A Study of a Proline Specific Seprase Activity from Mammalian Serum

Thesis submitted for the degree of
Doctor of Philosophy

by

Pamela O'Brien, B.Sc.

Supervised by

Dr. Brendan F. O'Connor

School of Biotechnology
Dublin City University
Ireland

August 2006

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Pamela O'Brien ID No.: 98399080

Pamela O'Brien

Date: 19th September 2006

Dedications

I would like to dedicate this thesis, the hard work and long hours that went into completing it to……

Brendan for giving me the opportunity to take on this challenge. Your enthusiasm helped me through those 'groundhog days' and made the whole experience an enjoyable and memorable one.

My parents and big sister Jen. Thank you for all your encouragement and support over the years. My student days are now over! Thank god I hear you say!

Daragh, thank you for being there to listen, encourage and support me when I needed you. Thank you for being you!

Zelda, for keeping me company in the lab and for our chats! It kept me going on those long days.

The DCU gang. Our coffee breaks, lunches and nights out were always entertaining – cheers! And for those who opened their labs up to me – I thank you!

To Paraic and Paul….what can I say boys. Thank you both for all your help in decoding the genetic code for me! You have both provided me with great entertainment, in and out of the lab! Paraic, your theories were enlightening.

To Roman and Pat. Your help and guidance when I started this project was greatly appreciated.

And to the girls, Aoife, Anita and Niamh. For your friendship and most importantly for introducing me to the joys of Harry Potter back in the BT days!

Thank you!

Publications

Papers

Collins PJ, McMahon G, O'Brien P, O'Connor B. (2004) Purification, identification and characterisation of seprase from bovine serum. *Int J Biochem Cell Biol.* **36**(11):2320-2333.

O'Brien, P and O'Connor B. Review: Seprase, A strategically important Matrix Serine Protease. *BBA Reviews.* [Manuscript in preparation].

O'Brien P, O'Donoghue A, McMahon G, Craik C, O'Connor B. Characterisation of the active site of Seprase, a novel post proline cleaving endopeptidase. *Proc Natl Acad Sci.* [Manuscript in preparation].

Posters

O'Brien P, Collins PJ, O'Donoghue A, Craik CS, and O'Connor B. *Identification and Characterisation of Seprase from Bovine Serum.* IACR Annual meeting, Galway. March 2006.

O'Brien P and O'Connor B. *Molecular Cloning and Characterisation of Seprase.* National IRCSET Symposium, Croke Park, Dublin. November 2005.

O'Brien P and O'Connor B. *The Purification and Characterisation of Seprase from Bovine Serum.* National IRCSET Symposium, Croke Park, Dublin. November 2004.

O'Brien P and O'Connor B. *The Purification and Characterisation of Seprase from Bovine Serum.* National IRCSET Symposium, Royal Hospital Kilmainham, Dublin. November 2003.

Presentations

O'Brien P and O'Connor B. (2005) "*Seprase – A potential Biomarker for Breast Cancer.*" Breast Check, Mater Misericordiae Hospital, Dublin.

O'Brien P and O'Connor B. (2006) "*Review of the Clinical Research Study on Seprase levels in IDC.*" Breast Check, Mater Misericordiae Hospital, Dublin.

Abbreviations

A_{600}	Absorbance at 600 nm
AA	Amino Acid
Ac	Acetyl
AEBSF	4-(2-aminoethyl)-benzenesulfonyl fluoride
ACC	7-amino-4-carbamoylmethylcoumarin
AFC	7-amino-4-trifluoromethylcoumarin
AFP	alpha-fetoprotein
α_2AP_{PRO}	α_2-antiplasmin, 'pro' form, 464 amino acid residues
α_2AP_{ACT}	α_2-antiplasmin, mature form, 452 amino acid residues
AMC	7-amino-4-methylcoumarin
AMCC	7-amino-4-methyl-3-carbamoylmethylcoumarin
AP	α_2-antiplasmin
APCE	Antiplasmin Cleaving Enzyme
APMSF	4-amidino-phenylmethane-sulphonyl fluoride
APSF	4-amidino phenylsulfonyl fluoride
BCA	Bicinchoninic Acid
BCIP	5-Bromo-4-chloro-3-indolyl phosphate, toluidine salt
Bisacryl	Bisacrylamide
BoroPro	Prolineboronic acid
bp	Base-pair
BSA	Bovine Serum Albumin
BT	Benzothiazol
Bz	Benzoyl
CA	Cancer antigen
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
cDNA	copy Deoxyribonucleic acid
CEA	Carcinoembryonic antigen
cmk	Chloromethyl-ketone
CMV	Cytomegalovirus
CN	2-nitrile
Conc.	Concentration
CPC	Calcium Phosphate Cellulose

DABCYL	4-(4-dimethylaminophenylazo)benzoyl
DAPI	4',6-diamidino-2-phenylindole
DEPC	Diethyl Pyrocarbonate
DIC	Differential Interference Contrast
DICI	Diisopropylcarbodiimide
DFP	Diisopropyl fluorophosphates
dH_2O	Deionised water
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPIV	Dipeptidyl peptidase IV
DPPIV	Dipeptidyl peptidase IV
DTT	Dithiothreitol
EBV	Epstein Barr Virus
EC	Enzyme commission
eds.	Editors
ECM	Extracellular matrix
EDANS	5-[(2-aminoethyl)amino]-naphthalene-1-sulfonic acid
EDTA	Ethylenediaminetetra acetic acid
EH	Eadie-Hofstee
EtOH	Ethanol
FAPα	Fibroblast Activation Protein-alpha
FGF-2	Fibroblast growth factor 2
FITC	Fluorescein isothiocyanate
FLAG	Octapeptide - DYKDDDDK
Fmoc	9-fluorenylmethyloxycarbonyl
FPLC	Fast Protein Liquid Chromatography
FRET	Fluorescence Resonance Energy Transfer
HATU	O-(7-Azabenzotriazole-1yl)-N,N,N'N'-tetramethyluronium hexafluorophosphate
HCl	Hydrochloric Acid
HCV	Hepatitis C virus
HIC	Hydrophobic Interaction Chromatography
HOBt	N-hydroxy benzotriazole

HPLC	High Performance Liquid Chromatography
HR	High Resolution
HW	Hanes-Woolf
IC_{50}	Inhibitor conc. resulting in 50% loss in enzymatic activity
IDC	Infiltrating ductal carcinoma
IEF	Isoelectric focusing
IgG	Immunoglobulin G
IMAC	Immobilised metal affinity chromatography
IPF	Idiopathic pulmonary fibrosis
IPTG	Isopropyl-β-$_D$-thiogalactopyranoside
JTP-4819	(S)-2-[[(S)-2-(hydroxyacetyl)-1-pyrrolidinyl]carbonyl]-N-[phenylmethyl]-1-pyrrolidinecar-boxamide
k_{cat}	Turnover number
K_i	Inhibition Constant
K_m	Michaelis constant
K_m^{app}	Apparent K_m
LB	Lineweaver-Burk
LB	Luria Bertani
Log	Logarithm
LOX	Human melanoma cell line
mAb	Monoclonal antibody
MCS	Multiple cloning site
MeOH	Methanol
MMP	Matrix metallo-proteinase
MOPS	3-(N-Morpholino)propanesulfonic acid
mRNA	Messenger RNA
MW	Molecular Weight
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NBT	Nitro blue tetrazolium chloride
ND	Not determined
NEM	N-ethylmaleimide
NPY	Neuropeptide Y
NTA	Nitrilotriacetic acid

OA	Osteoarthritis
OD	Optical density
ORF	Open reading frame
P	Statistical significance
PAGE	Polyacrylamide gel electrophoresis
PAP1	Pyroglutamyl Peptidase I
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PDB	Protein data bank
PE	Prolyl Endopeptidase
PEP	Prolyl Endopeptidase
pH	log of the reciprocal of the hydrogen ion concentration
pI	Isoelectric point
PIPES	Piperazine-1,4-bis(2-ethanesulfonic acid)
PMSF	Phenylmethylsulphonyl fluoride
PO	Prolyl oligopeptidase
POP	Prolyl oligopeptidase
PVDF	Polyvinylidene fluoride
Pyrr	Pyrrolidide
RBS	Ribosome binding site
RCO	Acyl
RET	Resonance energy transfer
RFU	Relative Fluorescence Units
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RPMI	Roswell Park Memorial Institute medium
RT-PCR	Reverse transcription PCR
S9	Prolyl oligopeptidase family of serine proteases
SDS	Sodium dodecyl sulphate
SEM	Standard Error Mean
Serpin	Serine protease inhibitor
SIMP	Serine Integral Membrane Protein
SMC	Smooth Muscle Cell
S-S	Disulphide bridges

Suc-	Succinyl
T_{ann}	Annealing Temperature
TAE	Tris-Acetate/EDTA solution
TBS	Tris Buffered Saline
TBS-T	Tris Buffered Saline solution containing Tween 20
TE	Tris/EDTA solution
T/E	Trypsin/EDTA solution
TEMED	N, N, N, N'-tetramethyl ethylenediamine
TIS	Triisopropylsilane
TFA	Trifluoroacetic acid
T_m	Temperature of melting point
TRH	Thyrotropin-releasing hormone
TRITC	Tetramethylrhodamine isothiocyante
Tris	Tris (hydroxymethyl) amino methane
tRNA	Transfer RNA
uPA	Urokinase plasminogen activator
uPAR	Urokinase plasminogen activator receptor
UV	Ultraviolet
v/v	Volume per volume
V_{max}	Maximal enzyme velocity
WGA	Wheat germ agglutinin
WG	Wire guided
WLE	Wide local excision
w/v	Weight per volume
Xaa	Any amino acid
X-Gal	5-bromo-4-chloro-3-indolyl-β-$_D$-galactopyranoside
Yaa	Any amino acid
Z-	*N*-benzyloxycarbonyl
Zaa	Any amino acid
ZIP	Z-Pro-prolinal insensitive Z-Gly-Pro-AMC degrading peptidase

Units

A	Ampere
Å	Angstrom (1.0×10^{-10} metres)
Da	Dalton
g	g-force (relative centrifugal force - rcf)
g	Gram
h	Hour
Hz	Hertz
k	Kilo
kb	Kilo base
L	Litre
m	Metre
M	Molar
min	Minute
mM	Millimolar
°C	Degrees Celsius
pH	Logarithm of reciprocal hydrogen-ion concentration
pI	Isoelectric point
rpm	Revolutions per minute
sec	Second
V	Volts

Prefixes

k	kilo (1×10^{3})
c	centi (1×10^{-2})
m	milli (1×10^{-3})
μ	micro (1×10^{-6})
n	nano (1×10^{-9})
p	pico (1×10^{-12})

Amino Acid Abbreviations

Ala	**A**	Alanine
Arg	**R**	Arginine
Asn	**D**	Asparagine
Asp	**N**	Aspartic acid
Cys	**C**	Cysteine
Gln	**Q**	Glutamine
Glu	**E**	Glutamic Acid
Gly	**G**	Glycine
His	**H**	Histidine
Ile	**I**	Isoleucine
Leu	**L**	Leucine
Lys	**K**	Lysine
Met	**M**	Methionine
Nle	n	Norleucine
Phe	**F**	Phenylalanine
Pro	**P**	Proline
Ser	**S**	Serine
Thr	**T**	Threonine
Trp	**W**	Tryptophan
Tyr	**Y**	Tyrosine
Val	**V**	Valine

Table of Contents

List of Figures

Chapter 6

Chapter 7

Chapter 8

List of Tables

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Abstract

Seprase *or Fibroblast activation protein (FAP) is an integral membrane serine peptidase, which has been shown to have gelatinase activity. It appears to act as a proteolytically active 170-kDa dimer, consisting of two 97-kDa subunits. It is a member of the group type II integral serine proteases, which include dipeptidyl peptidase IV (DPPIV/CD26) and related type II trans-membrane prolyl serine peptidases, which exert their mechanisms of action on the cell surface. DPPIV and Seprase exhibit multiple functions due to their abilities to form complexes with each other and to interact with other membrane-associated molecules. Localization of these protease complexes at cell surface protrusions, called invadopodia, may have a prominent role in processing soluble factors and in the degradation of extracellular matrix components that are essential to the cellular migration and matrix invasion that occur during tumour invasion, metastasis and angiogenesis.*

Seprase was isolated and purified from bovine serum, yielding a specific activity of 166.41 units/mg. The purified soluble Seprase was shown to have both exopeptidase and endopeptidase activity. The soluble form of the glycoprotein was shown to bind to the Wheat Germ Agglutinin (WGA)-Lectin Affinity Chromatography. Biochemical studies show that Seprase has a pH optimum of 8.0 and it is thermostable (up to 40 ^{o}C*). The second order rate constant,* k_2 *for DFP inhibition of Seprase was determined to be 3.31 x* 10^3 $M^{-1}s^{-1}$*. Positional scanning of the* P_1 *dipeptide library determined that Seprase has a marked preference for proline in the* S_1 *subsite. Data obtained during screening of the* P_2 *sub-library revealed a much broader specificity in the* S_2 *binding pocket of bovine Seprase. In the* S_2 *subsites, this study shows that Seprase has a preference for Norleucine, Alanine, Leucine, Glycine, Arginine, Methionine and does not tolerate aromatic, strongly basic or acidic residues. The kinetic constant* K_m *determined for purified Seprase was 82.10*μM*. Seprase was found to have an average* k_{cat} / K_m *ratio of 1.17 x* 10^5 $M^{-1}s^{-1}$ *for cleavage of the fluorimetric substrate Z-Gly-Pro-AMC. Tissue localisation studies identified Seprase activity in bovine large intestine, serum, kidney, liver and spleen and also in the breast cancer cell line Hs578T. Seprase was successfully cloned into prokaryotic and eukaryotic expression systems. Clinical studies on serum samples from breast cancer patients indicated that Seprase levels are elevated in patients with invasive ductal carcinoma.*

Chapter 1

Introduction

1.1 Cancer

Cancer is a class of diseases or disorders characterized by uncontrolled division of cells and the ability of these cells to invade other tissues, either by direct growth into adjacent tissue through *invasion* or by implantation into distant sites by *metastasis*. The word cancer comes from the Greek word 'Karkinos' meaning crab, which refers to the metaphorical claws reaching out to invade surrounding tissues. 'Cancer' generally refers to a malignant tumour or neoplasm. Malignant tumours possess potentially lethal abnormal characteristics enabling them to invade and to metastasise, or spread, to other tissues.

Each year there are over 19,000 new cancer cases and 11,000 cancer deaths in Ireland (The National Cancer Registry Ireland www.ncri.ie). The four commonest types of cancer are breast, colorectal, lung and prostate. For women in Ireland, incidence rates of breast, colorectal, lung, lymphoma cancers and melanoma of the skin are significantly higher than the rates for other women in the EU. For women, breast cancer is the leading type of cancer and the leading cause of cancer-related death. Breast cancer currently affects over 1,700 women in Ireland every year (The Irish Cancer Society www.cancer.ie).

The earlier breast cancer is diagnosed and treated, the better the long-term prospects for women with the disease. It is important when patients are diagnosed with cancer that doctors know whether the disease is local or has spread to other locations. It is this ability to spread to other tissues and organs that makes cancer a possibly life-threatening disease. Therefore it is of great interest to scientists and clinicians to understand what makes metastasis possible for a cancerous tumour.

1.1.1 Cancer Development

Cancer development is a multi-step process driven in most part by genetic change. Mutation of genes can be either inherited or acquired. Irreversible damage to DNA leads to the formation of a cell containing DNA errors, which can multiply to form additional mutated cells. Under normal conditions DNA repair may revert the cell back to a normal cell. In other cases programmed cell death or apoptosis is initiated. Failure in each case usually results in carcinogenesis.

However, it is becoming increasingly apparent that the tumourigenic phenotype must overcome the suppressive effects of the surrounding microenvironment. Solid tumours consist of *neoplastic cells* and *stroma*. The neoplastic cell population is embedded in and supported by a connective tissue framework called the stroma (from the Greek word meaning mattress), which provides mechanical support and nutrition to the neoplastic cells.

Under normal conditions tissues communicate through a complex network of interactions which are necessary for cellular differentiation and to create complex tissue structures. This may be (i) physically, through direct contact or through the intervening extracellular matrix (ECM) or (ii) biochemically through soluble/insoluble signalling molecules. These intracellular signals may become disrupted, for example, by the activation of fibroblasts due to wound healing (Liotta et al., 1991). These conditions are normally temporary and reversible. However, when inflammation is sustained, continuous upregulation of enzymes (such as matrix metalloproteinases MMPs) by stromal fibroblasts can disrupt the ECM, and invading immune cells can overproduce factors that promote abnormal proliferation. The transition from normal to invasive carcinoma is preceded by, or is concomitant with, activation of the local host stroma (see Figure 1.1.1) (Liotta and Kohn, 2001).

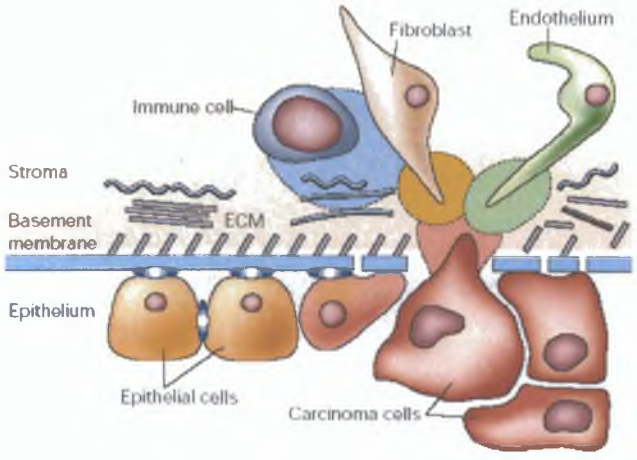


Figure 1.1.1 Microecology of the tumour-host invasion field

Transition of the invasive carcinoma is preceded by the activation of host fibroblasts, immune cells and endothelial cells. Invasion occurs in a localised area of cross-talk and cooperation between the stromal cells and the pre-malignant epithelium (depicted as zones demarked by dashed lines). Cytokine and enzyme exchange between the participating cells stimulates migration of both cell types towards each other and modifies the adjacent extracellular matrix/basement membrane (ECM). The result is the breakdown of normal tissue (Liotta and Kohn, 2001)

Wound healing and tumour development are dynamic progressive processes that involve the interaction of several tissue types and have many mechanistic similarities. The composition of tumour stroma markedly resembles that of wound granulation tissue, although a distinguishing feature of the tumour stroma is the absence of platelets and a lower density of inflammatory cells (Sieweke and Bissell, 1994). In cancer, these changes in the stroma drive invasion and metastasis (formation of secondary tumours), the hallmarks of malignancy (see Figure 1.1.2). Bissell and Radisky (2001) describe how tumours activate some of the normal wound-healing responses. Tumour cells produce many of the same growth factors that activate the adjacent stromal tissues as in wounding or fibrosis. Activated fibroblasts and infiltrating immune cells (macrophages) secrete proteases (MMPs) and cytokines such as FGF-2 (fibroblast growth factor-2). These factors potentiate tumour growth, stimulate angiogenesis and induce fibroblasts to undergo differentiation into myofibroblasts (cells that share the characteristics of fibroblasts) and into smooth muscle (Bissell and Radisky, 2001).

Dingemans et al., (1993) transplanted colon carcinomas into granulation tissue which induced an invasive phenotype, thereby demonstrating that cancer invasion is stimulated by wound-healing stroma.

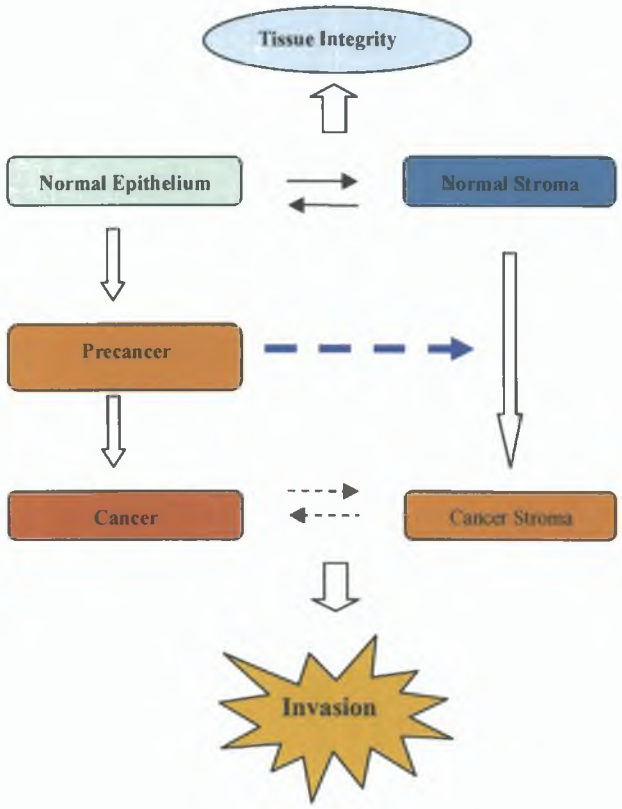


Figure 1.1.2 Summary of Intracellular Signalling

Signalling from the normal epithelial cells to the stroma and back (black thin arrows) maintains the integrity of the epithelial tissue (thick arrow). During epithelial carcinogenesis, the signalling changes (dotted blue arrow) and this causes cancer-associated changes in the stroma. The new cross-talk (dotted black arrows) between the cancer cells and the stroma leads to invasion (thick arrow). Adapted from De Wever and Mareel, (2003).

Cell surface proteases play an important role in facilitating cell invasion into the extracellular matrix. Proteases associate at plasma membrane protrusions, called invadopodia, which contact and dissolve the matrix. Invadopodia degrade a variety of immobilised substrates including fibronectin, laminin and type I collagen (Monsky et al., 1994). Integral membrane proteases may contribute significantly to ECM degradation by metastatic cells by virtue of their localisation at invadopodia which are in contact with

the ECM. Integral membrane proteases can be defined as a group of cell surface glycoproteins that contain extracellular domains of either metallo- or serine-proteases, a transmembrane domain and a short cytoplasmic tail. Examples of such transmembrane glycoproteins include meprin, matrix metalloproteinase, DPPIV, fibroblast activation protein α (FAPα) and Gbase. Abundant expression of these enzymes is associated with poor prognosis (Iwasa, 2005).

This study will take a detailed look at the serine protease Seprase / fibroblast activation protein α (FAPα), a 170kDa gelatinase, which belongs to the S9b peptidase family (Abbott and Gorrell, 2002; Barrett et al., 1998).

1.2 Fibroblast Activation Protein α / Seprase

Fibroblast activation protein α (FAPα) and/or Seprase (*s*urface *e*xpressed *p*rotease) are a 170kDa integral membrane gelatinase. FAPα is an inducible cell surface glycoprotein that was originally identified in cultured fibroblasts using monoclonal antibody F19 (Garin-Chesa et al., 1990; Rettig et al., 1993). Seprase was originally identified as a glycoprotein peptidase selectively expressed on the surface of invadopodia and was isolated from a human malignant melanoma cell line LOX by Pineiro-Sanchez et al., (1997). The expression of Seprase correlates with the invasiveness of human melanoma and carcinoma cells. Molecular cloning of FAPα and Seprase revealed that they are the same cell surface serine protease which is found on chromosome 2q23 (Chen et al., 2003; Chen and Kelly, 2003; Mathew et al., 1995; Pineiro-Sanchez et al., 1997; Scanlan et al., 1994). For the clarity of this review, the protease is referred to as Seprase throughout.

1.3 Classification of Seprase

Seprase (EC 3.4.21.-) belongs to the small family of serine integral membrane peptidases (SIMP). These peptidases are inducible, specific for proline-containing peptides and macromolecules and active on the cell surface (Chen et al., 2003). Post-prolyl peptidases modify bioactive peptides and change their cellular functions. This class of peptidases have important roles in cancer (Busek et al., 2004; Chen et al., 2003; Chen, 2003; Chen and Kelly, 2003; Rosenblum and Kozarich, 2003). This group of enzymes also includes prolyl endopeptidase, dipeptidyl peptidase 8 and dipeptidyl peptidase IV-β (Rosenblum and Kozarich, 2003). However, the best studied of this class of enzymes is dipeptidyl peptidase IV (DPPIV or CD26) (EC 3.4.14.5) (Rosenblum and Kozarich, 2003). Studies have shown the importance of DPPIV in regulating tumour cell behaviour and function (Bauvois, 2004). Seprase shows up to 52% homology to DPPIV, both being members of the S9b peptidase family (Goldstein et al., 1997). Seprase is, therefore, a member of the DPIV-like gene family (Abbott and Gorrell, 2002) grouped in subfamily S9B of the peptidase family S9 (prolyl oligopeptidase family), clan SC (Barrett et al., 1998). Even though all SIMP members are known to cleave prolyl peptide

(Pro-Xaa) bonds there are conflicting reports on possible dipeptidyl peptidase activity associated with Seprase but its main distinguishing feature is its gelatinase activity (Gorrell et al., 2001; Park et al., 1999). An early report had suggested that DPPIV had gelatinase activity (Bermpohl et al., 1998); however, more recent reports suggest otherwise (Gorrell et al., 2001; Park et al., 1999). The roles Seprase plays are only beginning to be understood but insights into the potential functions of Seprase can be obtained from the vast amount of work done on DPPIV.

1.4 Structure of Seprase and Biochemical Aspects

Active Seprase is a 170kDa homodimer that contains two N-glycosylated 97-kDa subunits. The 760-amino acid Seprase protein (GenBank GI 1888316) is a type II integral membrane protein with a large C-terminal extracellular domain. Seprase has been shown to be shed from the cell surface and recent studies by our group have identified a serum form of the protease (Collins et al., 2004). A second group has recently identified the soluble form of Antiplasmin Cleaving Enzyme (APCE) as Seprase (Lee et al., 2005a). The crystal structure of the extracellular domain of Seprase has recently been solved (see Figure 1.4.2) (Aertgeerts et al., 2005). This provides information for the substrate specificity of Seprase. The Seprase monomer has 5 potential N-glycosylation sites, 13 cysteine residues, 3 segments that correspond to highly conserved catalytic domains of serine proteases, a hydrophobic transmembrane segment and a short cytoplasmic tail (6 amino acids) (see Figure 1.4.1).

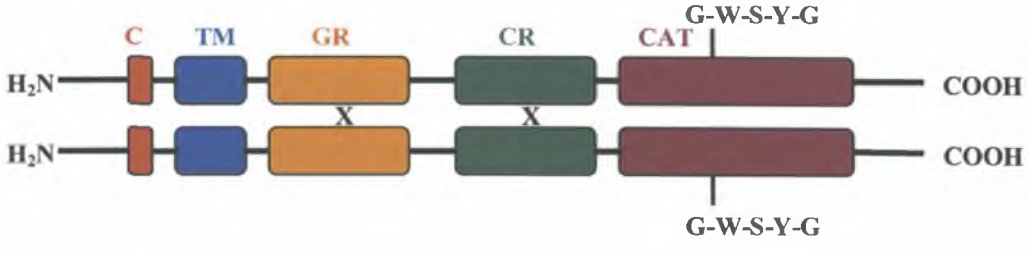


Figure 1.4.1 General structure of Serine-Integral Membrane Peptidases (SIMP)

SIMP prototypes such as Seprase and DPPIV are proteolytically active as dimers and have a cytoplasmic domain (C); transmembrane domain (TM); glycosylation-rich region (GR); cysteine-rich region (CR); and catalytic region (CAT). X represents the disulphide bonds. Adapted from Chen et al., (2003).

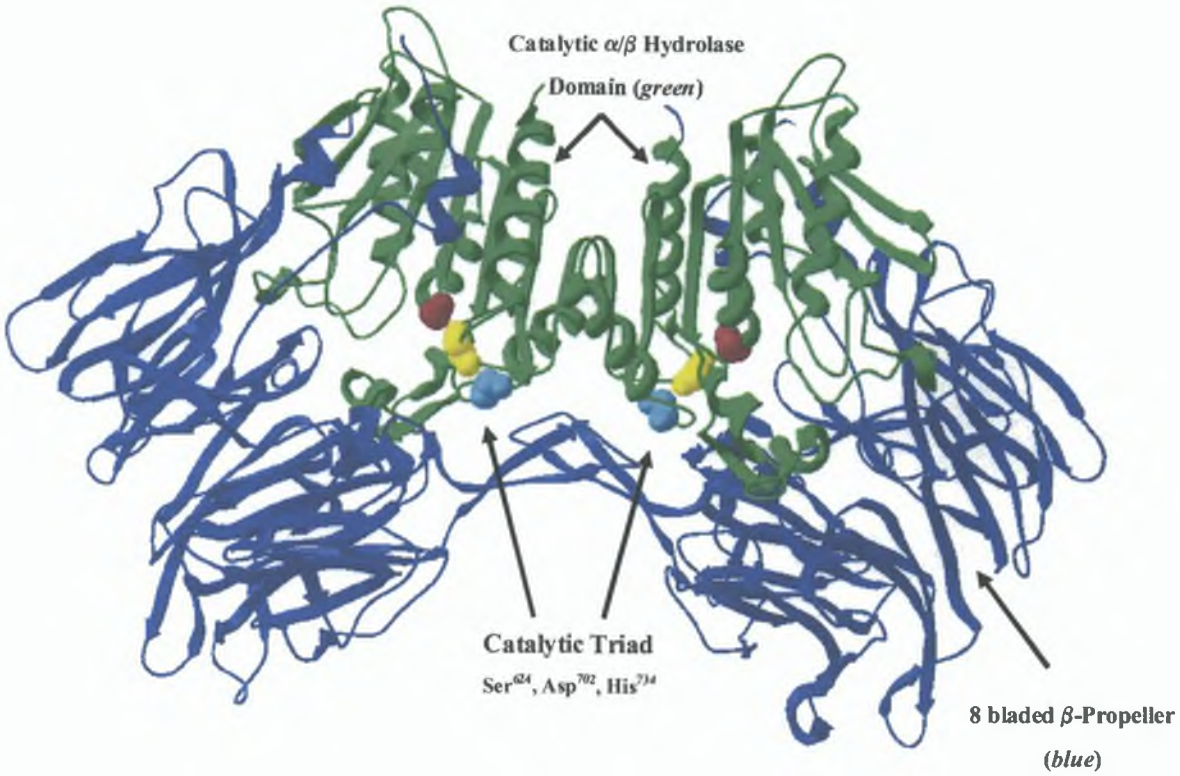


Figure 1.4.2 Three-Dimensional Structure of Seprase

The ribbon diagram illustrates the dimeric structure of Seprase (pdb accession code 1Z68). The extracellular domain consists of 2 domains, an eight bladed β-propeller domain (indicated in blue) and an α/β hydrolase domain (indicated in green) that contains the catalytic triad. The catalytic residues are shown, catalytic **Serine 624**, Aspartic Acid 702 and Histidine 734. (Aertgeerts et al., 2005)

Each subunit contains two topologically distinct domains; the β-propeller (residues 54-492) and the α/β-hydrolase domain (residues 27-53 and 493-760) (see Figure 1.4.2). The catalytic triad is located at the interface of the β-propeller and the α/β-hydrolase domain. The arrangement of the catalytic triad in the order nucleophile-acid-base is a characteristic of the α/β hydrolase domain (Ollis et al., 1992). This domain features mostly parallel β-sheets connected by α-helices on either surface of the sheet (see Figure 1.4.3). Each blade of the β-propeller domain comprises a three- or four-stranded anti-parallel β-sheet (see Figure 1.4.4). The sheets are twisted and radially arranged around their ventral tunnel. The eight bladed β-propeller domain is situated on top of the catalytic triad and may serve as a "gate" to selectively filter protein access to the catalytic triad (Aertgeerts et al., 2005; Cheng et al., 2002). The β-propeller domain in

Prolyl Oligopeptidase has been shown to regulate proteolysis (Fulop et al., 1998). The oscillating propeller blades have been shown to act as a gating filter during catalysis, letting small peptides substrates into the active site while excluding large proteins to prevent accidental proteolysis in the cytosol (Fulop et al., 2000). The active site is accessible in two ways, through a cavity formed between the β-propeller and the hydrolase domain. The side opening has a diameter of ~24Å in contrast to the narrower β-propeller opening (~14Å).

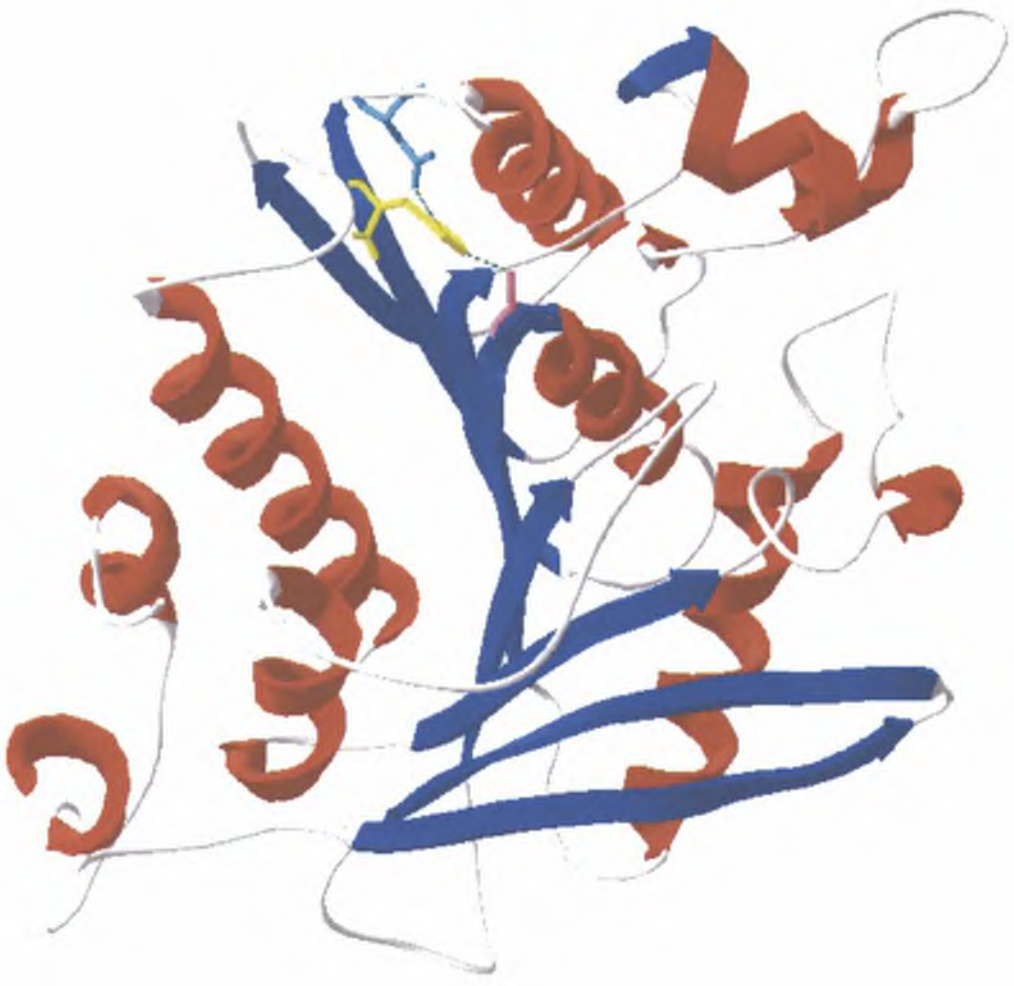


Figure 1.4.3 α/β hydrolase domain of Seprase

The ribbon diagram illustrates the α/β hydrolase domain (residues 27-53 and 493-760) of Seprase (pdb accession code 1Z68). The α/β hydrolase domain contains the catalytic triad. The catalytic residues are shown, catalytic **Serine 624**, **Aspartic Acid 702** and **Histidine 734** (Aertgeerts et al., 2005). α-helices are indicated in red, β-sheets are indicated in blue and the hydrogen bonds are indicated in green. For amino acid information see Appendix B. Generated using DeepView (Section 2.10.10).

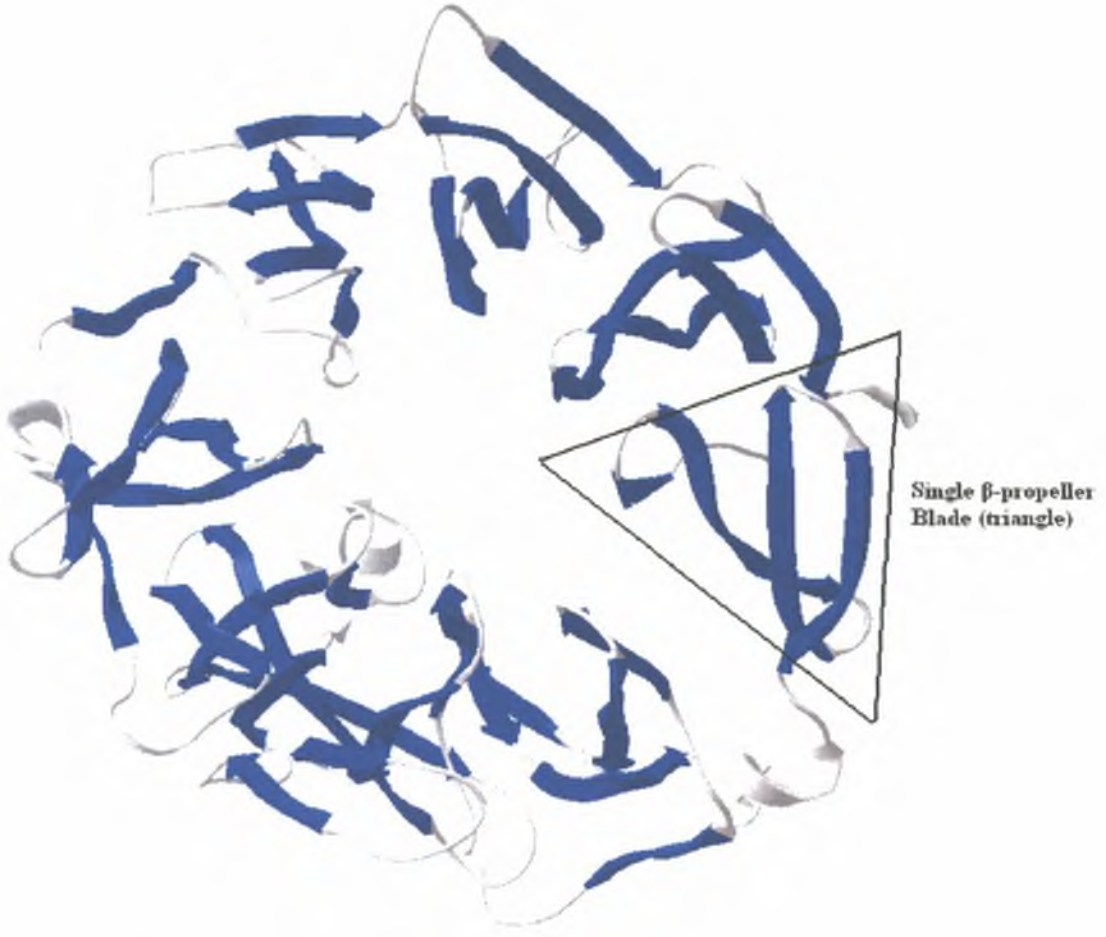


Figure 1.4.4 Eight bladed β-propeller domain of Seprase

The ribbon diagram illustrates the eight bladed β-propeller domain (residues 54-492) of Seprase (pdb accession code 1Z68). Generated using DeepView (Section 2.10.10).

DPPIV shows similar structure homology to Seprase (see Figure 1.4.1). In DPPIV the N-terminal hydrophobic sequence represents an uncleavable signal peptide that also functions as a membrane-anchoring domain (Barrett et al., 1998; Ogata et al., 1989). In Seprase the N-terminal domain possibly has a similar role as a signal peptide although there is no published data to support this. A highly conserved residue Asp^{599} in DPPIV has been shown to be important in enzyme processing such as proper folding, dimerisation and transport (Chen et al., 2003). A mutation in this residue (D599A) specifically decreased the cell surface expression of DPPIV in stably transfected mouse fibroblasts. Seprase also has this conserved Asp^{599} residue (see Figure 1.4.5); therefore, it is possible to conclude that this residue is also important in the processing of the Seprase enzyme.

The Seprase gene has been observed in several species (see Figure 1.4.5). A mouse homologue has been identified (Niedermeyer et al., 1998) as well as a *Xenopus laevis* homologue (Brown et al., 1996). The mouse Seprase gene spans approximately 60kb and contains 26 exons ranging in size from 46bp to 195bp. This genomic organisation is similar to that of the human Seprase gene (Niedermeyer et al., 1998). The catalytic serine residue arranged within the consensus sequence G-X-S-X-G is split between two exons. Gly-Trp is located at the very end of exon 21 and Ser-Tyr-Gly at the beginning of exon 22. This arrangement differs from the typical serine protease where the complete serine consensus site is encoded within one exon. The study of the mouse homologue has shown alternative splicing and 3 distinct Seprase splice variants have been detected in tissues (Niedermeyer et al., 1997). An alternative spliced Seprase was later identified in the human melanoma cell line LOX which encodes a novel truncated isoform (Goldstein and Chen, 2000). The splice variant encodes for a 239 amino acid polypeptide with a molecular weight of 27kDa that precisely overlaps the carboxyl-terminal catalytic region of the wild type Seprase.

An alignment of eukaryotic Seprase C-terminal amino acid sequences is shown in Figure 1.4.5. Sequence similarity is represented in gray scale shading, with black being the highest similarity. Figure 1.4.5 illustrates that the C-terminal catalytic region of Seprase is highly homologous throughout the different species.

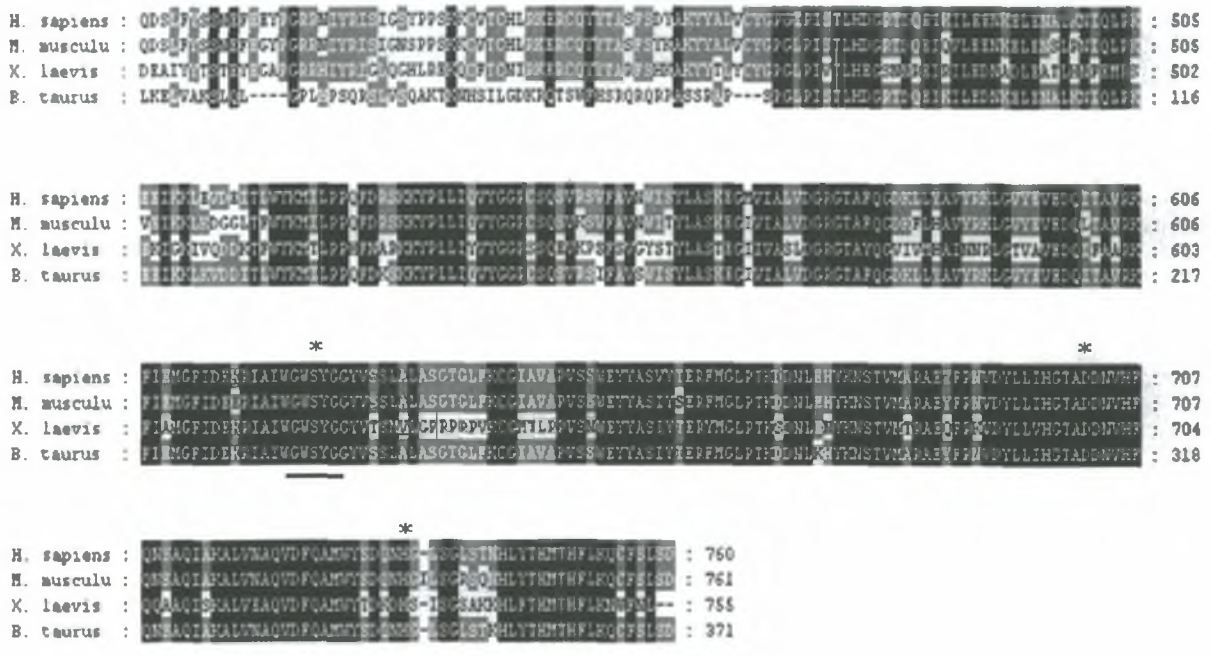


Figure 1.4.5 Amino acid sequence alignment of eukaryotic Seprase C-terminal

Alignment of amino acid sequences for the C-terminal of Seprase; deduced sequences *Homo sapiens* (AAC51668), *Mus musculus* (AAH19190), *Xenopus laevis* (AAC59872) and the putative sequence *Bos Taurus* (XP_603457). Alignment was performed using MultAlin and edited using GenDoc (Section 2.10.10). Sequence similarity is represented by grey scaling, with black being the highest similarity. The amino acids of the catalytic triad (Ser^{624}, Asp^{702}, His^{734}) are indicated (*). The serine protease consensus motif G-X-S-X-G is underlined. For amino acid information see Appendix B.

The charged N-terminal end of substrate peptides is recognised by two glutamates (Glu motif). Comparison of the crystal structure of Seprase and DPPIV revealed one major difference in the vicinity of the Glu motif (Glu^{203}-Glu^{204} for Seprase; Glu^{205}-Glu^{206} for DPPIV) within the active site of the enzyme (see Figure 1.4.6). The importance of the Glu motif in DPPIV catalysis has been confirmed by single point mutations that abolish the enzyme's aminopeptidase activity (Abbott et al., 1999). Detailed comparison of Seprase and DPPIV revealed that the Ala^{657} residue in Seprase, instead of Asp^{663} as in DPPIV, reduces the acidity in this pocket. This change could explain the lower affinity for N-terminal amines by Seprase (Aertgeerts et al., 2005). Mutant proteins were developed to determine the importance of these two residues. Studies have shown that in DPPIV, the replacement of Asp^{663} by an alanine Ala^{663} results in a ~4-fold decrease in

catalytic efficiency for N-terminal dipeptides, with a concomitant increase in efficiency to cleave Z-Gly-Pro-AMC (which has been shown not to be cleaved by DPPIV). This mutation caused the wild type DPPIV catalytic efficiency for Z-Gly-Pro-AMC to be increased from 9.0 $M^{-1}s^{-1}$ to 1.6 x 10^3 $M^{-1}s^{-1}$. On the other hand, mutation of Ala^{657} to an Asp^{657} in Seprase resulted in reducing the k_{cat}/K_m for cleavage of Z-Gly-Pro-AMC by ~350-fold (Aertgeerts et al., 2005).

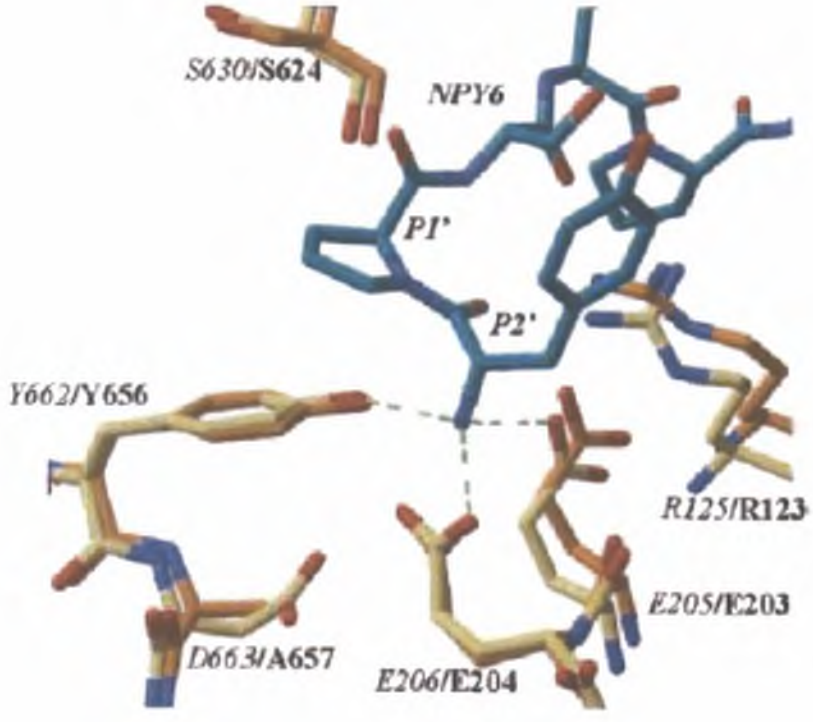


Figure 1.4.6 Superposition of Seprase with DPPIV showing a detailed view of the residues around the Glu motif.

The interactions in DPPIV between the active site residues and the N-terminus of the hexapeptide (NPY6, YPSKPD, *blue*), present in the crystal structure of DPPIV (*yellow*), are shown as *dashed lines*. Residues in Seprase are *orange*. Amino acid residues are labelled in *italic* and *bold* for DPPIV and Seprase, respectively (Aertgeerts et al., 2005).

Structural comparison of the active sites of Seprase and DPPIV revealed similar S2-S2' specificity pockets. The S_1' subsite (numbered according to Schechter and Berger (1967) see Figure 5.1.1) in Seprase is flat and could accommodate most amino acids. The S_2' active site pocket is lined by Trp^{623} and Tyr^{745}. These residues would be expected to interact with large aliphatic side chains of peptide substrates. The S_1 specificity pocket in Seprase is a well defined hydrophobic pocket lined by Tyr^{625}, Val^{650}, Trp^{653}, Tyr^{656}, Tyr^{660} and Val^{705}. This site optimally accommodates a proline residue. Large hydrophobic and aromatic residues can be modelled in the hydrophobic S_2 pocket, defined by residues Arg^{123}, Phe^{350}, Phe^{351}, Tyr^{541}, Pro^{544}, Tyr^{625} and Tyr^{660} (Aertgeerts et al., 2005).

1.4.1 Catalytic Classification

In its membrane form, the majority of Seprase, including its catalytic domain, is exposed to the extracellular environment (see Figure 1.4.1). The catalytic domain consists of the catalytic serine (S624) flanked by glycines in the classical consensus sequence for an active site serine, G-X-S-X-G. This conserved serine protease motif is present as G-W-S-Y-G. The catalytic serine in conjunction with aspartate (D702) and histidine (H734) comprises the catalytic triad (Goldstein et al., 1997; Pineiro-Sanchez et al., 1997; Scanlan et al., 1994) (see Figure 1.4.8). The orientation of these residues is similar to members of the prolyl oligopeptidase family and its structural organisation is similar to that of DPPIV. Therefore, this enzyme is classified as a non-classical serine protease. An interesting observation made in DPPIV is that a single substitution of either Gly residue in the motif resulted in the retention of the newly synthesised enzyme in the endoplasmic reticulum and rapid degradation (Barrett et al., 1998). This suggests that both residues are also essential for correct folding and transport of the enzyme to the cell surface. The three catalytic amino acids play an essential role in catalysis. The histidine acts as a general acid-base catalyst activating the nucleophilic group, the hydroxyl group of the serine acts as a nucleophile in the attack on the peptide bond while the aspartic acid stabilizes charged tetrahedral intermediates formed in the reaction (Fink, 1987; Rennex et al., 1991).

The mechanism of action of serine proteases (Figure 1.4.7) involves an acyl-enzyme intermediate. Both the formation and decomposition of the acyl-enzyme proceed through the formation of a negatively charged tetrahedral intermediate (Polgar, 2002).

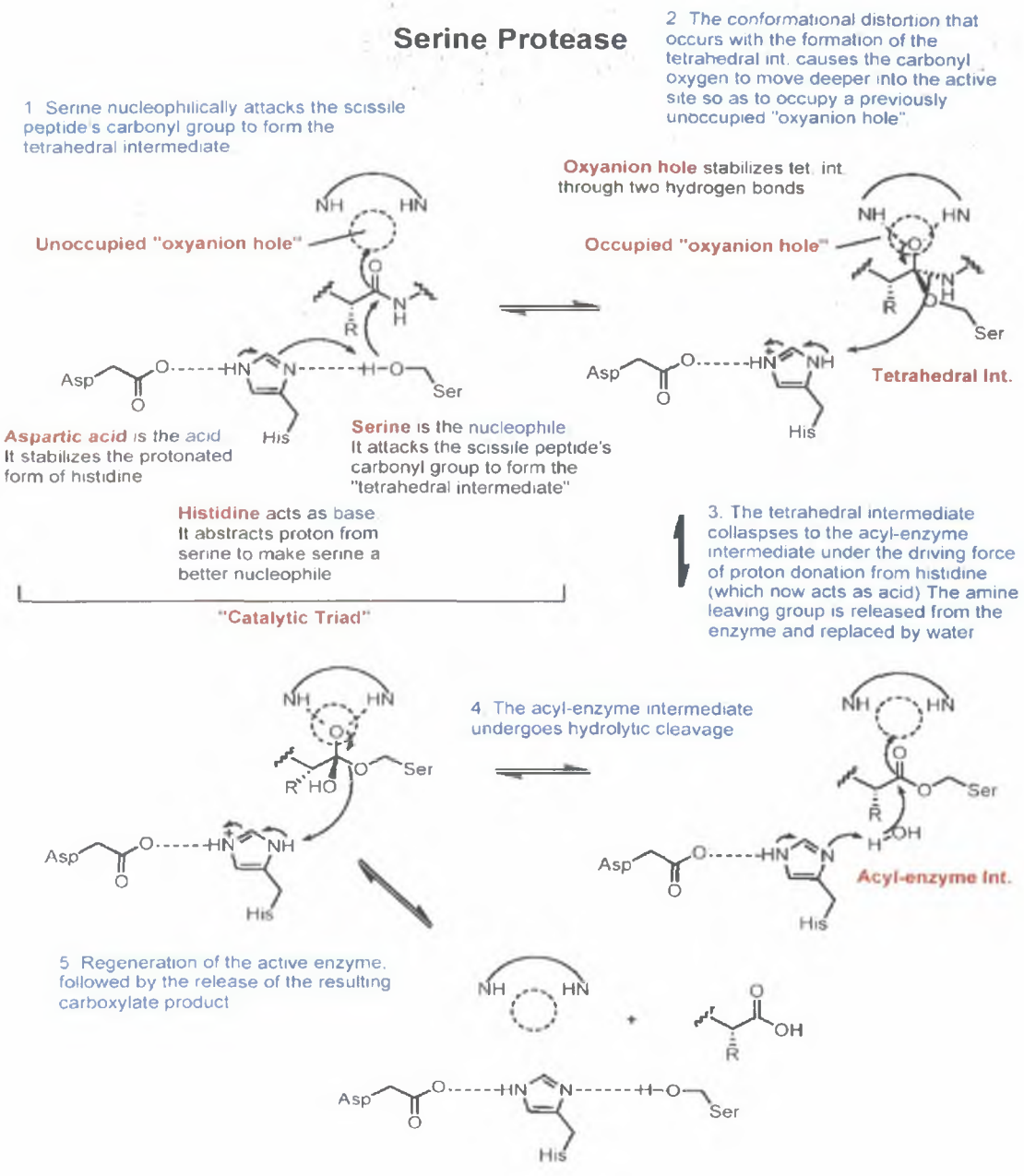


Figure 1.4.7 Schematic Representation of the Catalytic Mechanism of Serine Proteases

(Polgar, 2002)

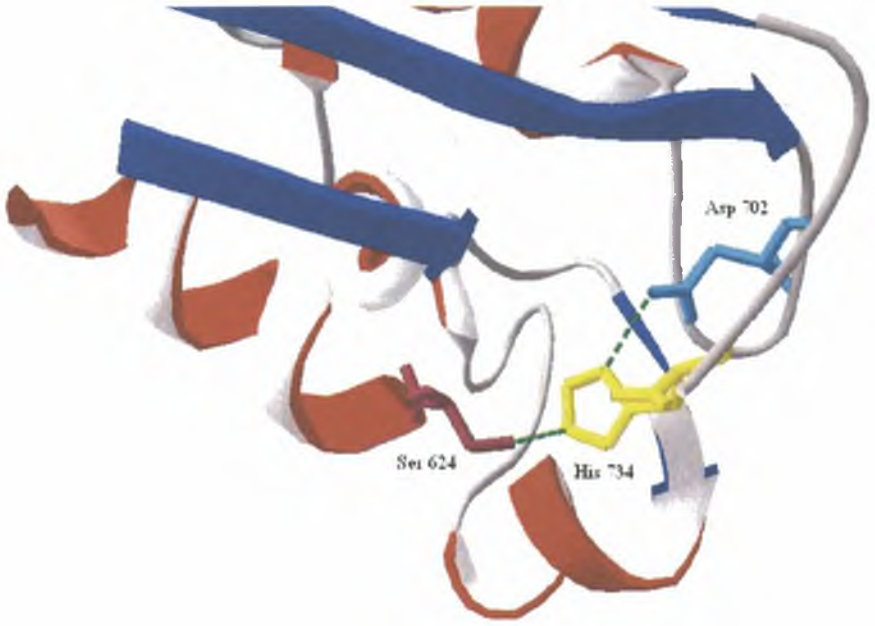


Figure 1.4.8 Catalytic Triad of Seprase

Ribbon diagram of the catalytic triad of Seprase (pdb accession code 1Z68). The catalytic triad residues are indicated **Serine 624**, Aspartic Acid 702 and Histidine 734. Hydrogen bonds are indicated in green. For amino acid information see Appendix B. Generated using DeepView (Section 2.10.10).

1.4.2 Biochemical Aspects of Seprase

Post translational modification of the protein such as N-glycosylation occurs and it is thought that the N-terminus may be blocked (Pineiro-Sanchez et al., 1997). The resolved crystal structure of Seprase shows that there are 5 potential glycosylation sites on the asparagine residues 49, 92, 227, 314 and 679. Four are located in the β-propeller domain and one is located in the hydrolase domain (Aertgeerts et al., 2005). The glycosylated form of Seprase has both post-prolyl dipeptidyl peptidase and gelatinase activities while the non-glycosylated form lacks any detectable activity (Sun et al., 2002).

Reports show that the gelatinase activity of Seprase was completely blocked by serine-protease inhibitors, including diisopropyl fluorophosphate (DFP) and phenylmethylsulfonyl fluoride (PMSF) (Aoyama and Chen, 1990; Collins et al., 2004). Seprase could be affinity labelled by [^{3}H]-DFP, but the proteolytically inactive 97kDa subunit could not (Pineiro-Sanchez et al., 1997). This confirmed the existence of a serine protease active site on the dimeric form of the enzyme. This was further demonstrated by the loss of proteolytic activity upon the dissociation of its 97kDa subunits following

treatment with acid, heat, or cysteine and histidine modifying agents (Pineiro-Sanchez et al., 1997). Therefore, it can be concluded from this that Seprase activity is determined by the association of the subunits to form a proteolytically active dimer (Aertgeerts et al., 2005). The proteolytic activity of membrane bound Seprase was found to be maximal at neutral pH and was enhanced by a mixture of 2mM EDTA and 2mM DTT (which inhibits metal-dependent proteases and activates cysteine proteases respectively) (Aoyama and Chen, 1990). However, a previous study of the soluble form of Seprase demonstrated that EDTA had no effect on the proteolytic activity and, contrary to previous reports DTT had a detrimental effect, with 5mM DTT causing a 10% loss of activity (Birney and O'Connor, 2001). Table 1.4 summarises some of the biochemical properties of Seprase.

Source	MW (kDa)	pI	pH optimum	Temperature optimum (°C)	Amino Acid Identity to Human Seprase (%)	GenBank accession number	Reference
Homo sapiens							
Malignant Melanoma cells; *LOX* cells	170 - dimer 97- subunit	5.0	7.0	-	100%	AAC51668	(Aoyama and Chen, 1990; Monsky et al., 1994; Pineiro-Sanchez et al., 1997)
Recombinant; cDNA from WI38 cells	95 – subunit	-	8.5	-	100%	-	(Sun et al., 2002)
Bovine Serum	96 - subunit	5.68	7.5	-	-	-	(Birney and O'Connor, 2001; Collins et al., 2004)
Chicken embryo	160 – dimer 97 – subunit	-	7.6	37	-	-	(Kelly, 1999)
Mus musculus	-	-	-	-	89%	AAH19190	(Niedermeyer et al., 1997)
Xenopus laevis	-	-	-	-	50%	AAC59872	(Brown et al., 1996)

Table 1.4 Biochemical Properties of Seprase

Summary of the published biochemical properties of Seprase. Those marked with (-) have not been published.

1.5 Purification and Activity Detection of Seprase

Seprase has been mainly purified from cell membranes and shed vesicles of LOX human amelanotic melanoma cells (Aoyama and Chen, 1990; Monsky et al., 1994). It has also been purified from 9-day old chicken embryos (Kelly, 1999). Size exclusion chromatography (S-200) and affinity chromatography using wheat germ agglutinin (WGA)-agarose are two of the most widely used resins for Seprase purification (Aoyama and Chen, 1990; Pineiro-Sanchez et al., 1997). Immunoaffinity purification of Seprase has also been utilised, with the mAb F19 pre-coated onto Sepharose CL-4B beads (Rettig et al., 1994). Soluble forms and isoforms of these membrane proteases are beginning to be found in biological fluids (Chen et al., 2003; Collins et al., 2004). Recently our group have published the purification scheme for the soluble form of Seprase from bovine serum (Collins et al., 2004). This purification procedure involved a combination of hydrophobic interaction chromatography, hydroxylapatite, and cibacron blue chromatography followed by size exclusion chromatography.

Until recently, the most sensitive assay available for Seprase detection involved gelatin zymography which exploits the established gelatinase activity associated with Seprase (Pineiro-Sanchez et al., 1997). However, this is not a quantitative assay. A semi-quantitative assay was developed based on the degradation of radiolabelled gelatin substrate and subsequent qualitative measurement of the released fragments (Kelly, 1999). Seprase was reported to possess prolyl dipeptidyl peptidase activity (Abbott et al., 1993; Chen et al., 2003; Niedermeyer et al., 1998; Sun et al., 2002) and Ala-Pro-AFC was seen as a potential sensitive fluorogenic substrate. Interestingly, conflicting results exist with reports that Seprase has no such prolyl dipeptidyl peptidase cleavage activity (Ghersi et al., 2002; Pineiro-Sanchez et al., 1997). Recently our group have also reported on the post prolyl cleaving endopeptidase activity of the soluble form of Seprase using the fluorimetric substrate Z-Gly-Pro-AMC (Collins et al., 2004).

1.6 Distribution of Seprase

Studies have shown that Seprase is transiently expressed in certain normal fetal mesenchymal tissues, during wound healing and in reactive stroma responding to epithelial cancers and some sarcomas (Garin-Chesa et al., 1990; Rettig et al., 1993). Normal adult tissues as well as malignant epithelial, neural and haematopoietic cells are generally Seprase-negative. The initial identification of Seprase involved the study of six surface glycoproteins that were differentially expressed during normal development, proliferative activation and malignant transformation of mesenchymal cells and tissues. The monoclonal antibody (mAb) F19 was used to define the human cell-surface glycoprotein Seprase (Rettig et al., 1988). The F19 antigen (now know as Seprase) was found to be expressed on cultured fibroblasts derived from various organs, several foetal mesenchymal tissues, scar tissue and a proportion of sarcoma cell lines. In normal adult tissues expression of the F19 antigen was restricted to occasional fibroblasts and to a set of pancreatic islet cells (Rettig et al., 1988). The pattern observed in this initial study suggests that Seprase is a cell-surface marker for proliferating mesenchymal cells and that its expression may be induced by normal growth factors or during malignant transformation (Rettig et al., 1988).

Another early study describes the induction of F19 in the reactive mesenchyme of epithelial tumours (carcinomas) (Garin-Chesa et al., 1990). Fibroblasts positive for the F19 antigen (Seprase) using immunohistochemical studies were found in primary and metastatic carcinomas including colorectal, breast, ovarian, bladder, and lung carcinomas. This study also analysed dermal incision wounds and found that F19 was strongly induced during scar formation. These studies suggest that the $F19^{+}$ phenotype correlates with specialised fibroblast functions in wound healing, inflammation and malignant tumour growth (Garin-Chesa et al., 1990). Another important observation from this study is that the cellular immunostaining patterns obtained with the tumour tissue suggests that Seprase is localised exclusively in the cell membrane/cytoplasm of fibroblastic cells. This cellular staining is consistent with the cell surface localisation of Seprase in cultured fibroblasts (Garin-Chesa et al., 1990).

Since these early reports, more studies have shown that Seprase is expressed in reactive human tumour stromal fibroblasts (Park et al., 1999). Studies have confirmed the expression of Seprase in primary breast infiltrating ductal carcinoma, colon adenocarcinoma and lung adenocarcinoma and also in metastatic colon adenocarcinoma in the hepatic system (see Figure 1.6.1) (Dolznig et al., 2005; Park et al., 1999).

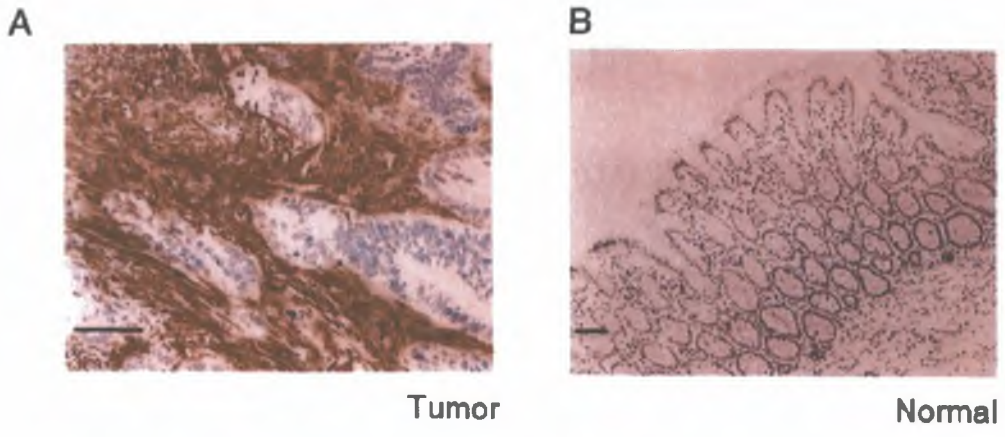


Figure 1.6.1 Detection of Seprase antigen and enzyme activity in human tissue

A and B, immunohistochemical detection of Seprase using mAb F19 in a colon carcinoma (A) or matched control normal colon tissue from the same patient (B) (bars = 100μm). (Park et al., 1999).

Several groups have shown Seprase to be expressed in the reactive stromal fibroblasts of human breast cancer and its absence from normal breast tissue (Ariga et al., 2001; Garin-Chesa et al., 1990; Rettig et al., 1994; Scanlan et al., 1994). In addition to this, Seprase is also expressed by infiltrating ductal carcinoma (IDC) cells in breast cancer patients while it is not expressed by normal breast epithelia (Dolznig et al., 2005; Kelly et al., 1998; Park et al., 1999). Stromal expression of Seprase in IDC of the breast was associated with longer survival of patients (Ariga et al., 2001). Further work has shown that Seprase expression is not confined to stromal fibroblasts but that the protease is also expressed in some types of malignant cells of epithelial origin (Chen et al., 2003; Iwasa et al., 2005; Jin et al., 2003; Okada et al., 2003). Reports differ in the cellular localisation of FAPα and Seprase depicted by immunohistochemistry (Chen et al., 2003; Kelly et al., 1998). The apparent difference is thought to be partially due to the use of

antibodies that recognise, with varying affinity, different epitopes exhibited by FAPα (derived from activated fibroblasts) and Seprase (derived from invasive cancer cells).

Expression patterns of Seprase were examined in cervical carcinoma and cervical intraepithelial neoplasm (Jin et al., 2003). This embraces both carcinoma in situ and the precursor lesions known as dysplasia or 'disordered differentiation'. Some micro-invasive carcinomas and all invasive carcinomas showed Seprase immunoreactivity in the cancer cells (Jin et al., 2003). The findings in this study show a direct correlation between gelatinase expression and the malignant phenotype. Thus Seprase may be an early marker of tumour progression characterising Seprase expression with invasive growth. A separate study to support this concept was performed by Iwasa et al., (2005) which examined Seprase expression in colorectal cancer specimens. Immunoblotting showed higher levels of Seprase protein in the cancer tissue than in normal colorectal tissue ($p<0.001$). The results also revealed a significant correlation between Seprase expression and lymph node metastasis ($p=0.033$).

Expression patterns of Seprase in human gastric cancer were investigated using immunohistochemistry and the study showed that there were distinct differences in its expression between the intestinal- and diffuse-type gastric cancer (Okada et al., 2003). Results also showed, as in Iwasa et al., (2005), a correlation between Seprase expression and depth of invasion. In intestinal cancer, the stromal expression of Seprase significantly correlated with liver metastasis ($p=0.0002$) and lymph node metastasis ($p<0.0001$). In contrast, in diffuse-type cancer there was no correlation between stromal Seprase expression and lymph node metastasis ($p=0.0821$) (Okada et al., 2003). A separate study looked at the expression of Seprase at the mRNA and protein level. This study found that Seprase expressing carcinoma tissues were more prominently found in the scirrhous type than in other types of gastric carcinoma (Mori et al., 2004).

Immunohistochemical studies have shown Seprase expression was induced in patients with idiopathic pulmonary fibrosis (IPF) (Acharya et al., 2006). Its expression pattern is restricted to fibroblasts in areas of ongoing tissue injury (see Figure 1.6.2).

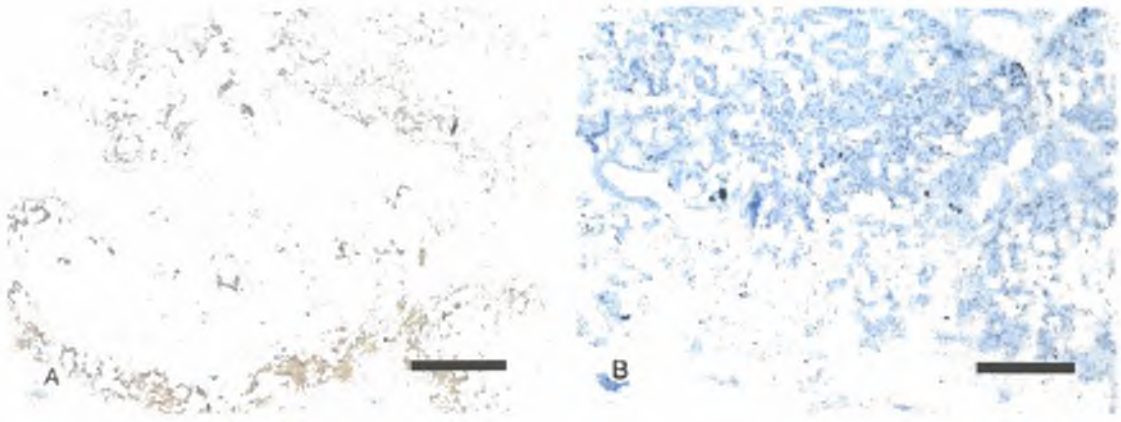


Figure 1.6.2 Detection of Seprase antigen expression in IPF

Serial lung tissue sections immunostained with Seprase (A) and an isotope-matched control antibody (B). Note areas of thickened interstitium interspersed with normal alveolar architecture. Seprase expression is only seen in areas of fibrosis (bar = 1mm for both A and B) (Acharya et al., 2006).

Seprase was recently identified for the first time on chondrocyte membranes under conditions that promoted cartilage resorption and elevated expression in cartilage from osteoarthritis patients (OA) (Milner et al., 2006). The results from this study supported a role for Seprase in the mechanisms leading to cartilage degeneration in OA. Gene expression profiling in the murine model showed a 7-fold increase in Seprase expression in inflamed, compared to non-inflamed paws.

A study by Huber et al. (2003) found that Seprase was expressed in benign and malignant melanocytic skin tumours. This is in contrast to the findings in benign epithelial tumours, in which little or no expression of Seprase was observed on stromal fibroblasts (Dolznig et al., 2005; Garin-Chesa et al., 1990). Normal adult skin, however, did not have any detectable Seprase activity. These contradictory results could be explained if melanocytic naevi (moles) are considered as precursor lesions for melanoma development, characterised by a constitutively-active tumour stroma (Huber et al., 2003; Huber et al., 2006). Gene expression studies have identified Seprase to be uniquely overexpressed in aggressive fibromatosis (Skubitz and Skubitz, 2004). Aggressive fibromatosis is locally invasive but rarely metastasises. There are histologic similarities between this disease and the proliferative phase of wound healing.

To further support the concept that Seprase is expressed during wound healing, studies have shown that stellate cells at the tissue remodelling interface of cirrhosis in humans, express Seprase (Levy et al., 1997; Levy et al., 1999). Hepatic stellate cells (HSCs) are a known major source of ECM degrading enzymes and protease inhibitors of hepatic tissue remodelling (Alcolado et al., 1997). It is thought that Seprase may contribute to the (HSC)-induced extracellular matrix changes (ECM) of cirrhosis. Seprase expression was also detected in the hepatic parenchyma of patients with chronic hepatitis C virus (HCV) infection. In fact, Seprase expression correlates with the degree of fibrosis in HCV ($p<0.0001$) (Levy et al., 2002). This adds considerable weight to the assertion that Seprase has a role in the pathogenesis of chronic liver disease.

Another study involving the *Xenopus laevis* Seprase homologue describes its increased expression during hormone-induced tail resorption, indicating a possible role in tissue remodelling (Brown et al., 1996). An extension of this concept is that Seprase has been found to be expressed during mouse embryo development (Niedermeyer et al., 2001). Seprase deficient mice showed no overt developmental defects and were viable. To analyse Seprase expression at various stages of embryonic development, mice deficient in Seprase but expressing β-galactosidase under the control of the Seprase promoter were developed (Seprase$^{-/-}$ *lac*Z). These Seprase$^{-/-}$ *lac*Z mice expressed β-galactosidase at active tissue remodelling during embryogenesis supporting a role for Seprase in tissue remodelling processes (Niedermeyer et al., 2001).

Overall, Seprase has been seen to be expressed in stromal fibroblasts of more than 90% of all epithelial tumours including lung, colorectal and breast carcinomas (primary and metastatic) (Garin-Chesa et al., 1990). It has also been shown to be expressed in post-natal non-tumour tissues (Chen et al., 2003), in a proportion of bone and soft tissue sarcoma tumour cells (Dolznig et al., 2005), in granulation tissue of healing wounds, and on pancreatic cells (Rettig et al., 1988). However, normal adult tissue, benign and premalignant epithelial lesions are generally Seprase negative. Seprase-positive cells were found in close proximity to the endothelial cells of the tumour capillaries and surround the tumour nodules (Garin-Chesa et al., 1990; Rettig et al., 1988; Welt et al., 1994). These results suggest that Seprase may have a critical role to play in altering the

micro-environment to promote tumour growth, thus enhancing the theory that metastatic effects may depend on proteolytic enzymes derived not only from tumour cells but also from stromal cells.

1.7 Biological and Pathological Roles

High levels of Seprase activity have been positively correlated with the invasive behaviour of many malignantly transformed cells but its function in malignancy is still unknown (Aoyama and Chen, 1990; Kelly et al., 1994; Monsky et al., 1994)

1.7.1 Tumour Suppressor

There is compelling evidence that Seprase may play a role as a tumour suppressor (Ramirez-Montagut et al., 2004; Welt and Ritter, 1999; Wesley et al., 1999). Initial studies have focused on DPPIV. Expression of DPPIV is lost as normal melanocytes progress to malignant melanoma. Experiments showed that the re-expression of DPPIV was able to change the mouse melanoma cells to a more differentiated and normal phenotype with a return to dependence on exogenous growth factors (Wesley et al., 1999). Wesley et al. (1999) observed that the re-establishment of dependence on exogenous growth factors occurred even when a catalytically inactive mutant of DPPIV was expressed. This group attributed this to the expression of endogenous Seprase which was co-induced when the mutant DPPIV was expressed. More recent studies have provided direct evidence for Seprase as a tumour suppressor. It was observed that the expression of Seprase decreased the tumourigenicity of mouse melanoma cells in animals and restored contact inhibition and growth factor dependence (Ramirez-Montagut et al., 2004). Another interesting finding from this study is that the catalytic mutant of Seprase contributed to the tumour suppression in the absence of an active protease. In addition to this, the study showed that while DPPIV expression induced Seprase expression, the converse was not so; i.e. Seprase expression did not induce DPPIV expression. Therefore, it was concluded that the tumour-suppressive activities of wild-type and mutant Seprase are most likely due to functions residing in Seprase.

1.7.2 Tumour Promoter

The role of Seprase in breast cancer was investigated using human breast cancer cell lines that normally express Seprase (MDA-MB-435 and MDA-MB-436) (Goodman et al., 2003). Anti-sense suppression of Seprase rendered these cells sensitive to serum starvation while control transfectants with high levels of Seprase expression grew well in the absence of serum. Goodman concluded that the breast cancer cells with high levels of Seprase expression are less dependent on exogenous serum factors for growth and have gained independence from normal growth regulatory controls. Independence from normal growth regulation is a key characteristic of malignantly transformed cells that distinguishes them from normal cells.

Seprase expression was also engineered in the human breast cancer cell line MDA-MB-231, which lacks normal Seprase expression (Huang et al., 2004). A mouse model was utilised to show that Seprase expressing tumours grew more rapidly and were highly vascular when compared to tumours of control transfectants which lacked Seprase expression (Huang et al., 2004). This group did find that when these cells were grown *in vitro*, those cells that did express Seprase grew at the same rate as those that did not. This indicated that Seprase had a more noticeable effect on tumour cell growth in the mouse mammary fat pad environment. For tumours to grow and integrate into surrounding normal tissue, they must attract a blood supply (angiogenesis), which sustains growth. This study was the first evidence for a pro-angiogenic function for Seprase and it can be concluded that Seprase expression promotes growth of breast cancer tumours at least in part by driving angiogenesis (Huang et al., 2004). This conclusion is supported by studies showing that Seprase mRNA is up-regulated by endothelial cells undergoing re-organisation and capillary morphogenesis (Aimes et al., 2003). Both these findings together suggest that Seprase expression by breast cancer cells favourably alters the tumour microenvironment to promote tumour growth.

