Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

SUBSTRATE SPECIFICITY AND STRUCTURAL INVESTIGATION INTO PEPO AND PEPW: TWO PEPTIDASES FROM *LACTOBACILLUS RHAMNOSUS*

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science
in
Biochemistry

at Massey University, Palmerston North
New Zealand.

Karen Maree Yates

2005

ABSTRACT

The proteolytic systems of lactic acid bacteria have important roles in the maturation and flavour development of cheese. Lactic acid bacteria pepetidases contribute to the taste of cheese through the production of low-molecular weight peptides and free amino acids. Although some lactic acid bacteria peptidases have been structurally and enzymatically characterised for their substrate specificity, there are some that are yet to be completely biochemically characterised. The aim of the present study was to investigate the substrate specificity and three-dimensional structure of two peptidases that could potentially be used as a tool to modify and control cheese bitterness and possibly other flavour attributes from *Lactobacillus rhamnosus*, PepO and PepW.

The *pepW* gene was successfully cloned from *L. rhamnosus* into an *E. coli* expression system. Recombinant PepW was purified to homogeneity and was shown to exist as a hexamer of 50 kDa subunits. Recombinant PepO was expressed from a previously established *L. lactis* expression system and purified to homogeneity. PepO was shown to exist as a 70 kDa monomer, and function as a metallopeptidase.

PepO and PepW were shown to selectively hydrolyse chymosin-derived bovine β- and κ-casein peptides, and casein peptides extracted from Cheddar cheese. One conclusive PepO cleavage site that had not been previously characterised was identified. This was the β-casein peptide bond between Leu_6-Asn_7. Several possible PepO and PepW cleavage sites in α_{s1}-, β- and κ-casein were identified, suggesting that PepO has a broad endopeptidase activity, whilst PepW has a specific exopeptidase activity.

PepO and PepW crystals were successfully grown for structure determination by x-ray crystallography. Native data sets were collected for both PepO and PepW, and derivative data were collected for PepO. Structure determination was attempted using Multiple Isomorphous Replacement and Molecular Replacement techniques.

Results from the substrate specificity and structural investigation of the *L. rhamnosus* peptidases, PepO and PepW, are presented in this thesis.

ACKNOWLEDGMENTS

I would like to acknowledge the following people for their assistance throughout the course of this project:

My supervisors Dr Gill Norris and Dr Julian Reid for their advise, guidance and encouragement.

Professor Geoffrey Jameson for his help collecting and processing x-ray diffraction data, and explanation of crystal twinning. Dr Brian Anderson for his help in processing x-ray diffraction data.

The Centre for Molecular Biodiscovery for setting up the sitting drop crystallisation screens. Dr Julian Reid and Mr Aveenash Bala for carrying out the liquid-chromatography tandem mass spectrometry experiments. Trevor Loo for his help with Edman N-terminal sequencing and Dr Rosemary Brown for supplying the pET32a-Trx-His-rTEV vector.

Members of the X-lab who made working on this project enjoyable (most of the time).

Finally, thank you to my family and friends who have always encouraged me to pursuit my interests and helped me overcome my obstacles.

This work was carried out in collaboration with the Fonterra Co-operative Group Limited, and supported by the Foundation for Research Science and Technology, New Zealand.

TABLE OF CONTENTS

ABSTRACT ii
ACKNOWLEDGMENTS iii
TABLE OF CONTENTS iv
LIST OF FIGURES vii
LIST OF TABLES x
ABBREVIATIONS xiii

PROTEOLYTIC ENZYMES: AN INTRODUCTION
1.1 PROTEOLYTIC ENZYMES 1
1.2 PROTEOLYTIC ENZYMES IN BIOLOGY 5
1.3 LACTIC ACID BACTERIA 6
1.4 *LACTOBACILLUS RHAMNOSUS* PEPO 12
1.5 *LACTOBACILLUS RHAMNOSUS* PEPW 15
1.6 PROJECT AIMS 18

PART 1:
SUBSTRATE SPECIFICITY OF THE *L. RHAMNOSUS* PEPTIDASES, PEPO AND PEPW.

CHAPTER 1 OBJECTIVES AND STRATEGIES
1.1 INTRODUCTION 19
1.2 OBJECTIVES AND STRATEGIES 19

CHAPTER 2 MOLECULAR CLONING
2.1 INTRODUCTION 20
2.2 METHODS 21
2.3 RESULTS AND DISCUSSION 25
2.3.1 *L. RHAMNOSUS* PEPW CLONING 25

CHAPTER 3 PROTEIN EXPRESSION AND PURIFICATION
3.1 INTRODUCTION 29
3.2 METHODS 31
3.3 RESULTS AND DISCUSSION 35
3.3.1 *L. RHAMNOSUS* PEPO EXPRESSION AND PURIFICATION 35
3.3.2 *L. RHAMNOSUS* PEPW EXPRESSION AND PURIFICATION 40

CHAPTER 4 PREPARATION OF SUBSTRATES
4.1 INTRODUCTION 44
4.2 METHODS 47

4.3 RESULTS AND DISCUSSION 50
4.3.1 *L. LACTIS* CELL-ENVELOPE PROTEINASE EXTRACTION 50
4.3.2 BOVINE CASEIN PEPTIDE SUBSTRATES 52
4.3.3 CHEESE EXTRACTED PEPTIDE SUBSTRATES 52

CHAPTER 5 SUBSTRATE SPECIFICITY
5.1 INTRODUCTION 53
5.2 METHODS 57
5.3 RESULTS AND DISCUSSION 59
5.3.1 *L. RHAMNOSUS* PEPO SUBSTRATE SPECIFICITY 59
5.3.2 *L. RHAMNOSUS* PEPW SUBSTRATE SPECIFICITY 68

CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS
6.1 CONCLUSIONS 71
6.2 FUTURE DIRECTIONS 72

PART 2:
THREE-DIMENSIONAL STRUCTURE DETERMINATION OF THE *L. RHAMNOSUS* PEPTIDASES, PEPO AND PEPW.

CHAPTER 1 OBJECTIVES AND STRATEGIES
1.1 INTRODUCTION 73
1.2 OBJECTIVES AND STRATEGIES 73

CHAPTER 2 CRYSTALLISATION
2.1 INTRODUCTION 74
2.2 METHODS 77
2.3 RESULTS AND DISCUSSION 80
2.3.1 *L. RHAMNOSUS* PEPO CRYSTALLISATION 80
2.3.2 *L. RHAMNOSUS* PEPW CRYSTALLISATION 88

CHAPTER 3 X-RAY DIFFRACTION OF CRYSTALS
3.1 INTRODUCTION 91
3.2 METHODS 96
3.3 RESULTS AND DISCUSSION 97
3.3.1 *L. RHAMNOSUS* PEPO X-RAY DIFFRACTION 97
3.3.2 *L. RHAMNOSUS* PEPW X-RAY DIFFRACTION 105

CHAPTER 4 SOLVING THE PHASE PROBLEM
4.1 INTRODUCTION 106

4.2 METHODS 112

4.3 RESULTS AND DISCUSSION 114

4.3.1 *L. RHAMNOSUS* PEPO PHASE PROBLEM 114

4.3.2 *L. RHAMNOSUS* PEPW PHASE PROBLEM 128

CHAPTER 5 CONCLUSIONS AND FUTURE DIRECTIONS

5.1 CONCLUSIONS 129

5.3 FUTURE DIRECTIONS 130

APPENDICES

APPENDIX I *L. RHAMNOSUS* PEPO AND PEPW SEQUENCE 131

APPENDIX II CHYMOSIN AND *L. LACTIS* SUBSP. *CREMORIS* STRATINS H2 AND SK11 PRTP CLEAVAGE SITES AND SOLUBLE PEPTIDE PRODUCTS OF α_{S1}-, β-, AND κ-CASEIN 134

APPENDIX III RP-HPLC ANALYSES OF α_{s1}-, β- and κ-CASEIN SUBSTRATE PEPTIDES 137

APPENDIX IV RP-HPLC ANALYSES OF *L. RHAMNOSUS* PEPO SUBSTRATE SPECIFICITY 142

APPENDIX V THE SOLUBLE PEPO SUBSTRATE PEPTIDES FROM α_{S1}-, β- AND κ-CASEIN AND THE PRODUCT PEPTIDES RESULTING FROM THEIR HYDROLYSIS BY PEPO 147

APPENDIX VI CRYSTALLISATION REAGENTS 150

REFERENCES 160

LIST OF FIGURES

PROTEOLYTIC ENZYMES: AN INTRODUCTION

Figure 1.1 Proteolytic Enzymes Catalyse the Hydrolysis of Peptide Bonds 1

Figure 1.2 The Covalent and Non-covalent Catalytic Mechanism of Peptidases 3

Figure 1.3 Nomenclature for Peptidase Substrate Specificity 5

Figure 1.4 The Lactic Acid Bacteria Proteolytic Network 8

Figure 1.5 Cleavage Specificity of *L. rhamnosus* PepO and *L. lactis* PepO for α_{s1}-casein(1-23) 12

Figure 1.6 Sequence Alignment of PepO from *L. rhamnosus*, *L. lactis* and *H. sapiens* NEP 13

Figure 1.7 The Three-Dimensional Structure of *H. sapiens* NEP 14

Figure 1.8 The Swiss-Model Predicted Three-Dimensional Structure of PepO 14

Figure 1.9 Sequence Alignment of PepW from *L. rhamnosus*, *L. lactis* PepC, *H. sapiens* hBH, and *S. cerevisiae* Gal6 16

Figure 1.10 The Three-Dimensional Structure of the *S. cerevisiae* Bleomycin Hydrolase 17

Figure 1.11 The Swiss-Model Predicted Three-Dimensional Structure of PepW 17

PART 1:

SUBSTRATE SPECIFICITY OF THE *L. RHAMNOSUS* PEPTIDASES, PEPO AND PEPW.

