URIFICATION AND CHARACTERIZATION OF PROTEASE(S) FROM BACILLUS SPECIES

MUSTAFA KAMAL

DEPARTMENT OF MICROBIOLOGY
UNIVERSITY OF KARACHI
1993
In the Name of ALLAH
the Beneficent, the Merciful
PURIFICATION AND CHARACTERIZATION OF PROTEASE(S) FROM BACILLUS SPECIES

BY

MUSTAFA KAMAL

THESIS SUBMITTED TO THE UNIVERSITY OF KARACHI IN FULFILLMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MICROBIOLOGY UNIVERSITY OF KARACHI (SEPTEMBER 1993)
Dedicated to my Parents
for their love and affections
CONTENTS

ACKNOWLEDGEMENTS I
SUMMARY II
LIST OF FIGURES IV
LIST OF TABLES VII
LIST OF ABBREVIATIONS VIII

1-0 INTRODUCTION
1-1 General introduction 1
1-2 Classification of proteases 2
1-3 Serine proteases 3
1-4 Subtilisin family 4
1-4-1 Subtilisin BPN' 6
1-4-2 Subtilisin Carlsberg 7
1-4-3 Subtilisin Novo 7
1-4-4 Subtilisin Amylosacchariticus 8
1-4-5 Subtilisin DY 8
1-4-6 Alkaline mesentericopeptidase 9
1-4-7 Glu/Asp-specific protease from Bacillus licheniformis 9
1-5 Site-directed mutagenesis of subtilisins 9
1-6 Substrate specificity of subtilisins 11
1-7 Specific proteases from other microorganisms 11
1-8 Metalloproteases 12
1-9 Protease production and spore formation 13
1-10 Cloning of genes encoding proteases 15
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-10-1</td>
<td>Cloning of subtilisin gene</td>
<td>15</td>
</tr>
<tr>
<td>1-10-2</td>
<td>Cloning of neutral protease</td>
<td>16</td>
</tr>
<tr>
<td>1-10-3</td>
<td>Neutral protease as a cloning and expression marker</td>
<td>17</td>
</tr>
<tr>
<td>1-11</td>
<td>Culture medium for the production of proteases</td>
<td>18</td>
</tr>
<tr>
<td>1-12</td>
<td>Industrial applications of proteases</td>
<td>19</td>
</tr>
<tr>
<td>2-0</td>
<td>MATERIALS AND METHODS</td>
<td></td>
</tr>
<tr>
<td>2-1</td>
<td>Isolation of microorganisms</td>
<td>21</td>
</tr>
<tr>
<td>2-2</td>
<td>Strain identification</td>
<td>21</td>
</tr>
<tr>
<td>2-3</td>
<td>Maintenance of bacteria</td>
<td>21</td>
</tr>
<tr>
<td>2-4</td>
<td>Formulation of medium for efficient and cost effective production of proteolytic enzymes</td>
<td>22</td>
</tr>
<tr>
<td>2-5</td>
<td>Determination of proteolytic activity in Lohlein Vollhard (L.V.) units</td>
<td>27</td>
</tr>
<tr>
<td>2-6</td>
<td>Cultivation of strain RT-5 for the production of subtilisin</td>
<td>28</td>
</tr>
<tr>
<td>2-7</td>
<td>Isolation of serine protease from strain RT-5</td>
<td>28</td>
</tr>
<tr>
<td>2-7-1</td>
<td>Ammonium sulphate fractionation</td>
<td>28</td>
</tr>
<tr>
<td>2-7-2</td>
<td>Desalting of precipitate</td>
<td>29</td>
</tr>
<tr>
<td>2-7-3</td>
<td>Fractionation of protein on DEAE-Sepharose</td>
<td>29</td>
</tr>
<tr>
<td>2-7-4</td>
<td>Chromatography on Mono-S FPLC</td>
<td>29</td>
</tr>
<tr>
<td>2-7-5</td>
<td>Rechromatography of the active fraction</td>
<td>29</td>
</tr>
<tr>
<td>2-8</td>
<td>SDS-polyacrylamide gel electrophoresis</td>
<td>30</td>
</tr>
<tr>
<td>2-9</td>
<td>Chemical cleavage of subtilisin RT-5</td>
<td>31</td>
</tr>
<tr>
<td>2-9-1</td>
<td>Separation of cyanogen bromide digested peptides</td>
<td>31</td>
</tr>
<tr>
<td>2-10</td>
<td>Enzymatic cleavage of CNBr digested peptides of subtilisin RT-5</td>
<td>31</td>
</tr>
<tr>
<td>2-11</td>
<td>Purification of peptides</td>
<td>31</td>
</tr>
</tbody>
</table>
2-12 Amino acid analysis
2-13 Amino acid sequence analysis
2-14 Carboxymethylation of alcohol dehydrogenase
2-15 Substrate specificity of subtilisin RT-5
2-16 Assay of enzymatic activity of subtilisin RT-5
2-17 Enzyme inhibition assay
2-18 Measurement of $k_{\text{cat}}/K_m$
2-19 pH dependence of $k_{\text{cat}}/K_m$
2-20 Thermostability of subtilisin RT-5
2-21 Cloning strategy for PCR-amplification & sequencing of the subtilisin RT-5 gene
2-21.1 Isolation of chromosomal DNA of B. subtilis RT-5
2-21.2 Phenol: Chloroform extraction
2-21.3 RNA elimination by Ribonuclease digestions
2-21.4 Removal of Phenol
2-21.5 Precipitation of DNA
2-22 Oligonucleotides used for PCR-amplification
2-23 Conditions for PCR-amplifications
2-24 Cloning of PCR-amplified subtilisin RT-5 gene
2-25 DNA sequence analysis
3-0 RESULTS
3-1 Initial screening of soil samples for proteolytic bacterial strains
3-2 Strain identification
3-3 Comparison of specific growth rates of Bacillus subtilis strain GBS01 on various media
3-4 Purification of proteases
3-4-1 Ammonium sulphate fractionation of cell free broth 42
3-5 Separation and characterization of proteases 42
3-6 Purification of alkaline protease (Subtilisin RT-5) 42
3-7 Estimation of molecular mass 43
3-8 N-terminal sequence analysis 43
3-9 Chemical cleavage of subtilisin RT-5 43
3-10 Enzymatic fragmentation and purification of peptides 43
3-11 Amino acid analysis 44
3-12 Primary structure of the subtilisin RT-5 44
3-13 Enzyme inhibition assay 44
3-14 Measurement of $k_{\text{cat}}/K_m$ 45
3-15 Thermostability of subtilisin RT-5 45
3-16 PCR-cloning and DNA sequence analysis 45
4-0 DISCUSSION 78
4-1 Preliminary characterization of subtilisin RT-5 from Bacillus subtilis strain RT-5 81
4-2 Enzymatic properties of the subtilisin RT-5 82
4-3 Primary structure of the subtilisin RT-5 83
4-4 Comparison of the subtilisin RT-5 with other subtilisins 84
4-4-1 Comparison of the subtilisin RT-5 and subtilisin Novo 86
4-4-2 Comparison of subtilisins and pancreatic proteases 86
4-5 Presence of repeated sequences 87
4-6 Structure-function relationship 88
4-7 Prediction of three-dimensional structure 90
4-8 Thermostability of subtilisin 91
4-9 Cloning of the subtilisin RT-5 gene 92
ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to Professor Tashmeen Razzaki, my supervisor, for her excellent guidance, continuous support, encouragement and for sharing her knowledge of molecular biology and bacterial physiology. She always gave me valuable suggestions whenever I asked.

I am grateful to Professor Zafar H. Zaidi, HEJ Research Institute of Chemistry, University of Karachi, for introducing me to protein chemistry research specially in the protease business. His support and generous encouragement made this work a great experience for me. I am also thankful for all the lab facilities and financial support that he has provided to complete this study.

Many thanks are due to Professor Hans Jornvall and Dr. Jan Olov Hoog, Karolinska Institute, Sweden for their guidance and many fruitful suggestions on how to proceed with the work and also for the excellent working facilities that they have provided during my stay in Stockholm.

My particular debts of gratitude and humble thanks are due to my parents, brother and sisters. Without their love, encouragement, unflinching care, patience and sacrifices this work would have never been completed. I hope that they will overlook the negligence of my duties toward them in the process of continuing my research.

My special thanks are due to Dr. Atiya Abbasi, HEJ Research Institute of Chemistry, University of Karachi, for teaching me the basic instrumentation during my infancy in the protein chemistry lab and since then helping me whenever I needed her.

Swedish Institute, Stockholm is gratefully acknowledged for their scholarship to support my studies in the laboratory of Professor Hans Jornvall.

I would like to thank my colleagues and friends at the HEJ Research Institute and Karolinska Institute, Stockholm for their help and nice company during this study.

MUSTAFA KAMAL
SUMMARY

The present study was undertaken (a) to formulate an efficient and cost effective medium for the production of proteolytic enzyme(s), and (b) to isolate, purify and characterize new proteolytic enzyme(s) from an indigenous strain of *Bacillus subtilis* by employing ammonium sulphate fractionation, anion exchange chromatography and cation exchange FPLC.

The work presented here reports the primary structure of a variant of subtilisin from *Bacillus subtilis* strain RT-5. Successive fractionations and chromatographies of cell free broth resulted in the isolation of a purified alkaline serine protease subsequently named subtilisin RT-5. Separation of alkaline and neutral proteases on anion exchanger led to the realization that this approach provides an efficient means of separating the two proteases from each other.

The complete primary structure of subtilisin RT-5 has been established by a combination of direct sequence analysis of the intact protein, digestion and peptide analysis and DNA sequence analysis of PCR-produced fragment corresponding to the segment not recovered in the protein digest. The enzyme consists of 275 amino acid residues. It shows 97% homology to mesentericopeptidase with 5 exchanges and 87% homology to subtilisin Novo with 40 amino acid replacements.

Though the physico-chemical properties of subtilisin RT-5 are similar to those of subtilisin Novo, its catalytic efficiency has been found to be 40 fold higher than that of subtilisin Novo. The catalytic prowess of subtilisin RT-5 might be due to the amino acid replacements in the region around the active site His64.

The isolated enzyme is closely related to mesentericopeptidase, but it is important to note that this closest homologue is thermolabile suggesting
that the replacements of Ala by Ser at positions 85, 89 & 183, and Ser by Ala and Asn at positions 88 & 259 respectively in subtilisin RT-5 might be responsible for its thermostability.

The enzyme shows preference toward hydrophobic residues like valine and phenylalanine at P4 position which might be due to the replacement of Ser130 by Thr130 at the entrance of the P4 binding pocket.
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Growth rate of <em>Bacillus subtilis</em> strain GBS01 in media containing different concentrations of beef extract.</td>
<td>47</td>
</tr>
<tr>
<td>3-2</td>
<td>Growth rate of <em>Bacillus subtilis</em> strain GBS01 in media containing tryptose and gelatin.</td>
<td>48</td>
</tr>
<tr>
<td>3-3</td>
<td>Growth rate of <em>Bacillus subtilis</em> strain GBS01 in media containing different concentrations of gelatin.</td>
<td>49</td>
</tr>
<tr>
<td>3-4</td>
<td>Growth rate of <em>Bacillus subtilis</em> strain GBS01 in media containing corn steep liquor and/or yeast extract.</td>
<td>50</td>
</tr>
<tr>
<td>3-5</td>
<td>Growth rate of <em>Bacillus subtilis</em> strain GBS01 in media containing M9 salts and/or B1.</td>
<td>51</td>
</tr>
<tr>
<td>3-6</td>
<td>Growth rate of <em>Bacillus subtilis</em> strain GBS01 in media containing different salts.</td>
<td>52</td>
</tr>
<tr>
<td>3-7</td>
<td>Growth rate of <em>Bacillus subtilis</em> strain GBS01 in media containing different concentrations of magnesium sulphate.</td>
<td>53</td>
</tr>
<tr>
<td>3-8</td>
<td>Schematic presentation of separation protocol for proteases.</td>
<td>54</td>
</tr>
<tr>
<td>3-9</td>
<td>Elution profile of subtilisin RT-5 on cation exchanger Mono-S FPLC.</td>
<td>55</td>
</tr>
<tr>
<td>3-10</td>
<td>Rechromatography of peak 2, fig.3-9 on Mono-S FPLC.</td>
<td>56</td>
</tr>
<tr>
<td>3-11</td>
<td>SDS-PAGE of peak 1, fig.3-10, showing purified subtilisin RT-5 on 12.5% gel.</td>
<td>57</td>
</tr>
<tr>
<td>3-12</td>
<td>Gel permeation chromatography of CNBr digested subtilisin RT-5 on Superose 12 FPLC column.</td>
<td>58</td>
</tr>
<tr>
<td>3-13</td>
<td>Elution profile of CNBr digested subtilisin RT-5 on Vydac C4 column.</td>
<td>59</td>
</tr>
<tr>
<td>Fig. 4-4</td>
<td>Comparison of the primary structure of subtilisin RT-5 with subtilisin Carlsberg.</td>
<td></td>
</tr>
<tr>
<td>Fig. 4-5</td>
<td>Short repeated sequences within the subtilisin RT-5.</td>
<td></td>
</tr>
<tr>
<td>Fig. 4-6</td>
<td>Genetic map of <em>Bacillus subtilis</em>.</td>
<td></td>
</tr>
<tr>
<td>Fig. 4-7</td>
<td>Structure of pEMBL 8 vector.</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 3-14  Elution profile of peak 1, fig.3-12 on Ultropac C<sub>18</sub> column.

Fig. 3-15  Elution profile of tryptic digested pool A of fig.3-13 on Ultropac C<sub>18</sub> column.

Fig. 3-16  Elution profile of tryptic digested pool B of fig.3-13 on Ultropac C<sub>18</sub> column.

Fig. 3-17  Elution profile of Lys-C digested peptides on Ultropac C<sub>18</sub> column.

Fig. 3-18  Elution profile of Asp-N digested peptides on Ultropac C<sub>18</sub> column.

Fig. 3-19  Primary structure of subtilisin RT-5 from Bacillus subtilis strain RT-5.

Fig. 3-20  Effects of CaCl<sub>2</sub> on the activity of subtilisin RT-5 towards succinyl Ala-Ala-Pro-Phe p-nitroanilide.

Fig. 3-21  Influence of pH on k<sub>catalytic</sub>/K<sub>m</sub> of subtilisin RT-5.

Fig. 3-22  Elution profile of subtilisin RT-5 digested horse liver alcohol dehydrogenase on Ultropac C<sub>18</sub> column.

Fig. 3-23  Cleavage sites of subtilisin RT-5 on horse liver alcohol dehydrogenase.

Fig. 3-24  Thermostability of the subtilisin RT-5 at 50°C.

Fig. 3-25  DNA sequence of the subtilisin RT-5 gene.

Fig. 4-1  Comparison of the primary structure of subtilisin RT-5 with other subtilisins.

Fig. 4-2  Replacements of the amino acid residues around the reactive site His64 between subtilisin RT-5 and other subtilisins.

Fig. 4-3  Comparison of the primary structure of subtilisin RT-5 with subtilisin Novo.
LIST OF TABLES

Table 1-1  Subtilisins and their characteristics.  5
Table 3-1  Biochemical characteristics of
Bacillus subtilis strains GBS01 and RT-5.  72
Table 3-2  Amino acid composition of subtilisin RT-5.  74
Table 3-3  Comparison of amino acid compositions of
subtilisin RT-5 with other subtilisins.  75
Table 3-4  Amino acid compositions of tryptic peptides
of subtilisin RT-5.  76
Table 3-5  Effects of EDTA, PMSF and 3,4-DCI on the
proteolytic activity of subtilisin RT-5.  77
Table 4-1  Growth rate constants and protease
production obtained by Bacillus subtilis GBS01 in
different media.  94
Table 4-2  Genetic markers of Bacillus subtilis.  101
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMBIC</td>
<td>Ammonium bicarbonate</td>
</tr>
<tr>
<td>CSL</td>
<td>Corn steep liquor</td>
</tr>
<tr>
<td>3,4-DCI</td>
<td>3,4-Dichloro isocumarin</td>
</tr>
<tr>
<td>DEAE</td>
<td>Diethyl amino ethyl</td>
</tr>
<tr>
<td>DFP</td>
<td>Diisopropyl fluorophosphate</td>
</tr>
<tr>
<td>DIP</td>
<td>Diisopropyl phosphate</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene diamine tetra acetic acid</td>
</tr>
<tr>
<td>FPLC</td>
<td>Fast protein liquid chromatography</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>PTH</td>
<td>Phenyl thiohydantoin</td>
</tr>
<tr>
<td>RP</td>
<td>Reversed phase</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N,N-Tetra methyl-ethylene diamine</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoro acetic acid</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxy methyl) amino methane</td>
</tr>
</tbody>
</table>
1-0 INTRODUCTION
1-1 General Introduction

Proteases constitute a group of enzymes that cleave the peptide bonds of proteins and peptides. In fact the very term enzyme was first used for the well-known proteolytic enzyme called trypsin. In the beginning of the 20th century, Northup et al. (1948) isolated and crystallized trypsin and chymotrypsin from the bovine pancreatic tissue extract. Following the discovery of the proteolytic enzymes of the mammalian digestive tract, proteases in various animal tissues and of vegetable origin were discovered. With the development of modern techniques of large-scale cultivation of bacteria the extracellular protease of Bacillus subtilis, named subtilisin became the first to be obtained in a crystalline form (Guntelberg and Ottensen, 1952). This was followed by the isolation and very thorough characterization of other members of the same class of proteases.

Proteases of animal origin are secreted in the form of inactive precursors called proenzymes or zymogens; activation of these is generally accomplished by limited proteolysis (Davie and Neurath, 1955). In contrast nearly all of the plant and microbial proteases are produced in their active form.