Similar growth-promoting effects of Seprase have been described (Cheng et al., 2002; Wang et al., 2005a). Murine Seprase was transfected into HEK293 human embryonic kidney cells and it was shown that tumours derived from Seprase expressing cells grew more rapidly than the control transfectants. This group also demonstrated in an animal model, that antibodies found to inhibit the dipeptidase activity of Seprase, also suppressed the growth of tumours which were derived from cells expressing Seprase. Wang et al. (2005b) found that the over-expression of Seprase in the human hepatic stellate cell (HSC) cell line LX-2 increased cell adhesion, migration and invasion. Interestingly, this group found that Seprase protease activity was not necessary for these functions. These findings further support a pro-fibrogenic role for Seprase by indicating that, in addition to its enzymatic functions, Seprase has important non-enzymatic functions (Wang et al., 2005b). Overall, it can be concluded from these studies that Seprase activity mediates in some way the increased tumour growth.

There is an obvious discrepancy between Seprase function in tumour promotion and tumour suppression. It is proposed by some that Seprase expression has a profound effect on cells that do not normally express Seprase and that the context in which Seprase is expressed determines the biological response to Seprase (promotion or inhibition of growth) (Kelly, 2005). A model has been created that explains all the findings so far. It describes how Seprase associates with membrane-bound signalling molecules causing transmission of growth stimulatory or inhibitory signals (see Figure 1.7.1 for an example of stimulatory signals). The factor that determines this must reside in the signalling molecules that are available for interaction with Seprase on the cells. Therefore it was concluded that Seprase carries out its biological functions in a cell-context dependent manner through a combination of its protease activity and its ability to form complexes with other cell surface molecules (see Section 1.7.3) (Kelly, 2005).

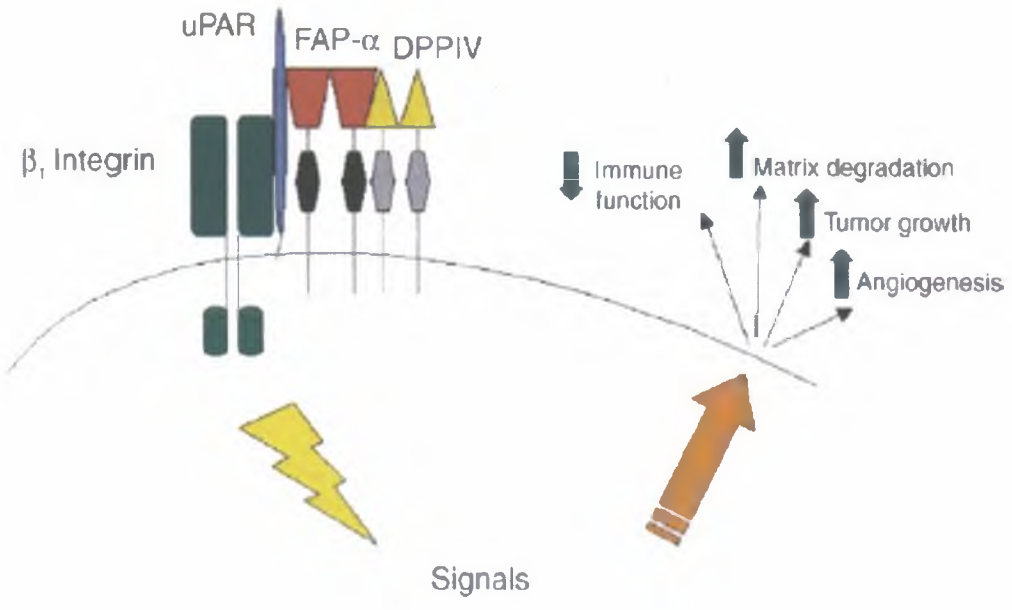


Figure 1.7.1 Model of Seprase supramolecular complexes

Model of a Seprase supramolecular complex with proteolytic, adhesive, and stimulatory signalling capabilities. Seprase dimers form heteromeric complexes with DPPIV dimers (Ghersi et al., 2002; Scanlan et al., 1994). Seprase can also form close association with β_1 integrins and uPAR (Artym et al., 2002). It is known that uPAR can bind to $\alpha_3\beta_1$ integrin through the α_3 chain (Zhang et al., 2003). However, the binding interactions that associate Seprase, uPAR, and the integrins are not defined and may not be as depicted. Model taken from (Kelly, 2005). See Section 1.7.3 for Seprase complexes in cellular invasiveness.

1.7.3 Seprase complexes in cellular invasiveness

Seprase associates with $\alpha_3\beta_1$ integrin, DPPIV, MMP-2, membrane-type 1 MMP and uPA at the invadopodia of human malignant cells and so may interact with these proteinases and receptors and with associated cascades (Artym et al., 2002; Milner et al., 2006; Monsky et al., 1994; Mueller et al., 1999).

The integrin family of transmembrane adhesion proteins have been shown to exhibit multiple functions, including adhesion to ECM and their localisation at invadopodia (Mueller et al., 1999). One theory is that integrins are involved in recruiting proteases to these sites of cell invasion. The $\alpha_3\beta_1$ integrin has been shown to associate with Seprase and it is thought that $\alpha_3\beta_1$ may participate in the formation of functional invadopodia by docking Seprase (see Figure 1.7.2).

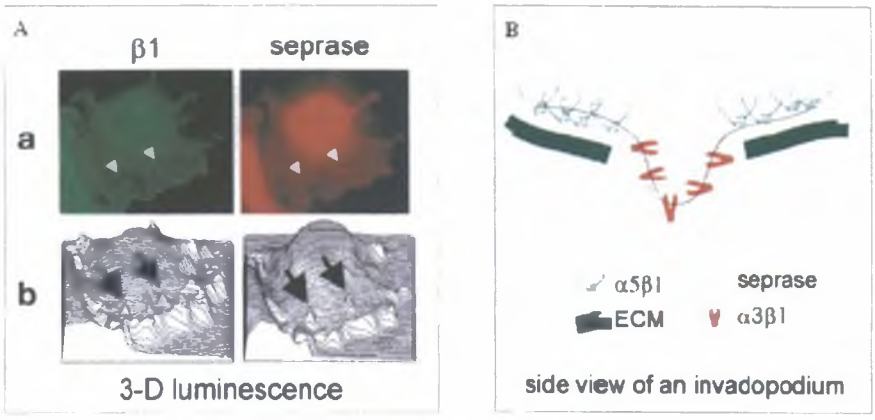


Figure 1.7.2 Co-localisation of Seprase and β_1 integrin

Panel A, Co-localisation of Seprase and β_1 integrin at invadopodia of LOX cells using immunofluorescence and image analysis. (a) directly labelled FITC (green)-mAb C27 against β_1 integrin and rhodamine (red)-mAb D28 against Seprase co-localise in the same invadopodia (*arrowheads*). (b) 3D luminescence profiles illustrate β_1 integrin and Seprase staining intensities in the micrograph shown above each profile. *Arrows* indicate invadopodia. ***Panel B***, diagrammatic representation of integrin $\alpha_3\beta_1$ and Seprase localisation as would be seen in a vertical section through an invadopodium. (Mueller et al., 1999)

Seprase and DPPIV can form a complex localised at invadopodia of fibroblasts on collagenous fibres, that has both gelatinolytic and gelatin binding activities, which facilitates cell migration (Chen, 2003; Chen and Kelly, 2003; Ghersi et al., 2003; Ghersi et al., 2006). Seprase and DPPIV are highly homologous with both enzymes having the same domain structure (see Figure 1.4.1). Seprase shows 50% amino acid identity with dipeptidyl peptidase IV over the entire sequence, with almost 70% identity in the catalytic domain (Goldstein et al., 1997; Pineiro-Sanchez et al., 1997; Scanlan et al., 1994). The Ser-Asp-His catalytic triad is conserved in both enzymes and their genes map to the same region of the human chromosome 2. Despite their shared structural features, Seprase and DPPIV differ in their tissue distribution and expression (Welt et al., 1994). Seprase can form heteromeric complexes with DPPIV and the formation of the DPPIV-Seprase complex, which exhibits both prolyl peptidase activity and gelatinase activity, is necessary for cell migration on a collagen substrate (Ghersi et al., 2002). It is possible that the two enzymes cooperate for efficient degradation of substrates and its consequent biological functions. The exact natures of the homodimer and heterodimer complexes of Seprase are poorly understood.

A recent study has shown that Seprase-Urokinase-type plasminogen activator receptor (uPAR) membrane complexes are associated with the invadopodia of LOX cells, suggesting their co-operative roles in tumour invasion (see Figure 1.7.3) (Artym et al., 2002). Urokinase plasminogen activator (uPA), a serine protease, converts the tissue zymogen plasminogen into plasmin. Plasmin has the ability to degrade most proteins in the ECM, including type IV collagen, laminin and fibronectin. In model systems, both the inhibition of uPA activity and the prevention of uPA binding to its receptor have been shown to limit the formation of metastasis (Iwasa et al., 2005). Thus, Seprase may also be a potential candidate for anti-metastatic therapies. The interaction of uPA and the integrins have been shown to regulate both adhesive and signalling activities of integrins (Mueller et al., 1999). The formation of the Seprase-uPAR membrane complexes is dependent upon both the cytoskeleton and integrins, specifically the β_1 integrin (Artym et al., 2002).

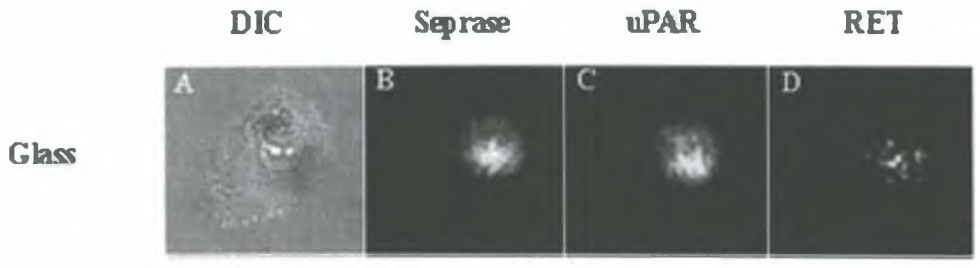


Figure 1.7.3 Co-localisation and RET of Seprase and uPAR on LOX cells

Lox cells were attached to glass coverslips and were examined by immunofluorescence microscopy. Cells were labelled with FITC-conjugated anti-Seprase D28 mAb (B) and TRITC-conjugated anti-uPAR mAb (C). Columns 1-4: Differential Interference Contrast (DIC), fluorescence of anti-Seprase, fluorescence of anti-uPAR and resonance energy transfer (RET) (Artym et al., 2002)

1.7.4 Antiplasmin Cleaving Enzyme (APCE)

As mentioned in Section 1.4 Lee et al. (2005a) have recently identified the soluble form of Antiplasmin Cleaving Enzyme (APCE) as Seprase. α_2-antiplasmin (AP) is a plasma glycoprotein of the serpin (*ser*ine *p*rotease *in*hibitor) superfamily and is the primary physiological inhibitor of plasmin, a key enzyme in fibrin degradation (Lee et al., 1997). The fibrinolytic system leads to the generation of plasmin from plasminogen through the action of Tissue or Urokinase Plasminogen Activator (t-PA or uPA). Studies into α_2-antiplasmin have shown that it is present in the blood in a larger "pro"-form, α_2AP_{PRO} and in a smaller mature form of α_2AP_{ACT}. α_2AP_{PRO} has 464 amino acid residues, with Met as the N-terminus and the mature form α_2AP_{ACT}, having 452 amino acid residues, with an N-terminal asparagine (Bangert et al., 1993; Lee et al., 2004). Both forms have been shown to be present in the blood (66% α_2AP_{ACT} ; 34% α_2AP_{PRO}) and they both form a stable complex with plasmin. However the "pro"-form has remarkably less capacity of cross-linking fibrin (see Figure 1.7.4) (Koyama et al., 1994). When α_2AP_{ACT} is cross-linked to fibrin (during clot formation), it retains the ability to effectively inhibit plasmin and protect fibrin from proteolysis. The processing of α_2AP_{PRO} to α_2AP_{ACT} was shown to involve a plasma protease that cleaved the Pro12-Asn13 bond of the α_2AP_{PRO} protein (Lee et al., 2005a). This protease was called Antiplasmin Cleaving Enzyme (APCE).

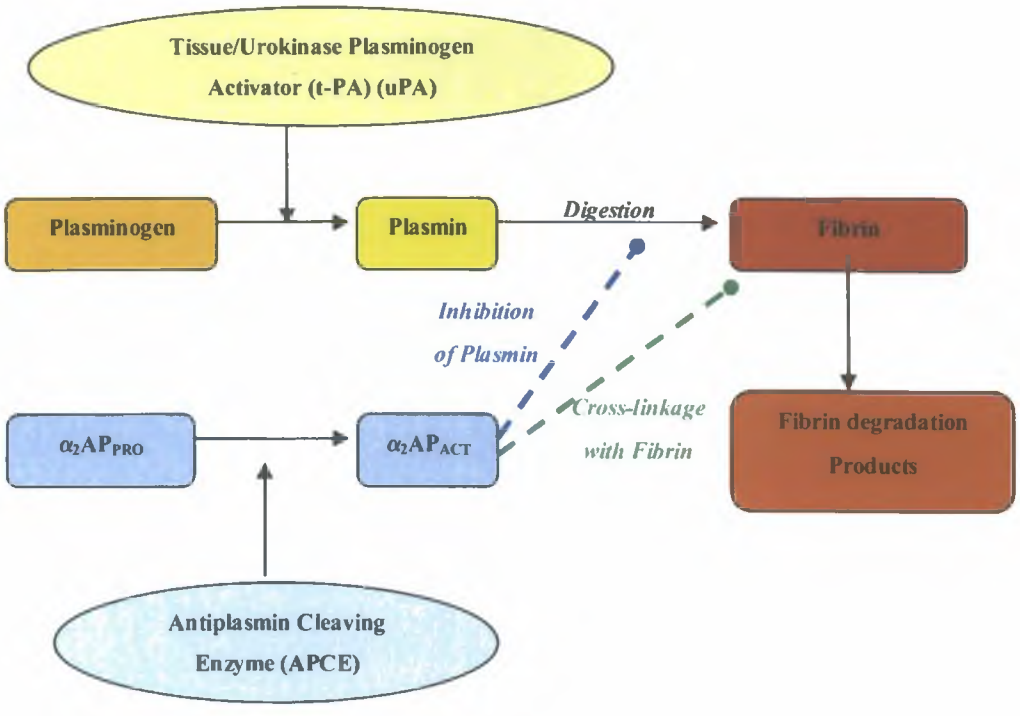


Figure 1.7.4 Summary of Fibrinolytic System

Plasmin is generated from plasminogen through the action of Tissue or Urokinase Plasminogen Activator (t-PA or uPA). Plasmin digests Fibrin giving rise to degradation products. α_2AP_{PRO} is processed to α_2AP_{ACT} by the action of APCE. α_2AP_{ACT} can cross-link with both Plasmin and Fibrin. During clot formation, α_2AP_{ACT} becomes cross-linked to fibrin, where it retains the ability to effectively inhibit plasmin and protect fibrin from proteolysis.

APCE was purified from human plasma using an initial ammonium sulphate fractionation, followed by hydrophobic, anion exchange, T-gel thiophilic, and immunoaffinity chromatographies (Lee et al., 2004). The fact that APCE is found in human plasma and cleaves α_2AP_{PRO} suggests that one of its physiologic functions is the regulation of α_2AP_{ACT} availability for plasmin inhibition within cross-linked fibrin (Lee et al., 2004). Plasmin is resistant to α_2-antiplasmin inhibition when it is associated with the plasma membrane (Del Rosso et al., 2002). A study analysing the regulation of liver regeneration by the plasmin/α_2-AP system suggested that this system plays an important role in the hepatic repair via proteolysis of the matrix elements and the clearance of cellular debris from the injured site (Okada et al., 2004). Altered fibrin turnover may be

responsible for tumour-promoting effects by various means (Gerner et al., 2001). In clinical situations where fibrin formation is likely, the development of an agent that inhibits APCE might result in decreased amount of α_2AP_{ACT} available for cross-linking to fibrin as thrombi develop or as inflammation progresses. Abnormalities of haemostasis during tumour growth and metastasis have been observed and it has been shown that the ECM of neoplastic tissue is abnormally rich in fibrin bundles and is mixed with other stromal components (Del Rosso et al., 2002). Degradation of fibrin, fibrinogen and the ECM proteins by plasmin, in addition to the concomitant activation of the Urokinase Receptor by Urokinase Plasminogen Activator, was suggested to facilitate tumour cells invasion and metastasis (Gerner et al., 2001). It has also been shown that elevated plasma levels of cross-linked fibrinogen indicated cancer-related fibrin deposition and fibrinolysis (Gerner et al., 2001).

1.8 Substrate Specificity

The identification of key substrates recognised by Seprase may provide a better understanding of the physiological role and of the pathways that control the activation of stromal fibroblasts during tumourigenesis. Seprase has been shown to have two proteolytic activities. Firstly it is a gelatinase (Aoyama and Chen, 1990; Monsky et al., 1994; Pineiro-Sanchez et al., 1997) and secondly an N-terminal post-prolyl amino peptidase that can cleave Ala-Pro-AFC. Table 1.8 illustrates the kinetic constants obtained in studies of dipeptide substrates. Most serine proteases have no absolute substrate specificity (Fulop et al., 1998). They can cleave peptide bonds with a variety of side chains adjacent to the scissile bond. The S_1-P_1 hydrogen bond is a very important factor in the proper substrate orientation and the transition state stabilisation. Proline residues do not possess main chain NH groups and therefore the architecture of the S_1 binding site is evolved to be much more specific in Seprase (Fulop et al., 1998).

A study by Collins et al. (2004) revealed that the soluble serum form of Seprase has a preference for N-blocked amino substrates of Gly-Pro-X and is most effective when X is Phe or Met and least effective when X is His or Glu. This study reported an affinity constant (K_m) of 270μM for bovine serum Seprase with the substrate Z-Gly-Pro-AMC. Extensive substrate specificity studies suggested that this peptidase has an extended substrate binding region in addition to the primary specificity S_1 (see Figure 5.1.1). The analysis also revealed that at least five subsites were involved in enzyme-substrate binding, with the smallest peptide cleaved being a tetrapeptide.

Substrate specificity studies using zymography have shown that Seprase degrades gelatin and thermally denatured type I and type IV collagens but not laminin, fibronectin, fibrin or casein (Monsky et al., 1994; Pineiro-Sanchez et al., 1997). The gelatinase activity distinguishes Seprase from DPPIV, which does not have gelatinase activity (Gorrell et al., 2001; Park et al., 1999). However, an earlier report by Bermpohl et al. (1998) suggested that DPPIV does have gelatinase activity. Both the gelatinase and dipeptidyl peptidase activities of Seprase are mediated by an active site serine^{624} (Park et al., 1999; Pineiro-Sanchez et al., 1997).

Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	References
Ala-Pro-AFC	ND	250	ND	Recombinantly expressed murine Seprase (Cheng et al., 2005)
Ala-Pro-AFC	2.0	200	1.0×10^4	Recombinantly expressed human Seprase with an N-terminal histidine tag (Sun et al., 2002)
Ala-Pro-AFC	ND	>200	2.1×10^4	Wild type human Seprase (Aertgeerts et al., 2005)
Ala-Pro-AFC	14.2	244	5.8×10^4	Recombinantly expressed human Seprase with an N-terminal Flag tag (Edosada et al., 2006b)
Ala-Pro-AFC	1.08	323	3.34×10^3	Recombinantly expressed human Seprase (Lee et al., 2005a)
Ala-Pro-AFC	0.99	272	3.64×10^3	Wild type human APCE (Lee et al., 2005a)
Gly-Pro-AFC	5.6	248	2.3×10^4	Recombinantly expressed human Seprase with an N-terminal Flag tag (Edosada et al., 2006b)
Gly-Pro-AFC	ND	>200	4.3×10^3	Wild type human Seprase (Aertgeerts et al., 2005)
Gly-Pro-AMC	ND	>1000	3.0×10^3	Wild type human Seprase (Aertgeerts et al., 2005)
Z-Gly-Pro-AMC	ND	270	ND	Soluble bovine form of Seprase (Collins et al., 2004)
Z-Gly-Pro-AMC	ND	ND	5.3×10^4	Wild type human Seprase (Aertgeerts et al., 2005)
Z-Gly-Pro-AMC	0.51	101	5.0×10^3	Recombinantly expressed human Seprase (Lee et al., 2005a)
Z-Gly-Pro-AMC	0.54	124	4.35×10^3	Recombinantly expressed human Seprase (Lee et al., 2005a)
Z-Gly-Pro-AMC	ND	ND	7.4×10^3	Recombinantly expressed human Seprase with an N-terminal Flag tag (Edosada et al., 2006b)
FRET peptide	1.20	29	4.14×10^4	Recombinantly expressed human Seprase (Lee et al., 2005a)
FRET peptide	1.05	26	4.03×10^4	Wild type human APCE (Lee et al., 2005a)

Table 1.8 Kinetic Constants for the hydrolysis of fluorogenic dipeptide substrates by Seprase

ND-not determined. FRET peptide: Arg-Lys(DABCYL)-Thr-Ser-Gly-Pro-Asn-Gln-Gln-Gln-Glu(EDANS)-Arg

1.8.1 Modifications of bioactive peptides

The endopeptidase activity of Seprase may function to regulate bioactive peptides. A proline near the end of a peptide produces a significant bend. The natural substrates of Seprase are unknown, although the serum form has been shown to cleave the Pro_{12}-Asn_{13} bond of α_2 antiplasmin (α_2AP_{PRO}) producing the Asn-α_2 antiplasmin (α_2AP_{ACT}) (Lee et al., 2004). The cleaved protein binds fibrin more efficiently and slows clot lysis by plasmin more effectively than full-length α_2 antiplasmin (α_2AP_{PRO}). Based on the Seprase cleavage site in α_2AP_{PRO}, the specificity of Seprase in the P_4-P_4' positions were defined (Edosada et al., 2006a). The study confirmed that Seprase requires a Proline in the P_1 position and Glycine (or D-amino acids) at the P_2 position (see Figure 1.8.1) (Collins et al., 2004; Edosada et al., 2006a). Seprase was shown to have a broader tolerance at the P_4, P_3, P_1', and P_2' positions. In the P_3 position, Seprase had a preference for Ala and Ser. Less activity was observed against charged and aromatic amino acid residues, suggesting a preference for small, uncharged amino acids at the P_3 position. Seprase seems to have a broader specificity at the P_4 position. Compared to the previous study mentioned above, it was found that Seprase had a preference for Phe or Met at the P_1' position and least tolerance of His or Glu (Collins et al., 2004; Edosada et al., 2006a).

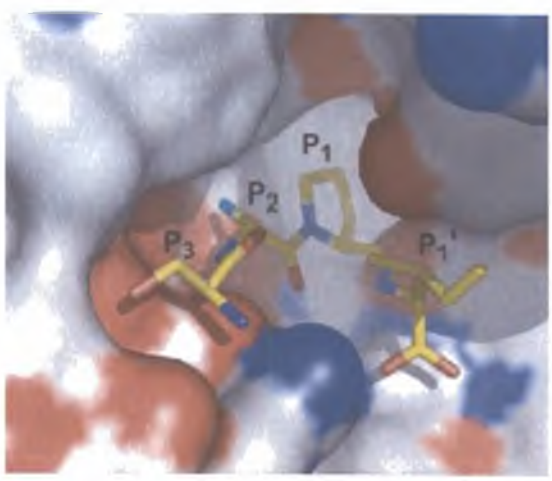


Figure 1.8.1 Seprase binding model

The model shows a peptide with the P_3-P_1' sequence Ser-D-Ala-Pro-Ile bound to Seprase. The surface depicts the structure of Seprase (pdb accession code 1Z68) with carbon atoms shown in white, oxygens in red, and nitrogens in blue. The carbon atoms of the peptide are in yellow and the side chain of the P_2 D-Ala residue is highlighted in cyan. (Edosada et al., 2006a)

Seprase has been shown to have a preference for N-blocked amino substrates. Studies have shown that Seprase cleaves formyl-, benzyloxycarbonyl- and biotinyl- blocked substrates, which DPPIV cleaved poorly (Edosada et al., 2006b). The substrate preferences as defined by Edosada et al. (2006b) were used to design a peptidyl-chloromethyl ketone (cmk) that inhibited Seprase but not DPPIV.

A fluorescence resonance energy transfer (FRET) peptide (see Table 1.8) was developed that contained the APCE-sensitive Pro12-Asn13 bond within the Thr9-Gln16 sequence of Met-α_2AP (α_2AP_{PRO}) (Lee et al., 2004). Hydrolysis of the Pro-Asn bond separates the fluorophore, EDANS, from the quenching group, DABCYL, to give an increase in fluorescence. Kinetic data using the FRET peptide has shown Seprase to have high affinity and kinetic efficiency, indicating that residues in the P_4-P_4' region contribute to the substrate specificity (Lee et al., 2005a). The k_{cat}/K_m value for cleavage of the FRET peptide is ~8-fold higher than for Z-Gly-Pro-AMC and ~12-fold higher than for Ala-Pro-AFC (see Table 1.8) (Lee et al., 2004).

Combined results from the Edosada et al. (2006a) study also provided a composite α_2AP-derived peptide substrate for Seprase namely, Ala-Ser-Gly-Pro-Ser-Ser. Comparing the kinetic parameters for the parental (TSGP-NQ) and composite α_2AP-derived peptide substrate (ASGP-SS) shows that the K_m value for the cleavage of the composite substrate was 3.3 fold higher than the parental substrate (K_m = 4.3μM and 1.3μM respectively). However, a concomitant 3.4 fold increase in k_{cat} was observed, yielding a nearly equivalent catalytic efficiency (k_{cat}/K_m) for each peptide (k_{cat}/K_m = 1.3 x 10^6 $M^{-1}s^{-1}$ and 1.2 x 10^6 $M^{-1}s^{-1}$).

Identification of the natural substrates of Seprase and the functions of Seprase activated peptides remains an important area to be investigated.

1.9 Possible Clinical Significance of Seprase

Invasion with or without metastatic spread of cancer to distant sites has already occurred in a significant number of patients by the time their disease is detected (Fidler, 2002). Significant improvement in a patient's survival may be achieved by halting the invasion process and containing the metastatic spread and growth of the disease. There are various possibilities on how this can be pursued.

1.9.1 Seprase provides target specificity to anti – tumour agents

Surface expressed Seprase represents an additional potential target for immunolocalisation and immunotherapy of epithelial cancers. Since a large proportion of carcinomas contain an abundance of Seprase positive stroma, it is possible that they would be accessible to circulating mAb. Several groups have developed antibodies directed against Seprase. Antibodies were developed using the Seprase specific mAbF19 as a model and humanising the antibody (Hofheinz et al., 2003; Mersmann et al., 2001; Samel et al., 2003; Scott et al., 2003; Tahtis et al., 2003).

A phase I study investigated the effects of the mAbF19 targeted against metastatic colon cancer (Welt et al., 1994). This study concluded that there were no toxic effects associated with intravenous administration of iodine^{131} labelled mAbF19 (^{131}I-mAbF19). Increased expression of Seprase in primary and metastatic colorectal carcinomas with limited expression in normal adult tissue was also observed during this study. This highly selective expression pattern allowed for imaging of carcinoma lesions as small as 1 cm in diameter. A logical concern with regards to the development of stromal targeting for cancer therapy is the notion that the stromal response to cancer (at the cellular and molecular level (Welt et al., 1994)) is similar to that of wound healing and Seprase is known to be expressed in healing dermal incisions (Rettig, 1988). However, the patients involved in this study would have had extensive scarring due to surgery but did not show any localisation of ^{131}I-mAbF19 at these sites. Another phase I trial and limited phase II trial have been conducted with an unconjugated, humanized version of monoclonal antibody F19 to Seprase called sibrotuzumb (Hofheinz et al., 2003; Scott et al., 2001).

Results from this trial found that sibrotuzumb was well tolerated and relatively safe and a pharmacokinetic model has been developed for it (Kloft et al., 2004). Similar to the results obtained in the previous phase I trial (Welt et al., 1994), trace labelling with ^{131}I and detection imaging revealed that sibrotuzumb specifically accumulated in the tumours and not in the normal tissues (Scott et al., 2003). These studies suggest that mAbF19 has diagnostic and therapeutic applications and could be constructed to provide novel immune and non-immune effector functions (Kelly, 2005; Welt et al., 1994).

1.9.2 Inhibition of Seprase protease activity

The protease activity of Seprase is another appealing target for inhibiting the promotion of tumour growth by Seprase. Seprase is inhibited by the serine protease inhibitors DFP (0.005mM), PMSF (1mM), AEBSF (5mM) and APSF (0.5mM) (Aoyama and Chen, 1990; Collins et al., 2004; Pineiro-Sanchez et al., 1997; Scanlan et al., 1994). The 170kDa protease was sensitive to others inhibitors, such as NEM and $HgCl_2$, that bind the sulfhydryl group of amino acid residues of the proteases (Aoyama and Chen, 1990). The inhibition studies performed suggest that Seprase contains a catalytically active serine residue.

A series of dipeptide proline diphenyl phosphonates were designed against DPPIV and Seprase (Gilmore et al., 2006). The most potent inhibitor for both enzymes was found to be Gly-$Pro^{P}(OPh)_2$, which exhibited overall second-order rate constants of inactivation of 5.24×10^5 $M^{-1}min^{-1}$ and 1.06×10^4 $M^{-1}min^{-1}$ against DPPIV and Seprase respectively. In the case of Seprase, it was found that the nature of the P_2 residue of the inhibitor had a less pronounced effect on the second order rate constants, in contrast to DPPIV. It was found for DPPIV that the second order rate constants varied by one order of magnitude between the most effective [Gly-$Pro^{P}(Oh)_2$] and the least effective [Ala-$Pro^{P}(Oh)_2$] inhibitors. This group also found that Gly-$Pro^{P}(Oh)_2$ and Tyr-$Pro^{P}(Oh)_2$ exerted a considerable dose dependent anti-invasive effect on the LOX melanoma cell line, *in vitro*.

The inhibition of the catalytic activity of Seprase with anti-catalytic antibodies has been investigated. It was found that the tumour-promoting effects of Seprase expression were inhibited by the use of these antibodies (Cheng et al., 2002; Ghersi et al., 2002). This group also observed the over-expression of Seprase, leading to increased tumourigenicity, compared to a mutated S624A Seprase (Cheng et al., 2005). In this same study, Seprase activity was shown to be inhibited by Val-boroPro, a boronic acid inhibitor, with an IC_{50} of 4 x 10^{-8}mol/L (Cheng et al., 2005). However, this inhibitor was also found to inhibit DPP IV , DPP-7, -8 and -9 (Edosada et al., 2006b; Flentke et al., 1991). N-acyl-Gly-Pro dipeptides were identified as Seprase selective substrate motifs and a second boronic acid inhibitor was designed, Ac-Gly-BoroPro (Edosada et al., 2006b) and it inhibited these prolyl peptidases with K_i values ranging from ~9- to 5400-fold higher than that for Seprase inhibition (K_i = 23nM for Seprase) (Edosada et al., 2006b). This indicated that the Ac-Gly-Pro motif conferred significant Seprase selectivity and suggests that this could possibly be used to test Seprase as a therapeutic target.

These studies together suggest that Seprase activity plays an important role in the promotion of tumour growth and that the Seprase protease is a good potential target for therapies designed to slow tumour growth.

1.9.3 Disruption of signalling

The targeting of antigens selectively expressed on the surface of tumour stromal fibroblasts or tumour capillary endothelial cells is currently being explored for the immunotherapy of cancer (Fassnacht et al., 2005; Lee et al., 2005b; Tahtis et al., 2003). By targeting or preventing the generation of tumour stroma or angiogenic blood vessels, tumour lesions may be deprived of the essential support services or nutrients required for survival and growth (Dvorak, 1986). For reasons discussed above, Seprase is an important antigen for targeted therapy of the tumour stroma.

Garin-Chesa et al., (1990) propose that radiolabelled or toxin-conjugated mAbs or inflammatory mAb isotypes detecting Seprase may be used to induce cell damage in the Seprase positive supporting tumour stroma. This would lead to tumour cell necrosis (death) and inflammatory cell infiltrates. Recruitment of additional Seprase positive reactive fibroblasts would renew the target cell population and aid formation of fibrous capsules enclosing and isolating epithelial tumour cells.

Studies in murine models have shown that immunologic targeting of the tumour vasculature, a key element of tumour stroma, can lead to protective immunity in the absence of significant pathology (Lee et al., 2005b). Targeting the tumour stroma, instead of tumour cells, reduces the incidence of immune evasion. This is due to the fact that stromal cells are diploid, unlike tumour cells, and are therefore more genetically stable and exhibit limited proliferative capacity.

1.10 Biomarkers

Genes that are expressed in a highly tissue- or disease-specific manner provide possible targets for (i) anti-cancer therapeutics (ii) the early detection of cancer and (iii) monitoring of disease during and after treatment (Brown et al., 2005). Tumour markers are substances occurring in blood, tissue, saliva or urine, that are associated with cancer and whose measurement or identification is useful in patient diagnosis or clinical management. These soluble molecules are usually glycoproteins detected by monoclonal antibodies. The ideal tumour marker should be (1) specific for the cancer for which it is testing; (2) not present in any other conditions; and (3) the concentration should change with the amount of malignant tissue present. An ideal tumour maker could be used for screening, diagnosis and the monitoring of disease progression and should be easily and reproducibly measured.

To date there is no serum or saliva tumour marker(s) for breast cancer screening. There are three main methods of screening for breast cancer: mammography, clinical breast examination and breast self-examination. A mammogram is a breast x-ray that is the most proven screening test for reducing the risk of dying from breast cancer. It is important to remember that breast cancer cannot be diagnosed by mammography alone. In addition, a patient may be asked to have further testing (e.g., ultrasound or biopsy) because something on the mammogram needs more evaluation. One study found that 11 percent of mammograms performed in the United States require additional evaluation and the lesion turns out to be benign in more than 90 percent of these cases (Brown et al., 1995).

Currently there are no ideal tumour markers which are recommended for screening of the general population and most are restricted to monitoring cancers once they have been detected and diagnosed using other methods. A new prognostic marker is also needed to identify patients who are at the highest risk for developing metastasis. Most tumour markers have too many false positives from benign conditions to make screening feasible. Many only clearly identify malignancy once the cancer is sufficiently advanced to make this detection of limited use. Table 1.10 shows the currently used tumour markers

Tumour Marker	Normal Value	Primary Tumour (s)	Additional Associated Malignancies	Benign Conditions	Sensitivity
CA 27.29	< 38 units per ml	***Breast Cancer***	***Colon, gastric***, hepatic, ***lung***, ***pancreatic***, ***ovarian***, prostrate cancers	***Breast***, Liver and kidney disorders, ovarian cysts	Elevated in about 33% of early stage breast cancers and about 67% of late-stage breast cancers
CEA	<2.5 ng per ml in non-smokers <5 ng per ml in smokers	***Colerectal Cancer***	***Breast, Lung, gastric, pancreatic***, bladder, medullary thyroid, head and neck, cervical, and hepatic cancers, lymphoma, melanoma	Cigarette smoking, peptic ulcer disease, inflammatory bowel disease, pancreatitis, hypothyroidism, cirrhosis, biliary obstruction	Elevated in less than 25% of early-stage colon cancers and 75% of late-stage colon cancers.
CA 19-9	< 37 units per ml	Pancreatic cancer, biliary tract cancers	***Colon***, esophageal, and hepatic	Pancreatitis, biliary disease, cirrhosis	Elevated in 80% to 90% of pancreatic cancers and 60% to 70% of the late-stage colon cancers.
CA 125	< 35 units per ml	Ovarian Cancer	Endometrial, fallopian tube, ***breast, lung***, esophageal, ***gastric,*** hepatic, and ***pancreatic*** cancers	Menstruation, pregnancy, fibroids, ovarian cysts, pelvic inflammation, cirrhosis, ascites, pleural and pericardial effusions, endometriosis	Elevated in about 85% of ovarian cancers; elevated in only 25% of early stage ovarian cancers
AFP	< 5.4 ng per ml	Hepatocellular carcinoma, non-seminomatous germ cell tumours	***Gastric***, biliary and ***pancreatic*** cancers	Cirrhosis, viral hepatitis, pregnancy	Elevated in 80% of heptocellular carcinomas

Table 1.10 Conditions Associated with Elevated Tumour Marker Levels in Serum

CA = cancer antigen; CEA = carcinoembryonic antigen; AFP = alpha-fetoprotein. (Perkins et al., 2003)

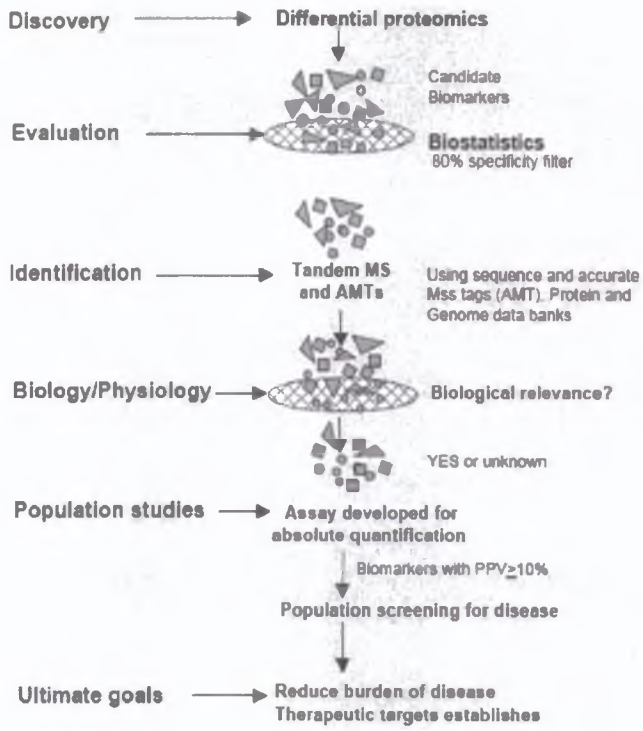


Figure 1.10.1 The roadmap for cancer biomarker development

(Plebani, 2005)

Researchers have proposed sequential phases for cancer biomarker development in the form of a road map (see Figure 1.10.1). This takes into consideration the different steps that gradually expose the structure, biological relevance and the predictive value of the novel biomarker for the diagnosis and screening of neoplastic disease (Plebani, 2005).

As described previously, tumour cells, in order to establish a metastasis, have to invade their surrounding host tissue, enter the circulatory system, arrest in capillary beds of distant organs, invade the host tissue and proliferate. The development of an assay to detect these cells before the manifestation of distant metastasis should therefore be useful for patient prognosis. There is substantial evidence that Seprase is involved in the degradation of the extracellular matrix and subsequent invasion of the surrounding host tissue by cancer cells. The cell surface localization of Seprase makes it an ideal candidate for targeted therapeutic agents to malignant breast cells, i.e. the selective inhibition of Seprase should prevent a critical part of tumour invasion. However, to date, the pathophysiologic significance of its expression remains poorly understood.

Malignant transformation from normal to cancerous tissue is associated with cell-surface glycoprotein modifications (Cordero et al , 2000) These proteins can be released in the circulation through increased cell turnover, secretion or shedding from the malignant cells and have been considered as potential tumour markers for helping in screening, diagnosis, staging, prognosis and monitoring of cancer therapy (Cordero et al , 2000) It may be possible that during the early stages of breast cancer, for example, Seprase activity becomes over-expressed and starts to appear in the serum in relatively low levels Therefore, Seprase would have a huge potential to be a highly specific 'biomarker' for human breast cancer

1 11 Aims of the project

The initial objective of this research project was to purify and identify the second Z-Gly-Pro-AMC degrading activity found in bovine serum in collaboration with Dr Patrick Collins, Dublin City University Previous studies were unsuccessful in purifying this second Z-Gly-Pro-AMC degrading activity from bovine serum to a level of purity and quantity required to identify the enzyme by protein sequencing Previous attempts at identifying this activity demonstrated that there were contaminating proteins which required removal before the enzyme could be identified This would involve developing a purification method that would remove these contaminating proteins and to develop a method which would allow the native enzyme to be detected in a PAGE gel Upon identification of the activity as Seprase, biochemical characterisations studies were designed, including investigating the gelatinase activity, the temperature and pH profiles of the soluble form of bovine Seprase More detailed substrate specificity studies were also designed using a dipeptide library specific for proteases These studies would provide valuable insight into the catalytic and structural properties of bovine Seprase and open the way for the better design of inhibitors and pharmacological agents This project would also involve the further development and optimisation of the specific assay for Seprase

Upon identification of the activity as Seprase, the project also focused on the cloning of the human Seprase gene Upon cloning of the gene, the project would focus on the development of an expression system capable of producing catalytically active recombinant human Seprase If expression of active recombinant human Seprase was possible in a prokaryotic system, this would enable the rapid production of large quantities of the enzyme Alternatively, a mammalian or yeast-based expression system would be explored The availability of purified recombinant human Seprase would allow a series of detailed molecular, biochemical and kinetic studies to be carried out Using the recombinant Seprase expression system, a series of site-specific mutagenesis experiments would be carried out, giving valuable insight into the catalytic and

structural properties of human Seprase and opening the way for the design of inhibitors and pharmacological agents

Collaboration with BreastCheck and the Symptomatic Breast Care Unit of the Mater Misericordiae Hospital, Dublin was established to analyse clinical samples from patients with confirmed breast cancer, in the hope of observing significantly elevated Seprase levels in the serum The results from this study could lead to a greater understanding of Seprase expression in breast cancer and possibly lead to the identification of Seprase as a specific biomarker and thus the development of a biomarker diagnostic test

To date, Seprase expression has been studied using immunohistochemistry and a specific sensitive assay for Seprase has not been developed Such studies as those described in this research thesis would significantly add to published data on this recently discovered protease, Seprase, and represent the first detailed study of a soluble form of the enzyme from bovine serum This research thesis would also provide a specific assay for Seprase which could be used to detect Seprase

Chapter 2

Materials and Methods

All chemicals and reagents were obtained from Sigma-Aldrich unless otherwise stated

2 1 Protein Concentration

2 1 1 Quantitative determination by Biuret Assay

The Biuret assay (Sigma) was used to monitor protein in post column chromatography fractions and also to quantify protein in crude serum and post chromatography active enzyme pooled sample containing greater than 2 mg/ml protein (Gornall et al , 1949) For protein quantification, all samples were dialysed into ultra-pure water and diluted appropriately to achieve a concentration suitably determinable by the assay Bovine serum albumin (BSA) standards (2-10mg/ml) and enzyme samples were prepared and assayed in triplicate Fifty microlitres of standard or sample was added in triplicate to 200 µl of Biuret reagent, and allowed stand at room temperature for 5 minutes Absorbances were determined at 540nm using a Tecan Spectra Plate Reader

2.1 2 Quantitative determination by Standard BCA Assay

The standard bicinchominic acid (BCA) assay (Sigma) was used for protein monitoring and quantification in column fractions and pooled samples which contained less than 2mg/ml of protein (Smith et al , 1985) Samples were prepared as outlined in section 2 1 1 and BSA standards in the range 0-2mg/ml were included Twenty-five microlitres of standard or sample was added in triplicate with 200 µl of BCA reagent, and incubated at 37°C for 30 minutes Absorbances were read at 570nm

2 1 3 Quantitative determination by Coomassie Plus Assay

The coomassie plus protein assay was used to monitor and quantify protein concentration in post column chromatography fractions and pooled samples containing between 2 5 and 25 µg/ml protein (Bradford, 1976) Samples were prepared as outlined in section 2 1 1 and BSA standards in the range 2 5-25 µg/ml were included One hundred and fifty microlitres of standard or sample was added in triplicate to 150 µl of coomassie plus reagent (Pierce), allowed stand at room temperature for 5 minutes and finally read at 595nm

2 2 Fluorescence Spectrometry using 7-Amino-4-Methyl-Coumarin (AMC)

2 2 1 AMC Standard Curves

100 µM stock AMC solution containing 4% v/v methanol was prepared in 100 mM potassium phosphate, pH 7 4 All lower AMC concentrations were obtained using 100mM potassium phosphate, pH 7 4 containing 4% v/v methanol as diluent Stock solution and standards were stored in the dark at 4°C for up to one month Standard curves in the range 0-5 µM and 0-20 µM AMC were prepared in triplicate by combining 25 µl (100 µl)* of 100 mM potassium phosphate pH 7 4, 100 µl (400 µl)*of appropriate AMC concentration and 175 µl (1 ml)* of 1 5 M acetic acid Fluorimetric analysis of these samples was achieved using a Perkin Elmer LS50 Fluoresence Spectrophotometer at excitation and emission wavelengths of 370nm and 440nm respectively Excitation slit widths were maintained at 10nm while emission slit widths were adjusted accordingly for the range being analysed

** Note Bracketed values represent the volumes used for the cuvette assay*

2 2 2 Inner Filter Effect

The inner filter or quenching effect of enzyme samples was determined by combining 25 µl (100 µl) of enzyme sample, 100 µl (400 µl) appropriate AMC dilution and finally 170 µl (1 ml) of 1 5 M acetic acid The filtering effect of crude serum samples was assessed in the presence and absence of 5 µl, 2 5x10^{-4} M JTP-4819* in 10% v/v MeOH The samples were all assayed in triplicate as described in section 2 2 1

** Note Even though the peptidase under study is named ZIP (Z-Pro-prolinal Insensitive Peptidase), JTP-4819 (see Figure 2 2 1), which is also a potent and specific inhibitor of Prolyl Oligopeptidase, is used throughout this work for distinguishing between these peptidases This is due to the commercial unavailability of Z-Pro-prolinal*

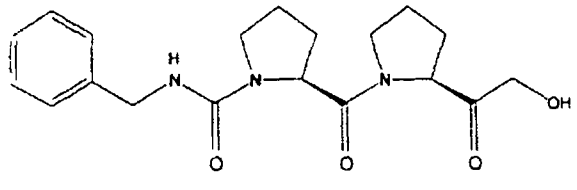


Figure 2 2 1 Chemical Structure of JTP-4819
(Toide et al , 1995)

2 3 Enzyme Assays

2 3 1 Substrate Preparation

Both Prolyl Oligopeptidase (PO) and Z-Pro-prolinal Insensitive Peptidase (ZIP) activities were determined using the fluorimetric substrate Z-Gly-Pro-AMC, with modification of the original protocol of Yoshimoto *et al* (1979) 10 mM Z-Gly-Pro-AMC stock was prepared in 100% methanol, aliquoted and stored at -20°C 100 µM substrate containing 4% v/v MeOH was prepared by slowly adding 300 µl MeOH and 100 µl Z-Gly-Pro-AMC stock to 9 6 ml of 100 mM potassium phosphate, pH 7 4 at 37°C Prolyl Oligopeptidase activity was determined with 10 mM DTT in the above substrate, while for determination of ZIP activity 500 mM NaCl was included

2 3 2 Quantitative Z-Gly-Pro-AMC Degrading Activity Measurements

2 3 2 1 Cuvette

Four hundred microlitres of the 100 µM substrate was added to 100 µl enzyme sample in triplicate and incubated at 37°C for 60 min Both samples and substrate were preincubated for 10 min at 37°C to allow them reach thermal equilibrium Reactions were terminated by the addition of 1 ml of 1 5 M acetic acid Blanks or negative controls were prepared by adding 1 ml of 1 5 M acetic acid to 100 µl of enzyme sample prior to substrate addition and incubation at 37°C for 60 min Formation of AMC was measured as described in section 2 2 1 End point measurements were allowed, as the enzyme assay was linear with respect to time and enzyme concentration up to 60 min (Birney and O'Connor, 2000) Fluorimetric intensities observed were converted to nanomole AMC released per minute per ml using the appropriate standard curve as outlined in section 2 2 2 One unit of activity is defined as the amount of enzyme which releases 1 nanomole of AMC per minute at 37°C (see Appendix A for calculations)

2 3 2 2 Microtitre Plate

One hundred microlitres of the 100 μM substrate was added to 25 μl enzyme sample in triplicate and the microtitre plate was incubated at 37°C for 60mins Both samples and substrate were pre-incubated at 37°C to allow them reach thermal equilibrium Reactions were terminated by the addition of 175 μl of 1 5 M acetic acid Blanks or negative controls were prepared by adding 175 μl of 1 5 M acetic acid to 25 μl of enzyme sample prior to substrate addition and incubation at 37°C for 60mins

Fluorimetric analysis of these samples was achieved using a Perkin Elmer LS50 Fluorescence Spectrophotometer with plate reader attachment, at excitation and emission wavelengths of 370nm and 440nm respectively Excitation slit widths were maintained at 10nm while emission slit widths were adjusted accordingly for the range being analysed

End point measurements were allowed, as the enzyme assay was linear with respect to time and enzyme concentration up to 60mins Fluorimetric intensities observed were converted to nanomole AMC released per minute per ml using the appropriate standard curve (Appendix A) Enzyme units were defined as nanomole of AMC released per minute at 37°C

2.3 3 Quantitative Z-Pro-prolinal Insensitive Z-Gly-Pro-AMC (ZIP) Degrading Activity Measurement

Section 2 3 2 describes the determination of ZIP activity (which shall henceforth be referred to as Seprase) in situations where it was most certainly separated from prolyl oligopeptidase activity In crude bovine serum, the following assay modifications were necessary to distinguish Seprase activity from PO activity, 25 μl (100 μl)* of enzyme sample was pre-incubated for 15 minutes at 37°C with 5 μl (20 μl)* of $2\,5x10^{-4}$ M JTP-4819 in 10% v/v MeOH prior to substrate addition Negative controls were also included, as in section 2 3 2, incorporating JTP-4819 and determinations were carried out in triplicate AMC released was determined fluorimetrically, with end point measurements taken as in section 2 3 2 Fluorimetric intensities obtained for each sample were converted to nanomole of AMC released

per minute per ml using the standard curves incorporating JTP-4819, prepared as outlined in section 2 2 1 Enzyme units were defined as nanomole of AMC released per minute at 37°C

** Note Bracketed values represent the volumes used for the cuvette assay*

2 3 4 Non-Quantitative Z-Gly-Pro-AMC Degrading Activity Measurements

A non-quantitative fluorimetric microtitre plate assay was developed to assist in the rapid identification of Z-Gly-Pro-AMC degrading activities in post-column chromatography fractions Two hundred microlitres of 100 µM Z-Gly-Pro-AMC in 4% MeOH containing 500 mM NaCl (ZIP) or 10 mM DTT (PO), at 37°C was added to 100 µl of sample in each well Post-phenyl sepharose fractions were assayed in the presence and absence of 5 µl 2 5x10^{-4} M JTP 4819 in 10% MeOH The microtitre plate was incubated at 37°C for 30 minutes AMC released was determined fluorimetrically as outlined in section 2 2 1, using the Perkin Elmer LS-50B plate reader attachment

2.4 An Optimised Purification Procedure for Bovine Serum Seprase

Purification steps were all carried out at 4°C apart from the final size exclusion column, which was performed using BioRad Biologic HR FPLC at room temperature based on the method of Collins and O'Connor (2003)

2 4 1 Bovine Serum Preparation

Bovine whole blood was collected from a freshly slaughtered animal and stored at 4°C over 24 hours to allow clot formation The remaining un-clotted whole blood was decanted and centrifuged at 6000 rpm for 1 hour at 4°C using a Beckman J2-MC centrifuge fitted with JL-10 5 rotor The supernatant and loose cellular debris was decanted and re-centrifuged at 20,000 rpm for 15min using a JL-20 rotor The final serum was collected and stored at -17°C in 20 ml aliquots

2 4 2 Phenyl Sepharose Hydrophobic Interaction Chromatography

A 20 ml Phenyl Sepharose CL-4B hydrophobic interaction column (2 5cm x 7cm) was equilibrated with 100 ml of 100 mM potassium phosphate containing 200 mM ammonium sulphate, pH 7 4 Solid ammonium sulphate was dissolved in 20 ml of bovine serum to give a final concentration of 200 mM This sample was then applied to the equilibrated column followed by a 100 ml wash with 100 mM potassium phosphate containing 200 mM ammonium sulphate, pH 7 4 The column was then washed with 100 ml of 100 mM potassium phosphate containing 50 mM ammonium sulphate, pH 7 4 Bound protein was eluted isocratically with a 100 ml ultra-pure water wash Equilibration of the column was carried out at a flow rate of 1 ml/min, while all other steps were performed at 2 ml/min Five millilitre fractions were collected throughout the procedure and were assayed for Seprase activity according to section 2 3 4 Protein content in each fraction was determined using the biuret assay according to section 2 1 1 Fractions containing Seprase activity were pooled to yield post-phenyl Sepharose Seprase Enzyme activity and protein content were quantified using the fluorimetric assay as outlined in section 2 3 2 and the biuret assay as in section 2 1 1 The phenyl Sepharose resin was regenerated with a 100 ml wash of pure ethanol at 0 5 ml/min, followed by 150 ml of equilibration buffer

2 4 3 Calcium Phosphate Cellulose Chromatography

2 4 3 1 Resin Preparation

One litre of 500 mM sodium hydrogen phosphate was added (at 6 4 ml/min) to 1 5 L of constantly stirred 500 mM calcium chloride at room temperature Following a 15 min agitation, 1 5 ml of concentrated ammonia solution was added and stirred for a further 10 min The precipitated gel was allowed to settle and the supernatant decanted and discarded 1 L of ultra-pure water was added and the stirring continued for 5 min Settling and decanting was performed and this procedure was repeated until the gel was washed with 10 L of ultra-pure water The washed calcium phosphate gel was stored in 1 L of ultra-pure water at 4°C Sigma cellulose type 50 (10g) was soaked overnight in 200 ml 500 mM potassium phosphate containing 150 mM potassium chloride, pH 6 8 The cellulose was washed eight times with ultra-pure water and dried overnight at 70°C Washed and dried cellulose (2 g) was dissolved in 20 ml 20 mM potassium phosphate, pH 7 4 and added to 24ml of evenly suspended calcium phosphate gel The calcium phosphate cellulose was poured into a column (2 5 x 7 0cm) and allowed to settle, yielding 15 ml of packed resin

2 4 3 2 Column Chromatography

The column was equilibrated at 1 ml/min with 10 mM potassium phosphate, pH 7 4 The post-phenyl Sepharose Seprase was concentrated to 10 ml and applied to the column followed by a 100 ml wash with 10 mM potassium phosphate, pH 7 4 The column was then washed with 100 ml of 170 mM potassium phosphate, pH 7 4 Bound protein was eluted with 100 ml of 500 mM potassium phosphate, pH 7 4 Loading and washing steps were performed at 1 ml/min, while elution was performed at 2 ml/min Five millilitre fractions were collected throughout the procedure, and again were assayed for Seprase activity according to section 2 3 4 Protein content in each fraction was determined using the standard BCA assay as outlined in section 2 1 2 Fractions containing Seprase activity were pooled to yield post calcium phosphate cellulose Seprase Enzyme activity and protein content were then quantified as outlined in section 2 3 2 and 2 1 2

2 4 4 Cibacron Blue 3GA Chromatography

100 ml of 20 mM potassium phosphate, pH 7 4 was used to equilibrate a 20 ml Cibacron blue 3GA resin The post calcium phosphate cellulose Seprase was concentrated and then dialysed overnight against 2 L of 20 mM potassium phosphate, pH 7 4 After sample application, the column was washed with 100 ml of 20 mM potassium phosphate, pH 7 4 to remove any unbound protein Elution was performed using a 100 ml linear 0-2M NaCl gradient in 20 mM potassium phosphate, pH 7 4 Loading, washing and elution were all performed at 1 ml/min Five millilitre fractions were collected and assayed for Seprase activity according to section 2 3 4 Protein content in each fraction was determined using coomassie plus protein reagent as in section 2 1 3 Fractions containing Seprase activity were pooled and enzyme activity and protein content were quantified as outlined in section 2 3 2 and 2 1 3

2 4 5 Sephacryl S-300 Size Exclusion Chromatography

A HiPrep 16/60 Sephacryl S-300 High Resolution (HR) size exclusion chromatography column (Amersham, GE Healthcare) was attached to a fast protein liquid chromatography system The column was equilibrated at 0 8ml/min with 250 ml of 100 mM potassium phosphate containing 150 mM NaCl, pH 7 4, which had been filtered and degassed The post cibacron blue 3GA Seprase pool was concentrated to 2 ml and applied to the column followed by a 150 ml wash with equilibration buffer at 0 8 ml/min Five millilitre fractions were collected and assayed for Seprase activity as outlined in section 2 3 4 Protein content in each fraction was monitored online at 280nm and also using coomassie plus protein reagent as outlined in section 2 1 3 Fractions containing Seprase activity were pooled and enzyme activity and protein content were quantified as described in section 2 3 2 and 2 1 3

2 4 6 Preparation of dialysis tubing

The required amount of tubing was placed in a 1 L glass beaker and rinsed thoroughly with distilled water The beaker was filled with distilled water and ~1 g of EDTA was added The beaker contents were brought to the boil and boiled for 2 min The water was allowed to cool and then poured off The beaker was re-filled with fresh distilled water and boiled again for 2 min The water was allowed to cool and poured off The tubing was thoroughly rinsed with distilled water and was stored at 4°C in distilled water

2 5 Polyacrylamide Gel Electrophoresis

SDS polyacrylamide gel electrophoresis based on the method of Laemmli (1970) was employed to determine the effectiveness of the purification procedure

2 5 1 Preparation of SDS Gels

10% resolving and 4 5% stacking gels were prepared as in Table 2 5 1 Gels were cast using an ATTO vertical mini electrophoresis system The resolving gel solution (Table 2 5 1) was degassed and filtered, TEMED added, mixed and the gel poured immediately An overlay of ethanol/water was placed over the resolving gel After polymerisation, the overlay was removed The stacking gel solution (Table 2 5 1) was degassed and filtered, TEMED added, mixed and the gel poured immediately A comb was placed into the top of the gel liquid to form loading wells, and the gel was allowed polymerise

Solution	10% Resolving Gel	15% Resolving Gel	4% Stacking Gel
1 5 M Tris-HCl, pH 8 8 Resolving Gel buffer	1 625 ml	1 625 ml	-
0 5 M Tris-HCl, pH 6 8 Stacking Gel buffer	-	-	0 625 ml
dH_2O	2 64 ml	1 56 ml	1 538 ml
Acrylamide/Bis-acrylamide 30%/0 8% (w/v)	2 17 ml	3 25 ml	0 335 ml
10% (w/v) Ammonium Persulphate	32 5 µl	32 5 µl	12 5 µl
20% (w/v) SDS	32 5 µl	32 5 µl	12 5 µl
TEMED	3 25 µl	3 25 µl	2 5 µl

Table 2 5 1 SDS PAGE Gel Preparation

2.5.2 Sample Preparation

Samples generated from the different stages of the purification of bovine serum Seprase (Post-phenyl Sepharose, post-calcium phosphate cellulose, post-cibacron blue 3GA and post-S-300 gel filtration) were extensively dialysed overnight into ultra-pure water. Each dialysed sample was added to an equal volume of 2x solubilisation buffer which consisted of 20% v/v glycerol, 4% w/v SDS, 10% v/v 2-mercaptoethanol, 0.004% w/v bromophenol blue and 125 mM Tris/HCl, pH 6.8. Samples were boiled for 2 minutes and stored on ice until application.

2.5.3 Sample Application

Twenty microlitres of each prepared sample from section 2.5.2 was applied to the 10% SDS PAGE gel. The Sigma high molecular weight standards used consisted of Myosin (205kDa), β-Galactosidase (116kDa), Phosphorylase B (97kDa), Fructose-6-phosphate kinase (84kDa), BSA (66kDa), Glutamic dehydrogenase (55kDa), Ovalbumin (45kDa) and Glyceraldehyde-3-phosphate dehydrogenase (36kDa). Coloured markers were also used (Figure 2.5.1). Twenty microlitres of the standards solution was also applied to the gel, which was then run at 125V for 2h at room temperature. Running buffer was 25 mM Tris/HCl, 192 mM glycine, 0.1% SDS, pH 8.3.

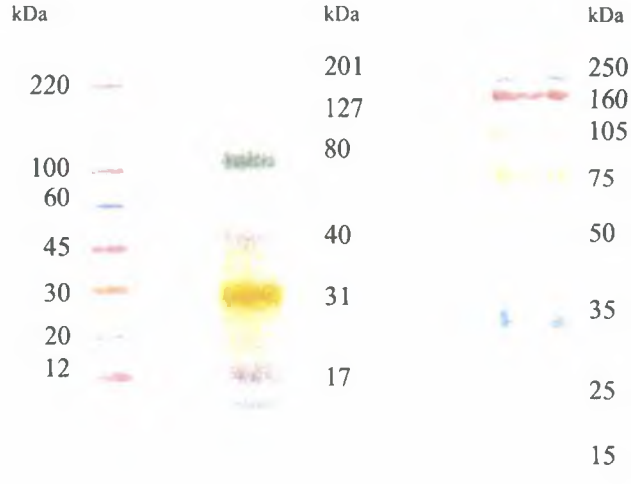


Figure 2.5.1 Coloured Markers
Coloured Markers used in SDS analysis. *Left*: Sigma ColourBurst Markers. *Middle*: Kaleidoscope Pre-stained Standards (BioRad). *Right*: Full Range Rainbow Markers (Amersham). Molecular weight markers are listed.

2 5 4 SDS PAGE Gel Staining

Polyacrylamide gels were removed from the electrophoresis chamber and washed with dH_2O for 5 min Routinely, gels were stained for 60 min in a 0 25% (w/v) Coomassie blue solution containing 10% (v/v) Acetic Acid/45% (v/v) methanol and then de-stained overnight using 10% (v/v) Acetic Acid/45% (v/v) methanol Subsequent soaking in dH_2O enhanced the protein bands further If this did not sufficiently stain the gel, the more intense silver staining method (Blum et al , 1987) was used as outlined in Table 2 5 2 An image of each stained gel was scanned or captured using a Casio Exilim Ex-Z40 digital camera coupled to Casio software

Step	Duration	Reagent
Fix	60 min	50% ethanol, 12% acetic acid, 0 05% formaldehyde (37% stock)
Wash	3 x 20 min	50% ethanol
Pre-Treat	1 min	200 µl of a 5% $Na_2S_2O_3$ x H_2O stock solution in 100 ml dH_2O
Rinse	2 x 20 s	dH_2O
Impregnate	20 min	0 1g $AgNO_3$, 70 µl formaldehyde in 100 ml dH_2O
Rinse	2 x 20 s	dH_2O
Development	10 min (max)	3g Na_2CO_3, 50 µl formaldehyde, 4 µl $Na_2S_2O_3$ x H_2O stock solution in 100 ml dH_2O
Stop	5 min	0 1 M EDTA

Table 2 5 2 Silver Staining Procedure

2 6 Identification of this Z-Pro-prolinal Insensitive Peptidase

2 6 1 Determination of N-Terminal Protein Sequence

N-terminal sequencing of ZIP was performed commercially by Bryan Dunbar's group at the Aberdeen Proteome Facility, University of Aberdeen, Scotland The enzyme sample had to be firstly separated by SDS PAGE, then electroblotted onto polyvinylidene difluoride (PVDF) membrane and stained The recommended method is shown below (Keen, 2006)

2 6 1 1 SDS Polyacrylamide Gel Electrophoresis

An SDS PAGE gel was precast according to Table 2 5 2 and section 2 5 1 and stored overnight at 4°C to prevent N-terminal blocking by free amines Fifty millilitres of bovine serum was purified as in section 2 4 and the final purified ZIP sample was dialysed extensively overnight against ultra-pure water and concentrated to 500 µl using a Thermo Savant DNA110 SpeedVac An equal volume of sample was added to an equal volume of solubilisation buffer (see section 2 5 2) and boiled for 2 minutes Coloured markers were employed as to aid visualising electrophoresis and electrotransfer The gel was pre-run at 50V for 30min with 200 µM thioglycolic acid in the upper reservoir in order to provide a scavenger for free radicals The chamber was emptied, rinsed with reservoir buffer and further electrophoresis was performed at a constant current of 20 mA for 2hr

2 6 1 2 PVDF Electroblotting

1 L of 10x CAPS buffer was prepared and adjusted to pH 11 using sodium hydroxide 1x electroblotting buffer in 10% MeOH was prepared and stored at 4°C On removal of the gel it was placed in electroblotting buffer for 1 h, while the PVDF membrane was dipped in MeOH for 10 sec and also equilibrated in electroblotting buffer for 1 h A transblotting sandwich was made using 8 layers of pre-soaked filter paper, cut to the exact size of the gel The PVDF membrane was placed on top of the filter paper followed by the gel Eight more layers of equilibrated filter paper were added to complete the sandwich Electroblotting was performed at 0 8mA/cm^2 for 1 h 10min Removal of the membrane was followed by a rapid membrane wash in

ultra-pure water It was then saturated in 100% MeOH for 10 s and stained with coomassie blue for 3 min Destaining in 1% acetic acid was repeated twice, then the blot was extensively washed with ultra-pure water and finally allowed to air dry The blot was wrapped in clingfilm and sent for sequencing

2 6 2 Protein G Affinity Chromatography

2 5 ml of protein G resin was equilibrated with 10 ml of 100mM potassium phosphate containing 200 mM NaCl, pH 7 4 One to two millilitres of concentrated purified ZIP sample was applied to the affinity resin and allowed to shake slowly for 30 mins at 4°C (Batch bind) The resin was allowed to settle and the supernatant removed The resin was then washed with 10 ml with equilibration buffer, allowed settle and the supernatant removed A second 10 ml wash of 100 mM potassium phosphate containing 500mM NaCl, pH 7 4 was then performed, the resin again allowed to settle and the supernatant removed Each supernatant was pooled and assayed for ZIP activity according to section 2 3 4 Bound protein was eluted using a 10 ml wash with 0 1 M glycine/HCl, pH 2 2

2 6 3 Gelatin Sepharose Chromatography

Five millilitres of gelatin sepharose was equilibrated with 15 ml 100 mM potassium phosphate containing 200 mM NaCl, pH 7 4 The post-protein G ZIP sample was concentrated, applied to the resin and allowed to batch-bind slowly for 1 hour at 4°C The resin was allowed to settle and the supernatant removed This was followed by a wash with 15 ml of equilibration buffer, the resin allowed to settle and the supernatant removed A second 15 ml wash of 100 mM potassium phosphate containing 1 M NaCl, pH 7 4 was performed, with the resin again allowed to settle and the supernatant removed Each supernatant was pooled and assayed for ZIP activity according to section 2 3 4 Bound protein was eluted with a 15 ml wash of 4 M urea, pH 7 0

2 6 4 UV Zymogram Development

2 6 4 1 Native Polyacrylamide Gel Electrophoresis

A 10% native polyacrylamide gel was prepared as in section 2 5 1 except that no SDS was present in the gel and the gels in this case were cast in the large ATTO electrophoresis system The cast gels were stored overnight at 4°C Purified ZIP sample was dialysed extensively overnight into ultra-pure water and concentrated to 500 µl using a ThermoSavant DNA 110 SpeedVac

10x (non-denaturing) solubilisation buffer was prepared as follows 3 2 ml 10% SDS, 2 ml 0 5M Tris/HCl pH 6 8, 1 6 ml glycerol, 0 05% (w/v) bromophenol blue and 1 2 ml of ultra-pure water, in a final volume of 8 ml This solution was stable for 4-6 weeks at 4°C, or for months if stored at -20°C

Running buffer was prepared as in section 2 5 3 without SDS being added 54 µl of concentrated ZIP sample was added to 6 µl 10x (non-denaturing) solubilisation buffer and applied directly to the gel The gel was run at a constant voltage of 125V at 4°C for 2 5 h

2 6 4 2 U V Zymogram Assessment

After native PAGE electrophoresis, the gel was rapidly removed and placed in 50 ml of 100 µM Z-Gly-Pro-AMC in 100mM potassium phosphate pH 7 4 containing 4% v/v MeOH and 500 mM NaCl It was incubated with shaking for 10-15 min at 37°C, after which excess substrate was removed and the gel placed under ultraviolet light in an image analyser for visualisation

2 6 5 Gelatin Zymogram Development

Gelatin zymography was performed to observe enzyme activity against the protein substrate, gelatin The gel was prepared by incorporating the protein substrate of interest (gelatin) within the polymerised acrylamide matrix The enzyme sample was resolved by 10% native PAGE gel in the presence of 1mg/ml gelatin The method of Laemmli (1970) was followed, excluding any reducing agents or boiling procedures

Samples were mixed 3 1 with 4x (non denaturing) solubilisation buffer, which consisted of 16% w/v SDS, 40% v/v Glycerol and 0 08% Bromophenol Blue The gels were run at 125V in running buffer (25 mM Tris, 192 mM Glycine) for up to 4 hours at 4°C After electrophoresis, the gel was washed for 30min in 2 5% Triton X-100 at room temperature, with one wash change The gel was then incubated overnight at 37°C in reaction buffer (100 mM potassium phosphate, pH 7 4, 500 mM NaCl) After staining with Coomassie stain (section 2 5 4) for 2 hours with shaking, the gel was destained in water until clear bands were visible Gelatin degrading activity was identified as a clear zone of lysis against a blue background

The identification of Seprase as a serine protease was also investigated by means of determining the sensitivity of the enzyme to serine protease inhibitors Prior to electrophoresis, enzyme samples were also incubated for 10 minutes in the presence of the serine protease inhibitors PMSF (5 mM), DFP (5µM) and JTP-4819 (2 5x10 ^{4}M - PO specific inhibitor) Samples were mixed as 3 1 with 4x solubilisation buffer as previously described

2 6 6 Wheat Germ Agglutinin (WGA) Lectin Affinity Chromatography

Twenty millilitres of 100 mM potassium phosphate, pH 7 4 containing 200 mM NaCl was used to equilibrate a 1 ml WGA lectin resin The purified post-FPLC Seprase sample was applied to the resin and allowed to batch-bind slowly for 3 hours, at 4°C The resin was allowed to settle and the supernatant removed The resin was put into a column and washed with 15 ml of equilibration buffer (100 mM potassium phosphate containing 200 mM NaCl, pH 7 4 Bound protein was eluted with an 11 ml wash of 100 mM potassium phosphate containing 0 5 M N-acetyl-D-glucosamine and 200 mM NaCl, pH 7 4 Washing and elution were all performed at 0 5 ml/min One millilitre fractions were collected and assayed for Seprase activity according to section 2 3 4 Protein content in each fraction was determined using Coomassie plus protein reagent (section 2 1 3) Fractions containing Seprase activity were pooled and enzyme activity and protein content were quantified as in section 2 3 2 and 2 1 3

2 7 Biochemical Studies

2 7 1 Temperature Studies

2 7 1 1 Determination of Optimum Assay Temperature

Purified Seprase was assayed in triplicate for 1 h at temperatures ranging from 4-70°C, as per section 2 3 2 with the temperature alteration Samples were pre-incubated for 15 minutes at the appropriate temperature prior to assay This ensured that the observed activity reflected the activity of the peptidase at the temperature under investigation A plot of Residual activity (%) versus Temperature (°C) was constructed Activity levels measured at 37°C were defined as 100%

2.7 1 2 Thermal Profile, (T_{50})

This method was used to determine the temperature at which the enzyme begins to unfold and lose catalytic activity Temperatures ranging from 4-75°C were used

Seprase was incubated at each temperature for 10 min and then 25 µl aliquots of enzyme were removed and stored on ice to prevent refolding, until assayed for remaining activity The enzyme solution was then placed in the next temperature for a further 10 min Activities were assayed in triplicate under standard conditions according to section 2 3 2 A plot of Residual activity (%) versus Temperature (°C) was constructed and the T_{50} (the temperature where half the residual enzymatic activity is lost) determined Activity levels measured at 37°C were defined as 100%

2 7 1 3 Thermal Stability

Purified Seprase was incubated at temperatures ranging from 4-70°C for up to 3 hours Aliquots of enzyme were removed at various time points and stored on ice All samples were brought to thermal equilibrium for 15 min at 37°C and assayed under standard conditions according to section 2 3 2 Activity levels were plotted as a percentage progress from initial activity (T_0)

2 7 1 4 Thermal Inactivation of Seprase

Purified Seprase was maintained at T_{50} (55°C) in a heated waterbath Aliquots were removed at appropriate time intervals and stored on ice All samples were brought to thermal equilibrium for 15 mins at 37°C and assayed under standard conditions according to section 2 3 2 A plot of Residual activity (%) versus Time (min) was constructed The first order rate constant of inactivation, k_i was determined by fitting the data to linear regression plots using Enzfitter® (Biosoft, Cambridge, U K) This allowed for the calculation of the half-life, $t_{½}$

2.7.2 Determination of pH optimum

The pH-activity profile of Seprase was determined by carrying out the standard activity assay as described in section 2 3 2 at pH range 6 0-10 5 This range was established using the following buffers 100 mM potassium phosphate for pH range 6 0-8 0, 100 mM Tris-HCl for pH range 7 5-9 5 and 100 mM NaOH/glycine for pH range 9 5-10 5 The pH ranges were chosen so that the points overlapped Seprase samples were pre-incubated in the above buffers for 10 min at 37°C prior to addition of the substrate solution, also prepared in the respective buffers

2 7 3 Determination of the Second Order Rate Constant, (k_2) for DFP Inhibition of Seprase

It has been shown that the irreversible and classic serine protease inactivator Diisopropylfluorophosphate (DFP) has a high specificity for the catalytic serine of tissue Seprase, thus catalytically classifying this enzyme as a possible serine protease The inhibition of Seprase by DFP results in an IC_{50} value of 100nM being obtained (Collins et al, 2004) It was decided that the second order rate constant should be determined for this inhibitor An enzyme concentration was chosen so that both the enzyme and DFP were in equimolar concentrations

A stock substrate of 200 µM Z-Gly-Pro-AMC in 8% v/v MeOH and containing 500 mM NaCl was prepared using 100 mM potassium phosphate, pH 7 4 An inhibitor stock of 10 µM DFP and containing 500 mM NaCl was prepared using 100 mM potassium phosphate, pH 7 4 Both stocks were pre-incubated at 37°C until completely dissolved and thermal equilibrium was reached Equal volumes of substrate stock and inhibitor stock were added together to give a final substrate concentration of 100 µM Z-Gly-Pro-AMC in 4% v/v MeOH containing 500 mM NaCl and 5 µM DFP Also, equal volumes of Seprase and inhibitor stock were added together to give a final DFP concentration of 5 µM Samples were assayed immediately after the addition of the inhibitor and labelled time zero (T_0) A 25 µl aliquot of enzyme inhibitor mixture was added to 225 µl of substrate inhibitor mixture and assayed in triplicate at 37°C for 1 hour Reactions were terminated by the addition of 50 µl of 1 5 M acetic acid Aliquots were removed from both inhibitor mixtures at 1 min intervals and assayed immediately A plot of 1/[AMC] Released ($\mu mole^{-1}$ L) versus Time (min) was constructed

2 7 4 Inhibition Studies using synthesised Dipeptidyl Phosphonate Esters

Serine protease inhibitors were provided by Dr Peter Kenny, School of Chemical Sciences, DCU These inhibitors were not designed specifically to inhibit Seprase They had been developed as part of a separate research project (Anderson, 2005) and they were tested for inhibitory properties using Seprase and Prolyl Oligopeptidase purified from bovine serum (see section 2 4 and 2 7 4 1 respectively) The partial purification procedure for bovine serum Prolyl Oligopeptidase was described previously by Collins and O'Connor (2003)

2 7 4 1 Partial Purification Procedure for Bovine Serum Prolyl Ohgopeptidase

2 7 4 1 1 Phenyl Sepharose Hydrophobic Interaction Chromatography I

A 20 ml Phenyl Sepharose CL-4B hydrophobic interaction column (2 5 x 7cm) was equilibrated with 100 ml of 100 mM potassium phosphate containing 200 mM ammonium sulphate, pH 7 4 Solid ammonium sulphate was dissolved in 20 ml of bovine serum to give a final concentration of 200 mM This sample was then applied to the equilibrated column followed by a 100 ml wash with 100 mM potassium phosphate containing 200 mM ammonium sulphate, pH 7 4 The column was then washed with 100 ml of 100 mM potassium phosphate containing 50 mM ammonium sulphate, pH 7 4 Bound protein was eluted with a distilled water wash Equilibration of the column was carried out at a flow rate of 1 ml/min, while all other steps were performed at 2 ml/min Five millilitre fractions were collected throughout the procedure and were assayed for PO activity according to section 2 3 4 Protein content in each fraction was determined using the Biuret assay according to section 2 1 1 Fractions containing PO activity were pooled to yield post-phenyl Sepharose fraction I Enzyme activity and protein content were quantified using the fluorimetric assay as outlined in sections 2 3 2 and 2 3 3 and section 2 1 1 respectively

2 7 4 1.2 Phenyl Sepharose Hydrophobic Interaction Chromatography II

The 20 ml resin used in section 2 7 4 1 1 was regenerated with a 100 ml wash of pure ethanol at 0 5 ml/min The column was then equilibrated with 100 ml of 100 mM potassium phosphate containing 1 M ammonium sulphate, pH 7 4 Solid ammonium sulphate was dissolved in the concentrated post-phenyl Sepharose I pool (pH was

maintained at 7 4 using 1 M NaOH), to give a final concentration of 1 M The sample was then applied to the column followed by a 100 ml wash of equilibration buffer Bound protein was eluted with a 100 ml linear gradient of 1-0 M ammonium sulphate, 100-0 mM potassium phosphate, pH 7 4 The column was then washed with 25 ml of ultra-pure water All steps were performed at a flow rate of 2 ml/min Five millilitre fractions were collected throughout the procedure, and again were assayed for PO activity according to section 2 3 4 Protein content in each fraction was determined using the standard BCA assay as outlined in section 2 1 2 Fractions containing PO activity were pooled to yield post phenyl sepharose fraction II Enzyme activity and protein content were then quantified as outlined in sections 2 3 2 and 2 1 2

2 7 4 1 3 Cibacron Blue 3GA Chromatography

The post phenyl sepharose II pool was concentrated and dialysed overnight against 2 L of 20 mM potassium phosphate One hundred millilitres of 20 mM potassium phosphate, pH 7 4 was used to equilibrate a 20 ml cibacron blue 3GA resin After sample application the column was washed with 100 ml of 20 mM potassium phosphate, pH 7 4 to remove any unbound protein Elution was performed using a 100 ml linear 0-2 M NaCl gradient in 20 mM potassium phosphate, pH 7 4 Loading, washing and elution were all performed at 1 ml/min Five millilitre fractions were collected and assayed for PO activity according to section 2 3 4 Protein content in each fraction was determined using coomassie plus protein reagent as in section 2 1 3 Fractions containing PO activity were pooled and enzyme activity and protein content were quantified as outlined in section 2 3 2 and 2 1 3