CHAPTER 2

Figure 2.1 Operon Containing the *pepW* Gene 25

Figure 2.2 Conserved Domains of the Putative Peptidase, PepW 25

Figure 2.3 Cloning Strategy for the *pepW* Gene 26

Figure 2.4 Agarose Gel Electrophoresis Analysis of *pepW* Cloning 27

CHAPTER 3

Figure 3.1 *Lactococcus lactis* Nisin-Inducible Expression System for *L. rhamnosus* PepO 35

Figure 3.2 SDS-PAGE Analysis of Nisin Induced Expression and Solubility of PepO 36

Figure 3.3 SDS-PAGE Analysis of PepO Purification 37

Figure 3.4 PepO Peptidase Activity During Protein Purification 38

Figure 3.5 *E. coli* IPTG-Inducible Expression System for *L. rhamnosus* PepW 40

Figure 3.6 SDS-PAGE Analysis of IPTG Induced PepW Expression 41

Figure 3.7 SDS-PAGE Analysis of PepW Solubility 41

Figure 3.8 SDS-PAGE Analysis of PepW Purification 42

CHAPTER 4

Figure 4.1 The LAB Proteolytic Network 44

Figure 4.2 SDS-PAGE Analysis of *L. lactis* subsp. *cremoris* Strains H2 and SK11 PrtP Following Autoproteolytic Release from the Cell-Envelope 50

Figure 4.3 Proteolytic Activity of the *L. lactis* subsp. *cremoris* Strains H2 and SK11 PrtP Enzymes 51

CHAPTER 5

Figure 5.1 *L. rhamnosus* PepO has Unique Cleavage Specificity for α_{s1}-casein (1-23) 53

Figure 5.2 Formation of the B-ion and Y-ion Series from Fragmentation of a Singly-Charged Precursor 56

Figure 5.3 RP-HPLC Analysis of PepO Substrate (Chymosin-Generated β-casein Peptides) and Product Peptides 61

Figure 5.4 RP-HPLC Analysis of PepO Substrate (Chymosin-Generated κ-casein Peptides) and Product Peptides 62

Figure 5.5 RP-HPLC Analysis of PepO Cheddar Cheese Extracted Substrate and Product Peptides 63

Figure 5.6 Bonds that must be Fragmented to Produce the a-, b-, c- and x-, y-, z-ions 65

Figure 5.7 RP-HPLC Analysis of PepW Substrate (Chymosin-Generated β-casein Peptides) and Product Peptides from β-casein 69

Figure 5.8 RP-HPLC Analysis of PepW Substrate (Chymosin-Generated κ-casein Peptides) and Product Peptides from κ-casein 70

PART 2:

THREE-DIMENSIONAL STRUCTURE DETERMINATION OF THE *L. RHAMNOSUS* PEPTIDASES, PEPO AND PEPW.

CHAPTER 2

Figure 2.1 Crystallisation Phase Diagram 74

Figure 2.2 PepO Crystals Obtained from Initial Crystallisation Screens at Room Temperature 80

Figure 2.3 PepO Crystals Grown in Conditions (d) and (e) 81

Figure 2.4 PepO Crystals Grown in Prasoedymium and Phosphoramidon 83

Figure 2.5 PepO Crystals Obtained from Secondary Crystallisation Screens 86

Figure 2.6 Dumb Luck PepO Crystals Obtained from Condition (b") 87

Figure 2.7 PepW Crystals Obtained from Initial Crystallisation Screens at Room Temperature 89

Figure 2.8 PepW Crystal Grown in Condition (a') 90

CHAPTER 3

Figure 3.1 Braggs Law of Diffraction 91
Figure 3.2 The Ewald Construction 92
Figure 3.3 Schematic Representation of X-ray Diffraction and Crystal Rotation 93

CHAPTER 4

Figure 4.1 The Patterson Map 107
Figure 4.2 The Molecular Replacement Technique 108
Figure 4.3 The MIR Harker Diagram 109
Figure 4.4 The MIRAS Harker Diagram 110
Figure 4.5 Wilson Plot for (left) PepO Native Data Set 1 and (right) PepO Native Data Set 2 114
Figure 4.6 Cumulative Intensity-Distribution Curves for (left) PepO Native Data Set 1 and (right) PepO Native Data Set 2 114
Figure 4.7 Hklview Analysis of (left) PepO Native Data Set 1 and (right) PepO Native Data Set 2 115
Figure 4.8 Swiss-Model Predicted Structure of PepO 117
Figure 4.9 Wilson Distribution and Cumulative Intensity-Distribution for PepO Native Data Set 3 123
Figure 4.10 Visual Analysis of PepO Native Data Set 3 Unit Cell 125
Figure 4.11 Visual Analysis of PepO Native Data Set 3 Unit Cell as Viewed Along the b- and c-axis (top) and a- and b-axis (below) 126
Figure 4.12 Swiss-Model Predicted Structure of PepW 127

LIST OF TABLES

PROTEOLYTIC ENZYMES: AN INTRODUCTION

Table 1.1 Peptide Classification According to Enzyme Commission and Reaction Catalysed 2
Table 1.2 Statistics for the MEROPS Database 5
Table 1.3 *Lactococcus lactis* Intracellular Peptidases 9

PART 1:
SUBSTRATE SPECIFICITY OF THE *L. RHAMNOSUS* PEPTIDASES, PEPO AND PEPW.

CHAPTER 3

Table 3.1 Protein Properties Exploited in Purification 29
Table 3.2 Purification Table of *L. rhamnosus* PepO 39

CHAPTER 4

Table 4.1 Properties of Bovine Casein 45
Table 4.2 Chymosin-derived Soluble Peptides of αs1-, β-, and κ-casein 46

CHAPTER 5

Table 5.1 Mass and Sequence Analysis of PepO Substrate and Product Peptides from β-casein........ 61
Table 5.2 Mass and Sequence Analysis of PepO Substrate and Product Peptides from κ-casein 62
Table 5.3 Mass and Sequence Analysis of PepO Cheddar Cheese Casein Substrate and Product Peptides 63
Table 5.4 LC-ESI-MSMS Analysis of PepO Substrate (Chymosin-Generated β-casein Peptides and Product Peptides from β-casein 66
Table 5.5 Mass and Sequence Analysis of PepW Substrate and Product Peptides from β-casein 69
Table 5.6 Mass and Sequence Analysis of PepW Substrate and Product Peptides from κ-casein 70

PART 2:
THREE-DIMENSIONAL STRUCTURE DETERMINATION OF THE *L. RHAMNOSUS* PEPTIDASES, PEPO AND PEPW.

CHAPTER 2

Table 2.1 Factors Affecting Crystallisation ... 75
Table 2.2 Heavy Atom Derivative Crystallisation Conditions ... 79

CHAPTER 3

Table 3.1 The Crystal Systems and Point Groups ... 94
Table 3.2 X-ray Diffraction Analysis of PepO Crystals Grown in Condition (d) and (e) ... 97
Table 3.3 Summary of Data Collection Statistics for PepO Native Data Sets 1 and 2 in Tetragonal Crystal Systems ... 98
Table 3.4 Summary of Data Collection Statistics for PepO Native Data Sets 1 and 2 in Monoclinic Crystal Systems ... 99
Table 3.5 X-ray Diffraction Analysis of PepO Derivative Crystals ... 100
Table 3.6 Summary of Data Collection Statistics for PepO Derivative Data Set Crystals ... 102
Table 3.7 X-ray Diffraction Analysis of PepO Crystals Grown in Condition (a") and (g) ... 103
Table 3.8 Summary of Data Collection Statistics for PepO Native Data Set 3 in an Orthorhombic Crystal System ... 104
Table 3.9 X-ray Diffraction Analysis of PepW Crystals Grown in Condition (b') ... 105

CHAPTER 4

Table 4.1 AMoRe Solutions for Data Set 1 Processed in Tetragonal using *H. sapiens* NEP as the Search Model ... 118
Table 4.2 AMoRe Solutions for Data Set 1 Processed in Monoclinic using *H. sapiens* NEP as the Search Model ... 118
Table 4.3 AMoRe Solutions for Data Set 1 Processed in Tetragonal using the Swiss-Model Predicted PepO Structure as the Search Model ... 119
Table 4.4 AMoRe Solutions for Data Set 1 Processed in Monoclinic using the Swiss-Model Predicted PepO Structure as the Search Model ... 119
Table 4.5 AMoRe Solutions for Data Set 2 Processed in Tetragonal using *H. sapiens* NEP as the Search Model ... 120
Table 4.6 AMoRe Solutions for Data Set 2 Processed in Monoclinic using *H. sapiens* NEP as the Search Model ... 120
Table 4.7 AMoRe Solutions for Data Set 2 Processed in Tetragonal using the Swiss-Model Predicted PepO Structure as the Search Model ... 121
Table 4.8 AMoRe Solutions for Data Set 2 Processed in Monoclinic using the Swiss-Model Predicted PepO Structure as the Search Model ... 121
Table 4.9 Scaleit Statistics for PepO Derivative Data Sets ... 122
Table 4.10 *AMoRe* Solutions for PepO Native Data Set 3 using *H. sapiens* NEP as the Search Model ... 124

Table 4.11 AMoRe Solutions for PepO Native Data Set 3 using the Swiss-Model Predicted PepO Structure as the Search Model 124

CHAPTER 5

Table 5.1 X-ray Diffraction Analysis of PepO and PepW Crystals 128

ABBREVIATIONS

GENERAL ABBREVIATIONS

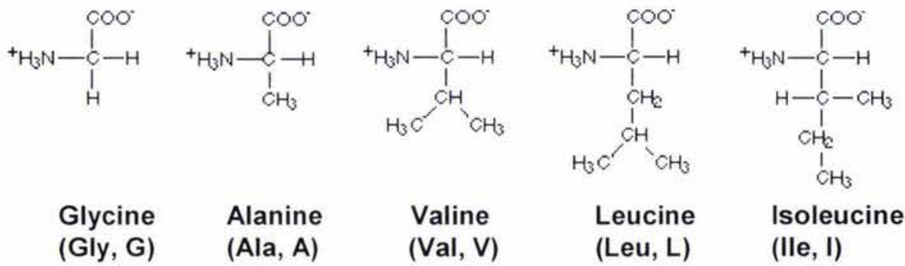
AA	Amino Acid
AC	Affinity Chromatography
ACE	Angiotensin Converting Enzyme
AIDS	Acquired Immune Disease Syndrome
AMoRe	Automated Molecular Replacement
AMP	Ampicillin
AMU	Atomic Mass Units
API	Atmospheric Pressure Ionisation
AS	Anomalous Scattering
ATP	Adenosine Tri-Phosphate
BH	Bleomycin Hydrolase
BLAST	Basic Local Alignment Search Tool
BTP	Bis-Tris Propane
CAD	Collect Assorted Data
CAPS	Cyclohexyl Amino Propane Sulfonic acid
CCP4	Collaborative Computational Project number 4
CCP4i	CCP4 graphical user Interface
CHES	2-Cyclo-Hexylamino Ethane Sulfonic acid
CID	Collision-Induced Dissociation
Da	Dalton
DNA	Deoxy-ribose Nucleic Acid
dNTP	Deoxyribo-Nucleotide Tri-Phosphate
Dpt	Di- and Tri-Peptide Transport System
DTT	Di-Thio-Threitol
EC	Enzyme Commission
EDTA	Ethylene-Di-amine-Tetra-Acetate
EPPS	4-(2-hydroxy-Ethyl)-1-Piperazine Propane Sulfonic acid
ESI	Electro-Spray Ionisation
FITC	Fluorescein Iso-Thio-Cyanate
FPLC	Fast Protein Liquid Chromatography
HCCA	α-Cyano-4-Hydroxy-trans-Cinnamic Acid
HEPES	2-(4-(2-Hydroxy-Ethyl)-1-Piperazinyl) Ethane Sulfonic acid
HIC	Hydrophobic Interaction Chromatography
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
IEX	Ion Exchange Chromatography

