751/pSG-BSP-MprT

4.3.8 Determination of Optimum pH and Temperature

The activity of the recombinant protease was tested under a variety of pH values. The enzyme had the highest activity from pH 7-8 (Figure 4-16). At pH 6 and pH 9, the protease retained 82% and 73% of the activity respectively. The enzyme activity was checked at temperature range of 50-90°C. The temperature optimum was determined to be 50°C (Figure 4-17). This is consistent with the optimal growth temperature (50°C) of the source organism. The enzyme activity was decreased to 87% at 60°C.
Figure 4-16 Activity of protease at various pH and 50°C after 20 min of incubation

Figure 4-17 Activity of protease at various temperatures and pH 8 after 20 min of incubation
The thermostability of the protease was tested at various temperatures and times (Figure 4.18). At 50°C, there is a slight drop in activity after 60 min and after 3 hours the activity decreased to 70%. At 60°C, there is a larger decrease to approximately 70% after 60 min. At 70°C, half the activity is lost at 60 min. Finally, at 80°C and 90°C, only 19% and 13% of the activity retained after 30 min incubation.

**4.3.9 Effect of Additives on Protease Activity**

The effects of numerous divalent cations on protease activity were tested (Fig 4-19). Ni^{2+}, Zn^{2+}, Mg^{2+} and Co^{2+} decreased enzyme activity to 55–76%, while Cu^{2+} reduced the activity to 22%. Ca^{2+}, Fe^{2+} and Mn^{2+} stimulated the protease activity to some extent i.e. 122%, 113% and 104% respectively of the unsupplemented enzyme level. These metal ions protected the enzyme from thermal denaturation and maintained its active conformation at the high temperature. In addition, protease
required a divalent cation like Ca\textsuperscript{2+} and Mn\textsuperscript{2+} or combination of these cations for its maximum activity [144]. In addition, these cations enhanced the stability of a Bacillus protease [145].

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4-19}
\caption{Effect of divalent metal ions on recombinant protease activity (at 50°C, for 30 min)}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4-20}
\caption{Effect of inhibitors on recombinant protease activity (at 50°C, for 30 min)}
\end{figure}
The effect of various inhibitors was studied at 5 mM final concentration. Our results indicated that the addition of 5 mM EDTA showed an inhibitory effect on protease activity at concentration of 5 mM. PMSF reduced activity to 48%. The β-ME and DTT reducing agents both induced inhibition to 54% and 60%, while IAA showed moderate inhibition.

### 4.4 Discussion

New industrial enzymes produced by *Bacillus* strains such as proteases have emerged as a result of cloning, with desired activity, specificity and stability properties. It provides a tremendous method for the control and manipulation of genes. More than 50% of the industrially important enzymes are now produced from genetically engineered microorganisms [11, 146]. By the same token, the cloning of a structural gene for an extracellular protease from the thermophilic *Bacillus* sp would be significant practically as well as academically. Since, the main objective of this study was to cloning protease gene of *Bacillus sp* in *E. coli* BL21 DE3 and *Bacillus subtilis* 1A751 deficient in protease production and studying gene expression in the new hosts.

There are different expression systems for recombinant proteins, two widely used ones being *E.coli* and *Bacillus* expression systems. *E.coli* expressions systems are commonly used because *E.coli* is a rapidly growing microorganism and it is compatible with many reagents and vectors however *Bacillus subtilis* serves as an attractive expression host for heterologous protein production because it is nonpathogenic and capable of secreting extracellular proteins directly to the culture medium [147].

Protease encoding gene was hardly secreted from *E. coli*, therefore *E. coli* would not be a good option if the secretory proteins were expected. Moreover, the secretion of proteins would simplify the extraction procedure. *B. subtilis* 1A751 and PY79 are effective hosts for secretory proteins since there are abundant of vectors and regulatory elements for *Bacillus* species to help proteins folding.

Many different expression systems have been developed for *B. subtilis*. Here we compare two widely used expression systems, the IPTG-inducible derivative of spac
system (hyper-spank) and the xylose-inducible (xyl) system. Expression was triggered from plasmids that were integrated into the chromosome at the amyE locus (xyl and hyper-spank system).

Heterologous expression of the BSP-MprT gene encoding a thermolysin like metalloprotease in *E. coli* was investigated. The chromosomal DNA from *Bacillus subtilis* BSP was extracted and used as template to amplify the open reading frame (ORF) encoding a protease protein. Approximately 1.7 kb size of DNA was amplified by using forward (BSP-F) and reverse primer (BSP-R). The restriction digested purified vector and insert were ligated using T4 DNA ligase (NEB) at 16 °C overnight in the circulatory water bath. The ligated product was used for transformation of *E. coli* JM109 which was grown overnight on LB agar plate with supplementation of kanamycin (30 µg/mL). Gohel and Singh [148] have reported of an alkaline protease gene amplified from genomic DNA of alkaliophilic actenomycetes by designing six pairs of primers based on nucleotide sequence of alkaline protease gene from organisms of haloalkaliphilic origin. Two sets of primers showed amplification of DNA of about 0.5 and 1.2 kb from *Bacillus subtilis* L010 which were cloned into pET 12a+ vector for expression studies. In another study, Li et al. [149] has reported the cloning of alkaline protease into pET 28a vector. The ORF of the protease was 1149 bp long and encoded a protein of 382 amino acids comprised with a 30-residual signal peptide, a 77 -residual propeptide, and a 275 residual mature protein, and the encoded protein was one of subtilisins—a member of serine proteases and designated as SprD. The precursor of SprD (proSprD) autoprocessed into active SprD mediated by the propeptide when pro-SprD was recombinant expressed in BL21 (DE3). Ali et al. [150] cloned and expressed an extracellular protease from thermophilic *Bacillus stearothermophilus* into *E. coli*. An alkaline protease from *B. subtilis* strain 168 in *E. coli* was also reported in order to characterize the enzyme but did not measure enzyme expression in the recombinant host [151].

A similar study on recombinant protease purification from transformed BL21 cells by affinity chromatography (HiTrap Q Hp column; GE Healthcare USA) using an Aktaprime system (GE Healthcare, USA) was reported [152]. BtsTLP1 protease was purified by two successive anion-exchange chromatography steps. a HiPrepTM Q XL 16/10 column (GE Healthcare, Pittsburgh, USA) pre-equilibrated with buffer A.
Fractions containing the target protein were pooled, concentrated and buffer exchanged. The resulting concentrate was then loaded onto a HiPrepTM DEAE FF 16/10 column (GE Healthcare, Pittsburgh, USA) equilibrated with buffer B and purified. Following the elution, BtsTLP1- containing fractions were pooled, concentrated via the 10K Amicon Ultra device and stored in 40% glycerol at -20°C.

Huesgen et al. [153] also reported recombinant protease purification by nickel affinity chromatography (HisTrap) using an Akta purifier fast protein liquid chromatography system (GE Healthcare). Other protease purification report involved ammonium sulfate precipitation, gel filtration by Sephadex G-100, and anion-exchange chromatography by DEAE-cellulose [154]. Yang et al. [155] purified thermolysin-like protease from *Halo Bacillus* sp. SCSIO 20089 by combination of ammonium sulfate precipitation, DEAE Sephadex ion exchange chromatography and Sephadex G-75 gel filtration. The AKTA FPLC system (Amersham Biosciences, Piscataway, NJ, USA) was used in these experiments.

The plasmid DNA from the clone obtained was analyzed for the nucleotide sequence of the insert. The sequence analysis revealed an open reading frame of 1638 bp encoding 546 amino acids which would then be processed to produce a mature enzyme of 36 kDa (318 amino acids). Berg et al. [156] have reported nucleotide sequence analysis and shown an ORF of 1632 bp encoding a protein of 544 amino acids, which is very close to the nucleotide sequence of protease gene of the present study.

The ligated plasmid was used to transform *E. coli* BL21 for the expression study of protease gene. It was cultured in LB media in the presence of 30 µg/mL kanamycin which was induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 6 h [157]. The samples collected were analyzed for the protease expression both extracellular and intracellular. There was no extracellular expression of protease or any other Intracellular expression of protease was observed after 12 hr and the molecular weight of protease was approximately 36 kDa. Lee et al. [158] reported a molecular weight of 38 kDa of an extracellular metal protease isolated from a *Pseudodalteromonas* sp. strain A28. Though the over-expression of protease was observed by SDS-PAGE the functional activity of protease was very less. The reduced functional expression of protease may be due to the formation of inclusion
bodies. Majority of proteins in *E. coli* are expressed as inclusion bodies [159]. Pushpam et al. [160] have also reported the expression of protease as inclusion body which was later solubilized with a Tris buffer containing urea, sodium chloride, imidazole and 2-mercaptopethanol to make it functional.

For extracellular expression, the BSP-MprT was successfully expressed in *Bacillus subtilis*. The safety and yield of the production of BSP-MprT in recombinant *B. subtilis* was found better and higher than that in recombinant *E. coli*. Meanwhile, the over-expression and secretion are able to reduce the cost of the large-scale industrial production. Gao et al. [138] cloned an aminopeptidase gene from *Bacillus subtilis* Zj016 (BSAP) into *B. subtilis* expression system and showed enhancement of enzyme expression in the recombinant host compared to the original strain and a molecular weight of 50 kDa. In another study an alkaline protease was cloned from a *B. pumilus* strain in *B. subtilis* and demonstrated over 6 fold enhanced expression in the recombinant strain [146]. We compared two *B. subtilis* expression systems, the IPTG-inducible (pDR111) and the xylose-inducible (pSG1154) system to analyze their applicability for production of extracellular protease. The *amyE* locus, coding for a nonessential α-amylase, is used in both cases for ectopic integration. This system has been developed by Hidenori and Henner [161] and contains in its simplest form an antibiotic resistance marker and a multiple cloning site sandwiched between the two halves of the *amyE* gene, designated *amyE*-front and *amyE*-back. Upon transformation of *B. subtilis* cells, both *amyE* sequences will recombine at their homologous sites, thereby stably inserting the DNA sequences in between *amyE*-front and *amyE*-back into the *B. subtilis* chromosome via a double-crossover event. The MprT gene from Bacillus spp was cloned and expressed in *B. subtilis* 1A751, deficient in proteases using the expression vector pDR111 and pSG1154. Our results showed a 2.6 and 1.7 fold higher expression of the cloned MprT gene in 1A751-pDR-BSP-MprT and 1A751-pSG-BSP-MprT compared to its natural host before any optimization studies.