2 7 4 2 Inhibitor Analysis

The structures of the various inhibitors can be seen in Table 2 7 and Figure 2 7

Compound	R_1	R_2	n	M W *g/mol*	Compound	R_1	R_2	n	M W *g/mol*
8a	c	a	1	554 53	*9a*	b	a	1	526 33
8b	c	a	2	568 56	*9b*	b	a	2	540 55
8c	c	a	3	582 58	*9c*	b	a	3	554 58
8d	c	b	1	568 56	*9d*	b	b	1	540 55
8e	c	b	2	582 58	*9e*	b	b	2	554 58
8f	c	b	3	596 61	*9f*	b	b	3	568 61
8g	c	c	1	492 46	*9g*	b	c	1	464 46
8h	c	c	2	506 49	*9h*	b	c	2	478 49
8i	c	c	3	520 51	*9i*	b	c	3	492 51
10a	a	a	1	492 39	*10d*	a	b	2	520 44
10b	a	a	2	506 42	*10e*	a	c	1	430 32
10c	a	b	1	506 42	*10f*	a	c	2	444 35

Table 2 7 Structure of the Dipeptidyl Phosphonate Esters

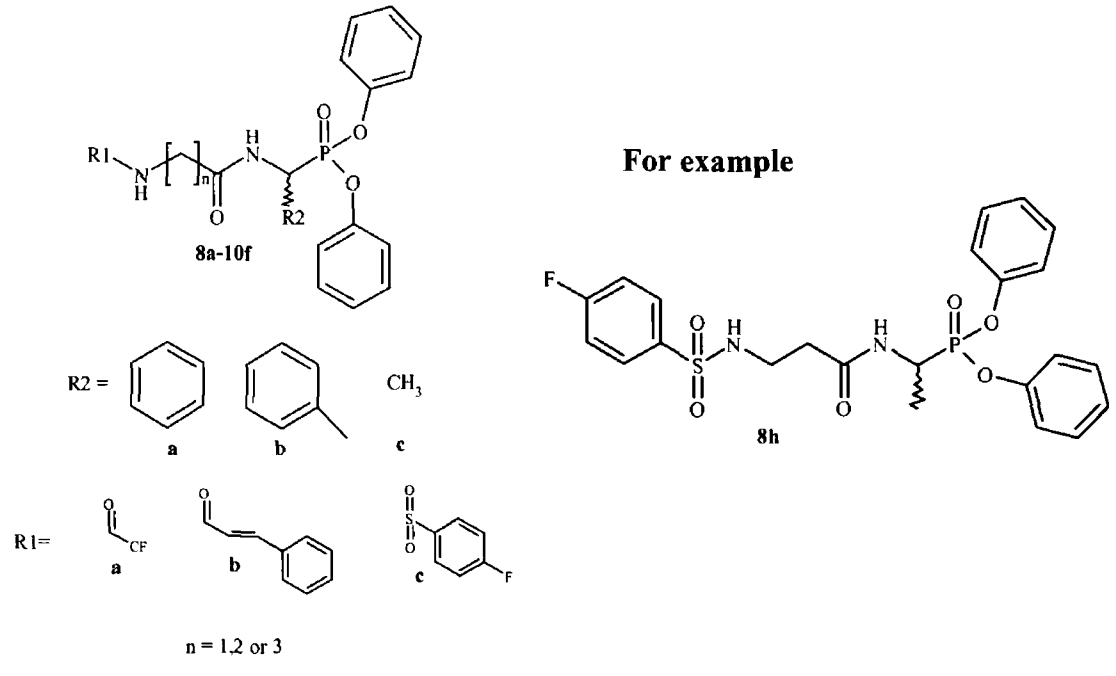


Figure 2 7 Chemical Structure of the Dipeptidyl Phosphonate esters

A range of inhibitor concentrations (0-50 μM) were prepared from a stock concentration of 5 mM in 5% v/v DMSO 100 μM Z-Gly-Pro-AMC in 4% v/v MeOH and containing 500 mM NaCl in 100 mM potassium phosphate was used as diluent The DMSO concentration was maintained at 4% to ensure the solubility of

the inhibitors during the assay Analysis showed that above 5% DMSO, the enzyme activity was inhibited Seprase and Prolyl Oligopeptidase from bovine serum were isolated and purified using a separation strategy described previously (see sections 2 4 and 2 7 4 1)

One hundred microlitres of the substrate inhibitor mixture was added to 25µl enzyme Seprase activity was assayed in triplicate as outlined in section 2 3 2 using these substrate inhibitor mixtures Suitable negative controls were prepared, excluding the inhibitor to determine the effect if any the DMSO and MeOH had on the enzyme activity The IC_{50} value of each inhibitor for Seprase/Prolyl Oligopeptidase was determined

2.8 Substrate Specificity Studies

2 8 1 Substrate Specificity using a Combinatorial Dipeptide Library

This work was carried out in collaboration with Dr Charles S Craik at the University of California, USA Dr Craik is involved in the study of substrate specificity, catalytic mechanism, structure and biological function of various proteases and their natural inhibitors The lab has developed a method for rapid and general profiling of protease specificity using combinatorial fluorogenic substrate libraries (Harris et al 2000)

The substrate specificity of Seprase was determined using this positional scanning synthetic combinatorial dipeptide substrate library The positional scanning library used in this study has the general structure Xaa-Xaa-ACC (where Xaa = any amino acid) Amide bond hydrolysis occurs between the Xaa-Xaa dipeptide and the fluorogenic 7-amino-4-methylcoumarin, resulting in an increase in fluorescence This library is composed of two sublibraries of 400 compounds each, utilizing the 20 naturally occurring amino acids, excluding cysteine and including norleucine (n) The P_1 sublibrary is composed of 20 separate mixtures in which the P_1 position is fixed, and the P_1 position contains equimolar mixture of all amino acids In the P_1 library, the P_1 residue is positionally defined, and the P_1 position contains an equimolar mixture of all other amino acids The results using these libraries provide a complete understanding of the specificities of dipeptidases in S_1 and S_2 subsites of the active site

2 8 1 1 Preparation of the Seprase sample

Seprase from bovine serum was isolated and purified using a separation strategy described previously (section 2 4) A sample of this peptidase was prepared for Dr Craiks laboratory, for functional characterization Prior to use, the Seprase sample was concentrated 80-fold to a protein concentration of 80 µg/ml in a Microcon YM-10 Centrifugal Filter Unit (Millipore) Proteolytic activity was measured by the continuous fluorimetric assay with 100 µM Z-Gly-Pro-AMC (Bachem), prepared in 100 mM potassium phosphate pH 7 4, containing 250 mM NaCl, 2% (v/v) DMSO Assays were carried out at 37°C in 96-well Microfluor 1 Black "U" bottomed plates

(Dynex Technologies) in a Molecular Devices SpectraMax Gemini microplate reader fitted with λ_{ex} 370 nm and λ_{em} 440 nm filters

2 8.1 2 Construction of a Dipeptide Substrate Library

A P_1-diverse library with bound ACC fluorophore was constructed as outlined by Harris and coworkers at the University of California, San Francisco (Figures 2 8 A-E) (Harris et al , 2000) To introduce the randomized P2 position, a mixture of Fmoc-amino acids (AnaSpec) [14 8 mmol per well] was pre-activated with diisopropylcarbodiimide (DICI) (390 µl, 2 5 mmol), and 1-hydroxybenzotriazole (HOBt) (340 mg, 2 5 mmol) in dimethylformamide (DMF) (10 ml) An aliquot (0 5 ml) of the pre-activated solution was added to each of the wells (Figures 2 8 F) The reaction block was agitated for 3 hours, filtered and washed with DMF (three times in 0 5 ml) The Fmoc of the P_2 amino acid was removed by filtration and the resin was washed with DMF (three times with 0 5 ml) and treated with 0 5 ml of a capping solution consisting of acetic anhydride (2 5 mmol), HOBt (2 5 mmol), and DICI (2 5 mmol) in DMF (10 ml) After agitation for 4 h at room temperature, the resin was washed with DMF (three times with 0 5 ml) and CH_2Cl_2 (three times with 0 5 ml), and treated with a solution of 95 2 5 2 5 TFA/TIS/H_2O (Figures 2 8 F-G) After incubation for a further 1 h period (room temperature), the reaction block was opened and placed on a 96-deep-well microplate and the wells were washed with additional TFA/TIS/H_2O solution (twice with 0 5 ml)

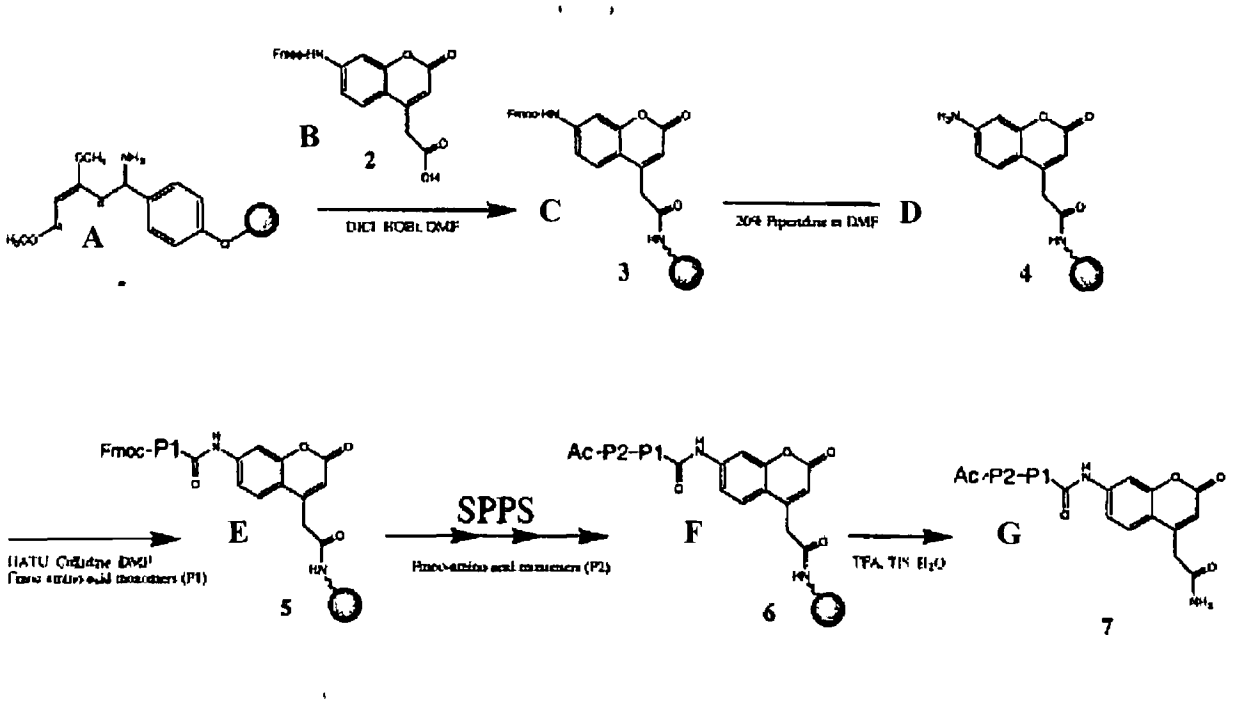


Figure 2 8 **Diagram illustrating 7-amino-4-carbamoylmethylcoumarin (ACC)-labelled dipeptide library construction using standard Fmoc protocols where SPPS corresponds to solid-phase peptide synthesis**

2 8 1 3 Dipeptide Substrate Library Screening

Each dipeptide compound in this library was prepared as a 50 μM stock solution in DMSO Seprase activity was analysed using 0 5 μM of each compound under the assay conditions described, with fluorescence determined using λ_{ex} 350 nm and λ_{em} 450 nm All assays were performed in triplicate and the formation of free ACC was calculated using the conversion (1 RFU sec^{-1} = 0 0007226 pM sec^{-1}) as determined by Harris and colleagues (Harris et al , 2000)

2 8 2 Kinetic Analysis

Substrate specificity studies on bovine serum Seprase, based on kinetic analysis were performed

2 8 2 1 K_m Determination For Z-Gly-Pro-AMC

A 300 µM stock solution of Z-Gly-Pro-AMC in 5% v/v MeOH containing 500 mM NaCl was prepared A range of substrate concentrations (0-300 µM) was prepared from this stock using 100 mM potassium phosphate, pH 7 4, containing 5% v/v MeOH and 500 mM NaCl as diluent Purified Seprase was assayed with each concentration in triplicate as outlined in section 2 3 2 The K_m of Seprase for the substrate Z-Gly-Pro-AMC was estimated when the data obtained was applied to various kinetic models (see Appendix A)

2 8 2 2 K_i^{app} Determination Using Selected Synthetic Peptides

The effect of a variety of selected synthetic peptides on the kinetic interaction between Seprase and the substrate Z-Gly-Pro-AMC was determined Substrate concentrations in the range 100-300 µM Z-Gly-Pro-AMC in 60% v/v MeOH containing 500 mM NaCl was prepared in the presence of 200 µM peptide, each in a final volume of 2 ml The assay MeOH concentration was maintained at 5% The peptides studied, and their preparation, are outlined in Table 2 8 Seprase activity was assayed in triplicate using these substrate mixtures as outlined in section 2 3 2 The data obtained was applied to the various kinetic models (see Appendix A), where the inhibition constant (K_i^{app}) and the type of inhibition observed were determined as outlined in Appendix A

Peptide	Stock Conc (mM)	Solubility	Assay Conc (μM)
Z-Gly-Met-Phe	5	50% MeOH*	200
Z-Gly-Pro-Phe	5	50% MeOH*	200
Z-Ala-Pro-Phe	5	50% MeOH*	200
Z-Ala-Met-Phe	5	50% MeOH*	200
Z-Ala-Nle-Phe	5	50% MeOH*	200
Z-Gly-Met-Phe-His	5	41% MeOH*	200
Z-Gly-Nle-Phe-His	5	10% MeOH*	200
Z-Ala-Pro-Phe-His	5	10% MeOH*	200
Z-Ala-Nle-Phe-His	5	50% MeOH*	200
Z-His-Pro-Phe-His	5	Ultrapure Water	200
Ala-Ser-Gly-Pro-Ser-Ser	5	10% MeOH*	200
Ala-Ser-Nle-Pro-Ser-Ser	5	10% MeOH*	200

Table 2 8 Peptide Preparation for K_i^{app} Determinations

* Sonication using an ultrasonic water bath was required for complete dissolution

2.9 Tissue Localisation Studies

Bovine tissues and mammalian cell lines in culture were obtained in order to determine the presence of Seprase and to create an expression profile for both Seprase and Prolyl Oligopeptidase

2 9 1 Tissue preparation

Tissue samples were obtained from freshly slaughtered cows (Kepak Meats, Clonee, Co Meath) and transported to the laboratory on ice Each organ was divided into approximately 7 g portions, which were stored at $-17^{o}C$ if not being used immediately A 7 g piece of each tissue was suspended in 30 ml 100 mM potassium phosphate buffer, pH7 4 These were homogenised at full speed for 15 s using a Sorvall Omni Mixer, followed by 15 s intermission to minimise heating and foaming effects Homogenisation of the samples was continued until the tissue was visually disrupted The homogenate (H) was then collected and centrifuged at 15,000 rpm (23,667 x g) for 45 min at $4^{o}C$ in a Beckman Coulter Allegra 21R refrigerated centrifuge fitted with a F0850 rotor The supernatant was decanted and retained (S_1) The pellet was resuspended in 30 ml 100 mM potassium phosphate buffer, pH7 4, and homogenised The second homogenate was then collected and centrifuged as previous The supernatant (S_2) was again decanted The pellet (P_2) was resuspended in 30ml 100 mM potassium phosphate buffer, pH 7 4 The pellet and 2 supernatants obtained from each tissue were assayed for Seprase and Prolyl Oligopeptidase activity as described in section 2 3 3 The total Z-Gly-Pro-AMC activity present was obtained in the absence of the potent inhibitor and PO activity was taken to be the residual non-Seprase activity The concentration of protein in each sample was measured using the Biuret assay (section 2 1 1) and specific activities of Seprase and PO were calculated for each tissue

	Tissue Sample
1	Liver
2	Large Intestine
3	Small Intestine
4	Heart
5	Lungs
6	Brain
7	Serum
8	Kidney
9	Spleen

Table 2 9 1 Bovine Tissue Samples

2 9 2 Preparation of cleared lysate from cultured mammalian cells

A cell pellet containing 6 x 10^6 cells obtained from the cell culture lab was washed in 100 mM potassium phosphate buffer, pH 7 4 followed by centrifugation at 2,000 rpm (421 x g) for 5 min (using a Beckman Coulter Allegra 21R refrigerated centrifuge fitted with a F0850 rotor) The supernatant was discarded and the cells were re-suspended in 2 ml 100 mM potassium phosphate buffer, pH 7 4, containing 500 mM NaCl The cells were disrupted on ice with a 3 mm micro-tip somcator (Sonics & Materials Inc) using 2 5 s, 40 kHz pulses for 15 s The cell debris was removed by centrifugation at 4,000 rpm for 15 min at 4°C The cleared lysate was transferred to a fresh universal and stored at 4°C The cleared lysate was tested for total protein, Seprase and Prolyl Oligopeptidase activity as described in section 2 1 2 and 2 3 3 respectively

Mammalian Cell Line	Source	Description
MDA-MB-435 SF	Mammary Gland, breast	Human breast cancer cell line derived from a female with ductal adenocarcinoma
Hs578T	Mammary Gland, breast	Human breast cancer cell line derived from a carcinoma
A549	Lung	Human lung carcinoma cell line derived from carcinomatous tissue
HeLa (Chang Liver)	Liver, HeLa Contamination	Originally thought to be derived from normal liver tissue, but it has now been found to be established from HeLa contamination
SK-N-FI	Brain, Neuroblast	Neuroblastoma cell line derived from a bone marrow metastasis
DG75	Abdominal Lymphoma	Lymphoid B cell line derived from an Israeli Burkitt-like lymphoma case (Ben-Bassat *et al*, 1977)
L428	Lymphatic tissue	Hodgkin's Lymphoma B-cell line, EBV negative
L591	Lymphatic tissue	Hodgkin's Lymphoma B-cell line, EBV positive
BAEC	Heart	Bovine aorta endothelial cell line
BASMC	Heart	Bovine aortic SMC cell line
SW620	Colon	Human colorectal adenocarcinoma cell line established from the lymph node metastasis of a male
SW480	Colon	Human colorectal adenocarcinoma cell line established from the lymph node metastasis of a male
HT29	Colon	Human colorectal adenocarcinoma cell line established from the primary tumour of a female patient

Table 2 9 2 Mammalian Cell Lines

2 10 Molecular Cloning of Seprase

2 10 1 Bacterial strains, primers and plasmids

The bacterial strains, primers and plasmids used in this work are listed in Tables 2 10 1, 2 10 2, 2 10 3 and 2 10 4 respectively

Strain	Genotype	Features/Uses	Source
Escherichia coli			
DH5α	F *endA*1 *recA*1 *relA*1 *gyrA*96 *sup*E44 *thi*-1 *hsdR*17(r_K, m_{K+}) Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deo*R *pho*A λ	High transformation efficiency	Bethesda Research Labs
XL10-Gold	Δ(*mcr*A)183 Δ(*mcr*CB-*hsd*SMR-mrr)173 *end*A1 *rec*A1 *rel*A1 *gyr*A96 *sup*E44 *thi*-1 *lac* Hte [F' *pro*AB *lacI*^{q}ZΔM15 Tn10(*tet*^{R}) Amy (*cam*^{R})]	High transformation efficiency Antibiotic resistance Expression host	Stratagene
BL21 (DE3)	F⁻ *dcm ompT hsd*S_B(r_B, m_B) *gal* λ(DE3)	Protease deficient Expression host	Novagen

Table 2.10 1 Bacterial strains

Name	Sequence (5' – 3')	T_m (°C)
Cloning/Analysis		
BamHI-Fap1	**AAAAAGGATCC**CCACGCTCTGAAGACAGAATT	57 9
XhoI-Fap6	**AAAAAACTCGAG**TCAGATTCTGATACAGGCT	52 3
Fap3	CCAGCAATGATAGCCTCAA	54 5
Fap4	ACAGACCTTACACTCTGAC	54 5
Fap7_for	ATGAAGACTTGGGTAAAAATCG	54 6
BamH1-A-Fap7_for	**AAAAAGGATCCA**ATGAAGACTTGGGTAAAAATCG	54 6
BamH1Fap7_for	**AAAAAGGATCC**ATGAAGACTTGGGTAAAAATCG	54 6
Fap8_rev	TTAGTCTGACAAAGAGAAACACTGC	59 7
Xho-1-Fap8_rev	**AAAAAACTCGAG**TTAGTCTGACAAAGAGAAACACTGC	59 7
Cla1-Fap9_rev	ACGCAGGGTAAGTGGTATCG	59 4
Cla1-Fap10_for	GCTAAGAATCCCGTTGTTCG	57 3
Nco-1-Fap7_for	**AAAACCATGGGG**ATGAAGACTTGGGTAAAAATCG	54 6
Bgl-II-Fap8_rev	**AAAAAAGATCT**TTAGTCTGACAAAGAGAAACACTGC	59 7
BamH1-Koz1-Fap7_for	**AAAAAGGATCC*GCCATG*GGG**ATGAAGACTTGGGTAAAAATCG	54 6
BamH1-Koz2-Fap7_for	**AAAAAGGATCC*GCCATG***AAGACTTGGGTAAAAATCG	54 6
BamH1-Koz3_for	**AAAAAGGATCC*GCCACCATGGGG***AAGACTTGGGTAAAAATCG	50 2
Not-1-Fap8_1rev	**AAAAAAGCGGCCGC**TTAGTCTGACAAAGAGAAACACTGC	59 7
Not-1-Fap8_2rev	**AAAAAAGCGGCCGC**AGTCTGACAAAGAGAAACACTGC	58 4
	DNA clamps are in bold type Restriction sites are underlined Kozak / Ribosome Binding Site sequences are in italics T_m takes into account only those bases that bind $T_m = [69\ 3 + 0\ 41(\%GC)] - 650/length$	
β-actin For	GAAATCGTGCGTGACATTAAGGAGAAGCT	65 3
β-actin Rev	TCAGGAGGAGCAATGATCTTGA	58 4

Table 2 10 2 Primers (obtained from MWG-Biotech AG)

Name	Sequence (5' – 3')	Source
Sequencing		
M13rev(-29)	CAGGAAACAGCTATGACC	(pCR2 1 forward primer)
M13uni(-21)	TGTAAAACGACGGCCAGT	(pCR2 1 reverse primer)
M13rev(-49)	GAGCGGATAACAATTTCACACAGG	(pCR2 1 forward primer)
pQEPfor	CCCGAAAAGTGCCACCTG	(pQE-30Xa forward primer)
pQEPrev	GGTCATTACTGGAGTCTTG	(pQE-30Xa reverse primer)
pPOBfor1	CCCAGACTGGGTTTATGAAGAGG	Internal sequence forward primer
pPOBfor2	CCTCAATTTGACAGATCAAAGAAGTATCCC	Internal sequence forward primer
pPOBrev1	GGGATACTTCTTTGATCTGTCAAATTGAGG	Internal sequence reverse primer
pPOBrev2	CCTCTTCATAAACCCAGTCTGGG	Internal sequence forward primer
CMVfor	CGCAAATGGGCGGTAGGCGTG	(pIRES forward primer)
T7	TAATACGACTCACTATAGGG	(pIRES forward primer)
SP6	CATTTAGGTGACACTATAG	(pIRES reverse primer)
pIRES-for	ATGGGCGGTAGGCGTGTA	(pIRES forward primer)
pIRES-rev	ATGCAGTCGTCGAGGAATTG	(pIRES reverse primer)

Table 2 10.3 Sequencing Primers (obtained from MWG-Biotech AG)

Plasmid	Description	Source
Vectors		
pCR2 1	TA cloning vector P*lac*, *amp*^{R}, *kan*^{R}, *lacZ*α, ColE1 origin Figure 2 10 1	Invitrogen
pQE-30-Xa	Expression vector T5 promoter/*lac* operon, *amp*^{R}, 6xHis sequence at 5' end of MCS, contains a Factor Xa Protease recognition site, ColE1 origin Figure 2 10 2	Qiagen
pBR322	Cloning vector *amp*^{R}, *tet*^{R}	Roche
pcDNA3	Mammalian Expression vector CMV promoter, *amp*^{R}, *neo*^{R}, ColE1 origin, f1 ori Figure 2 10 3	Gifted by Dr Walls, D C U
pcDNA-3-HA	pcDNA expression vector with a HA epitope at the 5' end of MCS, cloned *Hind* III *Bam*HI Figure 2 10 3	D'Souza 2000 / Gifted by Dr Walls, D C U
pIRES-hrGFP II	Mammalian Expression vector CMV promoter, *neo*^{R}, *kan*^{R}, internal ribosome entry site, hrGFP ORF, FLAG epitope Figure 2 10 4	Stratagene
Constructs		
pPOB12	pcDNA3-HA containing Seprase ORF	This Work
pPOB15	pcDNA3 containing Seprase ORF, subcloned from pPOB12, for expression Contains Eukaryotic Kozak sequence 2	This Work
pPOB16	pcDNA3-HA containing Seprase ORF, subcloned from pPOB12, for expression	This Work
pPOB17	pcDNA3 containing Seprase ORF, subcloned from pPOB12, for expression Contains Eukaryotic Kozak sequence 2	This Work
pPOB18	pcDNA3 containing Seprase ORF, subcloned from pPOB12, for expression Contains Eukaryotic Kozak sequence 1	This Work
pQPOB5	pQE-30Xa containing Seprase ORF with 5' 6xHis sequence fusion, for expression and subsequent purification	This Work
pIRPOB1	pIRES-hrGFP II containing Seprase ORF, for expression	This Work
pIRPOB2	pIRES-hrGFP II containing Seprase ORF with FLAG sequence fusion, for expression and subsequent purification	This Work

Table 2 10 4 Plasmids

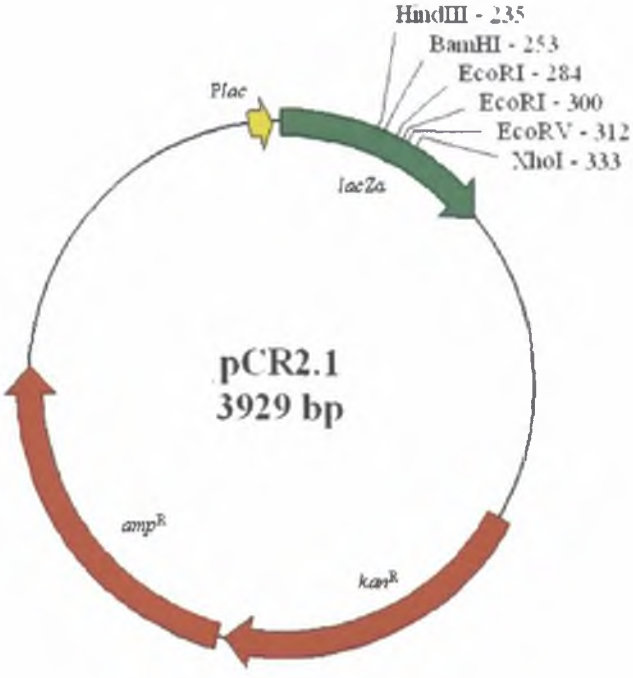


Figure 2.10.1 pCR2.1 vector
The 3926bp TA cloning vector pCR2.1 (Table 2.10.3). The multiple cloning site (MCS) enzymes are indicated. The TA cloning site is situated within the *LacZα* ORF (green), which is under the control of the P_{lac} promoter (yellow). Ampicillin and Kanamycin resistance genes (amp^R & kan^R) are shown in red. Generated using pDRAW32 (Section 2.10.10).

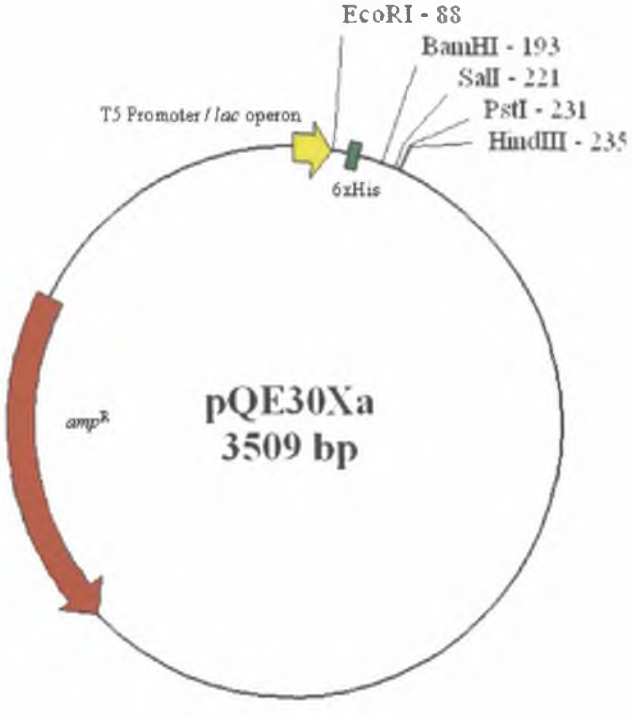


Figure 2.10.2 pQE-30Xa vector
The 3509bp cloning/prokaryotic expression vector pQE-30Xa (Table 2.10.3). The multiple cloning site (MCS) enzymes are indicated. The 6xHis coding sequence (green) is situated at the 5' end of the MCS, which in turn is situated downstream of the T5 promoter/*lac* operon (yellow). Ampicillin resistance gene (amp^R) is shown in red. Generated using pDRAW32 (Section 2.10.10).

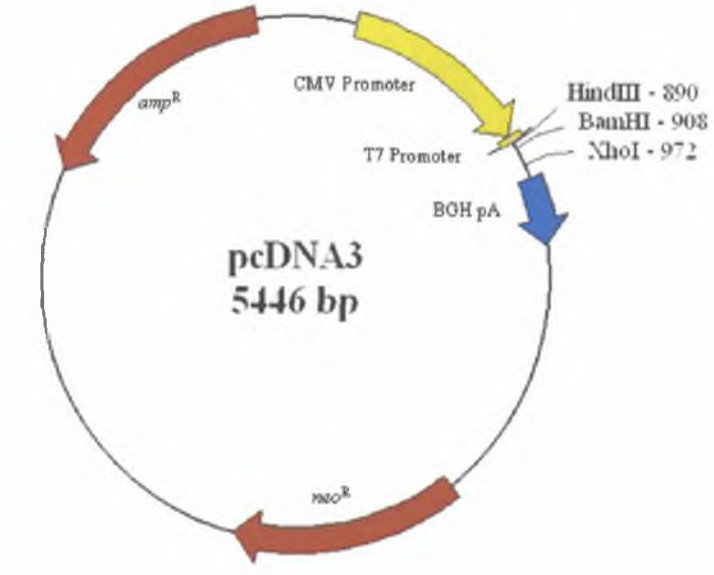


Figure 2.10.3 pcDNA3 and pcDNA3-HA vectors
The 5446bp cloning/mammalian expression vectors pcDNA3 / pcDNA3-HA (Table 2.10.3). The multiple cloning site (MCS) enzymes are indicated. The T7 promoter (yellow) is downstream of the CMV promoter (yellow). Ampicillin and neomycin resistance genes (*amp*^{R} and *neo*^{R}) are shown in red. HA epitope of the vector pcDNA3-HA was cloned in *Hind*III and *Bam*HI (D'Souza, 2000). Generated using pDRAW32 (Section 2.10.10).

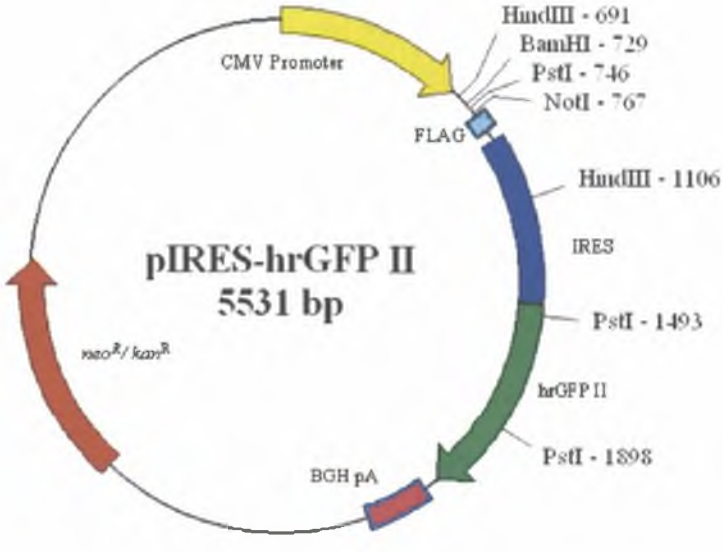


Figure 2.10.4 pIRES-hrGFP II vector
The 5531bp cloning/mammalian expression vector pIRES-hrGFP II (Table 2.10.3). The multiple cloning site (MCS) enzymes are indicated. The CMV promoter (yellow) is upstream of the MCS and the 3x FLAG epitope (cyan). Kanamycin and neomycin resistance genes (*kan*^{R} and *neo*^{R}) are shown in red. The internal ribosome entry site (IRES – blue) is upstream of the hrGFP II ORF (green). Generated using pDRAW32 (Section 2.10.10).

2 10 2 Media, solutions and buffers

All chemicals and reagents were obtained from Sigma-Aldrich unless otherwise stated All chemicals were analytical grade Microbiological media are obtained from Oxoid Sterilisation was achieved by autoclaving at 121°C and 15 lb/in^2 for 20 min, unless otherwise stated

Luria Bertani broth (LB)

Tryptone	10 g/L
NaCl	10 g/L
Yeast Extract	5 g/L

Adjusted to pH 7 0 with NaOH Sterilised by autoclaving For solid LB, 15 g/L Technical Agar No 3 (Oxoid) was included

TE buffer

Tris-HCl	10 mM
Na_2-EDTA	1 mM
pH	8 0

TAE buffer (50X)

Tris	242 g/L
Glacial Acetic Acid	57 1 ml/L
EDTA	100 ml/L (of 0 5 M stock EDTA, pH 8 0)
pH	8 0

Solution 1 of 1-2-3 method

Glucose	50 mM
Na_2-EDTA	10 mM (from 0 5 M stock EDTA, pH 8 0)
Tris-HCl	25 mM (from 1 M stock Tris-HCl, pH 8 0)

Solution 2 of 1-2-3 method

NaOH	200 mM
SDS	1% (w/v)

Solution 3 of 1-2-3 method

Potassium acetate	3 M
pH	4 8

To 60ml of 5M potassium acetate, 11 5ml of glacial acetic acid and 28 5ml of dH_2O was added The resulting solution was 3M with respect to potassium and 5M with respect to acetate

RF1 buffer

RbCl	100 mM
$CaCl_2$	10 mM
Potassium acetate	30 mM
Glycerol	15% (v/v)
pH (with HCl)	5 8

After the pH had been adjusted $MnCl_2$ was added to 50 mM The solution was filter sterilised through a 0 22 μm membrane and stored at 4°C in the dark ($MnCl_2$ is light sensitive)

RF2 buffer

RbCl	10 mM
MOPS	10 mM
$CaCl_2$	75 mM
Glycerol	15% (v/v)
pH	6 8

The solution was filter sterilised through a 0 22 μm membrane and stored at 4°C

Transfer buffer

Tris-HCl	25 mM
Glycine	150 mM
Methanol	10% (v/v)

TBS buffer

Tris-HCl	10 mM
NaCl	150 mM
pH	7 5

TBS-Tween buffer

Tris-HCl	20 mM
NaCl	500 mM
Tween 20	0 01% (v/v)
pH	7 5

Gel loading dye (6X)

Bromophenol Blue	0 25%
Xylene Cyanol	0 25%
Ficoll (Type 400)	15%

Bromophenol Blue and/or Xylene Cyanol were used as appropriate On a 1% agarose gel, bromophenol blue and xylene cyanol migrate approximately with the 300 bp and 4000 bp fragments respectively

Ethidium bromide stain

A 10 mg/ml stock solution in dH_2O was stored at 4°C in the dark For the staining of agarose gels, 100 µl of the stock solution was mixed into 1 L of dH_2O The staining solution was kept in a plastic tray and covered to protect against light Spent staining solution was collected and filtered through a deactivating filter and then treated as aqueous waste The filter was destroyed by incineration

Antibiotics

- Ampicillin was prepared in dH_2O at a concentration of 100 mg/ml and stored at –20°C The working concentration for *E coli* was 100 µg/ml
- Tetracycline was prepared in 50% v/v ethanol at a concentration of 10 mg/ml and stored at –20°C The working concentration for *E coli* was 10 µg/ml
- Kanamycin was prepared in dH_2O at a concentration of 100 mg/ml and stored at –20°C The working concentration for *E coli* was 50 µg/ml

2 10 3 Isolation and purification of DNA and RNA

2.10 3.1 Isolation of RNA

RNA was isolated using Trizol Reagent (Invitrogen), a mono-phasic solution of phenol and guanidine isothiocyanate This is a modification of the method developed by Chomczynski and Sacchi (1987) 100 mg of tissue was homogenised in 1 ml of Trizol reagent using a glass-Teflon homogeniser (treated with RNase AWAY, Molecular *Bio*-Products, *inc*) and transferred to a microfuge tube Alternatively, a pellet of cultured cells was lysed in 1 ml of Trizol reagent by repeated pipetting The sample was incubated at room temperature for 5 min to allow complete dissociation of nucleoprotein complexes This was followed by the addition of 200 µl of chloroform, mixed briefly by vortexing and incubated at room temperature for 3 min The phases were separated by centrifugation at 13,000 rpm (≥12,000 x g) for 15 min The upper aqueous layer was transferred to a fresh microfuge tube Then 500µl of isopropanol was added and mixed by inversion The sample was incubated at room temperature for 10 min to precipitate the RNA and then centrifuged at 13,000 rpm (≥12,000 x g) for 10 min The supernatant was discarded and the RNA pellet was washed with 1ml of 70% v/v ethanol The RNA pellet was air-dried and dissolved in 30 µl RNase-free water Any possible DNA contamination was eliminated by treatment with Deoxyribonuclease I (Section 2 10 7)

2 10 3 2 Isolation of plasmid DNA

Three procedures for the isolation of plasmid DNA were variably employed The 1-2-3 Method (Section 2 10 3 1 1) was used for convenient plasmid isolation from large numbers of samples, mostly for the purpose of screening The GenElute Plasmid Miniprep Kit (Sigma, Section 2 10 3 2 2) was used to prepare consistently pure and supercoiled plasmid DNA, mostly for the purpose of DNA sequencing The Qiagen Plasmid Mini Kit (2 10 3 2 3) was used to prepare ultrapure supercoiled transfection grade plasmid DNA with high yields

2 10.3 2 1 1-2-3 method

This method is adapted from the procedure described by Birnhoim and Doly (1979) Briefly, 1 5 ml of a bacterial culture in a microfuge tube was centrifuged at 13,000 rpm for 5 min to collect the cells The supernatant was discarded and the cell pellet re-suspended in 200 µl of Solution 1 Alternatively, bacterial growth was taken off an LB agar culture plate with a sterile loop and re-suspended in 200 µl of Solution 1 The re-suspension was left for 5 min at room temperature Then 200 µl of Solution 2 was added, the tube was mixed by inversion and placed on ice for 5 min Following the addition of 200 µl of Solution 3, the tube was mixed by inversion and placed on ice for 10 min A clot of chromosomal DNA was collected by centrifugation at 13,000 rpm for 10 min The supernatant was removed to a new microfuge tube with 400 µl of phenol chloroform isoamylalcohol (25 24 1) and mixed by brief vortexing Upon centrifugation at 13,000 rpm (≥12,000 x g) for 5 min the mixture is divided into an upper aqueous and lower organic layer The aqueous layer was removed to a new microfuge tube with an equal volume of isopropanol and mixed by inversion The tube was left at room temperature for 5 min and then centrifuged at 13,000 rpm (≥12,000 x g) for 20 min to pellet the plasmid DNA The pellet was washed with 70% v/v ethanol and then dried briefly in a SpeedVac (Savant) vacuum centrifuge The plasmid DNA was resuspended in 50 µl of TE buffer and 1 µl of Ribonuclease A (Section 2 10 7) was added to digest co-purified RNA Plasmid DNA was stored at 4°C

2 10 3 2 2 GenElute plasmid Mmiprep kit

The kit was used according to the manufacturer's instructions Briefly, 1 5 ml of a bacterial culture in a microfuge tube was centrifuged at 13,000 rpm (≥12,000 x g) for 5 min to collect the cells The supernatant was discarded and the cell pellet was completely re-suspended in 200 µl of re-suspension solution Alternatively, bacterial growth was taken off an LB agar culture plate with a sterile loop and re-suspended in 200 µl of Solution 1 Then 200 µl of lysis solution was added and mixed by inversion to lyse the cells Following this, 350 µl of neutralisation/binding buffer was added and mixed by inversion to precipitate cell debris, lipids, proteins and chromosomal DNA The precipitate was collected by centrifugation at 13,000 rpm (≥12,000 x g)

for 10 min The supernatant was transferred to a spin column in a microfuge tube and centrifuged at 13,000 rpm (≥12,000 x g) for 1 min to bind the plasmid DNA The flow through was discarded and 750 µl of washing solution was added followed by further centrifugation at 13,000 rpm (≥12,000 x g) for 1 min The flow through was discarded and the column was centrifuged at 13,000 rpm (≥12,000 x g) for 1 min to dry the spin column The spin column was transferred to a fresh microfuge tube and 50 µl TE buffer was added The DNA was eluted by centrifugation at 13,000 rpm (≥ 12,000 x g) for 1 min

2 10 3 2 3 Qiagen plasmid Mini kit

The kit was used according to the manufacturer's instructions Briefly, 3 ml of a bacterial culture in a microfuge tube was centrifuged at 13,000 rpm (≥12,000 x g) for 5 min to collect the cells The supernatant was discarded and the cell pellet was completely re-suspended in 300 µl of Buffer P1 Then 300 µl of Buffer 2 was added and mixed by inversion to lyse the cells Following this 300 µl Buffer P3 was added, mixed by inversion and incubated on ice for 5 min to precipitate cell debris, lipids, proteins and chromosomal DNA The precipitate was collected by centrifugation at 13,000 rpm (≥12,000 x g) for 10 min The supernatant was transferred to a Qiagen-tip and allowed to empty by gravity flow The DNA bound to the resin and therefore the flow through was discarded 2 x 2 ml of Buffer QC was applied to the column to wash the column 800µl Buffer QF was added to the column to elute the DNA, which was collected in a microfuge tube Then 560 µl of isopropanol was added The tube was left at room temperature for 5 min and then centrifuged at 13,000 rpm (≥ 12,000 x g) for 30 min to pellet the plasmid DNA The pellet was washed with 1 ml 70% v/v ethanol and centrifuged at 10,000 rpm (≥12,000 x g) for 10min The supernatant was removed and the pellet was then dried briefly in a SpeedVac (Savant) vacuum centrifuge The plasmid DNA was resuspended in 30 µl of TE buffer Plasmid DNA was stored at 4°C

2 10 3 3 Isolation of DNA from agarose gels

2 10 3 3 1 Glass Wool Gel Extraction

The DNA band to be isolated was excised from the agarose gel using a scalpel The gel piece was placed into a PCR tube which had been pierced at the end and contained a small piece of glass wool The PCR tube was then placed in a microfuge tube and centrifuged at 13,000rpm (≥12,000 x g) for 2 min The solution containing DNA was then purified according to section 2 10 3 4

2 12 3 3 2 PerfectPrep kit

The kit was used according to the manufacturer's instructions The desired DNA band was excised from the agarose gel using a scalpel The gel slice was weighed and placed in a microfuge tube The tube was incubated at 65°C for 10 min to completely dissolve the agarose in the chaotropic solution One gel slice volume of isopropanol was added and mixed by inversion The solution was transferred to a spin column in a microfuge tube and centrifuged at 13,000 rpm (≥12,000 x g) for 1 min to bind the plasmid DNA The flow through was discarded and 750 µl of washing solution was added followed by further centrifugation at 13,000 rpm (≥ 12,000 x g) for 1 min The flow through was discarded and the column was centrifuged at 13,000 rpm (≥12,000 x g) for 1 min to dry the spin column The spin column was transferred to a fresh microfuge tube and 30 µl TE buffer was added The DNA was eluted by centrifugation at 13,000 rpm (≥12,000 x g) for 1 min

2 10 3 4 Purification and concentration of DNA samples

The sample containing the DNA to be precipitated was brought to 500 µl with dH_2O Then 400 µl of phenol chloroform isoamylalcohol (25 24 1) was added and mixed by brief vortexing Upon centrifugation at 13,000 rpm (≥12,000 x g) for 5 min the mixture was separated into an upper aqueous and lower organic phase The aqueous phase was removed to a new microfuge tube and an equal volume of chloroform was added and mixed by brief vortexing The tube was centrifuged at 13,000 rpm (≥ 12,000 x g) for 5 min and the aqueous phase was transferred to a new microfuge tube A 1/10 volume of 3 M sodium acetate was added followed by an equal volume of isopropanol and mixed by inversion The tube was left at room temperature for 60

min and then centrifuged at 13,000 rpm (≥12,000 x g) for 20 min to pellet the DNA The pellet was washed with 70% v/v ethanol and then dried briefly in a SpeedVac (Savant) vacuum centrifuge The DNA was resuspended in 20-50 µl of TE buffer

2 10 3 5 Quantification of DNA and RNA

Nucleic Acid concentration was quantified by UV spectrophotometry (Maniatis, et al, 1982) A dilution of the sample (typically 1/50) in dH_2O was measured for absorbance at 260 nm A reading of 1 0 corresponds to 50 µg/ml of DNA or 40 µg/ml of RNA

2 10 4 Agarose gel electrophoresis

DNA was analysed by electrophoresis through agarose gels in a BioRad horizontal gel apparatus according to Maniatis et al (1982) Agarose was added to TAE buffer to the required concentration (typically 0 7-2%) and dissolved by boiling The agarose solution was poured into plastic trays and allowed to set with a plastic comb fitted to create sample wells TAE buffer was used as the running buffer Loading dye was mixed with the DNA samples to facilitate loading and to give a visual indication of migration distance during electrophoresis When RNA samples were being analysed, the 1X TAE solution was prepared using 0 1% (v/v) DEPC Gels were run at 140 volts for 20-40 min depending on size of gel Gels were stained for 15 min by immersion in an ethidium bromide staining solution Gels were visualised using a UV transilluminator coupled with an image analyser to capture the image to a PC On every gel, 0 5 µg of 1 Kb Plus or 1Kb DNA Ladder (Invitrogen, Figure 2 10 5) was run as a molecular size marker

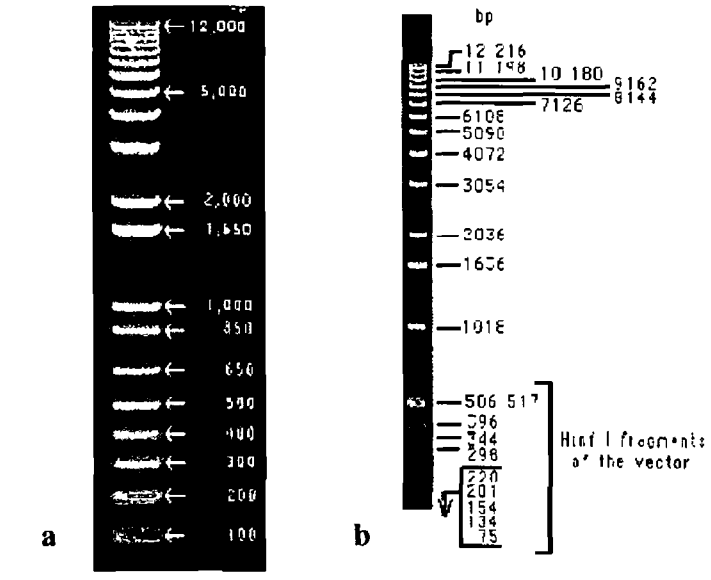


Figure 2 10 5 DNA Ladders
DNA Ladders used (a) 1Kb Plus (b) 1Kb DNA ladders

2 10 5 Competent cells

2 10 5 1 Rubidium chloride method

This is an adaptation of the method outlined by (Hanahan, 1985) Sterile conditions were used throughout A 10 ml sterile universal of LB broth containing the relevant antibiotics was inoculated with a single colony of the desired bacterial strain from a plate stock and cultured overnight at 37°C A 1 L flask with 200 ml of LB broth was inoculated with 2 ml of the overnight culture and incubated at 37°C with shaking at 225 rpm When the culture had reached an OD_{600} of ~0 5 (early-mid exponential phase) the flask was cooled in ice water All subsequent transactions took place at 4°C The culture was transferred to a sterile centrifuge bottle The cells were collected by centrifugation at 3,000 rpm (1380 x g) for 5 min (using a Beckman JA-14 rotor) The supernatant was decanted and the cells gently re-suspended in 60ml of chilled RF1 buffer The suspension was left on ice for 90 min The cells were again collected by centrifugation at 3,000 rpm for 5 min The supernatant was decanted and the cells gently re-suspended in 8ml of chilled RF2 buffer Aliquots of 800 µl were prepared in sterile 1 5 ml microfuge tubes and flash frozen using –80°C ethanol The competent cells were stored at –80°C Cells were routinely used within a few weeks

2 10 5 2 Transformation of competent cells

An aliquot of competent cells was thawed on ice A 200 µl portion of the cell suspension was mixed gently with 1-50 µl of plasmid DNA in a sterile 1 5 ml microfuge tube The mixture was left on ice for 30 min The cells were heat-shocked at 42°C for 30 seconds and placed back on ice for 2 min Then, 800 µl of LB broth was added to the cells followed by incubation at 37°C for 60 min Following this, 100 µl of the transformation suspension was spread on an LB agar plate containing the relevant antibiotics and incubated at 37°C overnight

2 10 5 3 Determining competent cell efficiency

Competent cell efficiency is defined in terms of the number of colony forming units obtained per µg of transformed plasmid DNA A 250 ng/µl stock of pBR322 plasmid DNA was diluted to 250 pg/µl, 25 pg/µl and 2 5 pg/µl Then 1µl of each dilution was transformed as described above The cell efficiency was calculated from the number of colonies obtained, taking into account the dilution factor and the fraction of culture transferred to the spread plate

$$\frac{No\ of\ Transformants}{\mu g\ of\ DNA} x \frac{Final\ Volume\ at\ recovery\ (ml)}{Volume\ Plated\ (ml)} = No\ Transformants\ per\ \mu g\ DNA$$

2 10 6 Bacterial storage

Bacterial strains were stored as 40% v/v glycerol stocks Mix 750 µl of an overnight culture with 750 µl sterile 80% v/v glycerol in a microfuge tube If the bacterial strains contained plasmids, the selective antibiotic was included in the culture Duplicate stocks were stored at –20°C and –80°C Working stocks streaked on LB agar plates, containing antibiotics where appropriate, were stored at 4°C

2 10 7 Enzymes

All restriction endonucleases, Antarctic Phosphatase, Phusion High Fidelity DNA Polymerase and T4 DNA ligase were obtained from Invitrogen Life Technologies or New England Biolabs Deoxyribonuclease I, Ribonuclease A (Deoxyribonuclease free), M-MLV reverse Transcriptase, RED*Taq* and REDAccu*Taq* LA DNA polymerases were obtained from Sigma-Aldrich Enzymes were used with their relevant buffers according to the manufacturers instructions

2 10 7 1 Reverse transcription

Combine 1 µg of RNA with 1 µl of specific reverse primer or ohgo$(dT)_{15}$ primer (final concentration 1-5 µM) and 1 µl of 10 mM dNTP mix The volume was made to 10 µl with RNAse free water The mixture was incubated at 70°C for 10 min and then placed on ice for 5 min To the mixture was added 1 µl M-MLV reverse transcriptase and 2 µl specific 10x enzyme buffer The volume was brought to 20 µl with RNAse free water The first strand was synthesised at 37°C for 50 min followed by inactivation of the transcriptase at 95°C for 10 min Two microlitres of the reaction was used as template for PCR

2.10 7 2 Polymerase chain reaction

PCR reactions (Mullis and Faloona, 1987) were carried out using a Hybaid PCR Express Thermocycler The standard PCR reaction volume was 50 µl containing 1 µl template (10-100 ng), 1-2 5 mM $MgCl_2$, 0 5 µM of each primer, 200 µM of each dNTP, 1 unit RED*Taq* or REDAccu*Taq* LA DNA polymerase and 5 µl specific 10x enzyme buffer The standard PCR program was

Stage 1	Step 1	95°C for 10 min
Stage 2 (30 cycles)	Step 1	95°C for 1 min
	Step 2	T_{ann} for 30 s
	Step 3	72°C for 1 min per kb to be synthesised
Stage 3	Step 1	72°C for 10 min

(T_{ann} was routinely 5°C below the T_m of the primers)

Phusion High-Fidelity DNA Polymerase was used for amplification of the Seprase gene and for subsequent cloning The standard Phusion PCR reaction volume was 50 µl containing 1 µl template (10-100 ng), 1 5mM $MgCl_2$, 0 5 µM of each primer, 200 µM of each dNTP, 1 unit High-Fidelity Phusion DNA polymerase and 10 µl specific 5x enzyme buffer The standard PCR program was

Stage 1	Step 1	98°C for 3 min
Stage 2 (30 cycles)	Step 1	98°C for 15 s
	Step 2	T_{ann} for 30 s
	Step 3	72°C for 30 s per kb to be synthesised
Stage 3	Step 1	72°C for 10 min

(T_{ann} was routinely 5°C below the T_m of the primers)

2 10 8 Gene manipulation

A variety of plasmid-based gene cloning approaches such as described by Maniatis et al (1987) were used, employing PCR techniques (Section 2 10 7 2), restriction endonucleases and DNA ligase (Section 2 10 7)

2 10 8 1 Antarctic Phosphatase Reaction

Antarctic Phosphatase catalyzes the removal of 5' phosphate groups from DNA and RNA Phosphatase treated fragments lack the 5' phosphoryl termini required by ligases, therefore, they cannot self-ligate This property was used to decrease the vector background in the cloning strategies used One microgram of DNA (restricted and extracted according to sections 2 10 7 and 2 10 3 4 respectively) was added to 1X Reaction Buffer and then 1 µl Antarctic Phosphatase added This mixture was mixed and incubated overnight at 37°C The mixture was extracted (section 2 10 3 4) and resuspended in 13 µl TE buffer for ligations

2 10 8 2 Cloning of PCR products

PCR products were cloned using both prokaryotic and mammalian expression vectors (1) pCR2 1 (2) pQE30Xa (3) pcDNA3 (4) pcDNA3-HA (5) pIRES (see Table 2 10 3, Figures 2 10 1 2 10 2, 2 10 3 and 2 10 4 respectively) The ligation was transformed into *E coli* cells (routinely DH5α or XL-10 Gold strains were used, see Table 2 10 1) as described in section 2 10 5 2 Then 100 µl of the transformation reaction was plated on an LB agar plate containing the appropriate antibiotic For cloning into pCR2 1 60µl of 40mg/ml X-gal, to test for α-complementation of β-galactosidase, was spread onto the agar plate X-gal stock was prepared in DMF and stored in the dark at -20°C The *lacZα* gene encodes the α-peptide of β-galactosidase, which cleaves the X-gal substrate yielding a blue product This allows for positive selection of transformants harbouring plasmids with PCR product inserts Routinely a small proportion of colonies are blue It is assumed that these arise as a result of re-circularisation of the pCR2 1 vector

Colonies arsing from the transformations were routinely screened using the 1-2-3 Method for Plasmid DNA Isolation (section 2 10 3 2 1), and restriction analysis identified transformants with the correct insert of DNA

2 10 9 DNA sequencing

Recombinant clones were verified by DNA sequencing Commercial sequencing services were provided by MWG Biotech AG Suitable sequencing primers (Table 2 12 2) for standard vectors were provided as part of the service Samples were invariably sent as dried plasmid DNA

2 10 10 Bio-Informatics

Nucleotide and Amino Acid sequences were analysed using a variety of web-based tools The BLAST programs (Altschul *et al*, 1997) at the National Centre for Biotechnology Information (http //www ncbi nlm nih gov) were used to identify homologous sequences deposited in GenBank (Benson *et al*, 1996) Protein sequences, structure files and tools to calculate protein molecular weights and isoelectric points were obtained from the Swiss-Prot database (Bairoch and

Apweiler, 1996) at (http //us expasy org) and the Protein Data Bank (PDB, Berman *et al*, 2000) at (http //www rcsb org/pdb) Alignments of DNA and Protein sequences were performed using the MultAlin program (Corpet, 1988) available at (http //prodes toulouse inra fr/multalin/multalin html) and edited using the GeneDoc program (Nicholas *et al*, 1997) available for download at (www psc edu/biomed/genedoc) Plasmid maps were constructed using the pDRAW32 program available for download at (http //www acaclone com) Chemical structures were drawn using the ACD/Labs ChemSketch program available for download at (http //www acdlabs com) DNA sequences were analysed for restriction enzyme sites using the Webcutter 2 0 tool at (http //rna lundberg gu se/cutter2/) Tertiary protein structure was analysed and visualised using the DeepView software (Guex and Peitsch, 1997) available for download at (http //ca expasy org/spdbv)

2 10 11 Prokaryotic Expression System

2 10 11 1 Protein expression

2 10 11 1 1 Standard expression culture

A 100ml portion of LB broth was inoculated with 1 ml of a stationary phase culture of *E coli* that had been transformed with an expression plasmid Selective antibiotics were included in the LB broth The culture was incubated at 37°C, shaking at 220 rpm, until an optical absorbance (A_{600}) of 0 3-0 5 was reached IPTG was added (from 10 mM stock) to a final concentration of 50 µM to induce expression (if required) The culture was allowed to incubate for 4 hours , then centrifuged at 5,000 rpm (3833 x g) for 5 min (using a Beckman JA-14 rotor) to pellet the cells The supernatant was discarded and the pellets were stored at -20°C

2 10 11 1 2 Preparation of cleared lysate

A cell pellet from a 100 ml expression culture was washed in potassium phosphate buffer, pH 7 4 followed by centrifugation at 4,000 rpm for 5 min (using a Beckman JA-20 rotor) The supernatant was discarded and the cells were re-suspended in 10 ml potassium phosphate buffer, pH 7 4 The cells were disrupted on ice with a 3 mm micro-tip sonicator (Sonics & Materials Inc) using 2 5 s, 40 kHz pulses for 30 s The cell debris was removed by centrifugation at 4,000 rpm for 20 min at 4°C (using a Beckman JA-20 rotor) The cleared lysate was transferred to a fresh universal container and stored at 4°C

2 10 11 2 Protein purification using the Standard IMAC procedure

Immobilised Metal Affinity Chromatography (IMAC) was used to purify recombinant human Seprase having an N-terminal 6xHis tag

2 10 11 2 1 Standard IMAC procedure

One millilitre of nickel-nitrilotriacetic acid resin (Ni-NTA, Qiagen) was gently mixed with 10 ml of cleared lysate for 60 min at 4°C The mixture was poured into a 0 7 x 15 cm column, allowing the resin to settle The column was washed three times with 10 ml potassium phosphate buffer, pH 7 4 containing 20 mM imidazole and then eluted with 5 ml potassium phosphate buffer, pH 7 4 containing 200 mM imidazole The eluate was dialysed overnight against 1 L potassium phosphate buffer, pH 8 0 Samples taken throughout the procedure were analysed by SDS-PAGE (Section 2 5) Protein concentration was determined by the standard BCA or Coomassie Plus assay (Section 2 1) and Seprase activity was determined by the fluorimetric assay (Section 2 3 3) The purified Seprase sample was stored at 4°C or –20°C with 40% glycerol

2 10 11 2 2 Recharging of Ni-NTA resin

This procedure was routinely used before re-using the Ni-NTA Resin The resin was poured into a column and washed with 2 column volumes (2cv) of distilled water followed by 2cv 50% v/v ethanol The resin was then stripped with 3cv 100 mM EDTA, pH 8 0 The resin was washed with 2cv 500 mM NaCl followed by 2cv distilled water The resin was re-charged with 2cv 100 mM $NiSO_4$ The resin was finally washed with 2cv distilled water, transferred to a plastic container and stored at 4°C in 20% v/v ethanol

2 10 12 Mammalian Expression System

2 10 12 1 Cell Culture Methods

All cell culture techniques were performed in a sterile environment using a Holten laminar air flow cabinet Cells were visualised with an Olympus CK2 inverted phase contrast microscope Cell culture media, supplements and related solutions were purchased from Sigma-Aldrich, Dublin, Ireland, unless otherwise stated

Phosphate buffered saline (PBS) (Lennox) was prepared by dissolving five tablets in 500ml ultra-distilled water (dH_2O) This was then autoclaved at 115°C and 15 lb/in^2 for 20 min PBS was stored at room temperature 1X Trypsin/EDTA solution (T/E) was made up as follows 50 ml of 10X Trypsin and 10 ml of 1% w/v EDTA were added to 440 ml PBS This was aliquoted into sterile universal containers and stored at -20°C A stock solution of 1% EDTA can be made up in advance and stored at 4°C

2 10 12 1 1 Culture of Cells in Suspension

The cell line DG75 was maintained in supplemented RPMI 1640 medium Additional supplements were added to the culture media 10% Fetal Calf Serum (FCS), 2 ml L-Glutamine (200 mM), 2 ml penicillin/streptomycin (10 mg/ml) Cultures were seeded at a density of 2 x 10^5 to 5 x 10^5 cells per ml in 25 cm^2 flasks and expanded in 75 cm^2 flasks Cells were sub-cultured two or three times per week by harvesting into a sterile centrifuge tube and centrifuged at 1000 x g for 5 min at room temperature The cell pellet was resuspended gently in an appropriate volume of fresh media and replaced into the tissue culture flask All cell lines were incubated in a humid 5% CO_2 atmosphere at 37°C in a Heraeus cell culture incubator

2 10 12 1 2 Culture of adherent cells

The cell lines Hs578T and SW480 were maintained in supplemented DMEM For Hs578T, high glucose DMEM was supplemented with 5% (v/v) FCS, 5ml penicillin/streptomycin (10 mg/ml), 4 mM L-Glutamine (200 mM stock) and 1 5 ml insulin (10 mg/ml) SW480 medium was supplemented with 10% (v/v) FCS, 5ml penicillin/streptomycin As the cells were strongly adherent, trypsinisation was

required for harvesting prior to sub-culturing For trypsinisation, the medium was decanted and the cells were washed with 2 ml of sterile 1X PBS to remove any residual FCS which contains a Trypsin-inhibitor activity (α_2 –macroglobulin) Then 2 ml of 1X T/E solution was placed in each flask and the flasks incubated at 37°C for 5 min or until all the cells could be visualized as having detached from the flask surface The cell suspension was then decanted into a sterile centrifuge tube containing 5 ml of sterile supplemented medium (FCS inhibits Trypsin) and centrifuged at 1000 x g for 5 min Cells were resuspended in supplemented medium at 2 to 5 x 10^5 cells/ml, using 5 ml per 25 cm^2 flask and 15 ml per 75 cm^2 flask Cells were then incubated as described in section 2 10 12 1 1

2 10 12 2 Cell Counts

Cell counts were performed using an improved Neubauer haemocytometer slide Trypan blue exclusion dye was routinely used to determine cell viability Ten microlitres of trypan blue was added to 90 µl of a cell suspension and mixed A sample of this mixture was added to the counting chamber of the haemocytometer and cells were visualized by light microscopy Viable cells appear clear and do not stain, whereas non-viable cells stain blue from the influx of trypan blue across breached membranes

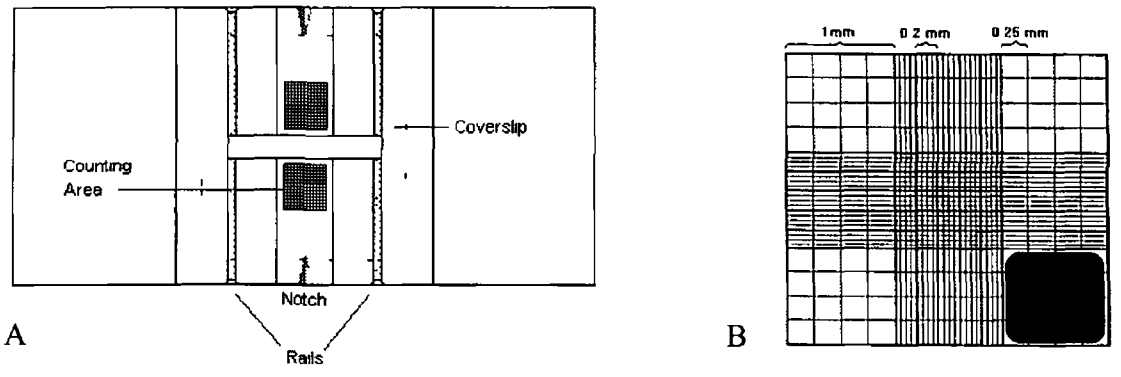


Figure 2 10.12.2 Diagram of Haemocytometer
(A) Diagram of haemocytometer with cover slip (B) Illustration of squares on a haemocytometer, showing one of the corner quadrants shaded The volume underneath the coverslip of this shaded area (or one square) is 0 1 mm^3 or 10^{-4} ml

Cells were counted from the four large corner quadrants and the centre square as observed under the 10X objective This total number was divided by 5 to give the average cell number per square This was multiplied by the dilution factor of 1 1 and then by 10^4, which results in the total cell number per ml

2 10 12 3 Cell Storage and Recovery

Cell stocks were prepared for long term storage as follows Suspension cells 1 x 10^7 cells in exponential phase were pelleted and resuspended in 800µl of supplemented RPMI to which 100µl of FCS was added, then placed on ice for 10 min DMSO was added to a final concentration of 10% (v/v), mixed gently and transferred to a sterile cryotube Adherent cells one confluent 75 cm^2 flask of adherent cells was used per cell stock Adherent cells were washed with 1X PBS followed by trypsinisation and resuspended in 900µl of FCS and 100µl of DMSO The cells were mixed gently and added to a sterile cryotube The cryotubes were slowly lowered into the gas phase of liquid nitrogen and immersed in liquid nitrogen in a cryofreezer (Cooper Cryoservices Ltd) Cells were recovered from liquid nitrogen by thawing rapidly at 37°C and transferred to a sterile centrifuge tube containing 5ml of pre-warmed supplemented medium The cells were centrifuged at 1000 x g for 5 min, the pellet was resuspended in 5-10ml of fresh supplemented medium, transferred to a culture flask and incubated at 37°C in 5% CO_2

2 10 12 4 Transfection of Cells using Electroporation

In all cases, cells were seeded at a density of 5 x 10^5 per ml of medium 24 h prior to transfection Total DNA for transfection was normally purified using Qiagen plasmid mini kit according to section 2 10 3 2 3, using the same total quantity of DNA per transfection

On the day of transfection, adherent cells were trypsinised as described in section 2 10 12 1 2 Suspension cells were centrifuged at 1,000 x g for 5 min Cells were washed twice with cold unsupplemented medium and resuspended at a density of 4 x 10^7 cells per ml of cold unsupplemented medium For each transfection, 10 µg DNA to be transfected (dissolved in 20 µl T E pH 7 4, section 2 10 2) was added to 200 µl

(1 x 10^7 cells) cold unsupplemented medium cell suspension This cell/DNA suspension was then dispensed into labelled pre-cooled cuvettes (BioRad, 0 4 mm) and placed on ice Each cell/DNA mix was then pulsed at (1) 250V for DG75 cell lines and (2) 100V for SW480 cell lines, with a capacitance of 960μF (with capacitance extender) in a BioRad electroporator, and the cuvettes returned to ice immediately after electroporation Cells must not stay longer than 10 min on ice before being transferred to medium Contents of cuvettes were transferred to culture dishes using a micropipette and yellow tip, treating the cells gently Cuvettes were washed with warmed supplemented medium from the culture dish, and placed at 37°C in a 5% CO_2 incubator for the required amount of time

2 10 12 5 Stable Transfections

Stable transfectants were selected by the addition of geneticin (G418) (Sigma) Geneticin is an aminoglycoside antibiotic similar in structure to gentamicin, neomycin and kanamycin It is a selective antibiotic, used in the selection of mammalian cells, that interferes with the function of 80S ribosomes and protein synthesis A stock solution of 50 mg/ml was made up in supplemented medium On complete dissolution of G418, the solution was sterile filtered and then stored at -20°C The DG75 stably transfected cell lines were maintained under permanent selection in supplemented RPMI 1640 containing 1 mg/ml G418 The SW480 stably transfected cell lines were maintained in supplemented DMEM containing 1 mg/ml G418

2 10.12 6 Western Blot

2 10 12 6 1 Preparation of Cellular Protein

Prior to protein isolation, the viability of the cells was examined and the viable cell count determined by trypan blue exclusion as before (section 2 10 12 2) This method was employed to isolate total cellular protein (i e cytoplasmic and nuclear proteins) and was used for the extraction of Seprase protein Suspensions cells were pelleted at 1,000 x g for 5 min and washed with 5 ml of ice-cold PBS The cells were then transferred to a microfuge tube in 1 ml of ice-cold PBS, pelleted at 5,000 x g for 5 min and all of the supernatant removed Adherent cells were trypsinised according to section 2 10 12 1 2 The cell pellet was washed twice with 5 ml of ice-cold PBS, and then pelleted by centrifugation at 5,000 x g for 5 min and all of the supernatant removed The cell pellet was dispersed in ice-cold suspension buffer (0 1 M NaCl, 0 01 M Tris-Hcl pH 7 6, 0 001 M EDTA pH 8 0, 1 μg/ml aprotinin, 1 μg/ml leupeptin and 100 μg/ml PMSF), using 200 μl of suspension buffer for every 5 x 10^6 cells An equal volume of 2X SDS gel loading buffer [100 mM Tris-HCl pH 7 6, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercaptoethanol, 0 2% (w/v) bromophenol blue] was immediately added to the cell suspension, after which the sample becomes extremely viscous The sample was then placed in a boiling water bath for 10 min and then subjected to sonication for 1 min on full power to shear the DNA [2 5 sec, 40kHz pulses] The lysate was clarified by centrifugation at 12,000 x g for 10 min at room temperature The supernatant was aliquoted and stored at – 20°C Samples were analyzed by SDS-PAGE (section 2 5), loading approximately 5 x 10^5 cells per lane

2 10 12 6.2 Western Blot Analysis

An SDS-PAGE gel was run as described in section 2 5 Coloured molecular weight markers (Sigma ColorBurst, Figure 2 5 1) were included on the gel 6 pieces of 3 mm filter paper (Whatman) and 1 piece of nitrocellulose membrane (Schleicher and Schuell) were cut to the size of the gel The filter paper sheets, membrane and gel were soaked in Transfer Buffer (section 2 10 2) for 15 min Avoiding air bubbles, 3 sheets of filter paper were placed on the cathode of a horizontal semi-dry electro-blotter (ATTO), followed by the membrane, the gel, 3 sheets of filter paper and

finally the anode The protein transfer was allowed to take place at a constant 15V for 20 min The membrane was blocked in 30 ml 5% (w/v) Milk Powder in 1X TBS-T buffer (0 1% (v/v) Tween-20 in TBS) for 60 min, followed by incubation in 20ml of Anti-HA (6E2) monoclonal antibody solution (prepared in blocking buffer) overnight at 4°C (see Table 2 10 12 6)

Primary Antibody	Dilution	Secondary Antibody	Dilution
HA-Tag (6E2) Mouse Monoclonal Antibody (Cell Signalling Technology)	1 1000	Ap-Conjugated Anti Mouse IgG (Promega)	1 5000
Mouse HA Monoclonal Antibody, Clone HA7 (Sigma)	1 1000	Alexa Fluor 488 anti-mouse fluorescent antibody (Molecular Probes)	1 400

Table 2.10 12 6 Incubation Conditions for Antibodies Used

After overnight incubation, the membrane was washed three times in TBS-T for 15 min The membrane was then incubated with the appropriate secondary antibody (see Table 2 10 12 6) for 90min at room temperature, followed by washing three times with TBS-T for 15 min each All of the above incubations were carried out with agitation Membranes were then placed in a clean container and covered with BCIP/NBT (Sigma) substrate in the case of the Ap-conjugated complexes The BCIP/NBT container was placed in the dark at room temperature without shaking for 30 min or longer if required The membranes were then rinsed in distilled water to stop the reaction, scanned or photographed, then wrapped in cling film and stored in the dark

2 10 12 7 Immunocytochemistry

In order to visually monitor the expression and/or subcellular localisation of proteins, cells were prepared for immunocytochemical analysis as described by Groarke et al (2001) with minor modifications Cells were washed twice in phosphate buffered saline (PBS) and fixed with 3% v/v formaldehyde for 15 min Cells were subsequently washed, permeabilised for 15 min with 0 2% Triton X-100 and blocked for 30 min in 5% w/v Marvel solution Following blocking, cells were incubated for 2 h with the primary Mouse HA mAb, Clone HA7 as indicated in Table 2 10 12 6 This was followed by 1 hour incubation with 1 400 dilution of Alexa Fluor 488 anti-mouse fluorescent secondary antibody (excitation/emission maxima of 495/519nm)

Nuclear DAPI staining was routinely performed by incubating cells with 0 5 x 10^{-6} µg/ml DAPI for 3 min Cells were sealed with coverslips using DAKO mounting medium (DAKO Cytomation, Cambridgeshire UK) and visualized by standard fluorescent microscopy (Olympus BK50) Suspension cells were centrifuged and washed with PBS between each step

2 11 Clinical Research Study of Human Serum

Twenty patients in total were recruited through BreastCheck and the Symptomatic Breast Care Unit of the Mater Misericordiae Hospital over a five month time period The study and consent forms were approved by the Research Ethics Committee of the Mater Misericordiae Hospital Ref 1/378/1014 Patients had histologically proven Invasive Ductal Carcinoma and were awaiting scheduled excision (10 patients in total) This form of breast cancer was chosen as immunohistochemistry studies have shown that Seprase is over expressed in these cases (see Section 1 6) Patients were informed of the study aims on the evening of admission (or at any appropriate pre-operative time-point) and signed the pertinent Consent Form The patient's basic demographics were recorded prior to blood sampling (see Table 8 1) Venous blood was withdrawn from the patient's ante-cubital fossa at the time of admission and prior to any surgical procedure A standard non-heparinised labelled blood bottle was used for this purpose The blood was then transferred to the Surgical Laboratory where it underwent centrifugation (15,000 rpm x 20min) to isolate the serum The serum was removed and transferred to a plastic labelled container which was stored in the Surgical Laboratory at -80°C

Clinical samples obtained were analysed for enzyme activity and protein content, using the fluorimetric assay as outlined in section 2 3 3 and the Biuret assay as in section 2 1 1 Statistical analysis was performed by Dr Michael Parkinson of the School of Biotechnology, D C U Normal distribution was assessed using the Kolmogorov-Smirnov test Variance homogeneity was evaluated by the Levene test The statistical significance of the difference in serum Seprase levels was assessed by Student's t-test performed in the SPSS program for Windows P values $\leq$0 05 were considered statistically significant (P = statistical significance)

All additionally relevant patient details were collected when available These included (but were not limited to)

(a) Tumour size

(b) Tumour grade

(c) Lymph node status

(d) Lympho-vascular invasion status

(e) ER/PR status

(f) Her2-neu status

Venous samples from a control population (age- and sex-matched) were also collected for comparative purposes This population comprised 10 patients admitted for elective varicose vein surgery When the blood samples from the entire complement of patients in the study cohort had been accrued, they were transferred to the School of Biotechnology, DCU It was important that all patients included in this pilot study be in good health (excluding the presence of breast cancer/varicose veins) prior to recruitment Exclusion criteria incorporated the following

(a) History of epithelial cancer (esp colonic/gastric/cervical/melanoma)

(b) Inflammatory disorders (esp rheumatoid arthritis/polyarteritis, etc)

(c) History of recent surgery or trauma

Chapter 3

Purification and Identification of Seprase from Bovine Serum

3 1 Protein Determination

In the Biuret assay, copper ions in alkaline solution react with peptide bonds to form a purple colour with an absorbance maximum at 540nm (Gornall et al , 1949) The linear relationship between protein concentration and absorbance at 540nm is observed up to 10mg/ml

The BCA assay is similar to the Lowry procedure (Lowry et al , 1951) but is more sensitive than either Biuret or Lowry procedures It also has less variability than the Bradford assay Cu^{2+} is reduced by protein to Cu^{1+}, and it is this reduction that is proportional to the amount of protein present in a sample BCA forms a purple-blue complex with Cu^{1+} thus allowing the reduction of alkaline Cu^{2+} by proteins to be monitored, at 570nm (Smith et al , 1985)

The Coomassie Plus protein assay is a modification of the Bradford method (Bradford, 1976) It was used for the detection of low levels of protein, typically in the range of 2 5 to 25µg/ml The coomassie molecule binds to protein in acidic conditions causing a colour change from brown to blue, which maximally absorbs at 595nm

Protein Standard curves incorporating BSA were prepared as outlined in sections 2 1 1, 2 1 2 and 2 1 3 Plots of protein absorbance versus BSA concentration are presented in Figures 3 1 1, 3 1 2, 3 1 3 for the Biuret, Standard BCA and Coomassie Plus assays respectively