Microorganisms most commonly used for the production of proteases include species of Bacillus, Aspergillus, Mucor and Rhizopus. The genus Bacillus produces a large number of extracellular enzymes, many of which are of considerable industrial importance (Priest, 1977).

Genes for a number of microbial proteases have been cloned to achieve increased enzyme production and to gain more detailed knowledge of mechanisms of secretion and control of gene expression (Sashihara et al., 1984; Bernier et al., 1983; Fujii et al., 1983; Cornelis et al., 1982; Palva, 1982; Gray and Chang, 1981).
Organisms of the genus Bacillus may be divided into three groups with respect to their protease production:

1) Species producing neutral protease alone.

2) Species producing alkaline protease alone.

3) Species producing both neutral and alkaline proteases.

Strains of *Bacillus subtilis* produce either alkaline protease alone or a mixture of neutral and alkaline enzymes (Fogarty et al., 1974).

1-2 Classification of proteases

Proteases may be classified with regards to their origin, mechanism of action and cleavage specificity. There is thus no single criterion to classify them but anyone of these or a combination of them can be used and each will result in a different grouping of these enzymes.

1) *Origin*: On the basis of origin they may be classified into:

a) animal, b) plant, and c) microbial proteases.

2) *Nature of catalytic residues of the active center*: Four different types can be distinguished on this basis:

a) *Serine proteases*: They have a unique serine residue in their active center and are specifically inhibited by diisopropylfluorophosphosphate (DFP) and other organophosphates.

b) *Thiol proteases or cysteine containing proteases*: Their activity depends on an intact SH group in their active center. They are inhibited by thiol reagents e.g., heavy metals, alkylating and oxidizing agents.

c) *Acid proteases or Aspartate proteases*: They contain aspartic acid in their active center. They are not inhibited either by chelating agents or by thiol group reagents.
d) Metalloproteases: Their activity depends on the presence of bound divalent cations at their active center. Zinc is the catalytically active metal. They can be inactivated by chelating agents like EDTA.

3) Point of cleavage: Proteases may be divided into three large groups on the basis of location of the cleavage site(s) within the target molecule.

a) Exopeptidases: They split off the terminal residue of a polypeptide chain. They are further classified as:

i) Carboxypeptidases: They split off the C-terminal residue.

ii) Aminopeptidases: They split off the N-terminal residue.

b) Endopeptidases: They split the peptide bond between residues present inside the polypeptide chain.

c) Dipeptidases: They require the presence of both a charged α-amino and a charged α-carboxyl group in the immediate vicinity of the peptide bond.

1-3 Serine proteases

Serine proteases are the most extensively studied and best understood enzymes. They are characterized by the presence of a unique serine residue at the active center that covalently binds with substrates and inhibitors. This class of proteases includes two evolutionary families which are not homologous in their amino acid sequences and 3-D structures but have very similar mechanisms of action. This is one of the best examples of convergent evolution of proteins at the molecular level. One family is typified by subtilisin. The other is the trypsin family which includes chymotrypsin, trypsin, mammalian elastases, and several bacterial proteases.
1-4 Subtilisin family

This family includes extracellular alkaline proteases from *Bacillus subtilis* and other species of Bacillus, which contain active site serine and are sensitive to DFP.

*History of subtilisins*

Subtilisin was first discovered by Linderstrom-Lang and Ottesen (1947) during an investigation of the conversion of ovalbumin to palkalbumin. Later on they isolated and crystallized the enzyme called subtilisin Carlsberg (Guntelberg, 1954; Guntelberg & Ottesen, 1954; Guntelberg & Ottesen, 1952) in the Carlsberg laboratory, Denmark. Hagihara et al. (1958) isolated and crystallized another type of subtilisin named subtilisin BPN'. He also characterized the properties of both subtilisin Carlsberg and BPN' (Hagihara, 1960).

Ottesen and Spector (1960) also reported another extracellular alkaline protease from *Bacillus subtilis* called subtilisin Novo which showed differences in chromatographic and kinetic properties from subtilisin Carlsberg. Tsuru et al. (1966a) purified subtilisin Amylosacchariticus from *Bacillus amylosacchariticus* by a combination of ammonium sulphate precipitation, chromatography on Duolite A-2, CM-cellulose, and DEAE-Sephadex. They crystallized the enzyme from Tris-HCl, calcium acetate buffer at pH 8.5. Enzymatic properties of this enzyme were found to be similar to other subtilisins.

Welker and Campbell (1967a, 1967b) showed that the large differences in the amino acid sequences of subtilisin BPN' and subtilisin Carlsberg as reported by Smith et al. (1966) were due to the fact that they were isolated from different species of Bacillus rather than from different strains of *Bacillus subtilis*. They also suggested that the organism from which subtilisin BPN' was isolated is a different species from *Bacillus subtilis* and
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Subtilisin</td>
<td>References</td>
</tr>
<tr>
<td>E. subtilis strain RT-5</td>
<td>Medkov et al. (1983)</td>
</tr>
<tr>
<td>E. subtilis strain I-168</td>
<td>Tsuru et al. (1966)</td>
</tr>
<tr>
<td>E. subtilis strain DY</td>
<td>Hershko et al. (1970)</td>
</tr>
</tbody>
</table>

**Table 1.1**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. subtilis strain RT-5</td>
<td>Medkov et al. (1983)</td>
</tr>
<tr>
<td>E. subtilis strain I-168</td>
<td>Tsuru et al. (1966)</td>
</tr>
<tr>
<td>E. subtilis strain DY</td>
<td>Hershko et al. (1970)</td>
</tr>
</tbody>
</table>

- Highly soluble in water at 50°C. Activity even after 24 h incubation at 50°C.
- Thrombins soluble in 60% of ethanol and ions enhanced by Ca²⁺ ions.
- E. subtilis strain RT-5
- The gene encoding the enzyme has been sequenced.
- The gene encoding the enzyme is 8.5 kb.
- Asp 23.
- Less soluble in water; sensitive to ionic strength.
- Similar to other subtilisins.
- In water, 50% ethanol, highly soluble.
- Resistant to 6 M urea and 2 M NaCl.
- E. amyloliquefaciens
- E. novaebrpn.
- Ph 7.8, Ca²⁺ and other ions.
- E. amyloliquefaciens
- Ph 9.4; more stable at acidic pH.
should be called *Bacillus amyloliquefaciens*. Keay and Moser (1970, 1969) prepared alkaline proteases from *Bacillus subtilis* strain NRRL 3411, *Bacillus licheniformis*, and *Bacillus pumilus*. On the base of amino acid compositions, serological cross reactions and ratios of esterase to protease activity they classified subtilisin Carlsberg and the enzymes of *Bacillus licheniformis* and *Bacillus pumilus* into Group-A while Group-B included subtilisins Novo, BPN', Amylosacchariticus and subtilisin from *Bacillus subtilis* NRRL B3411. These enzymatic groupings have also been referred to as Bacillopeptidases A and Bacillopeptidases B:

1-4-1 Subtilisin BPN'

Subtilisin BPN' was first isolated by Hagiwara et al. (1958). Matsubara and Nishimura (1958) reported that the enzyme had a molecular mass of 30.5 kDa with alanine as the N-terminal residue. The enzyme was found to be stable at acidic pH and stabilized by Ca\(^{++}\) and other salts (Matsubara et al., 1958). Matsubara et al. (1965) reported that the enzyme is a single polypeptide chain. They inferred a molecular mass of 27.6 kDa by its amino acid analysis.

The complete amino acid sequence of subtilisin BPN' was determined from tryptic, chymotryptic, peptic, and CNBr digests (Markland and Smith, 1967; Kasper and Smith, 1966a, 1966b; Markland et al., 1966). The polypeptide chain contains 275 residues with conspicuous absence of cysteine. The reactive serine residue is at position 221. It appeared that there is no homology between amino acid sequence of subtilisin BPN' and those of pancreatic proteases (Smith et al., 1966). Sanger and Shaw (1960) showed that the sequence around the reactive serine, 220Thr-Ser\(^*\)-Met-Ala223, is different from that around the reactive serine in the mammalian enzymes i.e., 194Asp-Ser\(^*\)-Gly196.

There are several repeated sequences in the polypeptide chain which are believed to have arisen from gene duplications during evolution of
microbial serine proteases and certain other enzymes e.g., trypsin from a common ancestor (Doolittle, 1979). There are many di, tri, and one tetra peptide repetitions of the same residue e.g., Ala (six Ala-Ala sequences, three of which are Ala-Ala-Ala), Ser (six Ser-Ser, including one Ser-Ser-Ser), Thr (one Thr-Thr-Thr), Gly (two Gly-Gly), Ile (one Ile-Ile) and Tyr (one Tyr-Tyr sequence).

1-4-2 Subtilisin Carlsberg

It was first isolated by Linderstrom and Ottesen (1947). Gunnelberg and Ottesen (1954) reported that the enzyme had an isoelectric point of 9.4 and was more stable in the acidic pH range of 5.3-6.5 than at alkaline pH; however, in the alkaline pH range it was more stable than trypsin or chymotrypsin. Molecular mass was found to be 27.27 kDa with alanine as the N-terminal residue.

The complete amino acid sequence was determined from tryptic (De Lange and Smith, 1968a, 1968b) and chymotryptic (Landon et al., 1968; Evans et al., 1968) peptides. The protein contains 274 residues in a single polypeptide chain. There are 84 replacements towards subtilisin BPN' and one deletion of proline at position 56. The presence of 11 lysine and 2 arginine residues in subtilisin BPN' and 9 lysines and 4 arginines in subtilisin Carlsberg at different positions in the two polypeptide chains results in different tryptic peptides of the two enzymes (Smith et al., 1968).

1-4-3 Subtilisin Novo

Smith et al. (1968) reported that the subtilisin BPN' isolated by Hagihara et al. (1958) and subtilisin Novo isolated by Ottesen and Spector (1960) are identical because the chromatographic behavior of both the enzymes and the amino acid compositions of the 14 expected tryptic peptides are similar. Both enzymes had identical enzymatic and chemical properties (Barel and Glazer, 1968; Glazer, 1967; Johansen and Ottesen, 1964; Hunt
and Ottesen, 1961). Like other subtilisins, subtilisin Novo was also found to be resistant to denaturing agents such as 6M urea and 50% ethanol (Gounaris and Ottesen, 1965), and stable in the presence of detergents like SDS and sodium tripolyphosphate (Tsuru, 1969).

1-4-4 Subtilisin Amylosacchariticus

Subtilisin Amylosacchariticus was found to be less soluble in neutral solution and water as compared to other subtilisins (Tsuru et al., 1966a, 1966b). At a high concentration the enzyme (0.5%) could be crystallized between pH 6.5 and 8.5 by dialysis against weak salt solutions. Molecular mass of the enzyme was reported to be 22.7 kDa with alanine as the N-terminal residue (Tsuru et al., 1967). It was later found that the molecular mass of DIP inactivated enzyme was 28.0 kDa, thus the lower values obtained by Japanese workers may have been due to autolysis. The isoelectric point of 7.8 and the amino acid composition of the enzyme resemble that of the subtilisin BPN' rather than the subtilisin Carlsberg.

Complete amino acid sequence of subtilisin Amylosacchariticus was reported by Kurihara et al. (1972). The determined sequence shows that the enzyme is similar to subtilisin BPN', although there are some regions which resemble subtilisin Carlsberg. There is one interesting replacement in subtilisin Amylosacchariticus around the reactive Ser221 i.e., Ser224 has been substituted by Thr 224 in subtilisin Amylosacchariticus.

1-4-5 Subtilisin DY

Nedkov et al. (1983) reported the complete amino acid sequence of subtilisin DY, which was isolated from Bacillus subtilis strain DY. This protein is closely related to subtilisin Carlsberg. There are 32 amino acid replacements between subtilisin DY and Carlsberg and 82 replacements and one deletion towards subtilisin Novo. The single polypeptide chain of subtilisin DY consists of 274 residues with N-terminal alanine.
1-4-6 Alkaline mesentericopeptidase

Another serine protease known as alkaline mesentericopeptidase was isolated and crystallized by Karadzhova et al. (1970) from a strain of *Bacillus mesentericus*. Svendsen et al. (1986) deduced the complete amino acid sequence of this enzyme. They classified it as a serine endopeptidase of the subtilisin Novo type. Comparison with subtilisins (Novo, Carlsberg, Amylosacchariticus, DY, and a subtilisin isolated from *Bacillus subtilis* strain 168), showed that the alkaline mesentericopeptidase is closely related to subtilisin Amylosacchariticus in physico-chemical properties. This is not surprising in view of the fact that there are only six amino acid replacements between the two enzymes (Svendsen et al., 1986).

1-4-7 Glu/Asp- specific protease from *Bacillus licheniformis*

Svendsen and Breddam (1992) isolated an endopeptidase from *Bacillus licheniformis* which cleaves specifically at the carboxyl side of glutamic acid. The isolated enzyme consists of a single polypeptide chain of 222 amino acids with a molecular mass of 23.6 kDa. The amino acid sequence analysis revealed similarities with the Glu/Asp-specific enzymes from *Staphylococcus aureus* V8, Actinomyces species and *Streptomyces thermovulgaris* (De Filippis and Fontana, 1990; Khaidorva et al., 1989).

1-5 Site-directed mutagenesis of subtilisins

Site directed mutagenesis is a well established method of protein engineering. It has been used extensively to develop enzymes with new activities or altered specificities. Wells et al. (1987) and Wells & Estell (1988) have used subtilisin BPN’ as a model system for protein engineering. Wright et al. (1999), Kraut (1977), and Bott et al. (1988) have reported high resolution crystal structure of subtilisin BPN’. Bode et al.
(1986), Mitsui et al. (1979) and McPhalen & James (1988) reported crystal structures of subtilisin complexes with various inhibitors like eglin C, streptomycines subtilisin inhibitor (SSI), and chymotrypsin inhibitor (CHI), respectively.

Furthermore, with the development of suitable cloning and expression vectors (Thomas et al., 1985; Wells et al., 1983), and establishment of accurate methods to quantitate the effect of mutations on proteolytic activity (Groen et al., 1992; Estell et al., 1985; Thomas et al., 1985), subtilisin BPN' became a good model system for site directed mutagenesis and protein engineering.

Carter et al. (1989) prepared variants of subtilisin BPN' by using combination of protein engineering and substrate optimization. The variant enzyme specifically cleaved a designed target sequence in a fusion protein. They replaced the catalytic histidine 64 with alanine to modify the specificity of the wild type subtilisin BPN' in such a way that certain histidine containing substrates are preferentially hydrolysed.

Carter and Wells (1987) reported a novel approach to engineer enzyme specificity in which enzyme was inactivated by removing catalytic group by site directed mutagenesis. Then the catalytic power was partially restored by the substrate containing the missing catalytic functional group. Catalytic efficiency of the mutant enzyme was reduced by a factor of $10^6$, as assayed with Succinyl phenylalanyl-alanyl-alanyl-phenylalanyl p-nitroanilide (sFAAF-pNA), when catalytic histidine-64 was replaced with alanine. They found that His64 to Ala64 mutant hydrolyzes His P2 substrate (sFAHP-pNA) upto 400 times faster than a homologous Ala P2 or Gln P2 substrate (sFAAF-pNA or sFACG-pNA) at pH 8.0. They named this process as substrate-assisted catalysis.
1-6 Substrate specificity of subtilisins

Subtilisins are non-specific proteases and have broad substrate specificity. They mainly use two specificity pockets, S1 and S4, which interact with P1 and P4 amino acid residues of the substrate, respectively. Estell et al. (1986) and Wells et al. (1987) generated mutations at the S1 pocket and thereby tried to transform a non-specific subtilisin to a specific one by changing the hydrophobic, steric, and electrostatic environment of the enzyme. Carter and Wells (1987) and Carter et al. (1991) developed a mutant subtilisin, by using the mechanism of substrate-assisted catalysis, which was highly specific. But the mutant enzyme had reduced turnover of substrate. By using peptide substrates of different sizes Groen et al. (1992) reported that S4 pocket is nearly as important as the S1 pocket in determining the substrate specificity of BPN'.

Rheinheimer et al. (1993) reported that single amino acid replacements at positions Tyr104 and Ile107 improved the specificity of subtilisin BPN' toward substrates with large hydrophobic P4 residues. They increased the size of the P4 binding pocket by (a) replacing Ile107 with Gly, Ala, and Val, and (b) eliminating the hydrogen bond between Tyr104 and Ser130 at the entrance of the P4 binding pocket by replacing Tyr104 by Phe. They showed that $k_{cat}/K_m$ was increased by 200 fold for Phe as compared to Ala at the P4 position.

1-7 Specific proteases from other microorganisms

As subtilisins are non-specific proteases, they do not require the presence of a particular amino acid residue at the P1 position but there is a group of other proteolytic enzymes which require a particular sequence around the scissile bond, and these are specific proteases. These enzymes are particularly convenient for specific cleavages and generation of longer peptides containing several cleavage sites.
Masaki et al. (1978) isolated lysine specific enzyme from Achromobacter. Gilles et al. (1979) reported the Arg specific enzyme called clostripain. The post-proline specific enzyme from Flavobacterium was isolated by Yoshimoto et al. (1980). The Glu/Asp specific proteases from various sources were isolated by Drapeau et al. (1972), Houmard (1976), Haumard and Drapeau (1972), Khaidorva et al. (1989), Yoshida et al. (1988) and Svendsen & Breddam (1992).