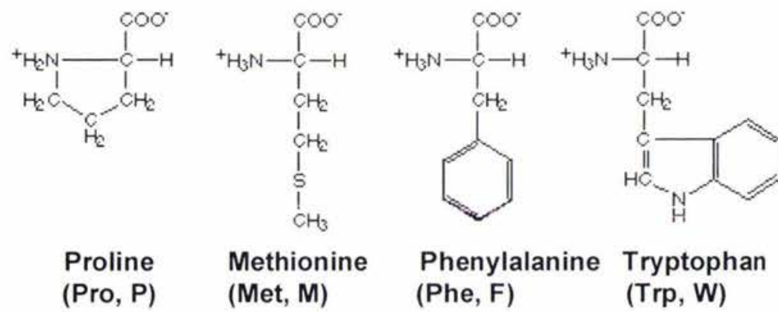
IMAC	Immobilised Metal Affinity Chromatography
IPTG	Iso-Propyl-beta-*D*-Thio-Galactopyranoside
I/σ	Signal-to-noise ratio
IUBMB	International Union of Biochemistry and Molecular Biology
LAB	Lactic Acid Bacteria
LB	Luria Bertani Media
LC	Liquid Chromatography
MALDI	Matrix Assisted Laser Desorption Ionisation
MCP	Micro-Channel Plate
MCS	Multiple Cloning Site
MES	2-(*N*-Morpholino) Ethane Sulfonic acid
MIR	Multiple Isomorphous Replacement
MIRAS	Multiple Isomorphous Replacement with Anomalous Scattering
MOPS	3-(*N*-MOrpholino) Propane Sulfonic acid
MPD	2-Methyl-2,3-Pentane-Diol
MR	Molecular Replacement
mRNA	Messenger Ribo-Nucleic Acid
MRS	Man, Rogosa and Sharpe Media
MS	Mass Spectrometry
MSMS	Tandem Mass Spectrometry
Mw	Molecular Weight
M/z	Mass-to-charge ratio
NEP	Neutral Endo-Peptidase
NCBI	National Centre for Biotechnology Information
NIZO	Netherlands Institute for Dairy Research
OD	Optical Density
Opp	Oligopeptide Transport System
ORF	Open Reading Frame
PAGE	Poly-Acrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Poly-Ethylene Glycol
Pep	Peptidase
pI	Isoelectric Point
PIPES	PIPerazine-1,4-bis 2-Ethane Sulfonic acid
PITC	Phenyl-Iso-Thio-Cyanate
PMSF	Phenyl-Methane Sulfonyl Fluoride
PPG	Poly-Propylene Glycol
Prt	Proteinase
PSI	Pound-force per Square Inch

PTH	Phenyl-Thio-Hydantoin
PVDF	Poly Vinyli-Dene Fluoride
RE	Restriction Endonuclease
RP	Reverse Phase
RNA	Ribo-Nucleic Acid
RNase	Ribo-Nuclease
RSM	Reconstituted Skim Milk
SDS	Sodium Dodecyl-Sulphate
SEC	Size Exclusion Chromatography
SIRAS	Single Isomorphous Replacement with Anomalous Scattering
TCA	Tri -Chloro-Acetic acid
TEV	Tobacco Etch Virus
TFA	Tri-Fluoro Acetic acid
TOF	Time Of Flight
TRIS	Tris-(hydroxymethyl)-aminomethane
Trx	Thioredoxin
UV	Ultra Violet

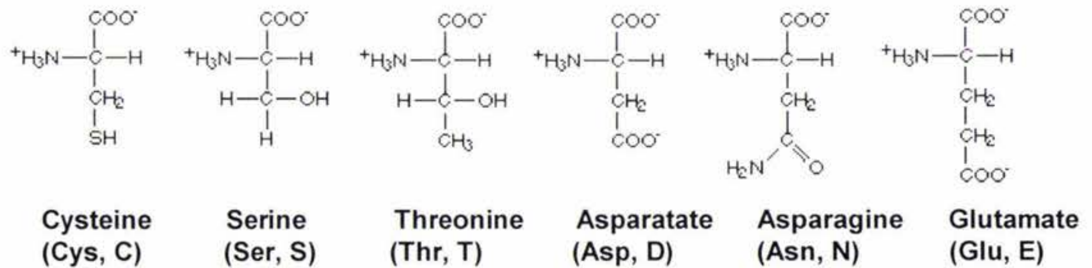
AMINO ACID ABBREVIATIONS

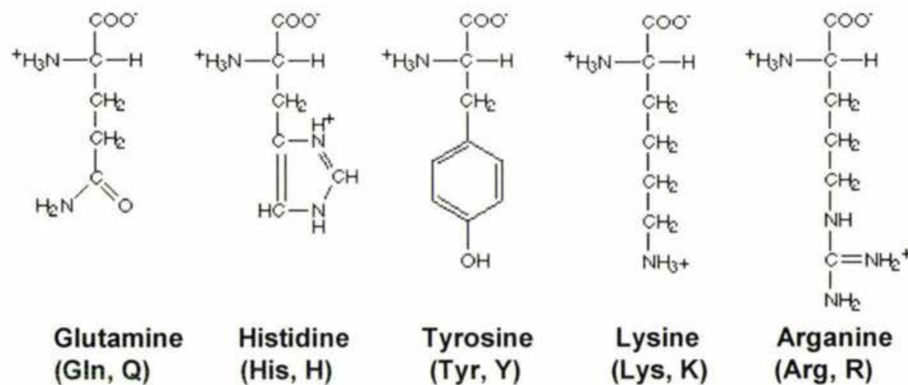
Amino acids with hydrophobic side groups:





Amino acids with hydrophilic side groups:





INTRODUCTION:

PROTEOLYTIC ENZYMES:
AN INTRODUCTION

PROTEOLYTIC ENZYMES: AN INTRODUCTION

1.1 PROTEOLYTIC ENZYMES

In recent years it has become clear that proteolytic enzymes mediate many important biological processes including: fertilization, cell division and differentiation, controlled cell death, blood coagulation and protein catabolism. Proteolytic enzymes are found in all organisms regardless of kingdom, and the importance of these enzymes is emphasized by the fact that approximately 2% of human genes encode proteolytic enzymes (Saklatvala 2003).

Proteolytic enzymes break peptide bonds linking the amino group of one amino acid with the carboxy group of an adjacent amino acid in a peptide chain (Figure 1.1).

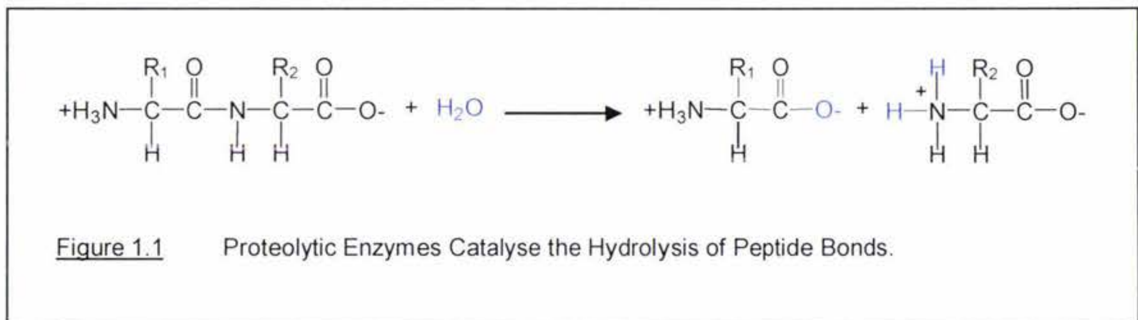


Figure 1.1 Proteolytic Enzymes Catalyse the Hydrolysis of Peptide Bonds.

1.1.1 CLASSIFICATION

Proteolytic enzymes require water to be present in the catalytic site to break the peptide bond and are thus classified as hydrolases (EC 3). All enzymes that catalyse the cleavage of any chemical bonds using water in the mechanism are termed hydrolases. This family of enzymes encompass esterases, lipases, peptidases and proteases. While the systematic name of a hydrolase enzyme always includes the term 'hydrolase', the common name is formed by the name of the substrate with the suffix -ase. The term "peptidase" is synonymous with "peptide hydrolase" and has been recommended by the International Union of Biochemistry and Molecular Biology (IUBMB) as the general term for all enzymes that hydrolyse peptide bonds (NC-IUBMB 1992).

Peptidases (EC 3.4) are primarily classified on the basis of their action as either exopeptidases or endopeptidases. Exopeptidases (EC 3.4.11-19) cleave peptide bonds near the amino terminus or the carboxy terminus of a peptide substrate, while endopeptidases (EC 3.4.21-99) hydrolyze peptide bonds distant from the amino and carboxy termini of the peptide substrate (NC-IUBMB 1992; Barrett 2004) (Table 1.1).

Table 1.1 Peptidase Classification According to Enzyme Commission and Reaction Catalysed.

Peptidase	EC	Reaction Catalysed [a]
Exopeptidases	3.4.11-19	
Aminopeptidases	3.4.11	X↓X(n)
Dipeptidases	3.4.13	X↓X
Dipeptidyl-peptidases, tripeptidyl-peptidases	3.4.14	X-X(X)↓X(n)
Peptidyl-dipeptidases	3.4.15	X(n)↓X-X
Serine-type carboxypeptidases	3.4.16	X(n)↓X
Metallocarboxypeptidases	3.4.17	X(n)↓X
Cysteine-type carboxypeptidases	3.4.18	X(n)↓X
Omega peptidases	3.4.19	Y↓X(n)
Endopeptidases	3.4.21-99	
Serine endopeptidases	3.4.21	X(n)↓X(n')
Cysteine endopeptidases	3.4.22	X(n)↓X(n')
Aspartic endopeptidases	3.4.23	X(n)↓X(n')
Metalloendopeptidases	3.4.24	X(n)↓X(n')
Threonine endopeptidases	3.4.25	X(n)↓X(n')
Endopeptidases of unknown catalytic mechanism	3.4.99	X(n)↓X(n')

[a] Reaction Catalysed: ↓, cleavage site; n, number of residues; X, amino acid; Y, amino acid that is cyclized, or linked by an isopeptide bond.