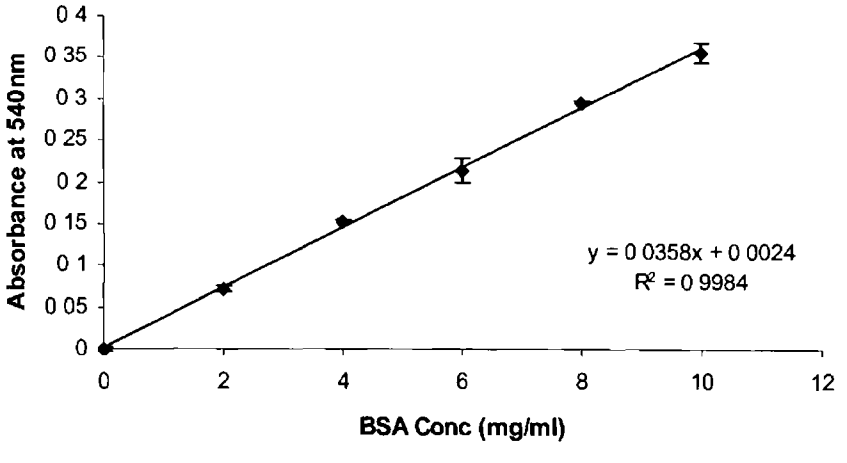


Figure 3 1 1 BSA Standard Curve

Plot of absorbance at 540nm versus BSA concentration obtained using the Biuret assay procedure as outlined in section 2 1 1 Error bars represent the SEM of triplicate readings

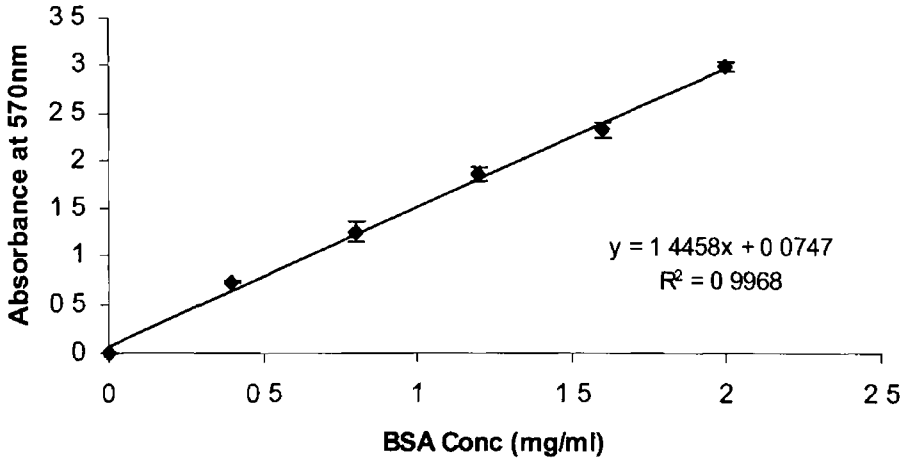


Figure 3 1 2 BSA Standard Curve

Plot of absorbance at 570nm versus BSA concentration obtained using the Standard BCA assay procedure as outlined in section 2 1 2 Error bars represent the SEM of triplicate readings

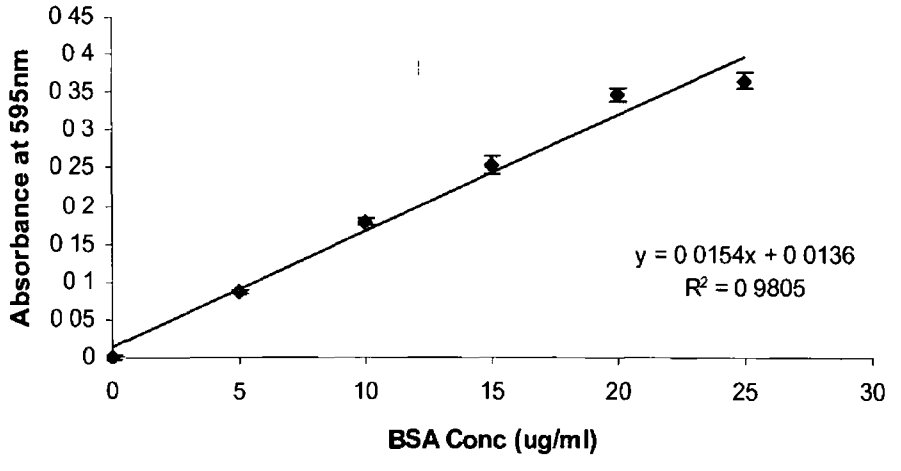


Figure 3 1 3 BSA Standard Curve

Plot of absorbance at 595nm versus BSA concentration obtained using the Coomassie Plus assay procedure as outlined in section 2 1 3 Error bars represent the SEM of triplicate readings

3 2 AMC Standard Curve and Inner Filter Effect

ZIP (Seprase) activity studied in this work was determined fluorimetrically by means of monitoring AMC release The linear relationship between fluorescence and AMC concentration allows for the accurate quantification of enzyme activity (Appendix A)

The synthetic substrate used in this work for the detection of Seprase activity was Z-Gly-Pro-AMC Previously described enzyme assays (Collins et al , 2004) were modified to a microtitre plate assay format This allowed for the use of minimal sample volume, while maintaining the sensitivity and accuracy of the enzyme assay While developing these assays it also had to be considered that Z-Gly-Pro-AMC was described as a specific Prolyl Oligopeptidase substrate (Yoshimoto et al , 1979) A specific Prolyl Oligopeptidase inhibitor, JTP-4819 (instead of Z-Pro-Prolinal) was incorporated into the assay as outlined in section 2 3 3 This allowed the identification of Seprase activity in samples that also contained Prolyl Ohgopeptidase activity, such as serum and tissue samples Section 2 3 2 describes the determination of Seprase activity in situations where it was most certainly free of Prolyl Ohgopeptidase activity

The non-quantitative microtitre plate assay described in section 2 3 4 allowed for the rapid identification of Seprase or Prolyl Oligopeptidase activity in post column chromatography fractions, using fewer sample and substrate volumes and requiring half the incubation period before analysis

The linearity of these discontinuous assays, with respect to time and enzyme concentration, has been shown previously by Birney and O'Connor (2001)

AMC standard curves were prepared as outlined in section 2 2 1 Plots of fluorescent intensity versus AMC concentration are presented in Figures 3 2 1 and 3 2 2

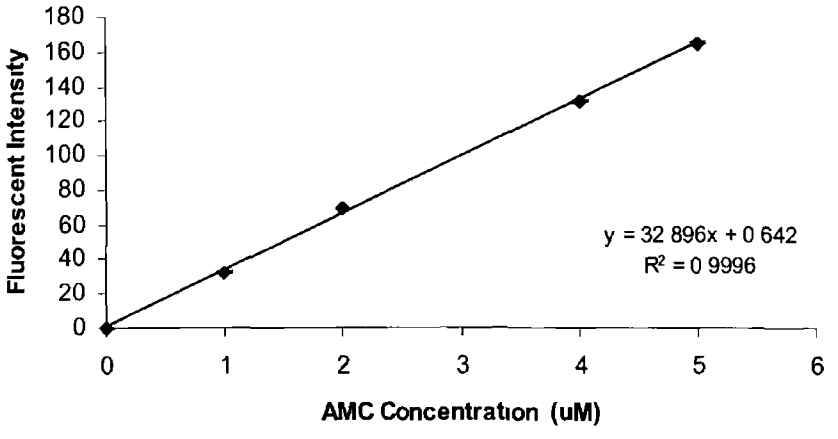


Figure 3 2 1 AMC Standard Curve

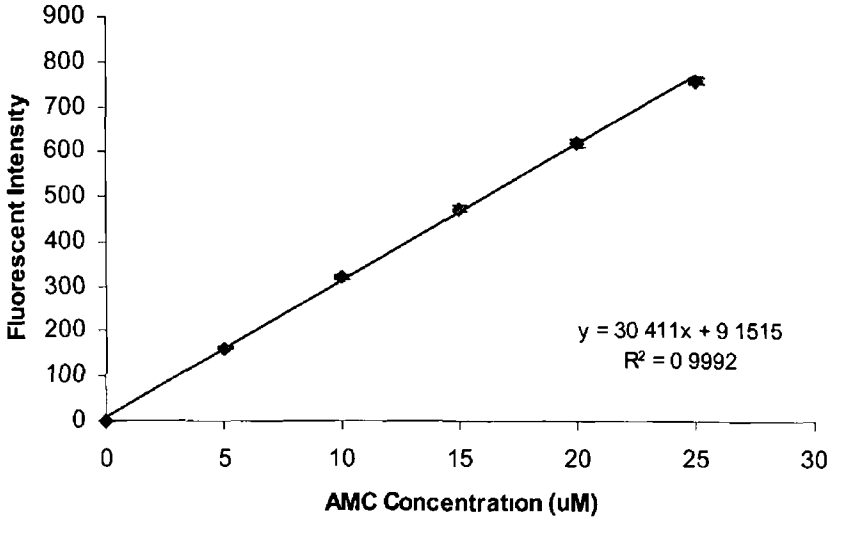


Figure 3 2 2 AMC Standard Curve

Plots of fluorescent intensity versus AMC concentration Excitation slit width was maintained at 10nm, while the emission slit width was 2 5nm

The inner filter effect (quenching) was also performed according to section 2 2 2 This was performed to observe the effect on fluorescence of including crude serum or post column pooled fractions in the assay mixture Figures 3 2 3 and 3 2 4 are plots of fluorescent intensity for crude serum and post-Phenyl Sepharose pool Plots

of the remaining post column pools are not included, as the inner filter effect was not observed for these samples

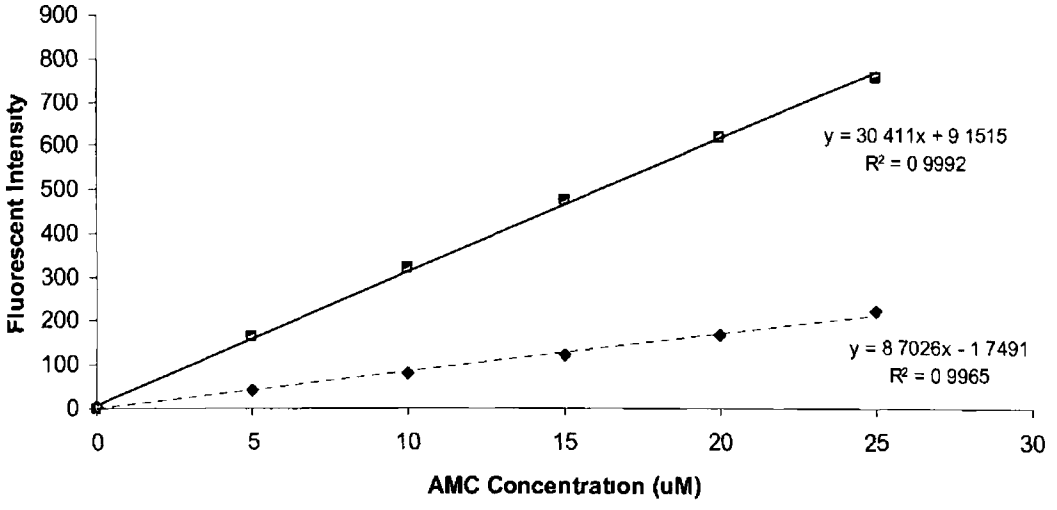


Figure 3 2 3 Serum Quenched AMC Standard Curve

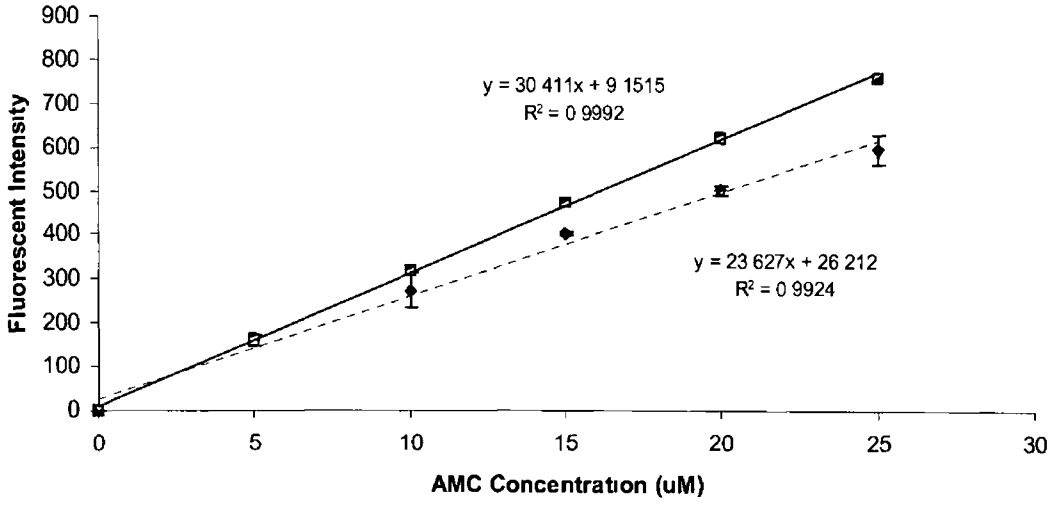


Figure 3 2 4 Post Phenyl Sepharose Quenched AMC Standard Curve

Plots of fluorescent intensity versus AMC concentration Figure 3 2 3 illustrates the filter effect for serum (---) in comparison to unfiltered for buffer (—) Figure 3 2 4 shows the filter effect for post-Phenyl Sepharose Seprase (---) in comparison to buffer (—)

3 3 An Optimised Purification Procedure for Bovine Serum Seprase (ZIP)

The purification of a target protein from a mixture involves the exploitation of a number of its physical, chemical and biological characteristics The purpose for which the protein is being purified and also the quantity required for future characterisation studies, have a significant influence on the purification strategy employed

The isolation of Seprase from bovine serum for biochemical studies required the purification of relatively large quantities of highly active enzyme The conservation of the biological activity of the enzyme was an important factor to consider when developing this purification procedure This purification process was modified by Collins and O'Connor (2003) from the original procedure (Birney and O'Connor, 2001) in order to obtain an increased level of purity This purification procedure was further enhanced in this study by the addition of affinity columns to remove contaminating protein that was interfering with the identification of the second Z-Gly-Pro-AMC degrading activity in serum

Bovine serum was chosen as the starting material for the purification protocol as a relatively large volume of purified enzyme was necessary for the biochemical studies Bovine serum was found to have a relatively high content of ZIP (Seprase) activity and was a reliable, commercially available source However, it was noted that there was batch to batch variation of Seprase and PO activity levels in the serum from different bovine sources

Appropriate precautions and considerations were taken throughout the purification procedure All steps were carried out at 4°C except the final column, which was carried out at room temperature This however didn't seem to affect the activity of the enzyme Concentration and dialysis steps were kept to the minimum, as these can be detrimental to the preservation of activity

3 3.1 Phenyl Sepharose Hydrophobic Interaction Chromatography (HIC)

Proteins are classified as hydrophobic due to the non-polar nature of their side-chains (see Appendix B) A minority of proteins have hydrophobic groups, such as alanine, methionine, tryptophan, and phenylalanine, on their surface Proteins vary with their degree of surface hydrophobicity and it is this characteristic that is exploited with HIC Protein fractionation depends on the target protein, the gel matrix and the surrounding environment, which is usually aqueous The addition of neutral salts, in this case ammonium sulphate, increases the ionic strength of the solution thereby increasing the hydrophobicity of the protein (Harris and Angel, 1989) Phenyl Sepharose was the resin of choice for the first purification step as it is classified as having a high binding capacity (20mg/ml), which is required when purifying protein from serum

Figure 3 3 1 illustrates the elution profile of Seprase activity from a phenyl Sepharose column which was run according to section 2 4 2 The profile shows the presence of two Z-Gly-Pro-AMC degrading activity peaks (as indicated by →) measured as per section 2 3 4 The first activity peak (Prolyl Oligopeptidase) is present in the run through wash, while the second major activity peak (Seprase) elutes with the ultra-pure water wash These activities are distinguished according to their sensitivity to JTP-4819 (a potent and specific inhibitor of Prolyl Oligopeptidase), which can be observed in Figure 3 3 2 Recently, another group has reported similar results in that they observed two distinct Z-Gly-Pro-AMC hydrolysing activities when purifying APCE from human plasma, one bound to the phenyl column, while the other did not (Lee et al , 2004) However, using a FRET peptide (see Sections 1 8 and 5 1) as a substrate they were able to distinguish between the two activities Only the phenyl-bound fraction possessed the proteinase activity that cleaved the FRET peptide, which is now known as Seprase

Fractions 42-46 were combined yielding 31ml post Phenyl Sepharose Seprase sample This sample was analysed quantitatively for Seprase activity (see section 2 3 3) and for protein (see section 2 1 1)

The phenyl sepharose resin was very successful in separating the bulk protein from the target protein, Seprase Figure 3 3 1 shows the bulk of the protein being eluted with Prolyl Oligopeptidase The HIC purification resulted in the substantial loss in protein, 1686mg to 64 9mg, a 5 6% loss in activity and a 24-fold purification of Seprase (see Table 3 3 1)

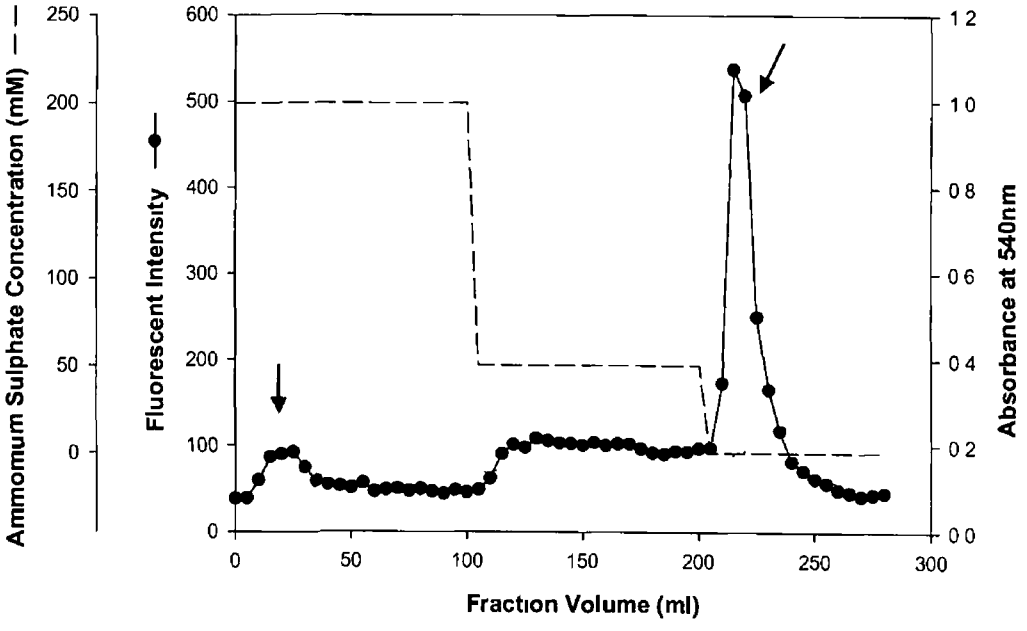


Figure 3 3 1 Fractionation of the Two Z-Gly-Pro-AMC Hydrolysing Activities in Bovine Serum by Phenyl Sepharose Hydrophobic Interaction Chromatography

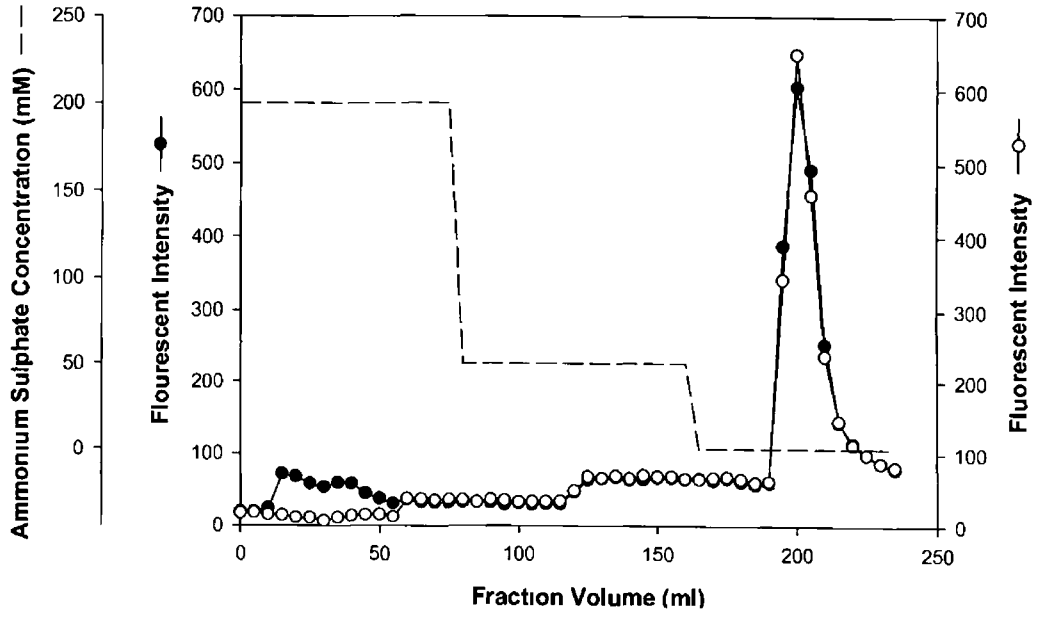


Figure 3 3 2 Differentiation of the Two Distinct Z-Gly-Pro-AMC Degrading Activities (•) in Bovine Serum by the Inclusion of JTP-4819 (○)

3 3.2 Calcium Phosphate Cellulose Chromatography

The Hydroxylapatite resin [$Ca_{10}(PO_4)_6(OH)_2$] for this purification was produced in the laboratory according to modifications to methods previously described ((Bernardi et al, 1972, Donlon and Kaufman, 1980, Tiselius et al, 1956) The resin was prepared in bulk, eliminating the possibility of batch to batch variation The mechanism of protein adsorption is thought to involve both Ca^{2+} and PO_4^{3} groups on the crystal surface (Bernardi et al, 1972)

After concentration of the post-phenyl Sepharose pool to 10ml, the sample was further purified using calcium phosphate cellulose as outlined in section 2 4 3 Figure 3 3 3 shows the elution profile of Seprase activity and protein, measured according to sections 2 3 4 and 2 1 2 Seprase bound to the column at a low phosphate concentration and eluted on application of 500mM phosphate The elution profile shows the removal of contaminating protein with 170mM wash step but more importantly separating Seprase from Gly-Pro-AMC and Z-Phe-Arg-AMC degrading peptidases (Collins and O'Connor, 2003) Fractions 42-45 were pooled yielding 23ml post calcium phosphate cellulose Seprase sample This sample was analysed quantitatively for Seprase activity (see section 2 3 2) and protein (see section 2 1 2)

This purification step was effective in removing contaminating protein, reducing protein from 64 89mg/31ml to 1mg/23ml This led to a 63% recovery of the overall biological activity of Seprase and purification factor increased to 997-fold (see Table 3 3 1)

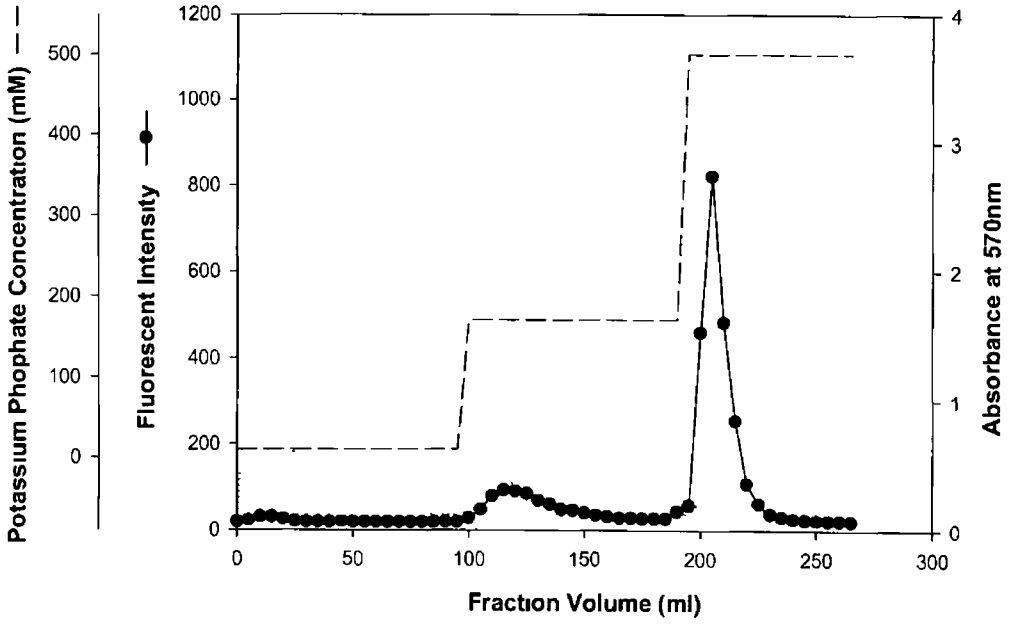


Figure 3 3 3 Elution Profile of Seprase during Calcium Phosphate Cellulose Chromatography

3 3 3 Cibacron Blue 3GA Chromatography

Cibacron Blue 3GA is a reactive dye or 'pseudo-ligand' used in affinity chromatography The resin is ideal for binding bovine serum albumin It was therefore used in this purification scheme to remove possible bovine serum albumin (a major component of bovine serum) contamination from the Seprase sample

Concentrated and dialysed post calcium phosphate cellulose Seprase was applied to a Cibacron blue 3GA column and further purified as outlined in section 2 4 4 Seprase bound to the resin and was eluted using an increasing linear gradient of sodium chloride Figure 3 3 4 illustrates the elution profile from the column

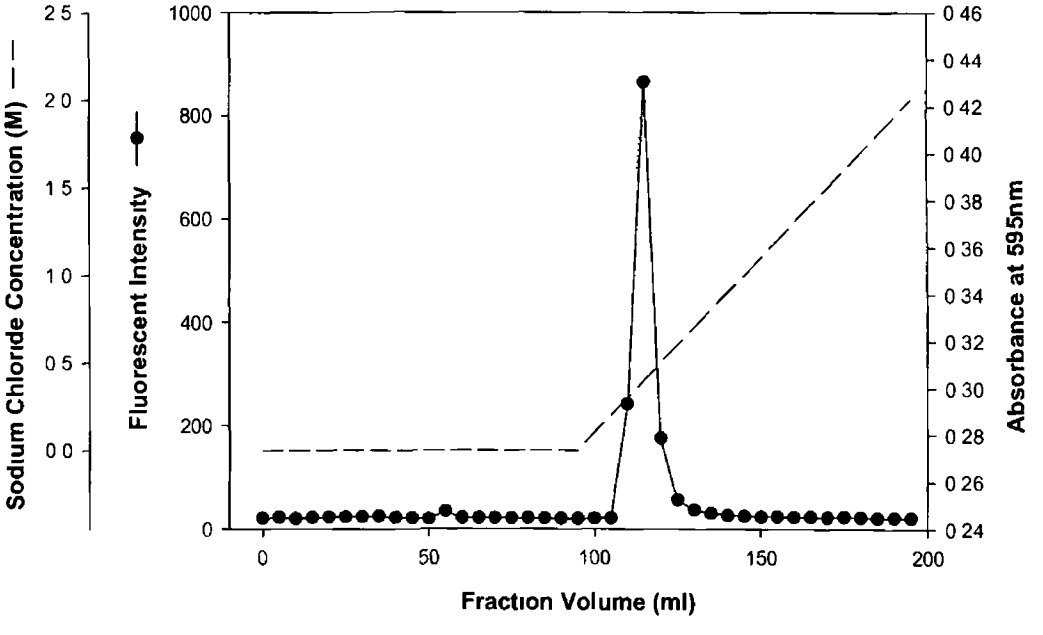


Figure 3 3 4 Elution profile of Seprase during Cibacron Blue 3GA Chromatography

Fractions 22-24 were pooled yielding the post Cibacron blue 3GA Seprase sample This sample was analysed for quantitative Seprase activity (section 2 3 2) and quantitative protein (section 2 1 3) determinations

This resin resulted in a further 7 5% loss of protein, which enabled a 7035-fold purification factor Specific Activity also increased to 84 8 Units/mg (Table 3 3 1)

3 3.4 Sephacryl S-300 HR Size Exclusion Chromatography

Gel permeation chromatography is a form of partition chromatography used for separating proteins of different sizes The HiPrep Sephacryl S-300 HR gel filtration column was chosen in this purification procedure due to its fractionation range of 10-1500kDa

2ml of concentrated post cibacron blue 3GA Seprase was applied to the S-300 column as outlined in section 2 4 5 Figure 3 3 5 shows the elution profile of Seprase from the gel filtration resin Fractions 11-13 were pooled, yielding the post-S-300 Seprase sample This sample was analysed for quantitative Seprase activity (section 2 3 2) and quantitative protein (section 2 1 3) determinations

The resin was very effective in removing larger and smaller protein contamination from the Seprase samples as shown by the protein peaks before and after the Seprase activity peak The gel filtration was an ideal purification step, with an 11% loss of the total biological activity, while reducing protein from 80µg to 30µg

Overall, the purification protocol was extremely successful, conserving 25% of the total biological activity of Seprase, while reducing the total protein from 1686mg/20ml to 0 03mg/14ml This protocol led to a specific activity 13,805-fold higher than that in crude bovine serum (Table 3 3 1)

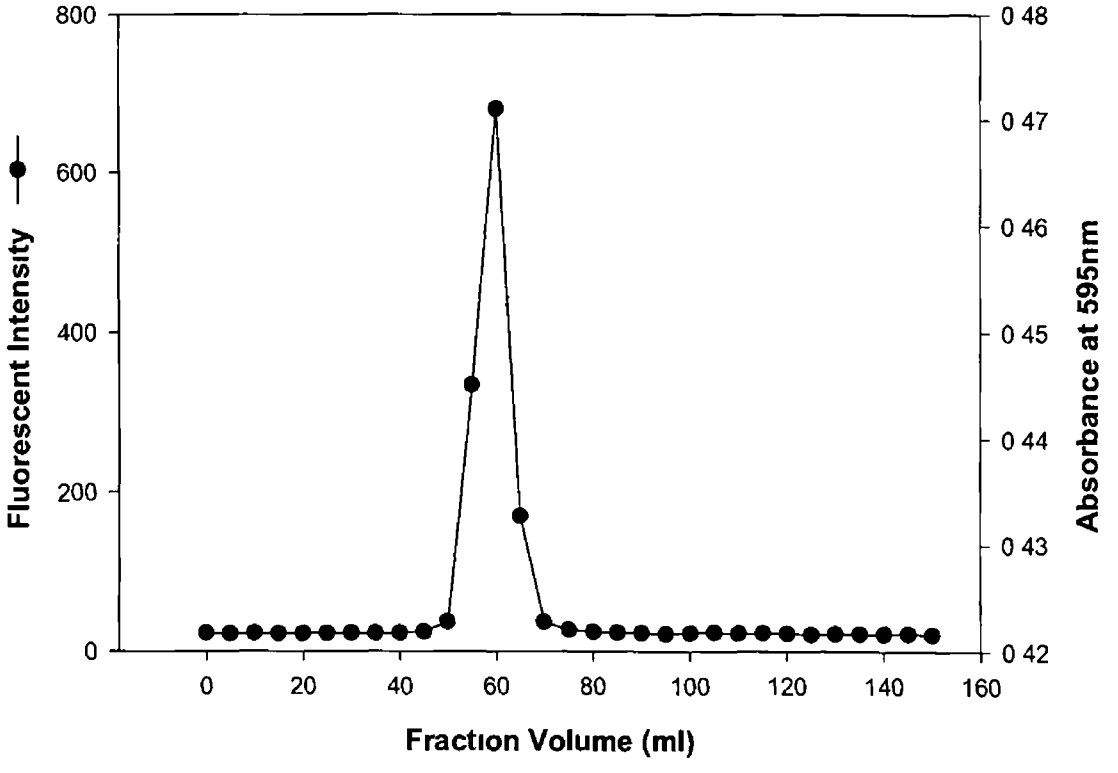


Figure 3 3 5 Elution Profile of Seprase Activity during S-300 Gel Filtration Chromatography

Purification Step	Volume	Total Activity [a]	Total Protein	Specific Activity	Purification Factor	Recovery
	ml	*Units*	*mg*	*Units/mg*		*%*
Crude Serum	20	20 33	1686 8	0 0121	1	100
Phenyl Sepharose	31	19 20	64 9	0 2960	24 6	94 45
CPC	23	12 20	1 07	12 021	997 2	63 31
Cibacron Blue	18	7 21	0 08	84 80	7,035	35 46
S-300	14	5	0 03	166 41	13,805	24 6

Table 3 3 1 Purification of Seprase from Bovine Serum

The purification table was constructed to assess the overall effectiveness of the purification procedure

[a] Based on the enzymatic activity using 100μM Z-Gly-Pro-AMC, where Units = nmol min^{-1}, i e Units are expressed as nanomol of AMC released per minute at 37°C

3 4 Purity Assessment

Purity assessment of ZIP using this purification procedure and fluorimetry-based analysis has been previously shown (Collins et al , 2004) The post S-300 purified ZIP was shown to have no hydrolysis activity on any substrate other than Z-Gly-Pro-AMC The absence of contaminating peptidases such as Prolyl Oligopeptidase and Dipeptidyl Peptidase IV was critical, as these would have interfered with the subsequent analysis such as inhibitor and substrate specificity studies

SDS-PAGE was also used to investigate the purity level of ZIP in the purified sample Figures 3 4 1 and 3 4 2 represent images of the coomassie stained and silver stained polyacrylamide gels respectively Figure 3 2 1 illustrates the progression of the purification procedure (lanes 2-5) and the concurrent loss of protein, indicated by the reduction of protein bands There is no visible band present for the target protein ZIP at 97kDa Silver staining this gel (Figure 3 4 2) shows some contaminating protein in the final sample (post S-300) However, a molecular weight band of 97kDa is still not visible in this final purified sample (lane 5)

Based on the previously obtained fluorimetry results and on the SDS-PAGE analysis, the purity of the post S-300 sample was considered to be acceptable for further studies

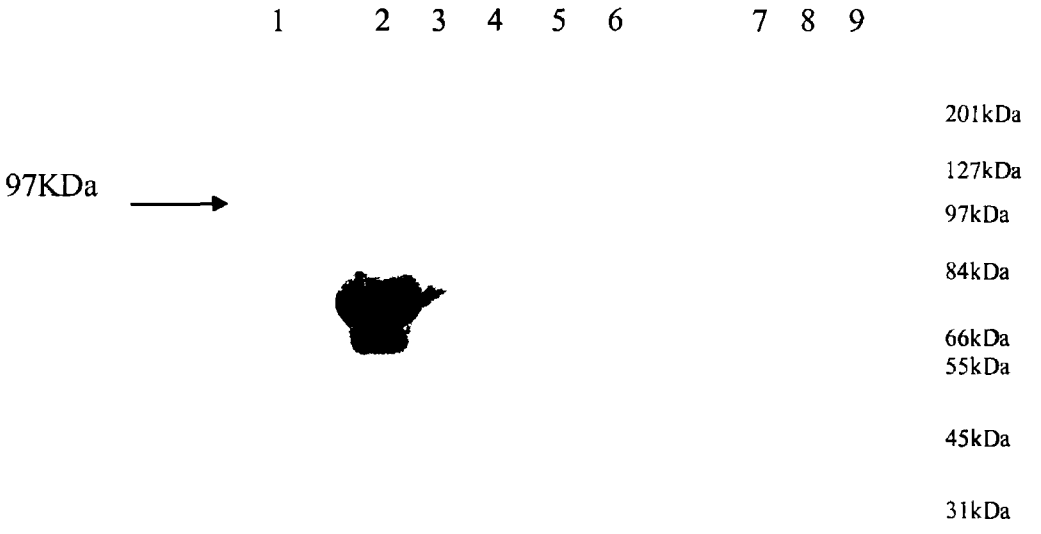


Figure 3 4 1 SDS-Gel of Purification Procedure

SDS gel stained using coomassie blue stain as per section 2 5 4 Lane (1) Sigma High Range Marker, (2) Crude Serum, (3) Post Phenyl Sepharose, (4) Post Calcium Phosphate, (5) Post S-300 Gel Filtration, (6) WGA Affinity Column, (7) Sigma High Range Marker, (8) Colourburst Marker, (9) Kaleidoscope Marker

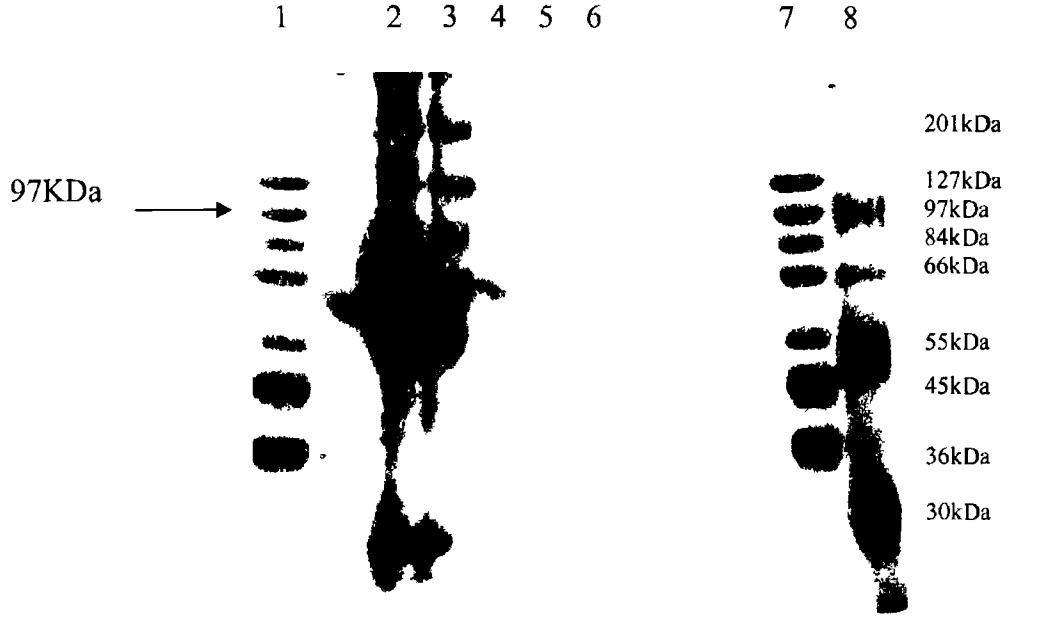


Figure 3 4 2 SDS-Gel of Purification Procedure

SDS gel stained using Silver stain procedure as per section 2 5 4 Lane (1) Sigma High Range Marker, (2) Crude Serum, (3) Post Phenyl Sepharose, (4) Post Calcium Phosphate, (5) Post S-300 Gel Filtration (6) WGA Affinity Column, (7) Sigma High Range Marker, (8) Colourburst Marker

3.5 Identification of the Z-Pro-prolinal Insensitive peptidase

The primary structure of a protein plays a major role in determining its three-dimensional structure. This in turn leads to the specific functionality of the protein i.e. its distinct biological activity. It was very important to this study to obtain the amino acid sequence data for ZIP, by N-terminal or internal sequencing. This would ultimately confirm exactly what protein ZIP is and thus finally identify the protein/protease responsible for the second Z-Gly-Pro-AMC hydrolysing activity in bovine serum.

50ml of bovine serum was purified to obtain sufficient ZIP sample for sequence analysis. SDS PAGE was performed (see section 2.6.1.1) and the gel electroblotted to a PVDF membrane (see section 2.6.1.2 and Figure 3.5.1). The blot was sent to Dr. Bryan Dunbar, Aberdeen Proteome Facility, University of Aberdeen, Scotland for N-terminal sequencing. This facility has 2 Applied Biosystem protein sequencers for Edman sequencing. The blot showed a possible faint band close to the previously reported molecular weight of ZIP, 95KDa (Birney and O'Connor, 2001). However, there were also a number of other minor contaminating bands present; possibly due to the requirement to purify relatively large amounts of starting material to observe a Seprase band.

Figure 3.5.1 Coomassie-Stained PVDF Membrane Blot

Lane (1) and (2), Purified post-S-300 bovine Seprase from 50ml serum; (3) Coloured marker

This sequencing attempt failed due to an inadequate amount of protein present For clarification purposes, the strong band seen just above the 52KDa was sequenced and determined to be BSA

Another larger sample was prepared and purified over the first three columns in the purification procedure This partially purified ZIP sample was concentrated and sent to the Proteomics Department, University of Dundee for internal sequencing Figure 3 5 2 shows that using the larger volume of starting materials and only partially purifying the sample led to increased contamination by other proteins It was thought that by removing a purification step, there would be more of the target protein present to obtain sequencing data Internal sequencing of the band below 55KDa marker identified seven proteins with the major protein identified being bovine heavy chain IgG (Collins and O'Connor, 2003) All other sequence data obtained had no link to the protease activity of ZIP

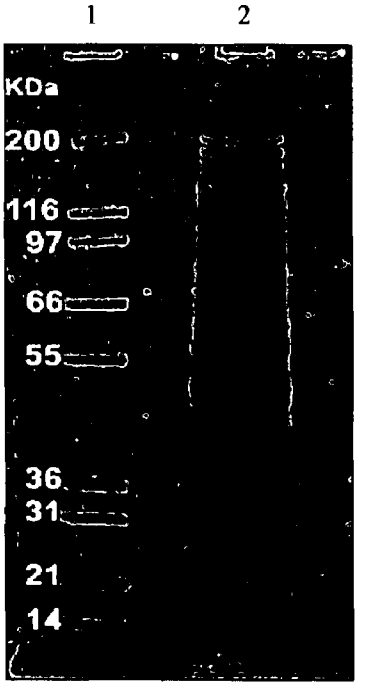


Figure 3 5 2 SDS PAGE Gel of partial purified ZIP Sample

Lane (1) Sigma High Range Molecular Weight Markers, (2) Partially purified post-Cibacron Blue bovine Seprase

3 5 1 UV Zymogram Development

These results showed that, without knowledge of the exact location of the ZIP band under denaturing conditions, it was not possible to sequence and identify the protein of interest As can be seen from the results, this was made more difficult when trying to purify large volumes of starting material in order to obtain enough target protein, as it resulted in increased amounts of minor protein contamination

In order to be sure that the correct protein of interest was being sequenced, an in-gel (zymogram) assay for ZIP was developed There were no reports of this technique having been performed previously using this enzyme A highly concentrated ZIP sample was separated on a 10% native polyacrylamide gel, prepared and run according to section 2 6 4 The presence of native ZIP hydrolysing the substrate Z-Gly-Pro-AMC is illustrated in Figure 3 5 3, which was observed under ultraviolet light

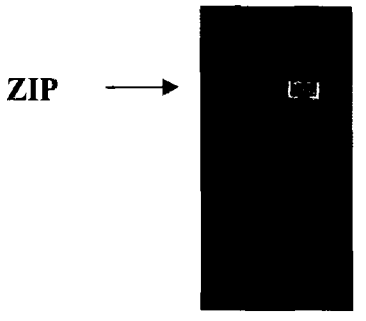


Figure 3 5 3 Native PAGE Analysis of ZIP Proteolytic Analysis

Zymogram clearly shows the fluorescent band (→) which appears in the region of Z-Gly-Pro-AMC substrate hydrolysis after visualisation under ultraviolet light

The fluorescent bands were excised and placed in ultra-pure water overnight at 4°C, so that the protein present would diffuse out of the bands into solution Due to the presence of a small amount of protein in the bands, this technique did not work Another fluorescent band which had been excised from the same gel was sent to the University of Dundee for internal sequence analysis Due to such a low level of target protein, it was difficult to identify the protein(s) present However, two proteins were identified by this approach, fibronectin and IgG Fibronectin was of interest at first, as its plasma form is known to have potential proteolytic activity (Unger and Tschesche, 1999) It was important to remove these two proteins from

the sample if they were contaminants, but if the activity was caused by fibronectin, to try and remove it by gelatin chromatography.

80ml of bovine serum was purified, in 20ml batchs as described in section 2.4. The enzyme activity from each purification was pooled and concentrated. Chromatography was performed using protein G affinity resin as outlined in section 2.6.2. This resin failed to bind ZIP and any IgG present in the sample was removed successfully. Next, the post-Protein G sample was applied to the gelatin Sepharose resin according to section 2.6.3. The ZIP activity failed to bind to this resin also, thus eliminating the involvement of fibronectin in our protease activity. Therefore, both IgG and Fibronectin were contaminating proteins co-migrating with ZIP on the native gel.

The final sample was prepared and run on a native PAGE gel, analysed using the zymogram assay and silver stained (see sections 2.6.4 and 2.5.4 respectively).

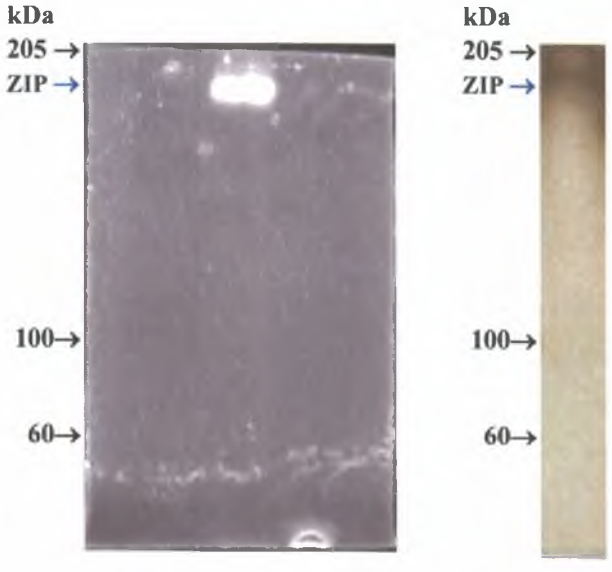


Figure 3.5.4 UV Zymogram and Silver Staining of native ZIP

The fluorescent band in Figure 3.5.4 was excised and sent to Harvard Microchemistry Facility where sequence analysis was performed by microcapillary reverse-phase HPLC nano-electrospray tandem mass spectrometry (μLC/MS/MS) on a Finnigan LCQ DECA XP Plus quadrupole ion trap mass spectrometer.

This procedure identified the ZIP protein as Seprase. Seprase is a known serine protease, associated with the small group of serine integral membrane peptidases. The peptides sequenced and identified are shown in Figure 3.5.5 and Figure 3.5.6. Table 3.7.1 compares some important biochemical properties of both ZIP and Seprase.

gi|16933540 **MS/MS Spectra:** 12 **Sum TIC:** 4.1e8 **Avg TIC:** 3.4e7
fibroblast activation protein, alpha subunit; integral membrane serine protease, seprase [Homo sapiens] gi|1924982|gb|AAC51668.1| integral membrane serine protease Seprase [Homo sapiens] gi|20072811|gb|AAH26250.1| fibroblast activation protein, alpha [Homo sapiens]

Sequence	Reference	TIC	Ions	Scan
(-) KLGVYEVEDQITAVR	gi|16933540 +3	5.4e7	22/28	899
(-) KLGVYEVEDQITAVR	gi|16933540 +3	1.1e8	35/56	897
(-) LAYVYQNNIYLK	gi|16933540 +5	3.1e7	17/22	858
(-) IFNGIPDWVYEEEM*LATK	gi|16933540 +3	1.7e7	20/34	1243
(-) FIEM*GFIDEKR	gi|16933540 +3	6.6e6	16/20	734
(-) NVDYLLIHGTADDNVHFQNSAQIAK	gi|16933540 +6	1.2e6	25/96	1260
(-) YALWWSPNGK	gi|16933540 +2	3.2e6	14/18	1019
(-) IFNGIPDWVYEEEMLATK	gi|16933540 +3	8.9e6	16/34	1263

Figure 3.5.5 Peptides sequenced and identified as Seprase

This is an example of the results obtained from the Harvard Microchemistry Facility and therefore is not an exhaustive list.

```
gi|1924982 : MKTWVKIVFGVATSAVLALLVMCIVLRPSRVHNSEENTMRALTLKDILNGTFSYKTFFPNWISGQEYLHQSADNNIVLYNIETGQSYTILSNRTMKSVNAS : 101

gi|1924982 : NYGLSPDRQFVYLESDYSKLWRYSYTATYYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVYQNNIYLKQRPGDPPFQITFNGRENKIFNGIPDWVYE : 202

gi|1924982 : EEMLATKYALWWSPNGKFLAYAEFNDTDIPVIAYSYYGDEQYPRTINIPYPKAGAKNPVVRIFIIDTTYPAYVGPQEVPVPAMIASSDYYFSWLTWVTDER : 303

gi|1924982 : VCLQWLKRVQNVSVLSICDFREDWQTWDCPKTQEHIEESRTGWAGGFFVSTPVFSYDAISYYKIFSDKDGYKHIHYIKDTVENAIQITSGKWEAINIFRVT : 404

gi|1924982 : QDSLFYSSNEFEEYPGRRNIYRISIGSYPPSKKCVTCHLRKERCQYYTASFSDYAKYYALVCYGPGIPISTLHDGRTDQEIKILEENKELENALKNIQLPK : 505

gi|1924982 : EEIKKLEVDEITLWYKMILPPQFDRSKKYPLLIQVYGGPCSQSVRSVFAVNWISYLASKEGMVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRK : 606

gi|1924982 : FIEMGFIDEKRIAIWGWSYGGYVSSLALASGTGLFKCGIAVAPVSSWEYYASVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVDYLLIHGTADDNVHF : 707

gi|1924982 : QNSAQIAKALVNAQVDFQAMWYSDQNHGLSGLSTNHLYTHMTHFLKQCFSLSD : 760
```

Figure 3.5.6 Seprase amino acid sequence

Seprase amino acid sequence (GI: 1924982) with identified sequencing peptides marked.

3.6 Gelatin Zymography

Seprase is known to exhibit gelatinase activity (Aoyama and Chen, 1990; Monsky et al., 1994; Pineiro-Sanchez et al., 1997); therefore, it was important that gelatin zymographic analysis be performed on purified ZIP/Seprase for this potential activity. Gelatin zymography was performed according to section 2.6.5. Figure 3.6.1 (A) illustrates a gelatin zymogram with a clear zone of lysis where the purified sample showed enzymatic activity towards gelatin.

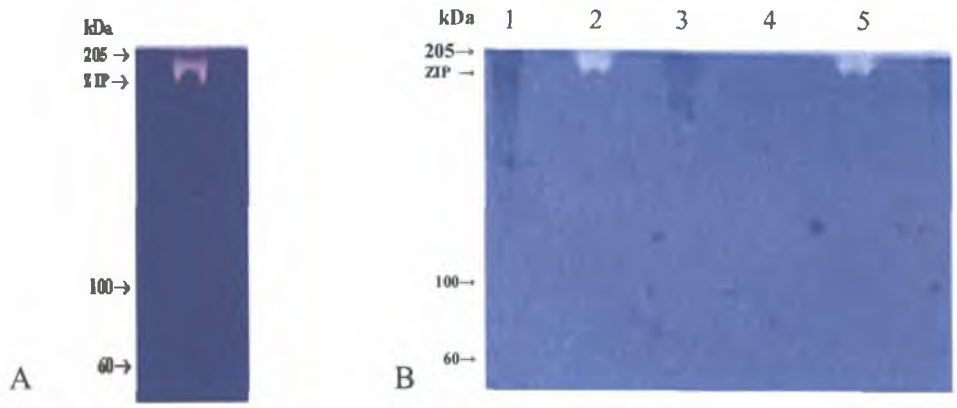


Figure 3.6.1 Gelatinolytic Activity of Seprase, detected by Zymography

(A) Purified ZIP/Seprase was loaded on a 10% native gel incorporating 1 mg/ml gelatin and activity subsequently detected using gelatin zymography. (B) Identification of Seprase as a serine protease. Purified ZIP/Seprase pre-incubated with the appropriate inhibitor was loaded on a 10% native gel incorporating 1 mg/ml gelatin and activity subsequently detected using gelatin zymography. Lane 1, Kaleidoscope Marker; Lane2, Seprase, Lane 3, Seprase incubated with PMSF; Lane 4, Seprase incubated with DFP; Lane 5, Seprase incubated with JTP-4819. Molecular masses (kDa) of standard proteins are indicated on the left.

This result also shows that Seprase, apart from having an exopeptidase activity (cleaves Z-Gly-Pro-AMC), also has an endopeptidase activity. To date, no physiological substrate has clearly been defined for this protease but its ability to effectively degrade gelatin suggests a candidate protein substrate in vivo and a possible role in extracellular matrix protein degradation. Complete inhibition of the soluble form of Seprase gelatinase activity was observed for the serine protease inhibitors PMSF and DFP (see Figure 3.6.1 (B)), as reported for the membrane-associated Seprase (Aoyama and Chen, 1990). The Prolyl Oligopeptidase-specific inhibitor JTP-4819 did not have any effect on this activity. This inhibition study suggests that Seprase contains a catalytically active serine residue.

3 7 Wheat Germ Agglutinin (WGA) Lectin Affinity Chromatography

Lectins are a family of carbohydrate binding proteins/glycoproteins They are capable of binding glycoproteins even in the presence of various detergents The agglutination activity of these highly specific carbohydrate-binding molecules is usually inhibited by a simple monosaccharide Wheat germ agglutinin has a sugar specificity for *N*-acetyl-β-D-glucosamine

Seprase has been shown to bind to the Wheat Germ Agglutinin lectin affinity resin (Aoyama and Chen, 1990, Pineiro-Sanchez et al , 1997) This affinity resin was investigated using purified ZIP/Seprase, with the thought that it would improve the standard purification procedure detailed in section 2 4

The post Sephacryl S-300 Seprase was applied to the WGA lectin column as per section 2 6 6 Bound Seprase was eluted using 0 5M N-acetyl-D-glucosamine Figure 3 7 1 illustrates the elution profile of Seprase Fractions 13-17 were pooled to yield the post WGA Seprase sample This sample was analysed for quantitative Seprase activity (section 2 3 2) and quantitative protein (section 2 1 3) determinations

The resin was very effective in removing more protein contamination from the Seprase sample, as shown by the protein peaks before and after the Seprase activity peak The affinity resin resulted in a 15% loss of the total biological activity, while reducing protein from 30µg to 7µg

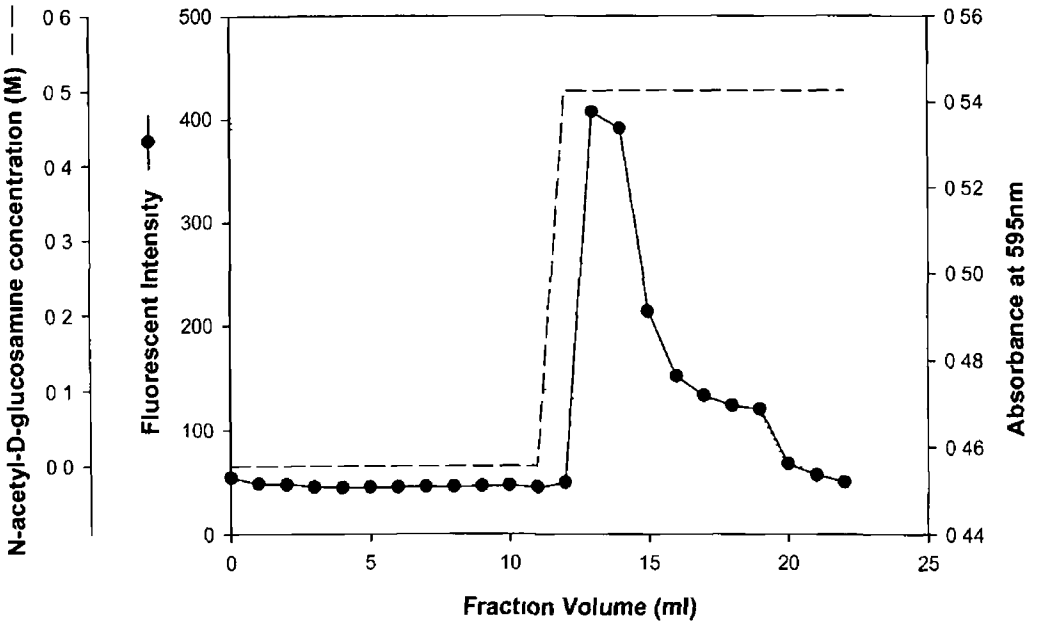


Figure 3 7 1 Elution Profile of Seprase Activity during WGA Lectin Affinity Chromatography

Overall, the purification protocol was extremely successful, conserving almost 10% of the total biological activity of Seprase, while reducing the total protein from 1686mg/ml to 0 007mg/ml This protocol led to a specific activity 21,483-fold higher than that in crude bovine serum (Table 3 7 1)

This column did show its effectiveness in binding Seprase and would be useful when purifying Seprase from other sources, such as mammalian expression systems that do not contain as much protein contamination as bovine serum

Purification Step	Volume	Total Activity [a]	Total Protein	Specific Activity	Purification Factor	Recovery
	ml	*Units*	*mg*	*Units/mg*		%
Crude Serum	20	20 33	1686 8	0 0121	1	100
Phenyl Sepharose	31	19 20	64 9	0 2960	24 5	94 45
CPC	23	12 20	1 07	12 021	997 2	63 31
Cibacron Blue	18	7 21	0 08	84 80	7,035	35 46
S-300	14	5	0 03	166 41	13,805	24 60
WGA Lectin	5 5	1 95	0 007	258 96	21,483	9 58

Table 3.7 1 Purification of Seprase from Bovine Serum with the addition of WGA Affinity Lectin Column

The purification table was constructed to assess the overall effectiveness of the purification procedure

[a] Based on the enzymatic activity using 100μM Z-Gly-Pro-AMC, where Units = nmol min^{-1} i e Units are expressed as nanomol of AMC released per minute at 37°C

	ZIP	Seprase
MW (Native) kDa	174*	170
MW (Monomer) kDa	95-96	97
pI	5 68*	~5
pH (Optimum)	7 4-8 0*	7 8
Assay Temp (°C)	37*	37
Catalytic Type	Serine protease*	Serine protease
Reaction Catalysed	Endopeptidase*, proline specific	Endopeptidase, proline specific
Inhibitors	DFP, AEBSF*	DFP, AEBSF
Gelatinase activity	Yes	Yes
Reference	*(Birney and O'Connor, 2001)	(Pineiro-Sanchez et al, 1997)

Table 3 7 2 Comparison of Important Biochemical Properties of both enzymes

Table 3 7 2 compares some of the important biochemical properties of both ZIP and Seprase This study presents the first identification of a serum/soluble form of this protease (Collins et al, 2004), as it displays a strong similarity with the primary structure (see Figure 3 5 6) and similar catalytic activity to the cell associated serine integral membrane peptidase Seprase A second group has recently identified the soluble form of Antiplasmin Cleaving Enzyme (APCE) as Seprase (Lee et al, 2005a)

Chapter 4

Biochemical Analysis of Seprase

4.1 Temperature Studies

Enzymes are sensitive to their environmental conditions Up to a point, the rate of the reaction will increase as a function of temperature because the substrate molecules will collide more frequently with the active site These sets of experiments were designed to investigate the effects of temperature on the enzymatic activity of Seprase

4 1 1 Determination of Optimum Assay Temperature

Each enzyme has an optimal temperature that favours the native conformation for maximum activity The influence of temperature on the activity of Seprase was investigated As outlined in section 2 7 1 1, the standard activity assay was carried out at a range of temperatures The resulting optimum assay temperature profile is shown in Figure 4 1 1

The physiological temperature of a healthy human, range from 36 2 to 37 2°C and for a bovine animal is 38 9°C (Prendeville et al, 2002) Most metabolic enzymes function with an optimal temperature near body temperature, but this is not always the case In this study, an increase in residual Seprase activity (%) was observed up to 40°C, which is not unexpected considering that the source of Seprase is bovine Despite this observation, it was decided to continue using 37°C as the standard assay temperature, considering the mammalian origins of this enzyme Seprase purified from the chicken embryo was reported to have a temperature optimum of 37°C (Kelly, 1999)

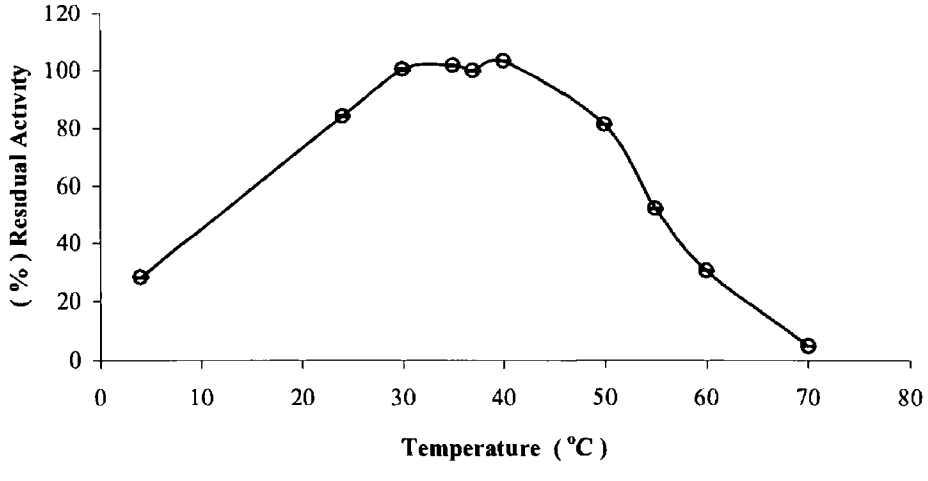


Figure 4.1.1 Determination of the Optimum Assay Temperature for Seprase

Activity levels measured at 37°C were defined as 100% Error bars shown represent SEM of triplicate readings

4 1 2 Thermal Profile, (T_{50})

Above the enzyme's optimum temperature, the enzyme degrades as the noncovalent interactions that stabilise the native conformation are disrupted This method was used to determine the temperature at which the enzyme begins to unfold and lose catalytic activity The assay was carried out at temperatures ranging from 4°C to 75°C according to section 2 7 1 2

The apparent activity increased with temperature up to 40°C, above which the activity started to decline The half-inactivation temperature of Seprase was estimated by inspection as 55°C

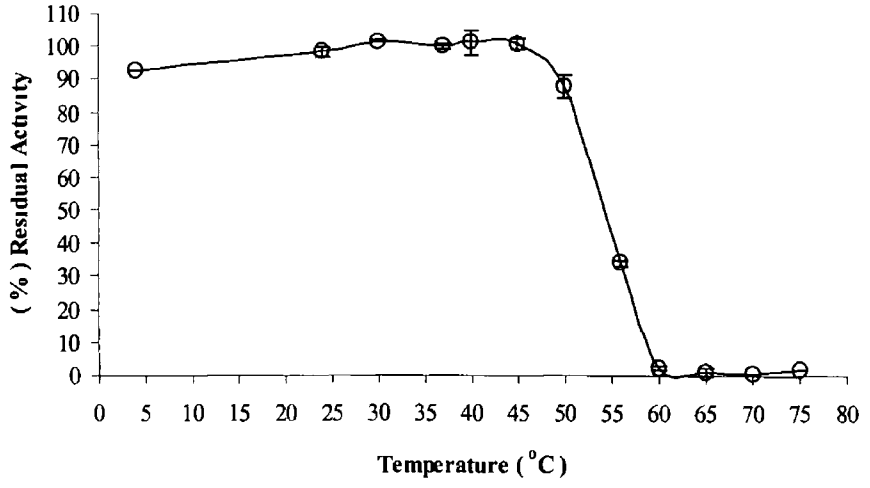


Figure 4 1 2 Thermal Profile of Seprase

Activity levels measured at 37°C were defined as 100% Error bars shown represent SEM of triplicate readings

4.1.3 Thermal Stability

Figure 4.1.3 illustrates the effect of temperature on the stability of Seprase. Stability of this mammalian enzyme would be expected at temperatures up to 37°C, and this was clearly the case. After 3 hours of incubation at 37°C, 90% of Seprase activity remained. It was found that Seprase remained stable up to 40°C during this time frame. Above this temperature, Seprase activity was found to decline rapidly. However, 60% of Seprase activity was found to remain after incubation at 50°C for 60 minutes. At physiological temperature Seprase seems to be activated, increasing activity.

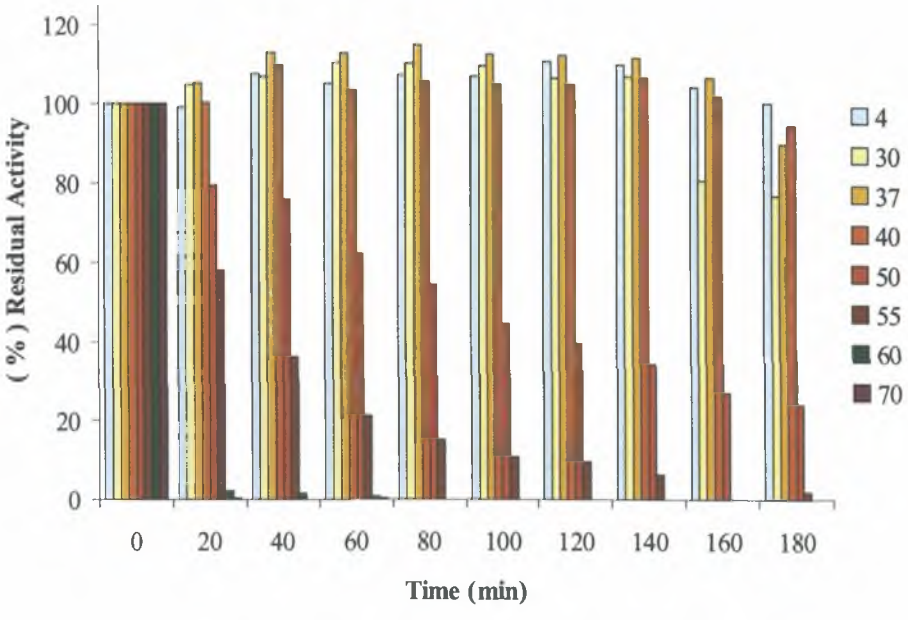


Figure 4.1.3 Determination of the Thermal Stability of Seprase

Dilute protein solutions are often unstable, due to adsorption to surfaces and dissociation of subunits. Relatively high thermostability can be attributed to the presence of other proteins 'protecting' the peptidase. These proteins increase interactions between water molecules, thus strengthening hydrophobic interactions. This explains why BSA is often added to purified enzyme samples (Harris and Angel, 1989). Progression of a purification scheme removes these stabilising proteins, thereby causing the protein of interest to become more heat-labile. Although the Seprase sample was extremely pure (13,805-fold, Table 3.3.1, section 3.3), there was no requirement for the addition of protein due to its excellent thermostability profile.

4 1 4 Thermal Inactivation of Seprase

Thermal inactivation of Seprase was carried out at 55°C (T_{50}), for up to 90mins Activity losses of 50% were obtained after 16 minutes of incubation as illustrated by Figure 4 1 4

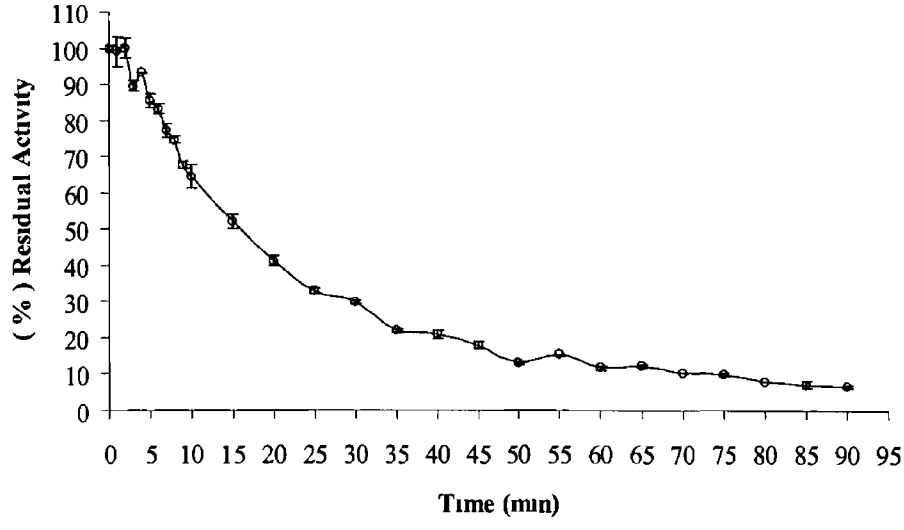


Figure 4 1 4 Thermal inactivation of Seprase

Error bars shown represent SEM of triplicate readings

The first 50 time points fitted satisfactorily to the first order process, showing that Seprase lost activity progressively according to first order kinetics (data were fitted using Enzfitter®) This allowed for an estimation of an apparent or pseudo half-life The first order rate constant of inactivation was determined to be 0 0436 ± 0 00143s^{-1} The half-life of this reaction was calculated using the equation $t_{\frac{1}{2}} = \frac{\ln 2}{k}$ to be 15 9s

It seems a pseudo-two-state unfolding mechanism is contributing to the loss of Seprase activity It seems that one inactivation process dominates over the first 40 minutes and then a second at longer time periods The active form of Seprase is a dimer and the loss of activity may proceed initially via the dissociation of the two subunits followed by their unfolding

4.2 Influence of pH

As with temperature, each enzyme has an optimal pH which is favourable to the native conformation for maximal activity. The pH effect results due to critical amino acids at the active site of the enzyme that participate in substrate binding and catalysis. The ionic or electric charge on the active site amino acids can enhance and stabilize interactions with the substrate. In addition, the ability of the substrate and enzyme to donate or receive a H^+ is affected by pH.

The influence of pH on the activity of Seprase was investigated as outlined in section 2.7.2. Figure 4.2 shows the pH profile for purified Seprase.

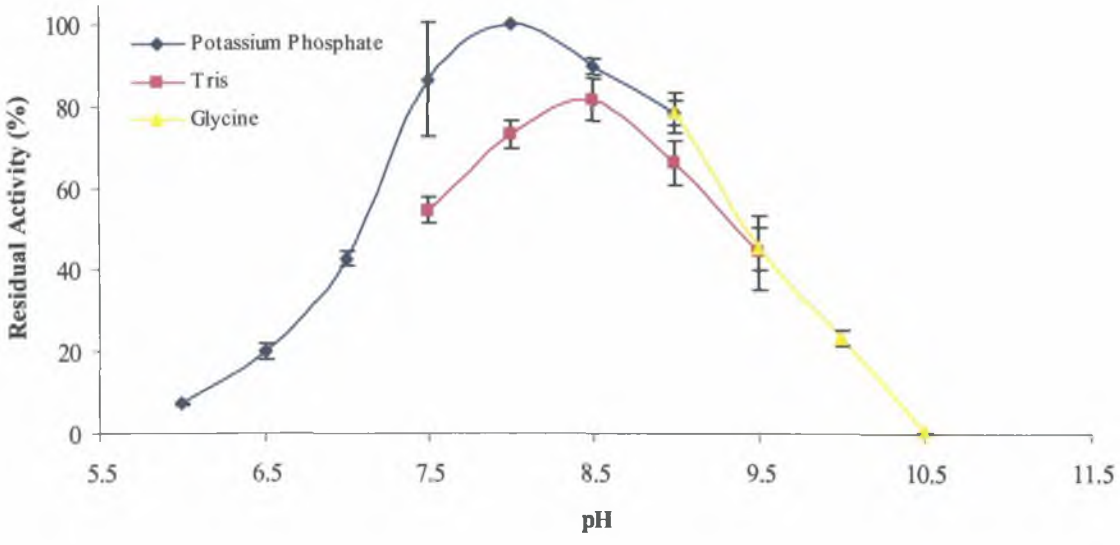


Figure 4.2 Influence of pH on Seprase activity

Residual enzyme activity over pH range 6.0 to 10.5. Different buffers used for specific ranges as described in Section 2.7.2. pH values of 8.0 were taken to be 100%. Error bars shown represent SEM of triplicate readings.

Seprase was found to remain stable over a pH range of 7.0 to 9.5 in the various buffering systems. It is clear from Figure 4.2 that the maximum Z-Gly-Pro-AMC hydrolysing activity occurred at pH 8.0 in the phosphate buffer system. By comparison, in the Tris buffer system, the relative activity was found to decrease by 27%. Previous studies have reported pH optima of Seprase to be 7.0, 8.5, 7.5 and 7.6 (see Table 1.4) (Aoyama and Chen, 1990; Birney and O'Connor, 2001; Kelly, 1999; Sun et al., 2002).