1-8 Metalloproteases

Fukumoto et al. (1951) isolated Bacillus amyloliquefaciens from soil. The strain produced neutral protease (BANP) and subtilisin Amylosacchariticus (subtilisin AM) besides saccharifying type alpha amylase. Tsuru et al. (1987, 1970, 1967, 1966a, 1966b) characterized the enzymatic and physico-chemical properties of these two proteolytic enzymes. Tsuru et al. (1965) and McConn et al. (1964) have identified the neutral protease as a Zn$^{+2}$-metalloenzyme with high specific activity, but the protein sequence analysis remains to be clarified.

The complete amino acid sequence of neutral protease from Bacillus subtilis var. amylosacchariticus was reported by Kobayashi et al. (1989). The enzyme consists of 300 amino acid residues with Ala and Leu as its amino and carboxyl termini. The molecular mass has been estimated to be 32.63 kDa. Comparison of the amino acid sequence of BANP with those of other Bacillus neutral proteases, deduced by nucleotide sequencing, and with that of thermolysin revealed that the putative active site amino acid residues, Zn-binding ligands, and two calcium binding sites are well conserved among them.

van den Burg et al. (1990) have identified the autodigestion target sites in Bacillus subtilis neutral protease under various conditions. At elevated temperatures, under non-inhibitory conditions, mature protein was rapidly
degraded but there was no accumulation of specific breakdown products. However, the specific peptides were accumulated by incubating purified enzyme on ice for a longer duration. They thus identified five fission sites in the neutral protease, three of which were identical with autodigestion target sites in thermolysin, as determined by Fontana et al. (1986). They suggested a high degree of similarity in the three-dimensional structures of Bacillus subtilis neutral protease and thermolysin.

At the amino acid level, the neutral protease from Bacillus subtilis shows 47.2% sequence identity with thermolysin. Morihara & Tsuzuki (1970) by using synthetic oligopeptides showed that the substrate specificities of both proteases are very similar. Pangburn & Walsh (1975), Hangauer et al. (1984) and Tronrud et al. (1986) did crystallographic studies of thermolysin-inhibitor complexes and suggested a catalytic mechanism in which the Zn$^{+2}$ ion and a limited number of residues play a critical role in enzyme catalysis. These residues are conserved in the Bacillus neutral proteases. Toma et al. (1989) replaced two of these residues in the Bacillus subtilis neutral protease by other amino acids. They strongly supported the notion that the catalytic mechanisms in both enzymes are highly similar.

1-9 Protease production and spore formation

The relationship between protease production and sporulation is still not well understood but there are few hypotheses.

At the end of logarithmic phase Millet (1970, 1969) and Prestidge et al. (1971) isolated three extracellular proteases i.e., a metalloprotease, a serine protease and an esterase from sporulating cultures of Bacillus subtilis. Priest (1977) reported that increased amount of proteases are secreted into the medium at the end of logarithmic phase and during the course of sporulation. In addition to extracellular proteases one intracellular protease was isolated by Reysset and Millet (1972), Stepnov
et al. (1977), Strongin et al. (1978) and Kerjan et al. (1979) at the end of logarithmic phase in sporulating cells.

Hoch & Spizzen (1969) and Schwaffer (1969) in their separate experiments found that the non-sporulating mutants of *Bacillus subtilis* also lack protease production. Prestidge et al. (1971) reported that specific activities of all three types of proteases rapidly increased in cells undergoing spore formation and established the correlation between protease production and sporogenesis. An increase in extracellular protease production in *Bacillus* during sporulation was also observed by Higerd et al. (1972). Leighton et al. (1972) reported that an extracellular serine protease is important for generation of spores. Similarly it was reported by Tanaka et al. (1987) that *spoO* mutation in *Bacillus subtilis* which inhibited spore formation also reduced extracellular production of serine and metalloproteases.

In contrast to the foregoing observations Millet (1970) and Shoer & Rappaport (1972) isolated extracellular protease deficient mutants which sporulated normally. Hageman and Carlton (1973) isolated a mutant which could sporulate normally but could not produce metalloprotease. They suggested that metalloproteases do not play any or at least a significant role in sporulation. Later on Yang et al. (1984) studied the mechanism of sporulation and protease production by using deletion mutants of a neutral protease. They confirmed the observations of Millet (1970), Shoer & Rappaport (1972), and Hageman & Carlton (1973) that deletions in genes encoding neutral and alkaline proteases had no effect on spore formation while the protease production was completely abolished.

As compared to extracellular proteases intracellular ones play an important role in protein turnover and modification; hence it is likely that they are more significant in sporogenesis. Possible role of intracellular proteases in sporulation has been studied from sporulating cultures of *Bacillus cereus* by Cheng and Aronson (1977a, 1977b), *Bacillus*
megaterium by Chaloupka et al. (1977), and Bacillus thuringiensis by Lacadet et al. (1977).

Szulmajster and Kere (1975), Cheng & Aronson (1977a, 1977b) and Kerjan et al. (1979) reported that intracellular proteases are responsible for protein turnover and processing of spore coat precursor proteins. Lacadet et al. (1977) have suggested that they might be involved in modification of DNA dependent RNA polymerase. Bernlohr and Clark (1971) indicated the scavenger action of intracellular proteases.

1-10 Cloning of genes encoding proteases

Several genes encoding proteolytic enzymes have been cloned with the purpose of enhancing protease production as well as to gain a better understanding of gene expression and mechanism(s) of protein secretion (Sashihara et al., 1984; Bernier et al., 1983; Fujii et al., 1983; Cornelis et al., 1982; Palva, 1982; Gray and Chang, 1981).

1-10-1 Cloning of subtilisin gene

The gene for serine protease subtilisin BPN' (subtilisin Novo) from Bacillus amyloliquefaciens has been cloned and expressed in Bacillus subtilis independently by Wells et al. (1983), Vasantha et al. (1984), and Thomas et al. (1985).

Wells et al. (1983) cloned and expressed the subtilisin gene from Bacillus amyloliquefaciens under its own promoter on a high copy number plasmid, pBS42, in Bacillus subtilis Marburg strain I-168. They deduced the complete nucleotide sequence of the cloned gene by the dideoxy sequencing method. By Bal31 exonuclease digestion studies at the 5' end of the gene they defined a 31 base pair stretch necessary for efficient expression of subtilisin. Beside this putative promoter region they deduced a sequence of unknown function between the signal sequence and the
mature enzyme, which codes for roughly 75 amino acid residues. They proposed that this region serves as a pro-peptide as is commonly found in secreted proteolytic enzymes of eukaryotes.

Vasanthan et al. (1984) isolated genes for alkaline protease (apr) and neutral protease (npr) from Bacillus amylo liquefaciens and expressed them in Bacillus subtilis. The DNA sequences of apr and npr revealed the presence of a large open reading frame. They also reported the presence of signal sequence and an additional polypeptide sequence ("pro"-sequence) before the mature protein. They identified the translation initiation site as the amino acid residue -107 for apr and -221 for npr. They suggested that the presence of a "pro" sequence may be common to all of the secreted proteases from bacilli.

Thomas et al. (1985) cloned the structural gene encoding subtilisin BPN' from Bacillus amylo liquefaciens in the vector pUB110. They named the plasmid which harbors the subtilisin gene on a 3.4 kilobase EcoR1 fragment as pPT30. They used a series of primers to deduce the DNA sequence of subtilisin gene by the dideoxy method. In these experiments Bacillus subtilis strain DB 104, which is deficient in extracellular alkaline and neutral proteases, was used as a recipient. Subtilisin BPN' was expressed and secreted at a high level by DB 104 transformed with pPT30. Furthermore, the DNA sequence of the structural gene of subtilisin was found to be identical to that published by Wells et al. (1983) and Vasanthan et al. (1984).

1-10-2 Cloning of neutral protease

Fujii et al. (1983) were the first to clone and express the gene encoding neutral protease from Bacillus stearothermophilus into Bacillus subtilis. The 4.5 MDa cloned fragment could express itself in both Bacillus stearothermophilus and Bacillus subtilis strains. Vasanthan et al. (1984) cloned in addition to alkaline protease gene, the npr gene from Bacillus

16
*amyloliquefaciens* in *Bacillus subtilis*. They deduced the complete nucleotide sequence of the npr gene. Neutral protease gene from one strain of *Bacillus subtilis* into another strain of the same species has been cloned by Yang et al. (1994). Genes encoding alpha amylase and neutral proteases from an industrial strain Bacillus Bs1 have been cloned into two plasmids by Reid et al. (1986) and the plasmids were introduced into *Bacillus subtilis*. Both plasmids remained stable and produced mainly the Bs1 alpha amylase and the neutral protease along with few other extracellular proteins.

Kubo and Imanaka (1988) cloned the nprM gene encoding a highly thermostable neutral protease from *Bacillus stearothermophilus* MK 232 in *Bacillus subtilis*. They also deduced the nucleotide sequence of nprM gene. They reported that the specific activity and thermostability of NprM protease is higher than that of thermolysin. Only two amino acid replacements towards thermolysin were found when the amino acid sequence of the extracellular protease was compared with that of thermolysin while the rest of the primary structure was identical.

**1-10-3 Neutral protease as a cloning and expression marker**

A *Bacillus subtilis* strain deficient in three proteases which can be used as a useful host to study the secretion of other proteins was developed by Nakamura et al. (1990).

Wu et al. (1991) by using pUB110 derived plasmid pWL 267 carrying a neutral protease gene as a screening marker (Wang et al., 1989) have developed a novel *Bacillus subtilis* cloning system which is simple and rapid. As compared to X-Gal/IPTG of the lacZ system developed by Haima et al. (1990a,1990b) this system is very economical because in this system detection is based on halo formation due to casein hydrolysis. Furthermore, a variety of hosts can be used and plasmids remain stable in host strains. A variety of restriction sites for different restriction enzymes
are also available. Due to flexibility in the propeptide region of nprE gene many more cloning sites can be inserted into the vector. Hence this system can be used as an expression/secretion vector when heterogenous target genes are cloned in frame to nprE gene.

Kawamura et al. (1979) and Jones & Errington (1987) have developed cloning systems of Bacillus subtilis, such as Q105 system. The selection mechanism of Wu et al. (1991) can be adapted to such cloning systems because no modification of the marker gene is necessary for its expression. It is also possible to apply the system with other microorganisms as demonstrated by cloning and expression of nprE gene from Bacillus subtilis in Escherichia coli (Wang et al., 1990).

1-11 Culture medium for the production of proteases

Since decades proteolytic enzymes are produced by the fermentation of natural media by microorganisms. The two major advantages of using natural media are: a) they are inexpensive, and b) they yield high quantities of proteolytic enzymes.

After 1960's various synthetic and semisynthetic media were used for the cultivation of Bacillus species to study their metabolic regulation and protease production. Glucose was used as the only carbon source. Yeast extract, sodium glutamate and inorganic ammonium salts were used as nitrogen sources (Verseveld et al., 1986; Frankena et al., 1985; Halon and Hodges 1981a, 1981b; Wouters and Buysman, 1977; Dawes and Thornley, 1970; Dawes and Mandelstam, 1970). These, however, were expensive and the yield of proteolytic enzymes obtained was not comparable to the natural media.

Presently proteases are generally produced by the cultivation of microorganisms in a liquid medium for two to seven days at 20°C to 37°C at a pH between 5 to 8 on a natural or synthetic medium containing
suitable carbon and nitrogen sources and inorganic salts. Any method of surface or shake culture ensuring aerobic conditions is employed.

1-12 Industrial applications of proteases

Detergents:

The enzymes used in detergents have been almost totally subtilisins, produced by the selected strains of *Bacillus subtilis* or *Bacillus licheniformis*. They meet the requirements of a detergent enzyme (Fogarty et al., 1974).

Leather Industry:

Enzymes may be used in all stages of leather preparation: i) Initial soaking, ii) dehairing, iii) bating i.e., removal of fibrillar material. Proteases of *Bacillus subtilis* are mainly used in the bating process. The major advantages of the use of enzymes in this industry are that a better quality of leather or animal hair is obtained and that considerable reduction in the problem of effluent disposal is also achieved (Grimm, 1958).

Baking Industry:

In the manufacture of biscuits and crackers, the dough should be very soft in order to prevent bending and wrinkling of biscuits in the oven. In this respect Bacillus enzyme particularly neutral proteases have shown excellent properties. The amount of enzyme added to the flour is extremely small (Bryce, 1966).

Brewing Industry:

In the manufacture of sugars of the malt replacement type, it is usually necessary to add protease to achieve sufficient breakdown of protein. Proteases of Bacillus origin have been found suitable for this application. In order to prevent undesirable haze development in ale and lager when they are cooled, proteases are added during the finishing operations to
chili proof these beverages (Fogarty et al., 1974).

Fish Industry:

Bacillus proteases are used in the fish industry where inedible and waste scrap fish are processed for meal, oil and solubles (Lovern, 1955).

Other uses:

It has been shown that administration of Bacillus subtilis protease:amylase mixture in the form of a mouthwash resulted in a 50% reduction of plaque deposition on the surface of teeth in humans (Fogarty et al., 1974).

Use of bacterial proteases and operating at a lower temperature, result in a significant improvement in protein quality for the production of animal fat.

Other existing or potential applications include recovery of silver from photographic film by digestion of the gelatin emulsion coating, spot removal in dry cleaning of clothing, meat tenderization, modification of protein in breakfast cereal and baby food (Fogarty et al., 1974).
2-0 MATERIALS AND METHODS
2-1 Isolation of microorganisms

In this study *Bacillus subtilis* strains GBS01 and RT-5 were used.

Strain GBS01 was provided by M/s. Ghani Son’s Pvt. Ltd. Strain RT-5 was isolated from a soil sample of the Tharparker desert of Pakistan. The strain was selected because of its high proteolytic activity, as judged by the emergence of large zones of hydrolysis around the colonies on casein agar plates.

2-2 Strain identification

RT-5 strain was identified as a species of the genus *Bacillus* on the base of its gram staining reaction, spore formation and catalase production according to the Bergey’s Manual of Determinative Bacteriology (Sneath, 1980). Subsequently it was confirmed to be a strain of *Bacillus subtilis* by using API 20 E system and CHB 50 system according to O’Donell et al. (1980).

For cataloging and maintenance of record in the Microbial Culture Collection of Pakistan (MCCP) strains GBS01 and RT-5 were respectively assigned numbers B-001 and B-002.

2-3 Maintenance of bacteria

The strains were maintained on nutrient agar slants in screw capped tubes and stored at 4°C for routine use.

For long term storage 0.5 ml of autoclaved 80% glycerol was mixed with 0.5 ml of overnight bacterial culture and stored at -70°C.
Formulation of medium for efficient and cost effective production of proteolytic enzymes.

Different ingredients such as beef extract, tryptose and gelatin were used in various concentrations and combinations to formulate a medium for the production of proteolytic enzymes from _Bacillus subtilis_ GBS01. Growth curve experiments were performed with each formulation.

Compositions of the media tested are as follows:

**Media containing various concentration of beef extract:**

1. Tryptose 5%
   Gelatin 10%
   Beef extract 3%
   pH 7.0

2. Tryptose 5%
   Gelatin 10%
   Beef extract 2%
   pH 7.0

3. Tryptose 5%
   Gelatin 10%
   Beef extract 1%
   pH 7.0

**Media containing tryptose and/or gelatin:**

4. Tryptose 5%
   Gelatin 5%
   pH 7.0

5. Tryptose 5%
   pH 7.0
Media containing different concentrations of gelatin:

6. Gelatin 20%  
   Glycerol 0.2%  
   pH 7.0

7. Gelatin 10%  
   Glycerol 0.2%  
   pH 7.0

8. Gelatin 5%  
   Glycerol 0.2%  
   pH 7.0

9. Gelatin 2%  
   Glycerol 0.2%  
   pH 7.0

Media containing corn steep liquor and/or yeast extract:

10. Gelatin 10%  
    Glycerol 0.2%  
    Corn steep liquor (CSL) 0.4%  
    Yeast extract 0.1%  
    pH 7.0

11. Gelatin 10%  
    Glycerol 0.2%  
    CSL 0.4%  
    pH 7.0

12. Gelatin 10%  
    Glycerol 0.2%  
    Yeast extract 0.1%  
    pH 7.0
13. Gelatin 10%
   CSL 0.4%
   pH 7.0

Media containing M9 salts and/or Vit B1:

14. Gelatin 10%
    Glycerol 0.2%
    CSL 0.4%
    M9 salts:
    Na$_2$HPO$_4$ 0.6%
    KH$_2$PO$_4$ 0.3%
    NaCl 0.05%
    NH$_4$Cl 0.1%
    CaCl$_2$ 0.1 mM
    MgSO$_4$.7H$_2$O 1 mM
    B1 1 µg/ml
    pH 7.0

15. Gelatin 10%
    Glycerol 0.2%
    CSL 0.4%
    M9 salts*
    pH 7.0

16. Gelatin 10%
    Glycerol 0.2%
    CSL 0.4%
    B1 1 µg/ml
    pH 7.0

* as described in medium No. 14.
Media containing different salts:

| 17. | Gelatin        | 10%  |
|     | Glycerol       | 0.2% |
|     | CSL            | 0.4% |
|     | Na$_2$HPO$_4$  | 0.6% |
|     | KH$_2$PO$_4$   | 0.3% |
|     | NaCl           | 0.05%|
|     | NH$_4$Cl       | 0.1% |
|     | MgSO$_4$·7H$_2$O | 1 mM |
|     | B1             | 1 µg/ml |
|     | pH             | 7.0  |

| 18. | Gelatin        | 10%  |
|     | Glycerol       | 0.2% |
|     | CSL            | 0.4% |
|     | Na$_2$HPO$_4$  | 0.6% |
|     | KH$_2$PO$_4$   | 0.3% |
|     | NaCl           | 0.05%|
|     | NH$_4$Cl       | 0.1% |
|     | CaCl$_2$       | 0.1 mM |
|     | B1             | 1 µg/ml |
|     | pH             | 7.0  |

| 19. | Gelatin        | 10%  |
|     | Glycerol       | 0.2% |
|     | CSL            | 0.4% |
|     | Na$_2$HPO$_4$  | 0.6% |
|     | KH$_2$PO$_4$   | 0.3% |
|     | NaCl           | 0.05%|
|     | MgSO$_4$·7H$_2$O | 1 mM |
|     | CaCl$_2$       | 0.1 mM |
|     | B1             | 1 µg/ml |
|     | pH             | 7.0  |
Media containing different concentration of MgSO₄·7H₂O:

20. Gelatin 10%
    Glycerol 0.2%
    CSL 0.4%
    MgSO₄·7H₂O 2 mM
    pH 7.0

21. Gelatin 10%
    Glycerol 0.2%
    CSL 0.4%
    MgSO₄·7H₂O 1mM
    pH 7.0

22. Gelatin 10%
    Glycerol 0.2%
    CSL 0.4%
    MgSO₄·7H₂O 0.5mM
    pH 7.0

Sources of chemicals and media:

Tryptose and beef extract were obtained from Oxoid, England. Gelatin and Corn steep liquor were gifts from M/s Ghani Sons, Pvt. Ltd. Yeast was obtained from Ravi Rayon Ltd (Pakistan). M9 salts and other chemicals were obtained from E.Merck (Germany) and Fluka (Switzerland).
Determination of proteolytic activity in Lohlein Volhard (L.V.) units

For the determination of L.V. units sample and blanks were prepared as follows:

Sample:

i. 8.0 ml of 5% casein solution at pH 8.2 was placed in a water bath at 37°C for 10 minutes.

ii. 4.0 ml of enzyme solution at pH 8.2 was added and the mixture was replaced in water bath.

iii. 4.0 ml of 0.2 N hydrochloric acid was added exactly after one hour.

iv. 4.0 ml sodium sulphate solution was added. The mixture was shaken thoroughly and replaced for 20 minutes in water bath.

v. The mixture was filtered through Whatman filter paper No. 1.