1.1.2 EXOPEPTIDASES

The exopeptidases can be further classified on the basis of the reaction catalysed. Exopeptidases that act at the amino-terminus of a peptide substrate can either release a single amino acid residue (aminopeptidases EC 3.4.11) or a dipeptide or a tripeptide (dipeptidyl-peptidases and tripeptidyl-peptidases EC 3.4.14). Those that act at the carboxy-terminus can also release either a single residue (carboxypeptidases EC 3.4.16-18) or a dipeptide (peptidyl-dipeptidases EC 3.4.15). Some exopeptidases are specific for certain dipeptides (dipeptidases EC 3.4.13) and some remove terminal residues that are cyclized, or linked by isopeptide bonds (peptide linkages other than those of α-amino to α-carboxyl groups) e.g. omega peptidases (EC 3.4.19) (Barrett 1992; NC-IUBMB 1992; Barrett 2004).

1.1.3 ENDOPEPTIDASES

Due to difficulties defining the substrate specificity of endopeptidases, an alternative classification system has been developed. In 1960 Hartley described four distinct types of catalytic mechanism used by peptidases; serine, cysteine, aspartic and metallo, and this observation developed into a practical system for peptidase classification (Hartley 1960; Barrett 1992). Peptidases are now classified under a hierarchical system built on the concepts of catalytic type, clan, family and peptidase (Barrett 2004; Rawlings 2004).

1.1.4 CATALYTIC TYPE

The concept of catalytic type of a peptidase depends upon the chemical nature of the groups responsible for catalysis. The currently recognized catalytic types are; Serine, Cysteine, Threonine, Aspartic, Metallo, Glutamic and some as yet unclassified. These catalytic types partition essentially into two catalytic mechanisms: covalent and non-covalent (general acid-base) catalysis (Salvesen 2003) (Figure 1.2).

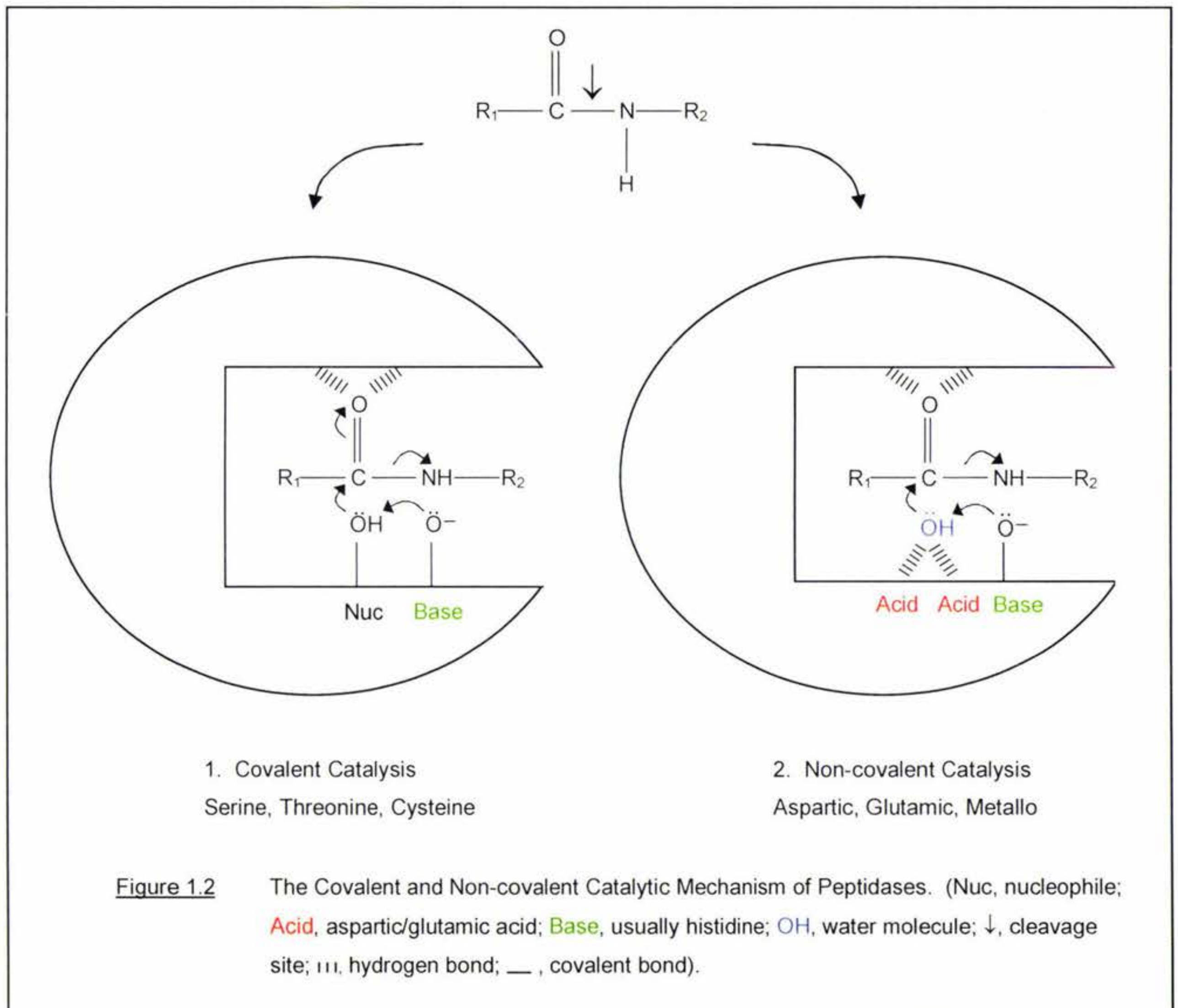


Figure 1.2 The Covalent and Non-covalent Catalytic Mechanism of Peptidases. (Nuc, nucleophile; Acid, aspartic/glutamic acid; Base, usually histidine; OH, water molecule; ↓, cleavage site; ııı, hydrogen bond; ___ , covalent bond).

The covalent catalytic mechanism of the serine, cysteine and threonine peptidases involves the nucleophilic attack of the scissile peptide bond of the substrate, by a nucleophilic side chain of the enzyme generating a covalent enzyme-substrate intermediate. The nucleophile that attacks the carbonyl carbon of the scissile peptide bond may be the sulfhydryl group of a cysteine (cysteine peptidases EC 3.4.22), or the hydroxyl group of a serine (serine peptidases EC 3.4.21) or threonine residue (threonine peptidases EC 3.4.25), with the nitrogen of the imidazole of histidine often functioning as a base (Barrett 2004).

The catalytic mechanism of the aspartic, glutamic and metallopeptidases involves a non-covalent intermediate and differs from the serine, cysteine and threonine peptidases in that the nucleophile that attacks the scissile peptide bond is an activated water molecule rather than the nucleophilic side chain of an amino acid (Barrett 2004).

The aspartic peptidases (EC 3.4.23) and recently discovered glutamic peptidases (also EC 3.4.23) have either an aspartic acid or a glutamic acid residue at the catalytic site, that activates the water molecule, while the metallopeptidases (EC 3.4.24) contain a divalent cation, usually zinc, at the active site centre which in conjunction with a glutamic acid residue activates the water molecule. The imidazole ring of histidine often functions as a base for these peptidases as it does for the serine, cysteine and threonine peptidases (Barrett 2004; Rawlings 2004).

The catalytic mechanism of a peptidase can often be predicted by amino acid sequence analysis; for example metallopeptidases often contain a conserved HEXXH pentapeptide that has been shown in crystallographic studies to coordinate a metal ion at the active centre (Barrett 1992; Barrett 2004). However, the catalytic mechanism can only be reliably determined by establishing the susceptibility of the peptidase to specific protease inhibitors.

1.1.5 CLAN, FAMILY AND PEPTIDASE

Peptidases are further classified under the concepts of clan, family and peptidase. Clans and families are groups of homologous peptidases; a clan contains one or more families that appear to have come from a common origin. The clearest kind of evidence for homology at the clan level is three-dimensional structural similarity, but amino acid sequence similarity can also be revealing. The name of a clan is formed from the initial letter of the catalytic type (S, serine; C, cysteine; T, threonine; A, aspartic; M, metallo; G, glutamic; U, unclassified) followed by a serial second capital letter. Each peptidase family is named with a letter denoting the catalytic type followed by a sequentially assigned number (Barrett 2004; Rawlings 2004).

A peptidase family is a group of peptidases, all of which display a particular kind of peptidase activity, and are closely related in amino acid sequence (Barrett 2004). Each member of the group is assigned a unique identifier constructed of two parts, each part being separated by a decimal point: these are the family name and an arbitrary three-digit serial number (Barrett 2004; Rawlings 2004).

Due to the rapid expansion of research on proteolytic enzymes an updatable reference resource containing summary data on all peptidases has been established. The MEROPS peptidase database (www.merops.ac.uk) was first published on the World Wide Web in 1996 and currently provides information on 177 peptidase families, 49 clans and 2018 unique peptidases (Rawlings 2004) (Table 1.2).

Table 1.2 Statistics for the MEROPS Database.

Catalytic Type	Total	Family	Clan	Peptidase	PDB Entry
Serine	9484	39	10	773	145
Threonine	977	4	1	57	37
Cysteine	4771	58	7	459	75
Aspartic	1501	12	6	159	36
Metallo	7581	53	15	552	84
Glutamic	10	1	1	4	1
Unknown	575	10	9	14	1
Total	24899	177	49	2018	379

1.1.6 NOMENCLATURE FOR PEPTIDASE SUBSTRATE SPECIFICITY

The system used to describe peptidase substrate specificity was introduced in 1970 by Berger and Schechter, who proposed that the peptidase catalytic site is flanked by specificity subsites, each able to accommodate the side chain of a single amino acid residue. These subsites are called S (for subsites) and are numbered from the catalytic site, S1 ... Sn towards the amino-terminus of the substrate, and S1′ ... Sn′ towards the carboxy-terminus. The peptide substrate residues that they accommodate are called P (for peptide) and are numbered P1 ... Pn, and P1′ ... Pn′, respectively (Berger 1970; Barrett 1992; Barrett 2004) (Figure 1.3).