The phosphate buffering system was chosen for most assay and storage protocols The Tris buffer system was less effective in maintaining optimum enzyme activity and has also been found to interfere with a number of protein assays due to primary amine presence

4 3 Determination of the Second Order Rate Constant, (k_2) for DFP Inhibition of Seprase

The activity of Seprase is irreversibly inhibited by the serine protease inhibitor diisopropyl fluorophosphate, which specifically and stoichiometrically reacts with Seprase at Ser624, demonstrating that this residue represents the active site nucleophile Inhibition of the soluble form of Seprase resulted in an IC_{50} of 100nM (Collins et al , 2004)

In many reactions, the rate of reaction changes as the reaction progresses Initially the rate of reaction is relatively large, while after a period of time the rate of reaction decreases to zero (at which point the reaction is complete) In order to characterise the kinetic behaviour of a reaction, it is desirable to determine how the rate of reaction varies as the reaction progresses

A rate law is a mathematical equation that describes the progress of the reaction There are two forms of rate law for chemical kinetics the differential rate law and then integrated rate law

The differential rate law describes how the rate of reaction varies with the concentration of various species, usually reactants, in the system The rate of reaction is proportional to the rates of change in concentrations of the reactants and the products, that is, the rate is proportional to the derivative of the concentration Each rate law contains a constant, k, called the rate constant For the second-order reaction, the rate of reaction (r) is directly proportional to the square of the concentration of one of the reactants $[A]^2$

The differential rate law for a second order reaction is,

$$r = k[A]^2$$

$$-\frac{d[A]}{dt} = k[A]^2$$

And the integrated rate law,

$$[A] = \frac{[A]}{1 + kt[A]_o}$$

The rate constant of inhibition for DFP was determined under second-order conditions based on the equation,

$$\frac{1}{[A]} = \frac{1}{([A]_0 + k)t} \text{ or } \frac{1}{[A]} = kt + \frac{1}{[A]_o}$$

$$(y = mx + b)$$

with the molar ratio or Seprase to DFP at 1 1

The second-order rate constant k, has units of L mole ^{1} s^{-1} or M ^{1} s ^{1}

Equimolar amounts of Seprase and DFP were incubated according to section 2 7 3 for varying periods of time Fluorescent intensity was converted to [AMC] released using the calculations according to Appendix A The second order rate constant for DFP was determined using the plot of 1 / [AMC] released (μmol) ^{1} L versus Time (min)

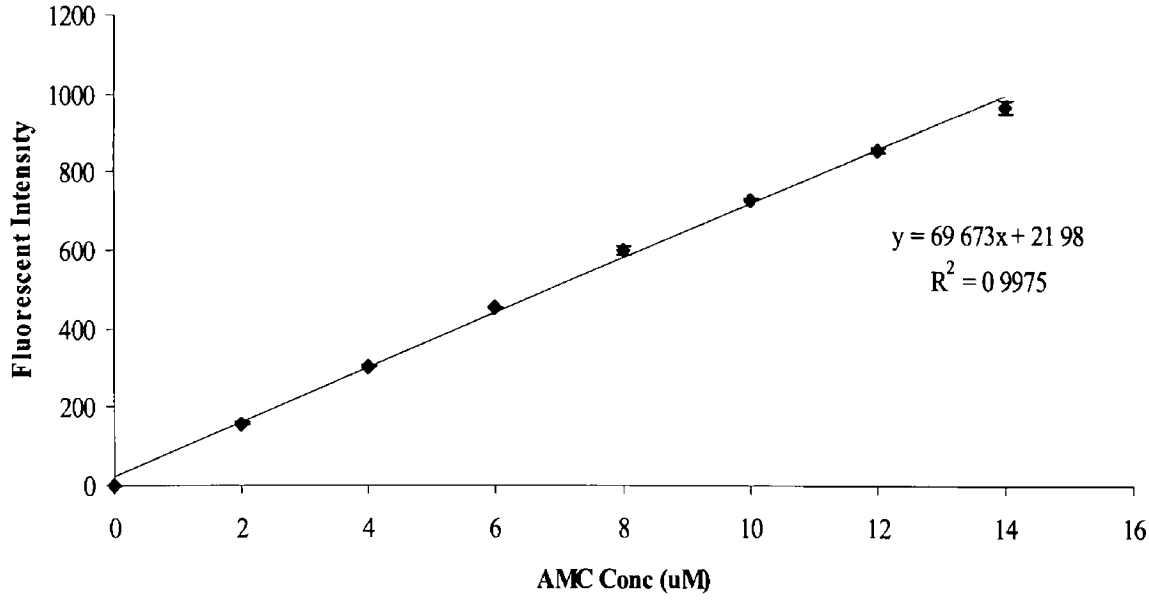


Figure 4.3 1 Standard Curve for DFP Analysis
Error bars shown represent SEM of triplicate readings

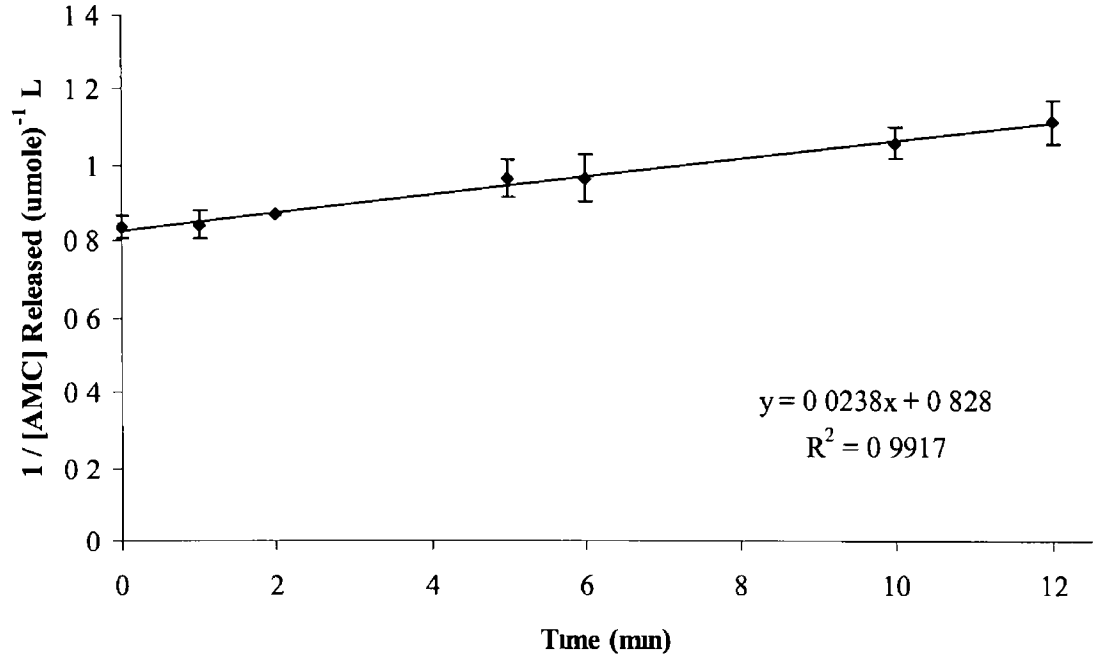


Figure 4 3 2 k_2 determination for DFP Inhibition of Seprase
Error bars shown represent SEM of triplicate readings

The second order rate constant (k_2), of inhibition of DFP against Seprase was calculated, from the slope of the line, to be 3 31x10^3 $M^{-1}s^{-1}$ This kinetic constant indicates very good inhibition of Seprase

The second order rate constant, k_2, determined for DFP inhibition of porcine brain prolyl endopeptidase (PO) was estimated to be 825 $M^{-1}s^{-1}$ (Rennex et al, 1991) Seprase displays a 4-fold higher sensitivity to DFP inhibition to that obtained for PO This indicates strongly that although both Seprase and PO cleave the substrate Z-Gly-Pro-AMC, the substrate specificity of both enzymes is different

Recent studies of dipeptide proline diphenyl phosphonates found that Gly-$Pro^{P}(Oh)_2$ exhibited an overall second order rate constant of inactivation of 177 $M^{-1}s^{-1}$ and 8 7 x 10^3 $M^{-1}s^{-1}$ against Seprase and DPPIV respectively (Gilmore et al, 2006) The data from this study suggests that both enzymes have a similar inhibition profile, and there exists subtle differences in their binding affinities This irreversible inhibitor displays an 18 7 fold lower inactivation rate of Seprase than DFP However, DFP is an extremely toxic serine protease inactivator and it is understandable that Seprase would have an increased sensitivity to it

4 4 Inhibition Studies

Serine protease inhibitors were provided by Dr Peter Kenny, School of Chemical Sciences, DCU They were tested for inhibitory properties using purified Seprase (section 2 4) and Prolyl Oligopeptidase (section 2 7 4 1) from bovine serum

4 4 1 Partial Purification Procedure for Bovine Serum Prolyl Ohgopeptidase

Prolyl Ohgopeptidase was partially purified from bovine serum as outlined in section 2 7 4 1 This enabled comparative inhibitory studies with Seprase Enzyme assays for Prolyl Ohgopeptidase were performed as described in section 2 3 2, with the addition of DTT in the substrate preparation as outlined in section 2 3 1

4 4 1 1 Phenyl Sepharose Hydrophobic Interaction Chromatography I

The first step in Prolyl Ohgopeptidase purification from bovine serum is the same step discussed in section 3 3 1 for the isolation of Seprase This step is extremely important for both purification procedures in that it critically separates the two Z-Gly-Pro-AMC hydrolysing activities found in bovine serum, from each other Figure 4 4 1 shows the elution profile of PO activity from Phenyl Sepharose column, run according to section 2 7 4 1 The first activity peak, which is present in the run through and is inhibited by JTP-4819, is Prolyl Ohgopeptidase

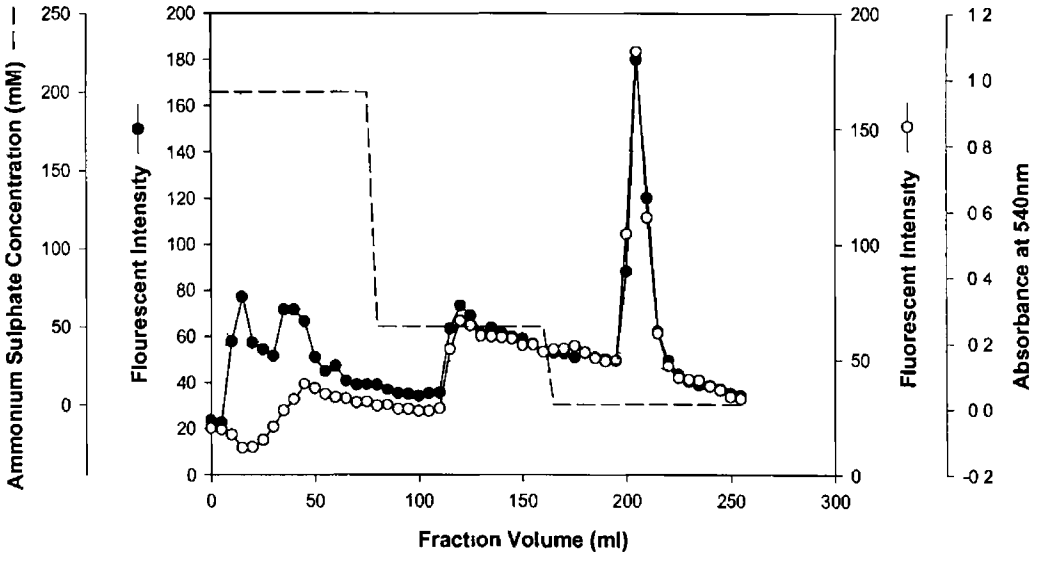


Figure 4 4 1 Elution profile of PO activity during Phenyl Sepharose Hydrophobic Interaction chromatography I

Elution profile shows the differentiation of the two distinct Z-Gly-Pro-AMC degrading activities (•) in bovine serum by the inclusion of JTP-4819 (○)

Fractions 3-9 were pooled yielding 30ml post phenyl sepharose I PO sample This sample was analysed for quantitative PO activity (section 2 3 2) and quantitative protein determinations (section 2 1 1)

83% of PO activity was retained with this purification step, most likely due to the fact that the enzyme did not bind and was present in the run through The co-elution of contaminating protein (793mg) explains the poor purification factor of 2 16-fold obtained

4 4 1 2 Phenyl Sepharose Hydrophobic Interaction Chromatography II

The first chromatography step in the purification of Prolyl Oligopeptidase was applied to separate Seprase and PO A second hydrophobic interaction column was employed to further purify PO A higher ammonium sulphate concentration (1M) was used during this second phenyl Sepharose column for binding of Prolyl Oligopeptidase Figure 4 4 2 illustrates the elution profile of PO, during a linear decreasing ammonium sulphate concentration gradient PO seems to leak off the column, prior to the elution gradient This was possibly due to the inefficient equilibration of the hydrophobic column

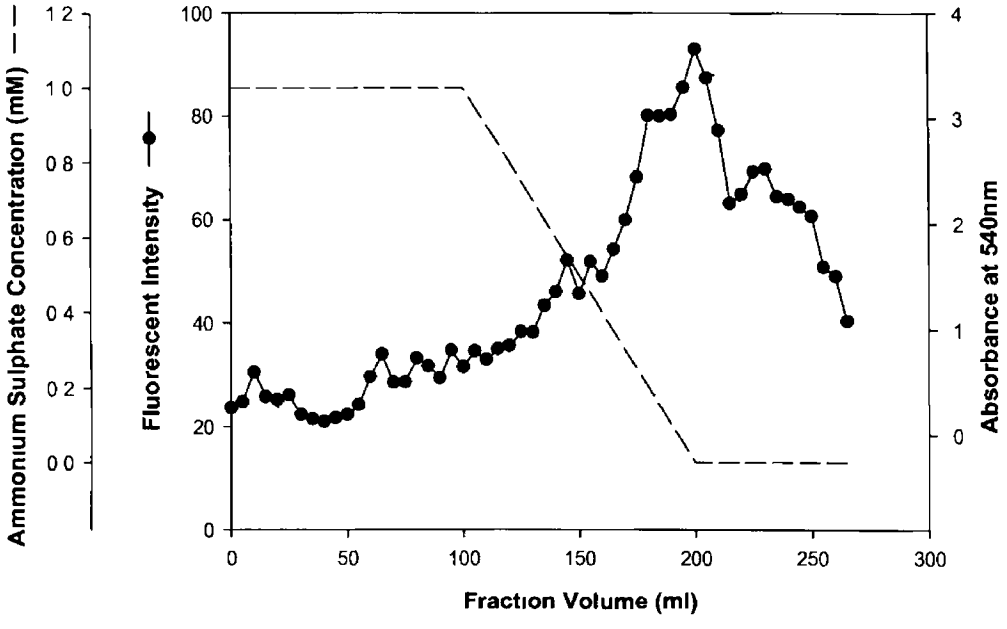


Figure 4 4 2 Elution profile of PO during purification using Phenyl Sepharose Hydrophobic Interaction Chromatography II

Fractions 37-42 were pooled yielding 35ml post Phenyl Sepharose II PO sample This sample was analysed quantitatively for PO activity (see section 2 3 2) and protein (see section 2 1 2)

Table 4 4 1 shows a loss of 99% of the total applied protein This was an extremely significant loss of protein, which resulted in a 67% decrease in the total applied enzyme activity This second purification step resulted in a 58-fold purification factor It is possible that, along with this removal of contaminating protein, as well as the concentration and addition of ammonium sulphate to the post-phenyl Sepharose I sample, has led to 32% of the original biological activity remaining Compared to the Seprase purification scheme, after two purification steps, 60% of biological activity remained (see Table 3 3 1) Another possible reason for the loss of PO activity is that it has been reported to be rather unstable during purification (Cunningham and O'Connor, 1997, Dowling, 1998)

4 4 1 3 Cibacron Blue 3GA Chromatography

This resin was employed as in the Seprase purification to remove BSA present in the sample (section 2 7 4 1 3) and to purify PO to a greater extent

Concentrated and dialysed post Phenyl Sepharose II PO was applied to the Cibacron Blue 3GA column Prolyl Oligopeptidase did not bind to the Cibacron Blue 3GA and was present in the run through as illustrated in Figure 4 4 3 This is the total opposite of Seprase, which was bound to this resin, as shown in section 3 3 3

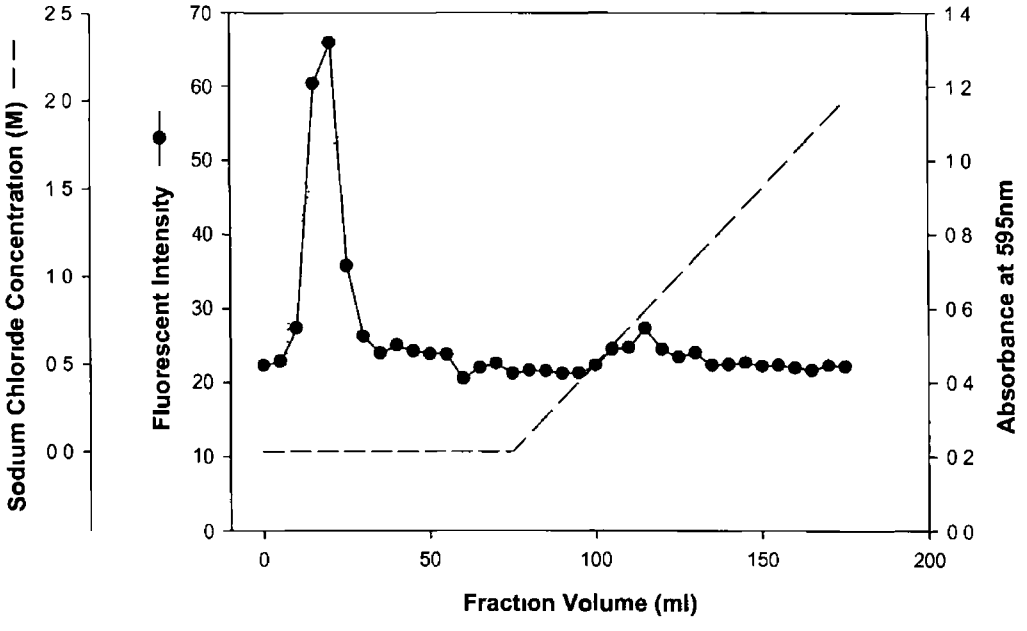


Figure 4 4 3 Cibacron Blue 3GA Chromatography of PO Activity

Fractions 4-6 were pooled yielding 9ml post-Cibacron Blue 3GA PO sample This sample was analysed for quantitative PO activity (section 2 3 2) and quantitative protein determinations (section 2 1 3)

This resin was successful in removing a further 11 1mg of contaminating protein resulting in an overall 96% reduction in protein A 313-fold purification factor along with 6 4% retention of biological activity was obtained This was deemed satisfactory for a partial purification of Prolyl Oligopeptidase from bovine serum

Purification Step	Volume	Total Activity [a]	Total Protein	Specific Activity	Purification Factor	Recovery
	ml	*Units*	*mg*	*Units/mg*		*%*
Crude Serum	25	16 145	2052 61	0 0079	1	100
Phenyl Sepharose I	30	13 493	793 52	0 0170	2 16	83 57
Phenyl Sepharose II	35	5 258	11 53	0 4561	58	32 57
Cibacron Blue	9	1 033	0 42	2 463	313	6 4

Table 4 4 1 Partial Purification of Prolyl Oligopeptidase from Bovine Serum

The purification table was constructed to assess the overall effectiveness of the purification procedure

[a] Based on the enzymatic activity using 100µM Z-Gly-Pro-AMC, where Units = nmol min^{-1}, i e Units are expressed as nanomol of AMC released per minute at 37°C

4 4 2 Inhibitor Analysis

It can be seen from Figures 4 4 5, 4 4 6 and 4 4 7 that the inhibitor compounds tested (as described in section 2 7 4 2) had little or no inhibitory effect on Seprase and PO A

activities over the concentration range used Compound 8h (Figure 4 4 5 and Table 4 4 2) showed some signs of inhibitory effects by reducing Seprase activity to 53% activity, while Prolyl Oligopeptidase activity was reduced to 85%

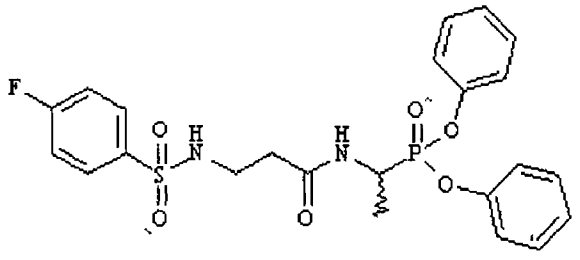


Figure 4 4.4 Structure of 8h Inhibitor

Due to the lack of inhibition, the IC_{50} value of each inhibitor for Seprase/PO could not be determined A higher concentration of inhibitor would be required for this analysis If, however, a higher concentration of inhibitor compound were to be tested, a higher concentration of DMSO would be required to maintain the inhibitor's solubility However the concentration of DMSO in the reaction had to be kept at a minimum in order to retain enzyme activity If further work were planned on testing these dipeptides, it would be required to make the inhibitor compounds more water-soluble

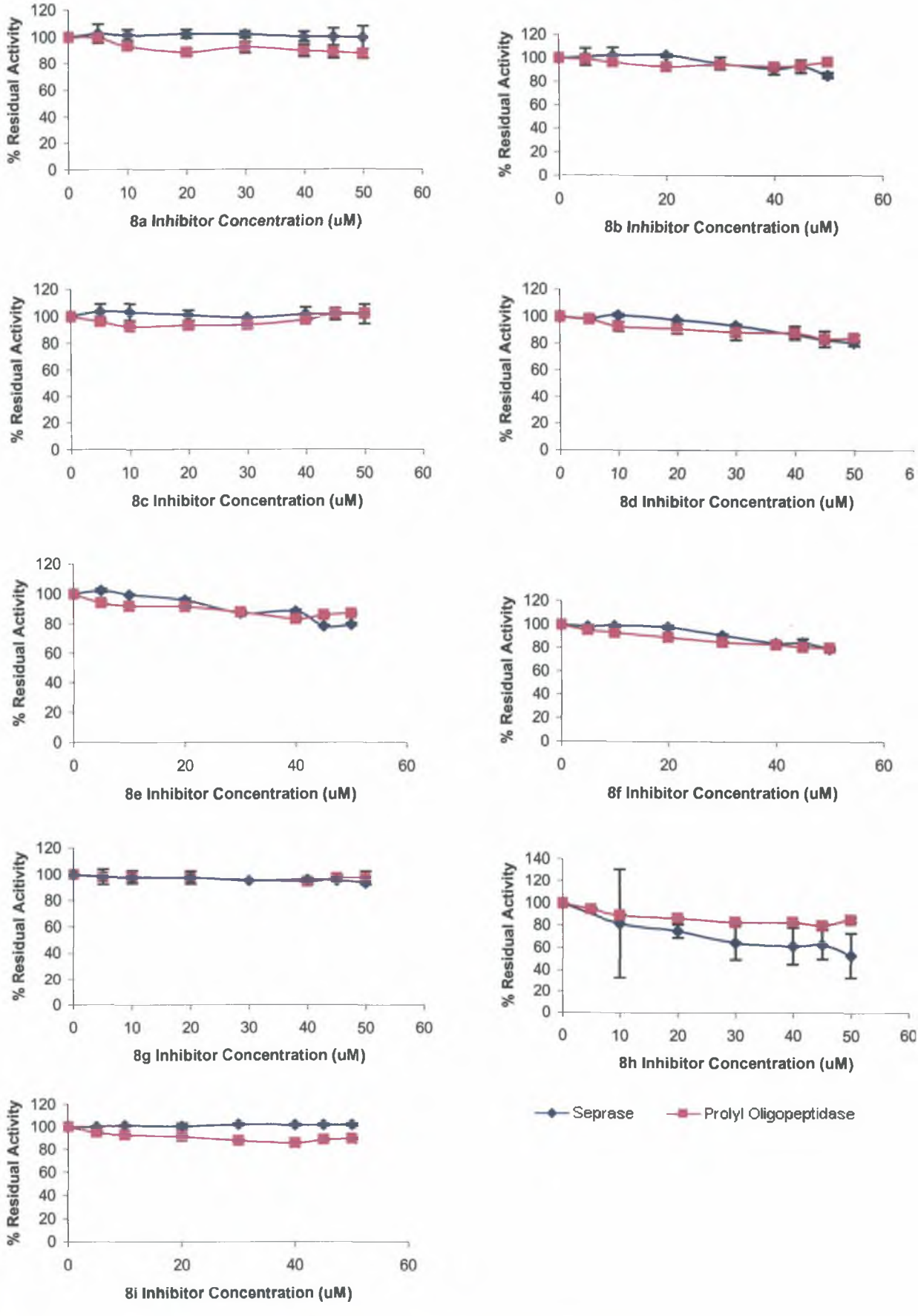


Figure 4.4.5 Inhibitor Profile Effect of the Compound 8 Series

Enzyme activities are expressed as a percentage of uninhibited enzyme. Error bars shown represent SEM of triplicate readings.

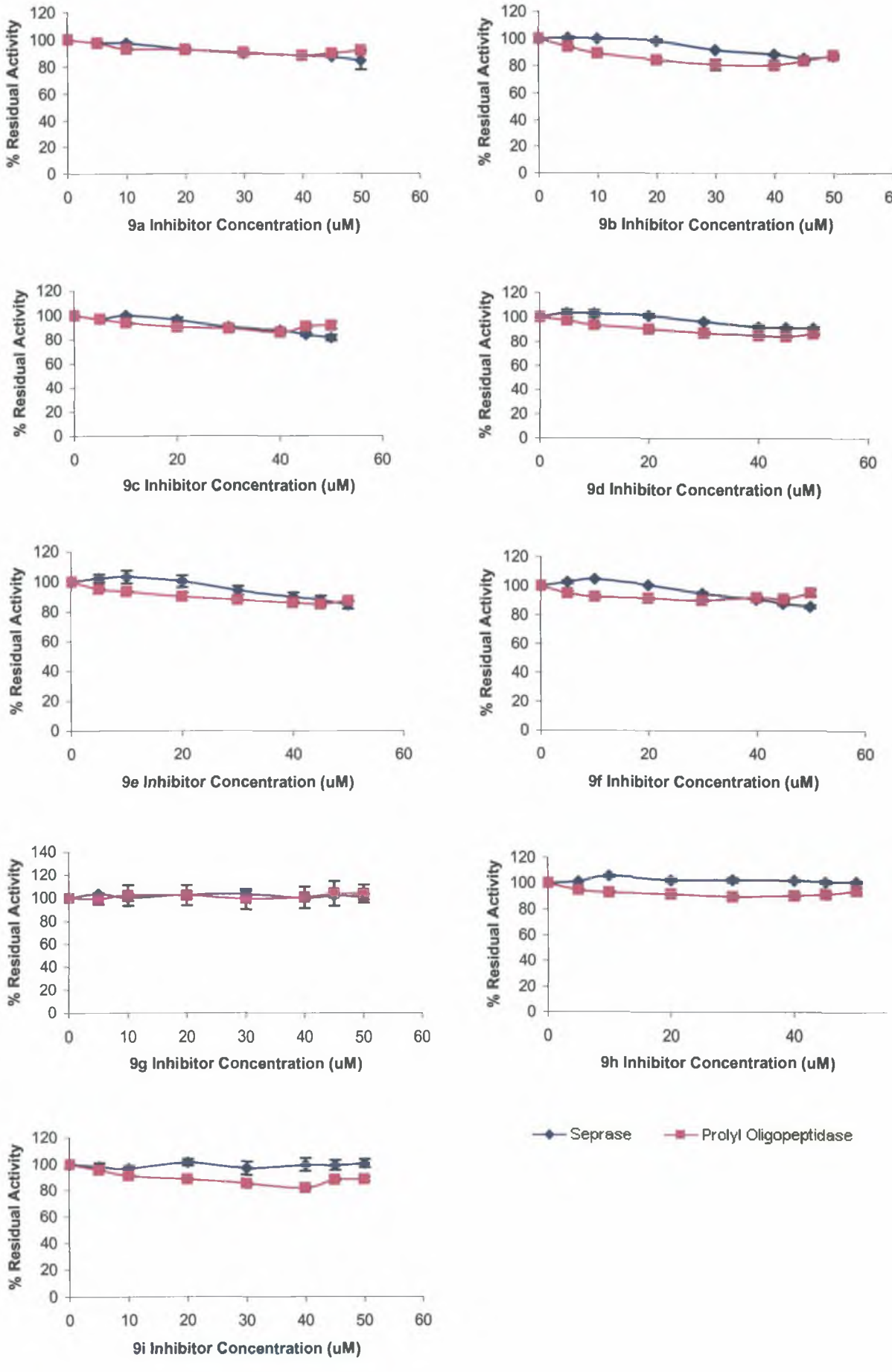


Figure 4.4.6 Inhibitor Profile Effect of the Compound 9 Series

Enzyme activities are expressed as a percentage of uninhibited enzyme. Error bars shown represent SEM of triplicate readings.

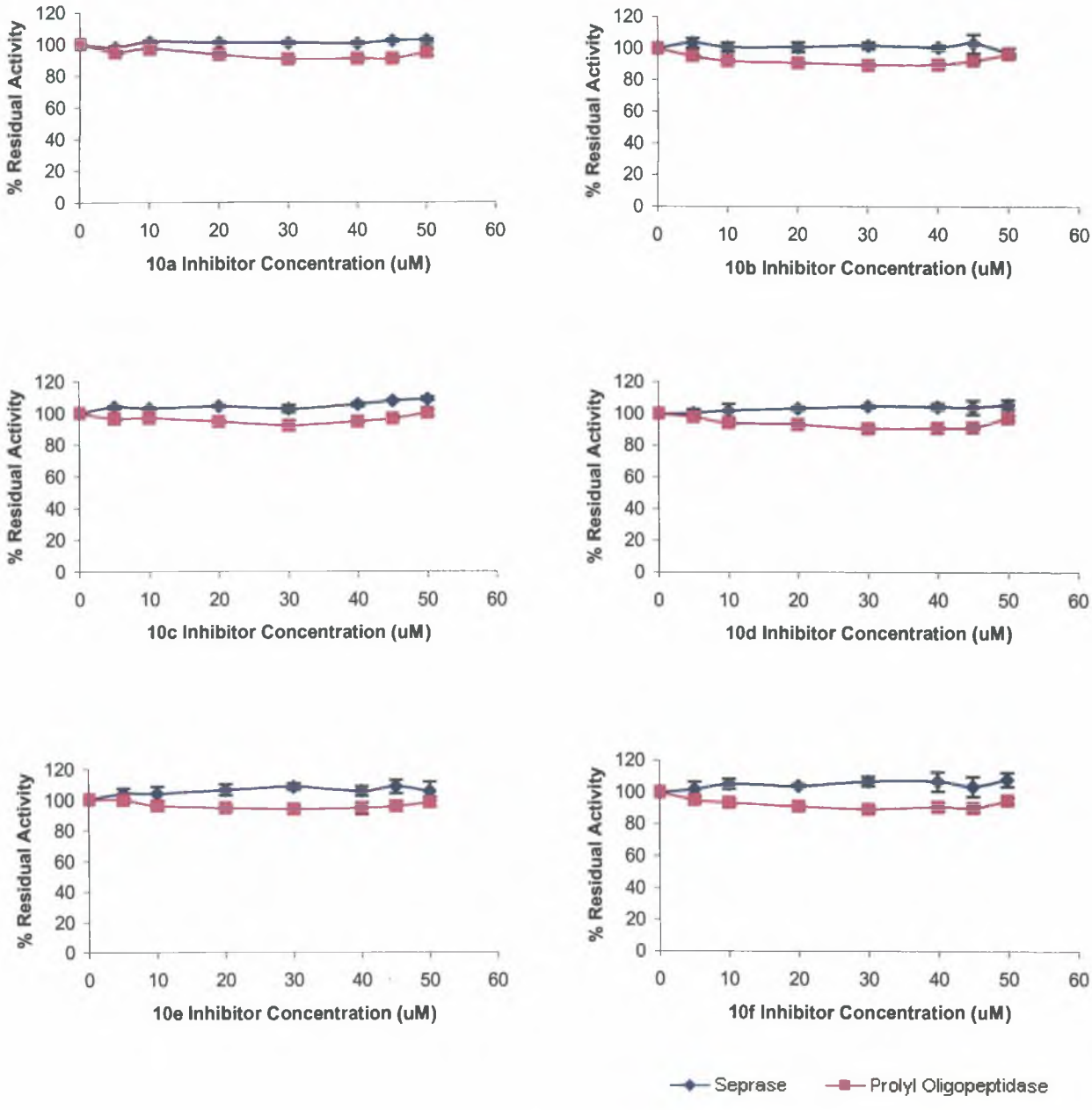


Figure 4.4.7 Inhibitor Profile Effect of the Compound 10 Series

Enzyme activities are expressed as a percentage of uninhibited enzyme. Error bars shown represent SEM of triplicate readings.

Compound	*% Residual Activity*		Compound	*% Residual Activity*		Compound	*% Residual Activity*	
	Seprase	**PO**		**Seprase**	**PO**		**Seprase**	**PO**
8a	99	87	**9a**	85	92	**10a**	102	95
8b	85	96	**9b**	87	87	**10b**	97	96
8c	101	101	**9c**	81	91	**10c**	109	100
8d	80	84	**9d**	91	87	**10d**	106	97
8e	79	87	**9e**	85	87	**10e**	105	98
8f	79	80	**9f**	86	95	**10f**	108	95
8g	97	93	**9g**	101	103			
8h	*53*	*85*	**9h**	101	94			
8i	102	90	**9i**	101	89			

Table 4 4 2 Overall Inhibitory Studies on Seprase and Prolyl Oligopeptidase

Enzyme activities are expressed as a percentage of uninhibited enzyme Residual Activities (%) are shown for the effect of each compound (50μM) on both Seprase and PO

Chapter 5

Substrate Specificity Studies

Substrate specificity studies of an enzyme are extremely important and can ultimately lead to an understanding of the possible physiological significance of an enzyme It is known that the substrate specificity of a protease is not just determined by the two amino acid residues at either side of the scissile bond of a substrate (S_1 and S_1') but can also be influenced by the other factors such as peptide length and conformation

Initial work on the substrate specificity of Seprase (Birney and O'Connor, 2001) studied a number of proline containing peptides as potential substrates Collins and colleagues (2004) determined that bovine Seprase releases Z-Gly-Pro from the synthetic substrate Z-Gly-Pro-AMC with high efficiency However, no cleavage of Gly-Pro-AMC or Pro-AMC occurs, indicating intolerance at the P_1 and P_2 -sites for free α-amine groups (see Figure 5 1 1 for nomenclature) Recent studies have shown that Seprase tolerates other N-blocking groups such as N-methyl-, formyl, and succinyl- (Edosada et al , 2006b)

Hydrolysis of Z-Gly-Pro-AMC by Seprase was profiled in the presence of proline-containing peptides with variations in the amino acid residues located at the *C*-terminal site of the scissile bond (position P_1') in addition to variations in conformation and length All peptides contained the constant sequence of Z-Gly-Pro-, which inhibited Seprase either in a mixed or non-competitive mode The lowest K_i^{app} values were obtained for peptides Z-Gly-Pro-Phe and Z-Gly-Pro-Met Highest K_i^{app} values occurred in substrates with His, Glu and Ala in the P_1' position Kinetic analysis of Seprase revealed greatest affinity for large hydrophobic residues in the P_1' position, and reduced affinity for acidic, basic and small amino acids in this position (Collins et al , 2004) More detailed substrate specificity studies were carried out on purified Seprase and were performed as outlined in section 2 8

5 1 Substrate Specificity using Combinatorial Dipeptide Library

Dr Craik's laboratory developed a method for rapid and general profiling of protease specificity using combinatorial fluorogenic libraries Initially it was proposed to use a positional scanning library that would analyse $P_1 - P_4$ position However, after numerous attempts, no results were obtained from this analysis It was then decided to use the positional scanning synthetic combinatorial dipeptide substrate library This dipeptide library was synthesised using the bifunctional fluorogenic leaving group 7-amino-4-carbamoylmethylcoumarin (ACC) similar to the tetrapeptide library constructed by Harris and colleagues (Harris et al , 2000) However, *N*-terminal acetylation of the P2 residue produced a library with the general composition, Ac-Xaa-Xaa-ACC (Figure 2 8), where Xaa is any amino acid Amide bond hydrolysis occurs between the Xaa-Xaa dipeptide and ACC, resulting in an increase in fluorescence

This library was composed of two sub-libraries (P_1 and P_2) consisting of 400 compounds each, utilizing the 20 naturally occurring amino acids, but excluding cysteine and including norleucine The P_1 sub-library was composed of 20 separate mixtures in which the P_1 position was fixed and the P_2 position contained an equimolar mixture of all of the other amino acids In the P_2 sub-library, the P_2 residue is positionally defined, and the P_1 position contains an equimolar mixture of all other amino acids The results using these libraries provide a complete understanding of the specificities of Seprase in S_1 and S_2 subsites of the active site

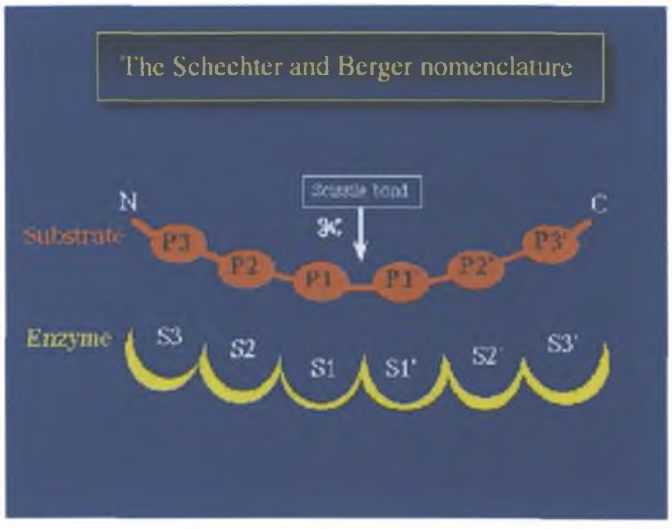


Figure 5.1.1 Schematic Representation of an Enzyme-Substrate Complex

Residues of the substrate and enzyme subsites are numbered according to their distance from the scissile bond and their location on the N- or C-terminal side of the cleavage site (Schechter and Berger, 1967).

In this study, positional scanning of the P_1 ACC library determined that Seprase has a marked preference for proline in the S_1 subsite, but appears to tolerate all other amino acids with the exception of tryptophan (Figure 5.1.2). When fluorescence was normalized as a percentage of P_1-proline released (i.e. Ac-Xaa-Pro-ACC = 100%), ACC was not released from peptides with amino acids other than proline in the P_1 position at a rate greater than 24%. Lowest activity was observed with P_1 aromatic residues, such as phenylalanine, tyrosine and tryptophan. These results indicate that the S_1 subsite of Seprase is therefore designed to specifically fit proline residues. A recent study based on the cleavage site in α_2AP_{PRO}, confirmed that Seprase requires a Proline in the P_1 position (Edosada et al., 2006a).

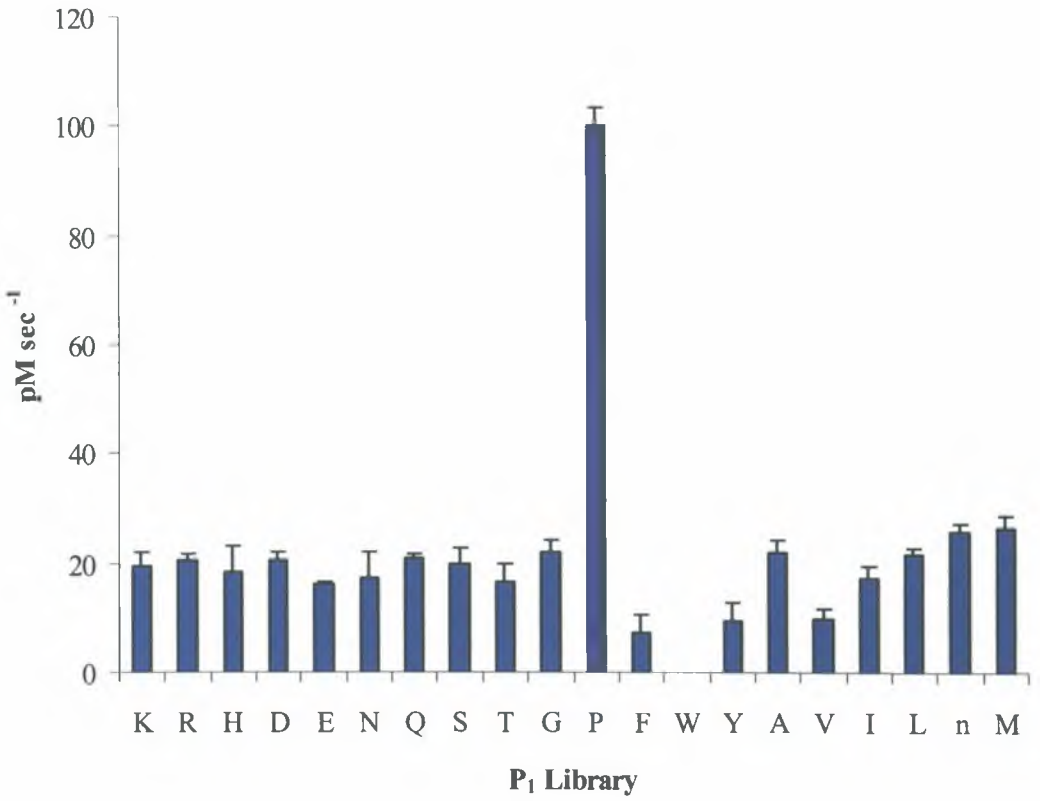


Figure 5.1.2 P_1 Positional Scanning Profile for Seprase

Positional scanning of bovine Seprase using a combinatorial dipeptide library assayed at the P_1 position. The X-axis gives the standard single letter code for amino acids where n = norleucine. Amino acids are arranged in the order of basic (K,R,H), acidic (D,E), polar (N,Q,S,T), special (G,P), aromatic (F,W,Y) and aliphatic (A,V,I,L,n,M) groups. Free ACC fluorescence of 1 RFU sec^{-1} was calculated to be equivalent to 0.0007226 pM sec^{-1}. (100% scale)

Data obtained during screening of the P_2 sub-library revealed a much broader specificity in the S_2 binding pocket of bovine Seprase. However, the catalytic rates are relatively low in comparison to cleavage after P_1-proline. Comparing Figure 5.1.3 and 5.1.4 together, it appears that the initial reaction at the P_2 position for each amino acid may be related to the presence in each well of P_1-Pro ($1/20^{th}$ of each substrate). Therefore, even if the enzyme does not favour a particular amino acid at P_2, cleavage will still occur due to the presence of P_1-Pro in the well. After the first 12 minutes, P_1-Pro has been used up and Seprase then begins to cleave the other substrates with much lower processing (see Figure 5.1.4).

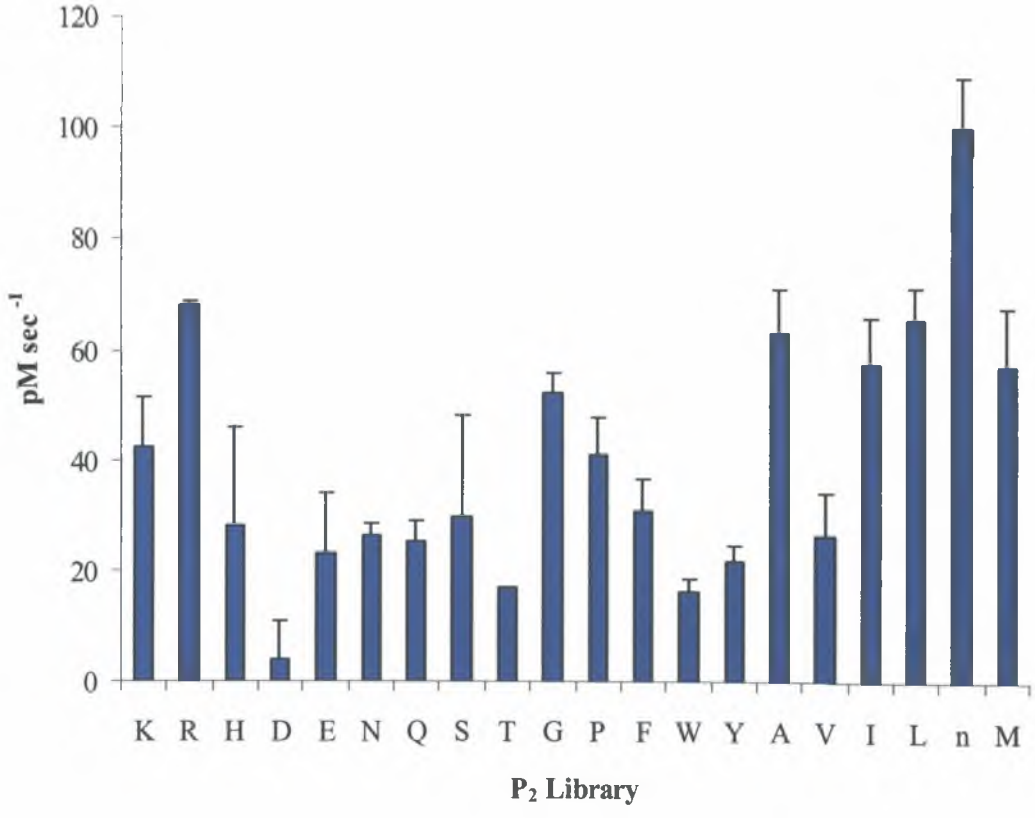


Figure 5.1.3 P_2 Positional Scanning profile from 0-60 minutes for Seprase

Positional scanning of bovine Seprase using a combinatorial dipeptide library assayed at the P_2 position. The X-axis gives the standard single letter code for amino acids where n = norleucine. Amino acids are arranged in the order of basic (K,R,H), acidic (D,E), polar (N,Q,S,T), special (G,P), aromatic (F,W,Y) and aliphatic (A,V,I,L,n,M) groups. Free ACC fluorescence of 1 RFU sec^{-1} was calculated to be equivalent to 0.0007226 pM sec^{-1}. (100% scale)

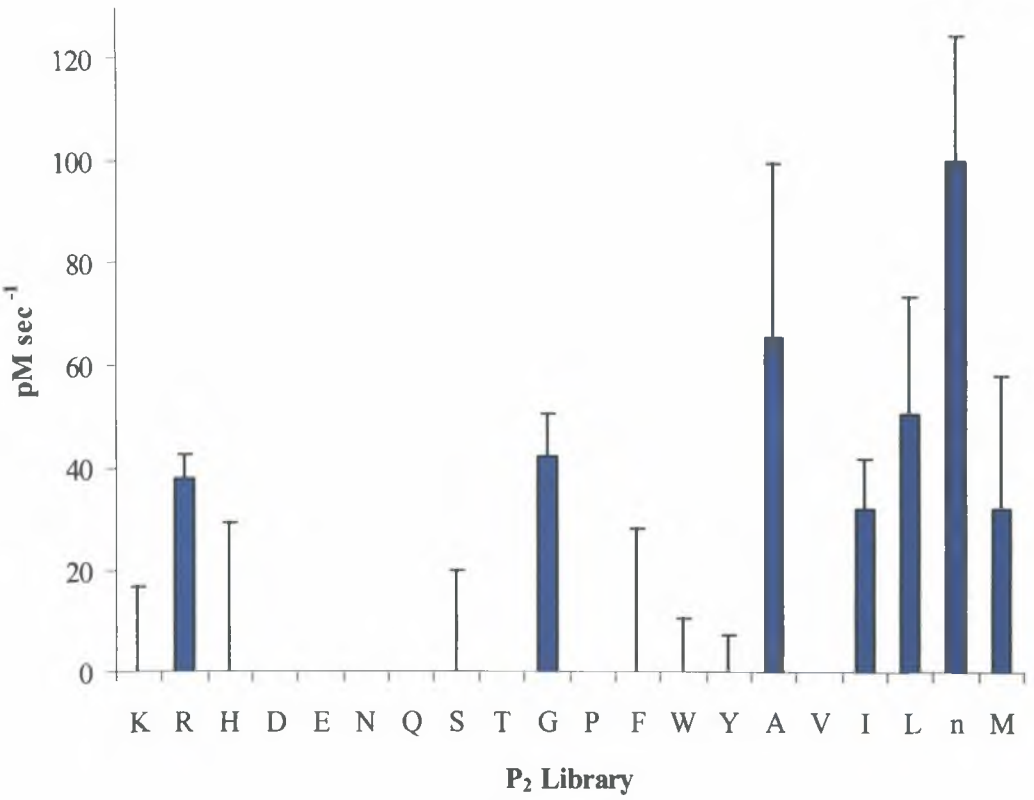


Figure 5.1.4 P_2 Positional Scanning profile from 12-60 minutes for Seprase

Positional scanning of bovine Seprase using a combinatorial dipeptide library assayed at the P_2 position. The X-axis gives the standard single letter code for amino acids where n = norleucine. Amino acids are arranged in the order of basic (K,R,H), acidic (D,E), polar (N,Q,S,T), special (G,P), aromatic (F,W,Y) and aliphatic (A,V,I,L,n,M) groups. Free ACC fluorescence of 1 RFU sec^{-1} was calculated to be equivalent to 0.0007226 pM sec^{-1}. (100% scale)

In the S_2 subsites, this study shows that Seprase has a preference for Norleucine (Nle), Alanine (Ala), Leucine (Leu), Glycine (Gly), Arginine (Arg), Methionine (Met), and does not tolerate aromatic, strongly basic or acidic residues. Using substrates that span the P_4-P_2' positions, recent studies showed that Seprase has a preference for Glycine at the P_2 position (Edosada et al., 2006a). Using longer peptides confers a certain amount of conformational restriction on the amino acid entering the S_2 site of the enzyme, as their positions will affect the position of the P_2 amino acid. Kinetic studies have shown that Seprase has a preference for Ala-Pro-AFC over Gly-Pro-AFC as a substrate with k_{cat}/K_m values of 5.8 x 10^4 $M^{-1}s^{-1}$ and 2.3 x 10^4 $M^{-1}s^{-1}$ respectively (Edosada et al., 2006b). However, Z-Gly-Pro-AMC has a 2-3-fold higher affinity for Seprase than does Ala-Pro-AFC (K_m 101μM and 323μM respectively) (see Table 1.8) (Lee et al., 2005a).

As discussed previously, Gly-Pro^{P}(Oh)$_2$ was most effective in inhibiting Seprase activity, exhibiting an overall second order rate constant of inactivation of 177 $M^{-1}sec^{-1}$ against Seprase (see Section 1.9.2) (Gilmore et al., 2006). This is in keeping with Seprase having a preference for P_2-Glycine. However, this group did not investigate the effect of Ala-Pro^{P}(Oh)$_2$ on Seprase activity. Seprase was also found to have a broader specificity at the P_4, P_3, P_1' and P_2' positions (see section 1.8) (Edosada et al., 2006a). Kinetic analysis of Seprase using a FRET peptide has shown this peptide to have high affinity and kinetic efficiency, indicating that residues in the P_4-P_4' region contribute to the substrate specificity (see Section 1.8) (Lee et al., 2005a).

A similar study to that reported in this work, used a P_2-Pro-AMCC library and an Ac-P_2-Pro-AMCC library to identify the peptide motifs for Seprase-selective inhibitor design (Edosada et al., 2006b). With the P_2-Pro-AMCC library, Seprase showed a preference for Ile, Pro and Arg at the P_2 position. By contrast, the second library Ac-P_2-Pro-AMCC found Seprase only cleaved Ac-Gly-Pro-AMCC. This library is of similar structure to that reported in this research thesis, however, the latter study shows that Seprase has a broader specificity at the P_2 position. The source of Seprase used in these studies may be a possible reason for the slight variation in the results. The study mentioned used recombinantly expressed human Seprase with an N-terminal FLAG tag, whereas our study used the novel bovine serum form of Seprase.

These results for Seprase can be compared to DPPIV (Figure 5 1 5), to which Seprase shows considerable sequence similarity at 52% (Goldstein et al , 1997) DPPIV also has a preference for Proline in the P_1 position However, Alanine is the next most preferred residue, albeit with greatly reduced efficiency DPPIV has a highly indiscriminate S_2 subsite, with all residues well tolerated (Leiting et al , 2003), provided that the terminal amino group is protonated (Gilmore et al , 2006) Inhibitory studies have shown, however, that the P_2 residue of the inhibitor has an effect on the second order rate constant (see Section 1 9 2) (Gilmore et al , 2006) Detailed comparison of Seprase and DPPIV revealed that the Ala^{657} residue in Seprase, instead of Asp^{663} as in DPPIV, reduces the acidity in this pocket, and this change could explain the lower affinity for N-terminal amines by Seprase (see Figure 1 4 4) (Aertgeerts et al , 2005) In contrast to Seprase, DPPIV has been shown to have a marked preference for a substrate with a free N-terminus (Edosada et al , 2006b)

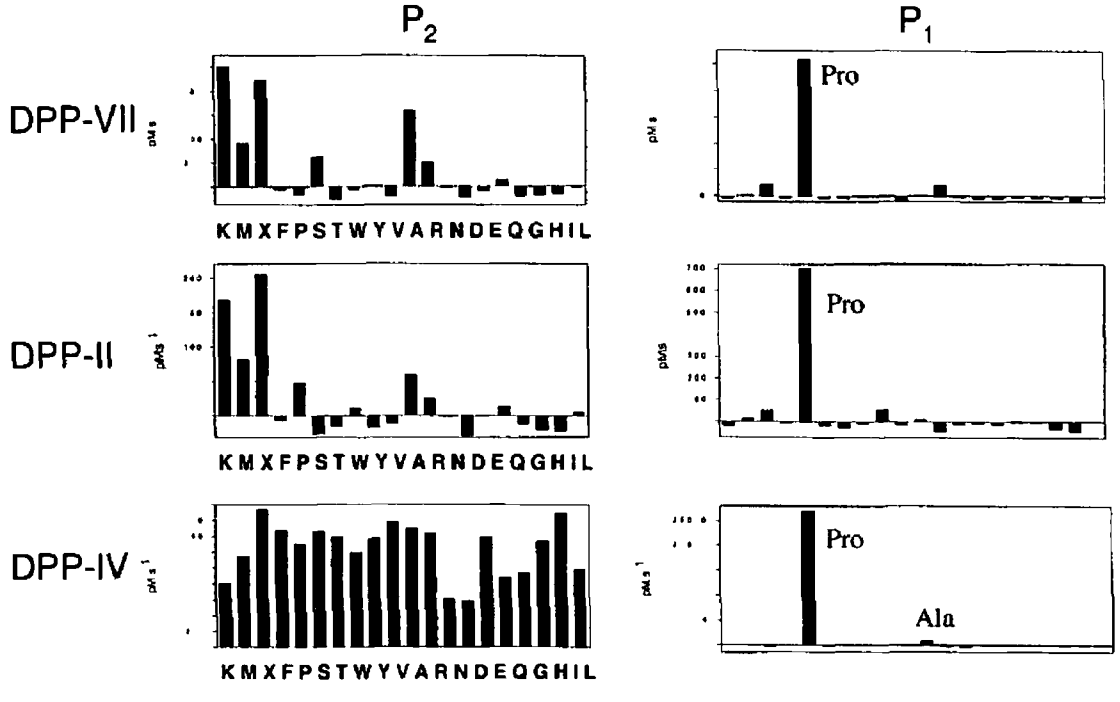


Figure 5 1 5 Positional Scanning combinatorial dipeptide library profiles

Positional scanning combinatorial dipeptide library data for recombinant Gly^{30} DPP-VII, human placenta DPP-II and human recombinant DPP-IV Amino acid residues are shown in single letter code X represents norleucine (Leiting et al , 2003)

5 2 Kinetic Analysis

Kinetic analysis can lead to a better understanding of the mechanism of action of an enzyme and the physiological importance of the enzyme in vivo Substrate specificity studies by Collins et al (2004) suggest that Seprase has an extended substrate-binding region in addition to the primary specificity site, S_1 It is possible that the substrate binding region is comprised of three sites located at the ammo-terminal site (S_1, S_2, S_3) and two sites at the carboxyl site from the scissile bond S_1' and S_2' Edosada et al (2006a) recently carried out specificity studies (P_4 – P_2) on Seprase They designed peptides based on the Serpase cleavage site of α-antiplasmin TSGP-NQ They determined that Seprase had a high affinity for the peptide ASGP-SS Using these studies and the results from the combinatorial dipeptide library, peptides were designed to fully profile the substrate-binding region of Seprase

5 2 1 K_m Determination for Z-Gly-Pro-AMC

The kinetic behaviour of purified Seprase was investigated using Z-Gly-Pro-AMC K_m values for the hydrolysis of the substrate Z-Gly-Pro-AMC were determined according to section 2 8 2 1 The resulting Michaelis-Menten curve for purified Seprase is shown in Figure 5 2 1 Data obtained was applied to the Lineweaver-Burk, Eadie-Hofstee and Hanes-Woolf kinetic models for analysis (Appendix A) These data were also fitted by linear regression using Enzfitter (a non-linear regression data analysis program) From these plots, the maximal enzyme velocity (V_{max}) and the Michaelis constant (K_m) were determined, which are given in Table 5 2 1 Also given in Table 5 2 1 is the turnover number (k_{cat}) for purified Seprase, which was determined as described in Appendix A High catalytic activity and high specificity for substrates, can be described kinetically by the constants such as k_{cat} and k_{cat} / K_m

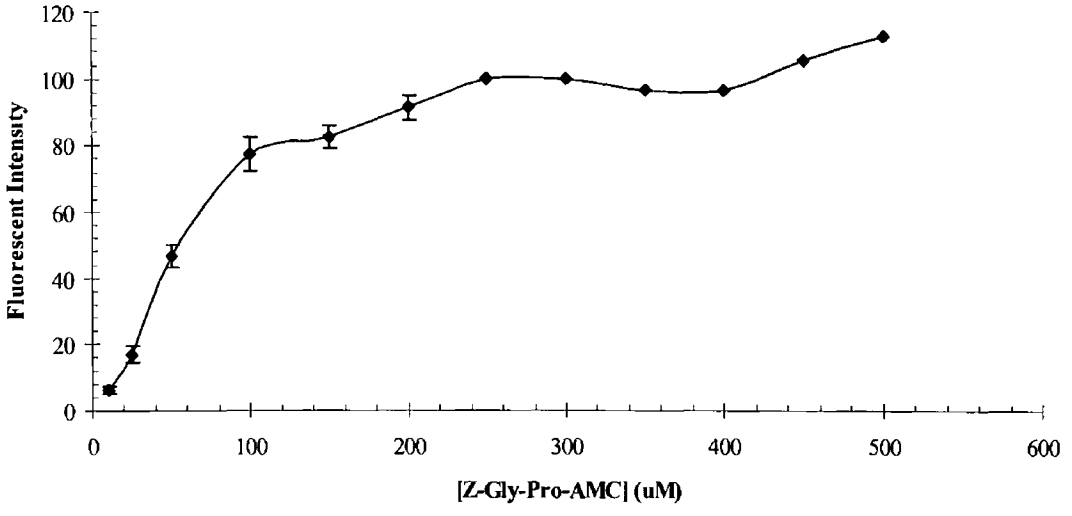


Figure 5 2 1 Michaelis-Menten curve for purified Seprase

Plot of substrate concentration [Z-Gly-Pro-AMC] (μM) versus reaction rate, represented by fluorescent intensity (FI), for purified Seprase Data obtained by the procedure described in Section 2 8 2 1 Error bars shown represent SEM of triplicate readings

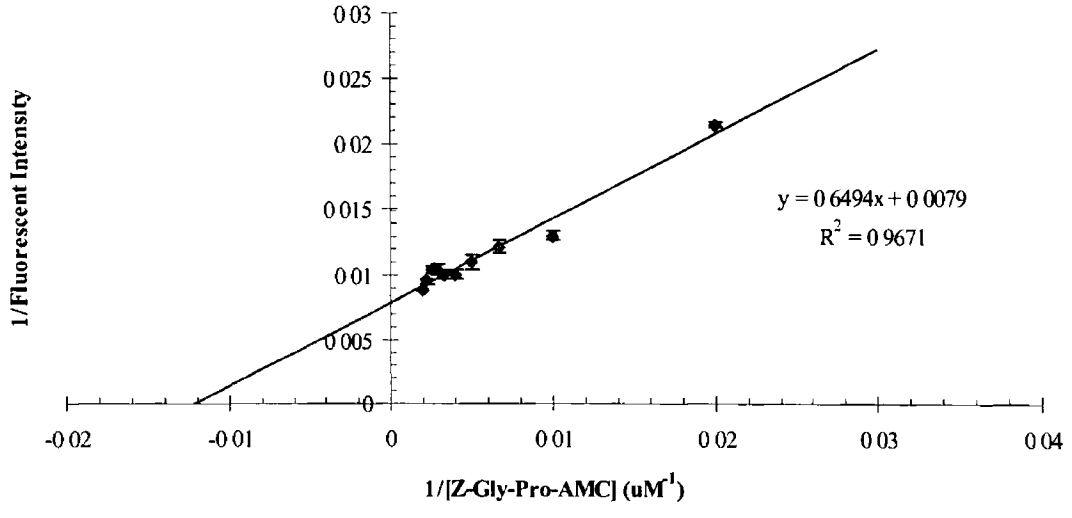


Figure 5 2 2 Lineweaver-Burk plot for purified Seprase

Data from Figure 5 2 1 fitted to the Lineweaver-Burk model for the determination of kinetic parameters, as outlined in Appendix A Equation and regression are shown K_m and V_{max} values given in Table 5 2 1 Error bars shown represent SEM of triplicate readings

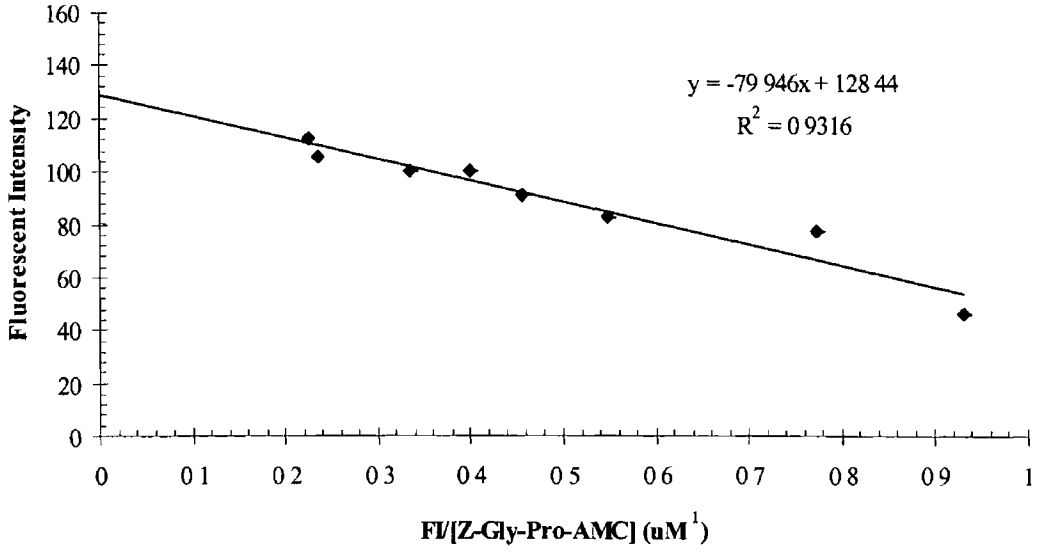


Figure 5 2 3 Eadie-Hofstee plot for purified Seprase

Data from Figure 5 2 1 fitted to the Eadie-Hofstee model for the determination of kinetic parameters, as outlined in Appendix A Equation and regression are shown K_m and V_{max} values given in Table 5 2 1 Error bars shown represent SEM of triplicate readings

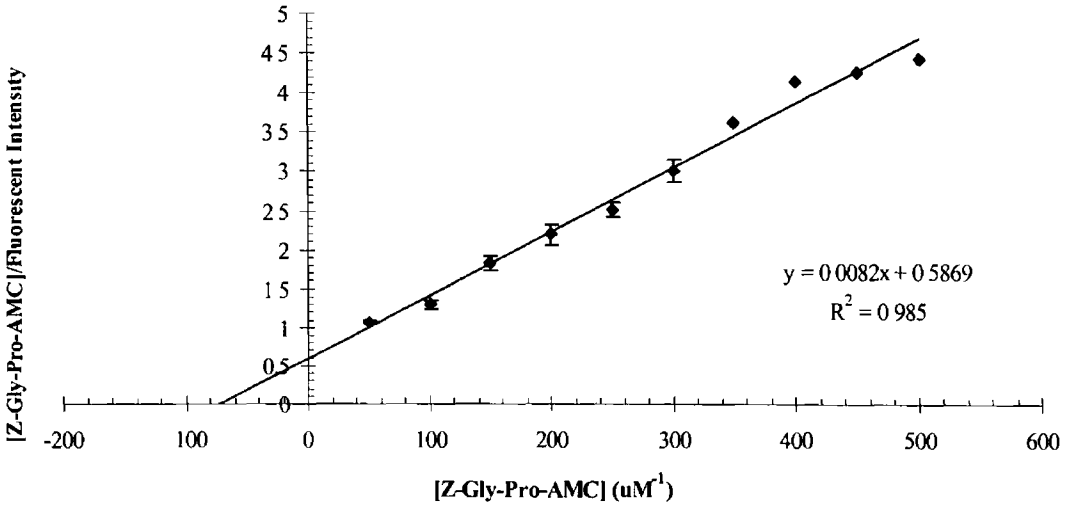


Figure 5 2 4 Hanes-Woolf plot for purified Seprase

Data from Figure 5 2 1 fitted to the Hanes-Woolf model for the determination of kinetic parameters, as outlined in Appendix A Equation and regression are shown K_m and V_{max} values given in Table 5 2 1 Error bars shown represent SEM of triplicate readings

	Lineweaver-Burk	Eadie-Hofstee	Hanes-Woolf	Enzfitter	Average
K_m (μM)	83 33	79 946	65	100 13	82 10
V_{max} (Units/ml)	125	128	121 95	135	127 49
k_{cat} (sec^{-1})	8 823	9 035	8 608	-	8 822
k_{cat} / K_m ($M^{-1}s^{-1}$)	$1\ 06 \times 10^5$	$1\ 13 \times 10^5$	$1\ 32 \times 10^5$	-	$1\ 17 \times 10^5$

Table 5 2 1 Kinetic parameters for Seprase

K_m and V_{max} values obtained from Figures 5 2 2 to 5 2 4 k_{cat} values were determined according to Appendix A

The Michaelis-Menten curve (Figure 5 2 1) shows that up to about 80μM Z-Gly-Pro-AMC, the Seprase substrate concentration is limiting to the enzymatic reaction For this reason, the standard activity assay used in this work (Section 2 3 2) used 100μM Z-Gly-Pro-AMC, ensuring a substrate excess A higher concentration of substrate could also have been used, but it was deemed unnecessary

The kinetic data in Table 5 2 1 shows that relatively little variance was observed between the three kinetic models plotted in Figures 5 2 1 to 5 2 4 and that using Enzfitter The average K_m determined for purified Seprase was 82 10μM This value is ~1 5-fold higher than the K_m value of 54μM reported previously for Seprase by Birney and O'Connor (2001) and ~3 3-fold lower than 270μM reported by Collins et al (2004) This work, unlike those mentioned, used the microplate assay (section 2 3 2) to analyse these kinetic parameters This led to a decrease in experimental error due to lapses in time initiating/terminating each reaction and the subsequent analysis of samples Recently published results for human APCE and Seprase show a K_m of 101μM and 124μM toward Z-Gly-Pro-AMC respectively (see Table 1 8) (Lee et al , 2005a) These results are on average ~1 4-fold higher than the K_m obtained in this study for bovine Seprase

While the K_m is independent of enzyme concentration, V_{max} is not Since the concentration of purified Seprase was known, the turnover number (k_{cat}), the number of substrate molecules converted into product by one enzyme molecule per second, could be calculated as 8 822s^{-1} k_{cat} is a measure of the maximum potential catalytic activity of an enzyme The ratio of k_{cat} / K_m is the second order rate constant for the reaction of enzyme and substrate to form products Therefore, k_{cat} / K_m can be described as a specificity constant for an enzyme and provides a measure of how rapidly an enzyme can work at low substrate concentration Since an enzyme and substrate cannot combine more rapidly than diffusion permits, there is what is called a '*limit of efficiency*', an upper limit on enzyme catalysis The value of k_{cat} / K_m cannot be greater than 1 x 10^9 s^{-1} M^{-1} Those enzymes that have values that approach the diffusion limit indicate extreme efficiency in binding substrate and in converting it to a product Seprase was found to have an average k_{cat} / K_m ratio of 1 17 x 10^5 M^{-1}s^{-1} for cleavage of Z-Gly-Pro-AMC Other groups have reported values of 5 3 x 10^4 M^{-1}s^{-1} and 7 4 x 10^3 M^{-1}s^{-1} (see Table 1 8) (Aertgeerts et al , 2005, Edosada et al , 2006a) The substrate Z-Gly-Pro-AMC is however, not cleaved by DPPIV (Aertgeerts et al , 2005)

5 2 2 Competitive inhibition of Seprase by synthetic peptides

A ligand-induced change in the effective value of K_m is one way of regulating the activity of an enzyme By measuring the effects of different compounds on K_m, it is possible to identify physiologically important inhibitors and activators The K_i^{app} values obtained in this study will be regarded as the affinity constants of the enzyme Seprase for that particular peptide

The competitive influence of selected tri- and tetra-peptide substrates (Table 2 8) on the Z-Gly-Pro-AMC degrading activity of Seprase was investigated as described in Section 2 8 2 2 The resulting data sets were applied to the aforementioned Lineweaver-Burk, Eadie-Hofstee and Hanes-Woolf kinetic models for analysis (Appendix A) From these plots the apparent K_m (K_m^{app}) and V_{max} (V_{max}^{app}) were determined in the presence of each peptide These K_m^{app} values were used to calculate the dissociation constants (K_i^{app}) as outlined in Appendix A, which are given in Table 5 2 2 For example the K_m^{app} for Z-Gly-Met-Phe, calculated from the kinetic model in Figure 5 2 5, was 95 26μM Using the formula described in Appendix A the K_i^{app} was calculated to be 1247 72μM

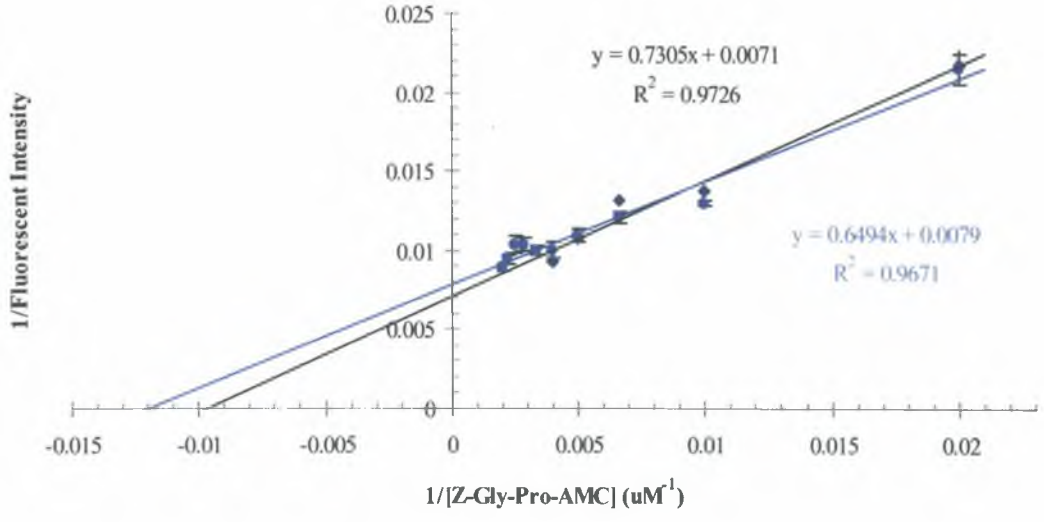


Figure 5.2.5 K_i^{app} determination for Z-Gly-Met-Phe using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the mixed inhibition of Seprase by Z-Gly-Met-Phe. Error bars shown represent SEM of triplicate readings.

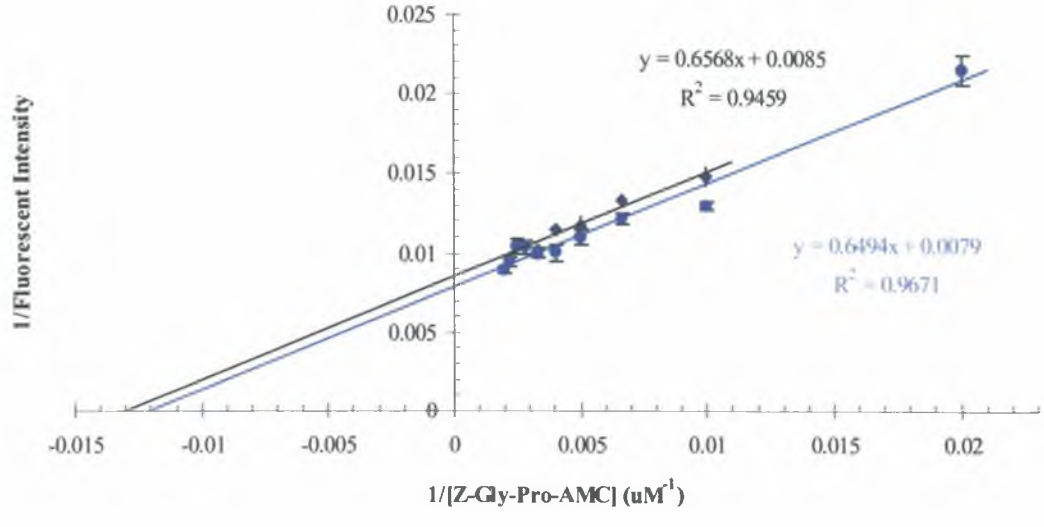


Figure 5.2.6 K_i^{app} determination for Z-Gly-Pro-Phe using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the uncompetitive inhibition of Seprase by Z-Gly-Pro-Phe. Error bars shown represent SEM of triplicate readings.

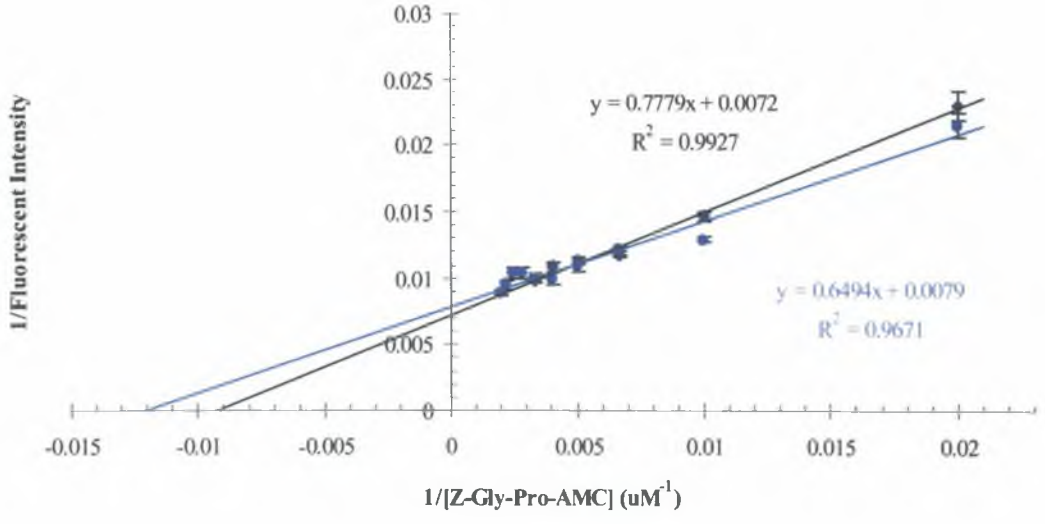


Figure 5.2.7 K_i^{app} determination for Z-Ala-Pro-Phe using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the mixed inhibition of Seprase by Z-Ala-Pro-Phe. Error bars shown represent SEM of triplicate readings.

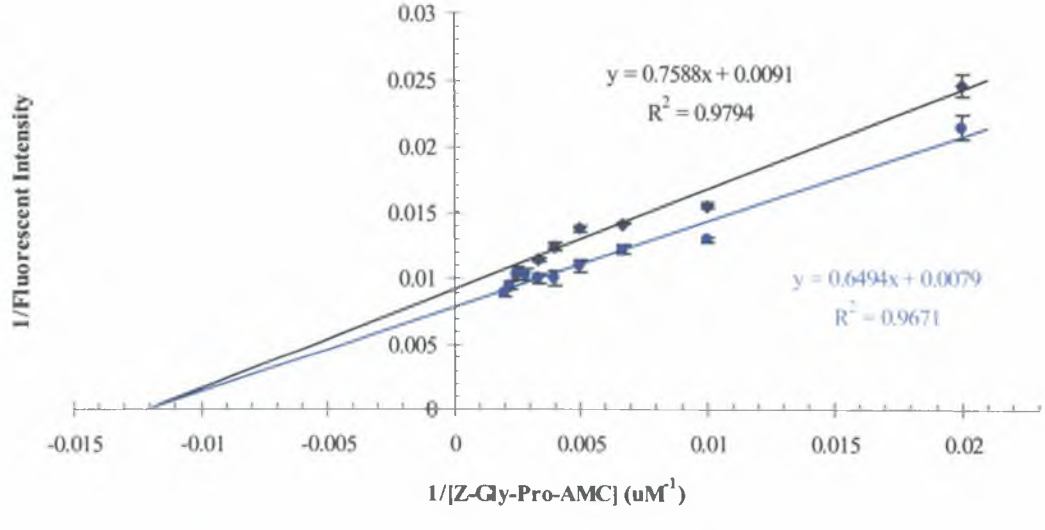


Figure 5.2.8 K_i^{app} determination for Z-Ala-Met-Phe using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the non-competitive inhibition of Seprase by Z-Ala-Met-Phe. Error bars shown represent SEM of triplicate readings.

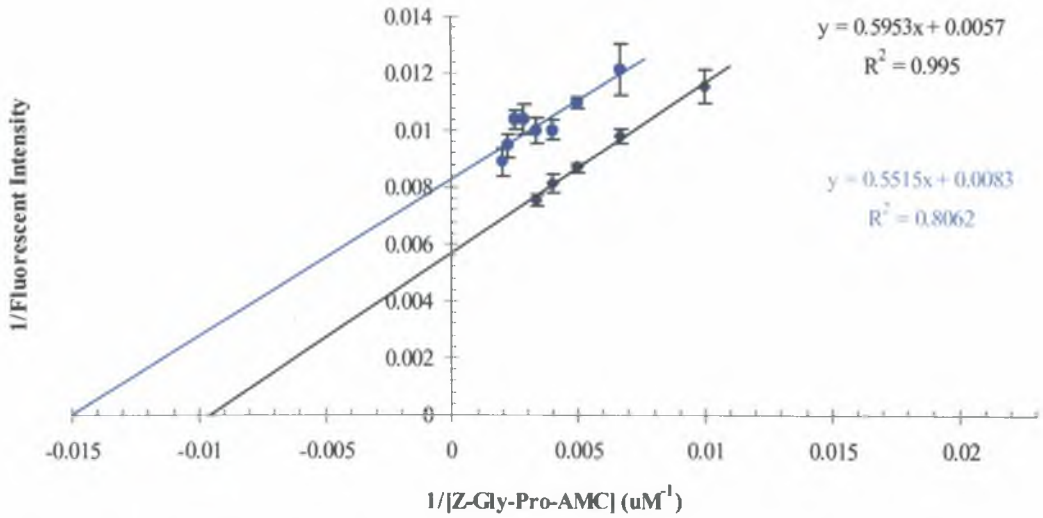


Figure 5.2.9 K_i^{app} determination for Z-Ala-Nle-Phe using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the uncompetitive inhibition of Seprase by Z-Ala-Nle-Phe. Error bars shown represent SEM of triplicate readings.

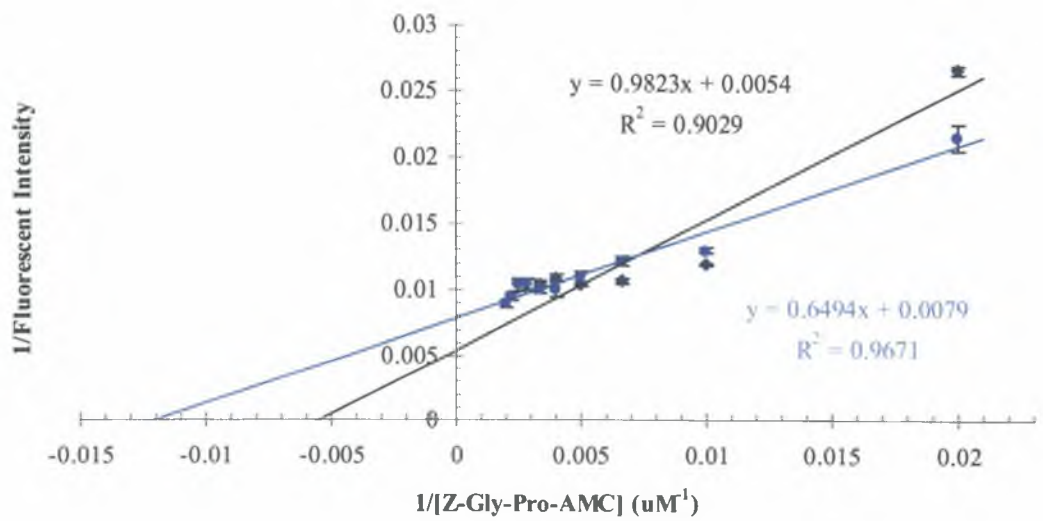


Figure 5.2.10 K_i^{app} determination for Z-Gly-Met-Phe-His using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the mixed inhibition of Seprase by Z-Gly-Met-Phe-His. Error bars shown represent SEM of triplicate readings.

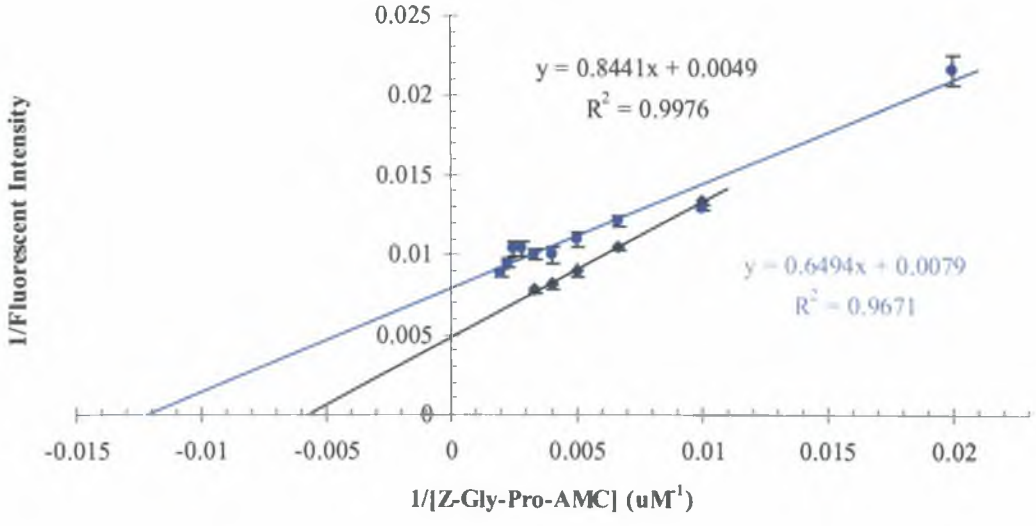


Figure 5.2.11 K_i^{app} determination for Z-Gly-Nle-Phe-His using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the mixed inhibition of Seprase by Z-Gly-Nle-Phe-His. Error bars shown represent SEM of triplicate readings.