Blank:

i. 8.0 ml of 5% casein solution at pH 8.2 was placed in a water bath at 37°C for 10 minutes.

ii. 4.0 ml 0.2 N hydrochloric acid was added.

iii. 4.0 ml 10% sodium sulphate solution was added and the mixture shaken thoroughly.

iv. 4.0 ml of enzyme solution at pH 8.2 was added. The mixture was shaken thoroughly and replaced in the water bath for 20 minutes.

v. The mixture was filtered through Whatman filter paper No. 1.
**Titration:**

The blank and the sample solutions were titrated as follows:

10 ml filtrate in a beaker was placed on a magnetic stirrer and titrated with 0.1N sodium hydroxide upto the pH 7.3. Volume of sodium hydroxide consumed was measured in ml. The pH was monitored by pH meter.

**Calculation:**

Proteolytic activity of the enzyme expressed in L.V. units per gram was calculated as under:

\[
A = a / b \times 17.391 \times 1000
\]

Where "A" is the proteolytic activity in L.V. units.

"a" is the consumption of 0.1N sodium hydroxide i.e., sample-blank

"b" is the concentration of enzyme in mg/ml.

**2-6 Cultivation of strain RT-5 for the production of subtilisin**

A purified colony of *Bacillus subtilis* RT-5 from Luria-casien agar plate was inoculated into medium containing 10% gelatin, 0.2% glycerol, 0.4% corn steep liquor and 5 mM CaCl₂ and incubated overnight at 37°C, after which 1.0 L of the same medium was inoculated with this culture and incubated at 37°C for 72 hrs in an orbital shaker at 220 oscillations/min.

**2-7 Isolation of serine protease from strain RT-5**

**2-7-1 Ammonium sulphate fractionation**

The 72 hours old culture broth was centrifuged at 10,000 rpm for 15 minutes to remove cells and debris. Proteins in the supernatant were precipitated with 35% and then with 75% ammonium sulphate and subsequently collected by centrifugation at 10,000 rpm for 30 minutes. The
resulting precipitates were stored at -20°C.

2-7-2 Desalting of precipitate

The crude enzyme precipitate was dissolved in 100 ml of 10 mM Tris-HCl (pH 7.5), 5.0 mM CaCl₂ and dialyzed overnight against 3 x 5 L of the same buffer at 4°C.

2-7-3 Fractionation of protein on DEAE-Sepharose

The dialyzed preparation was applied to DEAE-Sepharose fast flow (2x40 cms) column, equilibrated with 10 mM Tris-HCl, pH 7.5, 5.0 mM CaCl₂ and was eluted with the same buffer. The flow-through containing a serine protease, as checked by inhibition of the enzyme activity with 1.0 mM PMSF but not with 10.0 mM EDTA, was collected and concentrated in 10mM PO₄ buffer, pH 6.0 using amicon-PM 10 membrane filter.

2-7-4 Chromatography on Mono-S FPLC

The concentrated flow-through of DEAE-Sepharose fast flow was applied to Mono-S HR 5/5 FPLC column. The column was washed with 5 ml equilibration buffer (10 mM PO₄ buffer, pH 6.0) and eluted with a linear gradient of 0-1.0 M NaCl in the same buffer at a flow rate of 1 ml/min. The absorbance was recorded at 280 nm. The enzyme eluted at approximately 0.05 M NaCl, as revealed by proteolytic activity against casein.

2-7-5 Rechromatography of the active fraction

The active fraction was rechromatographed through the same column and in the same conditions but eluted with a shallow gradient of 0-0.2 M NaCl in equilibration buffer. The single fraction with high activity was concentrated by centricon-10 ultrafiltration.
2-8 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Molecular mass and purity of fractions after each purification step were checked by SDS-PAGE according to Laemmli (1970).

Pharmacia PhastSystem (Pharmacia, Sweden) was used to run precast 12.5 % SDS and IEF gels. Sample treatment, staining and destaining were carried out according to the manufacturer's recommendations (Phast System, Users Manual).

Composition of solutions used:

Sample treatment buffer: 10mM Tris-HCl, 1 mM EDTA, 2.5% SDS, 5% β-mercaptoethanol, pH 8.0. About 0.5-2.0 µg protein was dissolved in buffer and heated at 100°C for 5 min.

Electrode buffer: The buffer system in Phast SDS strips is of 0.2 M Tricine (trailing ion), 0.2 M Tris and 0.55% SDS, pH 7.5. The buffer strips are made up of 2% agarose.

Staining solution: 0.1% Phast Blue R solution in 30% methanol and 10% acetic acid in distilled water.

Destaining solution: 30% methanol and 10% acetic acid in distilled water.
2-9 Chemical cleavage of subtilisin RT-5

The purified enzyme was dissolved in 70% formic acid and cyanogen bromide was added to a concentration of 1mg/ml. The tube was capped and the reaction was carried out for 24 hours at room temperature (25°C).

2-9-1 Separation of cyanogen bromide peptides

Cyanogen bromide digested peptides of subtilisin RT-5 were separated and purified by Superose 12 FPLC column in 30% acetic acid and then by reversed phase HPLC on Vydac C4 column using 0.1% TFA with a linear gradient of acetonitrile containing 0.1% TFA.

2-10 Enzymatic cleavage of CNBr digested peptides of subtilisin RT-5

Cleavages with TPCK-trypsin, Lys-C protease from Achromobacter lyticus and Asp-N protease from Pseudomonas fragi were carried out as described by Kaiser et al. (1990).

About 6-8 nmols of cyanogen bromide digested peptides were dissolved in 200 µl of 20 mM ambic buffer, pH 8.15, followed by the addition of enzyme in a ratio of 1:100 (enzyme : substrate). The mixture was flushed with nitrogen and incubated at 37°C for 4 hrs for TPCK-trypsin and Lys-C proteases and for 20-22 hrs for Asp-N protease. Digestion was terminated by the addition of 10-30 µl concentrated acetic acid.

2-11 Purification of peptides

The peptides were separated by reversed phase HPLC on Ultropac C18 column. 0.1% trifluoroacetic acid (TFA) and 100% acetonitrile containing 0.1% TFA were used as solvent A and B, respectively with a gradient of 0-60% solvent B in 90 minutes.
2-12 Amino acid analysis

About 1-2 nmols of native protein/peptides were hydrolyzed with 5.7 M HCl containing 0.5% phenol at 110°C for 24 hrs under vacuum. After hydrolysis, samples were evaporated and analyzed on LKB model, Alpha Plus, amino acid analyzer.

2-13 Amino acid sequence analysis

The amino acid sequence analysis of the intact protein and peptides obtained after enzymatic/chemical digestion was performed with a Millipore solid phase sequencer 6600 and an Applied Biosystems 410A gas phase sequencer. 0.5-1.0 nmole of protein/peptide was applied to the glycine-polybrene precycled membrane and subjected to sequence analysis. The first cycle in all analyses included a double coupling step. Phenylthiohydantoin derivatives were identified by on-line reverse phase HPLC. Intact protein and peptides were applied on sequelon-AA membrane (for C-terminal coupling) for sequencing on solid phase sequencer.

2-14 Carboxymethylation of alcohol dehydrogenase

About 10 nmols of horse liver alcohol dehydrogenase was dissolved in carboxymethylation buffer (0.4M Tris-HCl, pH 8.15, 6M guanidine hydrochloride, 2 mM EDTA). 10 μl of 0.5 M DTT was added, flushed with nitrogen and incubated in dark at 37°C for 2 hours. Then 70 μl of 100mM iodoacetate was added, flushed with nitrogen and again incubated in dark at 37°C for 2 hours. At the end of the incubation period the mixture was passed through PD-10 desalting column, equilibrated with 30% acetic acid.
2-15 Substrate specificity of subtilisin RT-5

The peptide cleavage specificity of subtilisin RT-5 was determined by digestion of carboxymethylated EE chain of class 1 horse liver alcohol dehydrogenase. 5.0 nmole protein was dissolved in 70 \mu l 1-2 M urea, followed by 10 \mu l 1.0 M ammonium bicarbonate, pH 8.2, and 10 \mu l water. 0.1 nmole (10 \mu l) of enzyme was then added to an enzyme/substrate ratio of 1:50. The reaction mixture was incubated at 37°C for 4 hrs. The reaction was terminated by addition of acetic acid to a final concentration of 30%.

2-16 Assay of enzymatic activity of subtilisin RT-5

The hydrolytic activity was monitored by the Kunitz's casein method (1948). This has been modified by reducing the volume as follows:

The reaction mixture containing 330 \mu l casein in 100 mM Tris-HCl, pH 8.0, and 10-50 \mu l enzyme solution in a final volume of 660 \mu l of the same buffer was incubated at 37°C for 20 minutes. The reaction was terminated by the addition of 1.0 ml 0.33 M trichloroacetic acid. Supernatant was collected by centrifugation at 6,500 rpm for 10 minutes. Hydrolysis was determined spectrophotometrically at 280 nm with a spectrophotometer DU. 68 (Beckman, U.S.A.). An increase in absorbancy of 0.001 is equal to one Kunitz's units (KU).

2-17 Enzyme inhibition assay

Effects of 10 mM EDTA, 1.0 mM PMSF and 0.1mM 3-4 DCI were determined in 100 mM Tris-HCl, pH 8.0. The enzyme fractions were preincubated with the inhibitor for 30-60 minutes at 37°C and the remaining casein hydrolytic activity was measured by Kunitz's method.
2-18 Measurements of $k_{\text{cat}}$ and $K_m$

The kinetic parameters $k_{\text{cat}}$ and $K_m$ for the substrate succinyl-L-alanyl-L-alanyl-L-pronyl-L-phenylalanyl $p$-nitroanilide (Suc-AAPF-pNA) were monitored according to Wells et al. (1983) by the initial rate measurements of substrate hydrolysis by following absorption at 412 nm. The enzyme assays were performed in 1.0 ml of 0.1M Tris-HCl buffer, pH 8.6, at 25°C with substrate concentrations varying between 0.1 and 2 mM. Reactions were initiated by the addition of 5 $\mu$l of the enzyme solution.

2-19 pH dependence of $k_{\text{cat}}/K_m$

pH dependence of $k_{\text{cat}}/K_m$ for the hydrolysis of Suc-AAPF-pNA by subtilisin RT-5 at 25°C in 0.1 M Tris-HCl buffer, pH 8.6 was measured according to Thomas et al. (1985). The reaction was initiated by adding 5 $\mu$l of enzyme to 1.0 ml of buffered substrate (0.02 mM < $K_m$) in a cuvette maintained at 25°C in spectrophotometer DU. 68 (Beckman, U.S.A.). The increase in absorbance at 412 nm on the release of $p$-nitroanilide was monitored. The value of $k_{\text{cat}}/K_m$ at each pH value was determined from the first-order plots. The low concentration of substrate used in this method avoids product inhibition.

2-20 Thermostability of subtilisin RT-5

Thermostability of subtilisin RT-5 was monitored by measuring the residual enzyme activity at various temperatures using succinyl Ala-Ala-Pro-Phe $p$-nitroanilide as a substrate. Sample at a concentration of 500 $\mu$g/ml in 100 mM Tris-HCl, pH 8.5, 10mM CaCl$_2$, were heated from 40 to 80°C in water bath. After each 30 minutes 10$\mu$l enzyme aliquot was taken and mixed with 1.0 ml of substrate. Enzyme activity was measured at 440 nm for 5.0 minutes in a thermo-regulated chamber of DU. 68 spectrophotometer maintained at 25°C and residual activity was expressed as a percent of the initial activity.
Cloning strategy for PCR-amplification & sequencing of the subtilisin RT-5 gene

Isolation of chromosomal DNA of B. subtilis RT-5 by phenol:chloroform extraction method

Run PCR-amplification reaction with two degenerate primers

Purification of PCR fragment of 670 bp by low melting agarose gel

Inserted into Eco R1 digested pEMBL8 vector

Transformed into competent E.coli TG1 cells

Transformants were selected on ampicillin containing plates and identified by X-gal/IPTG screening system

Plasmid DNA containing inserted DNA segment was isolated and digested with Eco R1.

Run on gel to see the presence of same segment.

Direct DNA sequencing by Sanger's dideoxy chain termination method by using T7 DNA seq. kit(Pharmacia)
2-21.1 Isolation of chromosomal DNA of *B. subtilis* RT-5

Chromosomal DNA from *Bacillus subtilis* Strain RT was prepared according to the method of Thomas et al. (1966). To isolate DNA cells were grown in L.B medium for 18hr at 37°C. The cells were harvested and washed twice by SSC. Dilute suspension of cells in SSC was treated with lysozyme in a conc. of 4mg/ml. Proteinase K was added to the mixture to a final conc. of 50 μg/ml. Then 1% SDS was added to the mixture and incubated at 37°C for 2 hours.

2-21.2 Phenol : Chloroform extraction

Equal volume of 1:1 mixture of Phenol : Chloroform was added to the suspension and it was rolled for 30 minutes at room temperature at 60 rev/min. The mixture was chilled to 0°C and centrifuged at 5000 rpm for 15 minutes. The aqueous phase was removed from the phenol : chloroform mixture and extraction was repeated thrice.

2-21.3 RNA elimination by Ribonuclease digestions

Prior to dialysis the preparation was treated with 20μg/ml of pancreatic ribonuclease.

2-21.4 Removal of Phenol

To remove the phenol from the preparation it was dialyzed against 2x2L of T.C.buffer pH 8.0 for overnight at 4°C.
2-21.5 Precipitation of DNA

To 9.0 ml of aqueous phase 1.0 ml of 3.0M Na-acetate pH 5.0 was added. 2 volume of chilled ethanol was added and mixed. The mixture was kept at -30°C for 15 minutes. Centrifuged at 15,000 rpm for 15 min. The precipitate thus obtained was washed with 70% ethanol. Ethanol was carefully decanted and the pellet was air dried. The pellet was dissolved in 1.0 ml of T.E.buffer pH 8.0. The purity of the extracted DNA and its molecular mass were checked on agarose gel.

2-22 Oligonucleotides used for PCR-amplification

Based on the determined amino acid sequences two degenerate primers were designed: one spanning the region of the gene encoding Ala 1 to Tyr 6 and the other spanning the active site region from the residue Asn 218 to Ala 223. We decided to make the upstream primer at the active site region because this is the most conserved region in all the previously sequenced subtilisins. We decided to create an EcoR1 restriction sites on 5' ends of both primers to facilitate the cloning of PCR product to the vector. The sequences of downstream and upstream primers are as under:

downstream primer: Ala 1 to Tyr 6
5'-GG GAATTC GCG CAA TCT GTT CCT TAT -3'

upstream primer: Ala 223 to Arg 218
5' GG GAATTC CGC GAT GGA CGT ACC GTT-3'

37
1.0 μg of *Bacillus subtilis* strain RT-5 genomic DNA was used for PCR-amplification reaction in a 50 μl volume. The final conc. of Mg\(^{2+}\) in the reaction mixture was adjusted to 2.5. Other components of the reaction mixture were present in the following concentrations: 20mM Tris/HC1, pH 8.8, deoxynucleotides (200μM). 1.0μl(2.5u) of Taq polymerase was added to each tube. Mixed and then two drops of mineral oil was layered on the top of reaction mixture and subjected to 30 PCR cycles. The terminal cycling profile used was as follows: denaturation at 94°C for 45 seconds, primer annealing at 56°C for 45 seconds and primer extension at 72°C for 2 min.