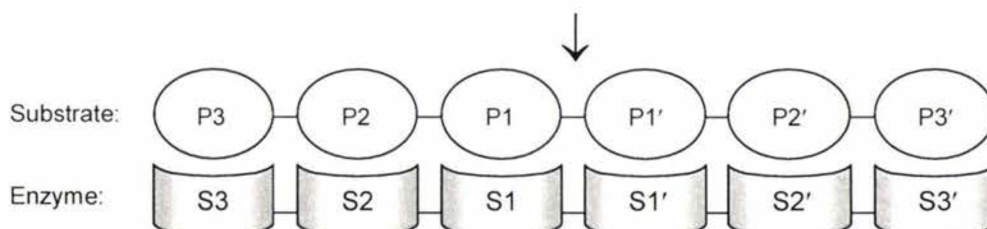


Figure 1.3 Nomenclature for Peptidase Substrate Specificity. (S, subsites; P, peptide residue; ↓, cleavage site)

1.2 PROTEOLYTIC ENZYMES IN BIOLOGY

Research into the proteolytic enzymes is rich and diverse. One of the earliest and best understood paradigms of proteolytic activity is zymogen activation. Many enzymes are synthesised as inactive precursors (zymogens) that are subsequently activated by cleavage of one or a few specific peptide bonds. Zymogen activation is responsible for a great variety of

physiological regulation, including protein digestion, and is particularly important when coupled to cascade activation as has been well documented for blood coagulation, apoptosis and the complement system (Neurath 1999).

Unlike the phosphorylation/dephosphorylation cascades, which are reversible, proteolytic cascades are irreversible, and therefore must be tightly regulated to prevent the serious deleterious effects of uncontrolled proteolytic activity (Amour 2004).

Deregulation of peptidase activity can have pathological effects and peptidases often have an important role in the development and manifestation of disease (Saklatvala 2003). The focus of current peptidase research is the search for inhibitors to a number of peptidases implicated in specific diseases. For example HIV peptidase inhibitors, and angiotensin converting enzyme (ACE) inhibitors have both been shown to be very effective drugs for the treatment of AIDS and cardiovascular disease, respectively (Pietzsch 2003).

By virtue of their ability to introduce restricted cleavages into a polypeptide substrate, peptidases are also valuable biochemical tools and have received a great deal of attention. In this context proteases and peptidases are also important in a number of industries particularly the dairy industry.

1.3 LACTIC ACID BACTERIA

Lactic acid bacteria (LAB) are gram-positive bacteria that are used in the dairy industry for the production of a variety of fermented milk products, including cheese and yoghurt. The principal role of LAB in the production of cheese and yoghurt is to ferment the lactose in milk to lactic acid, reducing the pH to approximately 4.5-5. This low pH prevents the growth of undesirable micro-organisms and contributes to casein precipitation, the initial step in cheese production (Reid 1994b).

It has been well established that many LAB, isolated from milk products, are multiple amino acid auxotrophs and therefore must obtain essential amino acids from the growth medium (Kunji 1996). The quantities of free amino acids and small peptides present in milk are not sufficient to support the growth of LAB to a high cell density. Subsequently, these bacteria have developed complex proteolytic systems to obtain essential amino acids from caseins, the primary proteins in milk.

The proteolytic systems of LAB have received an enormous amount of attention due to their importance in both the physiology of the bacteria themselves, and because of their involvement in the maturation and flavour development of cheese.

Proteolysis contributes to the taste of cheese through the production of low-molecular weight peptides and free amino acids. A number of free amino acids, mainly those with nonpolar or hydrophobic side chains, exhibit a bitter flavour (e.g. isoleucine, tryptophan, phenylalanine), as do many small peptides which exhibit flavours dependent on their hydrophobicity and amino acid composition (Lemieux 1992). Certain sequences in the caseins, including α_{s1}-casein(1-9) and β-casein(193-209), are particularly hydrophobic and, when excised by proteolytic enzymes result in bitter flavours in dairy products which are undesirable (Sousa 2001; Yvon 2001).

Although the direct contribution of peptides and amino acids to flavour is probably limited to a background flavour, they are precursors of many other flavour compounds. Conversion of amino acids through deamination, transamination, and decarboxylation to various amines, aldehydes, acids, and sulfur-containing compounds results in the development of different flavours (Sousa 2001). Flavour development in cheese is of major economic interest since the final flavour of cheeses partly determines consumer choice and because the development of flavour is a time consuming and an expensive process. An understanding of the proteolytic enzymes involved in LAB proteolysis is necessary to enable controlled synthesis of flavour compounds through judicious choice of LAB organisms. In the long term, the development of genetically modified LAB strains with improved fermentation properties, and substrate specificity will provide the ultimate tool for the control of flavour in various dairy products.

1.3.1 LACTIC ACID BACTERIA PROTEOLYTIC NETWORK

Research on the LAB proteolytic network has focussed to a large extent on the well characterised LAB, *Lactococcus lactis*. Most, if not all, of the components of the *L. lactis* proteolytic network have been identified and the majority of the enzymes have been purified and biochemically characterised. Furthermore the genetics of the corresponding genes have been studied (Stepaniak 2004).

The *L. lactis* proteolytic network is composed of several closely related but catalytically distinct cell-envelope proteinases (PrtP) which are involved in the initial cleavage of casein; cell membrane transport systems (Opp, DtpT, DtpP) that are involved in the uptake of small peptides and amino acids; and intracellular peptidases (Pep) that hydrolyse the casein-derived peptides (Kunji 1996; Christensen 1999) (Figure 1.4).

1.3.2 CELL ENVELOPE PROTEINASE

The *L. lactis* cell-envelope proteinases, PrtPs, are large monomeric serine endopeptidases with molecular weights of approximately 200 kDa. They are associated with the cell envelope through a membrane anchor and a cell wall-spanning region and are sometimes known as lactocepins (lactococcal cell envelope-associated proteinase) (Barrett 2004).

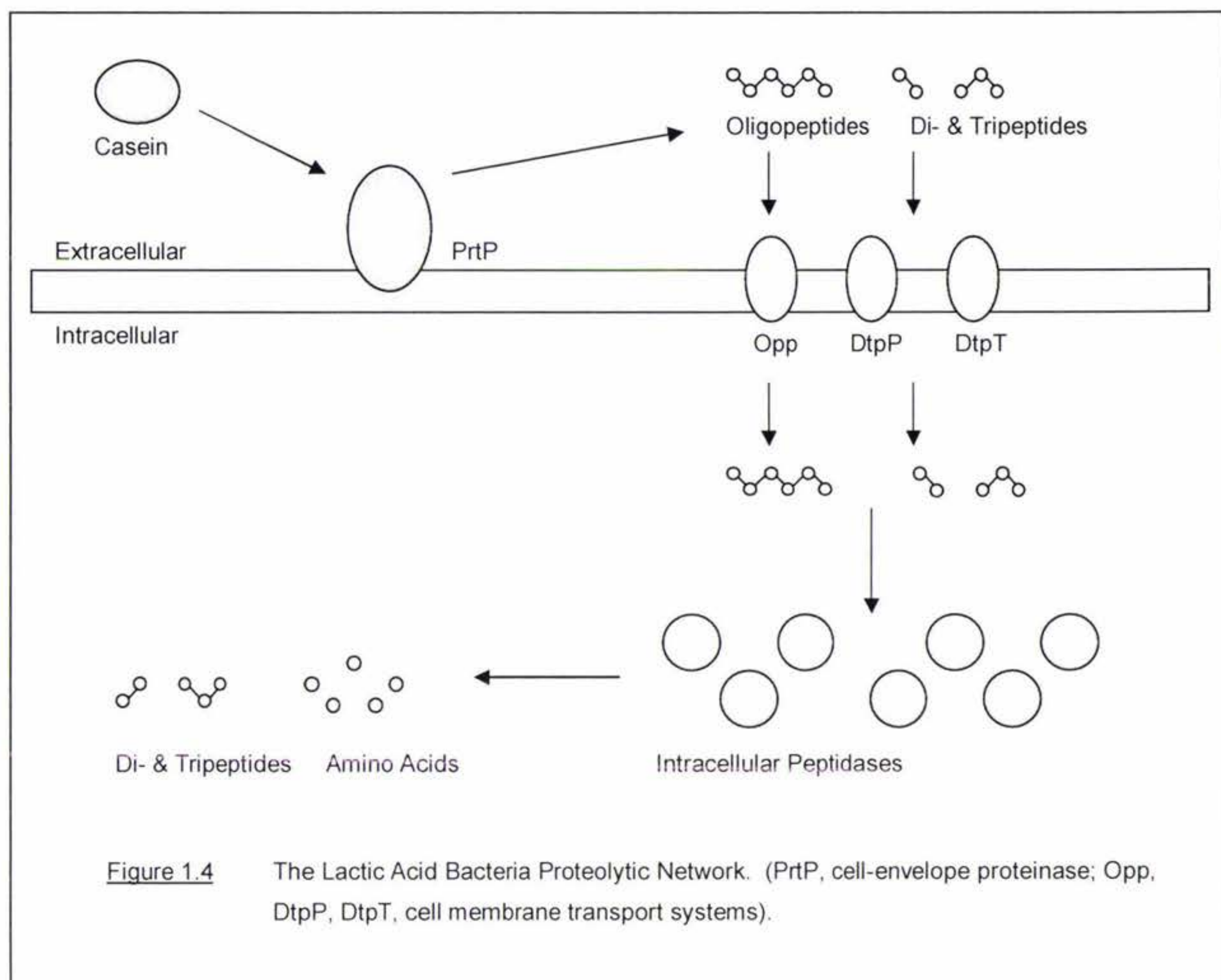


Figure 1.4 The Lactic Acid Bacteria Proteolytic Network. (PrtP, cell-envelope proteinase; Opp, DtpP, DtpT, cell membrane transport systems).

Individual strains of *L. lactis* produce only one form of cell-envelope proteinase, type I, type III or one of intermediate specificity. Both PI- and PIII-type peptidases hydrolyse β- and κ-caseins, although with different cleavage specificities, while α_{s1}-casesin is hydrolysed efficiently only by PIII-type peptidases (Barratt 2004). The enzyme is stabilised by Ca^{2+} (Exterkate 1999), and in the absence of Ca^{2+}, PrtP undergoes autoproteolysis and is released from the cell envelope. Whether associated with the cell envelope or not, PrtP is able to hydrolyse casein into small molecular weight peptides that can then be taken up by the cell via cell membrane transport systems (Barrett 2004).

1.3.3 CELL MEMBRANE TRANSPORT SYSTEMS

L. lactis has separate transport systems for amino acids, di- and tripeptides (DtpT and DtpP), and oligopeptides (Opp) in the cell membrane. The oligopeptide transport system (Opp) mediates the ATP-driven transport of peptides with four or more residues, and plays a central role in the proteolytic network. The di- and tripeptide transport systems DtpT and DtpP are also important components of this network, with DtpP preferentially transporting di- and tripeptides composed of hydrophobic, branched-chain amino acid residues, and DtpT transporting more-hydrophilic peptides (Foucaud 1995; Barrett 2004).

1.3.4 INTRACELLULAR PEPTIDASES

Following transport into the cell, the casein-derived peptides are further hydrolysed by intracellular peptidases. There are at least ten intracellular peptidases in the *L. lactis* proteolytic network, the majority of which are exopeptidases (Barrett 2004; Stepaniak 2004) (Table 1.3).