The molecular weight of the protease recovered from the recombinant strain was 36 kDa, similar to the enzyme obtained from the *Bacillus* parent strain. The expression vector pDR111 allows for better expression of the inserted gene (s) under the expressed P<sub>hyperspank</sub> promoter. On the other hand, protease production in *B. subtilis*
was inducible and occurred in the presence of IPTG. Following the same trend, maximum protease production occurred after 24 h in \textit{B. subtilis} 1A751/MprT compared to 72 h in the native strain. In a study Vavrová et al. [162] compared two expression systems, the IPTG-inducible derivative of spac system (hyper-spank) and the xylose-inducible (xyl) to the SURE (subtilin-regulated gene expression) system. Western blot analysis of the membrane protein SpoIISA together with its protein partner SpoIISB showed that the highest expression level of this complex is obtained using the SURE system. Measurement of β-galactosidase activities of the promoter-lacZ fusions in individual expression systems confirmed that the PspaS promoter of the SURE system is the strongest of those compared. Based on these results, it was concluded that the SURE system is the most efficient of these three \textit{B. subtilis} expression systems in terms of the amount of expressed product. Bhavsar et al. [163] compared the xylose-dependent expression system expression system to the isopropyl-\textit{b}-D-thiogalactopyranoside-induced \textit{spac} system using a thermostable β-galactosidase reporter (BgaB) and found the \textit{xyl} promoter-operator to have a greater capacity for modulated expression, a higher induction/repression ratio (279-fold for the \textit{xyl} system versus 24-fold with the \textit{spac} promoter), and lower levels of expression in the absence of an inducer. These findings do not support our study. So, in case of protease expression in \textit{Bacillus subtilis} hyper-spank expression system was found more efficient than xylose-dependent expression system.

Despite the number of thermostable proteases which have been introduced, development of recombinant organisms which can over express the enzyme is useful for the expanding industrial needs in different communities.

Recombinant 1A751-pDR-MprT was further biochemically characterized to check its stability under native enzyme conditions and it showed same behavior when compared to native protease. It is widely accepted that an excess amount of Zn$^{2+}$ or Cu$^{2+}$ can inhibit the proteolytic activities of thermolysin homologs. Although the exact mechanism was unclear, a few studies suggested that binding to these divalent metal ions could alter the conformation of the wild-type enzyme. For example, a 3D structural analysis of thermolysin saturated with exogenous Zn$^{2+}$ indicated that in addition to the catalytic Zn$^{2+}$, a second Zn$^{2+}$ could interact with the enzyme side chains,
thereby excluding the substrate from the active site [164]. Similarly, Fukasawa et al. [165] proposed that Cu$^{2+}$ substitution of the active-site Zn$^{2+}$ in thermolysin homologs might increase the rigidity of the catalytic site, limiting its interaction with both the substrate and solvent H$_2$O. It is very likely that such mechanisms could also account for the BtsTLP1 inhibition as the active sites are conserved within the thermolysin family. Our data show that both Zn$^{2+}$ and Cu$_2^+$ reduced the catalytic activity of BSP-MprT which could provide new insights into their inhibitory role.
5 Enhanced Thermostability of Thermolysin
Protease from Thermophilic *Bacillus subtilis*
BSP Strain

5.1 Introduction
The successful implementation of enzymes as industrial biocatalysts requires the availability of suitable enzymes with high activity, specificity and stability under process conditions. However, naturally occurring enzymes are often not optimized to fulfill these requirements. Within this context, directed evolution is very effective in closing these functional discrepancies [166].

Directed enzyme evolution has become a powerful alternative to rational approaches for engineering biocatalysts [167-169]. It can generate proteases with improved functions to meet the requirements of commercial applications [170]. Modern enzyme development relies to an increasing extent on strategies based on diversity generation followed by screening for variants with optimized properties. The directed evolution strategies might be used for optimizing any enzyme property, which can be screened for in an economically feasible way, even if the molecular basis of that property is not known. Stability is an interesting property of enzymes because (1) it is of great industrial importance, (2) it is relatively easy to screen for, and (3) the molecular basis of stability relates closely to contemporary issues in protein science such as the protein folding problem and protein folding diseases. Thus, engineering enzyme stability is of both commercial and scientific interest [171].

Generating and identifying enzymes adapted to non-natural (non-biological) conditions and achieving functional or economically efficient catalytic rates in such conditions is a particularly difficult task; despite being able to successfully modify protein sequence and change enzymatic properties, our understanding of the molecular mechanisms which will allow predicting amino acid sequence changes that would result
in a specific behavior remains incomplete. Successfully engineered enzymes for non-natural conditions enabled biocatalysts to be efficiently used in industrial processes which were traditionally performed chemically. Retroactively, in the past decade, the increasing demand of biocatalyst as replacement for chemical processes has progressively fueled the need for novel and improved enzymes applied to product-driven biocatalysis fields such as the synthesis of pharmaceutical precursors and the production fine chemicals [172].

Comparing studies of either naturally evolved homologous enzymes or artificially engineered enzymes have demonstrated that there is a trade-off between the rigidity required for stability and the flexibility necessary for activity in most enzymes [62]. Nevertheless, some variant enzymes coupling high activity and high stability have been engineered by directed evolution [76] or site-directed mutagenesis [173]. Thermostability has always been considered among the critical factors determining the feasibility of using proteases for industrial applications because thermal inactivation represents a common problem in the application of enzyme. Efforts are continuously being undertaken towards enhancing enzyme stability and functionality because naturally occurring enzymes often deficient in features necessary for commercial/industrial applications. In fact, several studies have taken these issues into account. The specific activity of the subtilisin E I31L mutant has, for instance, been substantially increased to bring about a prominent 26-fold increase in the \( k_{\text{cat}}/K_m \) values. The thermostability of subtilisin E has also been enhanced by substituting P239A and introducing of a disulfide bridge (Cys61/Cys98) using site-directed mutagenesis [174]. Furthermore, the thermostability and \( k_{\text{cat}}/K_m \) value of AprP, an alkaline protease from *Pseudomonas* sp., were enhanced by the introduction of a disulfide bond (G199C/F236C) [175]. Similarly, highly thermostable variants of mesophilic subtilisin E [176] and psychrophilic subtilisin S41 [177] have been developed by directed evolution, without sacrificing their low-temperature activities.

In directed evolution, enzymes with the desired property or set of properties are identified by selection or screening, and mutants may serve as starting points for additional rounds of mutagenesis to accumulate beneficial mutations. One of the most popular directed evolution methods to improve enzyme properties is error-prone polymerase chain reaction (PCR), which is based on the increasing frequency of
mismatched incorporation of nucleotides into newly synthesized PCR products [178-180].

The aim of this work was to improve an industrially important mutant protease by means of directed evolution using the error-prone PCR method towards higher thermostability while keeping its catalytic activity unaffected. A second objective was to study the amino acid substitutions identified in the improved variants in order to investigate the possible structure-function relationships involved in this adaptation. A micro titer plate expression and proteolytic assay systems were adapted for screening for increased residual activity at elevated temperature. The mutant Apr-BSP with the highest activity was selected and sequenced. Then, the enzyme was characterized biochemically and the results were compared with that of the parent enzyme.

5.2 Materials and Methods

5.2.1 Strains and Media

E. coli JM 109 [(recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, Δ(lac-proAB)/F′(traD36, proAB+, lacIq, lacZΔM15)] (Promega; WI, USA) and E. coli BL21-CodonPlus (DE3)-RIL ( B F- ompT hsdS( rB- mB- ) dcm+ Tetr E. coli gal λ (DE3) endA Hte [argU ileY leuW Camr] was obtained from Novagen (Madison, WI, USA) was used as host strains for DNA manipulation and gene expression.

E. coli strains and E. coli transformants used in this study for sub cloning and plasmid recovery was grown in Luria broth (LB) with 30 µg/mL of kanamycin at 37°C, 200 rpm for 24 h. For cultivation on agar plates, LB agar and casein agar was used.

5.2.2 Plasmids

pET29b(+) (Novagen) was used as vector for cloning, expression, and DNA sequencing. This vector has kanamycin resistant gene as selective marker and this gene can be expressed in E. coli.
5.2.3 Random Mutagenesis

In our previous study, the DNA encoding the thermostable metalloprotease of *Bacillus subtilis* BSP been identified and subcloned into pET29b(+), subsequently transformed and maintained in *E.coli* JM109 (chapter 4). The plasmid including the target gene was isolated by plasmid isolation kit (Qiagen CA, USA) from *E. coli* JM109 after overnight cultivation in LB broth, and used as a template for mutagenesis. Random mutagenesis was performed by error prone PCR (GeneMorph® II Random Mutagenesis Kit, CA, USA).