2-24 Cloning of PCR-amplified subtilisin RT-5 gene

The product DNA obtained from PCR-amplification reaction was analyzed by agarose gel electrophoresis. The amplified DNA fragment of approx. 670 bp was excised from the gel and purified using Promega Magic PCR Preps DNA purification kit. The purified DNA Fragment was inserted into the Eco R1 digested pEMBL8 vector and transformed into *E.coli* TG1 cells made competent with CaCl\(_2\). Transformants were selected on L.B agar plates containing X-gal, IPTG and ampicillin. The plasmid preparations from the transformants cells were obtained using Promega Miniprep. Presence of the desired insert was reconfirmed by Eco R1 digestion of an aliquot of plasmid prep and run on 1% agarose gel.
DNA sequencing was performed using the dideoxy nucleotide termination method of Sanger et al. (1977) using Pharmacia T7 DNA-sequencing kit. Determined nucleotide sequence was matched with the previously determined amino acid sequences.
3-0 RESULTS
Initial screening of soil samples for proteolytic bacterial strains

Isolation of strain RT-5 used in this study was done on the basis of colonial morphology and production of a large zone of hydrolysis around each colony on casein agar. The isolated bacteria were purified on nutrient agar plates.

3-2 Strain Identification

On the basis of its colonial morphology, Gram reaction and other cultural and biochemical characteristics strain RT-5 was identified as *Bacillus subtilis*. Identification was confirmed by using API 20 E and CHB 50 system. The strains selected for further study were GBS01 and RT-5.

The growth characteristics, colonial morphology, Gram reaction and other biochemical properties of the two strains are given in Table 3-1.

3-3 Comparison of specific growth rates of the *Bacillus subtilis* strain GBS01 on various media

1) Effect of varying concentrations of beef extract on the growth of *Bacillus subtilis* strain GBS01 is shown in fig. 3-1. It is observed that in presence of only 1% beef extract almost as much cell mass can be obtained as in a medium containing 3%. Cell numbers have been quantitated by calculating specific growth rate constants, "k" (See appendix for the definition and calculation of the specific growth rate constant, "μ"). Values of this parameter were identical i.e., 0.31⁻⁻ for all the three beef extract based compositions i.e., media # 1, 2 and 3.
2) Growth in tryptose and/or gelatin in the absence of beef extract is shown in fig. 3-2. Medium # 4 in which beef extract was replaced by 5% tryptose and 5% gelatin was found to be suitable for the cultivation of strain GBS01. Values of \( \mu \) were 0.30\(^{-h}\) and 0.26\(^{-h}\) for media # 4 and 5 respectively.

3) Growth rates achieved in presence of various concentrations of gelatin ranging from 20.0% to 2% are shown in fig. 3-3. It can be seen that medium # 7 containing 10% gelatin and 0.2% glycerol is as good as medium # 4 containing 5% tryptose and 5% gelatin. Values of \( \mu \) were 0.20\(^{-h}\), 0.19\(^{-h}\), 0.15\(^{-h}\) and 0.12\(^{-h}\) for media # 6, 7, 8 and 9 respectively.

4) Effect of addition of yeast extract and/or corn steep liquor to 10% gelatin medium on the growth rate is shown in fig. 3-4. When glycerol was omitted from the medium the stationary phase was shortened and a sharp decline in viable cell count was observed (as in medium # 13). Values of \( \mu \) were 0.21\(^{-h}\), 0.23\(^{-h}\), 0.22\(^{-h}\) and 0.22\(^{-h}\) for media # 10, 11, 12 and 13 respectively.

5) Effect of addition of vitamin B1 to 10% gelatin medium with or without M9 salts is shown in fig. 3-5. Almost equal number of cells were obtained in media # 14, 15 and 16, the values of \( \mu \) being 0.16\(^{-h}\) respectively. The highest proteolytic activity, however, was obtained when both the M9 salts and B1 were present (as in medium # 14).

6) Effect of presence of various salts on the growth of strain GBS01 is shown in fig. 3-6. It can be seen that the addition of divalent cations enhances the growth rate. Values of \( \mu \) were 0.20\(^{-h}\), 0.19\(^{-h}\) and 0.19\(^{-h}\) for media # 17, 18 and 19 respectively.
7) Effect of varying magnesium ions is shown in fig. 3-7. Magnesium sulphate decreased the growth rate of strain GBS01. Values of \( \mu \) were 0.23-\(^{-h}\), 0.23-\(^{-h}\) and 0.19-\(^{-h}\) for media # 20, 21 and 22 respectively.

3-4 Purification of proteases

3-4-1 Ammonium sulphate fractionation of cell free broth

Cell free broth was fractionated with 35% and then with 75% ammonium sulphate. Proteolytic activity was found to be in 75% ammonium sulphate precipitated fraction. Proteolytic activity was monitored against casein as a substrate and expressed in Kunitz's units.

3-5 Separation and characterization of proteases

Proteolytic enzymes in 75% ammonium sulphate precipitated fraction were separated in two portions through DEAE-Sepharose fast flow chromatography. Flow-through preparation contained serine protease as judged by the inhibition of proteolytic activity by 1.0 mM PMSF and not by 10mM EDTA while the bound fraction contained neutral protease, which was inhibited by 10 mM EDTA but not by 1.0mM PMSF (Fig.3-8).

3-6 Purification of alkaline serine protease (Subtilisin RT-5)

Concentrated unbound preparation from DEAE-Sepharose fast flow chromatography was further fractionated on Mono-S Fast Protein Liquid Chromatography (FPLC) column. Fig.3-9 shows the elution profile with 6 major peaks. Peak 2 contained the proteolytic activity as judged by the Kunitz's method by using casein as a substrate. Rechromatography of this peak on the same column and under the
same conditions but with shallow gradient resulted in 2 major
peaks (Fig.3-10). Proteolytic activity was found in peak 1 (Fig.
3-10) while peak 2 was not active.

3-7 Estimation of molecular mass

Purity of peak 1 of Fig.3-10 was checked by SDS-PAGE. The peak
was resolved as a single band on SDS-PAGE phast gel system and
was judged to be homogeneous (upto 95%) having a molecular
mass of approximately 30 kDa on 12.5% gel concentration
(Fig.3-11).

3-8 N-terminal sequence analysis

Intact protein was sequenced on gas phase and solid phase
sequencers. 25 N-terminal residues were identified on gas phase
sequencer, whereas on the solid phase sequencer the intact
protein was sequenced upto 43 residues. The N-terminal residues
identified from both the runs were identical.

3-9 Chemical cleavage of subtilisin

Chemical cleavage of intact protein was achieved by cyanogen
bromide (CNBr). Peptides were fractionated by gel permeation
chromatography through Superose 12 FPLC (Fig.3-12) column and
then rechromatographed by HPLC (Fig.3-13 and 3-14) using Vydac
C4 and Ultracap C18 reversed phase columns and subjected to
automatic solid phase sequencing. CNBr digested peptides were
pooled as pool A and pool B (Fig.3-13).

3-10 Enzymatic fragmentation and purification of peptides

Redigestion of CNBr digested pool A and pool B peptides with
TPCK-trypsin, Lys-C and Asp-N proteases and separation of the
resulting peptides by reversed phase HPLC on Ultrapac C18 column provided well resolved peptide map.

Elution profiles of tryptic digested pool A and pool B, Lys-C and Asp-N digested peptides are shown in Fig.3-15 and Fig.3-16, Fig.3-17 and Fig.3-18 respectively. The peptides were recovered and sequenced as shown in Fig.3-19.

3-11 Amino acid analysis

The amino acid composition of the intact protein shown in table 3-2, closely resembles those of the known subtilisins. Table 3-4 shows the amino acid compositions of tryptic peptides generated from pool A and pool B of CNBr digested peptides. The amino acid compositions of peptides are in good agreement with their sequence analyses.

3-12 Primary structure of the subtilisin RT-5

Primary structure of the subtilisin RT-5 was deduced by sequence analysis of the intact protein, and the peptides from CNBr, TPCK-trypsin, Lys-C and Asp-N protease digestions. Complete amino acid sequence of the protein is shown in Fig.3-19.

It consists of 275 amino acid residues. However, cystein residue and consequently disulphide bonds are missing in the enzyme. Residues of the active site triad i.e., Asp, His and Ser are located at positions 32, 64, 221 respectively. Four metionine residues are all conserved at positions 119,125,199 and 222.

3-13 Enzyme Inhibition assay

The results reported in Table 3-5 show that the subtilisin RT-5 is a serine protease which is inhibited by 1mM PMSF and 0.1mM
3,4-DCI but not by 10mM EDTA while the metalloprotease was inhibited by 10mM EDTA but not by 1mM PMSF.

3-14 Measurement of $k_{cat}/K_m$

The catalytic efficiency, $k_{cat}/K_m$, of subtilisin RT-5 was measured by using Suc-AAPF-pNA as a substrate. Results indicated that the higher catalytic efficiency of subtilisin RT-5 is due to the increase in $k_{cat}$ (catalytic rate constant) whereas $K_m$ (the binding efficiency) remains somewhat unaltered as compared to other subtilisins.

The hydrolysis of the synthetic substrate Suc-AAPF-pNA by subtilisin BPN' is characterized by values for $k_{cat}$ and $K_m$ of 57s$^{-1}$ and 0.15 mM respectively (at 25°C in 0.1 M Tris-HCl, pH 8.6). The values of $k_{cat}$ and $K_m$ for Carlsberg enzyme are 938s$^{-1}$ and 0.234 mM. Under the same conditions, the values for the $k_{cat}$ and $K_m$ of subtilisin RT-5 are 2000s$^{-1}$ and 0.10 mM.

3-15 Thermostability of subtilisin RT-5

Thermostability of the enzyme was checked by using succinyl Ala-Ala-Pro-Phe p-nitroanilide as substrate in 0.1 M Tris-HCl, pH 8.5, 10mM CaCl$_2$ at 40-80°C. Results indicate that the subtilisin RT-5 is thermostable (retains 70% activity) even after 3h of incubation at 50°C (fig. 3-24).

3-16 PCR-cloning and DNA sequence analysis

Using PCR amplification we were able to clone most of the subtilisin gene encoding residues Ala1 to Ala223. DNA sequence analysis was matched with the already determined primary structure of the enzyme. It provided the DNA sequence of that segment of the polypeptide chain which could not be recovered in protein digests. It revealed the presence of four methionine
residues at positions 119, 125, 199 & 222 and two Asn-Gly regions at Asn109-Gly110 and Asn218-Gly219 respectively. The determined DNA sequence of subtilisin gene is shown in fig. 3-25.
Fig. 3-1 Growth rate of *Bacillus subtilis* strain GBS01 in media containing different concentrations of beef extract. Culture was grown in media containing indicated quantities of beef extract at 37°C for 96 hours. Extent of growth was determined by measuring the titre of culture in CFU/ml.

- medium # 1 [(B.E.) 3%]
- medium # 2 [(B.E.) 2%]
- medium # 3 [(B.E.) 1%]
B.E. = Beef extract.
Fig. 3-2. Growth rate of Bacillus subtilis strain GBS01 in media containing tryptose and gelatin. Culture was grown in media containing indicated quantities of tryptose and gelatin at 37°C for 72 hours. Extent of growth was determined by measuring the titre of culture in CFU/ml.

- medium # 1 [(B.E.) 3%]
- medium # 4 [tryptose 5% + gelatin 5%]
- medium # 5 [tryptose 5%]
Fig. 3-3. Growth rate of *Bacillus subtilis* strain GBS01 in media containing different concentrations of gelatin. Culture was grown at 37°C for 72 hours. Extent of growth was determined by measuring the titre of culture in CFU/ml.

medium # 6 [gelatin 20%]
medium # 7 [gelatin 10%]
medium # 8 [gelatin 5%]
medium # 9 [gelatin 2%]
Fig. 3-4. Growth rate of *Bacillus subtilis* strain GBS01 in media containing corn steep liquor and/or yeast extract. Culture was grown in media containing indicated quantities of corn steep liquor or yeast extract at 37°C for 72 hours. Extent of growth was determined by measuring the titre of culture in CFU/ml.

medium # 10 [gelatin 10% + glycerol 0.2% + CSL 0.4% + Y.E. 0.1%]
medium # 11 [gelatin 10% + glycerol 0.2% + CSL 0.4%]
medium # 12 [gelatin 10% + glycerol 0.2% + Y.E. 0.1%]
medium # 13 [gelatin 10% + CSL 0.1%]
CSL = Corn steep liquor
Y.E. = Yeast extract
Fig. 3-5. Growth rate of *Bacillus subtilis* strain GBS01 in media containing M9 salts and/or B1. Culture was grown in media containing M9 salts and/or B1 at 37°C for 72 hours. Extent of growth was determined by measuring the titre of culture in CFU/ml.

medium # 11 [gelatin 10% + glycerol 0.2% + CSL 0.4%]
medium # 14 [medium 11 + M9 salts + B1]
medium # 15 [medium 11 + M9 salts]
medium # 16 [medium 11 + B1]
Fig. 3-5. Growth rate of *Bacillus subtilis* strain GBS01 in media containing different salts. Culture was grown in media containing salts at 37°C for 72 hours. Extent of growth was determined by measuring the titre of culture in CFU/ml.

medium # 17 [see section 2-4]
medium # 18 [see section 2-4]
medium # 19 [see section 2-4]
medium # 16
Fig. 3-7. Growth rate of *Bacillus subtilis* strain GBS01 in media containing different concentrations of magnesium sulphate. Culture was grown in media containing indicated quantities of magnesium sulphate at 37°C for 72 hours. Extent of growth was determined by measuring the titre of culture in CFU/ml.

medium # 20 [MgSO₄·7H₂O 2.0 mM]
medium # 21 [MgSO₄·7H₂O 1.0 mM]
medium # 22 [MgSO₄·7H₂O 0.5 mM]
medium # 11
Extracellular broth

→

Fractionation with 75% ammonium sulphate

→

Dialysis of precipitate against 10 mM Tris-HCl, pH 7.5, 5mM CaCl₂

→

DEAE-Sepharose chromatography

\[\begin{align*}
\text{Flow-through} \\
\text{SERINE PROTEASE} \\
(\text{Inhibited by 1mM PMSF} \\
\text{and 0.1 mM 3,4-DCI})
\end{align*}\]  

\[\begin{align*}
\text{Bound} \\
\text{METALLOPROTEASE} \\
(\text{Inhibited by 10 mM EDTA})
\end{align*}\]

→

Mono-S FPLC \\
(10 mM PO₄, pH 6.0)

→

Rechromatography on the same column and under the same conditions

→

Obtained pure enzyme

Fig. 3-8 Schematic presentation of purification protocol for proteases.
1 Elution profile of subtilisin RT-5 on cation exchanger, Mono-S FPLC. Peak 2 contained proteolytic enzyme. Fractions were eluted with 0-1.0 M NaCl in 10mM PO₄ buffer, pH 6.0.
Fig. 3-10  Rechromatography of peak 2, fig.3-9 on Mono-S FPLC, with a linear gradient of 0-0.2 M NaCl in 10mM PO₄ buffer, pH 6.0. Peak 1 contained proteolytic enzyme.
Fig. 3-11 SDS-PAGE of peak 1, fig.3-10, showing purified subtilisin RT-5 on 12.5% gel. Left lane: standard proteins Mr 94000 (top), 67000, 43000, 30000, 21000, 14400 (bottom); right lane: subtilisin RT-5.
Fig. 3-12 Gel permeation chromatography of CNBr digested subtilisin RT-5 on Superose 12 FPLC column in 30% acetic acid.
containing 0.1% TFA, CBI & CBS denatured cyanogen bromide peptides.

Reversed phase HPLC of CNIIR digested subtilisin RT-5 on Vydac C4 column in 0.1% TFA with a linear gradient of acetonitrile.

**Fig. 3.13**

**TIME (minutes)**

**ABSORBANCE AT 214 nm**
TF A. CB3 denatures cyanogen bromide peptide.

Fig. 3.14

Reversed phase HPLC of peak 1, Fig. 3.12 on Ultrace C18 column in 0.1% TF A with a linear gradient of acetonitrile containing 0
FIG. 3-15

Reversed phase HPLC of in vitro digested pool A of Hg-3-13 on Ultrasphere C18 column in 0.1% TFA with a linear gradient acetonitrile containing 0.1% TFA, 75:19 denotes typtic peptides.
Fig. 3-16
Reversed phase HPLC of tryptic digested pool B of IgG-3-13 on Ultrapac C18 column in 0.1% TFA with a linear gradient containing 0.1% TFA. T1-T4 denotes tryptic peptides.

TIME (minutes)

ABSORBANCE AT 214 nm
Reversed phase HPLC of Lys-C digested peptides on Ultrapac C18 column in 0.1% TFA with a linear gra
Fig. 3.18

Reversed phase HPLC of Asp-N digested peptides on Ultrapac C8 column in 0.1% TFA with a linear gradient.
Fig. 3-19 Primary structure of subtilisin RT-5 from *Bacillus subtilis* strain RT-5. The positions of peptide analyzed are also shown. "T", "K", "CND" and "CN" indicates peptides analyzed from trypsin, Lys-C, Asp-N and CNBr digestions. Unmarked continuous line shows the extent of N-terminal sequence analysis of the intact polypeptide chain. GAC AAA to CAA AGA are the nucleotide sequence of the peptides confirmed by DNA sequence analysis.
Fig. 3.20 Effect of CaCl₂ on the activity of subtilisin RT-5 towards succinyl Ala-Ala-Pro-Phe p-nitroanilide.
Fig. 3-21 Effect of pH on $k_{cat}/Km$ of subtilisin RT-5. Conditions are described in section 2-19.
Reversed phase HPLC of subtilisin RT-5 digested horse liver alcohol dehydrogenase on Ultrasphere C18 column in 0.1% TFA with linear gradient of acetonitrile containing 0.1% TFA.