Table 1.3 *Lactococcus lactis* Intracellular Peptidases.

Peptidase	Molecular Mass [a] (kDa)	Structure [b]	Class [c]	Substrate [d]
Expopetidase				
Aminopeptidase PepA	38 - 45	Hexamer, trimer	M	Asp/Glu↓X(n)
Aminopeptidase PepC	50 - 54	Hexamer	C	X↓X(n)
Aminopeptidase PepN	89 - 95	Monomer	M	X↓X(n)
Dipeptidase PepV	46 - 52	Monomer	M	X↓X
Tripeptidase PepT	23 - 55	Monomer, dimer	M	X↓X-X
Prolidase PepQ	41 - 43	Monomer	M	X↓Pro
Prolinase PepP	43	Monomer	M	X↓Pro-X(n)
X-prolyl dipeptidyl aminopeptidase PepX	82 - 117	Monomer, dimer	S	X-Pro↓X(n)
Endopeptidase				
Endopeptidase PepO	70 - 71	Monomer	M	X(n)↓X(n')
Endopeptidase PepF	70	Monomer	M	X(n)↓X(n')

[a] Molecular Mass: calculated from known amino acid sequence or SDS-PAGE.

[b] Structure: predicted from comparison of gel filtration value with either amino acid sequence or SDS-PAGE values, or x-ray crystal analysis.

[c] Catalytic class: C, cysteine; M, metallo; S, serine peptidases.

[d] Substrate: ↓, cleavage site; X, amino acid.

1.3.4.1 EXOPEPTIDASES

L. lactis exopeptidases include the aminopeptidases (PepA, PepC, PepN, PepT, PepP), the dipeptidases (PepV, PepQ) and the dipeptidyl peptidase (PepX). To date, no carboxypeptidases have been identified in this organism.

A distinguishing feature of PepA is its thermal stability. Independent research groups have shown that PepA from *L. lactis* subsp. *lactis* and subsp. *cremoris* retain greater than 50% activity after incubation at 70°C for 1 hour (Niven 1991; Bacon 1994, Barrett 2004).

In contrast to PepA, the aminopeptidases PepC and PepN have very broad substrate specificities. PepN although it has a very broad substrate specificity, shows a marked preference for substrates containing Arg as the N-terminal residue (Tan 1991; Barrett 2004).

PepC, a cysteine aminopeptidase that is inhibited by SH-group reagents such as iodoacetamide, is active only in a reducing environment (Barrett 2004). While exhibiting a broad specificity, it cannot hydrolyse peptide bonds in which a proline residue is involved. The crystallographic structure of *L. lactis* PepC has recently been solved (Barrett 2004; Mistou unpublished results), revealing a hexameric ring structure which has a prominent central channel large enough to accommodate a peptide. The active site of each monomer has a papain-like fold and is situated within the central channel (Barrett 2004). Insertion of the C-terminus of each PepC monomer into its own active site leaves sufficient space for only one substrate residue, the P1 residue of the substrate, to enter the active site, and appears to be the underlying molecular basis for the strict aminopeptidase activity of *L. lactis* PepC (Mata 1999; Barrett 2004). Site directed mutagenesis, in which the last four C-terminal residues were deleted, converted the enzyme into an endopeptidase (Mata 1999) to prove this hypothesis.

Recently, two peptidases that have about 55% identity in their amino acid sequences to *L. lactis* PepC and that have a natural deletion of these C-terminal residues have been identified in *Lactobacillus delbrueckii* subsp. *lactis*. These have been named PepG and PepW (Courtin 2002). *L. delbrueckii* PepG has been shown to preferentially hydrolyse Leu rich di- and tripeptides from the amino terminal end of peptides, PepW in contrast, has been shown to be a strict dipeptidyl-peptidase. To date only one PepW cleavage site has been identified, Gly-Leu↓Leu-Gly (Courtin 2002).

PepP is proline specific and exclusively hydrolyses oligopeptides of up to 10 residues containing X↓Pro-Pro-(X)n or X↓Pro-(X)n sequences at their N-termini. No hydrolysis of di- and tripeptides with similar sequences occurs (Barrett 2004).

PepT is capable of removing the N-terminal residue from tripeptides, although it is unable to hydrolyse the peptide bonds of di- or oligopeptides. PepT is capable of hydrolysing tripeptides with Pro in either position P1 or P2', but is unable to cleave peptides having Pro in the P1′ position (Barrett 2004).

The specificity of *L. lactis* PepV is strictly confined to dipeptides, and it prefers dipeptides with N-terminal hydrophobic residues (Christensen 1999).

PepQ activity is defined by a specificity for X-Pro peptides, where X can be any residue (Christensen 1999; Barrett 2004).

PepX, is a serine peptidase that is active as either a monomer or dimmer releasing X-Pro dipeptides from the N-terminus of oligopeptides. In addition to peptidase activity, PepX has esterase activity. The crystallographic structure of *L. lactis* subsp. *lactis* PepX has recently been solved showing the enzyme to be a dimmer related by 2-fold symmetry (Rigolet 2002).

The crystal structure of *L. lactis* PepX shows novel structural elements that could explain the selectivity of the enzyme for proline-containing peptides, and may help to explain its ability to function not only as a peptidase but also as an esterase (Rigolet 2002).

1.3.4.2 ENDOPEPTIDASES

To date only two distinct endopeptidases from *L. lactis*, PepF and PepO have been reported. Two copies of the *pepF* gene have been located in *L. lactis*; *pepF1* (located on a large plasmid of *L. lactis* subsp. *cremoris*) and *pepF2* (located on the chromosome of *L. lactis* subsp. *lactis*). They have 84% sequence identity and have a broad endopeptidase activity for peptides ranging in length from 7 to 23 amino acids (Barrett 2004).

PepO is able to cleave peptides ranging in length from 5 to at least 30 amino acids, although it is incapable of hydrolysing di- and tripeptides, or intact α-, β- or κ-casein (Tan 1991; Pritchard 1994; Baankreis 1995; Stepaniak 1995; Barrett 2004). While PepF has broad specificity, PepO has preference for peptide bonds with hydrophobic amino acids in the P1′ position (Barrett 2004).

Interestingly, the *pepO* gene is located immediately downstream of the genes encoding the oligopeptide transport system (Mierau 1993), which suggests that the two systems may be physiologically linked. However, culture of a *L. lactis pepO* mutant, generated by gene disruption, showed that PepO was not essential for the survival of the organism (Mierau 1993; Law 1997).

Single and multiple *L. lactis* peptidase knockouts have been constructed for the *pepA, pepC, pepN, pepT, pepX, pepV, pepO* and *pepF* genes. In general, no differences were reported for the growth rate of either single or multiple knockouts compared to the wild type strain when cells were grown in complex media (MRS or M17 broth) or complete amino acid defined media. When milk was the nutrient, however, there were noticeable reductions in growth rate for most of these mutants, which increased as the number of genes knocked out increased (Mierau 1996; Christensen 1999). This observation shows that individual peptidases are not essential for the survival of the organism, and also makes it possible to analyse the effect of each enzyme on cheese ripening and flavour development.

The effect of peptidases on cheese ripening and flavour development has also been evaluated by over-expressing individual peptidases. McGarry *et al.* (1994) and Christensen *et al.* (1995) studied the influence of *L. lactis* PepN on Cheddar cheese using *Lactococcus* strains engineered to overproduce PepN. No significant changes in cheese ripening or flavour development were observed (McGarry 1994; Christensen 1999). In contrast, Meyer and Spahni

(1998) who studied the influence of *L. lactis* PepX on Gruyere cheese found that PepX had an influence on cheese ripening and flavour development (Meyer 1998; Sousa 2001).

Flavour development in cheese is of major economic interest since the final flavour of cheeses partly determines consumer choice and because the development of flavour is a time consuming and an expensive process. For this reason, there is incentive to identify LAB peptidases with novel substrate specificity that could potentially be used to modify and control cheese bitterness and possibly other flavour attributes.

1.4 *LACTOBACILLUS RHAMNOSUS* PEPO

The *Lactobacillus rhamnosus pepO* gene is approximately 1.9-kbp long and encodes a monomeric 70.9 kDa enzyme. The amino acid sequence contains the HEXXH-pentapeptide characteristic of zinc metallopeptidases, and has 36% and 27% amino acid identity to *L. lactis* PepO and *H. sapiens* neutral endopeptidase (NEP), respectively (Christensson 2002) (Figure 1.6).

Although similar in sequence, the substrate specificity of *L. lactis* PepO and *L. rhamnosus* PepO is considerably different. *L. rhamnosus* PepO has been shown to hydrolyse peptide bonds located in the N-terminal part of α_{s1}-casein(1-23), whereas *L. lactis* PepO hydrolyses peptide bonds located within the C-terminal portion of this peptide (Christensson 2002) (Figure 1.5).

```
αs1-casein(1-23)        R-P-K-H-P-I-K-H-Q-G-L-P-Q-E-V-L-N-E-N-L-L-R-F
L. rhamnosus PepO                ↑   ↑ ↑ ↑
L. lactis PepO                                   ↑ ↑         ↑ ↑
```

Figure 1.5 Cleavage Specificity of *L. rhamnosus* PepO and *L. lactis* PepO for α_{s1}-casein(1-23). (↑, cleavage site).

The unique substrate specificity of *L. rhamnosus* PepO has ramifications for flavour development in cheese. The peptides α_{s1}-casein(1-13) and α_{s1}-casein(1-17) are known to exhibit a very bitter flavour, and have been found to accumulate during cheese ripening (Christensson 2002). Because PepO is able to hydrolyse peptide bonds within these regions, it has the potential to reduce bitterness by preventing the accumulation of the peptide.

In order to understand the molecular basis for the unique substrate specificity of PepO, it is necessary to determine the three-dimensional structure of the enzyme. The crystal structure of *H. sapiens* neutral endopeptidase (NEP), which shares 27% amino acid identity with PepO, has

been solved (Oefner 2000). NEP (EC 3.4.24.11) also known as neprilysin, enkephalinase, CALLA and CD10, is a monomeric 80-92 kDa integral membrane glycosylated metallopeptidase that is involved in the physiological degradation of the peptides modulating blood pressure.