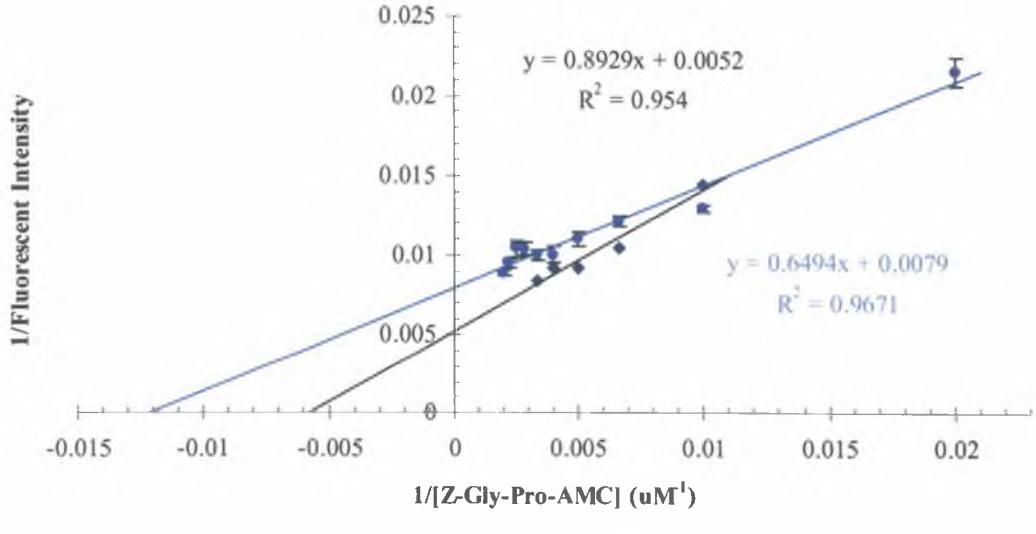


Figure 5.2.12 K_i^{app} determination for Z-Ala-Pro-Phe-His using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the mixed inhibition of Seprase by Z-Ala-Pro-Phe-His. Error bars shown represent SEM of triplicate readings.

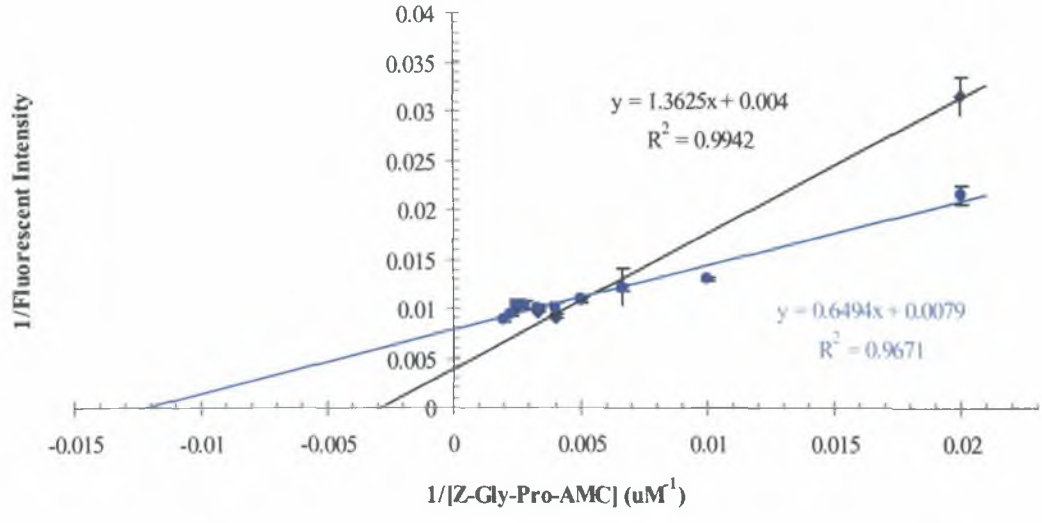


Figure 5.2.13 K_i^{app} determination for Z-Ala-Nle-Phe-His using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the mixed inhibition of Seprase by Z-Ala-Nle-Phe-His. Error bars shown represent SEM of triplicate readings.

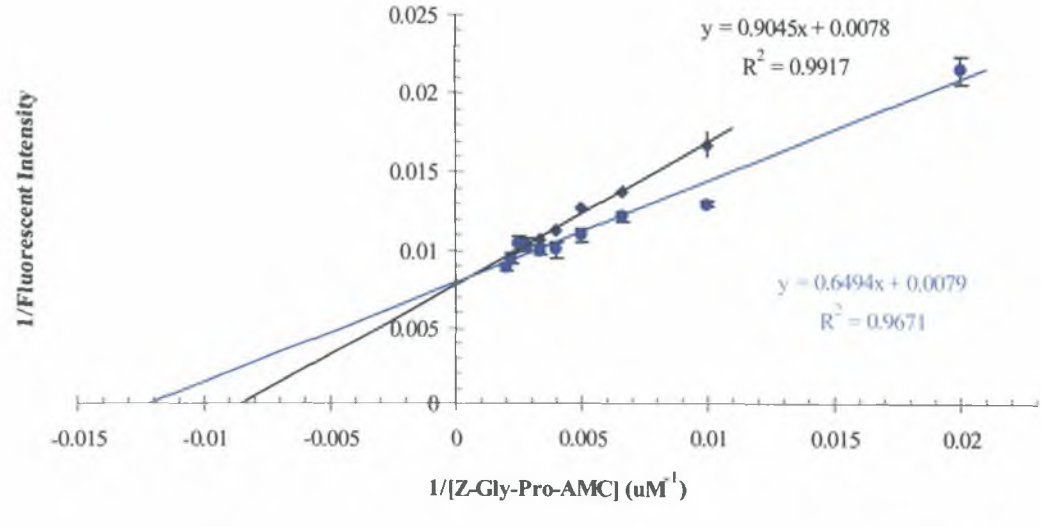


Figure 5.2.14 K_i^{app} determination for Z-His-Pro-Phe-His using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the competitive inhibition of Seprase by Z-His-Pro-Phe-His. Error bars shown represent SEM of triplicate readings.

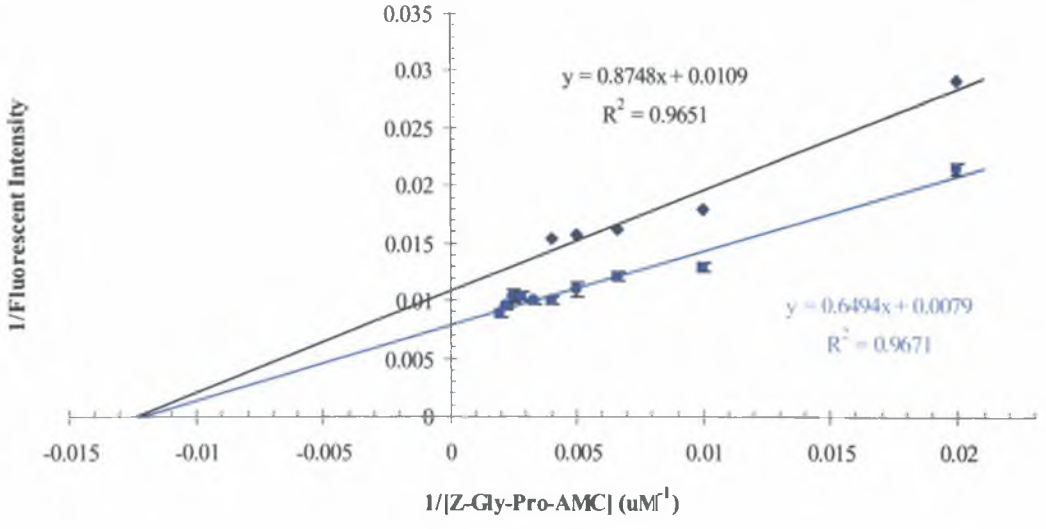


Figure 5.2.15 K_i^{app} determination for Ala-Ser-Gly-Pro-Ser-Ser using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the non-competitive inhibition of Seprase by Ala-Ser-Gly-Pro-Ser-Ser. Error bars shown represent SEM of triplicate readings.

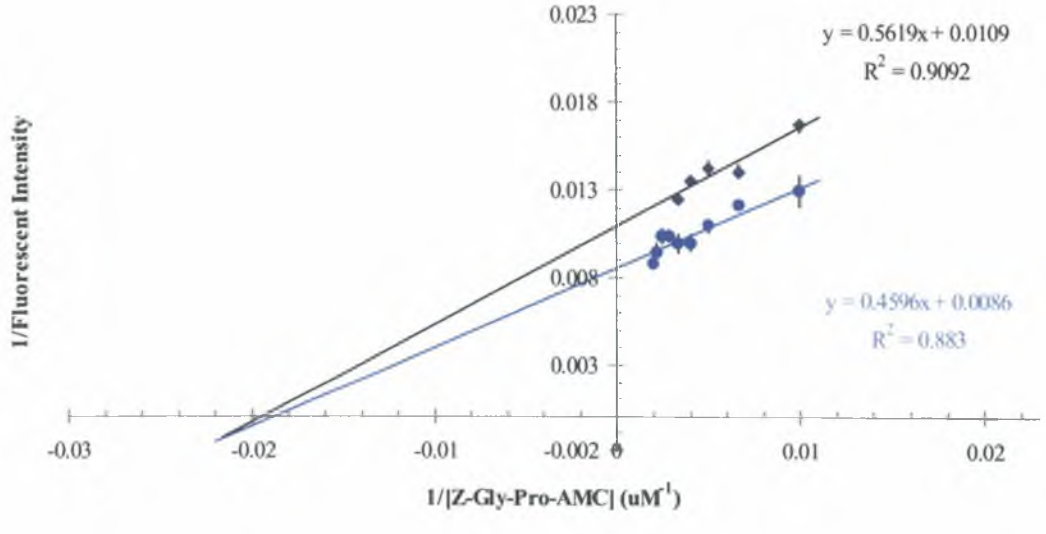


Figure 5.2.16 K_i^{app} determination for Ala-Ser-Nle-Pro-Ser-Ser using Lineweaver-Burk plot

Data from peptide competition assays (Section 2.8.2.2) fitted to Lineweaver-Burk model, as outlined in Appendix A, for the determination of apparent K_m (K_m^{app}). Non-inhibited data plotted in blue. Equations and regressions are shown. K_i^{app} values are given in Table 5.2.2. Plot illustrates the mixed non-competitive-uncompetitive inhibition of Seprase by Ala-Ser-Nle-Pro-Ser-Ser. Error bars shown represent SEM of triplicate readings.

Peptide	K_m^{app} (μM)			Average	K_i^{app} (μM)	V_{max}	Inhibition Type
	Lineweaver-Burk	Eadie-Hofstee	Hanes-Woolf			*Average*	
Z-Gly Pro-AMC	83 33	79 946	65	82 10*	-	127 49*	-
Z-Gly Pro-Phe	200	128 93	120	149 64	243 12	149 17	Uncompetitive
Z-Gly Met-Phe	105 26	90 52	90	95 26	1247 72	136 70	Mixed
Z-Ala-Pro-Phe	111 11	99 523	95	101 87	830 55	130 30	Mixed
Z-Ala-Met-Phe	83 33	75 983	90	83 10	16420	105 26	Non-competitive
Z-Ala-Nle-Phe	181 82	145 31	120	149 04	245 29	206 10	Uncompetitive
Z-Gly-Met-Phe-His	196 08	62 71	94	117 60	462 53	148	Mixed
Z-Gly Nle-Phe-His	178 57	168 9	160	169 16	188 60	202 03	Mixed
Z-Ala-Pro-Phe-His	172 41	136 74	159	156 05	222 04	183 14	Mixed
Z-Ala-Nle-Phe-His	344 82	45 434	159	183 08	162 60	173 08	Mixed
Z-His-Pro-Phe-His	357 14	117 85	125	200	139 27	162 86	Competitive
Ala-Ser-Gly-Pro-Ser-Ser	83 33	-	-	83 33	13349	90 90	Non-competitive
Ala-Ser-Nle-Pro-Ser-Ser	250	120	155 7	155 7	223 10	120 16	Mixed

Table 5 2 2 K_i^{app} values for selected peptides

K_m^{app} values obtained from Figures 5 2 5 to 5 2 17 K_i^{app} calculated as outlined in Appendix A
* Average values calculated as per Table 5 2 1

5 2 2 1 Effect of Tri- and Tetra peptides on Seprase Activity

Using the dipeptide substrate P_1 library, results show that Seprase will cleave Proline preferentially over Norleucine and Methionine respectively These residues will however fit into the S_1 site of Seprase The tri-peptide Z-Gly-Pro-Phe was altered using these results to see the effect of different conformations of amino acids on the affinity of Seprase for the particular peptide Changing the P_1 Proline to Methionine in the peptide Z-Gly-X-Phe (X being the substituted amino acid), decreased affinity 5-fold (243 12μM to 1247 72μM respectively) Methionine is obviously not in the right configuration for the binding pocket at this position Methionine, which is non-polar, is larger than Proline and it has quite a long side chain which extends away from the α-carbon of the amino acid This is quite a different configuration compared to that of the pyrrolidine ring of Proline While it is a bulky residue and can cause kinks in a polypeptide chain, the S_1 binding site in Seprase is capable of accommodating Proline

The change from Glycine to Alanine in the peptide Z-X-Pro-Phe affected the affinity by decreasing it 5 fold This demonstrates that P_2-S_2 is also an important site for the specificity of Seprase Both amino acids are similar in size and conformation but the decrease in affinity is possibly due to the change from a hydrophilic to a hydrophobic residue It is possible that the configuration and arrangement that Proline confers on the P_2 residue places it into close proximity to other residues that may or may not be favourable to hydrophobicity This result does show that for this peptide, Seprase has a preference for a hydrophilic residue in the P_2 position By altering this peptide further, changing the Proline to a Methionine (Z-Ala-X-Phe), led to an almost 20 fold decrease in affinity This is understandable when it is taken into consideration that Proline is the preferred residue in the S_1 subsite, and the conformation of Methionine as discussed previously The long side chain of Methionine may not allow the P_2 Alanine access to the S_2 subsite, thereby decreasing the affinity constant

A further alteration to this peptide, substituting Norleucine for Methionine (Z-Ala-X-Phe) led to a 67-fold increase in affinity (245 29μM and 16420μM respectively) Norleucine also has a long side chain, albeit not as long as Methionine, but it may be the orientation of the NH_2 group that is having an effect on the affinity of the substrate for accessing the binding pocket By substituting Proline back into this peptide (Z-Ala-Pro-Phe), the affinity constant is decreased 3-fold from 245 29μM to 830 55μM An extension of the peptide Z-Gly-Met-Phe by the addition of a Histidine to the P_2' position led to a 2 7-fold increase in the apparent affinity of Seprase for the substrate

Collins et al (2004) found that a 2-fold increase in affinity was observed in going from Z-Gly-Pro-Phe to Z-Gly-Pro-Phe-His By adding a charged Histidine residue in the P_2 position the affinity constant, K_m decreased from 461 53μM to 206 88μM, indicating increased affinity This study was unable to analyse the peptide Z-Gly-Pro-Phe-His as there wasn't sufficient peptide remaining from the previous study Therefore, in order to compare the effect of substituting Proline with Norleucine in the peptide Z-Gly-Pro-Phe-His, the K_i^{app} values obtained between this study and that of Collins et al (2004) had to be analysed to determine the degree of variation They reported a K_m of 270μM for Z-Gly-Pro-AMC which is over 3-fold higher than that obtained from the Enzfitter results in this study (82 10μM) Comparing the K_i^{app} results for Z-Gly-Pro-Phe from both studies shows that there is a ~2 fold difference (461 53μM and 243 12μM respectively) On average, there seems to be a ~2 5 fold difference between the two studies Collins et al (2004) reported a K_i^{app} of 206μM for Z-Gly-Pro-Phe-His Taking this ~2 5-fold difference into account, substituting Proline with Norleucine seemed to lead to a decrease in affinity [(188 60μM x 2 5) = 471 5μM from 206μM] This comparison comes with a caveat as it is not a direct comparison and it would be advisable to determine the $K_i^{(app)}$ for both peptides concurrently with the same batch of enzyme

The lowest K_i^{app} of 139 27μM was observed with the substitution of Glycine with Histidine in the peptide Z-X-Pro-Phe-His By replacing the P_2 residue with Alanine led to a decrease in affinity to 222 04μM This is possibly due to the Proline residue restricting the orientation of Alanine, thus not allowing access to the site or causing the

hydrophobic residue to react unfavourably in the binding pocket (there may be steric clashes between the different residues) It seems that Seprase tolerates a large positively charged residue at the P_2 position The increased affinity observed by substituting Proline with Norleucine in the peptide Z-Ala-X-Phe-His, may be due to Norleucine allowing Alanine to have greater access to the binding pocket It is possible that Norleucine gives greater flexibility to the peptide than proline-allowing for a greater tolerance of peptides that may not be as favourable to the Seprase binding pocket

Substitution of Glycine with Alanine at the P_2 position of the peptide Z-X-Nle-Phe-His, resulted in an increased K_i^{app} from 188 60μM to 162 60μM The addition of Norleucine at the P_1 position seems to allow a better tolerance of the hydrophobic residue Alanine at the P_2 position, as was discussed earlier with other peptides These results concur with those from the P_2 dipeptide library (Figures 5 1 3 and 5 1 4), in that Seprase shows a better tolerance for Alanine than Glycine This does seem to depend however, on the amino acid in the P_1 position

By altering the P_1 position in the peptide Z-Gly-X-Phe-His from Norleucine to Methionine the K_i^{app} increased over 2-fold from 188 60μM to 462 53μM This result confirms that although the P_2 substrate library results show that Seprase can tolerate Methionine in the S_1 subsite, the amino acid does not have the correct confirmation for the binding pocket and the extended substrate binding region

Edsouda et al (2006a) determined that Seprase had a high affinity for the peptide ASGP-SS (K_m 4 3μM and k_{cat}/K_m 1 3 x 10^6 $M^{-1}s^{-1}$) This peptide was analysed with the bovine Seprase along with a second peptide that had Norleucine instead of Glycine at the P_2 position By changing the Glycine residue the K_i^{app} increased 60-fold, from 13349μM to 233 10μM This peptide showed more flexibility and affinity in the binding pocket The results from the P_2 library (Figure 5 1 3 and 5 1 4) show that Norleucine is the preferred amino acid to Glycine This is shown in the markedly improved affinity constants described here Both peptides show competitive inhibition towards the hydrolysis of Z-Gly-Pro-AMC by Seprase (Figures 5 2 15 and 5 2 16) The reported results, using the

dipeptide library, show that the specificity of P_2 is not as strict as that reported by Edosada in 2006a Their library defined FAP as an endopeptidase of strict Gly-Pro cleaving specificity Another group looking at potent irreversible inhibitors of Seprase, report that the S_2 binding pocket appears to be capable of accommodating a wide variety of amino acid residues as our P_2 library results suggest (Gilmore et al , 2006)

It seems from these results and those of Collins et al (2004), that it is necessary for P_1-S_1, P_2-S_2 P_1'-S_1' and P_2'-S_2' to be filled for relatively good affinity There was better affinity when multiple substrate binding sites were occupied The results do show that Seprase did, however, have a lower affinity for these peptides than for Z-Gly-Pro-AMC The study of the P_1 site shows that it is designed to fit proline but it does tolerate other amino acids as does Prolyl Oligopeptidase, provided they do not exceed the size of the Prolines pyrrolidine ring (Edosada et al , 2006a) I would recommend the analysis of Z-Gly-Pro-Phe-His and Z-Ala-Pro-Phe-His with the same batch of enzyme to fully determine the P_2 preference, when Proline is in the P_1 position

Overall, from these results, Seprase shows that it has clear specificity for Proline in the P_1 position The peptidase has a preference for hydrophobic residues in the P_2 position and a bulky hydrophobic residue in the P_3 position Previous results by Collins et al (2004) have shown an evident preference for a hydrophobic residue at the C-terminal end of the scissile bond (P_1')

Chapter 6

Localisation Studies

6 1 Localisation Studies

Both bovine tissues and mammalian cell culture cell lines were used to determine the expression profile for Seprase and PO The result of which would ascertain the tissue of origin of the two serum Z-Gly-Pro-AMC degrading peptidases and/or establish the optimum source for future study of these enzymes A previous tissue localisation study (Birney and O'Connor, 2001) did not perform as extensive a profile as that described here

6 1 1 Bovine Tissue Studies

Bovine tissue samples were obtained from freshly slaughtered cows and prepared as outlined in section 2 9 1 The pellet (P_2) and both supernatants (S_1 and S_2) obtained from the tissues listed in Table 2 9 1 were analysed quantitatively for enzyme activity and protein content as outlined in section 2 3 3 and 2 1 1 respectively The total Z-Gly-Pro-AMC activity present was obtained in the absence of the potent inhibitor JTP-4819 and PO activity was taken to be the residual non-Seprase activity The inner filter effect as discussed in section 2 2 2 was observed with the coloured homogenised tissue samples, therefore, filtered AMC standard curves were prepared for each tissue extract (S_1, S_2 and P_2)

The Total Activity (Unit) and Specific Activity (Unit/mg) were calculated for each tissue fraction according to Appendix A Figures 6 1 1, 6 1 2, 6 1 3 and 6 1 4 illustrate these calculated activities for Seprase and PO obtained for each tissue

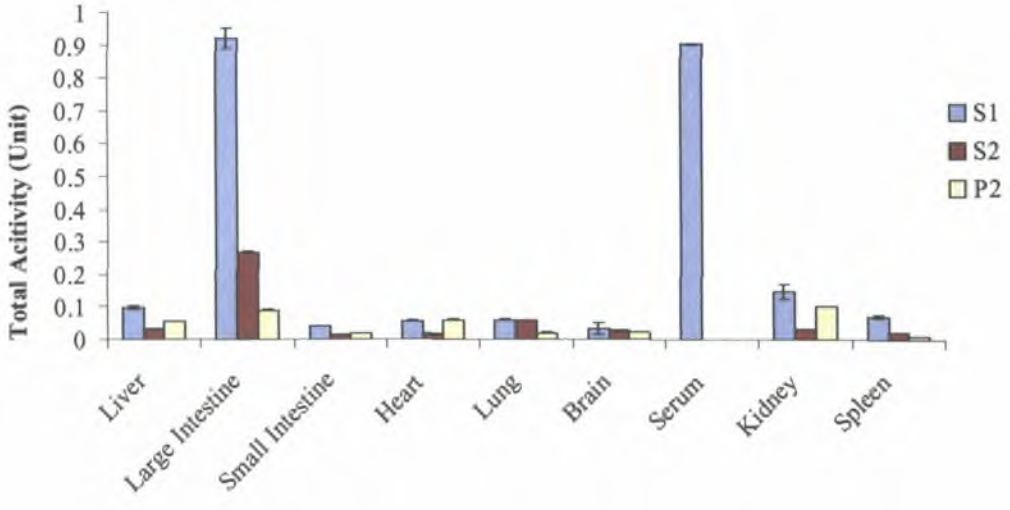


Figure 6.1.1 Total Seprase Activity in Bovine Tissue

Total Activity was calculated according to Appendix A. Error bars shown represent SEM of triplicate readings.

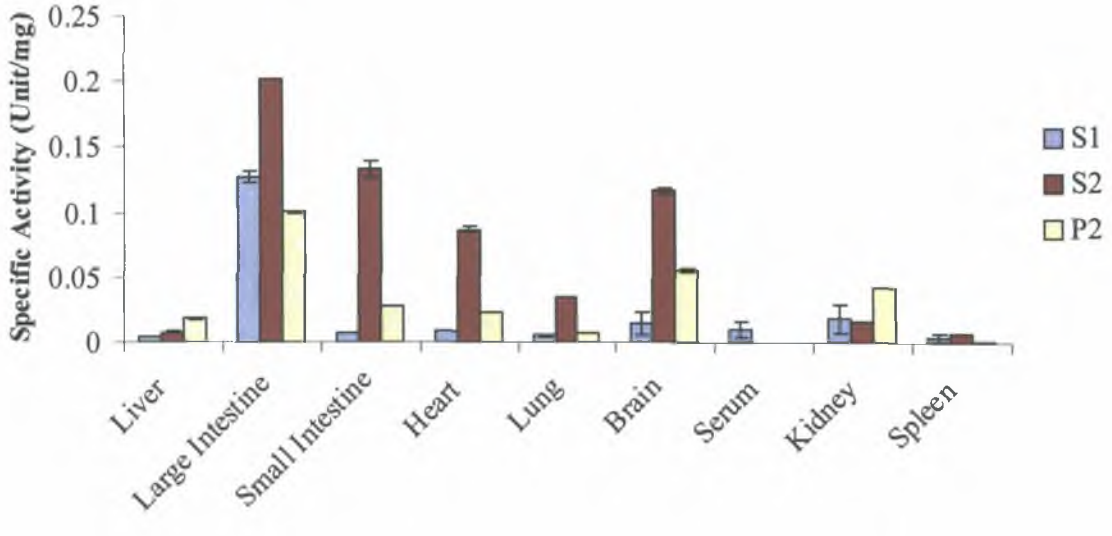


Figure 6.1.2 Specific Activity of Seprase in Bovine Tissue

Specific activity is a ratio of total enzyme activity to total protein and was calculated according to Appendix A. Error bars shown represent SEM of triplicate readings

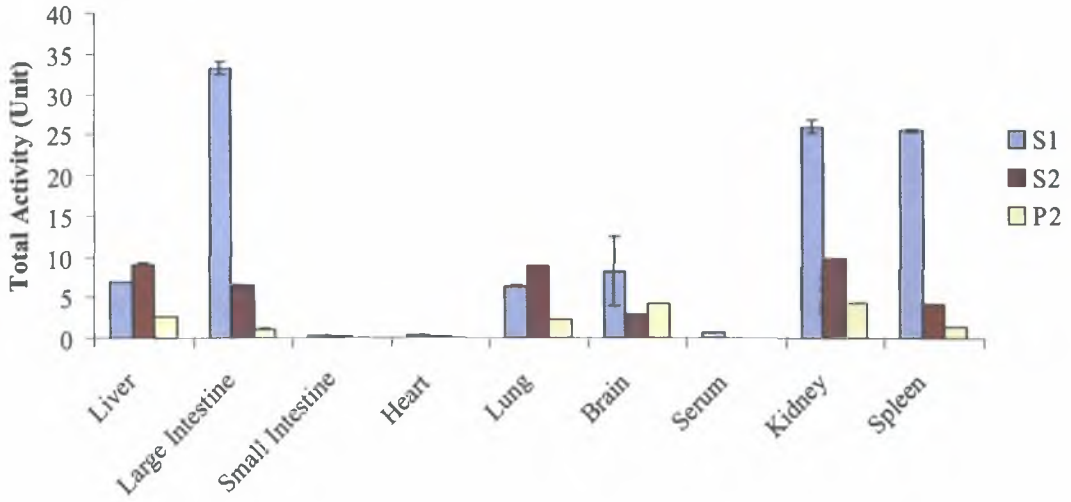


Figure 6.1.3 Total Prolyl Oligopeptidase Activity in Bovine Tissue

Total Activity was calculated according to Appendix A. Error bars shown represent SEM of triplicate readings.

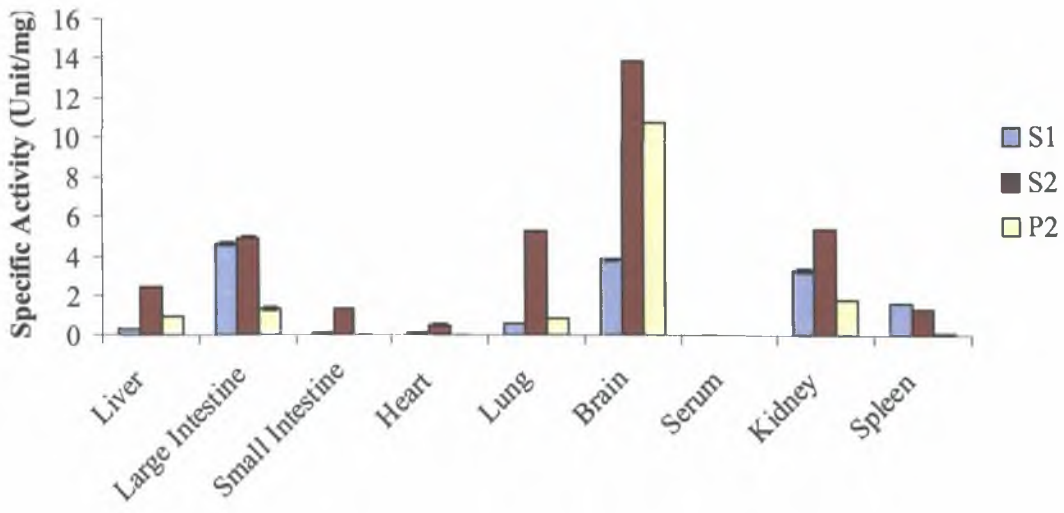


Figure 6.1.4 Specific Activity of Prolyl Oligopeptidase in Bovine Tissue

Specific activity is a ratio of total enzyme activity to total protein and was calculated according to Appendix A. Error bars shown represent SEM of triplicate readings.

Extraction of soluble proteins from animal tissues is relatively straightforward due to the weakness of the cell membranes and the absence of a cell wall. For all tissue samples, the bulk peptidase activity was determined to be in the S1 fraction (Figures 6.1.1 and 6.1.3) and, therefore, the focus will be on the specific activities calculated in these supernatants.

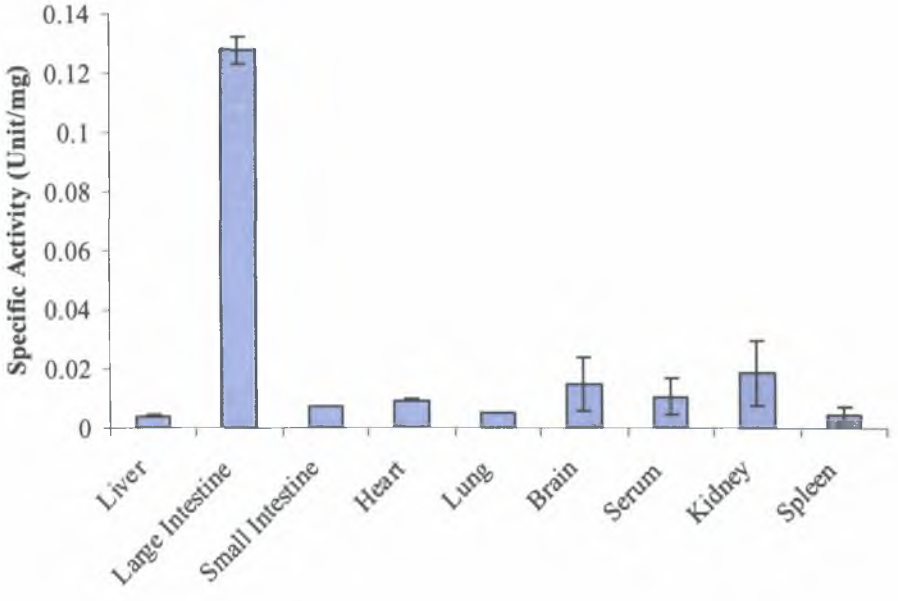


Figure 6.1.5 Specific Activity of Seprase in S_1 Fraction

Error bars shown represent SEM of triplicate readings.

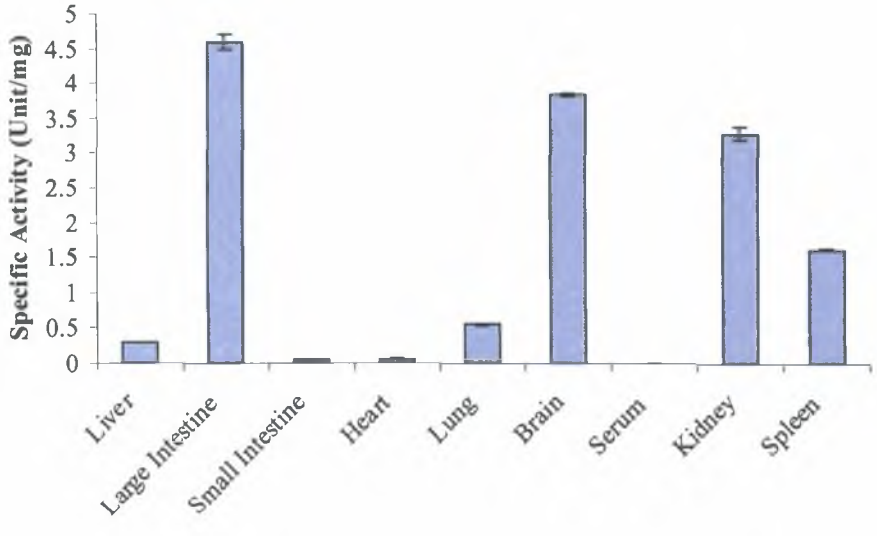


Figure 6.1.6 Specific Activity of Prolyl Oligopeptidase in S_1 Fraction

Error bars shown represent SEM of triplicate readings.

Figures 6.1.1 and 6.1.5 clearly illustrate the positive identification of Seprase activity in bovine large intestine, serum, kidney, liver and spleen extracts with total activities of 0.92, 0.91, 0.15, 0.1 and 0.07 Units respectively. Figure 6.15 illustrates a reduction in the apparent activity of Seprase in these extracts. This is due to the protein content in these extracts. Serum had the highest protein content, followed by liver and the spleen (85.19, 23.81 and 15.7 mg respectively). The large intestine contained 7.2 mg

protein, thereby allowing for quite a high specific activity of 0 128 units/mg Figures 6 1 2 and 6 1 6 illustrate the identification of PO activity in bovine large intestine, kidney, spleen, brain, liver and lung with total activities of 33 22, 25 85, 25 40, 8 19, 0 1 and 0 06 Units respectively

Using the developed highly specific fluorimetric microplate assay, both enzymes seem to be ubiquitously expressed, although PO to a greater extent These results are comparable to those reported by Birney and O'Connor (2001) Bovine large and small intestine and heart were included in this research study and, as can be seen in Figures 6 1 5 and 6 1 6, the large intestine has the highest specific activity for both Seprase and PO (0 128 units/mg and 4 61 units/mg respectively) Although Figure 6 1 5 shows that the large intestine has the highest specific activity (0 128 unit/mg) of Seprase, Table 6 1 1 shows that this activity only represents 13% of the total Z-Gly-Pro-AMC hydrolysing activities in the organ The large intestine reabsorbs water and any digested materials which have not already passed into the bloodstream The localisation of Seprase in the large intestine could indicate a role for Seprase in the digestive system, possibly proteolytically cleaving proline-containing peptides or proteins This is the first report of Seprase expression in the large intestine and kidney

Tissue *S_i Samples*	**Seprase Activity** *%*	**PO Activity** *%*
Liver	1	99
Large Intestine	13	87
Small Intestine	1	99
Heart	3	97
Lung	2	98
Brain	1	99
Serum	56	44
Kidney	10	90
Spleen	1	99

Table 6 1 1 Levels of Seprase and PO Activities in Various Bovine Samples

Activities are expressed as a percentage of the total Z-Gly-Pro-AMC degrading activity (100%)

In general, Seprase specific activity levels were considerably lower than those for PO in the same tissue extracts, except for serum This table shows that serum has the highest percentage of Seprase activity out of all the other tissue samples tested

As mentioned above, Figure 6 1 6 clearly shows that PO activity was most abundant in bovine large intestine extracts The second most abundant source of PO was the brain with 99% of the total Z-Gly-Pro-AMC hydrolysing activity Birney and O'Connor (2001) found that the brain had 94% PO activity

The large intestine is certainly a possible source for both Seprase and PO, with high specific activities and low protein content Table 6 1 1 also shows that another alternative source would be kidney However, extracting the enzymes from tissues involves a much more labour intensive process, employing a combination of mechanical and centrifugal forces

These results correlate well with those reported for Seprase in literature Seprase has been shown to be expressed in the liver (Levy et al , 1999) and also in intestinal– and diffuse-type gastric carcinomas (Okada et al , 2003) This could explain the high activity levels of Seprase found in the large intestine (Okada et al , 2003)

Bovine brain or kidney would also appear to be good sources of PO, being rich in enzyme activity and both are commonly employed in the extraction of prolyl oligopeptidase (Dehm and Nordwig, 1970, Yoshimoto et al , 1983)

6.1.2 Cell Localisation Studies

Mammalian cell lines were also used to determine the expression profile of Seprase and PO.

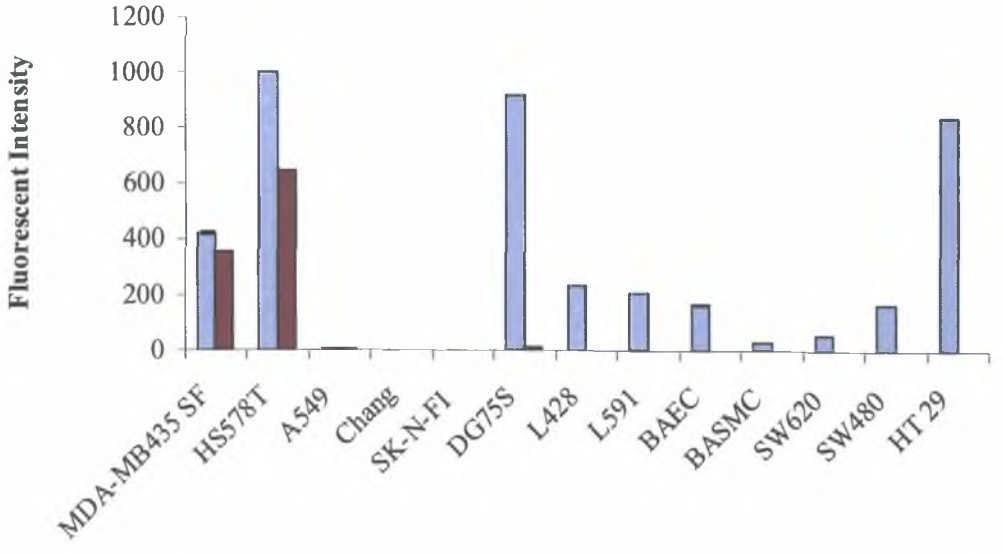


Figure 6.1.7 Z-Gly-Pro-AMC Hydrolysing Activity Profile for Mammalian Cells

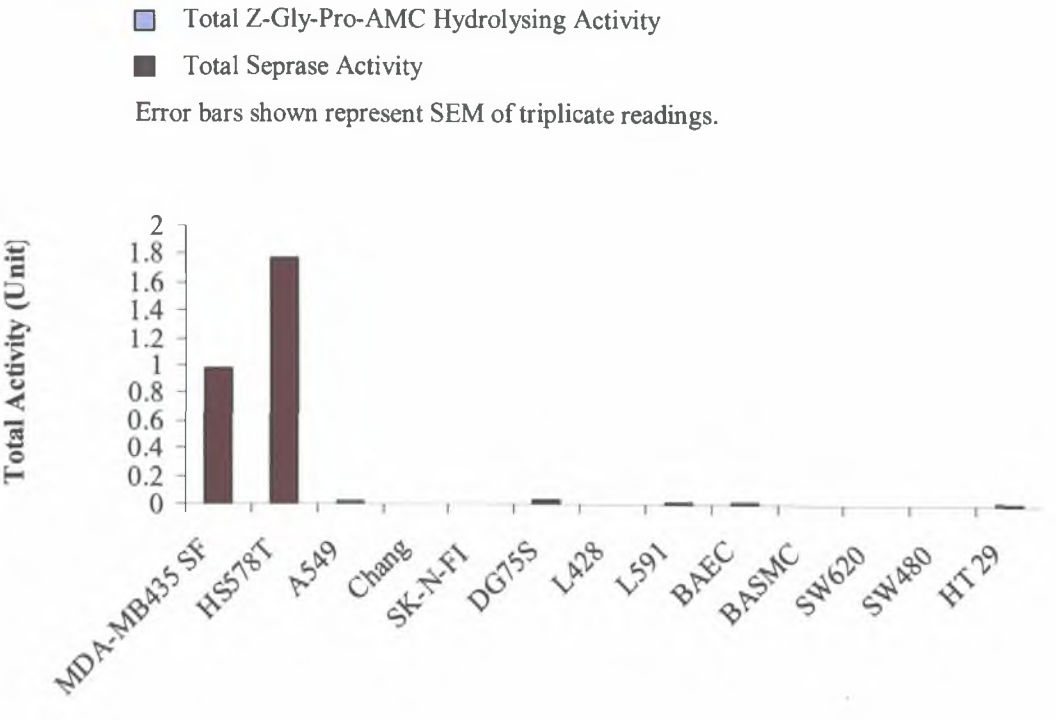


Figure 6.1.8 Total Seprase Activity in Mammalian Cell Lines

Error bars shown represent SEM of triplicate readings.

Figure 6 1 7 clearly shows that the breast cancer cell lines Hs578T and MDA-MB435 SF express Seprase Seprase activity represents 64% and 85% of the total Z-Gly-Pro-AMC hydrolysing activity of these cell lines, respectively Seprase activity has also been identified by others in breast cancer cell lines such as MCF-7 (Goodman et al , 2003) This is the first report of Seprase expression in the breast cancer cell lines Hs578T and MDA-MB435 SF The cell lines A549 (lung), Chang (liver) and SK-N-FI (brain) did not show any Z-Gly-Pro-AMC hydrolysing activity The remainder of the cell lines showed varying degrees of PO activity The breast cancer cell line Hs578T shows a high degree of Seprase activity and, therefore, was chosen for use in cloning the Seprase gene (Section 7 0)

Chapter 7

Molecular Cloning of Seprase

7 1.1 Cloning of human *Seprase* gene

The objective of this work was to clone the *Homo sapiens* (human) Seprase gene from cDNA (Section 2 10 3 1 and 2 10 7 1) and subsequently develop a recombinant expression system (Section 2 10 11 and 2 10 12), capable of producing catalytically active recombinant human Seprase at relatively high levels, which could be purified for further biochemical and functional studies

Cell pellets (~1 x 10^7 cells) of three human cancer cell lines, MDA-MB-435 SF, SW480 and Hs578T (American Type Culture Collection No HTB-129, CCL-228 and HTB-126 respectively) were a kind gift from the laboratories of Dr Rosaleen Devery and Dr Susan McDonnell (DCU, Ireland) MDA-MB-435 SF is derived from a metastatic ductal adenocarcinoma SW480 was derived from a primary adenocarcinoma of the colon Hs578T was derived from a carcinoma of the breast The SW480 cells were found to be negative for Seprase activity (0 0007 Units) as determined by the method described in Section 2 9 2 In contrast, the MDA-MB-435 SF and Hs578T cells were found to be positive for Seprase activity (0 97 and 1 76 Units respectively) Total RNA was isolated from a fresh Hs 578T cell pellet as described in Section 2 10 3 1 Figure 7 1 1 shows analysis of this RNA by agarose gel electrophoresis (Section 2 10 4) Sharp, clear 28S and 18S ribosomal RNA (rRNA) bands were an indication that the RNA preparation had not been degraded to any significant extent

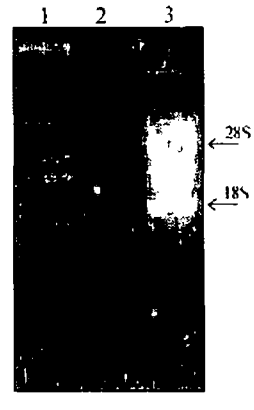


Figure 7 1 1 RNA Analysis

1% agarose gel Lane 1, DNA Ladder, Lane 2, RNA isolated from Hs578T cells (Section 2 10 3 1) treated with Ribonuclease A (Section 2 10 7), Lane 3, RNA isolated from Hs578T cells Prominent 18S and 28S rRNA bands are indicated using arrows

The sequence for human *Seprase* mRNA (NM_004460, GI 16933539, Figure 7.1.4) was obtained from GenBank. The 2788bp sequence contains a 2283bp open reading frame (ORF) for human *Seprase*. The initial aim of this work was to clone the *Seprase* gene into pCR2.1, thereby providing a DNA template for further cloning applications. cDNA was generated from the Hs578T RNA using the oligo$(dT)_{15}$ primer as described in section 2.10.7.1. The cDNA was used as template for PCR (Section 2.10.7.2) using primers *Bam*HI-Fap1 and *Xho*I-Fap6 and REDAccu*Taq* LA DNA polymerase. The *Bam*HI-Fap1 and *Xho*I-Fap6 (Table 2.10.2) were designed external to this ORF to amplify the *Seprase* gene (Figure 7.1.4) as a 2366bp fragment. The resulting PCR product, analysed by agarose gel electrophoresis (see Section 2.10.4), is shown in Figure 7.1.2 (A). A relatively faint band corresponding to the expected size of 2366bp was obtained. Some non-specific bands were also present. Several attempts were made to increase the specificity of the PCR and eliminate these bands by increasing the annealing temperature during the PCR reaction (see Figure 7.1.2. (A)). However, these attempts proved unsuccessful. The PCR Optimisation Kit (Sigma) containing PCR buffers with varying pH and Mg^{++} concentrations were used to further enhance the PCR. It was found that Buffer G of the kit increased the amount of PCR product formed although it was unable to increase the primer specificity, leading to the production of non-specific bands (Figure 7.1.2 (B)).

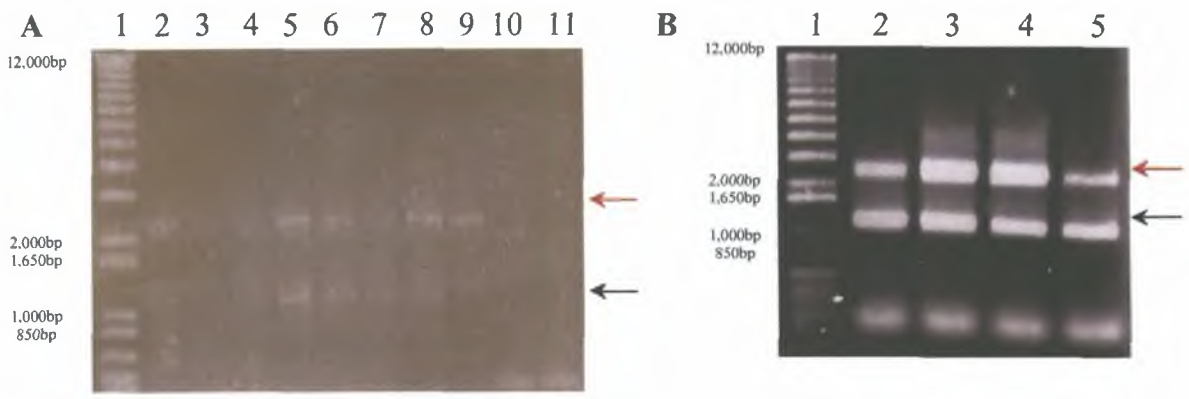


Figure 7.1.2 Cloning of human *Seprase* gene

0.7% agarose gels. (A) Lane 1, DNA Ladder; Lane 2-11, temperature gradient products of RT-PCR on Hs578T RNA using primers *Bam*HI-FAP1, *Xho*I-FAP6, REDAccu*Taq* LA for temperatures 52.1°C, 52.5°C, 53.3°C, 54.7°C, 56.4°C, 58.4°C, 60.3°C,62.3°C, 64.6°C, 66.0°C. (B) Lane 1, DNA Ladder; Lane 2-5, products of RT-PCR on Hs578T RNA using primers *Bam*HI-FAP1, *Xho*I-FAP6 REDAccu*Taq* LA and Buffer G (2.5mM $MgCl_2$). Annealing temperatures used were 54.6°C, 53.4°C, 52.5°C and 52.1°C respectively. A red arrow indicates the band corresponding to the expected 2344bp *Seprase* gene product, while a black arrow indicates non-specific bands.

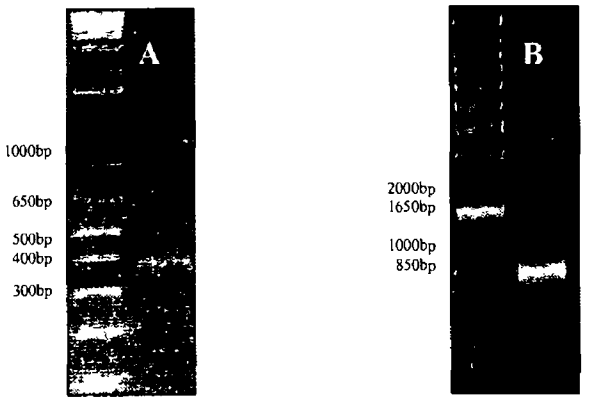


Figure 7 1 3 Control PCR Reactions

Control RT-PCR reactions performed on Hs578T RNA using (A) β-actin primers β-actin For and β-actin Rev (Table 2 10 2), yielded 360bp product, (B) Fap3 and Fap4 primers yielded the 836bp fragment of *Seprase* Control reactions were performed on all PCR reactions (Section 2 10 7 1) to determine that it was successful in making good quality cDNA

The PCR products shown in Figure 7 1 2 (B) were extracted from the gel according to Section 2 10 3 3 1, restricted with the respective enzymes (Section 2 10 7), and cloned into the cloning vector pCR2 1 as described in Section 2 10 8 2 Several white colonies were purified and plasmid DNA was isolated (Section 2 10 3 2 1) Screening for a clone having the desired ~2366bp recombinant *Seprase* insert, was carried out by enzymatic restriction analysis with *Eco*R1 (Section 2 10 7) The site for the insertion of PCR products in the pCR2 1 vector is bounded by *Eco*R1 sites that function as diagnostic restriction sites The clones were found to contain a smaller fragment than the ~2366bp expected fragment

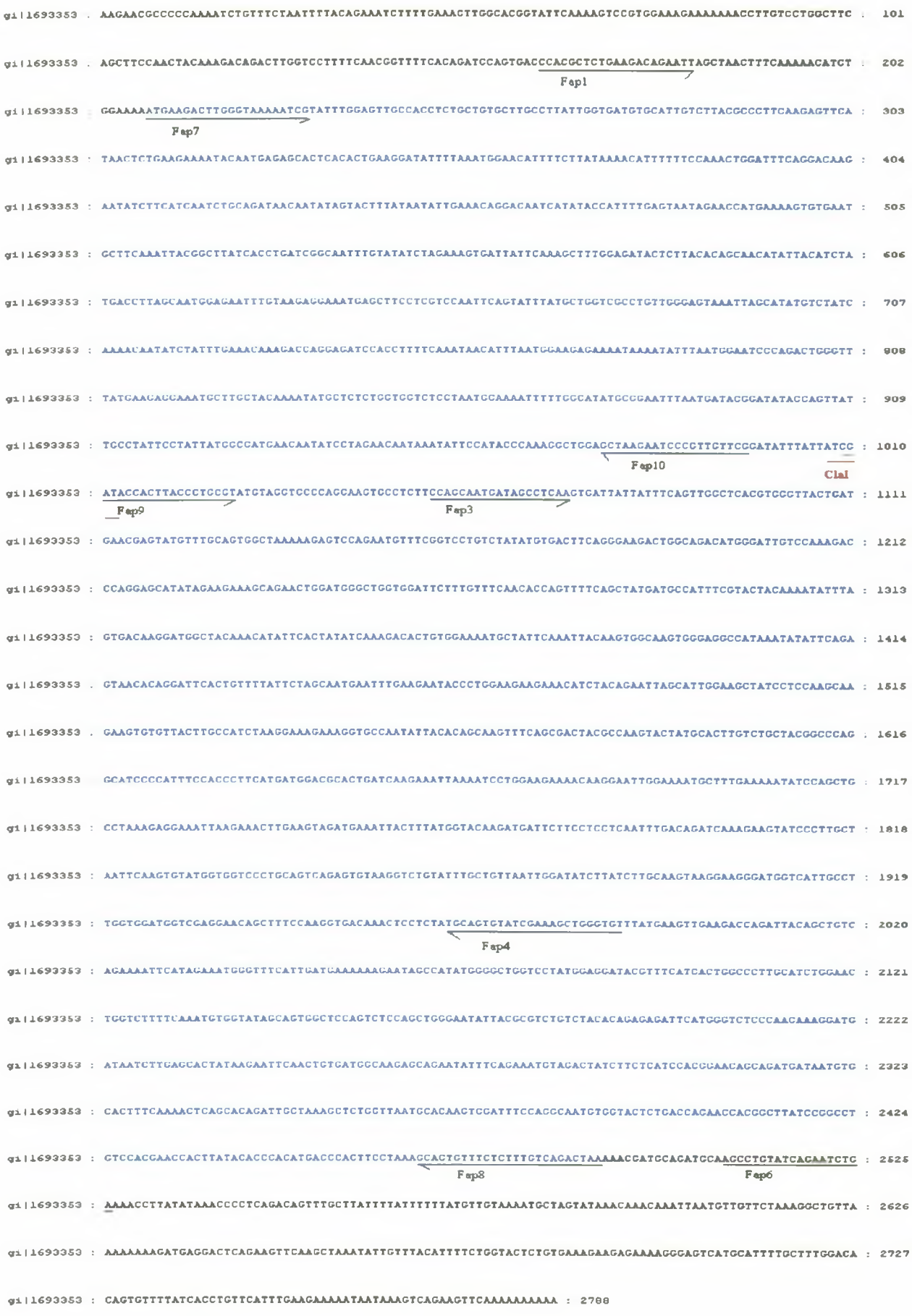


Figure 7.1.4 Human *Seprase* mRNA sequence

Nucleotide sequence GI 16933539: human *Seprase* mRNA. ORF is highlighted in blue. Binding locations of primers are indicated. *Cla*I recognition site indicated in red. Illustrated using GenDoc (Section 2.10.10).

Variations of different forward and reverse primers were used for RT-PCR reactions to try and amplify the full *Seprase* gene. These were unsuccessful (possibly due to the large size of the gene) and therefore it was decided to clone the *Seprase* gene in two parts.

The *Seprase* gene has a unique *Cla*I recognition site at 800bp from the start codon. Primers were designed to amplify *Seprase* in two parts. Figure 7.1.5 (A) illustrates the strategy employed. The *Seprase* gene was amplified using REDAccu*Taq* LA and the products formed were run on an agarose gel (Figure 7.1.5 (B)). The PCR products were concentrated using the method described in Section 2.10.3.4 and restricted with the respective endonuclease restriction enzymes according to Section 2.10.7. The vector pCR2.1 (Figure 2.10.1) was restricted using *Bam*HI and *Xho*I and treated using Antarctic Phosphatase (Section 2.10.8.1). The two restricted inserts were ligated with the treated vector and transformed into the *E. coli* strain XL-10 Gold according to Section 2.10.5.2. A candidate clone (pPOB1) was identified and further analysed by PCR using *Bam*HI-Fap7 and *Xho*I-Fap8 (Table 2.10.2). A band corresponding to the expected 2283bp recombinant *Seprase* (r*Seprase)* gene product was obtained, as shown in Figure 7.1.5. C.

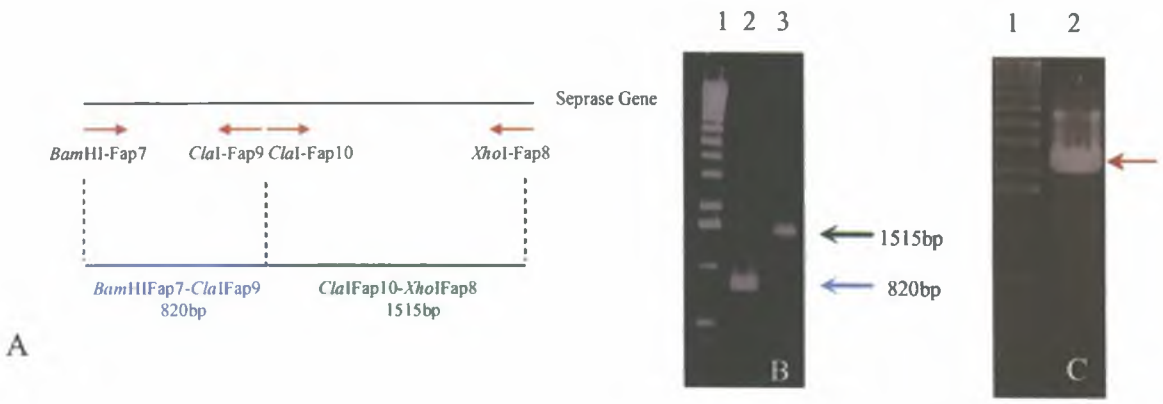


Figure 7.1.5 The Cloning of r*Seprase* in two parts strategy

The *Seprase* gene was cloned into the vector pCR2.1 in two parts utilising the unique *Cla*I recognition site in the gene (at 800bp) and amplification strategy in (A). RT-PCR was performed on RNA isolated from Hs578T. PCR was performed on this cDNA using REDAccu*Taq* LA DNA polymerase and the products visualised on a 1% agarose gel (B). Lane 2 shows the 820bp fragment (blue) and Lane 3 shows the 1515bp fragment (green). (C) 0.7% agarose gel; Lane 1: DNA Ladder; Lane 2; product of PCR on pPOB5 using primers *Bam*HI-Fap7 and *Xho*I-Fap8 (Table 2.10.2). The band corresponding to the expected 2283bp *Seprase* gene product is indicated by the red arrow.

A restriction digest of this clone (pPOB1) analysed by agarose gel electrophoresis is shown in Figure 7 1 6 Restriction with *Bam*HI and *Xho*I produces a band corresponding to the expected 2283bp fragment containing the r*Seprase* sequence Single restriction with *Bam*H1, *Xho*I and *Eco*RI produced bands corresponding to the expected 6132pb linearised plasmid Restriction with *Hind*III produced bands corresponding to the expected 424bp and 5708bp Restriction with *Pst*I produced bands corresponding to the expected 2947bp, 1765bp and 1420bp The sequence of pPOB1 was confirmed by DNA sequencing (Section 2 10 10) The sequencing data is given in Figure 7 1 8 A map of pPOB1 is shown in Figure 7 1 7

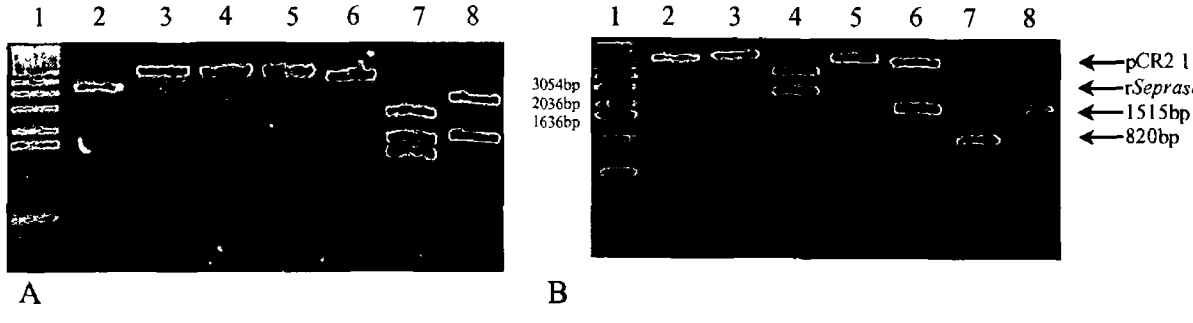


Figure 7 1 6 Verification Digest of pPOB1

Restriction digest of pPOB1 analysed on 1% agarose gel (A) Lane 1, 1kb DNA ladder, Lane2, pPOB1 uncut, Lane 3, *Bam*HI, Lane 4, *Xho*I, Lane 5, *Eco*RI, Lane 6, *Hind*III, Lane 7, *Pst*I, Lane 8, *Bam*HI and *Eco*RI (B) Lane 1, 1kb DNA ladder, Lane 2, pPOB1 cut with *Bam*HI, Lane 3, pPOB5 cut with *Xho*I, Lane 4, pPOB1 cut with *Bam*HI & *Xho*I, Lane 5, pPOB1 cut with *Bam*HI & *Cla*I, Lane 6, pPOB5 cut with *Cla*I & *Xho*I, Lane 7, PCR fragment *Bam*HI-Fap7--*Cla*I-Fap9, Lane 8, PCR fragment *Cla*I-Fap10--*Xho*I-Fap8

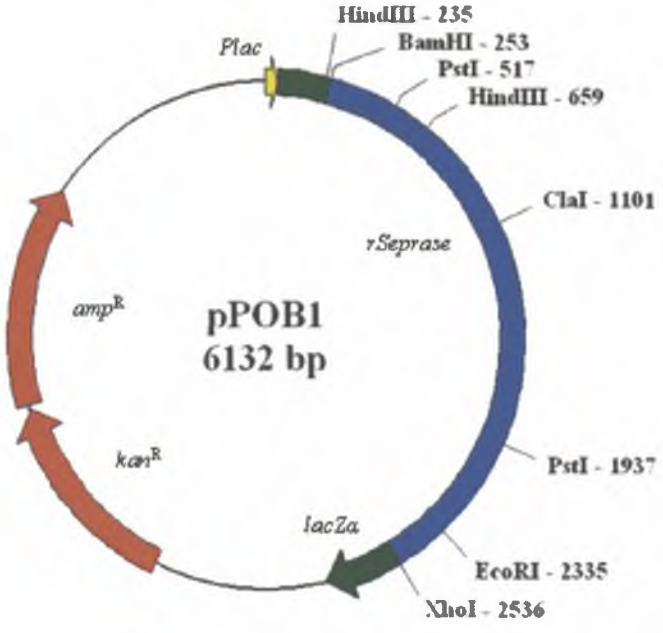


Figure 7.1.7 pPOB1 Plasmid Map

The *rSeprase* gene fragment (blue) is inserted into the vector pCR2.1. The *lacZα* ORF (green) is under the control of the *lac* promoter (yellow). Ampicillin and Kanamycin resistance genes (amp^R & kan^R) are shown in red. Generated by pDRAW32 (Section 2.10.10).

Figure 7.1.8 illustrates the alignment of pPOB1 with the FAP/Seprase mRNA sequence GI 16933539. The alignment illustrates a frame shift mutation at the 3' end of the *rSeprase* sequence, thereby altering the codon sequence and the last 3 amino acids.

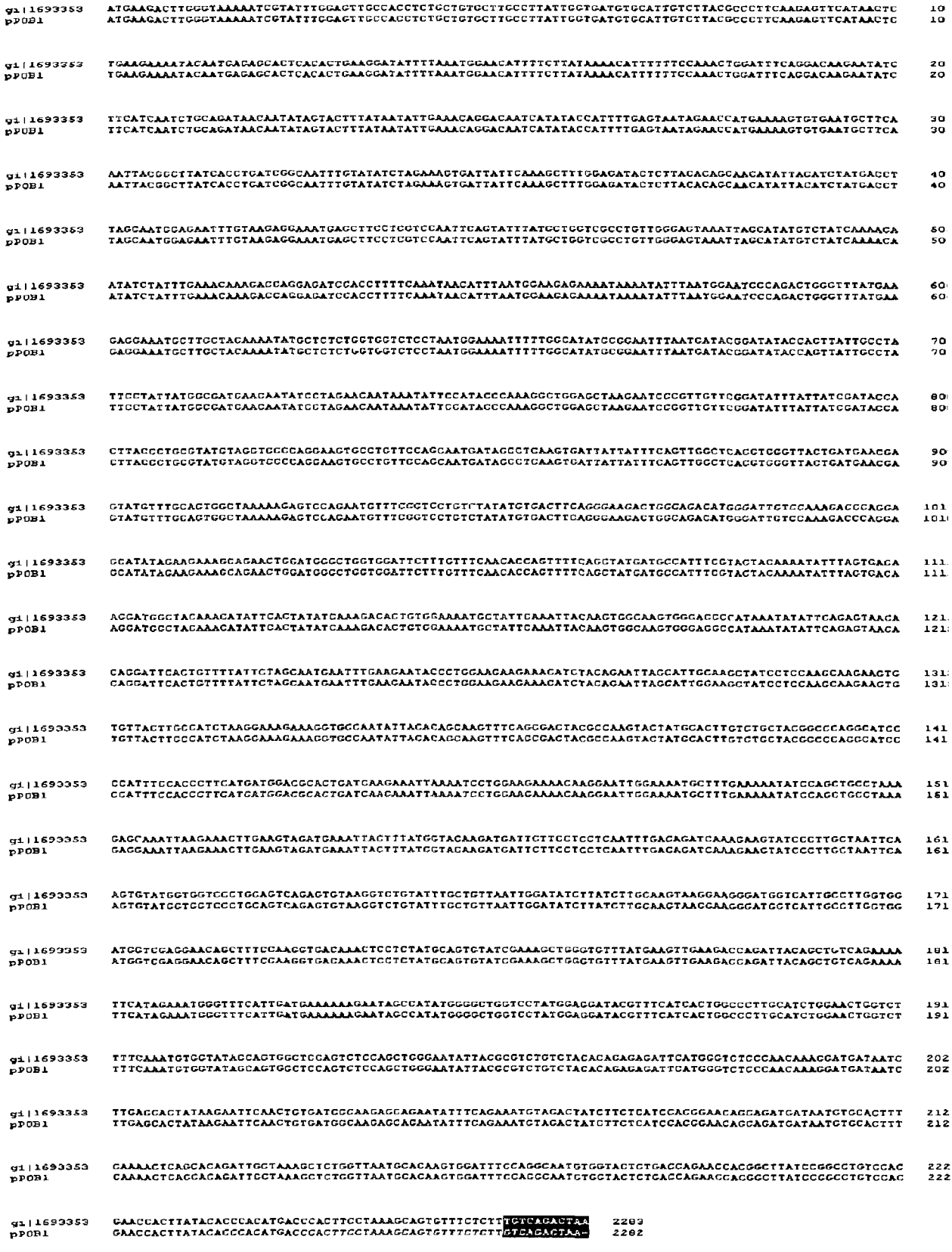


Figure 7 1 8 DNA Sequencing of pPOB1

Sequence of pPOB1 aligned with that of Seprase (GI 16933539) The alignment illustrates a frame shift at the 3' end of the sequence (highlighted in grey) Illustrated using GenDoc (Section 2 10 10)

While analysing the pPOB1 sequence (Figure 7 1 8), the cloned gene product was aligned with various FAP/Seprase published sequences, which led to some interesting observations Analysis of the various published FAP/Seprase mRNA sequences displayed differences between them at the mRNA levels

Figure 7 1 9 illustrates the differences at mRNA level between *FAP* (GI 16933539 and GI 1888315) and *Seprase* (GI 1924981) A total of 5 base changes were observed at mRNA level (see Figure 7 1 9 (A)), two of which conserved the amino acid residues at the protein level (see Figure 7 1 9 (B), marked *) The remaining 3 base changes altered Ala^{207}-Pro^{207}, Tyr^{229}–Lys^{229} and Tyr^{354}-Arg^{354} (marked as 1, 2, 3 respectively in Figure 7 1 9) The addition of Proline at position 207 does not change the hydrophobic nature of that position but it would cause a 'kink' in the sequence, possibly altering the tertiary structure of the protein Proline is also a larger amino acid than Alanine Tyrosine^{229} was replaced by the smaller charged residue Lysine Tyrosine altered to Arginine led to the addition of a much larger and charged residue at position 354 The long extended side chains of Lysine and Arginine obviously do not hinder the folding or the activity of the protein This can also be said for the addition of Proline^{207} Obviously adding a 'kink' in the sequence at that position does not hinder the folding or the activity of Seprase

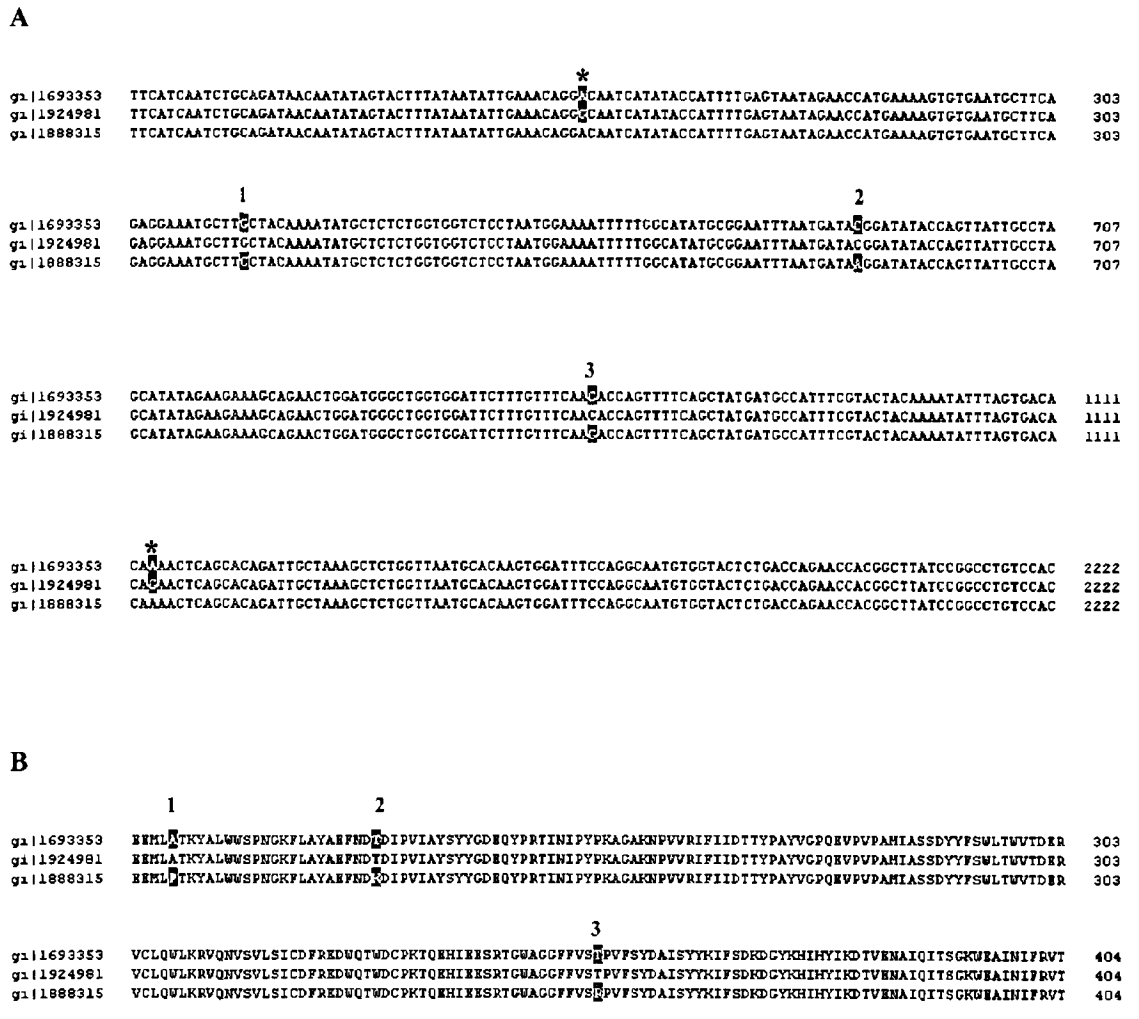


Figure 7 1 9 Alignments of Sequences

(A) *Seprase* (GI 1924981) and *FAP* mRNA (GI 16933539 and GI 1888315) sequences aligned This Figure illustrates the 5 altered bases between the sequences (shaded in grey) (B) The respective Seprase and FAP protein sequences for the mRNA sequences indicated in (A) are aligned The two base changes seen in Seprase GI 1924981 (marked *) conserved the amino acid residues of FAP (GI 16933539) The 3 base changes observed in GI 1888315 (marked 1, 2, 3) however altered the amino acid sequence Illustrated using GenDoc (Section 2 10 10)

A new high fidelity polymerase called Phusion High-Fidelity DNA Polymerase was supplied by NEB This polymerase generates PCR products with accuracy and speed even on more difficult and long templates, such as *Seprase* The error rate of Phusion DNA polymerase is 50 fold lower than that of *Thermus aquaticus* DNA polymerase The use of Phusion also has the added advantage of reduced extension and cycling times

Phusion DNA Polymerase, instead of REDAccu*Taq* LA, was therefore employed to clone the *Seprase* gene, as the PCR and cloning strategy to date illustrated the difficulty of the DNA template Cloning strategies were developed for cloning the *Seprase* gene into (1) pcDNA3, (2) pcDNA3-HA and (3) pQE30-Xa (Figures 2 10 3 and 2 10 2 respectively) A further cloning strategy was developed for cloning the Seprase gene into pIRES-hrGFP II (Figure 2 10 4)

7 1 2 Cloning of human *Seprase* gene into pcDNA3 and pcDNA3-HA

The mammalian expression vectors pcDNA3 and pcDNA3-HA (Table 2 10 3 and Figure 2 10 3) were chosen to clone and express the Seprase gene Both vectors were restricted with *Bam*HI and *Xho*I, treated with Antarctic Phosphatase according to Section 2 10 8 1 and concentrated according to Section 2 10 3 4 cDNA was generated from the Hs578T RNA using the oligo(dT)$_{15}$ primer as described in section 2 10 7 1 The primers *Bam*HI-A-Fap7 (5' end of the gene) and *Xho*I-Fap8 (3' end of the gene) were designed to amplify the full length *Seprase* gene (see Figure 7 1 4 and Table 2 10 2) in order to clone it into pcDNA3/pcDNA3-HA vectors The *Seprase* gene was amplified using this cDNA as a template for PCR with Phusion DNA Polymerase (Section 2 10 7 2) and the primers *Bam*HI-A-Fap7 and *Xho*I-Fap8 The products formed were run on an agarose gel (Figure 7 1 9 (B)) The PCR products were concentrated using the method described in Section 2 10 3 4 and restricted with the respective endonuclease restriction enzymes (*Bam*HI and *Xho*I) according to Section 2 10 7 The restricted PCR fragments were then ligated with the treated pcDNA3/pcDNA3-HA vector and transformed according to 2 10 5 2, into the *E coli* strain XL-10 Gold

Clones were screened according to Section 2 10 3 2 1 and a candidate clone was further analysed by restriction analyseis A restriction digest of this clone (pPOB12) analysed by agarose gel electrophoresis is shown in Figure 7 1 10 Restriction with *Bam*HI and *Xho*I produces a band corresponding to the expected 2283bp fragment containing the *rSeprase* (recombinant *Seprase*) sequence Single restriction with *Bam*H1 and *Eco*RI produced bands corresponding to the expected 7692pb linearised plasmid Restriction with *Pst*I produced bands corresponding to the expected 1420bp,

2007bp and 3983bp There also seems to be some uncut and partially cut vector DNA in Lane 7 (Figure 7 1 10 (A)) The successful cloning of *rSeprase* was confirmed by DNA sequencing (Section 2 10 10) The sequencing data is given in Figure 7 1 12 and Appendix C A map of pPOB12 is shown in Figure 7 1 11

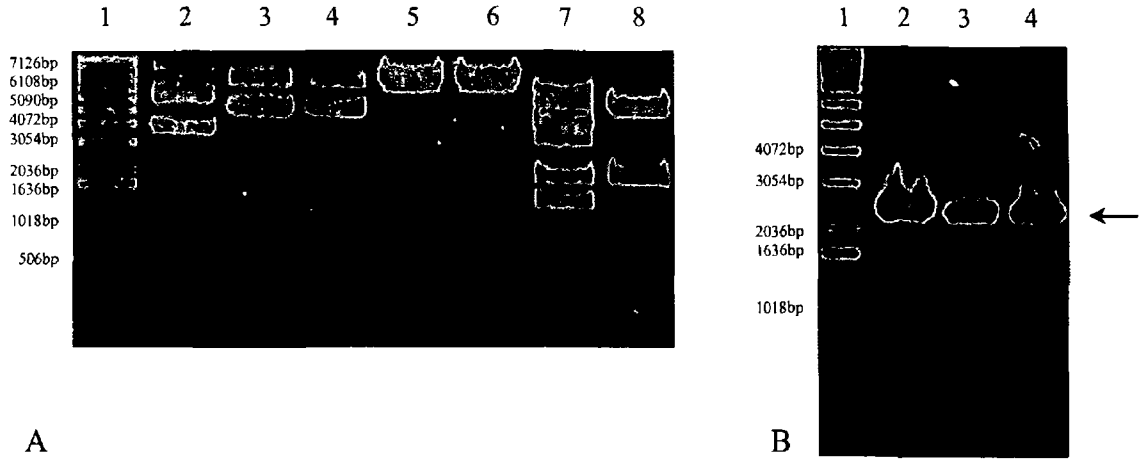


Figure 7 1 10 Verification Digest of pPOB12

Restriction digest of pPOB12 analysed on 0 7% agarose gel (A) Lane 1, 1kb DNA ladder, Lane2, pcDNA3-HA uncut, Lane 3 & 4, pPOB12 uncut, Lane 5, *Bam*HI, Lane 6, *Eco*RI, Lane 7, *Pst*I, Lane 8, *Bam*HI and *Eco*RI (B) Lane 1, 1kb DNA ladder, Lane 2, PCR fragment *Bam*HI-A-Fap7--*Xho*I-Fap8, Lane 3, PCR fragment *Bam*HI-Kos1-Fap7--*Xho*I-Fap8, Lane 4, PCR fragment BamHI-Kos2-Fap7--XhoI-Fap8

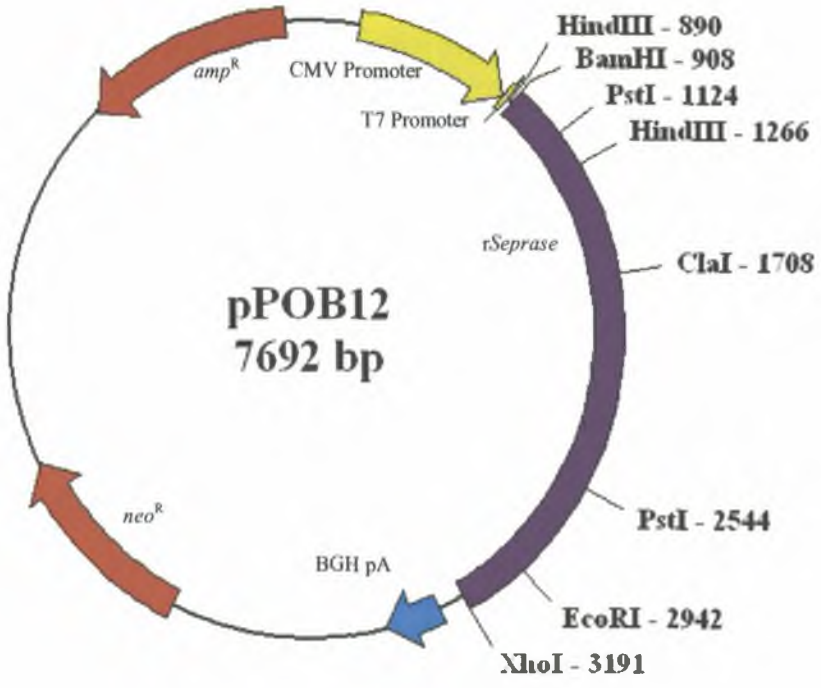


Figure 7.1.11 pPOB12 Plasmid Map

The *rSeprase* gene fragment (purple) is inserted into the vector pcDNA3-HA. The CMV and T7 promoter (yellow) are up stream of the inserted gene fragment. Ampicillin and Neomycin resistance genes (amp^R & neo^R) are shown in red. HA epitope of the vector is cloned in *Hind*III and *Bam*HI (D'Souza et al. 2000). Generated using pDRAW32 (Section 2.10.10).

