

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.

**EXPRESSION, PURIFICATION AND  
CHARACTERISATION OF RECOMBINANT  
PEPTIDE:N-GLYCOSIDASE F.**

A thesis presented in partial fulfilment of the requirements for  
the degree of Master of Philosophy in Biochemistry  
at Massey University, New Zealand.

**Trevor Stephen Loo**

**2000**

“I have not failed. I've just found 10,000 ways that won't work.”

**Thomas Edison.**

**ABSTRACT**

PNGase F (Peptide- $N^4$ -( $N$ -acetyl-D-glucosaminyI) asparagine amidase F) is an amidohydrolase isolated from the extracellular medium of the Gram-negative bacterium *Flavobacterium meningosepticum*. The 34.8-kDa enzyme catalyses the complete and intact cleavage of asparagine-linked oligosaccharide chains from their associated proteins. A T7 promoter-based *E. coli* expression system was developed in which PNGase F was expressed as a fusion protein with a leader sequence from the *ompA* gene. The hexa-histidine-tagged PNGase F was correctly processed and exported to the *E. coli* periplasm and had a calculated molecular weight of 36.2 kDa. A single step purification using immobilised metal affinity chromatography yielded 8 mg of pure protein per litre of culture.

The sequence of the PNGase F coding region from the CDC strain 3352 of *F. meningosepticum* was found to differ from a published sequence from another strain of the bacterium (ATCC 33958) in 57 positions. These differences between the two strains result in eight amino acid substitutions, which are mostly conservative in nature and are on the surface of the protein. Moreover, three potential  $N$ -glycosylation sites not present in the ATCC strain 33958 were detected in CDC strain 3352.

The recombinant enzyme has similar characteristics of the native enzyme with a pH optimum of 8.5 and is strongly inhibited by  $\text{Ag}^+$ ,  $\text{Cu}^{2+}$ , and  $\text{Fe}^{3+}$  ions but not by sulfhydryl-targeting agents such as DTT and NEM. This indicates inhibition by these ions is probably through interactions with a histidine residue at position 193 that may be involved in substrate recognition or catalysis. The specific activity of the native PNGase F is about four times that of the recombinant protein which may be contributed to inhibition by components of the Complete™ protease inhibitor tablets used in the enzyme preparation or due to modifications for cloning and purification. Using a discontinuous assay and a non-labelled 11-mer ovalbumin-derived glycopeptide as substrate, a rough estimate of the Michaelis constant ( $K_m$ ) for the recombinant PNGase F was determined to be 2.1  $\mu\text{M}$ . An intriguing observation with the activity assays was the apparent product inhibition of enzyme activity and the inhibitor may be either peptide and/or glycan components, which require further investigations into the cause of the inhibition.



## ACKNOWLEDGEMENTS

The many staff and students at the Institute of Molecular Biosciences to be thanked for their help and support. The following people deserve special mention:

My supervisor Dr. Gillian Norris for her invaluable advice, infinite patience and enthusiasm throughout the course of my study. Sincere thanks also to my co-supervisors Dr. Mark Patchett and Dr. Shaun Lott for their advice, encouragement and support.

I would also wish to thank my friends in X-lab: Deborah Frumau for her assistance, Julian Adams for his advice and modelling 3D structures, and F. Y. Chai for his humour and assistance. Dr. Cristina Weinberg for her encouragement as a caring friend who was forever patient with my questions, who kept me company late at night, and made me eat healthy.

Joanne Mudford for her friendship and giving me a “crash course” about her job before moving to Dunedin. Carmen Norris and David Elgar from NZDRI for giving me permission to use their HPLC system, Dr. Robert Norris for proof-reading my thesis, Associate Professor Geoff Jamerson for his generosity in equipment purchase and consumables, Associate Professor David Harding and Dick Poll for their valuable advice on chemistry and equipment loans, Dr. Catherine Day and Mrs. Carole Flyger for advice and assistance when I was working on my expression constructs, and Professor Pat Sullivan for his interest and encouragement for this work.

I should like to thank my long-suffering flatemate Desmond for enduring the chaos and his infallible efforts to keep the house habitable during house renovations and my long absences in writing of this manuscript.

Finally, thanks to my family who has always encouraged me to pursue my interests regardless how bizarre they seemed.

## TABLE OF CONTENTS

<b>ABSTRACT .....</b>	<b>i</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>ii</b>
<b>TABLE OF CONTENTS .....</b>	<b>iii</b>
<b>LIST OF FIGURES .....</b>	<b>xii</b>
<b>LIST OF TABLES .....</b>	<b>xvii</b>
<b>ABBREVIATIONS .....</b>	<b>xviii</b>
<b>ABBREVIATIONS OF AMINO ACIDS .....</b>	<b>xxi</b>
<b>ABBREVIATIONS OF SUGARS .....</b>	<b>xxiii</b>
 <b>Chapter 1 Introduction and Literature Review .....</b>	 <b>1</b>
1.1 Protein Glycosylation .....	1
1.1.1 Carbohydrates .....	1
1.1.2 Glycoproteins .....	3
1.1.2.1 The Nature of Glycoproteins .....	3
1.1.2.2 Types of Glycosidic Linkages .....	4
1.1.2.2.1 Glycosylphosphatidylinositol (GPI) Anchors .....	5
1.1.2.2.2 N-linked Glycosylation .....	5
1.1.2.2.3 O-linked Glycosylation .....	8
1.1.2.3 Complexity and Diversity of Glycans .....	9
1.1.2.4 Potential Roles of Protein Glycosylation .....	10
1.2 Deglycosylation .....	12
1.2.1 Tools to Study the Roles of Glycans in Glycoprotein Functions .....	12

1.2.1.1 Chemical Deglycosylation .....	12
1.2.1.2 Enzymatic Deglycosylation .....	12
1.2.2 Endo <i>N</i> -acetyl- $\beta$ -D-glucosaminidases (ENGases) .....	13
1.2.3 Peptide- <i>N</i> <sup>4</sup> -( <i>N</i> -acetyl- $\beta$ -D-glucosaminyI) asparagine amidase (PNGases) .....	15
1.2.3.1 The <i>in vivo</i> Functions of PNGases .....	19
1.2.4 Other PROXlases .....	20
1.2.5 Glycosidases from <i>Flavobacterium meningosepticum</i> .....	20
1.2.5.1 Endo F <sub>1</sub> .....	21
1.2.5.2 Endo F <sub>2</sub> and F <sub>3</sub> .....	21
1.2.5.3 PNGase F .....	21
1.2.5.3.1 Properties of PNGase F .....	21
1.2.5.3.2 Substrate Structure Requirements of PNGase F .....	22
1.2.5.3.3 The Three Dimensional Structure of PNGase F ..	23
1.2.5.3.4 The Active Site of PNGase F .....	25
1.2.6 Constructs of PNGase F .....	27
1.2.6.1 The First Clones of PNGase F .....	28
1.2.6.2 PNGase F Cloned as GST Fusions .....	29
1.2.6.3 A Clone of PNGase F from <i>F. meningosepticum</i> (CDC strain 3352) .....	29
1.2.6.4 Inclusion Bodies Formation in <i>E. coli</i> .....	29
1.3 Targeting and Assembly of Proteins in the Bacterial Periplasm .....	31
1.3.1 The Bacterial Periplasm .....	31
1.3.2 The Signal Peptide .....	32
1.3.3 The Sec Export Pathway .....	32
1.3.4 Protein Folding in the Periplasm .....	33

1.4 The Scope of this Project .....	35
<b>Chapter 2 Construction of an Expression Vector for the Production of PNGase F in <i>E. coli</i> .....</b>	<b>37</b>
2.1 Introduction .....	37
2.2 Experimental Objectives and Strategies .....	37
2.3 Materials .....	39
2.3.1 Chemicals and Enzymes .....	39
2.3.2 Plasmids and Bacterial Strains Used in This Study .....	40
2.3.3 Bacterial Growth Media .....	40
2.3.3.1 M9 with 0.5% Casamino Acids .....	41
2.3.3.2 LB Broth .....	41
2.3.3.3 SOB Medium .....	41
2.3.4 Storage and Propagation of Bacterial Cultures .....	41
2.4 Methods used for Cloning .....	42
2.4.1 Phenol:Chloroform Extraction of DNA from a DNA/Protein Mixture .....	42
2.4.2 Ethanol Precipitation of DNA .....	42
2.4.3 Size-Selective Polyethylene Glycol (PEG) Precipitation of DNA .....	42
2.4.4 Agarose Gel Electrophoresis .....	43
2.4.5 Quantitation and Size Determination of DNA Fragments ...	43
2.4.6 Amplification of PNGase F gene by PCR .....	43
2.4.7 Digestion of DNA with Restriction Endonucleases .....	44
2.4.8 Purification of DNA from Agarose Gels .....	44
2.4.9 Ligation of DNA Fragments .....	45

2.4.10 Preparation of Competent <i>E. coli</i> Cells .....	45
2.4.11 Transformation of Ligated Plasmids into Competent <i>E. coli</i> Cells .....	46
2.4.12 Small-Scale Preparation of Plasmid DNA .....	46
2.4.13 Sequence Analysis of DNA .....	47
2.5 Results .....	48
2.5.1 Amplification of PNGase F Gene from pT7-PNG by PCR ..	49
2.5.2 Restriction Enzyme Digest of the PNGase F DNA .....	53
2.5.3 Ligation of the PNGase F Gene into the Vector .....	55
2.5.4 Transformation of Competent <i>E. coli</i> Cells with Ligation Products .....	56
2.6 Discussion and Conclusions .....	59
 <b>Chapter 3 Expression and Purification of PNGase F ..</b>	 <b>63</b>
3.1 Experimental Objectives .....	63
3.2 Materials .....	63
3.3 Methods used for Expression and Analysis of PNGase F .....	66
3.3.1 General Methods .....	66
3.3.1.1 Determination of Protein Concentration .....	66
3.3.1.1.1 Alkaline Copper (Lowry) Protein Assay .....	66
3.3.1.1.2 Bicinchoninic Acid (BCA) Protein Assay .....	67
3.3.1.1.3 Coomassie Blue (Bradford) Protein Assay .....	67
3.3.1.1.4 UV Method .....	68
3.3.1.2 Calculation of PNGase F Extinction Coefficient .....	68

3.3.1.3 Determination of PNGase F Activity with a Reverse Phase High Performance Liquid Chromatography (RP-HPLC) Assay .....	69
3.3.1.3.1 Mechanism of Assay .....	69
3.3.1.3.2 Preparation of Samples .....	71
3.3.1.3.3 Determination of Specific Activity of PNGase F ...	73
3.3.1.4 Polyacrylamide Gel Electrophoresis (PAGE) .....	75
3.3.1.5 Electroblotting of Proteins from Acrylamide Gels .....	75
3.3.1.6 N-Terminal Sequencing .....	76
3.3.1.7 Isoelectric Focusing (IEF) of Native and Recombinant PNGase F .....	77
3.3.1.8 Electrospray Ionisation Mass Spectrometry .....	77
3.3.1.9 Cell Growth and Protein Production Studies .....	78
3.3.2 Expression of PNGase F .....	78
3.3.2.1 Growth and Induction of <i>F. meningosepticum</i> Culture .....	78
3.3.2.2 Growth and Induction of <i>E. coli</i> Culture .....	79
3.3.2.2.1 Preparation of Periplasmic Fraction from <i>E. coli</i> ..	79
3.3.3 Purification of PNGase F .....	80
3.3.4 Chromatographic Methods .....	79
3.3.4.1 Hydrophobic Interaction Chromatography (HIC) .....	81
3.3.4.1.1 Hydrophobic Interaction Chromatography with Phenyl Sepharose 6 Fast-Flow (Low Substitution) .....	82
3.3.4.1.2 HIC with <i>t</i> -butyl-TSK .....	83
3.3.4.1.3 HIC Tests with Various Hydrophobic Matrices ....	83
A. Pharmacia HiTrap™ Test Kit .....	83
B. Macro-Prep <i>t</i> -butyl HIC Econo-Pac Cartridges .....	84
C. TSK-butyl-Toyopearl 650M .....	84

D. Alkyl Sepharose HR 10/10 .....	84
3.3.4.1 Ion Exchange Chromatography (IEX) .....	84
3.3.4.2.1 IEX Test with various Ion Exchangers .....	86
A. Mono-Q HR5/5 & Uno-Q .....	86
B. Uno-S Polishing Column .....	86
C. Hydroxyapatite .....	87
3.3.4.3 Size Exclusion Chromatography (SEC) with Superdex 75 .....	87
3.3.4.4 Purification of Recombinant PNGase F using Immobilised Metal Affinity Chromatography (IMAC) ...	89
3.3.4.5 Size Exchange Chromatography with Superdex 75 ...	91
3.4 Results and Discussion .....	92
3.4.1 Verification of the Extinction Coefficient of Recombinant PNGase F .....	92
3.4.2 Comparison of Protein Quantitation Methods .....	93
3.4.3 Growth and Protein Production Studies of Transformed <i>E. coli</i> Cells .....	95
3.4.4 Purification of Native PNGase F .....	100
3.4.4.1 HIC with Phenyl Sepharose 6 Fast-Flow (Low Substitution) .....	100
3.4.4.2 HIC with Home-Made <i>t</i> -butyl Substituted TSK .....	102
3.4.4.3 HIC Tests with various Hydrophobic Matrices .....	103
3.4.4.3.1 Phenyl Sepharose 6 Fast-Flow (High Substitution) .....	103
3.4.4.3.2 Butyl Sepharose 4 Fast-Flow .....	105
3.4.4.3.3 Macro-Prep <i>t</i> -butyl Econo-Pac .....	106
3.4.4.3.4 TSK-butyl-Toyopearl 650M .....	107
3.4.4.3.5 Alkyl Sepharose .....	108

3.4.4.4 Ion Exchange Tests with various Ion Exchangers .....	111
3.4.4.4.1 Anion Exchange Chromatography .....	111
A. Mono-Q at pH 8.9 .....	111
B. Mono-Q at pH 9.8 .....	113
3.4.4.4.2. Uno-Q .....	114
A. Uno-Q at pH 8.9 .....	114
B. Uno-Q at pH 9.8 .....	115
3.4.4.4.3 Cation Exchange Chromatography: Uno-S at pH 7.2 .....	116
3.4.4.4.4 Hydroxyapatite Column .....	118
3.4.4.5 HIC with Newly Synthesised <i>t</i> -butyl TSK .....	120
3.4.4.6 SEC with Superdex 75 .....	121
A. Superdex 75 HR 16/60 Prep-Grade Resin .....	121
B. Superdex 75 HR 10/30 Super Fine Resin .....	122
C. Superdex 75 HR 10/30 Super Fine Resin in the presence of Thesit .....	123
3.4.5 Purification of Recombinant PNGase F .....	124
3.4.5.1 IMAC Test of various Chelating Resins .....	125
A. Boehringer Mannheim Poly-His Purification Resin Chelated with $\text{Zn}^{2+}$ .....	125
B. Boehringer Mannheim Poly-His Purification Resin Chelated with $\text{Ni}^{2+}$ .....	125
C. Pharmacia HiTrap™ Chelating Resin Chelated with $\text{Zn}^{2+}$ .....	126
D. Pharmacia HiTrap™ Chelating Resin Chelated with $\text{Ni}^{2+}$ .....	127
E. Elution Profile of Recombinant PNGase F from $\text{Ni}^{2+}$ Charged HiTrap™ Chelating Resin .....	128



3.4.5.2 SEC of Recombinant Enzyme with Superdex 75 in the presence of 0.006% Thesit .....	128
3.5 Conclusion .....	135
<b>Chapter 4 Characterisation of Recombinant PNGase F</b> .....	<b>136</b>
4.1 Introduction .....	136
4.2 Experimental Objectives .....	136
4.3 Materials .....	137
4.4 Methods used in Characterisation .....	138
4.4.1 Comparison of Substrate Specificity .....	138
4.4.1.1 Detection of Deglycosylated Substrate with Digoxigenin (DIG) Glycan/Protein Double Labelling kit .....	139
4.4.1.1.1 Principle of Detection .....	139
4.4.2 Characterisation of Recombinant PNGase F .....	140
4.4.2.1 Effects of Temperature .....	140
4.4.2.2 Effects of pH .....	141
4.4.2.3 Effects of Metal ions and Additives .....	141
4.4.2.4 Shelf-Life .....	142
4.4.3 Determination of the Michaelis Constant ( $K_m$ ) of Native and Recombinant PNGase F .....	142
4.5 Results and Discussion .....	144
4.5.1 Comparison of Substrate Specificity .....	144
4.5.2 Characterisation of Recombinant PNGase F .....	149
4.5.2.1 Effects of Temperature .....	149
4.5.2.2 Effects of pH .....	150
4.5.2.3 Effects of Metal Ions and Additives .....	151
4.5.2.5 Shelf-Life of Recombinant PNGase F .....	152

4.5.3 Determination of the Michaelis Constant ( $K_m$ ) of Native and Recombinant PNGase F .....	156
4.5.3.1 Experimental Strategy .....	156
4.5.3.2 Results .....	156
4.5.3.2 Shortcomings of the Discontinuous Activity Assay .....	159
4.5.3.2.1 Limited Availability of the Substrate .....	159
4.5.3.2.2 Low Detection Sensitivity of the Assay .....	159
4.5.3.2.3 Other Difficulties of the Assay .....	159
4.5.4 Alternative Methods to Measure Activity .....	160
4.5.4.1 Oxidation of NADPH by <i>L</i> -Glutamic Dehydrogenase .....	160
4.5.4.1.1 Principle of Detection .....	160
4.5.4.1.2 Sample Preparation .....	161
4.5.4.1.3 Results .....	161
4.5.4.2 Real Time Measurement of $K_m$ using Surface Plasmon Resonance (SPR) Technology .....	162
4.5.4.2.1 Principle of Detection .....	163
4.5.4.2.2 Application of SPR to the $K_m$ Determination of PNGase F .....	164
4.5.4.3 Kinetic Analysis of Ribonuclease B Products by ES-MS .....	164
4.5.4.4 Pre-Kinetic Analysis of PNGase F Deglycosylation Reaction by ES-MS .....	166
<b>4.6 Conclusions</b> .....	167
<b>Chapter 5 Final Conclusion</b> .....	170
<b>REFERENCES</b> .....	173
<b>APPENDIX</b> .....	193

## LIST OF FIGURES

Figure 1.1	The different configurations of monosaccharides .....	2
Figure 1.2	Condensation reaction between two monosaccharides .....	3
Figure 1.3	The $\beta$ -asparatylglucosylamine link of a glycosylated asparagine	5
Figure 1.4	Complex type N-linked glycan .....	7
Figure 1.5	High mannose type N-linked glycan .....	7
Figure 1.6	Hybrid type N-linked glycan .....	8
Figure 1.7	The glycosidic bond cleaved by ENGases .....	14
Figure 1.8	The two step cleavage reaction catalysed by PNGases .....	18
Figure 1.9	Topology of the PNGase F molecule .....	23
Figure 1.10	The PNGase F structure .....	24
Figure 1.11	The orientation of <i>N,N'</i> -diacetylchitobiose (CTB) inside the active site of PNGase F .....	26
Figure 1.12	Schematic diagram showing the intermolecular hydrogen bonding contacts between PNGase F, <i>N,N'</i> -diacetylchitobiose, and water molecules .....	27
Figure 1.13	Structure of PNGase F signal sequence .....	28
Figure 1.14	A model for protein translocation across the inner membrane ...	33
Figure 1.15	Model of disulfide formation catalysed by the various Dsb enzymes in the periplasm of <i>E. coli</i> .....	34
Figure 2.1	Map of pKS-OmpA3-His .....	38
Figure 2.2	An inverse image of a 12% SDS-PAGE with silver staining shows the solubility problem of recombinant PNGase F expressed in <i>E. coli</i> .....	48
Figure 2.3	Cloning strategy for PNGase F .....	50
Figure 2.4	Synthetic oligonucleotides used for PCR amplification .....	51
Figure 2.5	PNGase F amplified from pT7-PNG using PCR .....	52
Figure 2.6	Double digested vector and PNGase F DNA .....	53
Figure 2.7	Schematic diagram showing predicted PCR products using various primer pairs on a successfully ligated plasmid .....	54
Figure 2.8	PCR confirmation of successful ligation .....	55

Figure 2.9	Plasmid miniprep of XL1-Blue transformed with pOPH6 .....	58
Figure 2.10	Possible nucleotide misincorporation by <i>Taq</i> polymerase during the amplification of CDC strain 3352 PCR product .....	61
Figure 3.1	PNGase F cleaves the amide bond in the ovalbumin glycopeptide .....	68
Figure 3.2	Formation of the homoserine lactone (Hsl) ring .....	69
Figure 3.3	Chromatogram of a typical deglycosylation activity assay .....	71
Figure 3.4	A summary of the purification scheme for PNGase F .....	90
Figure 3.5	Growth curve of <i>E. coli</i> transformed with pOPH6 .....	94
Figure 3.6	A schematic diagram of the <i>E. coli</i> BL21(DE3) cell showing IPTG induction of PNGase F synthesis and subsequent secretion of the nascent polypeptide into the periplasm .....	96
Figure 3.7	PNGase F production in induced cells at different time points ..	97
Figure 3.8	An inverse image of a 12% SDS-PAGE with silver staining shows the protein solubility problem of pKS-PNG is resolved by secretion in pKS-OPH6 .....	98
Figure 3.9	Standard curved used to estimate the molecular weight of recombinant PNGase F .....	99
Figure 3.10	Chromatogram of the elution from Phenyl Sepharose6 Fast-Flow (low substitution) of <i>F. meningosepticum</i> culture medium .....	100
Figure 3.11	SDS-PAGE of the PNGase F fractions from Phenyl Sepharose 6 Fast-Flow (low substitution) .....	100
Figure 3.12	Chromatogram of PNGase F on <i>t</i> -butyl substituted TSK .....	101
Figure 3.13	SDS-PAGE of fractions from <i>t</i> -butyl substituted TSK .....	101
Figure 3.14	Chromatogram of elution of PNGase F from Phenyl Sepharose 6 Fast-Flow (high substitution) .....	103
Figure 3.15	SDS-PAGE of fractions form Phenyl Sepharose 6 Fast-Flow (high substitution) .....	103
Figure 3.16	Chromatogram of elution from Butyl Sepharose 4 Fast-Flow ...	105
Figure 3.17	SDS-PAGE analysis on active fractions from Butyl Sepharose 4 Fast-Flow .....	105
Figure 3.18	Chromatogram of elution from <i>t</i> -butyl Econo-Pac .....	106
Figure 3.19	SDS-PAGE of active fractions from Bio-Rad <i>t</i> -butyl Econo-Pac .....	106
Figure 3.20	Chromatogram of elution from TSK-butyl-Toyopearl .....	107
Figure 3.21	SDS-PAGE analysis of active fractions from TSK-butyl-Toyopearl .....	108

Figure 3.22	Chromatogram of elution of PNGase F from Alkyl Sepharose ..	109
Figure 3.23	SDS-PAGE analysis of fractions from Alkyl Sepharose .....	109
Figure 3.24	SDS-PAGE analysis of native and recombinant PNGase F used for IEF .....	110
Figure 3.25	Isoelectrical focusing of native and recombinant PNGase F ....	111
Figure 3.26	Chromatogram of elution from Mono-! At pH 8.9 .....	112
Figure 3.27	SDS-PAGE of active fractions form Mono-Q at pH 8.9 .....	112
Figure 3.28	Chromatogram of elution from Mono-Q at pH 9.8 .....	113
Figure 3.29	SDS-PAGE of active fractions from Mono-Q at pH 9.8 .....	113
Figure 3.30	Chromatogram of elution from Uno-Q at pH 8.9 .....	114
Figure 3.31	SDS-PAGE analysis of factions from Uno-Q at pH 8.9 .....	115
Figure 3.32	Chromatogram of elution from Uno-Q at pH 9.8 .....	115
Figure 3.33	SDS-PAGE analysis of the fractions from Uno-Q at pH 9.8 ..	116
Figure 3.34	Chromatogram of elution from Uno-S at pH 7.2 .....	117
Figure 3.35	SDS-PAGE analysis of fractions from Uno-S at pH 7.2 .....	117
Figure 3.36	Chromatogram of elution from hydroxyapatite .....	118
Figure 3.37	SDS-PAGE analysis of fractions from hydroxyapatite .....	118
Figure 3.38	Charged groups distribution on the surface of native PNGase F	119
Figure 3.39	Chromatogram of elution from new <i>t</i> -butyl substituted TSK ...	120
Figure 3.40	SDS-PAGE analysis of fractions from new <i>t</i> -butyl TSK .....	121
Figure 3.41	Chromatogram of elution from Superdex 75 HR 16/60 .....	121
Figure 3.42	SDS-PAGE analysis of fractions from Superdex 75 HR 16/60 .	122
Figure 3.43	Chromatogram of elution from Superdex 75 HR 10/30 .....	122
Figure 3.44	SDS-PAGE analysis of fractions from Superdex 75 HR 10/30 .	123
Figure 3.45	Chromatogram of elution from Superdex 75 HR 10/30 with Thesit .....	123
Figure 3.46	An inverse image of a silver stained SDS-PAGE of fractions from Superdex 75 HR 10/30 with Thesit .....	124
Figure 3.47	SDS-PAGE of fractions from Zn <sup>2+</sup> charged Poly-His resin .....	125
Figure 3.48	SDS-PAGE of fractions from Ni <sup>2+</sup> charged Poly-Hs resin .....	126
Figure 3.49	SDS-PAGE of fractions from Zn <sup>2+</sup> charged HiTrap™ chelating resin .....	127

Figure 3.50	An inverse image of a silver stained SDS-PAGE of fractions from $\text{Ni}^{2+}$ charged HiTrap™ chelating resin .....	127
Figure 3.51	Elution profile from $\text{Ni}^{2+}$ charged Pharmacia chelating resin ...	128
Figure 3.52	Chromatogram of elution from Superdex 75 HR 10/30 .....	129
Figure 3.53	An inverse image of a SDS-PAGE stained with silver staining method shows the purity of recombinant PNGase F from Superdex 75 HR 10/30 .....	129
Figure 3.54	The reciprocal relationship of substrate depletion and product formation in the activity assay .....	130
Figure 3.55	The linear relationship of product area and substrate concentration in the activity assay .....	130
Figure 4.1	Schematic representation of the detection principle of DIG glycan/protein double labelling kit .....	140
Figure 4.2	Lineweaver-Burk plot with error bars of $\pm 0.05V$ .....	142
Figure 4.3	Hanes-Woolf plot with error bars of $\pm 0.05V$ .....	142
Figure 4.4	A 12% SDS-PAGE of the deglycosylated products of fetuin by recombinant PNGase F under different buffering conditions after 48 hours of incubation at 37°C .....	145
Figure 4.5	The deglycosylated products of different glycoproteins used in DIG glycan/protein double labelling kit .....	146
Figure 4.6	The detection of glycans and proteins in deglycosylated products of various glycoproteins using DIG glycan/protein double labelling kit .....	147
Figure 4.7	The comparison of glycoprotein deglycosylation by native and recombinant PNGase F .....	148
Figure 4.8	Temperature profile of recombinant PNGase F .....	149
Figure 4.9	pH profile of recombinant PNGase F .....	150
Figure 4.10	Structures of Caps and Capso buffers .....	150
Figure 4.11	The effects of metal ions on recombinant PNGase F activity ...	151
Figure 4.12	The effects of additives on recombinant PNGase F activity .....	152
Figure 4.13	Initial estimate of the $K_m$ of recombinant PNGase F with ovalbumin-derived glycopeptide in small reaction volume .....	157
Figure 4.14	$K_m$ determination of dialysed recombinant PNGase F in a large reaction volume .....	158
Figure 4.15	The curve generated for calculation of $K_m$ by the program Enzfitter.....	158

Figure 4.16	A continuous calorimetric assay for measuring PNGase F activity through the oxidation of NADPH by <i>L</i> -glutamic dehydrogenase ( <i>L</i> -GLDH).....	160
Figure 4.17	Chromatogram from the <i>L</i> -GLDH coupling assay.....	162
Figure 4.18	Detection of biomolecular binding events by SPR.....	163
Figure 4.19	Understanding the sensorgram.....	163
Figure 4.20	The sensor chip NTA.....	164
Figure 4.21	The spectrum for ribonuclease B product quantitation by ES-MS	165
Figure 4.22	The spectrum for ovalbumin glycopeptide product quantitation by ES-MS.....	166
Figure 4.23	Structures of Capso, glycerol, and isopropanol.....	168



## LIST OF TABLES

Table 1.1	Some consequences of genetic defects or polymorphisms in Glycosylation .....	4
Table 1.2	Some biological roles of oligosaccharides .....	11
Table 1.3	The occurrence of some endo- <i>N</i> -acetyl- $\beta$ -D-glucosaminidases (ENGases) .....	14
Table 1.4	The occurrence of some peptide- <i>N</i> <sup>4</sup> -( <i>N</i> -acetyl- $\beta$ -D-glucosaminyl) asparagine amidase (ENGases) .....	15
Table 1.5	Characteristics of known purified PNGases .....	16
Table 1.6	The occurrence of other PROXIases .....	20
Table 2.1	Colony count from transformation of pOPH6 ligation mix into <i>E. coli</i> XL1-Blue cells .....	57
Table 2.2	Amino acid variation between ATCC strain 33958 and CDC strain 3352 .....	59
Table 3.1	Chromatographic supports used in purification .....	64
Table 3.2	PNGase activity assay gradient programme .....	70
Table 3.3	The nine glycoforms of hen egg white ovalbumin .....	73
Table 3.4	Summary of purification of native and recombinant PNGase F .....	132
Table 4.1	List of buffers used in the activity test .....	141
Table 4.2	Protease inhibitor set (Roche) was suspected to have components similar to those of the Complete™ protease inhibitor tablets .....	154
Table 4.3	The effects of long-term storage on recombinant PNGase F activity .....	155



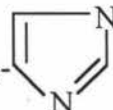
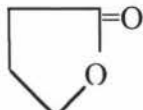
## ABBREVIATIONS


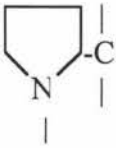
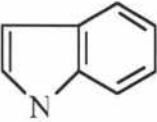
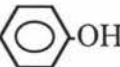
amu	Atomic mass units
Amp	Ampicillin
ATCC 33958	American Type Culture Collection No. 33958; sequence from this strain was published by Tarentino <i>et al.</i> , 1990.
AUFS	Absorbance units at full scale
BSA	Bovine serum albumin
BCA	Bicinchoninic acid
Caps	3-Cyclohexylamino-1-propanesulfonic acid
Capso	3-Cyclohexylamino-2-hydroxy-1-propanesulfonic acid
CDC strain 3352	United States Communicable Disease Centre culture collection strain 3352; the strain used in this study.
CDI	1,1'-carbonyldiimidazole
cfu	Colony forming units
CIAP	Calf intestinal alkaline phosphatase
CNBr	Cyanogen bromide
CTB	<i>N,N'</i> -diacetylchitobiose
Dabsyl	4-(dimethylamino)-azobenzene-4'-sulfonyl
DIG	Digoxigenin
DNA	Deoxyribose nucleic acid
DNTPs	Deoxyribose nucleotide triphosphates
DMSO	Dimethyl sulphoxide
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid (di-sodium salt)
EGCase	Endoglycoceramidase
ENGase	Endo- <i>N</i> -acetyl- $\beta$ -D-glucosaminidase or endoglycosidase
EPPS	N-[2-Hydroxyethyl]piperazine-N'[3-propanesulfonic acid]
ER	Endoplasmic reticulum
ES-MS	ElectroSpray Mass Spectrometry
EtBr	Ethidium bromide
FPLC	Fast protein liquid chromatography
L-GLDH	L-Glutamic dehydrogenase
GdmHCl	Guanidine hydrochloride

GPI	Glycosylphosphatidylinositol
HCl	Hydrochloric acid
IMAC	Immobilised metal ion affinity chromatography
IPTG	Isopropyl-1-thio- $\beta$ -D-galactopyranoside
kb	kilo base pairs
kDa	kilo daltons
$\alpha$ -KG	$\alpha$ -Ketoglutaric acid
LB	Luria broth
MW	Molecular weight
MWCO	Molecular weight cut-off
Mes	Morpholinoethane sulfonic acid
MOPS	3-[N-Morpholino] propanesulfonic acid
NaAc	Sodium acetate buffer
NaCl	Sodium chloride
NADPH	$\alpha$ -Nicotinamide adenine dinucleotide phosphate (reduced form)
NEM	<i>N</i> -Ethylmaleimide
NGase	$\beta$ -aspartyl- <i>N</i> -acetylglucosamine hydrolase
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Ammonium sulphate
NMR	Nuclear magnetic resonance
<i>O</i> -GlcNAcase	Cytoplasmic $\beta$ -GlcNAcase
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PI	Phosphatidyl inositol
PMF	Proton motive force
PMSF	Phenylmethylsulfonyl fluoride
PNase	Peptide- <i>N</i> <sup>4</sup> -( <i>N</i> -acetyl- $\beta$ -D-glucosaminy) asparagine amidase F
POGase	Peptide- <i>O</i> -glycanase
Psi	Pounds per square inch
PVDF	Polyvinylidene difluoride
RP-HPLC	Reverse phase high performance liquid chromatography
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SPR	Surface plasmon resonance

Taps	(N-tris[Hydroxy-methyl]methyl-3-amino propanesulfonic acid
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris (hydroxymethyl)-aminomethane
TFA	Trifluoroacetic acid
Thesit	Polyoxyethylene 9-laurylether
TSK	Chromatography matrix copolymer of oligoethylene glycol, glycidylmethacrylate and pentaeryhtrol-dimethacrylate
UV	Ultra violet
X-Gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside

## Amino Acid Abbreviations

Amino acid	Three Letter Symbol	One Letter Symbol	MW	Side Chain Structure
Alanine	Ala	A	89	-CH <sub>3</sub>
Arginine	Arg	R	174	$-(\text{CH}_2)_3-\text{NH}-\text{C}-\text{NH}_2$ $\parallel$ $\text{NH}$
Asparagine	Asn	N	132	$\text{O}$ $\parallel$ $-\text{CH}_2-\text{C}-\text{NH}_2$
Aspartic Acid	Asp	D	133	-CH <sub>2</sub> -COOH
Asparagine or Aspartic acid	Asx	B		
Cysteine	Cys	C	121	-CH <sub>2</sub> -SH
Glutamic Acid	Glu	E	147	-(CH <sub>2</sub> ) <sub>2</sub> -COOH
Glutamine	Gln	Q	146	$\text{O}$ $\parallel$ $-(\text{CH}_2)_2-\text{C}-\text{NH}_2$
Glutamine or glutamic acid	Glx	Z		
Glycine	Gly	G	75	-H
Histidine	His	H	154	
Homoserine	Hs	Hs	119	-CH <sub>2</sub> -CH <sub>2</sub> -OH
Homoserine Lactone	Hsl	Hsl	101	
Isoleucine	Ile	I	131	$-\text{CH}-\text{CH}_2-\text{CH}_3$ $ $ $\text{CH}_3$

Leucine	Leu	L	131	$  \begin{array}{c}  \text{CH}_3 \\    \\  -\text{CH}_2-\text{CH} \\    \\  \text{CH}_3  \end{array}  $
Lysine	Lys	K	146	$-(\text{CH}_2)_4-\text{NH}_2$
Methionine	Met	M	149	$-(\text{CH}_2)_2-\text{S}-\text{CH}_3$
Phenylalanine	Phe	F	165	$-\text{CH}_2-$ 
Proline	Pro	P	115	
Serine	Ser	S	105	$-\text{CH}_2\text{OH}$
Threonine	Thr	T	119	$  \begin{array}{c}  -\text{CH}-\text{CH}_3 \\    \\  \text{OH}  \end{array}  $
Tryptophan	Trp	W	204	$-\text{CH}_2-$ 
Tyrosine	Tyr	Y	181	$-\text{CH}_2-$ 
Valine	Val	V	117	$  \begin{array}{c}  \text{CH}_3 \\    \\  -\text{CH} \\    \\  \text{CH}_3  \end{array}  $

Sugar Abbreviations

Sugar	Three Letter Symbol
Fucose	Fuc
Galactose	Gal
Mannose	Man
N-acetylgalactosamine	GalNAc
N-acetylglucosamine	GlcNAc
N-acetylneuraminic (sialic) acid	NeuNAc

Note: Sugar linkages are described using conventional carbon ring numbers connected by a slash and anomericity is denoted by  $\alpha$  or  $\beta$ . For example, galactose  $\beta$ 1-4 linked to N-acetylglucosamine is written as Gal $\beta$ 1-4GlcNAc.

## **Chapter 1 Introduction and Literature Review**

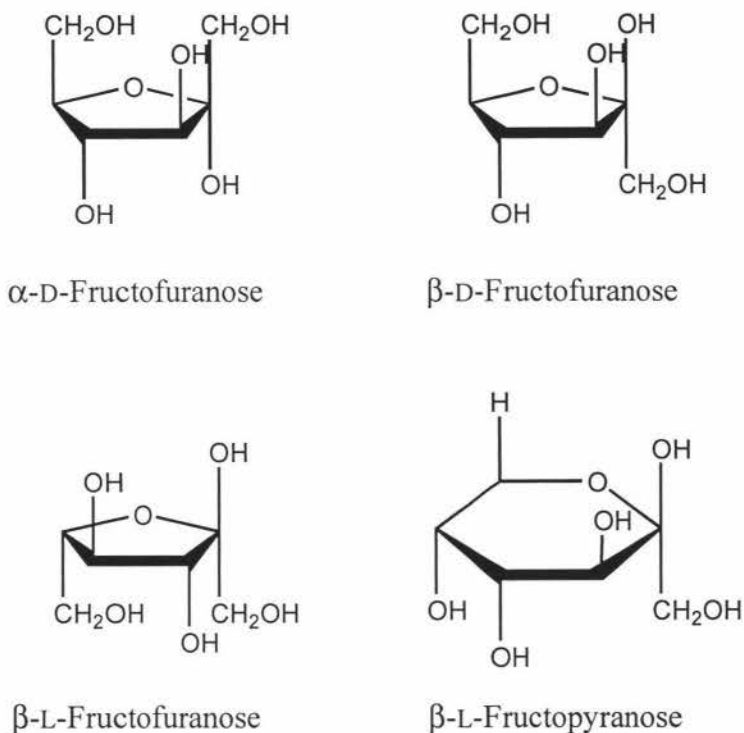
This chapter provides a brief overview of the relationship between protein glycosylation and the enzymes that remove them. The diversity of glycans and the manner in which they are linked to proteins along with their functional roles in plants and animals are reviewed. The properties and possible functions of the glycan-removing enzymes are discussed with the main focus centred on PNGase F.

### **1.1 Protein Glycosylation**

#### **1.1.1 Carbohydrates**

Carbohydrates are the most abundant biomolecules on earth. Simple carbohydrates are polyhydroxy aldehydes or ketones, while such complex carbohydrates that yield these compounds on hydrolysis. Certain carbohydrates such as sugar and starch are the main source of the human diet in most parts of the world, and the oxidation of carbohydrates (glycolysis of D-glucose) is the central energy-yielding pathway in most non-photosynthetic cells. Insoluble carbohydrate polymers feature as structural and protective elements in the cell walls of bacteria (glycosaminoglycans) and plants (cellulose), and in the connective tissues and cell coats of animals (proteoglycans). Other carbohydrate polymers lubricate skeletal joints (hyaluronates) and provide adhesion between cells. Complex carbohydrate polymers are found covalently linked to proteins or lipids where they serve as signals that determine the intracellular transport or the metabolic fate of these glyconjugates. There are three major classes of carbohydrates: Monosaccharides, oligosaccharides, and polysaccharides.

- (1) Monosaccharides consist of a single polyhydroxy aldehyde or ketone unit, the most abundant monosaccharide in nature being D-glucose. Each unit may take up either an anomeric ( $\alpha$  or  $\beta$ ) configuration, exist as (D or L) isoforms and exist in a furanose or pyranose form, resulting in eight possible configurations (figure 1.1).

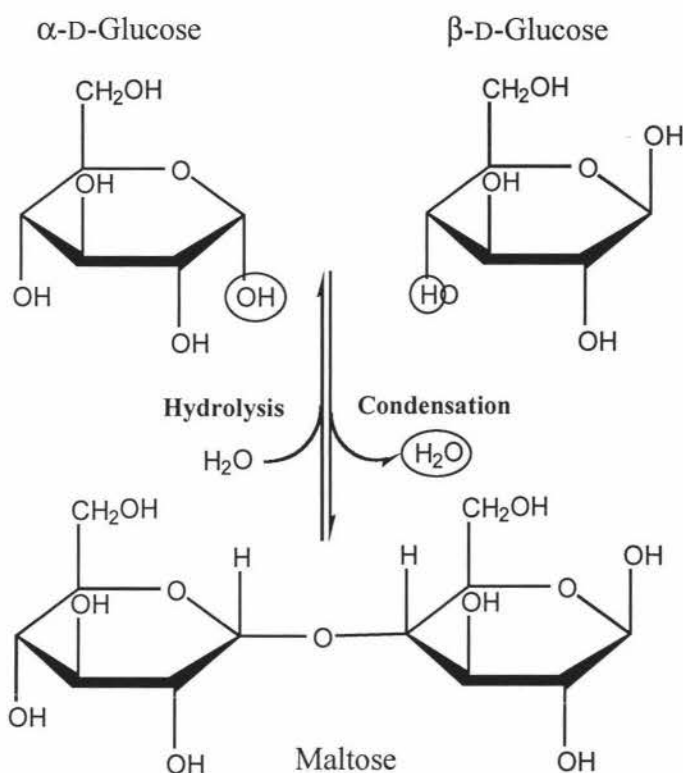


**Figure 1.1 The different configurations of monosaccharides.**

$\alpha$ -D-Fructofuranose, fucose in furanose form, is a ketone ( $\text{C2} \rightarrow \text{C5}$ ) in the anomeric  $\alpha$ -configuration ( $\text{CH}_2\text{OH}$  group on C1 is above the plane of the ring), whereas  $\beta$ -L-fructofuranose is a stereoisomer of  $\beta$ -D-fructofuranose.  $\beta$ -L-fructopyranose, fucose in pyranose form, is an aldehyde ( $\text{C1} \rightarrow \text{C5}$ ).

- (2) Oligosaccharides consist of a number of monosaccharide units linked through a condensation reaction between two hydroxyl groups from two sugar units to form a glycosidic bond (figure 1.2). They are generally short chains consisting of two to ten monosaccharides. When two different sugar units are linked together, there are four distinct isometric links that one hydroxyl group of one sugar unit can form with the hydroxyl group at the C2, C3, C4 or C6 positions of another monosaccharide. Taking the eight possible configurations of each monosaccharide unit into account, there are thirty-two possible configurations may be formed in the linkage of two monosaccharide units. Therefore, the number of possible permutations and combinations of monosaccharide types and glycosidic linkages is enormous.





**Figure 1.2 Condensation reaction between two monosaccharides.**

The hydroxyl group of  $\beta$ -D-Glucose condenses with the hemiacetal of the  $\alpha$ -D-Glucose, resulting in the elimination of  $H_2O$  and formation of the glycosidic bond. The reverse reaction is hydrolysis or attack by  $H_2O$  on the glycosidic bond. Adopted from Lehninger *et al.*, (1993).

- (3) Polysaccharides consist of long chains having hundreds to thousands of monosaccharide units. Some polysaccharides such as cellulose occur in linear chains, whereas others such as glycogen have branched chains.

## 1.1.2 Glycoproteins

### 1.1.2.1 The Nature of Glycoproteins

Glycosylation, where proteins possess a covalently linked oligosaccharide moiety, is the most widespread form of post-translational modification of proteins. These oligosaccharides are known to encode a large amount of structural and biochemical information. Genetic defects or polymorphisms in glycosylation are uncommon in higher animals because most are generally fatal. However, there are some that cause changes in the glycan primary structures with the consequences shown in table 1.1 (Varki, 1993).

**Table 1.1 Some consequences of genetic defects or polymorphisms in glycosylation**

Genetic defect/variation	Defect in glycosylation	Biological consequence(s)
Partial deficiency of xylosylprotein 4- $\beta$ Galactosyltransferase	Decreased production of glycosaminoglycan chains with core Gal $\beta$ 1 $\rightarrow$ 4Xyl linkage.	Progeroid syndrome with delayed mental development, and multiple connective tissue abnormalities.
Hereditary opsonic defect	Point mutation in serum mannose-binding protein.	Heterozygous state causes low opsoninisation of pathogens. Increased infections in childhood.
Haemophilia A variant	Point mutation creates a new N-linked glycosylation site.	Decreased function of Factor VIII, leading to bleeding disorder.
Deficiency of UDP-Gal:3- $\alpha$ Galactosyl transferase	Marked decrease of Gal $\alpha$ 1 $\rightarrow$ 3 Gal $\beta$ 1 $\rightarrow$ 4 GlcNAc sequences terminating glycoprotein and glycolipid oligosaccharides.	No obvious abnormality results. All humans have a natural antibody (up to 1% of circulating IgG) against Gal $\alpha$ 1 $\rightarrow$ 3Gal $\beta$ 1 $\rightarrow$ 4GlcNAc sequences.

### 1.1.2.2 Types of Glycosidic Linkages

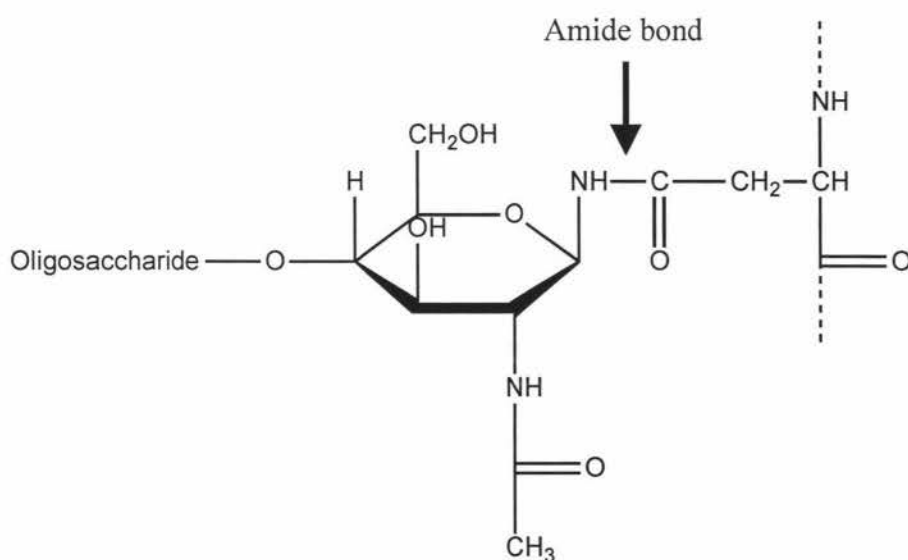
The cellular mechanisms for carrying out glycosylation are available to all proteins that carry the appropriate signals in the polypeptide chain. Firstly, the polypeptide must contain a signal for entry into the secretory pathway, so that, as it leaves the ribosome, the amino-terminus translocates into the lumen of the endoplasmic reticulum (ER), where it first encounters the glycosylation machinery. One or more sites on a given protein may be glycosylated in the ER and the Golgi apparatus with a large population of structurally related oligosaccharides. These oligosaccharides are produced by a sequential addition of monosaccharide units to a core structure and are classified by the way in which they are bound to the protein. The three main classes of covalent glycosidic linkages to proteins are: Glycosylphosphatidylinositol (GPI) anchors, O-linked, and N-linked glycans.

#### 1.1.2.2.1 Glycosylphosphatidylinositol (GPI) Anchors

GPI anchors serve to link some cell surface proteins to the lipid bilayer of membranes. The post-translational attachment of a GPI anchor to a fully folded protein, occurs in the ER and involves the transfer of a pre-assembled precursor onto a specific site on the protein through an ethanolamine phosphate linkage. GPI anchors are attached selectively to those proteins that contain a GPI signal sequence at the carboxy-terminus. This sequence is cleaved and replaced by a pre-assembled GPI anchor precursor, which may then be subsequently modified. In the fully formed anchor, the glycan contains a conserved backbone sequence (Man $\alpha$ 1-2Man $\alpha$ 1-6Man $\alpha$ 1-4GlcNH<sub>2</sub>) which is linked to the 6-position of the *myo*-inositol ring of phosphatidyl inositol (PI). All GPI anchors contain two lipids, normally one acyl and one alkyl, attached to the glycan through the phosphate and glycerol on the inositol ring

#### 1.1.2.2.2 N-linked Glycosylation

N-glycosylation is a co-translational modification available to, but not necessary used by, all proteins that contain the sequon Asn-X-Ser/Thr (where X is any amino acid except Asp and Pro). The oligosaccharides are N-linked to glycoproteins through an amide bond between the asparagine residue and a GlcNAc moiety at the reducing end of the oligosaccharide, commonly known as the  $\beta$ -aspartylglucosylamine link (figure 1.3).

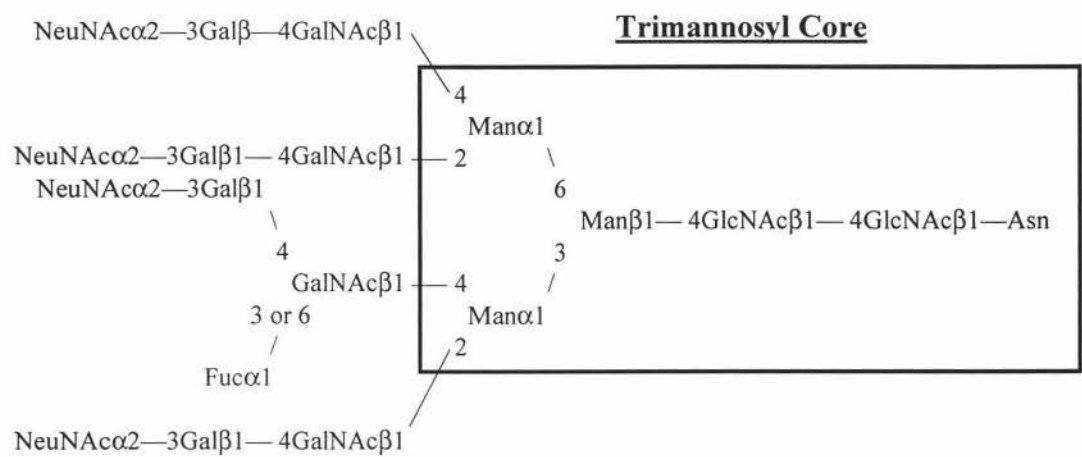


**Figure 1.3** The  $\beta$ -aspartylglucosylamine link of a glycosylated asparagine.

The initiation of *N*-glycosylation starts with the biosynthesis of a lipid linked dolichol-oligosaccharide ( $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{PPDoI}$ ) precursor that is transferred onto the asparagine in the sequon Asn-X-Ser/Thr in the nascent polypeptide chain. The transfer and subsequent trimming and elongation reactions are catalysed by glucosidases, mannosidases and a number of membrane bound oligosaccharyl transferases in the rough endoplasmic reticulum (RER) and the Golgi apparatus to produce a large population of glycans at each glycosylation site. The amino acids that surround the glycosylation site and within the sequon itself have been shown to exert an influence on the efficiency of core glycosylation. Studies with a series of peptides have shown that Asn-X-Thr sequons favour glycosylation compared to Asn-X-Ser sequons (Kaplan *et al.*, 1987) and interestingly, peptides with Asn-X-Cys sequons were reported to be glycosylated at admittedly low levels (Shakin-Eshleman, 1996). Site-directed mutagenesis studies have shown the Y amino acid in the Asn-Leu-Ser/Thr-Y sequon is another important determinant of glycosylation efficiency (Mellquist *et al.*, 1998). The Asn-Leu-Thr-Y sequons are more efficient sites for glycosylation than Asn-Leu-Ser-Y sequons and the efficiency decreases with aromatic residues, negatively charged residues or proline in the Y position. The identity of the X amino acid in the sequon was also shown to alter the efficiency of glycosylation. Negatively charged amino acids, aromatic amino acids, or proline in the X position appear to inhibit glycosylation. However, glycosylation is favoured when small amino acids or positively charged amino acids are in the X position of the sequon (Shakin-Eshleman, 1996).

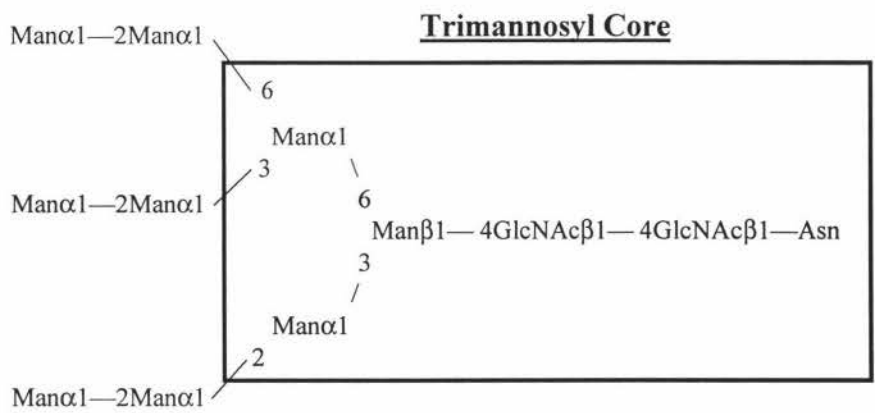
Mature N-linked oligosaccharides all have a common trimannosyl core, because they arise from the same biosynthetic precursor,  $\text{Glc}_3\text{Man}_9(\text{GlcNAc})_2$ . They fall into three general categories: Complex-, high-mannose-, and hybrid glycans.

- (1) Complex oligosaccharides have  $\beta$ -lactosamine substituted onto the core mannose termini and have the greatest structural variation (figure 1.4). The trimannosyl core can be linked up to five oligosaccharide chains and extended by the addition of a wide variety of sugars such as fucose, xylose, galactose and sialic acid. Further structural variation comes from the addition of fucose residues at the C3 or C6 positions of the  $\alpha\text{GlcNAc}$  proximal to the asparagine residue.



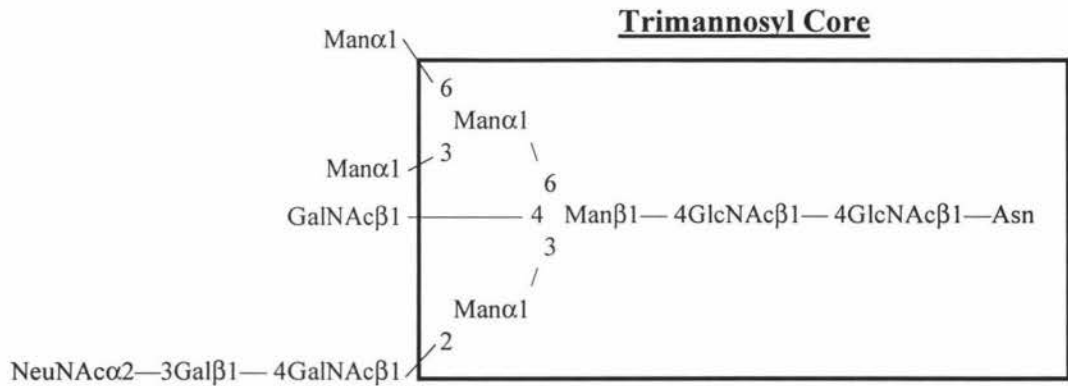
**Figure 1.4 Complex type N-linked glycan.**  
Modified from Mort and Pierce, (1995).

- (2) High-mannose type glycans are composed entirely of mannose, apart from the two GlcNAc residues of the trimannosyl core (figure 1.5).



**Figure 1.5 High mannose type N-linked glycan.**  
Adopted from Mort and Pierce, (1995).

- (3) Hybrid structures have features of complex and high mannose type oligosaccharides, and most contain a bisecting GlcNAc  $\beta$ 1-4 linked to  $\beta$ -mannose residue (figure 1.6).



**Figure 1.6 Hybrid type N-linked glycan.**

Adopted from Mort and Pierce, (1995).

#### 1.1.2.2.3 O-linked Glycosylation

In contrast to *N*-glycosylation and the addition of GPI anchors, *O*-glycosylation does not begin with the addition of a precursor to the polypeptide chain, but of a single monosaccharide, usually *N*-acetylgalactosamine (GalNAc) but sometimes *N*-acetylglucosamine (GlcNAc) or xylose. This is transferred to the side chain oxygen of a serine, threonine, hydroxyproline, or hydroxylysine in both fully folded proteins and nascent polypeptide chains. The link between an  $\alpha$ -*N*-acetylgalactosamine and serine or threonine ( $\alpha$ GalNAc-Ser/Thr) residue is known as a mucin type linkage in mammals. *O*-linked glycans can be either linear, formed by an extension of the oligosaccharide chain through a  $\beta$ 1,3-linked galactose to the  $\alpha$ GalNAc (Gal $\beta$ 1-3GalNAc-Ser/Thr), or branched where additional extensions are made from the GalNAc $\alpha$ -Ser/Thr core through a  $\beta$ 1,6-link between a GlcNAc and the  $\alpha$ GalNAc. Single monosaccharide units such as GalNAc or GlcNAc are commonly found *O*-linked to a growing number of proteins, although more commonly to nucleoplasmic and cytoplasmic proteins. There appear to be no defined sequon requirement for attachment (Haltiwanger *et al.*, 1991).

### 1.1.2.3 Complexity and Diversity of Glycans

The synthesis of the polypeptide chain of a glycoprotein is under genetic control. In contrast, oligosaccharides are attached to the protein then processed by a series of enzymes. While the populations of these enzymes are the result of a genetic template, their activities are governed by the availability of substrates, conditions in the cell and competition from other enzymes. Consequently, a single polypeptide that is glycosylated normally emerges from the biosynthetic pathway as a population of proteins that have the same polypeptide chain, but different oligosaccharide chains or glycans, known as glycoforms. Because the same glycosylation machinery is available to all proteins that enter the secretory pathway in a given cell, most glycoproteins emerge with characteristic glycosylation patterns but heterogeneous populations of glycans at each glycosylation site. This suggests that the proteins themselves direct the processing of their own glycan chains within the constraints imposed by a given array of enzymes and sugar nucleotides. Within the ER all glycoproteins contain the same limited range of oligomannose sugars, and it is later, within the Golgi, that the extensive heterogeneity develops. The factors that control the composition of the glycoform populations and the role that heterogeneity plays in the function of glycoproteins are important questions that are not yet fully understood. Different combinations of stereochemical configuration and monosaccharide composition lead to slight differences in physiochemical properties and hence the potential for diversity in function. This contrasts with the linear linking of amino acids or nucleic acids. Summarised below are the types of variation possible:

- Glycosylation is cell, tissue and species specific
- Glycoproteins from the same cell may contain different oligosaccharide chains
- Individual polypeptides may have identical glycosylation patterns at a particular glycosylation site
- Identical polypeptides may have different oligosaccharide structures at a particular glycosylation site
- Polypeptides may contain multiple glycosylation sites, each of which may have different glycosylation structure
- The pattern of oligosaccharide heterogeneity at a single glycosylation site under constant physiological conditions is reproducible and not random



#### 1.1.2.4 Potential Roles of Protein Glycosylation

Post-translational glycosylation is an important process for modifying the structure and function of the majority of eukaryotic secreted and membrane proteins. Covalently linked oligosaccharides are known to alter the physiochemical properties of their parent proteins such as viscosity, isoelectric point (pI), solubility and thermal stability. They also have an influence on the global physical properties of protein structure (Lehninger *et al.*, 1993): The conjugated clusters of hydrophilic carbohydrates may alter the polarity and solubility of glycoproteins. Oligosaccharide chains attached to newly synthesised proteins in the Golgi complex may also influence the sequence of polypeptide-folding events that lead to the native tertiary structure of the protein in the following ways: (i) Steric interactions between peptide and oligosaccharide may repress one folding route and favour another. (ii) When a population of negatively charged oligosaccharides chains are clustered in one region of a protein, the charge repulsion among them favours the formation of an extended, rod-like structure in that region. (iii) The bulkiness and negative charge of oligosaccharide chains may protect some proteins from attack by proteolytic enzymes.

Beyond these effects on tertiary structure, there are more specific biological functions ascribed to the oligosaccharide chains on glycoproteins depending on the protein to which they are attached. These include cell-cell recognition, intracellular sorting and targeting, and affinity to receptors. Some of these examples are listed in Table 1.2 (Varki, 1993; Lis and Sharon, 1993). However, while these observations suggest that oligosaccharides play important biological roles in cell growth and differentiation, and the fact that changes in oligosaccharide structure are observed in various disease states, no single common theory has emerged to explain the diversity of structures and the variety of roles. By studying the function of glycoproteins in their glycosylated and deglycosylated states, some of the biological roles played by oligosaccharides were elucidated. An example is the inhibition of the initial *N*-glycosylation of ecto-apyrase (HB6), a human brain E-type ATPase, expressed in COS cells, by tunicamycin. The non-glycosylated proteins are devoid of ATP and ADP hydrolysing activity and exist as a monomer unlike the glycosylated HB6 that exists as a homodimer (Smith *et al.*, 1999).



Table 2. Some biological roles of oligosaccharides.

Biological role of glycans	Example(s)
Structural, protective, and stabilising	<ul style="list-style-type: none"> <li>Maintenance of tissue structure, integrity and porosity.</li> <li>Protects polypeptides from recognition by proteases or antibodies.</li> <li>Initiation of correct folding in the rough endoplasmic reticulum.</li> </ul>
Organisational and barrier	<ul style="list-style-type: none"> <li>Oligosaccharide binding domains of glycoconjugates involved in the organisation of extra-cellular matrix.</li> <li>Chondroitin/dermatan sulphate chains of proteoglycan decorin is required for the deposition of fibronectin in Chinese Hamster Ovary cell matrix.</li> </ul>
Traitorous	<ul style="list-style-type: none"> <li>As specific receptors for variety of viruses, bacteria, parasites, plant and bacterial toxins.</li> <li>Antigens for autoimmune and alloimmune reactions.</li> <li>Recognition of terminal sialic acids on glycans is the first step in infectious process of the influenza virus.</li> </ul>
Masking and decoy	<ul style="list-style-type: none"> <li>Addition of specific monosaccharides masks the sequences recognised by microorganisms, toxins, or autoimmune antibodies.</li> <li>Addition of galactose and sialic acid to the Tn antigen abolishes its autoimmune reactivity.</li> </ul>
Symbiotic	<ul style="list-style-type: none"> <li>As specific receptors for microorganisms in symbiotic relationships.</li> <li>Certain gut bacteria in animals and some root-nodule forming bacteria in plants mediate their binding to host cell surfaces through specific sugar sequences.</li> </ul>
On-off and tuning	<ul style="list-style-type: none"> <li>Glycosylation can substantially modulate the interaction of peptides with their cognate ligands or receptors.</li> <li>Binding affinities and biological activities of hematopoietic growth factors such as erythropoietin changes substantially with differing degrees of N-linked glycosylation.</li> </ul>
Targeting and clearance	<ul style="list-style-type: none"> <li>Glycosylation can affect protein turnover and half-life in single cell.</li> <li>Exposed terminal <math>\beta</math>-Gal residues on mammalian plasma proteins are recognised by the asialoglycoprotein receptor and cleared from circulation.</li> </ul>
Hormonal action	<ul style="list-style-type: none"> <li>Free oligosaccharides can have biological effects in various systems.</li> <li>Various plant receptors recognise a specific <math>\beta</math>-glycan oligosaccharide of the <i>Phytophthora</i> fungal cell walls, causing the release of phytoalexins.</li> </ul>
Cell-cell and cell-matrix recognition	<ul style="list-style-type: none"> <li>Oligosaccharides closely spaced together on a polypeptide generate clustered sugars for specific recognition.</li> <li>Involvement of selectin family of receptors in response to tissue injury or infection. The ligands involved in recognition appear to be sialylated fucosylated Sialyl Lewis<sup>x</sup> and Sialyl Lewis<sup>a</sup>.</li> </ul>

## 1.2 Deglycosylation

The previous sections have reviewed the significance of glycosylation in altering the physiochemical properties of glycoproteins, the biochemical information encoded by the attached glycans, and their possible biological roles. Much of this information has been obtained by removing these oligosaccharides to determine both the structure and function of the glycan moieties.

### 1.2.1 Tools to Study the Roles of Glycans in Glycoprotein Functions

Approaches include enzymatic or chemical removal of glycan chains, inhibition of initial glycosylation with tunicamycin, changing the glycosylation pathway, prevention of glycan processing with inhibitors such as castanospermine, and elimination of specific glycosylation sites by site-directed mutagenesis (Varki, 1993).

#### 1.2.1.1 Chemical Deglycosylation

The most widely employed chemical procedure for releasing N-linked oligosaccharide chains is hydrazinolysis. This method is non-selective and therefore gives uniform release of unreduced N-linked oligosaccharides from glycoproteins in high yield. Hydrazinolysis requires relatively harsh reaction conditions that, in addition to cleavage of the asparatyl-*N*-acetylglucosamide bond, results in partial release of any *O*-linked sugars, de-*N*-acetylation, hydrazone formation, and in some cases, peeling and other side reactions (Takasaki *et al.*, 1982). Furthermore, not only the protein component of the glycoprotein is also destroyed in the reaction, but it is also necessary to re-*N*-acetylate each oligosaccharide and to regenerate its free-reducing terminal before proceeding with further characterisation (Hirani *et al.*, 1987).

#### 1.2.1.2 Enzymatic Deglycosylation

'Proximal glycanases' (PROXIases) have been defined as a class of enzymes involved in the deglycosylation of glyconjugates (Suzuki *et al.*, 1994c). These enzymes catalyse the cleavage of the linkage between proximal monosaccharide and core protein (ceramides) or between two proximal monosaccharide moiety to release free glycan and apo-glycoconjugates. PROXIases rapidly become the biochemical tools of choice by

researchers attempting to analyse the structure and function of the carbohydrate moiety and to aid in crystallisation of glycoproteins. In contrast to chemical deglycosylation, a simple incubation with PROXIases under mild conditions releases intact oligosaccharides and their protein counterpart in a form suitable for purification and characterisation. Furthermore, oligosaccharide structure can be deduced on basis of PROXIase specificity limitations (Hirani *et al.*, 1987; Maley *et al.*, 1989). Side reactions are unlikely unless contaminating enzymes such as exoglycosidases (enzymes that remove sugars from the reducing end) or phosphatases are present.

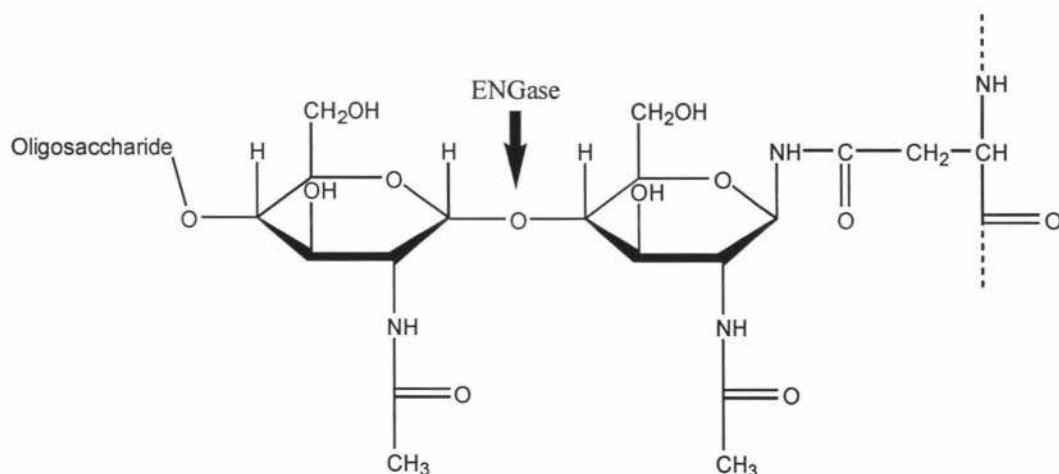
PROXIases are organised into five classes: Cytoplasmic  $\beta$ -GlcNAcase (*O*-GlcNAcase), Endoglycoceramidase (EGCase), Endo-*N*-glycanase (ENGase), Peptide-*N*-glycanase (PNGase), and Peptide-*O*-glycanase (POGase). The occurrences of these enzymes, their functions and possible biological roles are briefly reviewed in the next section.

### 1.2.2 Endo-*N*-acetyl- $\beta$ -D-glucosaminidases (ENGases)

ENGases (EC 3.2.1.96) have been found in various sources (table 1.3). The first ENGases detected was in a crude extract of fig (Ogata-Arakawa *et al.*, 1977) and since then more than 30 have subsequently been found and characterised. These enzymes hydrolyse the glycosidic linkage between the two *N*-acetylglucosamine (GlcNAc) residues within the asparagine-linked oligosaccharide core, leaving a GlcNAc residue on the glycoasparagine with the release of a free oligosaccharide with one less reducing end GlcNAc residue as shown in figure 1.7 (Tai *et al.*, 1977). The key structural determinants for ENGase activity are a polypeptide on either side of the asparagine residue and a dichitobiose core (two GlcNAc residues) with at least three mannose residues. There are also specific geometric configuration and monosaccharide composition requirement of the oligosaccharide chain (Maley *et al.*, 1989). For example, Endo F<sub>1</sub> and H only recognises high mannose structures and could not hydrolyse complex asparagine-linked oligosaccharides such as glycopeptides with a fucose linked  $\alpha$ 1-6 to the proximal *N*-acetylglucosamine. Other ENGases have their own different and restricted substrate specificities.