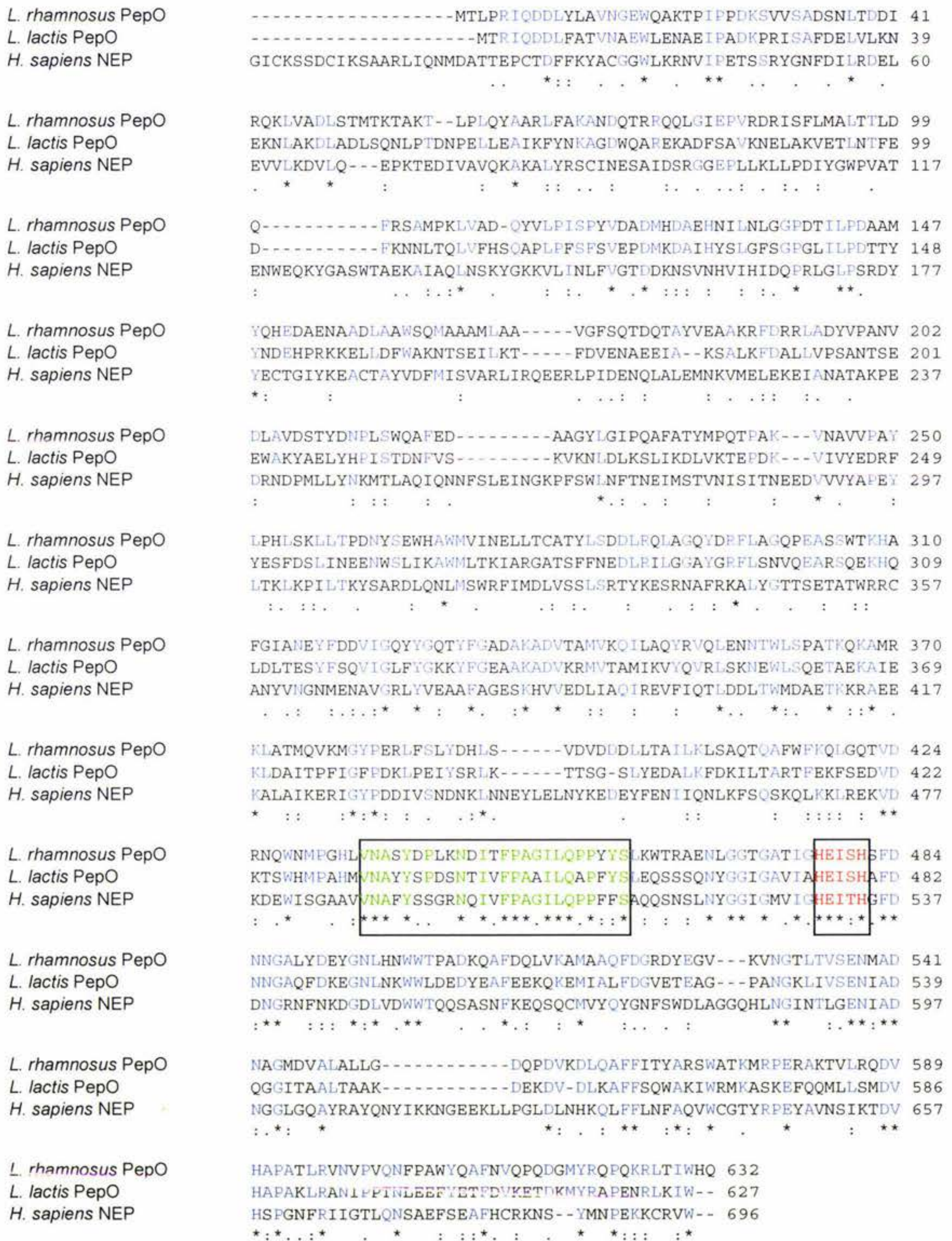


Figure 1.6 Sequence Alignment of PepO from *L. rhamnosus*, *L. lactis* and *H. sapiens* NEP. (Blue, identical residues; Green, highly conserved residues; Red, HEXXH petapeptide motif; Boxed, highly conserved domain; *, conserved residues; :, very similar residues;., similar residues. Generated using Clustal W, Chenna 2003).

The extracellular domain (residues 52-749), of NEP was crystallised in complex with the generic metallopeptidase inhibitor, phosphoramidon, and its structure solved to 2.1 Å resolution (Oefner 2000). The structure (Figure 1.7) reveals two largely α-helical domains which embrace a large central cavity containing the active site.

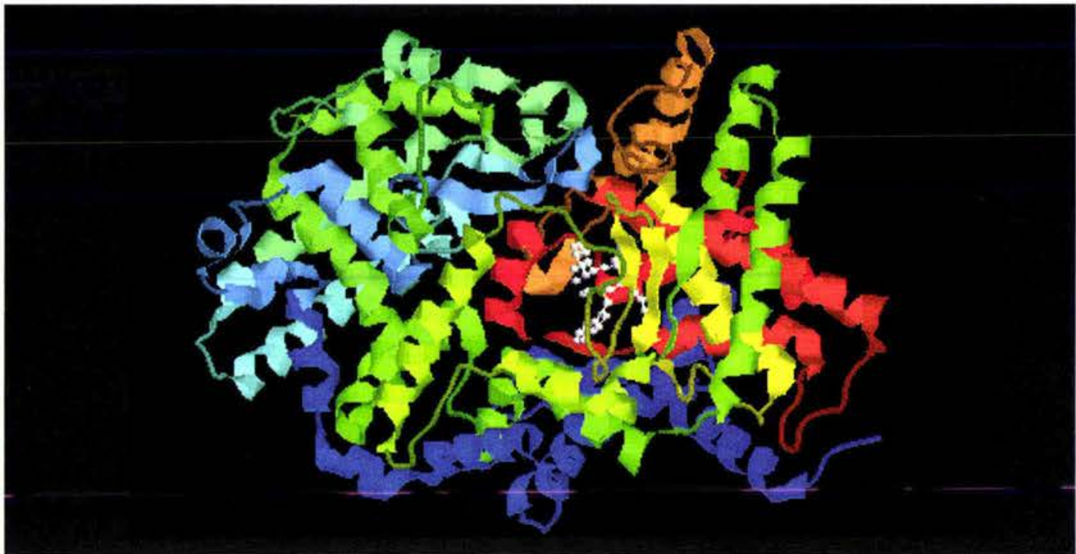


Figure 1.7 The Three-Dimensional Structure of *H. sapiens* NEP. (Cartoon representation coloured by residue from the N-terminus (blue) to the C-terminus (red). The phosphoramidon molecule is represented as a ball and stick model in white. Generated using Rasmol, Sayle 1995).

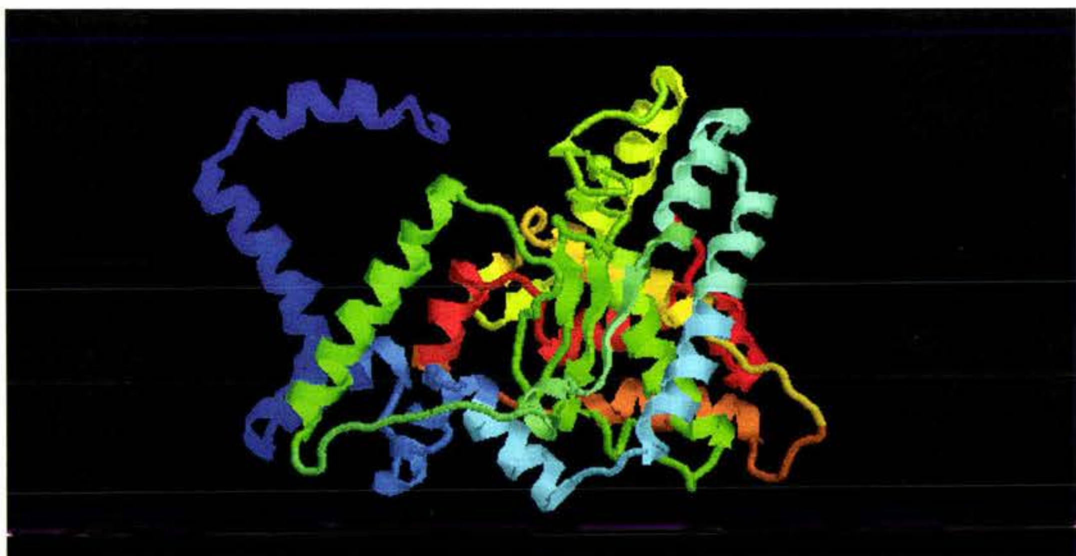


Figure 1.8 The Swiss-Model Predicted Three-Dimensional Structure of PepO. (Cartoon representation coloured by residue from the N-terminus (blue) to the C-terminus (red). Generated using Rasmol, Sayle 1995).

The inhibitor is bound to one side of this cavity and its binding mode provides a detailed understanding of the complex ligand-binding and substrate specificity determinants (Oefner 2000). The C-terminal domain of NEP shares 27% amino acid identity with PepO. Although structural investigations often require an atomic model with > 30% sequence identity, it is possible that this protein could aid in determining the three-dimensional structure of PepO. The Swiss-Model predicted three-dimensional structure of PepO using NEP as an atomic model is illustrated in Figure 1.8 (Schwede 2003).

1.5 *LACTOBACILLUS RHAMNOSUS* PEPW

Recently the *L. rhamnosus* HN001 genome has been sequenced, and analyses have revealed the presence of a 1.347 kb gene, *pepW*, encoding a putative protein of 448 amino acids, that has been named PepW (Lubbers unpublished results). PepW is 93%, 52%, 44% and 37% identical to aminopeptidase C (PepC) from *Lactobacillus casei*, *Lactococcus lactis*, and bleomycin hydrolase from *H. sapiens* (hBH) and *S. cerevisiae* (Gal6), respectively (Figure 1.9).

The name PepW was given due to incorrect sequence analysis of the *pepW* gene, which suggested that the putative peptidase, PepW, was missing the C-terminal residues needed to be classified as a PepC. Despite further analysis showing that the gene did indeed code these residues, the name of the gene and its protein product have been retained.

Although the substrate specificity of PepW is completely unknown, it is likely to have broad specificity, like the closely related LAB PepC peptidases (Barrett 2004). Substrate specificity, however, will only be unequivocally established experimentally using a number of different substrates. It may be informed by a three dimensional structure, which would allow comparisons with the structures of the yeast and human bleomycin hydrolases, which have been solved to 2.2 Å (Joshua-Tor 1995) and 2.6 Å resolution, respectively (O'Farrell 1999; Barrett 2004).

Bleomycin hydrolase (EC 3.4.22.40) is a highly conserved cysteine peptidase from bacteria to humans that hydrolyzes the anticancer drug bleomycin (Joshua-Tor 1995). The enzyme is an intracellular homohexameric cysteine peptidase of approximately 300 kDa (Barrett 2004). The structures of the *H. sapiens* and *S. cerevisiae* bleomycin hydrolases are very similar overall, adopting an unusual hexameric ring structure which has a prominent central channel of ~22 A diameter (Joshua-Tor 1995; O'Farrell 1999; Barrett 2004). The six papain-like active sites are situated within the central channel, in a manner resembling the organization of active sites in the proteasome (Joshua-Tor 1995). The three-dimensional structure of the yeast bleomycin hydrolase is shown in Figure 1.10.