Figure 7.1.12 illustrates the alignment of pPOB12 with FAP (GI 16933539). The sequencing results confirm the ORF of FAP with no base changes or frame shift errors. pPOB12 was used as a template for further cloning procedures.

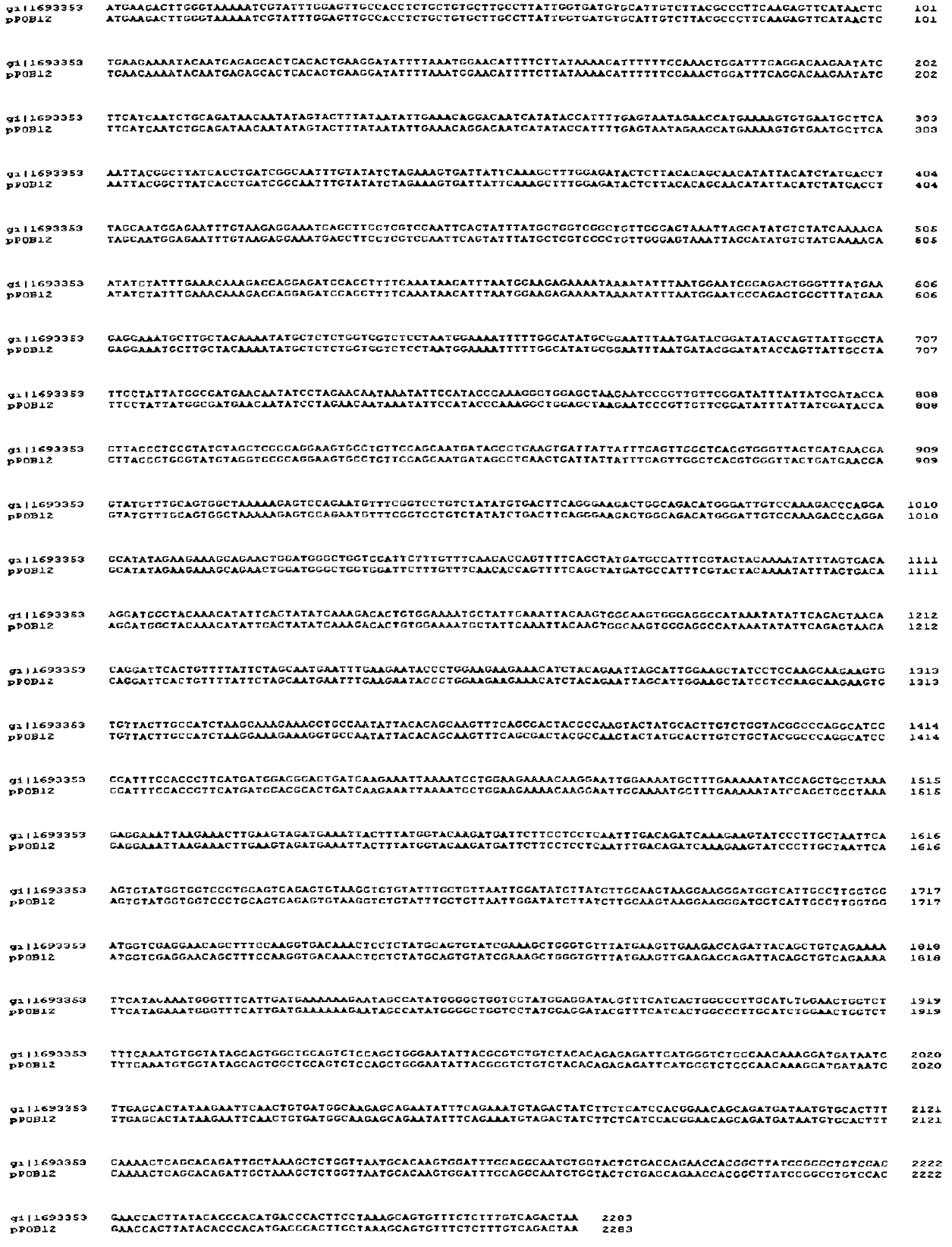


Figure 7.1 12 DNA Sequencing of pPOB12

Sequence of pPOB12 aligned with that of Seprase DNA sequencing confirms the successful cloning of *rSeprase* Illustrated using GenDoc (Section 2 10 10) Sequencing data is given in Appendix C

The primers *Bam*HI-Koz1-Fap7 and *Bam*HI-Koz2-Fap7 were designed (Table 2 10 2) so that they contained a Kozak sequence, 5'-*GCC(A/G)CCATGG*-3' for initiation of translation of the protein Figure 7 1 10 (B) shows the amplified *Seprase* gene fragments using pPOB12 as template, Phusion DNA Polymerase and the primers *Bam*HI-Koz1-Fap7, *Bam*HI-Koz2-Fap7 and *Xho*I-Fap8 The fragments were concentrated according to Section 2 10 3 4 and restricted with *Bam*HI and *Xho*I The restricted PCR fragments were then ligated with the treated pcDNA3-HA and pcDNA3 vectors and transformed according to 2 10 5 2, into the *E coli* strain XL-10 Gold The use of the two vectors would distinguish the effect, if any, the HA tag would have on the refolding and activity of the enzyme

Clones were screened according to Section 2 10 3 2 1 and candidate clones were further analysed by restriction analyses A restriction digest of these clones (pPOB15, 16, 17, 18) analysed by agarose gel electrophoresis is shown in Figure 7 1 13 Restriction with *Bam*HI and *Xho*I produces a band corresponding to the expected 2283bp fragment containing the *rSeprase* sequence Single restriction with *Bam*H1 and *Eco*RI produced bands corresponding to the expected 7692pb linearised plasmid Restriction with *Hind*III produced bands corresponding to the expected 376bp and 7316bp Restriction with *Pst*I produced bands corresponding to the expected 1420bp, 2007bp and 3983bp There also seems to be some uncut and partially cut vector DNA in Lane 7 (Figure 7 1 10) The successful cloning of *rSeprase* was confirmed by DNA sequencing (Section 2 10 10) The sequencing data is given in Appendix C A map of pPOB15-18 is shown in Figure 7 1 14

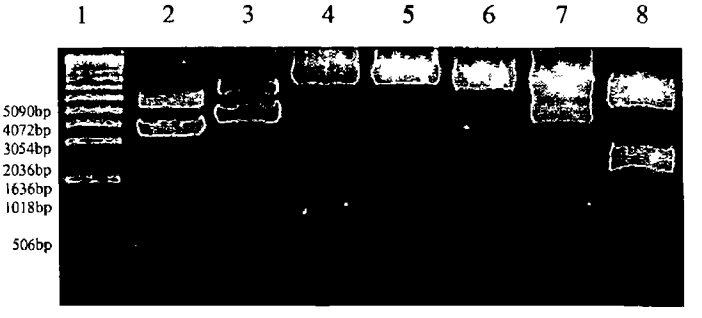


Figure 7 1 13 Verification Digest for pPOB15-18

Restriction digest of pPOB15-18 analysed on 0 7% agarose gel (A) Lane 1, 1kb DNA ladder, Lane2, pcDNA3-HA/pcDNA3 uncut, Lane 3, pPOB15-18 uncut, Lane 4, *Bam*HI, Lane 5, *Eco*RI, Lane 6, *Hind*III, Lane7, *Pst*I, Lane 8, *Bam*HI and *Eco*RI

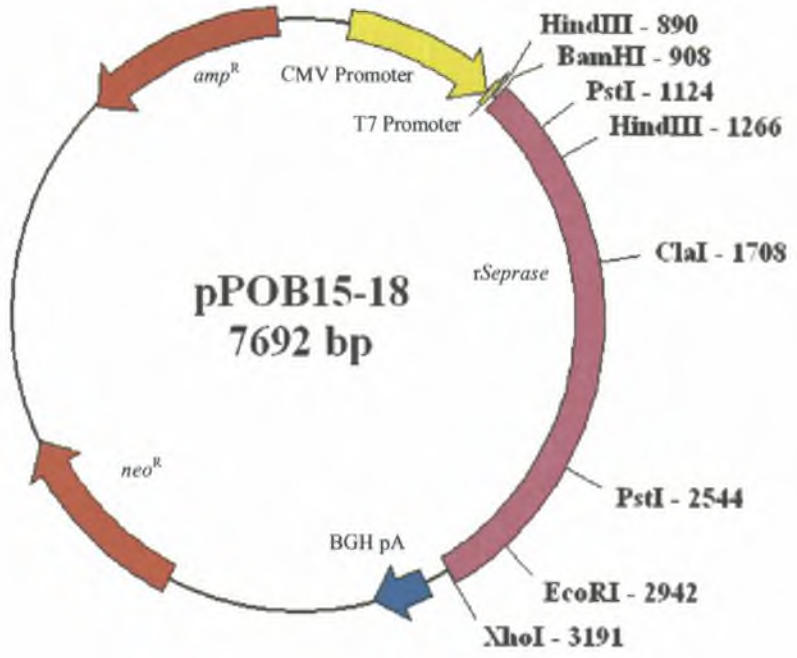


Figure 7.1.14 pPOB15-18 Plasmid Map

The *rSeprase* gene fragment (purple) is inserted into the vectors pcDNA3/pcDNA3-HA. The CMV and T7 promoter (yellow) are up stream of the inserted gene fragment. Ampicillin and Neomycin resistance genes (amp^R & neo^R) are shown in red. HA epitope of the vector is cloned in *Hind*III and *Bam*HI (D'Souza, 2000). Generated using pDRAW32 (Section 2.10.10).

The mammalian constructs pPOB12, pPOB15, pPOB16, pPOB17, pPOB18 and the vector pcDNA3 were transfected into the DG75 cell line according to Section 2.10.12.4. Stable transfectants were then tested for Seprase activity according to Section 2.9.2. Table 7.1.1 identifies the DG75 cell line and respective transfected construct. Figure 7.1.15 illustrates the expression profile of the various cell lines. Results show that none of the transfectants display Seprase activity, although they do express active Prolyl Oligopeptidase. It was necessary to reveal the reason why Seprase activity was not detected. The expression of Seprase was investigated firstly at the transcription level to ensure the efficient transcription of the *Seprase* gene and then at the translation level.

Clone	Cloning Vector	Cloning Sequence Forward Primer	Actual Cloning Vector	Actual Sequence 5'	Frame	Transfected Cell Line
pPOB12	pcDNA3-HA	*Bam*HI-A-Fap7	pcDNA3-HA	*Bam*HI-A-Fap7	In frame with HA tag	DG75-1
pPOB15	pcDNA3	*Bam*HI-Koz2-Fap7	pcDNA3-HA	*Bam*HI-Koz2-Fap7	Out of frame with HA tag	DG75-2
pPOB16	pcDNA3-HA	*Bam*HI-Koz1-Fap7	pcDNA3-HA	*Bam*HI-A-Fap7	In frame with HA tag	DG75-3
pPOB17	pcDNA3-HA	*Bam*HI-Koz2-Fap7	pcDNA3-HA	*Bam*HI-Koz2-Fap7	In frame with HA tag	DG75-4
pPOB18	pcDNA3	*Bam*HI-Koz1-Fap7	pcDNA3-HA	*Bam*HI-Koz1-Fap7	Out of frame with HA tag	DG75-5
-	pcDNA3	-	pcDNA3-HA	-	-	DG75-6

Table 7.1.1 Summary of cloning strategy of pcDNA3 and pcDNA3-HA vectors

Table shows the planned cloning procedure for each construct. The constructs listed were transfected into the DG75 cell line. Sequences containing varying Kozak sequences were cloned into the 2 vectors. pcDNA3 and pcDNA3-HA were used to determine if the HA tag effected the refolding and activity of the Seprase protein.

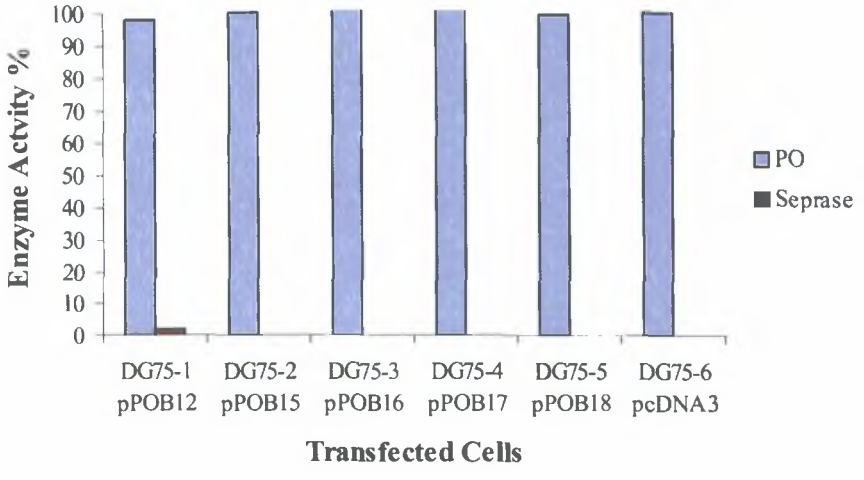


Figure 7.1.15 Expression profile of transfected cell lines

Transfected cell lines are indicated with their respective transfected plasmid. The transfected cell lines were tested for Seprase activity and Prolyl Oligopeptidase activity. PO activity was detected at the same level in all cell lines. DG75-1, containing the plasmid pPOB12, showed some Seprase activity but that may be residual PO activity not inhibited by JTP-4819. The levels of Seprase activity would lead to the conclusion that the stable transfectants did not express active Seprase activity.

In order to investigate the efficient transcription of r*Seprase*, RNA extracted from the stable transfectants (Section 2 10 3 1) was treated with Deoxyribonuclease I (DNase I) according to Section 2 10 7 to remove any DNA contamination PCR of β-actin and full length *Seprase* gene was performed on the treated RNA to ensure that the DNA was digested thoroughly No bands were present when the PCR products were run on an agarose gel RT-PCR was then performed on the treated mRNA and the resulting PCR fragments were run on an agarose gel illustrated in Figure 7 1 16 Figure 7 1 16 (A) illustrates the PCR of full length *Seprase* gene Lanes 2-6 show the PCR fragments corresponding to the *Seprase* gene (2283bp fragment) run on an agarose gel, illustrating that *Seprase* mRNA is present in the transfected cell lines DG75-1-5 Lane 7 is lacking a band corresponding to the Seprase fragment, illustrating that DG75-6 does not contain *Seprase* mRNA This was expected as DG75-6 only contains the vector pcDNA3-HA Figure 7 1 16 (B) illustrates the RT-PCR of β-actin, housekeeping gene Lanes 2-7 show the PCR fragments corresponding to the β-actin fragment (360bp) run on an agarose gel This confirms that the RT-PCR reaction was successful in producing good quality cDNA and that β-actin mRNA is present in the transfected cell lines DG75-1-6

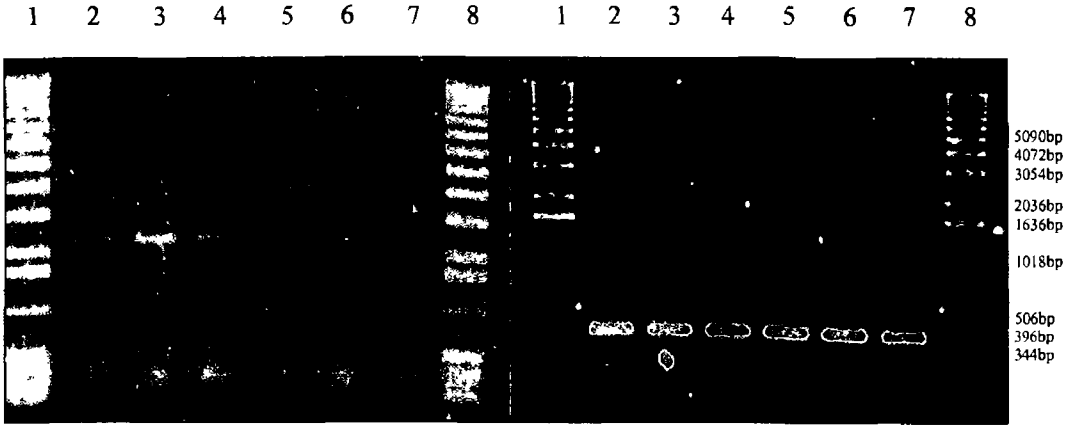


Figure 7 1 16 mRNA Analysis of stable transfected cell lines

Analysis of mRNA from the stable transfectants mRNA extracted from the cell lines was treated with DNase I to digest DNA (A) PCR of the full length Seprase gene using the DNase I treated mRNA PCR fragments produced were run on 0 7% agarose gel Lane 1, 1kb DNA ladder, Lane 2, DG75-1, Lane 3, DG75-2, Lane 4, DG75-3, Lane 5, DG75-4, Lane 6, DG75-5, Lane 7, DG75-6, Lane 8, 1 kb DNA ladder (B) PCR of β-Actin using the DNase I treated mRNA PCR fragments produced were run on 1% agarose gel Lane 1, 1kb DNA ladder, Lane 2, DG75-1, Lane 3, DG75-2, Lane 4, DG75-3, Lane 5, DG75-4, Lane 6, DG75-5, Lane 7, DG75-6, Lane 8, 1 kb DNA ladder

Analysis of Seprase protein expression was performed using western blot analysis. Western blot analysis was carried out on lysates from the transfectants (DG75 1-6), according to Section 2.10.12.6.2. The western blot would confirm if Seprase protein was being expressed, albeit in an inactive form. Figure 7.1.17 shows multiple protein bands reacting with the HA-Tag (6E2) Mouse Monoclonal antibody (Table 2.10.12.6). This was unexpected as the membrane had been washed to avoid any non-specific binding of the antibody. It was also observed that the transfectants containing the pcDNA3 vector (Lane 4, Lane 7, Lane 8) also reacted to the same extent as those containing the pcDNA3-HA vector (Lane 3, Lane 5, Lane 6).

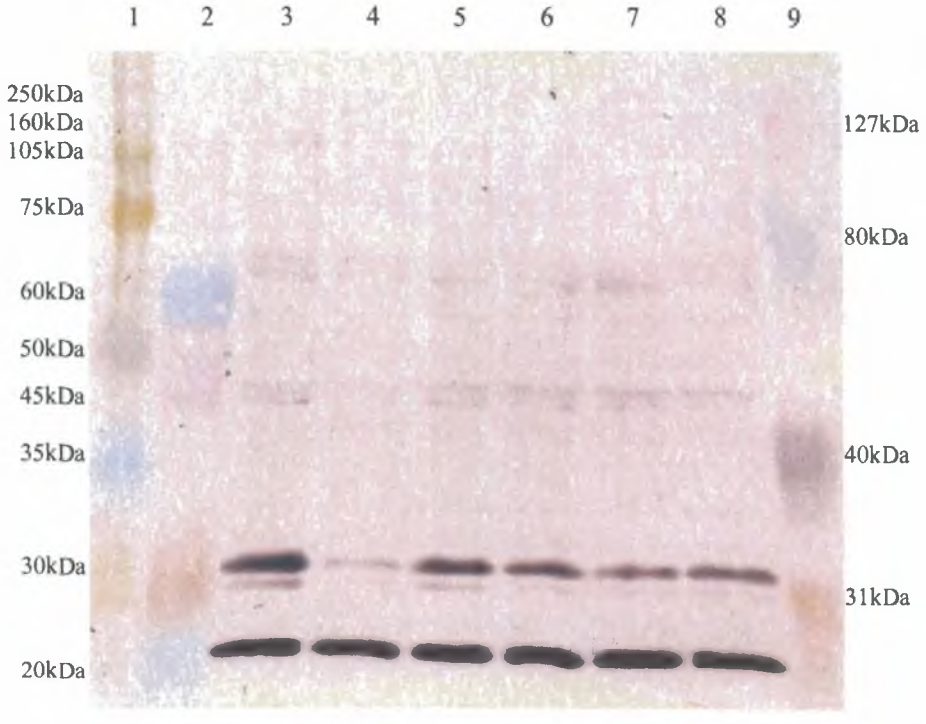


Figure 7.1.17 Western Blot Analysis of rSeprase expression

Western blot analysis of DG75 stable transfectants. Lane 1, Full Range Rainbow Marker; Lane 2, ColourBurst Marker; Lane 3, DG75-1; Lane 4, DG75-2; Lane 5, DG75-3; Lane 6, DG75-4; Lane 7, DG75-5; Lane 8, DG75-6; Lane 9, Kaleidoscope Pre-stained Standards.

The 5' region upstream of the cloned Seprase sequences of vectors pcDNA3, pcDNA3-HA, pPOB12-18 were analysed using DNA sequencing. This established that all vectors contained the HA epitope (Figure 7.1.18) and that the original vectors had been mislabeled prior to being gifted. Table 7.1.1 lists the actual sequences of the constructs transfected into the DG75 cell line. The sequencing also showed if the r*Seprase* sequences cloned into the pcDNA3-HA vector were in frame with the HA start codon and Kozak sequence.

Figure 7.1.18 also illustrates that all constructs were cloned into the pcDNA3-HA with either Kozak sequence (Koz1 or Koz2) as planned. However, the pPOB16 construct lacked the Kozak 1 sequence and the pPOB17 construct contained an extra Cytosine (C) base. Sequencing also confirmed the frame of the *Seprase* gene. The constructs pPOB12, pPOB16 and pPOB17 were found to be in frame with the HA epitope. Therefore, the respective stably transfected cell lines DG75-1, DG75-3 and DG75-4 should have expressed the Seprase protein fused to the HA tag. However, the stably transfected cell lines DG75-2 and DG75-5 contained constructs that were out of frame with the N-terminal HA tag. They did however contain a Kozak sequence, so therefore they should also have expressed the Seprase protein which was not fused to the HA epitope.

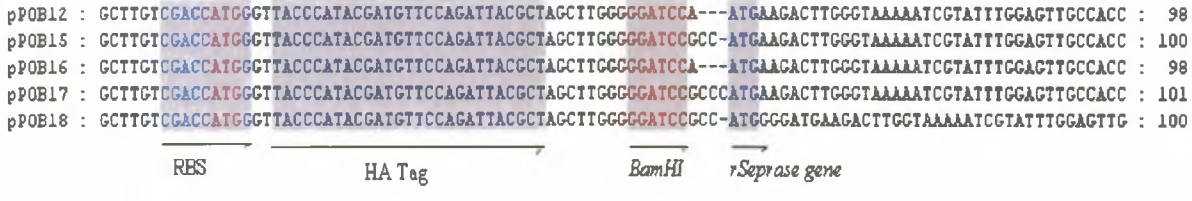


Figure 7.1.18 Alignment of Mammalian Expression Clones 5' Upstream Sequence

The 5' region of the *Seprase* cloned fragment was sequenced to determine the frame and sequence of the region. Sequencing determined that all constructs contain the HA tag (amino acid sequence YPYDVPDYA). The ribosome binding site (RBS – Kozak sequence) is indicated, containing the start codon (ATG) for the HA tag epitope. The *Bam*HI cloning restriction site is indicated in . The start codon of the *Seprase* gene is indicated in blue.

The cell lines DG75-1 and DG75-6 were chosen to be analysed using immunocytochemistry according to Section 2 10 12 7 and illustrated in Figure 7 1 19 DG75-1 and DG75-6 nuclei were stained with DAPI, which identified the cells under the fluorescent microscope and can be seen in Figure 7 1 19 (A) and (C) The same cells were then analysed for binding of the secondary Alexa Fluor 488 antibody Figure 7 1 19 (B) shows that the DG75-1 cells are fluorescing, indicating a positive reaction of the secondary antibody binding with the HA tag fused to protein Figure 7 1 18 (D) shows that the DG75-6 cells do not bind the secondary antibody These results suggest that the Seprase protein has been expressed in the DG75-1 cell line fused to the N-terminal HA tag, albeit at very low levels

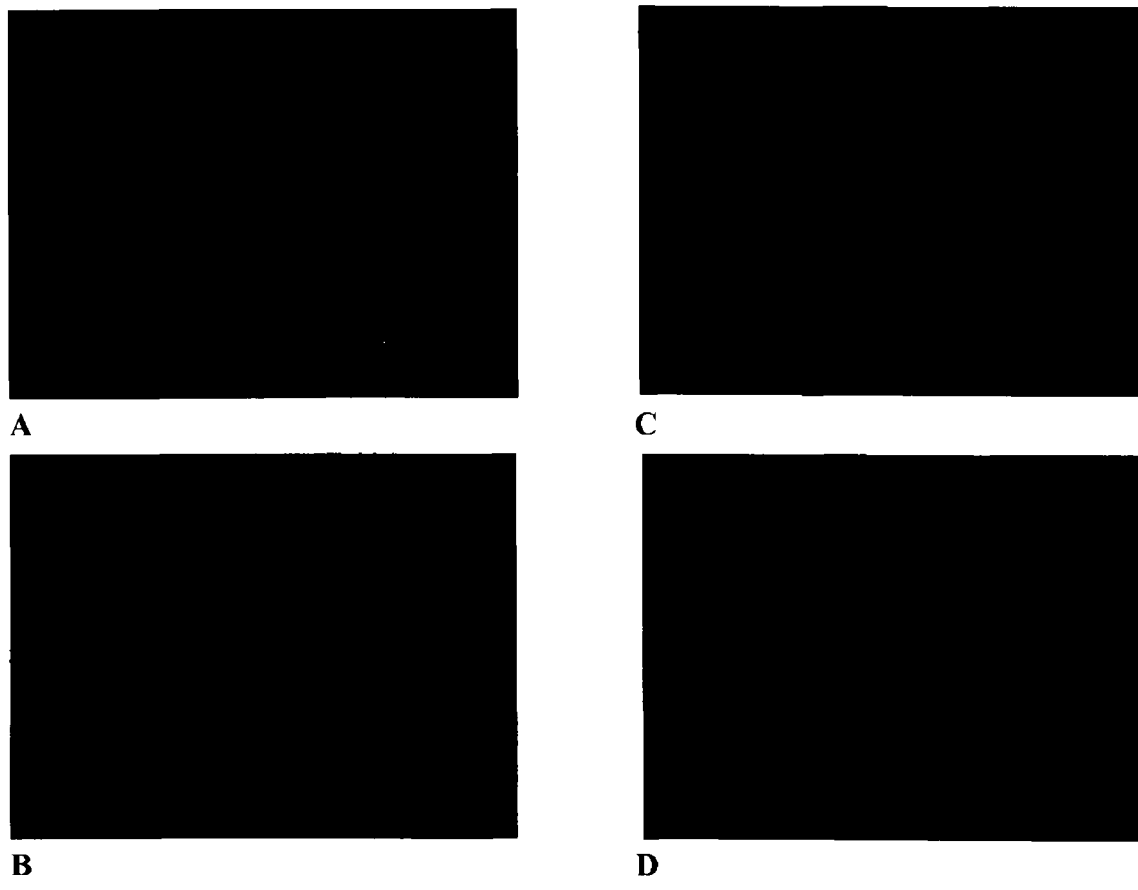


Figure 7 1 19 Immunocytochemistry of stably transfected cell lines

Immunocytochemistry was performed on the transfected cell lines DG75-1 and DG75-6 (A) and (B) relate to DG75-1 (A) DG75-1 cells nuclei stained with DAPI (B) DG75-1 cells fluorescing with the secondary Alexa Fluor 488 antibody Shows positive binding with the HA tag (C) DG75-6 cells nuclei stained with DAPI (D) DG75-6 cells fail to fluoresce with the secondary Alexa Fluor 488 antibody Shows negative binding with the HA tag

Overall, these results illustrate that transcription of *Seprase* mRNA from the 5 constructs is occurring Western blot analysis of the stable transfectants did not provide any conclusive results with regards to translation of the *Seprase* mRNA into protein Immunocytochemistry results do suggest, however, that the *Seprase* mRNA is translated into protein, albeit in an inactive form

7 1 3 Cloning of human *Seprase* gene into pQE30-Xa

Although there was some probability that catalytically active recombinant human Seprase (r*Seprase*) would not be expressed in *E coli*, it was decided to investigate this possibility Previous studies have shown that it is possible to express a catalytically active mammalian protein, PAP1, in *E coli* (Vaas, 2005) The prokaryotic expression vector pQE30-Xa (Table 2 10 3 and Figure 2 10 2) was chosen to clone and express the *Seprase* gene The vector pQE30-Xa features an optimised hybrid promoter-operator element consisting of the phage T5 transcriptional promoter and a *lac* operator sequence, which increases *lac* protein (LacI) binding, enabling repression of this strong promoter Associated with this operator region is a synthetic ribosome binding site (RBS) designed for high translation rates The vector encodes a Factor Xa Protease recognition site which is found 3' of the 6xHis-tag and 5' of the multiple cloning site Factor Xa Protease treatment results in a recombinant protein free of the vector-derived 6xHis-tag at the N-terminus

The vector was restricted with *Bam*HI and *Sal*I, treated with Antarctic Phosphatase according to Section 2 10 8 1 and concentrated according to Section 2 10 3 4 The primers *Bam*HI-Fap7 and *Xho*I-Fap8 were designed and were used to amplify the *Seprase* gene with *Bam*HI and *Xho*I restriction endonuclease sites and to clone the *Seprase* gene into pQE30-Xa vector The *Seprase* gene was amplified using pPOB12 as a template for PCR, with the primers *Bam*HI-Fap7 and *Xho*I-Fap8 and using Phusion DNA Polymerase (Section 2 10 7 2) The products formed were run on an agarose gel (Figure 7 1 20) The PCR products were concentrated using the method described in Section 2 10 3 4 and restricted with the respective endonuclease restriction enzymes (*Bam*HI and *Xho*I) according to Section 2 10 7 The restricted

PCR fragments were then ligated with the treated pQE30-Xa vector and transformed according to 2 10 5 2, into the *E coli* strain XL10-Gold

Clones were screened according to Section 2 10 3 2 1 and a candidate clone was further analysed by restriction analyses A restriction digest of this clone (pQPOB5) analysed by agarose gel electrophoresis is shown in Figure 7 1 20 Single restriction with *Bam*H1 and *Xho*I produced bands corresponding to the expected 5764pb linearised plasmid Restriction with *Eco*RI produced bands corresponding to the expected 2187bp and 3577bp Restriction with HindIII produced bands corresponding to the expected 1891bp and 3873bp Restriction with *Pst*I produced bands corresponding to the expected 609bp, 1420bp and 3735bp The successful cloning of *rSeprase* was confirmed by DNA sequencing (Section 2 10 10) A map of pQPOB5 is shown in Figure 7 1 21 The sequencing data is given in Figure 7 1 22 and Appendix C

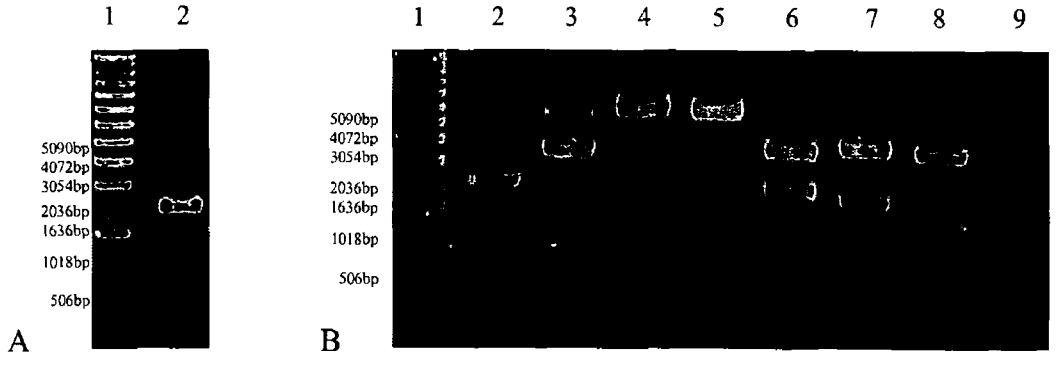


Figure 7 1 20 Verification digest of pQPOB5

(A) 0 7% agarose gel Lane 1, 1kb DNA ladder, Lane 2, PCR fragment *Bam*HI-Fap7--*Xho*I-Fap8 (B) 0 7% agarose gel Lane 1, 1 kb DNA ladder, Lane 2, pQE30-Xa uncut, Lane 3, pQPOB5 uncut, Lane 4, *Bam*HI, Lane 5, *Xho*I, Lane 6, *Eco*RI, Lane 7, *Hind*III, Lane 8, *Pst*I, Lane 9, 1 kb DNA ladder

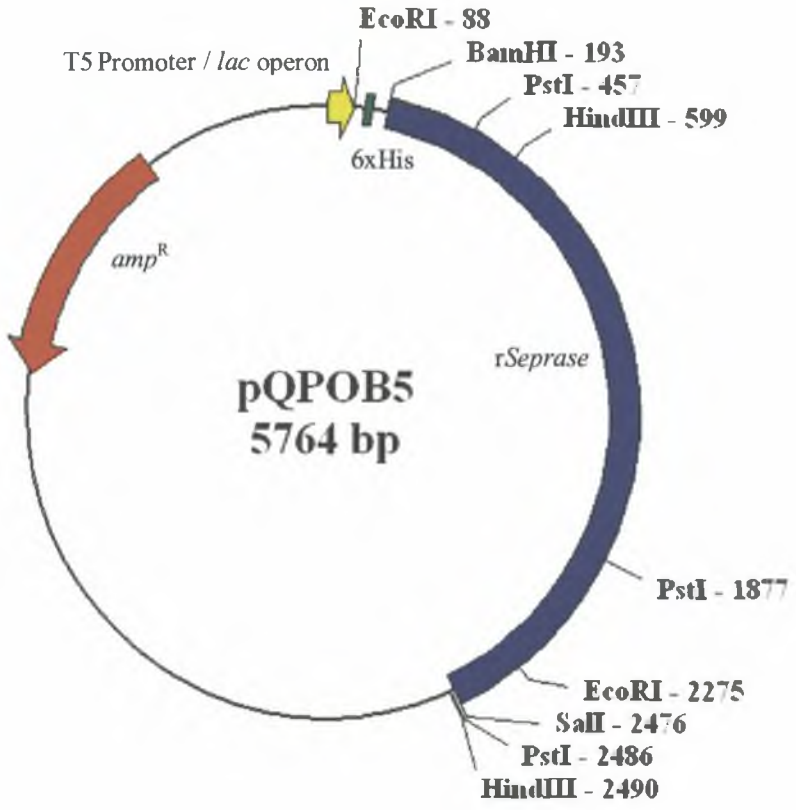


Figure 7.1.21 pQPOB5 Plasmid Map

The *rSeprase* gene fragment (purple) is inserted into the expression vector pQE30-Xa. T5 promoter (yellow) is up stream of the inserted gene fragment. The 6xHis coding sequence (green) is situated at the 5' end of the MCS. Ampicillin resistance gene (*amp*^{R}) is shown in red. Generated using pDRAW32 (Section 2.10.10).

```
gi|1693353 : ----------------------------------------------------------------------ATGAAGACTTGGGTAAA :  17
pQPOB5     : ATGAGACGATCGCATCACCATCACCATCACGGATCTGGCTCTGGATCTGGT                  TGGATCC..................... : 101
             Start Codon    6 x His Tag                        Factor Xa R.S  Factor Xa C.S.  BamHI  rSeprase gene

gi|1693353 : AATCGTATTTGGAGTTGCCACCTCTGCTGTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATAACTCTGAAGAAAATACAATGA :  118
pQPOB5     : ..................................................................................................... : 202
```

Figure 7.1.22 Sequence of pQPOB5 aligned with that of *Seprase*.

DNA sequencing confirms the successful cloning of *rSeprase* in the expression vector pQE30-Xa. Figure illustrates only the 5' region of the *Seprase* gene. Dots (...) indicate that the nucleic acid base at that position is conserved in the alignment. Start codon of pQE30-Xa vector is indicated (ATG) upstream of the N-terminal 6xHis Tag. The Factor Xa recognition site (R.S.) is highlighted in followed by the Factor Xa cleavage site (C.S.). The *Bam*HI recognition site used for cloning of the Seprase gene into the vector pQE30-Xa is indicated in blue. Illustrated using GenDoc (Section 2.10.10). Sequencing data is given in Appendix C.

Expression of recombinant Seprase protein (rSeprase) in both XL10-Gold and BL21 (DE3) was carried out (Section 2 10 11) and the enzyme was purified using IMAC (Section 2 10 11 2) The XL10-Gold strain contains F factors, which carry the *lacI*^{q} allele The *lacI*^{q} allele is a promoter mutation that expresses the LacI repressor protein at high levels, resulting in strong repression of the *lac* promoter unless the inducer IPTG is added, which binds and inactivates LacI and thus induces transcription from the *lac* promoter BL21 is a strain deficient in the OmpT87 protease, purportedly allowing higher recovery of heterologous recombinant proteins

The cleared lysate and purification fractions obtained from these expression cultures were analysed by SDS-PAGE (Section 2 5) and are shown in Figure 7 1 23 The clear lysate of both expression cultures was assayed for Seprase activity according to Section 2 3 3 Neither of the clear lysates had Seprase activity (data not shown) Figure 7 1 23 clearly shows that there was no expression of Seprase protein in either *E coli* strain The elution fractions in both purification schemes (see Figure 7 1 23) contains many proteins of varying molecular weights, none of which correspond to the 97kDa Seprase protein subunit

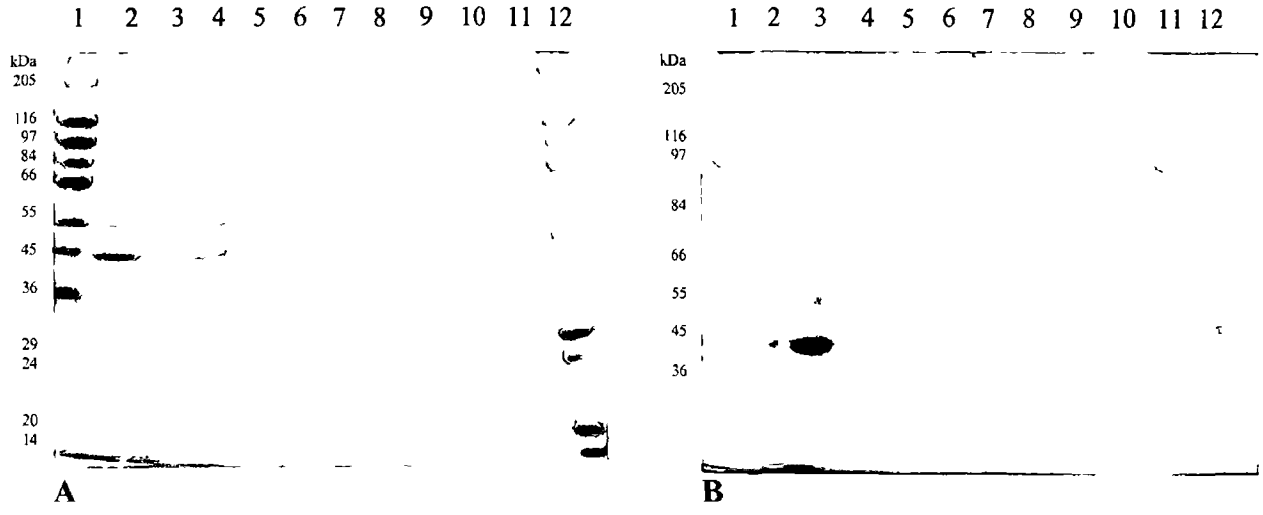


Figure 7 1 23 Expression and Purification in rSeprase in *E coli* XL10-Gold and BL21 (DE3)

Analysis of rSeprase expression and purification in *E coli* from pQPOB5 (Figure 7 1 21) by 10% SDS-PAGE (A) XL10-Gold Lane 1, High Range Sigma Marker, Lane 2, Clear lysate, Lane 3, Flow through, Lane 4, First wash, Lane 5, Second wash, Lane 6, Third wash, Lane 7, First elution, Lane 8, Second elution, Lane 9, Third elution, Lane 10, Fourth elution, Lane 11, Fifth elution, Lane 12, Wide Range Sigma Marker (B) BL21 (DE3) Lane 1, High Range Sigma Marker, Lane 2, Clear lysate, Lane 3, Flow through, Lane 4, First wash, Lane 5, Second wash, Lane 6, First elution, Lane 7, Second elution, Lane 8, Third elution, Lane 9, Fourth elution, Lane 10, Fifth elution, Lane 11, Wide Range Sigma Marker, Lane 12, Kaleidoscope Marker

Genes in both prokaryotes and eukaryotes exhibit a non-random usage of synonymous codons (a codon table is given in Appendix D) This means there is a bias toward one or two codons in most degenerate codon families The frequency of codon usage is usually reflected by the abundance of the associated tRNAs in different cell backgrounds This implies that recombinant heterologous genes enriched with codons rarely used in *E coli* may experience poor translation efficiency (Makrides, 1996) Arginine codons AGG and AGA are the least frequently used in *E coli* and the tRNAs that recognise them are among the least abundant Consecutive AGG or AGA codons can lead to a high level of frame shifting (Rosenberg et al, 1993) If the DNA sequence encoding the recombinant protein contains several such codons, a variety of truncated protein products, particularly from large recombinant proteins, is likely to be synthesised Figure 7 1 24 shows the nucleotide sequence coding for r*Seprase*, as it is found on the pQPOB5 construct Codons of minimal usage in *E coli* have been indicated, according to international DNA sequence databases (Nakamura, 2000)

Figure 7 1 24 clearly shows that the Seprase gene contains 82 rare codons and consecutive AGG or AGA codons are indicated (*) It is possible that Seprase was being expressed in truncated forms This could explain the numerous protein bands visualised in the elution fractions visualised in Figure 7 1 23 A possible solution would be to place the 6xHis affinity tag at the C-terminus, and therefore only full length proteins will bind to the Ni-NTA resin during the purification procedure, the truncated forms will not bind and will be removed in the flow-through and the wash fractions However, in the case of active Seprase expression, the C-terminal 6xHis tag could affect the folding of the protein and the subsequent dimerisation

E coli Rosetta is a derivative of the strain BL21 carrying the pRARE plasmid pRARE (Novy et al, 2001) encodes genes for six tRNAs that recognise codons considered rare in *E coli* The use of pRARE during the expression of heterologous proteins is one strategy to overcome a possible codon bias of *E coli* This system does not provide for all the rare codons but it could be used to express the 97kDa subunit of Seprase

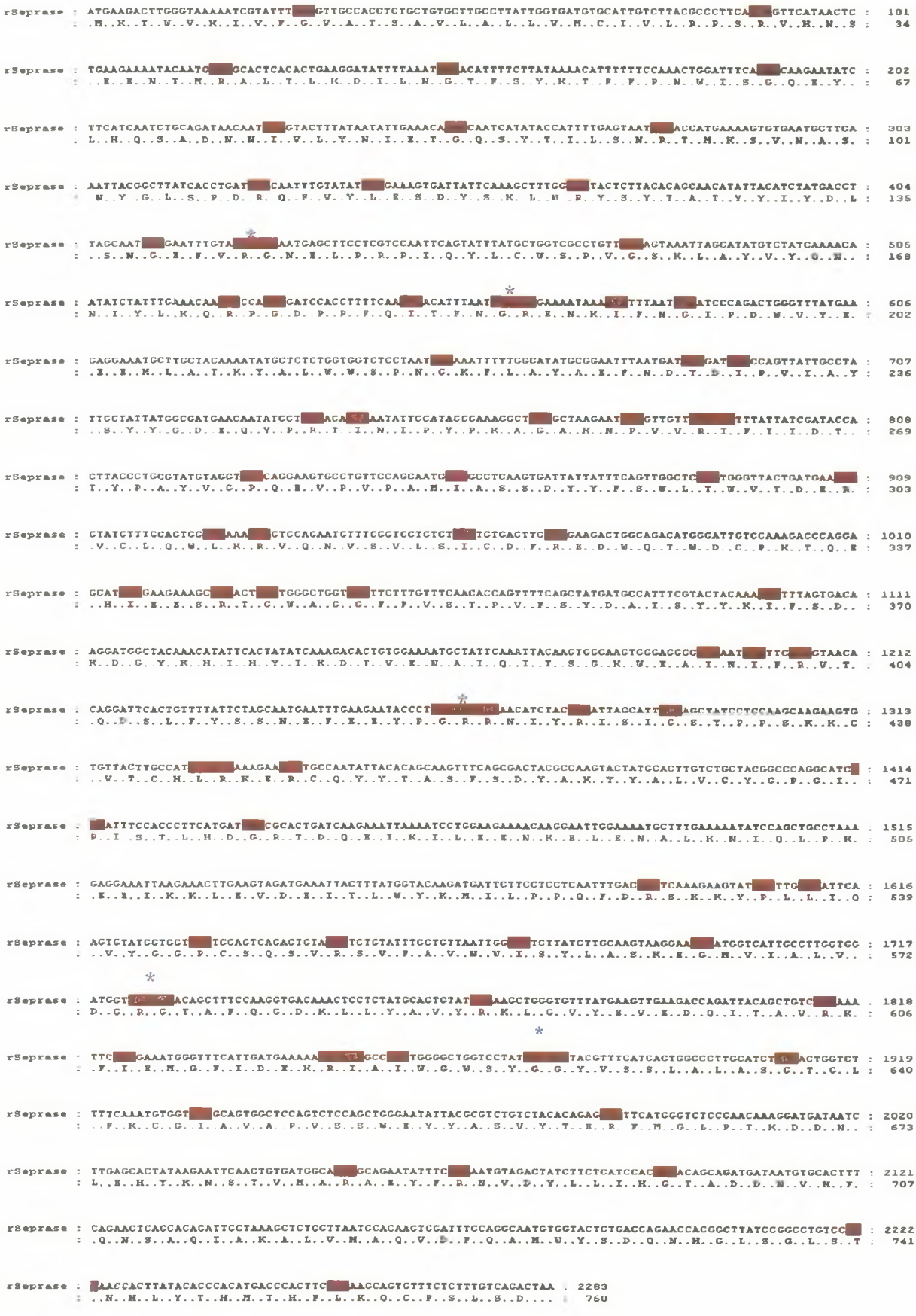


Figure 7.1.24 *E. coli* codon bias relative to r*Seprase*

Nucleotide sequence coding for r*Seprase*, as found on pQPOB5. It clearly shows that the *Seprase* gene contains 82 rare codons and consecutive AGG or AGA codons are indicated (*). Amino acids corresponding to each codon are specified. Codons considered rare in *E. coli* are highlighted red. For amino acid information see Appendix D. Illustrated using GenDoc (Section 2.11).

7 1 4 Cloning of human *Seprase* gene into pIRES-hrGFP II

The cloning strategy involving mammalian expression vector pcDNA3-HA failed to produce an active form of the protease Seprase Therefore the mammalian expression vector pIRES-hrGFP II (Table 2 10 3 and Figure 2 10 4) was chosen to re-clone and express Seprase The primer *Bam*HI-Koz3_for was designed (Table 2 10 2) so that it contained a Kozak sequence, 5'-*GCCACCATGG*-3' for initiation of translation of the protein Figure 7 1 23 (B) shows the amplified Seprase gene fragments using pPOB12 as template, Phusion DNA Polymerase and the primers *Bam*HI-Koz3_for, *Not*-I-Fap8_1rev and *Not*-I-Fap8_2rev The fragments were concentrated according to Section 2 10 3 4 and restricted with *Bam*HI and *Xho*I The restricted PCR fragments were then ligated with the treated pIRES-hrGFP II vector and transformed according to 2 10 5 2, into the *E coli* strain XL-10 Gold The primer *Not*-I-Fap8_1rev provides a stop codon (TAA) allowing for the Seprase gene to be expressed without a C-terminal fused FLAG tag The *Not*-I-Fap8_2rev primer however allows for the *Seprase* gene to be expressed fused to the FLAG tag The two reverse primers would distinguish the effect, if any, the C-terminal FLAG tag would have on the refolding and activity of the enzyme

Clones were screened according to Section 2 10 3 2 1 and candidate clones were further analysed by restriction analyses A restriction digest of these clones (pIRPOB1 and pIRPOB2) analysed by agarose gel electrophoresis is shown in Figure 7 1 25 (A) Single restriction with *Bam*H1, *Cla*I and *Eco*RI produced bands corresponding to the expected 7776bp linearised plasmid Restriction with *Hind*III produced bands corresponding to the expected 444bp, 3021bp and 4311bp Restriction with *Pst*I produced bands corresponding to the expected 405bp, 1420bp, 2130bp and 3821bp The successful cloning of *rSeprase* was confirmed by DNA sequencing (Section 2 10 10) The sequencing data is given in Figure 7 1 26 and Appendix C A map of pIRPOB1 and pIRPOB2 is shown in Figure 7 1 25

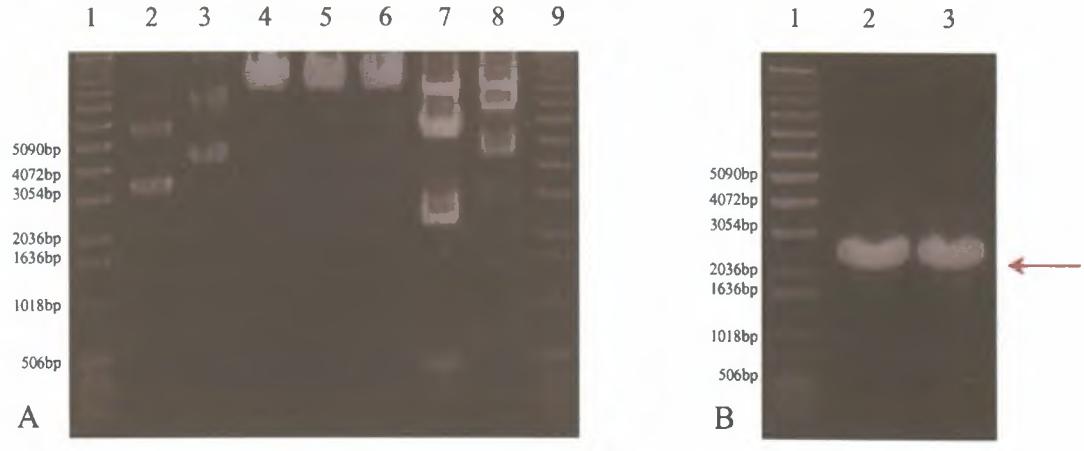


Figure 7.1.25 Verification Digest for pIRPOB1 and pIRPOB2

Restriction digest of pIRPOB1/2 analysed on 0.7% agarose gel: (A) Lane 1, 1kb DNA ladder; Lane2, pIRES-hrGFP II uncut; Lane 3, pIRPOB1/2 uncut; Lane 4, *Bam*HI; Lane 5, *Cla*I; Lane 6, *Eco*RI; Lane 7, *Hind*III; Lane8, *Pst*I; Lane 9, 1kb DNA ladder. (B) Lane 1, 1kb DNA ladder; Lane 2, PCR fragment *Bam*HI-Koz3_for--*Not*I-Fap8_1rev; Lane 3, PCR fragment *Bam*HI-Koz3_for--*Not*I-Fap8_2rev.

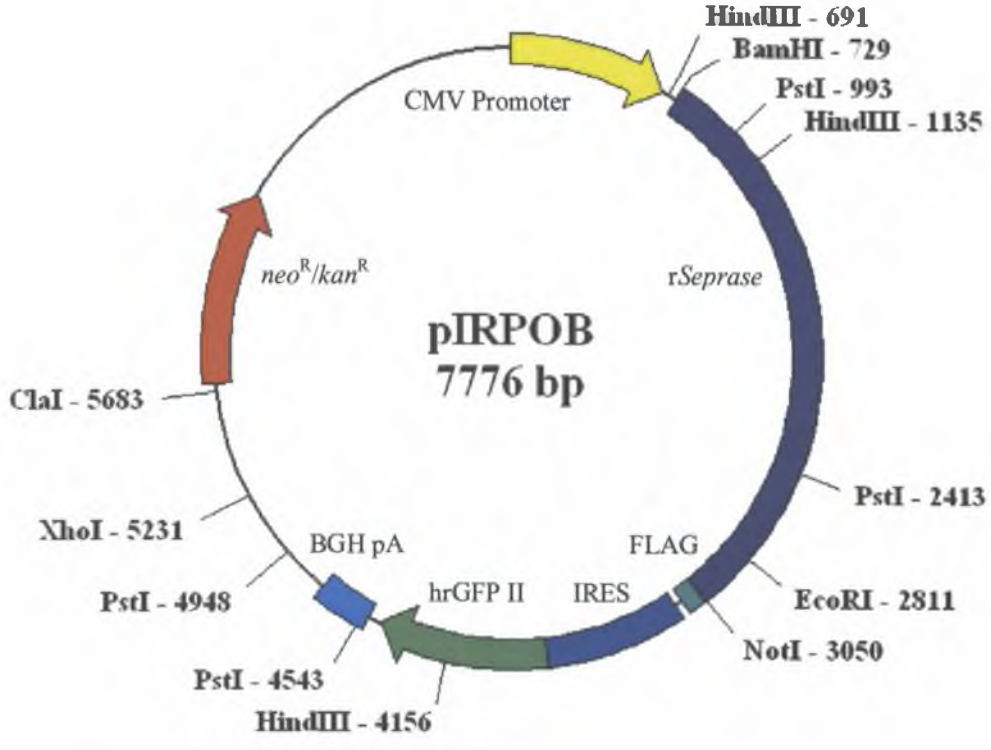


Figure 7.1.26 pIRPOB1 and pIRPOB2 Plasmid Map

The *rSeprase* gene fragment (purple) is inserted into the vector pIRPOB1 and pIRPOB2. CMV promoter (yellow) is up stream of the inserted gene fragment. The 3xFLAG epitope (green) is situated 3' of the inserted gene. Neomycin and Kanamycin resistance genes (neo^R and kan^R) are shown in red. Generated using pDRAW32 (Section 2.10.10).

Figure 7.1.27 Alignment of pIRES-hrGFP II Mammalian Expression Clones

The *Seprase* cloned fragment was sequenced to determine the frame and sequence of the region. The Figure illustrates only the 5' and 3' regions of the *Seprase* gene. Dots (...) indicate that the nucleic acid base at that position is conserved in the alignment.The *Bam*HI cloning restriction site is indicated in red. The ribosome binding site (RBS – Kozak sequence) is indicated (green), containing the start codon (ATG). Sequencing determined that the pIRPOB1 contained the stop codon and pIRPOB2 construct is in frame with the FLAG tag (amino acid sequence DYKDDDDK).

The mammalian constructs pIRES-hrGFP II, pIRPOB1 and pIRPOB2 constructs were transfected into the SW480 cell line according to Section 2.10.12.4. Seprase has been shown to be expressed in colon cancer therefore SW480 was chosen as an expression system as the cell line was derived from a colorectal carcinoma, and it has been found not to express Seprase (Figure 6.1.8). It was thought that by using a colon cell line, Seprase would be expressed in an active form. However, after numerous attempts at transfection by electroporation, no transfectants were obtained. Therefore the procedure for the transfection into the SW480 cell line must be optimised. It may also be necessary to use alternative methods such as transient transfection using lipofection. An alternative to the SW480 cell line should also be sought, as this cell line has recently shown difficulty in other transfection experiments performed by other research groups (data not shown).

Active recombinant human and mouse Seprase has been expressed in mammalian and insect expression systems (Niedermeyer et al., 1997; Park et al., 1999; Scanlan et al., 1994) and these systems have shown to produce low yields of Seprase (Lee et al., 2005a). Active recombinant Seprase proteins have been expressed with both fused N- and C-terminal tags. A soluble Seprase fusion protein was produced with an N-terminal murine CD8α extracellular domain (ECD) (Park et al., 1999). This fusion protein was shown to have both dipeptidyl peptidase and gelatinase activity. Another

group has also produced the extracellular domain of Seprase fused N-terminally to 6xHis tag and C-terminally to a FLAG tag (Cheng et al , 2002) It appears from these reports that the addition of an N- or C-terminal fusion tag to the Seprase protein does not affect the folding, dimerisation and, therefore, the activity of the enzyme

Chapter 8

Clinical Research Study of Human Serum

Currently there are no ideal tumour markers and most are restricted to monitoring cancers once they have been detected and diagnosed using other methods There are also no tumour markers currently recommended for screening of the general population Most tumour markers have too many false positives from benign conditions to make screening feasible Many only clearly identify malignancy once the cancer is sufficiently advanced to make this detection of limited use

There is currently no early detection tumour marker test for breast cancer The aim of this clinical research project was to determine if Seprase levels in serum are a useful breast cancer diagnostic marker By assaying the serum samples for Seprase activity it was possible to obtain strong data confirming that patients with breast cancer have an elevated level of this enzyme, as has been shown previously in studies using tissue

The test could be used during the initial assessment process of patients, in conjunction with the clinical investigation and mammogram/ultrasound It could also be very valuable in situations where mammography cannot detect the breast cancer By using this highly sensitive assay to detect this putative breast cancer serum marker (Seprase), it could be possible to identify patients with early stages of breast cancer, thereby increasing their chances of a better prognosis The initial assessment of patients would not require an invasive procedure such as a biopsy A minimal sample of blood (e g 1-2 ml serum) or a saliva sample is all that would be required to carry out the Seprase biomarker test

Seprase could be used as another important prognostic factor, indicating the possibility of metastasis, not only in breast cancer patients but also for the other epithelial cancers mentioned previously The test could also be used to monitor a patient's response to treatment

Patients were recruited from BreastCheck and the Symptomatic Breast Care Unit of the Mater Misericordiae Hospital (see section 2 11 and Appendix D) The study and consent forms were approved by the Research Ethics Committee of the Mater Misericordiae

Hospital Ref 1/378/1014 The two patient populations consisted of 20 patients 10 patients that have confirmed invasive ductal carcinoma (IDC) and a control cohort of 10 patients, who were undergoing varicose vein surgery Clinical samples were processed according to section 2 11 and were analysed for enzyme activity and protein content, using the fluorimetric assay as outlined in section 2 3 3 and the Biuret assay as in section 2 1 1 Additional relative patient information was collected and is shown in Table 8 1

Sample No	Age	Sex	Diagnosis	Size (cm)	Node status	ER/PR	Grade	Tumour Markers	Staging Liver US/ Bone Scan	Surgery	Follow Up
11	59	F	IDC, ILC	0 9	0/1 (-)	+/+	2	CEA = 1	No M D	W G WLE and sn biopsy	N A
12	61	F	IDC	1 9	0/1 (-)	+/+	2	CEA = 1 CA 15 3 = 10 4	No M D	W G WLE and sn biopsy	N A
13	58	F	IDC	1 1	2/2 (+)	+/+	1	N D	No M D	W G WLE and sn biopsy	Further Axillary Clearance
14	61	F	IDC	0 8	0/1 -	+/+	2	N D	N D	W G WLE and sn biopsy	N A
15	54	F	IDC	4	10/16 (+)	+/+	3	CEA = 3 2	Liver ok Metastatic deposit left femur	Mastectomy and axillary clearance	N A
16	49	F	IDC	2 5	6/23 (+)	N A	1	CEA = 1 0	No M D	W G WLE and sn biopsy	Further mastectomy and axillary clearance
17	58	F	IDC	1 1	(-)	+/+	2	N D	No M D	Mastectomy and axillary clearance	N A
18	52	F	IDC	7	0/10 (-)	+/+	2	N D	No M D	Mastectomy and axillary clearance	Reconstruction
19	65	F	IDC	0 7	0/1 (-)	+/+	1	N D	No M D	W G WLE and sn biopsy	N A
20	N A	F	N A	N A	N A	N A	N A	N A	N A	N A	N A

Table 8 1 Characteristics of the patient cohort supplying serum samples

Sample Numbers are listed as they appear in Figure 8 1 and Figure 8 2 N D not determined N A not available IDC infiltrating ductal carcinoma ILC infiltrating lobular carcinoma CEA normal level 0-5 CA15 3 normal level 0-40 Node status number of nodes positive for tumour cells/number of nodes examined ER estrogen receptor PR progesterone receptor No M D No metastatic disease WG WLE Wire guided WLE

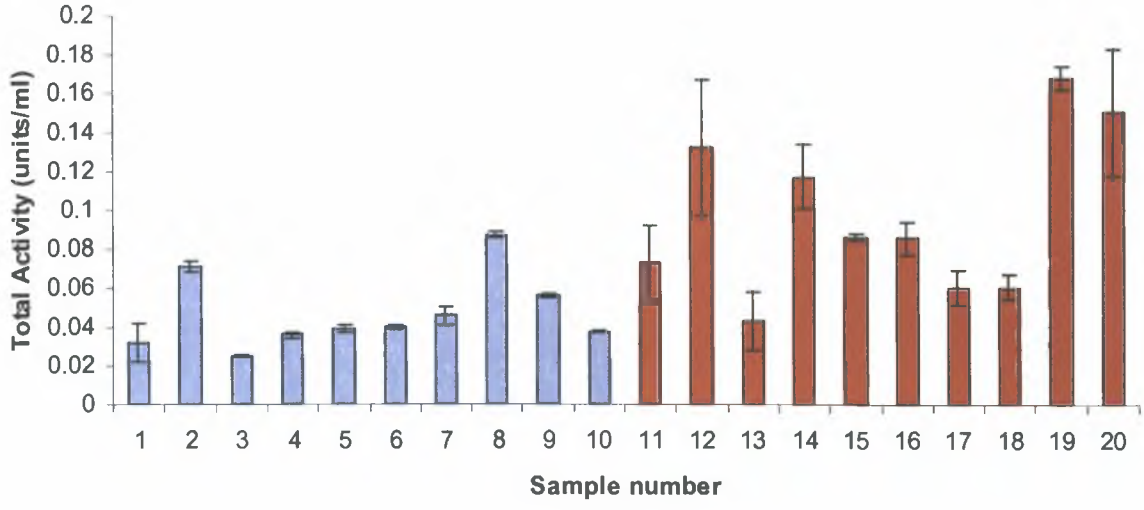


Figure 8.1 Total Activity of Serum Samples from full cohort of patients

The control samples are numbered 1-10 (blue). Those patients with confirmed cases of IDC are numbered 11-20 (red) (see Table 8.1).

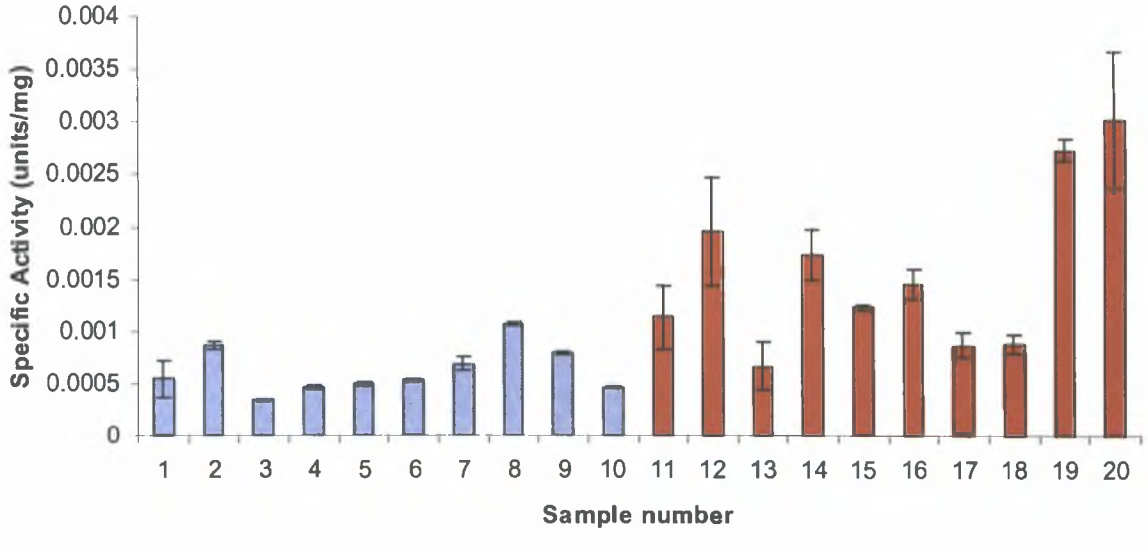


Figure 8.2 Specific Activity of Serum Samples from full cohort of patients

The control samples are numbered 1-10 (blue). Those patients with confirmed cases of IDC are numbered 11-20 (red) (see Table 8.1).

Statistical analysis of both population data sets show that is was normally distributed (see Table 8.2) and, therefore, the Student's t-test was chosen to determine significance. A summary of the statistical analysis of the clinical data is shown in Tables 8.2, 8.3 and 8.4. Analysis shows that there is a highly significant difference between the control and cancer patients for both Total Activity ($p = 0.005$) and Specific Activity ($p = 0.004$) of Seprase in serum (see Table 8.4). Therefore, it can be concluded from this that Seprase levels are elevated in cancer patients with invasive ductal carcinoma and that the specific Seprase assay described was able to detect this difference.

	Diagnosis	Kolmogorov-Smirnov(a)		
		Statistic	df	Sig.
Total Activity	**Cancer**	0.211	10	0.200
	Control	0.239	10	0.109
Specific Activity	**Cancer**	0.165	10	0.200
	Control	0.243	10	0.097

Table 8.2 Test of Normality of Clinical Samples

	Diagnosis	N	Mean	Std. Deviation	Std. Error Mean
Total Activity	**Cancer**	10	0.09752	0.042259	0.013364
	Control	10	0.04695	0.019200	0.006072
Specific Activity	**Cancer**	10	0.00157058	0.000796097	0.000251748
	Control	10	0.00062619	0.000221251	0.000069966

Table 8.3 Group Statistics of Clinical Samples

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. error Mean	95% Confidence Interval of the Difference	
							Lower	Upper
Total Activity	***Equal variances not assumed***	3.445	12.564	***0.005***	0.05057	0.01468	0.0004	0.0015
Specific Activity	***Equal variances not assumed***	3.614	10.382	***0.004***	0.00094	0.00026	0.0004	0.0015

Table 8.4 Independence Samples Test

To date there is no serum or saliva tumour marker(s) for breast cancer screening The current methods of screening for breast cancer, as described in section 1 10, have shown that they are not without disadvantages However until a suitable biomarker for breast cancer is discovered and developed, these screening methods remain at the forefront of patient diagnosis

The assay described in this research thesis is the only highly sensitive and specific assay developed for Seprase The assay has the unique ability to distinguish between the two Z-Gly-Pro-AMC hydrolysing activities found in serum Now that the assay has shown that Seprase levels are elevated in breast cancer patients, the study should be expanded to investigate and monitor the expression pattern of Seprase in serum and biopsy samples from patients with various types (e g normal, benign, *in situ* cancer, invasive ductal and lobular) and in various stages of breast cancer Serum samples would be taken from patients prior to surgery and at various stages post surgery This will give a more detailed analysis on the expression levels of Seprase at all stages of the cancer This assay was subject to invention disclosure and has been submitted for patent application A further expansion of the study would extend beyond breast cancer and would encompass the other cancers in which Seprase levels have been shown to be elevated (including colorectal, ovarian, bladder and lung carcinomas) This assay has the potential to be used initially as kit for breast cancer (1) screening, (2) diagnosis, and (3) for monitoring the progression of the disease It is envisaged that this test would be available in all hospitals and cancer clinics

Chapter 9

Summary and Recommendations

This research thesis describes the successful purification of the second Z-Gly-Pro-AMC hydrolysing activity in bovine serum using a highly effective purification procedure The development of the UV zymogram enabled the identification of this activity as Seprase (Chapter 3) Biochemical studies performed using the soluble serum form of Seprase have shown that the enzyme has both exopeptidase (cleaves Z-Gly-Pro-AMC) and endopeptidase (cleaves gelatin) activity This soluble form of Seprase was also shown to bind to the WGA lectin affinity column

Further biochemical analysis of Seprase (Chapter 4) and substrate specificity profiling provided valuable insight into the catalytic and structural properties of bovine Seprase (Chapter 5) The expression profile study identified the large intestine and the breast cancer cell line Hs578T as possible sources of Seprase (Chapter 6) This also allowed for the gene for human Seprase to be successfully cloned into both a prokaryotic and mammalian expression system (Chapter 7) The clinical study confirmed that Seprase levels were elevated in serum samples from patients with invasive ductal carcinoma (Chapter 8) These studies have greatly added to the growing amount of published data on this recently discovered protease, Seprase

The role of Seprase in cancer still awaits experimental proof Some additional studies on Seprase that could provide more clarity on the enzyme and also on the relationship between the structure and function are hereby outlined

Further work is required to express Seprase in both the prokaryotic and mammalian expression systems The expression of Seprase in the prokaryotic system requires optimisation to express the full length protein The transfection of mammalian cells also requires further optimisation This could involve varying the electroporation conditions or possibly using an alternative method such as transient transfection using lipofection An alternative to the SW480 cell line should also be sought, as this cell line has recently shown difficulty in other transfection experiments performed by other research groups (data not shown) This may involve cloning and expressing the soluble form of Seprase, lacking the transmembrane domain

Once an active form of Seprase is expressed it would be possible to perform mutational analysis Comparative biochemical and structural studies using the recombinant protein would provide information on the relationship between structure and function of the enzyme These studies would include those mentioned in this thesis and also the effects of glycosylation on both the exopeptidase and endopeptidase activity of Seprase The role of Seprase in cancer could be elucidated using function/knockdown approach (i e by specific inhibitors or RNAi) and results of quantitative expression profiling studies (using real-time PCR)

The identification of Seprase in the bovine large intestine has created the opportunity to clone the bovine Seprase gene and to determine the amino acid sequence of the bovine protease A recombinant bovine form of the enzyme could be expressed and compared to the human Seprase recombinant protein

The substrate specificity studies could be expanded by analysing the hydrolysis of the synthetic peptide substrates by Seprase using HPLC and identifying the products using LC-MS

The clinical study should also be expanded to investigate and monitor the expression pattern of Seprase in serum and biopsy samples from patients with various types (e g normal, benign, *in situ* cancer, invasive ductal and lobular and in various stages of breast cancer Serum samples would be taken from patients prior to surgery and at various stages post surgery This will give a more detailed analysis on the expression levels of Seprase at all stages of the cancer A further expansion of the study would extend beyond breast cancer and would encompass the other cancers in which Seprase levels have been shown to be elevated including colorectal, ovarian, bladder and lung carcinomas

Finally, this work now provides some useful tools for the continued research of Seprase and its *in vivo* role The clone presents a probe, which can be used in various tissue and array-hybridisation studies and to investigate possible tissue-specific forms and expression of the enzyme Also, antibodies can be raised against the purified recombinant protein and subsequently used in immunohistochemical investigations and in applications such as high-resolution tissue microscopy coupled with immunocytochemical amplification

References

-A-

Abbott, C A , and M D Gorrell 2002 The Family of CD26/DPIV and Related Ectopeptidases The DPIV Family *Ectopeptidases CD13/Aminopeptidase N and CD26/Dipeptidylpeptidase IV in Medicine and Biology* Kluwer Academic/Plenum Publishers, New York 171-195 pp

Abbott, C A , D Koorey, and G W McCaughan 1993 An EcoRI polymorphism within the dipeptidyl peptidase IV (DPPIV) gene *Hum Mol Genet* 2 1507

Abbott, C A , G W McCaughan, and M D Gorrell 1999 Two highly conserved glutamic acid residues in the predicted [beta] propeller domain of dipeptidyl peptidase IV are required for its enzyme activity *FEBS Letters* 458 278-284

Acharya, P S , A Zukas, V Chandan, A L Katzenstein, and E Pure 2006 Fibroblast activation protein a serine protease expressed at the remodeling interface in idiopathic pulmonary fibrosis *Hum Pathol* 37 352-360

Aertgeerts, K , I Levin, L Shi, G P Snell, A Jennings, G S Prasad, Y Zhang, M L Kraus, S Salakian, V Sridhar, R Wijnands, and M G Tennant 2005 Structural and kinetic analysis of the substrate specificity of human fibroblast activation protein alpha *J Biol Chem* 280 19441-19444

Aimes, R T , A Zijlstra, J D Hooper, S M Ogbourne, M L Sit, S Fuchs, D C Gotley, J P Quigley, and T M Antalis 2003 Endothelial cell serine proteases expressed during vascular morphogenesis and angiogenesis *Thrombosis and Haemostasis* 89 561-572

Alcolado, R , M J Arthur, and J P Iredale 1997 Pathogenesis of liver fibrosis *Clin Sci (Lond)* 92 103-112

Altschul, S F , Gish, W , Miller, W , Myers, E W and Lipman, D J 1990 Basic Local Alignment Search Tool *Journal of Molecular Biology* 215 403-410

Anderson, F P 2005 The synthesis, structural characterisation and biological evaluation of potential chemotherapeutic agents PhD Thesis, Dublin City University

Aoyama, A , and W T Chen 1990 A 170-kDa membrane-bound protease is associated with the expression of invasiveness by human malignant melanoma cells *Proc Natl Acad Sci U S A* 87 8296-8300

Ariga, N , E Sato, N Ohuchi, H Nagura, and H Ohtani 2001 Stromal expression of fibroblast activation protein/seprase, a cell membrane serine proteinase and

gelatinase, is associated with longer survival in patients with invasive ductal carcinoma of breast *Int J Cancer* 95 67-72

Artym, V V, A L Kindzelskii, W T Chen, and H R Petty 2002 Molecular proximity of seprase and the urokinase-type plasminogen activator receptor on malignant melanoma cell membranes dependence on beta1 integrins and the cytoskeleton *Carcinogenesis* 23 1593-1601

-B-

Bairoch, A and Apweiler, R 1996 The SWISS-PROT protein sequence data bank and its new supplement TREMBL *Nucleic Acids Research* 24 21-25

Bangert, K, A H Johnsen, U Christensen, and S Thorsen 1993 Different N-terminal forms of alpha 2-plasmin inhibitor in human plasma *Biochem J* 291 (Pt 2) 623-625

Barrett, J C, N D Rawlings, and J F Woessner 1998 Handbook of Proteolytic Enzymes Academic Press, London

Bauvois, B 2004 Transmembrane proteases in cell growth and invasion new contributors to angiogenesis? *Oncogene* 23 317-329

Benson, D A, Boguski, M ,Lipman, D J and Ostell, J 1996 GenBank *Nucleic Acids Research* 24 1-5

Berman, H M, Westbrook, J, Feng, Z, Gilliland, G, Bhat, T N, Weissig, H, Shindyalov, I N and Bourne, P E (2000) The Protein Data Bank *Nucleic Acids Research* 28 235-242

Bermpohl, F, K Loster, W Reutter, and O Baum 1998 Rat dipeptidyl peptidase IV (DPP IV) exhibits endopeptidase activity with specificity for denatured fibrillar collagens *FEBS Letters* 428 152-156

Bernardi, G, M -G Giro, and C Gaillard 1972 Chromatography of polypeptides and proteins on hydroxyapatite columns some new developments *Biochimica et Biophysica Acta* 278 409-420

Birnboim, H C, and J Doly 1979 A rapid alkaline extraction procedure for screening recombinant plasmid DNA *Nucleic Acids Research* 7 1513-1523

Birney, Y A and B O'Connor 2000 The Purification and Characterisation of a Z-Pro-Prolinal Insensitive Z-Gly-Pro-MCA degrading peptoidase from bovine serum PhD Thesis, Dublin City University

Birney, Y A , and B F O'Connor 2001 Purification and characterization of a Z-pro-prolinal-insensitive Z-Gly-Pro-7-amino-4-methyl coumarin-hydrolyzing peptidase from bovine serum--a new proline-specific peptidase *Protein Expr Purif* 22 286-298

Bissell, M J , and D Radisky 2001 Putting tumours in context *Nat Rev Cancer* 1 46-54

Blum, H , H Beir, and H J Gross 1987 Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels *Electrophoresis* 8 93-99

Bradford, M M 1976 A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding *Anal Biochem* 72, 248-254

Brown, D D , Z Wang, J D Furlow, A Kanamori, R A Schwartzman, B F Remo, and A Pinder 1996 The thyroid hormone-induced tail resorption program during Xenopus laevis metamorphosis *Proc Natl Acad Sci U S A* 93 1924-1929

Brown, M L , F Houn, E A Sickles, and L G Kessler 1995 Screening mammography in community practice positive predictive value of abnormal findings and yield of follow-up diagnostic procedures *Am J Roentgenol* 165 1373-1377

Brown, N M , T T Stenzel, P N Friedman, J Henslee, G Huper, and J R Marks 2005 Evaluation of expression based markers for the detection of breast cancer cells *Breast Cancer Res Treat* 97 41-47

Busek, P , R Malik, and A Sedo 2004 Dipeptidyl peptidase IV activity and/or structure homologues (DASH) and their substrates in cancer *The International Journal of Biochemistry & Cell Biology* 36 408-421

-C-

Chen, W -T , T Kelly, and G Ghersi 2003a DPPIV, seprase, and related serine peptidases in multiple cellular functions *In* Current Topics in Developmental Biology Vol Volume 54 Academic Press 207-232

Chen, W T 2003b DPPIV and seprase in cancer invasion and angiogenesis *Adv Exp Med Biol* 524 197-203

Chen, W T , and T Kelly 2003c Seprase complexes in cellular invasiveness *Cancer Metastasis Rev* 22 259-269

Cheng, J D , R L Dunbrack, Jr , M Vahanou, A Rogatko, R K Alpaugh, and L M Weiner 2002 Promotion of tumor growth by murine fibroblast activation protein, a serine protease, in an animal model *Cancer Res* 62 4767-4772

Cheng, J D , M Vahanou, A A Canutescu, E K Jaffe, H O Lee, H Wang, J H Lai, W W Bachovchin, and L M Weiner 2005 Abrogation of fibroblast activation protein enzymatic activity attenuates tumor growth *Mol Cancer Ther* 4 351-360

Chomczynski, P , and N Sacchi 1987 Single-Step method of RNA Isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform Extraction *Analytical Biochemistry* 162 156-159