**Table 1.3 The occurrence of some endo-*N*-acetyl- $\beta$ -D-glucosaminidases (ENGases).**

Kingdom	Organism	Name	Reference(s)
Animals	Hen Oviduct	-	Tarentino <i>et al.</i> , 1976
	Human Kidney	-	DeGasperi <i>et al.</i> , 1989
	Human saliva	HS	Ito <i>et al.</i> , 1993
	Rat liver	-	Fujisaki <i>et al.</i> , 1991
Bacteria	<i>Arthrobacter protophormiae</i>	A	Takegawa <i>et al.</i> , 1989
	<i>Clostridium perfringens</i>	CI, CII	Ito <i>et al.</i> , 1975
	<i>Streptococcus pneumoniae</i>	D	Muramatsu <i>et al.</i> , 1971
	<i>Flavobacterium meningosepticum</i>	F <sub>1</sub> , F <sub>2</sub> , F <sub>3</sub>	Plummer <i>et al.</i> , 1991
	<i>Streptomyces plicatus</i>	H	Tarentino <i>et al.</i> , 1972
	<i>Streptomyces plicatus</i>	L	Tarentino <i>et al.</i> , 1974
	<i>Pseudomonas</i> sp	PI, PII	Takegawa <i>et al.</i> , 1991
	<i>Stigmatella aurantiaca</i>	St	Bourgerie <i>et al.</i> , 1993
Fungi	<i>Sporotrichum dimorphosporum</i>	B	Bouquelet <i>et al.</i> , 1980
	<i>Mucor heimalis</i>	M	Kadowaki <i>et al.</i> , 1990
Mould	<i>Dictyostelium discoideum</i>	S	Freeze <i>et al.</i> , 1984
Plants	Fig	FI, FII	Chien <i>et al.</i> , 1977
	<i>Canavalia ensiformis</i> (Jackbean)	J	Yet <i>et al.</i> , 1988
	<i>Phyllostachys heterocycla</i> (Bamboo shoots)	P	Nishiyama <i>et al.</i> , 1991
	<i>Raphanus sativus</i> (Radish)	R	Berger <i>et al.</i> , 1995b
	<i>Silene alba</i> (White campion)	Se	Lhernould <i>et al.</i> , 1995

**Figure 1.7 The glycosidic bond cleaved by ENGases.**

### 1.2.3 Peptide $N^4$ -(*N*-acetyl- $\beta$ -D-glucosaminyI) asparagine amidase (PNGases)

PNGases (EC 3.5.1.52) differ from ENGases because they specifically hydrolyse the  $\beta$ -asparatyl-glucosaminyIamine bond between the asparagine residue in peptide linkage and the GlcNAc moiety at the reducing end of the oligosaccharide chain. This cleavage results in the conversion of the asparagine residue to an aspartic acid and the concomitant liberation of free and intact oligosaccharides. These enzymes are therefore more correctly described as amidases (amidohydrolyases) rather than endoglycosidases which cleave the glycosidic bond between two GlcNAc sugars. The first PNGase discovered was from almond emulsin and designated as PNGase A (Takahashi *et al.*, 1977). Other PNGases have subsequently been discovered from many sources (table 1.4) but only a few have been purified to homogeneity and characterised (table 1.5).

**Table 1.4 The occurrence of peptide  $N^4$ -(*N*-acetyl- $\beta$ -D-glucosaminyI) asparagine amidase (PNGases).**

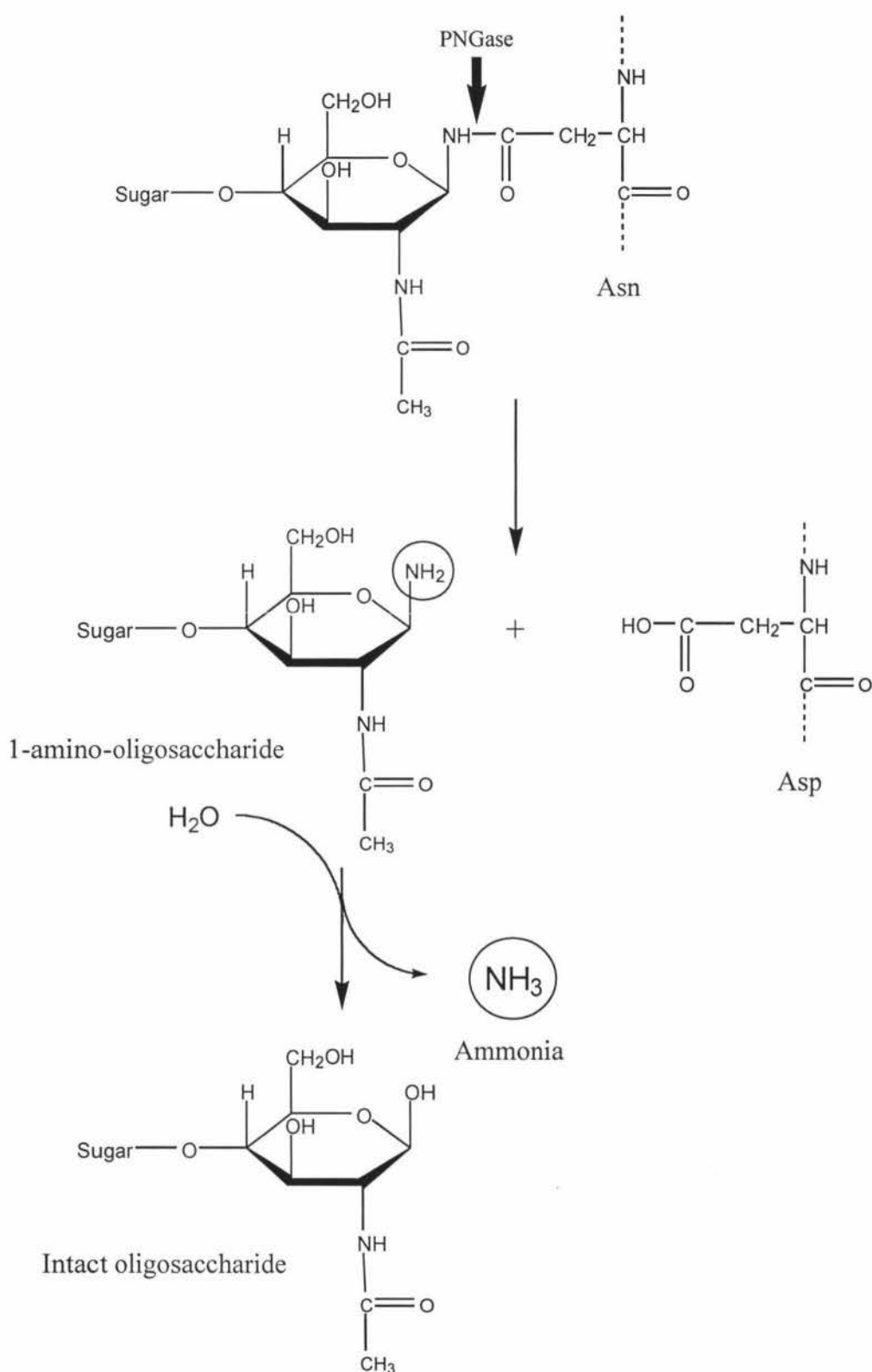
Kingdom	Organism	Name	Reference(s)
Animals	Hen oviduct	HO	Suzuki <i>et al.</i> , 1997
	Mouse (liver ER)	-	Weng <i>et al.</i> , 1997
	Mouse (L-929 fibroblast cells)	L-929	Chang <i>et al.</i> , 1997
	Humans, chickens	-	Suzuki <i>et al.</i> , 1995a
	Various mouse organs	-	Kitajima <i>et al.</i> , 1995
	<i>Oryzias latipes</i> (Medaka fish embryo)	M	Seko <i>et al.</i> , 1991, 1999
Bacteria	<i>Flavobacterium meningosepticum</i>	F	Plummer <i>et al.</i> , 1984
	<i>Saccharomyces cerevisiae</i> (Yeast ER)	-	Suzuki <i>et al.</i> , 1998
Fungi	<i>Aspergillus tubigenesis</i>	At	Paquin <i>et al.</i> , 1997
Plant	<i>Prunus amygdalusi</i> (Almond emulsin)	A	Takahashi, 1977
	<i>Glycine max</i> (Soybean seeds)	GM	Kimura <i>et al.</i> , 1998
	<i>Canavalia ensiformis</i> (Jack beans)	J	Sugiyama <i>et al.</i> , 1983
	<i>Hordeum vulgare</i> (Barley seeds)	JIP60	Dunaeva <i>et al.</i> , 1999
	<i>Oryza sativa</i> (Rice seeds)	Os	Chang <i>et al.</i> , 2000
	<i>Pisum sativum</i> (Split pea)	P	Plummer <i>et al.</i> , 1987
	<i>Raphanus sativus</i> (Radish)	R	Beger <i>et al.</i> , 1994
	<i>Silene alba</i> (White campion)	Se	Lhernould <i>et al.</i> , 1992
	Various plant seeds	-	Plummer <i>et al.</i> , 1987

Table 1.5 Characteristics of known purified PNGases.

	A	At	F	GM	J	L-929	M (Acidic)	Os
pH optimum	4.5	5.0	8.5	5.0	5.0	7.0	4.0	5.0
Subunit	Hetero-dimer	Mono-mer	Mono-mer	Mono-meric	Multi-meric	Homo-dimeric		Mono-meric
Molecular weight (kDa)	75.5	78	34.8	93	69	212	150	80
-SH group requirement for activity	No	No	No	ND	No	Yes	No	No
Action on plant complex type <i>N</i> -glycan-peptide <i>e.g. Bromelain or ricin D glycopeptides that contains fucose <math>\alpha</math>1-3 linked to the proximal GlcNAc</i>	Yes	Yes	No	Yes	ND	No	No	ND
Action on sialylated glycopeptide <i>e.g. Fetuin glycopeptides</i>	Yes (strong)	ND	Yes (weak)	Yes (weak)	No	Yes	Yes	ND
Action on peptides containing a single GlcNAc residues	Yes	Yes (weak)	No (almost)	No (almost)	ND	No	ND	ND

Results from  $^1\text{H}$  NMR and kinetic studies have shown that the cleavage of the  $\beta$ -asparatyl-glucosaminylamine bond occurs in two steps (Risley and Van Etten, 1985). The two-step reaction is presented in figure 1.8. In the first step, hydrolysis of the amide bond generates an aspartic acid on the polypeptide and liberates the carbohydrate moiety as a 1-amino-oligosaccharide intermediate, the amino group from the asparagine being retained. Non-enzymatic decomposition of the 1-amino-oligosaccharide to release ammonia is very slow at pH 8.6 and above, and very fast below pH 8 (Risley and Van Etten, 1985; Tarentino *et al.*, 1982). This reaction is mechanistically analogous to that of a lysosomal enzyme,  $\beta$ -aspartyl-*N*-acetylglucosamine hydrolase (NGases), but differs in that the lysosomal enzyme cannot hydrolyse the amide bond when the amino or carboxyl end of the asparagine is substituted (Tarentino *et al.*, 1969). Conversely, PNGases cannot cleave oligosaccharides on a single asparagine residue (Plummer and Tarentino, 1981; Suzuki *et al.*, 1994a and 1994c).

The structural determinants for recognition of the substrate appear to include both the polypeptide chain and the glycan core. Studies have shown that the location of the oligosaccharide on the peptide backbone and its chain length are major determinants for enzymatic activity (Plummer and Tarentino, 1981). Glycosylated asparagine residues are hydrolysed less favourably if present at the carboxyl- or amino-terminal position of a peptide chain. In the same paper, the authors showed that the activity was inhibited when substrates contained long polypeptide chains (<20 residues), and in general, glycopeptides with a dipeptide on either side of the N-linked asparagine are good substrates for PNGases. Substrates with a negative charge in position X in the *N*-glycosylation sequon -Asn-X-Ser/Thr- have been shown to effect a negative influence on the rate of hydrolysis (Tarentino *et al.*, 1993a).



**Figure 1.8 The two step cleavage reaction catalysed by PNGases.**

The hydrolysis of the  $\beta$ -asparatylglucosaminylamine bond converts Asn to Asp and releases oligosaccharide intermediates that slowly degrade to intact oligosaccharides and free ammonia.



### 1.2.3.1 The *in vivo* Functions of PNGases

Although PNGases are widely used in research to remove sugar chains for structural and functional studies of the oligosaccharide moieties and to aid in protein crystallisation, the functional roles of PNGases *in vivo* are unknown. Studies of the occurrence of PNGases in various living organisms lead to suggestion that these enzymes play many different roles depending on their cellular location and the developmental stage of the cell. For example, two PNGases with distinct enzymatic properties and two different pH optimum values are expressed during embryogenesis of medaka fish (Seko *et al.*, 1999). Neutral PNGase M from medaka fish was suggested to control the biological functions of L-hyosophorin that is involved in the regulation of cell-to-cell interactions during early development. Whereas the acidic PNGase M is responsible for the production of glycoposphoprotein-type free glycans that may contribute to the degradation and the absorption of glycoposphoprotein by the developing embryos. Some PNGases are believed to be responsible for quality control of *de novo* synthesised proteins inside the cell such as the neutral PNGases from the ER of rats, hen oviduct and in the cytosolic fraction of the yeast ER (Weng and Spiro, 1997; Suzuki *et al.*, 1997, 1998). PNGase L-929 is proposed to have a dual role as both a glycosidase and a lectin-like receptor protein *in vivo* (Suzuki *et al.*, 1994b). This enzyme possesses a unique carbohydrate-binding site that is separate from the catalytic site, and regulates the catalytic site by feedback inhibition at high concentrations of free oligosaccharides. PNGase J from jack bean converts the glycosylated concanavalin A precursor into an active form of the lectin (Bowles *et al.*, 1986) and releases the unconjugated glycans (UNGs) which have a specific role in the plant cell (Priem *et al.*, 1994).

Detailed knowledge of the characteristics of many PNGases is limited because of the extremely small quantities of PNGases present in living organisms. This makes localisation, isolation, purification, and characterisation of these enzymes very difficult. Furthermore, some of the PNGases are expressed only during certain stage of growth and development. Creation of knock-out strains that are deficient in PNGase may not have an obvious effect on the organisms involved as some enzymatic pathways have redundancy systems for important processes needed for survival. In addition, creation of such strains requires detailed knowledge of the sequence encoding the PNGases and the genomes of the organisms involved. Antibodies raised against different PNGases

can facilitate studies on the cellular localisation and development profiles of PNGases. However, the generation of these antibodies again requires the protein to be purified in relatively large quantities and antibodies raised against one PNGase may not detect other PNGases, if the epitopes on those PNGases are not recognised. Nevertheless, the functional roles played by PNGases will become clear as more of these enzymes are characterised.

#### 1.2.4 Other PROXlases

Other enzymes involved in deglycosylation of glycoconjugates are listed in table 1.6. They are, however, beyond the scope of this review and would not be discussed further.

**Table 1.6 The occurrence of other PROXlases.**

Enzyme	Kingdom	Organism	Reference(s)
EGCase	Animal	Earthworm	Li <i>et al.</i> , 1987
		Leech	Li <i>et al.</i> , 1986
		Rabbit	Basu <i>et al.</i> , 1990
	Bacteria	<i>Corynebacterium</i> sp	Ashida <i>et al.</i> , 1992
		<i>Rhodococcus</i> sp.	Ito and Yamagata, 1986
Endo- $\beta$ -Xylase	Animal	Mollusc	Takagaki <i>et al.</i> , 1990
		Rabbit	Takagaki <i>et al.</i> , 1988
O-GlcNAcase	Animal	Rat	Dong and Hart, 1994
POGase	Bacteria	<i>Alcaligenes</i> sp.	Fan <i>et al.</i> , 1988
		<i>Diplococcus pneumoniae</i>	Bhavanandan <i>et al.</i> , 1976
		<i>Streptomyces</i> sp.	Ishii-Karakasa <i>et al.</i> , 1992

#### 1.2.5 Glycosidases from *Flavobacterium meningosepticum*

The aerobic, rod-shaped, Gram-negative bacteria secretes an amidase, PNGase F, and three endoglycosidases, Endo F<sub>1</sub>, Endo F<sub>2</sub>, and Endo F<sub>3</sub>. (Elder *et al.*, 1982; Plummer *et al.*, 1984; Trimble *et al.*, 1991). The properties and specificities of the endoglycosidases from *F. meningosepticum* are briefly examined with the main focus being on PNGase F.

#### 1.2.5.1 Endo F<sub>1</sub>

Endo F<sub>1</sub> shows maximum activity between pH 5 and 6. It retains 65% activity at pH 7.0 but below pH 5, the activity drops off rapidly (Maley *et al.*, 1989). The Endo F<sub>1</sub> gene has been cloned and sequenced, and codes for a mature protein of 289 amino acids with a molecular mass of 31, 667 Da (Tarentino *et al.*, 1992). Endo F<sub>1</sub> has a very similar substrate specificity to Endo H from *Streptomyces plicatus*. Both enzymes are nearly identical in their ability to hydrolyse high-mannose oligosaccharides and differ only in their specificity for a core-substituted fucose that impedes hydrolysis by Endo F<sub>1</sub> (Trimble *et al.*, 1991). Both Endo F<sub>1</sub> and Endo H do not hydrolyse any complex type N-glycans.

#### 1.2.5.2 Endo F<sub>2</sub> and F<sub>3</sub>

Endo F<sub>2</sub> and Endo F<sub>3</sub> are more active between pH 4.0 and 4.5, retaining 70% of their activity at pH 3.0 but above pH 6, the activity of both glycosidases sharply decreases. Endo F<sub>2</sub> preferentially hydrolyses biantennary complex glycans and although it can hydrolyse high mannose oligosaccharides, it does so at a greatly diminished rate. Endo F<sub>2</sub> does not cleave fucose-containing hybrid structures, or tri- or tetra-antennary oligosaccharides (Trimble *et al.*, 1991). Endo F<sub>3</sub> also hydrolyses bi- and tri-antennary glycans but at rates slower than Endo F<sub>2</sub>. However, if the core asparagine-proximal N-acetylglucosamine is substituted with an  $\alpha$ 1-6 fucose residue, then Endo F<sub>3</sub> will hydrolyse the glycan at a much greater rate (Tarentino *et al.*, 1994a).

#### 1.2.5.3 PNGase F

PNGase F is the best-characterised PNGase with a known tertiary structure (Norris *et al.*, 1994b; Tarentino *et al.*, 1994b), and the gene has been cloned, sequenced, and expressed in *E.coli* by several groups (Lemp *et al.*, 1990; Tarentino *et al.*, 1990; Barsomian *et al.*, 1990; Grueninger-Leitch *et al.*, 1996).

##### 1.2.5.3.1 Properties of PNGase F

While it is most active at pH 8.5, the enzyme is at least 80% active between pH 7.5 and 9.5. The mature enzyme comprises 314 amino acids, has a molecular weight of 34, 779 Da and contains a relatively high number of nine tryptophan residues (Tarentino *et al.*, 1990). PNGase F is compatible with a wide variety of inorganic and organic buffers

( $\leq 0.1$  M), including sodium phosphate, lithium carbonate, ammonium bicarbonate, Tris-HCl, glycylglycine, HEPES, triethylamine acetate, but sodium borate is inhibitory (Tarentino *et al.*, 1994a). It is stable in protein denaturants such as 2.5 M urea at 37°C for 2 hours and still possesses 40% of its activity in 5 M urea. The activity is not inhibited by high concentrations of the chelating agents such as EDTA and 1,10-orthophenanthroline or by the serine protease inhibitor PMSF (Maley *et al.*, 1989). The enzyme is stable at least 6 months at 4°C and indefinitely at -80°C with 50% (v/v) glycerol, and should not be exposed to repeated freezing and thawing. The presence of glycerol (5%) is inhibitory and should not exceed 0.1% for maximal activity (Dr. G. E. Norris, personal communication).

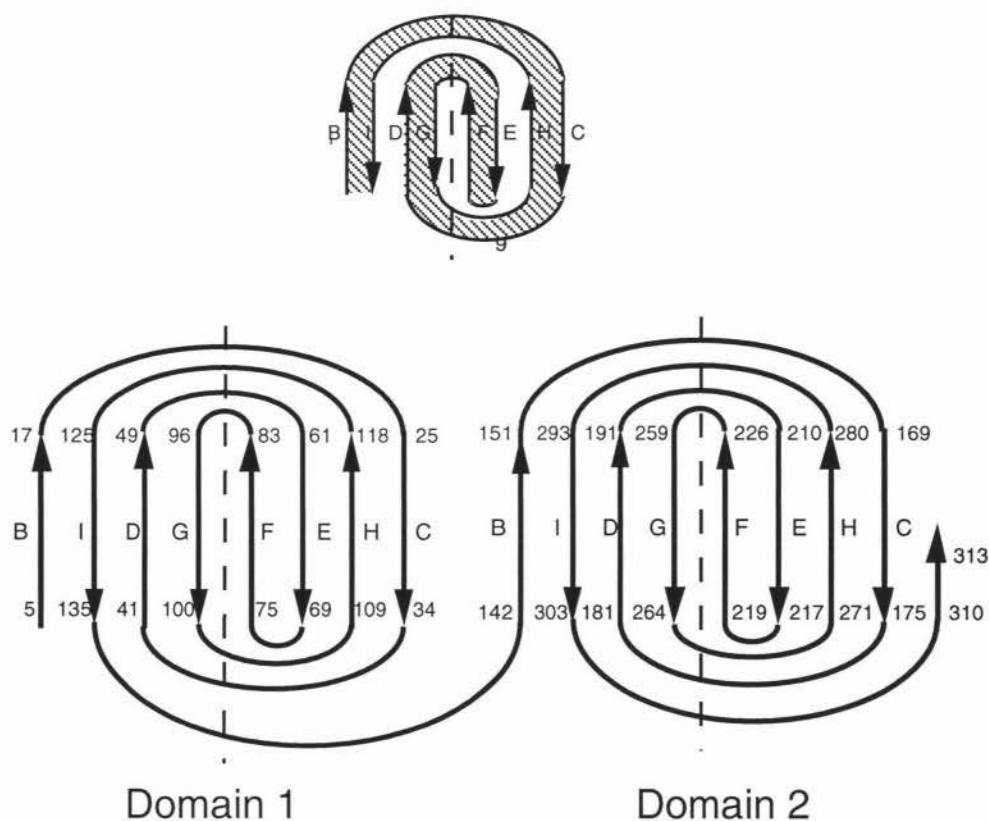
#### 1.2.5.3.2 Substrate Structure Requirements of PNGase F

PNGase F activity requires both the amino and carboxyl groups of the asparagine residue to be in a peptide linkage, while the minimum the oligosaccharide must consist at least two GlcNAc residues. The enzyme is also highly sensitive to modifications of the sugar core, as an  $\alpha 1-3$ -fucose substituent on the asparagine-proximal GlcNAc found in glycoproteins from plants and insects completely blocks PNGase F activity, but an  $\alpha 1-6$ -fucose substituent has no effect (Tretter *et al.*, 1991). However, PNGase F shows no selectivity for the outer carbohydrate structure, giving it a broad specificity for N-linked glycoproteins. Studies with a stepwise degradation of a biantennary glycopeptide with exoglycosidases have shown that the size of the carbohydrate moiety on the substrate has little influence on enzyme activity and that the hydrolysis rate may be primarily determined by the length of the peptide (Altman *et al.*, 1995). Recently, Fan *et al.*, (1997) carried out detailed studies on substrate structure requirements of PNGase A and F, using more than 30 glycopeptides of varying peptide lengths and oligosaccharide moieties, and made several interesting observations:

- (i) Neither enzyme cleaves cellobiose nor lactose substituted glycopeptides, indicating that the 2-acetamido group on the Asn-linked GlcNAc is important in substrate recognition.
- (ii) PNGase A efficiently cleaves a Gln-bound CTB glycopeptide (Gln replaces the Asn), but the action of PNGase F on this peptide is minimal.
- (iii) PNGase A can act on CTB dipeptides whereas PNGase F prefers a tripeptide or longer.

### 1.2.5.3.3 The Three Dimensional Structure of PNGase F

Structures from two strains (ATCC 33958 and CDC strain 3352) of *F. meningosepticum* were published (Kuhn *et al.*, 1994; Norris *et al.*, 1994a). Although both structures were crystallised under different conditions (at pH 4.3 and pH 8.5), they are essentially identical. The structure consists of two tightly associated all- $\beta$  domains, an amino-terminal domain comprising residues 1-135 and a carboxyl-terminal domain comprising residues 142-314. The  $\beta$ -barrels in each domain are arranged in a 4+4 jelly roll arrangement (figure 1.9) that closely resembles viral capsid proteins (Rossmann *et al.*, 1983).



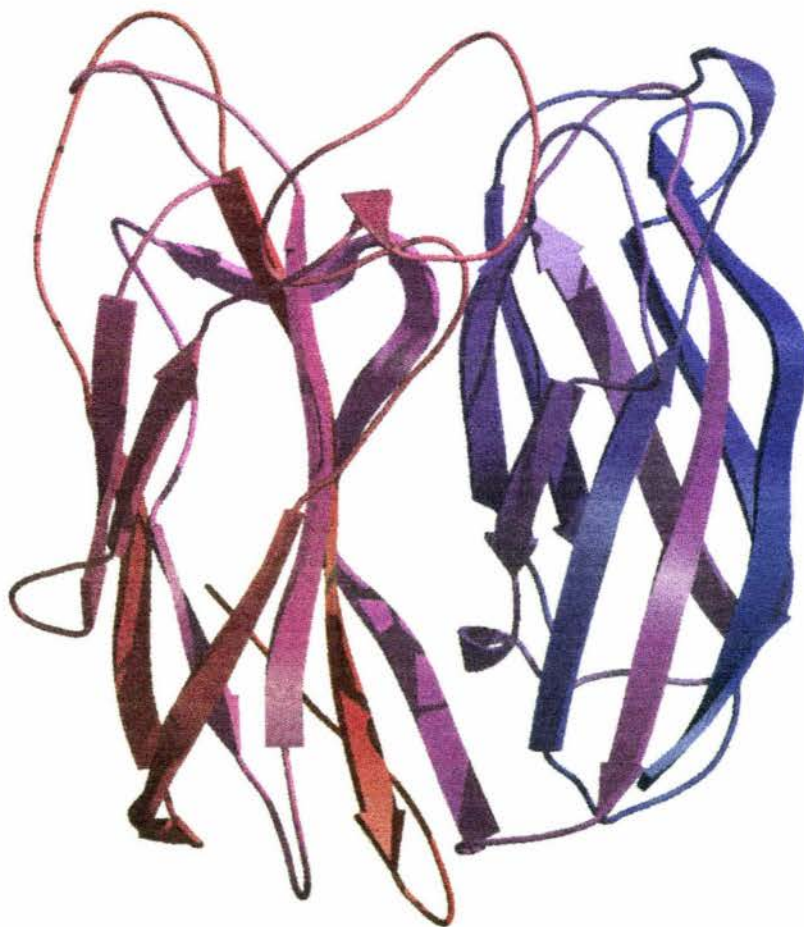
**Figure 1.9 Topology of the PNGase F molecule.**

The  $\beta$ -strands are identified using the convention adopted for the viral coat proteins and the residue number associated with each strand are given. The inset shows the classic eight-stranded  $\beta$ -jelly roll motif for comparison. (Figure from Norris *et al.*, 1994a.)

A short extended piece of polypeptide (residues 136-141) links the two domains at the bottom of the molecule. At the top of the molecule, several connecting loops from domain 2 reach across to interact closely with domain 1 to tie the domains together (figure 1.10). The most important of these comprises residues 227-257 that links



strands F and G of domain 2 and forms a double loop in which residues 227-249 form the first part. This loop extends to and back from domain 1 and includes a wide  $\Omega$  loop between residues 231-245. The second part of the loop is formed by residues 250-257 and is tied to the first part by a disulfide bridge, 231-252. At the back of the molecule, the loop 151-159, which links strands B and C of domain 2, also stretches across towards domain 1, with residues 151-155 making a number of hydrogen bonds. These loops provide most of the inter-domain interactions and also play a major role in forming the active site. The domains are packed back to back to create an approximately rectangular molecule of overall dimensions 50 x 45 x 30Å (Norris *et al.*, 1994a).

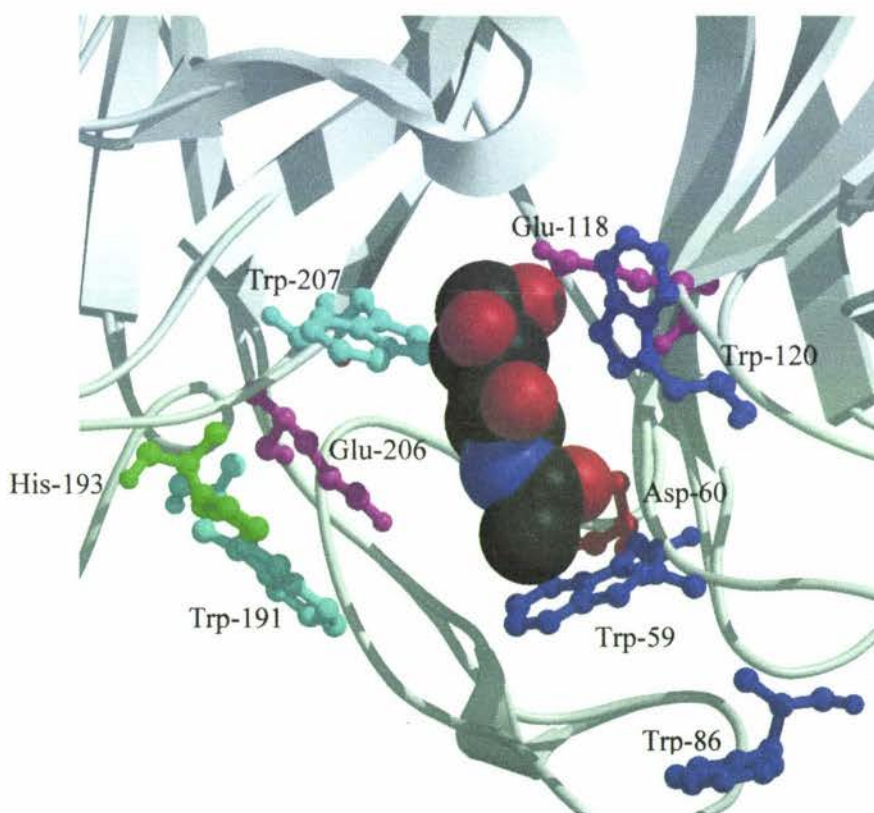


**Figure 1.10 The PNGase F structure.**

The protein is folded into two domains, each with an eight-stranded, antiparallel  $\beta$ -jelly roll configuration similar to structures found in lectins. Domain 1 is on the right and domain 2 is on the left. This figure is produced by the program MolScript (Kraulis *et al.*, 1991) and viewed with Raster3D (Merritt and Bacon, 1997).

#### 1.2.5.3.4 The Active site of PNGase F

Khun *et al.*, (1995) found that a disaccharide, GlcNAc- GlcNAc, acted as an inhibitor to the enzyme and was able to obtain a crystal structure with the inhibitor bound (Figure 1.11). The active site is located in a deep cleft at the interface between the two domains at the top of the molecule. This cleft, formed by long loops that connect the  $\beta$ -strands between the  $\beta$ -sheets, is lined by His-193 (light green) and five exposed tryptophan residues, Trp-59, Trp-86 and Trp-120 on domain 1 (blue), and Trp-191 and Trp-207 on domain 2 (cyan). Three acidic residues Asp-60 (red), Glu-206 and Glu-118 (magenta) located at the bottom of the cleft have been shown to be essential for activity by site-directed mutagenesis studies (Khun *et al.*, 1995). The D60N mutant has no detectable activity while E206Q and E118Q have less than 0.01 and 0.1% of the wild type activity, respectively.



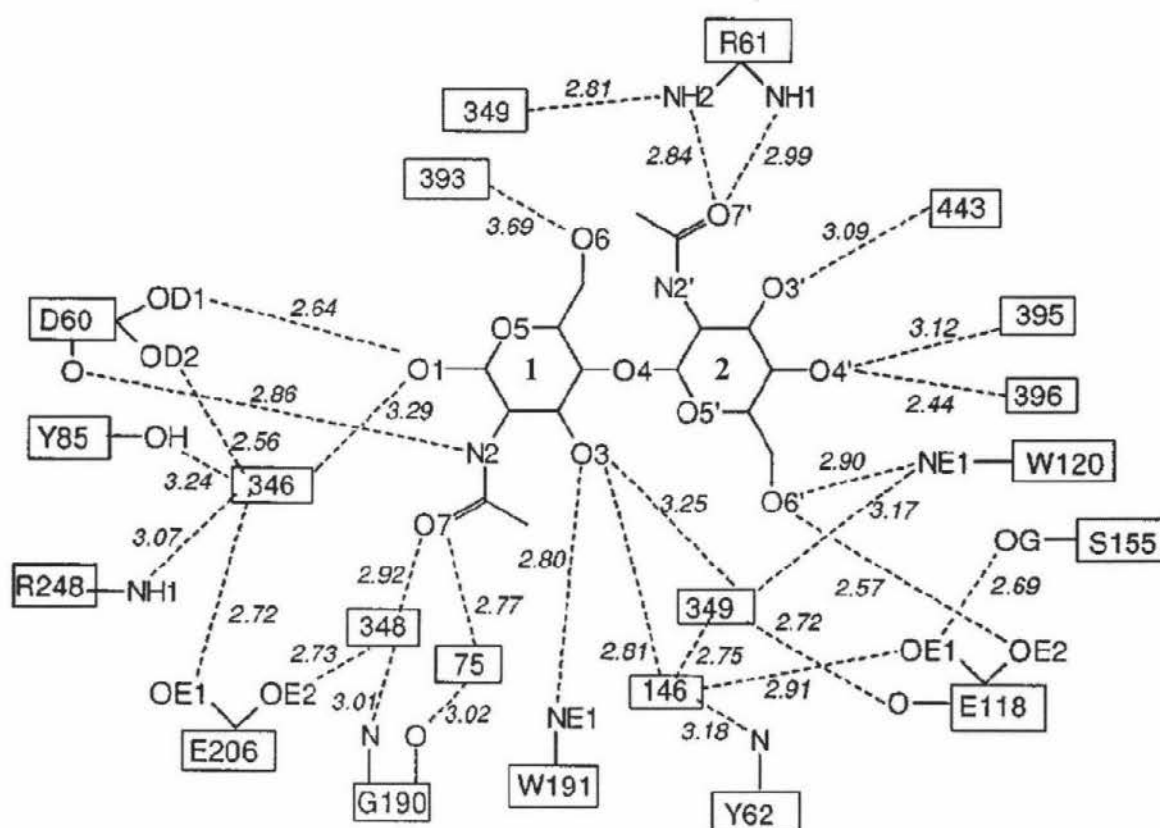
**Figure 1.11 The orientation of *N,N'*-diacetylchitobiose (CTB) inside the active site of PNGase F.**

The CTB, in  $\alpha$ -configuration, is positioned with the non-reducing end pointing into the cleft. The active site is lined by one basic (His-193 in light green), three acidic (Asp-60 in red and Glu-118 and Glu-206 in magenta), and five tryptophan residues (blue on domain 1 and cyan on domain 2). This figure is generated by MolScript (Kraulis *et al.*, 1991) and viewed with Raster3D (Merritt and Bacon, 1997).

Figure 1.12 is a schematic diagram of the contacts made between the dichitobiose inhibitor and the enzyme. It shows that the oxygen (OD1) of Asp-60 forms hydrogen bonds with O<sub>1</sub> of the reducing-end of the disaccharide substrate and the water molecule, Wat<sup>346</sup> that connects Asp-60 to Glu-206. This indicates Asp-60 is the primary catalytic residue and since Glu-206 is not in direct contact with the substrate, it may be important for stabilisation of the reaction intermediates or interaction with the O<sub>8</sub> of the substrate asparagine. Glu-118 forms a hydrogen bond with the O<sub>6</sub> of the second *N*-acetylglucosamine residue of the substrate and the low activity of the E118Q mutant is probably due to its reduced ability to bind the oligosaccharide. Thus it is probably responsible for positioning the CTB in the active site.

The contacts of the second GlcNAc are much weaker and only O<sub>6</sub> and O<sub>7</sub> are involved in direct hydrogen bonds with the protein and these may assist in CTB positioning. Two aromatic residues that line the active site also probably aid in the orientation of CTB in the active site. Trp-191 is positioned nearly perpendicular to the disaccharide and forms a hydrogen bond with the O<sub>3</sub> of the reducing-end *N*-acetylglucosamine residue of the substrate, while Trp-120 is positioned in a way to be able to make a hydrophobic contact with the first mannose residue in the trimannosyl core. If this structure represents the true binding of the dichitobiose core when it is linked to the polypeptide, it is clear that steric hindrance would prevent N-glycans with a fucose  $\alpha$ 1-3 substituted on the proximal GlcNAc binding as shown in this diagram.





**Figure 1.12.** Schematic diagram showing the intermolecular hydrogen bonding contacts between PNGase F, *N,N'*-diacetylchitobiose and water molecules.

Protein residues are indicated with single letter amino acid code and sequence number in boxes, and water molecules are indicated by a number analogous to their number in the Protein Data Bank. The reducing-end GlcNAc residue is on the left. Hydrogen bonding distances, in Å, are shown in *italics*. Note that Wat<sup>349</sup> is present twice, once in contact with O<sub>3</sub> and with R61. (Figure reproduced from Khun *et al.*, 1995.)

### 1.2.6 Constructs of PNGase F

*N*-glycosylation is cell- and species-specific and the oligosaccharide moieties play important roles in the cell. It is therefore, of considerable interest to biochemists studying the structure and function of such oligosaccharide moieties to be able to remove the intact oligosaccharides from glycoproteins, without compromising the integrity of the protein. Several constructs for the *E. coli* based expression of recombinant PNGase F have been reported in the literature and are briefly discussed.



### 1.2.6.2 PNGase F Cloned as GST Fusions

In 1995, a Swedish group cloned PNGase F (ATCC 33958) into a pGEX-3X vector and expressed the enzyme as a fusion protein with a 26 kDa glutathione-S-transferase to give the plasmid, GST-PNGase F (Grueninger-Letich *et al*, 1995). Approximately 5 mg of pure GST-PNGase F was recovered per litre of culture giving a yield of 2.5 mg L<sup>-1</sup>. The fusion enzyme was active on a range of glycoprotein substrates including phytase, acid phosphatase and human renin.

### 1.2.6.3 A Clone of PNGase F from *F. meningosepticum* (CDC strain 3352)

The PNGase F DNA with no leader signal sequence was isolated from the genomic DNA of *F. meningosepticum* (CDC strain 3352) and cloned into the plasmid pT7-7 to produce the construct, pT7-PNG. The PNGase DNA cloned contained an extra 27 base pairs (bp) downstream from the initiation codon of pT7-7. The recombinant enzyme was mainly produced in inclusion bodies, although enough soluble enzyme was produced to demonstrate it was active on both glycoproteins and glycopeptides (Dr. G. E. Norris, personal communication). In order to produce a high copy number expression vector containing an f1 origin for single stranded DNA production, deletion PCR was carried out on pT7-PNG to move the DNA to start at the initiation codon at the unique *Nde* I site in pT7-7. The *Xba* I/*Pst* I fragment containing the PNGase F gene and ribosome binding site of pT7-PNG was subcloned into pBluescript KS(-) downstream of the T7 promoter, to produce the plasmid pKS-PNG. As with the previous construct, the protein was expressed at high levels in an insoluble state.

### 1.2.6.4 Inclusion Bodies Formation in *E. coli*

Despite extensive knowledge of the genetics of *E. coli*, not all genes can be expressed efficiently in this organism. This may be due to the subtle structural features of the gene sequence, the stability and translational efficiency of the mRNA, the ease of protein folding, degradation of the protein by host cell proteases, major differences in the codon usage between the foreign gene and native *E. coli* and the potential toxicity of the protein to the host (Makrides, 1996). Proteins expressed in *E. coli* often accumulate intracellularly in the form of inclusion bodies (Marston, 1986). Biologically active protein can be recovered from such aggregates by denaturation and refolding *in vitro* (Hockney, 1994).

A statistical analysis of the composition of 81 proteins that do and do not form inclusion bodies in *E. coli* showed that there are six parameters that are correlated with inclusion body formation. They are average charge, fraction of turn-forming residues, cysteine content, proline content, hydrophobicity and the total number of residues (Wilkinson *et al*, 1991). While there are some advantages in producing recombinant proteins as inclusion bodies, the process of denaturation and refolding into biologically active conformation is very empirical and, in most cases, not applicable for the efficient reconstitution of biologically active proteins (Rudolfg and Lilie, 1996).

It has been shown that the post-translational folding of proteins, the assembly of polypeptides into oligomeric structures and the localisation of proteins are mediated by specialised host proteins known as molecular chaperones, such as GroES and GroEL proteins in *E. coli* (Backer and Craig, 1994; Clarke, 1996; Ellis and Hartl, 1996). Several of these chaperones are heat-shock proteins whose synthesis is induced in response to stress. The exact mechanism of chaperone-assisted protein folding is still unclear, but it has been suggested chaperones assist protein folding by preventing unproductive side reaction such as aggregation (Rudolf and Lilie, 1996). However, the experiments where chaperones were used to assist protein folding have been inconclusive with the effects of chaperone co-production on gene expression in *E. coli* being protein specific (Wall and Pluckthun, 1995). One solution to the solubility problem is to export proteins to the periplasm.

## 1.3 Targeting and Assembly of Proteins in the Bacterial Periplasm

All cells sequester biological activities in subcellular compartments that are bound by lipid bilayers. If lipid bilayers are considered as compartments, then Gram-negative bacteria such as *Escherichia coli* can be divided into four regions: Cytoplasm, inner membrane, periplasm, and outer membrane. The latter three compartments can also be viewed as one structure that envelops and separates the cytoplasm from the external environment, and are collectively known as the bacterial envelope. While compartmentalisation is essential for viability, it also poses a problem for the cell. For example in *E. coli*, all proteins are synthesised in the cytoplasm and proteins destined for the periplasm and the outer membrane must be translocated across the inner membrane and then exported to their specific targets. This process involves the proteins being synthesised as precursor proteins. These proteins contain an amino-terminal signal sequence that directs the precursors to a collection of proteins that catalyse precursor translocation across the inner membrane. During translocation, the signal sequence is proteolytically cleaved, generating a mature form of the protein, which is then transported to its appropriate destination.

### 1.3.1 The Bacterial Periplasm

The periplasm is the region between the inner and outer membranes of Gram-negative bacteria and constitutes about 30% of the total cell volume of *E. coli* (van Wielink and Duine, 1990). It has a gel-like structure, filled with a peptidoglycan matrix that is progressively more tightly cross-linked towards the outer membrane (Hobot *et al.*, 1984), and has been shown to be impermeable to whole nucleotides. The non-reducing environment of the periplasm favours disulfide bridge formation in proteins thus is of particular interest in the heterologous periplasmic expression of recombinant proteins, such as PNGase F, that contain multiple disulfide bonds which contribute to their stability and in some cases, their catalytic activity.

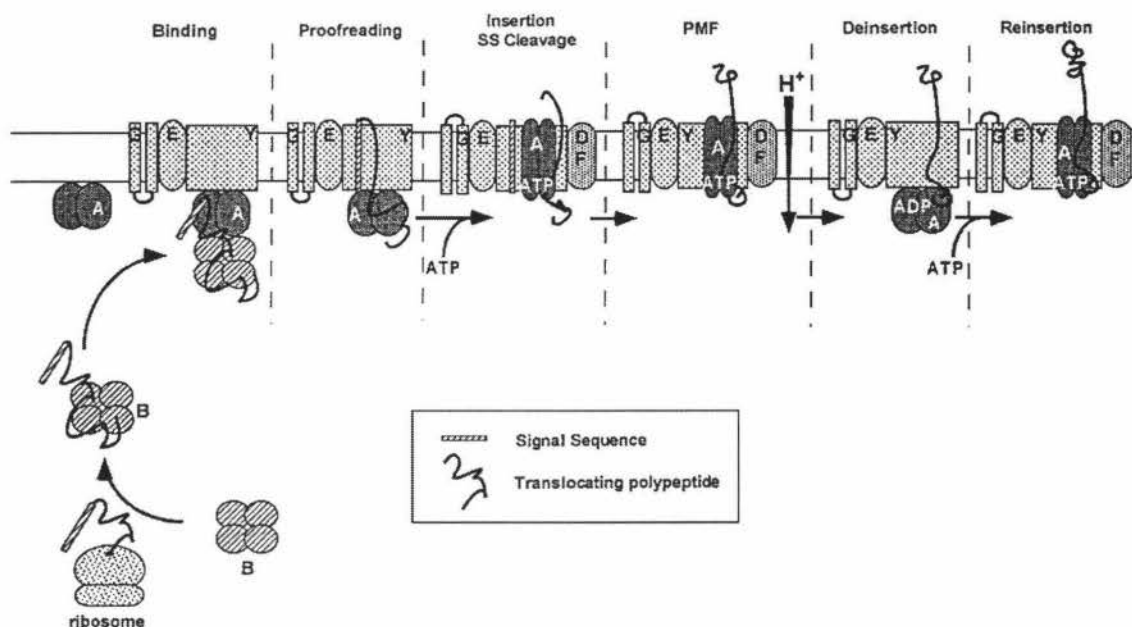
### 1.3.2 The Signal Peptide

Proteins destined for export are synthesised as precursor proteins that contain amino-terminal signal sequences. These signal peptides have been recognised to possess three distinct regions based on length, hydrophobicity and conformation (von Heijne, 1990). The amino-terminal region of the signal peptide is about 5-8 amino acids long and is characterised by the presence of basic residues. The net positive charge is essential for interaction with the negatively charged surface of the inner membrane (Inouye *et al.*, 1982). The central region is about 8-12 non-polar amino acids long and has a high inclination for  $\alpha$ -helical formation that may facilitate translocation across the bilayer (Engelman and Steitz, 1981). The carboxyl cleavage region is typically about 3-9 amino acids long (Jain *et al.*, 1994), and involved in signal peptidase recognition and cleavage during folding and localisation of the exported protein. Many recombinant proteins when expressed intracellularly in *E. coli* at high levels, become inclusion bodies. The introduction of an appropriate signal sequence to the 5' end of the sequence encoding the gene of interest usually results in the polypeptide chain being exported into the periplasm of the bacteria. One example is the addition of a 63 base pair long signal sequence from a major outer membrane in *E. coli*, OmpA, to  $\beta$ -lactamase (Ghrayeb *et al.*, 1984). Upon induction of gene expression,  $\beta$ -lactamase was secreted into the periplasm with a correctly processed amino-terminus.

### 1.3.3 The Sec Export Pathway

Compartmentalisation inside *E. coli* demands nascent extracytoplasmic proteins to be actively transported to destinations such as the periplasm and outer membrane. To perform this duty, *E. coli* employs a series of Sec (secretion) proteins arranged in a hexameric complex, known as the preprotein translocase, to catalyse the translocation of various polypeptides through the inner membrane. The multi subunit enzyme is composed of an integral membrane domain, SecYEGDFyajC, and a peripheral membrane domain, SecA (Wickner and Rice-Leonard, 1996). Genetic and biochemical studies have identified the core subunits of this enzyme to be the SecY, SecE and SecA proteins (Schatz and Bechwith, 1990; Akimaru *et al.*, 1991; Duong and Wicker, 1997). Translocation depends upon the energy of ATP hydrolysis by SecA (Chen and Tai, 1985) and is strongly stimulated by proton motive force (PMF) across the membrane (Geller *et al.*, 1986). A model for protein translocation across the inner membrane by preprotin translocase is shown in figure 1.14.





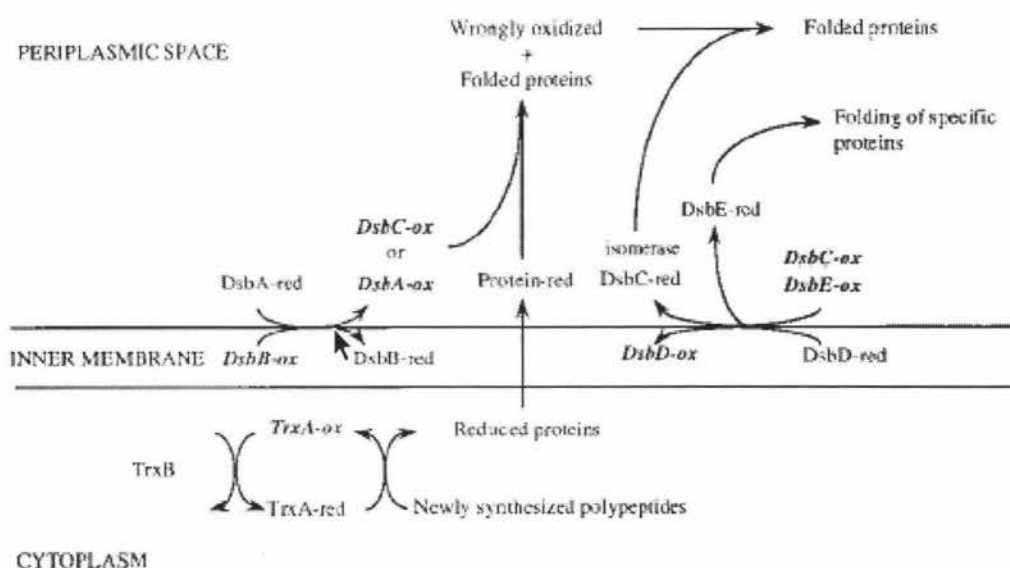
**Figure 1.14** A model for protein translocation across the inner membrane.

Beginning from the lower left, a precursor protein emerges from the ribosome and interacts with SecB. The SecB/precursor complex then interacts with membrane-associated SecA at the cytoplasmic face of the SecEGY complex (Binding). The signal sequence of the precursor is verified, and SecB is released (Proof-reading). Binding of ATP to the SecA high-affinity site causes insertion of SecA and bound 20-30 residues of the precursor into the membrane. The signal sequence is cleaved (Insertion/SS Cleavage). SecD/F stabilise the membrane-inserted form of SecA and the proton motive force stimulates additional translocation of the SecA-bound polypeptide (PMF). ATP hydrolysis causes SecA to deinsert from the membrane (Deinsertion). Additional ATP binding causes reinsertion of SecA into the membrane. This reinsertion is accompanied by additional stepwise protein translocation. (Figure reproduced from Danese and Silhavy, 1998.)

### 1.3.4 Protein Folding in the Periplasm

Protein folding in the cytoplasm is catalysed by the two families of chaperones, Hsp 60 and Hsp70, also known as the GroEL and DnaK proteins, respectively in *E. coli* (Georgopoulos, 1992; Gething and Sambrook, 1992). These chaperones stabilise unstable conformers of proteins during folding, translocation and assembly, and are expressed at higher levels at elevated temperatures to prevent drastic protein aggregation and proteolysis of damaged proteins (Hendrick and Hartl, 1993). However, these chaperones have not been located in the periplasm, probably because ATP, which is essential for chaperone activity, is absent. Instead, two types of folding catalysts have been found in the periplasm that accelerate only a specific rate-limiting step of the

folding reaction. One type of slow conformational rearrangement, caused by thiol-disulfide exchanges during disulfide bond formation or rearrangement, is catalysed by protein disulfide isomerases (PDI). An example of the PDI in *E. coli* is the Dsb family of proteins that belongs to the thioredoxin superfamily. Dsb proteins do not share overall sequence homology with thioredoxins but they share at least one common active site, which is Cys-X-X-Cys (Missiakas and Raina, 1997). A proposed model of disulfide formation catalysed by various Dsb proteins in the periplasm of *E. coli* is shown in figure 1.15.



**Figure 1.15** Mode of disulfide formation catalysed by the various Dsb enzymes in the periplasm of *E. coli*.

(Figure reproduced from Missiakas and Raina, 1997.)

Peptide bonds display a partial double-bond character that force the carbonyl and amino group into a planar structure. The  $\alpha$ -carbons attached, to the carboxyl carbon and the amide nitrogen can assume a *cis* or *trans* conformation with respect to each other. For most dipeptides, these groups are predominately in the *trans* conformation with the exception of peptide bonds formed from X-proline dipeptides (X is any residue) which can be in either conformation (Levitt, 1981). The *cis/trans* isomerisation of prolyl peptides is a slow process that impedes protein folding in proteins such as in bovine pancreatic trypsin inhibitor derivative RCAM(14-38) (Jullien and Baldwin, 1981). Peptidyl prolyl isomerases (PPI) catalyse the isomerisation around the X-Pro peptidyl bonds. Four PPIases have been found in *E. coli* periplasm, RotA, FklB, FkpA, and SurA.



## 1.4 The Scope of this Project

PNGases are widely used in research for both investigations into the function and structural characterisation of the glycan moieties on glycoproteins and in the removal of heterogeneous sugar chains from glycoproteins to aid their crystallisation. Studies of the sugar moieties of glycoproteins are particularly important in the biotechnology industry as many recombinant proteins designed for therapeutic use are in fact glycosylated and many disease states exhibited changes in oligosaccharide structures. Therefore, it is of great interest to be able to produce relatively large quantities of the enzymes to aid in such studies. The focus of this project is on PNGase F that is one of the best-characterised PNGases but several important aspects of the enzyme have not been addressed yet:

- (1) The protein is secreted in small amount by *F. meningosepticum* and the purification of the secreted enzyme is a tedious multi step task with a low yield (at best  $< 0.5 \text{ mg L}^{-1}$  culture medium).
- (2) The biochemical properties and the kinetic aspects of the protein have not been determined in great detail. The  $K_m$  of the PNGase F from ATCC 33958 strain of *F. meningosepticum* on a ovalbumin-derived glycopeptide has been reported as being  $31 \text{ } \mu\text{M}$  (Altman *et al.*, 1995).
- (3) The three-dimensional structures of PNGase F has been determined and site-directed mutagenesis studies have identified the residues in the active site responsible for oligosaccharide recognition and asparagine-sugar amide bond cleavage. However, the mechanism of hydrolysis remains unclear, as does the question of recognition of the polypeptide component of the substrate that requires a minimum of a tripeptide for activity.

In order to answer these questions, sufficient enzyme must be available and crystallographic studies require ultra pure proteins. The primary aim of the project was therefore to design an expression system for the production of large quantities of pure PNGase F. Such a system is necessary to pursue further studies on the enzyme such as the complete biochemical characterisation of the recombinant enzyme including its

substrate specificity and the molecular determinants for this specificity. A long-term aim is to develop an immobilised enzyme that will deglycosylate proteins without the need to separate the enzyme from deglycosylated protein.

This thesis is divided into three parts: (i) The engineering of an expression vector for the high-yield production of PNGase F in *E. coli* described in chapter two. (ii) The expression and purification of both native and recombinant PNGase F outlined in chapter three. (iii) Detailed characterisation of recombinant PNGase F.

## **Chapter 2 Construction of an Expression Vector for the Production of PNGase F in *E. coli***

### **2.1 Introduction**

Purification of PNGase F from the culture medium of *F. meningosepticum* is a tedious multi step task with very low yield (typically  $< 0.5 \text{ mg L}^{-1}$  culture). The access to relatively large quantities will allow more in depth studies of both the hydrolysis mechanism and the molecular determinants for substrate recognition, as well as facilitate general studies on the biological role of glycans. The advantages of producing a recombinant enzyme over native include the increase of production and increased ease of purification. It also eliminates the need for purifying the enzyme from the mildly infectious microorganism, *F. meningosepticum*.

### **2.2 Experimental Objectives and Strategies**

The primary aim of the work was to remove the coding region of PNGase F from the pT7-PNG vector and insert it into a vector compatible with a high-yield *E. coli* based expression system. The immediate goal was to engineer and optimise a high-yield expression system along with a simple purification scheme to provide an abundant and convenient source of pure enzyme for use in biochemical characterisation and crystallisation studies. Because of the difficulties encountered with purification of the enzyme, it was decided to add DNA coding for a poly-histidine tag (His-tag) to the vector in order to simplify the purification of the recombinant protein (Hochuli *et al.*, 1987). Such a tag, usually five to six histidines long on either terminus of the enzyme, will preferentially bind to immobilised divalent metal ions and can be eluted virtually contamination free. His-tags are most commonly added to the N-terminus of recombinant proteins and they are usually designed to incorporate a specific protease site that allows the removal of the tag. In most proteins, the N-terminus is exposed on the surface instead of buried inside the core of the protein and as such should not interfere with protein folding. However, in the case of PNGase F, structural analysis (Norris *et al.*, 1994b; Tarentino *et al.*, 1994b) shows that both N- and C-terminal residues are surface-exposed and well removed from the active site. Hence, the addition of a His tag to the C-terminus in this study should not affect the activity of the enzyme.

The vector chosen for the cloning of PNGase F was pKS-OmpA3-His, a pBluescript KS(-) derived plasmid that contains an *OmpA* signal peptide downstream from a T7 promoter, a ribosome-binding site (rbs) and sequence encoding a hexa-histidine tag (Figure 2.1). The RNA polymerase of bacteriophage T7 is very selective for specific promoters, and efficient termination signals are rare, so that T7 RNA polymerase is able to make complete transcripts of almost any open reading frame that is placed under control of a T7 promoter. Furthermore, T7 RNA polymerase is able to elongate about five times faster than *E. coli* RNA polymerase, making T7 RNA polymerase ideal as a promoter for heterologous protein expression in *E. coli* (Studier *et al.*, 1990). The *OmpA* signal peptide is from an *E. coli* membrane protein and it has been shown that this addition of the signal sequence to the 5' end of the sequence encoding the gene of interest results in the polypeptide chain being exported into the periplasm of the bacteria (Takaara *et al.*, 1985; Favconnier *et al.*, 1998; Pines *et al.*, 1999). Such strategy often prevents the formation of inclusion bodies that is a common consequence of using the T7 expression system. The appropriately positioned ribosome-binding site in the plasmid enables efficient translation of mRNA, while the histidine tag provides for a simple one-step purification of the expressed protein. Colony selection is provided by ampicillin whose antibacterial property is neutralised by the 31.5 kDa *bla* gene product,  $\beta$ -lactamase.

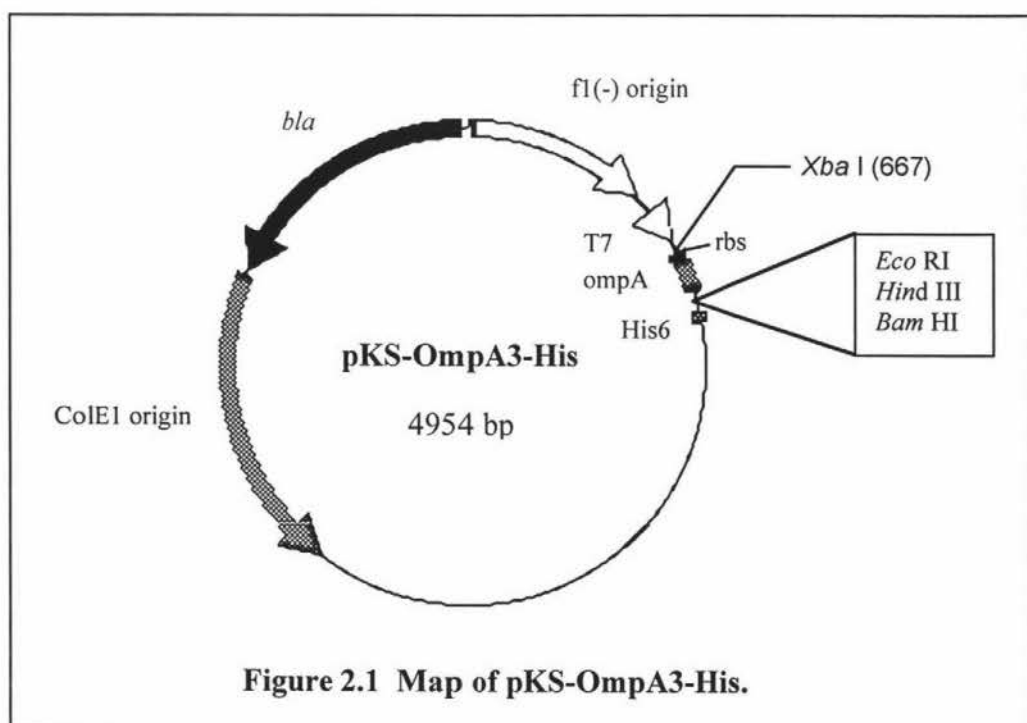


Figure 2.1 Map of pKS-OmpA3-His.

## 2.3 Materials

### 2.3.1 Chemicals and Enzymes

The chemicals listed below were used throughout the course of this project and were of analytical grade or better obtained from the following companies:

Ajax Chemicals (Sydney, Australia):

Ammonium acetate, potassium chloride, manganese chloride.

BDH Laboratory Supplies (Poole, England):

Sodium chloride, sodium hydroxide, sodium acetate, methanol, acetic acid, phenol, chloroform, hydroxyquinoline, hydrochloric acid, PEG8000, glycerol, calcium chloride.

Boehringer Mannheim (West Germany):

*Taq* and *Pwo* polymerase, T4 DNA ligase, calf intestinal alkaline phosphatase (CIAP), ribonuclease A, lysozyme.

Bio-Rad Laboratories (California, USA):

Quantum Prep plasmid miniprep kit, SDS.

Difco Laboratories (Detroit, Michigan, USA):

LB broth concentrate.

GibcoBRL (Scotland, UK):

Synthetic oligonucleotides, 1kb DNA ladders, restriction endonucleases, yeast extract, tryptone, bactopectone, bacteriological agar.

United States Biochemical Corporation (Ohio, USA):

Bromophenol blue, casamino acids.

Serva Feinbiochemica GmbH & Co. (Heidelberg, West Germany):

Tris base, EDTA.

Sigma Chemical Company (St. Louis, USA):

Ampicillin, tetracycline, type I-A: low electroendosmosis (EEO) agarose, ethidium bromide (EtBr), dimethyl sulphoxide (DMSO), and 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside (X-Gal).

Ultra pure water ( $>16 \text{ M}\Omega \text{ cm}^{-1}$ ) used in this study was supplied from a Sybron/Barnstead NANOpure II filtration system (MA, USA) containing two ion-exchange and two organic filtration cartridges. This water will be referred to as Milli-Q water throughout this study. Sterile Millex-GS  $0.22 \mu\text{m}$  syringe filters were supplied by the Millipore Corporation (MA, USA).

### 2.3.2 Plasmids and Bacterial Strains used in This Study

The *Flavobacterium meningosepticum* culture used in this study was sourced from the Massey University Microbiology Department culture collection (480A), which in turn was obtained from the New Zealand National Health Institute culture collection (NHI 847). This culture came from the United States Communicable Disease Centre culture collection (CDC strain 3552) and was originally collected from the spinal fluid of an infant with meningitis (King, 1959).

*Escherichia coli* BL21(DE3) (Studier *et al.*, 1990) and XL1-Blue (Stratagene) were laboratory strains (IMBS, Massey University). The plasmid pT7-PNG was a kindly donated by my supervisor Dr. G. E. Norris. The plasmids pKS-OmpA3-His, and pBluescript KS(-) were generous gifts from Dr. J. S. Lott (School of Biological Sciences, University of Auckland, N.Z.).

### 2.3.3 Bacterial Growth Media

*F. meningosepticum* was grown in M9 media (Davis *et al.*, 1980) with 0.5% casamino acids added. Luria broth (LB) broth was used for growth of *E. coli*. All media were sterilised by autoclaving for 20 minutes at  $121^\circ\text{C}$  and 15 psi. Agar media was prepared from solidifying liquid media by the addition of  $15 \text{ g L}^{-1}$  agar prior to autoclaving. Stock solutions of ampicillin ( $100 \text{ mg mL}^{-1}$  in Milli-Q water) were filter sterilised using Millex-GS  $0.22 \mu\text{m}$  syringe filters and stored at  $-20^\circ\text{C}$ . Tetracycline stock solution ( $20 \text{ mg mL}^{-1}$ ) was made up in 100% ethanol, filtered and stored.

#### 2.3.3.1 M9 with 0.5% Casamino Acids

5 g L<sup>-1</sup> Casamino acids, 6 g L<sup>-1</sup> Na<sub>2</sub>HPO<sub>4</sub>, 3 g L<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub>, 0.5 g L<sup>-1</sup> NaCl and 1 g L<sup>-1</sup> NH<sub>4</sub>Cl. 1 mL each of sterile 100 mM CaCl<sub>2</sub> and 1 M MgSO<sub>4</sub> were added per litre of medium post-sterilisation to avoid salt precipitation.

#### 2.3.3.2 LB Broth

10 g L<sup>-1</sup> bacto-tryptone, 5 g L<sup>-1</sup> bacto-yeast extract and 10 g L<sup>-1</sup> NaCl (pH 7.0) or pre-mixed LB broth base (Difco Laboratories) was used. 100 µg mL<sup>-1</sup> ampicillin or 12.5 µg mL<sup>-1</sup> tetracycline was added after autoclaving when the temperature of the medium was less than 50°C.

#### 2.3.3.3 SOB Medium

20 g L<sup>-1</sup> bacto-tryptone, 5 g L<sup>-1</sup> bacto-yeast extract, 0.5 g L<sup>-1</sup> NaCl, 2.5 mM KCl and 10 mM MgCl<sub>2</sub>, pH 7.0

### 2.3.4 Storage and Propagation of Bacterial Cultures

For long term storage, all strains were stored at -70°C as frozen glycerol stocks. Small 5mL cultures were grown in media containing appropriate antibiotic if necessary (M9 with 0.5% casamino acids for *F. meningosepticum*; LB with 100 µg mL<sup>-1</sup> ampicillin for *E. coli*) until the optical density at 600 nm reached ~0.6. The cultures were mixed with 30% (v/v) glycerol before being dispensed as 300 µL aliquots, snap frozen in liquid nitrogen and stored at -70°C.

The propagation of frozen cultures involved the cultures to be streaked onto appropriate agar plates and incubated at 37°C for 12-16 hours. Liquid cultures were grown from single colonies, taken from agar plates, at 37°C for 12-16 hours with shaking (200 rpm). Plasmids were selected and maintained in *E. coli* in the presence of suitable antibiotics at all stages.



## 2.4 Methods used for Cloning

Most procedures were carried out as described by Sambrook *et al.*, (1989). All glassware, plastics, media and solutions were sterilised by autoclaving as described in section 2.3.3.

### 2.4.1 Phenol:Chloroform Extraction of DNA from a DNA/Protein Mixture

This standard method to remove protein contaminants from nucleic acid solutions is described in appendix E.3 of Sambrook *et al.*, 1989. An equal volume of phenol:chloroform [50% (v/v) phenol, 48% (v/v) chloroform, 2% (v/v) isoamyl alcohol, 0.1% (w/v) hydroxyquinoline, pH8.0] was added to the DNA sample in a microfuge tube and mixed by inversion before being centrifuged at 14,000 g for 15 seconds in a bench-top microfuge at room temperature. The aqueous phase was transferred to a fresh tube and the extraction was repeated until no protein was visible at the interface of the organic and aqueous phases. Any traces of phenol remaining in the aqueous phase were removed by extraction with an equal volume of chloroform.

### 2.4.2 Ethanol Precipitation of DNA

DNA was recovered by precipitation with 0.2 volumes of 3 M ammonium acetate (pH 5.2) and 2 volumes of ethanol (-20°C) followed by centrifugation (14,000 g, 10 minutes) at 4°C. The supernatant was removed by aspiration and the DNA pellet rinsed with 70% ethanol (-20°C) to remove any residual chloroform. The pellet was then centrifuged at 4°C (14,000 g, 2 minutes) and dried under vacuum in a Speedvac (Savant, NY, USA) before being resuspended in a minimum volume (~20 µL) of TE buffer, pH 8.0 (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

### 2.4.3 Size-Selective Polyethylene Glycol (PEG) Precipitation of DNA

This protocol was adapted from Lis *et al.*, (1980) and is based on the size-selected precipitation of DNA fragments. The PNGase F DNA from either an endonuclease digestion or a PCR amplification could not be precipitated by conventional methods as described in section 2.4.2. Therefore a final concentration of 10% PEG (w/v) in the presence of 10 mM magnesium chloride was successfully used to precipitate the DNA with a recovery of about 90%.