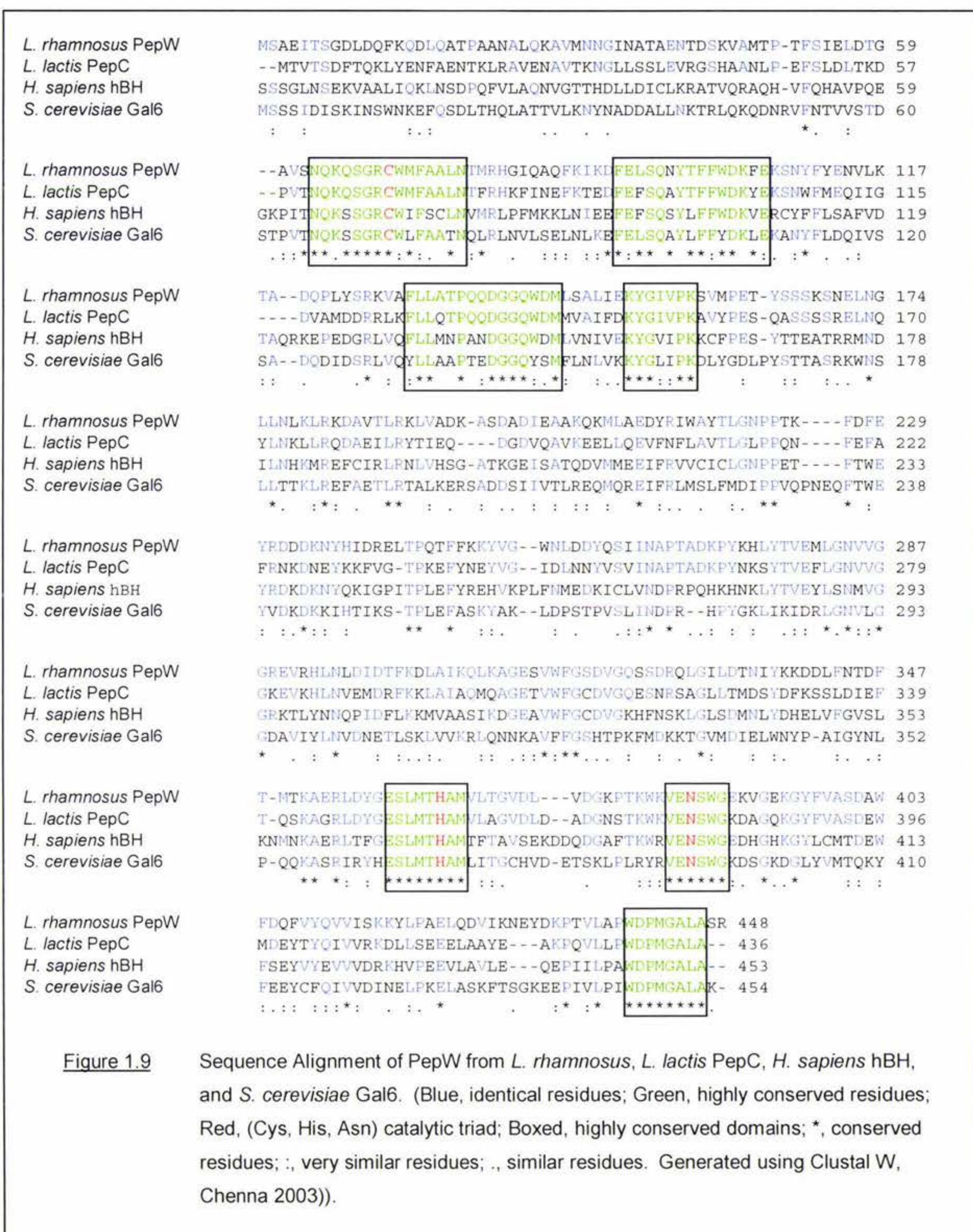


Figure 1.9 Sequence Alignment of PepW from *L. rhamnosus*, *L. lactis* PepC, *H. sapiens* hBH, and *S. cerevisiae* Gal6. (Blue, identical residues; Green, highly conserved residues; Red, (Cys, His, Asn) catalytic triad; Boxed, highly conserved domains; *, conserved residues; :, very similar residues; ., similar residues. Generated using Clustal W, Chenna 2003)).

A striking feature of the molecule is the projection of the C-terminus of the protein into the active-site cleft. Intriguingly, the yeast bleomycin hydrolase has DNA-binding activity, and has been shown to bind to the upstream activating sequences (UAS) of the GAL system (hence the designation Gal6) (Joshua-Tor 1995). *L. lactis* PepC has also been shown to bind DNA (Xu 1994; Barrett 2004; Xu unpublished results), although human bleomycin hydrolase does not have this nucleic acid binding activity (Barrett 2004).

The Swiss-Model predicted three-dimensional structure of PepW obtained using a combination of native and mutant *H. sapiens* and *S. cerevisiae* bleomycin hydrolase structures as atomic models is illustrated in Figure 1.11.

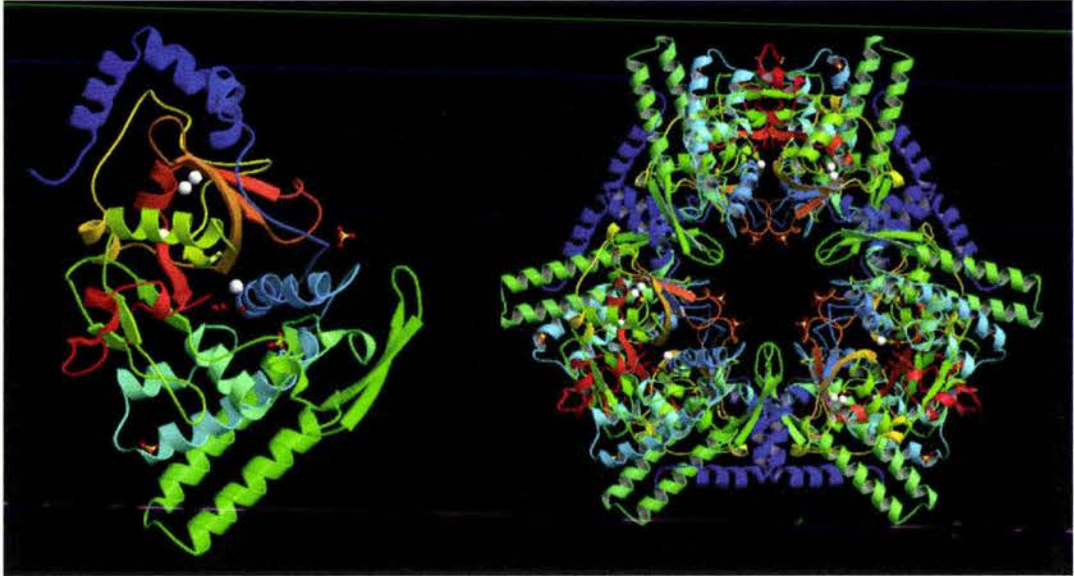


Figure 1.10 The Three-Dimensional Structure of the *S. cerevisiae* Bleomycin Hydrolase. (Cartoon representation coloured by residue from the N-terminus (blue) to the C-terminus (red). The mercury atoms (Hg) are represented as space-fill models in white. A monomer is shown on the left, with the hexamer being modelled based on crystallographic symmetry on the right, Berman 2002).

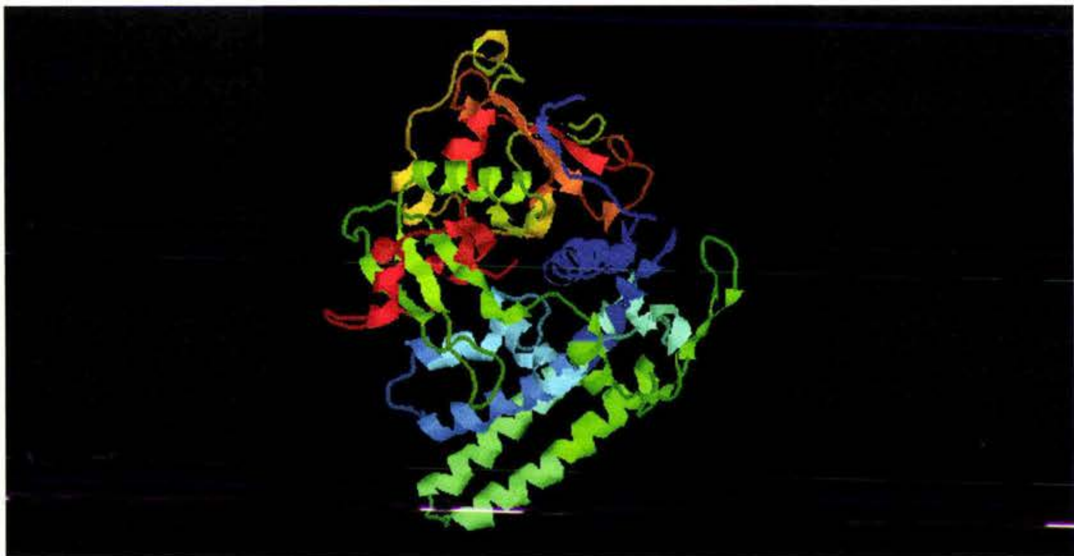


Figure 1.11 The Swiss-Model Predicted Three-Dimensional Structure of PepW. (Cartoon representation coloured by residue from the N-terminus (blue) to the C-terminus (red). Generated using Rasmol, Sayle 1995).

1.6 PROJECT AIMS

This research aims to investigate the substrate specificity and three-dimensional structure of the *L. rhamnosus* peptidases, PepO and PepW.

Aims:

1. Clone the *pepW* gene into a suitable expression system using standard molecular cloning techniques.

2. Isolate and purify recombinant PepO and PepW using traditional chromatographic methods for the former, and methods that exploit the purification tag of the latter.

3. Investigate the substrate specificity of PepO and PepW towards casein-derived substrates, using reverse-phase high pressure liquid chromatography, mass-spectrometry and protein sequencing techniques.

4. Grow suitable crystals of PepO and PepW for structure determination using x-ray diffraction.

5. Solve the three-dimensional structure of PepO and PepW using molecular replacement and/or multiple isomorphous replacement techniques.

Information obtained from this research will contribute to our fundamental understanding of the role of specific structural motifs for peptidases in general and in particular the *L. rhamnosus* peptidases, PepO and PepW. The structural and functional characterisation of *L. rhamnosus* PepO and PepW could in the future be used to inform the genetic engineering of proteolytic enzymes designed to develop new cheese flavours.