Random mutant gene library was constructed by error prone PCR [181]. In the present study, the thermostability of an organic solvent tolerant *B. subtilis* BSP protease has been improved by two rounds of error-prone PCR (eP-PCR). The DNA fragment encoding the mature enzyme was amplified from plasmid pET29b (+)-BSP-Mpr with primer F, 5' cgcctATGAACAAACGGCAGT 3' (underlined section is the *Nde*I restriction site) and primer R, 5' cgactcgagTTAATACACTCCAACC 3' (underlined section is the *Xho*I restriction site).

The PCR conditions included 41.5 µL of water, 5 µL of 10× Mutazyme II reaction buffer, 1 µL of 40 mM dNTP mix (200 µM each final), 0.5 µL of primer mix (250 ng/µL of each primer), 1 µL of Mutazyme II DNA polymerase (2.5 U/µL), 1 µL pET29b-BSP-MprT plasmid template (750 ng/µL).

PCR was carried out as 95°C for 2 min and 30 cycles of 95°C for 30 sec, 62°C for 30 s and 72°C for 2 min and then extension at 72°C for 10 min.

5.2.4 Cloning of Mutant-BSP-MprT Gene

All gene cloning and manipulation steps were carried out according to the molecular cloning method [182]. Transformation was carried out as standard heat shock transformation according to Promega (WI, USA) as described in chapter 2 section 2.11.14.2. Amplified BSP-MprT with its pre-pro-sequence including the *Bacillus* promoter was cloned into pET29b (+) by using *Nde*I and *Xho*I restriction sites. The generated construct was named pET-BSP-Mutant-MprT (7009 bp).
Plasmids were verified by PCR, restriction enzyme digestion and sequence analysis. Sequence results were analyzed by Vector NTI 9 (Invitrogen). Plasmid isolation, gel purification, PCR purification and Mini prep kits were purchased from Qiagen (CA, USA).

5.2.5 Construction of Expression Vector and Sub-cloning into *E. coli*

The amplified products were digested with *Nde*I and *Xho*I, and then ligated into the *Nde*I–*Xho*I restriction site of pET29b (+) with T4 DNA ligase (as described in chapter 2 section 2.11.13) to construct the random mutant library. Ligation products of random mutant or recombinant gene library were transferred into *Escherichia coli* BL21 (DE3), and the cells were spread on Luria-Bertani (LB) plates containing 30ug/mL of kanamycin and 1% of casein. After incubation at 37°C overnight to form colonies, the plates were placed at 60 °C. The colonies that formed clear-zone faster than the parental plasmids (pET29b (+)-BSP-Mpr) were selected. Double-stranded plasmid was isolated and sequenced to ascertain that the appropriate mutation(s) was introduced.

5.2.6 Sequencing Results and Analysis

Sequencing results of the selected variants were analysed using Vector NTI 9 (Invitrogen). The variant showed the amino acid substitution at the position 347. Amino acid sequences of both the wild type BSP-Mpr and Mutant-BSP-Mpr were compared by CLUSTAL W (1.83) multiple sequence alignment and the homology structural modeling of protease was carried out on the Swiss-Model Server [183]. The structural modeling was viewed and analyzed using the software RasMol coupled to the PDB (protein data bank) [184].

5.2.7 Expression of Mutant-BSP-MprT Gene

The pET29-BSP-MprT-Mutant plasmid was transformed into BL21(DE3)RIL host cells. The candidate strains were individually cultured in 15 mL BD falcons (Company) containing 3 mL of LB medium with 30 ug/mL of kanamycin at 37°C,
until the OD$_{600}$ reached the level of 0.6. Recombinant protein expression was induced by the addition of 1.0 mM isopropyl-thiogalactopyranoside (IPTG) into the medium. After 3 h of additional growth, the cells were harvested by centrifugation and lysed with CellLytic™B (Sigma) according to the manufacturer's protocol.

**5.2.8 Indicator Plates for Pre-screening**

Casein was used as a substrate for proteolytic activity detection [185]. LB agar plates supplemented with 1 % (w/v) casein and 30 µg/mL kanamycin were used as pre-screening plates.

**5.2.9 Enzyme Activity Detection**

Protease activity on solid media was detected by visible clearings around microorganisms growing on LB agar plates containing 1% casein substrate. Proteolytic activity in liquid was determined using the EnzChek assay kit (Molecular Probes, Eugene, Oreg.) essentially according to the manufacturer's instructions and read using a with an excitation/emission maxima at 590/645 nm. Values from controls without added enzyme was subtracted from sample values. Unit of enzyme activity for expression studies was determined as described in chapter 2 section 2.11.17.

**5.2.10 Screening for Improved Thermostability**

Thermal inactivation was determined by incubating the enzyme variants (5 µL) in the range of 50–90°C. After 30 min, the enzyme was immediately placed on ice for 10 min to decrease the temperature of the sample, and then the remaining activity was measured. The remaining activity was recorded as percentage of the original activity.

**5.2.11 Thermal Inactivation Profile**

Cell culture supernatants (1 mL) were incubated at different temperatures for 30 min. The time course of the heat inactivation of the enzyme was examined at 60°C, 70°C, 80°C and 90°. At various time intervals, aliquots were withdrawn and cooled on ice immediately, and then subjected to caseinolytic activity assay at 60°C, and residual
activity was calculated. Residual activity (%) was defined as the activity of the treated sample in absorbance unit per min (AU/min) divided by the activity of the untreated sample, multiplied by 100.

### 5.2.12 Comparison of Cultivation Profiles of Wild and Mutants in Shake Flasks

The wild type and transformants with the highest protease enzyme secreting ability at tested higher temperature were selected and used for shake flask cultivation. The media consisting of 0.5% yeast extract, 1.0% peptone and 1.0% NaCl was used. The pH was adjusted to 8 and the media were sterilized by autoclave for 20 min at 121°C. For pre-culturing, 10 mL of the above described media in 100 mL baffled Erlenmeyer flask was inoculated with one colony of each mutant and wild type culture in each flask and operated on a rotary shaker overnight at 200 rpm. These cultures were used to inoculate 30 mL of the growth media in 250 mL baffled Erlenmeyer flasks to a starting OD 600 of 0.1. Protein expression was induced by the addition of 1.0 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) into the media and cultivation was carried out for 24 h. Biomass was harvested by centrifugation from both the media and protease activity of the cell lysates was determined as discussed above.

### 5.2.13 Kinetics of Thermal Inactivation and Estimation of Inactivation Energy

In order to study the thermal inactivation kinetics of wild type protease and Mutant-protease, both enzymes were incubated at different temperatures in the absence of the substrate. At periodic intervals, aliquots were withdrawn and assay was performed. The residual activity was expressed as percent of the initial activity. The inactivation rate constants (k_d) were calculated from slopes of a semi logarithmic plot of residual activity versus time and apparent half-lives were estimated using Eq. (1). The time where the residual activity reaches 50 % is known as the half-life

\[ T_{1/2} = \frac{\ln (2)}{k_d} \]  

(1)
The temperature dependence of $k_d$ was analyzed using the Arrhenius plot. The inactivation energy was calculated from the Arrhenius equation as

\[
\ln(k_d) = \ln(K_0) - (E_d/R)1/T \quad \text{(2)}
\]

The values of $E_d$ and $k_0$ were estimated from the slope and intercept of the plot of $\ln(kd)$ versus $1/T$, respectively.

### 5.2.14 Estimation of Thermodynamic Parameters

The thermal stability of protease was determined by the inactivation rate constant ($k_d$) as a function of temperature at 50°C, 60°C, 70°C, 80°C, and 90°C. The temperature dependence of $k_d$ was analyzed from Arrhenius plot (natural logarithm of $k_d$ versus reciprocal of the absolute temperature); the activation energy ($E_d$) was obtained from the slope of the plot. Activation enthalpy ($\Delta H^0$) was calculated according to the equation

\[
\Delta H^0 = E_a - RT \quad \text{(3)}
\]

where, $T = \text{absolute temperature}$ and $R$ is the universal gas constant. The values for free energy of inactivation ($\Delta G^0$) at different temperatures were calculated from the equation

\[
\Delta G^0 = -RT \ln(k_d h/kT) \quad \text{(4)}
\]

Where, $k = \text{Boltzmann constant}$ and $h$ is the Planck constant. Activation entropy ($\Delta S^0$) was obtained from equation

\[
\Delta S^0 = (\Delta H^0 - \Delta G^0)/T \quad \text{(5)}
\]

### 5.3 Results

Protein engineering approach, random mutagenesis was chosen for creating protease variants which are more stable at temperature higher than 60°C. This approach was chosen based on our limited ability to predict a mutation's effects on enzyme stability and activity at higher temperature. The library was created and screened for enhanced
stability at higher temperature. The thermostability of protease was improved by the error prone PCR technique of random mutagenesis.

5.3.1 Plasmid DNA Isolation and Purification

This procedure was carried out using pET29b-BSP-MprT plasmid as target for mutation. It was derived from ligation of a BSP-MprT gene into pET29b(+) cloning vector as described in previous chapter. After completion of plasmid DNA isolation, the pET-BSP-MprT plasmid was run on a 0.9% agarose gel electrophoresis and visualized under ultraviolet light using the Biorad Gel-Doc system. Upon visualization the present bands confirms the plasmid pET-BSP-MprT.

5.3.2 Random Mutagenesis

Random mutagenesis was carried out using the GeneMorph® II Random Mutagenesis Kit (Stratagene). The GeneMorph protocol was followed and the initial target amount of 100 ng (equivalent to 480 ng plasmid template) with 25 PCR cycles was used to achieve a low mutation frequency (0–4.5 nucleotides/kb). The product was analyzed on an agarose gel (Figure 5-1) with 100ng of 1 kb DNA standard (New England Biolab). The random mutagenesis reaction produced product of the correct size (1.638 kb) and had a good yield. The 1.638 kb corresponds to the expected size of the PCR insert. The mutant PCR product and the vector pET29b+ were digested with _NdeI_ and _XhoI_. pET29b(+) was undergone dephosphorylation before ligation as described in chapter 2 section 2.11.10.