A DNA mixture containing the PNGase F DNA was mixed with an equal volume of 20mM MgCl<sub>2</sub> / 20% PEG-8000 then centrifuged at 14,000 *g* for 10 minutes in a microfuge. The supernatant containing small non-precipitated DNA fragments was discarded and the DNA pellet with the PNGase F DNA was suspended in a minimum volume (~20 µL) of TE buffer, pH 8.0 (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

#### 2.4.4 Agarose Gel Electrophoresis

DNA fragments were separated on the basis of size by agarose gel electrophoresis in 1x TAE buffer (40 mM Tris-HCl, 20 mM acetic acid and 2 mM EDTA, pH 8.0). 1% agarose gels were prepared by melting 0.4 g agarose in 40 mL 1x TAE buffer. The gels were run in a Bio-Rad “Mini-Sub” electrophoresis unit tank using 1x TAE buffer following a standard protocol (Sambrook *et al.*, 1989). 10-20 µL samples were loaded using 0.2% (w/v) bromophenol blue in 50% (v/v) glycerol as 6x loading buffer. Electrophoresis was performed at 3-5 V cm<sup>-1</sup> (typically 50V) until the dye front had migrated to the other end of the gel. The gel was stained in a 0.5 µg mL<sup>-1</sup> ethidium bromide bath and the DNA bands were visualised by ultraviolet (UV) light illumination at 302 nm. Permanent records of gels were obtained with a Gel Doc system using Alpha Imager 2000 software (model TM 20, Alpha Innotech Corporation, USA).

#### 2.4.5 Quantitation and Size Determination of DNA Fragments

Quantitative DNA estimation was done by comparing relative intensities of the bands with known concentrations of a plasmids or fragments of similar size following agarose gel electrophoresis. The 1 kb DNA ladder from Gibco BRL was used in a similar fashion to determine fragment size by comparing the relative migration distance of the standards versus sample.

#### 2.4.6 Amplification of PNGase F gene by PCR

Two oligonucleotide PCR primers PNG1 [ 5' ATG GAA TTC CAG CTC CGG CAG ATA AA CCG TAA AT 3' ] and PNG2 [ 3' TAA TCA CG GGC CAT CAA TGT TTG AAC CTA GGC CT 5' ] synthesised by Life Technologies were designed from the known sequence of the American variant of the PNGase gene (ATCC strain 33958; Tarentino *et al.*, 1990).

The optimum conditions for amplification were obtained by trial PCR experiments in which both the concentration of  $Mg^{2+}$  and the annealing temperature were varied. Conditions for a typical PCR reaction were as follows: 100 ng of pT7-PNG template was amplified with 50 pmol of both primers using 2.5 units of *Taq* polymerase in 50  $\mu$ L total reaction volume inside a 500  $\mu$ L Eppendorf tube. The mixture contained a final concentration of each of the 200  $\mu$ M deoxyribo-nucleotide triphosphate (dNTP), 1x standard  $Mg^{2+}$  free polymerase buffer and 1.5 mM  $MgCl_2$ . The thermo-cycling program that successfully amplified PNGase F was initial denaturation at 95°C for 5 minutes and this was followed by 30 cycles of primer annealing at 40°C for 1 minute with extension at 72°C for 1 minute.

#### 2.4.7 Digestion of DNA with Restriction Endonucleases

200 ng of the PNGase F gene and vector were digested with 5 units of *Eco* RI and *Bam* HI in a final volume of 50  $\mu$ L of 1x REACT3™ buffer (50 mM Tris-HCl, 10 mM  $MgCl_2$ , 100 mM NaCl, pH 8.0, Gibco BRL) at 37°C for 2-3 hours. For larger amounts of DNA, quantities were proportionally scaled up and the mixtures were incubated for longer periods of time to ensure complete digestion as determined by gel electrophoresis.

#### 2.4.8 Purification of DNA from Agarose Gels

After digestion by restriction endonucleases or amplification by PCR, DNA fragments were separated by electrophoresis in agarose gels of the appropriate concentration. DNA fragments were purified from agarose gels by electroblotting onto DEAE-cellulose membranes as described in section 6.24 of Sambrook *et al.*, (1989). DNA fragments were separated by agarose gel electrophoresis, stained with ethidium bromide and the bands of interest were located under UV illumination at 302 nm. An incision was made with a razor blade in the gel directly in front of those bands and about 2 mm wider than the band on each side. A piece of DEAE-cellulose membrane (NA-45, Schleicher and Schuell) the same length as the incision was cut and soaked in Milli-Q water for 10 minutes. Using blunt-ended forceps, the membrane was inserted into the slit and the incision closed so no air bubbles were trapped. Electrophoresis was then resumed until the band of DNA had migrated onto the membrane. The membrane was rinsed with 5-10 mL of low-salt wash buffer (50 mM Tris-HCl, 150 mM NaCl, 10 mM

EDTA, pH 8.0) at room temperature to remove any agarose adhering to the membrane. The membrane piece was then transferred to a clean microfuge tube and the bound DNA was eluted by incubation with a minimal volume (~300  $\mu\text{L}$ ) of high-salt buffer (50 mM Tris-HCl, 1 M NaCl, 10 mM EDTA, pH 8.0) for 30 minutes at 65°C. The membrane was transferred to a fresh tube of high salt buffer and the incubation was repeated for 15 minutes. The membrane was checked for residual DNA before disposal. The two aliquots of eluted DNA in high-salt buffer were combined and extracted once with phenol:chloroform as outlined in section 2.4.1 and then precipitated as described in section 2.4.3.

### 2.4.9 Ligation of DNA Fragments

In order to prevent self-ligation and hence increase the probability of ligation between the insert DNA and the plasmid, the digested plasmid was 5'-dephosphorylated using calf intestinal alkaline phosphatase (CIAP). Ligations with bacteriophage T4 DNA ligase were carried out in the smallest volume possible (typically 10  $\mu\text{L}$ ) and incubated at 12-16°C overnight or at room temperature for 3-4 hours (Sambrook *et al.*, 1989).

### 2.4.10 Preparation of Competent *E. coli* Cells

Ultra-competent BL21(DE3) or XL1-Blue cells were produced using the protocol of Inoue *et al.*, (1990). The desired *E. coli* strain was streaked from a frozen stock onto LB agar plates containing an appropriate antibiotic if necessary (12.5  $\mu\text{g mL}^{-1}$  tetracycline for XL1-Blue) and incubated at 37°C for 12-16 hours. 10-12 isolated colonies were used to inoculate 300 mL SOB medium in a 2-liter conical flask. The culture was then incubated at 18 or 22°C with shaking (200 rpm) and grown until the optical density at 600 nm reached approximately 0.6 absorbance units. The cells were placed on ice for 10 minutes before being pelleted by centrifugation (3,000  $g$  for 10 minutes at 4°C). The cell pellet was gently resuspended in 80 mL of ice-cold transformation buffer (10 mM PIPES, 55 mM  $\text{MnCl}_2$ , 15 mM  $\text{CaCl}_2$ , 250 mM KCl, pH 6.7) and incubated on ice for 10 minutes. The cells were again centrifuged as before and the resulting pellet was gently resuspended in 24 mL of transformation buffer. Cold dimethyl sulphoxide (1.8 mL of DMSO) was added drop-wise and mixed with gentle swirling. The cells were placed on ice for 10 minutes before being dispensed into 300  $\mu\text{L}$  aliquots, snap frozen in liquid nitrogen and stored at -70°C.

#### 2.4.11 Transformation of Ligated Plasmids into Competent *E. coli* Cells

Transformation into competent *E. coli* cells was performed as described by Inoue *et al.*, (1990). 100  $\mu\text{L}$  of competent cells were thawed and incubated on ice for 30 minutes with 1-10  $\mu\text{L}$  of plasmid DNA. After this time, the cells were heat-shocked at 42°C for 30 seconds then returned to ice for 10 minutes before 0.9 mL of LB broth (without ampicillin) was added. The cells were then incubated at 37°C for one hour with shaking (200 rpm), followed by an appropriate dilution ( $1\text{-}10^5$ ) with sterile LB broth. The diluted cells were plated onto LB-agar plates containing the appropriate antibiotic and incubated at 37°C for 12-16 hours. When necessary, the  $\beta$ -galactosidase activity of transformed cells was assessed by plating the diluted cells onto LB plates containing 40  $\mu\text{g mL}^{-1}$  IPTG as the inducer and 200  $\mu\text{g mL}^{-1}$  X-Gal as the chromogenic substrate. Under these conditions, colonies containing the plasmid with a DNA insert that disrupts the *lac Z* gene, appear white. Colonies carrying the plasmid with no disruption of the *lac Z* gene are induced by IPTG to hydrolyse X-Gal and appear pale blue. This is another convenient selection process, in addition to the use of antibiotics, that ensures the transformed cells carrying the plasmid with the desired DNA are picked for expression.

#### 2.4.12 Small-Scale Preparation of Plasmid DNA

This procedure was performed to produce small amounts of the desired plasmid for sequencing, protein expression or storage after a successful ligation and transformation. The alkaline method was used to isolate plasmids from *E. coli* as described in section 1.25 of Sambrook *et al.*, (1989). Alternatively, a Quantum Prep plasmid mini-prep kit (Bio-Rad) was used following manufacturer's instructions.

Single, isolated *E. coli* colonies from agar plates were inoculated into 5 mL LB broth containing the appropriate antibiotic to maintain the plasmid, and incubated at 37°C for 12-16 hours with shaking (200rpm). Cells from 1.5 mL of this culture were pelleted using a bench-top microfuge (7,000 g, 1 minute) and resuspended in 200  $\mu\text{L}$  of Solution I (50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA, 50  $\mu\text{g mL}^{-1}$  RNase A, 0.8 mg  $\text{mL}^{-1}$  lysozyme, pH 8.0) by vigorous vortexing. The suspension was incubated at room temperature for 5 minutes before 400  $\mu\text{L}$  of Solution II [200 mM NaOH, 1% SDS

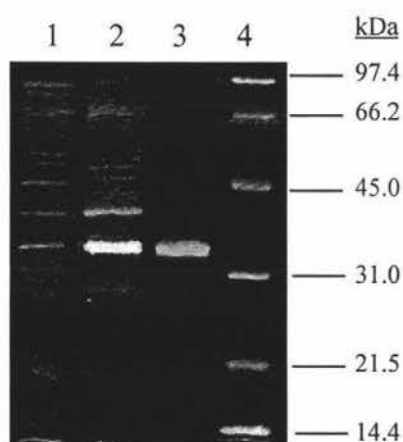
(w/v)] was added and the resulting solution was mixed by inversion. After another 5 minute incubation at room temperature, 300  $\mu$ L of Solution III (7.5 M ammonium acetate) was added and mixed thoroughly by inversion. The mixture was placed on ice for 10 minutes before being centrifuged in a microfuge at full speed (14,000 g, 10 minutes) at 4°C to pellet the cell debris. The supernatant was transferred to a fresh tube and the plasmid DNA precipitated by the addition of 2 volumes of ice-cold isopropanol. The DNA was pelleted by centrifugation (14,000 g, 10 minutes) at 4°C, rinsed with 70% ethanol (-20°C), dried and redissolved in minimal volume (~20  $\mu$ L) of TE buffer, pH 8.0 (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

#### **2.4.13 Sequence Analysis of DNA**

Confirmation of the DNA sequence was carried out using double-stranded plasmid as a template on an ABI Prism 377 automated DNA sequencer using the dye-termination method .

## 2.5 Results

*E. coli* strain BL21(DE3) cells transformed with the first plasmid construct, pKS-PNG, produced very high levels of protein, but unfortunately about 85% of the expressed PNGase F was found in inclusion bodies (figure 2.2). Experiments designed to optimise the growth and induction conditions to increase the production of correctly folded protein were undertaken with the following changes: (a) The culture growth temperature was reduced from 37 to 20°C, and (b) the concentration of the inducer, IPTG, was reduced from 1 to 0.2 mM. These changes improved the yield of soluble, active enzyme to ~25% of the expressed protein. Attempts to solubilise and refold these inclusion bodies following a protocol adapted from Hayes., (1998) were only moderately successful and the amount of material produced for the work involved made this procedure impractical for large-scale production of the enzyme. It was therefore decided to clone the gene into a vector that would export the nascent polypeptide in order to try to resolve the solubility problem of the previous construct.



**Figure 2.2** An inverse image of a 12% SDS-PAGE with silver staining shows the solubility problem of recombinant PNGase F expressed in *E. coli*.

Lane

- (1) Supernatant of sonicated *E. coli* carrying the pKS-PNG construct.
- (2) Cell pellet of sonicated *E. coli* carrying the pKS-PNG construct.
- (3) Native PNGase F
- (4) Bio-Rad molecular weight markers



The construction of a new PNGase F expression vector was divided into four steps. The first step involved amplifying the gene from the pT7-PNG vector. Step two required digestion with appropriate restriction endonucleases and in step three, the digested fragment was ligated into the pKS-OmpA3-His vector to form the pOPH6 expression plasmid. Finally, the plasmid was transformed into the *E. coli* strain BL21 (DE3) for expression of the recombinant protein.

Figure 2.3 summarises the overall cloning strategy. The region coding for PNGase F was amplified from the pT7-PNG vector by the polymerase chain reaction (PCR), double digested with *Eco* RI and *Bam* HI, and ligated into pKS-OmpA3-His to produce the expression plasmid pOPH6. The ligated product was transformed into *E. coli* XL1-Blue host cells. Transformants carrying the pOPH6 plasmid were identified by digestion of isolated plasmid DNA with diagnostic restriction endonucleases and by DNA sequencing of both strands using an ABI-Prism 377 automated DNA sequencer. Finally, pOPH6 was transformed into the *E. coli* strain BL21(DE3) for protein expression.

### 2.5.1 Amplification of PNGase F Gene from pT7-PNG by PCR

Two restriction enzyme sites in the multiple cloning site of pKS-OmpA3-His (*Eco* RI and *Bam* HI ) were chosen to cut the vector for insertion of the PNGase gene as these sites are not present in the PNGase gene itself. Digestion by *Eco* RI and *Bam* HI, produces different cohesive ends that ensure the gene is ligated in the correct orientation within the plasmid. Two oligonucleotide primers were designed to introduce these sites into the gene by PCR using pT7-PNG as template and *Pwo* polymerase. *Pwo* polymerase was used instead of *Taq* polymerase for this experiment because of its superior fidelity due to its 3'→5' exonuclease proof-reading activity that is not present in *Taq*. The primer PNG1 introduced an *Eco* RI site to the 5' end of the coding sequence and the primer PNG2 replaced the original stop codon at the 3' end of the PNGase F gene with a *Bam* HI site. Both primers have 26 nucleotides that are complementary with the PNGase F gene (figure 2.4) with a calculated melting temperature of about 60°C and an expected annealing temperature of 50°C for the PCR reaction.



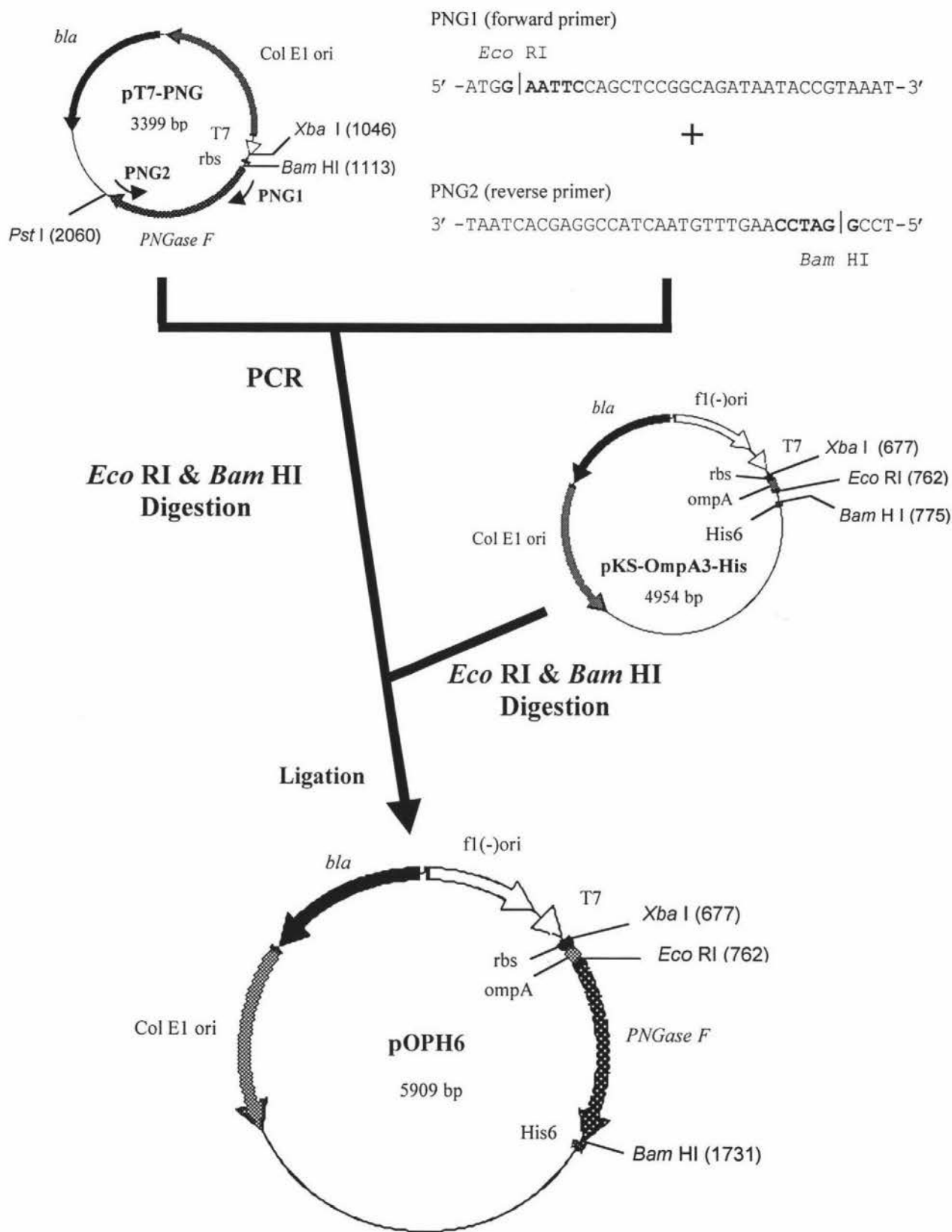
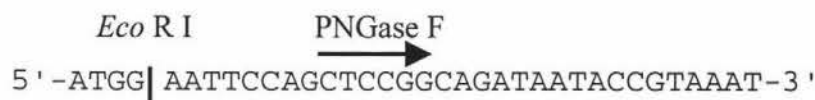


Figure 2.3 Cloning strategy for PNGase F.

**PNG1** (forward primer): A 35 mer with 26 nt complementary and a  $T_m$  of 58°C



**PNG2** (reverse primer): A 35 mer with 26 nt complementary and a  $T_m$  of 59°C

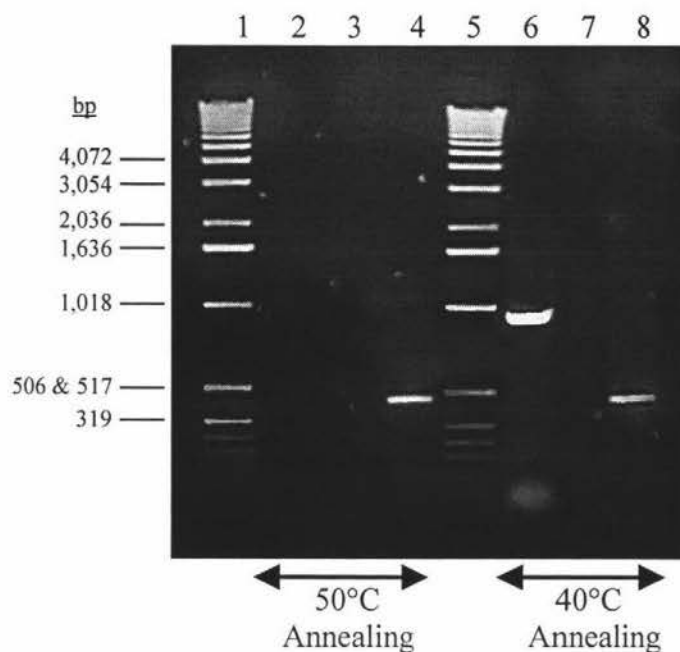


**Figure 2.4 Synthetic oligonucleotides used for PCR amplification.**

PNG1 added an *Eco* RI site to the 5' end of the PNGase F gene and PNG2 substituted the original stop codon at the 3' end of the PNGase F gene with a *Bam* HI site. The endonuclease cleavage sites introduced are shown and the  $T_m$  values are provided by the manufacturer.

The first attempt at amplification using *Taq* polymerase failed as shown in figure 2.5 (lane 2). The negative control without pT7-PNG as template in lane 3 showed no bands indicating there was no contamination in the reaction mixture that could result in a false PCR reaction. The primers OX1 and NO4, which were known to amplify a fragment of the PNGase F coding sequence to give a product of 503 base pairs (bp), were used as a positive control. In lane 4 where these primers were used, there was a product of about 500 bp suggesting all components of the mixture other than the new primers were functional. This result could have been due to non-optimum  $\text{Mg}^{2+}$  concentration or annealing temperature or both.

In a second attempt with *Taq* polymerase, the annealing temperature was lowered to 40°C while the other parameters and reaction mixture composition were unchanged. Agarose gel analysis of the PCR products showed a fragment of about 970 bp (lane 6), the expected length of the modified PNGase F gene. As the original PNGase gene was 942 bp long, the extra 11 base pairs added either side to create restriction enzyme sites should give a 964 bp fragment. A clean PCR product meant there was no need to further refine the reaction conditions. Using these parameters, the experiment analysed in lane 6 was repeated with high fidelity *Pwo* polymerase (results not shown).



**Figure 2.5** PNGase F amplified from pT7-PNG using PCR.

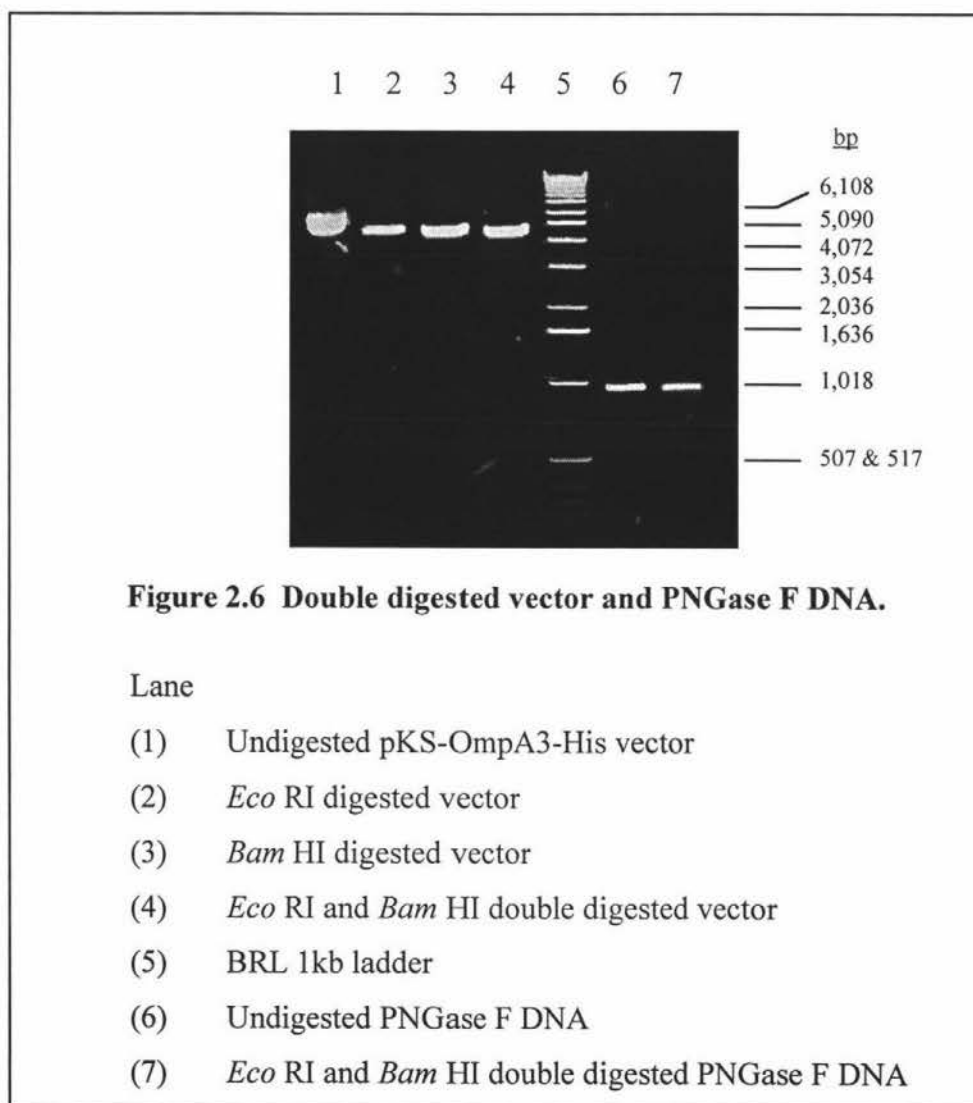
Lane

- (1) BRL 1kb ladder
- (2) No PCR product obtained at annealing temperature of 50°C
- (3) Negative control (pT7-PNG template absent)
- (4) Positive control (primers known to amplify PNGase F)
- (5) BRL 1kb ladder
- (6) Amplified PNGase F coding sequence obtained at annealing temperature of 40°C
- (7) Negative control (pT7-PNG template absent) on 2<sup>nd</sup> attempt
- (8) Positive control (primers known to amplify PNGase F) on 2<sup>nd</sup> attempt

## 2.5.2 Restriction Enzyme Digest of the PNGase F DNA

The amplified PNGase F DNA was purified from an agarose gel as described in section 2.4.8. PEG precipitation detailed in section 2.4.3 was used to precipitate the purified DNA fragment.

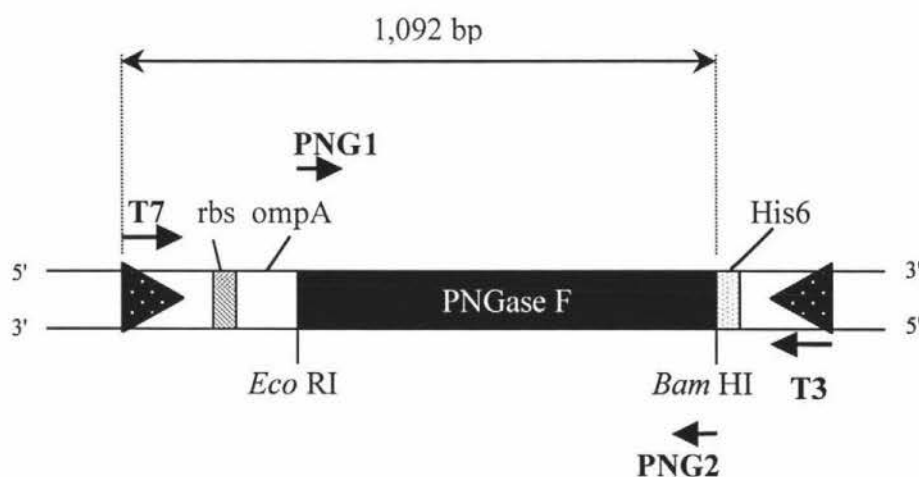
Digestion of the pKS-OmpA3-His vector by either *Eco* RI or *Bam* HI or both resulted in a linear fragment of ~5,000 bp as shown in figure 2.6 (lanes 2 to 4), indicating that the restriction enzymes were functional either individually or together. Lanes 6 and 7 shows the undigested PNGase F PCR product and the product digested by *Eco* RI and *Bam* HI respectively. As expected, agarose gel electrophoresis showed no difference in these fragments because these bands differ by only a few base pairs. Although it is impossible to verify the creation of cohesive ends using these restriction endonucleases, the double digest for the vector (lane 4) was assumed to be complete.



### 2.5.3 Ligation of the PNGase F gene into the Vector

The digestion products in lanes 4 and 7 were purified as described in section 2.4.8 and the vector was 5'-dephosphorylated with calf intestinal alkaline phosphatase to prevent subsequent self-ligation. The vector and the PNGase F DNA were then ligated together using three molar ratios of insert to vector (1:1, 1:3, 3:1) to maximise chances of a successful ligation.

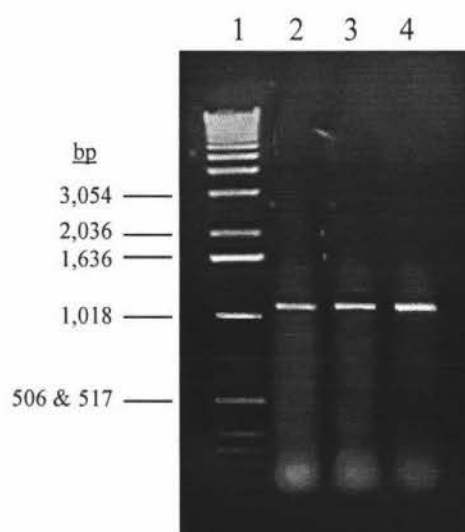
Authentication of a successful ligation could be obtained by PCR amplification with a small amount of the ligation mix as template using the T7 promoter sequencing primer as the forward primer and PNG2 primer as the reverse primer. The only PCR product obtained should be from an insert properly ligated to the vector because the T7 primer is complementary to the T7 promoter of the vector whereas PNG2 is complementary to the 3' end of the PNGase F gene (figure 2.7).



**Figure 2.7** Schematic diagram showing predicted PCR products using various primer pairs on a successfully ligated plasmid.

- T7 + PNG2: PNGase gene with N-terminal ligation junction.
- PNG1 + T3: PNGase gene with C-terminal ligation junction.
- T7 + T3: PNGase gene with both terminal ligation junctions and may also amplify a self-ligated plasmid with no insert.

Figure 2.8 shows a single PCR product the size expected for the PNGase gene with the upstream elements (1,092 bp) indicated in figure 2.7, showing that the ligation had been successful. It is important to note that only the 5' ligation junction of PNGase gene could be verified by this amplification. To verify the 3' end would require PNG1 as the forward primer and the T3 primer as the reverse primer. If the T7 and T3 primers were used in a PCR reaction both the PNGase gene from the vector plus the insert as well as the sequence due to a self-ligated plasmid would be amplified. The 3' junction was not confirmed by PCR due to a lack of T3 promoter sequencing primer. Instead, insertion of the PNGase gene into the vector was confirmed by *Eco* RI and *Bam* HI digests of the plasmid after it has been transformed into *E. coli*.



**Figure 2.8 PCR confirmation of successful ligation**

Lane

- (1) BRL 1kb ladder
- (2) Vector : Insert (3:1) ligation mixture as template
- (3) Vector : Insert (1:1) ligation mixture as template
- (4) Vector : Insert (1:3) ligation mixture as template

### 2.5.4 Transformation of Competent *E. coli* Cells with Ligation Products

Before transformation into *E. coli* BL21(DE3) for protein expression, the ligation mixture was transformed into *E. coli* XL-1 Blue cells to select the correct clones by analysing plasmid minipreps. This was necessary because of the small amount of ligated pOPH6 available and because of the much lower transformation efficiency of BL21(DE3) cells ( $\sim 10^{4.5}$  cfu  $\mu\text{g}^{-1}$  DNA) compared to XL1-Blue ( $\sim 10^{6.7}$  cfu  $\mu\text{g}^{-1}$ ). The *E. coli* XL-1 Blue strain is optimised for preparation of plasmid DNA with its abolished non-specific endonuclease I activity (EndA<sup>-</sup>). A drawback of using BL21(DE3) cells as the initial transformation strain is that it contains the DE3 lysogen that may lead to plasmid instability due to an over-expression of the T7 RNA polymerase (Studier *et al.*, 1990).

The procedure for transformation is described in section 2.4.11. Transformations into XL1-Blue for plasmid minipreps were performed using 2 and 6  $\mu\text{L}$  of ligation mix and the results obtained are shown in Table 2.1. The two positive controls for this transformation were: (i) the original uncut vector, and (ii) competent cells on LB plates without ampicillin to show the cells were viable. The four negative controls were: (i) The competent cells transformed with sterile Milli-Q water on LB-Amp plates, (ii) ligation mix without ligase on LB-Amp plates, (iii) cut vector with ligase but no insert grown on LB-Amp plates, and (iv) digested pKS-OmpA3-His with no ligase on LB-Amp plates. An unexpectedly large number of colonies developed from cells transformed with the digested vector with ligase, one of the one negative controls. This implied that incompletely digested vector was present due to an inefficient double digest by *Eco* RI and *Bam* HI. However, this could not be verified using agarose gel analysis (figure 2.6) as the difference between a single cleavage and double digested vector is too small to be seen on the gel ( $\sim 15$  bp).

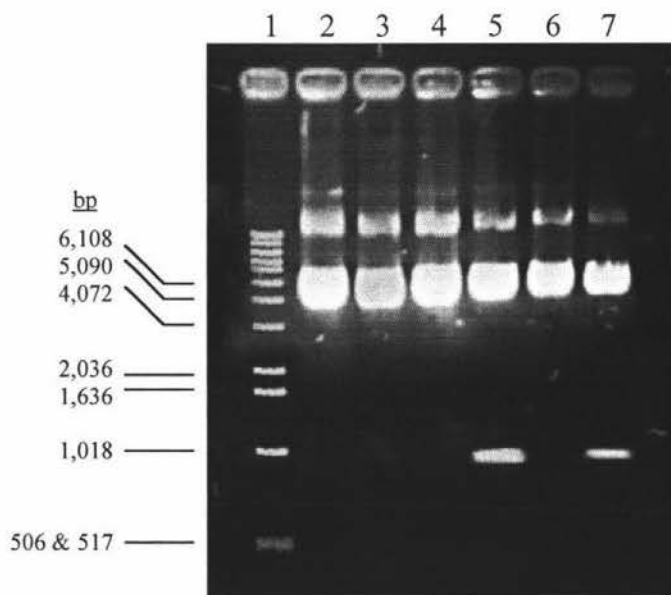


**Table 2.1 Colony count from transformation of pOPH6 ligation mix into *E. coli* XL1-Blue cells.**

Transformant	cfu
2 $\mu$ L vector : insert (3:1)	3
6 $\mu$ L above	28
2 $\mu$ L vector : insert (1:1)	14
6 $\mu$ L above	42
2 $\mu$ L vector : insert (1:3)	36
6 $\mu$ L above	65
+ve control (ii): Cells + sterile water on LB without Amp	>500
+ve control (i): pKS-OmpA3-His only	>300
-ve control (i): Cells + sterile water	0
-ve control (ii): 2 $\mu$ L vector : insert mix, no ligase	0
-ve control (iii): Digested pKS-OmpA3-His with ligase	56
-ve control (iv): Digested pKS-OmpA3-His, no ligase	1

The one blue colony that grew from cells transformed with cut and dephosphorylated vector may have resulted from contamination of air-borne particles or more likely, the presence of some incompletely digested plasmid. Twelve colonies containing the putative recombinant plasmid were selected from vector: insert (1:1) and plasmids were prepared from each using the method outlined in section 2.4.12. These plasmid minipreps were digested with *Eco* RI and *Bam* HI to verify that the PNGase DNA was present. Figure 2.9 shows the results of three of the twelve plasmids digests using these restriction enzymes. The linearised plasmid alone is expected to be about 4.9 kbp whereas the linearised ligated plasmid plus insert should be about 5.9 kbp and the insert, 942 bp in length. With an unusually large number of colonies carrying the singly digested vector ligated by T4 ligase (Table 2.1), the background resulting from these cells was expected to be high (two out of twelve minipreps carried this vector and were discarded, lanes 2 & 3). One miniprep that carried the insert was selected and the sequence around the PNGase DNA was determined using automated sequencing of both strands. This would detect any errors that may have been introduced by *Pwo* polymerase during PCR. The entire sequence of the gene was confirmed by automatic

and manual reading of the sequence. Once the sequence had been verified, the pOPH6 plasmid was transformed into BL21(DE3) cells, as described in section 2.4.11, for protein expression.



**Figure 2.9** Plasmid miniprep of XL-1 Blue transformed with pOPH6.

Lane

- (1) BRL 1kb ladder
- (2) Plasmid miniprep of colony 1
- (3) *Eco* RI and *Bam* HI double digest of colony 1
- (4) Plasmid miniprep of colony 2
- (5) *Eco* RI and *Bam* HI double digest of colony 2
- (6) Plasmid miniprep of colony 3
- (7) *Eco* RI and *Bam* HI double digest of colony 3

## 2.6 Discussion and Conclusions

In this chapter, the engineering of an expression vector for the high-yield production of PNGase F in *E. coli* is described: PCR primers were designed based on a published sequence (ATCC 33958; Tarentino *et al.*, 1990) in such a way as to introduce restriction endonuclease sites to the gene necessary for its insertion into the plasmid. The annealing temperature for the PCR amplification of the PNGase F gene from PT7-PNG vector was optimised. The product and the pKS-OmpA-His vector were double digested with both *Eco* RI and *Bam* HI, cleaned and ligated to produce the pOPH6 vector. The insert was sequenced and verified to be correct (appendix 2) before being transformed into *E. coli* BL21(DE3) cells for recombinant DNA protein expression.

The sequence of the PNGase F coding sequence from the *F. meningosepticum* (CDC strain 3352) differs from the previously published nucleotide sequence (ATCC 33958, Tarentino *et al.*, 1990) in 57 positions (appendix 3), resulting in eight amino acid substitutions (Table 2.2).

**Table 2.2 Amino acid variations between ATCC strain 33958 and CDC strain 3352.**

Amino acid substitution	Change in MW	Change in charge
T39A	-30	-
I149V	-14	-
G168A	+14	-
A219S	+16	-
I243N	+1	-
A245T	+30	-
T269I	+12	-
S281N	+27	-
Net = +56 Da		Net = 0

These substitutions resulted in a protein that has a predicted molecular weight of 34836 Da compared with 34780 Da for the PNGase F from ATCC 33958 as calculated by ExPASy ProtParam Tools ([www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html)). The eight predicted amino acid changes detected in between the two strains could be due to either (i) misincorporation errors by *Taq* polymerase during PCR amplification of the genomic template (CDC strain 3352), or (ii) genuine inter-strain variation. Three major factors can affect insertion error frequencies: Firstly, the type of mispair formed, for example G(primer)•T and T(primer)•G mispairs can generally be made more easily than C•C or G•G mispairs (Mendelman *et al.*, 1989). Secondly, the sequence context where the mispair occurs (Mendelman *et al.*, 1989), and thirdly, the identity of the polymerase as different polymerase appear to exhibit different mutational characteristics (Kunkel and Bebenek, 1988). For example, T4 DNA polymerase induces misincorporations of G•C to A•T, and A•T to G•C for *Taq* polymerase (Keohavong, P., and Thilly, W. G., 1989). When *Taq* polymerase, which lacks 3'→5' exonuclease proof-reading activity, is used to amplify DNA, the T→C and A→G mismatches, occur about once every  $10^3 - 10^4$  base pairs copied. Pur•Pur and Pyr•Pyr mismatches are likely to occur one to two orders of magnitude less frequently at about one per  $10^5$  to  $10^6$  bp.

The fidelity of a PCR reaction depends on a number of factors besides the template and the type of polymerase used. These include the buffer composition, especially dNTP and  $Mg^{2+}$  concentrations, thermal cycling parameters such as denaturation and extension times and cycle numbers (Innis *et al.*, 1995). The overall fidelity in a PCR reaction (the cumulative error frequency) depends not only on the misincorporation rate (errors per nucleotide polymerised) but also on the number of doublings. There should be no significant proportion of PCR products containing the same error in a PCR reaction unless misincorporations occur in the first few cycles which might lead to an ambiguity.

There are two mutations at positions 728 and 733 in the CDC strain 3352. The T728A mutation is not a due to the more common G•T or A•C mismatch and is thus unlikely to be a PCR error, whereas the G733A mutation is possibly a G•T mismatch by *Taq* polymerase. Figure 2.10 illustrates this argument.



There are eleven T→C and six A→G substitutions out of the fifty-seven nucleotide differences between the two sequences. Hence in the worst case scenario, these seventeen differences could be attributed to the mispairings typical of *Taq* while the other forty would seem to be due to inter-strain variation. In amino acid terms, only two (T39A and I149V) out of the eight amino acid differences between ATCC 33958 and the PCR product of CDC strain 3352 could have been caused by the common *Taq* mispairings. Both amino acid changes are due to two base changes within single codons, which makes these mutations unlikely to be due to *Taq* misincorporation. In theory, one way to verify these eight changes as genuine inter-strain differences is to obtain an exact mass of the CDC strain 3352 native PNGase F and compare with calculated mass from ATCC 33958 by Electrospray Mass Spectrometry (ES-MS) as described in the next chapter. Other alternatives include:

- (i) Direct sequencing of a PCR product amplified from the genomic DNA (CDC strain 3352) using high fidelity *Pfu* polymerase to minimise misincorporation to compare with the published sequence for the ATCC 33958 strain.
- (ii) Sequence two or more plasmids from independent transformants and look for changes that are common to all.
- (iii) Sequence of the native protein by Edman degradation.

Most of the amino acid differences observed between the two strains are conservative in nature and are on the surface of the protein. However, two differences occur near the active site of the enzyme (I243N and A245T), and both make the  $\Omega$  loop between the two domains of the protein more hydrophilic. Such changes could lead to changes in biological activity. Moreover, three potential N-glycosylation sites (at positions 217-219, 243-245 and 281-282) not seen in the ATCC strain 33958 were seen in the expression plasmid for the CDC strain 3352 prepared by PCR. One site in particular (positions 243-245) involves two amino acid substitutions and would seem unlikely to occur by chance. There is however no evidence that native PNGase F is glycosylated, although self-deglycosylation cannot be ruled out. Glycosylation of endoglycosidases is not uncommon: Endo F<sub>2</sub> and F<sub>3</sub> were reported to be O-glycosylated at serine residues (three sites in Endo F<sub>2</sub> and two sites in Endo F<sub>3</sub>; Plummer *et al.*, 1995). These amino acid changes in CDC strain 3352 PNGase F may affect enzymatic activity and/or substrate recognition, and this will be investigated in chapter 4.

## Chapter 3 Expression and Purification of PNGase F

### 3.1 Experimental Objectives

The goal of this work was to express and purify PNGase F from *Flavobacterium meningosepticum* and from transformed *Escherichia coli* to homogeneity for further studies. In order to compare the biochemical properties of the native and recombinant PNGase F, both enzymes were isolated at the same time. The purification protocol for native bacteria was adapted from Flaus (1989) who established the growth and purification of PNGase F from the medium of *F. meningosepticum* (CDC strain 3352). A similar purification protocol was also described by Tarentino *et al.*, (1993a) with strain ATCC 33958.

### 3.2 Materials

This inventory of chemicals and materials is in addition to those listed in section 2.3.

Amersham Pharmacia Biotech (Uppsala, Sweden):

PhastGel™ IEF (3-9).

Amicon Inc. (Beverly, USA):

10 kDa MWCO Centricons™ and Microcons™, stirred ultrafiltration cells (models 8010, 8050 and 8400), Diaflo™ YM-10 membranes and 0.22 µm Micropure™ separators.

Applied Biosystems (CA, USA):

ProBlot™ PVDF membranes and Trifluoro Acetic Acid (TFA).

BDH Laboratory Supplies (Poole, England):

Far UV grade acetonitrile and HPLC grade methanol.



Bio-Rad Laboratories (CA, USA):

Low range molecular weight protein standards for SDS-PAGE, SDS and TEMED.

Millipore Corporation (MA, USA):

Sterile Millex-GS 0.2  $\mu\text{m}$  syringe filters and Durapore<sup>®</sup> membrane filters (0.22  $\mu\text{m}$ , type: GV).

New Zealand Industrial Gases

Technical grade helium and nitrogen.

Roche (West Germany):

Complete<sup>™</sup> protease inhibitor cocktail tablets.

Sigma (St. Louis, USA):

IEF standards (3.6-9.3), chicken egg albumin (grade II).

Spectrum (CA, USA):

Molecular porous dialysis membranes (3 and 10 kDa MWCO).

Vivascience Limited (Lincoln, UK):

Vivaspin 500  $\mu\text{l}$  concentrators (10 kDa MWCO).

Chromatographic media used in this study are listed in Table 3.1.

**Table 3.1 Chromatographic supports used in purification.**

<b>Chromatographic Method</b>	<b>Trade Name and Manufacturer</b>	<b>Matrix Base</b>
Hydrophobic Interaction	Phenyl Sepharose 6 Fast-Flow (Low substitution) Amrad Pharmacia	Cross-linked agarose
	Phenyl Sepharose 6 Fast-Flow (High substitution; 1mL) Amrad Pharmacia	Cross-linked agarose
	TSK-butyl-Toyopearl 650M Tosoh Corporation	Vinyl polymer
	<i>t</i> -butyl HIC Econo-Pac (5 mL) Bio-Rad	Polyacrylic
	Alkyl Sepharose HR10/10 Amrad Pharmacia	Cross-linked agarose
	Butyl Sepharose 4 Fast-Flow Amrad Pharmacia	Cross-linked agarose
	In-house <i>t</i> -Butyl substituted TSK HW55(S) Fractogel base resin from Merck (Synthesis protocol described in Appendix 6)	Vinyl polymer
Ion Exchange	Mono-Q HR 5/5 Amrad Pharmacia	Hydrophilic polyether resin
	Uno-Q Bio-Rad	Continuous bed matrix (polyacrylamide)
	Uno-S (polishing column) Bio-Rad	Continuous bed matrix (polyacrylamide)
	Hydroxyapatite Bio-Rad	Ceramic beads
Size Exclusion	Superdex 75 HR 10/30; Hi-Load 16/60 (prep grade) Amrad Pharmacia	Cross-linked agarose and dextran
Immobilised Metal Affinity	HiTrap chelating (50mL) Amrad Pharmacia	Cross-linked agarose
	Poly-His purification Boehringer Mannheim	Cross-linked agarose
Reverse Phase HPLC	Jupiter C18 reverse-phase column (5 $\mu$ , 250 x 4.6 mm) with Security Guard column Phenomenex	Silica

### 3.3 Methods used for Expression and Analysis of PNGase F

#### 3.3.1 General Methods

##### 3.3.1.1 Determination of Protein Concentration

Four assay methods were tested to determine which was most suitable for estimating protein concentrations. They were the Bicinchoninic acid (BCA), Bradford, Lowry, and UV methods. The first three methods required the construction of a standard curve with bovine serum albumin to a maximum of 0.5 mg mL<sup>-1</sup>.

##### 3.3.1.1.1 Alkaline Copper (Lowry) Protein Assay

This method described by Lowry *et al.*, (1951) is based on both the Biuret reduction of Cu<sup>2+</sup> to Cu<sup>+</sup> by proteins under alkaline conditions and the reaction of Folin-Ciocalteu reagent with phenols. It results in the development of an intense blue colour that depends partly on the tyrosine and tryptophan content of the protein sample and as such is dependent on protein composition. It is normally used in the range 0.01-1.0 mg mL<sup>-1</sup> of protein.

Folin-Ciocalteu reagent contains 100 g sodium tungstate, 25 g sodium molybdate, 150 g lithium sulfate, 50 mL 85% phosphoric acid, 100 mL concentrated hydrochloric acid and a few drops of bromine made up to 1 litre with Milli-Q water. An appropriate volume of this reagent was diluted four fold immediately before use. Complex-forming reagent was also prepared immediately before use by mixing 100 volumes of 2% (w/v) sodium carbonate, 1 volume of 1% copper sulfate and 1 volume of 2% sodium potassium tartrate.

For each assay, 100 µL of sample or standard was mixed with 100 mL of 2M NaOH and boiled for 10 minutes before being cooled to room temperature. This was followed by the addition of 1 mL of complex-forming reagent and after 10 minutes of incubation at room temperature, 100 µL of Folin-Ciocalteu reagent was then added. The resulting solutions were mixed using a vortex mixer and incubated at room temperature for 30-60 minutes before their absorbances were measured at 550 nm.

### 3.3.1.1.2 Bicinchoninic Acid (BCA) Protein Assay

The BCA assay was first described by Smith *et al.*, (1985) and has the same dependency on the reduction of copper as the Lowry assay. The reduced copper then combines with BCA (1:2) to form a purple coloured complex. While both assays have similar sensitivity, BCA is generally more tolerant to compounds that interfere with the Lowry assay such as detergents and denaturing agents. The BCA assay has several advantages over the Lowry assay. Firstly, it requires that only one reagent be added to the sample. Secondly, there is less protein-to-protein variability, and thirdly, it is linear over a greater protein range from 0.2 to 50  $\mu\text{g mL}^{-1}$  (Stoscheck, C. M., 1990).

The standard reaction mixture was prepared immediately before use by mixing 50 mL of premixed BCA solution (Sigma) with 1 mL of freshly prepared 4% (w/v) copper sulfate solution. 100  $\mu\text{L}$  of sample or standard was added to 1.9 mL of reagent, incubated at 60°C for 30 minutes then cooled to room temperature before the absorbance at 562 nm was measured.

### 3.3.1.1.3 Coomassie Blue (Bradford) Protein Assay

The dye-binding method using Coomassie Blue G-250 was performed as described by Bradford (1976). When the blue dye is dissolved in a strong acid, it turns red-brown in colour as it protonated. The blue colour is restored when it binds to a positively charged protein because of a shift in the  $\text{pK}_a$  of the bound dye. The Bradford assay is simpler, faster, more sensitive and subject to less interference by common reagents and non-protein components of biological samples than the Lowry and BCA assays. However, Kruger (1994) noted that the dye binds most readily to arginyl residues of proteins (but not to the free amino acid) which can lead to variation in the response for different proteins. Another problem Kruger observed with the Bradford assay is that BSA displays an unusually large dye response, and when used as a standard may result in underestimating the protein content in a sample. In that paper, Kruger suggested using bovine  $\gamma$ -globulin as an alternative standard, as it gives results that are closest to the mean concentration of 16 proteins compared using the Bradford assay.

The Bradford reagent contained 100 mg of Coomassie Blue G-250 dissolved in 50 mL 95% ethanol, mixed with 100 mL 85% phosphoric acid, diluted to 1 litre with Milli-Q water and was filtered before use. 100  $\mu$ L samples were mixed with 1 mL of reagent and left to stand for 20-30 minutes at room temperature before being measured at 595 nm.

#### 3.3.1.1.4 UV method

As the extinction coefficient for PNGase F has been reported,  $\epsilon_{280}^{1\%} = 18$  (Tarentino *et al.*, 1990), this assay was used for determining the concentration of the more pure protein solutions. The extinction coefficient for the ATCC 33958 strain of the recombinant enzyme was determined by a protocol described in section 3.3.1.2, and this figure was then used to calculate PNGase F concentration using the following equation:

$$\text{PNGase F Concentration (mg mL}^{-1}\text{)} = A_{280} / 1.8 \times \text{dilution factor}$$

#### 3.3.1.2 Calculation of PNGase F Extinction Coefficient

The theoretical calculation of the extinction coefficient of PNGase F was determined using a method described by Gill and von Hippel (1989) which has an accuracy of  $\pm 5\%$ . The first step was to calculate the extinction coefficient of a 1% solution of the protein in 6 M guanidine hydrochloride (Gdn.HCl) denatured protein using the following equation:

$$\epsilon_{\text{M.Gdn.HCl}} = A \epsilon_{\text{M.Tyr}} + B \epsilon_{\text{M.Trp}} + C \epsilon_{\text{M.Cys}}$$

Where  $\epsilon_{\text{M.Tyr}}$ ,  $\epsilon_{\text{M.Trp}}$  and  $\epsilon_{\text{M.Cys}}$  are the molar extinction coefficients of tyrosine, tryptophan and cysteine residues at  $A_{280}$  with values of 1280, 5690, and 120  $\text{M}^{-1} \text{cm}^{-1}$  respectively. A, B, and C are the number of each type of residues per molecule of protein. The next step was to measure the absorbance at 280 nm of identical concentrations of native and Gdn.HCl-denatured solutions of PNGase F. From these data, the extinction coefficient of native protein was calculated using the equation:

$$\epsilon_{\text{M.native}} = (A_{280.\text{native}}) (\epsilon_{\text{M.Gdn.HCl}}) / (A_{280.\text{Gdn.HCl}})$$

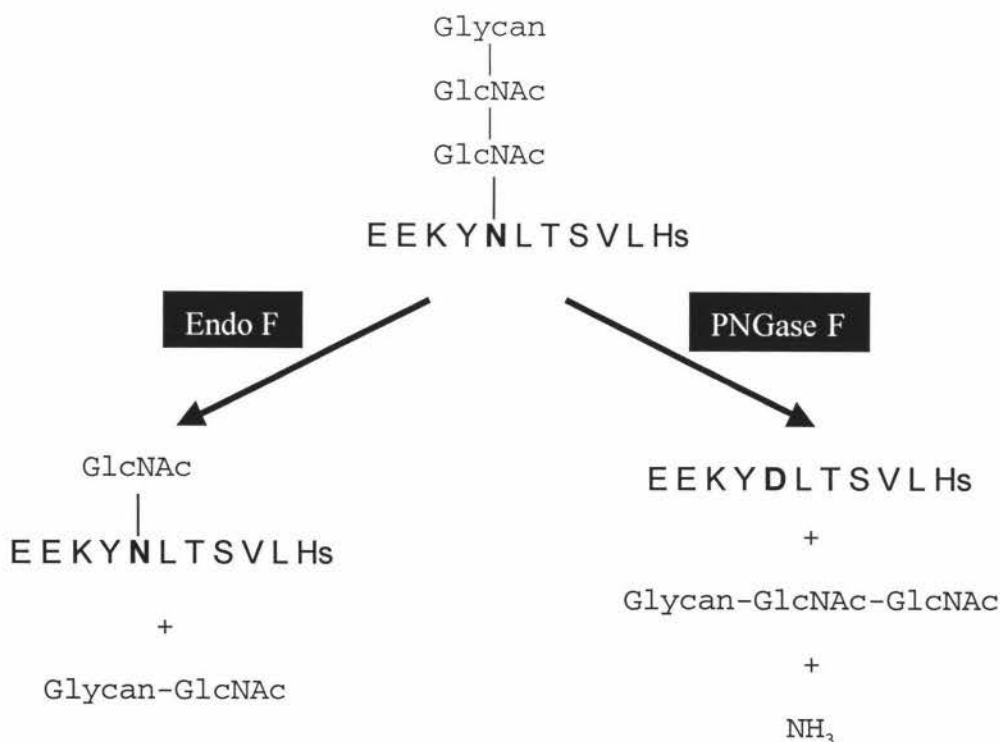
Where  $\epsilon_{\text{M.native}}$  is the molar extinction coefficient of the non-denatured PNGase F.

### 3.3.1.3 Determination of PNGase F Activity with a Reverse Phase High Performance Liquid Chromatography (RP-HPLC) Assay

A sensitive, discontinuous activity assay based on separation of the natural and deglycosylated forms of a glycopeptide using RP-HPLC was used to detect PNGase F activity at various stages of purification Norris et al., (1994a).

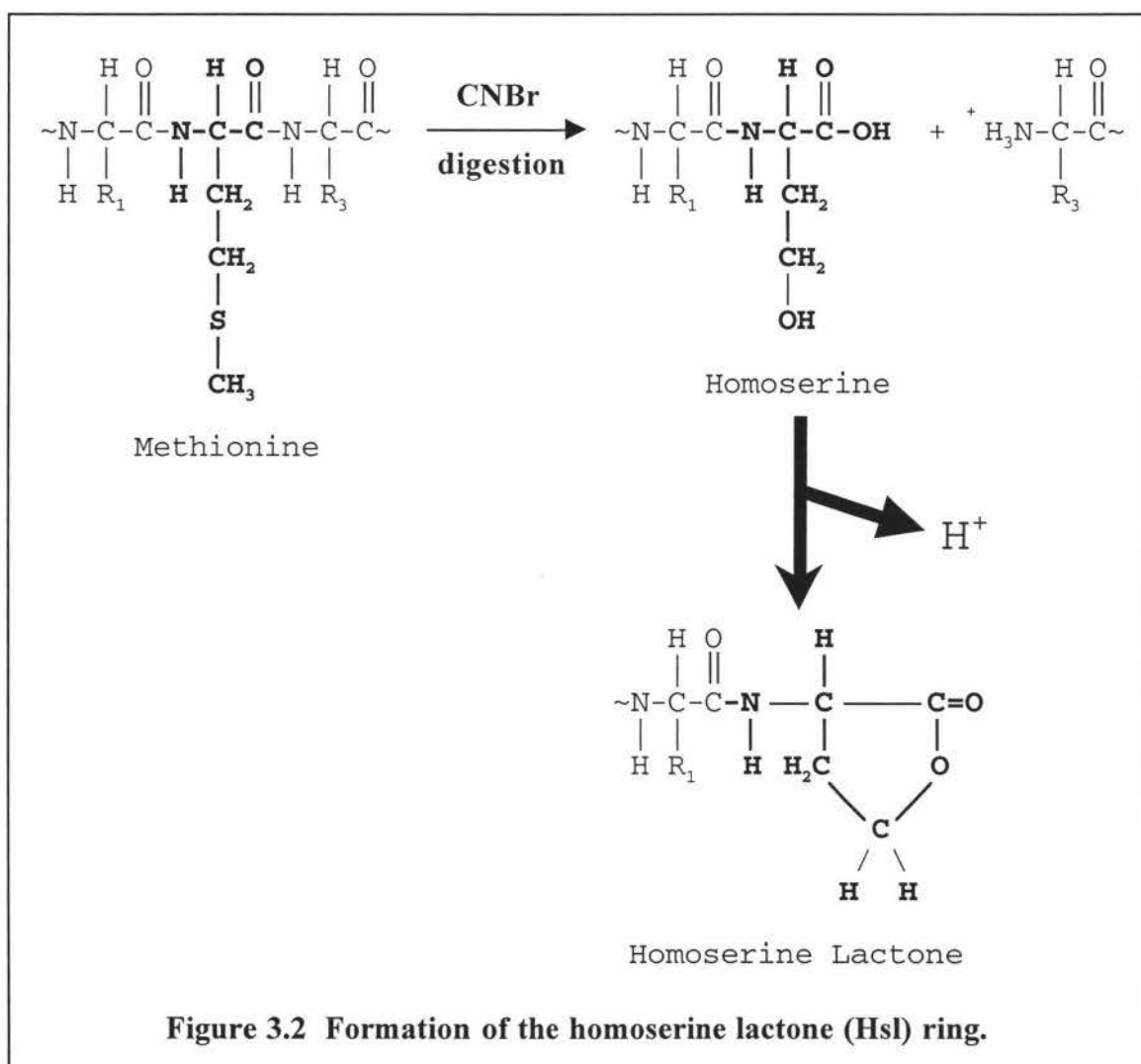
#### 3.3.1.3.1 Mechanism of Assay

The substrate used for the assay is a 11-mer glycopeptide containing a complex, biantennary oligosaccharide purified from the cyanogen bromide (CNBr) digestion of hen egg white ovalbumin outlined in Appendix 4. The bacterium *F. meningosepticum* secretes three endoglycosidases (Endo F1, Endo F2 and Endo F3) and one amidase (PNGase F). Endoglycosidases cleave the glycosidic bond in the dichitobiose glycan core whereas the amidase cleaves the  $\beta$ -Aspartylglucosylamine bond between asparagine and the sugar, converting the N-linked asparagine (N) to an aspartic acid (D), and releasing the intact glycan and ammonia (figure 3.1).



**Figure 3.1** PNGase F cleaves the amide bond in the ovalbumin glycopeptide.

As the peptide was prepared from CNBr treatment, the methionine (M) residue of the glycopeptide was converted to homoserine (Hs) during the cleavage. The calculated molecular weight for the peptide product was 1296 Da. However, because 10% formic acid was used to stop the reaction, the homoserine was converted to homoserine lactone (Hsl), resulting in a peptide with a molecular weight at 1278 Da (figure 3.2).



The assay is able to identify both the endoglycosidase and amidase activities because of the difference in hydrophobicity between a mono-glycosylated and a completely deglycosylated peptide. Therefore, in the order of hydrophobicity: Peptide > Peptide-GlcNAc > Peptide-oligosaccharide. The product peak of PNGase F being more hydrophobic elutes further away from the formic acid reference peak than the product peak of an endoglycosidase in an elution gradient of increasing solvent hydrophobicity.



### 3.3.1.3.2 Preparation of Samples

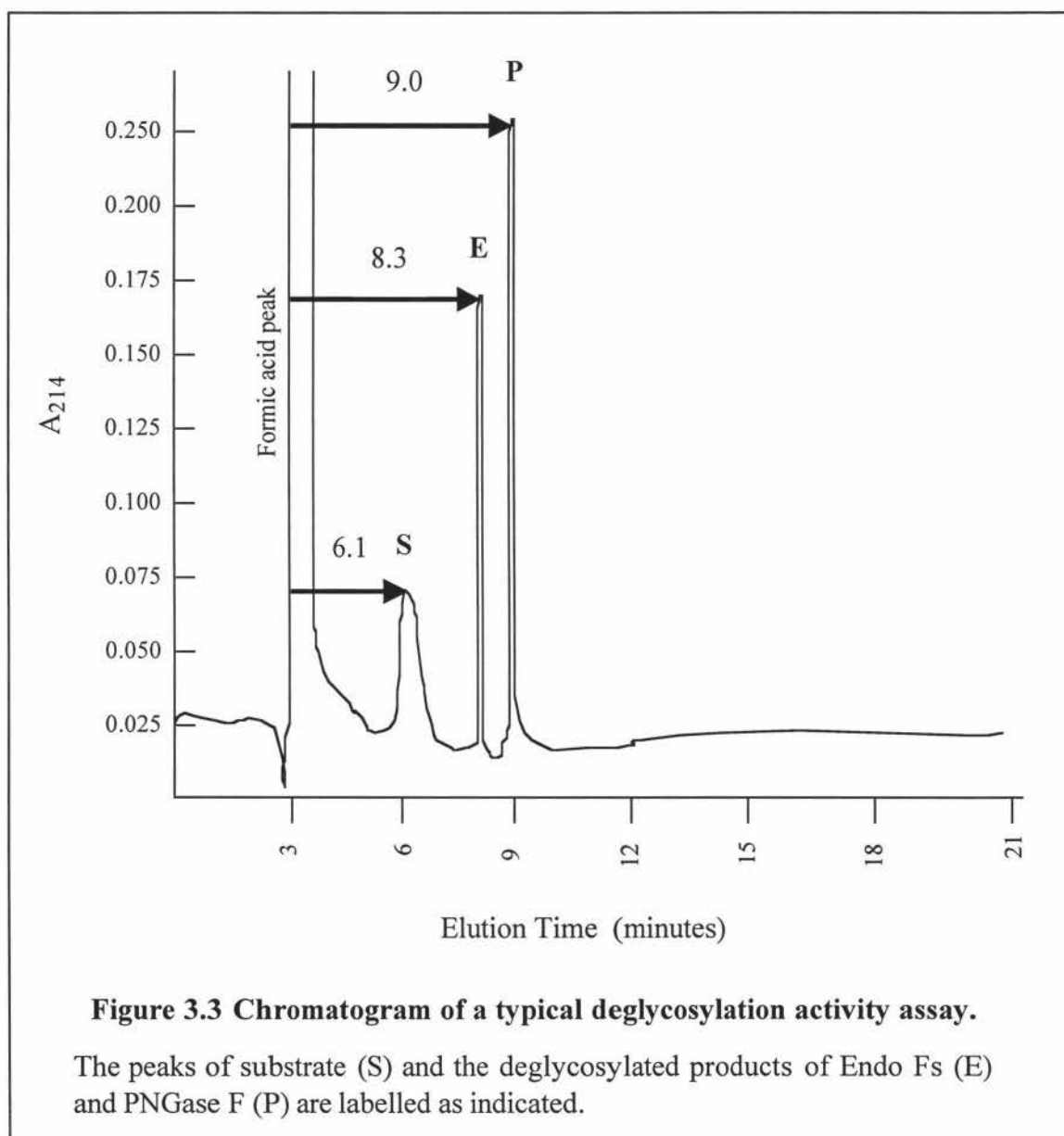
The ovalbumin glycopeptide substrate was prepared as a 1.0 mg mL<sup>-1</sup> solution in 5 mM EPPS buffer (pH 8.0) and Complete™ protease inhibitor (Roche) was added following the manufacturer's instructions. The substrate was dispensed as 50 µl aliquots into 500 µL Eppendorf tubes by weight (50 mg) to which 40 µl of distilled water (40 mg) was added and kept at -20°C before use. The enzyme was stored on ice and was diluted to 1 µg mL<sup>-1</sup> with 5 mM EPPS (pH 8.0) for use in the assay. Both the enzyme and substrate solutions were equilibrated at the incubation temperature of 37°C before the addition of the enzyme. Substrate and deglycosylated product peptide separations were performed on a Hewlett Packard Agilent 1100 series HPLC equipped with binary pumps, photodiode array detector, temperature controlled column oven and autosampler. Quantification of the separated peaks was performed by approximation of the peak areas by the accompanying software ChemStation® version A.07.01.

Generally, 5 µL of enzyme was incubated with the substrate for 5 minutes and the deglycosylation reaction terminated by addition of 5 µL of formic acid and rapid heating to 100°C for 5 minutes. The reaction mixtures were filtered through 0.22 µm Micropure™ separators at 14,000 g in a microfuge before being transferred to micro autosampler tubes (Sun catalogue # 500 900). Injections of 70 µL were injected into a C<sub>18</sub> reversed phase column (250 x 4.6 mm, Phenomenex) using a flow rate of 1 mL min<sup>-1</sup>. Substrate and deglycosylated product peptides were eluted in a 15-minute linear gradient outlined in Table 3.2 and the eluent was monitored at 214 nm.

**Table 3.2 PNGase activity assay gradient programme.**

Time (minutes)	%A (0.1% TFA in H <sub>2</sub> O)	%B (0.08% TFA in CH <sub>3</sub> CN)	Flow Rate (mL min <sup>-1</sup> )
0	80	20	1.0
15	60	40	1.0
20	80	20	1.0
25	80	20	1.0
500	80	20	0.1

Figure 3.3 shows the chromatogram of a typical assay. Using the formic acid as the reference point, the substrate elutes at 6.1 minutes in a peak that is broad because of the presence of different glycoforms of the glycopeptide. The deglycosylated peptide produced by Endo Fs retains a GlcNAc sugar is slightly more hydrophilic than the product of PNGase F having a retention time of 8.3 minutes. The deglycosylated peptide produced by the action of PNGase F is more hydrophobic and elutes at 9.0 minutes.



### 3.3.1.3.3 Determination of Specific Activity of PNGase F

The progress of the purification was monitored by measuring the specific activity of PNGase F throughout the purification. Protein concentrations of pooled fractions from each purification step were determined using the Bradford Assay. For measuring activities, fractions were diluted by an amount chosen to result in the hydrolysis of 5-15% of the substrate in 5 minutes at 37°C to ensure the rate measured falls in the linear portion of the rate curve. In some cases, the enzyme was serially diluted up to 200 fold with 5 mM EPPS (pH 8.0) and the assays were carried out as described in section 3.3.1.3.2. Each assay was performed in duplicate and if the peak area discrepancy between the duplicates was not within 5% of each other, the assays were repeated.

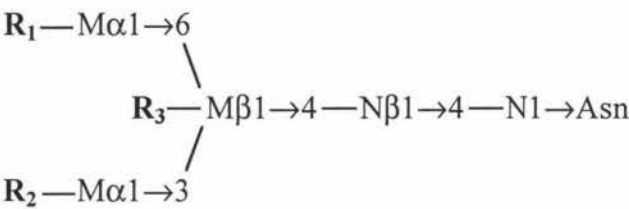
The substrate has nine uniformly distributed hybrid and high-mannose glycoforms all with different molecular weights between 2,351 and 3,284.8 Da (table 3.3) and an average molecular weight of approximately 2864.2 Da. There are approximately 17.5 nmol of substrate in a total reaction volume of 95 µL:

$$\frac{50 \mu\text{g}}{2864.2} = 17.5 \text{ nmol}$$

However, a more accurate measurement of the activity is to measure the appearance of the product of the reaction that has a molecular weight of 1278 Da, assuming the extent of substrate disappearance is equal to the appearance of the product. Activity of the reaction was measured in milliunits and was defined as 1 µmol of ovalbumin glycopeptide hydrolysed per minute at pH 8.0 and 37°C.

Table 3.3 The nine glycoforms of hen egg white ovalbumin.

Core glycan:



MW of core glycan = 892.82 Da

Variable groups:	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	MW
Group-I	Mα1→6 Mα1→3	Gβ1→4Nβ1→4 Nβ1→2	Nβ1→4	1096.02
Group-IIA	Mα1→3	Gβ1→4Nβ1→4	Nβ1→4 Nβ1→2	933.88
Group-IIB	Mα1→6 Mα1→3	Nβ1→4	Nβ1→4 Nβ1→2	933.88
Group-IIIA	Mα1→6 Mα1→3	Nβ1→2	Nβ1→4	730.68
Group-IIIB	Mα1→2 Mα1→3	Mα1→6	Mα1→2	648.56
Group-IIIC	Mα1→3	Nβ1→6 Nβ1→2	Nβ1→4	771.74
Group-IV	Mα1→6 Mα1→3		Mα1→2	486.42
Group-V	Mα1→6 Mα1→3			324.28
Group-VI	Mα1→3			162.14
Average =				676.4

Adapted from Sharon and Lis, 1982.

The average molecular weight of the glycopeptide substrate is therefore:

Peptide + glycan = 1295 + 892.82 + 676.4  
= 2864.2 Da.