Collins, P J and O'Connor B 2003 The study of a new and proline cleaving peptidase from bovine serum PhD Thesis, Dublin City University

Collins, P J , G McMahon, P O'Brien, and B O'Connor 2004 Purification, identification and characterisation of seprase from bovine serum *Int J Biochem Cell Biol* 36 2320-2333

Cordero, O J , D Ayude, M Nogueira, F J Rodriguez-Berrocal, and M Paez de la Cadena 2000 Preoperative serum CD26 levels diagnostic efficiency and predictive value for colorectal cancer *British Journal of Cancer* 83 1139-1146

Corpet, F 1988 Multiple Sequence Alignment with Hierarchical clustering *Nucleic Acids Research* 16 10881-10890

Cunningham, D F , and B O'Connor 1997 Proline specific peptidases *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1343 160-186

-D-

Dehm, P , and A Nordwig 1970 The cleavage of proyl pepides by Kidney Peptidases Partial purification of a "X-Prolyl-Aminopeptidase" from Swine Kidney Microsomes *European Journal of Biochemistry* 17 364-371

D'Souza, B , M Rowe, and D Walls 2000 The *bfl*-1 gene is transcriptionally upregulated by the Epstein-Barr Virus LMP1, and its expression promotes the survival of a Burkitt's Lymphoma Cell Line *J of Virology* 74,6652-6658

De Wever, O , and M Mareel 2003 Role of tissue stroma in cancer cell invasion *J Pathol* 200 429-447

Del Rosso, M , G Fibbi, M Pucci, S D'Alessio, A Del Rosso, L Magnelli, and V Chiarugi 2002 Multiple pathways of cell invasion are regulated by multiple families of serine proteases *Clin Exp Metastasis* 19 193-207

Dingemans, K P , I M Zeeman-Boeschoten, R F Keep, and P K Das 1993 Transplantation of colon carcinoma into granulation tissue induces an invasive morphotype *Int J Cancer* 54 1010-1016

Dolznig, H , N Schweifer, C Puri, N Kraut, W J Rettig, D Kerjaschki, and P Garin-Chesa 2005 Characterization of cancer stroma markers in silico analysis of an mRNA expression database for fibroblast activation protein and endosialin *Cancer Immun* 5 10-19

Donlon, J , and S Kaufman 1980 Relationship between the multiple forms of rat hepatic phenylalanine hydroxylase and degree of phosphorylation *J Biol Chem* 255 2146-2152

Dowling, O 1998 PhD Thesis Dublin City University

Dvorak, H F 1986 Tumors wounds that do not heal Similarities between tumor stroma generation and wound healing *N Engl J Med* 315 1650-1659

-E-

Edosada, C Y , C Quan, T Tran, V Pham, C Wiesmann, W Fairbrother, and B B Wolf 2006a Peptide substrate profiling defines fibroblast activation protein as an endopeptidase of strict Gly(2)-Pro(1)-cleaving specificity *FEBS Letters* 580 1581-1586

Edosada, C Y , C Quan, C Wiesmann, T Tran, D Sutherlin, M Reynolds, J M Elliott, H Raab, W Fairbrother, and B B Wolf 2006b Selective inhibition of fibroblast activation protein protease based on dipeptide substrate specificity *J Biol Chem* 281,7437-7444

-F-

Fassnacht, M , J Lee, C Milazzo, D Boczkowski, Z Su, S Nair, and E Gilboa 2005 Induction of CD4(+) and CD8(+) T-cell responses to the human stromal antigen, fibroblast activation protein implication for cancer immunotherapy *Clin Cancer Res* 11 5566-71

Fidler, I J , 2002 The pathogenesis of cancer metastasis revisited Nature Reviews Cancer 3, 1-6 doi 10 1038/nrc1098

Fink, A L 1987 Acyl group transfer-The Serine Proteases Enzyme Mechanisms, Royal Society of Chemistry 159-177 pp

Flentke, G R , E Munoz, B T Huber, A G Plaut, C A Kettner, and W W Bachovchin 1991 Inhibition of dipeptidyl aminopeptidase IV (DP-IV) by Xaa-boroPro dipeptides and use of these inhibitors to examine the role of DP-IV in T-cell function *Proc Natl Acad Sci U S A* 88 1556-1559

Fulop, V , Z Bocskei, and L Polgar 1998 Prolyl Ohgopeptidase An unusual beta-propeller domain regulates proteolysis *Cell* 94 161-170

Fulop, V , Z Szeltner, and L Polgar 2000 Catalysis of serine oligopeptidases is controlled by a gating filter mechansim *EMBO Reports* 1 277-281

-G-

Garin-Chesa, P , L J Old, and W J Rettig 1990 Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers *Proc Natl Acad Sci U S A* 87 7235-7239

Gerner, C , W Steinkellner, K Holzmann, A Gsur, R Grimm, C Ensinger, P Obrist, and G Sauermann 2001 Elevated plasma levels of crosslinked fibrinogen gamma-chain dimer indicate cancer-related fibrin deposition and fibrinolysis *Thromb Haemost* 85 494-501

Ghersi, G , H Dong, L A Goldstein, Y Yeh, L Hakkinen, H S Larjava, and W T Chen 2002 Regulation of fibroblast migration on collagenous matrix by a cell surface peptidase complex *Journal of Biological Chemistry* 277 29231-29241

Ghersi, G , H Dong, L A Goldstein, Y Yeh, L Hakkinen, H S Larjava, and W T Chen 2003 Seprase-dPPIV association and prolyl peptidase and gelatinase activities of the protease complex *Adv Exp Med Biol* 524 87-94

Ghersi, G , Q Zhao, M Salamone, Y Yeh, S Zucker, and W T Chen 2006 The protease complex consisting of dipeptidyl peptidase IV and seprase plays a role in the migration and invasion of human endothelial cells in collagenous matrices *Cancer Res* 66 4652-4661

Gilmore, B F , J F Lynas, C J Scott, C McGoohan, L Martin, and B Walker 2006 Dipeptide proline diphenyl phosphonates are potent, irreversible inhibitors of seprase (FAPalpha) *Biochem Biophys Res Commun* 346, 436-446

Goldstein, L A , and W T Chen 2000 Identification of an alternatively spliced seprase mRNA that encodes a novel intracellular isoform *J Biol Chem* 275 2554-2559

Goldstein, L A , G Ghersi, M L Pineiro-Sanchez, M Salamone, Y Yeh, D Flessate, and W T Chen 1997 Molecular cloning of seprase a serine integral membrane protease from human melanoma *Biochim Biophys Acta* 1361 11-19

Goodman, J D , T L Rozypal, and T Kelly 2003 Seprase, a membrane-bound protease, alleviates the serum growth requirement of human breast cancer cells *Clin Exp Metastasis* 20 459-470

Gornall, A G , C J Bardawill, M M David 1949 Determination of serum proteins by means of the Biuret reagent *J Biol Chem* 177, 751

Gorrell, M D , V Gysbers, and G W McCaughan 2001 CD26 a multifunctional integral membrane and secreted protein of activated lymphocytes *Scand J Immunol* 54 249-264

Groarke DA, Drmota T, Bahia DS, Evans NA, Wilson S, Milligan G 2001 Analysis of the C-terminal tail of the rat thyrotropin-releasing hormone receptor-1 in interactions and cointernalization with ß-arrestin 1-green fluorescent protein *Mol Pharmacol* 59 375–385

Guex, N and Peitsch, M C 1997 SWISS-MODEL and the Swiss-Pdb Viewer An environment for comparative protein modelling *Electrophoresis* 18 2714-2723

-H-

Hanahan, D 1985 Techniques for transformation of *E coli* IRL Press, Oxford 109-135 pp

Harris, J L , B J Backes, F Leonetti, S Mahrus, J A Ellman, and C S Craik 2000 Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries *Proc Natl Acad Sci U S A* 97 7754-7759

Harris, E L V , and S Angel 1989 Protein Purification Methods - A Practical Approach IRL Press at Oxford University press

Hofheinz, R D , S E al-Batran, F Hartmann, G Hartung, D Jager, C Renner, P Tanswell, U Kunz, A Amelsberg, H Kuthan, and G Stehle 2003 Stromal antigen targeting by a humanised monoclonal antibody an early phase II trial of sibrotuzumab in patients with metastatic colorectal cancer *Onkologie* 26 44-48

Huang, Y , S Wang, and T Kelly 2004 Seprase promotes rapid tumor growth and increased microvessel density in a mouse model of human breast cancer *Cancer Res* 64 2712-2716

Huber, M A , N Kraut, J E Park, R D Schubert, W J Rettig, R U Peter, and P Garin-Chesa 2003 Fibroblast activation protein differential expression and serine protease activity in reactive stromal fibroblasts of melanocytic skin tumors *J Invest Dermatol* 120 182-188

Huber, M A , N Kraut, N Schweifer, H Dolznig, R U Peter, R D Schubert, K Scharffetter-Kochanek, H Pehamberger, and P Garin-Chesa 2006 Expression of stromal cell markers in distinct compartments of human skin cancers *J Cutan Pathol* 33 145-155

-I-

Iwasa, S , K Okada, W T Chen, X Jin, T Yamane, A Ooi, and M Mitsumata 2005 'Increased expression of seprase, a membrane-type serine protease, is associated with lymph node metastasis in human colorectal cancer' *Cancer Lett* 227 229-236

-J-

Jin, X , S Iwasa, K Okada, M Mitsumata, and A Ooi 2003 Expression patterns of seprase, a membrane serine protease, in cervical carcinoma and cervical intraepithelial neoplasm *Anticancer Res* 23 3195-3198

-K-

Keen, J 2006 Protein sequncing facility http //bmb leeds ac uk/bmaf/proseq1 html

Kelly, T 1999 Evaluation of seprase activity *Clin Exp Metastasis* 17 57-62

Kelly, T 2005 Fibroblast activation protein-[alpha] and dipeptidyl peptidase IV (CD26) Cell-surface proteases that activate cell signaling and are potential targets for cancer therapy *Drug Resistance Updates* 8 51-58

Kelly, T , S Kechelava, T L Rozypal, K W West, and S Korourian 1998 Seprase, a membrane-bound protease, is overexpressed by invasive ductal carcinoma cells of human breast cancers *Mod Pathol* 11 855-863

Kelly, T , S C Mueller, Y Yeh, and W T Chen 1994 Invadopodia promote proteolysis of a wide variety of extracellular matrix proteins *J Cell Physiol* 158 299-308

Kloft, C , E U Graefe, P Tanswell, A M Scott, R Hofheinz, A Amelsberg, and M O Karlsson 2004 Population pharmacokinetics of sibrotuzumab, a novel therapeutic monoclonal antibody, in cancer patients *Investigational New Drugs* 22 39-52

Koyama, T , Y Koike, S Toyota, F Miyagi, N Suzuki, and N Aoki 1994 Different NH2-terminal form with 12 additional residues of alpha 2-plasmin inhibitor from human plasma and culture media of Hep G2 cells *Biochem Biophys Res Commun* 200 417-422

-L-

Laemmli, U K 1970 Cleavage of Structural proteins during assembly of the head of structural bacteriophage T4 *Nature* 227, 680-685

Lee, K N , K W Jackson, V J Christiansen, C S Lee, J G Chun, and P A McKee 2005a Antiplasmin-cleaving enzyme is a soluble form of fibroblast activation protein *Blood* 107 1397-1404

Lee, J , M Fassnacht, S Nair, D Boczkowski, and E Gilboa 2005b Tumor immunotherapy targeting fibroblast activation protein, a product expressed in tumor-associated fibroblasts *Cancer Res* 65 11156-11163

Lee, K N , K W Jackson, V J Christiansen, K H Chung, and P A McKee 2004 A novel plasma proteinase potentiates alpha(2)-antiplasmin inhibition of fibrin digestion *Blood* 103 3783-3788

Lee, S C , K N Lee, D G Schwartzott, K W Jackson, W C Tae, and P A McKee 1997 Purification of human alpha 2-antiplasmin with chicken IgY specific to its carboxy-terminal peptide *Prep Biochem Biotechnol* 27 227-237

Leiting, B , K D Pryor, J K Wu, F Marsilio, R A Patel, C S Craik, J A Ellman, R T Cummings, and N A Thornberry 2003 Catalytic properties and inhibition of proline-specific dipeptidyl peptidases II, IV and VII *Biochem J* 371 525-532

Levy, M T , M D Gorrell, C A Abbott, J E Park, W J Rettig, and G W McCaughan 1997 Fibroblast activation protein alpha (Fap) A membrane protease expressed on stellate cells at the tissue remodelling interface in human cirrhosis *Hepatology* 26 1283-1283

Levy, M T , G W McCaughan, C A Abbott, J E Park, A M Cunningham, E Muller, W J Rettig, and M D Gorrell 1999 Fibroblast activation protein a cell surface dipeptidyl peptidase and gelatinase expressed by stellate cells at the tissue remodelling interface in human cirrhosis *Hepatology* 29 1768-1778

Levy, M T , G W McCaughan, G Marinos, and M D Gorrell 2002 Intrahepatic expression of the hepatic stellate cell marker fibroblast activation protein correlates with the degree of fibrosis in hepatitis C virus infection *Liver* 22 93-101

Liotta, L A , and E C Kohn 2001 The microenvironment of the tumour-host interface *Nature* 411 375-379

Liotta, L A , P S Steeg, and W G Stetler-Stevenson 1991 Cancer metastasis and angiogenesis an imbalance of positive and negative regulation *Cell* 64 327-336

Lowry, O H and N J Rosenbrough, A L Farr, R J Randall (1951) Protein Measurement with the Folin Phenol Reagent *J Biol Chem* 193,265-275

-M-

Makrides, A C 1996 Strategies for achieving high-level expression of genes in *Escherichia coli Microbiological Reviews* 60 512-538

Mamatis, T , Fritsch, E F and Sambrook, J (1982) Molecular Cloning, A Laboratory Manual Cold Spring Harbour Laboratory

Mathew, S , M J Scanlan, B K Mohan Raj, V V Murty, P Garin-Chesa, L J Old, W J Rettig, and R S Chaganti 1995 The gene for fibroblast activation protein alpha (FAP), a putative cell surface-bound serine protease expressed in cancer stroma and wound healing, maps to chromosome band 2q23 *Genomics* 25 335-237

Mersmann, M , A Schmidt, J F Rippmann, T Wuest, B Brocks, W J Rettig, P Garin-Chesa, K Pfizenmaier, and D Moosmayer 2001 Human antibody derivatives against the fibroblast activation protein for tumor stroma targeting of carcinomas *Int J Cancer* 92 240-248

Milner, J M , L Kevorkian, D A Young, D Jones, R Wait, S T Donell, E Barksby, A M Patterson, J Middleton, B F Cravatt, I M Clark, A D Rowan, and T E Cawston 2006 Fibroblast activation protein alpha is expressed by chondrocytes following a pro-inflammatory stimulus and is elevated in osteoarthritis *Arthritis Res Ther* 8 R23

Monsky, W L , C Y Lin, A Aoyama, T Kelly, S K Akiyama, S C Mueller, and W T Chen 1994 A potential marker protease of invasiveness, seprase, is localized on invadopodia of human malignant melanoma cells *Cancer Res* 54 5702-5710

Mori, Y , K Kono, Y Matsumoto, H Fujii, T Yamane, M Mitsumata, and W T Chen 2004 The expression of a type II transmembrane serine protease (Seprase) in human gastric carcinoma *Oncology* 67 411-419

Mueller, S C , G Ghersi, S K Akiyama, Q X Sang, L Howard, M Pineiro-Sanchez, H Nakahara, Y Yeh, and W T Chen 1999 A novel protease-docking function of integrin at invadopodia *J Biol Chem* 274 24947-24952

Mullis, K B and Faloona, F A (1987) Specific Synthesis of DNA *in Vitro* via a Polymerase-Catalyzed Chain Reaction *Methods in Enzymology* **155** 335-350

-N-

Nakamura, Y 2000 Codon usage tabulated from international DNA sequence databases status for the year 2000 *Nucleic Acids Research* 28 292

Nicholas, K B , Nicholas, H B Jr and Deerfield II, D W 1997 GeneDoc Analysis and Visualization of Genetic Variation *EMBnet News* 4 14

Niedermeyer, J , B Enenkel, J E Park, M Lenter, W J Rettig, K Damm, and A Schnapp 1998 Mouse fibroblast-activation protein--conserved Fap gene organization and biochemical function as a serine protease *Eur J Biochem* 254 650-654

Niedermeyer, J , P Garin-Chesa, M Kriz, F Hilberg, E Mueller, U Bamberger, W J Rettig, and A Schnapp 2001 Expression of the fibroblast activation protein during mouse embryo development *Int J Dev Biol* 45 445-447

Niedermeyer, J , M J Scanlan, P Garin-Chesa, C Daiber, H H Fiebig, L J Old, W J Rettig, and A Schnapp 1997 Mouse fibroblast activation protein molecular cloning, alternative splicing and expression in the reactive stroma of epithelial cancers *Int J Cancer* 71 383-389

Novy, R , D Drott, K Yaeger, and R Mierendorf 2001 Overcoming the codon bias of *E coli* for enchanced protein expression *Innovations* 12 1-3

-O-

Ogata, S , Y Mitsumi, and Y Ikehara 1989 Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA and identification of the NH_2-terminal signal sequence as the membrane-anchoring domain *Journal of Biological Chemistry* 264 3596-3601

Okada, K , W T Chen, S Iwasa, X Jin, T Yamane, A Ooi, and M Mitsumata 2003 Seprase, a membrane-type serine protease, has different expression patterns in intestinal- and diffuse-type gastric cancer *Oncology* 65 363-370

Okada, K , S Ueshima, M Imano, K Kataoka, and O Matsuo 2004 The regulation of liver regeneration by the plasmin/alpha 2-antiplasmin system *J Hepatol* 40 110-116

Ollis, D , E Cheah, M Cygler, B Dijkstra, F Frolow, S Franken, M Harel, J Remington, I Silman, J Schrag, J Sussman, K Verschueren, and A Goldman 1992 The alpha/beta hydrolase domain *Protein Engineering* 5 197-211

-P-

Park, J E , M C Lenter, R N Zimmermann, P Garin-Chesa, L J Old, and W J Rettig 1999 Fibroblast activation protein, a dual specificity serine protease expressed in reactive human tumor stromal fibroblasts *J Biol Chem* 274 36505-36512

Perkins, G L , E D Slater, G K Sanders, and J G Prichard 2003 Serum tumor markers *Am Fam Physician* 68 1075-1082

Pineiro-Sanchez, M L , L A Goldstein, J Dodt, L Howard, Y Yeh, H Tran, W S Argraves, and W T Chen 1997 Identification of the 170-kDa melanoma membrane-bound gelatinase (seprase) as a serine integral membrane protease *J Biol Chem* 272 7595-7601

Plebani, M 2005 Proteomics the next revolution in laboratory medicine? *Clin Chim Acta* 357 113-122

Polgar, L 2002 The prolyl oligopeptidase family *Cell Mol Life Sci* 59 349-362

Prendeville, D J , J Lowe, B Earley, C Spahr, and P Kettlewell 2002 Radiotelemetry systems for measuring body temperature *Beef Production Series No 57* Teagasc Report

-R-

Ramirez-Montagut, T , N E Blachere, E V Sviderskaya, D C Bennett, W J Rettig, P Garin-Chesa, and A N Houghton 2004 FAPalpha, a surface peptidase expressed during wound healing, is a tumor suppressor *Oncogene* 23 5435-5446

Rennex, D , B A Hemmings, J Hofsteenge, and S R Stone 1991 cDNA cloning of porcine brain prolyl endopeptidase and identification of the active-site seryl residue *Biochemistry* 30 2195-2203

Rettig, W J , P Garin-Chesa, H R Beresford, H F Oettgen, M R Melamed, and L J Old 1988 Cell-surface glycoproteins of human sarcomas differential expression in normal and malignant tissues and cultured cells *Proc Natl Acad Sci U S A* 85 3110-3114

Rettig, W J , P Garin-Chesa, J H Healey, S L Su, H L Ozer, M Schwab, A P Albino, and L J Old 1993 Regulation and heteromeric structure of the fibroblast activation protein in normal and transformed cells of mesenchymal and neuroectodermal origin *Cancer Res* 53 3327-3335

Rettig, W J , S L Su, S R Fortunato, M J Scanlan, B K Raj, P Garin-Chesa, J H Healey, and L J Old 1994 Fibroblast activation protein purification, epitope mapping and induction by growth factors *Int J Cancer* 58 385-392

Rosenberg, A H , E Goldman, J J Dunn, F W Studier, and G Zubay 1993 Effects of consecutive AGG codons on translation in *Escherichia coli* demonstrated with a versatile codon test system *J Bacteriol* 175 716-722

Rosenblum, J S , and J W Kozarich 2003 Prolyl peptidases a serine protease subfamily with high potential for drug discovery *Current Opinion in Chemical Biology* 7 496-504

-S-

Samel, D , D Muller, J Gerspach, C Assohou-Luty, G Sass, G Tiegs, K Pfizenmaier, and H Wajant 2003 Generation of a FasL-based proapoptotic fusion protein devoid of systemic toxicity due to cell-surface antigen-restricted Activation *J Biol Chem* 278 32077-32082

Scanlan, M J , B K Raj, B Calvo, P Garin-Chesa, M P Sanz-Moncasi, J H Healey, L J Old, and W J Rettig 1994 Molecular cloning of fibroblast activation protein alpha, a member of the serine protease family selectively expressed in stromal fibroblasts of epithelial cancers *Proc Natl Acad Sci U S A* 91 5657-5661

Schechter, I , and A Berger 1967 On the size of the active site in proteases I Papain *Biochem Biophys Res Commun* 27 157-162

Scott, A M , G Wiseman, S Welt, A Adjei, F T Lee, W Hopkins, C R Divgi, L H Hanson, P Mitchell, D N Gansen, S M Larson, J N Ingle, E W Hoffman, P Tanswell, G Ritter, L S Cohen, P Bette, L Arvay, A Amelsberg, D Vlock, W J Rettig, and L J Old 2003 A Phase I dose-escalation study of sibrotuzumab in patients with advanced or metastatic fibroblast activation protein-positive cancer *Clin Cancer Res* 9 1639-1647

Scott, A M , G Wiseman, S Welt, F T Lee, W Hopkins, P Mitchell, A Adjei, C R Divgi, S M Larson, E W Hoffman, P Tanswell, P Bette, D Vlock, A Amelsberg, W Rettig, and L J Old 2001 A phase I dose-escalation study of BIBH-1 in patients with advanced or metastatic fibroblast activation protein positive cancer *Journal of Nuclear Medicine* 42 311P

Sieweke, M H , and M J Bissell 1994 The tumor-promoting effect of wounding a possible role for TGF-beta-induced stromal alterations *Crit Rev Oncog* 5 297-311

Skubitz, K M , and A P N Skubitz 2004 Gene expression in aggressive fibromatosis *Journal of Laboratory and Clinical Medicine* 143 89-98

Smith, P K , R I Krohn, G T Hermanson, A K Mallia, F H Gartner, M D Provenzano, E K Fujimoto, N M Gorke, B J Olsen, D C Klenk (1985) Measurment of protein using Bicinchoninic Acid *Analytical Biochemistry* 150, 76-85

Sun, S , C F Albright, B H Fish, H J George, B H Selling, G F Hollis, and R Wynn 2002 Expression, Purification, and Kinetic Characterization of Full-Length Human Fibroblast Activation Protein *Protein Expression and Purification* 24 274-281

-T-

Tahtis, K , F T Lee, J M Wheatley, P Garin-Chesa, J E Park, F E Smyth, Y Obata, E Stockert, C M Hall, L J Old, W J Rettig, and A M Scott 2003 Expression and targeting of human fibroblast activation protein in a human skin/severe combined immunodeficient mouse breast cancer xenograft model *Mol Cancer Ther* 2 729-737

Tiselius, A , S Hjerten, and O Levin 1956 Protein Chromatography on Calcium Phosphate Columns *Archives of Biochemistry and Biophysics* 65 132-155

Toide, K , K Okamiya, Y Iwamoto, and T Kato 1995 Effect of a novel Prolyl Endopeptidase Inhibitor, JTP-4819, on Prolyl Endopeptidase Activity and Substance P- and Arginine-Vasopression-Like Immunoreactivity in the brains of aged rats *Journal of Neurochemistry* 65 234-240

-U-

Unger, J , and H Tschesche 1999 The Proteolytic Activity and Cleavage Specificity of Fibronectin-Gelatinase and Fibronectin-Lamininase *Journal of Protein Chemistry* 18 403-411

-V-

Vaas, P-R 2005 Molecular Characterisation of Recombinant Human Pyroglutamyl Peptidase (type I) PhD Thesis, Dublin City University

-W-

Wang, X M , D M Yu, G W McCaughan, and M D Gorrell 2005a Fibroblast activation protein increases apoptosis, cell adhesion, and migration by the LX-2 human stellate cell line *Hepatology* 42 935-945

Wang, X M , D M Yu, A Schnapp, D Marguet, S Cordoba, W Rettig, G W McCaughan, and M D Gorrell 2005b The role of fibroblast activation protein (FAP) in cell adhesion, migration and liver fibrosis *Hepatology* 42 738A

Welt, S , C R Divgi, A M Scott, P Garin-Chesa, R D Finn, M Graham, E A Carswell, A Cohen, S M Larson, L J Old, and et al 1994 Antibody targeting in metastatic colon cancer a phase I study of monoclonal antibody F19 against a cell-surface protein of reactive tumor stromal fibroblasts *J Clin Oncol* 12 1193-1203

Welt, S , and G Ritter 1999 Antibodies in the therapy of colon cancer *Seminars in Oncology* 26 683-690

Wesley, U V , A P Albino, S Tiwari, and A N Houghton 1999 A role for dipeptidyl peptidase IV in suppressing the malignant phenotype of melanocytic cells *Journal of Experimental Medicine* 190 311-322

-Y-

Yoshimoto, T , K Ogita, R Walter, M Koida, and D Tsuru 1979 Post proline cleaving enzyme, Synthesis of a new fluorogenic substrate and distribution of the endopeptidase in rat tissues and body fluids of man *Biochimica et Biophysica Acta* 569 184-192

Yoshimoto, T , T Nishimura, T Kita, and D Tsuru 1983 Post Proline Cleaving Enzyme (Prolyl Endopeptidase) from Bovine Brain *Journal of Biochemistry* 94 1179-1190

-Z-

Zhang, F , C C Tom, M C Kugler, T T Ching, J A Kreidberg, Y Wei, and H A Chapman 2003 Distinct ligand binding sites in integrin alpha3beta1 regulate matrix adhesion and cell-cell contact *J Cell Biol* 163 177-188

Appendix A

Activity & Kinetics

Quantification of Seprase Activity (Microplate Method 1)

One unit of activity is defined as the amount of enzyme which releases 1 nanomole of AMC per minute at 37°C (unit = nmoles min^{-1})

AMC released (X) is defined as fluorescent intensity (Fi) divided by the slope of the appropriate AMC standard curve is

$$\frac{Fi}{m} = X\ \mu moles\ L^{-1} = X\ \mu M$$

The reaction volume is $125x10^{-6}$ L and duration is 60 min

$$\text{AMC released} = \frac{X(125x10^{-6})}{60}\ \mu moles\ L^{-1}\ L\ min^{-1}$$

Reaction uses 25 $x10^{-6}$ L enzyme

$$\text{AMC released by enzyme} = \frac{X(125\ x10^{-6})}{60(25\ x10^{-6})}\ \mu moles\ min^{-1}\ L^{-1}$$

$$= \frac{X(125\ x10^{-6})(1000)}{60(25\ x10^{-6})(1000)}\ nmoles\ min^{-1}\ ml^{-1}$$

$$= \frac{X}{12}\ units\ ml^{-1}$$

$$= \frac{fli}{12m}\ units\ ml^{-1}$$

When expressing units as μmoles min^{-1} this formula changes to $\frac{fli}{12000m}\ units\ ml^{-1}$

Quantification of Seprase Activity (Microplate Method 2)

One unit of activity is defined as the amount of enzyme which releases 1 nanomole of AMC per minute at 37°C (unit = nmoles min^{-1})

AMC released (X) is defined as fluorescent intensity (Fi) divided by the slope of the appropriate AMC standard curve is

$$\frac{Fi}{m} = X\ \mu moles\ L^{-1} = X\ \mu M$$

The reaction volume is $250x10^{-6}$ L duration is 60 min

$$\text{AMC released} = \frac{X(250 \times 10^{6})}{60}\ \mu moles\ L^{-1}\ L\ min^{-1}$$

The reaction uses 25×10^{-6} L enzyme

$$\text{AMC released by enzyme} = \frac{X(250 \times 10^{6})}{60(25 \times 10^{-6})}\ \mu moles\ min^{-1}\ L^{-1}$$

$$= \frac{X(250 \times 10^{6})(1000)}{60(25 \times 10^{-6})(1000)}\ nmoles\ min^{-1}\ ml^{-1}$$

$$= \frac{X}{6}\ units\ ml^{-1}$$

$$= \frac{fli}{6m}\ units\ ml^{-1}$$

When expressing units as μmoles min^{-1} this formula changes to $\frac{fli}{6000m}\ units\ ml^{-1}$

Calculation of the Second Order Rate Constant, k_2 for DFP

Characteristic Kinetic Plot for the second order rate constant

$$\frac{1}{[A]} vs\ t$$

Slope of Kinetic Plot is k_2, L mole ^{1} sec ^{1}

[A] = [AMC] is calculated from

One unit of activity is defined as the amount of enzyme which releases 1 nanomol of AMC per minute at 37°C (unit = nmol min^{-1})

AMC released (X) is defined as fluorescent intensity (Fi) divided by the slope of the appropriate AMC standard curve

$$\frac{Fi}{m} = X\ \mu moles\ L^{1} = X\ \mu M$$

The reaction volume is $250x10^{-6}$ L

AMC released $= X\ (250 \times 10^{-6})\ \mu moles\ L^{-1}\ L$

The reaction uses 25 $x10^{6}$ L enzyme

AMC released by enzyme $= \frac{X(250 \times 10^{6})}{(25 \times 10^{-6})} \mu moles\ L^{-1}$

$$= 10X\ \mu moles\ L^{-1}$$

$$= \frac{10fli}{m} \mu moles\ L^{-1}$$

From Kinetic Plot

$$= \frac{(\mu mole)^{-1}\ L}{min}$$

$$= \mu M^{-1}\ min^{-1}$$

$$= \frac{Slope \times 10^{6}}{60} M^{-1} sec^{-1}$$

$$= k_2\ M^{-1}\ sec^{-1}$$

Purification Table Calculations

Total activity (unit) Units of enzyme activity, calculated from fluorescent intensity as described in section 6 1

Total protein (mg) mg/ml of protein estimated from appropriate BSA standard curve x volume of sample in mls

Specific activity (unit/mg) Total activity/Total protein

Purification factor Specific activity of sample/Specific activity of starting sample (serum)

Recovery (%) (Total activity of sample/Total activity of starting sample) x 100

Error Bars

Error bars on all enzyme activity graphs represent the standard error of the mean of triplicate fluorescence values The standard error (SE) of the mean is defined as

$$SE = \sigma/\sqrt{n}$$

Where

σ is the standard deviation i e $\sqrt{\text{variance}}$ (variance = $(a^2 + b^2 + c^2)/3$, where a, b & c are the triplicate values)

n is the number of repeat values measured (three in this case)

Kinetic Analysis

K_m Determination

The Michaelis constant, or K_m as it is usually known, is defined as the *substrate concentration which gives rise to a velocity equal to half the maximal velocity* This constant can be analysed by measuring reaction velocity (V) at various concentrations of substrate, the data giving rise to the well-known Michaelis-Menten hyperbola curve In the case of Seprase, activity is determined by measuring fluorescent intensity (Fi) resulting at various concentrations of the substrate Z-Gly-Pro-AMC, as described in Section 2 8 2 1 A plot of Fi versus [Z-Gly-Pro-AMC] yields the Michaelis-Menten curve Once an enzyme-catalysed reaction follows normal Michaelis-Menten kinetics, data can be applied to a number of kinetic models such as

Lineweaver-Burk Plot of 1/fluorescent intensity versus 1/substrate concentration The intercept of the line on the y-axis gives a direct readout of $1/V_{max}$ and the intercept of the line on the x-axis gives $-1/K_m$ The slope of the line is equal to K_m/V

Eadie-Hofstee Plot of fluorescent intensity versus fluorescent intensity/substrate concentration The intercept on the y-axis represents $1/V_{max}$, while the slope is given as $-K_m$

Hanes-Woolf Plot of substrate concentration/fluorescent intensity versus substrate concentration The intercept on the x-axis gives $-K_m$, and on the y-axis K_m/V_{max} The slope is 1/V

Determination of k_{cat}

The turnover constant (k_{cat}) of an enzyme can be determined if the V_{max} and total enzyme (E_t) are known

$$k_{cat} = \frac{V_{max}}{E_t} \qquad \frac{mol\ min^{-1}\ ml^{-1}}{mol\ ml^{-1}} = min^{-1}\ (s^{-1})$$

In the case of Seprase, V_{max} (nmoles min^{-1} ml^{-1} or units ml^{-1}) is determined experimentally as described above E_t (moles ml^{-1}) can be calculated from the molecular weight (g $mole^{-1}$) and amount of enzyme used (g ml^{-1}) The molecular weight of Seprase is 170,000 g $mole^{-1}$, as deduced from the amino acid sequence

K_i Determinations

The inhibition constant (K_i), or dissociation constant for the enzyme-inhibitor complex, can be determined by measuring the apparent K_m (K_m^{app}) observed in the presence of an inhibitor at a specific concentration ([I]), as described in Section 2 8 2 2 and applied to the following equation

K_i for competitive inhibition is calculated as follows,

$$K_i = \frac{K_m[I]}{K_m^{app} - K_m}$$

Where

K_i is the inhibition constant

K_m is the Michaelis constant with no inhibitor present

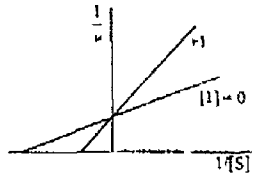
[I] is the inhibitor concentration

K_m^{app} means the new or "apparent" value of K_m in the presence of I, the inhibitor Measured as for K_m

Types of Reversible Inhibition

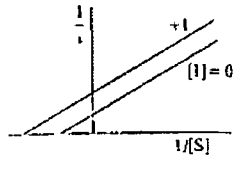
Inhibitors can be divided into two groups, reversible and irreversible Reversible inhibition involves noncovalent forces that bind inhibitors to enzymes There are several types of reversible inhibition competitive, uncompetitive, non competitive and mixed Distinguished, using the Lineweaver-Burk plot, by determining the effect of inhibitor on K_m, V_{max} and K_m/V_{max}

Competitive Inhibition Involves the binding of an inhibitor to the enzyme forming an enzyme-inhibitor complex The inhibitor bearing a structural and chemical similarity to the substrate binds to the active site thus competing with it for substrate



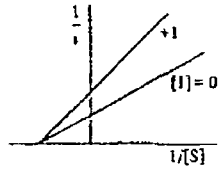
In competitive inhibition, the slope of the Lineweaver-Burk plot increases The K_m value also increases but the V_{max} value is unchanged This type of inhibition is recognised by plots intersecting at a common point on the positive y-axis

Uncompetitive inhibition Involves binding only to the enzyme-substrate complex to prevent catalysis



This type of inhibition causes a decrease in K_m and V_{max} There is no change in the slope of the Lineweaver-Burk plot It is identified by the presence of parallel lines in the Lineweaver-Burk plot

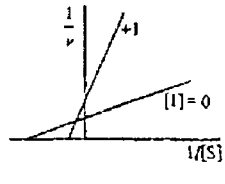
Non-competitive inhibition Involves binding equally to enzyme and enzyme-substrate complex



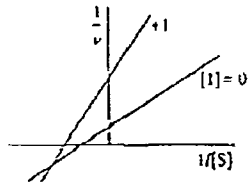
This type of inhibition causes a decrease in V_{max} along with K_m remaining the same The slope of the Lineweaver-Burk plot increases and is recognisable by plots intersecting at a common point on the negative x-axis

Mixed Inhibition· These inhibitors bind unequally to both enzyme and enzyme-substrate complex There can be two cases with these inhibitors,

Mixed competitive-non-competitive increase K_m and decrease V_{max} They cause an increase in the Lineweaver-Burk slope and are identified by plots intersecting above the negative portion of the x-axis

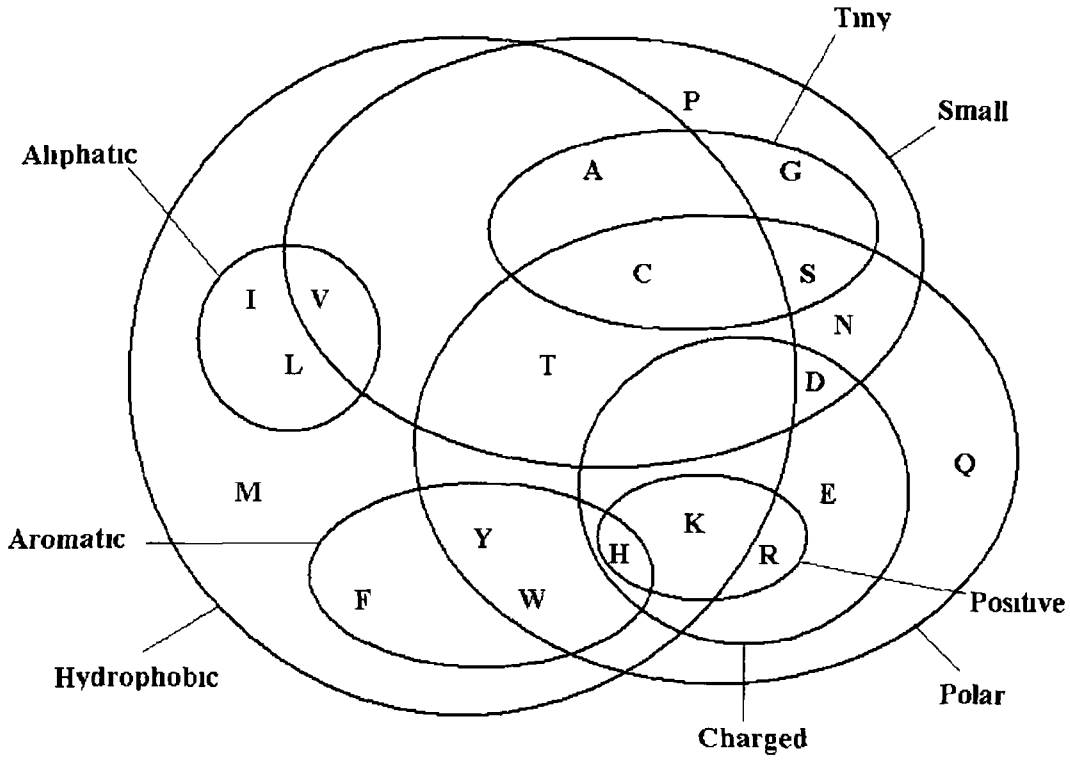


Mixed non-competitive-uncompetitive decrease both K_m and V_{max} The Lineweaver-Burk plot slope is increased and plots intersecting below the negative portion of the x-axis identify this type of inhibition

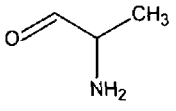
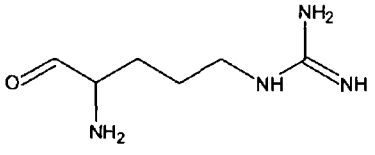
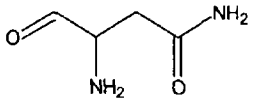
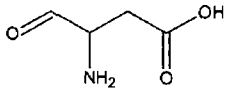
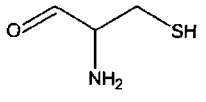
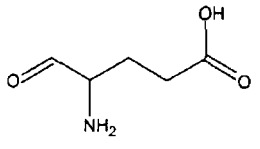
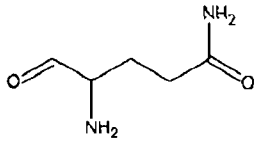


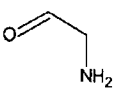
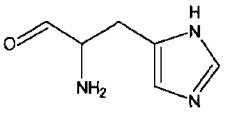
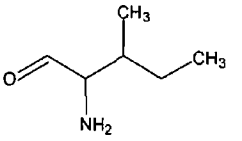
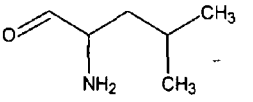
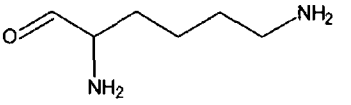
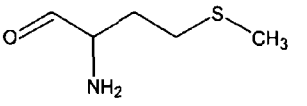
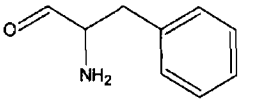
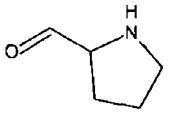
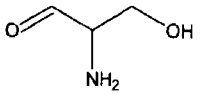
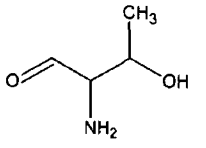
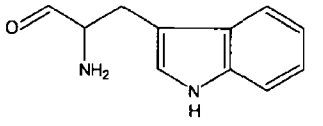
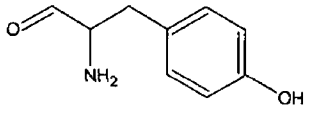
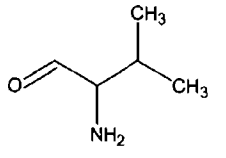
Appendix B

Amino Acids



Amino acid properties
Venn diagram presenting properties of the 20 amino acids

Alanine	Ala	A	
Arginine	Arg	R	
Asparagine	Asn	N	
Aspartic acid	Asp	D	
Cysteine	Cys	C	
Glutamic acid	Glu	E	
Glutamine	Gln	Q	

Glycine	Gly	G	
Histidine	His	H	
Isoleucine	Ile	I	
Leucine	Leu	L	
Lysine	Lys	K	
Methionine	Met	M	
Phenylalanine	Phe	F	
Proline	Pro	P	
Serine	Ser	S	
Threonine	Thr	T	
Tryptophan	Trp	W	
Tyrosine	Tyr	Y	
Valine	Val	V	

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

Genetic code
Amino acid codons applicable to both *H sapiens* and *E coli*

Appendix C

Sequence Data

>pPOB1 (labelled pPOB5 in fridge)

ATGAAGACTTGGGTAAAAATCGTATTTGGAGTTGCCACCTCTGCTGTGCTTGCCTTATTG
GTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATAACTCTGAAGAAAATACAATGAGA
GCACTCACACTGAAGGATATTTTAAATGGAACATTTTCTTATAAAACATTTTTTCCAAACT
GGATTTCAGGACAAGAATATCTTCATCAATCTGCAGATAACAATATAGTACTTTATAATA
TTGAAACAGGACAATCATATACCATTTTGAGTAATAGAACCATGAAAAGTGTGAATGCTT
CAAATTACGGCTTATCACCTGATCGGCAATTTGTATATCTAGAAAGTGATTATTCAAAGC
TTTGGAGATACTCTTACACAGCAACATATTACATCTATGACCTTAGCAATGGAGAATTTG
TAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTATTTATGCTGGTCGCCTGTTGGGAGTA
AATTAGCATATGTCTATCAAAACAATATCTATTTGAAACAAAGACCAGGAGATCCACCTT
TTCAAATAACATTTAATGGAAGAGAAAATAAAATATTTAATGGAATCCCAGACTGGGTTT
ATGAAGAGGAAATGCTTGCTACAAAATATGCTCTCTGGTGGTCTCCTAATGGAAAATTTT
TGGCATATGCGGAATTTAATGATACGGATATACCAGTTATTGCCTATTCCTATTATGGCG
ATGAACAATATCCTAGAACAATAAATATTCCATACCCAAAGGCTGGAGCTAAGAATCCC
GTTGTTCGGATATTTATTATCGATACCACTTACCCTGCGTATGTAGGTCCCCAGGAAGTGC
CTGTTCCAGCAATGATAGCCTCAAGTGATTATTATTTCAGTTGGCTCACGTGGGTTACTGA
TGAACGAGTATGTTTGCAGTGGCTAAAAAGAGTCCAGAATGTTTCGGTCCTGTCTATATG
TGACTTCAGGGAAGACTGGCAGACATGGGATTGTCCAAAGACCCAGGAGCATATAGAAG
AAAGCAGAACTGGATGGGCTGGTGGATTCTTTGTTTCAACACCAGTTTTCAGCTATGATG
CCATTTCGTACTACAAAATATTTAGTGACAAGGATGGCTACAAACATATTCACTATATCA
AAGACACTGTGGAAAATGCTATTCAAATTACAAGTGGCAAGTGGGAGGCCATAAATATA
TTCAGAGTAACACAGGATTCACTGTTTTATTCTAGCAATGAATTTGAAGAATACCCTGGA
AGAAGAAACATCTACAGAATTAGCATTGGAAGCTATCCTCCAAGCAAGAAGTGTGTTAC
TTGCCATCTAAGGAAAGAAAGGTGCCAATATTACACAGCAAGTTTCAGCGACTACGCCA
AGTACTATGCACTTGTCTGCTACGGCCCAGGCATCCCCATTTCCACCCTTCATGATGGAC
GCACTGATCAAGAAATTAAAATCCTGGAAGAAAACAAGGAATTGGAAAATGCTTTGAAA
AATATCCAGCTGCCTAAAGAGGAAATTAAGAAACTTGAAGTAGATGAAATTACTTTATG
GTACAAGATGATTCTTCCTCCTCAATTTGACAGATCAAAGAAGTATCCCTTGCTAATTCA
AGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTCTGTATTTGCTGTTAATTGGATATCT
TATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTGGTGGATGGTCGAGGAACAGCTTTC
CAAGGTGACAAACTCCTCTATGCAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGAC
CAGATTACAGCTGTCAGAAAATTCATAGAAATGGGTTTCATTGATGAAAAAAGAATAGC
CATATGGGGCTGGTCCTATGGAGGATACGTTTCATCACTGGCCCTTGCATCTGGAACTGG
TCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTACGCGTCTGTC
TACACAGAGAGATTCATGGGTCTCCCAACAAAGGATGATAATCTTGAGCACTATAAGAA
TTCAACTGTGATGGCAAGAGCAGAATATTTCAGAAATGTAGACTATCTTCTCATCCACGG
AACAGCAGATGATAATGTGCACTTTCAAAACTCAGCACAGATTGCTAAAGCTCTGGTTAA
TGCACAAGTGGATTTCCAGGCAATGTGGTACTCTGACCAGAACCACGGCTTATCCGGCCT
GTCCACGAACCACTTATACACCCACATGACCCACTTCCTAAAGCAGTGTTTCTCTTGTCA
GACTAA

>pPOB12

ATGAAGACTTGGGTAAAAATCGTATTTGGAGTTGCCACCTCTGCTGTGCTTGCCTTATTG
GTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATAACTCTGAAGAAAATACAATGAGA
GCACTCACACTGAAGGATATTTTAAATGGAACATTTTCTTATAAAACATTTTTTCCAAACT
GGATTTCAGGACAAGAATATCTTCATCAATCTGCAGATAACAATATAGTACTTTATAATA
TTGAAACAGGACAATCATATACCATTTTGAGTAATAGAACCATGAAAAGTGTGAATGCTT
CAAATTACGGCTTATCACCTGATCGGCAATTTGTATATCTAGAAAGTGATTATTCAAAGC
TTTGGAGATACTCTTACACAGCAACATATTACATCTATGACCTTAGCAATGGAGAATTTG
TAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTATTTATGCTGGTCGCCTGTTGGGAGTA
AATTAGCATATGTCTATCAAAACAATATCTATTTGAAACAAAGACCAGGAGATCCACCTT
TTCAAATAACATTTAATGGAAGAGAAAATAAAATATTTAATGGAATCCCAGACTGGGTTT
ATGAAGAGGAAATGCTTGCTACAAAATATGCTCTCTGGTGGTCTCCTAATGGAAAATTTT
TGGCATATGCGGAATTTAATGATACGGATATACCAGTTATTGCCTATTCCTATTATGGCG
ATGAACAATATCCTAGAACAATAAATATTCCATACCCAAAGGCTGGAGCTAAGAATCCC
GTTGTTCGGATATTTATTATCGATACCACTTACCCTGCGTATGTAGGTCCCCAGGAAGTGC
CTGTTCCAGCAATGATAGCCTCAAGTGATTATTATTTCAGTTGGCTCACGTGGGTTACTGA
TGAACGAGTATGTTTGCAGTGGCTAAAAAGAGTCCAGAATGTTTCGGTCCTGTCTATATG
TGACTTCAGGGAAGACTGGCAGACATGGGATTGTCCAAAGACCCAGGAGCATATAGAAG
AAAGCAGAACTGGATGGGCTGGTGGATTCTTTGTTTCAACACCAGTTTTCAGCTATGATG

CCATTTCGTACTACAAAATATTTAGTGACAAGGATGGCTACAAACATATTCACTATATCA
AAGACACTGTGGAAAATGCTATTCAAATTACAAGTGGCAAGTGGGAGGCCATAAATATA
TTCAGAGTAACACAGGATTCACTGTTTTATTCTAGCAATGAATTTGAAGAATACCCTGGA
AGAAGAAACATCTACAGAATTAGCATTGGAAGCTATCCTCCAAGCAAGAAGTGTGTTAC
TTGCCATCTAAGGAAAGAAAGGTGCCAATATTACACAGCAAGTTTCAGCGACTACGCCA
AGTACTATGCACTTGTCTGCTACGGCCCAGGCATCCCCATTTCCACCCTTCATGATGGAC
GCACTGATCAAGAAATTAAAATCCTGGAAGAAAACAAGGAATTGGAAAATGCTTTGAAA
AATATCCAGCTGCCTAAAGAGGAAATTAAGAAACTTGAAGTAGATGAAATTACTTTATG
GTACAAGATGATTCTTCCTCCTCAATTTGACAGATCAAAGAAGTATCCCTTGCTAATTCA
AGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTCTGTATTTGCTGTTAATTGGATATCT
TATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTGGTGGATGGTCGAGGAACAGCTTTC
CAAGGTGACAAACTCCTCTATGCAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGAC
CAGATTACAGCTGTCAGAAAATTCATAGAAATGGGTTTCATTGATGAAAAAAGAATAGC
CATATGGGGCTGGTCCTATGGAGGATACGTTTCATCACTGGCCCTTGCATCTGGAACTGG
TCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTACGCGTCTGTC
TACACAGAGAGATTCATGGGTCTCCCAACAAAGGATGATAATCTTGAGCACTATAAGAA
TTCAACTGTGATGGCAAGAGCAGAATATTTCAGAAATGTAGACTATCTTCTCATCCACGG
AACAGCAGATGATAATGTGCACTTTCAAAACTCAGCACAGATTGCTAAAGCTCTGGTTAA
TGCACAAGTGGATTTCCAGGCAATGTGGTACTCTGACCAGAACCACGGCTTATCCGGCCT
GTCCACGAACCACTTATACACCCACATGACCCACTTCCTAAAGCAGTGTTTCTCTTTGTCA
GACTAA

>pPOB15
TTACGCTAGCTTGGGGGATCCGCCATGAAGACTTGGGTAAAAATCGTATTTGGAGTTGCC
ACCTCTGCTGTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATA
ACTCTGAAGAAAATACAATGAGAGCACTCACACTGAAGGATATTTTAAATGGAACATTTT
CTTATAAAACATTTTTTCCAAACTGGATTTCAGGACAAGAATATCTTCATCAATCTGCAG
ATAACAATATAGTACTTTATAATATTGAAACAGGACAATCATATACCATTTTGAGTAATA
GAACCATGAAAAGTGTGAATGCTTCAAATTACGGCTTATCACCTGATCGGCAATTTGTAT
ATCTAGAAAGTGATTATTCAAAGCTTTGGAGATACTCTTACACAGCAACATATTACATCT
ATGACCTTAGCAATGGAGAATTTGTAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTATT
TATGCTGGTCGCCTGTTGGGAGTAAATTAGCATATGTCTATCAAAACAATATCTATTTGA
AACAAAGACCAGGAGATCCACCTTTTCAAATAACATTTAATGGAAGAGAAAATAAAATA
TTTAATGGAATCCCAGACTGGGTTTATGAAGAGGAAATGCTTGCTACAAAATATGCTCTC
TGGTGGTCTCCTAATGGAAAATTTTTGGCATATGCGGAATTTAATGATACGGATATACCA
GTTATTGCCTATTCCTATTATGGCGATGAACAATATCCTAGAACAATAAATATTCCATAC
CCAAAGGCTGGAGCTAAGAATCCCGTTGTTCGGATATTTATTATCGATACCACTTACCCT
GCGTATGTAGGTCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTGATTATTAT
TTCAGTTGGCTCACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAAAAAGAGTC
CAGAATGTTTCGGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACATGGGATTGT
CCAAAGACCCAGGAGCATATAGAAGAAAGCAGAACTGGATGGGCTGGTGGATTCTTTGT
TTCAACACCAGTTTTCAGCTATGATGCCATTTCGTACTACAAAATATTTAGTGACAAGGA
TGGCTACAAACATATTCACTATATCAAAGACACTGTGGAAAATGCTATTCAAATTACAAG
TGGCAAGTGGGAGGCCATAAATATATTCAGAGTAACACAGGATTCACTGTTTTATTCTAG
CAATGAATTTGAAGAATACCCTGGAAGAAGAAACATCTACAGAATTAGCATTGGAAGCT
ATCCTCCAAGCAAGAAGTGTGTTACTTGCCATCTAAGGAAAGAAAGGTGCCAATATTAC
ACAGCAAGTTTCAGCGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCCAGGCATC
CCCATTTCCACCCTTCATGATGGACGCACTGATCAAGAAATTAAAATCCTGGAAGAAAAC
AAGGAATTGGAAAATGCTTTGAAAAATATCCAGCTGCCTAAAGAGGAAATTAAGAAACT
TGAAGTAGATGAAATTACTTTATGGTACAAGATGATTCTTCCTCCTCAATTTGACAGATC
AAAGAAGTATCCCTTGCTAATTCAAGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTC
TGTATTTGCTGTTAATTGGATATCTTATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTG
GTGGATGGTCGAGGAACAGCTTTCCAAGGTGACAAACTCCTCTATGCAGTGTATCGAAA
GCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTCAGAAAATTCATAGAAATGG
GTTTCATTGATGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGATACGTTTCAT
CACTGGCCCTTGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTC
CAGCTGGGAATATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGG
ATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTCAGA
AATGTAGACTATCTTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTCAAAACTCA
GCACAGATTGCTAAAGCTCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCT

GACCAGAACCACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACATGACCCAC
TTCCTAAAGCAGTGTTTCTCTTTGTCAGACTAACTCGAGCATGCATCTAGAGGGC

>pPOB16

TTACGCTAGCTTGGGGGATCCAATGAAGACTTGGGTAAAAATCGTATTTGGAGTTGCCAC
CTCTGCTGTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATAAC
TCTGAAGAAAATACAATGAGAGCACTCACACTGAAGGATATTTTAAATGGAACATTTTCT
TATAAAACATTTTTTCCAAACTGGATTTCAGGACAAGAATATCTTCATCAATCTGCAGAT
AACAATATAGTACTTTATAATATTGAAACAGGACAATCATATACCATTTTGAGTAATAGA
ACCATGAAAAGTGTGAATGCTTCAAATTACGGCTTATCACCTGATCGGCAATTTGTATAT
CTAGAAAGTGATTATTCAAAGCTTTGGAGATACTCTTACACAGCAACATATTACATCTAT
GACCTTAGCAATGGAGAATTTGTAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTATTTA
TGCTGGTCGCCTGTTGGGAGTAAATTAGCATATGTCTATCAAAACAATATCTATTTGAAA
CAAAGACCAGGAGATCCACCTTTTCAAATAACATTTAATGGAAGAGAAAATAAAATATT
TAATGGAATCCCAGACTGGGTTTATGAAGAGGAAATGCTTGCTACAAAATATGCTCTCTG
GTGGTCTCCTAATGGAAAATTTTTGGCATATGCGGAATTTAATGATACGGATATACCAGT
TATTGCCTATTCCTATTATGGCGATGAACAATATCCTAGAACAATAAATATTCCATACCC
AAAGGCTGGAGCTAAGAATCCCGTTGTTCGGATATTTATTATCGATACCACTTACCCTGC
GTATGTAGGTCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTGATTATTATTT
CAGTTGGCTCACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAAAAAGAGTCCA
GAATGTTTCGGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACATGGGATTGTCC
AAAGACCCAGGAGCATATAGAAGAAAGCAGAACTGGATGGGCTGGTGGATTCTTTGTTT
CAACACCAGTTTTCAGCTATGATGCCATTTCGTACTACAAAATATTTAGTGACAAGGATG
GCTACAAACATATTCACTATATCAAAGACACTGTGGAAAATGCTATTCAAATTACAAGTG
GCAAGTGGGAGGCCATAAATATATTCAGAGTAACACAGGATTCACTGTTTTATTCTAGCA
ATGAATTTGAAGAATACCCTGGAAGAAGAAACATCTACAGAATTAGCATTGGAAGCTAT
CCTCCAAGCAAGAAGTGTGTTACTTGCCATCTAAGGAAAGAAAGGTGCCAATATTACAC
AGCAAGTTTCAGCGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCCAGGCATCCC
CATTTCCACCCTTCATGATGGACGCACTGATCAAGAAATTAAAATCCTGGAAGAAAACA
AGGAATTGGAAAATGCTTTGAAAAATATCCAGCTGCCTAAAGAGGAAATTAAGAAACTT
GAAGTAGATGAAATTACTTTATGGTACAAGATGATTCTTCCTCCTCAATTTGACAGATCA
AAGAAGTATCCCTTGCTAATTCAAGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTCT
GTATTTGCTGTTAATTGGATATCTTATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTG
GTGGATGGTCGAGGAACAGCTTTCCAAGGTGACAAACTCCTCTATGCAGTGTATCGAAA
GCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTCAGAAAATTCATAGAAATGG
GTTTCATTGATGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGATACGTTTCAT
CACTGGCCCTTGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTC
CAGCTGGGAATATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGG
ATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTCAGA
AATGTAGACTATCTTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTCAAAACTCA
GCACAGATTGCTAAAGCTCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCT
GACCAGAACCACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACATGACCCAC
TTCCTAAAGCAGTGTTTCTCTTTGTCAGACTAACTCGAGCATGCATCTAGAGGGCCCTATT
C

>pPOB17

TTACGCTAGCTTGGGGGATCCGCCCATGAAGACTTGGGTAAAAATCGTATTTGGAGTTGC
CACCTCTGCTGTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCAT
AACTCTGAAGAAAATACAATGAGAGCACTCACACTGAAGGATATTTTAAATGGAACATT
TTCTTATAAAACATTTTTTCCAAACTGGATTTCAGGACAAGAATATCTTCATCAATCTGCA
GATAACAATATAGTACTTTATAATATTGAAACAGGACAATCATATACCATTTTGAGTAAT
AGAACCATGAAAAGTGTGAATGCTTCAAATTACGGCTTATCACCTGATCGGCAATTTGTA
TATCTAGAAAGTGATTATTCAAAGCTTTGGAGATACTCTTACACAGCAACATATTACATC
TATGACCTTAGCAATGGAGAATTTGTAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTAT
TTATGCTGGTCGCCTGTTGGGAGTAAATTAGCATATGTCTATCAAAACAATATCTATTTG
AAACAAAGACCAGGAGATCCACCTTTTCAAATAACATTTAATGGAAGAGAAAATAAAAT
ATTTAATGGAATCCCAGACTGGGTTTATGAAGAGGAAATGCTTGCTACAAAATATGCTCT
CTGGTGGTCTCCTAATGGAAAATTTTTGGCATATGCGGAATTTAATGATACGGATATACC
AGTTATTGCCTATTCCTATTATGGCGATGAACAATATCCTAGAACAATAAATATTCCATA
CCCAAAGGCTGGAGCTAAGAATCCCGTTGTTCGGATATTTATTATCGATACCACTTACCC

TGCGTATGTAGGTCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTGATTATTA
TTTCAGTTGGCTCACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAAAAAGAGT
CCAGAATGTTTCGGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACATGGGATTG
TCCAAAGACCCAGGAGCATATAGAAGAAAGCAGAACTGGATGGGCTGGTGGATTCTTTG
TTTCAACACCAGTTTTCAGCTATGATGCCATTTCGTACTACAAAATATTTAGTGACAAGG
ATGGCTACAAACATATTCACTATATCAAAGACACTGTGGAAAATGCTATTCAAATTACAA
GTGGCAAGTGGGAGGCCATAAATATATTCAGAGTAACACAGGATTCACTGTTTTATTCTA
GCAATGAATTTGAAGAATACCCTGGAAGAAGAAACATCTACAGAATTAGCATTGGAAGC
TATCCTCCAAGCAAGAAGTGTGTTACTTGCCATCTAAGGAAAGAAAGGTGCCAATATTAC
ACAGCAAGTTTCAGCGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCCAGGCATC
CCCATTTCCACCCTTCATGATGGACGCACTGATCAAGAAATTAAAATCCTGGAAGAAAAC
AAGGAATTGGAAAATGCTTTGAAAAATATCCAGCTGCCTAAAGAGGAAATTAAGAAACT
TGAAGTAGATGAAATTACTTTATGGTACAAGATGATTCTTCCTCCTCAATTTGACAGATC
AAAGAAGTATCCCTTGCTAATTCAAGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTC
TGTATTTGCTGTTAATTGGATATCTTATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTG
GTGGATGGTCGAGGAACAGCTTTCCAAGGTGACAAACTCCTCTATGCAGTGTATCGAAA
GCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTCAGAAAATTCATAGAAATGG
GTTTCATTGATGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGATACGTTTCAT
CACTGGCCCTTGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTC
CAGCTGGGAATATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGG
ATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTCAGA
AATGTAGACTATCTTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTCAAAACTCA
GCACAGATTGCTAAAGCTCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCT
GACCAGAACCACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACATGACCCAC
TTCCTAAAGCAGTGTTTCTCTTTGTCAGACTAACTCGAGCATGCATCTAGAGGGC

>pPOB18
GATTACGCTAGCTTGGGGGATCCGCCATGGGGATGAAGACTTGGGTAAAAATCGTATTTG
GAGTTGCCACCTCTGCTGTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAG
AGTTCATAACTCTGAAGAAAATACAATGAGAGCACTCACACTGAAGGATATTTTAAATG
GAACATTTTCTTATAAAACATTTTTTCCAAACTGGATTTCAGGACAAGAATATCTTCATCA
ATCTGCAGATAACAATATAGTACTTTATAATATTGAAACAGGACAATCATATACCATTTT
GAGTAATAGAACCATGAAAAGTGTGAATGCTTCAAATTACGGCTTATCACCTGATCGGCA
ATTTGTATATCTAGAAAGTGATTATTCAAAGCTTTGGAGATACTCTTACACAGCAACATA
TTACATCTATGACCTTAGCAATGGAGAATTTGTAAGAGGAAATGAGCTTCCTCGTCCAAT
TCAGTATTTATGCTGGTCGCCTGTTGGGAGTAAATTAGCATATGTCTATCAAAACAATAT
CTATTTGAAACAAAGACCAGGAGATCCACCTTTTCAAATAACATTTAATGGAAGAGAAA
ATAAAATATTTAATGGAATCCCAGACTGGGTTTATGAAGAGGAAATGCTTGCTACAAAAT
ATGCTCTCTGGTGGTCTCCTAATGGAAAATTTTTGGCATATGCGGAATTTAATGATACGG
ATATACCAGTTATTGCCTATTCCTATTATGGCGATGAACAATATCCTAGAACAATAAATA
TTCCATACCCAAAGGCTGGAGCTAAGAATCCCGTTGTTCGGATATTTATTATCGATACCA
CTTACCCTGCGTATGTAGGTCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTG
ATTATTATTTCAGTTGGCTCACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAA
AAAGAGTCCAGAATGTTTCGGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACAT
GGGATTGTCCAAAGACCCAGGAGCATATAGAAGAAAGCAGAACTGGATGGGCTGGTGG
ATTCTTTGTTTCAACACCAGTTTTCAGCTATGATGCCATTTCGTACTACAAAATATTTAGT
GACAAGGATGGCTACAAACATATTCACTATATCAAAGACACTGTGGAAAATGCTATTCA
AATTACAAGTGGCAAGTGGGAGGCCATAAATATATTCAGAGTAACACAGGATTCACTGT
TTTATTCTAGCAATGAATTTGAAGAATACCCTGGAAGAAGAAACATCTACAGAATTAGCA
TTGGAAGCTATCCTCCAAGCAAGAAGTGTGTTACTTGCCATCTAAGGAAAGAAAGGTGC
CAATATTACACAGCAAGTTTCAGCGACTACGCCAAGTACTATGCACTTGTCTGCTACGGC
CCAGGCATCCCCATTTCCACCCTTCATGATGGACGCACTGATCAAGAAATTAAAATCCTG
GAAGAAAACAAGGAATTGGAAAATGCTTTGAAAAATATCCAGCTGCCTAAAGAGGAAAT
TAAGAAACTTGAAGTAGATGAAATTACTTTATGGTACAAGATGATTCTTCCTCCTCAATT
TGACAGATCAAAGAAGTATCCCTTGCTAATTCAAGTGTATGGTGGTCCCTGCAGTCAGAG
TGTAAGGTCTGTATTTGCTGTTAATTGGATATCTTATCTTGCAAGTAAGGAAGGGATGGT
CATTGCCTTGGTGGATGGTCGAGGAACAGCTTTCCAAGGTGACAAACTCCTCTATGCAGT
GTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTCAGAAAATTCAT
AGAAATGGGTTTCATTGATGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGAT
ACGTTTCATCACTGGCCCTTGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGC

TCCAGTCTCCAGCTGGGAATATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCC
AACAAAGGATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAAT
ATTTCAGAAATGTAGACTATCTTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTC
AAAACTCAGCACAGATTGCTAAAGCTCTGGTTAATGCACAAGTGGATTTCCAGGCAATGT
GGTACTCTGACCAGAACCACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACA
TGACCCACTTCCTAAAGCAGTGTTTCTCTTTGTCAGACTAACTCGAGCATGCATCTAGAG
GGCCCTATTC

> pQPOB5

ATGAGAGGATCGCATCACCATCACCATCACGGATCTGGCTCTGGATCTGGTATCGAGGGA
AGGCCTTATAATGGAACTGGATCCATGAAGACTTGGGTAAAAATCGTATTTGGAGTTGCC
ACCTCTGCTGTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATA
ACTCTGAAGAAAATACAATGAGAGCACTCACACTGAAGGATATTTTAAATGGAACATTTT
CTTATAAAACATTTTTTCCAAACTGGATTTCAGGACAAGAATATCTTCATCAATCTGCAG
ATAACAATATAGTACTTTATAATATTGAAACAGGACAATCATATACCATTTTGAGTAATA
GAACCATGAAAAGTGTGAATGCTTCAAATTACGGCTTATCACCTGATCGGCAATTTGTAT
ATCTAGAAAGTGATTATTCAAAGCTTTGGAGATACTCTTACACAGCAACATATTACATCT
ATGACCTTAGCAATGGAGAATTTGTAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTATT
TATGCTGGTCGCCTGTTGGGAGTAAATTAGCATATGTCTATCAAAACAATATCTATTTGA
AACAAAGACCAGGAGATCCACCTTTTCAAATAACATTTAATGGAAGAGAAAATAAAATA
TTTAATGGAATCCCAGACTGGGTTTATGAAGAGGAAATGCTTGCTACAAAATATGCTCTC
TGGTGGTCTCCTAATGGAAAATTTTTGGCATATGCGGAATTTAATGATACGGATATACCA
GTTATTGCCTATTCCTATTATGGCGATGAACAATATCCTAGAACAATAAATATTCCATAC
CCAAAGGCTGGAGCTAAGAATCCCGTTGTTCGGATATTTATTATCGATACCACTTACCCT
GCGTATGTAGGTCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTGATTATTAT
TTCAGTTGGCTCACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAAAAAGAGTC
CAGAATGTTTCGGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACATGGGATTGT
CCAAAGACCCAGGAGCATATAGAAGAAAGCAGAACTGGATGGGCTGGTGGATTCTTTGT
TTCAACACCAGTTTTCAGCTATGATGCCATTTCGTACTACAAAATATTTAGTGACAAGGA
TGGCTACAAACATATTCACTATATCAAAGACACTGTGGAAAATGCTATTCAAATTACAAG
TGGCAAGTGGGAGGCCATAAATATATTCAGAGTAACACAGGATTCACTGTTTTATTCTAG
CAATGAATTTGAAGAATACCCTGGAAGAAGAAACATCTACAGAATTAGCATTGGAAGCT
ATCCTCCAAGCAAGAAGTGTGTTACTTGCCATCTAAGGAAAGAAAGGTGCCAATATTAC
ACAGCAAGTTTCAGCGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCCAGGCATC
CCCATTTCCACCCTTCATGATGGACGCACTGATCAAGAAATTAAAATCCTGGAAGAAAAC
AAGGAATTGGAAAATGCTTTGAAAAATATCCAGCTGCCTAAAGAGGAAATTAAGAAACT
TGAAGTAGATGAAATTACTTTATGGTACAAGATGATTCTTCCTCCTCAATTTGACAGATC
AAAGAAGTATCCCTTGCTAATTCAAGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTC
TGTATTTGCTGTTAATTGGATATCTTATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTG
GTGGATGGTCGAGGAACAGCTTTCCAAGGTGACAAACTCCTCTATGCAGTGTATCGAAA
GCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTCAGAAAATTCATAGAAATGG
GTTTCATTGATGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGATACGTTTCAT
CACTGGCCCTTGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTC
CAGCTGGGAATATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGG
ATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTCAGA
AATGTAGACTATCTTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTCAAAACTCA
GCACAGATTGCTAAAGCTCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCT
GACCAGAACCACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACATGACCCAC
TTCCTAAAGCAGTGTTTCTCTTTGTCAGACTAA

> pIRPOB1

GGATCCGCCACCATGGGGAAGACTTGGGTAAAAATCGTATTTGGAGTTGCCACCTCTGCT
GTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATAACTCTGAAG
AAAATACAATGAGAGCACTCACACTGAAGGATATTTTAAATGGAACATTTTCTTATAAAA
CATTTTTTCCAAACTGGATTTCAGGACAAGAATATCTTCATCAATCTGCAGATAACAATA
TAGTACTTTATAATATTGAAACAGGACAATCATATACCATTTTGAGTAATAGAACCATGA
AAAGTGTGAATGCTTCAAATTACGGCTTATCACCTGATCGGCAATTTGTATATCTAGAAA
GTGATTATTCAAAGCTTTGGAGATACTCTTACACAGCAACATATTACATCTATGACCTTA
GCAATGGAGAATTTGTAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTATTTATGCTGGT
CGCCTGTTGGGAGTAAATTAGCATATGTCTATCAAAACAATATCTATTTGAAACAAAGAC

CAGGAGATCCACCTTTTCAAATAACATTTAATGGAAGAGAAAATAAAATATTTAATGGA
ATCCCAGACTGGGTTTATGAAGAGGAAATGCTTGCTACAAAATATGCTCTCTGGTGGTCT
CCTAATGGAAAATTTTTGGCATATGCGGAATTTAATGATACGGATATACCAGTTATTGCC
TATTCCTATTATGGCGATGAACAATATCCTAGAACAATAAATATTCCATACCCAAAGGCT
GGAGCTAAGAATCCCGTTGTTCGGATATTTATTATCGATACCACTTACCCTGCGTATGTA
GGTCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTGATTATTATTTCAGTTGG
CTCACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAAAAAGAGTCCAGAATGTT
TCGGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACATGGGATTGTCCAAAGACC
CAGGAGCATATAGAAGAAAGCAGAACTGGATGGGCTGGTGGATTCTTTGTTTCAACACC
AGTTTTCAGCTATGATGCCATTTCGTACTACAAAATATTTAGTGACAAGGATGGCTACAA
ACATATTCACTATATCAAAGACACTGTGGAAAATGCTATTCAAATTACAAGTGGCAAGTG
GGAGGCCATAAATATATTCAGAGTAACACAGGATTCACTGTTTTATTCTAGCAATGAATT
TGAAGAATACCCTGGAAGAAGAAACATCTACAGAATTAGCATTGGAAGCTATCCTCCAA
GCAAGAAGTGTGTTACTTGCCATCTAAGGAAAGAAAGGTGCCAATATTACACAGCAAGT
TTCAGCGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCCAGGCATCCCCATTTCC
ACCCTTCATGATGGACGCACTGATCAAGAAATTAAAATCCTGGAAGAAAACAAGGAATT
GGAAAATGCTTTGAAAAATATCCAGCTGCCTAAAGAGGAAATTAAGAAACTTGAAGTAG
ATGAAATTACTTTATGGTACAAGATGATTCTTCCTCCTCAATTTGACAGATCAAAGAAGT
ATCCCTTGCTAATTCAAGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTCTGTATTTGC
TGTTAATTGGATATCTTATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTGGTGGATGG
TCGAGGAACAGCTTTCCAAGGTGACAAACTCCTCTATGCAGTGTATCGAAAGCTGGGTGT
TTATGAAGTTGAAGACCAGATTACAGCTGTCAGAAAATTCATAGAAATGGGTTTCATTGA
TGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGATACGTTTCATCACTGGCCCT
TGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGA
ATATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGGATGATAATCT
TGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTCAGAAATGTAGACT
ATCTTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTCAAAACTCAGCACAGATTG
CTAAAGCTCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCTGACCAGAACC
ACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACATGACCCACTTCCTAAAGC
AGTGTTTCTCTTTGTCAGACTAAGCGGCCGCGACTACAAGGATGACGATGACAAGGATTA
CAAAGACGACGATGATAAGGACTATAAGGATG

> pIRPOB2

GGATCCGCCACCATGGGGAAGACTTGGGTAAAAATCGTATTTGGAGTTGCCACCTCTGCT
GTGCTTGCCTTATTGGTGATGTGCATTGTCTTACGCCCTTCAAGAGTTCATAACTCTGAAG
AAAATACAATGAGAGCACTCACACTGAAGGATATTTTAAATGGAACATTTTCTTATAAAA
CATTTTTTCCAAACTGGATTTCAGGACAAGAATATCTTCATCAATCTGCAGATAACAATA
TAGTACTTTATAATATTGAAACAGGACAATCATATACCATTTTGAGTAATAGAACCATGA
AAAGTGTGAATGCTTCAAATTACGGCTTATCACCTGATCGGCAATTTGTATATCTAGAAA
GTGATTATTCAAAGCTTTGGAGATACTCTTACACAGCAACATATTACATCTATGACCTTA
GCAATGGAGAATTTGTAAGAGGAAATGAGCTTCCTCGTCCAATTCAGTATTTATGCTGGT
CGCCTGTTGGGAGTAAATTAGCATATGTCTATCAAAACAATATCTATTTGAAACAAAGAC
CAGGAGATCCACCTTTTCAAATAACATTTAATGGAAGAGAAAATAAAATATTTAATGGA
ATCCCAGACTGGGTTTATGAAGAGGAAATGCTTGCTACAAAATATGCTCTCTGGTGGTCT
CCTAATGGAAAATTTTTGGCATATGCGGAATTTAATGATACGGATATACCAGTTATTGCC
TATTCCTATTATGGCGATGAACAATATCCTAGAACAATAAATATTCCATACCCAAAGGCT
GGAGCTAAGAATCCCGTTGTTCGGATATTTATTATCGATACCACTTACCCTGCGTATGTA
GGTCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTGATTATTATTTCAGTTGG
CTCACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAAAAAGAGTCCAGAATGTT
TCGGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACATGGGATTGTCCAAAGACC
CAGGAGCATATAGAAGAAAGCAGAACTGGATGGGCTGGTGGATTCTTTGTTTCAACACC
AGTTTTCAGCTATGATGCCATTTCGTACTACAAAATATTTAGTGACAAGGATGGCTACAA
ACATATTCACTATATCAAAGACACTGTGGAAAATGCTATTCAAATTACAAGTGGCAAGTG
GGAGGCCATAAATATATTCAGAGTAACACAGGATTCACTGTTTTATTCTAGCAATGAATT
TGAAGAATACCCTGGAAGAAGAAACATCTACAGAATTAGCATTGGAAGCTATCCTCCAA
GCAAGAAGTGTGTTACTTGCCATCTAAGGAAAGAAAGGTGCCAATATTACACAGCAAGT
TTCAGCGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCCAGGCATCCCCATTTCC
ACCCTTCATGATGGACGCACTGATCAAGAAATTAAAATCCTGGAAGAAAACAAGGAATT
GGAAAATGCTTTGAAAAATATCCAGCTGCCTAAAGAGGAAATTAAGAAACTTGAAGTAG

ATGAAATTACTTTATGGTACAAGATGATTCTTCCTCCTCAATTTGACAGATCAAAGAAGT
ATCCCTTGCTAATTCAAGTGTATGGTGGTCCCTGCAGTCAGAGTGTAAGGTCTGTATTTGC
TGTTAATTGGATATCTTATCTTGCAAGTAAGGAAGGGATGGTCATTGCCTTGGTGGATGG
TCGAGGAACAGCTTTCCAAGGTGACAAACTCCTCTATGCAGTGTATCGAAAGCTGGGTGT
TTATGAAGTTGAAGACCAGATTACAGCTGTCAGAAAATTCATAGAAATGGGTTTCATTGA
TGAAAAAAGAATAGCCATATGGGGCTGGTCCTATGGAGGATACGTTTCATCACTGGCCCT
TGCATCTGGAACTGGTCTTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGA
ATATTACGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGGATGATAATCT
TGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTCAGAAATGTAGACT
ATCTTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTCAAAACTCAGCACAGATTG
CTAAAGCTCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCTGACCAGAACC
ACGGCTTATCCGGCCTGTCCACGAACCACTTATACACCCACATGACCCACTTCCTAAAGC
AGTGTTTCTCTTTGTCAGACTGCGGCCGCGACTACAAGGATGACGATGACAAGGATTACA
AAGACGACGATGATAAGGACTATAAGG