Error prone PCR was conducted as described above. The experiment was designed such that not to change the sequence such that it bears no similarity to the donor sequence but rather to simulate natural selection in introducing fewer random mutations. The correct error rate is essential in a directed evolution project because protein functionality has to be maintained to some extent. In establishing the ideal/optimum ep-PCR condition in terms of mutations introduced, 0–4.5 mutations/kb was targeted for this study. This would result in a rather moderate probability of amino acid change.
The ep-PCR products were ligated into pET29b(+) vector and transformed into *E. coli* BL21 and the clones produced were checked for protease activity.

**Figure 5-1** Error prone PCR for BSP-MprT gene (A) and (B), Lane 1 represent ep-PCR for BSP-MprT (C) Purified PCR product of BSP-MprT (D) Restriction digestion of plasmid indicating the 1638 bp insert, M: 1kb DNA ladder (NEB, USA)
5.3.3 Screening of the Error-Prone Library

A random library of BSP-MprT mutants was created under the conditions described resulting in the production of hundreds clones. Screening for protease active colonies was performed on agar plates containing 1% casein (Figure 5-2).

![Screening of error prone mutants of BSP-MprT protease on casein (1 % w/v) agar plate. Protease hydrolyses casein to free and soluble peptides. This is visible in clear halos around colonies on casein agar plates.](image)

5.3.4 Screening of Proteolytic Activity and Thermal Resistance

Regarding protease activity, the only few of the generated variants had a comparable residual activity to wild type protease from BSP strain (Wt-BSP-MprT) (±15 %) at 60°C. About 100 colonies of the consensus library were screened for improved activity following incubation for 30 min at temperature 50-90°C. The residual activity of each variant was calculated by measuring the activity before and after incubation at 80°C in order to avoid false-positive results caused by protease
overexpression. From the generate library only 15% of the transformants showed clear zone of protease activity on 1% casein plates whereas rest of the transformants were inactive.

5.3.5 Sequence Comparison of Wild Type and Mutant BSP-MprT

The nucleotide sequence of best thermostable mutant protease gene after two rounds of ep-PCR was compared with the wild type BSP-MprT. The amino acid sequence obtained in the sequence analysis of genes of best thermostable mutant protease after two rounds of ep-PCR was compared with wild type BSP-MprT by CLUSTAL W (1.83) multiple sequence alignment.

It was found that the highest heat stable variant had the single substitution of amino acid Gly at position 347 with Cys residue. This thermal resistant variant was named as Mutant-BSP-MprT and was selected for detailed characterization of thermostable protease whereas the proteolytic activity of this variant remained unchanged when compared with wt-BSP-MprT. Both the wild-type and Gly347Cys mutant enzymes were similar with respect to expression levels and yield.

5.3.6 Postulations Based on Molecular Model

The protein structures of wild type-BSP-MprT and Mutant-BSP-MprT were modeled using swiss model (Figure 5-4). The proteases exhibit compact hydrolase fold with six-stranded parallel sheets and seven helices. His 142, His 146, Glu 166 form the catalytic triad. The mutation in the best mutant protease was found away from the catalytic active site of the enzyme. From the comparison of the three-dimensional structure between wild type BSP-MprT and Mutant-BSP-MprT, it was assumed that the higher thermal stability of Mutant-BSP-MprT arises from the presence of disulfide bridges. These two enzymes are homotetramers, and their three-dimensional structures are similar. However, an intersubunit disulfide bond is formed between Cys347 and Cys518, while it did not exist in wild type BSP-MprT. Amino acid sequence alignment of both the proteases is shown in Figure 5-3.
Figure 5-3 Alignment between wild type-BSP-MprT (Wt) and Mutant-BSP-MprT (Mt) obtained through clustalW

Figure 5-4 Comparison of three dimensional structures of wild type BSP-MprT and Mutant-BSP-MprT
5.3.7 Effect of Temperature on Activity and Stability

The ability of an enzyme to resist thermal unfolding in the absence of substrates is termed as thermostability. The thermostability of the wt-BSP-MprT and mutant-BSP-MprT was measured using the residual activity of the enzyme after incubation at various temperatures ranging from 50°C – 90°C for 30 to 180 min (Figure 5-5). The native and BSP-mutant enzyme was stable at 50°C but significant changes were observed at higher temperature of incubation when kept for 3 h of incubation. After 30 min of incubation at 60°C, 70°C, 80°C and 90°C, mutant-BSP-MprT showed 1.07, 1.3, 4.3 and 3.47 folds higher activity respectively to that of wt-BSP-MprT enzyme. The residual activity of the wt-BSP-MprT enzyme was lost up to 15% and 10% at 80°C and 90°C respectively whereas mutant-BSP-MprT retained 78% and 44% activity at 80° and 90°C respectively after 30 min of incubation.

When activity of wt-BSP-MprT and mutant- BSP-MprT proteases were compared at 80°C up to 3 hrs it was observed that the residual activity of the wild type enzyme was completely lost whereas mutant- BSP-MprT protease retained activity even after 3 h of incubation. mutant- BSP-MprT protease showed 19 folds increase in residual activity as compared to that of wild enzyme. These results confirmed that mutant- BSP-MprT protease favors higher temperatures as that can be observed from the stability data shown in Figure 5-5.
Figure 5 Thermal stability profiles: (A) Residual activity vs Temperature for wild type BSP-MprT and Mutant- BSP-MprT (30 min at 50°C to 90°C), (B) Thermal stability of wild type BSP-MprT and Mutant-BSP-MprT (at 80°C for 30-120 min incubation).
Figure 5-6 Kinetics of Thermal Inactivation: (A) Effect of Thermal Inactivation of wild type BSP-apr (B) Mutant-BSP-apr assayed at various temperatures (50°C to 90°C) with time of incubation 30 to 180 min
5.3.8 Kinetics of Thermal Inactivation and Estimation of the Inactivation Energy

Inactivation is defined as process in which secondary, tertiary or quaternary structure of a protein changes without breaking covalent bonds. Inactivation rate constants ($k_d$) of native and BSP-Mutant-MprT at 50°C, 60°C, 70°C, 80°C and 90°C (Table 5.1), were calculated as per “Materials and Methods” section 5.2.13. Likewise, the half-life values estimated using these constants and Eq. (1), are presented in the same (Table 5.1). The half-life of BSP-Mutant-MprT at 50°C, 60°C, 70°C, 80°C and 90°C was 2.6, 1.6, 3.7, 7.8 and 2.7 times higher than the half-life values of native protease. Inactivation energy ($E_d$) of the BSP-Mutant-MprT and native protease was determined from the slopes of the linear curve plotted by $1/T$ versus $\ln(k_d)$ using Eq. (2) and was found to be as 121.96 kJ/mol and 108.547 kJ/mol. The higher $E_d$ value of thermal denaturation of BSP-Mutant-MprT showed that the enzyme is more stable at higher temperature. The increasing slope of the Arrhenius plot in the BSP-Mutant-apr indicated enhancement in the thermal stability of the enzyme.

<table>
<thead>
<tr>
<th>Type of Enzyme</th>
<th>Temperature (°C)</th>
<th>Inactivation rate constant ($k_d$)</th>
<th>Half life (Min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wt-BSP-MprT</td>
<td>50</td>
<td>0.00017</td>
<td>738</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.00114</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>0.0038</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>0.0114</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>0.0134</td>
<td>9.3</td>
</tr>
<tr>
<td>Mutant-BSP-MprT</td>
<td>50</td>
<td>0.00013</td>
<td>1932</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.0014</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>0.002</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>0.0029</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>0.0098</td>
<td>26</td>
</tr>
</tbody>
</table>

It is clear that the $E_d$ (Table 5.2) of thermal denaturation reaction is sensitive to changes in temperature and provides a more quantitative thermodynamic approach to monitor the thermal stabilization of proteins in different conditions. These results
showed that Mutant-BSP-MprT may be considered as a potential candidate for various industrial applications.

5.3.9 Estimation of Thermodynamic Parameters

The changes in enthalpy ($\Delta H^o$) and entropy ($\Delta S^o$) are calculated using transition state theory according to Eqs. (3) and (5) for the thermal inactivation of Mutant-BSP-MprT (Table 5.2). A positive $\Delta H^o$ and $\Delta S^o$ were determined in the temperature ranges studied. With increase of temperature, decrease in $\Delta H^o$ and $\Delta S^o$ were observed in native enzyme and as well as Mutant-BSP-MprT, whereas values of $\Delta H^o$ and $\Delta S^o$ of Mutant-BSP-MprT were higher than that of native enzyme. This suggested decrease in the surface hydrophobicity of mutant enzyme stabilizes the reaction and is a major determinant of the extent of stabilization by a protein. Therefore this indicates that the hydrophobic interactions of mutant enzyme when heated at elevated temperatures does not favor the denatured state and hence stabilizes the enzyme. With decrease in entropy decrease in unfolding of the enzyme structure and disorder of inactivation was confirmed. The decrease in $\Delta S^o$ also indicated decrease in number of protein molecules in transition activated stage, resulting in higher values of $\Delta G^o$. The numerical values of $\Delta H^o$ and $\Delta S^o$ are reported to be influenced by structural changes. Predominantly, $\Delta S^o$ values are known to provide information regarding the degree of solvation and the degree of compactness of protein molecule. Also the increase in the $\Delta H^o$ with respect to temperature reveals that the conformation of the enzyme was altered. In enzymatic reaction first step would be enzyme substrate binding, it is known to be affected by temperature which has an effect on the higher order of protein structure and their interaction with substrate. The second step of the catalytic process is the conversion of the enzyme-substrate complex into an "activated complex," a high-energy complex that can decay into product and free enzyme. The free energy of activation ($\Delta G^o$) must be added to the enzyme-substrate complex to generate the activated complex. The magnitude of $\Delta G^o$ is, in effect, the "energy barrier" to the reaction, and by significantly reducing the $\Delta G^o$ values of chemical reactions, enzymes enable metabolic reactions to occur at high rates at biological temperatures. Where as in our study we found that the values of $\Delta G^o$ of modified enzyme was found to be higher than that of native enzyme. Significant increase in the
Table 5.2 Thermal Inactivation kinetics of wt-BSP-MprT and Mutant-BSP-MprT proteases