#### 3.3.1.4 Polyacrylamide Gel Electrophoresis (PAGE)

Proteins were separated analytically by native and sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis using a Bio-Rad Mini Protean II Dual Slab Cell system and the method described by Laemmli *et al* (1970). Stock solutions, 12% resolving and 4% stacking gels (80 x 100 x 0.75 mm) were made as described in Bio-Rad Mini Protean II instruction manual. A routine procedure for SDS-PAGE is summarised below. 10-20  $\mu$ L protein samples were mixed with sample buffer and heated to 100°C in a boiling water bath for 2 minutes before being loaded onto a gel. A constant voltage of 200 volts (current  $\leq$  50 mA) was applied across electrodes until the dye front reached the bottom of the resolving gel. The gels were removed, fixed and stained in a solution of 0.1% (w/v) Coomassie Brilliant Blue R<sub>250</sub> (40% methanol / 10% acetic acid in water) for 30 minutes. The background was removed by destaining the gel with an identical solution without the dye. For native gels, the same procedure was followed except the samples were not heated, and the separating and running buffers did not contain either SDS or a reducing agent. In some cases, silver staining was needed to assess the purity of the protein sample and the procedure is described in Appendix 5.

Molecular weights of proteins were estimated by comparison with known standards. Low molecular weight protein markers (Bio-Rad) were routinely used in all gels. Stained gels were vacuum dried (80°C, 2 hours) using a Bio-Rad gel drier by sandwiching the gels between two sheets of cellulose acetate membranes or placing onto Whatman No.3 filter papers after soaking in a solution of 2% glycerol / 30% acetic acid for 1 hour. The inclusion of glycerol prevents gel shatter on drying.

#### 3.3.1.5 Electroblotting of Proteins from Acrylamide Gels

For the purpose of confirming protein identity with N-terminal sequence analysis, proteins separated by PAGE were electroblotted onto PVDF membranes. The 12% gels were left to polymerise overnight as free acrylamide may block the N-terminus of some proteins and prevent efficient N-terminal sequencing (Walker and Gasstra, 1987).

Samples were loaded (~10 µg of protein) onto a SDS-polyacrylamide gel in a symmetrical pattern, so that once the gel was run it could be cut in half down the middle to give two mirror images. Electrophoresis was carried out with 1 mM sodium thioglycolate (or 10 mM reduced glutathione) in the cathode running buffer. Thioglycolate was included in the cathode buffer to scavenge any free radicals that may induce acetylation at the N-terminus of proteins (Walker and Gasstra, 1987).

After electrophoresis, one half of the gel was stained with Coomassie Blue R<sub>250</sub> for reference while the other was soaked in analytical grade methanol for 10 seconds then immersed in transfer buffer (10 mM CAPS, 20% methanol, pH 11) for 10 minutes to equilibrate for electroblotting. The gel was assembled in an Bio-Rad electroblotter between sheets of Whatman No. 3 filter paper soaked in transfer buffer, sandwiched next to a sheet of PVDF membrane that had been wetted in methanol and then equilibrated in transfer buffer. Proteins were transferred at a constant current of ~100 mA for 60 to 90 minutes. When electroblotting was completed, the membrane was removed and soaked in water for 10 minutes with shaking before being stained briefly in Coomassie Blue to visualise the protein bands. The blotted gel was also Coomassie-stained to assess the degree of transfer in comparison with the reference gel. The membrane was rinsed in Milli-Q water for 5 minutes and air-dried thoroughly before the desired bands were cut out, wrapped in clean plastic and stored at -20°C for N-terminal sequencing.

#### 3.3.1.6 N-Terminal Sequencing

N-terminal sequencing was based on the Edman degradation method (Hewick *et al.*, 1981), followed by reverse phase HPLC separation of the phenylthiohydantoin (PTH) derivatives of amino acids performed on an automated Applied Biosystems model 476A protein sequencer.

### 3.3.1.7 Isoelectric Focusing (IEF) of Native and Recombinant PNGase F

Isoelectric focusing (electrofocusing) involves migration of proteins in a pH gradient on a polyacrylamide gel by an electric current to a point in the system where the pH equals their isoelectric points (pI). The pH gradient is established with polymeric buffer compounds known as ampholytes that have large numbers of both positive and negative charges and possess isoelectric points that define the pH range across a gel slab or column. Application of an electric field drives both the ampholytes and proteins to their isoelectric positions. The protein being larger will be retarded more so that the pH gradient is established before the proteins reached their pI. Diffusion and zone mixing are countered in IEF (hence the so called “focusing” effect) because as a protein molecule diffuses away from its isoelectric zone, it becomes charged and migrates back again.

The isoelectric points of both forms of PNGase F were determined using the PhastSystem (Amrad Pharmacia) with Phast IEF gel (pH 3-9). The separated protein bands were visualised with the silver staining technique and the pI determined by comparing its position with that of standard proteins (IEF standards, Sigma). The samples and staining reagents were prepared according to the PhastSystem user guide.

### 3.3.1.8 Electrospray Ionisation Mass Spectrometry (ES-MS)

Atmospheric pressure ionisation (API) is a gentle ionisation technique that produces ionised forms of fragile molecules from a dilute solution by neubilisation in the presence of a strong electric field. A nebuliser gas carries the sample solution from the capillary inside the ion spray source at high voltage and disperses the sample into an aerosol of highly charged droplets. These droplets, which carry excess positive or negative charges depending on the polarity of the electric field, evaporate rapidly in the presence of a curtain gas and/or heat until a coulombic explosion releases multiple charged ions into the gas phase. The ions formed are attracted into the mass analyser (usually a quadrupole) under high vacuum where they are separated and analysed according to their mass to charge ratio ( $m/z$ ). For a typical protein, extensive multiple charging results in a family of ions with different charged states or  $m/z$  values.



ES-MS is extremely useful in the analysis of large, labile biomolecules up to 200 kDa with a precision of better than 0.01% every 10 kDa (Fenn *et al.*, 1989). It is possible to identify post-translational modifications such as phosphorylation and glycosylation and to establish or confirm the presence of variants resulting from amino acid substitution or deletion by mass changes in proteins or peptides within the range of 100 to 100 kDa. Samples are usually made to 50% acetonitrile / 49% water / 1% formic acid.

### 3.3.1.9 Cell Growth and Protein Production Studies

The growth and PNGase F production from *F. meningosepticum* was carried out using reported methods (Norris *et al.*, 1994a). The growth and protein production of *E. coli* BL21(DE3) cells transformed with pOPH6 was monitored by removing 5 mL aliquots from a 200 mL culture every hour beginning from 0 to 8 hours after induction. Growth was observed by measuring cell density at 600 nm and pH of the cell culture whereas PNGase F production was visualised by SDS-PAGE of the whole cell extract of induced cells. Since the cell density is higher at later stages of induction, a decreasing volume of cells were loaded on SDS-PAGE to give a consistent picture of enzyme production in the 8 hour period. The volume of cells used was corrected by the optical density of the culture at 600 nm and was calculated by the following formula:

$$\text{Volume of culture to spun down} = \frac{0.5 \text{ ml culture}}{\text{OD}_{600}}$$

## 3.3.2 Expression of PNGase F

### 3.3.2.1 Growth and Induction of *F. meningosepticum* Culture

M9 agar-plates streaked with either frozen stocks of *F. meningosepticum* or sterile Milli-Q water as a negative control were incubated at 25°C for 12-16 hours. Two single colonies from these plates were inoculated in 50 mL M9 media at 25°C for 12-16 hours with shaking (200 rpm). One of these cultures was transferred to 1 litre of M9 media (1:20 dilution) and was incubated as before for a further 12-16 hours. All of this culture was used to inoculate the main 10-litre culture (1:10 dilution) which was divided equally in five 5-litre conical flasks and grown for another 30 hours.

PNGase F is produced when the bacteria enters stationary phase (~30 hours) and is under stress due to a lack of nutrient and accumulation of metabolites. The yield of the enzyme was increased substantially when the culture was induced with 0.1% (w/v) bacto-yeast extract during this stage and incubated for a further 16 hours. 10mM EDTA (pH 8.0) was added immediately to the culture to halt metalloprotease activity before the cells were spun down at 13,800 g for 20 minutes at 4°C. The cell pellet was discarded and the clear medium retained for purification.

### 3.3.2.2 Growth and Induction of *E. coli* Culture

New transformations of the pOPH6 plasmid into *E. coli* BL21(DE3) were plated onto LB-Amp agar plates and incubated at 37°C for 12-16 hours. Three single colonies were inoculated in 5 mL LB at 37°C for 5 hours with shaking (200 rpm). One mL from one of these cultures was then transferred to 50 mL of LB broth (1:50 dilution). This culture was incubated for a further 3 hours and was used to seed the principal 5-litre culture (1:100 dilution). Ampicillin (100 mg mL<sup>-1</sup>) was present at all times to maintain the plasmid and the media was only a fifth of the volume of the conical flasks used to allow for adequate aeration of the culture.

The cultures were induced with 1 mM IPTG when optical density at 600 nm reached 0.5-0.6. Fresh ampicillin was added at this stage to replenish depleted antibiotic destroyed by the 31.5 KDa  $\beta$ -lactamase in order to maintain selection pressure. Growth was continued at 37°C for 3-4 hours after which the cells were harvested for purification by centrifugation at 3,000 g for 15 minutes at 4°C.

#### 3.3.2.2.1 Preparation of Periplasmic Fraction from *E. coli*

As the recombinant enzyme is exported into the periplasm, the number of contaminating proteins is considerably less (~100 proteins versus ~2,000 in cytoplasm; Janson *et al.*, 1996). Secreted PNGase F was isolated from the periplasm by gentle osmosis using a variation on a published protocol (Witholt *et al.*, 1976).

Cells were pelleted by centrifugation at 3,000 g for 20 minutes, then resuspended in 5% culture volume of ice-cold high osmotic strength buffer (100 mM Tris-HCl, 500 mM sucrose, 1 mM EDTA, pH 8.0). The cells were again pelleted by centrifugation before being resuspended in ice-cold water for 10 minutes.  $\text{MgCl}_2$  (1 mM final concentration) was added to the suspension which was left at 0°C for another 10 minutes before being centrifuged at 13,000 g. Supernatant carrying the periplasmic fraction was retained for purification.

### 3.3.3 Purification of PNGase F

The purification of native PNGase F from the medium of *F. meningosepticum* exploits the hydrophobicity of the protein that is mainly due to a high content of hydrophobic and aromatic residues near the active site of the protein (Norris *et al.*, 1994b). The purification scheme used in this study was similar to that used by Norris *et al.* (1994a) except as noted in the following sections. Recombinant PNGase F was expressed with a poly-histidine tag that facilitates a simple purification procedure based on immobilised metal ion affinity chromatography (IMAC).

### 3.3.4 Chromatographic Methods

Chromatography is the separation of solutes according to their different distributions between two phases, usually a solid matrix as the stationary phase and a liquid as the mobile phase. Separation is achieved through the partitioning of the analyte between these two phases. Proteins adsorb to a variety of solid phases in a selective manner and using a series of different chromatographic matrices, a protein can be purified from a complex biological mixture. The chromatographic steps performed in this study used both low and medium pressure chromatography systems. Low pressure chromatography generally relies on gravity feed or a peristaltic pump and uses resins with a larger particle size (20 – 40  $\mu\text{m}$ ) compared to those used in medium pressure systems (10 – 20  $\mu\text{m}$ ).

Smaller particle sizes result in much better resolution than larger particle sizes because the plate number (the number of uniform zones in a column able to accommodate the solutes) is proportional to the particle size. Thus, using smaller particle size increases the area of contact between the stationary and mobile phases, resulting in adsorption-desorption processes occurring in narrower bands. However, the pressure needed to obtain the same flow rate increases with the inverse square of the particle size. As the particle size becomes smaller the flow rate of the mobile phase through the column decreases necessitating the use of specialised pumping systems to achieve separations in reasonable times. Initial separations were carried out at low pressure using a Bio-Rad Econo System and subsequent steps were carried out using a Waters 650 Fast Performance Liquid Chromatography (FPLC) system with a Rheodyne injector. Samples for FPLC were filtered through 0.22  $\mu\text{m}$  syringe filters or Micropure™ separators to remove all particulate matter before injection onto the column and buffer solutions were thoroughly degassed and filtered (0.22  $\mu\text{m}$ ) before use. The purification of the native enzyme is described first followed by that of the recombinant protein. All steps except those noted were carried out at 4°C.

#### 3.3.4.1 Hydrophobic Interaction Chromatography (HIC)

HIC separates proteins by exploiting the variability of exposed hydrophobic amino acid residues on different proteins. It relies on the interaction between the hydrophobic pockets or patches on the surface of the protein with the aliphatic or aromatic ligands of the resin matrix. This interaction can be controlled through appropriate choice of both the ligand and the mobile phase. Van der Waals interactions between non-polar amino acid side chains play a role in both protein folding (Privalov, 1992) and HIC (Ståhlberg *et al.*, 1992). In HIC chromatography, the ligands are sparsely distributed on the surface so that proteins will not experience steric interactions with one another. The spacer arms used to link the functional groups to the matrix are longer and it is possible for the solute to reach the surface of the gel. The charged surface of the protein thus rests against the polar surface of the gel while the hydrophobic chains are able reach into hydrophobic grooves in the protein (Scopes, 1994).

Hydrophobic matrices usually contain alkyl, aryl or phenyl functional groups covalently linked to an inert matrix such as agarose or sepharose, and hydrophobic interactions are affected by several important factors:

1. The type of ligand and the degree of substitution are crucial as more hydrophobic proteins will bind to shorter alkyl chains while less hydrophobic ones preferentially bind to longer alkyl chains and phenyl groups.
2. Hydrophobic interactions are promoted by the presence of high concentrations of salts in the mobile phase. The concentration and type of salt modify these interactions where they are strengthened by high concentrations of lyotropic salts such as  $(\text{NH}_4)_2\text{SO}_4$  and weakened by high concentrations of chaotropic salts such as NaCl. Hence, HIC samples are usually loaded in high salt solutions and eluted in decreasing salt gradients.
3. Because hydrophobic interactions have a large entropy component, they are greatly affected by temperature. Hence the strength of the interactions is increased with increasing temperature.
4. Hydrophobic interactions are unsurprisingly weakened in the presence of detergents, or organic solvents (Scopes, 1994).

#### 3.3.4.1.1 Hydrophobic Interaction Chromatography (HIC) with Phenyl Sepharose 6 Fast-Flow (Low Substitution)

In this study, the clear medium from *F. meningosepticum* was made 1 M in ammonium sulfate and the pH adjusted to 5.5. It was then loaded directly onto a column (4.6 cm x 40 cm; Pharmacia) packed with Phenyl Sepharose 6 Fast-Flow (Low Substitution, Pharmacia) at room temperature equilibrated in 50 mM NaAc, 1 M  $(\text{NH}_4)_2\text{SO}_4$  (pH 5.5). After loading was completed, the column was washed extensively with loading buffer to remove any unbound proteins. The bound proteins were eluted in a decreasing gradient from 1.0 M to 0.0 M  $(\text{NH}_4)_2\text{SO}_4$  over 8 hours at  $2.5 \text{ mL min}^{-1}$ . Eluent was monitored at 280 nm, collected in 10 mL fractions, assayed for activity and by SDS-PAGE. Active fractions were pooled and concentrated to 1% original volume (~10 mL)

by ultrafiltration using a stirred cell concentrator (Amicon 8400 with an Amicon YM-10 membrane) both of which have been pre-washed with 0.1% Thesit to reduce protein loss via surface adsorption. The concentrate was then subjected to HIC on tertiary-butyl (*t*-butyl) substituted TSK HW55(S) Fractogel™ using Fast Performance Liquid Chromatography (Waters model 650 FPLC).

#### 3.3.4.1.2 HIC with *t*-butyl-TSK

Previous work had shown that HIC chromatography using an in-house *t*-butyl TSK resin was the most effective step in the purification of native PNGase F, giving 99+% purity (Norris *et al.*, 1994a). Using standard cleaning procedures, the resin was regenerated and packed into a FPLC column (10 mm x 10 cm; Pharmacia) before being equilibrated in 50 mM NaAc, 5 mM EDTA, 3M NaCl (pH 5.5) at 4°C. Two mL of stirred cell concentrate was made 3 M in NaCl before being loaded. The column was washed with the equilibration buffer until the absorbance at 280 nm dropped to less than 0.05. The bound protein was then eluted by reducing the salt concentration to zero in a linear gradient over 50 minutes using a flow rate of 1 mL min<sup>-1</sup>. Fractions of 1.5 mL were collected, analysed by SDS-PAGE and assayed for activity. Unfortunately, the column had lost resolution and the protein was still contaminated with other proteins. Active fractions were pooled and concentrated to 1% original volume (~2 mL) with a stirred ultrafiltration cell and membrane as before and used to test the effectiveness of a number of other hydrophobic matrices to purify the protein.

#### 3.3.4.1.3 HIC Tests with Various Hydrophobic Matrices

##### A. Pharmacia HiTrap™ Test Kit

The test kit was used to test a number of different ligands for their efficiency in separating PNGase F from contaminating proteins. The 1mL pre-packed HIC cartridges were connected to a Bio-Rad Econo System and equilibrated in 50 mM NaAc, 5 mM EDTA, 3 M NaCl (pH 5.5) at a flow rate of 0.5 mL min<sup>-1</sup> at 4°C. The sample was equilibrated in the same buffer before being loaded onto the column. Unbound proteins were removed by washing with loading buffer until the absorbance at 280 nm returned to



10% of full deflection. Proteins were eluted from the column by applying a gradient of 3.0 to 0.0 M NaCl over 50 minutes. 1-mL fractions were collected and assayed for activity. Cartridges used were Phenyl Sepharose 6 Fast-Flow (High Substitution) and Butyl Sepharose 4 Fast-Flow.

#### B. Macro-Prep *t*-butyl HIC Econo-Pac Cartridges

The 5 mL pre-packed column from Bio-Rad was again attached to the Econo system using a similar buffering system and gradient as above except the salt concentration was raised to 4 M NaCl in order to ensure binding of PNGase F.

#### C. TSK-butyl-Toyopearl 650M

This resin was used extensively by Tarentino *et al.*, (1993a) in the purification of PNGase F to a purity of 90+%. Fines were removed according to the manufacturer's instructions and the resin was packed into a FPLC column (10 mm x 10 cm, Pharmacia). The column was run using the same conditions as described in section 3.3.4.1.2.

#### D. Alkyl Sepharose HR 10/10

A pre-packed column (10 mm x 10 cm; Pharmacia) was connected to a FPLC and equilibrated using the same buffer (NaCl as the salt) as described in section 3.3.4.1.2. As the protein failed to bind, the column was re-equilibrated in 50 mM NaAc, 5 mM EDTA, 2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (pH 5.5). Ammonium sulfate is a more lyotropic salt than NaCl and therefore should promote binding of proteins to the hydrophobic matrix. Using this buffer, difficulty was experienced in removing the protein from the column, so it was washed with 50 mM TAPS, 5 mM EDTA, 50% (v/v) ethylene glycol (pH 8.5) after the salt concentration was reduced to zero to elute the proteins.



### 3.3.4.2 Ion Exchange Chromatography (IEX)

The isoelectric point of the native PNGase F was determined using the PhastSystem (Amrad Pharmacia) as described in section 3.3.1.7. Using this information IEX can be exploited to purify PNGase F. IEX separates proteins and molecules based on their different net charges at a given pH. The side chains of surface amino acids may be protonated or deprotonated depending on the pH of the environment, so that their surface distribution and overall net charge dictate the unique behaviour of a protein in an ion exchange environment. The functional groups on the matrix carry either positive or negative charges that can interact with proteins primarily through ion pairing. These interactions depend not only on the net charge of the protein but also on the ionic strength and pH of the buffer ions, the nature of these ions, and properties of the functional ligands. Stationary phase matrices carrying charged groups that retain their charge over a narrow pH range are regarded as weak ion exchangers whereas those stable over a wide pH range are considered strong exchangers. The charged groups may either be cationic and carry negatively charged groups such as the weak carboxymethyl (CM) and the strong sulfopropyl (SP) groups, or anionic with positively charged groups such as diethylaminoethyl (DEAE) and the strong hydroxypropyl diethylaminoethyl quaternary amine (Q) groups. The charges on these columns are balanced by counter ions such as  $\text{Cl}^-$  in the case of anion exchange and  $\text{Na}^+$  for cation exchange. The net charge on the protein is the same as those of the counter ions and thus the protein displaces the counter ions to bind to the ligand. Proteins carrying no charge, an opposite charge or the same net charge as the column pass through the column in the mobile phase.

Proper degassing of buffers is an important aspect for successful IEX as atmospheric carbon dioxide dissolves easily in buffers at alkaline pH changing the pH. Proteins that bind to IEX columns are generally eluted using three methods. The most commonly used method is to increase the concentration of the counter ion in the mobile phase, thereby weakening the electrostatic interaction between the protein and ligand by competition with an excess concentration of counter ions.

The second method is to alter the pH of the buffer so that the overall charge on the protein changes to become the same as the charge on the resin. Thus the pH is decreased for anion exchange and increased for cation exchange. This method is less commonly used however, because of difficulties in bringing about sharp changes in pH with a gradient, which results in very poor separation of proteins that have similar properties. The third method is displacement of the bound protein with a solute that binds more strongly to the functional groups on the resin than the protein.

#### 3.3.4.2.1 IEX Test with various Ion Exchangers

##### A. *Mono-Q HR5/5 & Uno-Q*

Two different buffers (Tris-HCl at pH 8.9 and CAPSO at pH 9.8), one below and the other above the pI of PNGase F were used with both Mono-Q (Amrad Pharmacia) and Uno-Q (Bio-Rad) in hopes of purifying PNGase F to homogeneity. While these resins have the same functional group, their matrices are very different which may affect the separation because of non-specific interactions with the matrix itself (Tarentino *et al.*, personal communication). The sample was dialysed extensively in ten times the sample volume of loading buffer with several buffer changes to ensure that it was at an identical ionic strength and pH to the column. Either one of the pre-packed columns was connected to an FPLC (Waters model 650) equilibrated in either 20 mM Tris-HCl, 5 mM EDTA (pH 8.9) or 20 mM CAPSO, 5 mM EDTA (pH 9.8) at 4°C. The sample was loaded at 0.2 mL min<sup>-1</sup> to allow time for interactions to occur and bound proteins were eluted in a gradient of 0.0 to 1.0 M NaCl over 50 minutes at a flow rate of 1 mL min<sup>-1</sup>. Fractions of 1 mL (50 drops) were collected, assayed for activity and analysed by SDS-PAGE.

##### B. *Uno-S polishing column*

In order to produce an overall net negative charge on the protein, a buffer with a pH below the pI of PNGase F was utilised. The sample was dialysed in 20 mM MOPS, 5 mM EDTA (pH 7.2). The column was equilibrated with the same buffer and eluted using the same gradient program as described above.

### C. *Hydroxyapatite*

A crystalline form of calcium hydroxyphosphate known as hydroxyapatite, is basically a cross-linked series of  $\text{Ca}^{2+}$  and  $\text{PO}_4^{2-}$  ions, coupled onto porous ceramic beads (Bio-Rad). Owing to the presence of the phosphate oxygens, the surface of hydroxyapatite is largely negative and the positive groups on proteins can interact through ion pairing. On the other hand, some  $\text{Ca}^{2+}$  sites near the surface of the gel can attract carboxylate groups of proteins. Therefore, the gel can bind both acidic and basic proteins and the type of interactions can be enhanced by the choice of the equilibrium buffer. For example, if the mobile phase is phosphate buffer, the  $\text{Ca}^{2+}$  sites tend to be covered and more phosphate sites are exposed. When the mobile phase is  $\text{CaCl}_2$ , there are more cationic sites and fewer partial negative sites. Acidic proteins are adsorbed by the  $\text{Ca}^{2+}$  sites on the gel surfaces, and anions such as  $\text{Cl}^-$  are not effective in displacing them as they have little affinity for  $\text{Ca}^{2+}$ . Basic proteins binding more through  $\text{PO}_4^{2-}$  sites, can be displaced by monovalent cations reasonably effectively, and very effectively by  $\text{Ca}^{2+}$  ions which have high affinity for the  $\text{PO}_4^{2-}$  sites (Bernardi, 1971).

The dry resin was hydrated according to the manufacturer's instructions, packed into a HR 5/5 FPLC column (5 mm x 5 cm; Pharmacia) and equilibrated in 10 mM  $\text{K}_2\text{PO}_4$  buffer (pH 6.8). The sample was equilibrated in this same phosphate buffer before being loaded onto the column. A gradient of 10 to 400 mM  $\text{K}_2\text{PO}_4$  (pH 6.8) was applied over 50 minutes at 1 mL min<sup>-1</sup> to elute bound proteins. Fractions of 1 mL were collected, assayed for activity and analysed by SDS-PAGE.

#### 3.3.4.3 Size Exclusion Chromatography (SEC) with Superdex-75

SEC, gel filtration or gel permeation separates proteins or peptides on basis of their size. The matrix used is commonly a hydrated agarose or dextran gel and consists of an open, cross-linked, three-dimensional molecular network with pores between 40-120  $\mu\text{m}$  in diameter (Scopes, 1994). The pores within the beads act like a molecular sieve that completely exclude large proteins while proteins small enough to penetrate the pores are retained to differing extents depending on their size. Consequently, proteins that are

completely excluded from the pores are eluted in the void volume ( $V_0$ ) while other molecules are eluted between the  $V_0$  and the total column volume ( $V_t$ ) in order of decreasing molecular weight. To improve resolution in SEC, the following conditions are required:

1. Non-specific interactions between the matrix and the proteins. Polar interactions can be minimised by the addition of a small concentration of salt, for example, 0.15 M NaCl to the buffer.
2. The sample size should be within 0.5% of the bed volume of the column and smaller than the separation volume between peaks.
3. The distribution of pore sizes should be as even as possible.
4. The bead size where the smaller the bead size, the greater the resolution.
5. Column dimensions where a long column with a small diameter operating at a low flow rate gives much better resolution than a short and squat column.

SEC was used as a final purification step and served to exchange the enzyme into a more suitable buffer for long term storage. Both pre-packed super fine Superdex-75 resin in a HR10/30 column (10 mm x 31 cm; Pharmacia) or prep-grade Superdex-75 resin in a HR16/60 column (16 mm x 60 cm; Pharmacia) were used for SEC. For the Superdex-75 HR10/30, 200  $\mu$ l of 2 mg mL<sup>-1</sup> protein solution was isocratically eluted in 20 mM EPPS buffer (pH 8.0) containing 2 mM EDTA at 0.2 mL min<sup>-1</sup>. The load could be increased to 500  $\mu$ l of sample when using the prep grade HR16/60 column. Flow rates of 0.5 mL min<sup>-1</sup> were used with the column due to the larger bead size and the higher capacity of this column. Fractions of 1 mL were collected, analysed for activity and by SDS-PAGE. The best fractions were concentrated and stored at 4°C for use in subsequent biochemical characterisation studies.

#### 3.3.4.4 Purification of Recombinant PNGase F using Immobilised Metal Affinity Chromatography (IMAC)

In IMAC, a metal chelator such as imino diacetate (IDA), nitrilo triacetic acid (NTA) or tris(carboxymethyl) ethylene diamine (TED) is immobilised onto a polysaccharide matrix through a spacer arm and loaded with a particular divalent metal ion. The metal ion is then able to form complexes with proteins possessing suitable donor ligands such as histidine, and to a lesser extent cysteine and tryptophan at neutral pH (Scopes, 1994). Stable protein-metal ion complexes are formed only when the histidine residues are in a suitable grouping that satisfies the steric requirements of the metal ion. Such spatial configurations are rarely found in proteins however, so proteins that have similar properties with respect to charge, MW and amino acid composition that differ in conformation can be separated. More recently this technique has been exploited in the purification of recombinant protein through the addition of a poly-histidine tag to either the N- or C- termini of protein allowing for one-step purification using IMAC. Apart from the spatial configuration of the donors within the protein, five other factors exert a considerable influence on the binding and selectivity of the matrix (Scopes, 1994): the nature of the immobilised ligand, the length of the spacer arm from matrix to chelate, the metal ion itself, the pH, and the nature of buffers.

The most commonly used metals in IMAC are copper ( $\text{Cu}^{2+}$ ), nickel ( $\text{Ni}^{2+}$ ), zinc ( $\text{Zn}^{2+}$ ) and iron ( $\text{Fe}^{3+}$ ).  $\text{Fe}^{3+}$  has been shown to be selective for phosphorylated proteins;  $\text{Cu}^{2+}$  ions can inactivate proteins through forming covalent interactions with cysteines.  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  ions do not chelate with the IDA group (Janson *et al.*, 1996);  $\text{Ni}^{2+}$  ions leave three vacant sites that can be occupied by either solvent or buffer molecules when the three-dentate IDA is used to chelate them (Davankov *et al.*, 1977). Immobilised  $\text{Ni}^{2+}$  has been found to have a lower capacity for histidine ligands compared to  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$  matrices and is susceptible to reaction with reducing agents such as dithiothreitol (DTT). Boehringer Mannheim recommend using zinc as the preferred metal for their Poly-His purification resin claiming none of the above disadvantages.

Metal chelating columns can possess an overall negative charge, as the chelating ligand is negatively charged while the metal ions bound to it are positive and the opposite charges do not always cancel out. Consequently, to avoid ion exchange effects, IMAC is operated at high ionic strength with up to 1 M NaCl present in the buffers. Such high salt conditions can modify binding capacities by adding hydrophobic interaction to the metal liganding properties (Scopes, 1994). Elution of bound proteins can be achieved using other metal chelators such as EDTA, increasing the ionic strength, changing the net charge on the protein through controlling the pH, or as in this study displacement of the bound protein with imidazole.

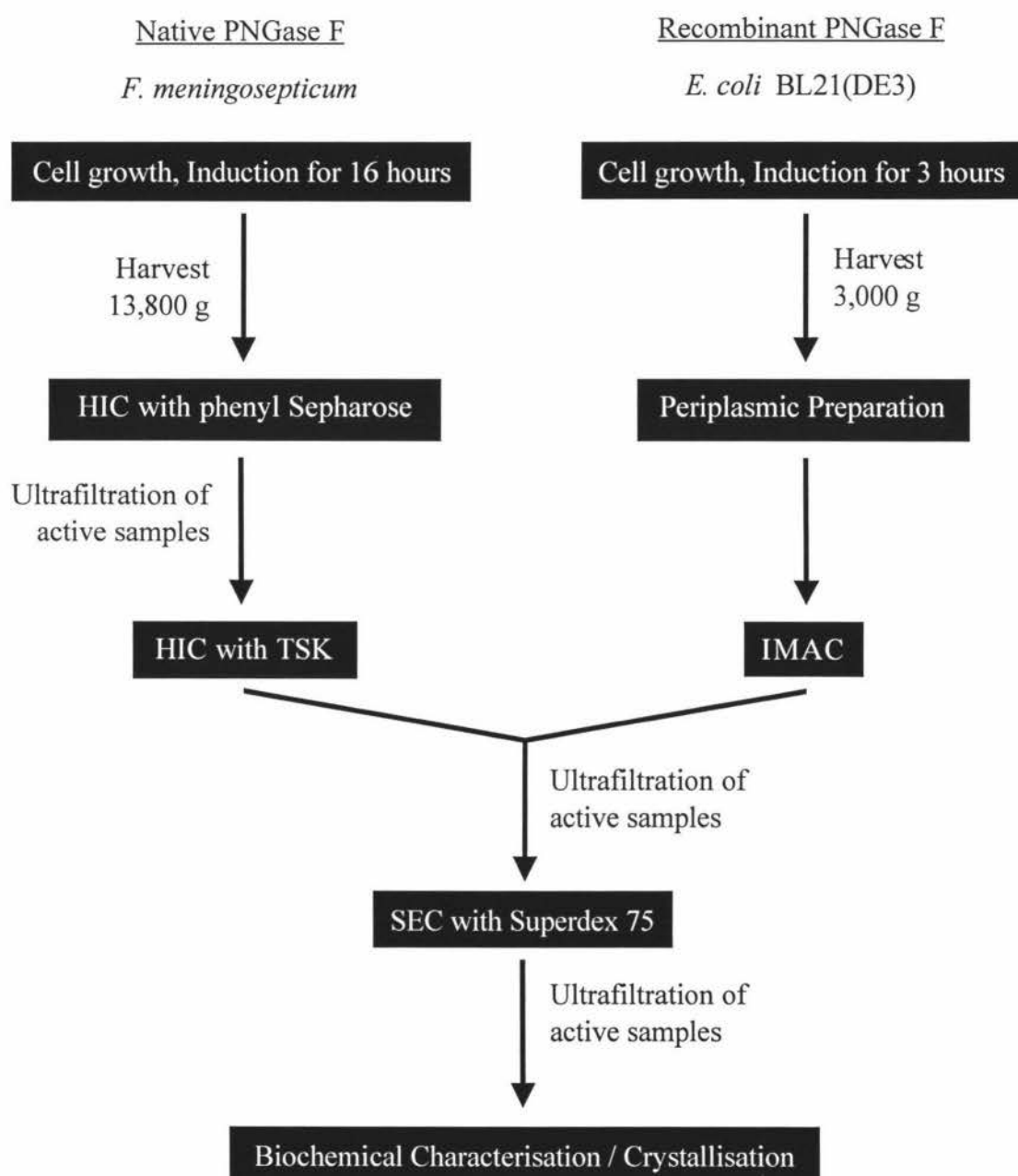
Two commercially available metal chelating matrices supplied by Boehringer Mannheim and Amrad Pharmacia were tested in this experiment as well as two different metal ions  $\text{Ni}^{2+}$  and  $\text{Zn}^{2+}$  ions.  $\text{Cu}^{2+}$  and  $\text{Fe}^{3+}$  were not used for reasons mentioned above. A standard procedure for IMAC used in this study is described below. The metal chelating resin was washed first with Milli-Q water to remove all traces of ethanol from storage then packed into a column (1 x 7 cm; Bio-Rad Econo-column). One column volume (~6 mL) of a 0.1 M metal solution ( $\text{ZnCl}_2$  or  $\text{NiCl}_2$ ) was applied to the column at a flow rate of 1 mL min<sup>-1</sup> and excess, unbound metal ions were removed from the column by extensive washing with Milli-Q water. To avoid the bleeding of metal ions from the column during purification, the column was washed with the elution buffer (20 mM MOPS, 0.5 M NaCl, 500 mM imidazole, pH 8.0) before being equilibrated with loading buffer (20 mM MOPS, 0.5 M NaCl, pH 7.5). The periplasmic fraction was dialysed against the loading buffer before being loaded and unbound proteins were removed by elution with the loading buffer. The bound proteins were eluted using increasing concentrations of imidazole in steps of 20, 50 or 500 mM imidazole in the elution buffer. Ten column volumes of buffer were applied at each elution step and the active fractions assayed by SDS-PAGE before being pooled and concentrated. The resin was regenerated by stripping chelated metal ions with stripping buffer (loading buffer with 100 mM EDTA, pH 8.0), then washed thoroughly with Milli-Q water.



## 3.3.4.5 Size Exclusion Chromatography with Superdex 75

The pooled fractions from IMAC were concentrated to a minimal volume (~1 mL) by ultrafiltration in a stirred cell as in section 3.3.4.1.1. The procedure for gel filtration of recombinant PNGase F was the same for native enzyme described in section 3.3.4.3.

A summary of the purification scheme for both native and recombinant enzyme is presented in figure 3.4.



**Figure 3.4** A summary of the purification scheme for PNGase F.



### 3.4 Results and Discussion

#### 3.4.1 Verification of the Extinction Coefficient of Recombinant PNGase F

Tarentino *et al.*, (1990) reported the molar extinction coefficient for PNGase F from *F. meningosepticum* (ATCC 33958) at 280 nm ( $A_{280}$ ) as being 18 for a 10 mg mL<sup>-1</sup> solution. However, the recombinant PNGase F from this study has eight amino acid substitutions and thus may have a different coefficient value. The first step in determining the extinction coefficient was to calculate the coefficient for the guanidine hydrochloride-denatured PNGase F, assuming aromatic amino acids Tyr, Trp as well as Cys all contribute to absorbance at a particular wavelength. The next step was to measure the  $A_{280}$  for both the native and denatured proteins. The DNA sequence revealed that PNGase F have 16 Tyr, 9 Trp and 6 Cys residues. These numbers together with the known coefficients for Tyr, Trp and Cys were substituted into the following equation to calculate the extinction coefficient for the guanidine hydrochloride-denatured recombinant enzyme:

$$\begin{aligned}\epsilon_{M.Gdn.HCl} &= A \epsilon_{M.Tyr} + B \epsilon_{M.Trp} + C \epsilon_{M.Cys} \\ &= 16(1280) + 9(5690) + 6(120) \\ &= 72410 \text{ M}^{-1} \text{ cm}^{-1}.\end{aligned}$$

The  $Abs_{280.Native}$  and  $Abs_{280.Gdn.HCl}$  were measured experimentally and were used to calculate the coefficient of the intact protein in the following equation:

$$\begin{aligned}\epsilon_{M.Native} &= (A_{280.Native}) (\epsilon_{M.Gdn.HCl}) / (Abs_{280.Gdn.HCl}) \\ &= 0.30765 (72410) / 0.32703 \\ &= 68.12 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}.\end{aligned}$$

The  $A_{280}$  of a 1% solution of recombinant PNGase F was calculated by:

$$\begin{aligned}
 A_{280} \text{ of a 1\% solution} &= \epsilon c l \\
 &= (\epsilon_{M, \text{Native}}) (\text{concentration}) (\text{length of light path}) \\
 &= (\epsilon_{M, \text{Native}}) (10 \text{ mg mL}^{-1} / \text{MW}) (\text{length of light path}) \\
 &= (68.12 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}) (10 / 36251 \text{ M}) (1 \text{ cm}) \\
 &= 18.79 \pm 0.9.
 \end{aligned}$$

The calculated molecular weight of the recombinant protein is 36251 Da and the method has an accuracy of  $\pm 5\%$ . The value found for recombinant PNGase F related closely to the figure quoted by Tarentino *et al.*, (1990). In addition, the extinction coefficient for the pure enzyme was verified by lyophilising a known amount of protein in a volatile buffer solution (10 mM ammonium acetate) as calculated by the coefficient and the dried sample was weighted.

### 3.4.2 Comparison of Protein Quantitation Methods

In order to calculate the percentage recovery and the degree of purification, two measurements have to be followed throughout the purification procedure. The recovery of PNGase F was monitored by measuring the deglycosylation activity of the enzyme using the discontinuous assay described in section 3.3.1.3. The total protein in each fraction was measured using Bicinchoninic Acid, Bradford, Lowry and UV methods, then compared to dry weight to assess the best assay to use.

Results obtained from both the Lowry and BCA assays indicated that these methods suffered a significant interference from interfering substances such as salts, buffer ions and detergents present in the sample (not shown). On the other hand, the Bradford assay was found to be quick and relatively free from interference by biochemical reagents. However, this method consistently underestimated (about 14%) the concentration of PNGase F in a solution when compared with the values obtained by the UV method, possibly due to the large dye response with BSA. As the method was

an equilibrium-binding process, the response was not linear and depended on the exact pH of the mixture. The method was only reliable when the protein was dissolved in low concentrations of buffer at neutral pH (Scopes, 1994).

Although the UV method is undoubtedly the most convenient and accurate (when using a known extinction coefficient), it is only suitable in the final stages of purification when the number of contaminating proteins is small. The Bradford method therefore was used extensively throughout the purification scheme despite its shortcomings as it was least prone to interference. The UV method relies on absorption at 280 nm of aromatic groups in the protein and is therefore dependent on the amino acid composition of the protein. Most proteins have an extinction coefficient that falls in the range of one absorbance unit being equivalent to of 0.7 to 2.5 mg mL<sup>-1</sup>. However, there are anomalies such as some parvalbumins that completely lacking either tyrosine or tryptophan have  $\epsilon_{280}=0$  and a tryptophan-rich lysozyme which has an  $\epsilon_{280}=2.65$ . The absorption in the far-VU around the peptide bond (205 nm) is far more sensitive as it is due to absorption by the peptide bond and hence is much less affected by the amino acid composition. By employing the formula below, the extinction coefficient can be calculated with less than 2% error (Scopes, 1994):

$$\epsilon_{205}^{1 \text{ mg mL}^{-1}} = 27 + 120 \times \frac{A_{280}}{A_{205}}$$

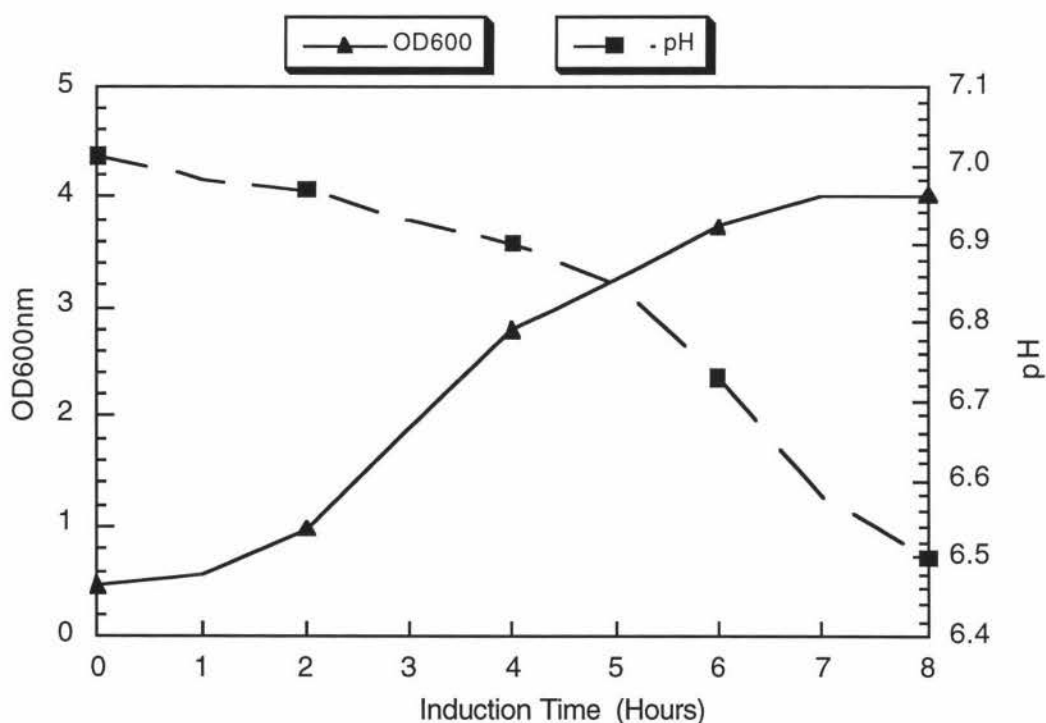
However, this method has several disadvantages:

1. Virtually all commonly used salts and buffers absorb strongly at 205 nm causing interference and the only common anions that do not are sulfate and perchlorate.
2. The wavelength is close to the useful limits of routine spectrophotometers
3. Measurements at 205 nm require high quality quartz cuvettes and a relatively new deuterium lamp.

While this method is theoretically the most accurate non-destructive method for determining protein concentration, it was not used because of the difficulties in its use. As the extinction coefficient at 280 nm had been both theoretically and empirically determined to within  $\pm 5\%$ , it was used to estimate PNGase concentration towards the end of the purification.

### 3.4.3 Growth and Protein Production Studies of Transformed *E. coli* Cells

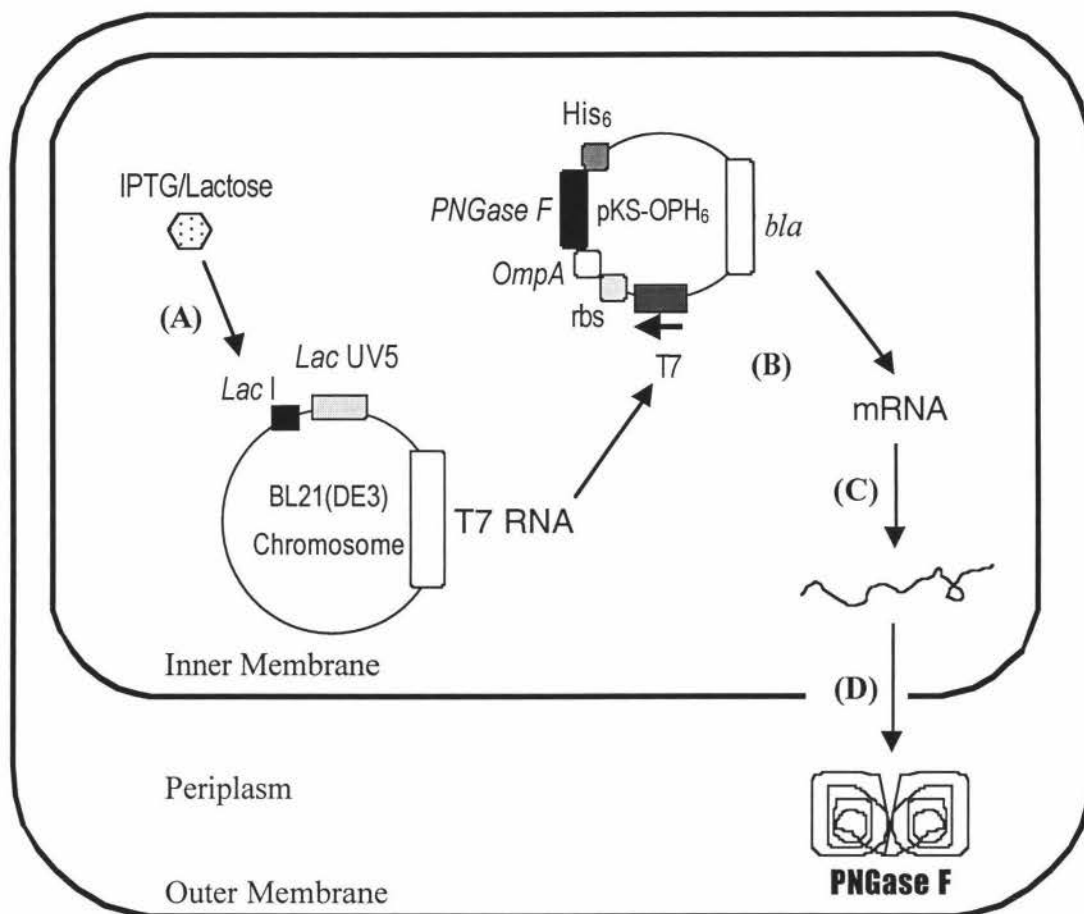
The growth of *E. coli* cells transformed with the pOPH6 expression vector was monitored over a 8 hour period after induction by the increase in optical density at 600 nm and the decrease in pH of culture medium due to the accumulation of metabolites (figure 3.5).



**Figure 3.5** Growth curve of *E. coli* transformed with pOPH6.

The cell growth has reached a plateau (stationary phase) at 7 hours after induction. A range of factors could contribute to this observation: the accumulation of metabolites to a toxic level (indicated by a decrease in pH); depletion of nutrients; decrease in aeration due to over-crowding; possible release of strain-specific toxins; and possible cell fatality due to the over-efficient transcription by T7 polymerase that drastically reduces overall transcription of essential proteins needed for survival by the host polymerase. The latter factor may lead to plasmid instability in the host BL21(DE3) cells (Studier *et al.*, 1990). This phenomenon was observed when some frozen glycerol stocks did not produce PNGase F after induction. However, this observation may also have been the result of the high concentration of glycerol (30% v/v) in the frozen stocks. It has been reported that glycerol becomes increasingly toxic to the cells as they are thawed before being diluted for growth (Studier *et al.*, 1990). To avoid this complication, new preparations of recombinant PNGase F were made using a fresh transformation of host cells with the expression vector as described in section 2.4.9 in chapter 2.

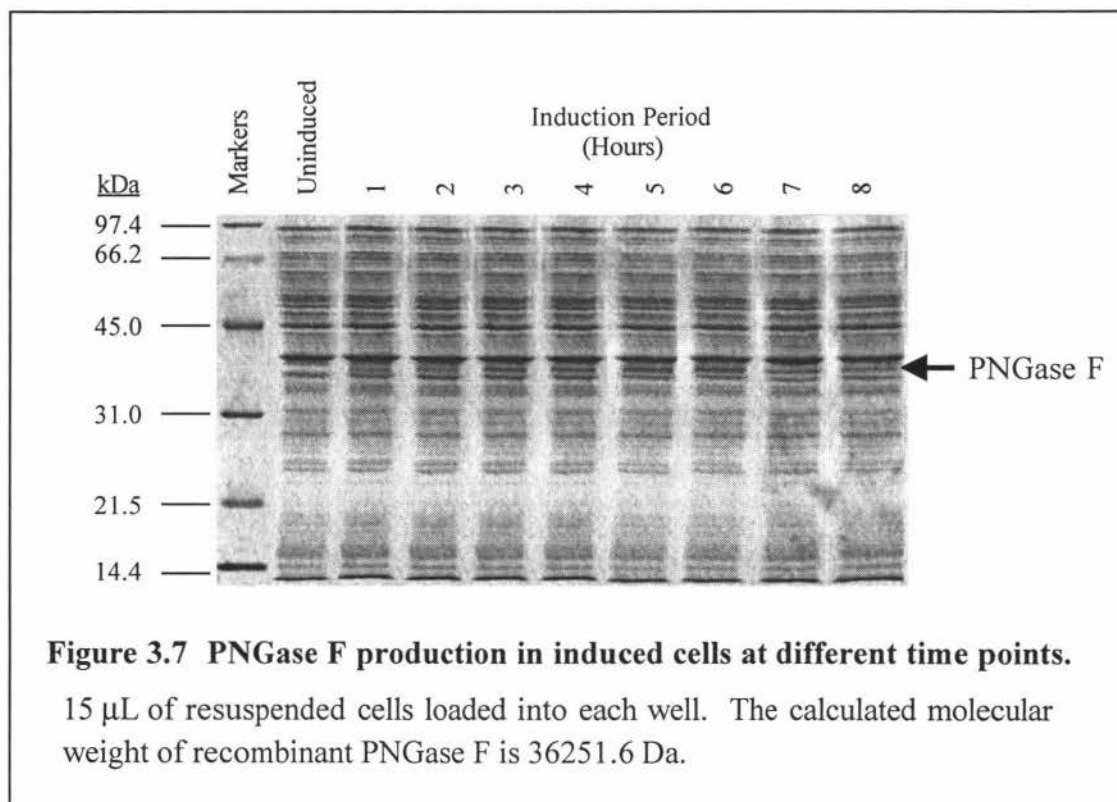
Figure 3.6 shows the induction of PNGase F synthesis by IPTG or lactose and the impending export of the nascent polypeptide into the periplasm. The *E. coli* BL21(DE3) chromosome contains a *LacI* gene which produces the *lacI* repressor protein that binds to the *LacUV5* promoter and prevents transcription of the T7 RNA polymerase gene. When IPTG or lactose is added to the cell media, it binds to the *lac* repressor allowing the transcription of the T7 RNA polymerase gene present on BL21(DE3) chromosome (A). The T7 RNA polymerase recognises the T7 promoter on pKS-OPH<sub>6</sub> plasmid, producing a messenger RNA (mRNA) containing a ribosome binding site (rbs), an *ompA* signal peptide, the PNGase F gene, and a hexa-histidine tag (B). Ribosomes in *E. coli* recognise the rbs and translate the mRNA into a polypeptide (C). The secretory mechanism inside *E. coli* recognises the *ompA* signal peptide and exports the nascent PNGase F into the periplasmic space where the protein folds into its native conformation (D).



**Figure 3.6** A schematic diagram of the *E. coli* BL21(DE3) cell showing IPTG induction of PNGase F synthesis and subsequent secretion of the nascent polypeptide into the periplasm.

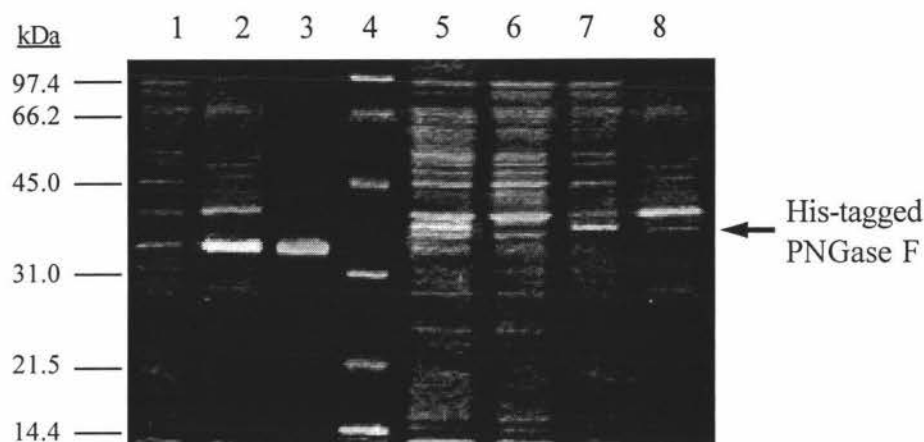
The expression of recombinant PNGase F was also monitored by measuring enzyme activity and 12% SDS-PAGE of the whole cell extracts (figure 3.7). These showed that the amount of recombinant enzyme produced reached a maximum after a three-hour induction period and appeared to decrease after six hours. This is possibly due to plasmid instability and/or proteolysis of the enzyme by host proteases. There are a large number of proteases in *E. coli* that are localised in the cytoplasm, the periplasm, and in the inner and outer membranes. These proteolytic enzymes participate in a number of metabolic activities, including the highly selective and efficient removal of abnormal proteins such as those expressed by a cloned foreign DNA (Goldberg and Goff, 1986). These mechanisms of protein degradation are poorly understood and it is unlikely that all proteolytic pathways or enzymes operating in *E. coli* have been

identified (Makrides, 1996). Thus, proteolysis inside the *E. coli* cell might explain the reduction of the recombinant enzyme observed after longer periods of induction when the bacteria are under stress. The putative PNGase F band was electroblotted onto a PVDF membrane and submitted for N-terminal protein sequencing to confirm the identity of the recombinant enzyme.



The recombinant protein expressed in *E. coli* BL21(DE3) cells transformed with the first construct, pKS-PNG, were produced in a largely insoluble state as shown in figure 3.8 (lane 2). However, when expressed as a fusion with the leader sequence from the *ompA* gene, hexa-histidine-tagged PNGase F was efficiently processed and exported to the *E. coli* periplasm (lane7).



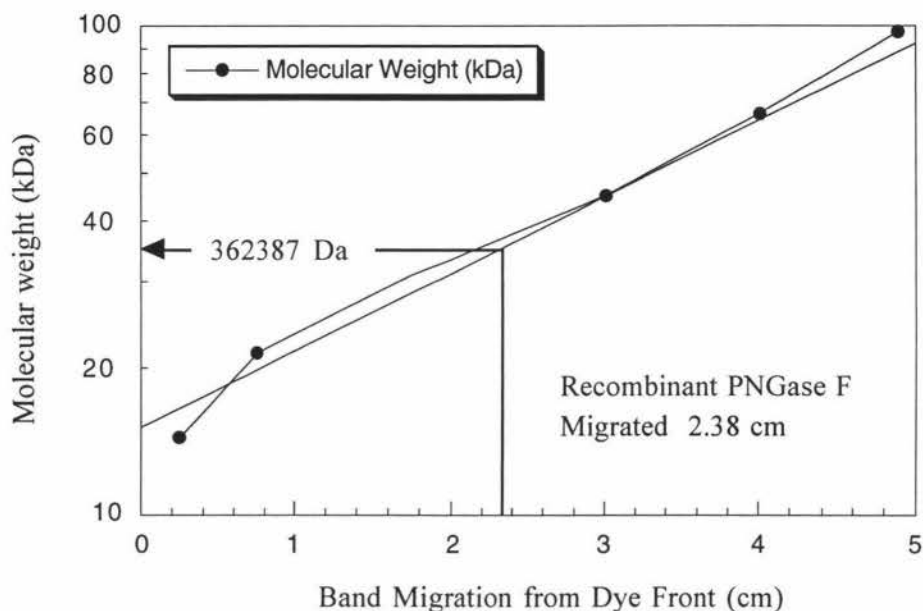


**Figure 3.8** An inverse image of a 12% SDS-PAGE with silver staining shows the protein solubility problem of pKS-PNG is resolved by secretion in pKS-OPH6.

Lane

- (1) Supernatant fraction of sonicated *E. coli* carrying the pKS-PNG construct.
- (2) Pellet fraction of sonicated *E. coli* carrying the pKS-PNG construct.
- (3) Native PNGase F.
- (4) Bio-Rad molecular weight markers.
- (5) Whole cell extract of induced *E. coli* carrying the pOPH6 construct.
- (6) Whole cell extract of *E. coli* carrying the pKS-OmpA3-His host vector (no PNGase F gene).
- (7) Supernatant fraction of sonicated *E. coli* carrying the pOPH6 construct.
- (8) Pellet fraction of sonicated *E. coli* carrying the pOPH6 construct.

A curve generated by plotting the molecular weight of the Bio-Rad markers run on a 12% SDS PAGE versus the distance migrated from the dye front (in centimetres) was used to estimate the molecular weight of recombinant PNGase F (figure 3.9). The molecular weight of the protein was found to be 36387 Da from the curve compared to a weight of 36251 Da by Expasy Tools ([www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html)). There are two possible explanations for the difference in molecular weight: The degree of SDS binding to proteins varies with different proteins and the cross-linking of acrylamide is not necessary uniform across the gel.

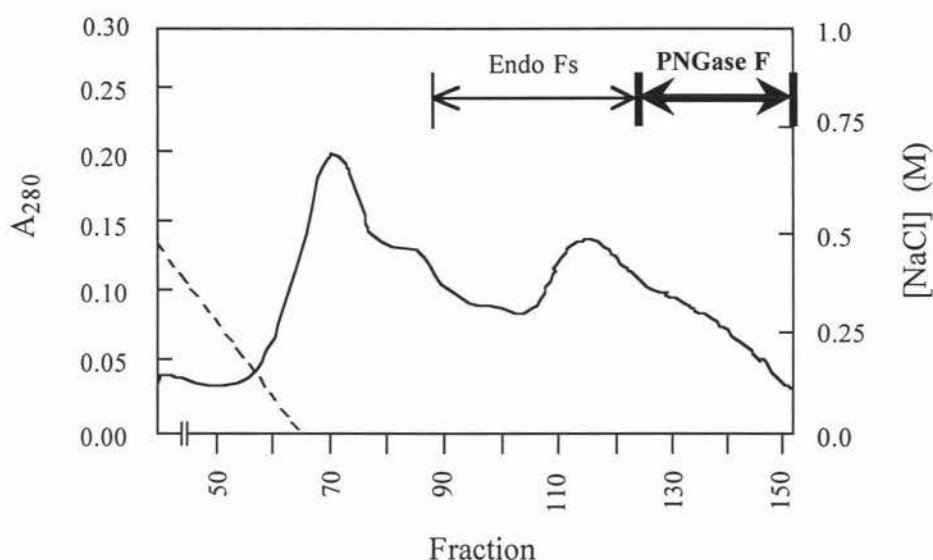


**Figure 3.9** Standard curve used to estimate the molecular weight of recombinant PNGase F.

### 3.4.4 Purification of Native PNGase F

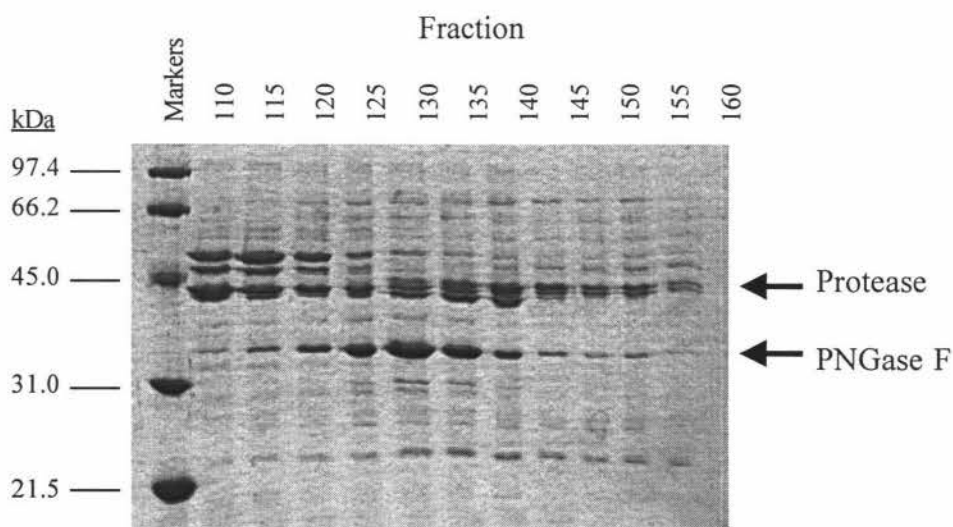
#### 3.4.4.1 HIC with Phenyl Sepharose 6 Fast-Flow (Low Substitution)

This first step in the purification removed the bulk of the contaminants and separated PNGase F from the three Endo Fs produced by this bacterium. Figure 3.10 shows the elution profile after extensive washes removed unbound proteins. Because of the high hydrophobicity of PNGase F, Endo Fs were eluted from the column at the end of the salt gradient (0% NaCl) with PNGase F being the most hydrophobic and eluting last. Fractions containing PNGase F activity were analysed by SDS-PAGE to assess the purity of the enzyme (figure 3.11). Native PNGase F has a molecular weight of 34.8 kDa and peaked between fractions 125 to 140. The fractions were pooled and concentrated by ultrafiltration in a stirred cell for the next step which was to remove the major contaminant, a 45 kDa protein reported to be a zinc-aspartyl endoprotease termed P45 protease (Tarentino *et al.*, 1995). P45 protease was later renamed P40 protease to reflect its true molecular weight of 40,086 Da. Interestingly, this protease has a single O-linked complex glycan (Plummer *et al.*, 1995).



**Figure 3.10** Chromatogram of the elution from Phenyl Sepharose 6 Fast-Flow (low substitution) of *F. meningosepticum* culture medium.

Arrows indicate active fractions and the dashed line represents the NaCl gradient.

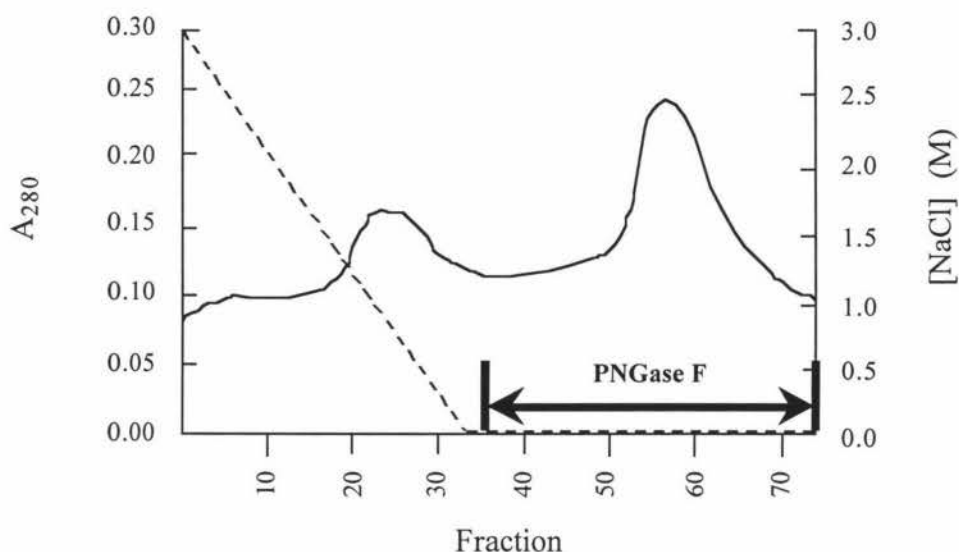


**Figure 3.11** SDS-PAGE of the PNGase F fractions from Phenyl Sepharose 6 Fast-Flow (low substitution).

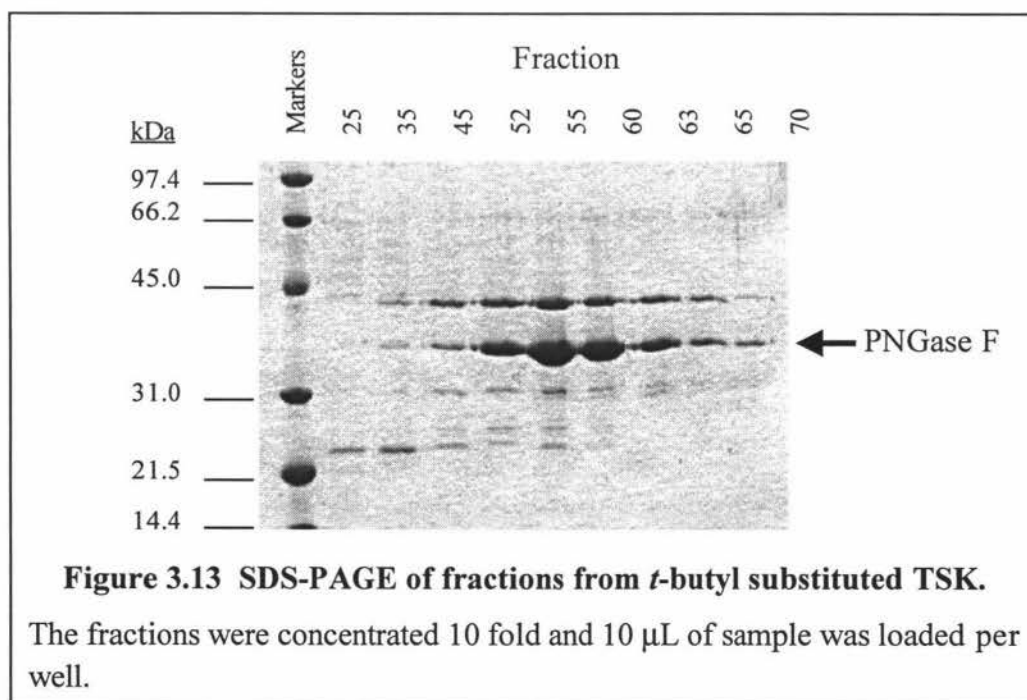
2 mL samples were concentrated 10 fold and 10  $\mu$ L of concentrate was loaded per well.

### 3.4.4.2 HIC with Home-Made *t*-butyl Substituted TSK

The second step is the most effective step in the purification scheme as it removes the troublesome 44 kDa protease band that tends to co-elute with native PNGase F. Figure 3.12 shows the elution of the enzyme from the TSK column where two main peaks are visible. While the second peak contained most of the PNGase F activity it was still contaminated with the 44 kDa band (figure 3.13).



**Figure 3.12** Chromatogram of PNGase F on *t*-butyl substituted TSK. Arrows indicate active fractions and the dashed line represents the NaCl gradient.



**Figure 3.13** SDS-PAGE of fractions from *t*-butyl substituted TSK. The fractions were concentrated 10 fold and 10  $\mu$ L of sample was loaded per well.

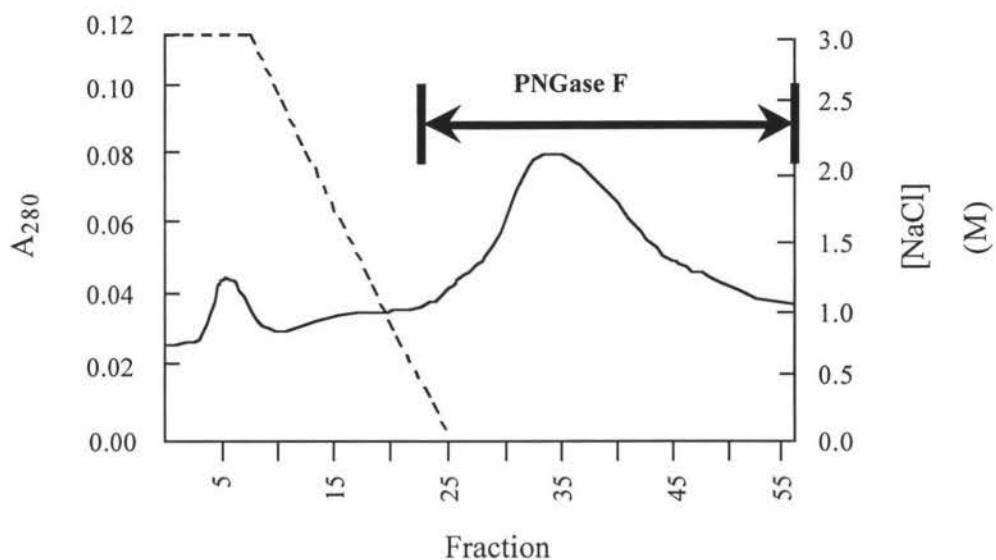
This was probably due to the resin losing the functional butyl group over the 8 long years of service. Two batches of substituted TSK resin were made during this period. One batch totally failed to bind PNGase F even with 4 M NaCl, and was presumed to have no substitution possibly due to degradation of the CDI that is used to replace the hydroxyl group on the resin with an imidazole which is used for the *t*-butyl group substitution. In contrast, a second batch had such a high substitution that once the enzyme was bound, it could not be eluted unless acetonitrile was added to the mobile phase. This resin was made back in 1995 and a quick test was performed to assess its usability as a resin in 1999.

The test involved transferring 500 µl of resin (~0.12 g) into a 1.5 mL microfuge tube, washing it several times in high salt loading buffer (50 mM NaAc, 5 mM EDTA, with 2, 4 M NaCl or 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>), before equilibrated samples were added. The mixture was gently agitated at 4°C for 15 minutes by taping the tube to an orbital shaker (IRA). After removal of loading buffer by centrifugation, the solid matrix was washed with 3 volumes of loading buffer before being eluted by several washings with non-salt buffer. No activity was detected in either the loading buffer washes or in the elution fractions (results not shown) indicating that the resin bound the enzyme too tightly to be suitable to use for purification. Obviously, the correct degree of substitution in the resin was difficult to replicate as no titration to determine the degree of substitution had been recorded for the original preparation. It was therefore decided to try and find a commercially available product that could achieve the separation desired.

### 3.4.4.3 HIC Tests with various Hydrophobic Matrices

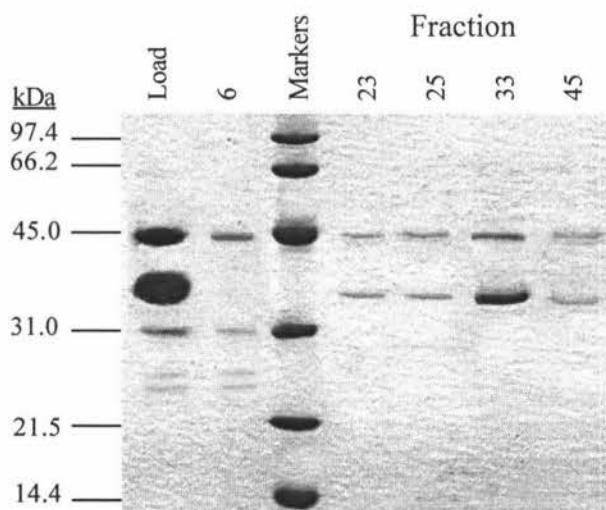
#### 3.4.4.3.1 Phenyl Sepharose 6 Fast-Flow (High Substitution)

The enzyme eluted from the column in one very broad peak (figure 3.14) and SDS-PAGE analysis of the collected fractions showed that while the enzyme bound to the column there was only a minor reduction in the contaminating bands associated with PNGase F (figure 3.15).



**Figure 3.14** Chromatogram of elution of PNGase F from Phenyl Sepharose 6 Fast-Flow (high substitution).

Arrows indicate active fractions and the dashed line represents the NaCl gradient.

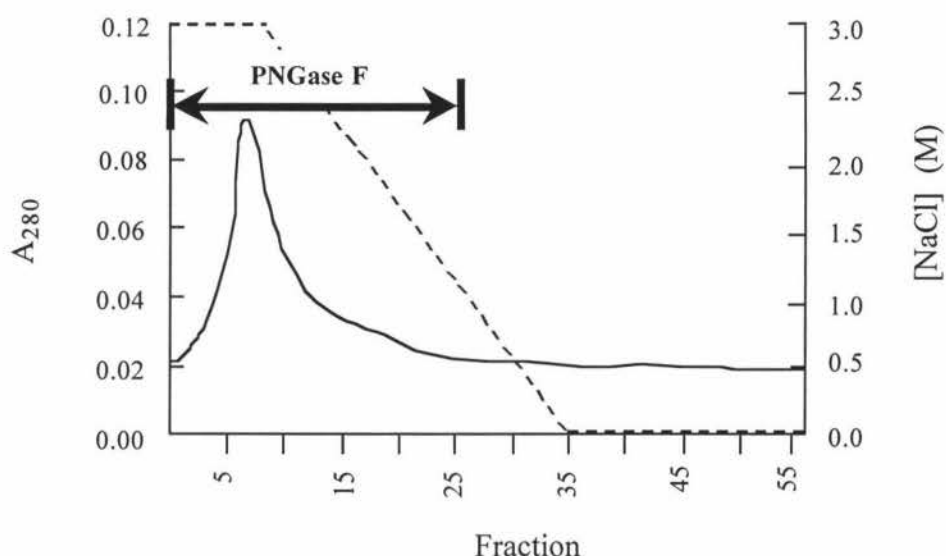


**Figure 3.15** SDS-PAGE of fractions from Phenyl Sepharose (high substitution).

The samples were concentrated 10 fold and 10  $\mu$ L of concentrate was loaded in each well.

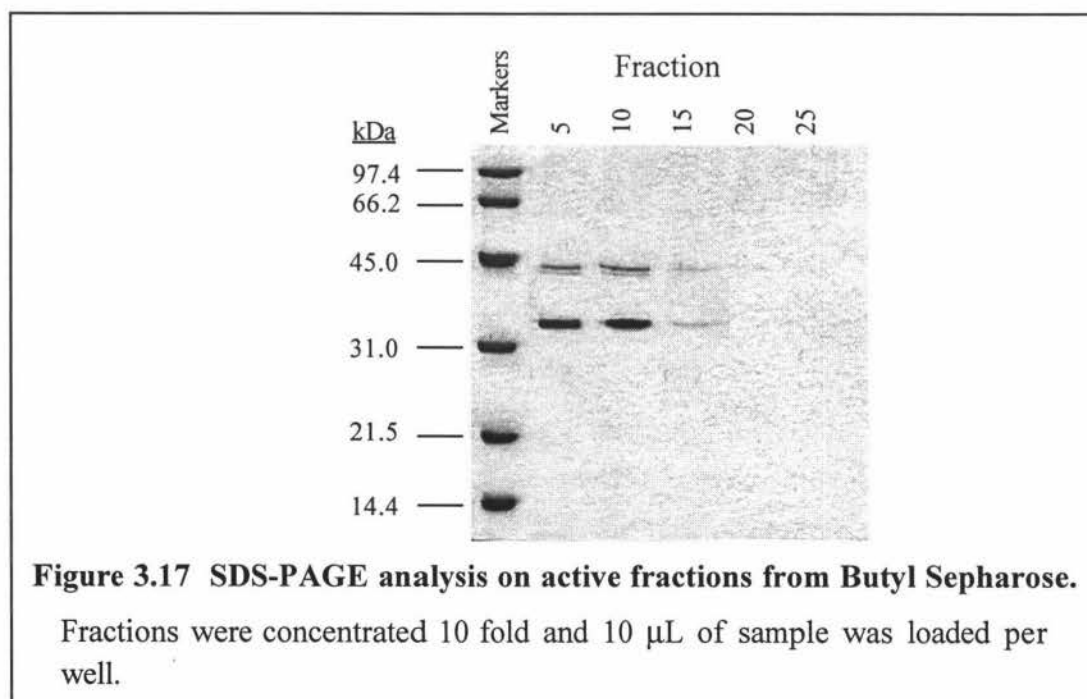
### 3.4.4.3.2 Butyl Sepharose 4 Fast-Flow

Figure 3.16 showed that the enzyme did not bind to the substituted resin even with high initial NaCl concentrations. Both the activity assay and SDS-PAGE analysis (figure 3.17) confirmed this observation. In hindsight, other concentrations of ammonium sulfate (a more lyophobic salt) should have been attempted.



**Figure 3.16** Chromatogram of elution from Butyl Sepharose 4 Fast-Flow.

Arrows indicate active fractions and the dashed line represents the NaCl gradient.



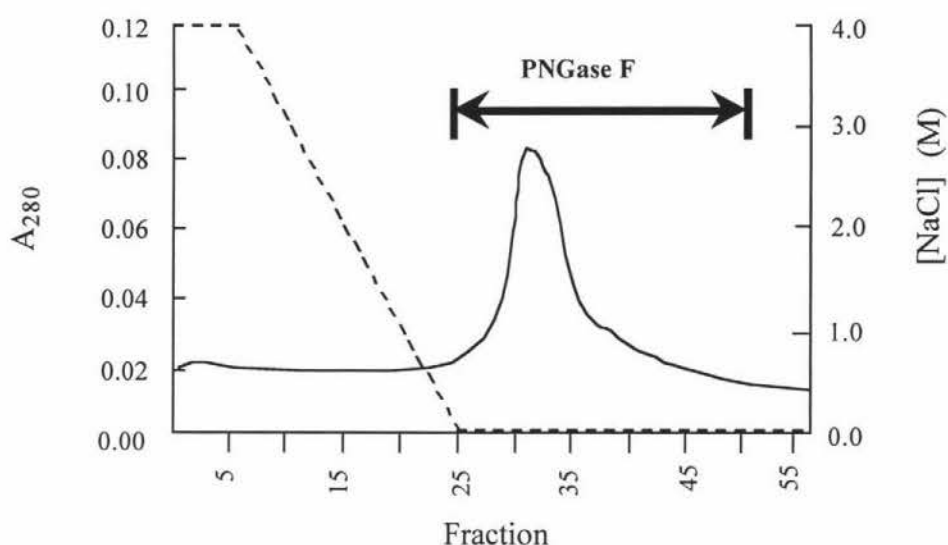
**Figure 3.17** SDS-PAGE analysis on active fractions from Butyl Sepharose.

Fractions were concentrated 10 fold and 10  $\mu$ L of sample was loaded per well.



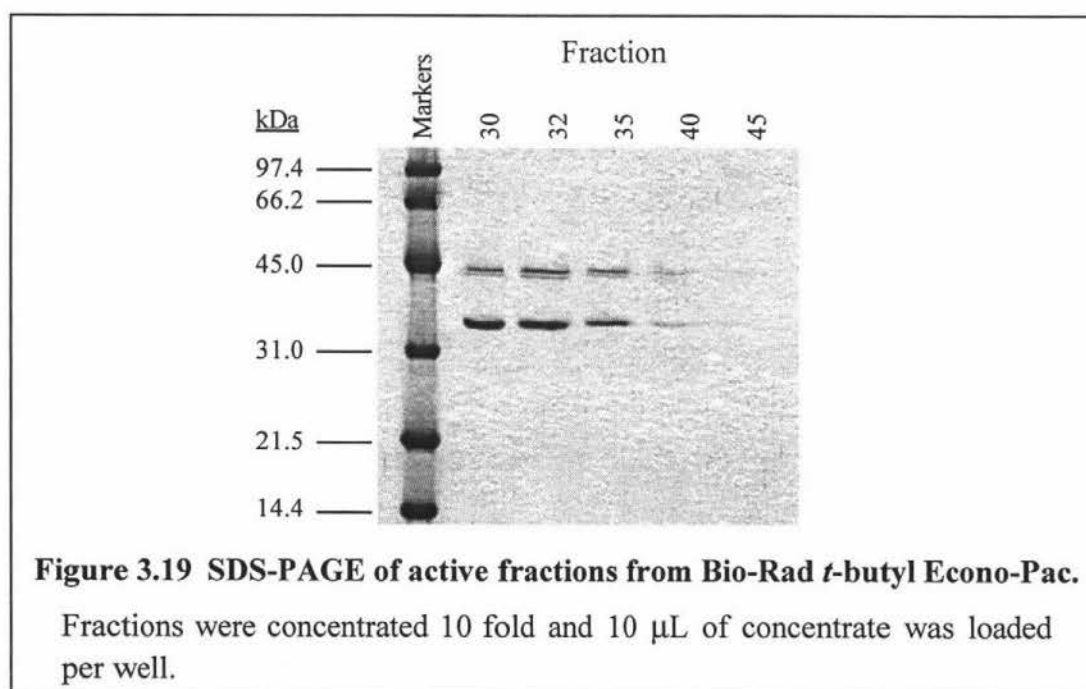
### 3.4.4.3.3 Macro-Prep *t*-butyl Econo-Pac

This resin has the same substitution group as the in-house TSK, the only difference being in the matrix, vinyl polymers in in-house versus polyacrylic in Macro-Prep. The enzyme bound well to the *t*-butyl resin as it was eluted at 0% NaCl of the gradient indicating strong binding (figure 3.18). Unexpectedly, no separation was achieved (figure 3.19) indicating that the substitution group on the resin and the nature of the resin matrix both play a role in protein binding.



**Figure 3.18** Chromatogram of elution from *t*-butyl Econo-Pac.

Arrows indicate active fractions and the dashed line represents the NaCl gradient.

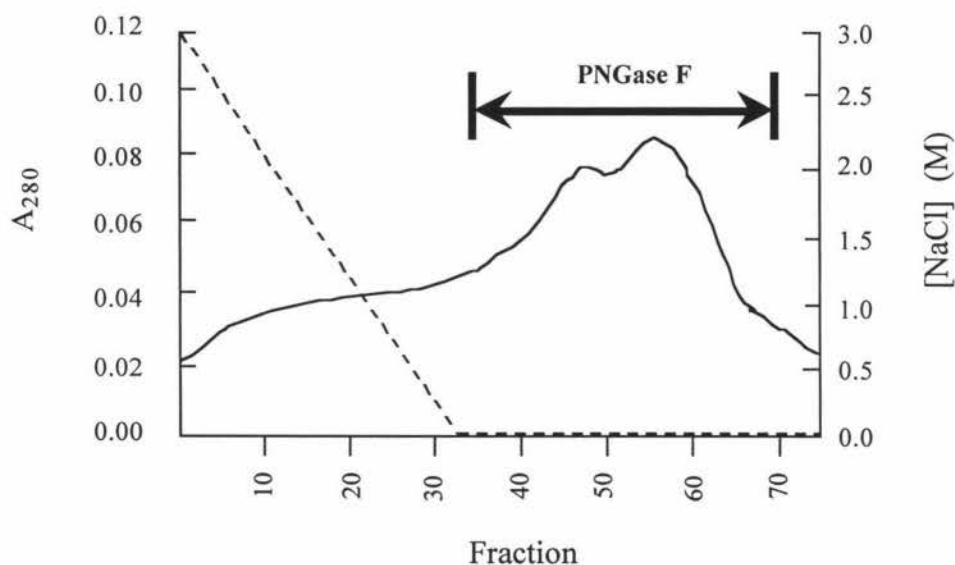


**Figure 3.19** SDS-PAGE of active fractions from Bio-Rad *t*-butyl Econo-Pac.

Fractions were concentrated 10 fold and 10  $\mu$ L of concentrate was loaded per well.

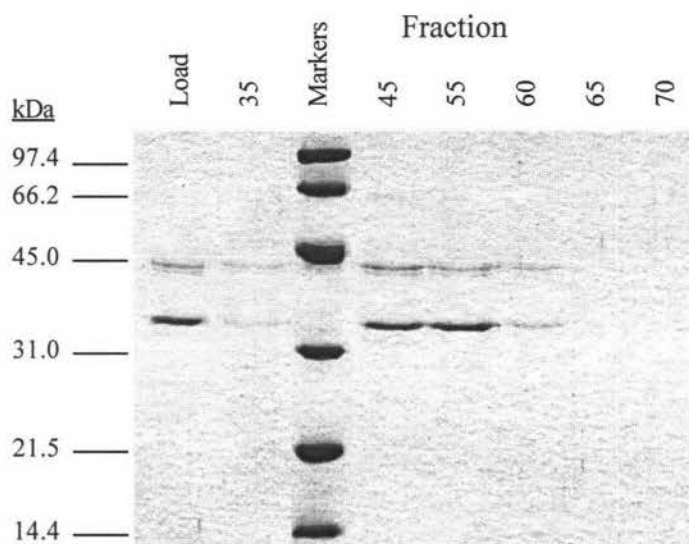
#### 3.4.4.3.4 TSK-butyl-Toyopearl 650M

This resin has been used in great success in the purification of several PNGases including PNGase At (*Aspergillus tubigensis*), F (*Flavobacterium meningosepticum*), GM (*Glycine max*), OV (hen oviduct) and mouse-derived L929 fibroblast cells (Ftouhi-Paquin *et al.*, 1997; Tarentino *et al.*, 1993c; Kimura *et al.*, 1998; Suzuki *et al.*, 1997; Suzuki *et al.*, 1994a). However the separation obtained with this resin was disappointing. Two unresolved peaks were eluted near the end of the gradient indicating poor separation of proteins (figure 3.20). SDS-PAGE analysis of the fractions showed however that there was some reduction of the 44 kDa band (figure 3.21). Unfortunately, extending the gradient from 50 to 100 minutes did not improve separation (results not shown). This indicates the amino acid changes in PNGase F from *F. meningosepticum* (CDC strain 3352) may have some influence on the hydrophobic properties of the protein.



**Figure 3.20** Chromatogram of elution from TSK-butyl-Toyopearl.

Arrows indicate active fractions and the dashed line represents the NaCl gradient.



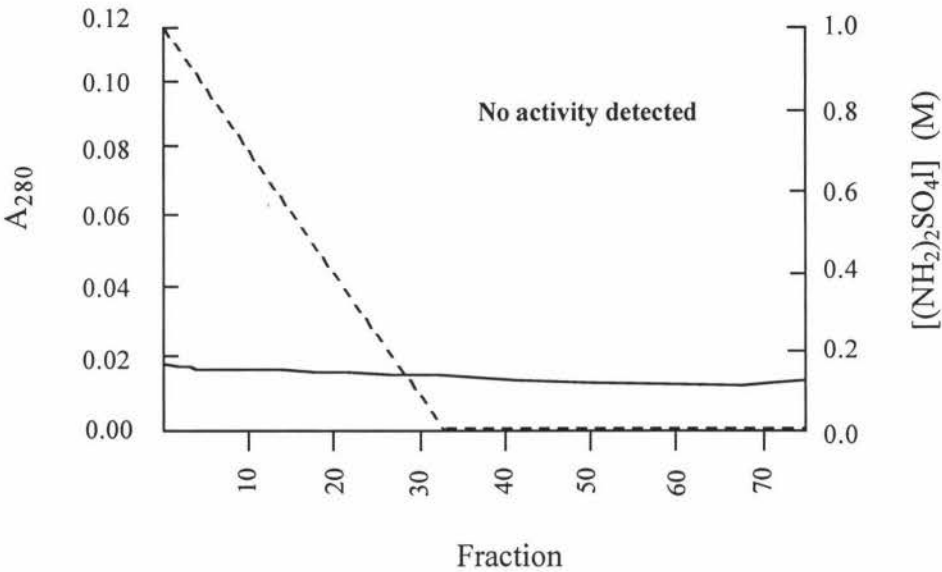
**Figure 3.21 SDS-PAGE analysis of active fractions from TSK-butyl-Toyopearl.**

The fractions were concentrated 10 fold and 10  $\mu$ L of concentrate was loaded per well.

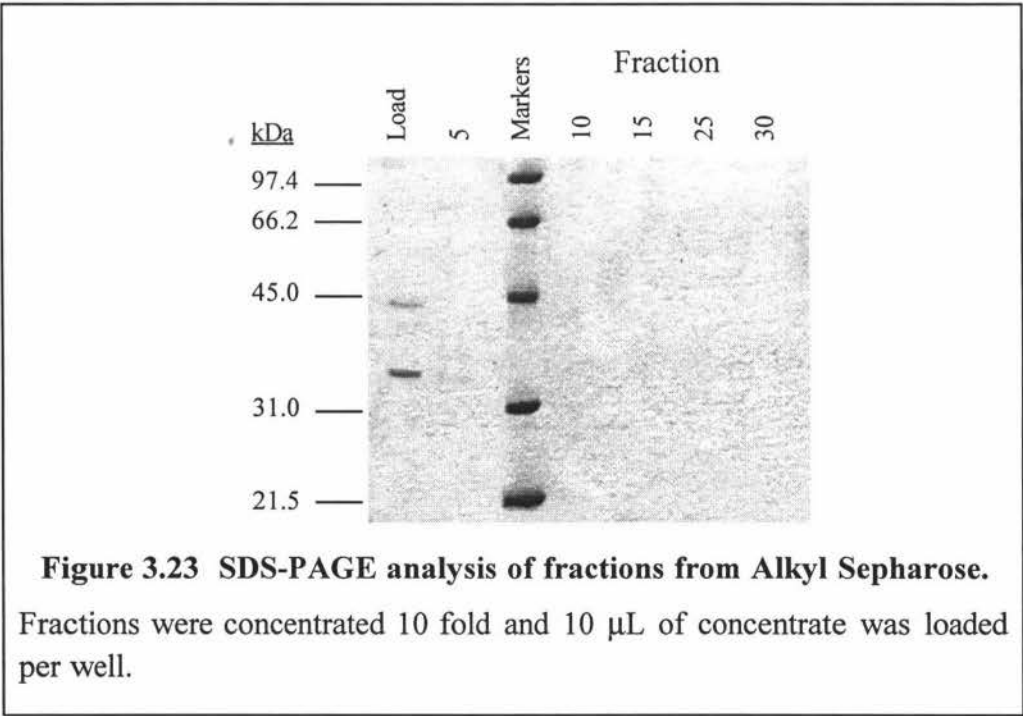
#### 3.4.4.3.5 Alkyl Sepharose

The sample was initially loaded at 3 M NaCl but no protein bound to the column (results not shown). A more lyotropic salt (1 M  $(\text{NH}_4)_2\text{SO}_4$ ) was therefore used to promote protein binding to the hydrophobic matrix. Although  $\text{K}_3\text{PO}_4$  is a more lyotropic salt than  $(\text{NH}_4)_2\text{SO}_4$ , it was not used as possible salt precipitation is a problem at 4°C. Figures 3.22 and 3.23 show that no protein was eluted and activity tests showed that there was no PNGase activity in the eluent. It was therefore assumed that the enzyme remains bound to the column. In order to elute the enzyme from the column, a buffer containing 50% (v/v) ethylene glycol, an effective surface tension-reducing agent that lessens the hydrophobic interactions with functional ligands (Scopes, 1994), was used. However, no enzyme was eluted (results not shown). Stronger chaotropic salts such as thiocyanate, iodide and lithium bromide were also not used because the concentration needed can cause denaturation of the displaced protein. The column was then washed with 0.1 M NaOH to remove any bound proteins. Interestingly, SDS-PAGE analysis of the NaOH wash showed no proteins were eluted.

As this treatment should remove any bound proteins, it can only be assumed that the protein must therefore have remained bound to the column because of unusually strong hydrophobic interactions, or have become hydrolysed or leached from the matrix slowly.

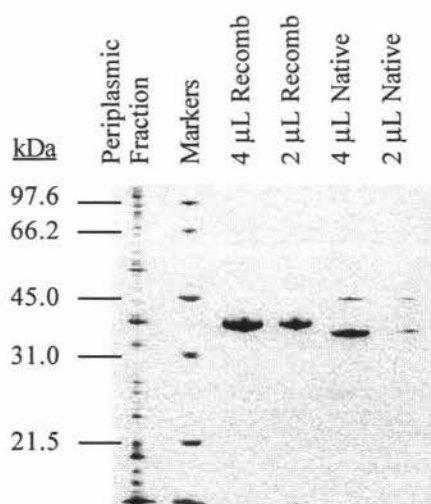


**Figure 3.22** Chromatogram of elution of PNGase F from Alkyl Sepharose. The dashed line represents the NaCl gradient.

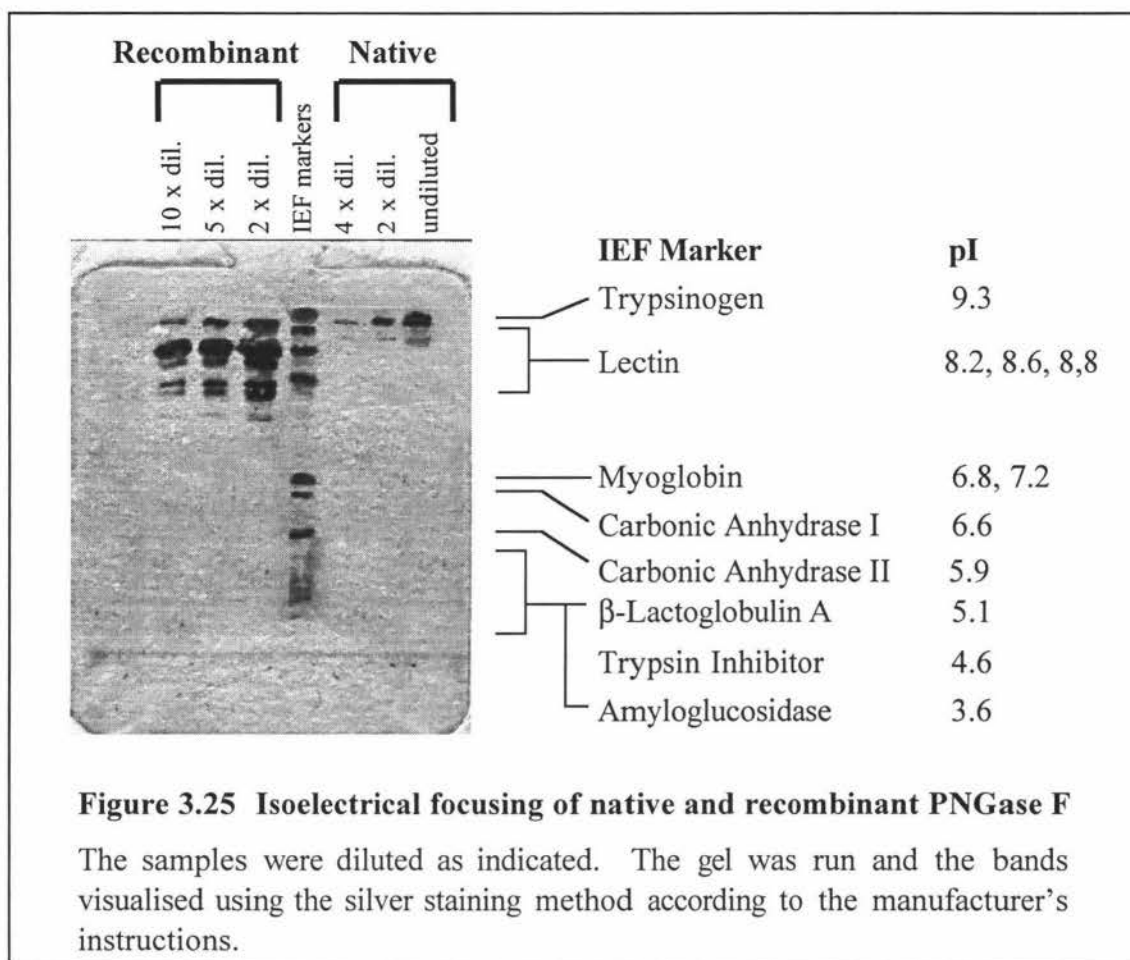


**Figure 3.23** SDS-PAGE analysis of fractions from Alkyl Sepharose. Fractions were concentrated 10 fold and 10  $\mu\text{L}$  of concentrate was loaded per well.

As none of the hydrophobic matrices tried successfully separated the protease at 44 kDa from PNGase F, it was decided to evaluate ion exchange chromatography as a purification step. To maximise the effect and chances of success with isoelectrical focusing (IEF), the isoelectric point (pI) of the enzyme was determined using the PhastSystem following the manufacturer's instructions. The purity of the samples used for IEF were analysed by SDS-PAGE using the silver staining method described in Appendix 5 (figure 3.24). The pI for the native enzyme (major band) was found to be 9.2 versus a calculated value of 8.14 from the amino acid sequence by ExPASy ProtParam Tools ([www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html)) and the 44 kDa band, the major contaminating band was found to be 8.7 (figure 3.25). Interestingly, the recombinant PNGase F appears as one band on SDS-PAGE (figure 3.24) but contains several bands on the IEF gel indicating that there may be several isoforms of the enzyme that are not present with the native protein. These bands could also be breakdown products of the recombinant protein. The pI value of a protein cannot be predicted from its amino acid composition because some of the acidic and basic groups are buried inside the protein and their  $pK_a$ 's are altered by the local dielectric constant. There are also hydrogen-bonding interactions, salt-bridging and ion-pairing interactions inside the protein that prevent the amino acid side chains from behaving as predicted.



**Figure 3.24** SDS-PAGE analysis of native and recombinant PNGase F used for IEF.



#### 3.4.4.4 Ion Exchange Tests with various Ion Exchangers

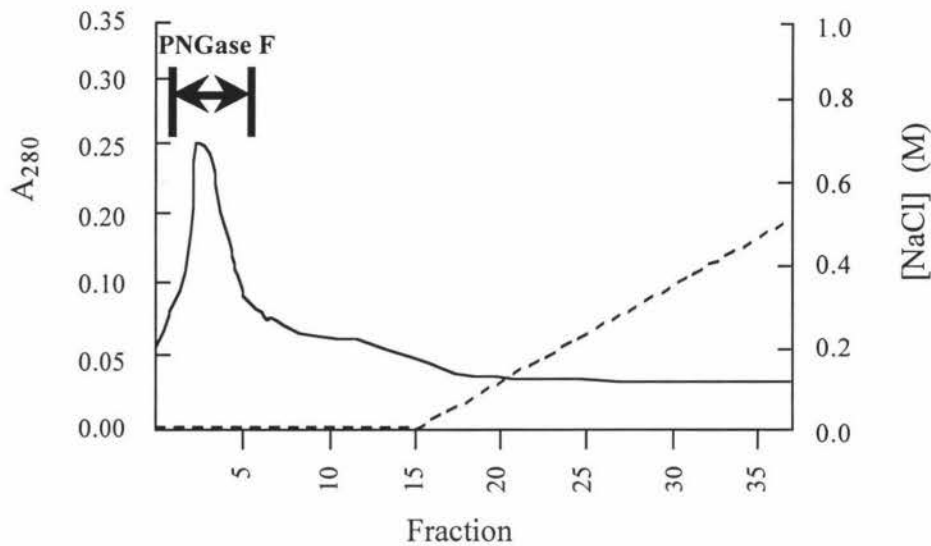
##### 3.4.4.4.1 Anion Exchange Chromatography

Mono-Q is a strong anion exchanger that will remain positively charged at a pH of 10.0. Two experiments designed to separate the protease from native PNGase F were carried out using buffers at pH 8.9 and 9.8.

##### A. Mono-Q at pH 8.9

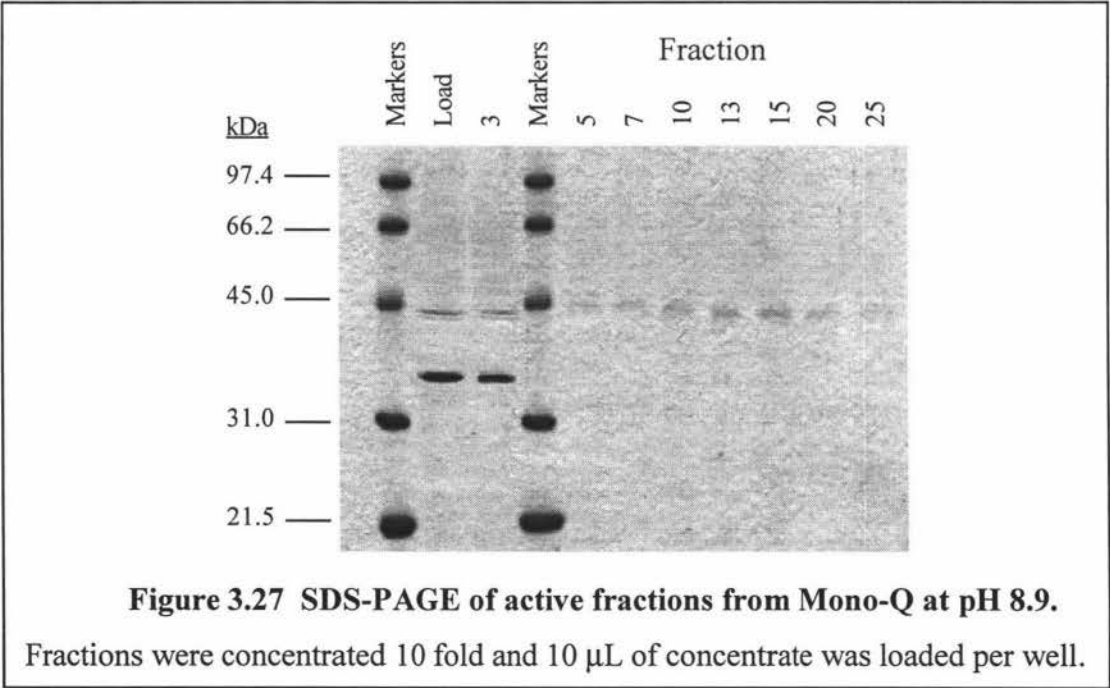
In the first experiment, the protein sample and column were equilibrated with 20 mM Tris-HCl buffer at pH 8.9. At this pH, the protease with a pI of 8.7 should carry a slight negative charge or be neutral whereas PNGase F with a pI of 9.2 should retain a positive net charge. The 44 kDa band should bind to the positively charged resin and thus be retained on the solid phase while the enzyme having the same charge as the functional groups on the resin should pass straight through the column.

Figure 3.26 shows that a single peak was eluted when the sample was loaded which extended to a shoulder when the column was washed, indicating that the protein did not bind to the resin. SDS-PAGE analysis confirmed that a small amount of the 44 kDa band was retarded by the matrix (figure 3.27). It was therefore decided to increase the pH to induce both proteins to bind to the column.



**Figure 3.26 Chromatogram of elution from Mono-Q at pH 8.9.**

Arrows indicate active fractions and the dashed line represents the NaCl gradient.



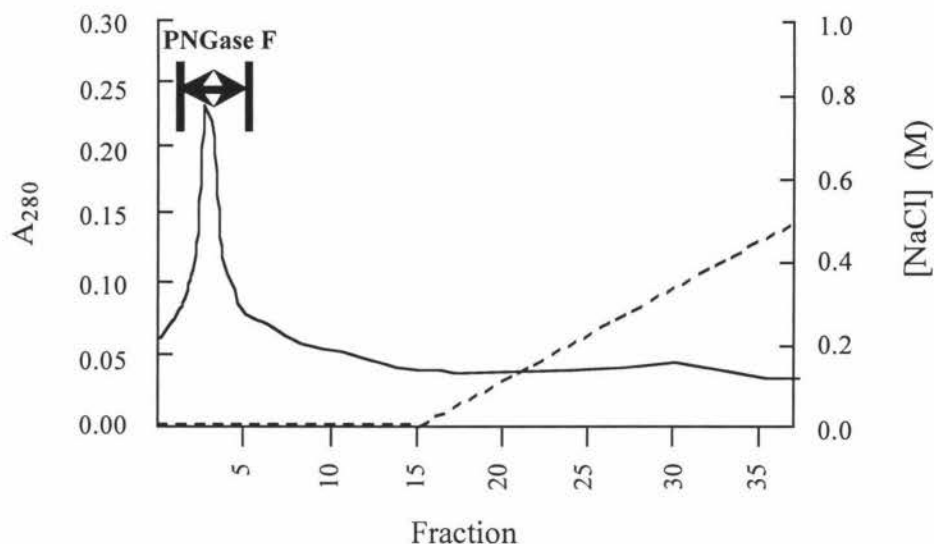
**Figure 3.27 SDS-PAGE of active fractions from Mono-Q at pH 8.9.**

Fractions were concentrated 10 fold and 10  $\mu$ L of concentrate was loaded per well.



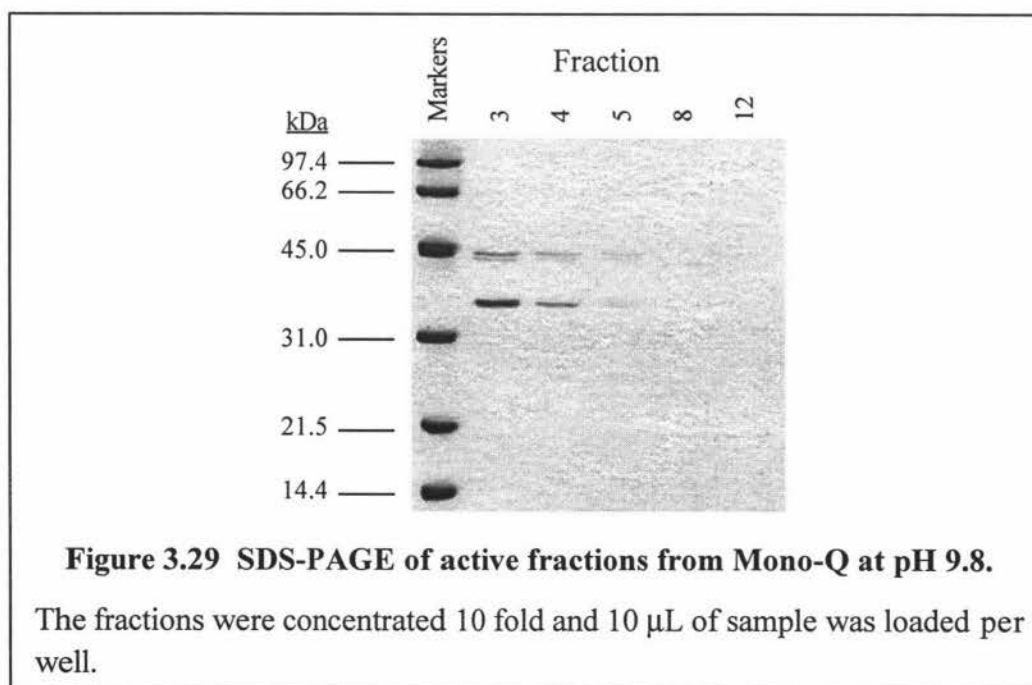
### B. Mono-Q at pH 9.8

In the second experiment, a 20 mM CAPSO buffer (pH 9.8) was used to equilibrate both the column and the protein sample. Both proteins should be negatively charged at this pH and should bind to the column at strengths dependent on the distribution of surface charge. However once more, neither protein bound to the column (figures 3.28 and 3.29). As matrix effects can influence binding, a Uno-Q column was also tested.



**Figure 3.28 Chromatogram of elution from Mono-Q at pH 9.8.**

Arrows indicate active fractions and the dashed line represents the NaCl gradient.



**Figure 3.29 SDS-PAGE of active fractions from Mono-Q at pH 9.8.**

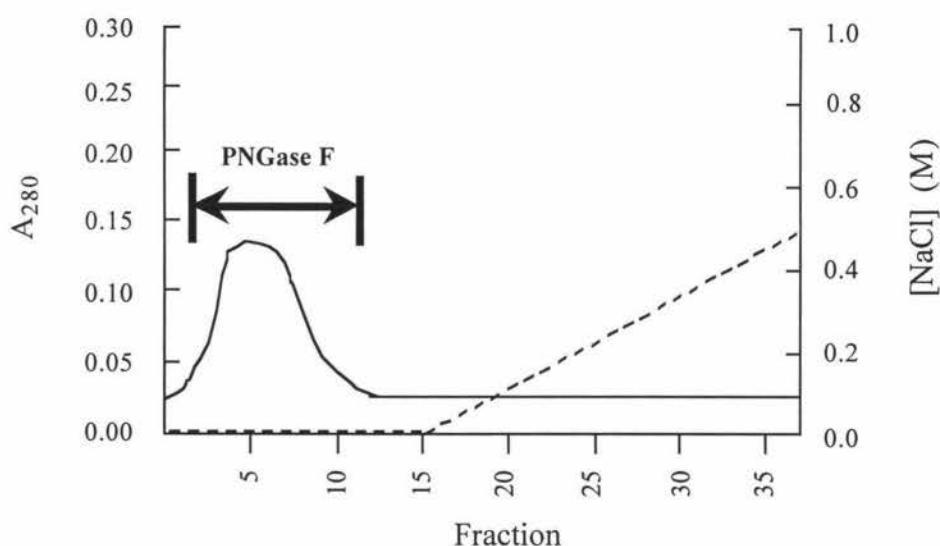
The fractions were concentrated 10 fold and 10  $\mu$ L of sample was loaded per well.

### 3.4.4.4.2 Uno-Q

Uno-Q is an unique continuous polyacrylamide gel matrix developed by Bio-Rad that allows a more uniform solvent flow and is able to withstand higher operating pressures. A buffering system and elution gradient identical to that used for Mono-Q was used in an attempt to isolate the two proteins.

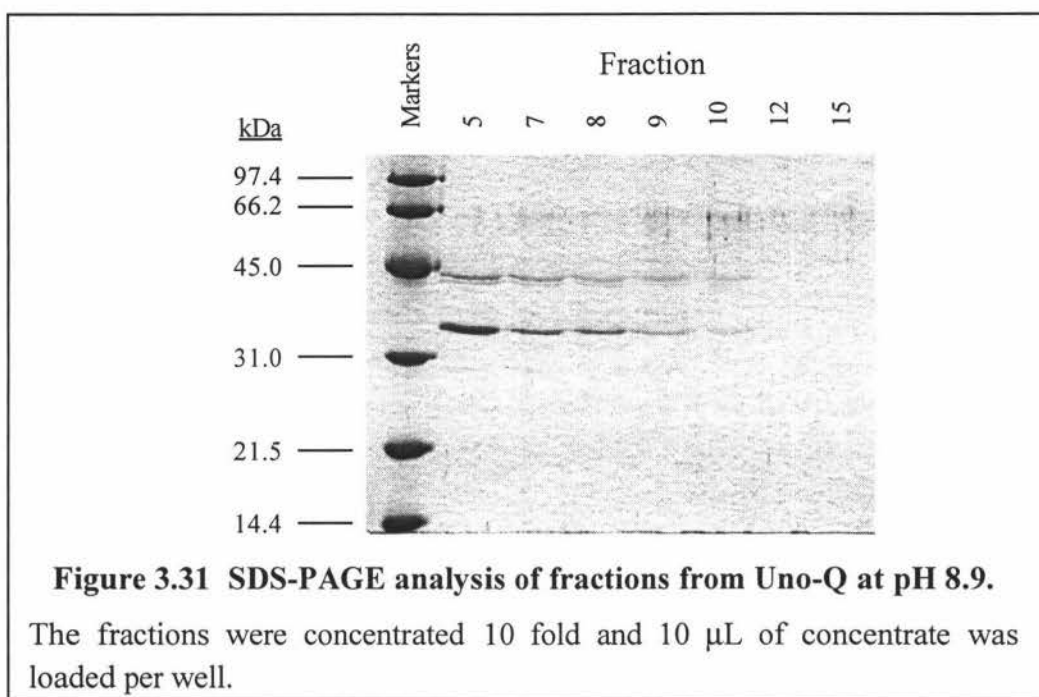
#### A. *Uno-Q at pH 8.9*

Figure 3.30 shows that a broad peak was eluted as the sample was loaded, with nothing being eluted during the subsequent salt gradient. The fact that no protein bound is shown by SDS-PAGE analysis (figure 3.31). The two bands at ~66 kDa are due to keratin contamination of the sample buffer used.



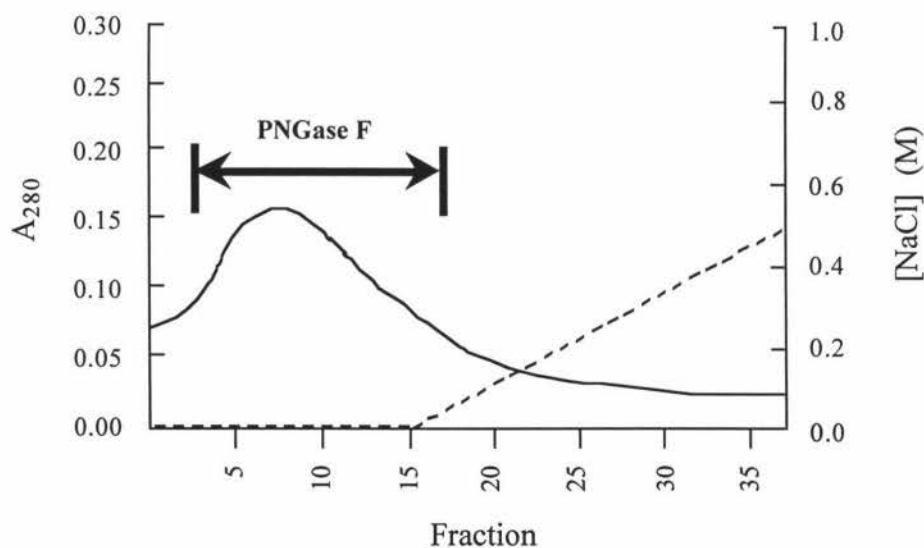
**Figure 3.30** Chromatogram of elution from Uno-Q at pH 8.9.

Arrows indicate active fractions and the dashed line represents the NaCl gradient.



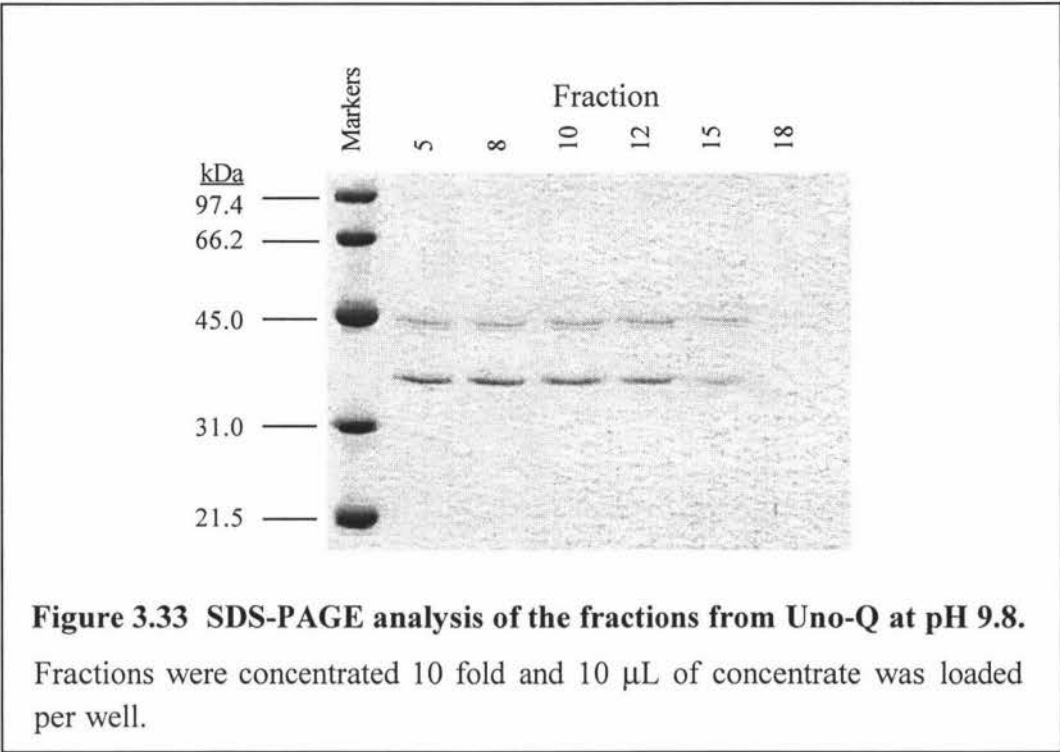
### B. *Uno-Q at pH 9.8*

Again, increasing the pH of the equilibration buffer did not promote binding, although the position of the peak indicated a slight increase in retention at pH 9.8 compared to pH 8.9 (figure 3.32). SDS-PAGE analysis of the collected fractions shows no separation (figure 3.33).



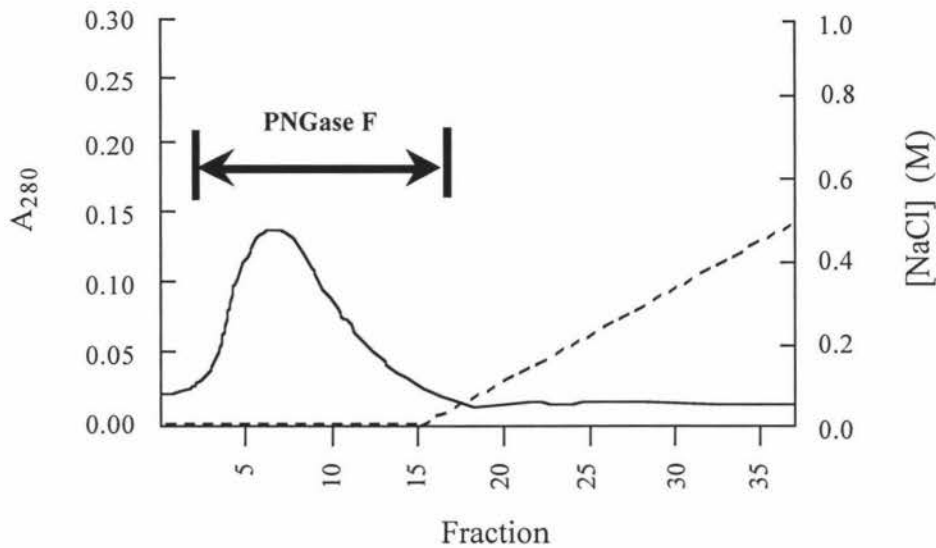
**Figure 3.32 Chromatogram of elution from Uno-Q at pH 9.8.**

Arrows indicate active fractions and the dashed line represents the NaCl gradient.



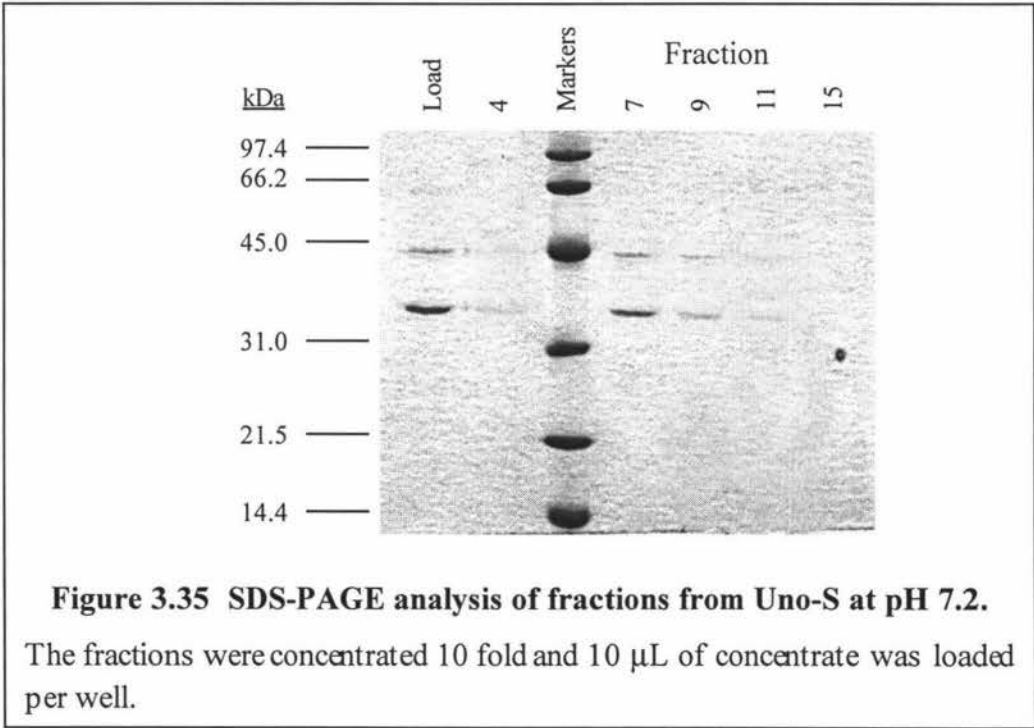
3.4.4.4.3 Cation Exchange Chromatography: Uno-S at pH 7.2

As anion exchangers do not separate the protease from PNGase F, cation exchange chromatography was attempted. Uno-S has a continuous polyacrylamide gel matrix and strong cationic functional groups that remains negatively charged below pH 4.0. Although the calculated and measured values for the pI of PNGase F were 8.1 and 9.2 respectively, it was decided to carry out the experiment at pH 7.2 to ensure binding of both proteins. However, the protein once more did not bind to the resin as shown in figure 3.34, and this was confirmed by SDS-PAGE analysis (figure 3.35).



**Figure 3.34** Chromatogram of elution from Uno-S at pH 7.2.

Arrows indicate active fractions and the dashed line represents the NaCl gradient.



**Figure 3.35** SDS-PAGE analysis of fractions from Uno-S at pH 7.2.

The fractions were concentrated 10 fold and 10  $\mu$ L of concentrate was loaded per well.

3.4.4.4.4 Hydroxyapatite column

Figure 3.36 shows that neither PNGase F nor the contaminating protease binds to the calcium hydroxyphosphate resin, a result confirmed by SDS-PAGE (figure 3.37).

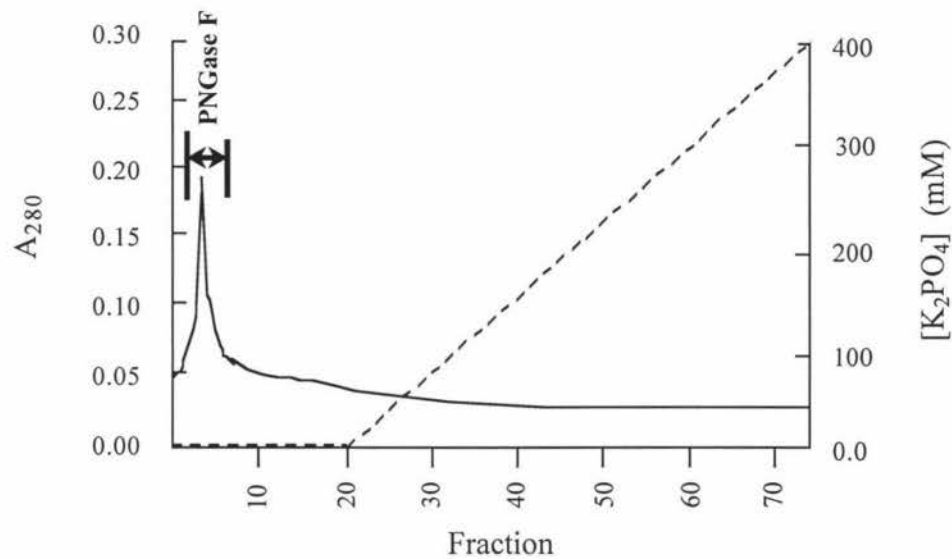


Figure 3.36 Chromatogram of elution from hydroxyapatite.

Arrows indicate active fractions and the dashed line represents the NaCl gradient.

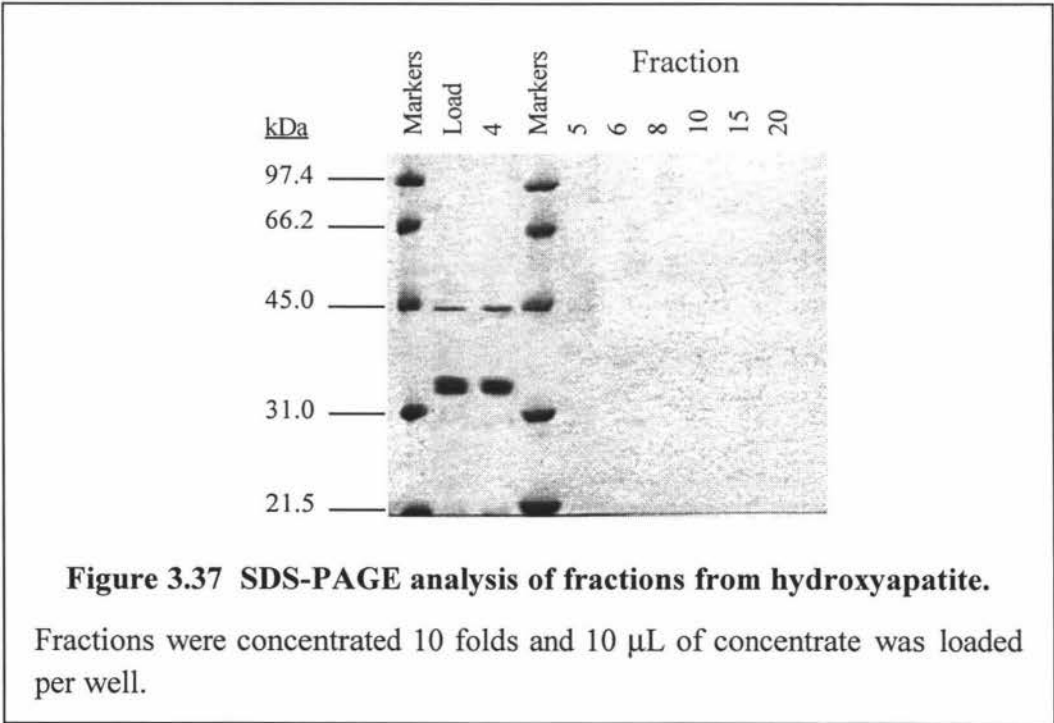
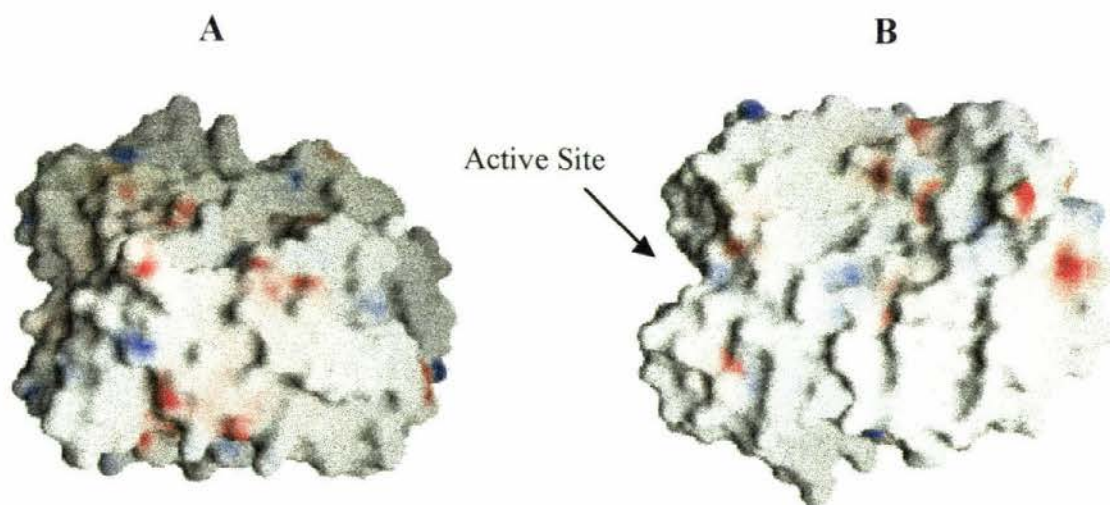


Figure 3.37 SDS-PAGE analysis of fractions from hydroxyapatite.

Fractions were concentrated 10 folds and 10  $\mu$ L of concentrate was loaded per well.

In the case of anion exchange chromatography, operating at pH 8.9, the contaminant may not be fully deprotonated which may account for it not binding to the resin. However, at pH 9.8, over one pH unit above the pI of the contaminant and half a pH unit above the pI of PNGase F, neither of the proteins bound to the positively charged ligands of the anion exchange column either. This is highly unexpected as the two pHs were deliberately chosen to be above the isoelectric point of the protease contaminant (pI of 8.7). No binding of either protein also occurred when samples and matrix were pre-equilibrated at pH 7.2, more than one pH unit below the pI of both proteins. This suggests the proteins may be shielding each other from functional groups on the resin, or that the distribution of charge on the surface is such that the positive and negative charges neutralise each other out. An examination of the structure of the enzyme shows that surface charge is evenly distributed in a way that wherever there is a positive charge, it is balanced by a nearby negative charge (figure 3.38). Hence there are no “patches” of charged groups on the surface of the protein for ion exchangers to interact with.



**Figure 3.38 Charged groups distribution on the surface of native PNGase F.**

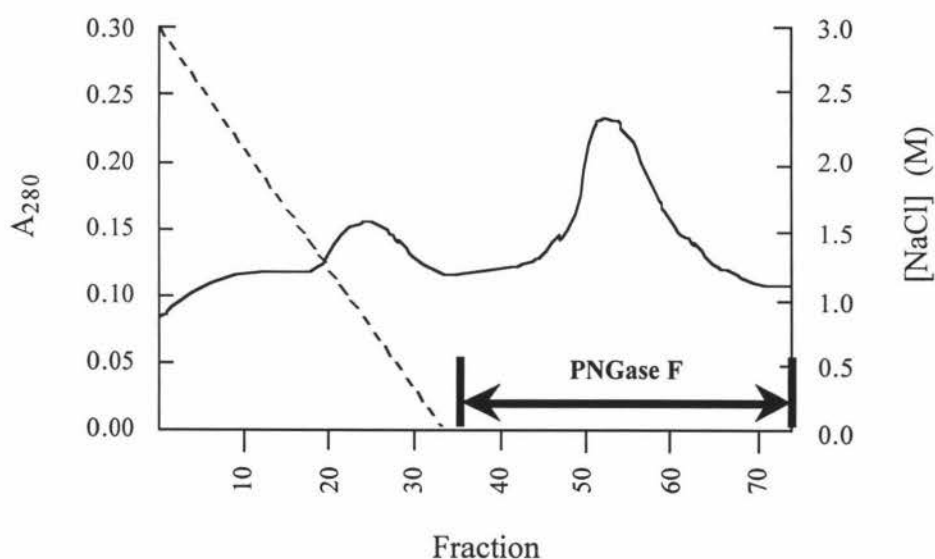
This representation of the structure was generated by GRASP (Nicholls *et al.*, 1995). Acidic (negatively) charged groups are denoted in blue and basic groups with positive charge are shown in red. Figure B is a 180° clockwise rotation of figure A, showing the opposite side of the molecule.



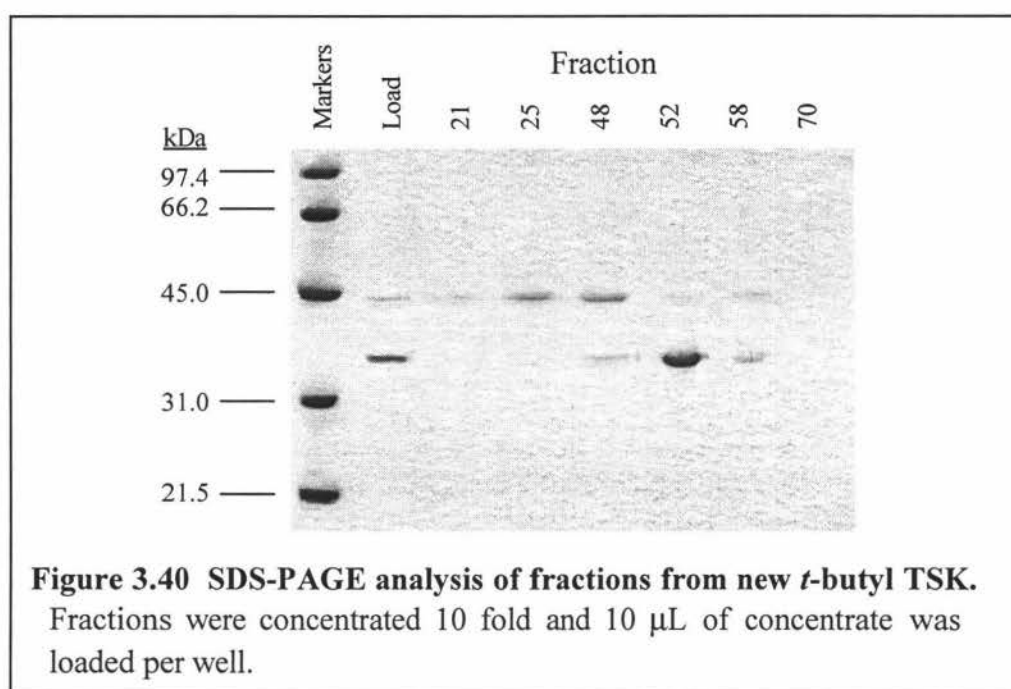
The separation of the 44 kDa band from the native PNGase F enzyme could not be achieved by using various commercial HIC or IEC adsorbants which warranted the synthesis of new *t*-butyl TSK resin in an attempt to reproduce the original substitution and hence separation of the proteins.

#### 3.4.4.5 HIC with Newly Synthesised *t*-butyl TSK

A new *t*-butyl TSK was synthesised following a protocol described in appendix 6. This resin was used to purify PNGase F using the same buffering system as used with the old TSK column described in section 3.3.4.1.2. Figure 3.39 shows two fused peaks. SDS-PAGE analysis revealed that the first peak contained the majority of the 44 kDa band and the second peak was the enzyme with a trace of the 44 kDa contaminant (figure 3.40). This was not an unexpected result, as at the time of resin synthesis only half of the needed CDI cross-linker was available thus the substitution was not as complete as the old TSK.



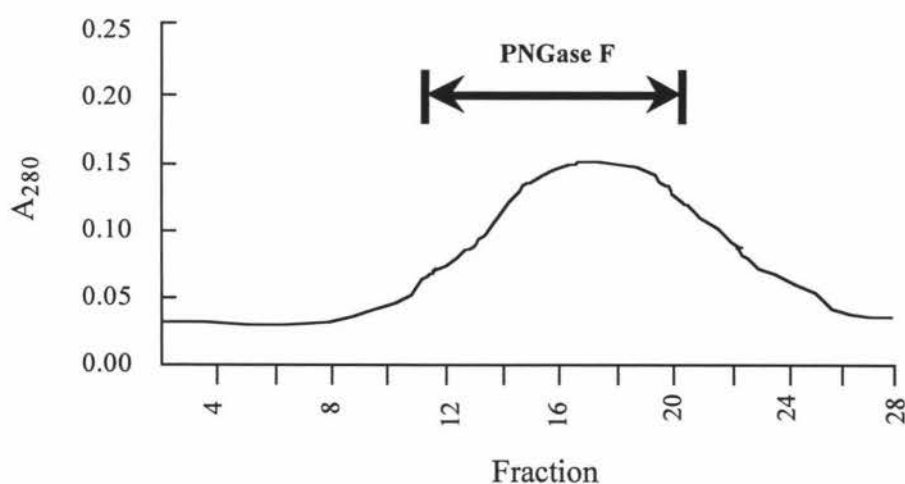
**Figure 3.39** Chromatogram of elution from new *t*-butyl substituted TSK. Arrows indicate active fractions and the dashed line represents the NaCl gradient.



#### 3.4.4.6 SEC with Superdex 75

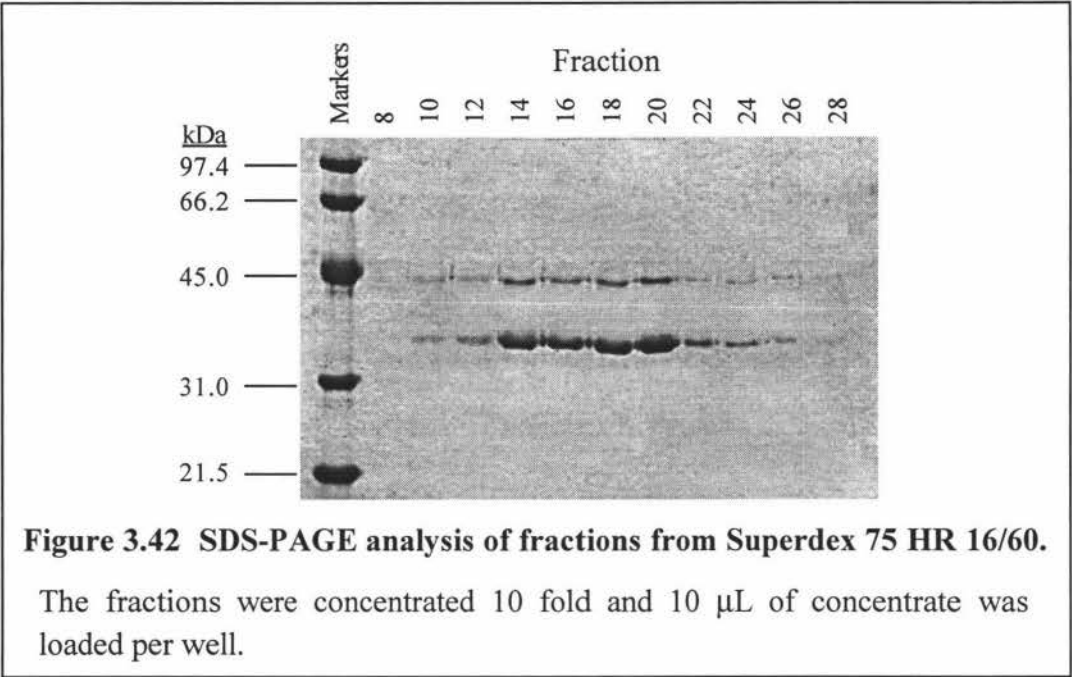
##### A. *Superdex 75 HR 16/60 Prep-Grade Resin.*

The prep-grade column was used initially due to the larger capacity of the column compared to the HR10/30 to reduce the number of runs. However, the column failed to separate PNGase F from the 44 kDa band as shown in figure 3.41. SDS-PAGE analysis of the collected fractions confirmed that no protein separation had occurred (figure 3.42).