**Figure 3.22**
Fig. 3-23 Cleavage sites of subtilisin RT-5 with horse liver alcohol dehydrogenase as a substrate.
Fig 3-24 Thermostability of the subtilisin RT-5 at 50°C
3-25 DNA sequence of subtilisin gene.
<table>
<thead>
<tr>
<th>Biochemical tests</th>
<th>Strain GBS01</th>
<th>Strain RT-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalase</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Voges-Proskauer</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Egg yolk</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Growth at pH 5.7</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Growth in 5% NaCl</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Growth in 7% NaCl</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Growth in 10% NaCl</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Growth at 45°C</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Growth at 50°C</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Growth at 55°C</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Growth at 60°C</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Acid from:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>L-Arabinose</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>D-Xylose</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>D-Mannitol</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Gas from glucose</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hydrolysis of starch</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Citrate utilization</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Nitrate reduction</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Casein hydrolysis</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gelatin liquefaction</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Table 3-1 A Test of the API system used to identify the species of Strain RT-5

<table>
<thead>
<tr>
<th>Test</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Nitrophenyl-b-D-galactopyranoside</td>
<td>+</td>
</tr>
<tr>
<td>Arginine dihydrodase</td>
<td>-</td>
</tr>
<tr>
<td>Lysine decarboxylase</td>
<td>-</td>
</tr>
<tr>
<td>Ornithine decarboxylase</td>
<td>-</td>
</tr>
<tr>
<td>Simmons citrate</td>
<td>+</td>
</tr>
<tr>
<td>Hydrogen sulphide</td>
<td>-</td>
</tr>
<tr>
<td>Urease</td>
<td>+</td>
</tr>
<tr>
<td>Tyrotophan deaminase</td>
<td>-</td>
</tr>
<tr>
<td>Indole</td>
<td>-</td>
</tr>
<tr>
<td>Voges-Proskauer</td>
<td>+</td>
</tr>
<tr>
<td>Gelatin liquefaction</td>
<td>+</td>
</tr>
<tr>
<td>Nitrate reduction</td>
<td>+</td>
</tr>
<tr>
<td>Control</td>
<td>+</td>
</tr>
<tr>
<td>Glycerol</td>
<td>+</td>
</tr>
<tr>
<td>Erythritol</td>
<td>-</td>
</tr>
<tr>
<td>D-Arabinose</td>
<td>-</td>
</tr>
<tr>
<td>L-Arabinose</td>
<td>-</td>
</tr>
<tr>
<td>Ribose</td>
<td>+</td>
</tr>
<tr>
<td>D-Xylose</td>
<td>-</td>
</tr>
<tr>
<td>L-Xylose</td>
<td>-</td>
</tr>
<tr>
<td>Adonitol</td>
<td>-</td>
</tr>
<tr>
<td>β-Methyl-xyloside</td>
<td>-</td>
</tr>
<tr>
<td>Galactose</td>
<td>-</td>
</tr>
<tr>
<td>D-Glucose</td>
<td>+</td>
</tr>
<tr>
<td>D-Fructose</td>
<td>+</td>
</tr>
<tr>
<td>D-Mannose</td>
<td>+</td>
</tr>
<tr>
<td>L-Sorbose</td>
<td>-</td>
</tr>
<tr>
<td>Rhamnose</td>
<td>-</td>
</tr>
<tr>
<td>Dulcitol</td>
<td>-</td>
</tr>
<tr>
<td>Inositol</td>
<td>-</td>
</tr>
<tr>
<td>Mannitol</td>
<td>+</td>
</tr>
<tr>
<td>Sorbitol</td>
<td>+</td>
</tr>
<tr>
<td>α-Methyl-D-mannoside</td>
<td>-</td>
</tr>
<tr>
<td>α-Methyl-D-glucoside</td>
<td>+</td>
</tr>
<tr>
<td>N-Acetyl-glucosamine</td>
<td>-</td>
</tr>
<tr>
<td>Amygdalone</td>
<td>+</td>
</tr>
<tr>
<td>Arbutin</td>
<td>+</td>
</tr>
<tr>
<td>Esculine</td>
<td>+</td>
</tr>
<tr>
<td>Salicin</td>
<td>+</td>
</tr>
<tr>
<td>Cellubiose</td>
<td>+</td>
</tr>
<tr>
<td>Maltose</td>
<td>+</td>
</tr>
<tr>
<td>Lactose</td>
<td>-</td>
</tr>
<tr>
<td>Melibiose</td>
<td>+</td>
</tr>
<tr>
<td>Saccharose</td>
<td>+</td>
</tr>
<tr>
<td>Trehalose</td>
<td>+</td>
</tr>
<tr>
<td>Inuline</td>
<td>+</td>
</tr>
<tr>
<td>Melezitose</td>
<td>-</td>
</tr>
<tr>
<td>D-Raffinose</td>
<td>+</td>
</tr>
<tr>
<td>Amidon</td>
<td>+</td>
</tr>
<tr>
<td>Glycogene</td>
<td>+</td>
</tr>
<tr>
<td>Xylitol</td>
<td>-</td>
</tr>
<tr>
<td>β-Gentiobiose</td>
<td>+</td>
</tr>
<tr>
<td>D-Turanose</td>
<td>+</td>
</tr>
<tr>
<td>D-Lybose</td>
<td>-</td>
</tr>
<tr>
<td>D-Tagatose</td>
<td>-</td>
</tr>
<tr>
<td>D-Fucose</td>
<td>-</td>
</tr>
<tr>
<td>L-Fucose</td>
<td>-</td>
</tr>
<tr>
<td>D-Arabinol</td>
<td>-</td>
</tr>
<tr>
<td>L-Arabinol</td>
<td>-</td>
</tr>
<tr>
<td>Gluconate</td>
<td>+</td>
</tr>
<tr>
<td>2-ceto-gluconate</td>
<td>-</td>
</tr>
<tr>
<td>5-ceto-gluconate</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 3-2  Amino acid composition of the subtilisin RT-5.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>mole/mole</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>26.1 (26)</td>
</tr>
<tr>
<td>Thr</td>
<td>18.5 (19)</td>
</tr>
<tr>
<td>Ser</td>
<td>30.2 (39)</td>
</tr>
<tr>
<td>Glx</td>
<td>15.9 (16)</td>
</tr>
<tr>
<td>Pro</td>
<td>11.0 (11)</td>
</tr>
<tr>
<td>Gly</td>
<td>32.5 (33)</td>
</tr>
<tr>
<td>Ala</td>
<td>34.1 (34)</td>
</tr>
<tr>
<td>Val</td>
<td>24.5 (25)</td>
</tr>
<tr>
<td>Met</td>
<td>3.8 (4)</td>
</tr>
<tr>
<td>Ile</td>
<td>15.7 (16)</td>
</tr>
<tr>
<td>Leu</td>
<td>14.8 (15)</td>
</tr>
<tr>
<td>Tyr</td>
<td>11.6 (12)</td>
</tr>
<tr>
<td>Phe</td>
<td>3.2 (3)</td>
</tr>
<tr>
<td>His</td>
<td>5.8 (6)</td>
</tr>
<tr>
<td>Lys</td>
<td>8.3 (8)</td>
</tr>
<tr>
<td>Arg</td>
<td>4.2 (4)</td>
</tr>
<tr>
<td>Trp</td>
<td>n.d.</td>
</tr>
</tbody>
</table>

Values represent moles of a given amino acid present per mole of protein. These values have been obtained by the standard acid hydrolysis procedure. However, values within parenthesis are from the sum of the protein sequence data.
Table 3-3  Comparison of amino acid compositions of subtilisin RT-5 with other subtilisins.

<table>
<thead>
<tr>
<th></th>
<th>RT-5</th>
<th>BPN'</th>
<th>S 168</th>
<th>SCB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>26</td>
<td>26</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Thr</td>
<td>19</td>
<td>13</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>Ser</td>
<td>39</td>
<td>37</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>Glx</td>
<td>16</td>
<td>15</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>Pro</td>
<td>11</td>
<td>14</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Gly</td>
<td>33</td>
<td>33</td>
<td>32</td>
<td>35</td>
</tr>
<tr>
<td>Ala</td>
<td>34</td>
<td>37</td>
<td>30</td>
<td>41</td>
</tr>
<tr>
<td>Val</td>
<td>25</td>
<td>30</td>
<td>22</td>
<td>31</td>
</tr>
<tr>
<td>Met</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Ile</td>
<td>16</td>
<td>13</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Leu</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Tyr</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Phe</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>His</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Lys</td>
<td>8</td>
<td>11</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Arg</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Trp</td>
<td>n.d.</td>
<td>3</td>
<td>n.d.</td>
<td>1</td>
</tr>
</tbody>
</table>

RT-5 = subtilisin RT-5  
BPN' = subtilisin BPN'/Novo  
S 168 = subtilisin from Marburg strain 1-168  
SCB = subtilisin Carlsberg
<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T6</th>
<th>T8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>-</td>
<td>1.2(1)</td>
<td>4.0(4)</td>
<td>3.7(4)</td>
<td>-</td>
<td>2.1(2)</td>
</tr>
<tr>
<td>Thr</td>
<td>-</td>
<td>1.3(1)</td>
<td>-</td>
<td>2.5(3)</td>
<td>1.0(1)</td>
<td>1.7(2)</td>
</tr>
<tr>
<td>Ser</td>
<td>2.4(2)</td>
<td>2.4(2)</td>
<td>2.8(3)</td>
<td>7.6(8)</td>
<td>1.4(1)</td>
<td>2.2(2)</td>
</tr>
<tr>
<td>Glx</td>
<td>2.4(2)</td>
<td>1.4(1)</td>
<td>-</td>
<td>2.1(2)</td>
<td>-</td>
<td>1.3(1)</td>
</tr>
<tr>
<td>Pro</td>
<td>0.8(1)</td>
<td>0.7(1)</td>
<td>0.6(1)</td>
<td>2.5(3)</td>
<td>0.6(1)</td>
<td>-</td>
</tr>
<tr>
<td>Gly</td>
<td>1.4(1)</td>
<td>2.3(2)</td>
<td>1.4(1)</td>
<td>6.5(7)</td>
<td>1.3(1)</td>
<td>2.4(2)</td>
</tr>
<tr>
<td>Ala</td>
<td>1.1(1)</td>
<td>1.8(2)</td>
<td>1.0(1)</td>
<td>5.6(6)</td>
<td>4.5(5)</td>
<td>1.3(1)</td>
</tr>
<tr>
<td>Val</td>
<td>0.9(1)</td>
<td>1.3(1)</td>
<td>2.6(3)</td>
<td>4.5(5)</td>
<td>1.2(1)</td>
<td>-</td>
</tr>
<tr>
<td>Met</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ile</td>
<td>1.8(2)</td>
<td>-</td>
<td>1.5(2)</td>
<td>1.7(2)</td>
<td>1.0(1)</td>
<td>-</td>
</tr>
<tr>
<td>Leu</td>
<td>-</td>
<td>1.2(1)</td>
<td>1.2(1)</td>
<td>2.5(3)</td>
<td>1.8(2)</td>
<td>1.8(2)</td>
</tr>
<tr>
<td>Tyr</td>
<td>1.0(1)</td>
<td>1.1(1)</td>
<td>-</td>
<td>1.5(2)</td>
<td>-</td>
<td>2.5(3)</td>
</tr>
<tr>
<td>Phe</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.7(1)</td>
<td>-</td>
<td>0.5(1)</td>
</tr>
<tr>
<td>His</td>
<td>-</td>
<td>0.8(1)</td>
<td>1.2(1)</td>
<td>1.6(2)</td>
<td>1.0(1)</td>
<td>-</td>
</tr>
<tr>
<td>Lys</td>
<td>1.0(1)</td>
<td>1.0(1)</td>
<td>-</td>
<td>1.0(1)</td>
<td>1.0(1)</td>
<td>1.0(1)</td>
</tr>
<tr>
<td>Arg</td>
<td>-</td>
<td>1.0(1)</td>
<td>-</td>
<td>-</td>
<td>0.9(1)</td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Values represent moles of a given amino acid present per mole of protein. These values have been obtained by the standard acid hydrolysis procedure. However, values within parenthesis are from the sum of the protein sequence data.
Table 3-5  Effects of EDTA, PMSF and 3,4-DCI on the proteolytic activity of subtilisin RT-5.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Serine protease</th>
<th>Metalloprotease</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>PMSF (1mM)</td>
<td>2</td>
<td>93</td>
</tr>
<tr>
<td>3,4-DCI (0.1mM)</td>
<td>1</td>
<td>n.d</td>
</tr>
<tr>
<td>EDTA (10mM)</td>
<td>80</td>
<td>3</td>
</tr>
</tbody>
</table>

The enzyme preparations were preincubated in the presence of 1mM PMSF, 0.1 mM 3,4-Dichloroisocumarin or 10 mM EDTA for 30 minutes at 37°C, after which remaining casein hydrolytic activities were monitored.
4-0 DISCUSSION
Bacterial proteases have multifarious industrial as well as laboratory applications ranging from non-specific degradation of proteins in leather tanning and brewing to specific cleavages of particular peptide bonds in protein sequencing.

The genesis of the work described here lies in our efforts to develop a cost-effective medium for industrial scale production of extracellular proteases from *Bacillus subtilis* by a group of local entrepreneurs who have set up the first Pakistani factory for commercial production of bating enzymes used in leather processing.


As the factory started initial production the medium cost became prohibitive from a commercial point of view. It transpired that the medium being used by the factory contained imported beef extract and tryptose along with gelatin. The problem was further compounded by the fact that the enzyme yield was not particularly high either. Moreover, the factory had availed only one bacterial strain producing a neutral protease.

Hence this project was initiated with the objectives to (a) formulate a medium for efficient and cost effective production of proteolytic enzymes on industrial scale, (b) to isolate and develop bacterial strains producing proteolytic enzyme(s) with industrially useful attributes, and c) to purify and characterize the enzyme(s).

As microbiologists we recognized at the very outset that *Bacillus subtilis*
being a saprophyte can grow well in a variety of nutritional milieux with diverse sources of proteins. Hence we argued that elimination of beef extract and tryptose from the medium should not notably curtail enzyme production as long as gelatin was available. Subsequently, a variety of media were formulated from which beef extract and tryptose were omitted but gelatin was retained as it is available locally at a reasonable price. In order to test the efficacy of these formulations, a series of growth curve experiments were done to monitor protease production and cell growth. As can be seen from the data in Table 4-1, media # 10 & 11 containing gelatin supported a level of cell growth and enzyme production that was favorably comparable to that obtained in media # 1, 2, & 3.

Further reduction in cost was achieved by replacing tryptose with corn steep liquor as it is a cheap and locally produced on a mass scale.

Being aware of the fact that certain enzymes of *Bacillus subtilis* are subject to catabolite repression by glucose or its metabolic products, e.g. organic acids (Haavik, 1974a,b; Laishley and Bernlohr, 1966; Martin and Demain, 1980) we used glycerol as a carbon source in the medium. Glycerol is not easily degradable and needs a longer period of time to be exhausted from the culture medium; hence it keeps the level of catabolite repressors lower and promotes protease production. Glycerol also serves as a buffering agent and thereby helps in maintaining the enzyme activity.

Halon and Hodges (1981a) have reported that production of bacitracin and protease are dependent on the growth rate of the culture. Similarly it has been shown that the production of gramicidin S by *Bacillus brevis*, and of $\alpha$-amylase and protease from *Bacillus licheniformis* is related to the specific growth rate of the culture (Matteo et al., 1976; Meers, 1972; Wouters and Buysman, 1977). It has also been indicated (Halon and Hodges, 1981b) that catabolite repression by glucose is due to the increased growth rate resulting in glucose depletion and accumulation of toxic products in the growth medium. These findings are confirmed by our
results as well: culture growing in medium # 14 containing glycerol, M9 salts and vitamin B1 had a low specific growth rate i.e. 0.16⁻¹⁻ while it had the highest proteolytic activity.

Wouters and Buysman (1977) reported that the production of proteases and other degradative enzymes like alkaline phosphatase and ribonuclease from Bacillus licheniformis 749/C is under the control of some kind of induction-repression system. They suggested that a specific catabolite or an end product repressible system is responsible for the production of exoenzymes under certain conditions of nutritional limitations. They also indicated that there is an inverse correlation between the specific growth rate and the production of enzymes susceptible to catabolite repression and that the production of proteases and other degradative enzymes was derepressed only at the lower growth rates of Bacillus licheniformis 749/C.

In our experiments as well a correlation has been noted between the specific growth rates and the protease production. It is indicated by the observation that cultures with low specific growth rates have high proteolytic activity.

It has been reported (Frankena et al., 1985) that at the specific growth rate of 0.22⁻¹⁻ in glucose-limited chemostat cultures maximum rate of protease production was obtained. Conversely, at the specific growth rate values of above 0.22, protease production was decreased by glucose repression. Similar results were obtained by Mao et al. (1992) indicating that the lowest initial concentration of glucose (6g/l) yields the highest protease production. They also indicated that a certain quantity of cell mass is necessary to obtain high protease production. But even in the presence of high amount of cell mass lower protease production was obtained indicating that high enzyme production depends not only upon a certain number of cells at a particular time period in the culture medium but also upon the appropriate fermentation conditions and the composition of the
medium used.

During the course of our experiments it was noted that the yield of proteolytic enzyme from *Bacillus subtilis* strain GBS01 was rather low. We, therefore, initiated a screening programme to isolate new strains which could yield higher quantities of enzymes. Hence soil samples were collected from the Tharparkar desert of Pakistan and highly proteolytic strains were screened for and isolated on the basis of production of large zones of hydrolysis around the colonies on casein agar plates. A thermotolerant *Bacillus subtilis* strain RT-5 displaying high proteolytic activity was chosen for further characterization. This strain was found to produce an alkaline serine protease as well as a neutral protease in addition to an amylase. The present study was focussed on the isolation, characterization and sequence analysis of the alkaline serine protease which has been named subtilisin RT-5.