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Inactivation energy $E_d$ (KJ/mol)</th>
<th>Temperature (°C)</th>
<th>$\Delta H^\circ$ (KJ/mol)</th>
<th>$\Delta G^\circ$ (KJ/mol)</th>
<th>$\Delta S^\circ$ (J/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wt-BSP-MprT</td>
<td>108.547</td>
<td>50</td>
<td>105.9</td>
<td>102.6</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>105.8</td>
<td>100.6</td>
<td>15.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70</td>
<td>105.7</td>
<td>100.3</td>
<td>15.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>105.6</td>
<td>100.1</td>
<td>15.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90</td>
<td>105.5</td>
<td>102.5</td>
<td>8.3</td>
</tr>
<tr>
<td>Mutant-BSP-MprT</td>
<td>121.961</td>
<td>50</td>
<td>119.3</td>
<td>103.3</td>
<td>49.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>119.3</td>
<td>100</td>
<td>57.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70</td>
<td>119.2</td>
<td>102</td>
<td>49.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>119.09</td>
<td>104</td>
<td>42.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90</td>
<td>119</td>
<td>103</td>
<td>42.8</td>
</tr>
</tbody>
</table>

$\Delta G^\circ$ values implies that it requires more free energy to form activated enzyme substrate complex which is gained by increased heat content due to increase in temperature. This indicates that the Mutant-BSP-MprT does not favor the denatured state and hence stabilizes the enzyme.

5.4 Discussion

Engineering of proteins for improved thermal stability is emerging as an effective alternative to search for proteins with similar activity in hyperthermophiles; in addition, it is the only option available if the protein of interest has no homologs in the thermophilic organisms [186]. Many studies on protein engineering of proteases have been conducted, with a focus on improving the catalytic activity or thermal stability of the enzymes [187-189].

Thermostability is one of the main requirements for commercial enzymes, because thermal inactivation represents a common problem in the application of enzyme so tailoring the thermostability of catalytic activity is one of the targets of protein engineering, especially for enzymes in industrial use.

In this study, the thermostability of *Bacillus subtilis* BSP protease was improved by two rounds of error-prone PCR mutagenesis. A random mutant library of *Bacillus subtilis* BSP protease was generated by ep-PCR. The mutant proteases obtained in this mutant library were screened for increased protease thermostability. It
was possible to identify the contribution of individual amino acid substitutions to thermostability by introducing an average of only one amino acid substitution per gene. In this study, we enhanced the thermostability from 60°C to 80°C by the single mutation Gly347Cys which has a stabilizing effect on the irreversible thermal inactivation. The half-life of the mutant-BSP-MprT was 2.6 folds than that of the wild type enzyme at 50°C. Sequence analysis and comparison of both wt-BSP-MprT and Mutant-BSP-MprT showed that wt-BSP-MprT had only one cysteine residue at Cys518 position while another cysteine has substituted Gly347 in Mutant-BSP-MprT. Introduction of another cysteine resulted in the formation of disulfide bridge between the two cysteine residues which play important role in the stability of protein tertiary structure. The intersubunit disulfide bond was estimated to be one of the factors for thermal stability [190]. Takagi et al. [174] reported the enhancement of the thermostability of subtilisin E by introduction of a disulfide bond by site directed mutagenesis. Another previous study showed the enhancement of the thermal stability of pyroglutamyl peptidase I by introduction of an intersubunit disulfide bond [190].

Disulfide bridges are believed to stabilize proteins mostly through an entropic effect, by decreasing the entropy of the protein's unfolded state. The entropic effect of the disulfide bridge increases in proportion to the logarithm of the number of residues separating the two cysteines bridged. Because of the susceptibility of cysteines and disulfide bridges to destruction at high temperatures, 100°C was believed to be the upper limit for the stability of proteins containing disulfide bridges [191].

The thermal stability of the wild-type and mutant enzymes was investigated in terms of the activity remaining after incubation at various temperatures. The mutant enzymes showed enhanced thermal stability as compared to the wild-type. The mutant-BSP-MprT retained 50% activity even at higher temperature 90°C.

A common parameter used in the characterization of enzyme stability is the half-life ($t_{1/2}$). Long half-life means that the enzyme has high stability. Half-lives during thermal denaturation were calculated at different temperatures. Mutant-BSP-MprT showed higher values of half-lives at all the temperature studied when compared to wt-BSP-MprT. At 90°C the half life of mutant-BSP-MprT was 26 min which was 2.8 folds higher than wt-BSP-MprT enzyme. A similar data reported [189]
showed that the half-life of thermal inactivation of the mutant subtilisin protease was 2-4 times longer than that of the wild-type enzyme.

The activation energy changes in the denaturation of the wild type and mutant proteases (Ed) were calculated from the natural logarithm of \( k_d \) values (ln \( k_d \)) against the reciprocal of heating temperatures (1/T). The \( E_d \) for mutant-BSP-MprT protease was determined to be higher than that of wild type enzyme. Higher \( E_d \) indicated that a higher temperature change is required to inactivate the enzyme [192]. Hernández-Martínez et al. [193] reported a thermophilic serine protease from \( Asp. fumitagrus \) had the inactivation energy of 69 kJ/mol. In our study higher values of inactivation energy suggested the protease stability.

A protein can be tolerant of thermal treatment through hydrophobic interaction, hydrogen bonds, disulphide bridges and electrostatic interactions [191]. Enthalpy (\( \Delta H \)) and entropy (\( \Delta S \)) are commonly used indicators which could describe protein stability. Lesser values of entropy (\( \Delta S \)) with increase in temperature for mutant-BSP-MprT indicated more thermal stability of enzyme.

The value of \( \Delta G^\circ \) as a function of temperature was calculated for both wild type and mutant protease. The linear relationship of \( \Delta G^\circ \) vs. temperature (\( \Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \)) allows us to calculate the thermodynamic variables, enthalpy (\( \Delta H^\circ \)), and entropy (\( \Delta S^\circ \)). Accordingly, the standard molar Gibbs free energy of activation (\( \Delta G^\circ \)) governs the thermostability. Gibbs free energy (\( \Delta G^\circ \)) includes both enthalpic and entropic contributions and could be served as a more accurate and reliable indicator for enzyme stability. Increased value of the \( \Delta G^\circ \) of mutant protease at higher temperature (50-80°C) suggested the stability of mutant protease substantially [194].

In conclusion, single amino acid mutation (Gly347) of the BSP-MprT from \( B. subtilis \) BSP lead to a systematic increase of catalytic efficiency under a wide range of temperatures. Thermodynamic analysis indicated that an increase of the free energy of activation (\( \Delta G^\circ \)) for thermal denaturation may account for the enhanced thermostability of Mutant-BSP-MprT as comparison to the wild type. With the help of the structural modeling, we can interpret that the formation of disulfide bonds between Cys347 and Cys518 residue can make considerable contributions to stability of the enzyme which may affect entropy of the denatured protein [195]. Many
attempts have been made to increase protein stability by introducing ‘intramolecular’ disulfide bonds [42, 174]. In the present study we succeeded to get the thermostable error prone mutant of the BSP-MprT by the introduction of disulfide bonds at higher temperature (2.8 folds).
6 Expression and Characterization of 
*Coprothermobacter proteolyticus* Alkaline Serine Protease

6.1 Introduction
Thermostable proteases like other enzymes from thermophilic and hyperthermophilic microorganisms (often called thermozymes) are quite compelling in some applications because they are steady and dynamic at temperatures over 60 - 70°C. Furthermore, they are more resistant than their mesophilic counterparts to organic solvents, detergents, extreme pH range and other denaturing agents [16, 196].

The utilization of proteases at higher temperatures is beneficial on the grounds that unfolded type of proteinaceous substrate is better vulnerable for the still dynamic thermostable protease. It brings about higher particular activities for proteases sourced from thermophiles and improvement of some mechanical methodologies, especially with the compounds that are dynamic at temperatures almost 100°C. The resistance against organic solvents makes the thermostable proteases valuable for the synthesis of high molecular weight peptides, completed in reaction media with low water content [19]. Performing enzyme reactions at higher temperatures take into account higher lower viscosity, substrate concentrations, reduction of microbial contamination risk and high reaction rates [20]. Extra advantage of some thermostable enzymes is diminished activity at low temperatures and this highlight empowers end of the reaction just by cooling the process [191]. In this manner the utilization of inhibitors or lavish enzyme evacuating methods after consummation of the reaction process can be evaded. Main flaws of thermozymes application are the expanded plausibility of byproducts and the degradation of thermolabile substrates or items. Currently, many heat resistant proteases have been reported but most of them are still not available in the market.
Thermophilic and hyperthermophilic microorganisms have been studied during the recent years fulfill with the aim to meet the demand on different proteolytic enzymes that have optimum pH (9.0-12.0 and temperatures ranged from 45-110°C. At present, thermophilic strains belonging to the genus *Bacillus* are the dominant origin for the production of industrially important thermostable proteases [144, 197-198]. Indeed, much ongoing effort is directed at the isolation of alkaline serine proteases from *Bacillus* sp. [199-202]. One example of well characterized heat resistant proteases are subtilisins of *Bacillus* origin. They exhibit broad substrate specificity and have similar properties such as optimal temperature of 60°C and an optimal pH of 10 [11]. However, there is great value in the characterization of proteases from other bacteria to obtain enzymes with diverse activity profiles. *Coprothermobacter proteolyticus* is an anaerobic bacterium that was isolated from a thermophilic digester fed with tannery waste and cattle manure [12, 203]. *C. proteolyticus* has an elevated optimum growth temperature of 63°C and secretes high levels of protease activity. It ferments protein more readily than casamino acids into acetate, hydrogen, and carbon dioxide. *C. proteolyticus* has recently been used in conjunction with a methanogen in the syntrophic degradation of proteinaceous substrates to produce methane [204].

At present, just couples of thermostable proteases are commercially accessible. One of them is alcalase confined from *Bacillus licheniformis*. The significant element of this arrangement is subtilisin, which is an endoprotease of serine sort, displaying highest activity at 60°C and pH of 8.3. Alcalase established numerous applications in the food industry; it is utilized as part of protein-braced soda pops and dietetic nourishment [16].

Capricious development states of numerous hyperthermophiles, moderately low cell and enzyme yields, and once in a while era of poisonous and destructive metabolites cause innovative challenges amid substantial scale protein generation [205]. Moreover, the composition of the growth media and cultivation conditions controlling protease production in hyperthermophiles are still not fully recognized. These problems can be eliminated by expression of the genes responsible for synthesis of the desired enzyme in mesophilic host.
We report the first expression and characterization of a recombinant *C. proteolyticus* protease. The enzyme is demonstrated to be a serine protease with an alkaline pH optimum (8–10) and functions at an elevated temperature (60°C). The protease also has the desirable property of retaining high activity in the presence of a wide variety of surfactants.

### 6.2 Material and Methods

#### 6.2.1 Bacterial Strains, Plasmid, and Reagents

The strains used were *Escherichia coli* JM109 (Promega; WI, USA) and protease deficient *Bacillus subtilis* 1A751 [206] from the *Bacillus* Genetic Stock Center (BGSC; OH, USA). The plasmid used for *B. subtilis* expression was pDR111 (a gift from David Rudner). Bacteria were propagated in Luria-Bertani (LB) broth (*E. coli* and *B. subtilis*) or tryptose blood agar base (TBAB) (*B. subtilis*) media at 37°C. All chemicals were purchased from Sigma-Aldrich (MO, SA) unless otherwise specified.

**Table 6.1 Primer Sequences used in this study**

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5’→3’)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copro-F</td>
<td>CGCCATAATGAAAAAGATC TTG TTGACCCCTGGTTATT</td>
<td>Forward primer for <em>Copro-Apr</em> gene cloning in pET29b(+). <em>NdeI</em> site is underlined</td>
</tr>
<tr>
<td>Copro-R</td>
<td>CGACTCGAGTTAAGGCCTG CCA ATTCACCTGCTG</td>
<td>Reverse primer for <em>Copro-Apr</em> gene cloning in pET29b(+). <em>XhoI</em> site is underlined</td>
</tr>
<tr>
<td>Copro-5-H</td>
<td>CGCAAGCTTACATAAGGA GGAACCTACTATGAAAAAG ATCTTGTGGACCCTGGTG GTTATT</td>
<td>Forward primer for <em>Copro-Apr</em> gene cloning in pDR111. <em>HindIII</em> site is underlined</td>
</tr>
<tr>
<td>P29-26-sp</td>
<td>GCGGGCATGCTCACAAGAAA ACCCTCAAGACCCGTTA G</td>
<td>Reverse primer for <em>Copro-Apr</em> gene cloning in pDR111. <em>SphI</em> site is underlined</td>
</tr>
</tbody>
</table>
6.2.2 Vector Construction and Gene Expression in *E. coli*

The *C. proteolyticus* aprE was resynthesized for optimal codon usage by *E. coli* (DNA 2.0; CA, USA). The 1323 bp gene was amplified and ligated into pET29b(+) at *Nde*1 and *Xho*1 as described in chapter 2 section 2.11.13. Plasmid pET-copro-Apr was transformed in either in *E. coli* or used as a template in an *in vitro* transcription/translation reaction (*E. coli* T7 S30 Extract System; Promega). *Escherichia coli* JM109 was employed in this study for subcloning and plasmid recovery. It was grown in Luria broth (LB) and *E. coli* transformants were grown LB agar medium with 30 µg/mL of kanamycin. *E. coli* BL21 (DE3) was used as the expression host and the transformants were again screened on LB agar containing 30 µg/mL of kanamycin. Expression was induced by IPTG as described in chapter 2 section 2.11.17.

6.2.3 *In Vitro* Transcription/Translation

Copro-aprE was expressed in coupled in vitro transcription/translation systems using The S30 T7 High-Yield Protein Expression System (Promega, WI, USA). It is an *E. coli* extract-based cell-free protein synthesis system which simplifies the transcription and translation of DNA sequences cloned in plasmid or lambda vectors containing a T7 promoter by providing an extract that contains T7 RNA polymerase for transcription and all necessary components for translation (Figure 6.1). By using this system high levels of recombinant proteins can be obtained (up to hundreds of micrograms of recombinant protein per milliliter of reaction) within an hour. We supplied only cloned pET-copro-aprE DNA containing a T7 promoter and a ribosome-binding site (RBS) as shown in table 6.2.

Reaction mixture was mixed thoroughly by pipetting several then centrifuged in a microcentrifuge for 5 seconds to force the reaction mixture to the bottom of the tube. Reaction was incubated at 37°C with vigorous shaking for 1 h and after that reaction was stopped by placing the tubes in an ice-water bath for 5 min. The results of the reaction were analyzed on SDS-PAGE as described in chapter 2 section 2.9.
Table 6.2 Coupled Transcription/Translation Procedure

<table>
<thead>
<tr>
<th>Component</th>
<th>Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid DNA (pET-copro-apr) template</td>
<td>1 µg</td>
</tr>
<tr>
<td>S30 Premix Plus</td>
<td>18 µL</td>
</tr>
<tr>
<td>T7 S30 Extract, Circular</td>
<td>20 µL</td>
</tr>
<tr>
<td>Nuclease-Free Water to a final volume of</td>
<td>50 µL</td>
</tr>
</tbody>
</table>

Figure 6-1 Schematic diagram showing the process of in vitro transcription/translation process

6.2.4 Expression of Copro-Apr Gene in Bacillus subtilis Expression System

For expression of copro-Apr gene in Bacillus subtilis expression system the gene was amplified by PCR using Copro-5-H and P29-26-sp primers (primer sequences in table 6-1) and subcloned into the pDR111 vector at HindIII and SphI restriction enzyme sites that were engineered into the 5’ and 3´ends of the gene, respectively, to create the expression plasmid pDR111-copro-apr in E. coli. The plasmids pDR111a (vector control) and
pDR111-copro-apr were transformed into *B. subtilis* and integrated into the bacterial chromosome using standard protocols found on the BGSC website (http://www.bgsc.org/catalogs/Catpart4.pdf). Protein was expressed by inoculating a fresh liquid culture (OD600 = 0.5) with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and growing the culture at 225 rpm at 37°C for 24 h.

### 6.2.5 Protease Assay-Solid Phase

*B. subtilis* colonies transformed with either pDR111 (vector control) or pDR111-copro-apr expression plasmid were spotted onto LB agar plates containing 1% casein, 1 mM IPTG, and 100 µg/mL spectinomycin antibiotic. The plates were incubated overnight at 37°C and then transferred to a higher temperature (70°C) at which the protease was active.

### 6.2.6 Protease Assay Liquid

The proteolytic activity was determined using casein as the substrate as described in chapter 2 section 2.5.

### 6.2.7 Determination of pH Optimum

The protease was preincubated at pH range 5–12 at 60°C for 30 min. The reaction was then initiated by adding an equal volume of 1% casein at the same pH as the enzyme. Time points were collected at 0 and 20 min, and the reactions were stopped with TCA and assayed at 275 nm as described above. 50 mM sodium succinate buffer (pH 5-6) and 50 mM sodium phosphate (pH 7.0–12.0) were used for the experiments.

### 6.2.8 Determination of Optimum Temperature

The enzyme and casein substrate were pre-incubated separately at different temperatures (50–90°C) at pH 9 for 30 min. Then, the enzyme and substrate were combined, and the
reaction proceeded at the pre-incubation temperature. Time points were collected at 0 and 20 min, and the level of proteolysis was assayed as described above.

### 6.2.9 Determination of Thermal Stability

The protease was heated at different temperatures (50–90°C) at pH 9 for 0, 30, and 60 min. After the heat challenge, the enzyme was pre-incubated for 30 min at 60°C, and then the reaction was initiated by the addition of 1% casein substrate. The reaction proceeded at 60°C and was assayed as described above.

### 6.2.10 Determination of Additives Effect

Various metal cations (Ca^{2+}, calcium chloride; Co^{2+}, cobalt chloride; Fe^{2+}, ferrous chloride; Mg^{2+}, magnesium chloride; Mn^{2+}, manganese sulfate; Ni^{2+}, nickel chloride; and Zn^{2+}, zinc sulfate) were pre-incubated with the protease for 30 min at 60°C. Various protease inhibitors β-Mercaptoethanol β-ME; dithiothreitol, DTT; ethylenediaminetetraacetic acid, EDTA; iodoacetate, IAA; and phenylmethylsulfonyl fluoride, PMSF) were added to the protease for 30 min at 37°C and then pre-incubated at 60°C for 30 min. Surfactants (sodium dodecyl sulfate, Triton X-100, and Tween-20) and oxidant (H_2O_2) were added to the protease for 60 min at 37°C and then pre-incubated at 60°C for 30 min. All additives were used at a final concentration of 5 mM. All incubations occurred at pH 9. After pre-incubations, casein was added to all the enzymes, the reactions proceeded at 60°C, and residual protease activities were assayed as described above.

### 6.3 Results

#### 6.3.1 Gene Cloning and Expression

A putative protease gene (aprE) from the thermophilic *C. proteolyticus* DSM 5265 was chosen as the target of study. The gene sequence (NCBI NC 011295) was resynthesized to conduct codon optimization for *E. coli* expression (Figure 6-2). The
The amino acid sequence of the protein has the predicted conserved catalytic triad residues of known serine proteases [207]. When the sequence was compared to the NCBI database by BLAST analysis, a peptidase from *Caldisericum exile* AZM16c01 (NCBI YP 005473527.1) was the closest match with 74% identity [208].

Figure 6-2 aprE resynthesized gene sequence and translation. Highlighted residues indicate the predicted catalytic triad of known serine proteases.
Figure 6-3 Cloning of Copro-aprE into pET29b(+). (A) Agarose gel showing amplified copro-aprE gene (B) Purified insert and vector after restriction digestion (C) Colony PCR of selected 30 colonies (D) Confirmation of ligated insert release from the pET29b vector

Th aprE gene was cloned into an E. coli expression vector (pET29b+; Novagen, WI, USA). This plasmid was either transformed into an E. coli expression strain (BL21 (DE3) pLysE; Novagen) (Figure 6.3) or used as a template in an in vitro transcription/translation reaction (E. coli T7 S30 Extract System; Promega). Only the in vitro transcription/translation reaction yielded recombinant enzyme, and no protease activity was detected (Figure 6-4).
Therefore, the gene was subcloned into a *B. subtilis* expression vector downstream of an IPTG inducible promoter (Figure 6-5). The expression plasmid was transformed into a protease-deficient *B. subtilis* strain, and an active recombinant enzyme was secreted from the cells (Figure 6-6). *B. subtilis* transformed with the expression vector produced high levels of enzyme activity in liquid culture medium 24 hours after IPTG induction, whereas control *B. subtilis* carrying only vector DNA resulted in no significant protease activity (data not shown).
Figure 6-5 Cloning of Copro-aprE into pDR111 (A) Amplification of copro-aprE, Lane 1 represent PCR for Copro-aprE. (B) Restricted digestion of Copro-aprE PCR product with HindIII and SphI. (C) Restricted digestion of pDR111 with HindIII and SphI (D) Double digested plasmid pDR-Copro-aprE.. M: 1kb DNA ladder (NEB, USA).
Figure 6-6 *B. subtilis* protease expression. Colonies transformed with either vector control (left or *aprE* expression construct (right) were spotted onto LB agar plates containing 1% casein and 1 mM IPTG

### 6.3.2 Confirmation of Cloning by PCR using Genomic DNA of 1A751/pDR-Copro-aprE as Template

The genomic DNA of 1A751/pDR-Copro-aprE was isolated from overnight 1mL culture of the strain by using FastDNA kit from Bio101 following the protocol as described in chapter 2 section 2.11.2. The clone was further confirmed by PCR using using genomic DNA of 1A751/pDR-Copro-aprE as template and forward and reverse primers designed for protease gene. PCR conditions and primers were same used for PCR amplification of copro protease gene. PCR product was run on 0.9% gel gel along with 1Kb NEB DNA ladder to estimate size of amplicon. Figure 6-7 showed the amplification of 1323 bp gene fragment of protease confirming the integration of gene in *Bacillus subtilis* 1A751.
6.3.3 pH and Temperature Optima

The activity of the recombinant protease was tested under a variety of pH values. The enzyme had the highest activity from pH 8 to pH 10 (Figure 6-8). At pH 7 and pH 11, the protease retained 70% of the activity. The temperature optimum was determined to be 60°C (Figure 6-9). This is consistent with the optimal growth temperature (63°C) of the source organism. When the reaction was conducted at 70°C, the activity decreased to 80%. Under optimal conditions (pH 9 and 60°C), there was 66 U/mL of activity. The thermostability of the protease was tested at various temperatures and times (Figure 6-10). At 50°C, there is a slight drop in activity after 60 min. At 60°C and 70°C, there is a larger decrease to approximately 70% after 60 min. At 80°C, half the activity is lost at 60 min. Finally, at 90°C, only 30% of the activity remains after 30 min.
Figure 6-8 Activity of protease at various pH and 60°C

Figure 6-9 Activity of protease at various temperatures and pH 9
Figure 6-10 Activity of protease after heat treatment. Residual activity was assayed at 60°C and pH 9

6.3.4 Effect of Additives on Protease Activity

The effects of numerous divalent cations on protease activity were tested (Figure 6-11). Fe$^{2+}$ greatly stimulated the protease activity to 248% of the unsupplemented enzyme level. Ca$^{2+}$ and Co$^{2+}$ both stimulated enzyme activity to approximately 150%. Cu$^{2+}$, Mg$^{2+}$, Mn$^{2+}$, and Ni$^{2+}$ decreased enzyme activity to 49–83%, while Zn$^{2+}$ had little effect. Addition of PMSF dramatically reduced enzyme activity, thus supporting the identity of this enzyme as a serine protease (Figure 6-12). EDTA reduced activity to 40% which could be a reflection of the cation influence on protease activity. The β-ME and DTT reducing agents both induced moderate inhibition, while IAA had no effect.

The protease was highly stable in a variety of surfactants and oxidant (Figure 6-13). All these additives resulted in either no effect or increased activity. SDS induced the greatest stimulation at 146% relative to unsupplemented enzyme.
Figure 6-11 Activity of protease after incubation with various metal cations

Figure 6-12 Activity of protease after incubation with various inhibitors
Figure 6-13 Activity of protease after incubation with various surfactants and oxidants

6.4 Discussion

Enzymes from hyperthermophiles offer extreme (thermo) stability and a number of enzymes have been studied to explore their potential for various biotechnological processes [209]. The abundance of microbial genome succession data permits the investigation of extremophiles for qualities encoding proteins with expected high thermostability. In this work, we report the expression in *Bacillus subtilis* of a novel thermostable serine protease identified by genome mining. We also describe the biochemical characterization of alkaline serine protease, one of the two putative extracellular proteases in the genome of the extreme thermophile *C. proteolyticus* [210]. *C. proteolyticus* was isolated from a thermophilic digester (55°C) fermenting tannery waste and cattle manure [211], and it assumes an essential part in biogas production by discharging hydrogen that is utilized by methanogenic organisms.

This is the first report of the biochemical characterization of a recombinant protease from the thermophilic *C. proteolyticus*. The enzyme was demonstrated to be an alkaline serine
protease that was active at elevated temperatures and resistant to many surfactants, thus indicating potential utility of this enzyme in detergent applications. Alkaline serine protease aprE is active at temperatures very close to optimum growth temperature of its *C. proteolyticus* host (63°C) and it can hydrolyze the protein casein at temperatures as high as 90°C. At elevated temperature (70 to 80°C), the enzyme is stable and displays a high catalytic activity. Thermostability is regularly connected with high resistance to organic solvents [212]. Alkaline serine protease indeed tolerated organic solvents in hydrolytic reactions efficiently.

In addition, the new protein sequence of this enzyme will be of great value in the continued efforts to develop protease activity improvements [213-214]. One of the key strategies of enzyme engineering is directed evolution through DNA shuffling between different family members [215-216]. The availability of more unique amino acid sequences from proteases whose activities have been experimentally validated is critical to such projects.

In conclusion, the convenient production in *Bacillus subtilis*, the high thermostability, the broad pH range, and its high tolerance to surfactants and organic solvents make this enzyme an attractive candidate for proteolysis under demanding reaction conditions
7 General Discussion

7.1 Aim of Thesis

Proteases represent one of the largest groups of the most important industrial enzymes, accounting for nearly 65% of aggregate overall enzyme deals [11]. Proteases from bacterial sources are the most significant in comparison with animal and fungal proteases and among various bacterial microorganisms, *Bacillus* strains are most significant producers of commercially applicable proteases. It was accounted for that around 35% of the total microbial enzymes utilized as a part of detergent industry are the proteases from microorganisms sources [48]. These proteases are widely used in pharmaceutical, laundry, leather, food and waste processing industries. Among proteases, studies on thermostable proteases from thermophilic microorganisms have been paid more attention with the expectation to produce thermostable enzymes [217]. Thermostable proteases are favorable in some industrial applications in light of the fact that at higher processing temperatures the faster reaction rates is possible along with the increased in solvency of nongaseous reactants and products and reduced chances of microbial contamination from mesophilic organisms [15]. Thus it is desirable to search for new thermostable proteases with novel properties from as many different sources as possible. These enzymes can be produced from the thermophiles through either optimized fermentation of the micro-organisms or using recombinant DNA technology to express the thermostable protease genes in the fast growing mesophilic hosts [218]. Another approach to obtain enzymes of increased thermal is protein engineering or searching for homologs in hyperthermophiles. In fact, thermostability enhancement is one of the most important aims of protein [174]. For this purpose, several strategies have been adopted: e.g. introduction of hydrophobic interaction or disulfide bond(s) into protein molecules by means of site-directed or random mutagenesis [186]. The more commonly used mutagenesis technique to increase stability are the *In vitro* directed evolution strategies. In this technique usually the genes encoding non-thermostable enzymes with desired
activities are used for development of better thermostability, and the temperature of the screening assay is used as selection pressure.

The present thesis describes the optimization of the conditions for the production of thermostable protease (thermolysin) from the thermophilic Bacillus subtilis BSP strain. The work described was primarily focused on understanding and improving the thermal stability of these enzymes, and extremely stable proteases (Thermolysin and serine) were obtained. Thermolysin and serine thermostable proteases genes were used for molecular characterization studies. Thermolysin metalloprotease from thermophilic Bacillus subtilis BSP strain was cloned, expressed of into E. coli and Bacillus subtilis expression systems. Serine protease gene from Coprothermobacter proteolyticus was expressed into Bacillus subtilis expression system. Directed evolution studies were done to increase the thermostability.

The major part of the thesis concerns experiments aimed at (1) understanding the factors governing protease stability and (2) engineering highly stable BSP-Mpr variant. The results presented provide a basis for the development of stability profile of thermolysin and serine.

7.2 Main Findings

7.2.1 Thermophilic Bacillus subtilis BSP Strain is the Potential Candidates for the Production of Extracellular Thermostable Protease

This part of study was an attempt to enhance protease production through optimization of nutritional/physical parameters by conventional (one-variable-at-a-time) and statistical approach (RSM) at shake flask level. Further, the effect of substrate concentration, inoculum density, incubation time and initial media pH on production of a thermostable protease from Bacillus subtilis BSP in bench-scale fermenter (under optimized nutritional and cultural parameters) was also envisaged in order to produce thermostable protease with maximum activity from this strain. Response surface methodology was found to be a
useful tool to elucidate the different culture conditions as it enabled us to study four variables at a time with changing two variables and keeping one constant thus developing a polynomial equation. This equation can predict optimum production of protease secreted in culture media for different levels of variables and can also be helpful for the future researchers and scientists to predict amount of enzyme produced. RSM is a three-dimensional (3-D) graphical representation of the data which looks beautiful enough to capture the attention of viewers and develop an immediate understanding of the trend of the enzyme response (Enzyme Unit in this case). Another significant aspect of this strategy is that it shifted the trend of laborious and tedious one-factor-at-a-time approach towards the study of multiple factors at a time. In this way it is a time effective approach with maximum understanding of interaction of multiple variants on the response. Moreover, the characteristic properties and stability profile of the thermostable protease of Bacillus subtilis BSP strain were also studied. EDTA inhibited the enzyme activity, indicating it to be a metalloprotease. Some of the novel features of this metalloprotease enzymes, such as stability over the wide range of pH, thermo stability of enzyme at higher temperatures, stability of enzyme in presence of high concentrations of various organic solvents, surfactants and oxidants make them attractive candidates for future studies.

7.2.2 Bacillus subtilis Expression System Efficiently Secretes Proteases in the Culture Media which has Higher Protease Activity than Source Microorganism

Cloning of thermostable protease gene from the genomic DNA Bacillus subtilis BSP strain was carried out in an over-expression vector, pET29b+ (Novagen) and an integration vector pDR111. Vector constructs were transformed in Escherichia coli strain BL21 and Bacillus subtilis 1A751. With respect to over-expression of protease genes the optimum production of recombinant proteases was obtained with 1mM IPTG. The molecular weights of the protease was estimated 36 kDa, which were quite comparable to thermolysin produced in native organisms. High level of purified enzymes obtained as
evident from the specific activity and yield. Enzymes were characterized for physico-chemical properties, as for their native counterparts and recombinant enzyme was found stable under native conditions. Molecular characterization of recombinant enzymes; distribution of amino acids, 3D structure analysis and protein secondary structure analysis was carried out. It contains a catalytically important zinc ion located in the active site cleft positioned between an amino-terminal domain that consists mainly of beta strands and a carboxy-terminal domain that is mainly α-helical that confirmed the metalloprotease nature of the enzyme.

7.2.3 Directed Evolution is a useful Approach to Increase Enzyme Thermostability

The discovery of novel enzymes with desired properties from a microbial source takes anywhere between months to years, and therefore, there is an increasing interest in modifying the existing microbial enzymes to meet the present day needs of the industry. The redesigning of industrially useful enzymes for improving their performance enzymes has become a trend for fine tuning of biocatalysts in the biotech industry. But it has been a challenge because bioinformatics databases reveal new facts on a day-to-day basis. Many rational design, semi-rational approaches and directed evolution approaches are being used for designing and constructing novel proteins. Prior knowledge of the sequence, 3D-structure and the knowledge of the structure-function aspects of the target proteins is required for the successful implementation of the rational methods. So far, 95,052 protein structures have been revealed that are available in the protein data bank (PDB) [13]. Unlike the rational design methodology, directed evolution is advantageous because it does not need prior knowledge of the primary sequence, function or structure of proteins. This method mimics natural selection in vitro and reduces the time required for evolution from millions of years to few weeks or months.

The goal of this work was to shift the thermal resistance of the *Bacillus subtilis* BSP strain protease (BSP-MprT) in order to make thermolysin proteases BSP-Mpr
applicable for industrial applications. We developed a directed evolution strategy to reengineer the BSP-MprT in order to adapt the thermal resistance.

The thermostable mutant proteases of Bacillus subtilis BSP strain were isolated following two rounds of directed evolution using error-prone PCR. The best mutant protease obtained after first and second round of error-prone PCR was characterized. The thermal stability of the wild-type and mutant enzymes was investigated in terms of the activity remaining after incubation at various temperatures. The mutant enzymes showed enhanced thermal stability and retained catalytic activity compared to the wild-type. The best mutant protease (Mutant-BSP-MprT) showed 5.2-fold increase in percentage thermal stability (% remaining activity after incubation of enzyme at 80°C for 30 min) than wild-type protease. We also investigated the reaction kinetics of the proteases. High values of inactivation energy ($E_a$), $\Delta H^\circ$, $\Delta S^\circ$ and $\Delta G^\circ$ favored the highly thermostable nature of mutant protease when compared to wild type protease.

Nucleotide and amino acid sequence analysis of mutant-BSP-MprT indicated the substitution of one amino acid (Glycine) by cysteine. Wild type protease gene contained only one cysteine while in mutant type of protease another cysteine was introduced by random mutagenesis. Intramolecular interactions such as hydrogen bonds, electrostatic interactions and disulfide bridges are well known as the dominant structural factors responsible for protein thermostability [219]. The data presented here showed that stabilization of the B. subtilis BSP protease is possible by the introduction of disulfide bond between the two cysteine residues within the molecule which allows the enzyme to be active at elevated temperatures.

Enzyme thermostability has been enhanced by protein engineering but it is usually not clear which amino acid substitutions should be made. The idea of stabilizing protein structure using non-native disulfide bridges was first explored by Villafranca et al. [220] with the enzyme dihydrofolate reductase (DHFR). Since then, thermal stability of many other proteins has been increased in enzyme variants by the introduction of disulfide bridges [42, 221-223].
In conclusion the disulfide mutant showed substantially improved thermostability with 7.8-fold increase in the $t_{1/2}$ value at 80°C compared with the parent enzyme. The investigation of enzymatic properties of the disulfide variant revealed that the additional disulfide bond did not interfere with the catalytic efficiency. This is an important parameter, as it can be used directly to calculate the economic feasibility of a biocatalytic process.

7.2.4 A Novel Highly Thermostable and Detergent/Organic Solvent Stable Alkaline Protease from Thermophilic Bacterium *Coprothermobacter proteolyticus*

Among proteases the largest group of commercial proteolytic enzymes are the serine proteases (subtilisins) [224] that account for more than half of the world total sales of enzymes [225] being extensively used in textile, food, pharmaceutical, detergent, and leather industries [15, 125, 226-228]. A few thermostable subtilisin homologs (subtilases) have been isolated and characterized e.g., pyrolysin [229], stetterlysin [230], pernisine [231], Tk-SP [232], Tk-1689, and Tksubtilisin [233-234] and from thermophilic bacteria fervidolysin [235], Ak.1 protease [236], Rt41A proteinase [237-238], aqualysin I [239] and islandisin [240]. Since cultivation of extremophiles is associated with potential difficulties, cloning and expression of protease genes into a mesophilic host that is easy to grow (*Escherichia coli*, *Bacillus subtilis*, and yeast) is of importance for biochemical investigation, protein engineering studies and practical enzyme production. Another possible advantage of using a mesophilic host is that enzyme isolation can be based on the difference in thermostability between host proteins and the target thermozyme [241].

In order to discover novel thermostable serine proteases that resist harsh process conditions, we explored available genome sequences from thermophiles, hyperthermophiles, and bioremediation organisms for homologs of subtilisin E. Through genome mining, we identified a gene encoding a putative serine protease in the thermophilic bacterium *Coprothermobacter proteolyticus*.
The gene was functionally expressed in *Bacillus subtilis* and characterized. Alkaline serine protease has a broad pH tolerance and is active at temperatures of up to 80°C. In addition, the enzyme shows good activity and stability in the presence of organic solvents, detergents, and oxidants. Based on its stability and activity profile, this enzyme can be an excellent candidate for applications where resistance to harsh process conditions is required.

### 7.3 Future Recommendations Specific to this Thesis

- Novel protein engineering techniques of directed evolution like error prone PCR and gene shuffling will help to improve the existing enzyme properties and yield.
- The generated knowledge about the structure-function relationship of improved variants of proteases promises a fair chance of success in the near future, in evolving proteases that were never made in nature.
- Production of the enzyme in large scale with the developed strain which would meet the requirements of the multitude of protease applications to expand the commercial protease markets.
- Bioinformatic and crystallographic studies will explore the more information of these novel thermostable proteases for their commercial exploitation.