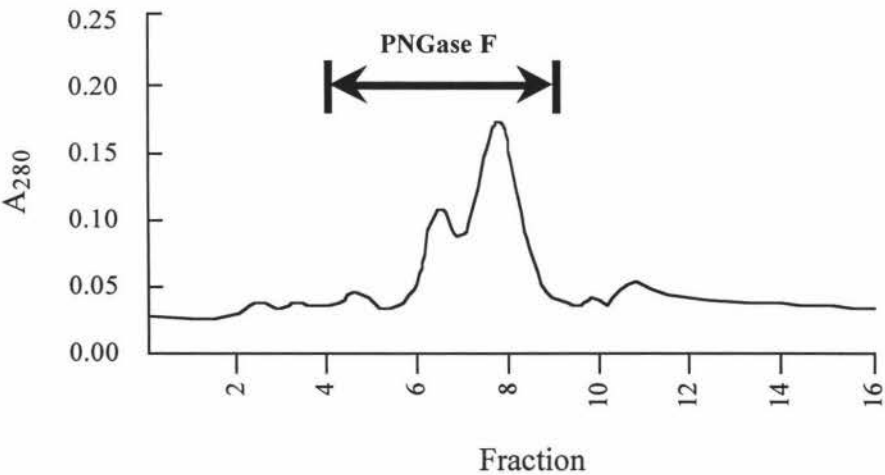
**Figure 3.41 Chromatogram of elution from Superdex 75 HR 16/60.**

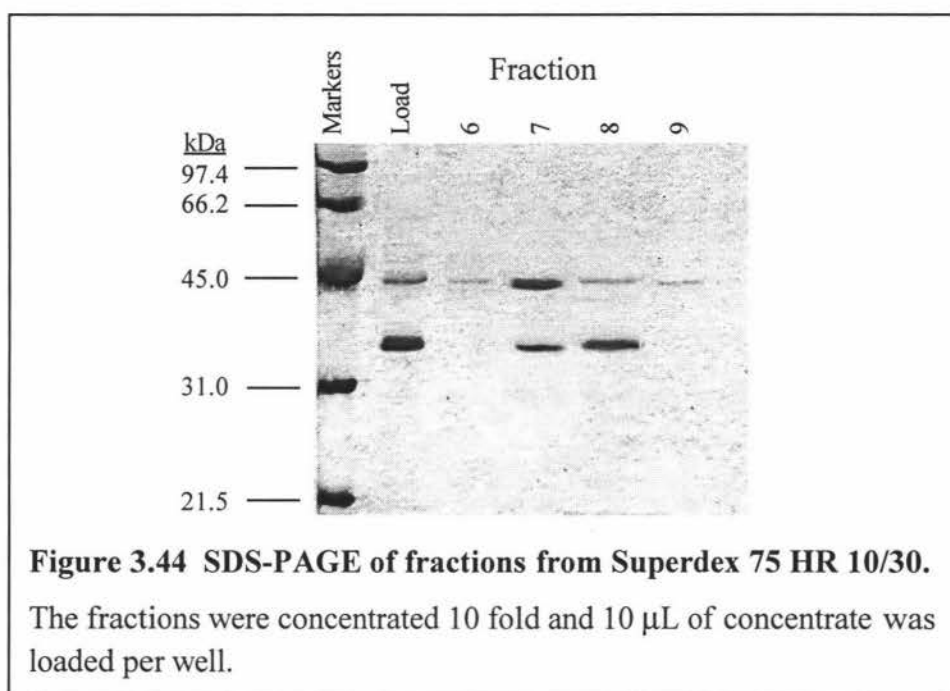
Arrows indicate active fractions.



*B. Superdex 75 HR 10/30 Super Fine Resin.*

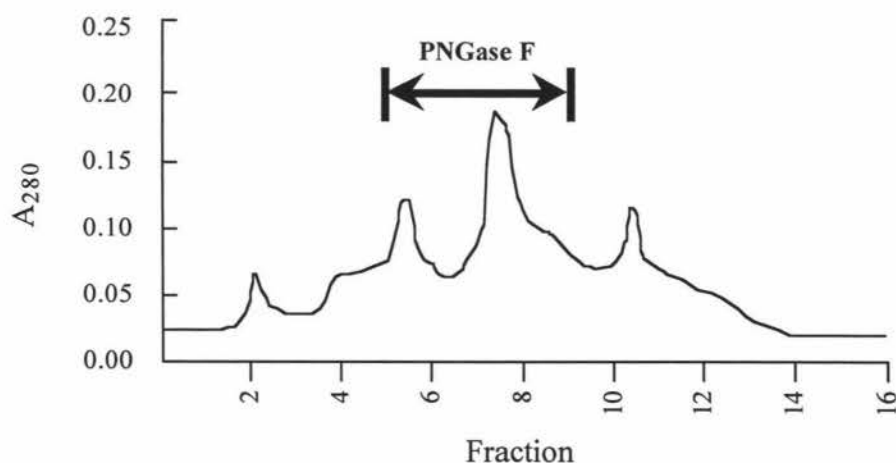
Two unresolved peaks were eluted in the run (figure 3.43) and SDS-PAGE shows a moderate separation of the contaminant from PNGase F (figure 3.44). It was decided to enhance this separation by adding a small amount (0.006%) of Thesit to the buffer to disrupt any potential hydrophobic interaction with the resin. Thesit is a gentle non-ionic detergent that had been shown not to denature the enzyme, is UV transparent and has a relatively high critical micelle concentration.





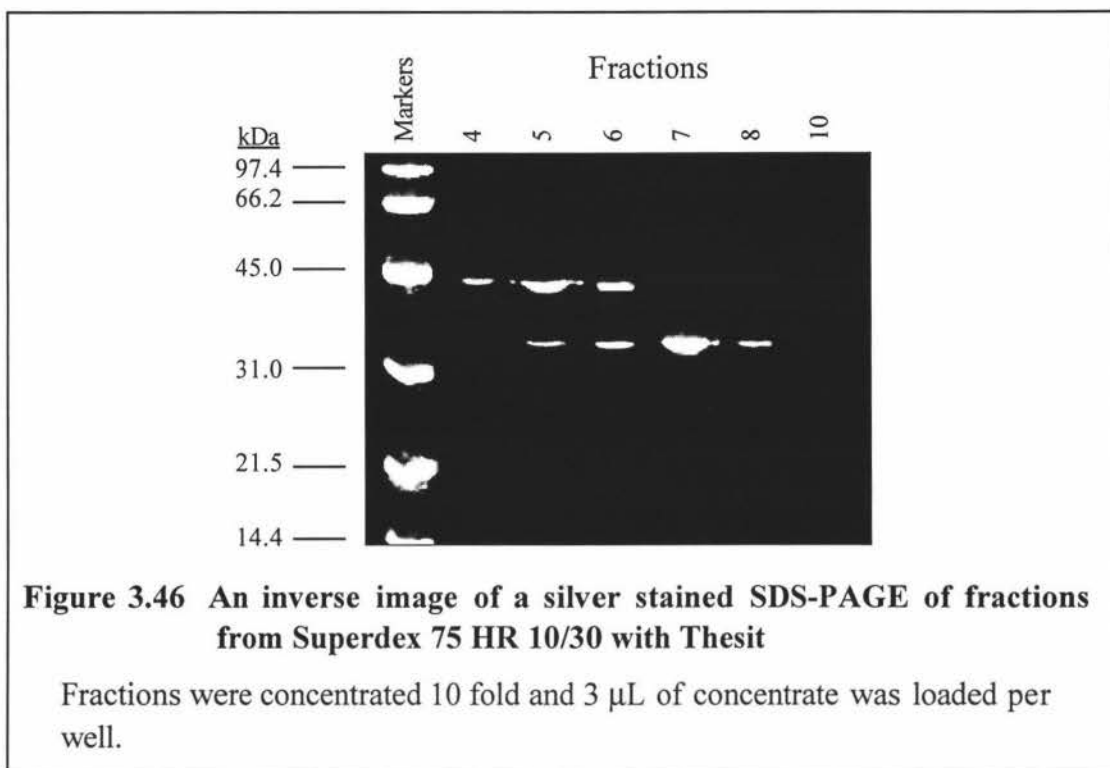
**C. Superdex 75 HR 10/30 Super Fine Resin in the presence of Thesit.**

The addition of 0.006% Thesit to the buffer markedly changed the elution profile of the sample, the two unresolved peaks from the previous run separated into four distinct peaks (figure 3.45). SDS-PAGE analysis of the active fractions shows that PNGase F in fractions 7 and 8 is free of the contamination (figure 3.46). These samples were pooled for biochemical studies. The peak eluted later in the run is possibly Thesit because no protein bands could be detected in that fraction using SDS-PAGE.



**Figure 3.45 Chromatogram of elution from Superdex 75 HR 10/30 with Thesit.**

Arrows indicate active fractions.



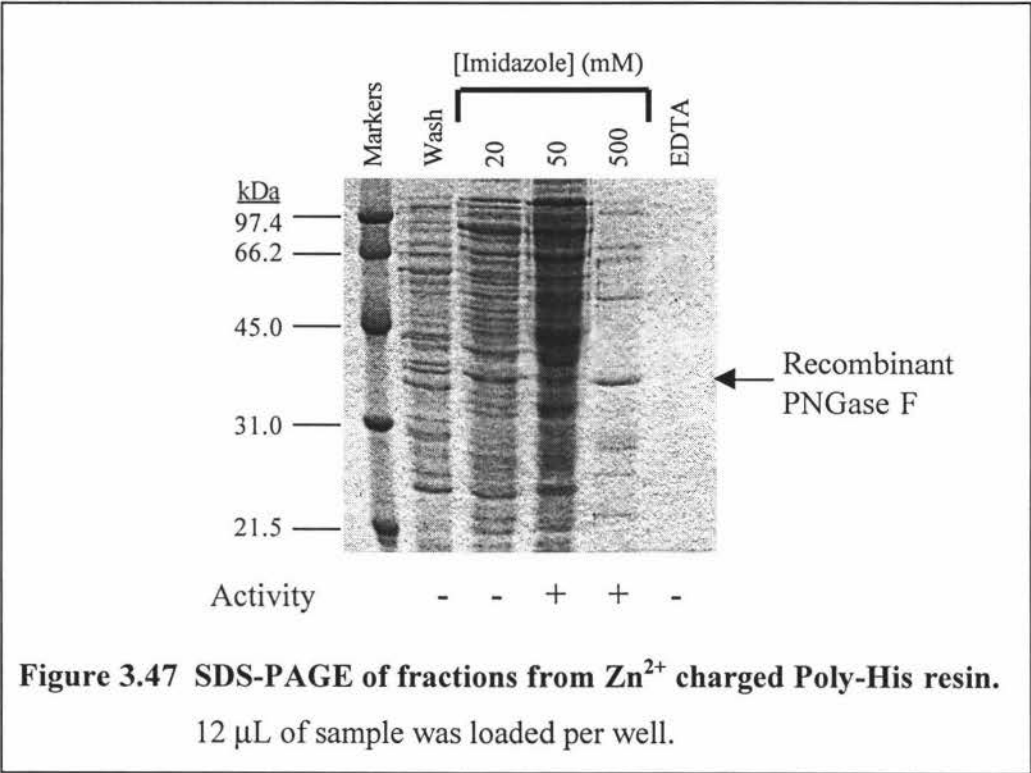
### 3.4.5 Purification of Recombinant PNGase F

In the initial stages in the purification of the hexa-histidine tagged recombinant PNGase F, efforts were focused on determining which metal ion ( $\text{Ni}^{2+}$  or  $\text{Zn}^{2+}$ ) and commercially available metal chelating matrix (Boehringer Mannheim or Pharmacia) combination gave the cleanest separation. The periplasmic fraction containing PNGase F was dialysed against the loading buffer before being applied to the column and unbound proteins were removed by extensive washing with loading buffer. The bound proteins were eluted in three steps of increasing imidazole concentrations (20, 50, 500 mM) in the elution buffer. Ten column volumes of buffer were applied at each elution step and the column was regenerated by removal of the immobilised metal ions with stripping buffer containing 100 mM EDTA.

3.4.5.1 IMAC Test of Various Chelating Resins

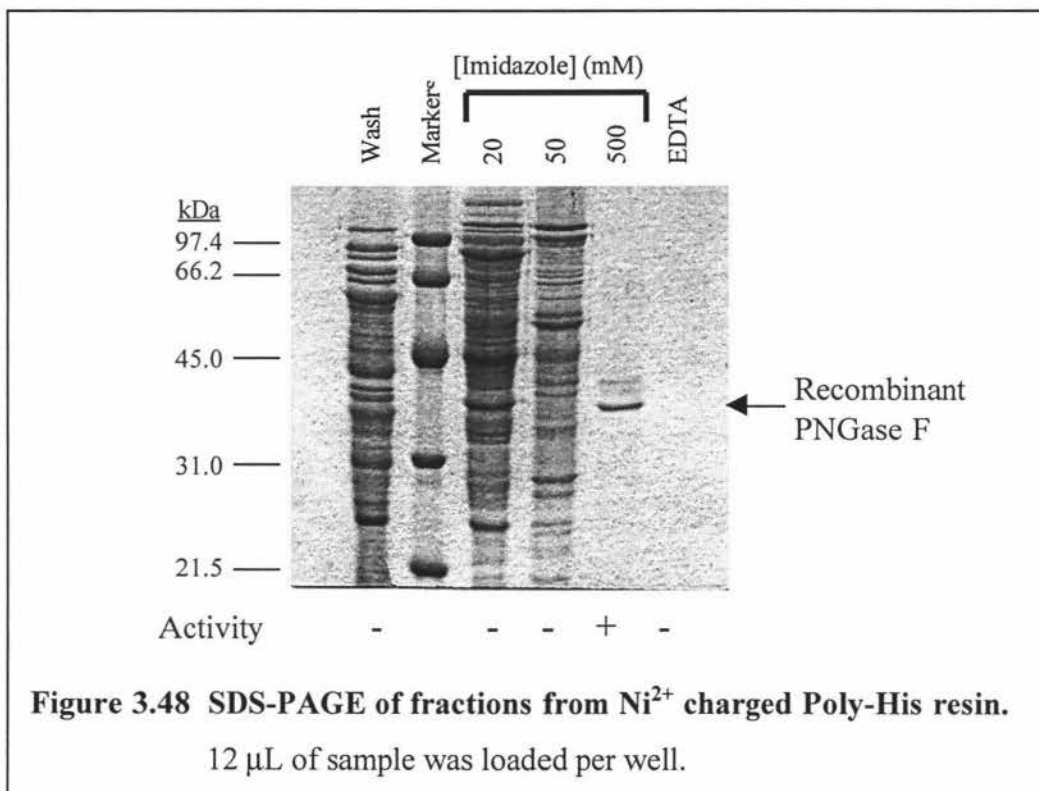
A. *Boehringer Mannheim Poly-His Purification Resin Chelated with Zn<sup>2+</sup>*

Zn<sup>2+</sup> is the metal ion recommended by the manufacturer for use with this resin. SDS-PAGE analysis shows however, that the final elution step containing the recombinant protein also contains several other protein bands (figure 3.47). PNGase F activity was also detected in the 50 mM step indicating weak binding.



B. *Boehringer Mannheim Poly-His Purification Resin Chelated with Ni<sup>2+</sup>*

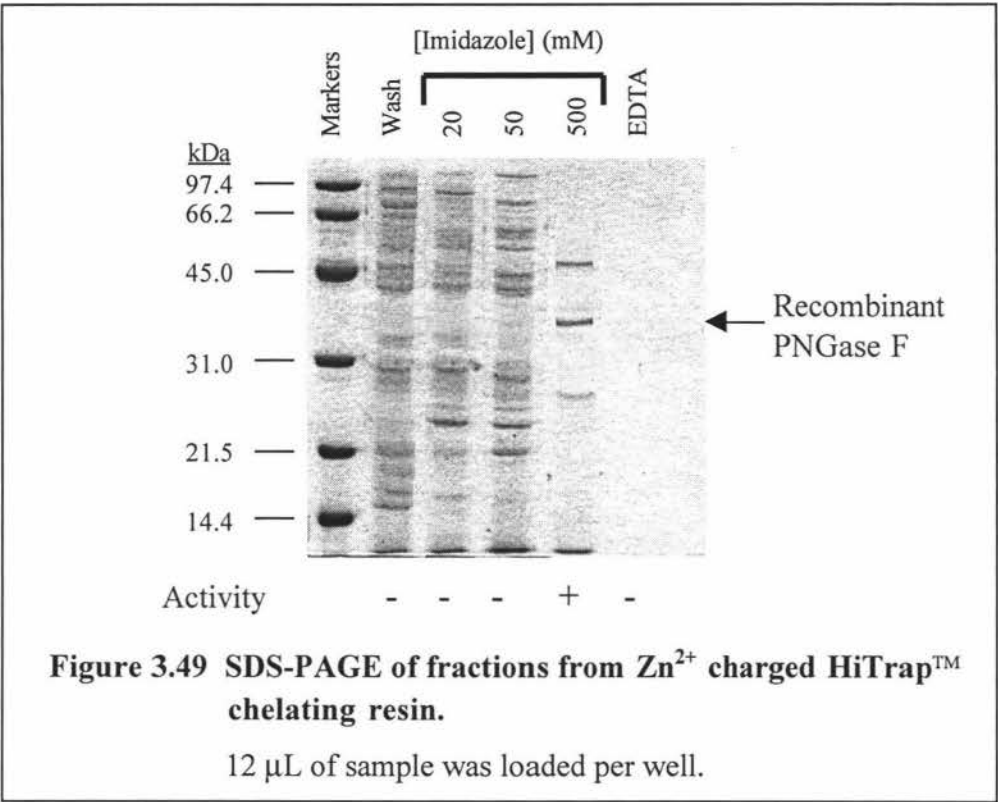
Deglycosylation activity of PNGase F was detected in the final step and SDS-PAGE showed that two weak contaminants were present (figure 3.48). In hindsight, elution buffers containing slightly higher concentrations of imidazole (100, 200 mM) should have been used in an attempt to remove the contaminants. At the time of the experiment, however, the trial runs with the Pharmacia chelating resin that followed next gave better results than this.



### C. Pharmacia HiTrap™ Chelating Resin Chelated with $\text{Zn}^{2+}$

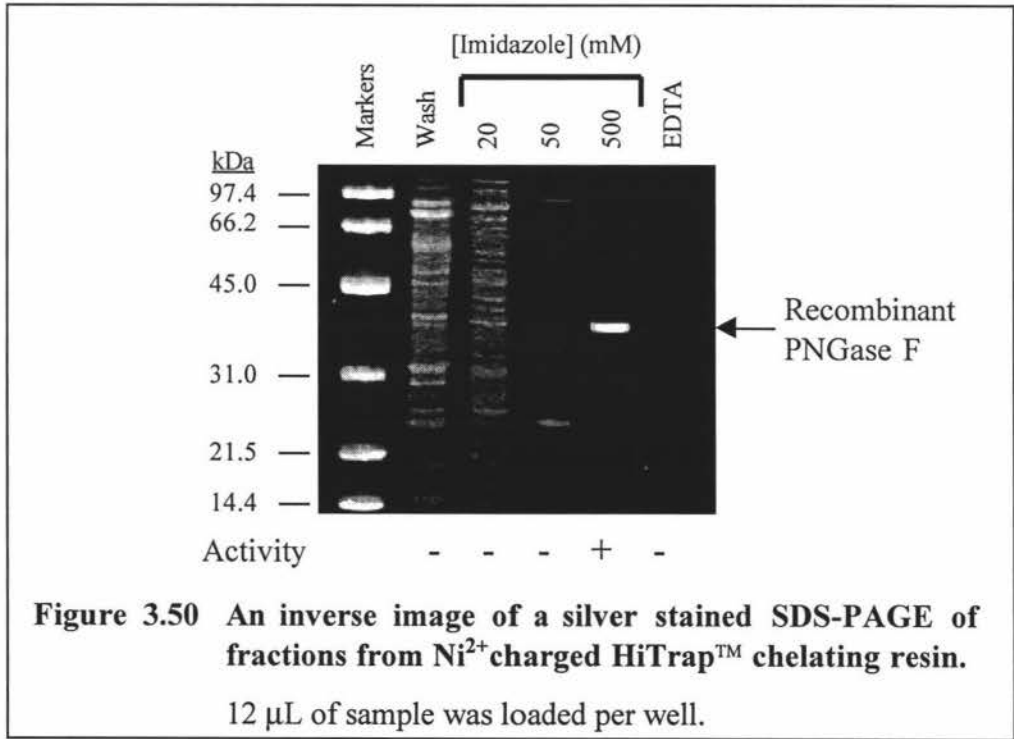
The pre-packed 1 ml HiTrap™ metal chelating columns with a binding capacity of about 5 mg of His-tagged protein were used in the initial trial runs. No bands corresponding to the molecular weight of the recombinant protein were detected in the two low imidazole concentration steps (figure 3.49). The final elution step contains PNGase F activity. Although there are several contaminating bands present, these can be removed by increasing the number of steps in imidazole concentration in the washing procedure before the final imidazole concentration of 500 mM is reached.





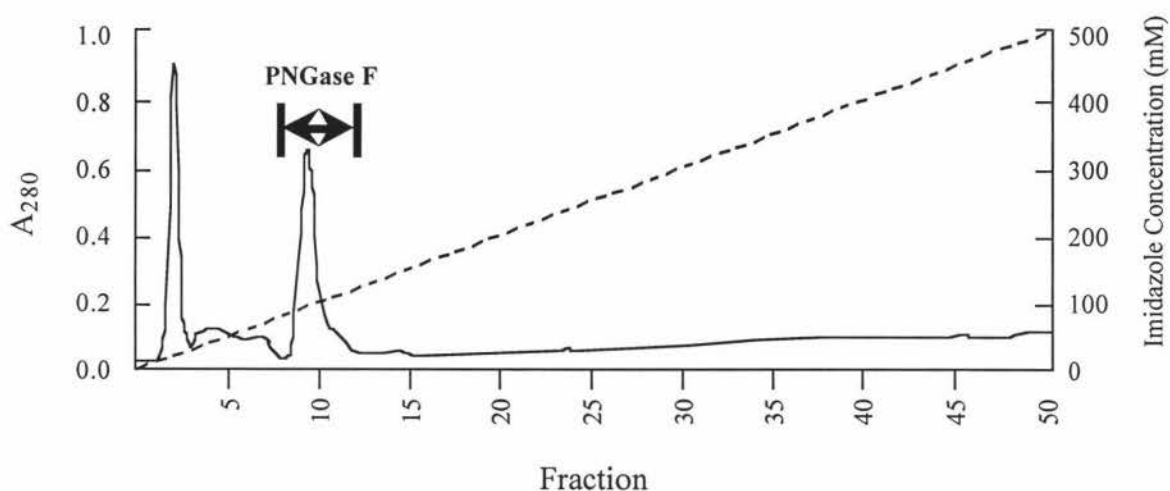
D. Pharmacia HiTrap™ Chelating Resin Chelated with  $\text{Ni}^{2+}$

One single band was detected in the final elution step (figure 3.50) and the activity assay showed the fraction contained PNGase F activity.



### E. Elution Profile of Recombinant PNGase F from $\text{Ni}^{2+}$ Charged HiTrap™ Chelating Resin

To find the concentration of imidazole required to displace recombinant PNGase F from the resin, a pre-packed 1 ml HiTrap™ column was connected to a Bio-Rad Econo System running a gradient from 0 to 500 over 500 minutes. PNGase F eluted from the resin around 85 mM imidazole indicating a fairly weak association possibly due to steric hindrance of C-terminal His-tag (figure 3.51). The gradual increase in the absorbance at 280 nm later in the gradient is the result of increasing imidazole concentration in the mobile phase.

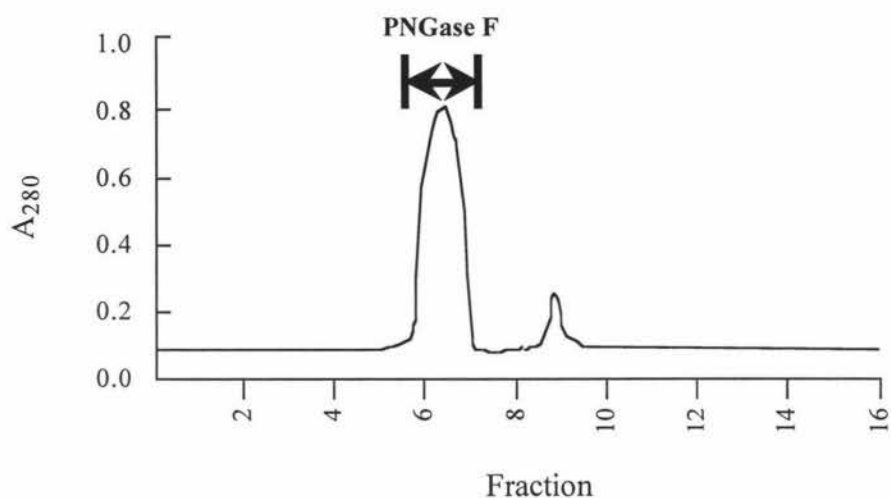


**Figure 3.51** Elution profile from  $\text{Ni}^{2+}$  charged Pharmacia chelating resin.

Arrows indicate active fractions and the dashed line indicates the imidazole concentration in mM. Unbound proteins were removed by washing with loading buffer before the gradient was initiated.

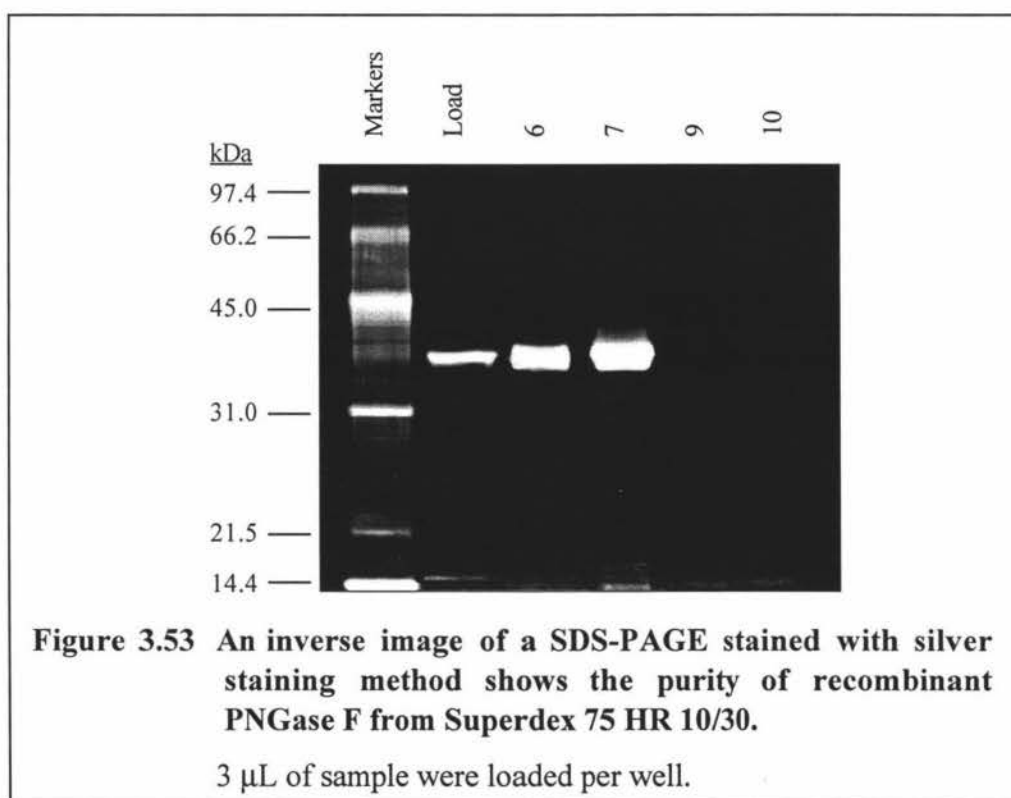
#### 3.4.5.2 SEC of Recombinant Enzyme with Superdex 75 in the presence of 0.006% Thesit

Figure 3.52 shows that two peaks of greatly differing magnitude were eluted and SDS-PAGE analysis with silver staining showed that the first and much larger peak contained pure recombinant PNGase F (figure 3.53). The smaller second peak contained no visible bands and had no activity.



**Figure 3.52** Chromatogram of elution from Superdex 75 HR 10/30.

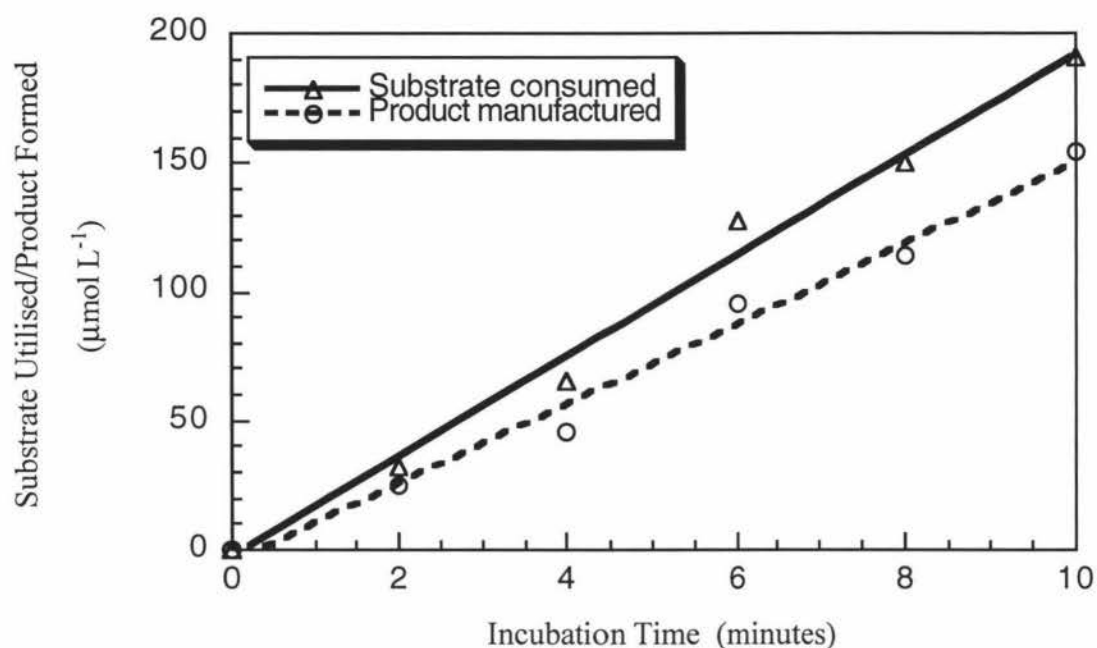
Arrows indicate active fractions.



**Figure 3.53** An inverse image of a SDS-PAGE stained with silver staining method shows the purity of recombinant PNGase F from Superdex 75 HR 10/30.

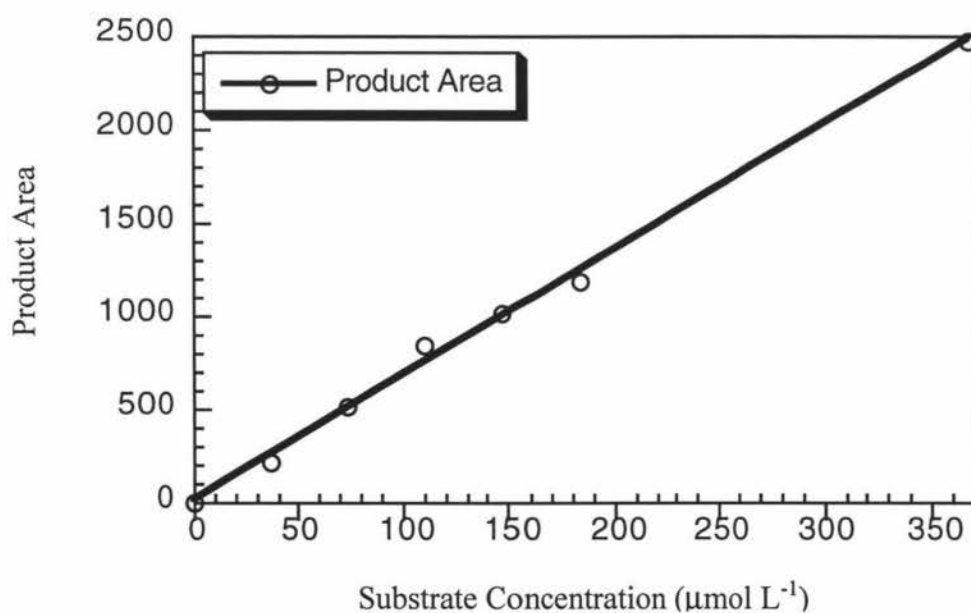
3 µL of sample were loaded per well.

The calculation of the rate of deglycosylation in the activity assay is based on the assumption that substrate depletion is equal to product formation and this reciprocal relationship is shown in figure 3.54. Although the product detected is lower than that of the substrate, the relationship is fairly linear and consistent. A typical peak integration report of an assay is shown in appendix 7.



**Figure 3.54** The reciprocal relationship of substrate depletion and product formation in the activity assay.

Figure 3.55 shows the relationship between the product area and substrate concentration. As the relationship is linear, this plot was used as a standard curve for determination of product formed in an assay from product area, based on the assumption that the substrate utilisation is equal to production formation.



**Figure 3.55** The linear relationship of product area and substrate concentration in the activity assay.

The yields from the purification of native and recombinant PNGase F are shown in Table 3.4. Native PNGase F was purified to homogeneity from *F. meningosepticum* culture to approximately  $0.15 \text{ mg L}^{-1}$  of protein. The recombinant enzyme has a yield of about  $3.2 \text{ mg L}^{-1}$  of pure protein and the preparation suffered massive losses in ultrafiltration of the pooled fractions despite pre-wash of the stirred cell with Thesit and the inclusion of 0.006% Thesit in the samples. The difference in the degree of protein loss between native and recombinant possibly depends on the ultrafiltration membrane. The membrane used for the native PNGase F was utilised in a previous purification of the same enzyme and thus most adsorption surfaces were coated with protein whereas for the recombinant preparation the membrane was brand new. The stirred cell was thoroughly washed before the first ultrafiltration and was used without being cleaned throughout the purification procedure in order to minimise protein adsorption. The protein loss decreased in the second ultrafiltration of the solution by 12% presumably due to the interior surface of the stirred cell and membrane becoming “coated” with protein. No precipitate of protein was observed in any of the concentration steps and if indeed present, it should have redissolved as minimal volumes of the filtrate were used to rinse the stirred cell three times and pooled. Pre-treatment of a new membrane with proteins such as BSA before use is undesirable because of the possibility of the bound BSA slowly desorbing from the membrane during use. The use of Centricons (Amicon Inc.) composed of a polymethylpetene body and a regenerated cellulose membrane improves the recovery of proteins by about 16% compared to the Vivaspin concentrators (Vivascience) that are made of polycarbonate with a polyethersulfone membrane of the same MWCO (results not shown). However, the poor recovery of the Vivaspin concentrators may be attributed to the way the concentrate was retrieved at the end of concentration. By aspirating the concentrate from a narrow well in the body, not all of the sample is collected whereas the concentrate in Centricons are spun down from the holder into a polycarbonate collector. Polyethersulfone membranes exhibit no hydrophobic or hydrophilic interactions with proteins whereas membranes made of regenerated cellulose are highly hydrophilic. This suggests that the non-specific adsorption is mainly of hydrophobic nature.

**Table 3.4 Summary of purification of native and recombinant PNGase F.**

Approximately 0.15 mg L<sup>-1</sup> of native PNGase F was isolated from 8-litres of *F. meningosepticum* culture and about 3.2 mg L<sup>-1</sup> of recombinant PNGase F was purified from 3-litres of transformed *E. coli* BL21(DE3) cells. Summary for native PNGase F is shown first.

Purification Step	Volume (mL)	Total Activity <sup>a</sup> (munits)	Total Protein (mg)	Specific Activity (munits mg <sup>-1</sup> )	Yield <sup>b</sup> (%)	Purification <sup>c</sup> (Folds)
Culture Medium	8000	292.28	8796.14	0.03	100	1
Phenyl Sepharose	596	103.50	61.72	1.68	35.4	51
Stirred Cell Concentrate	4.6	ND	53.08	-	-	-
<i>t</i> -butyl TSK	42	39.29	2.43	16.17	38.0	10
Stirred Cell Concentrate	1.5	ND	2.28	-	-	-
Superdex 75	17.3	28.36	1.36	20.85	72.2	-
Stirred Cell Concentrate	0.2	25.13	1.21	20.77	-	-
Periplasmic Fraction	138	157.12	70.65	2.22	100	-
IMAC	62	135.42	24.01	5.64	86.2	3
Stirred Cell Concentrate	1.1	ND	13.98	-	-	-
Superdex 75	13	76.67	13.74	5.58	56.7	-
Stirred Cell Concentrate	0.3	52.76	9.54	5.53	-	-

<sup>a</sup>1 munit of PNGase F is defined as 1  $\mu$ mol of ovalbumin glycopeptide hydrolysed per minute at pH 8.0 and 37°C.

<sup>b</sup>Yield was calculated by dividing the total units of the current step by units from the first step and expressed as percentage.

<sup>c</sup>Folds of purification was calculated by the specific activity of the current step by specific activity from the previous step.

To concentrate recombinant PNGase F, the volume of the final elution was minimised by using six times the concentration of imidazole (500 mM) required to elute the protein (85 mM). Increasing the salt concentration in the buffer from 0.5 to 0.75 M NaCl reduced the extent of non-specific binding of contaminating proteins. Alternative methods of concentration were tried in an effort to reduce the loss of protein. These included osmosis, lyophilisation and ammonium sulfate precipitation:

- Osmosis involved placing the pure sample in a treated 3-kDa MWCO dialysis tubing and drawing out the liquid by immersion in a stirring solution of 30 % (w/v) polyethylene glycol 20,000. However, this method has similar extent of protein loss to ultrafiltration with a new membrane (~40%) possibly due to surface adsorption to the interior of the dialysis tubing (results not shown).
- Lyophilisation involved snap freezing the pure protein solution in glass vials with liquid nitrogen and freeze drying overnight. This method retained 87% biological activity of the enzyme and 23% of the protein was lost through adsorption to the inner surface of the vials (results not shown). The biological activity of PNGase F may be better preserved by adding a disaccharide  $\alpha,\alpha$ -trehalose that has been reported to stabilise proteins during lyophilisation and long-term storage at ambient temperatures (Scopes, 1994).
- Ammonium sulfate precipitation involved adding solid ammonium sulfate to 80% saturation (561 g L<sup>-1</sup>) and equilibrating at 4°C overnight before centrifugation. The precipitate was redissolved in minimal volume of a buffer and dialysed extensively against the same buffer to remove the ammonium sulfate. Precipitation resulted in a fairly low degree of concentration (10 fold) and a fairly significant protein loss of about 28% (results not shown).



- Another method is low temperature (4°C) boiling of the protein solution with stirring under vacuum. This was not attempted because of foaming of the solution and the potential denaturation of the proteins.

To confirm the eight amino acid changes in the New Zealand variant (CDC strain 3352) of PNGase F are not PCR errors, both purified samples of native and recombinant proteins were analysed by ElectroSpray Mass Spectrometry (ES-MS) to determine accurate masses of the proteins. These masses are then used to compare with the molecular weight calculated from nucleic acid sequences. Despite a lot of different sample preparations and using all known techniques in ESI-MS, no mass near the calculated masses of both proteins were found even though the samples can be seen clearly and pure on a Coomassie Blue stained SDS-PAGE (results not shown). Several of these preparations were also submitted for Matrix Assisted Laser Desorption Ionisation with Time Of Flight (MALDI-TOF) Mass Spectrometry and again no near masses were obtained. However, a reverse-phase HPLC preparation of a sample of native PNGase F stored at 4°C for two years gave a mass of 34866 Da. This corresponds to the predicted mass plus 30 Da, and may correspond to the two methionine residues in the sequence being oxidised. It is possible to convert the suspected methionines on PNGase F to the reduced form (Cullwell, 1987; Houghton and Li, 1979), but the procedure requires milligram quantities of material which we do not have. This leaves the option of sequencing the protein using micro-Edman or mass spectrometry. Both methods require reduction, alkylation and digestion of the protein and separation of the cleavage products by reverse-phase HPLC. A protocol was devised but was not carried out in this study due to the lack of material.

### 3.5 Conclusion

The purification of both native and recombinant PNGase Fs are described in this chapter. Several unusual results were encountered during purification of the native protein:

- (1) No protein was recovered from the Alkyl Sepharose column despite using 50% (v/v) ethylene glycol or even 0.1 *N* NaOH to remove it.
- (2) No binding of the native protein to both anionic and cationic exchangers under conditions that should have promoted binding because the charged groups on the surface of the protein are evenly distributed and cancel each other out.
- (3) Superfine Superdex 10/30 with a small amount of non-ionic detergent in the buffer quite efficiently separated the two proteins. This confirms the fact that the two proteins probably aggregated through hydrophobic interactions, and explains why hydrophobic interaction chromatography failed to separate the two proteins.
- (4) There is a significant loss of PNGase F during the ultrafiltration and dialysis steps due to interactions with the plastic/cellulose material making up the concentrators/membranes. From these experiments, lyophilisation would appear to be the best method of concentration. However, lyophilisation was not used for this preparation because the enzyme had to be fully active for the biochemical characterisation studies.

The eight amino acid changes in the PNGase F used for this work, (CDC strain 3352), cannot be confirmed in this study due to the difficulties mentioned in the discussion. Efforts to purify sufficient quantity of native PNGase F for cleavage, fragment purification and identification, and subsequent confirmation of sequence changes are in place and the findings will be reported in the near future.

## Chapter 4 Characterisation of Recombinant PNGase F

### 4.1 Introduction

The biochemical properties of native PNGase F are well characterised as mentioned in the introduction. The pH optimum for deglycosylation activity is 8.5 and the enzyme is compatible with a wide variety of inorganic and organic buffers ( $\leq 0.1$  M). The activity is not inhibited by high concentrations of chelating agents such as EDTA or by the serine protease inhibitor PMSF (Marley *et al.*, 1989). The enzyme is stable for approximately two years at 4°C and indefinitely at -70°C with 50% (v/v) glycerol, but the activity decreases by repeated freezing and thawing.

### 4.2 Experimental Objectives

In chapter three, a high yield *E. coli* expression system was developed and the resulting recombinant protein was pure. There were however two questions that needed to be addressed:

- (1) Is the substrate specificity of the recombinant enzyme the same as that of the native enzyme?
- (2) Is the rate of deglycosylation similar to that of the native enzyme?

This chapter reports the kinetic characterisation of recombinant PNGase F. Thus, rates of deglycosylation of the ovalbumin derived glycopeptide using the discontinuous assay described in chapter 3 were compared for the recombinant and native enzymes under a variety of conditions. In order to compare their activities with glycoproteins as substrates, SDS-PAGE was used to monitor the ability of both enzymes to remove the glycans from a number of glycoproteins that are known to possess different glycans.

### 4.3 Materials

The materials used were of the highest grade available in the laboratory.

Ajax Chemicals (Sydney, Australia):

Ammonium bicarbonate, calcium sulfate, potassium chloride, and sodium chloride.

BDH Laboratory Supplies (Poole, England):

Zinc chloride and dithiothreitol (DTT).

Boehringer Mannheim (West Germany):

Leupeptin (hydrogen sulfate salt).

Reidel-de Haen (Germany):

Di-sodium ethylene diamine tetraacetic acid (EDTA), and magnesium sulfate.

May and Baker Limited (Dagenham, England):

Manganese sulfate, and nickel chloride.

Millipore Corporation (MA, USA):

Nitrocellulose (Type MF)

Roche (West Germany):

DIG glycan/protein double labelling kit (a discontinued product).

Serva Feinbiochemica GmbH & Co. (Heidelberg, West Germany):

Phenylmethylsulfonyl fluoride (PMSF).

Sigma (St. Louis, USA):

Fetuin, chicken egg albumin, ribonuclease B, L-Glutamic dehydrogenase (Type III),  $\alpha$ -ketoglutaric acid (disodium salt), reduced form of  $\alpha$ -nicotinamide adenine dinucleotide phosphate (NADPH), and pepstatin A.

## 4.4 Methods Used in Characterisation

### 4.4.1 Comparison of Substrate Specificity

To address the first question of substrate specificity of the recombinant PNGase F, three well-characterised glycoproteins with different N-linked glycan were used to compare the deglycosylation activity of the recombinant protein to the native enzyme. Fetuin contains sialylated triantennary N-linked complex sugar chains, chicken egg albumin has a single biantennary N-linked complex glycan, and ribonuclease B has multiple N-linked high mannose biantennary glycans. In the initial stages of comparison, all substrates were made up to  $1 \text{ mg mL}^{-1}$  in Milli-Q water, except fetuin. Because of its multiple glycoforms, a concentration of  $2 \text{ mg mL}^{-1}$  fetuin was required to visualise all the bands on an SDS-polyacrylamide gel clearly. Native and recombinant PNGase F ( $0.4 \text{ mg mL}^{-1}$ ) were incubated with an equal volume of substrate at  $37^\circ\text{C}$  to a final volume of  $500 \mu\text{L}$ . Aliquots were taken at 24-hour intervals, analysed by separation on a 12% polyacrylamide gel and stained with Coomassie Brilliant Blue R<sub>250</sub> as described in section 3.3.1.4. The decreases in molecular weight due to the loss of glycan chains from the substrates by the action of PNGase F were visualised as band shifts on SDS-PAGE.

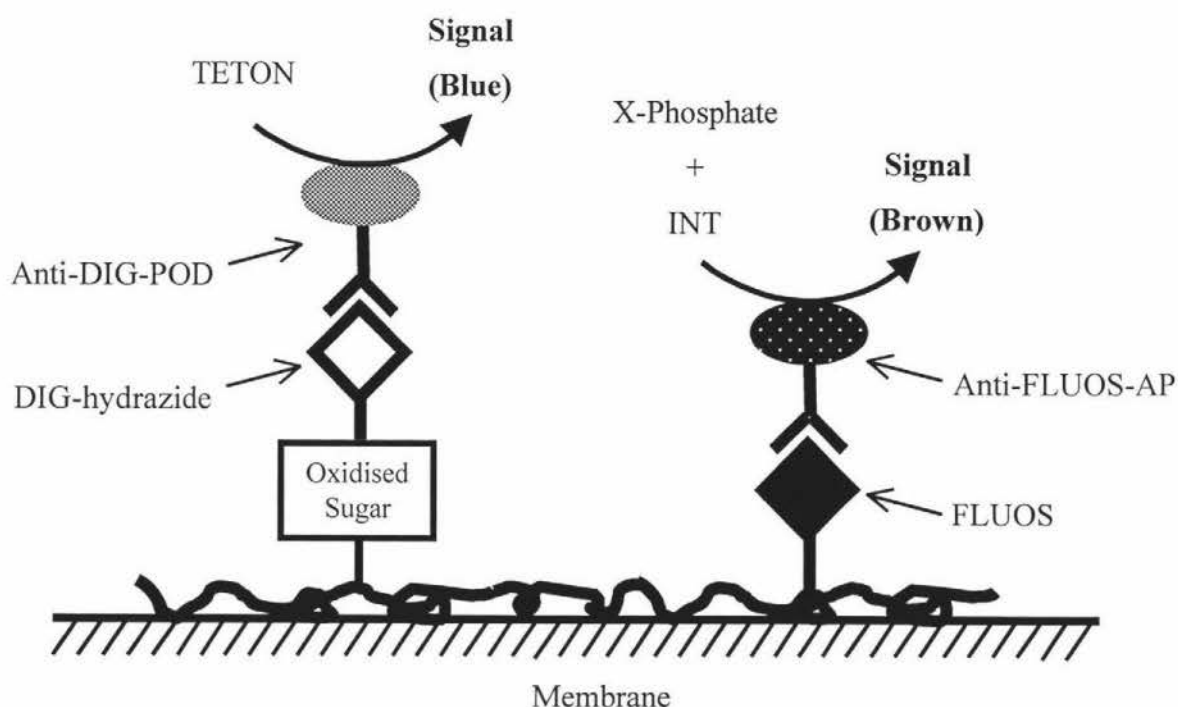
In the initial stages of substrate specificity comparison, efforts were focused on determining the optimum for recombinant enzyme activity. Fetuin was chosen for these experiments because it was rapidly deglycosylated and a characteristic “ladder” representing the partially deglycosylated products was seen on a polyacrylamide gel. The reaction mixture contained 5 mM buffer (10 mM EPPS or  $\text{Na}_2\text{HPO}_4$ ),  $0.2 \text{ mg mL}^{-1}$  recombinant PNGase F and  $1 \text{ mg mL}^{-1}$  fetuin. The mixture was incubated at  $37^\circ\text{C}$  for 48 hours and analysed by SDS-PAGE. In the later stages of comparison, the substrate contained Complete™ protease inhibitor (Roche) to halt any protease activity that may have been present in the enzyme preparation.

#### 4.4.1.1 Detection of Deglycosylated Substrate with Digoxigenin (DIG) Glycan/Protein Double Labelling Kit

The Digoxigenin glycan/ protein double labelling kit (Roche), referred to as the DIG kit in this study, can be used to detect glycosylated and deglycosylated proteins on polyacrylamide gels and on blots. Proteins that had been incubated with both recombinant and native PNGase F were first analysed by SDS-PAGE. The separated proteins were then first immobilised onto a nitrocellulose membrane (Millipore) before glycan/protein detection was performed. The kit has a detection limit of about 50 ng for proteins and can detect glycoproteins at concentrations of 10 ng and upwards. The principle of detection is discussed below.

##### 4.4.1.1.1 Principle of Detection

The sugar and protein portions of a glycoprotein are individually labelled with a different hapten that is then detected with the corresponding antibody coupled to a different reporter enzyme. A schematic representation of the detection principle is shown in figure 4.1. The labelling of the sugar side chains involves specific oxidation of the sugar hydroxyl groups with periodate to form aldehydes followed by covalent coupling of the steroid hapten, DIG-hydrazide, to the aldehyde groups created. The DIG-hydrazide is specifically detected in a reaction with anti-digoxigenin-POD antibody conjugate that contains peroxidase (POD) as a reporter. The substrate for the peroxidase is TETON (4-Triethylenetrioxo-1-naphtnol) and a blue coloured precipitation is obtained from the reaction indicating the presence of sugar side chains on the electroblotted protein. Similarly, free amino groups on the non-glycosylated or deglycosylated proteins are labelled by fluorescein (FLUOS) and are specifically detected by anti-fluorescein-AP antibody conjugate. The anti-fluorescein-AP conjugate has alkaline phosphatase that uses 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyl-tetrazolium chloride (INT) and X-phosphate as substrates to give a reddish-brown precipitation indicating non-glycosylated or deglycosylated protein is present.



**Figure 4.1** Schematic representation of the detection principle of DIG glycan/protein double labelling kit

#### 4.4.2 Characterisation of Recombinant PNGase F

In this section, the effects of pH, temperature, metal ions, storage period, and other additives on the deglycosylation activity of recombinant enzyme are reported. The rate of deglycosylation of an ovalbumin-derived glycopeptide using the discontinuous HPLC assay (section 3.3.1.3) was used to test these effects. Metal ions, buffer salts and other reagents that may have an effect on enzyme activity even at minute levels will interfere with the outcome of such investigations. In order to standardise conditions, attempts were made to remove any associated ions so that any differences in activity observed for a particular additive could be reliably assumed to be due to that additive alone. Dialysis tubing was boiled in a solution of 100 mM ammonium bicarbonate and 10 mM EDTA for 15 minutes, then boiled in Milli-Q water for another 15 minutes and stored in 20% ethanol (v/v) in a plastic container. The water used for dialysis, buffers and stock solutions was HPLC-grade double distilled Milli-Q water.

##### 4.4.2.1 Effects of Temperature

The assays in this test were set up as described in section 3.3.1.3.2 except the incubation temperature used were 15, 25, 30, 35, 40, 45, 50, 55, 60, and 70°C instead of 37°C.



#### 4.4.2.2 Effects of pH

In this section, the effects of both the pH and the type of buffer salt on PNGase F activity were investigated. The buffers used were made up as 100 mM stock solutions. In order to negate any possible effects being due to buffer ions rather than pH, the pH ranges were overlapped as shown in Table 4.1. The enzyme solution was dialysed against Milli-Q water to remove traces of EPPS buffer from SEC that may interfere with the investigation. The assay set up was essentially identical to those above but the reaction mixture now containing a final concentration of 10 mM of a listed buffer.

**Table 4.1 List of buffers used in the activity test.**

Buffer	pH
NaAc	4.5, 5.5
Mes	5.5, 6.5
Mops	6.5, 7.5
Taps	7.5, 8.5
Capso	8.5, 9.5
Caps	9.5, 10.5

#### 4.4.2.3 Effects of Metal Ions and Additives

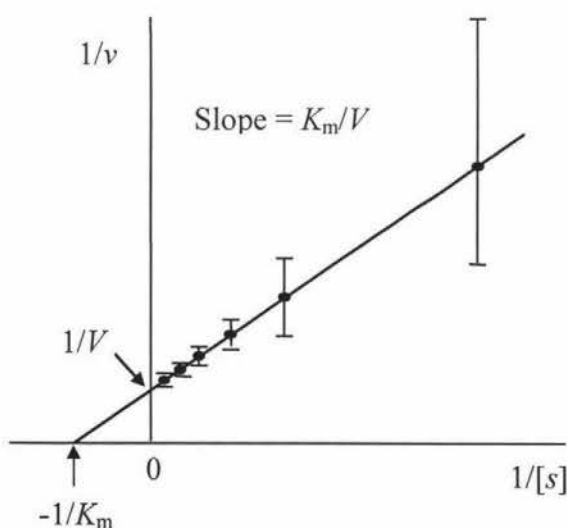
The enzyme was extensively dialysed against Milli-Q water containing 10 mM EDTA with two changes per day for 48 hours in order to completely remove traces of metal ions in the protein preparation. This was followed by dialysis against 5 mM EPPS (pH 8.0) with several changes for 24 hours to eliminate any EDTA present in the preparation. The enzyme was then diluted with the same buffer to a concentration of  $1 \mu\text{g mL}^{-1}$  used in the assay. The metal salts used in this investigation were AgCl, CaCl<sub>2</sub>, CuSO<sub>4</sub>, FeCl<sub>3</sub>, MgCl<sub>2</sub>, MgSO<sub>4</sub>, MnCl<sub>2</sub>, NaCl, NiCl<sub>2</sub>, KCl, and ZnCl<sub>2</sub>. As can be seen, efforts were made to eliminate differences that could be due to the anion. All metal ions solutions were made up in 5 mM EPPS (pH 8.0). The assays were set up as section 3.3.1.3.2 except that the assay mixture contained metal ions at a final concentration of either 5  $\mu\text{M}$  or 5 mM. Other additives included in the assay were SDS, 0.01% (0.3 mM); urea, 2.5 M; EDTA, 5 mM; DTT, 2 mM; NEM, 2 mM; PMSF, 1 mM; leupeptin,  $1 \mu\text{g mL}^{-1}$ ; and pepstatin,  $1 \mu\text{g mL}^{-1}$ . All assays were done in duplicate.

#### 4.4.2.4 Shelf-life

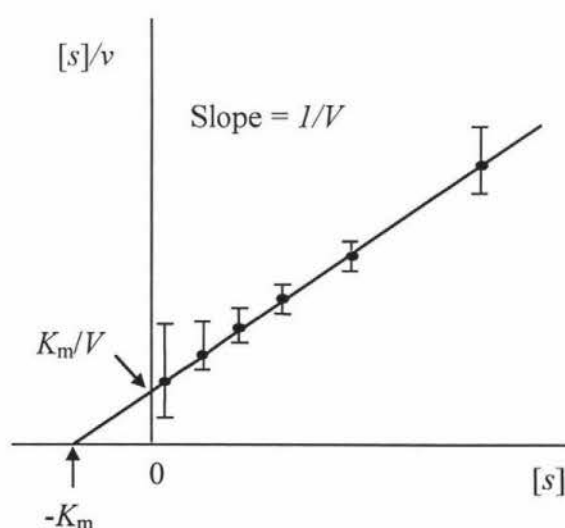
Old stocks of pure recombinant PNGase F stored at 4°C from various points in the project were tested to investigate the effects of long-term storage on the activity of the enzyme.

#### 4.4.3 Determination of the Michaelis Constant ( $K_m$ ) of Native and Recombinant PNGase F

To compare substrate bind and rates of reaction for the recombinant and native enzymes,  $K_m$  and  $V_{max}$  were determined for both enzymes using a modification of the Lineweaver-Burk plot known as the Hanes-Woolf plot. The Hanes-Woolf plot was chosen because the data values in the Lineweaver-Burk plot are unevenly weighted in the analysis (figure 4.2). Maximum weight is placed on data points obtained at very high substrate concentrations, where the highest degree of experimental error is associated. Hence the Lineweaver-Burk method gives greatest weight to the least precise measurements and minimises the scatter appearance of low substrate concentrations. The Hanes-Woolf plot with more evenly chosen substrate concentrations provides a more faithful reflection of the errors in the data (figure 4.3).



**Figure 4.2** Lineweaver-Burk plot with error bars of  $\pm 0.05V$ .



**Figure 4.3** Hanes-Woolf plot with error bars of  $\pm 0.05V$ .

A recent paper by Altman *et al.*, (1995) reported a  $K_m$  value of  $0.3 \text{ mg mL}^{-1}$  substrate for PNGase F using a similar ovalbumin peptide as substrate. The enzyme was serially diluted to  $1 \text{ } \mu\text{g mL}^{-1}$  with 5 mM EPPS (pH 8.0) and the glycopeptide substrate was used at concentrations of 0.25-, 0.5-, 1-, 2-, 4-, 8-  $\times K_m$  (where  $K_m = 0.3 \text{ mg mL}^{-1}$ ) in the same buffer. For measuring deglycosylation rates, a fixed volume of enzyme ( $5 \text{ } \mu\text{L}$ ) was added to the different concentrations of substrate in a total reaction volume of  $95 \text{ } \mu\text{L}$ . The reaction was incubated for 10 minutes at  $37^\circ\text{C}$ . At the end of this time, the reaction was terminated by addition of  $10 \text{ } \mu\text{L}$  formic acid and rapid heating to  $100^\circ\text{C}$  in a boiling water bath. Assays were performed in duplicate and repeated if the peak area discrepancy between the duplicates was not within 5%. A second set of reactions was carried out with more diluted substrate concentrations, five times diluted, to a final reaction volume of  $495 \text{ } \mu\text{L}$ . After 10 minutes of incubation, these samples were snap frozen in liquid nitrogen, and lyophilised to dryness. The sample was then redissolved in  $0.1 \text{ g}$  of 10% formic solution and  $70 \text{ } \mu\text{L}$  of each mixture was injected into the HPLC for analysis.

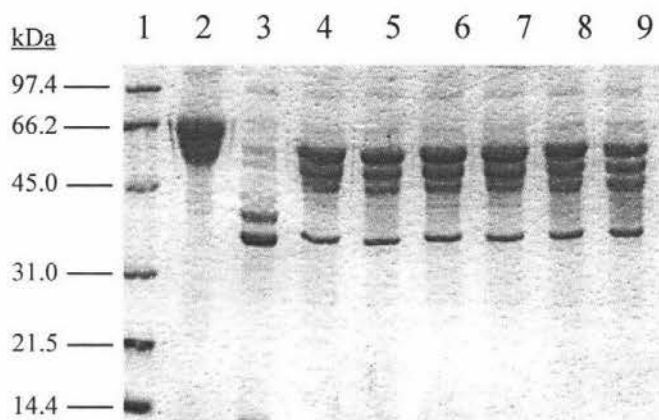
The problem of sample loss through adsorption onto the surface of Eppendorf tubes during lyophilisation was addressed by adding a known amount of internal standard into the reaction mixtures immediately before lyophilisation. The internal standard used was  $40 \text{ } \mu\text{L}$  of a  $1 \text{ mg mL}^{-1}$  dansyl-labelled ovalbumin-derived glycopeptide solution that was made as a substrate for Endo F. This standard has the sequence Dansyl-Y N<sup>CHO</sup> L T S V L Hs (where N<sup>CHO</sup> is the glycosylated asparagine) and absorbs at  $436 \text{ nm}$ . Thus it is not a substrate for PNGase F as the enzyme requires at least two amino acids on either side of the glycosylated asparagine for activity. Because it may act as an inhibitor of PNGase F, it was not added till the completion of incubation. It was chosen because of its similarity to the substrate with the rationale that it would as nearly as possible mimic the adsorption characteristics of the substrate to the Eppendorf tubes used in the lyophilisation process. It is assumed that adsorption of the internal standard is the same as the substrate/product in the reaction, losses due to adsorption can be corrected by a factor calculated from quantitation of the peak because of this glycopeptide. The correction factor was calculated by dividing peak area of the internal standard before lyophilisation by the area of standard recovered.

These experiments established the range of substrate concentrations that could be practically used to determine  $K_m$ . A set of five concentrations was chosen representing 3.75 to 120  $\mu\text{g}$  of substrate in total volume of 495  $\mu\text{L}$ . Next, the optimum incubation time for the assay was established by stopping the reactions at 1-minute intervals between 1 and 10 minutes. For discontinued assays, it is crucial that the reaction is first order, which means that hydrolysis should be between 5 and 10% completion. The rate of each substrate concentration was calculated from these results and was analysed using a Hanes plot to obtain the  $K_m$  value.

## 4.5 Results and Discussion

### 4.5.1 Comparison of Substrate Specificity

Initial comparison of substrate specificity involved determining which buffer (10 mM EPPS or  $\text{Na}_2\text{PO}_4$ ) at which pH (7, 8, and 9) gave the optimum conditions for recombinant enzyme activity on fetuin. The result of these experiments is shown in figure 4.4. An interesting result is shown lane 2 where it would appear that recombinant PNGase F has completely deglycosylated fetuin in the absence of any buffer ions or additives. Fetuin has 3 N- and 3 O-linked glycans and the result of the gel appeared to show that both types of glycans could be removed by the recombinant enzyme. The molecular weight of the deglycosylated protein appeared to be equivalent to the molecular weight of the polypeptide chain of fetuin, calculated to be 38,418 Da. There are two possible explanations for the observation: Firstly, the unlikely possibility that the recombinant PNGase F was able to hydrolyse O-linked glycans, or secondly, the more likely explanation that there were trace amounts of protease in the enzyme preparation. This result was reproducible with new substrate but only occurred in an unbuffered system suggesting that if a protease was present, it may be pH sensitive. To test these hypotheses, the deglycosylation experiment was carried out in the presence of a protease inhibitor cocktail, and the products were analysed for carbohydrate using the double labelling DIG kit.



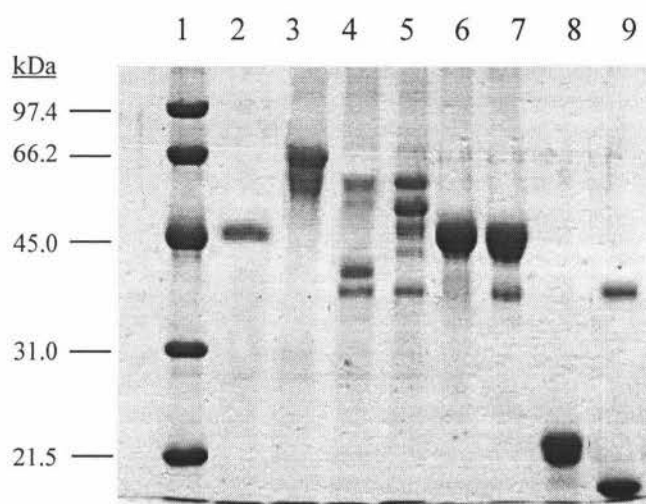
**Figure 4.4** A 12% SDS-PAGE of the deglycosylated products of fetuin by recombinant PNGase F under different buffering conditions after 48 hours of incubation at 37°C.

12  $\mu$ L reaction mixture was loaded per well.

Lane:	$\text{Na}_2\text{PO}_4$ Buffer pH	EPPS Buffer pH	Recombinant PNGase F
1. Bio-Rad MW markers	-	-	-
2. Fetuin	-	-	-
3. Fetuin	-	-	+
4. Fetuin	7	-	+
5. Fetuin	8	-	+
6. Fetuin	9	-	+
7. Fetuin	-	7	+
8. Fetuin	-	8	+
9. Fetuin	-	9	+

For the detection of glycans with the DIG kit, ovalbumin and ribonuclease B were incubated for 24 hours with PNGase F as before in the presence of 25  $\mu$ L Complete™ protease inhibitor. Fetuin was incubated with and without protease inhibitor. Two identical 12% SDS polyacrylamide gels were run and the bands from one gel were electroblotted onto a nitrocellulose membrane for glycan/protein detection. The gel showing the same pattern as before is shown in (figure 4.5). The addition of complete protease inhibitor to the mixture yielded the normal deglycosylation pattern of fetuin (lane 5) confirming that the lower band at around 38, 400 Da (lane 4) was most probably the result of proteolytic cleavage of the peptide chain. It is perhaps surprising

that no smaller breakdown products were obtained. The double labelling kit was used to confirm this observation.

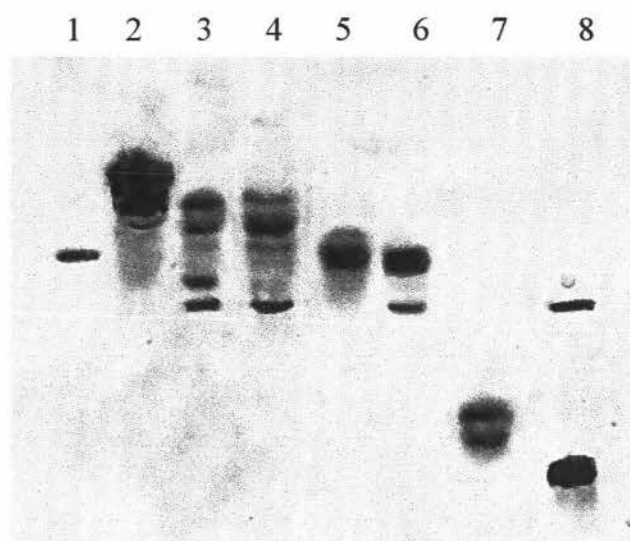


**Figure 4.5** The deglycosylated products of different glycoproteins used in DIG glycan/protein double labelling kit.

12  $\mu$ L reaction mixture was loaded per well.

Lane:		Recombinant PNGase F	Complete Protease Inhibitor
1.	Bio-Rad molecular weight markers	-	-
2.	Creatinase (DIG Kit protein standard)	-	-
3.	Fetuin	-	-
4.	Fetuin	+	-
5.	Fetuin	+	+
6.	Hen egg ovalbumin	-	-
7.	Hen egg ovalbumin	+	+
8.	Ribonuclease B	-	-
9.	Ribonuclease B	+	+

Figure 4.6 shows the result of the protein/glycan detection by the DIG kit. The band at 38,400 Da in lane 3 has a blue/green colour indicating the presence of glycans and thus confirming that the reduction in size is due to proteolytic activity (in combination with endoglycosidase activity) rather than endoglycosidase activity alone.



**Figure 4.6** The detection of glycans and proteins in deglycosylated products of various glycoproteins using DIG glycan/ protein double labelling kit.

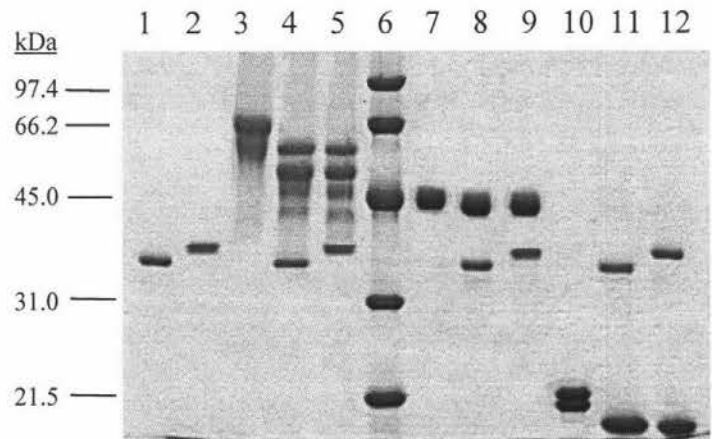
A blue-green precipitation indicates glycans present and a brown precipitation denotes protein detected.

Lane:	Recombinant PNGase F	Complete Protease Inhibitor
1. Creatinase (DIG Kit protein standard)	-	-
2. Fetuin	-	-
3. Fetuin	+	-
4. Fetuin	+	+
5. Hen egg ovalbumin	-	-
6. Hen egg ovalbumin	+	+
7. Ribonuclease B	-	-
8. Ribonuclease B	+	+

The protease is obviously pH sensitive and is presumably not active at pH 8.0. This result does however, explain the presence of multiple bands seen in the IEF gel (figure 3.25). It is possible that this protease is able to sequentially hydrolyse the histidines from the C-terminus of recombinant PNGase F which could result in the pattern seen on the IEF gel. The other glycoprotein substrates were deglycosylated as expected giving a brown signal with the polypeptide products and a blue signal for the glycosylated substrates.



A comparison of the activity of native and recombinant PNGase F with the different glycoprotein substrates is shown in figure 4.7. Incubations were carried out using the same concentration of enzyme for each incubation, and identical conditions for a 48 hours. Both forms of the enzyme deglycosylated their substrates at a similar rate with no discerable differences in the proportion or pattern of the products.



**Figure 4.7** The comparsion of glycoprotein deglycosylation by native and recombinant PNGase F.

12 µL of reaction mixture was loaded per well.

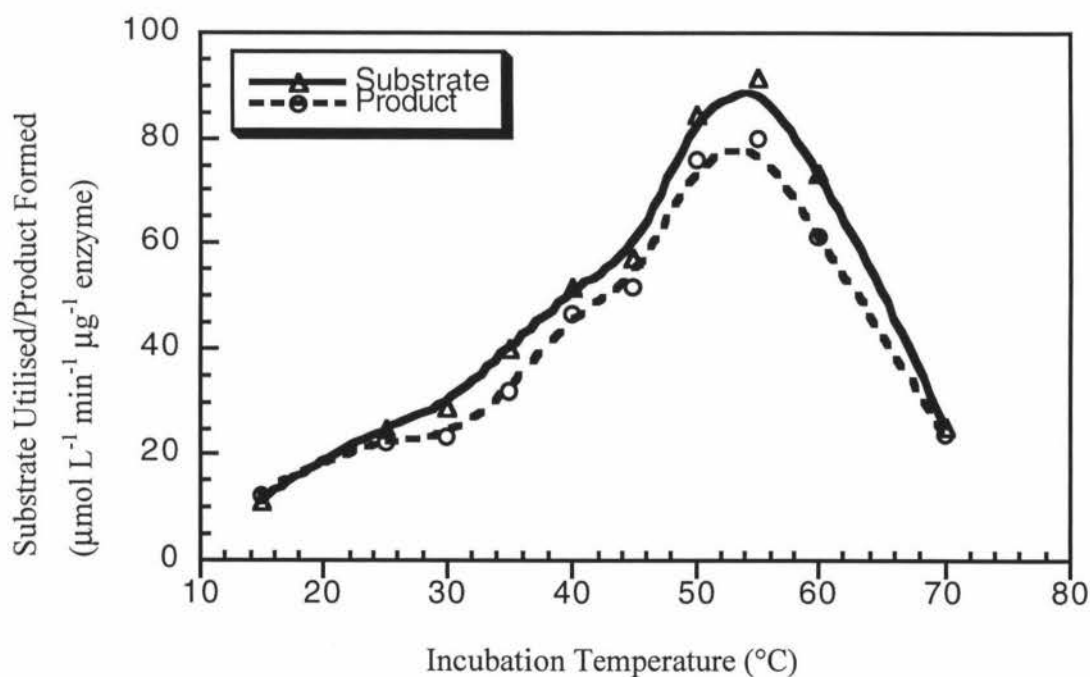
Lane:		Native PNGase F	Recombinant PNGase F
1.	Native PNGase F	+	-
2.	Recombinan PNGase F	-	+
3.	Fetuin	-	-
4.	Fetuin	+	-
5.	Fetuin	-	+
6.	Bio-Rad molecular weight markers	-	-
7.	Hen egg ovalbumin	-	-
8.	Hen egg ovalbumin	+	-
9.	Hen egg ovalbumin	-	+
10.	Ribonuclease B	-	-
11.	Ribonuclease B	+	-
12.	Ribonuclease B	-	+

### 4.5.2 Characterisation of Recombinant PNGase F

In this section, the results from experiments to investigate the effects of temperature, metal ions, storage, and additives are presented and discussed.

#### 4.5.2.1 Effects of Temperature

The temperature profile for recombinant PNGase F activity is shown in figure 4.8 and tabulated in appendix 10. Both the amount of substrate utilized and product formed were used to show the correlation between the two variables. The integrated area of product formed is generally higher than that of substrate consumed, which gives a slightly faster rate of activity when the product peak area is used in calculations. The rate of activity increases with an increase in incubation temperature to about 53°C where maximum activity is attained. This profile is similar to those of native PNGase F reported in the literature.



**Figure 4.8** Temperature profile of recombinant PNGase F.

### 4.5.2.2 Effects of pH

The buffers chosen for these experiments are listed in table 4.1 where two solutions were made using each buffer at  $\pm 0.5$  pK<sub>a</sub>. The optimum pH for maximum recombinant PNGase F activity is 8.5 as shown in figure 4.9. Note that at pH 8.5 and 9.5, there is only one data point for each of these entries on the plot. This is because the activity of recombinant PNGase F was somewhat inhibited by the Capso buffer compared with reactions containing Taps at pH 8.5 (11% less activity than Taps) or Caps at pH 9.5 (23% slower than Caps). As the only structural difference between Caps and Capso is the hydroxyl group on Capso making it slightly more hydrophilic than Caps (figure 4.10), it is interesting to speculate on how this could account for the inhibition effect. The results are also tabulated in appendix 11.

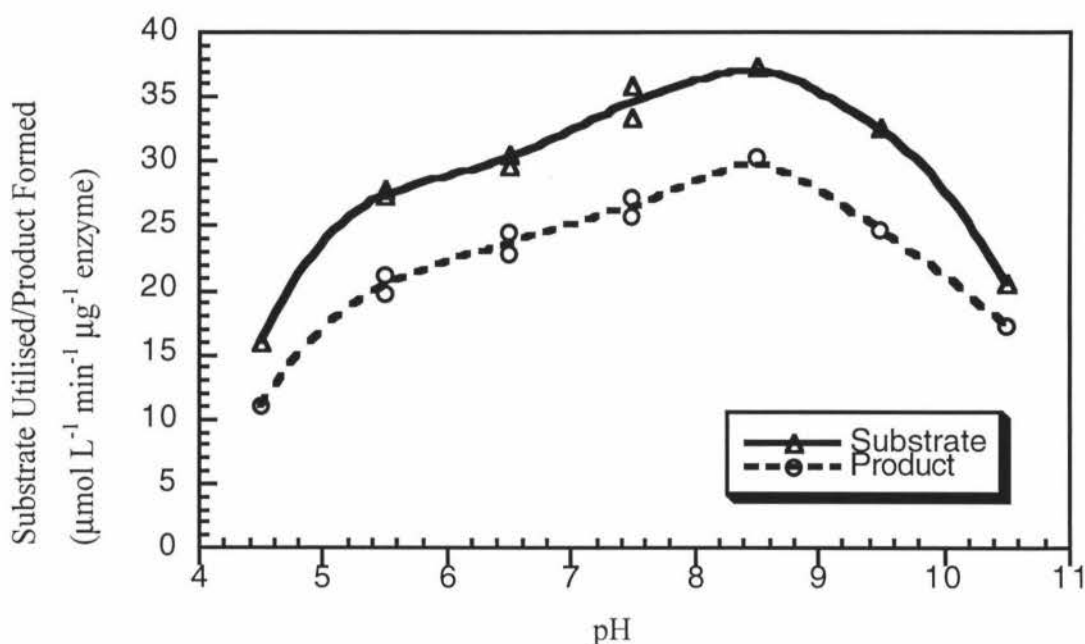


Figure 4.9 pH profile of recombinant PNGase F.

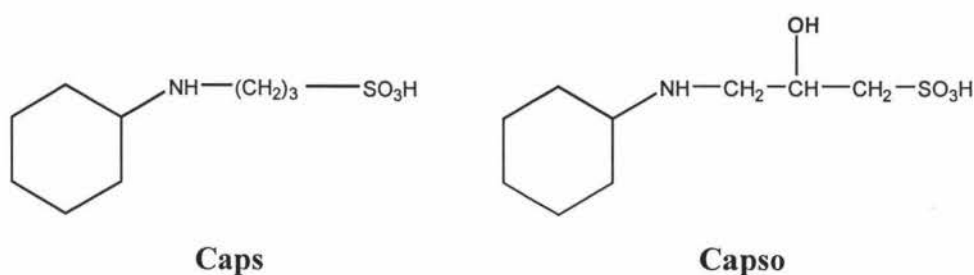
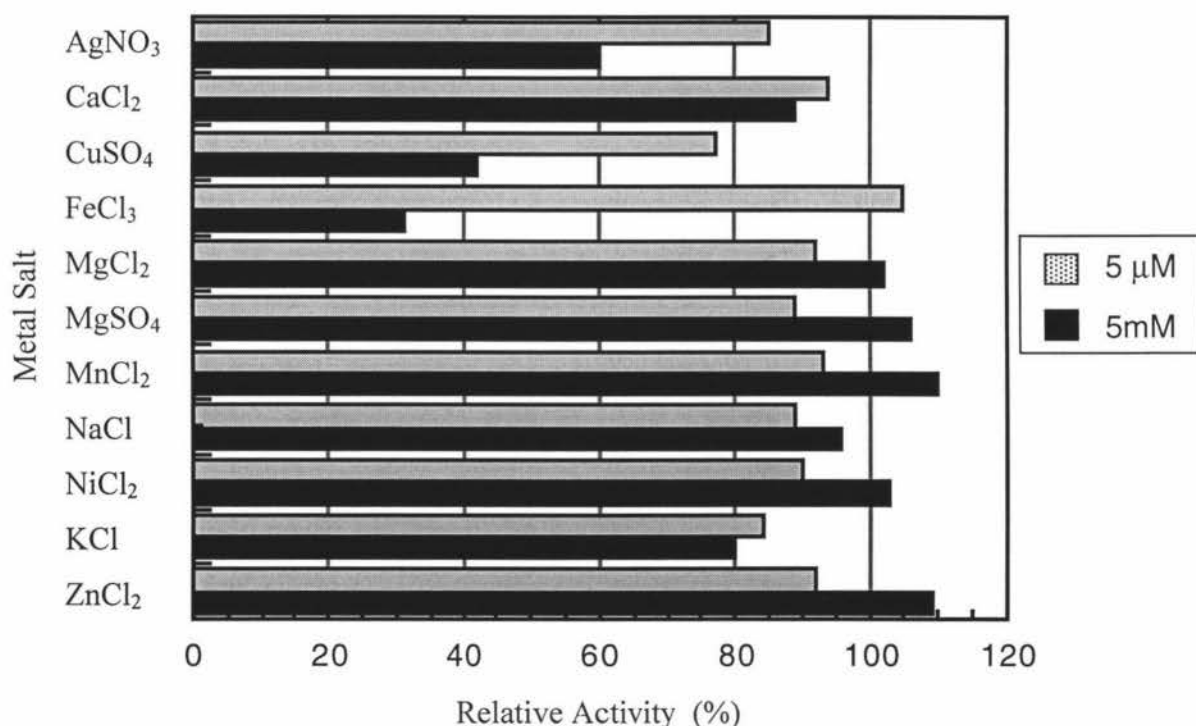


Figure 4.10 Structures of Caps and Capso buffers.

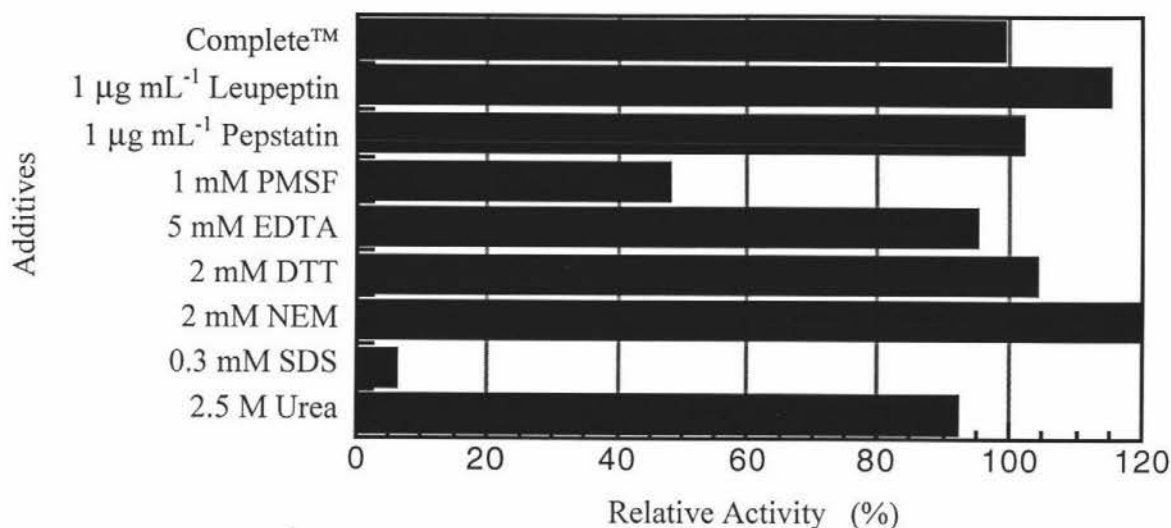
Structural difference between Caps and Capso is highlighted in bold.

### 4.5.2.3 Effects of Metal Ions and Additives

The results were presented as figures 4.11 and 4.12, and tabulated in appendix 12. In these figures, the activity in absence of metal ions was taken as 100% and the assay has a  $\pm 10\%$  error. The activities were expressed as percentage of the activity in the absence of metal ions and modifiers.  $\text{Ag}^+$ ,  $\text{Cu}^{2+}$  and  $\text{Hg}^+$  ions are known to inactivate proteins by forming covalent interactions with deprotonated cysteine ( $> \text{pH } 7$ ) and histidine ( $> \text{pH } 6$ ) residues. In some enzymes, cysteine and histidine residues have the ability to act as nucleophiles, or have a role in metal ion binding. They may also participate in hydrogen bonding or act as proton acceptors/donors in general acid-base catalysis. This ability requires the side chain nitrogen or sulfur atoms to be protonated, and the local environment of the side chain atoms affects the degree of protonation. At pH values below pH 6, histidine is protonated and is not reactive as a nucleophile whereas cysteine is predominantly in an unreactive, protonated state below pH 7. The reaction of  $\text{Ag}^+$  and  $\text{Cu}^{2+}$  ions with cysteine is sensitive to the protonation state of the thiol group and the reactivity with these ions is reduced when the pH is below 7 (Gregory, 1985).



**Figure 4.11** The effects of metal ions on recombinant PNGase F activity.



**Figure 4.12** The effects of additives on recombinant PNGase F activity.

While  $\text{Hg}^+$  was not used in this test, the presence of excess levels of  $\text{Ag}^+$ ,  $\text{Cu}^{2+}$  and  $\text{Fe}^{3+}$  ions in the assay mix has caused a significant inhibition of activity (figure 4.11). This indicates the reaction mechanism or structural stability is dependent on a cysteine and/or histidine. At the pH optimum of the enzyme (pH 8.5), cysteine and histidine should be normally deprotonated. N-ethylmaleimide (NEM) is a sulfhydryl modifying agent that react irreversibly with thiol groups to form thioesters. Dithiothreitol (DTT) is a reducing agent that keeps free thiol groups in a reduced state and is commonly added to protein purification preparations to prevent protein aggregation by disulfide bond formation. As expected, the addition of small concentration of thiol group specific agents such as DTT or NEM in the reaction had little effect on activity. Because the six cysteine residues in PNGase F are all involved in disulfide bonds (Norris *et al.*, 1994b). Since there are no free thiols in the enzyme, inhibition by these ions is therefore more likely to be due to interaction with an essential histidine (via interactions with the nitrogens of the active site histidine residue). Examination of the structure of PNGase F showed that there is a histidine residue near the active site at position 193. This residue is fairly exposed to the solvent and projects into the cleft where it could interact with bound substrates as a general acid or base. Inhibition could also be simply due to the effect of ionic strength. If this is the case, the presence of similar concentrations of other ions would be expected to show similar effects which they obviously do not. While  $\text{Na}^+$  and  $\text{K}^+$  ions at  $I > 0.05$  have been shown to inhibit PNGase F activity (Gosselin *et al.*, 1992), this is not obvious from the results reported in this work.

The catalytic residues of PNGase F, reported to be Asp-60, and Glu-206, are located in a cleft between the two domains (Kuhn *et al.*, 1995). However, for hydrolysis to occur, PNGase F must recognise both the oligosaccharide and peptide components of the substrate, and the residues for peptide recognition have yet to be identified. The results presented here suggest that His-193 may be involved in catalysis, most likely in substrate recognition.  $\text{Cu}^{2+}$  ions have also been found to be strongly inhibitory to PNGases from several animal and plant sources such as mouse L-929 fibroblast cells (Suzuki *et al.*, 1994a), and medaka fish (Seko *et al.*, 1999), and rice seeds (Chang *et al.*, 2000). Interestingly, inhibition of PNGases from medaka fish and rice seeds by  $\text{Cu}^{2+}$  ions is also probably due to an essential histidine as sulfhydryl modifying agents had no effect on them.

The unexpected decrease in enzyme activity caused by excess  $\text{Fe}^{3+}$  ions is most likely the result of precipitation of the enzyme. However,  $\text{Fe}^{3+}$  ions have been reported to inhibit PNGases from mouse L-929 fibroblast cells (Suzuki *et al.*, 1994a) and medaka fish (Seko *et al.*, 1999). No effects were observed with  $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$ , and  $\text{Zn}^{2+}$  ions where differences in activity were within the errors of the assay. While a partially purified PNGase from the rat liver endoplasmic reticulum was strongly inhibited by the presence of  $\text{Mn}^{2+}$ , and  $\text{Zn}^{2+}$  ions (Weng *et al.*, 1997), this was not observed for recombinant PNGase F. The activity of recombinant PNGase F was not affected by protein denaturants such as 2.5 M urea, nor by metal chelating agents such as 5 mM EDTA (figure 4.12), but near complete inhibition by 0.3 mM SDS. This is in agreement with reported characteristics of native PNGase F (Marley *et al.*, 1989; Tarentino *et al.*, 1994a).

During the determination of Michaelis Constant ( $K_m$ ), an increase in activity was observed when the final reaction mixture volume was increased five fold and this led to the speculation that some molecules in the enzyme preparation were inhibiting enzyme activity in a reversible fashion. The suspicion was some component in the Complete™ protease inhibitor added to the enzyme preparation was the cause of this inhibition. Unfortunately, no information was provided about the individual components of the inhibitor tablet but Roche supplies a protease inhibitor set (catalogue # 1 206 893) that was suspected to have the same components as the tablets (table 4.2). Two of these inhibitors (leupeptin and pepstatin) are readily available in the laboratory

and phenylmethanesulfonyl fluoride (PMSF) is similar to the water-soluble Pefabloc<sup>®</sup> SC. PMSF irreversibly inhibits serine proteases through the sulfonation of active serine residues in the active site of the protease and the formation of covalent bond depends on the protease being in its active site (i.e. it does not react with enzymatically inactive proteins or zymogens). The presence of protease inhibitors in the reaction had no significant effect on activity except with 1 mM PMSF (figure 4.12). This inhibition was later found due to the presence of a final concentration of 5% (v/v) isopropanol in the reaction mixture where isopropanol was used to dissolve PMSF (results not shown). Thus the component in the assay mixture thought to cause the inhibition has not been located yet.

**Table 4.2 Protease inhibitor set (Roche) was suspected to have components similar to those of the Complete<sup>™</sup> protease inhibitor tablets.**

Inhibitor	Type of Protease Inhibited	Mechanism of Action
Antipain-HCl	Serine & thiol	A tetrapeptide that forms a hemiacetal adduct between the aldehyde group of the inhibitor and the active serine of protease.
Aprotinin	Serine	A monomeric peptide of 38 residues held in conformation by three disulfide bridges and competitively forms a tight complex that blocking the active site.
Bestatin	Exoproteases	A dipeptide that competitively forms a complex and blocks the protease active site.
Chymostatin	(Chymotrypsin-like) Serine and most thiol.	A tetrapeptide that forms a hemiacetal adduct between the aldehyde group of the inhibitor and the active serine of protease.
EDTA	Metalloprotease	A metal chelator that binds bivalent cations.
E-64	Thiol	An irreversible and non-competitive inhibitor that forms a thioether bond with active thiol of the protease.
Leupeptin	Serine and thiol	A tripeptide that acts as a reversible competitive inhibitor and blocks protease active sites.
Pepstatin	Aspartic	A potent inhibitor of aspartic proteases by forming a tight 1:1 transition state analogue with acid proteases
Phosphoramidon	Metalloprotease	A competitive inhibitor that blocks protease active sites.
Pefabloc <sup>®</sup> SC	Serine	An irreversible inhibitor that modifies the active serine residue.



#### 4.5.2.4 Shelf-Life of Recombinant PNGase F

Recombinant PNGase F is stable at 4°C up to 6 months without a significant loss of activity (table 4.3). The enzyme from a 15 month storage period has 2.6% activity is partly due to natural denaturation of the protein and partly because of the action of protease where a number of degradation products can be seen on SDS-PAGE (results not shown). The enzyme in 40% (v/v) glycerol stored at -70°C retained 66% of activity compared to 45% activity from samples at 4°C in the same period. Lyophilisation of the enzyme results in the retention of most of the biological activity of the enzyme (87%) over a three-month period. Although the storage period for lyophilised PNGase F is the shortest in this investigation, there is no reason to question the fact that lyophilisation is the best long-term storage solution, especially when  $\alpha,\alpha$ -trehalose was added during lyophilisation to stabilise the protein.

**Table 4.3 The effects of long-term storage on recombinant PNGase F activity.**

Storage Period (Months)	Relative Activity* (%)
3	92
6	88
12	45
15 <sup>a</sup>	3
12 <sup>b</sup>	66
3 <sup>c</sup>	87

\*Activity of enzyme prepared on March 2000  $\pm 10\%$  error.

<sup>a</sup>Complete protease inhibitor is absent throughout the storage period.

<sup>b</sup>A 3.1 mg mL<sup>-1</sup> enzyme in 40% (v/v) glycerol stored at -70°C.

<sup>c</sup>A 3.7 mg mL<sup>-1</sup> enzyme in 2 mM EDTA, 5 mM EPPS (pH 8.0) was lyophilised and stored at -20°C.

### 4.5.3 Determination of the Michaelis Constant ( $K_m$ ) of Native and Recombinant PNGase F

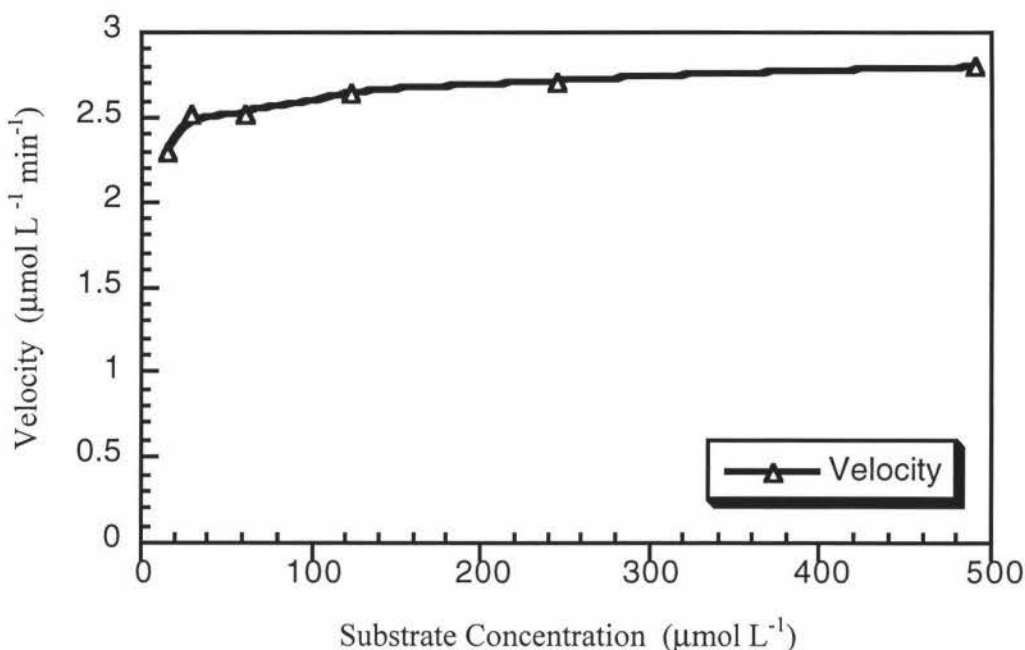
The results and problems encountered during the investigation of substrate binding and rates of reaction for both the native and recombinant enzymes are reported in this section.

#### 4.5.3.1 Experimental Strategy

Under ideal conditions, a broad range of substrate concentrations ( $0.1 - 10 K_m$ ) should be used to determine  $K_m$  and  $V_{max}$ . However, there are usually practical limits on the range to substrate concentrations over which such measurements can be performed. A compromise for this is to cover a minimum range of substrate concentrations ( $0.33 - 2.0 K_m$ ). As the kinetic constants were unknown prior to these experiments, it was thought that it was best to perform initial experiments with a limited number of data points that spanned a broad range of substrate concentration and to obtain a rough estimate of  $K_m$  from these data. This was to be followed by narrowing the substrate concentration range between  $0.33$  and  $2.0 K_m$  to obtain a larger number of data points for more accurate  $K_m$  determination.

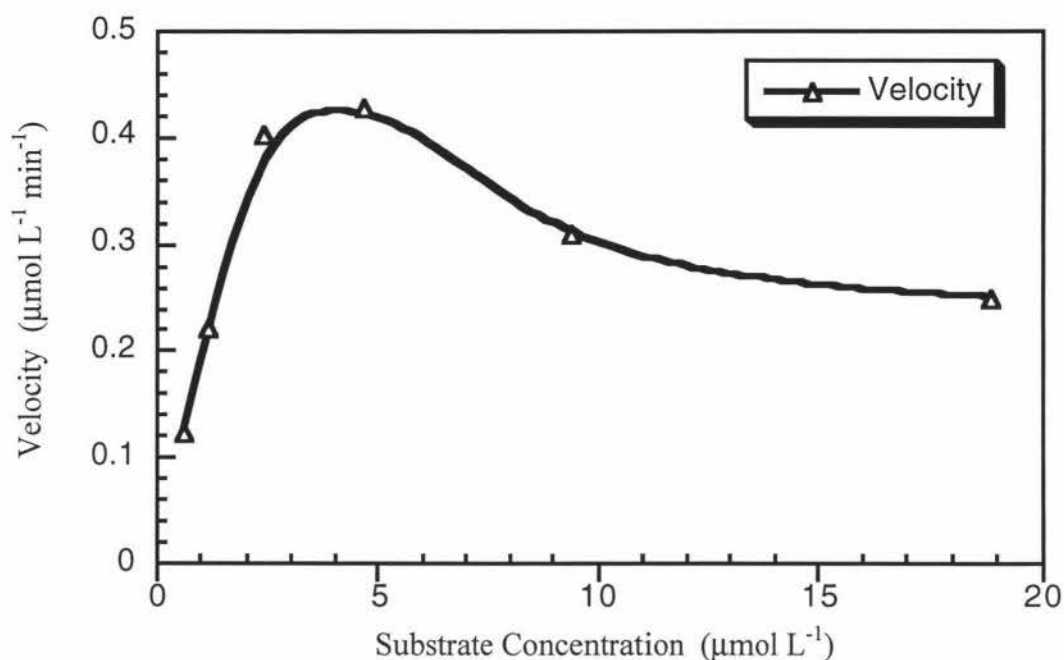
#### 4.5.3.2 Results

In the initial estimation of the  $K_m$ , six substrate concentrations ( $0.25 - 8 K_m$ ) were chosen based on the  $K_m$  value of  $0.3 \text{ mg mL}^{-1}$  for a similar ovalbumin peptide as substrate by Altman *et al.*, (1995). The results were plotted in figure 4.13 (also tabulated in appendix 13A) and it showed the substrate concentrations chosen were too high ( $[S] \gg K_m$ ). In this saturating substrate concentration range, the rate of reaction appears to be independent of substrate concentration (figure 4.13). Only the  $V_{max}$  can be determined from these data but a rough estimate of the upper limit for the  $K_m$  value can be deduced from the lowest substrate concentration which is  $15 \mu\text{mol L}^{-1}$ .



**Figure 4.13** Initial estimate of the  $K_m$  of recombinant PNGase F with ovalbumin-derived glycopeptide in small reaction volume.

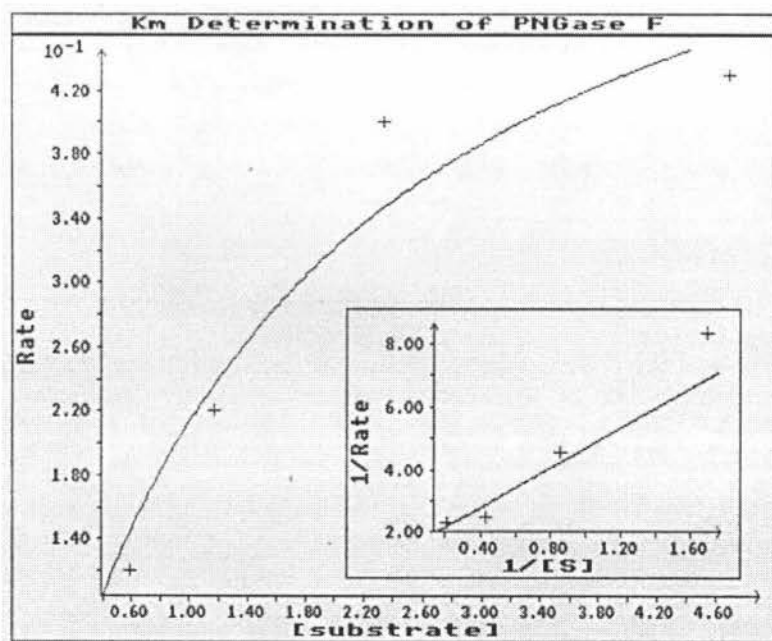
In order to overcome this problem, the volume of the reaction volume was increased five fold to decrease the substrate concentration. The results from this experiment showed an unexpected increase in activity. One explanation for this is that there is some inhibitor present in the assay mixture that is being diluted. Hence the hypothesis about the Complete™ inhibitor tablets as mentioned in section 4.5.2.3. However, that investigation was inconclusive and failed to find the component responsible for the apparent inhibition. Subsequent assays were carried out with enzyme solution that was dialysed extensively against Milli-Q water to remove any inhibitors. The results from two separate rounds of assays using this enzyme solution are shown in figure 4.14. The measured rate for the low lowest substrate concentration ( $0.25 K_m = 0.59 \mu\text{mol L}^{-1}$ ) was still above the targeted hydrolysis level of 10% substrate. Even more intriguing was the apparent product inhibition effect at higher substrate concentrations ( $4 - 8 K_m$ ) as indicated by the decrease in activity. One explanation for this could be product inhibition of the enzyme which could be investigated by collecting the separated products from the HPLC assays to determine whether this inhibition is genuine (free glycans elute in the void volume of the column with the formic acid). Further work needs to be done to determine the relative contributions of the products, glycan and/or peptide, to this inhibition by introducing the product(s) into the reaction mixture in a controlled fashion.



**Figure 4.14**  $K_m$  determination of dialysed recombinant PNGase F in a large reaction volume.

The results from the large volume reactions except the last two  $K_m$  were used to obtain a rough estimate of the  $K_m$  of 2.1  $\mu\text{M}$  and a  $k_{\text{cat}}$  of 0.65 using the program Enzfitter (version 1.04M, Elsevier-Biosoft) based on Michaelis-Menten kinetics (figure 1.15):

$$\text{Velocity} = k_{\text{cat}} \times [\text{S}] / (K_m + [\text{S}])$$



**Figure 4.15** The curve generated for calculation of  $K_m$  by the program Enzfitter.

#### 4.5.3.2 Shortcomings of the Discontinuous Activity Assay

Aside from the reaction was not being monitored real time, the assay has a number of limitations that are difficult to resolve:

##### 4.5.3.2.1 Limited Availability of the Substrate

The substrate is commercially available at a very high price and synthesis from crude material is time consuming, difficult, and with a low yield (~30 mg per preparation). Alternative glycopeptide substrates such as from fetuin and ribonuclease B are also available at similar cost and synthesis of the peptides presents similar difficulties to the synthesis of the ovalbumin peptide.

##### 4.5.3.2.2 Low Detection Sensitivity of the Assay

The ability to distinguish between the baseline noise and the small substrate and product peaks are the major limiting factor of this assay. The wavelength used (214 nm) in the assay is sensitive to solvent baseline and other compounds such as trifluoroacetic acid (TFA). Labelling the substrate with a fluorophore such as resorufin or dansyl chloride or with a radioactive tag e.g. [ $^{14}\text{C}$ ], could significantly improve the sensitivity of the assay. However, because of time constraints and the cost of the radioactive material, the substrate labelling was not performed.

##### 4.5.3.2.3 Other Difficulties of the Assay

1. Technical problems such as consistency of solution delivery by autopipettors was addressed by measuring the solution delivered by weight to  $\pm 0.0001\text{g}$ .
2. Significant absorption of the peptides onto the column matrix itself so that not all components of the assay were eluted at the end of the gradient. For example, PNGase F itself does not elute from the column until the mobile phase goes above 50% whereas in the assay it is only up to 40%. As the assays were run continuously, the column gradually becomes "coated" with components from the assay resulting in less peptide being bound to the matrix during any one run. This problem was overcome to some extent by using an internal standard described earlier in the chapter at the beginning and the end of the sample queue.

3. To determine the  $K_m$  with the Hanes-Woolf plot using data generated by this assay, it would require at least 7 data points on the plot itself at 5 different substrate concentrations performed in duplicates. Assuming no discrepancy between the duplicates, seventy assays are required.

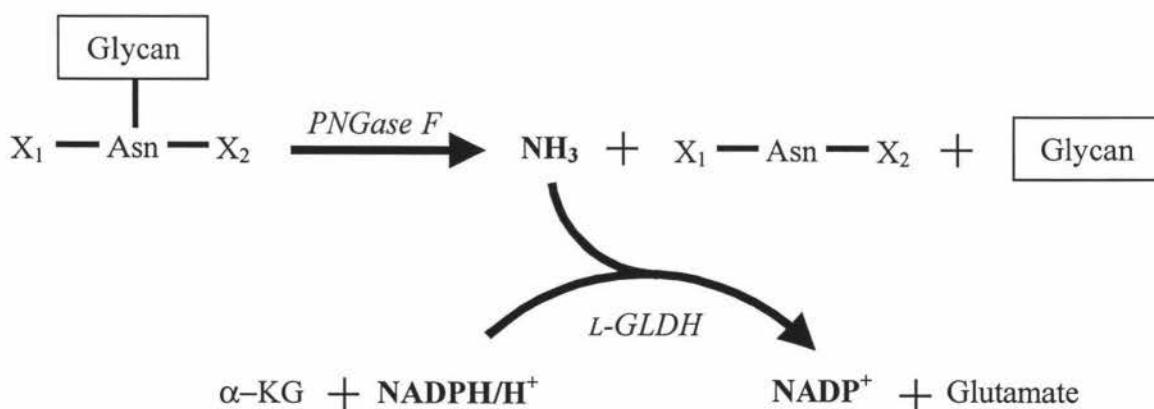
#### 4.5.4 Alternative Methods to Measure Activity

Since the lowest substrate concentration in a large volume is still above the targeted hydrolysis rate and the substrate and product peaks are barely distinguishable from the baseline noise, an alternative assay is needed to measure the reaction rate for  $K_m$  determination. Four alternative assays were designed and three were performed in an attempt to address the problem.

#### 4.5.4.1 Oxidation of NADPH by *L*-Glutamic Dehydrogenase (*L*-GLDH)

#### 4.5.4.1.1 Principle of Detection

The ammonia released from the deamination reaction of PNGase F is coupled to the conversion of an intermediate of the citric acid cycle,  $\alpha$ -ketoglutaric acid ( $\alpha$ -KG), to glutamic acid by *L*-GLDH. The conversion involves the oxidation of NADPH to NADP<sup>+</sup> that can be followed by the decrease in absorbance at 340 nm in real time (figure 4.16). An excess of *L*-GLDH, NADPH, and  $\alpha$ -KG is needed to immediately remove any ammonia produced by PNGase F, ensuring the coupling reaction does not become the bottleneck of the assay.



**Figure 4.16** A continuous colourmetric assay for measuring PNGase F activity through the oxidation of NADPH by *L*-glutamic dehydrogenase (*L*-GLDH).

#### 4.5.4.1.2 Sample preparation

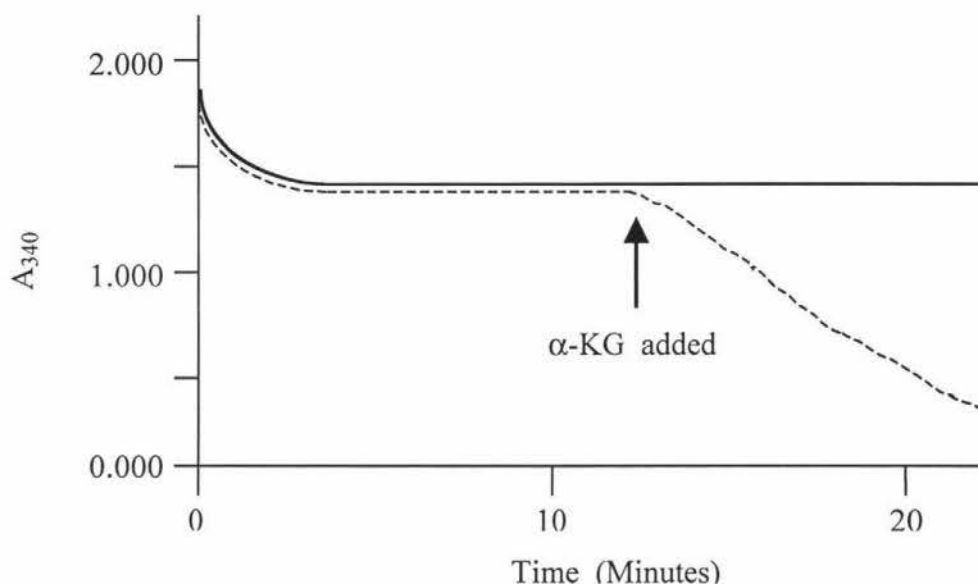
The following stock solutions were made in 10 mM  $\text{Na}_2\text{HPO}_4$  (pH 7.0). The substrate chosen for PNGase F in this assay was chicken egg ovalbumin and was made as a 20  $\text{mg mL}^{-1}$  solution whereas the substrate for L-GLDH,  $\alpha$ -ketoglutaric acid was prepared as a 1 M solution. The cofactor NADPH was made as a 30 mM solution and the enzyme L-GLDH was prepared as a 1.5  $\text{mg mL}^{-1}$  (60 units  $\text{mL}^{-1}$ ) solution in the same buffer. Recombinant PNGase F was dialysed in the same buffer overnight and concentrated to 3.0  $\text{mg mL}^{-1}$  by ultrafiltration in Thesit-treated Centricons. The reaction mixture was prepared in a 1 mL quartz cuvette by mixing 500  $\mu\text{L}$  of ovalbumin substrate, 25  $\mu\text{L}$  of L-GLDH, 100  $\mu\text{L}$  of PNGase F, with 265  $\mu\text{L}$  of Milli-Q water. This mixture was used as blank in a temperature controlled Varian® model Cary 100 UV-visible spectrometer with the detection set at 340 nm and chamber temperature controlled at 30°C. The mixture was incubated for 5 minutes before 10  $\mu\text{L}$  of NADPH was added and was incubated again until a steady state was reached before 100  $\mu\text{L}$  of  $\alpha$ -ketoglutaric acid was added to start the reaction.

#### 4.5.4.1.3. Result

In the preliminary stages, the assay was performed as a discontinuous assay to test whether the method is viable for further investigation. PNGase F was first incubated for 5 minutes before  $\alpha$ -ketoglutaric acid was added to generate a pool of ammonia (pKa 9.3). Thus the rate of absorbance decrease measured was actually the rate of L-GLDH activity. However, the rate of ammonia release from a large glycoprotein substrate such as ovalbumin was so low that even when the incubation period was elongated to 15 minutes, no L-GLDH activity was detected (figure 4.17, solid line). Ovalbumin was chosen as substrate because of the availability and to avoid using the scarce ovalbumin-derived glycopeptide. Fetuin was known to be a much faster substrate for PNGase F but the rate of ammonia release was still too low to be measured with the assay (results not shown). Thus a more sensitive detection system for ammonia detection is required, such as the quantitation of ammonia by reaction of hypochlorite and phenol producing indophenol which can be measured at 530 nm down to the  $\mu\text{M}$  range (not performed). It also seemed a glycopeptide substrate is required for PNGase F to release ammonia in a reasonable rate for measurement. Protocols for preparing glycopeptide substrates



from fetuin has been reported (Plummer *et al.*, 1987). However, because of the cost of the material and low yield, this was not done.



**Figure 4.17** Chromatogram from the L-GLDH coupling assay.

Solid line is the experimental result while the dotted line is the expected result.

#### 4.5.4.2 Real Time Measurement of $K_m$ using Surface Plasmon Resonance (SPR) Technology

Traditional real time measuring technique spectrophotometry has an enormous impact on the understanding and description of enzyme kinetic properties. However, substances that participate in signalling and the control of biological systems but lack catalytic properties have not yielded functional information, such as association and dissociation rate constants, affinity and *in vivo* concentration, easily. Optical biosensors such as BIAcore® (Pharmacia Biosensor; Uppsala, Sweden) are now widely used in the characterisation of biological macromolecular interactions. BIAcore® uses surface plasmon resonance (SPR) that allows label-free, real time monitoring of binding interactions of biomolecules. A typical experiment for examining the binding interactions between two biomolecules using BIAcore® can be summarised as a four-step process: (a) One reactant is immobilised on the gold-dextran sensor chip surface; (b) Another reactant (in solution) is injected over the sensor surface and a real time interaction curve is recorded; (c) Perform step (b) using a series of concentrations of the analyte; and (d) Choose an appropriate kinetic model from the software library and fit raw data to extract rate-constant estimates.

4.5.4.2.1 Principle of Detection

SPR is a phenomenon that occurs when surface plasmon waves are excited by biomolecular-binding events at a metal/liquid interface. These result in changes in the refractive index at the surface layer that are detected as changes in the particular angle where SPR creates extinction of light (figure 4.18). This change is measured continuously to form a sensorgram that provides a complete record of the progress of association or dissociation of the interactions (figure 4.19).

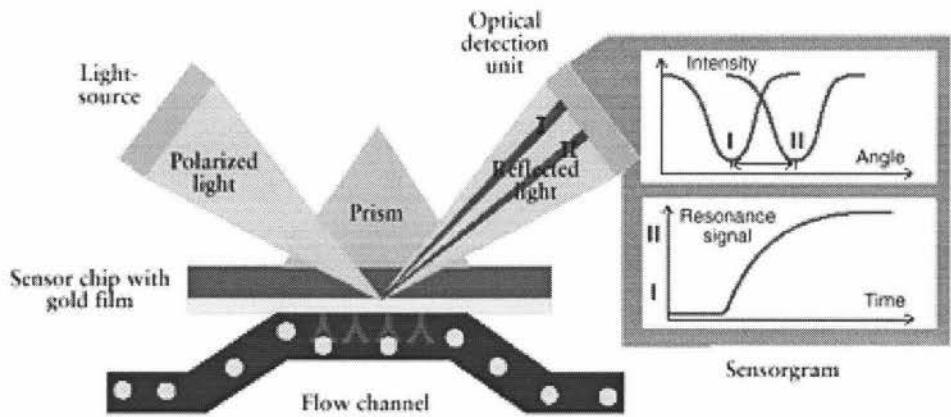


Figure 4.18 Detection of biomolecular binding events by SPR.

To perform an analysis, one reactant is captured on the sensor chip surface and the sample containing other reactant(s) is injected over this surface. Fixed wavelength light, in a fan-shaped form, is directed at the sensor chip surface where biomolecular-binding events cause a change in angle of the reflected light (I to II) by SPR. The decrease in light intensity is recorded as a resonance signal and is measured continuously to form a sensorgram.

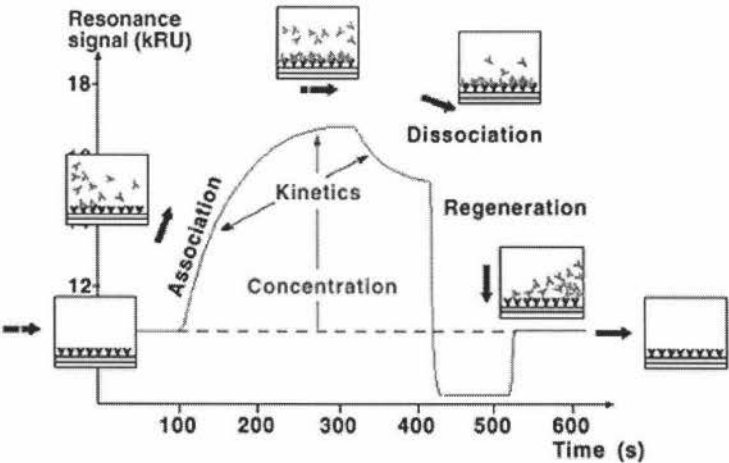
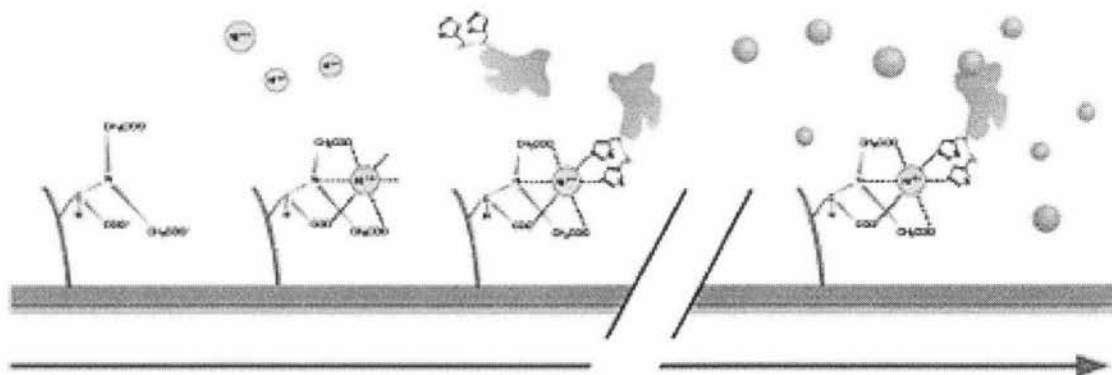


Figure 4.19 Understanding the sensorgram

The progress of an interaction is monitored as a sensorgram. Analyte binds to the surface-attached ligand during sample injection, resulting in an increase in signal. At the end of the injection, the sample is replaced by a continuous flow of buffer and the decrease in signal now reflects dissociation of interactant from the surface-bound complex.

#### 4.5.4.2.2 Application of SPR to $K_m$ determination of PNGase F

BIAcore® supplies a gold-dextran sensor chip to which NTA is pre-immobilised that allows the binding of histidine-tagged molecules such as recombinant PNGase F (figure 4.20). By following the simple four step process described earlier, a wealth of information can be generated such as the association and dissociation rate constants ( $K_m$ ). It is unfortunate that this experiment is not carried out because of the costs of the experiment were beyond the resources of the project.



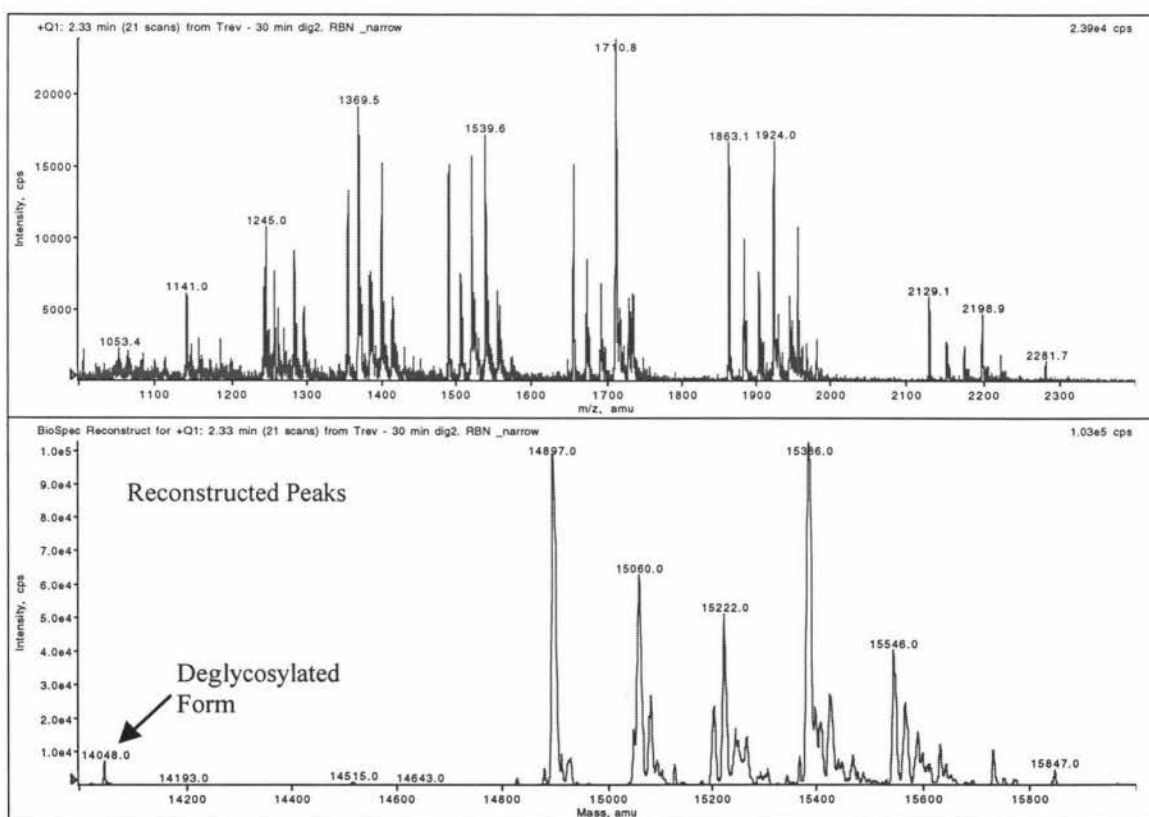
**Figure 4.20 The sensor chip NTA.**

NTA forms a quadri-dentate chelate with nickel, leaving two coordination sites on the nickel atom unoccupied. A histidine-tagged ligand will bind to these sites, schematically shown here with two histidine side chains depicted.

#### 4.5.4.3 Kinetic Analysis of Ribonuclease B Products by ES-MS

Another possible way of assuming the appearance of product and disappearance of substrate is to use electrospray mass spectrometry. There are a few examples of continuous assays that have been analysed using this technique (Bunk *et al.*, 1997; Hsieh *et al.*, 1995; Mirza *et al.*, 1995). In the initial stage of method assessment, the entire reaction mixture was injected into the mass spectrometer without separation of the substrate from the products by liquid chromatography (LC). The assay was performed as follows: Recombinant PNGase F was dialysed against double distilled Milli-Q water (final concentration of  $2.4 \text{ mg mL}^{-1}$ ) and ribonuclease B was prepared as a  $2 \text{ mg mL}^{-1}$  solution in the same water. The reaction mixture contained  $100 \text{ }\mu\text{L}$  of enzyme solution and an equal volume of ribonuclease B and was incubated at  $37^\circ\text{C}$  for intervals of 15 minutes up to 90 minutes. The reaction was stopped by rapid heating to  $100^\circ\text{C}$  in a boiling water bath and subjected to ES-MS in positive-ion mode (Sciex model API 300) at room temperature ( $22^\circ\text{C} \pm 1$ ). The parameters optimised to

maximise the intensity for ribonuclease B products are listed in appendix 14A. A typical protein would have extensive multiple charging during ionisation, resulting in a family of ions with different charged states or more precisely, mass to charge ( $m/z$ ) values which can be used to reconstruct the mass of the protein. Unfortunately, the accompanying software for the instrument has no option of calculating the area of a reconstructed peak. However, the software is capable of quantitating peaks in the  $m/z$  spectrum. Unfortunately there are many overlapping peaks obtained from the enzyme, substrate and product that cannot be distinguished easily (figure 4.21). One solution is to separate the three proteins by reverse phase HPLC before analysis by ES-MS. The isolation and purification steps needed prior to analysis necessitate the use of a protein internal standard to minimise the impact of sample loss on the accuracy of the determination. A simpler alternative is to use a glycopeptide substrate that is readily quantitated and the details of this experiment are reported in the following section.



**Figure 4.21** The spectrum for ribonuclease B product quantitation by ES-MS.

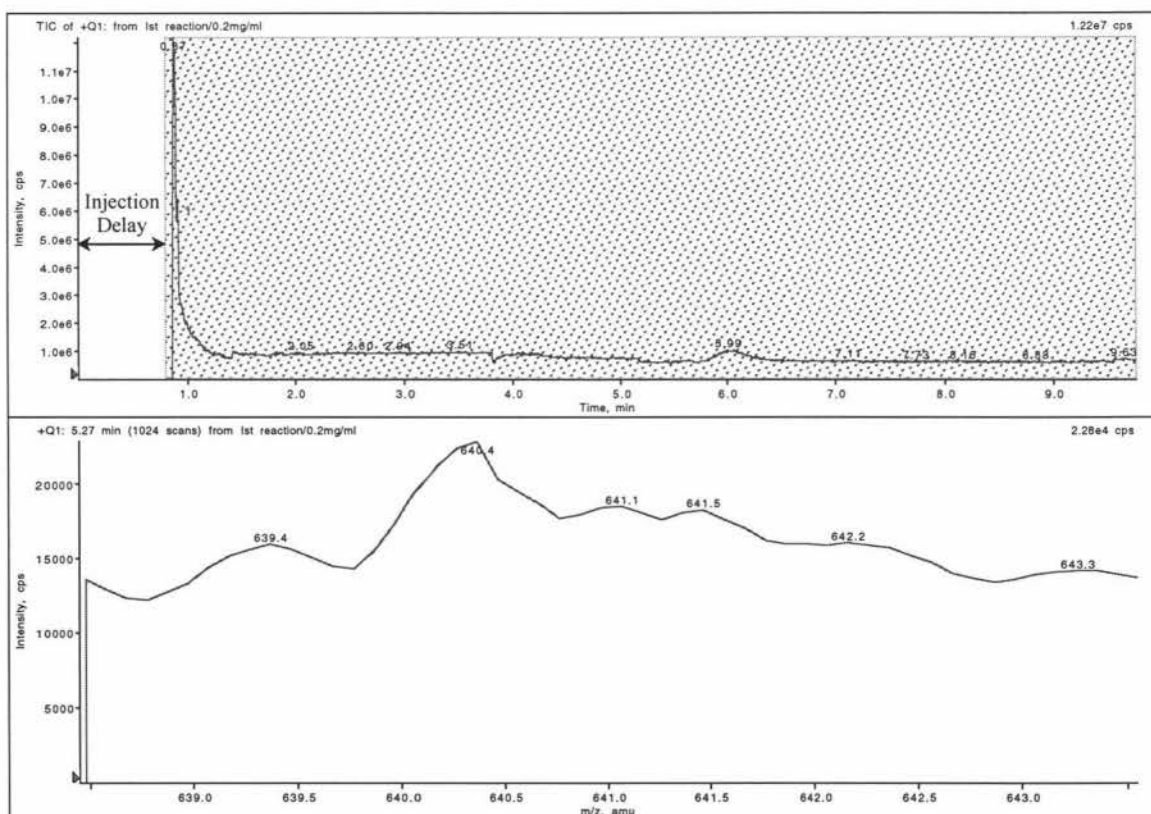
The five glycoforms of ribonuclease B differ from each other by one mannose residue (162 Da) whereas the deglycosylated form has a mass of 14048.0 Da.

#### 4.5.4.4 Pre-Kinetic Analysis of PNGase F Deglycosylation Reaction by ES-MS

This assay was a modification of the method described in the previous section, using an easily quantitated ovalbumin-derived glycopeptide (substrate of the HPLC activity assay) to determine  $K_m$  in real time (Zechel *et al.*, 1998). In the initial evaluation stages, 95  $\mu\text{L}$  of 0.2  $\text{mg mL}^{-1}$  substrate in double distilled Milli-Q water was added to 5  $\mu\text{L}$  of recombinant PNGase F (1  $\mu\text{g mL}^{-1}$ ) that was previously dialysed against Milli-Q water. The solution was mixed by aspiration in a 150  $\mu\text{L}$  glass syringe (Hamilton; internal diameter of 1.46 mm) and immediately injected into the mass spectrometer at a rate of 30  $\mu\text{L min}^{-1}$  by a syringe pump. The optimised mass spectrometer parameters for the single deglycosylated peptide product are listed in appendix 14B. The peptide product has a molecular weight of 1278 Da and was detected as a doubly charged ion at 640.4 (figure 4.22). The mass ( $M$ ) of the product is given by:

$$M = n \times m/z - nH$$

Where  $m/z$  is the mass to charge ratio of an ion,  
 $n$  is the number of protons (charges), and  
 $H$  is the mass of a proton (1.00794 Da)



**Figure 4.22** The spectrum for ovalbumin peptide product quantitation by ES-MS.

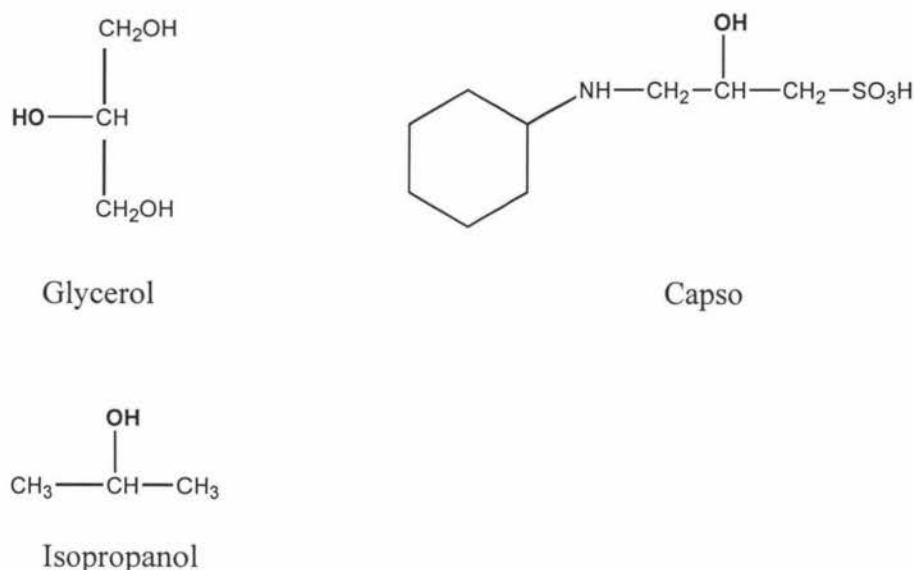
This experiment has established the optimum conditions for detecting the peptide product. However, in the small delay (~50 seconds) from the mixing of enzyme and substrate to start of product detection, over half of the substrate was deglycosylated, necessitating a new strategy. Zechel *et al.*, (1998) described a method to study the pre-steady state kinetics of xylanase that can be adapted for PNGase F. Briefly, two syringes were advanced simultaneously by a syringe pump. One syringe contained the enzyme and the other contained the substrate. The reaction was initiated by mixing the solutions from both syringes in a reaction tee that was connected to an ES ion source by a capillary of varied lengths. For a given substrate concentration, the reaction time can be varied by changing the length of capillary provided the flow rate is kept constant. At this stage, time ran out for further experimental work and this method remains to be investigated further.

## 4.6 Conclusions

Two questions were asked in the beginning of the chapter that relate to the functionality of the recombinant protein versus the native enzyme. The first question of whether the recombinant enzyme act on the same substrate as the native protein was answered by qualitatively showing on a SDS-PAGE that the recombinant enzyme deglycosylate a range of glycoproteins in a similar pace as the native protein. The second question of whether a difference in activity exist between both forms of the enzyme due to modifications made to the recombinant protein for cloning and purification was addressed by comparing the rate of deglycosylation of a glycopeptide in a discontinuous HPLC assay described in chapter three. The recombinant protein has an optimum pH of 8.5 and has similar characteristics as the native enzyme in the presence of metal ions and additives. The shelf life of recombinant PNGase F is about six months at 4°C and infinite when stored in a lyophilised state at -20°C. The susceptibility of the enzyme by  $\text{Ag}^+$ ,  $\text{Cu}^{2+}$ , and  $\text{Fe}^{3+}$  ions, together with information obtained from other additives, indicates that histidine is likely to be involved in either substrate recognition or catalysis. A probable candidate is His193 that projects into the active site and interacts with a bound substrate. Inhibition was also observed when the reaction mixture contained one or more components from the Complete™ protease inhibitor tablets.



Interestingly, recombinant PNGase F seems to be inhibited by low concentrations of glycerol, Capso, and isopropanol which all contained a hydroxyl group in the C2 position of the carbon backbone (Figure 4.23). This implies possible recognition of the hydroxyl group at this position by PNGase F and requires further investigation into the mechanism. Another intriguing result is the apparent inhibition of the enzyme by product(s) of the deglycosylation reaction during the determination of  $K_m$ . This inhibition by product(s) is not uncommon with PNGases, one example is the PNGase L-929 isolated from mouse L-929 fibroblast cells that is proposed to have a carbohydrate-binding domain, which down-regulates the deglycosylation activity of the enzyme in high concentrations of oligosaccharides (Suzuki *et al.*, 1994b).



**Figure 4.23 Structures of Capso, glycerol, and isopropanol.**

The hydroxyl groups in these structures are highlighted in bold.

The comparison with native protein by determining the Michaelis Constant of both enzymes using the HPLC assay was inconclusive because of limited substrate availability, sensitivity of the detection system and other technical problems. However, an imperfect estimate of the  $K_m$  for recombinant PNGase F was estimated to be 2.1  $\mu\text{M}$  and is ten times lower than a reported value on a similar glycopeptide substrate (Altman *et al.*, 1995). This difference may due to the eight amino acid substitutions between the two strains of *F. meningosepticum* mentioned in chapter 2, or inhibition by components of the Complete™ protease inhibitor.



Interestingly, the specific activity of a sample of the recombinant enzyme recently evaluated by Roche exceeds the specifications of the recombinant enzyme (26478 units  $\text{mg}^{-1}$  protein versus the specification of 25,000 units  $\text{mg}^{-1}$  enzyme) sold by the company (personal communication from Roche to Dr. G. Norris). The limits of the activity assay prevents the determination of a more accurate  $K_m$  value for the enzyme, thus needed alternative methods of measuring activity and they were:

- (1) The L-GLDH assay. It failed because of the slow release of ammonia from ovalbumin and fetuin.
- (2) Real time Measurement of  $K_m$  using SPR technology. Not performed because of the costs of the experiment.
- (3) Kinetic analysis of deglycosylated ribonuclease products by ES-MS. The software is incapable of quantitating reconstructed peaks but it can quantitate peaks from the  $m/z$  spectrum. However, the overlapping peaks cannot be distinguished on the  $m/z$  spectrum.
- (4) Pre-steady state kinetic analysis of PNGase F deglycosylation reaction by ES-MS. The preliminary stages established the conditions for the assay and future work following the strategy by Zechel *et al.*, (1998) may be able to determine the  $K_m$  of both enzymes.

## Chapter 5 Final Conclusion

There are three parts to this project and the procedures, problems encountered and their solutions were discussed in detail in chapters two to four. This final chapter provides a brief account of the intriguing findings and conclusions from those chapters.

A T7 promoter-based *E. coli* expression system was developed and PNGase F was expressed as a fusion protein with an ompA leader sequence on the amino-terminus and a hexa-histidine-tag on the carboxyl terminus. The recombinant enzyme was correctly processed and exported by the Sec pathway to the *E. coli* periplasm. The periplasmic PNGase F was purified to homogeneity in a single step using immobilised metal affinity chromatography, yielding 8 mg of pure protein per litre of culture. The recombinant enzyme is fully active on a range of glycoproteins and glycopeptide substrates. In addition, the presence of the hexa-histidine tag can facilitate the removal of the enzyme from the deglycosylation reaction or PNGase F to be immobilised on a matrix.

The sequence of the PNGase F coding region from the CDC strain 3352 of *F. meningosepticum* was found to differ from a previously published sequence from another strain of the bacterium (ATCC 33958) in 57 positions, resulting in eight amino acid substitutions. These nucleotide changes resulted in a protein with a theoretical molecular weight of 34836 Da compared to 34780 Da for the same enzyme from the ATCC 33958 strain. There are eleven T→C and six A→G substitutions out of the fifty-seven nucleotide differences between the two sequences. Hence at most, seventeen differences could be attributed to the mispairings typical of *Taq* while the other forty would seem to be genuine inter-strain variations. In amino acid terms, only two (T39A and I149V) out of the eight amino acid differences between the two strains could have been caused by the common *Taq* mispairings. Both amino acid changes are due to two base changes within single codons, which makes these mutations unlikely to be due to *Taq* misincorporation. Most of the amino acid differences observed are conservative in nature and are on the surface of the protein. However, two differences occur near the active site of the enzyme (I243N and A245T), and both make the  $\Omega$  loop between the two domains of the protein more hydrophilic and could lead to changes in biological activity. Three potential N-glycosylation sites (at positions 217-219, 243-245 and 281-

282) unique to the expression plasmid for the CDC strain 3352 prepared by PCR were seen. One site in particular (positions 243-245) involves two amino acid substitutions and would seem unlikely to be PCR errors, even though there is no evidence for the native enzyme to be *N*-glycosylated. In the future, this problem can be verified indirectly by comparing the discrepancies between the Edman and cDNA sequences, as a deglycosylated *N*-linked asparagine would be detected as an aspartic acid by Edman sequencing. Verification of the eight amino acid substitutions through comparison of the calculated mass from ATCC 33958 with an exact mass of the CDC strain 3352 native PNGase F obtained by Electrospray Mass Spectrometry (ES-MS) was inconclusive. These potential inter-strain substitutions thus require using one the alternative methods described in the conclusions of chapter 2.

Nearly two thirds of the expressed enzyme was lost during ultrafiltration and dialysis steps due to interactions with the concentrators and membranes themselves, despite these being described as low protein binding. Lyophilisation was shown to result in the least protein loss and most of the protein's biological activity is preserved indefinitely at  $-20^{\circ}\text{C}$ . The recombinant enzyme has near identical characteristics of the native enzyme with a pH optimum of 8.5 and is strongly inhibited by  $\text{Ag}^+$ ,  $\text{Cu}^{2+}$ , and  $\text{Fe}^{3+}$  ions but not by sulfhydryl-targeting agents such as DTT and NEM. Inhibition by these ions is thus probably due to their interaction with an essential histidine residue such as that at position 193 which could be involved in substrate recognition or catalysis. The specific activity of a sample of the recombinant CDC strain 3352 enzyme evaluated by Roche surpasses the specifications of a recombinant enzyme sold by the company which is  $>25,000$  units  $\text{mg}^{-1}$  protein. The definition of one unit of enzyme activity by Roche is the hydrolysis of 1 nmol dabsyl fibrin glycopeptide within 1 minute at  $37^{\circ}\text{C}$  in pH 7.8. However, a comparison of the specific activities of the native and the recombinant CDC strain 3352 PNGase Fs, showed that the former is about four times that of the latter. This difference may be due to inhibition by components of the Complete™ protease inhibitor tablets used in the enzyme preparation or due to the sequence changes observed in the recombinant enzyme. Interestingly, a recombinant enzyme marketed by Roche has a molecular weight of 35.5 kDa versus 34.8 kDa of the native ATCC 33958 PNGase F.

Using the discontinuous activity assay, a rough estimate of the Michael's constant ( $K_m$ ) for recombinant PNGase F was determined to be 2.1  $\mu\text{M}$ . This value is ten times lower than the figure determined by Altman *et al.*, (1995) using a shorter (9-mer) ovalbumin-derived glycopeptide substrate and PNGase F supplied by Roche. The difference is likely to be the poor data used for calculation of  $K_m$  because the specific activity reported by Roche on the recombinant CDC strain 3352 PNGase F exceeds that of the recombinant ATCC 33958 PNGase F. This issue can be clarified in the future using SPR technology or pre-steady state kinetic analysis by ES-MS as described in chapter 4.

One intriguing observation with the activity assays was the apparent product inhibition of enzyme activity. The inhibitor may either be the peptide and/or glycan components of the substrate and further investigations will be required to define the mechanism of inhibition. This inhibition by product(s) of the deglycosylation reaction was also seen in mammalian a PNGase, PNGase L-929 that down-regulates the activity of the enzyme at high oligosaccharide concentrations by binding the oligosaccharide products through a separate lectin-like receptor domain. Interestingly in legume lectins, the sugar binding site is formed by loops at one end of a  $\beta$ -sandwich structure similar to PNGase F. It is therefore possible that high concentrations of product may bind in the active site to inhibit hydrolysis. Another interesting observation is the apparent inhibition of PNGase F by glycerol, isopropanol, and Capso. These compounds contain a hydroxyl group in the C2 position of the carbon backbone that may be recognised by PNGase F.

There are still a lot of question to be addressed: (i) Substrate recognition and hydrolysis that would require both site-directed mutagenesis of active site residues and co-crystallisation of an inhibitor inside the active site of PNGase F. (ii) Potential inter-strain sequence differences, especially the potential *N*-glycosylation sites, needs to be verified by direct sequencing of genomic DNA and Edman sequencing of purified native PNGase F. (iii) A more accurate determination of the  $K_m$  of PNGase F using the pre-steady state kinetic analysis by ES-MS, following the strategy by Zechel *et al.*, 1998. (iv) The mechanism of inhibition by products of the deglycosylation reaction and compounds containing a hydroxyl group in the C2 position of the carbon backbone.

## REFERENCES

- Altmann, F., Schweiszer, S., and Weber, C. (1995) Kinetic comparison of peptide:N-glycosidases F and A reveals several differences in substrate specificity. *Glyconj. J.* **12**, 84-93.
- Akimaru, J., Matsuyama, S., Tokuda, H., and Mizushima, S. (1991) Reconstitution of a protein translocation system containing purified SecY, SecE, and SecA from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* **88**, 6545-6549.
- Ashida, H., Yamamoto, K., Kumagai, H., and Tochikura, T. (1992) Purification and characterization of membrane-bound endoglycoceramidase from *Corynebacterium* sp. *Eur. J. Biochem.* **205**, 729-735.
- Barsomian, G.D., Johnson, T.L., Borowski, M., Denman, J., Ollington, J.F., Hirani, S., McNeilly, D.S., and Rasmussen, J.R. (1990) Cloning and expression of Peptide-N<sup>4</sup>-(N-acetyl- $\beta$ -D-glucosaminyl) asparagine amidase F in *Escherichia coli*. *J. Biol. Chem.* **265**(12), 6967-6972.
- Basu, S., Ghosh, S., Basu, M., Hawes, J.W., Das, K.K., Zhang, B.J., Li, Z., Weng, S.A., and Westervelt, C. (1990) Carbohydrate and hydrophobic-carbohydrate recognition sites (CARS and HY-CARS) in solubilized glycosyltransferases. *J. Biol. Chem.* **27**, 386-395.
- Becker, J., and Craig, E.A. (1994) Heat-shocked proteins as molecular chaperons. *Eur. J. Biochem.* **219**, 11-23.
- Berger, S., Menudier, A., Julien, R., and Karamanos, Y. (1994) Endo-N-acetyl- $\beta$ -D-glucosaminidase and peptide N<sup>4</sup>-(N-acetyl-glucosaminyl) asparagine amidase activities during germination of *raphanus sativus*. *Phytochemistry.* **39**(3), 481-487.
- Berger, S., Menudier, A., Julien, R., and Karamanos, Y. (1995a) Do de-N-glycosylation enzyme have an important role in plant cells? *Biochimie.* **77**, 751-560.

- Berger, S., Menudier, A., Julien, R., and Karamanos, Y. (1995b) De-*N*-glycosylation enzymes from *Raphanus sativus*. *Phytochemistry*. **39**, 418-487.
- Bernardi, G. (1971). Chromatography of proteins on hydroxyapatite. *Methods Enzymol.* **22**, 325-339.
- Bhavanandan, V.P., Umemoto, J., and Davidson, E.A. (1976) Characterization of endo- $\alpha$ -*N*-acetylglucosaminidase from *Diplococcus pneumoniae*. *Biochem. Biophys. Res. Commun.* **70**, 738-745.
- Bowles, D.J., Marcus, S.E., Pappin, D.J.C., Findlay, J.B.C., Eliopoulos, E., Maycox, P.R., and Burgess, J. (1986) Post translational processing of concanavalin A precursors in jackbean cotyledons. *J. Cell Biol.* **102**, 1284-1297.
- Bouquelet, S., Strecker, G., Montreuil, J., and Spik, G. (1980) Characterization of a novel endo- $\beta$ -*N*-acetylglucosaminidase from the culture filtrate of a Basidiomycete (*Sporotricum dimorphosporum*), active on biantennary mono- and asialoglyco-asparagines of the *N*-acetyllactosaminic type. *Biochimie.* **62**, 43-49.
- Bourgerie, S. (1993) Evaluation des activités endoglycosidases: Caractérisation d'une endo- $\beta$ -D-acetylglucosaminidase de *Stigmatella aurantiaca* DW4. *Thèse de Doctorat*, Université de Limoges. Page 159.
- Bradford, M.M. (1976) A rapid and sensitive assay for the quantitation of microgram quantities of protein utilising the principle of protein dye-binding. *Anal. Biochem.* **72**, 248-254.
- Bunk, D.M., and Welch, M.J. (1997) Electrospray ionization mass spectrometry for the quantitation of albumin in human serum. *J. Am. Soc. Mass Spect.* **8**, 1247-1254.
- Chang, T., Suzuki, T., Inoue, S, and Inoue, Y. (1997) Change of expression of different peptide:*N*-glycanases at different stages of cell growth. *FASEB J.* **11**, A1247.

- Chang, T., Kuo, M.C., Khoo, K.H., Inoue, S., and Inoue, Y. (2000) Developmentally regulated expression of a Peptide:N-glycanase during germination of rice seeds (*Oryza sativa*) and its purification and characterisation. *J. Biol. Chem.* **275**(1), 129-134.
- Chen, L., and Tai, P.C. (1985) ATP is essential for protein translocation into *Escherichia coli* membrane vesicles. *Proc. Natl. Acad. Sci. USA.* **82**, 4384-4388.
- Chien, S.F., Weinburg, R., Li, S.C., and Li, Y.T. (1977) Endo- $\beta$ -D-acetylglucosaminidase from fig latex. *Biochem. Biophys. Res. Commun.* **76**, 317-323.
- Chen, J., Stambrook, P.J., and Tishfield, J.A. (1991) Polymerase chain reaction amplification and sequence analysis of human mutant adenine phosphoribosyl transferase genes: The nature and frequency of errors caused by *Taq* polymerase. *Mutat. Res.* **249**, 169-176.
- Clark, A.R. (1996) Molecular chaperones in protein folding and translocation. *Curr. Opin. Struct. Biol.* **6**, 43-50.
- Copeland, R.A. (1996) *Enzymes: A practical introduction to structure, mechanism, and data analysis*. Wiley-Vch, Inc., NY, USA.
- Cornish-Bowden, A. (1994) *Fundamentals of enzyme kinetics*. Revised edition, Portland Press, Cambridge, UK.
- Cullwell, A. (1987) Reduction of methionine sulfoxide in peptides using N-methylmercaptoacetamide. *Applied Biosystems User Bulletin*, Model 430, **No. 17**.
- Cumming, D.A. (1991) Glycosylation of recombinant protein therapeutics: control and functional implications. *Glycobiology.* **1**(2), 115-130.
- Danese, P.N., and Silhavy, T.J. (1998) Targeting and assembly of periplasmic and outer-membrane proteins in *Escherichia coli*. *Annu. Rev. Genet.* **32**, 59-94.



- Davankov, V.A., and Semechkin, A.V. (1977) Ligand-exchange chromatography. *J. Chromato.* **141**, 313-353.
- Davis, R.W., Boststein, R., and Roth, J.R. (1980) *Advanced bacterial genetics*. Cold Spring Harbour Laboratory, NY, USA.
- DeGasperi, R., Li, Y.T., and Li, S.C. (1989) Presence of two endo- $\beta$ -N-acetylglucosaminidases in human kidney. *J. Biol. Chem.* **264**(16), 9329-9334.
- Dong, D.L.Y., and Hart, G.W. (1994) Purification and characterization of an O-GlcNAc selective N-acetyl- $\beta$ -D-glucosaminidase from rat spleen cytosol. *J. Biol. Chem.* **269**, 19321-19330.
- Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., and Smith, F. (1956) Colorimetric method for determination of sugars and related substances. *Anal. Chem.* **28**, 350-355.
- Dunaeva, N., Goebel, C., Wasternack, C., Parthier, B., and Goerschen, E. (1999) The jasmonate-induced 60 kDa protein of barley exhibits N-glycosidase activity *in vivo*. *FEBS Letters.* **452**, 263-266.
- Duong, F., and Wickner, W. (1997) Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. *EMBO J.* **16**, 2756-2768.
- Elder, J.H., and Alexander, S. (1982) Endo- $\beta$ -N-Acetylglucosaminidase F: Endoglycosidase from *Flavobacterium meningosepticum* that cleaves both high-mannose and complex glycoproteins. *Proc. Natl. Acad. Sci. USA.* **79**, 4540-4544.
- Ellis, R.J., and Hartl, F.U. (1996) Protein folding in the cell: Competing models of chaperone function. *FASEB J.* **10**, 20-26.
- Engelman, D.M., and Steitz, T.A. (1981) The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell.* **23**(2), 411-22.

- Erikson, J.W., and Gross, C.A. (1989) Identification of the sigma E subunit of *Escherichia coli* RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression. *Genes Dev.* **3**(9), 1462-71.
- ExPASy ProtParam Tools. [www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html).
- Fan, J.Q., Kadowaki, S., Yamamoto, K., Kumagai, H., and Tochikura, T. (1988) Purification and characterization of endo- $\alpha$ -N-acetylglucosaminidase from *alcaligenes* sp. *Agri. Biol. Chem.* **52**, 1715-1723.
- Fan, J.Q., and Lee, Y.C. (1997) Detailed studies on substrate structure requirements of glycoamidases A and F. *J. Biol. Chem.* **272**(43), 27058-27064.
- Favconier, A., Van Elsen, A., and Bollen, A. (1998) Construction and sequencing of a new derivative of the pIN-III *ompA* secretion vector. *Genet. Anal.* **14**(4), 129-131.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, C.F., and Whitehouse, C.M. (1989) Electrospray ionisation for mass spectrometry of large biomolecules. *Science.* **246**, 64-71.
- Fenton, D.E. (1995) *Biocoordination Chemistry*. Oxford University Press, Oxford, England.
- Flaus, A.J. (1989) Deglycosylating enzymes from *Flavobacterium meningosepticum*. *BSc(Hons) Report*, Massey University, New Zealand.
- Freeze, H.H., and Etchinson, J.R. (1984) Presence of a non lysosomal endo- $\beta$ -N-acetylglucosaminidase in the cellular slime mould *Dictyostelium discoideum*. *Arch. Biochem. Biophys.* **232**, 414-421.
- Fujisaki, M., Tsuchida, T., Kadowaki, S., Yamamoto, K., and Tochikura, Y. (1991). Purification and properties of *Acinetobacter* sp. Endo- $\beta$ -N-acetylglucosaminidase acting on complex oligosaccharides of glycoproteins. *J. Ferment. Bioengin.* **71**, 242-247.

- Gennity, J., Goldstein, J., and Inouye, M. (1990) Signal peptide mutants of *Escherichia coli*. *J. Bioenerg. Biomembr.* **22**, 233-268.
- Georgopoulos, C. (1992) The emergence of the chaperone machines. *Trends Biochem. Sci.* **17**, 295-299.
- Gething, M.J., and Sambrook, J. (1992) Protein folding in the cell. *Nature.* **355**, 33-45.
- Ghrayeb, J., Kimura, H., Takahara, M., Hsiung, H., Masui, Y., and Inouye, M. (1984) Secretion cloning vectors in *Escherichia coli*. *EMBO J.* **3**(10), 2437-2442.
- Gill, S.C., and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.
- Goldberg, A.L., and Goff, S.A. (1986) *The selective degradation of abnormal proteins in bacteria*, Butterworths, Boston, USA.
- Gosselin, S., Martin, B.M., Murray, G.J., and Viswanatha, T. (1992) *Flavobacterium meningosepticum* peptide:N-glycosidase: influence of ionic strength on enzymatic activity. *J. Biochem. Biophys. Methods.* **24**, 71-79.
- Gregory, A.P. (1985) Preparation of isomorphous heavy-atom derivatives. *Methods Enzymol.* **114**, 147-156.
- Harley, J.L., and Bowen, H. (1996) PEG precipitation for selective removal for small DNA fragments. *Focus.* **18**(1), 27-28.
- Haltiwanger, R.S., Kell, W.G., Roquemore, E.P., Blomberg, M.A., Dong, L.Y.D., Kreppel, L., Chou, T.Y., and Hart, W. (1991) Glycosylation of nuclear and cytoplasmic proteins is ubiquitous and dynamic. *Biochem. Soc. Trans.* **20**, 264-269.

- Hayes, M.M. (1998). The refolding of recombinant human liver methylmalonyl-CoA mutase from inclusion bodies produced in *E. coli*. *MSc. Thesis*, Massey University, New Zealand.
- Hendrick, K.P., and Hartl, F.U. (1993) Molecular chaperone functions of heat-shock proteins. *Annu. Rev. Biochem.* **62**, 349-384.
- Hewick, R.M., Hunkapillar, M.W., Hood, L.E., and Dreyer, W.J. (1981) A gas-liquid solid phase peptide and protein sequencer. *J. Biol. Chem.* **256**, 7990-7997.
- Hirani, S., Bernasconi, R.J., and Rasmussen, J.R. (1987) Use of *N*-glycanase to release asparagine-linked oligosaccharides for structural analysis. *Anal. Biochem.* **162**, 485-492.
- Hobot, J.A., Carlemalm, E., Villiger, W., and Kellenberger, E. (1984) Periplasmic gel: new concept resulting from the reinvestigation of bacterial cell envelope ultrastructure by new methods. *J. Bacteriol.* **160**(1), 143-152.
- Hochuli, E., Dobeli, H. and Schacher, A. (1987) New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr.* **411**, 177-184.
- Hockney, R.C. (1994) Recent developments in heterologous protein production in *Escherichia coli*. *Trends Biotechnol.* **12**(11), 456-463.
- Houghton, R.A., and Li, C.H. (1979) Reduction of sulfoxides in peptides and proteins. *Anal. Biochem.* **98**, 36-46.
- Hsieh, F.Y.L., Tong, X., Wachs, S., Ganem, B., and Henion, J. (1995) Kinetic monitoring of enzymatic reactions in real time by quantitative high-performance liquid chromatography-mass spectrometry. *Anal. Biochem.* **229**, 20-25.
- Innis, M.A., Gelfand, D.H., and Sninsky, J.J. (1995) *PCR Strategies*. Academic Press, London, UK.

- Inoue, H., Nojima, H., and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene*. **96**, 23-28.
- Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K., and Inouye, M. (1982) Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane. *Proc. Natl. Acad. Sci. U.S.A.* **79(11)**, 3438-3441.
- Ishii-Karakasa, I., Iwase, H., Hotta, K., Tanaka, Y., and Omura, S. (1992) Partial purification and characterization of an endo- $\beta$ -*N*-acetylglucosaminidase from the culture medium of *Streptomyces* sp. *Biochem. J.* **288**, 47-482.
- Ito, K., Okada, Y., Ishida, K., Minamiura, N. (1993) Human salivary endo- $\beta$ -D-acetylglucosaminidase HS specific for complex type sugar chains of glycoproteins. *J. Biol. Chem.* **268**, 16074-16081.
- Ito, M., and Yamagata, T. (1986) A novel glycosphingolipid-degrading enzyme cleaves off the linkage between the oligosaccharide and ceramide of neutral and acidic glycosphingolipids. *J. Biol. Chem.* **261**, 14278-14282.
- Ito, S., Muramatsu, T., and Kobata, A. (1975). Endo- $\beta$ -*N*-acetylglucosaminidases acting on carbohydrate moieties of glycoproteins: purification and properties of the two enzymes with different specificities from *Clostridium perfringens*. *Arch. Biochem. Biophys.* **171**, 78-86.
- Jain, R.G., Rusch, S.L., and Kendall, D.A. (1994) Signal peptide cleavage regions. *J. Biol. Chem.* **269(23)**, 16305-16310.
- Janson, J.C., and Ryden, L. (1996) *Protein Purification: Principles, High Resolution, Methods and Applications*, second Ed., Wiley-Liss.

- Kadowaki, S., Yamamoto, K., Fujisaki, M., Izumi, K., Tochikura, T., and Yokayama, T. (1990) Purification and characterisation of a novel fungal Endo- $\beta$ -*N*-acetylglucosaminidase acting on complex oligosaccharides of glycoproteins. *Agric. Biol. Chem.* **54**, 97-106.
- Kaplan, H.A., Welply, J.K., and Lennarz, W.J. (1987) Oligosaccharyl transferase: the central enzyme in the pathway of glycoprotein assembly. *Biochem. Biophys. Acta* **906**, 161-173.
- Keohavong, P., and Thilly, W.G. (1989) Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA.* **86**(23), 9253-9257.
- Kimura, Y., and Ohno, A. (1998). A new peptide- $N^4$ -(acetyl-beta-glucosaminyl) asparagine amidase from soybean (*Glycine max*) seeds: Purification and substrate specificity. *Biosci., Biotechnol. Biochem.* **62**(2), 412-418.
- King, E.O. (1959) Studies on a group of previously unclassified bacteria associated with meningitis in infants. *J. Clin. Pathol.* **31**(3), 241-247.
- Kitajima, K., Suzuki, T., Kouchi, Z., Inoue, S., and Inoue, Y. (1995) Identification and distribution of Peptide-*N*-Glycanase (PNGase) in mouse organs. *Arch. Biochem. Biophys.* **319**(2), 393-401.
- Kruger, N.J. (1994) *Methods in Molecular Biology: Basic Protein and Peptide Protocols*. Volume 32. Humana Press Inc., N.J., USA.
- Kuhn, P., Tarentino, A.L., Plummer, T.H. Jr., and Van Roey, P. (1994) Crystal structure of Peptide- $N^4$ -(*N*-acetyl- $\beta$ -D-glucosaminyl)asparagine Amidase F at 2.2-Å resolution. *Biochemistry.* **33**, 11699-11706.
- Kuhn, P., Guan, C., Cui, T., Tarentino A.L., Plummer, T.H. Jr., and Van Roey, P. (1995) Active site and oligosaccharide recognition residues of peptide- $N^4$ -(*N*-acetyl-beta-D-glucosaminyl) asparagine amidase F. *J. Biol. Chem.* **270**(49), 29493-29497.

- Kunkel, T.A., and Bebenek, K. (1988) Recent studies of the fidelity of DNA synthesis. *Biochim. Biophys. Acta.* **951**, 1-15.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**, 680-685.
- Lehninger, A.L., Nelson, D.L., and Cox, M.M. (1993) *Principles of Biochemistry*. Second edition, Worth Publishers, N.Y., USA.
- Lemp, D., Haselbeck, A., and Klebl, F. (1990) Molecular cloning and heterologous expression of *N*-glycosidase F from *Flavobacterium meningosepticum*. *J. Biol Chem.* **265**(26), 15606-15610.
- Levitt, M. (1981) Effect of proline residues on protein folding. *J. Mol. Biol.* **145**, 251-263.
- Lhernould, S., Karamanos, Y., Boiurgerie, S., Strecker, G., Julien, M. R., Morvan, H. (1992) PNGase activity could explain the occurrence of extracellular xylomannosidases in a plant cell suspension. *Glycoconj. J.* **9**, 191-197.
- Lhernould, S., Karamanos, Y., Lerouge, P., and Morvan, H. (1995) Characterisation of the peptide *N*<sup>4</sup>-(*N*-acetylglucosaminyl)asparagine amidase (PNGase Se) from *Silene alba* cells. *Glycoconj. J.* **12**, 94-98.
- Li, S.C, DeGarespi, R., Muldrey, J.E., and Li, Y.T. (1986) A unique glycopingolipid-splitting enzyme (ceramide-glycanase from leech) cleaves the linkage between the oligosaccharide and the ceramide. *Biochem. Biophys. Res. Commun.* **141**, 346-352.
- Li, Y.T., Ishikawa, Y., and Li, S.C. (1987) Occurrence of ceramide-glycanase in the earthworm, *Lumbricus terrestris*. *Biochem. Biophys. Res. Commun.* **149**, 67-172.
- Lill, R., Dowhan, W., and Wickner, W. (1990) The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. *Cell.* **60**, 259-269.



- Lis, H., and Sharon, N. (1993) Protein glycosylation: structure and functional aspects. *Eur. J. Biochem.* **218**, 1-27.
- Lis, J.T. (1980) Fractionation of DNA fragments by Polyethylene Glycol Induced Precipitation. *Methods Enzymol.* **65**, 347-353.
- Loo, T.S., Patchett, M.L., Norris, G.E. and Lott, J.S. (2000) Using secretion to solve a solubility Problem: High-yield expression in *E. coli* and purification of the bacterial glycoamidase PNGase F. *Protein Expr Purif* (publication in manuscript).
- Lowry, O.H., Rosebough, N.J., Farr, A.L. and Randall, R. (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265-275.
- Makrides, S.C. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbio.l Rev.* **60(3)**, 512-518.
- Maley, F., Trimble, R.B., Tarentino, A.L., and Plummer, T.H. Jr. (1989) Characterisation of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal Biochem.* **180**, 195-204.
- Marston, F.A (1986) The purification of eukaryotic polypeptides synthesised in *Escherichia coli*. *Biochem. J.* **240**, 1-12.
- Mellquist, J.L., Kasturi, L., Spitalnik, S.L., and Shakin-Eshleman, S H. (1998) The amino acid following an Asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochemistry.* **37**, 6833-6837.
- Mendelman, L.V., Boosalis, M.S., Petruska, J., and Goodman, M.F. (1989) Nearest neighbour influences on DNA polymerase insertion fidelity. *J. Biol. Chem.* **264**, 14415-14423.
- Merrit, E.A., and Bacon, D.J. (1993) Raster3D: Photorealistic Molecular Graphics. *Methods Enzymol.* **277**, 505 524.

- Mirza, U.A., Chait, B.T., and Landers, H.M. (1995) Monitoring reactions of nitric oxide with peptides and proteins by electrospray ionization-mass spectrometry. *J. Biol. Chem.* **270**(29), 17185-17188.
- Missiakas, D., and Raina, S. (1997) Protein folding in the bacterial periplasm. *J. Bacteriol.* **179**(8), 2465-2471.
- Mort, A.J., and Pierce, M.L. (1995) Preparation of carbohydrates for analysis by HPLC and HPCE. *J. Chromato.* **58**, 3-5.
- Muramatsu, T. (1971) Demonstration of an endo-glycosidase acting on a glycoprotein. *J. Biol. Chem.* **246**(17), 5335-5537.
- Nicholls, A., Sharp, K.A., and Honig, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins.* **11**(4), 281-296.
- Nishiyama, A., Nishimoto, T., and Yamaguchi, H. (1991) A novel endo- $\beta$ -D-acetylglucosaminidase from the shoots of *Phyllostachys heterocycla* var pubescens. *Agric. Biol. Chem.* **55**, 115-1158.
- Norris, G.E., Flaus, A.J., Moore, C.H. and Baker, E.N. (1994a) Purification and crystallisation of the endoglycosidase PNGase F, a Peptide:N-glycosidase from *Flavobacterium meningosepticum*. *J. Mol. Biol.* **241**, 624-626.
- Norris, G.E., Stillman, T.J., Anderson, B.F., and Baker, E.N. (1994b) The three-dimensional structure of PNGase F, a glycosyl-asparaginase from *Flavobacterium meningosepticum*. *Structure.* **2**(11), 1049-1059.
- Ogata-Arakawa, M., Muramatsu, T., and Kobata, A. (1977) Partial purification and characterisation of an endo-N-acetyl- $\beta$ -D-glucosaminidase from fig extract. *J. Biol. Chem.* **82**, 611-614.

- Paquin, N.F., Haucer, C.R., Stock, R.F., Tarentino, A.L., and Plummer, T.H., Jr. (1997). Molecular cloning, primary structure and properties of a new glycoamidase from fungus *Aspergillus tubigenes*. *J. Biol. Chem.* **272(36)**, 22960-22965.
- Paquin, N.F., Tarentino, A.L., and Plummer, T.H. Jr. (1998) Overexpression of PNGase At from Baculovirus-Infected Insect Cells. *Protein Expr. Purif.* **14**, 302-308.
- Pines, O., and Inouye, M. (1999) Expression and secretion of proteins in *E. coli*.. *Mol. Biotechnol.* **12(1)**, 25-34.
- Plummer, T.H. Jr., Elder, J.H., Alexander, S., Phelan, A.W. and Tarentino, A.L. (1984). Demonstration of Peptide:N-Glycosidase F Activity in Endo- $\beta$ -N-acetylglucosamine F Preparations. *J. Biol Chem.* **259(17)**, 10700-10704.
- Plummer, T.H. Jr, Phelan, A.W., and Tarentino, A.L. (1987) Detection and quantification of peptide-N<sup>4</sup>-(N-acetyl- $\beta$ -glucosaminyl) asparagine amidases. *Eur. J. Biochem.* **163**, 163-173.
- Plummer, T.H. Jr., Tarentino, A.L. (1991) Purification of the oligosaccharide-cleaving enzymes of *Flavobacterium meningosepticum*. *Glycobiology.* **1**, 257-263.
- Plummer, T.H. Jr., Tarentino, A.L., and Hauer, C.R. (1995) Novel, specific O-glycosylation of secreted *Flavobacterium meningosepticum* proteins. *J. Biol. Chem.* **270(2)**, 13192-13196.
- Priem, B., Morran, H., and Gross, K.C. (1994) Unconjugated N-glycans as a new class of plant oligosaccharides. *Biochem. Soc. Trans.* **22**, 398-402.
- Privalov, P. (1992) *Protein Folding*. W. H. Freeman, N.Y., USA, Page 83.
- Risley, J.M., and Van Etten, R.L. (1985) <sup>1</sup>H nmr evidence that almond "Peptide:N-glycosidase" is an amidase. *J. biol. Chem.* **260(29)**, 15488-15494.

- Rouvère, P.E., and Gross, C.A. (1996) SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes Dev.* **10**, 3170-3182.
- Rudd, P.M., and Dwek, R.A. (1997) Glycosylation: Heterogeneity and the 3D structure of proteins. *Crit. Rev. Biochem. Mol. Biol.* **32**(1), 1-100.
- Rundolff, R. and Lilie, H. (1996) *In vitro* folding of inclusion body proteins. *FASEB Journal*, **10**, 49-56.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Second edition. Cold Spring Harbor Press, N.Y., USA.
- Sanger, F., Coulson A.R., Hong, G.F., Hill, D.F., and Peterson, G.B. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.
- Schatz, P.J., and Beckwith, J. (1990) Genetic analysis of protein export in *Escherichia coli*. *Annu. Rev. Genet.* **24**, 215-248.
- Scopes, R.K. (1994) *Protein purification: Principles and Practice*. Second edition, Springer-Verlag, N. Y., USA.
- Seko, A., Kitajima, K., Iwamatsu, T., Inoue, Y., and Inoue, S. (1999) Identification of two discrete peptide:N-glycanases in *Oryzias latipes* during embryogenesis. *Glycobiology.* **9**(9), 887-895.
- Shakin-Eshleman, S.H. (1996) Regulation of N-linked core glycosylation. *Trends Glycosci. Glycotechnol.* **8**, 115-130.
- Sharon, N., and Lis, H. (1982) *Glycoproteins in the enzymes*. Academic Press, N.Y., USA.

- Sheldon, P., Keen, J.N., Bowles, D.J. (1998) Purification and characterization of *N*-glycanase, a concanavalin a binding protein from jackbean (*Canavalia ensiformis*). *Biochem. J.* **330**, 13-20.
- Smith, P.K., Krohn, R.I, Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K. Goeke, N.M., Olson, B.J. and Klenk, D.C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**, 76-85.
- Smith, T.M., and Kirley, T.L. (1999) Glycosylation is essential for functional expression of a human brain ecto-apyrase. *Biochemistry.* **38**, 1509-1516.
- Ståhlberg, J., Jönsson, B. and Horvath, C. (1992) Combined effect of coulombic and van der Waals interactions in the chromatography of proteins. *Anal. Chem.* **64**, 3118-3124.
- Stoscheck, C.M. (1990) Quantitation of protein. *Methods Enzymol.*, **182**, 50-68.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60-89.
- Sugiyama, K., Ishihara, H., Tejima, S., and Takahashi, N. (1983) Demonstration of a new glycopeptidase, from Jack-bean meal, acting on aspartyl-glucosamine linkages. *Biochem. Biophys. Res. Commun.* **112**, 155-160.
- Suzuki, T., Seko, A., Kitajima, K., Inoue, Y., and Inoue, S. (1994a). Purification and enzymatic properties of peptide:*N*-glycanase from C3H mouse-derived L-929 fibroblast cells. *J. Bio. Chem.* **269**(26), 17611-17618.
- Suzuki, T., Kitajima, K., Inoue, S., and inoue, Y. (1994b) Does an animal peptide:*N*-glycanase have the dual role as an enzyme and a carbohydrate-binding protein? *Glycoconj. J.* **11**, 469-476.
- Suzuki, T., kitajima, K., Inoue, S., and Inoue, Y. (1994c) Occurrence and biological roles of 'proximal glycanases' in animal cells. *Glycobiology.* **4**(6), 777-789.

- Suzuki, T., Kitajima, K., Inoue, S., and Inoue, Y. (1995a) *N*-glycosylation/deglycosylation as a mechanism for the post-translational modification/remodification of proteins. *Glycoconj. J.* **12**(3), 183-193.
- Suzuki, T., Kitajima, K., Inoue, Y., and Inoue, S. (1995b) Carbohydrate-binding property of Peptide:*N*-Glycanase from mouse fibroblast L-929 cells are evaluated by inhibition and binding experiments using various oligosaccharides. *J. Biol. Chem.* **270**(25), 15181-15186.
- Suzuki, T., Kitajima, K., Emori, Y., Inoue, Y., and Inoue, S. (1997). Site-specific de-*N*-glycosylation of diglycosylated ovalbumin in hen oviduct by endogenous peptide:*N*-glycanase as a quality control system for newly synthesised proteins. *Proc Natl. Acad. Sci. USA.* **94**(12), 6244-6249.
- Suzuki, T., Park, H., Kitajima, K., and Lennarz, W. J. (1998) Peptides glycosylated in the endoplasmic reticulum of yeast are subsequently deglycosylated by a soluble peptide:*N*-glycanase activity. *J. Biol. Chem.* **273**(34), 21526-21530.
- Tai, T., Yamashita, K., and Kobata, A. (1977) The substrate specificities of endo-*N*-acetylglucosaminidases C11 and H. *Biochem. Biophys. Res. Comm.* **78**, 434-441.
- Takahashi, N. (1977) Demonstration of a new amidase acting on glycopeptides. *Biochem. Biophys. Res. Commun.* **76**, 1194-1201.
- Takasaki, S., Mizuochi, T., and Kobata, A. (1982) Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. *Methods Enzymol.* **83**, 263-268.
- Takahara, M., Hibler, D. W., Barr, P. J., Gerlt, J. A., and Inoue, M. (1985). The *ompA* signal peptide directed secretion of Staphylococcal nuclease A by *Escherichia coli*. *The J. Biol. Chem.* **260**(5), 2670-2674.
- Takegawa, K., Nakoshi, M., Iwahara, S., Yamamoto, K., and Tochikura, T. (1989) Induction and purification of Endo- $\beta$ -*N*-acetylglucosaminidase from *Arthrobacter protophormiae* grown in ovalbumin. *Appl. Environ. Microbiol.* **55**, 3107-3112.

- Takegawa, K., Naitoh, T., and Iwahara, S. (1991) Isolation of two endo- $\beta$ -N-acetylglucosaminidase with different specificities from *Pseudomonas* sp. *Biochem. Int.* **24**, 793-799.
- Tarentino, A.L., and Maley, F. (1969) The purification and properties of a beta-aspartyl N-acetylglucosylamine amidohydrolase from hen oviduct. *Arch. Biochem. Biophys.* **130**(1), 295-303.
- Tarentino, A.L., and Maley, F. (1972) A reevaluation of the oligosaccharide sequence associated with ovalbumins. *J. Biol. Chem.* **247**, 2626-2631.
- Tarentino, A.L., and Marley, F. (1974) Purification and properties of an endo- $\beta$ -N-acetylglucosaminidase from *Streptomyces griseus*. *J. Biol. Chem.* **249**, 811-817.
- Tarentino, A.L., and Maley, F. (1976) Purification and properties of an Endo- $\beta$ -N-acetylglucosaminidase from hen oviduct. *J. Biol. Chem.* **251**, 6537-6543.
- Tarentino, A.L., and Plummer, T.H. Jr. (1982) Oligosaccharide accessibility to Peptide:N-glycosidase as promoted by protein unfolding reagents. *J. Biol. Chem.* **257**, 10776-10780.
- Tarentino, A.L., Quinones, G., Trumble, A., Changchien, L.M., Duceman, B., Marley, F. and Plummer, T.H. Jr. (1990) Molecular cloning and amino acid sequence of Peptide-N<sup>4</sup>-(N-acetyl- $\beta$ -D-glucosaminy) asparagine Amidase from *Flavobacterium meningosepticum*. *J. Biol. Chem.* **265**(12), 6961-6966.
- Tarentino, A.L., Quinones, G., Schrader, W.P., Changchien, L.M., and Plummer, T.H. Jr. (1992) Multiple endoglycosidase (Endo) F activities expressed by *Flavobacterium meningosepticum*. *J. Biol. Chem.* **267**(6), 3868-3872.
- Tarentino, A.L., and Plummer, T.H. Jr. (1993a) Deglycosylation of asparagine-linked glycans by PNGase F. *Trends Glycosci. Glycotechnol.* **5**, 163-170.



- Tarentino, A.L., Phelan, A.W., and Plummer, T.H. Jr. (1993b) 2-Iminothiolane: a reagent for the introduction of sulphydryl groups into oligosaccharides derived from asparagine-linked glycans. *Glycobiology*. **3**(3), 279-285.
- Tarentino, A.L., Quinones, G., Changchien, L.M., Plummer, T.H. Jr. (1993c). Multiple endoglycosidase F activities expressed by *Flavobacterium meningosepticum* endoglycosidase F<sub>2</sub> and F<sub>3</sub>. *J. Biol. Chem.*. **268**(13), 9702-9708.
- Tarentino, A.L., and Plummer, T.H. Jr. (1994a) Enzymatic deglycosylation of asparagine-linked glycans: Purification, properties, and specificity of oligosaccharide-cleaving enzymes from *Flavobacterium meningosepticum*. *Methods Enzymol.* **230**, 44-59.
- Tarentino, A.L., and Plummer, T.H. Jr. (1994b). Crystal structure of Peptide N<sup>4</sup>-(N-acetyl- $\beta$ -D-glucosaminyl) asparagine amidase F a 2.2Å resolution. *Biochemistry*. **33**, 11699-11706.
- Tarentino, A.L., Quinones, G., Grimwood, B.G., Hauer, C.R., Plummer, T.H. Jr. (1995) Molecular cloning and sequence analysis of flavastacin: an O-glycosylated prokaryotic zinc metalloendopeptidase. *Arch. Biochem. Biophys.* **319**(1), 281-285.
- Trimble, R.B., and Tarentino, A.L. (1991) Identification of distinct endoglycosidase (Endo) activities in *Flavobacterium meningosepticum*: Endo F<sub>1</sub>, Endo F<sub>2</sub>, and Endo F<sub>3</sub>. *J. Biol. Chem.* **266**, 1646-1651.
- Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*. **3**, 97-130.
- von Heijne, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.*. **133**(1), 17-21.
- von Heijne, G. (1990) The signal Peptide. *J. Membr Biol.* **115**, 195-201.

- Van Wielink, J.E., and Duine, J.A. (1990) How big is the periplasmic space? *Trends Biochem. Sci.* **15**(4), 136-137.
- Walker, J.M., and Gasstra, W. (1987) *Techniques in Molecular Biology*. Volume 2. Croom Helm, Kent.
- Wall, J.G., and Pluckthun, A. (1995) Effects of overexpressing folding modulators on the *in vivo* folding of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* **6**(5), 507-16.
- Warburg, O., and Christian, W. (1941) Isolierung und kristallisation des G rungsferments Enolase. *Biochem. Z.* **310**, 384-421.
- Weng, S., and Spiro, R.G. (1997) Demonstration of a peptide:N-glycosidase in the endoplasmic reticulum of rat liver. *Biochem. J.* **322**, 655-661.
- Wetlaufer, D.R. (1962) Ultraviolet spectra of proteins and amino acids. *Adv. Protein Chem.* **17**, 303-390.
- Wickner, W., and Rice-Leonard, M.R. (1996) *Escherichia coli* preprotein translocase. *J. Biol. Chem.* **271**, 29514-29516.
- Witholt, B., Boekhout, M., Brock, M., Kingma, J., Heerikhuizen, H.V. and Leij, L.D (1976) An efficient and reproducible procedure for the formation of spheroplasts from variously grown *Escherichia coli*. *Anal. Biochem.* **74**, 160-170.
- Wilkinson, P.M., Lind, L.K., Berg, D.E., and Bjork, G.R. (1991) Predicting the solubility of recombinant proteins in *Escherichia coli*. *Biotechnology*, **9**, 443-448.
- Wülfing, C., and Plückthun, A. (1994) Protein folding in the periplasm of *Escherichia coli*. *Mol. Microbiol.* **12**(5), 685-692.

Yamamoto, K., Kadowaki, S., Takegawa, K., Kumagai, H., and Tochikura, T. (1986) Purification and characterisation of Endo- $\beta$ -*N*-acetylglucosaminidase from a *Flavobacterium* sp. *Agr. Biol. Chem.* **50**(2): 421-429.

Yet, M.G., and Wold, F. (1988) Purification and characteization of two glycopeptide hydrolases from jack beans. *J. Biol. Chem.* **263**, 118-122.

Zechel, D.L., Konermann, L., Withers, S.G. (1998) Pre-steady state kinetic analysis of an enzymatic reaction monitored by time-resolved electrospray ionization mass spectrometry. *Biochemistry*, **37**, 7664-7669.

Appendix 1

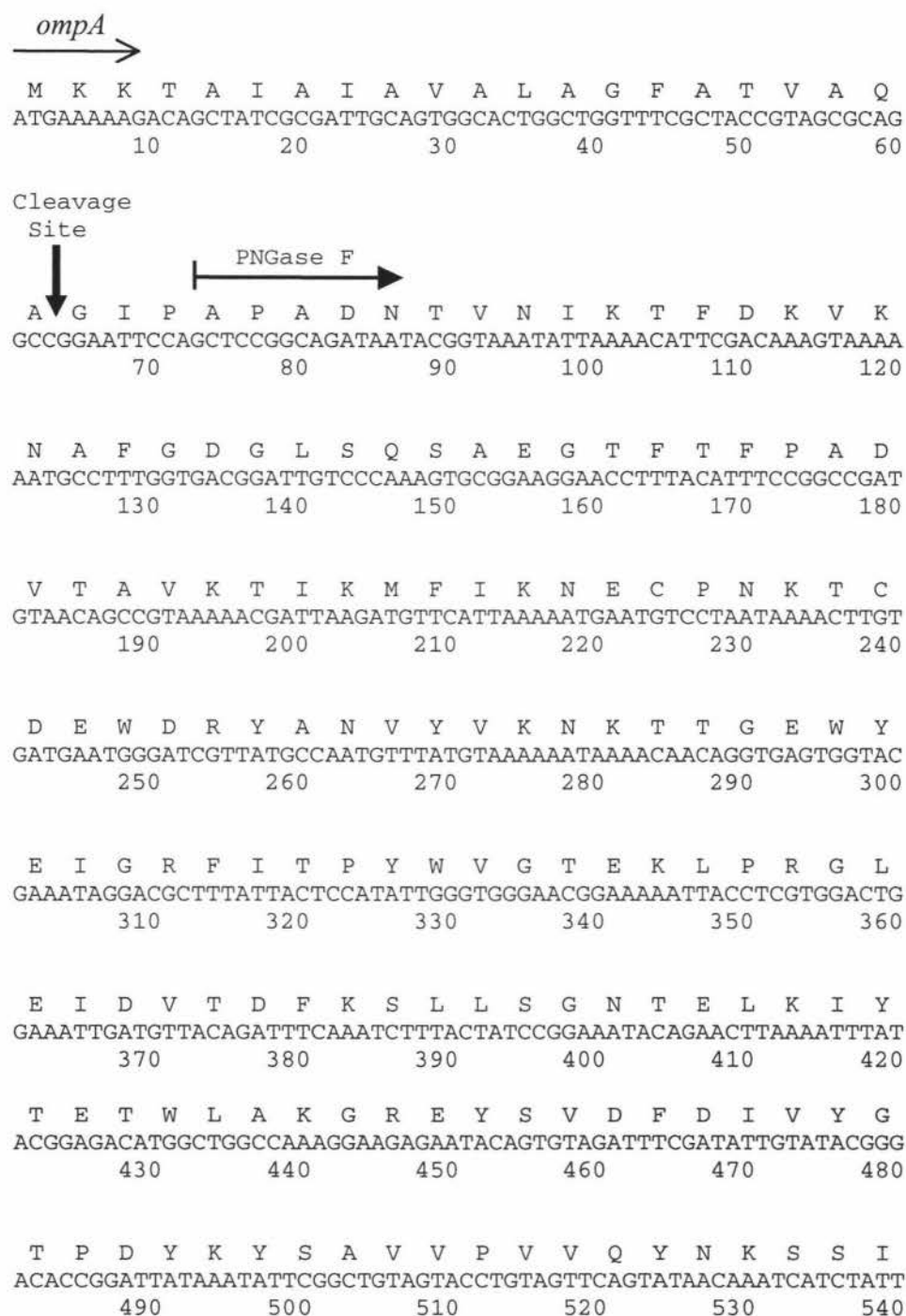
The Genetic Code

A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
GCA	CGA	AAC	GAC	UGC	CAA	GAA	GGA	CAC	AUA	CUA	AAA	AUG	UUC	CCA	UCA	ACA	UGG	UAC	GUA
C	C	U	U	U	G	G	C	U	C	C	G		U	C	C	C		U	C
G	G						G		U	G				G	G	G			G
U	U						U			U				U	U	U			U
	or									or					or				
	AGA									UUA					AGC				
	G									G					U				

		Second Position				
		U	C	A	G	
First Position (5' end)	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G
						Third Position (3' end)

## Appendix 2      Sequence of Recombinant PNGase F

The sequence of OPH6 construct is shown with the N-terminal *ompA* secretory signal and the C-terminal His tag. The asterisk at the end of the gene denotes the stop codon.



D G V P Y G K A H T L A L K K N I Q L P  
GACGGAGTCCCTTATGGTAAAGCACATACATTGGCTTTGAAAAAGAATATCCAGTTACCA  
550 560 570 580 590 600

T N T E K A Y L R T T I S G W G H A K P  
ACAAACACAGAAAAAGCTTATCTTAGAACTACTATTTCCGGATGGGGACATGCTAAGCCA  
610 620 630 640 650 660

Y D A G S R G C A E W C F R T H T I A I  
TATGATGCGGGAAGCAGAGGTTGTGCAGAATGGTGCTTCAGAACACACACTATAGCAATA  
670 680 690 700 710 720

N N S N T F Q H Q L G A L G C S A N P I  
AATAATTCGAATACTTTCCAGCATCAGCTGGGTGCTTTAGGATGTTTCAGCAAACCCCTATC  
730 740 750 760 770 780

N N Q S P G N W T P D R A G W C P G M A  
AATAATCAGAGTCCGGGAAATTGGACTCCCGACAGAGCCGGTTGGTGCCCGGGAATGGCA  
790 800 810 820 830 840

V P T R I D V L N N S L I G S T F S Y E  
GTTCCAACACGTATAGATGTACTGAATAATTCCTTTAATAGGCAGTACTTTTAGTTATGAA  
850 860 870 880 890 900

Y K F Q N W T N N G T N G D A F Y A I S  
TATAAATTCAGAACTGGACAAATAACGGAACCAATGGAGATGCTTTTATGCAATTTCC  
910 920 930 940 950 960

S F V I A K S N T P I S A P V V T N L D  
AGTTTTGTGATTGCAAAAAGTAATACACCTATTAGTGCTCCGGTAGTTACAAACTTGGAT  
970 980 990 1000 1010 1020

His Tag  
P H H H H H H \*  
CCGCATCACCATCACCACCATTTGA  
1030 1040

Mature PNGase F with the signal peptide removed have 326 amino acids (978 bp) with a theoretical MW of 36251.61 Da and a theoretical pI of 7.75 calculated by ExPASy ProtParam Tools ([www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html)).

Appendix 3 Sequence Comparison of Two Strains of PNGase F

The nucleic acid sequence of PNGase F from CDC strain 3352 is shown with the translated amino acid placed on top of the codon. Base changes resulting in amino acid changes versus the strain ATCC 33958 are indicated in bold letters and those that do not change the amino acid sequence were underlined. The amino acid substitutions that result in changes in potential N-glycosylation sites are enclosed in a box. The asterisk at the end of the gene denotes the stop codon.

A	P	A	D	N	T	V	N	I	K	T	F	D	K	V	K	N	A	F	G
GCT	CCG	GCA	GATA	ATA	ACGG	TAA	ATAT	TAAA	ACATT	CGA	CAA	AGT	AAAA	ATGC	CTT	TGG	T		
					<u>C</u>														
10					20			30			40			50					60
D	G	L	S	Q	S	A	E	G	T	F	T	F	P	A	D	V	T	<b>A</b>	V
GAC	GG	ATT	GT	CC	CAA	AGT	GC	AGA	AGG	AA	CTT	TAC	ATT	TCC	GG	CC	GAT	GTA	<b>ACCG</b>
																		<b>A</b>	<b>T</b>
70					80			90			100			110					120
K	T	I	K	M	F	I	K	N	E	C	P	N	K	T	C	D	E	W	D
AAA	AC	GAT	TA	AG	AT	GT	TCA	TAAAA	ATGA	ATG	TC	TA	ATA	AA	CTT	GT	GAT	GA	TGG
						<u>C</u>									<u>C</u>				
130					140			150			160			170					180
R	Y	A	N	V	Y	V	K	N	K	T	T	G	E	W	Y	E	I	G	R
CGT	TAT	G	CCA	ATG	TTT	ATG	TAAAA	AA	TAAAA	CAAC	AGG	TG	AGT	TGG	TAC	GAA	ATAG	GAC	G
												<u>A</u>	<u>A</u>		<u>T</u>				
190					200			210			220			230					240
F	I	T	P	Y	W	V	G	T	E	K	L	P	R	G	L	E	I	D	V
TTT	ATA	TAC	CC	AT	ATT	GGG	TGG	GA	ACG	GAAAA	ATTAC	CTC	GTG	GACT	TGG	AA	ATTG	AT	GTT
						<u>T</u>													
250					260			270			280			290					300
T	D	F	K	S	L	L	S	G	N	T	E	L	K	I	Y	T	E	T	W
AC	AG	ATT	TCA	AAT	CTT	TACT	AT	CCG	GAA	ATAC	AGAA	CTT	AAAA	TTT	TAC	GGA	GAC	AT	G
	<u>C</u>					<u>G</u>	<u>G</u>											<u>T</u>	
310					320			330			340			350					360



L A K G R E Y S V D F D I V Y G T P D Y  
 CTGGCCAAAGGAAGAGAATACAGTGTAGATTTCGATATTGTATACGGGACACCGGATTAT  
T 370 380 390 C T 400 T T 410 420

K Y S A V V P V **I** V Q Y N K S S I D G V P  
 AAATATTTCGGCTGTAGTACCTGTAGTTCAGTATAACAAATCATCTATTGACGGAGTCCCT  
A 430 **A C** A 440 450 460 C T T T 470 480

Y G K A H T L **G** L K K N I Q L P T N T E  
 TATGGTAAAGCACATACATTGGCTTTGAAAAAGAATATCCAGTTACCAACAAACACAGAA  
C **GA** A 490 500 510 520 530 G 540

K A Y L R T T I S G W G H A K P Y D A G  
 AAAGCTTATCTTAGAACTACTATTTCCGGATGGGGACATGCTAAGCCATATGATGCGGGA  
C 550 560 570 580 590 600

S R G C A E W C F R T H T I A I **A** N N S N  
 AGCAGAGGTTGTGCAGAATGGTGCTTCAGAACACACACTATAGCAATAAATAATTCGAAT  
G C 610 620 630 T 640 650 **G** 660

T F Q H Q L G A L G C S A N P I N N Q S  
 ACTTTCCAGCATCAGCTGGGTGCTTTAGGATGTTTCAGCAAACCTATCAATAATCAGAGT  
A C 670 680 690 700 T 710 720

P G **I** **A** N W T P D R A G W C P G M A V P T R  
 CCGGGAAATTGGACTCCGACAGAGCCGGTTGGTGCCCGGGAATGGCAGTTCCAACACGT  
**T** **G** T A G T G 730 740 750 760 770 780

I D V L N N S L **T** I G S T F S Y E Y K F Q  
 ATAGATGTACTGAATAATTCTTTAATAAGGCAGTACTTTTAGTTATGAATATAAATTCCAG  
GT C **CG** T 790 800 810 820 830 G 840

<b>S</b>																			
<b>N</b>	<b>W</b>	<b>T</b>	N	N	G	T	N	G	D	A	F	Y	A	I	S	S	F	V	I
AACTGGACAAATAACGGAACCAATGGAGATGCTTTTATGCAATTTCCAGTTTTGTGATT																			
<b>GT</b>			<b>C</b>																
			850			860			870			880			890			900	

A	K	S	N	T	P	I	S	A	P	V	V	T	N	*
GCAAAAAGTAATACACCTATTAGTGCTCCGGTAGTTACAAACTAA														
			910			920			930			940		

The DNA encoding PNGase F from *F. meningosepticum* (CDC strain 3352) contains a total of 942 base pairs with 57 base pair differences from the published sequence for strain ATCC 33958. Thirteen of these changes resulted in eight amino acid changes as indicated in bold. PNGase F from CDC strain 3352 has a theoretical molecular weight of 34836.07 Da and a pI of 8.14 versus PNGase F from strain ATCC 33958 which has a theoretical molecular weight of 34780 Da and a pI of 8.14 as calculated by ExPASy ProtParam Tools ([www.expasy.ch/tools/protparam.html](http://www.expasy.ch/tools/protparam.html)). Three potential N-glycosylation sites were gained in the CDC strain 3352 (positions 649-657, 727-735 and 841-849 bp). Interestingly, the site at position 727-735 bp involves two substitutions which are unlikely to be PCR errors.

## Appendix 4    Synthesis of Ovalbumin Glycopeptide Substrate

The protocol for glycopeptide synthesis was adapted from Norris *et al* (1994) with some modifications as noted below.

### Materials:

Grade II chicken egg albumin (Sigma); analytical grade ether (Ajax Chemicals), sulphuric acid (Rhône-Poulenc Chemicals Limited), phenol (BDH), and TCA (Ajax Chemicals). Cyanogen bromide was made up into 10 mg mL<sup>-1</sup> stock solution with analytical grade acetonitrile, flushed with nitrogen, and stored at -70°C till use. Chromatographic media used includes: Sephadex G-25 (Pharmacia), Jupiter C18 reverse-phase column (250 x 4.6 mm) with Security Guard column both from Phenomenex.

### Method:

1. Approximately 6 g of ovalbumin was digested with 3 g of CNBr in 150 mL of 70% formic acid for 24 hours at room temperature under nitrogen and with gentle stirring.
2. Toxic CNBr was removed by addition of an equal volume of water before the volume was reduced to 10% of the original by rotary evaporation.
3. Undigested protein and large insoluble peptides were precipitated by the addition of Trichloroacetic acid (TCA) to a final concentration of 5% and removed by centrifugation (3,000 g for 10 minutes).
4. Excess TCA and lipids were removed by extraction with ether and the aqueous phase containing the mixture of peptides was reduced to about 2 mL by rotary evaporation.
5. Acetic acid was added to the concentrated mixture to 0.5% (v/v) and insoluble material was pelleted with brief centrifugation in a microfuge.

1. Half of the mixture was subjected to gel filtration in a column (100 x 2.5 cm, Pharmacia) packed with Sephadex G-25 and equilibrated in 0.1 M acetic acid. Fractions containing the glycopeptide were identified using the phenol-sulphuric acid test (Dubois *et al.*, 1956). Briefly, the Dubois test for free reducing sugars was performed by adding 500  $\mu\text{L}$  of sample to 12.5  $\mu\text{L}$  of 80% (w/v) phenol and 1.25 mL of concentrated sulphuric acid. An orange to brown colour developed within 15 minutes and the intensity was proportional to the amount of reducing sugar present in the sample.
7. Fractions tested positive with sugar were pooled, lyophilised and reconstituted in 10 mL of water to give  $\sim 12 \text{ mg mL}^{-1}$  solution.
8. 200  $\mu\text{L}$  aliquots of reconstituted glycopeptide were subjected to further purification by reverse-phase HPLC on a C18 column running in 80% water / 20% acetonitrile / 0.1% TFA at  $1 \text{ mL min}^{-1}$ . Bound peptides were eluted by a 15 minute linear gradient from starting conditions to 60% water / 40% acetonitrile / 0.08% TFA. The eluent was monitored at 220 nm with a detector AUFS range at 2.0. The substrate peak was collected, lyophilised and stored at  $-20^{\circ}\text{C}$  till use.

## Appendix 5 Silver Staining of SDS-Polyacrylamide Gels

This method was a personal communication from Dr. J. S. Lott.

### Materials:

Analytical grade silver nitrate, sodium carbonate, potassium ferricyanide, sodium thiosulphate, acetic acid, methanol, glycerol and formaldehyde

### Method:

Due to the high sensitivity of this staining method, all glassware and plastics were either new or thoroughly washed with detergent; gloves, laboratory coats were worn at all times to prevent keratin contamination.

1. The gel was fixed in 30% ethanol / 10% acetic acid for 3 hours.
2. Soaked in 10% ethanol for 5 minutes to remove acetic acid.
3. Ethanol was removed by washing in Milli-Q water for 30 minutes with three changes of water.
4. The gel was soaked in freshly prepared Farmers reagent (3 g L<sup>-1</sup> sodium thiosulfate, 1.5 g L<sup>-1</sup> potassium ferricyanide and 0.5 g L<sup>-1</sup> sodium carbonate in Milli-Q water) for 1 minute to eliminate non-specific binding.
5. Washed with Milli-Q water three times for 10 minutes to remove the reagent.
6. The gel was stained with 0.1% silver nitrate for 30 minutes.
7. Excess silver nitrate was removed by rinsing in Milli-Q water for 10-20 seconds.
8. Protein bands was visualised by soaking in freshly made 2.5% sodium carbonate / 0.02% formaldehyde until desired amount of staining was achieved.
9. Staining was ceased by 1% acetic acid for 5 minutes.

10. Washed with Milli-Q water for 30 minutes with three changes of water to remove acetic acid.
11. Soaked in Farmers reagent for 10-30 seconds to reduce background.
12. The reagent was eliminated by washing with Milli-Q water for 30 minutes with three changes of water.
13. Gel was soaked in 2 % glycerol / 30% acetic acid before drying.

## Appendix 6     Synthesis of *t*-butyl TSK Resin

This protocol was a personal communication from Mr. Tony Bland.

### Materials:

1, 1'-carbonyl diimidazole (CDI) from Sigma, was stored desiccated at  $-20^{\circ}\text{C}$  and warmed to room temperature before opening. Analytical grade dioxan from Ajax Chemicals was dehydrated with a few spatulas full of calcium hydride in the bottle and filtered through Whatman filter paper before use. Fractogel™ TSK HW55 resin was obtained from Merck and analytical grade *t*-butylamine from BDH Laboratory Supplies.

### Method:

The preparation was carried out in fumehood. The reaction involved CDI activation where the hydroxyl group on the resin was replaced by an imidazole group which was a much better leaving group for the *t*-butyl group to come in.

1. Approximately 13 g resin in 50 mL of 20% ethanol was placed in a sintered glass funnel over a 2 L side-arm conical flask and the resin bed was protected from subsequent washings with a filter paper on top.
2. The resin was dehydrated by sequential washing with 2 volumes (100 mL) of the following solutions at a flow rate of roughly 1 volume per 10 minutes:

Wash #	Ethanol	Dry Dioxan	Milli-Q water
1	20%	-	80%
2	10%	25%	65%
3	-	50%	50%
4	-	75%	25%
5	-	100%	-



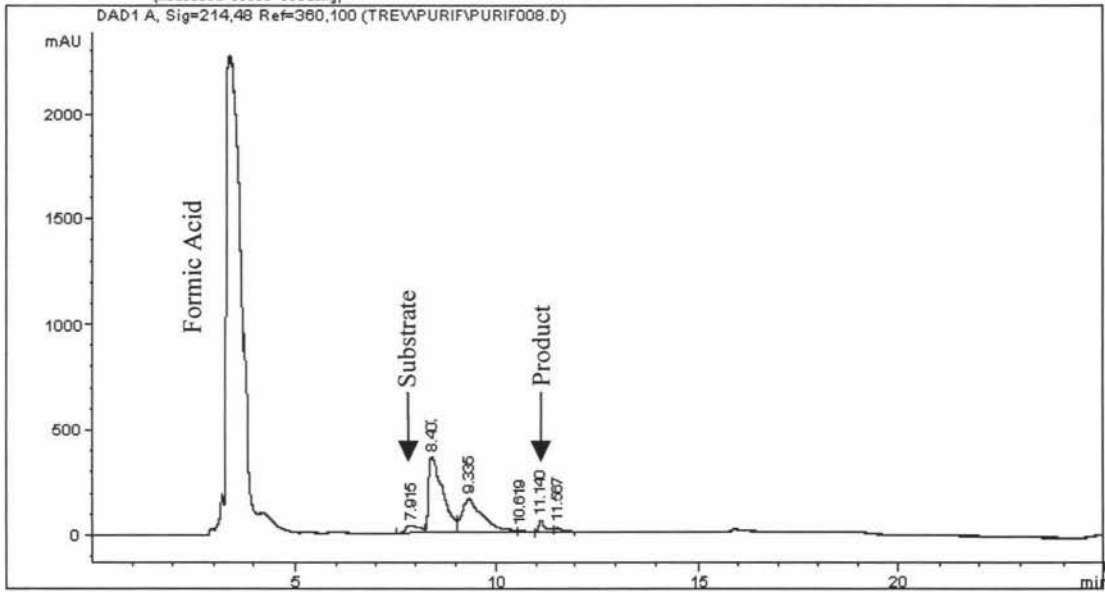
3. The resin was transferred to a 250-mL Schott bottle wrapped in aluminum foil where CDI (15 g per 10 g resin) was added and left on a shaker (200 rpm) at room temperature overnight.
4. The resin was transferred to a sintered funnel and washed with 5 volumes (250 mL) of dry dioxan to eliminate excess CDI.
5. The resin was again transferred to a clean 250-mL Schott bottle wrapped in foil before 20 mL of *t*-butylamine was added and left on a shaker (200 rpm) at room temperature overnight (~ 16 hours).
6. Excess of *t*-butylamine was removed by sequential washings of the following solutions at 1 volume every 10 minutes (1 volume = 50 mL):

Volumes of	Solvent
2	100% Dioxan
2	66% Dioxan
2	33% Dioxan
5	Milli-Q water
2	50 mM HCl
5	Milli-Q water

7. The resin was loaded into a FPLC column ((10 mm x 10 cm, Pharmacia) by gravity for chromatography or stored with 20% ethanol at 4°C.

Appendix 7 A Typical Integration Report for the Activity Assay.

=====  
Injection Date : 5/11/00 4:48:23 PM Seq. Line : 5  
Sample Name : S75 W Pool x2 Vial : 10  
Acq. Operator : Icev Inj : 1  
Inj Volume : 70 µl  
Acq. Method : D:\HPCHEM\1\METHODS\TRIV\OV-ASSAY.M  
Last changed : 5/11/00 4:47:02 PM by Icev  
(modified after loading)  
Analysis Method : D:\HPCHEM\1\METHODS\REF\_LC.M  
Last changed : 7/6/00 4:41:25 PM by Icev  
(modified after loading)



=====  
Area Percent Report  
=====

Sorted By : Signal  
Multiplier : 1.0000  
Dilution : 1.0000

Signal 1: DAD1 A, Sig=214,48 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	7.915	BV	0.3064	913.75305	37.21935	5.5414
2	8.407	VV	0.3438	8807.58691	359.49554	53.4128
3	9.335	VV	0.4626	5791.86523	159.89996	35.1242
4	10.619	VP	0.1859	94.67382	7.27054	0.5741
5	11.140	VV	0.1686	635.35162	53.49183	3.8530
6	11.567	VBA	0.1728	246.42067	19.31788	1.4944

Totals : 1.64897e4 636.69511

Results obtained with enhanced integrator!

## Appendix 8 Substrate Concentration and Product Area Relationship

This data was used to construct a standard curve for determining the amount of product created in the activity assay from product area, based on the assumption that substrate consumption is equal to product formation.

[Substrate] (mg mL <sup>-1</sup> )	[Substrate] (μmol L <sup>-1</sup> )	Product Area
0.2	36.73	211.9
0.4	73.47	511.4
0.6	110.21	844.8
0.8	146.95	1,016.9
1.0	183.68	1,192.6
2.0	367.37	2,475.7

## Appendix 9      Summary of Results in the Purification of PNGase F.

Purification Step	Product Area	[Product] ( $\mu\text{mo L}^{-1}$ )	Dil. Factor (fold)	Volume (mL)	Total Activity* (munits)	Total Protein (mg)	Specific Activity (munits $\text{mg}^{-1}$ )
Culture Medium	25.46	182.71	1	8000	292.28	8796.14	0.03
Phenyl Sepharose	127.48	868.29	1	596	103.50	61.72	1.68
<i>t</i> -butyl TSK	172.26	1,169.10	4	42	39.29	2.43	16.17
Superdex 75	120.24	819.64	10	17.3	28.36	1.36	20.85
Stirred Cell Concentrate	185.25	1,256.51	500	0.2	25.13	1.21	20.77
Periplasmic Fraction	210.05	1,423.17	4	138	157.12	70.65	2.22
IMAC	160.79	1,092.14	10	62	135.42	24.01	5.64
Superdex 75	107.97	737.18	40	13	76.67	13.74	5.58
Stirred Cell Concentrate	129.13	879.38	1,000	0.3	52.76	9.54	5.53

\*Total Activity = [Product] x Dilution Factor x Volume (in L) / 5 minutes incubation.

### Appendix 10    The Effect of Temperature on Recombinant PNGase F Activity

Temperature	Substrate Area	Substrate ( $\mu\text{mol L}^{-1} \mu\text{g}^{-1}$ )	Substrate Utilised ( $\mu\text{mol L}^{-1} \text{min}^{-1} \mu\text{g}^{-1}$ )	Product Area	Product ( $\mu\text{mol L}^{-1} \mu\text{g}^{-1}$ )	Product Formed ( $\mu\text{mol L}^{-1} \text{min}^{-1} \mu\text{g}^{-1}$ )
15	384.61	55.50	11.10	411.45	59.50	11.90
25	840.78	123.39	24.68	757.30	110.96	22.19
30	983.90	144.68	28.94	796.06	116.73	23.35
35	1,350.62	199.25	39.85	1,079.30	158.88	31.78
40	1,735.23	256.49	51.30	1,568.27	231.64	46.33
45	1,934.99	286.21	57.24	1,750.14	258.71	51.74
50	2,841.37	421.09	84.22	2,552.16	378.05	75.61
55	3,085.85	457.47	91.49	2,695.28	399.35	79.87
60	2,468.68	365.63	73.13	2,072.14	306.62	61.32
70	861.65	126.49	25.30	810.97	118.95	23.79

\*Incubation time was 5 minutes.

### Appendix 11 The Effect of pH on Recombinant PNGase F Activity

Buffer	Substrate Area	Substrate ( $\mu\text{mol L}^{-1} \mu\text{g}^{-1}$ )	Substrate Utilised ( $\mu\text{mol L}^{-1} \text{min}^{-1} \mu\text{g}^{-1}$ )	Product Area	Product ( $\mu\text{mol L}^{-1} \mu\text{g}^{-1}$ )	Product Formed ( $\mu\text{mol L}^{-1} \text{min}^{-1} \mu\text{g}^{-1}$ )
NaAc 4.5	545.61	79.46	15.89	381.63	55.06	11.01
NaAc 5.5	930.23	136.70	27.34	676.80	98.98	19.80
Mes 5.5	945.14	138.92	27.78	724.50	106.08	21.22
Mes 6.5	1,037.56	152.67	30.53	831.84	122.06	24.41
Mops 6.5	1,007.75	148.23	29.65	781.15	114.51	22.90
Mops 7.5	1,132.97	166.87	33.37	876.56	128.71	25.74
Taps 7.5	1,216.45	179.29	35.86	924.27	135.81	27.16
Taps 8.5	1,267.14	186.83	37.37	1,028.62	151.34	30.27
Capso 8.5	945.14	138.92	27.78	706.62	103.42	20.68
Capso 9.5	939.17	138.03	27.61	685.75	100.32	20.06
Caps 9.5	1,106.14	162.87	32.57	840.78	123.39	24.68
Caps 10.5	703.64	102.98	20.60	590.34	86.12	17.22

\* Incubation time was 5 minutes.

## Appendix 12 The Effects of Metal Ions and Additives on Recombinant PNGase F Activity.

Metal Salt/Additive	5 $\mu$ M		5 mM	
	Product Area	Rel. Activity (%)	Product Area	Rel. Activity (%)
AgNO <sub>3</sub>	987.1	85	696.8	60
CaCl <sub>2</sub>	1091.7	94	1033.6	89
CuSO <sub>4</sub>	894.2	77	487.8	42
FeCl <sub>3</sub>	1219.4	105	360.0	31
MgCl <sub>2</sub>	1068.4	92	1184.6	102
MgSO <sub>4</sub>	1033.6	89	1231.0	106
MnCl <sub>2</sub>	1080.1	93	1277.5	110
NaCl	1033.6	89	1114.9	96
NiCl <sub>2</sub>	1045.2	90	1196.2	103
KCl	975.5	84	929.1	80
ZnCl <sub>2</sub>	1068.4	92	1265.9	109

	Product Area	Rel. Activity (%)
Complete™ Protease Inhibitor Tablet	1149.7	99
1 $\mu$ g mL <sup>-1</sup> Leupeptin	1335.6	115
1 $\mu$ g mL <sup>-1</sup> Pepstatin	1184.6	102
1 mM PMSF	557.4	48
5 mM EDTA	1103.3	95
2 mM DTT	1207.8	104
2 mM NEM	1393.6	120
0.01% SDS	998.8	86
2.5 M Urea	1068.4	92
Control	1161.4	100

\*Activity in absence of metal ions was taken as 100% and the assay has a  $\pm$  10% error.



### Appendix 13 Summary of Results for $K_m$ Determination

(A) Initial estimate of  $K_m$  of recombinant PNGase F in 95  $\mu\text{L}$  reaction mixture.

$K_m^*$	[Substrate] ( $\mu\text{mol L}^{-1}$ )	Product Area	[Product ] ( $\mu\text{mol L}^{-1}$ )	Velocity ( $\mu\text{mol L}^{-1} \text{ min}^{-1}$ )
0.25	15.37	77.54	11.49	2.30
0.50	30.63	82.70	12.26	2.45
1.00	61.26	85.81	12.72	2.54
2.00	122.11	88.93	13.18	2.64
4.00	245.26	92.00	13.64	2.73
8.00	490.53	94.08	13.95	2.79

\* Estimated  $K_m$  value of  $0.3 \text{ mg mL}^{-1}$  based on Altman *et al.*, (1995).

## Appendix 13 Summary of Results for $K_m$ Determination

(B)  $K_m$  determination of dialysed recombinant PNGase F in 495  $\mu\text{L}$  reaction mixture.

$K_m^*$	[Substrate] ( $\mu\text{mol L}^{-1}$ )	Product Area	Internal Standard Area†	[Product] ( $\mu\text{mol L}^{-1}$ )	Corrected Product†	Velocity ( $\mu\text{mol L}^{-1} \text{ min}^{-1}$ )
0.25	0.59	3.97/3.87	3639.8/3557.8	0.581/0.567	0.58	0.12
0.50	1.17	7.64/7.73	3683.0/3767.3	1.126/1.139	1.11	0.22
1.00	2.34	13.37/13.42	3572.1/3626.8	1.976/1.983	2.01	0.40
2.00	4.69	14.39/14.26	3658.4/3596.6	2.127/2.108	2.13	0.43
4.00	9.41	10.67/10.52	3773.0/3620.8	1.576/1.553	1.55	0.31
8.00	18.83	5.01/5.071	3580.5/3613.7	0.734/0.745	0.75	0.25

\* Estimated  $K_m$  value of 0.3  $\text{mg mL}^{-1}$  based on Altman *et al.*, (1995).

Correction factor = Average of internal standard areas / Internal standard Area in  $K_m$  determination  
 = 3652.7 / Internal standard Area in  $K_m$  determination

Appendix 14 Parameters for ES-MS Quantitation of Deglycosylated Products.

A. Scanning parameters used for kinetic analysis of deglycosylated ribonuclease products by ES-MS in section 4.5.4.3.

Parameter	Value		Parameter	Value
NEB	11		RO <sub>1</sub>	-11
CUR	8		IQ <sub>2</sub>	-50
CAD	0		RO <sub>2</sub>	-100
IS	5,000		IQ <sub>3</sub>	-150
OR	115		DF	-100
RNG	360		CEM	4,900
Q <sub>0</sub>	-10		Scan Type	Pos Q <sub>1</sub>
IQ <sub>1</sub>	-11		Peak Hopping	Disabled
ST	-15		Pause Time	5.0 msec

B. Scanning parameters used for pre-steady state kinetic analysis of PNGase F deglycosylation reaction by ES-MS in section 4.5.4.4.

Parameter	Value		Parameter	Value
NEB	11		RO <sub>1</sub>	-11
CUR	8		IQ <sub>2</sub>	-20
CAD	0		RO <sub>2</sub>	-100
IS	5,000		IQ <sub>3</sub>	-150
OR	20		DF	-100
RNG	118.5		CEM	4,900
Q <sub>0</sub>	-10		Scan Type	Pos Q <sub>1</sub>
IQ <sub>1</sub>	-11		Peak Hopping	Disabled
ST	-15		Pause Time	5.0 msec