4-1 Subtilisin RT-5 from *Bacillus subtilis* strain RT-5

4-1-1 Preliminary characterization

Cell free concentrate obtained from fermentation of *Bacillus subtilis* RT-5 was investigated for the presence of proteolytic activity. Crude enzyme preparation was fractionated with 35% ammonium sulphate and then with 75% ammonium sulphate. The active proteolytic fraction was obtained in the 75% ammonium sulphate precipitate. Alternatively, crude enzyme preparation was precipitated with chilled acetone. In this case active proteolytic enzymes were found in 70% acetone precipitated fraction.

Ammonium sulphate or acetone precipitated preparation was subjected to ion-exchange chromatography on DEAE-Sepharose fast flow. Alkaline protease containing fraction did not bind to the resin while the neutral protease containing fraction bound to the column. Neutral protease was then eluted with salt gradient. This provides an efficient chromatographic
step for separating the alkaline protease from the neutral protease containing fractions. A subsequent chromatography of flow-through preparation of DEAE-Sepharose on Mono-S FPLC followed by rechromatography on the same column and under the same conditions but with a shallow salt gradient provided an enzyme preparation which was pure as judged by SDS-PAGE (Fig.3-11) and N-terminal amino acid analysis on gas and solid phase sequencers.

The molecular mass of the purified enzyme has been estimated to be around 30 kDa by SDS-PAGE under reducing and non-reducing conditions. The molecular mass calculated from the amino acid sequence analysis is 29.4 kDa. The amino acid composition of the intact protein obtained by acid hydrolysis is shown in Table 3-2, and its comparison with the compositions of other subtilisins and is presented in Table 3-3. The total amino acid composition of the subtilisin RT-5 is closely related to the subtilisin Novo indicating that it is a Novo type enzyme.

4-2 Enzymatic properties

The effect of addition of CaCl₂ on the activity of subtilisin RT-5 towards Succinyl Ala-Ala-Pro-Phe p-nitroanilide at pH 8.5 was investigated. When CaCl₂ was added to the assay medium the activity was increased. The activity increases linearly with concentration of CaCl₂ up to about 20mM (Fig. 3-20). A similar influence of calcium ions on proteolytic activity has been noted in previously isolated subtilisins (Stahl and Ferrari, 1984; Nedkov, 1983; Kurihara et al., 1972; Olaitan et al., 1968; Smith et al., 1968).

The influence of pH on kₐ₀/Kₘ for the hydrolysis of Succinyl Ala-Ala-Pro-Phe- p-nitroanilide is shown in Fig.3-21. The optimum pH is around 8.5, and it may be dependent on the protonation and deprotonation of groups with pKₐ around 9.3 and 6.6, respectively (Svendsen and Breddam, 1992).
The cleavage specificity data presented in Fig. 3-23 suggest that the enzyme exhibits a strong preference for the peptide bond at the carboxyl end of lysine and arginine and thus bears resemblance to trypsin. In addition, it also possesses chymotryptic like secondary activity which is detectable after digestion for a longer period (4 hrs) or at higher enzyme concentration (50 µM). A similar observation has been reported by Svendsen and Breddam (1992) for Glu/Asp specific endopeptidase from *Bacillus licheniformis*. At an enzyme concentration of 0.05µM they identified the following two cleavage points in ribonuclease: Glu9-Arg10 and Glu49-Ser50. At high enzyme concentration, however, the following additional cleavage sites were identified: Glu86-Thr87, Phe46-Val47, Glu111-Gly112, and Asp121- Ala122. It has, therefore, been suggested that high enzyme concentration and/or long reaction time led to the hydrolysis of other types of peptide bonds.

Subtilisin RT-5 is inhibited by 1.0 mM PMSF and 0.1mM 3-4 DCI hence it can be classified as a serine protease. Furthermore, as is the case with other alkaline proteases from closely related species of Bacillus (Stahl and Ferrari, 1984; Nedkov et al., 1983; Kurihara et al., 1972; Olaitan et al., 1968; Smith et al., 1968), the presence of 10 mM EDTA in the assay medium did not inhibit the enzyme activity. This suggests that the metal ions are important but not essential for the proteolytic activity of subtilisin (Table 3-5).

4-3 Primary structure of the subtilisin RT-5

The complete amino acid sequence of subtilisin RT-5 was established by N-terminal sequencing upto 43 residues of the intact protein and of peptides derived by cleavages with CNBr, trypsin, Lys-C and Asp-N proteases. The deduced primary structure of subtilisin RT-5 is shown in Fig.3-19. Primary structure analysis reveals the presence of 275 amino acid residues in this subtilisin with conspicuous absence of cysteine.
Catalytically active residues i.e., Asp32, His64 and Ser221 are conserved. Four methionine residues are also present in positions identical to those in other subtilisins.

4-4 Comparison with other subtilisins

Amino acid sequence of subtilisin Novo (Olaitan et al., 1968) has been included in Fig.4-1 to show the similarity between the two subtilisins. Comparison of the two enzymes suggested that they are homologous (237 residues are in identical positions). In addition, the known sequences of subtilisin DY (Nedkov et al., 1983) and mesentericopeptidase (Svendsen et al., 1986) are also shown. Sequence alignments show that all these enzymes belong to the same family of serine proteases but with a wide spread in variability i.e., from 5-83 replacements towards the enzyme from other Bacillus species.

Extensive homology is observed between the mesentericopeptidase and the subtilisin RT-5. There are only 5 amino acid replacements between these two enzymes. The substitutions are observed in positions Ala85 to Ser, Ser88 to Ala, Ala89 to Ser, Ala183 to Ser, and Ser259 to Asn. Significantly, with the exception of subtilisin RT-5 and mesentericopeptidase, all other subtilisins have a seryl residue at position 130 which is close to the important Tyr104 of the substrate binding region. A more radical change is observed in position 259 where the basic asparagine in subtilisin RT-5 has been replaced by uncharged serine in mesentericopeptidase and by aspartic acid in subtilisin Novo and subtilisin DY.

The sequence Ala88-Ser89 of subtilisin RT-5 is in inverse order in comparison to the corresponding sequence in mesentericopeptidase. Similarly the sequence Asn56-Pro57 is also in inverse order of the respective sequence in subtilisin Novo. Svendsen et al. (1986, 1983) and Wells et al. (1983) have also shown that in case of the Novo enzyme there
has been an error in the determination of the sequence in these positions.

With regards to the physico-chemical properties subtilisin RT-5, like subtilisin Novo, is highly soluble in water as compared to mesentericopeptidase and subtilisin Amylosacchariticus which have very low solubility in water. Svendsen et al. (1986) reported that the low solubility of mesentericopeptidase and subtilisin Amylosacchariticus cannot be explained from a comparison of their three-dimensional structures with that of subtilisin Novo, since most of the differences in amino acid residues on the enzyme surface are very conservative and do not cause any major change in the net charge of the enzyme molecule.

When the sequences of subtilisin RT-5 and other subtilisins are aligned for optimal homology, a total of 175 residues occupy identical positions in all the primary structures. According to these amino acid sequence homologies subtilisin RT-5 is closely related to mesentericopeptidase (97% homology) and to subtilisin Novo (86% homology). As expected, the most striking homologies are observed in the region around the essential amino acid residues of the active site i.e., Asp32 and Ser221. The only differences in these regions occur around His64 of the active site, where Phe58, Asp61 and Asn62 of subtilisin Novo are replaced in subtilisin RT-5 by Tyr, Gly and Ser, respectively. Thomas et al. (1985) have pointed out that because of a sequencing error there is an alanine at position 99 in the early crystallographic models of subtilisin Novo and that it should be Asp99. They have pointed out that Asp99 is some 14-15Å apart from the imidazole of His64 and that mutation of Asp99 to Ser99 decreased the k_cat/K_m by 20%. They have thus shown that the modification of a single charge in the vicinity of the active site of an enzyme can have a significant effect on the pH dependence of the catalytic reaction. It is also indicated that large effects on both the catalytic rate constants and pH dependence can be observed by replacements of multiple charges in the vicinity of the active site of an enzyme.
4-4-1 **Comparison of subtilisin RT-5 and subtilisin Novo**

Out of the 275 amino acid residues in a single polypeptide chain there are only 40 differences between subtilisin RT-5 and subtilisin Novo (Fig.4-3). At the nucleotide level, 25 of the replacements can be ascribed to single base substitution and 15 to double base substitutions.

As the primary structures of subtilisin RT-5 and Novo are very similar (differ only in 13% of the structure) there are close similarities in the enzymatic properties of the two proteins. These similarities also suggest that the polypeptide chain conformations of both the enzymes are very similar, if not identical, and that most substitutions would be conservative in nature, involving amino acid residues of the same type. When these substitutions are superimposed on the three-dimensional structure of subtilisin Novo it is observed that all substitutions occur in the exterior chain segments of the protein molecule.

Although, substitutions of amino acid residues are evenly distributed in the linear sequence of molecule there are several large regions of the sequence where no substitutions have occurred. One such region is around the reactive Ser221 and extends from Tyr214 to Asn240, and another is around reactive Asp32 extending from Ser9 to Leu42. There is also a conservative substitution of Ser224 by Thr in subtilisin RT-5.

4-4-2 **Comparison of subtilisins and pancreatic proteases**

Though subtilisins resemble pancreatic proteases in many ways e.g., inhibition with organophosphates, pH optima, involvement of histidine residue in enzyme catalysis, the primary amino acid sequences are very different. Presence of a high content of glycine, alanine, serine and the absence of cysteine residues are characteristic of subtilisins. Sequences of the amino acid residues around the active site histidine and serine of the
subtilisins do not resemble those present in the vertebrate enzymes. Furthermore, the position of and sequence around the aspartic acid residue involved in the active site of the subtilisins are completely different from the counterpart mammalian proteases (Blow et al., 1969) indicating that the bacterial and animal proteases have evolved independently, although they have similar mechanisms of action.

### 4-5 Presence of repeated sequence

Complete amino acid sequence of subtilisin shows the presence of similar sequences in different portions of the protein molecule. These sequences include His-Val-Ala-Gly (residues 67-70 and 226-229), Ala-Ala-Leu (residues 73-75 and 231-233)[Fig. 4-5]. Apart from these, there are a number of tripeptide repetitions, e.g., Ala-Ala-Ala (residues 151-153, 231-233, and 272-274), Val-Lys-Val (residues 26-28 and 93-95) and Val-Leu-Asp/Gly (residues 81-83 and 95-97)[Fig.4-4].

It is well established that the presence of such repetitive sequences in other proteins is responsible for extension of shorter peptide chains by a process of gene duplication. This type of gene duplication is well documented for ferredoxins (Matsubara et al., 1967; Eck and Dayhoff, 1966) and heptaglobulins (Hill et al., 1966). Hence it is reasonable to infer that the long polypeptide chain of 275 amino acid residues comprising the subtilisin molecule may have evolved by duplications leading to elongation of the genomic DNA.

The primary structure of subtilisin RT-5 also reveals that all the repeated sequences are not arranged in an exact linear order in the molecule and that some segments are present in inverse order. For example, at the N-terminal part of the molecule the duplicated sequences are in the order: residues 39-42, 67-75, and 82-94, but at the C-terminal part of the molecule they are present in inverse order as residues 238-241, 226-233, and 126-136 (Fig.4-5).
The varying number of residues in repeated sequences and the presence of repeated segments in inverse order indicate that the subtilisins may have evolved by a complex mechanism with several duplications of the particular sequences.

Presence of short repeated sequences has also been reported in rabbit brain ubiquitin by Wajih et al. (1992). Doolittle (1989, 1979) has emphasized the structural and functional importance of such sequences for many proteins. For example, the presence of repeats of Gly-Pro-X sequence in collagen and their role in its structural regularity and high tensile strength are well known.

Johnson et al. (1987) have also indicated the presence of repeated pentapeptide, Gly-Tyr-Asp-Lys-Tyr in egg shell structural proteins of schistosomes. These short peptidyl repeats are also found in non-structural proteins; for instance, an alternate arrangement of negatively charged and non-polar amino acid residues is responsible for building a cation binding site, as has been shown for the Pro-Glu repeated sequences in the iron binding protein of E.coli (Postle and Good, 1983).

4-6 Structure function relationship

Subtilisins like other proteolytic enzymes catalyze the hydrolysis of peptide bonds in proteins and peptides. They catalyze this reaction in a manner analogous to their mammalian counterparts, but the two classes differ in the preference of amino acid residues at the P1 position. In contrast to trypsin which cleaves peptide bonds at the carboxylic site of Lys and Arg, subtilisins, like chymotrypsin, cleave after bulky amino acid residues such as Phe, Tyr, Leu (Morihara et al., 1969). They thus resemble chymotrypsin in their mode of action.

The active site serine is considered to be a potent nucleophile due to its unique reactivity with substrate and inhibitors. Apart from serine, an
unprotonated His64 is also involved in the catalytic activity. This residue interacts as a base with the hydroxyl group of the Ser221 to promote its nucleophilicity (Russell and Fersht, 1987; Thomas et al., 1985).

Studies of the catalytic activities of subtilisins and their broad substrate specificities are useful in investigating the effects of different substrates and inhibitors upon their $k_{\text{cat}}$ and $K_m$. Subtilisins are well characterized, with respect to their active site residues which are involved in proteolysis. They are classified as serine protease due to the presence of a serine residue at the active site of the molecule. Serine plays a pivotal role in catalysis by forming covalent bonds with substrates and also with inhibitors.

The catalytic efficiency $k_{\text{cat}}/K_m$ of subtilisins depends upon the formation of an acyl-enzyme complex between the enzyme and the substrate. Though subtilisins are non-specific with respect to the P1 position of the substrate, the residues at the neighboring positions like P2, P3, P4 or P1', P2' are somewhat more important for the cleavage specificity of the enzyme (Morihara et al., 1970). The non-specific cleavage by subtilisins is due to the broader substrate binding pockets of these enzymes (Groen et al., 1992; Morihara et al., 1969).

The amino acid residues which form the catalytic triad are also conserved in subtilisin RT-5. The sequences around the active site Asp32 and Ser221 are also identical with the other subtilisins (Fig.4-1). The only significant difference between subtilisin RT-5 and subtilisin Novo is found in the sequence around the active site His64 where Tyr58, Gly61 and Ser62 are replaced with Phe, Asp, and Asn, respectively (Fig.4-2 & Fig.4-3).

The presence of Tyr, Gly, and Ser residues at these positions might be responsible for the increased catalytic activity of subtilisin RT-5 against Succinyl Ala-Ala-Pro-Phe p-nitroanilide. Notably, in subtilisin Carlsberg there is Gly at position 63 as compared to Ser in subtilisin Novo. The
higher activity of Carlsberg enzyme has been ascribed to the presence of Gly63 (Markland and Smith, 1970).

4-7 Prediction of three-dimensional structure of the subtilisin RT-5

Initial structural characterization of the subtilisin BPN' was done by Wright et al. (1969) and that of the Novo enzyme by Drenth et al. (1972) and Bott et al. (1988). Structure of the Carlsberg molecule is patterned along the same lines.

The general picture which has emerged is that of a globular protein which folds into a heart shaped structure with a shallow surface depression accommodating the active site. The molecule is comprised of a seven-stranded \( \beta \)-sheet and nine \( \alpha \)-helices. The \( \beta \)-sheet occupies a central position and divides the molecule into two distinct domains. The larger domain contains a total of seven \( \alpha \)-helices which are bundled together and are huddled against one face of the \( \beta \)-sheet. The remaining two helices reside in the smaller domain. Five of the \( \alpha \)-helices of the larger domain and the two of the smaller one run almost antiparallel to the strands of the \( \beta \)-sheet.

The crystal structure of subtilisin Novo (subtilisin BPN') resolved at 1.3\( ^\circ \)A has revealed the presence of two \( \text{Ca}^{+2} \) binding sites (Gallagher et al., 1993; Pantoliano et al., 1988; McPhalen and James, 1988). One of these sites (site A) binds calcium with high affinity and is located near the N-terminus. A-site involves amino acid residues: Gln2, Asp41, Leu75 and Asn77 in both the Novo enzyme and the subtilisin RT-5. The other site (site-B, located 32 \( ^\circ \)A away from the site-A) binds calcium and other cations much more weakly. It involves: Gly169, Lys170, Tyr171, Val174 and Glu195 in the Novo enzyme and Ala169, Lys170, Tyr171, Thr174 and Glu195 in the subtilisin RT-5. The high affinity calcium binding sites are responsible for the relatively high stability of subtilisins (Pantoliano et al.,
1988; Voordouw et al. 1976). This view is supported by the case of thermitase, a thermostable serine protease produced by *Thermoactinomyces vulgaris*. This enzyme contains three calcium binding sites and the unusually tight binding of the third calcium ion by thermitase is believed to be the most likely cause of its increased thermostability (Frommel and Sander, 1989).

The substrate binding pockets of subtilisins include a region of antiparallel β-structure consisting of hydrophobic residues Leu126, Gly127, Gly128 as substrate binding subsites S1-S3. Tyr104 acts as subsite S4 of the substrate binding pocket.

Comparison of the primary structure of subtilisin RT-5 with other subtilisins has revealed that all the amino acid substitutions are conservative in nature. They are present on the surface of the molecule and are not involved in any intramolecular contacts. More importantly catalytically active residues are conserved. From these facts and from the distinct sequence homology between subtilisin Novo and subtilisin RT-5 it can be surmised that the three-dimensional structure of subtilisin RT-5 and subtilisin Novo are probably identical. This contention is strongly supported by the predicted secondary and three-dimensional structure of thermitase, which is 44% homologous to subtilisin Carlsberg. Again most of the substitutions, insertions and deletions involve the surface of the molecule; hence the three-dimensional structures of the two enzymes are identical (Meloun et al., 1985).

**4-8 Thermostability of the subtilisin RT-5**

Subtilisin RT-5 significantly differs in thermostability from its mesophilic homologues mesentericopeptidase and subtilisin Novo. Comparative sequence analysis of these enzymes suggests that the replacements of Ala by Ser at positions 85, 89, and 183, and Ser by Ala and Asn at positions 88 and 259 respectively in subtilisin RT-5 are responsible for its
increased thermostability. This inference is corroborated by the findings of Frommel and Sander (1989) which indicated that replacements of Ala by Ser, Val by Ala, and Ser by Thr in the subtilisin family are responsible for the thermostability of these proteins.

Similarly, in comparison of the structures of subtilisin Novo and subtilisin RT-5 the replacements are noted: Ala by Ser, Ser by Thr, Val by Ala, Asp by Asn, Val by Ile and Gly by Ala. Again, all these replacements fit well with the previous report (Frommel and Sander, 1989) indicating that they are responsible for the increased thermostability of different enzymes.

4-9 Cloning of the subtilisin RT-5 gene

Subtilisin gene is present on the chromosome at position of $91^3$ from the origin of replication, as shown in Fig. 4-6. The N-terminal portion of the mature protein was amplified by using PCR. Since the DNA sequence of the subtilisin gene encoding subtilisin RT-5 was hitherto unknown, two degenerate oligonucleotide primers corresponding to residues 1-6 and 218-223 of the protein were designed. The DNA amplification reaction using these primers and the *Bacillus subtilis* RT-5 genomic DNA produced a single discrete band of approximately 670 bp on 1% agarose gel. This fragment was eluted from the band, cloned into pEMBL 8 vector containing M13 mp cloning and sequencing box (Fig.4-7) and sequenced. The availability of the DNA sequence made it possible to infer the amino acid sequence of the sector of the protein spanning residues 140-185 which could not be recovered from the protein digests. It confirmed the presence of 4 methionine residues and also the presence of two Asn-Gly regions at Asn109- Gly110 and Asn218-Gly219 positions respectively. The codon usage of subtilisin RT-5 gene is different from that of previously isolated genes from *Bacillus licheniformis* and *Bacillus amylo liquefaciens*. Analysis of the sequence reveals that the open reading frame of the cloned segment consists of 669 bp. The sequence showed high identity
with the nucleotide sequence of subtilisin gene of *Bacillus subtilis* Marburg strain I-168 (Stahl and Ferrari, 1984). Furthermore, the amino acid sequence inferred from the DNA sequence completely matched the amino acid sequence determined by protein sequencing.
Table 4-1 Growth rate constants and protease production obtained by *B. subtilis* GBS01 in different media.

<table>
<thead>
<tr>
<th>Media</th>
<th>Growth rate constant</th>
<th>Protease production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium#1</td>
<td>0.31</td>
<td>217</td>
</tr>
<tr>
<td>Medium#2</td>
<td>0.31</td>
<td>209</td>
</tr>
<tr>
<td>Medium#3</td>
<td>0.31</td>
<td>209</td>
</tr>
<tr>
<td>Medium#4</td>
<td>0.30</td>
<td>165</td>
</tr>
<tr>
<td>Medium#5</td>
<td>0.26</td>
<td>105</td>
</tr>
<tr>
<td>Medium#6</td>
<td>0.20</td>
<td>178</td>
</tr>
<tr>
<td>Medium#7</td>
<td>0.19</td>
<td>148</td>
</tr>
<tr>
<td>Medium#8</td>
<td>0.15</td>
<td>104</td>
</tr>
<tr>
<td>Medium#9</td>
<td>0.12</td>
<td>61</td>
</tr>
<tr>
<td>Medium#10</td>
<td>0.21</td>
<td>200</td>
</tr>
<tr>
<td>Medium#11</td>
<td>0.23</td>
<td>244</td>
</tr>
<tr>
<td>Medium#12</td>
<td>0.22</td>
<td>122</td>
</tr>
<tr>
<td>Medium#13</td>
<td>0.22</td>
<td>191</td>
</tr>
<tr>
<td>Medium#14</td>
<td>0.16</td>
<td>313</td>
</tr>
<tr>
<td>Medium#15</td>
<td>0.16</td>
<td>122</td>
</tr>
<tr>
<td>Medium#16</td>
<td>0.16</td>
<td>209</td>
</tr>
<tr>
<td>Medium#17</td>
<td>0.20</td>
<td>139</td>
</tr>
<tr>
<td>Medium#18</td>
<td>0.19</td>
<td>83</td>
</tr>
<tr>
<td>Medium#19</td>
<td>0.19</td>
<td>139</td>
</tr>
<tr>
<td>Medium#20</td>
<td>0.23</td>
<td>308</td>
</tr>
<tr>
<td>Medium#21</td>
<td>0.23</td>
<td>226</td>
</tr>
<tr>
<td>Medium#22</td>
<td>0.19</td>
<td>170</td>
</tr>
</tbody>
</table>

For media composition see material and methods.

Protease production was monitored in terms of L.V. units.
Fig. 4-1 Comparison of the primary structure of subtilisin RT-5 with other subtilisins. Continuous sequence is of subtilisin RT-5 and for other subtilisins only replacements are shown.
Subtilisin RT-5  | DGSSH*GTHV
Subtilisin Novo | DDNSH*GTHV
Subtilisin Carlsberg | DGNCH*GTHV

Fig. 4-2  Replacements of amino acid residues around the reactive site His64 between subtilisin RT-5 and other subtilisins.

H* = active site His64.
Fig. 4-3  Comparison of the primary structure of subtilisin RT-5 with subtilisin Novo.
Fig. 4.4  Comparison of the primary structure of subtilisin RT-5 with subtilisin Carlsberg.
His-Val-Ala-Gly-Thr-Ile-Ala-Ala-Leu
His-Val-Ala-Gly- - -Ala-Ala-Ala-Leu
226

82
Leu-Gly-Val-Ser-Pro-Ser-Ala-Ser-Leu-Tyr-Ala-Val-Lys
Leu-Gly- - -Gly-Pro-Thr-Gly-Ser-Thr- - -Ala-Leu-Lys
126 127 129 132 134 136

77
Asn-Ser-Ile-Gly-Val-Leu-Gly-Val-Ser-Pro-Ser-Ala
Ser-Thr-Val-Gly-Tyr-Pro-Ala-Lys-Tyr-Pro-Ser-Thr
163 166 172 173

39 42
His-Pro-Asp-Leu
His-Pro-Thr-Trp
238 241

45 49
Arg-Gly-Gly-Ala-Ser
Leu-Gly-Gly-Pro-Thr
126 130

73 77 167 171
Ala-Ala-Leu-Asn-Asn
Tyr-Pro-Ala- - -Lys-Tyr
Thr-Val-Val-Asp-Lys
137 141 209 214
Leu-Pro-Gly-Gly-Thr-Tyr

Fig. 4-5 Short repeated sequences within the subtilisin RT-5.
Fig. 4-6  Genetic map of *Bacillus subtilis* [Piggot, P. J. & Hoch, J. A. (1985) Microbiol. Rev. 49, 158-179].
Table 4.2 Genetic markers of *Bacillus subtilis* [Piggot, P. J. & Hoch, J. A. (1985) Microbiol. Rev. 49, 158-179].

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Monogenic Character</th>
<th>Map Position (Degrees)</th>
<th>Phenotype, enzyme deficiency, or other characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>abrA</em></td>
<td>Antibiotic resistance</td>
<td>325</td>
<td>Partial suppressor of stage 0 phenotypes, may be same as revA</td>
</tr>
<tr>
<td><em>abrB</em></td>
<td>Antibiotic resistance</td>
<td>3</td>
<td>Partial suppressor of stage 0 phenotypes, ribosome alterations</td>
</tr>
<tr>
<td><em>abrC</em></td>
<td>Antibiotic resistance</td>
<td></td>
<td>Weak intragenic suppressors of <em>spmA</em></td>
</tr>
<tr>
<td><em>abrD</em></td>
<td>Antibiotic resistance</td>
<td></td>
<td>Partial suppressor of stage 0 phenotypes; see <em>abrB</em></td>
</tr>
<tr>
<td><em>abrH</em></td>
<td>Antibiotic resistance</td>
<td></td>
<td>Partial suppressor of stage 0 phenotypes; see <em>abrB</em></td>
</tr>
<tr>
<td><em>acrA</em></td>
<td>Acrithinacitin</td>
<td>126</td>
<td>Pyruvate dehydrogenase defect, defective in El (pyruvate dehydrogenase) component of pyruvate dehydrogenase</td>
</tr>
<tr>
<td><em>acrB</em></td>
<td>Acrithinacitin</td>
<td>230</td>
<td>Resistant to acrithinacitin, also to ethidium bromide and distamycin; sensitive to streptomycin</td>
</tr>
<tr>
<td><em>acrF</em></td>
<td>Acrithinacitin</td>
<td>215</td>
<td>Same as <em>acrA</em></td>
</tr>
<tr>
<td><em>adrC</em></td>
<td>Adrenergic</td>
<td>NM</td>
<td>Adrenergic deaminase</td>
</tr>
<tr>
<td><em>adrF</em></td>
<td>Adrenergic</td>
<td>NM</td>
<td>Adrenergic phospholipid transferase</td>
</tr>
<tr>
<td><em>adnA</em></td>
<td>Aminobutylylcysteine</td>
<td>250</td>
<td>Aminobutylylcysteine resistance, regulation of <em>aspartokinase II</em></td>
</tr>
<tr>
<td><em>adnB</em></td>
<td>Aminobutylylcysteine</td>
<td>290</td>
<td>Aminobutylylcysteine resistance, structural gene for <em>aspartokinase II</em></td>
</tr>
<tr>
<td><em>ahbA</em></td>
<td>Arginine dehydrogenase</td>
<td>342</td>
<td>Arginine dehydrogenase resistance; linked to <em>epsA</em></td>
</tr>
<tr>
<td><em>ahbB</em></td>
<td>Arginine dehydrogenase</td>
<td>321</td>
<td>Arginine dehydrogenase resistance</td>
</tr>
<tr>
<td><em>ahbC</em></td>
<td>Arginine dehydrogenase</td>
<td>215</td>
<td>Arginine dehydrogenase resistance</td>
</tr>
<tr>
<td><em>ahbD</em></td>
<td>Arginine dehydrogenase</td>
<td>99</td>
<td>Arginine dehydrogenase resistance</td>
</tr>
<tr>
<td><em>alaA</em></td>
<td>Alanine</td>
<td>281</td>
<td>Alanine deaminase</td>
</tr>
<tr>
<td><em>alaB</em></td>
<td>Alanine</td>
<td>280</td>
<td><em>a</em>-Alanine dehydrogenase</td>
</tr>
<tr>
<td><em>alaC</em></td>
<td>Alanine</td>
<td>266</td>
<td>Acetolactate synthase</td>
</tr>
<tr>
<td><em>alaD</em></td>
<td>Alanine</td>
<td>279</td>
<td>Constitutive acetolactate synthase</td>
</tr>
<tr>
<td><em>aomA</em></td>
<td>Aminomycin</td>
<td>250</td>
<td>Glutaminate requirement</td>
</tr>
<tr>
<td><em>aomB</em></td>
<td>Aminomycin</td>
<td></td>
<td>3-Aminomycin resistance, part of or very close to <em>argC</em> locus</td>
</tr>
<tr>
<td><em>aomC</em></td>
<td>Aminomycin</td>
<td></td>
<td>Control of arginine synthesis; probably identical to <em>asnC</em> and <em>parC</em>; see <em>asnC</em></td>
</tr>
<tr>
<td><em>alnA</em></td>
<td>Alanine</td>
<td>25</td>
<td>Alanine dehydrogenase, gene also called <em>aomB</em></td>
</tr>
<tr>
<td><em>alnB</em></td>
<td>Alanine</td>
<td>25</td>
<td>Control of alanine synthesis, also called <em>aomC</em></td>
</tr>
<tr>
<td><em>aprF</em></td>
<td>Alkaline phosphatase</td>
<td>913</td>
<td>Structural gene for <em>subtilis</em> F, map order <em>thr</em> → <em>gly</em> → <em>asp</em> → <em>esp</em> (formerly <em>spel</em>)</td>
</tr>
<tr>
<td><em>aspA</em></td>
<td>Aspartate</td>
<td>294</td>
<td>Aspartate utilization</td>
</tr>
<tr>
<td><em>aspB</em></td>
<td>Aspartate</td>
<td>296</td>
<td>Aspartate utilization</td>
</tr>
<tr>
<td><em>aspC</em></td>
<td>Aspartate</td>
<td>172</td>
<td>Aspartate utilization</td>
</tr>
<tr>
<td><em>aspD</em></td>
<td>Aspartate</td>
<td>100</td>
<td>Asparagine resistance, identified by complementation of corresponding locus in <em>E. coli</em></td>
</tr>
<tr>
<td><em>aspE</em></td>
<td>Asparagine</td>
<td>100</td>
<td>Same as <em>argT</em></td>
</tr>
<tr>
<td><em>aspF</em></td>
<td>Asparagine</td>
<td>100</td>
<td>Same as <em>argT</em></td>
</tr>
<tr>
<td><em>aspG</em></td>
<td>Asparagine</td>
<td>100</td>
<td>Same as <em>argT</em></td>
</tr>
<tr>
<td><em>aspH</em></td>
<td>Asparagine</td>
<td>260</td>
<td>Same as <em>argF</em>, homologous gene</td>
</tr>
<tr>
<td><em>aspI</em></td>
<td>Asparagine</td>
<td>260</td>
<td>Same as <em>argI</em></td>
</tr>
<tr>
<td><em>aspJ</em></td>
<td>Asparagine</td>
<td>115</td>
<td>Asparagine:arginase or putrescine requirement</td>
</tr>
<tr>
<td><em>aspL</em></td>
<td>Asparagine</td>
<td>265</td>
<td>3-Keto-4-deoxy-d-arabinose-2-phosphate synthase</td>
</tr>
<tr>
<td><em>aspM</em></td>
<td>Asparagine</td>
<td>210</td>
<td>3-Keto-4-deoxy-d-arabinose-2-phosphate synthase</td>
</tr>
<tr>
<td><em>aspN</em></td>
<td>Asparagine</td>
<td>210</td>
<td>3-Keto-4-deoxy-d-arabinose-2-phosphate synthase</td>
</tr>
<tr>
<td><em>aspO</em></td>
<td>Asparagine</td>
<td>210</td>
<td>3-Keto-4-deoxy-d-arabinose-2-phosphate synthase</td>
</tr>
<tr>
<td><em>aspP</em></td>
<td>Asparagine</td>
<td>270</td>
<td>Carboxylase resistance, <em>oxyrimy</em> 1</td>
</tr>
<tr>
<td><em>aspQ</em></td>
<td>Asparagine</td>
<td>210</td>
<td>Carboxylase resistance, <em>oxygen</em> 1 and 2</td>
</tr>
<tr>
<td><em>aspR</em></td>
<td>Asparagine</td>
<td>25</td>
<td>Shikimate kinase</td>
</tr>
<tr>
<td><em>aspS</em></td>
<td>Asparagine</td>
<td>230</td>
<td>Lactate dehydrogenase, phenylalanine; see <em>hall</em></td>
</tr>
<tr>
<td><em>aspT</em></td>
<td>Asparagine</td>
<td>230</td>
<td>Asparagine dehydrogenase</td>
</tr>
<tr>
<td><em>aspU</em></td>
<td>Asparagine</td>
<td>17</td>
<td>200-kd protein, see <em>aspT</em></td>
</tr>
<tr>
<td><em>aspV</em></td>
<td>Asparagine</td>
<td>17</td>
<td>Constitutive asparagine</td>
</tr>
<tr>
<td><em>aspX</em></td>
<td>Asparagine</td>
<td>17</td>
<td>Aspartate transporter</td>
</tr>
<tr>
<td><em>aspY</em></td>
<td>Asparagine</td>
<td>17</td>
<td>Asparagine-threonine requirement</td>
</tr>
<tr>
<td><em>attA</em></td>
<td>Attachment</td>
<td>198</td>
<td>Integration site for phase SPB</td>
</tr>
<tr>
<td><em>attB</em></td>
<td>Attachment</td>
<td>10</td>
<td>Integration site for phase SPB</td>
</tr>
<tr>
<td><em>attP</em></td>
<td>Attachment</td>
<td>10</td>
<td>Integration site for phase SPB</td>
</tr>
</tbody>
</table>

101
Fig. 4-7  Structure of pEMBL 8 vector [Old, R. W. & Primrose, S. B. in "Principles of Gene Manipulation" (1989) pp. 81].
5-0 REFERENCES


Formula used for the calculation of specific growth rate constant:

\[ N_t = N_0 \cdot 2^n = N_0 \cdot 2^{t/t_d} \]

Similarly,

\[ N_t/N_0 = 2^{t/t_d} \]

and taking logarithms,

\[ \ln(N_t/N_0) = (\ln 2)t/t_d, \]

or,

\[ (\ln N_t - \ln N_0)/t = 0.693/t_d \]

\[ \mu = \ln 2/t_d = 0.693/t_d \]

where,

\[ \mu = \text{specific growth rate constant, expressed in reciprocal hours (h}^{-1}) \]

\[ \ln 2 = \text{natural log of 2} \]

\[ t_d = \text{doubling time} \]

\[ t = \text{time of incubation} \]

\[ n = \text{number of generation} \]

\[ N_t = \text{number of cells at time 't'} \]

\[ N_0 = \text{number of cells at zero time.} \]

Reference: