

H24/3817

No Restrictions.

MONASH UNIVERSITY
THESIS ACCEPTED IN SATISFACTION OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ON... [REDACTED] 14 December 2004.....

Sec. Research Graduate School Committee

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing for the purposes of research, criticism or review. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

ERRATA

p 48 section 1.5.3 lines 3, 11 and 14: "Barbour *et al* (2002 (a))" for "Barbour *et al* (2002)"

p 49 lines 9, 11 and 14: "Barbour *et al* (2002 (a))" for "Barbour *et al* (2002)"

p 49 paragraph 2 lines 7, 10 and 12: "Barbour *et al* (2002(b))" for "Barbour *et al* (2002)"

p 63 line 1: "AD494(DE3)" for "AD494"

p 81 section 4.2.3 line 2: "AD494(DE3)" for "AD494"

Figure 4.4 figure legend: "sequences" for "sequecnes"

p 106 paragraph 3 line 5: "decreases" for "increases"

Figure 6.3 figure legend line 6: "anti-His antibody" for "anti-His"

p 148: insert line between references by Stein *et al* (1995) and Stephen *et al* (2002).

p 168: line 6: "(2002(a))" for "(2002)"

p168: line 10: "(2002(b))" for "(2002)"

ADDENDUM

Figure 1.1, figure legend: delete last two sentences and add

"The formation of cleaved serpin (I*) and active protease, with the rate constant k_4 , results from complete proteolysis. However, if the serpin kinetically traps the protease, the inhibitory complex (E-I*) is formed with the rate constant k_3 ."

p 81 section 4.2.3: delete second sentence and replace with

"The AD494(DE3) strain carries the DE3 lysogen and a mutation in the thioredoxin reductase gene facilitating disulfide bond formation in the cytoplasm, providing the potential to produce properly folded active proteins."

p 96 paragraph 1: delete last sentence and add

"The 90kDa entity which co-purified with His-MMCM2 was separated on 12.5% SDS-PAGE and the band excised and prepared for mass spectroscopy. 18 fragments were identifiable using MS-Fit search. The 18 peptides matched MMCM2 and no other high probability matches with any other proteins were identified. Mass spectroscopy demonstrated that the only protein present in the 90kDa entity was His-MMCM2."

p 108: delete sentence 3 and replace with

"The native His-MMCM2 demonstrates a sharp decrease in ellipticity commencing at 55°C consistent with loss of secondary structure."

p 109 paragraph 4: Add after sentence 2

"Because His-MMCM2 did not undergo a sigmoidal unfolding transition the use of the first derivative plot makes the calculation of the melting temperature error prone."

p 137: Add at the end of paragraph 1

"The proteases that inhibited the serpin3 members are consistent with the predicted P₁. From the data presented in this thesis, definitive identification of the P₁ residue will require either N-terminal sequencing or mass spectroscopy of serpin-protease complexes. These studies will be performed when functionally active recombinant 6C28 and 3E46 become available. "

MOLECULAR CHARACTERISATION OF THE MURINE
 α_1 -ANTICHYMOTRYPSIN-LIKE SERPINS

Anita Julieanne Horvath

Monash University
Department of Medicine
Box Hill Hospital

Thesis Submitted for the Degree of Doctor of Philosophy

April 2004

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University or other institution. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of thesis.

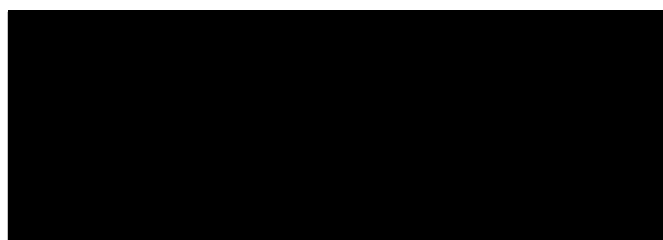
This thesis includes a manuscript accepted for publication in a peer reviewed journal. The ideas, development and writing of this manuscript were the principal responsibility of myself, the candidate, working within the Department of Medicine under the supervision of Dr Paul Coughlin.

Chapter 3 is presented as:

Horvath, A.J., Forsyth, S.L. and Coughlin, P.B (2004). Expression patterns of murine antichymotrypsin-like genes reflect evolutionary divergence at the *serpina3* locus. *J Mol Evol* (in press)

My contribution to the work involved the following:

- All methods in the manuscript, including: the design of primers, purification of RNA and gDNA and PCR.
- Wrote the first draft of the manuscript and worked on subsequent revisions.



ACKNOWLEDGEMENTS

Foremost, I would like to thank Dr Paul Coughlin. I could not have asked for a better supervisor; supportive, encouraging and always able to make time for me. Thank you for your guidance, wisdom and for providing a challenging yet friendly work environment.

I am extremely grateful for the constant help and support from Melinda Missen and Dr Sharon Forsyth. Their friendship has made working in the lab an absolute pleasure.

The analysis of serpin2A expression in haemopoietic progenitors would not have been possible without the generous guidance and advice of Dr Ivan Bertoncello and help from Brenda Williams and Kate Sells. Thank you for providing reagents, the animal handling and harvesting of progenitors, and for assisting in the analysis of transplanted animals. Many thanks to Assoc Prof Rob Pike and the members of his laboratory, in particular Dr Noelene Quinsey, for their assistance with kinetic analysis, use of reagents and for allowing me to work in their laboratory and making my visits so much fun. I am indebted to Dr Steve Bottomley, who taught me how to use the CD machine and provided expert advice on kinetic analysis.

I would also like to thank the people who have provided advice, help and technical support throughout my studies; Corrine Hitchen, Mythily Sachchithananthan, Andre Samson, Dr Stan Stasinopoulos, Dr Rob Medcalf, Dr Hong Yu and in particular, Dr Simone Schoenwaelder, for coming to my rescue on more than one occasion and for being a great friend. Most importantly they have all made the Department of Medicine a great place to work in. Thank you to Professor Hatem Salem for giving me the opportunity to complete my PhD in the Department of Medicine.

Thank you to family and friends for their words of encouragement, especially to all four parents for always being there and my sister Monika for trying to keep me balanced.

To my husband Mathew, thank you for your love, patience, encouragement and for making me laugh when I needed to most. I can accomplish anything with you by my side.

DEDICATION

*For my parents, Ibolyka and Peter,
For always wanting and providing the best for me
and for being my inspiration.
Thank-you for your love and faith in me.*

TABLE OF CONTENTS

PUBLICATIONS.....	11
ABBREVIATIONS	12
TABLE OF FIGURES.....	16
ABSTRACT	17

CHAPTER 1: THE SERPIN SUPERFAMILY: STRUCTURE, FUNCTION

AND EVOLUTION.....	18
1.1. SERPIN CONFORMATION AND MECHANISM OF ACTION	19
1.1.1. Mechanism of inhibition	20
1.2. SERPIN CONFORMATIONS	21
1.2.1. The native structure.....	21
1.2.2. Cleaved serpins.....	23
1.2.3. Latent serpin.....	24
1.2.4. δ -conformation.....	24
1.2.5. Serpin – protease structure	25
1.2.6. Serpin polymerisation	25
1.2.7. Regulation of conformational change	27
1.2.8. Co-factor binding	28
1.3. REACTIVE CENTRE LOOP AND PROTEASE SPECIFICITY	30
1.3.1. Importance of reactive centre loop composition	30
1.3.2. Importance of reactive centre loop length.....	32
1.3.3. Protease specificity.....	32
1.3.4. Cross-class inhibition.....	33
1.4. PHYLOGENY AND EVOLUTION OF THE SERPIN SUPERFAMILY	34
1.4.1. The A clade serpins.....	36
1.4.1.a. α_1 -Antitrypsin (SERPIN A1).....	37
1.4.1.b. α_1 -Antichymotrypsin (SERPINA3).....	38

1.4.2. The B clade serpins.....	39
1.4.2.a. The C-D interhelical loop.....	41
1.4.3. Clades C-P.....	42
1.5. THE MURINE MULTIGENE SERPIN FAMILIES	42
1.5.1. Comparison of the serpin clade A between human and mice.....	43
1.5.2. Serpina3 family of serpins	43
1.5.2.a. Gene organisation of the mouse <i>serpina3</i> locus	45
1.5.2.b. Expression of the <i>serpina3</i> genes.....	45
1.5.2.c. Serpin2A (<i>serpina3g</i>)	46
1.5.3. The <i>serpina1</i> family	48
1.5.4. Murine expansion of intracellular serpins.....	50
1.6. SERPINS AND THEIR ROLE IN HAEMOPOIESIS.....	51
1.6.1. Haemopoiesis.....	51
1.6.2. Haemopoiesis and proteolysis.....	52
1.6.3. Intracellular serpins and haemopoiesis.....	54
1.6.4. Extracellular serpins and haemopoiesis	56
 CHAPTER 2: MATERIALS AND METHODS	58
2.1. GENERAL REAGENTS AND EQUIPMENT	58
2.1.1. Reagents.....	58
2.1.2. Enzymes.....	58
2.1.3. Proteases	59
2.1.4. Equipment	59
2.2. MOLECULAR BIOLOGY	60
2.2.1. DNA purification.....	60
2.2.1.a. PCR purification and gel extraction of DNA.....	60
2.2.1.b. Mini Preparation	60
2.2.1.c. Maxi Preparation	60
2.2.2. DNA electrophoresis.....	61
2.2.3. Polymerase chain reaction (PCR)	62
2.2.4. DNA Sequencing.....	62
2.2.5. Preparation and transformation of chemically competent cells.....	63
2.2.6. Transformation of circular plasmid DNA	63

2.2.7. Transformation of ligations	64
2.2.8. pGEMT-easy cloning	64
2.2.9. Colony PCR.....	64
2.2.10. RNA extraction and purification	65
2.2.10.a. Tissue RNA	65
2.2.10.b. Cell line RNA.....	66
2.2.11. RNA integrity and concentration.....	66
2.2.12. Reverse transcription.....	66
2.2.13. RT-PCR.....	67
2.2.14. Semi-quantitative RT-PCR	67
2.2.15. Genomic DNA extraction.....	67
2.3. PROTEIN BIOCHEMISTRY	68
2.3.1. Generation of recombinant proteins by <i>in vitro</i> transcription/translation	68
2.3.2. SDS-PAGE.....	69
2.3.3. Non-dentauring polyacrylamide gel electrophoresis	70
2.3.4. Transverse urea gradient (TUG) gel electrophoresis.....	70
2.3.5. Western Blotting	71
2.3.6. Antisera	72
2.3.7. Protein concentration determination	72
2.3.7.a. Bio Rad Protein Assay	72
2.3.7.b. Absorbance at 280nM.....	73
2.4. CELL CULTURE	73
2.4.1. Culture conditions.....	73
2.4.2. WEHI-conditioned media	74
2.4.3. Storage of cells.....	74
2.4.4. Preparation of cells for RNA extraction.....	74
Declaration for thesis Chapter 3	75
CHAPTER 3: GENE EXPRESSION STUDIES OF THE SERPINA3 AND SERPINA1 LOCUS.....	77

CHAPTER 4 : PRODUCTION AND PURIFICATION OF BACTERIAL

RECOMBINANT MOUSE SERPINA3 PROTEINS	78
4.1. INTRODUCTION	78
4.2. METHODS	79
4.2.1. RT-PCR of mouse tissue cDNA and construction of expression plasmids	79
4.2.2. Trial inductions of <i>serpina3</i> proteins in BL21(DE3)pLysS.....	80
4.2.3. Trial inductions of His-MMCM2 and His-3E46 in AD494 cells	81
4.2.4. Analysis of trial inductions	82
4.2.5. Large scale-production of <i>serpina3</i> proteins	82
4.2.6. Purification of <i>serpina3</i> recombinants on nickel-agarose.....	83
4.2.7. Purification of recombinant His-tagged <i>serpina3</i> proteins by anion- exchange chromatography	84
4.2.8. Purification of His-MMCM2 by gel filtration	85
4.2.9. Mass spectroscopy.....	85
4.2.10. Multiple sequence alignments and prediction of cleavage point, molecular weight, extinction co-efficient and <i>pI</i>	86
4.3. RESULTS	87
4.3.1. Isolation of <i>serpina3</i> cDNA transcripts.....	87
4.3.2. Small-scale induction of <i>serpina3</i> proteins	90
4.3.3. Optimisation of expression	91
4.3.4. Large scale expression and purification of <i>serpina3</i> proteins	92
4.3.4.a. Purification of His-EB22.4	92
4.3.4.b. Purification of His-MMCM2	93
4.3.4.c. Gel filtration of His-MMCM2	94
4.3.4.d. Purification of His-3E46	94
4.3.4.e. Purification of His-6C28	95
4.3.5. Mass Spectroscopy of His-MMCM2 90 kDa complex.....	96
4.3.6. Comparison of purified <i>serpina3</i> proteins	96
4.4. DISCUSSION	96
4.5. CONCLUSION	100

CHAPTER 5: BIOPHYSICAL CHARACTERISATION OF RECOMBINANT

SERPINA3 PROTEINS	101
5.1. INTRODUCTION	101
5.2. METHODS.....	102
5.2.1. <i>Reactive centre loop cleavage studies</i>	102
5.2.2. <i>TUG electrophoresis of native recombinant serpina3 proteins</i> ...	102
5.2.3. <i>TUG electrophoresis of cleaved recombinant serpina3 proteins</i>	103
5.2.4. <i>Circular dichroism (CD) analysis of thermal stability</i>	103
5.2.5. <i>Polymerisation Studies</i>	104
5.2.6. <i>In vitro transcription/translation</i>	104
5.3. RESULTS	104
5.3.1. <i>Conformational analysis of native recombinant serpina3 proteins</i>	104
5.3.2. <i>Cleavage studies of His-MMCM2 and His-EB22.4</i>	106
5.3.3. <i>Circular dichroism analysis of recombinant serpina3 proteins</i>	108
5.3.4. <i>Polymerisation studies</i>	110
5.3.5 <i>Conformational studies of in vitro transcribed/translated His-3E46 and His-6C28</i>	111
5.4. DISCUSSION	112
5.5. CONCLUSION	116

CHAPTER 6: PROTEASE INTERACTIONS OF MUIRNE SERPINA3

SERPINS.....	117
6.1. INTRODUCTION	117
6.2. METHODS.....	118
6.2.1. <i>Analysis of complex formation of recombinant His-MMCM2 and His-EB22.4 using SDS-PAGE</i>	118
6.2.1.a. <i>Serine proteases</i>	118
6.2.1.b. <i>Cysteine proteases</i>	119
6.2.2. <i>Complex formation of in vitro transcribed/translated serpina3 proteins</i>	119
6.2.3. <i>Enzymes, assay buffers and substrates</i>	120

6.2.4. Active site titration of proteases.....	120
6.2.5. BSA coated-microtitre plates.....	121
6.2.6. Determination of the stoichiometry of inhibition	121
6.2.7. Kinetics of inhibitory reactions	122
6.2.7.a. Continuous method	122
6.2.7.b. Discontinuous method	123
6.3. RESULTS.....	123
6.3.1 The RCL of the mouse <i>serpina3</i> proteins predict a varied protease inhibitory spectrum.....	123
6.3.2. Protease interactions of His-EB22.4.....	125
6.3.3. Protease interactions of His-MMCM2	126
6.3.4. Protease interactions of His-3E46	127
6.3.5. Protease interactions of His-6C28	128
6.3.6. Kinetics of interactions.....	129
6.3.7. Inhibitory kinetics of recombinant His-EB22.4 with chymotrypsin, cathepsin G and elastase.....	129
6.3.8. Inhibitory kinetics of recombinant His-MMCM2 with trypsin.....	131
6.4. DISCUSSION	132
6.5. CONCLUSION.....	137

CHAPTER 7: ESTABLISHMENT OF A RETROVIRAL-MEDIATED

TRANSDUCTION SYSTEM TO INTRODUCE SERPIN2A INTO

MURINE HAEMOPOIETIC PROGENITORS 138

7.1. INTRODUCTION	138
7.2. METHODS.....	139
7.2.1. Cell lines and vector construction	139
7.2.2. Generation of ecotropic producer cells.	140
7.2.3. Western blot analysis.....	141
7.2.4. Mice	141
7.2.5. Bone marrow harvesting and sorting	142
7.2.6. Negative Immunomagnetic Cell Separation.....	143
7.2.7. Isolation of Lin ⁻ Sca-1 ⁺ HPC	143

7.2.8. Fluorescence activated cell sorting (FACS).....	144
7.2.9. Retroviral transduction of haemopoietic progenitors.....	144
7.2.10. Flow cytometry of retrovirally transduced progenitor cells.....	145
7.2.11. High proliferative potential colony forming (HPP-CFC) assay....	145
7.2.12. Competitive repopulation transplant assay.....	147
7.2.13. Transplant analysis.....	147
7.3. RESULTS	148
7.3.1. Establishment of a retroviral producer cell line	148
7.3.2. Retroviral transduction of murine progenitors by co-culture.....	149
7.3.3. Optimisation of retroviral transduction protocol.....	151
7.3.4 Retroviral transduction of Lin ⁻ Sca ⁺ progenitors using retroviral supernatant.....	154
7.4. DISCUSSION	155
7.5. CONCLUSION.....	158
CHAPTER 8 : GENERAL DISCUSSION	159
8.1. THE MOUSE SERPINA3 LOCUS	159
8.2 MOLECULAR CHARACTERISATION OF EXTRACELLULAR SERPINA3 MEMBERS.....	161
8.3 MOLECULAR CHARACTERISATION OF INTRACELLULAR SERPINA3 MEMBERS	162
8.4. FUTURE DIRECTIONS.....	165
BIBLIOGRAPHY.....	167

PUBLICATIONS

Forsyth, S., Horvath, A. and Coughlin, P. (2003). "A review and comparison of the murine alpha1-antitrypsin and alpha1-antichymotrypsin multigene clusters with the human clade A serpins." *Genomics* 81(3): 336-45.

Winkler, I. G., Hendy, J., Coughlin, P., Horvath, A. and Levesque, J. P. (2004). "Down-Regulation of the Serine-Protease Inhibitors, Serpina1 and Serpina3, in the Bone Marrow During Hematopoietic Progenitor Cell Mobilization Results in Increased Neutrophil Serine-Protease Activity." (Submitted).

Horvath, A.J., Forsyth, S.L. and Coughlin, P.B (2004). Expression patterns of murine antichymotrypsin-like genes reflect evolutionary divergence at the *serpina3* locus. *J Mol Evol* (in press) (presented as Chapter 3)

A reprint or copy of the first two manuscripts are located at the rear of this thesis.

ABBREVIATIONS

α_1 -ACT	α_1 -antichymotrypsin
APS	ammonium persulphate
Anb	antibiotic/antimycotic solution
α_1 -AT	α_1 -antitrypsin
BM	bone marrow
BSA	bovine serum albumin
CatG	cathepsin G
CatL	cathepsin L
CatV	cathepsin V
CBG	cortisol binding globulin
CD	circular dichroism
CFU	colony forming unit
CrmA	cowpox response modifier A
CSF	colony stimulating factor
DEPC	diethyl pyrocarbonate
dH ₂ O	distilled water
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
E.coli	Eschericia coli
EDTA	ethylenediaminetetraacetic acid
FACS	fluorescence-activated cell sorting
FCS	fetal calf serum
FITC	fluorescein isothiocyannate
Flt-3-L	Fms-like tyrosine kinase 3-ligand
5-FU	5-fluorouracil
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
G-CSF	granulocyte colony-stimulating factor
gDNA	genomic DNA

GFP	green fluorescent protein
GM-CSF	granulocyte-macrophage colony-stimulating factor
HI-FCS	heat-inactivated fetal calf serum
His-tag	polyhistidine tag
HIV	human immunodeficiency virus
HLE	human leukocyte elastase
HPP	high proliferative potential progenitor
HPC	haemopoietic progenitor cell
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
HSC	haemopoietic stem cell
IFN- γ	interferon- γ
IL	interleukin
IPTG	isopropyl- β -D-thiogalactopyranoside
kDA	kilodalton
k_a	second order association rate constant for inhibition
k_{obs}	observed first order association rate constant for inhibition
LB	Luria Broth
LD-BM	low density bone marrow
Lin ⁻	lineage negative
LPP	low proliferative potential
MALDI	matrix assisted laser desorption ionisation
α MEM	alpha modification of Eagles Minimal Essential Medium
MNEI	monocyte/neutrophil elastase inhibitor
mRNA	messenger ribonucleic acid
MSCV	murine stem cell virus
MMF	mégakaryocyte maturation factor
MW	molecular weight
Ni-agarose	nickel-agarose
OD	optical density
Ov-serpin	ovalbumin (intracellular) serpin family

PAGE	polyacrylamide gel electrophoresis
PAI-1	plasminogen activator inhibitor type 1
PAI-2	plasminogen activator inhibitor type 2
PB	peripheral blood
PBS	phosphate buffered saline
PCI	protein C inhibitor
PCR	polymerase chain reaction
PEG	polyethylene glycol
PI	protease inhibitor
PMSF	phenylmethyl sulphonyl fluoride
PSA	prostate specific antigen
PVDF	polyvinylidene difluoride
Rb	retinoblastoma susceptibility protein
RCL	reactive centre loop
RNA	ribonucleic acid
RNase	ribonuclease
RT-PCR	reverse transcribed-polymerase chain reaction
SCA	stem cell antigen
SCCA-1	squamous cell carcinoma antigen-1
SCCA-2	squamous cell carcinoma antigen-2
SCF	stem cell factor
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SI	stoichiometry of inhibition
S→R	stressed to relaxed (transition)
T _m	melting temperature
TAE	Tris-acetate-EDTA buffer
TBE	Tris-boric acid-EDTA buffer
TE	Tris-EDTA buffer
TEMED	N,N,N',N'-Tetramethylethylenediamine
tPA	tissue plasminogen activator
TNF- α	tumour necrosis factor- α

TUG	transverse urea gradient
UTR	untranslated region
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Z-Phe-Arg-AMC	(benzyloxycarbonyl-L-phenylalanyl-L-arginine-7-amino-4-methylcoumarin).
ZPI	protein Z-dependent protease inhibitor

TABLE OF FIGURES

Figure Number	Between Pages	Figure Number	Between Pages
1.1	20-21	5.5	108-109
1.2	22-23	5.6	108-109
1.3	23-24	5.7	109-110
1.4	43-44	5.8	110-111
1.5	51-52	5.9	111-112
4.1	79-80	5.10	111-112
4.2	87-88	6.1	124-125
4.3	88-89	6.2	125-126
4.4	88-89	6.3	126-127
4.5	90-91	6.4	127-128
4.6	90-91	6.5	128-129
4.7	91-92	6.6	129-130
4.8	92-93	6.7	130-131
4.9	92-93	6.8	130-131
4.10	93-94	6.9	131-132
4.11	93-94	7.1	139-140
4.12	93-94	7.2	147-148
4.13	94-95	7.3	148-149
4.14	94-95	7.4	149-150
4.15	95-96	7.5	149-150
4.16	95-96	7.6	150-151
4.17	96-97	7.7	151-152
5.1	105-106	7.8	151-152
5.2	106-107	7.9	152-153
5.3	106-107	7.10	154-155
5.4	107-108		

ABSTRACT

One of the major serpins circulating in human plasma is α_1 -antichymotrypsin. It is predominantly expressed in liver, and participates in inflammatory responses through the regulation of the granular protease cathepsin G. Unlike the human genome which contains a single copy of the α_1 -antichymotrypsin gene, analysis of the syntenic murine *serpina3* locus revealed an extensive expansion to fourteen genes. Although the murine genes share a high degree of overall sequence homology, the reactive centre loop domain, which determines target protease specificity, is markedly divergent. The studies presented in this thesis describe the characterisation of the *serpina3* locus.

Using a reverse transcription PCR technique designed to distinguish between the highly homologous *serpina3* genes, nine members were shown to be expressed in a tissue specific manner. As expected, mouse *serpina3* genes were strongly expressed in the liver, although prominent expression of some serpins was observed in brain, immune and haemopoietic tissues.

Functional characterisation was carried out on four of the most abundantly expressed *serpina3* members, EB22.4, MMCM2, 3E46 and 6C28. Biochemical and biophysical analysis of these serpins showed that they were able to undergo the conformational change required for protease inhibition. As predicted, they differed in their ability to inhibit the serine and cysteine proteases as demonstrated by their ability to form covalent complexes and in kinetic assays.

Previous studies indicated that *serpina3g* is strongly expressed in haemopoietic stem cells and down-regulated upon differentiation. In order to explore the biological significance of this observation a method was developed to express *serpina3g* constitutively in murine haemopoietic progenitors *in vivo*. This method provides a platform for further work on the biology of *serpina3* genes.

CHAPTER 1

THE SERPIN SUPERFAMILY: STRUCTURE, FUNCTION AND EVOLUTION

Serpins belong to a superfamily of proteins which share a highly ordered 3-dimensional tertiary structure (Carrell and Travis 1985). They were first described as serine protease inhibitors, however a multitude of alternative functions have been ascribed to them. The serpin scaffold is present in all members and is fundamental to their inhibitory activity. Despite the conservation of their structure, the amino acid sequence identity between members can be as low as 25%, suggesting that they are quite distinct from each other and possess variable biological functions (van Gent *et al.* 2003). The presence of extensions at both C- and N- termini and the degree of glycosylation also allows for variation in function (Perkins *et al.* 1990; Frank *et al.* 2003; McCoy *et al.* 2003). Some members of the superfamily possess no protease inhibitory function and have developed alternate roles such as the hormone trafficking serpins, corticosteroid binding globulin (Hammond *et al.* 1987), thyroxine-binding globulin (Flink *et al.* 1986) and the protein folding chaperone serpin, HSP-47 (Hirayoshi *et al.* 1991).

Approximately 700 members have been identified, ranging across species from prokaryotes to multicellular eukaryotes (Irving *et al.* 2002 (a)). With the continued sequencing of new genomes and the completion of existing ones, it is expected that even more members will be identified and additional functions attributed to this structurally similar yet functionally diverse superfamily of proteins.

1.1. SERPIN CONFORMATION AND MECHANISM OF ACTION

Serpins are single chain polypeptides of ~400 residues which fold into the conserved serpin structure, consisting of three β -sheets (A-C) and nine α -helices (Huber and Carrell 1989), although a few serpins possess only 8 α -helices, such as the viral serpin CrmA (Simonovic *et al.* 2000). Approximately 30-40 residues from the C-terminus is an exposed, mobile loop structure containing the reactive site residues which, in inhibitory serpins, interacts with the cognate protease. This region is termed the reactive centre loop (RCL). Cleavage of the RCL results in a dramatic conformational change in the serpin molecule and is characterised by insertion of the N-terminal portion of the cleaved RCL into β -sheet A (Carrell and Travis 1985). This conformational change is accompanied by a dramatic increase in the molecular stability of the protein. The denaturation temperature (T_m) of the protein is increased from a typical 55°C in the native state to 100°C in the loop-inserted form (Carrell and Owen 1985). This change in molecular stability is termed the stressed to relaxed (S→R) transition. The conformational change results in irreversible inhibition of the target protease as loop insertion translocates the protease and distorts its active site.

In addition to its importance in the inhibition of target proteases, the conformational flexibility of the serpin molecule is also utilised as a mechanism to regulate activity. The non-inhibitory hormone binding serpins, cortisol binding globulin and thyroxine binding globulin, undergo conformational change on cleavage of their RCL to release hormones at sites of inflammation (Pemberton *et al.* 1988). Therefore, the serpin structure is important for both inhibitory and non-inhibitory activity. However, not all serpins undergo conformational change. The non-inhibitory serpins, ovalbumin, maspin, angiotensinogen and PEDF are unable to undergo the S→R conformational change on loop cleavage (Mast *et al.* 1991; Becerra *et al.* 1995; Pemberton *et al.* 1995).

1.1.1. Mechanism of inhibition

Serpins inhibit serine proteases by a unique process known as the "suicide substrate inhibition mechanism" (Potempa *et al.* 1994). Irreversible protease inhibition is achieved through the formation of a covalent complex in which both serpin and protease are rendered inactive. This method is distinct from other classes of protease inhibitors, Kunitz and Kazal type, which inhibit through a reversible, tight, non-covalent lock and key mechanism (reviewed in (Bode and Huber 2000)). Serpin-protease complexes can be observed on SDS-PAGE as their covalent nature prevents dissociation.

The scheme shown in Figure 1.1 represents the kinetic mechanism of inhibition (Gettins 2002 (a)). The initial interaction between the serpin (I) and protease (E) forms a reversible non-covalent Michaelis complex (EI) (Stone and LeBonniec 1997). Cleavage of the peptide bond at P_1 - P_1' occurs by nucleophilic attack by the catalytic serine of the protease on the carbonyl carbon of the P_1 residue. The attack proceeds through the standard covalent tetrahedral intermediate of a substrate cleavage reaction, with the subsequent formation of an ester bond between the hydroxyl group of the Ser residue and the carbonyl carbon of the P_1 residue. This forms the covalent acyl-enzyme intermediate (EI^*) (Lawrence *et al.* 1995; Wilczynska *et al.* 1995). With the peptide bond between the serpin P_1 - P_1' residues broken, the N-terminal strand of the RCL is released and the loop can insert into the A- β sheet.

The mechanism becomes branched at this point with the rate of loop insertion determining whether the serpin behaves as a substrate or an inhibitor (Lawrence *et al.* 2000). If the loop insertion process is rapid ($k_3 > k_4$), the inhibitory pathway proceeds towards full protease inhibition ($E-I^*$) by trapping the acyl intermediate before the protease can complete the deacylation reaction. The rapid loop insertion, accompanied by conformational change in the serpin and translocation of the protease, distorts the protease active site and results in complete inhibition (Huntington *et al.* 2000; Lawrence *et al.* 2000). However, if loop insertion proceeds too slowly ($k_4 > k_3$), then the

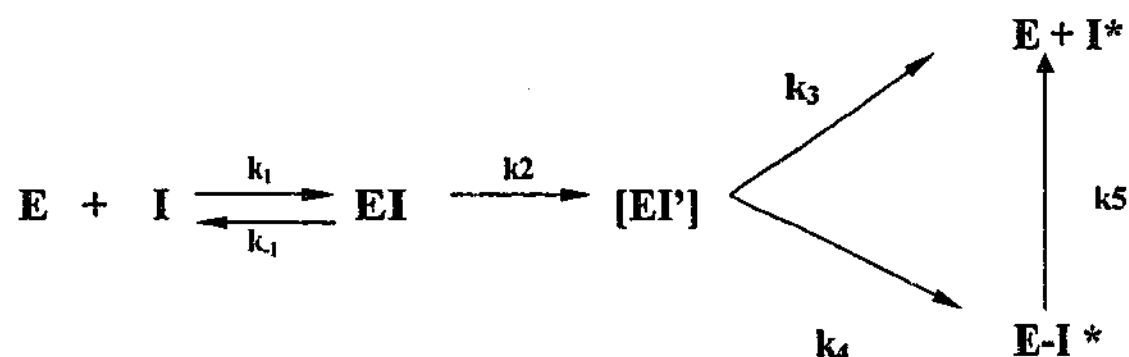


Figure 1.1: Schematic representation showing the mechanism of inhibition of a protease by a serpin. The Michaelis-like complex (EI) is formed after the initial interaction of the protease (E) and serpin (I), with the forward rate constant k_1 and the back constant of k_{-1} . This proceeds to formation of the covalent acyl-intermediate (EI'), with the k_2 . From this point the reaction can proceed down two different pathways which primarily depends on how rapidly the cleaved serpin loop is able to insert into the A β -sheet. The formation of cleaved serpin (I*) and active protease, with the rate constant k_3 , results from complete proteolysis. However, if the serpin kinetically traps the protease, the inhibitory complex (E-I*) is formed with the rate constant k_4 .

deacylation reaction continues, the serpin is cleaved without protease inhibition ($E + I$) and active protease is released. The inhibited serpin-protease complex is stable for days-weeks, depending on the serpin-protease pair (Calugaru *et al.* 2001; Zhou *et al.* 2001; Plotnick *et al.* 2002). The effectiveness with which the enzyme is trapped is reflected in the rate constant (k_5). It has been estimated that k_5 is at minimum 5-7 orders of magnitude less than (k_3). This mechanism of protease inhibition was confirmed by the elucidation of the crystal structure of a covalent serpin-protease complex (Huntington *et al.* 2000).

1.2. SERPIN CONFORMATIONS

Extensive investigations have been carried out on the three dimensional serpin structure utilising X-ray crystallography and other structural techniques. The first serpin crystal structure to be elucidated was that of cleaved α_1 -antitrypsin (α_1 -AT) (Loebermann *et al.* 1984). This has become the archetypal structure against which subsequent serpin structures are compared. A number of crystal structures have been solved for a wide variety of serpins in various conformations over the years. Structures have been described for native inhibitory and non-inhibitory serpins, RCL cleaved, latent serpin, polymerised serpin, inactive dimers and serpin-protease complex. These studies have emphasised the important relationship between structure and function.

1.2.1. The Native Structure

The first crystal structure of a native serpin was that of the non-inhibitory serpin, ovalbumin, by Stein *et al.* (Stein *et al.* 1990; Stein *et al.* 1991). Subsequently, the first native crystal structure of an inhibitory serpin was determined for antithrombin (Carrell *et al.* 1994; Schreuder *et al.* 1994). To date, a further 9 serpins have been crystallised in their native conformation. Comparison revealed tight conservation of both secondary and tertiary

structure between all members, including inhibitory and non-inhibitory serpins. They share approximately 51 residues, mostly buried, which are important for correct packaging of the serpin core (Irving *et al.* 2000).

The native structure highlights the exposure of the RCL-region outside the central serpin molecule and shows its attachment at the N-terminus to s5A of β -sheet A and at the C-terminus to s1C of β -sheet C (Figure 1.2). Of the three β -sheets, the A-sheet is the largest and is composed of five strands which are essentially antiparallel except for the two central strands. Insertion of the cleaved RCL occurs between these central strands making β -sheet A fully antiparallel. Seven of the eight α -helices are located at the rear of the serpin molecule, relative to β -sheet-A, whilst helix-F lies across the front. It has been proposed that helix-F may participate in stabilisation of the five-stranded conformation and opening of the A-sheet (Gettins 2002 (b)).

One of the striking features highlighted by the comparison of native structures is the variety of conformations the RCL adopts. The crystal structure of the archetypal serpin α_1 -AT revealed a RCL in extended β -conformation whereas the RCL of both ovalbumin and α_1 -antichymotrypsin (α_1 -ACT) are helical (Stein *et al.* 1991; Wei *et al.* 1994; Elliott *et al.* 1998 (a)). Native ovalbumin consists of a 3 turn α -helix and native α_1 -ACT is a distorted two-turn α -helix with strand s1C partially dissociated from the C-sheet. In the PAI-2 structure the RCL was shown to have no regular secondary structure (Harrop *et al.* 1999). Given the sequence variability of the RCL it is no surprise that its conformation differs between serpins.

The conformation of the RCL may play a role in the regulation of serpin activity. The native crystal structures of antithrombin and heparin cofactor II are unusual as they display an insertion of two residues of the RCL, P₁₅ and P₁₄, into β -sheet A. This is made possible due to a partial expansion in the A- β -sheet (Carrell *et al.* 1994). In their native conformations, antithrombin and heparin cofactor II, interact with their target proteases inefficiently. Binding of the polysaccharide heparin to antithrombin and heparin cofactor II expels the

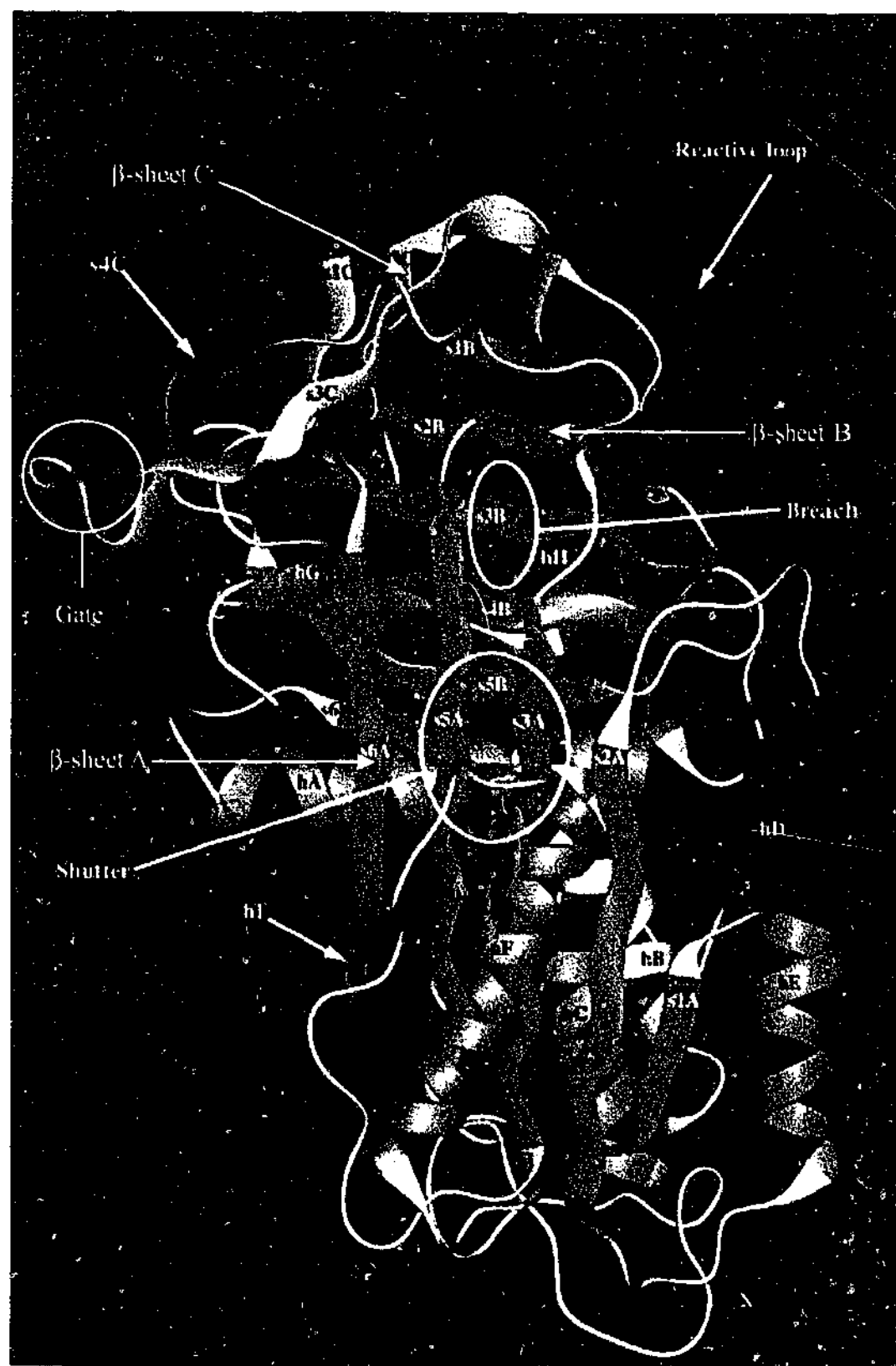


Figure 1.2: Structure of native α_1 -antitrypsin. The α -helices are labelled hA-hI and β -sheets A-C are indicated. The positions of the RCL, shutter, breach and gate are shown. Modified from Whisstock *et al* (2000).

residues from the β -sheet and increases the kinetics of interaction with factor Xa and thrombin (Jin *et al.* 1997; Baglin *et al.* 2002).

1.2.2. Cleaved serpins

α_1 -AT was the first serpin to have its structure determined in the cleaved conformation (Figure 1.3(A)). The most striking observation was the separation of the P_1 and P_1' residues by 67Å to opposite sides of the molecule (Loebermann *et al.* 1984; Mottonen *et al.* 1992). The cleaved structure showed that the N-terminal portion of the cleaved reactive centre loop, from 17 residues prior to the reactive centre (P_{17} - P_1), had inserted itself into the 5 stranded β -sheet A to form the middle (fourth) strand between strands 3 and 5. This resulted in a six-stranded antiparallel A- β -sheet. The observed differences between the native and cleaved structures suggested that protease cleavage of the RCL gave rise to a major conformational change in the serpin molecule. The incorporation of the cleaved RCL was made possible by an opening of the β -sheet based on a shift of the first three strands of the sheet (Steir, and Chothia 1991). This gave rise to further structural changes in the serpin molecule and allowed the β -sheet A to fully open and accept the inserting reactive centre loop.

The ability to incorporate the cleaved RCL into β -sheet A and the conformational change that ensues, is a characteristic feature of inhibitory serpins, and is supported by all reported crystal structures of cleaved inhibitory serpins, α_1 -ACT (Baumann *et al.* 1991), antithrombin (Mourey *et al.* 1993), and PAI-1 (Aertgeerts *et al.* 1995). However, this feature is not preserved in all non-inhibitory serpins. Elucidation of the cleaved structure of ovalbumin (plakalbumin) revealed that the RCL had not inserted into the A β -sheet (Wright *et al.* 1990). The absence of RCL insertion by some non-inhibitory serpins was further supported by the absence of an increase in thermostability upon loop cleavage, as observed with the non-inhibitory serpins angiotensinogen and PEDF (Stein *et al.* 1989; Becerra *et al.* 1995).

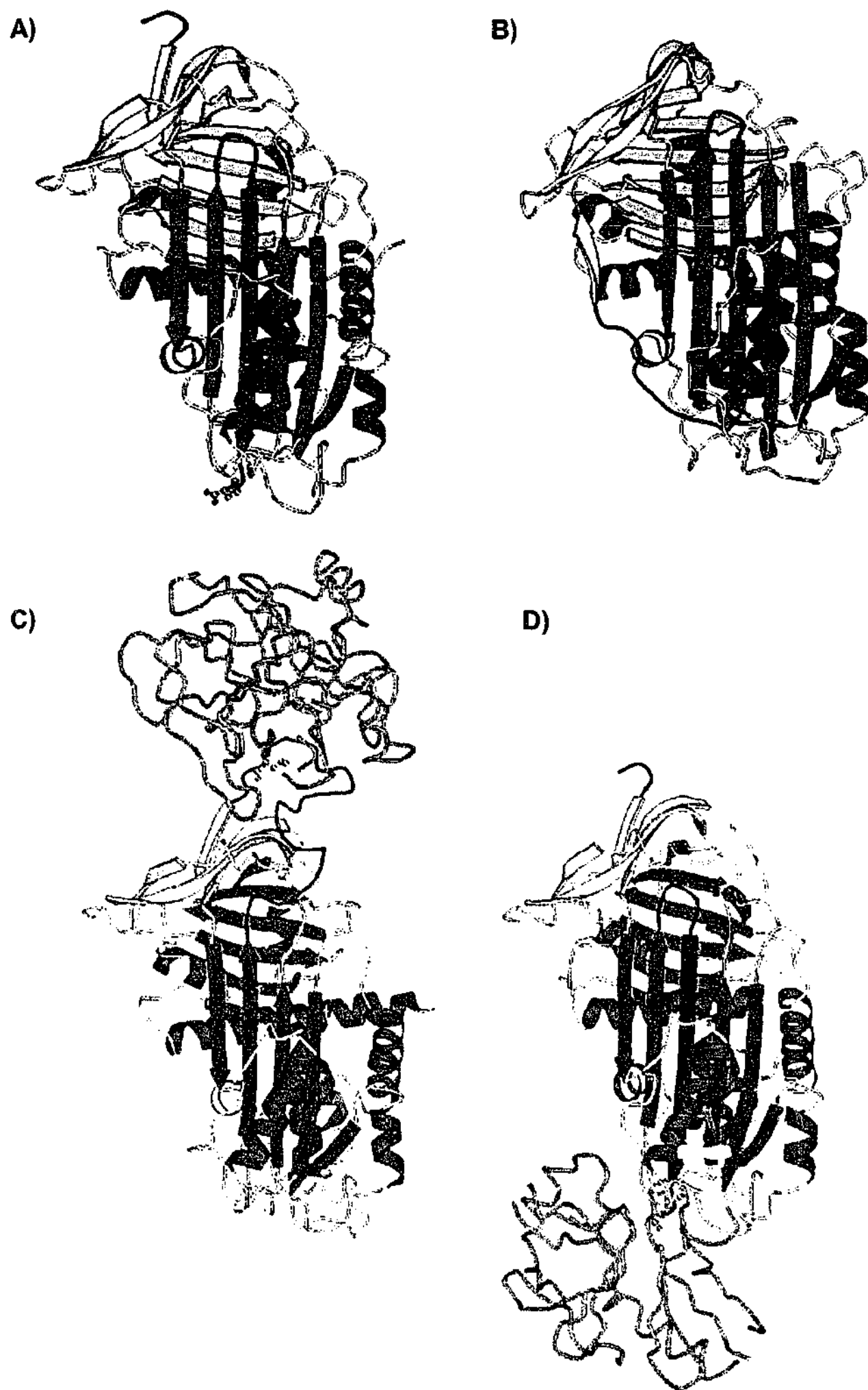


Figure 1.3: Serpin Conformations. A) Cleaved α_1 -antitrypsin, B) latent PAI-1, C) initial serpin-protease complex and D) final serpin-protease complex. Modified from Silverman *et al* (2000).

1.2.3. Latent serpin

An alternative native structure confined to a small number of serpins is the latent conformation. The latent state is analogous to the cleaved conformation and is similarly inactive and stable (Figure 1.3(B)). It was first observed in the crystal structure of PAI-1, in which the reactive centre loop from P₁₇-P₄ had inserted into β -sheet A in the absence of loop cleavage (Mottonen *et al.* 1992). Loop insertion was permitted by the release of the first strand of β -sheet-C, C-terminal to the reactive site residues, which provided a return for the RCL from the base of sheet-A back to sheet-B. Consequently, latent PAI-1 is unable to interact with target proteases as the reactive site residues are inaccessible. This conformation has also been described for the antithrombin mutant Rouen VI (187Asn→Asp) and results in the loss of inhibitory activity leading to thrombosis (Bruce *et al.* 1994).

PAI-1 is the only known wild-type serpin to spontaneously adopt the latent, inactive conformation under physiological conditions. The latent conformation can be induced in antithrombin and in α_1 -AT by mild denaturation at 60°C in the presence of citrate (Bruce *et al.* 1994; Lomas *et al.* 1995 (a)).

1.2.4. δ -conformation

A naturally occurring variant of α_1 -ACT (Leu55→Pro) was crystallised in which the RCL had partially inserted into β -sheet A, up to residue P₁₂ (Gooptu *et al.* 2000). The level of loop insertion observed was greater than the few hinge residues observed for antithrombin but less than the full loop insertion observed for latent PAI-1. The δ - α_1 -ACT crystal structure provided evidence for the presence of conformational intermediates between the formation of the initial Michaelis complex and the final irreversible loop inserted serpin.

1.2.5. Serpin – protease structure

The crystal structure of α_1 -AT in complex with trypsin confirmed that, upon full loop-insertion, the protease is translocated from one pole of the serpin molecule to the other (Huntington *et al.* 2000) (Figure 1.3 (C&D)). The structure further demonstrated that steric conflict between the two molecules, as the protease is forced over the serpin body on loop insertion, results in protease inhibition. The α_1 -AT molecule was unaffected, as its X-ray structure remained ordered, presumably due to the increased stability of the serpin molecule achieved upon loop insertion. However, the structure of the complexed protease was extensively disordered indicating that loop insertion caused destabilisation of the protease. The primary event in this destabilisation was the observed repositioning of the active serine residue (Ser195) more than 6Å away from its catalytic partner (His57) thus destroying the catalytic triad. The movement of the Ser residue also destroys the adjacent oxyanion hole which then prevents the continued catalytic deacylation of the ester bond with the serpin.

The extensive distortion of the catalytic site further destabilised interactions within the complexed protease. The X-ray structure showed a number of disordered regions in the complexed protease including the trypsin activation domain and the hinge region connecting the two large domains of the molecule. These observations explain why proteases in complex have an increased vulnerability to proteolysis (Kaslik *et al.* 1995; Stavridi *et al.* 1996). Destruction of the catalytic site therefore prevents the release of the protease from the complex and the structural disorder allows for its proteolytic destruction (Huntington *et al.* 2000).

1.2.6. Serpin polymerisation

One of the consequences of conformational flexibility of the serpin molecule is the formation of polymers. Native inhibitory serpins readily form polymers when heated to 50-60°C and their presence can be observed on non-denaturing gels

where a typical laddering pattern is observed due to the formation of non-covalent dimers, trimers and higher multimeric forms (Schulze *et al.* 1990; Patston *et al.* 1995). Serpin polymers can form following insertion of the RCL of one serpin molecule into the A β -sheet of another to form a continuous polymeric serpin strand. Alternatively, polymers can form by a loop-C-sheet mechanism where the insertion of a serpin's RCL into its own A-sheet liberates strand 1C which is then replaced by the RCL of a donor molecule. The C-sheet mechanism of polymerisation was highlighted in the crystal structure of an antithrombin native/latent dimer, where the RCL of the native molecule had inserted into the position of strand 1 in the C-sheet of the latent molecule (Carrell *et al.* 1994). In addition, serpin polymers have also been shown to form between the cleaved RCL of one serpin into the A β -sheet of another serpin molecule (Huntington *et al.* 1999; Dunstone *et al.* 2000).

Polymerisation occurs *in vivo* when a mutation affects the stability of the native serpin. This leads to the formation of polymeric aggregates within the cell of expression and as a consequence deficiency in plasma serpin activity. The α_1 -AT variant, Z-antitrypsin (Glu342 \rightarrow Lys), is one of the best characterised examples of polymerisation induced serpin deficiency and disease. The variant is present in 4% of Northern Europeans and perturbs the structure of the α_1 -AT molecule such that polymerisation proceeds through the insertion of an uncleaved RCL of one molecule into the A-sheet of the another (Sivasothy *et al.* 2000). The resultant phenotype is formation of inclusion bodies in liver and decreased levels of circulating α_1 -AT (Lomas *et al.* 1992). Deficiency of α_1 -AT activity results in uncontrolled proteolysis by the serine proteases elastase and cathepsin G and subsequent destruction of lung tissue leading to emphysema (Elliott *et al.* 1998 (b)). The hepatocyte polymeric inclusion bodies are associated with liver cirrhosis, neonatal hepatitis and hepatocellular carcinoma (Sveger 1976; Eriksson *et al.* 1986). Loop-sheet polymerisation of the α_1 -ACT variant (Pro229 \rightarrow Ala) is also associated with the onset of emphysema due to uncontrolled cathepsin G activity, via a similar mechanism to that of α_1 -AT deficiency (Faber *et al.* 1993; Poller *et al.* 1993).

Loop-sheet polymerisation has been described for variants of C1-inhibitor, antithrombin and neuroserpin associated with angioedema, thrombosis and dementia respectively (reviewed in (Stein and Carrell 1995) and (Lomas and Carrell 2002)). A number of point mutations in neuroserpin are associated with polymer formation and inclusion body accumulation in neurones. The Ser49→Pro or Ser52→Arg mutations in neuroserpin are located at the start of B-helix and destabilise the sheet-opening mechanism, leading to the formation of loop-sheet polymers and a loss of inhibitory activity towards tPA (Davis *et al.* 1999; Belorgey *et al.* 2002).

1.2.7. Regulation of conformational change

Regions important for regulating and modulating conformational change in serpins include: the reactive centre loop, particularly the hinge region (discussed in section 1.3.1), the breach, the shutter and the gate, and the F-helix, as indicated in Figure 1.2 (Stein and Carrell 1995; Whisstock *et al.* 2000). In addition, some serpins such as antithrombin and heparin-co-factor II require exosites to modulate their activity. Their importance in regulating serpin conformational mobility and activity is highlighted through the loss of functional activity caused by mutations in these regions (Stein and Carrell 1995).

The breach region is located at the top of the A- β sheet where initial loop insertion occurs. The pattern of amino acid conservation and the hydrogen bond network formed by these residues is likely to be critical for effective loop insertion (Whisstock *et al.* 2000).

The shutter region controls the opening of the A-sheet and is positioned at the centre of the serpin molecule towards the middle of the A-sheet (Whisstock *et al.* 2000). Mutations in the conserved residues of the shutter region renders the A-sheet more open and susceptible to incorporate a RCL, either its own or that of another serpin molecule. Subsequently, these mutations have an increased tendency to convert to the latent or polymeric conformation and usually result in serpin deficiency and loss of inhibitory activity, as discussed

previously for the α_1 -AT variant Z-antitrypsin (Section 1.2.6) (Stein and Carrell 1995).

The gate is located in the distal RCL hinge region and is a loop that is formed by the connection of strands s3C and s4c. Mutations in the gate region of antithrombin result in loss of inhibitory activity and the development of thrombosis (Stein and Carrell 1995).

Helix-F is positioned across the lower part of β -sheet A and its location makes it spatially impossible for the loop to insert past P9-P8 without movement of the helix (Gettins 2002 (b)). For the cleaved RCL to insert into the A-sheet helix F needs to be displaced from its position. It has been suggested that helix F moves from the front of the A-sheet to facilitate both loop-insertion and to produce the favourable serpin-protease end complex. This displacement is thermodynamically coupled to the energy of loop insertion and is required to make the final protease translocation and inhibition step an energetically favourable one.

1.2.8. Co-factor binding

The serpins antithrombin, heparin cofactor II, protease nexin 1, PAI-1 and protein C inhibitor (PCI) can bind sulphated polysaccharides, such as heparin, to increase their inhibitory activity towards proteases involved in the coagulation and fibrinolysis cascades ((reviewed by Huntington 2003) (Keijer *et al.* 1991; Cunningham *et al.* 1992; Rovelli *et al.* 1992)). Heparin binding to antithrombin causes a 1000-fold increase in activity towards thrombin and 300-fold activity towards factor Xa (Olson and Björk 1992; Olson *et al.* 1992). The sulphated polysaccharides line the walls of the vasculature and, in the case of antithrombin, binding also aids in localising the serpin to sites requiring inhibitory activity (Olson and Björk 1991).

There are two mechanisms of activation used by sulphated polysaccharides, (i) the bridging mechanism and (ii) a conformational change mechanism. In the

bridging mechanism the sulphated polysaccharide binds both serpin and protease to bring the two together in an appropriate orientation for interaction. This mechanism is used by all the aforementioned serpins. A cluster of positively charged residues located at the base of helix-A to the upper half of helix D, constitutes the heparin binding site for all these serpins (Huntington 2003), except PCI which utilises basic residues from helices A and H (Kuhn *et al.* 1990; Neese *et al.* 1998).

As discussed in Section 1.2.1, the RCL of heparin cofactor II and antithrombin are partially inserted into β -sheet A. The binding of sulphated polysaccharides to antithrombin and heparin cofactor II induces a conformational change and results in the expulsion of the partially inserted RCL from β -sheet A to an exposed position more suitable for interaction with the target proteases (Olson and Björk 1991; Baglin *et al.* 2002). In addition, the conformational change in heparin cofactor II is associated with displacement of the hirudin-like N-terminal extension which is sequestered in the native state. The N-terminal extension tethers to thrombin and confers increased specificity of heparin cofactor II towards the protease (Baglin *et al.* 2002). The conformational changes in antithrombin and heparin cofactor II dramatically contribute to an increase in the rate of inhibition, with the inhibition of factor Xa by antithrombin increased 300-fold (Olson *et al.* 1992) and the rate of heparin cofactor II inhibitory activity towards thrombin enhanced approximately 16,000-fold with either heparin or dermatan sulfate (Van Deerlin and Tollefsen 1991; Olson *et al.* 1992; Sheehan *et al.* 1994).

In addition to binding heparin, PAI-1 is also able to bind vitronectin. PAI-1 spontaneously adopts the inactive, latent conformation and binding of vitronectin stabilises the active conformation. This enables efficient protease interactions and may localise PAI-1 to specific tissues requiring inhibitory activity (Lawrence *et al.* 1994 (a); Lawrence *et al.* 1997).

As the name suggests, protein Z-dependent protease inhibitor (ZPI) binds protein Z which dramatically enhances the serpin's inhibitory activity towards

factor Xa (Tabatabai *et al.* 2001). Protein Z is a vitamin-K dependent plasma protein, similar in structure to factors VII, IX, X and protein C but without any properties of a serine protease zymogen (Sejima *et al.* 1990; Tabatabai *et al.* 2001). In the presence of phospholipid and calcium the ZPI-protein Z complex inhibits factor Xa at a rate 1000-fold faster than ZPI on its own.

1.3. REACTIVE CENTRE LOOP AND PROTEASE SPECIFICITY

The reactive centre loop (RCL) plays a pivotal role in serpin and protease interactions. The reactive site, designated P_1 - P_1' , comprises the bait peptide bond for protease interaction. The residues N-terminal to the scissile bond are designated the P-residues (P_{16} - P_1) and the residues on the C-terminal end are designated the P'-residues (P_1' - P_{10}') (as defined by (Schechter and Berger 1967)).

1.3.1. Importance of reactive centre loop composition

The amino acid composition of the RCL not only determines target protease specificity for inhibitory serpins but is also critical for the mechanism of inhibition. The residues occupying P_{15} - P_9 , the proximal hinge region, determine the ability of the loop to insert upon cleavage (Hopkins *et al.* 1993; Potempa *et al.* 1994). Sequence comparisons between inhibitory and non-inhibitory serpins indicate a conservation of short-chain residues, mainly alanine, between residues P_{12} - P_9 for inhibitory serpins. Non-inhibitory serpins have a high substitution rate to other residues, with few or no alanines (Hopkins *et al.* 1993).

The pattern of conservation in the proximal hinge is thought to facilitate the rapid insertion of the RCL once the P_{14} side chain has incorporated into the A-sheet (Hopkins *et al.* 1993; Irving *et al.* 2000). The conservation of P_{14} is critical as this is the first side-chain to insert into the A-sheet and stabilises the inserted loop conformation. For this reason P_{14} is typically a non-charged, short-chain residue, either threonine or serine. Site-directed mutagenesis of

the non-inhibitory serpin ovalbumin at P₁₄ Arg→Ser resulted in spontaneous insertion of the RCL on cleavage (Huntington *et al.* 1997). The P₁₅ residue is almost always occupied by glycine in inhibitory serpins and P₁₃ is usually glutamic acid (Hopkins *et al.* 1993). The proximal hinge has a requirement of non-proline residues at P₁₂ and P₁₀. This is supported by the P₁₀ antithrombin mutation, Ala (384) →Pro, which is associated with a loss of inhibitory activity and substrate-like behaviour with thrombin (Perry *et al.* 1989). Angiotensinogen also contains a Pro residue at P₁₂ and an Arg at P₁₄ and is unable to undergo the S→R transition presumably due to the inability to loop insert. P₈ also shows a preference for the small threonine side chain.

The residues from P₈-P₅' modulate target protease specificity as demonstrated by the absence of residue conservation between serpins and the ability of mutations to alter inhibitory activity (Djje *et al.* 1996; ChaillanHuntington *et al.* 1997; Djje *et al.* 1997). Natural variants and site-directed mutagenesis studies have shown P₁ as the most important determinant of protease specificity ((Travis *et al.* 1985; Jallat *et al.* 1986; Derechin *et al.* 1990; Stein and Carrell 1995)). The α_1 -AT Pittsburgh variant has a Met→Arg mutation at P₁ and alters the serpin specificity from elastase to thrombin (Owen *et al.* 1983). A number of antithrombin mutants at P₁ (Arg→His, Cys or Pro) result in complete loss of inhibitory activity towards factor Xa. C1-inhibitor P₁ mutants (Arg→His, Cys, Ser, Leu) are associated with angioedema due to the loss of inhibitory activity towards the serine protease protein C (reviewed in (Stein and Carrell 1995)).

The ability to change protease specificity by swapping the RCL from one serpin to another also illustrates the importance of RCL composition. Rubin *et al.* (1990) demonstrated that swapping the loop of α_1 -ACT for the RCL residues P₃-P₃' of α_1 -AT produced a modest inhibitor of neutrophil elastase (Rubin *et al.* 1990). Both α_1 -ACT and α_1 -AT were converted into chimeras with the viral serpin RCL. Although loop swapping generated changes in protease specificity, α_1 -ACT/serpin and α_1 -AT/serpin varied in their ability to inhibit the proteases examined (Bottomley and Stone 1998). This indicated that the serpin scaffold was also an important factor in determining the effectiveness of inhibition.

1.3.2. Importance of reactive centre loop length

The length of the RCL can determine the effectiveness of target protease inhibition. The RCL length shows very little variation in inhibitory serpins, usually consisting of 17 residues, with antipiasmin, C1-inhibitor and crmA being 16 residues (Simonovic *et al.* 2000; Gettins 2002 (a)). Variation on loop length, either by the addition of extra residues or removal of residues N-terminal to the P₁-P₁' bond, can result in loss of inhibitory activity or a decrease in the kinetics of interaction. RCL length is critical for generating the correct tension to remove the active serine residue within the catalytic triad of the protease (Zhou *et al.* 2001). The natural variant, α_2 -antipiasmin Enschede, has an alanine insertion at P₈ taking the RCL from the normal length of 16 residues to 17, and results in weak inhibitory activity towards plasmin (Holmes *et al.* 1987; Rijken *et al.* 1988). The increased RCL length decreases the force exerted on the protease active site residue and a loss of protease destabilisation and inhibition (Zhou *et al.* 2001).

1.3.3. Protease specificity

Proteases catalyse the hydrolytic cleavage of peptide bonds between amino acids. Different proteases utilise different strategies to generate a nucleophilic attack on the targeted peptide bond and this mechanism is used to classify them (Chapman *et al.* 1997). Important subgroups include the serine, cysteine, aspartic and metallo- proteases. The cleavage preference of a protease can often be used to predict the inhibitory profile of a serpin (Table 1.1).

Serine proteases are named after the nucleophilic Ser residue present at the active site which mediates binding to substrates during catalysis (Hedstrom 2002). They can be sub-divided into groups based on their substrate specificity (Table 1.1). Chymotrypsin-like proteases cleave after non-polar and aromatic amino acid residues such as Leu or Phe. The trypsin-like proteases have a preference for charged, basic residues such as Arg and Lys, while elastase has a preference for small non-polar residues such as Ala and Val and can

Table 1.1: Serine and cysteine proteases: Specificity and their serpin inhibitors

	PROTEASE	PREFERRED CLEAVAGE RESIDUE	INHIBITORY SERPIN	P ₁ RESIDUE
SERINE PROTEASES	Chymotrypsin	Leu, Phe	α_1 -ACT, SCCA-2 MNEI α_1 -AT, PI-6	Leu Phe Met
	Chymase	Leu, Phe	α_1 -ACT MNEI	Leu Phe
	Cathepsin G	Leu, Phe	α_1 -ACT MNEI PI-6 α_1 -AT	Leu Phe Arg Met
	Trypsin	Arg, Lys	PI-6, PI-8, yukopin α_1 -AT,	Arg Met
	Thrombin	Arg	Antithrombin, PI-6, PCI, protease nexin 1	Arg
	Factor Xa	Arg	Antithrombin, protease nexin 1	Arg
	Granzyme B	Asp	crmA PI-9	Asp Cys
	Elastase	Ala, Val	α_1 -AT MNEI	Met Cys
CYSTEINE PROTEASES	Cathepsin L	Leu, Ile (P ₂)	Hurpin, MENT SQN-5 SCCA-1	Thr Glu
	Cathepsin V	Leu, Ile (P ₂)	Hurpin, MENT, SQN-5 SCCA-1	Thr Glu
	Caspases	Asp	crmA, PI-9	Asp

P₁ residues for inhibitory serpins were adapted from Gettins (2003 (a)) and additional information derived from Zhou *et al.* 1997 (crmA), Welss *et al.* 2003 (hurpin), Irving *et al.* 2002 (MENT), and Al-Khunaizi *et al.* 2002 (SQN-5). Information on protease specificity was derived from Perona and Craik 1995 (serine proteases) and Chapman *et al.* 1997 (cysteine proteases).

cleave after polar residues Met and Cys (Perona and Craik 1995; Hedstrom 2002). The serine protease granzyme B cleaves after the acidic residue Asp.

All members of the cysteine protease family utilise the sulphhydryl moiety of the Cys residue at the active site for catalysis and they primarily reside within the intracellular compartment. The cysteine proteases can be grouped into two families, the caspase and the papain-like family (Chapman *et al.* 1997). The caspases cleave their substrates after Asp residues and participate in one of two distinct pathways, the activation of proinflammatory cytokines and promotion of apoptotic cell death (reviewed (Salvesen and Dixit 1997)). The papain-like subfamily is made up of the lysosomal cathepsins and the calcium dependent proteases, the calpains. The calpains are thought to play a role in cellular signalling through the proteolysis of intracellular proteins which regulate intracellular calcium concentration (Croall and DeMartino 1991; Du *et al.* 1995).

Substrate specificity of cathepsins is primarily determined by the S2 subsite rather than the S1 cleavage site as seen in the serine proteases (Chapman *et al.* 1997; Turk *et al.* 2001). There are 11 different human cathepsins which can be divided into two groups based on their amino acid preferences. Cathepsin B-like proteases cleave proteins after basic amino acids Arg-Arg or Arg-Lys while cathepsin L-like enzymes have a strong preference for hydrophobic or branched chain amino acids, such as Ile or Lys in the S2 subsite. The lysosomal cathepsins are associated with extracellular matrix degradation during bone remodelling, MHC II antigen presentation and prohormone processing (Chapman *et al.* 1997; Turk *et al.* 2001).

1.3.4. Cross-class inhibition

A number of serpins have been reported to possess inhibitory activity towards proteases outside of the serine protease family. These include proteases from the caspase and papain cysteine protease subfamilies and subtilisins. The intracellular serpin PI-8 is an efficient inhibitor of the endoproteinase furin and

serpin PI-9 is a potent inhibitor of subtilisin A (Dahlen *et al.* 1997; Dahlen *et al.* 1998 (a)). The plasma serpin α_1 -AT inhibits both subtilisin *Carlsberg* and proteinase K. α_1 -ACT has also been shown to display inhibitory activity towards proteinase K (Komiyama *et al.* 1996).

The viral serpin crmA inhibits caspase1 (interleukin-1 β converting enzyme), caspase3 and caspase 8, whilst PI-9 is a weak inhibitor of caspase1 (Annand *et al.* 1999; Young *et al.* 2000). Human SCCA1 and chicken MENT are efficient but relatively non-selective inhibitors of several lysosomal cathepsins namely, cathepsin L, V and K (Schick *et al.* 1998; Irving *et al.* 2002 (b)). However, the human serpin hurpin/headpin is a selective and efficient inhibitor of cathepsin L, with limited activity towards cathepsin V and cathepsin K (Welss *et al.* 2003).

The ability of the serpin scaffold to support cysteine protease inhibition was shown by substituting the RCL of SCCA-1, a known cysteine protease inhibitor, into the serpin scaffold of α_1 -AT (Irving *et al.* 2002 (c)). Inhibitory activity towards the cysteine proteases, cathepsin L, V and K, could be induced in the chimeras. Inhibitory activity of the α_1 -AT/SCCA-1 chimera was abolished by pre-incubating the chimera with an antithrombin RCL peptide which inserted into the A β -sheet. This suggested that loop insertion was required for cysteine protease inhibition and could occur via a mechanism similar to that of the serine proteases (Irving *et al.* 2002 (c)).

1.4. PHYLOGENY AND EVOLUTION OF THE SERPIN SUPERFAMILY

The evolutionary origin of the ancestral gene to the serpin superfamily is unclear. The presence of serpins in all eukaryotes suggests that serpins arose before animal and plant divergence (~2.7 billion years ago). The recent isolation of serpins in prokaryotes suggests they may have evolved even before eukaryote development (>3.9 billion years ago). However, with the sporadic distribution of serpins in prokaryotes and yeast, it is possible that the serpins are in fact a younger gene family and their presence in prokaryotes

has occurred through horizontal transfer from eukaryotes (Irving *et al.* 2002 (a)).

Many of the serpin gene families have conserved intron-exon structures, indicating that they share a phylogenetic history. A number of studies over the years have highlighted that many serpins are located in clusters (Remold-O'Donnell 1993; Rollini and Fournier 1997; Scott *et al.* 1999 (a)). In the human genome serpin clusters are located at chromosomes 1q, 6p25, 11q, 14q32.1, 17p13, 18q21.3-23. The genes within these clusters often share the same intron-exon pattern although the intron-exon pattern varies between serpin clusters on different chromosomes. The exception is the human intracellular serpins which are clustered at 6p25 and 18q21.3-23, as the genes located at 6p25 have one less exon (Scott *et al.* 1999 (a)).

Ragg *et al.* (2001) assigned the vertebrate serpins to six gene families based on their exon-intron patterns, diagnostic amino acid sites, and rare indels. This clade assignment was subsequently modified by tree reconstruction analyses, integrating amino acid sequences with intron-exon structure and diagnostic amino acids. Viral serpins were also included in this analysis (Atchley *et al.* 2001). These two alignments were analogous, however differed from the clade organisation proposed by Irving *et al.* (2000), where serpins were assigned to clades by consensus analysis based on patterns of sequence conservation. The genomic structure of the serpins was not included in this analysis. The phylogenetic trees of Ragg *et al.* (2001) and Irving *et al.* (2000) are similar, with exceptions in the evolutionary classification of the human serpins PAI-1, protease nexin 1 and neuroserpin and assignment of the viral serpins (Irving *et al.* 2000; Ragg *et al.* 2001).

Based on the clade structure of Irving *et al.* (2000), a nomenclature for the entire serpin family was proposed which organised the serpin superfamily into sixteen phylogenetic clades (Irving *et al.* 2000; Silverman *et al.* 2001). Currently the clades are designated A to P, with individual members within each clade numbered from one onwards (Silverman *et al.* 2001) (Table 1.2). The majority of the human serpins are dispersed between clades A-I.

Table 1.2: The serpin clades (Modified from (Irving *et al.* 2000; Silverman *et al.* 2001).

CLADE NAME	CLADE LETTER
α 1-proteinase inhibitor	A
Intracellular, ov-serpin	B
Antithrombin	C
Heparin cofactor II	D
Proteinase nexin, PAI-1	E
α 2-antiplasmin, PEDF	F
C1 inhibitor	G
HSP47	H
Neuroserpin	I
Horseshoe crab	J
Insect	K
Nematode	L
Blood fluke	M
Viral SPI1-2/CrmA	N
Viral SPI-3-like	O
Plant	P
Unclassified (orphans)	-

1.4.1. The A clade serpins

The A clade contains the archetypal human serpins α ₁-AT(SERPINA1) and α ₁-ACT(SERPINA3), and also three non-inhibitory serpins, the hormone binding serpins corticosteroid binding globulin (CBG ;SERPINA6), thyroxine binding globulin (TBG;SERPINA7) and angiotensinogen (AGT; SERPINA8). The majority of the serpins are clustered together on human chromosome 14q32.1 (Forsyth *et al.* 2003) (Figure 1.4) with the exception of TBG and AGT which are located on chromosomes Xq22.2 (Mori *et al.* 1995) and 1q42-q43

(Ilsa *et al.* 1990), respectively. Many of the A-clade serpins are expressed in liver (Marsden and Fournier 2003), except for centerin (SERPINA10) and some also have secondary sites of expression (Krebs *et al.* 1999; Frazer *et al.* 2000; Uhrin *et al.* 2000). All A-clade serpins appear to have N-terminal secretion peptides and are predictably secreted into extracellular compartments (Table 1.3).

1.4.1.a. α_1 -Antitrypsin (SERPIN A1)

The most abundant serpin in human plasma (1.5 mg/ml) α_1 -AT, is a 50-55kDa glycoprotein which is synthesised primarily in liver (Travis and Salvesen 1983). Originally identified as the major trypsin inhibitor in plasma, it was given the name α_1 -AT. It is also known as α_1 -proteinase inhibitor because of its broad spectrum of inhibitory activities towards chymotrypsin-like and trypsin-like proteases together with its primary target, leukocyte elastase (Beatty *et al.* 1980). The pairing of Met-Ser at P₁-P₁' allows efficient inhibition of elastase and α_1 -AT activity can be regulated through the oxidation of methionine residues both at the reactive site and at other positions throughout the serpin molecule (Johnson and Travis 1978). The oxidation of α_1 -AT leads to the loss of inhibitory activity and subsequent proteolysis of the serpin (Taggart *et al.* 2000). Deficiencies in α_1 -AT are associated with emphysema and liver cirrhosis, as already outlined in section 1.2.6.

Lower levels of α_1 -AT are also expressed in extra-hepatic compartments suggesting a role in regulating protease activity in these microenvironments. These compartments include haemopoietic cells, as well as gastric and breast carcinoma cells (Geboes *et al.* 1982; Kroll and Briand 1988; Missen *et al.* 2004). The exact function of α_1 -AT in these cell types is yet to be fully understood, however, expression in monocytes is implicated in their activation which may contribute to the formation of atherosclerotic plaques (Janciauskiene and Lindgren 1999; Dichtl *et al.* 2000). Expression in the bone marrow compartment is likely to be linked to the mobilisation of haemopoietic

Table 1.3: Human clade A serpins. Adapted from Gettins (2003 (a))

GENE	COMMON NAME	CHROMOSOME	RCL	TARGET PROTEINASES	FUNCTION
SERPINA1	α_1 -proteinase inhibitor (α_1 -PI), α_1 -antitrypsin (α_1 -AT)	14q32.1	GTEAAGAMFLEAIPM-SIPPE	Neutrophil elastase	Inflammatory response
SERPINA2	α_1 -antitrypsin related protein, proteinase inhibitor like	14q32.1	GTEATGAPHLEEKAW-SKYQT	-	Uncharacterised-truncated gene possible pseudogene
SERPINA3	α_1 -antichymotrypsin, α_1 -ACT	14q32.1	GTEASAATAVKITLL-SALVE	CathepsinG, chymase, protease specific antigen, chymotrypsin	Inflammatory response, binds β -amyloid peptide-implicated in Alzheimer's disease. Diagnostic marker in complex with PSA
SERPINA4	Kallistatin, kallikrein inhibitor, PI4	14q32.1	GTEAAAATTFAIKFF-SAQTN	Kallikrein	<i>In vivo</i> function not established
SERPINA5	Protein C inhibitor (PCI), plasminogen activator inhibitor (PAI-3)	14q32.1	GTRAAAATGTIFTFR-SARLN	Activated protein c, uPA,tPA, acrosin	Possible role in male reproduction
SERPINA6	Corticosteriod-binding globulin (CBG)	14q32.1	GVDTAGSTGVTLNLT-SKP II	Non-inhibitory	Transport of corticosteriod
SERPINA7	Thyroxine-binding globulin (TBG)	Xq22.2	GTEAAAVPEVELSDQ-PENTF	Non-inhibitory	Transport of thyroxine
SERPINA8	Angiotensinogen (AGT)	1q42-q43	EREPTSTQQLNKPE-VLEV	Non-inhibitory	Peptide fragment increases blood pressure
SERPINA9	Centerin	14q32.1	GTEATAATTTKFIVR-SKDPG	Uncharacterised	Maturation of B-cells
SERPINA10	Protein-Z dependent protease inhibitor (PZI)	14q32.1	GTEAVAGILSEITAY-SMPPV	Factor Xa and XIa	Binds proteinZ for factor Xa inhibition
SERPINA11	-	14q32.1	GTEAGAASGLLSQPP-SLNTM	Uncharacterised	Uncharacterised
SERPINA12	Vaspin	14q32.1	GTEGAAGTGAQTLPM-EPLVV	Uncharacterised	Uncharacterised

stem cells from the bone marrow to peripheral blood, as discussed in section 1.6.4 (Winkler *et al.* 2004).

1.4.1.b. α_1 -Antichymotrypsin (SERPINA3)

α_1 -ACT is plasma glycoprotein of ~66kDa and is primarily synthesised in the liver (Travis and Salvesen 1983) with lower levels of expression seen in bronchial epithelia, prostate and breast (Cichy *et al.* 1995). It is an acute phase protein whose plasma levels rise 2-fold within 8-16 hr following an inflammatory stimulus (Aronsen *et al.* 1972). Its primary function is the regulation of the chymotrypsin-like granular proteinases cathepsin G and mast cell chymase released by leukocytes at sites of inflammation (Travis and Salvesen 1983). α_1 -ACT also regulates neutrophil chemotaxis and superoxide production (Schuster *et al.* 1992; Lomas *et al.* 1995 (b)). Two mutations in α_1 -ACT, Leu55→Pro and Pro229→Ala alter the conformational stability of α_1 -ACT (section 1.2.6) and render the serpin inactive. Both mutations are associated with serum deficiency of α_1 -ACT and contribute to the development of chronic obstructive pulmonary disease (Poller *et al.* 1993). In addition, α_1 -ACT and the serine protease, prostate specific antigen, circulate as a complex in plasma and this complex is used as a diagnostic marker for prostate cancer (Mikolajczyk *et al.* 2002; Stephan *et al.* 2002).

A number of reports have suggested that α_1 -ACT may have biological functions in addition to protease inhibition. α_1 -ACT expression is higher in breast cancer tissue than normal breast tissue, where it is also up regulated by steroid hormones (Laursen and Lykkesfeldt 1992; Confort *et al.* 1995). In a study investigating the differential gene expression between cells expressing the Raf-1 oncogene, α_1 -ACT was identified as a gene whose expression increases when Raf-1 expression was down-regulated (Patel *et al.* 1997). This suggested that α_1 -ACT is associated with the Raf-1 signalling pathway and may modulate cell growth and proliferation.

A number of inflammatory cytokines have been reported to up-regulate α_1 -ACT expression in sites other than the liver (Cichy *et al.* 1995; Kordula *et al.* 1998; Nilsson *et al.* 2001). α_1 -ACT is known to be expressed by astrocytes and may play a permissive role in amyloid peptide polymerisation. *In vitro* evidence demonstrates that α_1 -ACT binds the amyloid β_{1-42} peptide by insertion into the serpin A β -sheet (Ma *et al.* 1994). The subsequent induction of fibrillogenesis and generation of amyloid plaque deposits is associated with the progression of Alzheimer's disease (Ma *et al.* 1994; Abraham 2001). Over expression of human α_1 -ACT in a murine model of Alzheimer's disease caused an increased rate of disease progression (Mucke *et al.* 2000).

Nuclear localisation of α_1 -ACT has been observed in cells and it has been suggested that α_1 -ACT may bind DNA through an unusual DNA binding motif of three lysine residues, located at residues 212-214 and a C-terminal peptide at residues 390-398 (Takada *et al.* 1986; Naidoo *et al.* 1995). This feature may reflect its reported inhibition of DNA primase and DNA polymerase (Tsuda *et al.* 1986; Takada *et al.* 1988). Although α_1 -ACT does not contain a typical nuclear localisation signal, it was recently shown to interact with the DnaJ-like protein 1 which does possess a putative nuclear localisation signal (Kroczyńska *et al.* 2003). The DnaJ-like protein activates the molecular chaperone BiP by its membrane associated J-domain. Binding to α_1 -ACT is mediated through the SANT2 motif. The exact role of the SANT2- α_1 -ACT complex is unclear, however, the complex may be involved in gene transcription and regulation. SANT motifs are important for acetylation of histone N-terminal tails and are suggested to play a role in delivering histone substrates to chromatin-modifying enzymes (Yu *et al.* 2003).

1.4.2. The B clade serpins

Members of the B clade constitute the subgroup of serpins known as the ovalbumin family. Based on the common characteristics of five serpins, ovalbumin, MNEI, PAI-2, SCCA and gene Y, Remold-O'Donnell *et al.* (1993)

named the "Ov-serpin family" after the oldest known member, ovalbumin. Some of their identifiable features are the absence of N-terminal and C-terminal extensions, (in particular the absence of an N-terminal secretion peptide), an almost identical gene organisation and the presence of oxidisable residues in or adjacent to the reactive centre residues P₁-P₁' (Remold-O'Donnell 1993).

In comparison to the archetypal serpin α_1 -AT, ov-serpins begin and end at the first and last buried residues at positions 23 and 391 respectively (Remold-O'Donnell 1993). The absence of an N-terminal cleavable signal sequence renders most ov-serpins intracellular, however, some ov-serpins may be secreted. For example, PAI-2 (SERPINB2) is both a cytoplasmic non-glycosylated protein and a glycosylated protein secreted from stimulated monocytes and placenta (Belin *et al.* 1989). It is believed that ov-serpins utilise non-cleavable internal signal sequences for secretion. A hydrophobic region at the N-terminus of both PAI-2 and ovalbumin has been linked to secretion (Ye *et al.* 1988; von Heijne *et al.* 1991).

The ov-serpin family currently consists of 13 members which share a higher amino acid identity to each other, than with other members of the serpin superfamily (Remold-O'Donnell 1993). The serpins are clustered on two chromosomes, 6p25 and 18q21.3 and share a conserved intron-exon structure, except that the genes located on chromosome 6p25 have one less exon. Their evolution from the ancestral "ov-serpin" gene is thought to have occurred by one or two interchromosomal duplications and several intrachromosomal duplications (Scott *et al.* 1999 (a)).

The B clade serpins are involved in a diverse array of functions which they mediate through the regulation of intracellular proteases as well as non-inhibitory roles. These functions include: the regulation of inflammatory response via the inhibition of neutrophil proteases, cathepsin G and elastase, by PI-6 and MNEI (Scott *et al.* 1999 (a); Cooley *et al.* 2001); involvement in tumourgenesis by SCCA-I, SCCA-2 and hurpin (Abts *et al.* 1999; Welss *et al.*

2003) and cell differentiation/apoptosis by PAI-2 and PI-9 (Dickinson *et al.* 1995; Bird *et al.* 1998; Darnell *et al.* 2003). The non-inhibitory B clade serpin, maspin (SERPINB5), is associated with breast cancer, is an inhibitor of metastasis and functions as a tumour suppressor (Zou *et al.* 1994). Maspin has also been reported to possess potent anti-angiogenic activity in the absence of serpin inhibitory activity (Zhang *et al.* 2000).

1.4.2.a. The C-D interhelical loop

A common feature of the intracellular serpins is the presence of an interhelical loop between helices C and D (Remold-O'Donnell 1993). The presence of this C-D loop has been described for B-clade members encoded on human chromosome 18q21.3, including the serpins PAI-2, bomapin and hurpin. The CD loop varies in length between members and is located at the opposite side of the serpin molecule from the RCL, suggesting that other proteins can bind to the serpin via the loop without affecting inhibitory activity. Its presence is associated with auxiliary functions to serpin inhibitory activity and may provide a means of localising serpin activity to specific intracellular locations (Jensen *et al.* 1994).

The human serpins PAI-2 and bomapin have 33-residue and 25-residue C-D loop insertions, respectively. This insertion in PAI-2 is linked with cell survival, and is thought to protect cells from TNF α induced apoptosis (Dickinson *et al.* 1998). Resistance to apoptosis is conferred by the presence of a retinoblastoma protein (Rb) binding domain within the C-D interhelical region (Darnell *et al.* 2003). Rb is a transcriptional regulator involved in cellular activities including cell cycle control, apoptosis, differentiation and tumour suppression. The degradation of Rb is prevented by binding to PAI-2 which results in the retention of Rb activity.

The C-D interhelical residues in bomapin are important for nuclear localisation. The insertion consists of a series of basic amino acid residues (Lys-Lys-Arg-Lys) which exhibit homology to the known nuclear targeting signal for

SV40 antigen. Substitution of these residues in bomapin completely abrogates nuclear localisation (Chuang and Schleef 1999). The chicken intracellular serpin MENT also possesses a 30 amino acid interhelical C-D loop. MENT is the predominant non-histone protein to associate with heterochromatin DNA in avian haemopoietic cells (Grigoryev and Woodcock 1998). The interhelical insertion contains two critical functional motifs, a nuclear localisation signal required for nuclear import and an AT-hook motif which is associated with chromatin and DNA binding (Grigoryev *et al.* 1999).

1.4.3. Clades C-P

The remaining human clades (clades C-I) consist of small groups containing 1-3 members and include both inhibitory and non-inhibitory serpins. They have a diverse array of functional activities including regulation of coagulation, by antithrombin (SERPINC1), heparin cofactor II (SERPIND1) (Huntington 2003) and α_2 -antiplasmin (SERPINF2) (Lijnen and Collen 1985) and protein chaperoning by HSP 47 (SERPINH1) (Dafforn *et al.* 2001). The non-inhibitory serpin PEDF (SERPINF2) has been shown to be anti-angiogenic and neurotrophic (Becerra *et al.* 1995), and there is also the neuronally expressed neuroserpin (SERPINI1) (Hastings *et al.* 1997). Several additional phylogenetic clades (clades J-P) have been described in a variety of animal, plant and viral species.

1.5. THE MURINE MULTIGENE SERPIN FAMILIES

The serpin clusters on human chromosomes appear to be conserved in the mouse. Syntenic regions have been reported for both A-clade and B-clade serpins (Borriello and Krauter 1991; Sun *et al.* 1997; Kaiserman *et al.* 2002; Forsyth *et al.* 2003). However, a number of the murine loci have expanded to multigene clusters encoding more than the one gene present in humans. Multigene families have been described for the murine equivalents of human A-clade serpins α_1 -AT and α_1 -ACT, and for B-clade members PI-6, PI-9 and

MNEI (Hill *et al.* 1985; Borriello and Krauter 1991; Sun *et al.* 1995 (a); Sun *et al.* 1997; Kaiserman *et al.* 2002). Whilst murine homologues have been identified for serpins from clades C-I, the murine serpin expansion has only been reported for clades A and B. This thesis will focus on the expansion of the *serpina3* family.

1.5.1. Comparison of the serpin clade A between human and mice

The clade A genes located on human chromosome 14q.32.1 have a syntenic region on chromosome 12F1 in the mouse (Forsyth *et al.* 2003). The organisation of clade A between human and mouse is largely conserved, as depicted in Figure 1.4. All human *SERPINA3* genes are present at syntenic mouse loci except the mouse homologue of the human serpin gene, *SERPINA4*, the kallikrein inhibitor, which appears to be a pseudogene due to the presence of stop codons (Forsyth *et al.* 2003). The two murine loci encoding the homologous genes to human *SERPINA1* (α_1 -AT) and *SERPINA3* (α_1 -ACT) are the only two mouse a-clade genes to have undergone multiplication. All other A clade members are represented by a single murine homologue (Forsyth *et al.* 2003).

1.5.2. *Serpina3* family of serpins

The first murine orthologue of human α_1 -ACT was isolated from mouse plasma in 1982 by Takahara and Sinohara. However, this serpin displayed greater inhibitory activity towards trypsin than chymotrypsin and was named contrapsin (*contra-trypsin*) (Takahara and Sinohara 1982). It was shown to possess inhibitory activity towards mouse plasmin and trypsin-like proteases from the mouse submaxillary gland (Takahara and Sinohara 1983). Sequence analysis revealed that the closest human orthologue to contrapsin was α_1 -ACT, with 70% nucleotide and 60% amino acid identity, although the RCL had significantly diverged (Hill *et al.* 1984). Contrapsin was subsequently shown to be a member of a multigene family termed the *Spi-2* locus by Inglis and Hill in

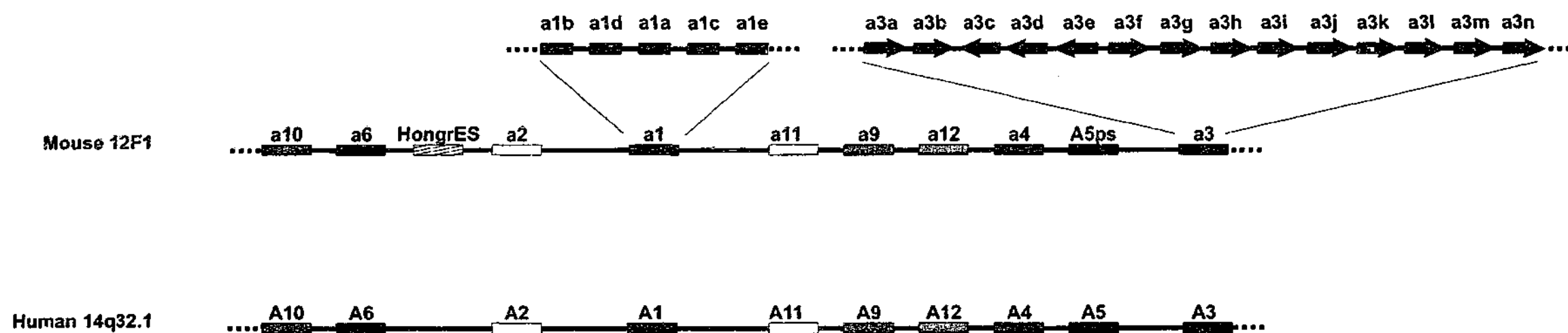


Figure 1.4: Gene organisation of the human A-clade serpins on chromosome 14q32.1 in comparison to the syntenic a-clade region on mouse chromosome 12F1. The human and mouse serpins are labelled according to the serpin nomenclature; A1: α_1 -AT, A2: α_1 -AT pseudogene, A3: α_1 -ACT, A4: kallistatin, A5: protein C inhibitor, A6: corticosteriod-binding globulin, A9: centerin, A10: protein Z-dependent inhibitor, A11 (un-named) and A12: vaspin. The murine homologues of α_1 -AT (a1) are shown in red. The murine homologues of α_1 -ACT (a3) are shown in blue and the arrows indicate the orientation of the genes.

1991. They demonstrated the existence of a locus of 10 closely linked members of a proteinase inhibitor multigene family on chromosome 12 (Inglis and Hill 1991). A further four members have since been identified (Forsyth *et al.* 2003). All orthologues have greatest sequence homology to α_1 -ACT (Table 1.4) and are believed to be derived from a single common ancestor. The murine *Spi-2* locus has since been renamed, in keeping with current serpin nomenclature, as the *serpina3* gene locus (Forsyth *et al.* 2003).

Comparative nucleotide and predicted protein sequence analysis of the mouse *serpina3* members revealed a high level of homology to human α_1 -ACT, particularly in sequences encoding the serpin scaffold (Inglis and Hill 1991). It appears however, that accelerated divergence has occurred in the predicted RCL region, with a high number of amino acid substitutions in this area (Table 1.4). The cDNA of the *serpina3* genes encode full open reading frames indicating the potential to give rise to functional proteins (Inglis and Hill 1991; Forsyth *et al.* 2003). The evolutionary pressures driving the expansion are unknown. Nevertheless, the murine orthologues of α_1 -ACT have evolved to preserve the serpin scaffold with divergence restricted to the RCL. This suggests they may have an array of target proteinase specificity.

The RCL of two mouse *serpina3* members resemble the human α_1 -ACT with a P_1 and P_1' pairing of Leu and Ser (Table 1.4). Another two *serpina3* members have predicted anti-tryptic behaviour with a P_1 residue of Arg. Interestingly, four of the mouse *serpina3* members have a Cys residue at the predicted P_1' position and furthermore, three members also have a cysteine residue at the P_1 position. Cysteine residues at the P_1 - P_1' positions have not previously been described in naturally occurring serpins therefore a target protease is difficult to predict. Recent work with the mouse member serpin2A (*serpina3g*) suggests cysteine proteases as targets (Liu *et al.* 2003). The unusual motifs in the RCL indicates the potential for novel target specificities (Inglis and Hill 1991; Forsyth *et al.* 2003). Another interesting feature of the mouse *serpina3* locus is the lack of predicted secretion signals in some of the genes indicating the possibility of both intracellular and extracellular serpins (Forsyth *et al.* 2003).

Table 1.4: Predicted RCL sequences of mouse α_1 -ACT homologues in comparison with human α_1 -ACT. Adapted from Forsyth *et al* (2003)

SERPIN NOMENCLATURE	COMMON NAME	RCL (P ₁ -P ₅)	% INDENTITY TO HUMAN α_1 -ACT
SERPINA3(human)	α_1 -ACT	GTEASAATAVKITLL-SALVE	
Serpina3a	UNK1	GTEADVITIARYNFQ-SAKIK	51%
Serpina3b	6A1	GTEGDAITIVGYNFM-SAKLK	51%
Serpina3c	1A1	GTEGVAATGVNFRIL-SRRTS	59%
Serpina3d	2B1	GTEADAATRFKIAPL-SAKFD	53%
Serpina3e	2B2 ^(a)	GTEAAAATGVKVNLR-CGKIY	58%
Serpina3f	2A1	GTEAAAGTGYQNLQC-CQGVI	59%
Serpina3g	2A2 or serpin2A	GTEAAAATGMAGVGC-CAVFD	55%
Serpina3h	6C28	GTEAAAATGVKLILC-CEKIY	60%
Serpina3i	2B2	GTEAAAATGVKVNLR-CGKIY	59%
Serpina3j	UNK2	GTEARATTRDKYDFL-STKSN	53%
Serpina3k	MMCM2	GTEAAAATGVIGGIR-KAVLP	58%
Serpina3l	3E2	DTEVDATSRAIYNFQ-SSKMY	51%
Serpina3m	3E46	GTEAAAATGFIFGFR-SRRIQ	61%
Serpina3n	EB22.4	GTEAAAATGVKFVPM-SAKLY	61%

(a) predicted pseudogene

1.5.2.a. Gene organisation of the mouse *serpina3* locus

Gene organisation of the mouse *serpina3* locus on chromosome 12F1 was determined from sequence information generated by the Mouse Genome Sequencing Consortium and from arrangements determined by Inglis and Hill (1991) (Forsyth *et al.* 2003) (Figure 1.4). There appears to be no ordered orientation of the genes although the predicted intracellular members cluster together at the centre of the locus flanked on each side by their extracellular counterparts. All intron/exon boundaries are conserved and correspond with those in the human *SERPINA3* clade (Silverman *et al.* 2001).

1.5.2.b. Expression of the *serpina3* genes

Prior to the data presented in this thesis limited information was available on expression of the mouse *serpina3* locus. Two *serpina3* cDNA transcripts, termed EB22/3 and EB22/4, were isolated from a mouse chondrocytic teratocarcinoma cell line EB22 (Inglis *et al.* 1991). The 2035bp EB22/4 cDNA contained the entire coding region, with a complete 3'-UTR and 60 bp of 5'-UTR, and was expressed specifically in liver. It has subsequently been identified in murine pancreatic β -cells where it is regulated by interferon- γ (IFN- γ) (Pavlovic *et al.* 1999). The EB22/3 transcript, which is also known as *serpin2A* and *serpina3g*, has since been further characterised and is described in section 1.5.2.c.

Whilst a number of reports cite the expression of mouse α_1 -ACT-like genes it is difficult to ascertain the exact member identified, given the high homology between them. Many of the cDNA probes used to identify the genes are designed against conserved serpin regions and would not discriminate between the highly homologous *serpina3* members. For example, Chiang *et al.* reported the up-regulation of a mouse *serpina3* gene(s) in the brains of IL-6 over-expressing mice and this was associated with the development of reactive gliosis (Chiang *et al.* 1994). Furthermore, a member from this locus

was implicated in the progression and pathogenesis of Alzheimer's disease in mouse models (Licastro *et al.* 1999). The probes used in both studies were designed to conserved regions of the serpin molecule therefore it is impossible to predict which mouse *serpina3* gene is responsible for the reported expression.

A number of reports have implicated the mouse *serpina3* locus in cellular growth, differentiation and transformation. Whitehead *et al.* isolated an antisense transcript of a *serpina3* gene which produced cellular transformation in mouse fibroblasts (Whitehead *et al.* 1995). Using DNA microarray to analyse gene expression following malignant transformation of mouse PB-3c mast cells with v-H-ras oncogene, a *serpina3* gene was shown to be down-regulated in tumour cell lines (Brem *et al.* 2001). Furthermore, in a study to identify differentially expressed genes on IFN- γ induced maturation of pre-B cells, a mouse *serpina3* cDNA transcript was implicated in the pre-B to immature B cell transition (Patrone *et al.* 2001). However, these studies may reflect the expression of one or more members of the *serpina3* family.

1.5.2.c. *Serpin2A (serpina3g)*

Serpin2A is a unique and intriguing member of the murine *serpina3* gene family. Subsequent to its partial cDNA isolation from the teratocarcinoma cell line EB/22, serpin2A was separately identified by subtractive hybridisation between immature and differentiated haemopoietic stem cells (Hampson *et al.* 1997). Hampson *et al.* utilised the primitive murine, pluripotent, haemopoietic cell line, FDCP-Mix A4, which can be stimulated to differentiate down different lineages under the influence of growth factors. Serpin2A was found to be highly expressed in undifferentiated cells and was down-regulated upon differentiation. Constitutive serpin2A expression in FDCP-Mix A4 cells also delayed differentiation and increased clonogenicity. Additionally, they demonstrated that high level serpin2A expression in bone-marrow derived bipotent granulocyte/macrophage-colony forming cells (GM-CFC) was down-regulated in their maturing progeny (Hampson *et al.* 1997). Northern blot

analysis showed that the principal tissues expressing serpin2A were thymus, lung, spleen, testis and bone marrow (Hampson *et al.* 1997). At the time it was believed a specific probe had been designed to serpin2A but given the highly homologous nature of the mouse serpin3 family it is likely the probe would have cross reacted with other members of the family. However, other groups have confirmed serpin2A to be differentially expressed in the haemopoietic stem cell compartment when compared with total cells of the bone marrow (Terskikh *et al.* 2001). The prominent expression of serpin2A in haemopoietic progenitors and its down regulation during differentiation suggests a role in haemopoietic differentiation. The biological role of serpin2A in haemopoiesis is a matter of investigation in this thesis.

Serpin2A is also up-regulated during macrophage and T-cell activation. Expression was stimulated in activated but not in resting T-cells (Hampson *et al.* 1997) and in macrophages in response to IFN γ and macrophage activation *in vivo* (Hamerman *et al.* 2002). Conjugation of serpin2A to the 15kDa protein ISG15, an interferon stimulated gene product which is homologous to ubiquitin, was also observed in macrophages (Hamerman *et al.* 2002). The conjugation process is thought to be similar but distinct to that of ubiquitination and is not believed to target proteins for degradation. Its significance to serpin2A function in activated macrophages is yet to be elucidated.

Despite sharing 60% amino acid identity with human α_1 -ACT, serpin2A has features which are distinct from its closest human orthologue. It has a Cys-Cys pairing at P₁-P₁' and an unusual C-terminal extension of approximately 30 residues which contains a further two cysteine residues. Another cysteine residue is also present in the N-terminus (Morris *et al.* 2003). The exposed nature of these residues suggests that they partake in redox reactions. In fact, recombinant serpin2A shows decreased structural stability under reducing conditions suggesting that serpin activity may be regulated by a redox environment. Furthermore, the absence of a conventional N-terminal signal sequence suggests it would be an intracellular protein. This was confirmed by

Morris *et al* where subcellular localisation studies of serpin2A demonstrated prominent nuclear localisation (Morris *et al*. 2003).

The biological role of serpin2A is unknown. It is able to undergo the stressed→relaxed conformational change and therefore has the capacity to function as an inhibitory serpin. Recently, serpin2A was demonstrated to possess inhibitory activity against the serine protease cathepsin G and the cysteine proteases, cathepsin L, V, K H and B (Liu *et al*. 2003). The ability to inhibit such a wide range of proteases which possess different mechanisms is surprising. No such inhibitory activity is observed by any of the cystatins which are the regulators of cysteine protease activity *in vivo*. Furthermore, these results are in direct contrast to our own data, in which no inhibitory activity towards cysteine proteases was observed (Morris *et al*. 2003). Therefore, the exact nature of this inhibition requires further investigation to fully understand the mechanisms involved.

1.5.3. The *serpina1* family

The second mouse a-clade locus to have undergone expansion, in comparison to its human counterpart, is the mouse *serpina1* (α_1 -AT) locus. Up to seven different isoforms have been identified in mice (Barbour *et al*. 2002). The number of genes present in mice is both species and strain specific, with *Mus caroli* and *Mus cookie* expressing only one isoform, while *Mus saxicola* and *Mus domesticus* possess four and five genes respectively (Goodwin *et al*. 1997). Furthermore, the inter-strain variation is highlighted in *Mus domesticus* where some strains express between 3-5 different genes. For example, C57BL/6J and A/J express five isoforms (serpina1a, serpina1b, serpina1c, serpina1d, and serpina1e) whilst BALB/cJ express only three (serpina1a, serpina1b and serpina1f) (Barbour *et al*. 2002). The inter-species and inter-strain variation observed in mice is the result of a differing number of genes present at the locus and not merely differences in gene expression, as shown by gene specific PCR (Barbour *et al*. 2002).

All the *serpina1* genes are expressed in liver as demonstrated by Northern blot analysis and by RT-PCR and all appear competent to encode functional proteins (Borriello and Krauter 1991). The sequence between the members is highly conserved throughout the core serpin structure with variations between the isoforms and their human homologue restricted to the reactive centre loop. The Met-Ser at P₁-P₁' is identical to that of their human orthologue α_1 -AT and this pairing is conserved in all mouse isoforms but two where the P₁ residue has been substituted by either a leucine or tyrosine (Borriello and Krauter 1991; Barbour *et al.* 2002; Forsyth *et al.* 2003). The Leu-Ser pairing suggests an α_1 -ACT like inhibitory specificity, whereas the Tyr-Ser was shown to inhibit trypsin and chymotrypsin (Barbour *et al.* 2002). The variations in expression observed between species and strains appears to involve only the unorthodox members (those with a substitution at P₁) as all mice express at least one gene with the orthodox Met-Ser at P₁-P₁' (Barbour *et al.* 2002). Other variations in sequence flank the predicted P₁-P₁' residues in the RCL. The proline residues at P₂ are conserved in four of the isoforms and all but one isoform has the Pro-Pro residue at P₃' and P₄'. However, between P₂ and P₆ there is extensive divergence (Table 1.5). This may reflect either a broader target protease specificity or protection from cleavage by non-cognate proteases.

The amino acid variations in the RCL are thought to be driven by positive Darwinian selection to provide adaptive functional activity between the different isoforms of mouse α_1 -AT. Analysis of the protease specificity of the α_1 -AT isoforms is consistent with this suggestion. Biochemical analysis showed that those isoforms with orthodox methionine at P₁ were efficient inhibitors of elastase, whilst the three unorthodox members with tyrosine at P₁ were able to inhibit either trypsin, chymotrypsin or both proteases (Barbour *et al.* 2002). Furthermore, by using RCL-swapping chimeras, it was demonstrated that the variation in target protease activity was directly attributable to the reactive centre sequence (Barbour *et al.* 2002). In addition, the different α_1 -AT isoforms from *Mus. saxicola* differed in their ability to bind proteinases in crude snake venom (Barbour *et al.* 2002). This is consistent with the theory that the array of exogenous proteases, introduced by a pathogen or predator, may drive the

Table 1.5: Predicted RCL sequences of mouse α_1 -AT homologues from *Mus.domesticus* (DOM) and *Mus. Saxicola* (SAX) in comparison with human α_1 -AT. Adapted from Goodwin *et al* (1997)

SERPIN	RCL
Human α_1 -AT	AMFLEAIPM-SIPPEVKFNK
DOM1	VTVLQMVPM-SMPPIILRFDH
DOM2	ATVFEAVPM-SMPPIILRFDH
DOM3	ATVLLAVPY-SMPPIVRFDH
DOM4	ATVLQVATY-SMPPIVRFDH
DOM5	ATVLQAGFL-SMPPIILHFNR
DOM6	ATVLLAVPY-SMPPIILRFDH
SAX1	TTIVEAVFM-SLPPIILHFNH
SAX2	TTIVEGVFM-SLPPIILNFNH
SAX3	TTVLAGSYM-SAPPILNFNC
SAX4	TTVLGSTLY-SAPAILHFNH
SAX5	TTVLAGTFT-SWPPILNFNR

functional diversification of the α_1 -AT's during evolution (Hill and Hastie 1987; Borriello and Krauter 1991).

1.5.4. Murine expansion of intracellular serpins

When comparing the human B-clade serpins MNEI, PI-6 and PI-9 with their murine counterparts a similar gene expansion was observed (Kaiserman *et al.* 2002). Analogous to their human counterparts the murine ov-serpins are clustered on two chromosomes in the mouse on chromosomes 1 and 13 syntenic to human chromosomes 6p25 and 18q21 (Copeland *et al.* 1993; Sun *et al.* 1995 (a)). The serpins at human 6p25 include the genes MNEI, PI-6, and PI-9 and the syntenic region on mouse chromosome 13 has expanded to encode approximately 16 genes (Sun *et al.* 1998; Kaiserman *et al.* 2002). Each of these loci has expanded to 4, 5, and 7 genes respectively. They are highly homologous to their human equivalents but, like the mouse a-clade expanded members, the murine b-clade expanded serpin genes show considerable variation in the functionally important RCL domain. Each of the mouse b-clade clusters encodes both functional genes and pseudogenes. The second syntenic region at 18q21 appears not to have undergone this expansion (Kaiserman *et al.* 2002).

Gene expression studies showed that at least one member from each expanded b-clade serpin locus has similar tissue distribution to their human counterpart (Benarafa *et al.* 2002; Kaiserman *et al.* 2002). In addition, at least one member shared the same target protease(s) as the human serpin, whilst other members showed either varied or reduced activity towards these proteases (Benarafa *et al.* 2002). This is most likely the result of the variation in RCL sequence.

The generation of multiple genes in mice is most likely to have occurred post divergence of the mice and men. The selective pressures driving the expansion are unknown. Whilst there is an increased number of granule proteases in mice these are only in the subclasses of granzymes and cysteine

proteases (Smyth and Trapani 1995; Deussing *et al.* 2002). It has been suggested that the increased copy number of serpin genes contributes to resistance to parasites and other pathogens (Hill and Hastie 1987; Inglis and Hill 1991). This selective pressure may be stronger in rodents due to increased exposure to parasitic infection and pathogens. Inhibition of the NS3 serine protease of the hepatitis C virus by the serpins C1 inhibitor and α_2 -antiplasmin *in vitro* demonstrates that serpins may have a role in regulating pathogen proteases (Drouet *et al.* 1999). However, no direct evidence supports the theory of exogenous proteinases driving serpin loci expansion. Therefore, it is equally likely that these serpins play a role in endogenous biochemical or cellular responses.

1.6. SERPINS AND THEIR ROLE IN HAEMOPOIESIS

1.6.1. Haemopoiesis

Throughout life the normal haemopoietic system generates approximately 4×10^6 mature blood cells per second in humans (Moore 1996). These mature cells have a limited life span and therefore a balance must be achieved between the production and loss of cells to maintain homeostasis. Haemopoiesis is sustained in adult mammals by primitive multipotential haemopoietic stem cells (HSC) which reside within the bone marrow. The key features of primitive HSC's are their ability to proliferate (self-renewal) and to undergo differentiation to replenish the mature blood cells of at least nine lineages (Szilvassy 2003) (Figure 1.5). As haemopoiesis proceeds from the primitive HSC's to their more mature progeny, there is a loss in self-renewal potential associated with increased lineage restriction.

Primitive HSC's differentiate into three branches of committed progenitor cells which give rise to the vascular, myeloid and lymphoid lineages. The vascular progenitors produce cells responsible for the maintenance of the circulatory system: erythrocytes, megakaryocytes and macrophages. The myeloid lineage

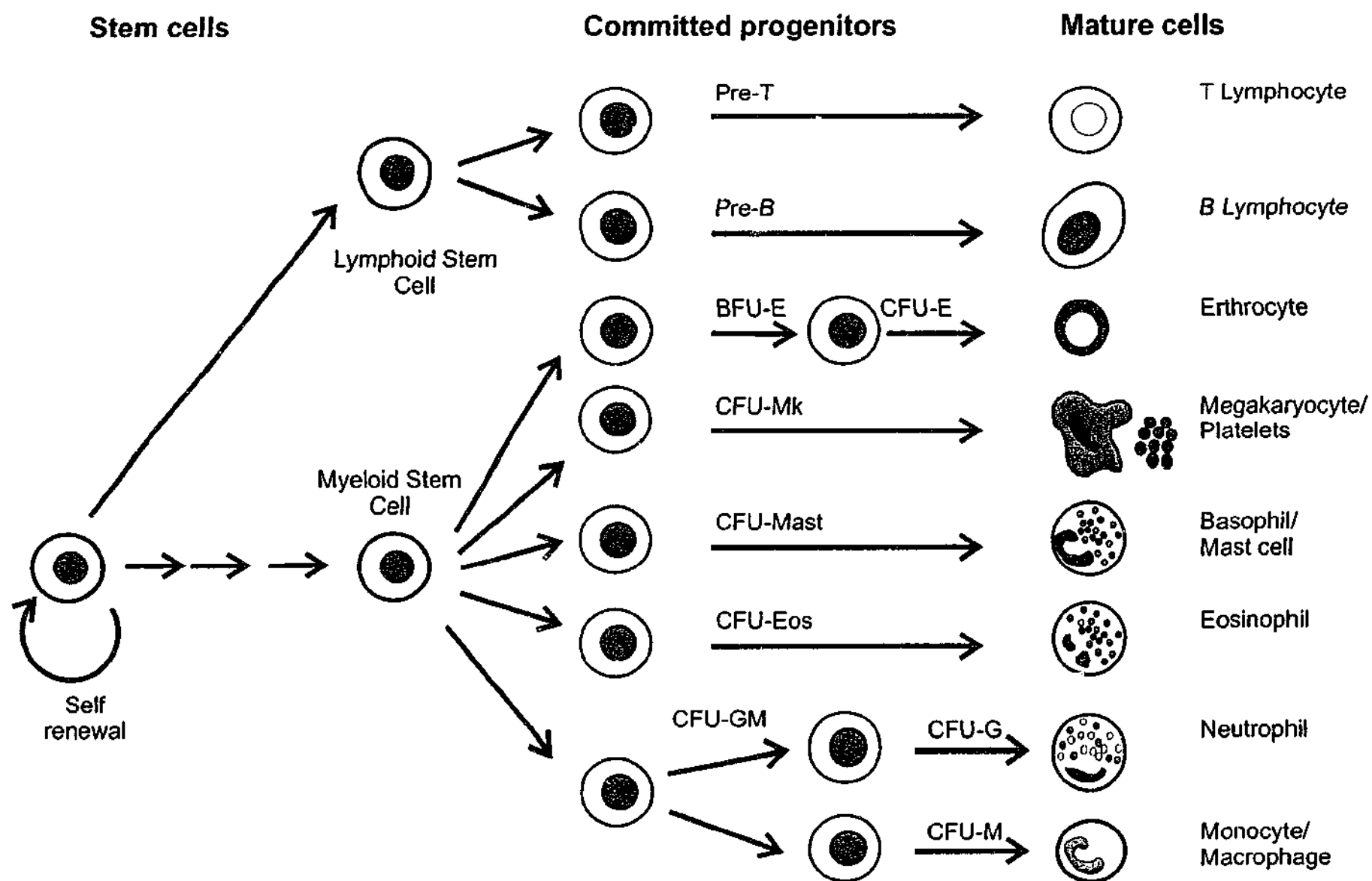


Figure 1.5: Schematic representation of the differentiation of haemopoietic stem cells into the mature progeny located in the circulation. Progenitor cells are represented as: colony-forming cells (CFC) or colony forming units (CFU) and specific lineages defined as GM:granulocyte/macrophage, G: granulocyte, M: macrophage, Eos: eosinophil, Mast: Mast or basophil and Mk: megakaryocyte. The terms BFU-E and CFU-E describe blast- and colony-forming units of the erythroid lineage. Adapted from Szilvassy (2001)

progenitors give rise to cells responsible for the innate immunity system: granulocytes, macrophages and mast cells. The cells generated from the lymphoid lineages are associated with antigen-specific acquired immunity: the B and T cells (reviewed in (Cooper and Spangrude 1999)). Differentiation of the HSC's to their more mature progeny is achieved by a number of stimulatory haemopoietic cytokines and these include the factors, colony stimulating factor-1 (CSF-1), granulocyte colony-stimulating factor (G-CSF) and erythropoietin.

In the steady-state bone marrow, the primitive HSC's are quiescent whilst more mature progenitors are actively cycling. A number of cytokines have been identified to promote survival and induce quiescent cells to cycle and include stem cell factor (SCF), Flt3-ligand, IL-6 and IL-11. Cytokines shown to induce the proliferation of primitive HSC's include IL-3, granulocyte-macrophage colony-stimulating factor (GM-CSF), and mpl ligand (thrombopoietin) (reviewed in (Ogawa 1993)) (Sitnicka *et al.* 1996).

The functional characterisation of a primitive HSC *in vivo* is the ability to reconstitute haemopoiesis in a marrow-ablated recipient (Lorenz *et al.* 1951), whilst the *in vitro* phenotype is characterised by the expression of specific cell surface markers. Cells that express Thy-1.1 and stem cell associated (Sca) antigen and the c-kit receptor, but lack lineage markers (Lin⁻), are viewed as primitive HSC's (Weissman *et al.* 2001).

1.6.2. Haemopoiesis and proteolysis

A number of proteases are associated with haemopoiesis. Meyer *et al.* identified the Stat5 serine protease in primitive haemopoietic cells (Meyer *et al.* 1998). This nucleus-associated protease cleaves the Stat5 protein in both active and inactive forms, to produce a carboxyl truncated STAT5 β (Lee *et al.* 1999). Stat5 is a signal transduction molecule, involved in the JAK/STAT signalling pathway and is stimulated by the cytokines that regulate the proliferation of and differentiation of the myeloid and lymphoid lineages,

including IL-3, IL-2, IL-15 and GM-CSF. In its cleaved form the Stat5 protein is unable to promote the induction of target genes. In the absence of Stat5 proteins there are significant defects in the development of CFU-Mix, CFU-Eos and CFU-GM colonies (Teglund *et al.* 1998), suggesting that the protease may regulate differentiation and self-renewal kinetics in the haemopoietic compartment (Lee *et al.* 1999).

The azurophilic granules of neutrophils contain the serine proteases elastase, cathepsin G and proteinase3/myeloblastin (Borregaard and Cowland 1997). These are released at sites of inflammation where they are responsible for the degradation of matrix components such as elastin, fibronectin and collagen (Owen *et al.* 1995; Shapiro 2002). Elastase and cathepsin G have been shown to cleave the cytokine receptors for IL-2 and IL-6, respectively (Bank *et al.* 1999). Mice lacking elastase are impaired in their ability to kill Gram-negative bacteria but are otherwise normal, whereas cathepsin G deficient mice are phenotypically normal (Belaaouaj *et al.* 1998; MacIvor *et al.* 1999).

The mapping of the gene responsible for autosomal dominant cyclical neutropenia to chromosome 19p13.3, the location of the human neutrophil elastase gene, highlights the importance of the neutrophil proteases in regulating haemopoiesis (Horwitz *et al.* 1999). Subsequent studies showed that patients with congenital neutropenia also have mutations in the elastase gene (Dale *et al.* 2000).

The neutrophil proteases have been implicated in the mobilisation of HSC's. Mobilisation of HSC's from the bone marrow into the peripheral circulation occurs in response to a wide variety of stimuli including polyanions, chemokines and haemopoietic growth factors (To *et al.* 1997). The cytokine G-CSF promotes the release of HSC's and is used therapeutically to reconstitute haemopoiesis following high-dose chemotherapy (Korbling and Anderlini 2001). HSC mobilisation in response to G-CSF has been shown to coincide with the accumulation of high concentrations of neutrophil elastase and cathepsin G in the bone marrow extracellular fluid (Levesque *et al.* 2001; Levesque *et al.* 2002). The two proteases disrupt the adhesive and

chemotactic interactions by which HSC's home to and remain in the bone marrow. The cleavage of VCAM-1 by the proteases reduces their adhesion to the bone marrow stroma and chemotaxis of the HSC's is altered due to the cleavage of the chemokine receptor CXCR4 expressed on HSC's (Levesque *et al.* 2001; Levesque *et al.* 2003). The exact mechanism by which neutrophil proteases accumulate in bone marrow is unknown, however, it is associated with a decrease in the expression and concentration of their cognate inhibitors, the serpins α_1 -AT and α_1 -ACT (Winkler *et al.* 2004).

Proteinase 3 (myeloblastin) was originally identified as a gene down-regulated in differentiating HL60 cells. Furthermore, down-regulation of proteinase3 using an anti-sense oligonucleotide resulted in the growth arrest and induction of monocytoid differentiation (Bories *et al.* 1989). Proteinase 3 down-regulates DNA synthesis of normal granulocytic progenitor cells and its expression is upregulated by G-CSF, a cytokine that promotes survival and proliferation of granulocytic progenitors (Lutz *et al.* 2000). The constitutive over-expression of proteinase 3 confers factor independent growth to Ba/F3 cells expressing the G-CSF receptor.

1.6.3. Intracellular serpins and haemopoiesis

The two intracellular serpins, MNEI and PI-6, are efficient inhibitors of the neutrophil granule proteases. PI-6 was originally isolated from the cytosolic fraction of the leukemic cell line K562 and from human placental extracts (Coughlin *et al.* 1993 (a,b)). More recently, its expression has been demonstrated in monocytes, granulocytes and myelomonocytic cell lines. PI-6 is an extremely rapid inhibitor of cathepsin G and is suggested to protect cells by sequestering the cathepsin G activity that leaks from leukocyte granules during biosynthesis or phagocytosis (Scott *et al.* 1999 (a)).

MNEI was initially purified from the cytoplasmic fraction of monocytic cells (Remold-O'Donnell 1985). The presence of Cys-Met residues at P₁-P₁' suggested an oxidation sensitive molecule. MNEI is an efficient inhibitor of

both elastase-like and chymotrypsin-like proteases and the reactions are mediated through the use of two reactive sites. The rapid inhibition rate constants ($>10^7 \text{ M}^{-1}\text{sec}^{-1}$) seen with neutrophil elastase and proteinase 3 is mediated through Cys (344), whilst inhibition of cathepsin G occurs through the preceding residue Phe (343) (Cooley *et al.* 2001). MNEI is thought behave similarly to PI-6 in controlling proteolysis in the cytosol to protect monocytes and granulocytes from their own granule contents (Remold-O'Donnell *et al.* 1992).

The intracellular serpins PI-9 and PI-10 were first identified by their expression in bone marrow (Riewald and Schleef 1995). Further examination of PI-9 showed predominant expression in natural killer cells and T and B lymphocytes (Sun *et al.* 1996). The presence of the Glu residue at P_1 makes PI-9 a rapid inhibitor of granzyme B. Transfection of PI-9 protected cells from apoptosis induced by either granzyme B or perforin. The degree of protection correlated with the level of PI-9 expression (Bird *et al.* 1998). Its exact role in the regulation of granzyme B is yet to be fully elucidated, however it is thought to protect cytotoxic lymphocytes and bystander cells from granzyme B-mediated apoptosis.

Sequence analysis of PI-10 predicts an Arg-Ile at P_1 - P_1' and it has been demonstrated to form stable inhibitory complexes with trypsin and thrombin (Riewald and Schleef 1995). It possesses a CD interhelical loop, as previously described (section 1.4.2) and this loop is required for its nuclear localisation (Chuang and Schleef 1999). Riewald *et al.* (1998) showed that PI-10 (bomapin) expression is higher in bone marrow than in peripheral blood. Furthermore, the levels of PI-10 were higher in the peripheral blood of patients with acute monoblastic leukaemia and chronic myelomonocytic leukaemia when compared with normal controls. From these results they inferred that PI-10 is expressed in haemopoietic precursors, although no direct evidence of PI-10 expression in haemopoietic progenitors was demonstrated (Riewald *et al.* 1998). The functional role for PI-10 in haemopoietic cells has not yet been determined, although HeLa cells transfected with PI-10 are protected from $\text{TNF}\alpha$ -induced cell death. This was associated with the formation of a higher-

molecular weight complex, which was suggested to be an interaction between PI-10 and a putative target protease (Schleef and Chuang 2000).

The serpin megakaryocyte maturation promoting factor (MMF) was initially identified as a factor able to induce megakaryocyte differentiation of murine bone marrow (Tsujimoto *et al.* 1997). The presence of Lys at P₁ predicts a trypsin-like inhibitory specificity, however, a cognate protease is yet to be elucidated. Recombinant MMF is able to promote megakaryocyte growth from human bone marrow *in vitro*, in the presence of IL-3, and its activity is similar to the cytokine IL-11. The injection of mice with recombinant MMF for 10 days resulted in a 40% increase in platelet count compared to control mice (Tsujimoto *et al.* 1997). These preliminary studies suggest MMF has a megakaryocyte potentiating activity, although further investigation is required to fully understand the role of MMF in haemopoiesis. In separate investigations, MMF has been identified as a serpin produced by mesangial cells in kidney and given the additional name, megsin (Suzuki *et al.* 1999).

1.6.4. Extracellular serpins and haemopoiesis

Whilst α_1 -AT is primarily expressed in the liver, it is also expressed in monocytes and myelomonocytic cells (Perlmutter *et al.* 1985; Krugliak *et al.* 1986; Missen *et al.* 2004). The role of α_1 -AT in the haemopoietic compartment is poorly characterised. Extracellular protease inhibitors have been suggested to be important in the protection of cytokines, growth factor receptors and other proteins essential to the HSC microenvironment from degradation by proteases secreted by myeloid cells in the bone marrow (Goselink *et al.* 1996; Winkler *et al.* 2004). Goselink *et al.* (1996) demonstrated that addition of α_1 -AT to growth media was able to support growth of CD34⁺ haemopoietic progenitor cells (HPC) *in vitro*. They suggested that the presence of protease inhibitors, such as α_1 -AT, protected the mediators of HPC proliferation, such as cytokines and cell surface receptors from degradation (Goselink *et al.* 1996).

In a recent study, gene products from the mouse a-clade loci of α_1 -AT (*serpina1*) and α_1 -ACT (*serpina3*) were suggested to play a role in the homeostasis of the haemopoietic microenvironment by determining the fate of HPC's (Winkler *et al.* 2004). Serine proteases have been shown to accumulate in the bone marrow after G-CSF treatment which promotes mobilisation of HSC's into the periphery (as described in Section 1.6.2). Gene expression of both mouse *serpina1* and *serpina3* loci was down-regulated in murine bone marrow in response to G-CSF and cyclophosphamide treatment without a significant increase in transcription of neutrophil elastase and cathepsin G. Therefore, the accumulation of neutrophil elastase and cathepsin G in the HPC microenvironment and subsequent mobilisation of stem cells was suggested to be a result of *serpina1* and *serpina3* down-regulation. This suggests that disruption of the serpin/protease balance may contribute to the mobilization of HSC's (Winkler *et al.* 2004).

At the commencement of this study little was known about the murine orthologues of α_1 -ACT. The aim of this study is to characterise the mouse *serpina3* locus by investigating gene expression and to determine whether the prominently expressed genes represent functional serpins. The latter part of this study aims to investigate the biological role of serpin2A in haemopoietic progenitors which may provide insights into one of many possible roles of the *serpina3* genes *in vivo*.

CHAPTER 2

MATERIALS AND METHODS

2.1. GENERAL REAGENTS AND EQUIPMENT

2.1.1. Reagents

All chemicals and reagents were obtained from either Sigma Aldrich (Castle Hill, Australia) or BDH Chemicals (Dorset, UK). Bacterial growth media constituents were obtained from Oxoid (Hampshire, England). Mammalian cell culture media and L-glutamine was obtained from Invitrogen (Melbourne, Australia) and fetal calf serum obtained from JRH Biosciences (Kansas, USA). Antibiotic Antimycotic Solution (100x) was purchased from Sigma Aldrich (Castle Hill, Australia). The radiochemicals ^{35}S Methionine and $\alpha\text{-}^{32}\text{P}$ -dATP were obtained from Perkin Elmer (Rowville, Australia). The ^{35}S -Met was aliquoted and stored at -80°C , and $\alpha\text{-}^{32}\text{P}$ -dATP stored at 4°C . Commonly used buffers and media are presented in Table 2.1.

2.1.2. Enzymes

Taq polymerase was purchased from Geneworks (Adelaide, Australia). Superscript II reverse transcriptase and DNase A were purchased from Invitrogen (Melbourne, Australia). Restriction endonucleases, T4 DNA ligase and Vent polymerase were purchased from New England Bio Labs (Beverly, USA). RNasin[®] Ribonuclease Inhibitor was obtained from Promega (Annendale, Australia).

Table 2.1: Commonly used buffers and media

Buffers	
10 X PBS	1.37 M NaCl, 27 mM KCl, 43 mM Na ₂ HPO ₄ ·7H ₂ O, 14 mM KH ₂ PO ₄ , pH 7.2
PBST	1 X PBS/ 0.1% Tween 20
TE	10 mM Tris-Cl, pH 7.4, 1 mM EDTA
TAE	40 mM Tris, 50 mM sodium acetate, 2mM EDTA
TBE	89m M Tris, 89mM boric acid, 2 mM EDTA
Native anode buffer (10x)	1M Tris-HCl, pH 7.8
Native cathode buffer (10x)	529 mM Tris-HCl, pH 8.2, 684 mM glycine
10 X SDS-PAGE electrophoresis buffer	0.250 M Tris base, pH 1.92 M glycine, 1%(w/v) SDS
10 x Western transfer buffer	0.250 M Tris base, 1.92 M glycine, For 1 X buffer: stock was diluted 1/10 with dH ₂ O and 10% methanol (final)
Coomassie Stain	50% (v/v) Methanol, 0.05% Coomassie Brilliant Blue R-250, 10% (v/v), acetic acid, 40 % dH ₂ O
Destain	10% (v/v) glacial acetic acid, 40% (v/v) methanol
DNA loading	1.5 mM Tris-Cl, pH 7.6, 2% glycerol (v/v), 10 mM EDTA, 0.2% SDS (w/v), 0.02% bromophenol blue (w/v)
TBE Gel fixative	20 % methanol, 7% acetic acid
5 X SDS-PAGE sample buffer	125 mM Tris-HCl, pH 6.8, 25% glycerol (v/v), 10% SDS (w/v), 0.5% bromophenol blue, 100µl of 14.2 M β-mercaptoethanol per ml loading dye
2 X SDS-PAGE sample buffer	50mM Tris-HCl, pH 6.8, 10% glycerol (v/v), 4% SDS (w/v), 0.20 % bromophenol blue (w/v), 25µl of 14.2 M β-mercaptoethanol per ml of loading dye
5 X Native-PAGE sample buffer	235 mM Tris, pH 6.8, 10% glycerol (v/v), 0.04% bromophenol blue (w/v)
Media	
Luria Broth (LB)	10g tryptone, 10g yeast extract , 5g NaCl, 1ml IM NaOH make up to 1 litre with dH ₂ O
2 X TY	16 g tryptone, 10 g yeast extract, 5g NaCl make up to 1 litre with dH ₂ O
LB/agar	15g of agar per litre of LB media

2.1.3. *Proteases*

Bovine chymotrypsin and trypsin were purchased from Sigma Aldrich (Castle Hill, Australia). Trypsin and chymotrypsin were dissolved in 50 mM Tris-HCl, pH7.4, 20 mM CaCl₂ at a concentration of 10 mg/ml and divided into 50 µl aliquots and stored at -80°C. Elastase and cathepsin G were obtained from Athens Research Technology (Athens, USA). Recombinant human cathepsin L and cathepsin V expressed in *Pichia pastoris* (GS115 host strain) and purified by Poh Chee Ong and Wei Wen Dai (Department of Biochemistry and Molecular Biology, Monash University). Thrombin and factor Xa were obtained from Dr Noelene Quinsey (Department of Biochemistry and Molecular Biology, Monash University). All proteases were aliquoted and stored at -80°C.

2.1.4. *Equipment*

The DNA gel electrophoresis apparatus, the Wide Mini-Sub Cell GT, and both the Mini-Protean III electrophoresis system and Mini Trans blot module for Western blotting were purchased from Bio Rad (California, USA). The Chemi Genius used for the visualisation of ethidium-bromide stained DNA gels, and Coomassie stained SDS-PAGE gels was obtained from Syngene (Fredrick, USA).

The DU 530 UV/Vis spectrophotometer (Beckman: Gladesville, Australia) was used for 1) protein estimation at 595nm and 280nm, 2) DNA qualification and quantification at OD 260nm and 280nm, and 3) bacterial density at 600nm. The Polstar Optima from BMG (Offenburg, Germany) was used for fluorogenic assays and for chromogenic assays Thermomax microplate reader from Molecular Devices (California, USA). The shaker-incubator for variable settings (20-37°C) was from Ratek (Boronia, Australia), and the 37°C CO₂ incubator was from Sanyo (Bensenville, IL, USA). Centrifuges used included; the Sorvall RC5B Plus with the SS34 and SLA-1500 rotors and the Heraeus Biofuge

Centrifuge which were purchased from Kendro (Lane Cove, Australia) and the Beckman GS-6R centrifuge from Beckman-Coulter (Gladesville, Australia).

2.2. MOLECULAR BIOLOGY

2.2.1. DNA purification

2.2.1.a. PCR purification and gel extraction of DNA

PCR products were purified using the PCR purification kit from Qiagen (Melbourne, Australia) according to manufacturer's instructions and DNA was eluted into sterile dH₂O. DNA bands excised from agarose gels were purified using the Gel Extraction Kit (Qiagen) according to manufacturer's instructions.

2.2.1.b. Mini Preparation

Mini Prep DNA was routinely purified from 1.5-5 ml overnight culture of *E. Coli* using the Qiagen Miniprep kit, and according to manufacturer's instructions. DNA bound to the silica-membrane of the spin-column was eluted into sterile Milli-Q dH₂O.

2.2.1.c. Maxi Preparation

Maxi Prep DNA was purified from 400-500 ml overnight *E. Coli* cultures using the Wizard Plus Maxipreps DNA Purification System (Promega). DNA was eluted into sterile Milli-Q dH₂O and concentration quantified by measuring absorbance at 260nm and calculated as follows:

DNA concentration (µg/ml) = Absorbance 260nm (units) x 50 x dilution factor.

2.2.2. DNA Electrophoresis

Agarose was dissolved into 1 x TAE buffer to a final concentration of either 1% or 2%. Ethidium bromide was added to a final concentration of 0.5 µg/ml. Molecular weight markers used were either λ/HindIII marker or φX/HaeIII marker from New England Bio Labs (Beverly, USA) or 100bp ladder from Promega (Annandale, Australia). DNA gels were electrophoresed at 80V in 1 x TAE buffer. DNA bands were visualized under a UV illuminator.

6% polyacrylamide/TBE electrophoretic gels were used to analyse DNA smaller than 200bp. Gels were poured between two glass plates with 1mm spacers and electrophoresed using the Mini-Protean System III in 1 x TBE at 120V for 50 mins.

TBE gel composition

<i>6% polyacrylamide gel</i>	
40% acrylamide (29:1)	1.5 ml
10 x TBE	1 ml
Deionised H ₂ O	7.5 ml
30% APS	70 µl
TEMED	5 µl

To visualise DNA gels were stained with ethidium bromide (0.2 µg/ml)/TBE for 5 min. For visualisation of ³²P-dATP labelled DNA the gels were fixed in TBE gel fixative (20% methanol, 7% acetic acid, in dH₂O) and dried under vacuum at 80°C for 1 hr.

2.2.3. Polymerase Chain Reaction (PCR)

PCR was used to amplify DNA sequences of interest. PCR primers were designed to anneal to specific regions of template DNA as outlined in Table 2.1. Primers were obtained from Sigma Genosys (Castle Hill, Australia). PCR was carried out using the Gene Amp® PCR System 9700 from Perkin Elmer (Rowville, Australia). The basic constituents of the PCR reaction consisted of: 10 x PCR buffer, 0.2 mM dNTP mix, 1.5 mM MgCl₂, 40 pmol of each primer, DNA polymerase (0.05 units/μl) and DNA template ~10 ng per 100 μl reaction. Specific primer pairs and cycle conditions are described in the subsequent chapters. PCR products were analysed by ethidium bromide agarose gel electrophoresis.

2.2.4. DNA Sequencing

Sequencing of DNA was performed using the Big Dye Terminator Cycle Sequencing kit Version 3 and 3.1 (ABI Prism, Perkin Elmer: Rowville, Australia) and performed according to manufacturers instructions. A PCR based sequencing reaction consisted of 200-500 ng of plasmid DNA, 3.2 pmol of sequencing primer and 8 μl of Big Dye sequencing reagent version 3. For reactions with Big Dye sequencing reagent version 3.1, 2 μl of sequencing dye was used in the presence of 3 μl of 5 x PCR buffer. Sequencing PCR cycles were: 94°C 30 sec, [94°C 5 sec, 50°C 5 sec, 60°C 4min] x 25 cycles. The 20 μl sequencing reactions were ethanol precipitated using 62.5 μl of 96% ethanol, 14.5 μl of dH₂O and 3 μl of 3M Na-acetate pH 4.6, for 15 min. The samples were centrifuged at 16,000 x g for 20 min. The pellet was washed with 250 μl of 70% ethanol and centrifuged at 16,000 x g for 5 min. The ethanol was removed and pellet air dried. Samples were sent to Australian Genome Research Facility (WEHI, Melbourne, Australia) for DNA sequencing.

Table 2.2: Primer sequences

Primer Number	5'→3'	Template	Orientation
254	ATGGATCCTTCCCAGATGGCAC	MMCM2/EB22.4 at position 61	Sense
255	ATGGATCCTTCTCAGATGACACATG	3E46 at position 1254	Sense
256	ATGGATCCTACTTGGGGTATTG	MMCM2 at position 1242	Antisense
257	ATGGATCCTCATTTGGGGTTGGCT	EB22.4 at position 1254	Antisense
258	ATGGATCCTCACTTGGGGTTAGTG	3E46 at position 61	Antisense
260	ATGGATCCTCACTTGGGATTTG	6C28 at position 1214	Antisense
265	ATGGATCCATGGCTGGTATCTCCCTG	6C28 to amplify from start Met	Sense
334	GATGGAGAACTTGGGCTG	Sequencing serpina3 family at position 934	Antisense
335	GTCTCAAGTTCAATCTCACAG	Sequencing serpina3 family at position 318	Sense
T7	TAATACGACTCACTATAGG	Sequencing from T7 promoter	Sense
oligo(dT)	TTTTTTTTTTTTTT	Poly A-tail of RNA	Antisense

2.2.5. Preparation and transformation of chemically competent cells

BL21(DE3)pLysS, DH5 α , XL1-blue and AD494 competent cells were prepared by chemical methods. Cells were streaked out onto LB/agar plates and incubated overnight at 37°C. A single colony was used to inoculate 5 ml LB. The culture was incubated overnight with shaking at 225 rpm and the following day diluted into 1:10 into 100 ml LB. The cells were grown until the OD_{600nm} reached 0.5-0.6 (approximately 2hr). The cells were harvested, placed on ice for 15 min and centrifuged at 3,000 x g for at 4°C. The cells were resuspended in 40 ml ice cold rubidium chloride buffer 1 (30 mM , 100 mM rubidium chloride, 10 mM CaCl₂, 50 mM MgCl₂ and 15% glycerol (v/v), pH 5.8), and placed on ice 15 min. The cells were centrifuged at 3,000 x g for 15 min at 4°C. The pelleted cells were resuspended in 2 ml ice cold rubidium chloride buffer 2 (10 mM MOPS, 75 mM CaCl₂, 10 mM rubidium chloride, 15% glycerol (v/v), pH6.5) and incubated on ice for 15 min. The cells were either used immediately or divided into 100 μ l aliquots on dry ice and then stored at -80°C.

2.2.6. Transformation of circular plasmid DNA

1 μ l of plasmid DNA was added to 50 μ l of competent cells, and incubated on ice for 10 min. The cells were then heat shocked at 42°C for 1 min and returned to ice for 2 min. 950 μ l of LB was added to the cells and incubated at 37°C for 10 min. The cells were then centrifuged at 500 x g for 5 min and cells resuspended in 100 μ l of LB. The cells were streaked out onto a LB/agar/ampicillin (50 μ g/ml) plates (for DH5 α and BL21(DE3)pLysS) or LB/agar/ampicillin (50 μ g/ml)/kanamycin (30 μ g/ml) plates (for AD494 cells) with a sterile loop. Plates were incubated at 37°C overnight.

2.2.7. Transformation of ligations

3 μ l of ligation reaction was added to 50 μ l of DH5 α competent cells and incubated on ice for 30 min. The cells were then heat shocked for 90 sec and returned to ice for 2 min. 950 μ l of LB was added to the reaction and incubated at 37°C for 45-60 min. The cells were then plated out onto LB/agar/ampicillin (50 μ g/ml) plates and incubated overnight at 37°C.

2.2.8. pGEMT-easy cloning

PCR products were ligated into the pGEMT-easy vector (Promega, Annandale, Australia) for sequencing. 25 ng of pGEMT-easy vector was ligated with 3.5 μ l of purified PCR product and ligated with units of ligase. Ligations were incubated overnight at 4°C. Ligations were transformed into XL1-blue competent cells and plated onto LB/agar/ampicillin/IPTG/X-gal plates.

LB/agar/ampicillin/IPTG/X-gal plates were prepared by plating 100 μ l of 100mM IPTG and 20 μ l of X-gal (50 μ g/ml) onto the surface of LB/ampicillin (50 μ g/ml) plates and allowed to absorb at 37°C for 30 min prior to use.

2.2.9. Colony PCR

This method was used to screen DH5 α cells transformed with ligations to rapidly identify bacterial colonies containing a cloned cDNA insert in a plasmid. 20 μ l PCR reactions were set up containing either two gene specific primers or a plasmid specific primer (*ie.* T7 promoter primer) and a gene specific primer, Taq polymerase, 0.2 mM dNTP's and 1.5 mM MgCl₂. Corresponding 2 ml cultures of LB/ampicillin (50 μ g/ml) were also prepared. Each colony was touched with a sterile toothpick and the colony was first transferred to the PCR tube and then used to inoculate the 2 ml LB/ampicillin (50 μ g/ml). PCR was performed using conditions which varied according to the properties of primers and template. Reaction products were analysed on ethidium bromide agarose

gels. Positive colonies were cultured overnight and mini-prep DNA extracted the following day. The presence of cDNA of interest was confirmed by restriction digests and DNA sequencing.

2.2.10. RNA extraction and purification

2.2.10.a. Tissue RNA

RNA was extracted from mouse tissue by the method of Chomczynski and Sacchi. The tissue was homogenized in 5 ml of 4 M guanidinium per 1 g of tissue. 1 ml of the homogenate was removed for RNA extraction and the remainder frozen in liquid N₂ and stored at -80°C. To 1 ml of homogenate 7.2 µl of 14.2 M β-mercaptoethanol and 25 µl of 20% (w/v) sarcosyl was added. The RNA was phenol-chloroform extracted by adding the following reagents in order: 100 µl 2M sodium-acetate pH 4.0, 1 ml water-saturated phenol, 200 µl chloroform/isoamyl alcohol 49:1 and the solution was mixed vigorously for 10 sec and incubated on ice for 30 min. The solution was centrifuged for 15 min at 16,000 x g. The upper aqueous phase was transferred to a new microfuge tube and an equal volume of ice cold 100% isopropanol added and precipitated at -20°C 1 hr. The sample was centrifuged at 16,000 x g at 4°C for 15 min. The pellet was resuspended in 300 µl of 4 M guanidinium solution and 300 µl 100% ice cold isopropanol and precipitated at -20°C for 1hr. The sample was centrifuged at 16,000 x g at 4°C for 15 min, and the pellet washed in ice-cold 70% ethanol. The pellet was air dried and resuspended in 25-50 µl of DEPC-dH₂O and stored at -20°C (Chomczynski and Sacchi 1987).

DEPC-treated dH₂O was prepared by adding 1 ml DEPC to 500 ml of dH₂O, shaking the water vigorously for 30 min and then autoclaved. All buffers for RNA isolation were prepared using DEPC-treated dH₂O.

2.2.10.b. Cell line RNA

1-5 x 10⁶ cells were harvested and washed in 1 x PBS. RNA from cell lines was extracted using the Qiagen RNA extraction kit according to manufacturer's instructions. RNA was eluted in DEPC-treated dH₂O and stored at -20°C.

2.2.11. RNA integrity and concentration

RNA was visualized on 1% agarose/ethidium bromide gels and electrophoresed in 1 x TAE. RNA concentration determined by absorbance at 260nm.

RNA concentration(μg/ml) = Absorbance units at 260nm x 40 x dilution factor

2.2.12. Reverse Transcription

Each RNA source was treated with DNase prior to reverse transcription. 2 μg of RNA was treated with 1 μl of DNase, 1 μl 10 x reaction buffer in a final reaction volume of 10 μl. The reaction was incubated at room temperature for 15 min and 1 μl of 25 mM EDTA added. The DNase was heat inactivated at 65°C for 10 min.

For reverse transcription 2 μl of the DNase-treated RNA was combined with 1 μl of 100 μM oligo(dT) and 9 μl of sterile dH₂O and incubated at 72°C for 10 min. The reaction was cooled and the following reaction components were added: 2 μl of dNTP's, 4 μl of 5 x Superscript II reaction buffer, 1 μl of RNasin (40 u/μl). The reaction was incubated at 42°C for 2 min and then 0.5 μl of Superscript II added. A control reaction minus reverse transcriptase was set up in parallel to rule out genomic DNA contamination. The reaction was then incubated at 42°C for 50 min. The reverse transcriptase was heat inactivated at 72°C for 10 min.

2.2.13. RT-PCR

1 μ l of cDNA (RT-reaction) and 1:5 and 1:20 diluted cDNA was added to separate PCR reactions consisting of 40 pmol gene specific primers, 0.2 mM dNTP, 1.5 mM $MgCl_2$, 1 x PCR reaction buffer and Taq polymerase. For the isolation of genes for cloning Vent polymerase was used.

2.2.14. Semi-quantitative RT-PCR

A one-twentieth volume of cDNA was used for PCR. 30 cycles (94°C 30 sec, 47°C 30 sec, 72°C 30 sec) were performed with 0.5 U Taq polymerase, in the presence of 10 μ Ci/100 μ l 32 P-dATP and 40 pmol of sense and antisense primers. The mRNA expression levels of GAPDH were used as a control. GAPDH primers were used in RT-PCR for 15-35 cycles and determined that 30 cycles exponentially amplified cDNA, thus a semi-quantitative estimation of the products was possible. PCR products were separated on 6% polyacrylamide/TBE gels. Gels were fixed in 20% methanol, 7% acetic acid, dried and analysed by autoradiography.

2.2.15. Genomic DNA extraction

Mouse genomic DNA (gDNA) was isolated from C57/Blk6, PT and NMRI liver and the cell line 32D. Single cell suspensions were prepared by homogenising liver in PBS and passaging through a 40 μ m cell strainer. The cells were lysed on ice for 15 min in 0.32 M sucrose, 1% Triton X-100 (v/v), 5 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.5. The lysate was centrifuged at 1,000 x g for 10 min. The pellet was resuspended in 200 μ L of TE buffer, and then 400 μ l of nuclear lysis buffer (0.32 M lithium acetate, 2% lithium dodecyl sulphate (w/v), 10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and mixed gently. An equal volume of phenol/chloroform was added and mixed gently. This was then centrifuged at 1,000 x g, 10 min at 4°C. gDNA was precipitated with 10 μ l of 3 M sodium

acetate, pH 5.5 in 500 μ l 100% ethanol on ice for 30 mins, centrifuged at 16,000 x g for 15 min and supernatant removed. The pellet was washed in 500 μ l of 70% dH₂O, air dried and resuspended in 40-50 μ l of sterile dH₂O.

2.3. PROTEIN BIOCHEMISTRY

2.3.1. Generation of recombinant proteins by in vitro transcription and translation

In vitro transcription/translation experiments were performed using the T_NT[®] T7 Coupled Reticulocyte Lysate System according to manufacturer's instructions (Promega, Annandale, Australia). The reticulocyte lysate was aliquoted into 25 μ l aliquots and stored with all reaction components at -70°C. When required the reticulocyte lysate was rapidly thawed and immediately stored on ice. All reaction components and 0.5-1 μ g of circular plasmid DNA were added to the 25 μ l of reticulocyte lysate on ice. The following reaction components were added: 2 μ l of T_NT[®] reaction buffer, 1 μ l of T_NT[®] T7 RNA polymerase, 1 μ l 1mM Amino Acid Mixture minus Methionine, 2 μ l of [³⁵S] methionine, 1 μ l RNasin[®] Ribonuclease Inhibitor (40 u/ μ l) and the volume made up to 50 μ l with sterile dH₂O. The reaction was incubated for 90 min at 30°C. 2.5 μ l of the reaction of then used to incubate with either 20 or 100 ng of each protease as outlined in Chapter 6. 5 μ l of 5 x SDS-PAGE loading dye was added to each sample and loaded onto 10% SDS-PAGE and electrophoresed at 200 V for 50 min or until the dye front had run off. The gel was soaked in fixing solution (20% methanol, 10% acetic acid) for 30 min followed by soaking in the fluorographic enhancement solution Amplify[™] Reagent (Amersham: New Jersey, USA) for 30 min at room temperature. The gel was dried under vacuum at 80°C for 1 hr and exposed to film overnight at room temperature.

2.3.2. SDS-PAGE

Electrophoretic separation of proteins was carried out by Tris-glycine buffered SDS-PAGE based on the method of Laemmli and using the Mini-PROTEAN electrophoresis system (Bio Rad; California, USA) (Laemmli 1970). Composition of resolving and stacking gel were as follows:

	<i>Resolving gel</i>		<i>Stacking gel</i>
	12.5%	10%	4%
40% acrylamide (37.5:1)	3.1 ml	2.5 ml	1 ml
1.5M Tris-HCl pH 8.9	2.5 ml	2.5 ml	-
0.5 M Tris pH 6.8	-	-	2.5
deionised H ₂ O	4.25 ml	4.85 ml	6.34 ml
10% SDS	100 µl	100 µl	100 µl
10% APS	50 µl	50 µl	50 µl
TEMED	5 µl	5 µl	10 µl

Samples were diluted 1:1 or 1:5 with either 2 x or 5 x SDS-PAGE sample buffer, and heated at 95°C for 5 min and centrifuged to remove condensation. The samples were then loaded into the wells, and electrophoresed in 1 x SDS-PAGE electrophoresis buffer at 200 V for 45-50 min, or until dye front had reached the end of the gel. Pre-stained markers, low molecular weight standards (Bio Rad; California, USA) were also run on the gel. The gel was either stained with Coomassie stain or transferred onto Immobilon-P PVDF membrane (Millipore, North Ryde, Australia) for Western blotting.

2.3.3. Non-denaturing Polyacrylamide Gel Electrophoresis

(Native PAGE)

Electrophoretic separation of proteins on native PAGE gels was carried out by the method of Goldenberg (1989 (a)) on 7% native resolving gels and 4% native stacking gels, using the Mini-PROTEAN electrophoresis system (Bio Rad; California, USA). Composition of resolving and stacking gels were as follows:

	Resolving gel	Stacking gel
	7%	4%
40% acrylamide (37.5:1)	1.75 ml	1 ml
1.5M Tris-HCl pH 8.9	2.5 ml	-
0.5 M Tris pH 6.8	-	2.5
deionised H ₂ O	5.7 ml	6.34 ml
10% APS	50 µl	50 µl
TEMED	5 µl	10 µl

Samples were diluted in 5 x Native-PAGE loading dye. The gels were electrophoresed in 1 x cathode buffer (inner) and 1 x anode buffer (outer) at 10 mA per gel until sample had migrated through the stacking gel, and then at 25 mA per gel until the dye front had run to the end of the gel. To visualise proteins the gel was stained with Coomassie stain.

2.3.4. Transverse Urea Gradient (TUG) Gel Electrophoresis

The TUG gel consisted of 7% polyacrylamide non-denaturing PAGE buffer system and a linear gradient of 0-8M urea was established by using a gradient maker and a pump (Goldenberg 1989 (b); Mast *et al.* 1992).

TUG gel composition

	0M Urea	8M Urea
40% Acrylamide	2 ml	2 ml
Separating buffer	2.5 ml	2.5 ml
Urea	-	4.8 g
dH ₂ O	5.5 ml	to 10 ml

3 ml of each solution was added to each side of the gradient maker, and 3 μ l of TEMED and 12 μ l of 10% APS were added immediately prior to pouring of the gel. Once polymerized, the gel was rotated 90°C and a 1 ml stacker layer (using non-denaturing stacker gel) was poured across the top of the gel. Protein samples (20-100 μ g) were combined with 5 x native-loading dye, and samples loaded across the entire length of the stacking gel. The gels were run using the discontinuous buffer system described for non-denaturing PAGE. The gels were ran at 15 mA per gel whilst the protein migrated through the stacker, and then at 30 mA per gel through the urea gradient until the dye front reached the end of the gel. The gel was then either stained for protein with Coomassie stain or transferred to PVDF for Western blotting.

2.3.5. Western Blotting

Proteins separated by electrophoresis were transferred to Immobilon-P PVDF membrane (Millipore; North Ryde, Australia) using the Mini Protean II Western Blot apparatus from Bio Rad. The membrane was pre-wetted with 100% methanol, prior to "sandwiching" of the membrane and gel between two sheets of filter paper on each side. The tank was filled with 1 x Western transfer buffer/10% methanol and the proteins electroblotted at 80 V for 60 min. After transfer the membrane was blocked with 5% skim milk powder/PBST for a minimum of 30 min at room temperature or 4°C overnight. Primary antibody was added at the dilutions specified in Section 2.3.6 using 0.5% skim milk/PBST and incubated at either room temperature for 1 hr, or 4°C overnight.

The membrane was then washed 3 x 5 min in PBST and secondary antibody added at a dilution of 1:5000 in PBST and incubated for 1 hr. The membrane was washed in PBST 3 x 5min, and developed using the chemiluminescence ECL reagents (Amersham: New Jersey, USA). Protein was detected by exposure to X-ray film from AGFA (Nunawading, Australia).

2.3.6. Antisera

Rabbit polyclonal anti-serpin2A was produced against recombinant serpin2A as described in Morris (2003) *et al.* The antibody was used at a dilution of 1:1000. Mouse monoclonal poly-His-tag antibody was purchased from Qiagen (Melbourne, Australia) and used at a dilution of 1:2000. HRP-conjugated anti-rabbit raised in sheep and HRP-conjugated anti-mouse raised in goat was obtained from Chemicon (California, USA) and used at a dilution of 1:5000.

2.3.7. Protein concentration determination

2.3.7.a. Bio Rad Protein Assay

To determine the concentration of a mixture of proteins the Bio Rad Protein Assay system was used. The dye-based assay binds Coomassie Brilliant Blue G-250 to protein and colour change occurs according to concentration of protein. The absorbance maximum shifts from 465nm to 595nm on protein binding. Protein samples were diluted between 1:1000 to 1:20 in dH₂O to 800 µl and 200 µl of Bio Rad assay reagent was added. Absorbance was measured at 595nm and protein concentration calculated as follows:

$$\text{Protein concentration (mg/ml)} = \text{Absorbance at 595nm} \times \text{dilution factor} \times 31.25$$

where the value 31.25 represents absorbance units/mg based on a standard curve using bovine serum albumin.

2.3.7.b. Absorbance at 280nm

The protein concentration of purified recombinant proteins was measured by taking absorbance readings at 280nm. The extinction co-efficient and the molecular weight of each recombinant protein was determined using the full amino acid sequence and the Peptide Properties Calculator (<http://www.basic.nwu.edu/biotools/proteincalc.html>).

Concentration was calculated as follows:

$$\text{Protein concentration (mg/ml)} = \text{Absorbance (units/cm)} \times \frac{\text{Molecular weight (g/mol)}}{\text{Extinction co-efficient (U.ml.mg}^{-1}\text{.cm}^{-1}\text{)}}$$

2.4. CELL CULTURE

2.4.1. Culture conditions

All cell lines were cultured in 10cm tissue culture dishes in a humidified incubator, at 37°C, 5% CO₂. Cell lines and specific culture conditions used in this investigation are summarised in Table 2.3.

The adherent cells were passaged once cells had reached approximately 80-90% confluence. The cells were washed with 1 x PBS and 1 ml 0.05% trypsin/0.1 mM EDTA/PBS added and cells incubated at 37°C for 5 min or until the cells detached from the dish. To neutralize the trypsin 10 ml media was added and cells resuspended to single cell suspension and passaged 1/10. Non-adherent cells were passaged as they approached 10⁶ cells/ml and diluted to 2 x 10⁵ cells/ml.

Table 2.3: Description of mammalian cell lines and culture conditions

BaF3	Mouse, peripheral blood pro-B cell, BALB/c, non-adherent; cultured in DMEM, 10% heat inactivated FCS (HI-FCS), 10% WEHI-conditioned media, 2mM L-glutamine, 1 x antibiotic/antimycotic solution (1 x Anb) (Sigma)
B16-F10	Mouse epithelial melanoma, C57Blk/6 adherent; cultured in RPMI 1640, 2mM L-glutamine, 10% HI-FCS, 1 x Anb
CH1	Mouse B-cell lymphoma, non-adherent; RPMI 1640, cultured in 5% HI-FCS, 2mM L-glutamine, 1 x Anb
COS-7	African green Monkey kidney, fibroblast, adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb
EMT	Mouse, mammary adenocarcinoma/ sarcoma, adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb
EJ6	Mouse fibroblast, NIH/3T3 cells transformed with DNA from the human EJ bladder carcinoma. Cultured in DMEM, 10% HI-FCS, 2mM L-glutamine, 1 x Anb
FDCP-1	Mouse bone marrow (IL-3 dependent), non-adherent; cultured in DMEM, 10% HI-FCS, 10% WEHI-conditioned media, 2mM L-glutamine, 1 x Anb
FDCP Mix A4	Mouse primitive bone marrow non-adherent: cultured in Iscove's media, 20% horse serum and 10% WEHI-3BD ⁺ conditioned media, 1 x Anb
J558L	Mouse myeloma, lymphoblast BALB/c, adherent; cultured in RPMI 1640, 10% HI-FCS, 2mM L-glutamine, 1 x Anb
M1	Mouse, myeloid leukemia bone marrow from SL mouse, non-adherent; cultured in RPMI 1640, 10% HI-FCS, 2mM L-glutamine, 1 x Anb
NIH-3T3	Mouse fibroblast from Swiss NIH embryo, adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb
Swiss-3T3	Mouse, fibroblast Swiss embryo, adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb
WEHI-3BD⁺	Mouse Balb/c, leukemia, granulocytic myelocytic, non-adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb
32D	Mouse bone marrow (IL-3 dependent), non-adherent; cultured in DMEM 10% HI-FCS, 10% WEHI-3BD ⁺ conditioned media, 2mM L-glutamine, 1 x Anb
4T1.2	Mouse, mammary high metastatic potential, adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb
66cin14	Mouse, mammary intermediate metastatic, adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb
67NR	Mouse, mammary carcinoma low metastatic, adherent; cultured in DMEM 10% HI-FCS, 2mM L-glutamine, 1 x Anb

2.4.2. WEHI-conditioned media

Approximately 80% confluent WEHI-3BD⁺ cells were split 1:20 in 2 x T₇₅ flasks in 30 ml DMEM, 10% FCS, 2mM L-glutamine, 1 x antibiotic/antimycotic solution (1 x Anb)(Sigma). Cells were cultured until approximately 3-4 days or until confluence, and then media harvested and centrifuged at 1,500 x g to pellet the cells. The media was carefully removed to not disrupt the cell pellet, and then filter sterilized through a 0.2µm filter. The WEHI-conditioned media was then either diluted to 10% in culture media or stored at -20°C until required.

2.4.3. Storage of cells

At approximately 80% confluence the cells were split 1:5 into 5 x 10 cm tissue culture dishes. The cells were harvested the following day and centrifuged at 500 x g for 10min. The cell pellet was resuspended at 10⁶ cells/ml in FCS/10% DMSO (v/v) and 1ml cell suspension per vial was aliquoted into cryovials. The cells were frozen at -80°C in an insulated container and then transferred to liquid nitrogen for long-term storage.

To recover cells from long-term storage, the cells were rapidly thawed at 37°C and diluted in 10ml culture media. Adherent cells were plated directly into a 10cm tissue culture dish. The media was replaced with fresh culture the next day. Non-adherent cells were centrifuged at 500 x g for 10 min, and the cell pellet resuspended in fresh culture media and the cells plated into a 10cm tissue culture dish.

2.4.4. Preparation of cells for RNA extraction

At approximately 80% confluence the cells were split 1:5 into 5 x 10cm tissue culture dishes. The cells were harvested the following day for RNA preparation.

Declaration for thesis Chapter 3

In the case of Chapter 3, the following authors have made contributions to the manuscript:

Anita J Horvath

Contribution: 85%

All methods in the paper, including: the design of primers, purification of RNA and gDNA and PCR. Wrote the first draft and worked on subsequent revisions.

Sharon L Forsyth

Contribution: 5%

Background on mouse genome by database analysis. Proofreading of manuscript.

Paul B Coughlin


Contribution: 10%

Supervision: General advice, ideas and facilitation of manuscript.

The undersigned hereby certify that:

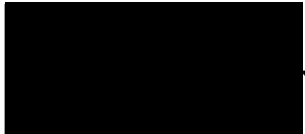
- (1) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
- (2) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
- (3) there are no other authors of the publication according to these criteria;
- (4) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
- (5) the original data are stored at the following location(s) and will be held for at least five years at Department of Medicine, Alfred Medical Research Precinct, Monash University

Author Name: **Anita J Horvath**

Signature: .. 

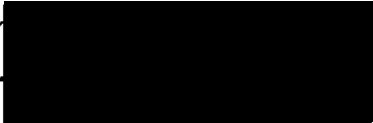
Date:19/4/04.....

Author Name: **Sharon L. Forsyth**

Signature: 

Date:19/4/04.....

Author Name: **Paul B. Coughlin**

Signature: 

Date:19/4/04.....

CHAPTER 3

Gene expression studies of the *serpina3* and *serpina1* locus

Presented as:

Horvath, A.J., Forsyth, S.L. and Coughlin, P.B (2004). Expression patterns of murine antichymotrypsin-like genes reflect evolutionary divergence at the *serpina3* locus. *J Mol Evol* (in press)

Expression patterns of murine antichymotrypsin-like genes reflect evolutionary divergence at the *serpina3* locus.

Anita J. Horvath, Sharon L. Forsyth and Paul B. Coughlin

Institution: Australian Centre for Blood Diseases, Monash University, Melbourne, Australia

Corresponding author: Paul Coughlin
Australian Centre for Blood Diseases
Alfred Medical Research Precinct
Monash University
Commercial Road, Prahran,
AUSTRALIA, 3181

Fax: +61-3-98950332

Email: [REDACTED]

Running title: Expression pattern of murine 'a3' serpins

Abstract

Members of the serpin (serine protease inhibitor) superfamily of genes are well represented in both human and murine genomes. In many cases it is possible to identify a definite ortholog on the basis of sequence similarity and by examining the surrounding genes at syntenic loci. We have recently examined the murine serpin locus at 12F1 and observed that the single human $\alpha 1$ -antichymotrypsin gene is represented by 14 paralogs. It is also known that the single human $\alpha 1$ -antitrypsin gene has 5 paralogs in the mouse. The forces driving this gene multiplication are unknown and there are no data describing the function of the various serpin gene products at the $\alpha 1$ -antichymotrypsin multigene locus. Examination of the predicted amino acid sequences shows that the serpins are likely to be functional protease inhibitors but with differing target protease specificities. In order to begin to address the question of the problem presented by the murine $\alpha 1$ -antichymotrypsins we have used RT-PCR to examine the expression pattern of these serpin genes. Our data show that the divergent reactive center loop sequence, and predictably variable target protease specificity, is reflected in tissue-specific expression for many of the family members. These observations add weight to the hypothesis that the antichymotrypsin-like serpins have an evolutionary importance which has led to their expansion and diversification in multiple species.

Keywords

Serpina1, *serpina3*, protease, antichymotrypsin, antitrypsin.

Introduction

The predominant protease inhibitors present in human plasma are α_1 -antitrypsin and α_1 -antichymotrypsin. Both are acute phase reactants produced in the liver and play roles in regulation of granular proteases released by leukocytes during the inflammatory response. The importance of this function is highlighted by the occurrence of destructive lung disease in patients with deficiency of α_1 -antitrypsin or α_1 -antichymotrypsin. The expression of α_1 -antichymotrypsin in other tissues such as brain, lung, breast and prostate has also raised the possibility of other important roles for this serpin (Higashiyama et al. 1995; Kanemaru et al. 1996; Laursen and Lykkesfeldt 1992; Wu et al. 1998).

The human genome contains single copies of α_1 -antitrypsin and α_1 -antichymotrypsin genes at chromosome 14q32.1. The corresponding *serpina1* and *serpina3* loci on mouse chromosome 12F1 are dramatically expanded. The *serpina1* locus (equivalent to human α_1 -antitrypsin) contains up to 5 closely related genes (Barbour et al. 2002; Borriello and Krauter 1991). The first thorough examination of the murine *serpina3* locus (equivalent to human α_1 -antichymotrypsin) was performed by Inglis who identified 10 closely related genes on two cosmid contigs spanning 220 kb (Inglis and Hill 1991). Sequencing of the terminal exons demonstrated hypervariability of the predicted reactive center loop (RCL) region hinting at the possibility of diverse functions for these serpins. Subsequent examination of the mouse genome (Mouse Genome Sequencing Consortium, http://www.ensembl.org/Mus_musculus/) has shown 14 *serpina3* genes of which 13 are predicted to encode full length, functional serpins (Forsyth et al. 2003). In contrast, 7 other serpin genes present in the same region of human chromosome 14q32.1 have only a single orthologous gene at mouse chromosome 12F1.

Although it is known that murine liver expresses antichymotrypsin-like serpins (often referred to as mouse contrapsin or Spi-2) and that murine plasma contains antichymotrypsin-like activity there is little data to indicate which genes are responsible for this (Hill et al. 1984; Inglis et al. 1991). There is also minimal published information on the expression of the 'a3' serpins in other tissues. The murine 'a3' and 'a1' serpins present a fascinating challenge because of the potential for functional diversity mediated by their hypervariable reactive center loops. Expression studies on these gene clusters using Northern blot or immunological reagents are complicated by their high degree of conservation at nucleotide and amino acid levels outside of the RCL region.

In order to examine the expression of the 'a3' serpins and begin to address their functions and understand the forces driving their duplication we have developed a gene-specific PCR method. This has allowed us to assess individual gene expression across a range of tissues and cell lines.

Materials and Methods

Murine tissues and cell lines

Tissue for RNA extraction was isolated from 7 week NMRI mice. The cell lines NIH-3T3, Swiss-3T3 (fibroblast), EJ6 (transformed fibroblast), WEHI-3BD⁺ (myelomonocytic) and the mammary carcinoma cell lines EMT6.5, 4T1.2 (highly metastatic), 66cl4 (intermediate metastatic) and 67NR (non-metastatic), were cultured in DMEM (Gibco) media supplemented with 10% heat-inactivated fetal calf-serum, 2mmol/L L-glutamine (Gibco) and 1% antibiotic/antimycotic (Sigma). The hematopoietic cell lines FDCP-1 (myeloid), 32D (myeloid) and BAF-3 (Pro-B-cell) were cultured as above with the addition of 10% WEHI-3BD⁺ conditioned media as a source of IL-3. The remaining hematopoietic cell lines M1 (myeloid), CH1 (B-cell), J558L (myeloma) and B16-F10 (melanoma) were cultured in RPMI 1640 (Gibco) and supplemented with 10% fetal calf serum and 1% antibiotic/antimycotic (Sigma). Cultures were maintained in a 5% CO₂, 95% air mixture at 37°C.

Mouse genomic DNA extraction

Mouse genomic DNA was isolated from C57/Black6, PT and NMRI liver and the cell line 32D. Single cell suspensions were prepared by homogenizing liver in PBS and passaging through a 40 µm cell strainer. Cells were lysed on ice for 15 minutes in 0.32M sucrose, 1% Triton X-100 v/v, 5mM MgCl₂, 10mM Tris-HCl, pH 7.5. Lysate was centrifuged at 1000xg for 10 minutes. The pellet was resuspended in 200µL of TE buffer, and then 400µL of nuclear lysis buffer (0.32M lithium acetate, 2% w/v lithium dodecyl sulphate, 10mM Tris-HCl pH 8.0, 1mM EDTA, and mixed gently. An equal volume of phenol/chloroform (1:1) was added and mixed gently. This was then centrifuged at 1000xg for 10 minutes at 4°C. DNA was precipitated by adding a 1/10 volume of 3M sodium acetate, pH 5.5 and 2 volumes of ethanol.

RNA Extraction

Total RNA was extracted from cell lines using Qiagen RNeasy mini-kit and from murine tissue using the acid-guanidinium/phenol/chloroform method (Chomczynski and Sacchi 1987). 2µg of total RNA was treated with 1U DNase (Invitrogen) to remove genomic DNA contamination prior to reverse transcription.

Reverse transcription-PCR.

2µL of each DNase-treated RNA was reverse transcribed using Superscript II reverse transcriptase (Invitrogen), 100 pmol oligo(dT) and RNase inhibitor (Promega) in a 20 µL reaction. A control reaction minus reverse transcriptase was set up in parallel to rule out genomic DNA contamination. Each 20 µL PCR reaction was made up of 1 µL of reverse transcription reaction (diluted 1/20), 8 pmol of each gene specific primer pair (see Fig. 1), 200 µM dNTP, 10 mM Tris-HCl, pH 8.8, 1.5 mM MgCl₂, 50 mM KCl, 0.1% Triton X-

100, 1.0U Taq polymerase (Geneworks) and 2 μ Ci α^{32} P-dATP(Perkin Elmer). 30 cycles (94°C 30sec, 47°C 30sec, 72°C 30sec) were performed. The primer sequences and positions are shown in 1. PCR for GAPDH was included as an internal control for each template analysed (primers: forward, ACGGATTGCGCGTA and reverse, ACGTCAGATCCACGA). PCR products were separated on 6% polyacrylamide/TBE gels. Gels were fixed in 20% methanol/7% acetic acid (v/v), dried and exposed to X-ray film (Agfa) overnight at -80°C.

Genomic PCR

PCR was carried out as for RT-PCR, using the *serpina3* primers with mouse genomic DNA as template. Products were analysed on 2% ethidium bromide agarose gels. The products from each serpin primer pair were subcloned into pGEMT-easy vector (Promega) and nucleotide sequencing performed using Big Dye Terminator (Applied Biosystems) mix to confirm that the correct serpins were amplified with each primer pair.

Results

Establishment of a semi-quantitative and specific RT-PCR assay.

In order to examine the expression of *serpina3* genes in a variety of tissues and cell lines we needed to use a method which was at least semi-quantitative and which allowed us to distinguish between closely related sequences. As Northern blotting is problematic because of the high degree of nucleotide sequence conservation outside the RCL we used gene-specific RT-PCR. To improve the quantitative accuracy and sensitivity of this method we added α^{32} P-dATP to the PCR so that the isotope would be incorporated into the product. We found incremental intensity of radioactive product between 25 and 35 cycles (data not shown). Therefore we selected 30 cycles as the standard for RT-PCR and used the intensity of GAPDH product as an internal control for quantity of template to allow comparison between various tissues and cell lines.

In order to make the PCR gene specific we designed a series of primer pairs in which the forward primer was based on the non-conserved RCL while the design of the reverse primer exploited minor sequence variations at the 3' end of the serpin open reading frame (Fig. 1). As all the PCR products corresponded to a portion of a single exon we were able to assess the specificity of the process by amplifying from genomic DNA. When this was done an appropriate sized product was seen on electrophoretic gels and authenticity was confirmed by nucleotide sequencing.

Expression pattern of $\alpha 3$ serpins in murine tissues.

In *Homo sapiens* $\alpha 1$ -antitrypsin (*SERPINA1*) and $\alpha 1$ -antichymotrypsin (*SERPINA3*) are expressed at high level in the liver. Fig. 2 shows that representatives of the $\alpha 1$ and $\alpha 3$

serpins are also strongly expressed in the murine liver. Of the 13 murine $\alpha 3$ genes, 3 are expressed prominently in the liver, namely *EB22.4* (*serpina3n*), *MMCM2* (*serpina3k*) and *3E46* (*serpina3m*). When the reactive center loops of the liver expressed murine $\alpha 3$ serpins are compared with human $\alpha 1$ -antichymotrypsin (Table 1) they appear to have markedly different amino acid sequences and predicted protease inhibitory activities. Human $\alpha 1$ -antichymotrypsin (SERPINA3) possesses a P_1 - P_1' of Leu-Ser making it an effective inhibitor of the chymotrypsin-like protease cathepsin G. In contrast, *EB22.4* (*serpina3n*) has a reactive center sequence with a predicted P_1 - P_1' of Met-Ser, reminiscent of human $\alpha 1$ -antitrypsin. However, a consistent feature of human $\alpha 1$ -antitrypsin and the murine $\alpha 1$ -antitrypsins is a P_3 '- P_4 ' PP motif which is absent from *EB22.4* (*serpina3n*) and this difference may affect the kinetics of interaction with target proteases. The other $\alpha 3$ serpins expressed strongly in the liver (*serpina3k* and *serpina3m*) have reactive center P_1 - P_1' residues of R-K and R-S respectively, suggesting preferential inhibition of trypsin-like proteases. Clearly the prediction of target protease specificity is at best approximate and more precise interpretations await biochemical and kinetic data.

Of all the murine $\alpha 3$ serpins, *serpina3n* (*EB22.4*) has the widest distribution with high level expression in brain, testis, lung, thymus and spleen and low level expression in bone marrow, skeletal muscle and kidney (Fig. 2A). The grouping of lung, thymus and splenic tissue is consistent with expression in lymphoid tissue and has been reported for human $\alpha 1$ -antichymotrypsin (Krugliak et al. 1986). The actual cells which express *serpina3n* and the role it plays in these tissues awaits further investigation.

Serpina3n (*EB22.4*) is the only member demonstrating significant expression in the brain (Fig. 2A). Similarly, human antichymotrypsin is expressed by astrocytes in the central nervous system and may play a permissive role in the progression of Alzheimer's Disease and cerebral amyloid angiopathy (Abraham 2001; Kanemaru et al. 1996; Yamada 2002). *In vitro* evidence demonstrates that $\alpha 1$ -antichymotrypsin binds the amyloid β_{1-42} peptide by insertion into the serpin A β -sheet (Janciauskiene et al. 1998). Also, overexpression of human $\alpha 1$ -antichymotrypsin in a murine model of Alzheimer's disease increases the rate of disease progression (Licastro et al. 1999; Mucke et al. 2000). The expression of an endogenous ' $\alpha 3$ ' serpin is also increased in the apolipoprotein E (apoE) knockout mouse model of Alzheimer's disease and this effect was reversed when apoE3 was re-expressed (Licastro et al. 1999). Our data demonstrate that *serpina3n*, being the only member of the $\alpha 3$ and $\alpha 1$ groups expressed in murine brain under resting conditions, is likely to be the functional ortholog of human antichymotrypsin in the brain.

Serpina3m (*3E46*) is expressed in liver and testis (Fig. 2). Its reactive center (P_5 - P_3' IFGFRSRR) bears a striking resemblance to human Protein C inhibitor (SERPINA5) (P_5 - P_3' IFTFRSAR) although the overall level of sequence identity is relatively low at 47%. Murine Protein C inhibitor (*serpina5*) (P_5 - P_3' IFTFRSAR) shares this conserved bait region strongly suggesting a similar target protease specificity and possible overlap of function.

Serpina3m (*3E46*) and *serpina3n* (*EB22.4*) are expressed in both liver and testis (Fig. 2A) while *serpina3a* (*Unknown 1*) is primarily expressed in the testis (Fig. 2B). Testicular

expression of human α 1-antichymotrypsin has not previously been noted although several other members of the human serpin family including PI-9 (SERPINB9), CBG (SERPINA6), PEDF (SERPINF1), SERPINB12 and Protein C inhibitor (SERPINA5) are produced in this tissue (Askew et al. 2001; Bladergroen et al. 2001; Hammond et al. 1987; Hirst et al. 2001; Uhrin et al. 2000). The known role of proteases in fertilization and implantation together with the abundance of proteases and serpins expressed in the testis and prostate suggests that control of proteolysis is important in reproductive biology (Hirst et al. 2001; Mikolajczyk et al. 1999; Stephan et al. 2002). Further evidence for the significance of serpins in the testis is highlighted by the finding of infertility related to defective spermatogenesis in the Protein C inhibitor knockout mouse (Uhrin et al. 2000).

Serpina3k (MMCM2), which possesses a predicted P₁-P₁' of RK, is expressed exclusively in the liver (Fig. 2A). The protease specificity of this serpin is difficult to predict as it possesses an unusual di-basic RK sequence at its reactive center. Furthermore, the RCL is shortened by 2 amino acids with respect to the other a3 members and this has been shown to affect inhibitory activity and susceptibility to polymerization in serpins (Bottomley and Chang 1997; Bottomley and Stone 1998). The expression by murine liver of 3 *serpina3* genes plus at least one *serpinal* gene potentially provides a much broader protease inhibitor spectrum in plasma compared to man. The reason for this is unclear as there is no evidence that mice possess a greater range of leukocyte proteases than humans. Furthermore, there is no direct evidence to support the proposal that serpins are involved in defence against parasites.

Serpina3h (6C28), like *serpina3n* (EB22.4), is expressed in murine thymus, spleen and lung (Fig. 2A) again suggesting a role in immune function. However, in contrast to *serpina3n*, *serpina3h* is predicted to be intracellular and the presence of a CC motif at the reactive center is consistent with this localization. Low level expression of *serpina3g* (2A2, *serpin2A*) is evident in bone marrow, spleen, thymus and lung as previously reported (Hampson et al. 1997). *Serpina3g* is known to be dramatically upregulated upon macrophage stimulation or T cell activation and this may account for its modest intensity in unstimulated spleen and thymus in our RT-PCR screen (Hamerman et al. 2002; Hampson et al. 1997). Similarly, only low level expression of *serpina3g* is seen in bone marrow but published data indicate that the gene is strongly activated in hematopoietic precursors which make up a small proportion of total cell mass (Hampson et al. 1997; Terskikh et al. 2001).

We also examined the expression pattern of the murine *serpinal* genes using PCR primers which were designed to anneal to all 5 known members (Fig. 2) (Borriello and Krauter 1991). As expected strong expression was noted in the liver but only low levels were seen in other sites. This is generally consistent with the pattern seen in humans but does not exclude the possibility of *serpinal* expression in sub-populations of cells within tissues.

Expression pattern of $\alpha 3$ serpins in murine cell lines.

In order to try to address the possibility of isolated gene expression within particular cell types we studied a panel of cell lines for expression of the $\alpha 3$ serpins. Fig. 3A shows the results of analysis of hematopoietic cell lines in which *serpina3g* (2A2) is predominant. Low levels of *serpina3f* (2A1) were also observed. Previous work has demonstrated *serpina3g* (2A2, *serpin* 2A) is one of the most abundant transcripts in hematopoietic stem cells and is down-regulated upon induction of differentiation (Hampson et al. 1997; Terskikh et al. 2001). Constitutive, low level expression of *serpina3g* by retroviral transduction of FDCP Mix-A4 cells was associated with delayed differentiation and increased clonogenicity (Hampson et al. 1997). Using gene specific primers we were able to assess whether other members of the $\alpha 3$ family were also expressed in these cells. Fig. 3A confirms expression of *serpina3g* in FDCP Mix-A4 cells. Similar levels of expression were also noted in the IL-3 dependent cell lines FDCP-1, BaF3 and 32D while the IL-3 independent lines M1 and CH1 showed only low level *serpina3g* expression. The 32D cells and BaF3 also showed low level expression of *serpina3f* which is closely related to *serpina3g* with a similar reactive center P₁-P₁' of Cys-Cys. *Serpina3f* and *serpina3g* lack secretion signal peptides and intracellular localization has been confirmed for *serpina3g* (Morris et al. 2003).

Fig. 3B shows results from fibroblast cell lines NIH 3T3, Swiss 3T3 and EJ6. Both *serpina3n* and *serpina3h* were expressed in Swiss 3T3 and NIH 3T3 cell lines although the absolute level was higher in Swiss 3T3 cells. In contrast, the transformed fibroblast cell line EJ6, which is derived from NIH 3T3 cells, expressed *serpina3n* almost exclusively, with minimal *serpina3h* expression. The role of these serpins in fibroblasts is obscure however Whitehead *et al* (Whitehead et al. 1995) previously showed that the expression of an antisense *serpina3* cDNA in NIH 3T3 cells induced transformation. It may be that protease-inhibitor balance plays a role in modulating the interaction between cell surface proteins and extra-cellular matrix.

In addition to the hematopoietic and fibroblast cell lines investigated we have also tested mammary (4T1.2, 66d4, 67NR and EMT6.5) and melanoma (B16-F10) cell lines but did not detect any significant expression of $\alpha 3$ serpins (data not shown). This was a surprising as $\alpha 1$ -antichymotrypsin is known to be expressed by human mammary cell lines (Laursen and Lykkesfeldt 1992).

We were aware that the number of antitrypsin genes vary between mouse species and it has even been shown that gene number is different between strains of *Mus domesticus* (Goodwin et al. 1997). In order to address this possibility for the *serpina3* locus we performed gene specific PCR using genomic DNA from four different murine laboratory strains. Fig. 4 shows that all 13 of the *serpina3* genes predicted to encode full length serpin cDNAs are present in all strains examined. We also performed PCR using cDNA from liver of C57/Bl6 and PT mice to assess strain dependence of gene expression. Fig. 5 shows an identical pattern of *serpina3* expression in C57/Bl6 and PT to that seen in the NMRI mice.

Discussion

The simplest hypothesis to explain the spectacular multiplication of murine *serpinal* and *serpina3* genes is that it has been driven by positive selection. If, as seems likely, it occurred through duplication and subsequent gene conversion events, then it follows that there was some advantage in preserving the basic serpin structure while expanding the range of potential protease targets through hypervariability in the specificity-determining RCL region. The expansion of the *serpina3* genes is clearly not restricted to the mouse. There are six rat antichymotrypsin-like serpins (contrapsins) identifiable in the rat genome (<http://www.ncbi.nlm.nih.gov/genome/guide/rat/>) (Table 2). Strikingly the rat serpin Spin2c has a RCL which is identical to murine EB22.4 (*serpina3n*) from P₁₅ to P₄'. The remaining five rat a3 serpins are divergent and bear little resemblance to the murine a3 serpins in the RCL domain. Similarly, multiple bovine antichymotrypsins have been identified and three of these cloned and shown to have variable reactive center loops and tissue specific expression (Hwang et al. 1994; Hwang et al. 1995). Three distinct porcine antichymotrypsins have also been reported (Stratil et al. 1995). It therefore seems likely that *Homo sapiens* with only one *antichymotrypsin* gene is the exception rather than the rule. The formal possibility exists that the single human *SERPINA3* is a result of gene loss, however, there is no evidence of related pseudogenes or gene fragments close by on chromosome 14. Examination of the 'A' clade cluster in other primates may also help to address this question.

Without evolutionary advantage driving this process most of these genes would be expected to fall into disuse and degenerate into pseudogenes. On the contrary, examination of the 14 murine *serpina3* genes shows only one (*serpina3e*, 2B2(b)) which is truncated at the 5' end and is therefore almost certainly non-functional. The remaining 13 genes encode serpins with full open reading frames and reactive center loop proximal hinge sequences indicative of inhibitory activity (Table 1). Five of the 13 genes have atypical proximal hinge sequences (*a3a*, *a3b*, *a3d*, *a3j* and *a3l*) with a P₁₁ charged residue (Glu in 4 cases and Arg in 1). One of the rat serpins (CPI-26) also has a P₁₁ Glu. The significance of this is uncertain as it is not present in other serpins except the non-inhibitory serpin, Hsp47. This comparison is not really informative as Hsp47 has other proximal hinge variations which abolish inhibitory function.

Four serpin genes (*a3f*, *a3g*, *a3h* and *a3i*) lack predicted N-terminal secretion signal peptides suggesting that they would be produced as intracellular proteins similar to members of the ov-serpin subfamily. In keeping with this observation they possess cysteine residues at the reactive site scissile bond which in other serpins confers functional sensitivity to oxidation. Our own experimental data confirms this prediction for *serpina3g* (Morris et al. 2003). The presence of intracellular serpins within the a3 murine cluster is most unexpected and raises the possibility of convergent evolution in which genes related to antichymotrypsin have lost the secretion signal peptide and acquired oxidisable residues in the RCL akin to the human clade B serpins. The recent demonstration that *serpina3g* (*serpin2A*) plays a key role in NF- κ B mediated control of TNF activity highlights the potential importance of a3 intracellular serpins in murine biology (Liu et al. 2003).

Further support for the functional importance of the expanded *serpina3* and *serpina1* genes comes from examination of the other 'a' clade serpins in the surrounding region of murine chromosome 12 F1. All of the remaining seven serpin genes in this locus are present as a single copy orthologous to their corresponding genes at the syntenic human chromosome 14q32.1 (Fig. 6). Whatever force has driven the *serpina3* and *serpina1* expansion was clearly a specific process not involving the surrounding serpin genes.

It should be noted that some members of the intracellular serpin family display a similar expansion in the mouse. Human PI-6, PI-9 and MNEI are represented in the mouse by a total of 15 genes in the mouse (Kaiserman et al. 2002). We are not aware of any unique structural features of the serpins or any flanking repeats which would increase the likelihood of gene duplication. One common theme linking the human intracellular serpins with the SERPINA1 and SERPINA3 proteins is that they play important roles in controlling leukocyte proteases. PI6 and MNEI are efficient intracellular inhibitors of elastase, cathepsin G and proteinase 3 while PI-9 inhibits granzyme B. SERPINA1 and SERPINA3 regulate the same leukocyte enzymes in the extracellular environment. These observations do not, however, help to explain the murine serpin expansions as there is no corresponding multiplication of the proteases.

The data presented in this paper gives the first comprehensive description of the expression pattern of the murine antichymotrypsin-like serpins. The observed hypervariability of the reactive center loop of the murine a3 serpins, in the context of striking conservation of the remaining structural elements of the proteins, argues strongly in favour of evolutionarily advantageous functional diversity. The fact that expansion of the a3 serpin locus has occurred in other species gives these genes a special status. Our observations on the expression pattern of these genes add further support to this concept. We are currently producing recombinant murine *serpina3* proteins in order to examine their biochemical and biophysical properties and target protease specificity. Ultimately we expect this work to provide insights into murine and human serpins whose complete repertoire of functions is only beginning to be understood.

Abbreviations

GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RCL, reactive center loop

Figure legends

Figure 1. Position of PCR primers for *serpina3* genes.

The nucleotide sequences of *serpina3* genes around the reactive center loop (left side) and close to the stop codon (right side) are shown. Sites for 5' and 3' primers for the *serpina3* genes are indicated by arrows below each sequence. Primer design exploited differences in nucleotide sequence between *serpina3* genes in order to make the PCR gene specific. Individual 5' gene specific primers were used as shown. Some 3' primers were used for two genes (*serpina3f* and *serpina3g*, *serpina3h* and *serpina3i*, *serpina3k* and *serpina3m*). Note that only part of the *serpina3a* and *serpina3l* primers are shown. The full sequences are (*serpina3a*) AGATGTCATCACAATAGCCCG and (*serpina3l*) CTAGTTTTCTAAAGGATCCAC. The numbering above the sequences refers to *serpina3n* (EB22.4) cDNA sequence (Inglis et al. 1991). *Serpina1* primers were designed to amplify all members of the murine antitrypsin family equally (5' GAAGCTGCAGCAGCTACAGTC and 3' TGTGGGATCTACCACTTTTCC).

Figure 2. RT-PCR of from murine tissues.

Polyacrylamide gel electrophoresis of ^{32}P labelled RT-PCR products from a panel of murine tissues are shown. Panels A and B represent different experiments in which distinct sets of reverse transcription reactions for the various tissues were used. The tissue sources of RNA for each set of RT-PCRs are shown on the left (Panel A) and right (Panel B). The mouse gene nomenclature names are shown across the top of the figure while the common names are shown across the bottom. For the sake of clarity molecular weight markers are not shown however the PCR products shown are correct size based on the known cDNA sequences and position of PCR primers. In some cases (*a3n* and *a3k*) a secondary band is evident with the predicted product size being the lower, dominant band. The significance of the upper band is uncertain as primers are sited within a single exon making alternatively spliced templates unlikely. It should be noted that there is some variation in the intensity of the GAPDH product and this should be taken into consideration when assessing serpin PCR products. Some RT templates (heart, kidney and skeletal muscle) yielded smears in some lanes rather than clearly defined bands and we regarded this as negative for specific product.

Figure 3 RT-PCR from (A) murine hematopoietic cell lines and (B) murine fibroblast cell lines.

Polyacrylamide gel electrophoresis of ^{32}P labelled RT-PCR products from murine hematopoietic cell lines and murine fibroblast cell lines are shown. The tissue sources of RNA for each set of RT-PCRs are shown on the left (Panel A) and right (Panel B). The mouse gene nomenclature names are shown across the top of the figure while the common names are shown across the bottom. For the sake of clarity molecular weight markers are not shown however the PCR products shown are correct size based on the known cDNA sequences and position of PCR primers. As indicated for Fig. 2, some primer pairs (*a3g* and *a3n*) yielded secondary higher molecular weight products whose significance is uncertain. The lower, dominant bands correspond to the predicted molecular weights.

Figure 4. PCR from genomic DNA derived from different mouse strains.

Agarose gel electrophoresis of RT-PCR products from the liver of different strains of mice is shown. Products were detected by ethidium bromide staining and UV transillumination. The strain of mouse from which RNA was derived is shown on the left. Mouse gene nomenclature names are shown across the top and common names are shown across the bottom. For the sake of clarity molecular weight markers are not shown however the PCR products shown are correct size based on the known cDNA sequences and position of PCR primers. RT-PCR products from NMRI mice were sub-cloned and nucleotide sequencing performed in order to ensure authenticity.

Figure 5. RT-PCR from liver tissue derived from mouse strains C57/Bl6 and PT.

Polyacrylamide gel electrophoresis of ^{32}P labelled RT-PCR products from the livers of C57/Bl6 and PT mice. The mouse gene nomenclature names are shown across the top of the figure while the common names are shown across the bottom. Strong expression of *serpina3m*, *serpina3n* and *serpina3k* is demonstrated which is identical to the pattern of expression seen in the liver of strain NMRI (Fig. 2).

Figure 6. Schematic representation of mouse, rat and human A clade clusters.

The A clade clusters for mouse, rat and human at chromosome loci 12F1, 6q32 and 14q32.1, respectively, are shown. Individual genes are represented by grey bars with the gene symbol (or accession number) shown above. The mouse and rat a1 and a3 expansions are shown above the main chromosomal representations. It can be seen that, with the exception of a1, a3 and HongrES1, all genes are conserved in single copies. HongrES1 (Hu et al. 2002), an epididymis specific serpins appears to be absent from the human A clade cluster. *SERPINA4* is represented by a pseudogene in the mouse. The common names for the serpins are as follows: A10, Protein Z-dependent protease inhibitor; A6, cortisol binding globulin; A2, α 1-antitrypsin-related protein; A1, α 1-antitrypsin; A11, unnamed; A9, centerin; A12, OL-64 or visceral adipose specific serpin; A4, kallistatin; A5, protein C inhibitor; A3, α 1-antichymotrypsin.

Table 1: Reactive centre loop sequences of the murine *serpina3* genes and tissue expression.

The sequences shown in this table are derived from the publicly available mouse genome sequence (<http://www.ncbi.nlm.nih.gov/genome/guide/mouse/>). The P_1/P_1' site is shown in bold and the predicted scissile bond is indicated by the symbol ∇ . Human α 1-antichymotrypsin and α 1-antitrypsin and murine antitrypsin (*serpina1a*) are shown below for comparison. Tissue expression of the murine a3 serpins is based on the results presented in Fig. 2 and Fig. 3. Expression levels are: +, low level; ++, intermediate and +++, high. Note that results from haematopoietic cell lines are summarized under "Blood cells".

Acknowledgements

PC is a Wellcome Trust Senior Research Fellow. We are indebted to Dr Robert Medcalf, Dr Hong Yu and Ms Melinda Missen for helpful discussions and technical assistance.

References

- Abraham CR (2001) Reactive astrocytes and alpha1-antichymotrypsin in Alzheimer's disease. *Neurobiol Aging* 22:931-6.
- Askew YS, Pak SC, Luke CJ, Askew DJ, Cataltepe S, Mills DR, Kato H, Lehoczy J, Dewar K, Birren B, Silverman GA (2001) SERPINB12 is a novel member of the human ov-serpin family that is widely expressed and inhibits trypsin-like serine proteinases. *J Biol Chem* 276:49320-30
- Barbour KW, Goodwin RL, Guillonneau F, Wang Y, Baumann H, Berger FG (2002) Functional diversification during evolution of the murine alpha(1)-proteinase inhibitor family: Role of the hypervariable reactive center loop. *Mol Biol Evol* 19:718-27.
- Bladergroen BA, Strik MC, Bovenschen N, van Berkum O, Scheffer GL, Meijer CJ, Hack CE, Kummer JA (2001) The granzyme B inhibitor, protease inhibitor 9, is mainly expressed by dendritic cells and at immune-privileged sites. *J Immunol* 166:3218-25
- Borriello F, Krauter KS (1991) Multiple murine alpha 1-protease inhibitor genes show unusual evolutionary divergence. *Proc Natl Acad Sci U S A* 88:9417-21.
- Bottomley SP, Chang W (1997) The effects of reactive centre loop length upon serpin polymerisation. *Biochem Biophys Res Comm* 241:264-269
- Bottomley SP, Stone SR (1998) Protein engineering of chimeric serpins: an investigation into effects of the serpin scaffold and reactive centre loop length. *Protein Eng* 11:1243-7
- Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162:156-159
- Forsyth S, Horvath A, Coughlin P (2003) A review and comparison of the murine alpha(1)-antitrypsin and alpha(1)-antichymotrypsin multigene clusters with the human clade A serpins. *Genomics* 81:336-45
- Goodwin RL, Barbour KW, Berger FG (1997) Expression of the alpha 1-proteinase inhibitor gene family during evolution of the genus *Mus*. *Mol Biol Evol* 14:420-7.
- Hamerman JA, Hayashi F, Schroeder LA, Gygi SP, Haas AL, Hampson L, Coughlin P, Aebersold R, Aderem A (2002) Serpin 2a is induced in activated macrophages and conjugates to a ubiquitin homolog. *J Immunol* 168:2415-23
- Hammond GL, Smith CL, Goping IS, Underhill DA, Harley MJ, Reventos J, Musto NA, Gunsalus GL, Bardin CW (1987) Primary structure of human corticosteroid binding globulin, deduced from hepatic and pulmonary cDNAs, exhibits homology with serine protease inhibitors. *Proc Natl Acad Sci U S A* 84:5153-7
- Hampson IN, Hampson L, Pinkoski M, Cross M, Heyworth CM, Bleackley RC, Atkinson E, Dexter TM (1997) Identification of a serpin specifically expressed in multipotent and bipotent hematopoietic progenitor cells and in activated T cells. *Blood* 89:108-118

- Higashiyama M, Doi O, Yokouchi H, Kodama K, Nakamori S, Tateishi R (1995) Alpha-1-antichymotrypsin expression in lung adenocarcinoma and its possible association with tumor progression. *Cancer* 76:1368-76
- Hill RE, Shaw PH, Boyd PA, Baumann H, Hastie ND (1984) Plasma protease inhibitors in mouse and man: divergence within the reactive centre regions. *Nature* 311:175-7
- Hirst CE, Buzza MS, Sutton VR, Trapani JA, Loveland KL, Bird PI (2001) Perforin-independent expression of granzyme B and proteinase inhibitor 9 in human testis and placenta suggests a role for granzyme B-mediated proteolysis in reproduction. *Mol Hum Reprod* 7:1133-42
- Hu ZH, Liu Q, Shang Q, Zheng M, Yang J, Zhang YL (2002) Identification and characterization of a new member of serpin family- HongrES1 in rat epididymis. *Cell Res* 12:407-10
- Hwang SR, Kohn AB, Hook VY (1994) Molecular cloning reveals isoforms of bovine alpha 1-antichymotrypsin. *Proc Natl Acad Sci U S A* 91:9579-83
- Hwang SR, Kohn AB, Hook VY (1995) Unique reactive site domains of neuroendocrine isoforms of alpha 1-antichymotrypsin from bovine adrenal medulla and pituitary revealed by molecular cloning. *FEBS Lett* 368:471-6
- Inglis JD, Hill RE (1991) The murine Spi-2 proteinase inhibitor locus: a multigene family with a hypervariable reactive site domain. *EMBO J* 10:255-261
- Inglis JD, Lee M, Davidson DR, Hill RE (1991) Isolation of two cDNAs encoding novel antichymotrypsin-like proteins in a murine chondrocytic cell line. *Gene* 106:213-220
- Janciauskiene S, Rubin H, Lukacs CM, Wright HT (1998) Alzheimer's peptide Abeta1-42 binds to two beta-sheets of alpha1-antichymotrypsin and transforms it from inhibitor to substrate. *J Biol Chem* 273:28360-4
- Kaiserman D, Knaggs S, Scarff KL, Gillard A, Mirza G, Cadman M, McKeone R, Denny P, Cooley J, Benarafa C, Remold-O'Donnell E, Ragoussis J, Bird PI (2002) Comparison of human chromosome 6p25 with mouse chromosome 13 reveals a greatly expanded ov-serpin gene repertoire in the mouse. *Genomics* 79:349-62.
- Kanemaru K, Meckelein B, Marshall DC, Sipe JD, Abraham CR (1996) Synthesis and secretion of active alpha 1-antichymotrypsin by murine primary astrocytes. *Neurobiol Aging* 17:767-71
- Krugliak L, Meyer PR, Taylor CR (1986) The distribution of lysozyme, alpha-1-antitrypsin, and alpha-1-antichymotrypsin in normal hematopoietic cells and in myeloid leukemias: an immunoperoxidase study on cytocentrifuge preparations, smears, and paraffin sections. *Am J Hematol* 21:99-109
- Laursen I, Lykkesfeldt AE (1992) Purification and characterization of an alpha 1-antichymotrypsin-like 66 kDa protein from the human breast cancer cell line, MCF-7. *Biochim Biophys Acta* 1121:119-29
- Licastro F, Campbell IL, Kincaid C, Veinbergs I, Van Uden E, Rockenstein E, Mallory M, Gilbert JR, Masliah E (1999) A role for apoE in regulating the levels of alpha-1-antichymotrypsin in the aging mouse brain and in Alzheimer's disease. *Am J Pathol* 155:869-75
- Liu N, Raja SM, Zazzeroni F, Metkar SS, Shah R, Zhang M, Wang Y, Bromme D, Russin WA, Lee JC, Peter ME, Froelich CJ, Franzoso G, Ashton-Rickardt PG

- (2003) NF-kappaB protects from the lysosomal pathway of cell death. *EMBO J* 22:5313-22
- Mikolajczyk SD, Millar LS, Marker KM, Rittenhouse HG, Wolfert RL, Marks LS, Charlesworth MC, Tindall DJ (1999) Identification of a novel complex between human kallikrein 2 and protease inhibitor-6 in prostate cancer tissue. *Cancer Res* 59:3927-30
- Morris EC, Dafforn TR, Forsyth SL, Missen MA, Horvath AJ, Hampson L, Hampson IN, Currie G, Carrell RW, Coughlin PB (2003) Murine serpin 2A is a redox-sensitive intracellular protein. *Biochem J* 371:165-73
- Mucke L, Yu GQ, McConlogue L, Rockenstein EM, Abraham CR, Masliah E (2000) Astroglial expression of human alpha(1)-antichymotrypsin enhances alzheimer-like pathology in amyloid protein precursor transgenic mice. *Am J Pathol* 157:2003-10
- Stephan C, Jung K, Diamandis EP, Rittenhouse HG, Lein M, Loening SA (2002) Prostate-specific antigen, its molecular forms, and other kallikrein markers for detection of prostate cancer. *Urology* 59:2-8.
- Stratil A, Cizova-Schrofflova D, Gabrisova E, Pavlik M, Coppieters W, Peelman L, Van de Weghe A, Bouquet Y (1995) Pig plasma alpha-protease inhibitors PI2, PI3 and PI4 are members of the antichymotrypsin family. *Comp Biochem Physiol B Biochem Mol Biol* 111:53-60
- Terskikh AV, Easterday MC, Li L, Hood L, Kornblum HI, Geschwind DH, Weissman IL (2001) From hematopoiesis to neurogenesis: evidence of overlapping genetic programs. *Proc Natl Acad Sci U S A* 98:7934-9.
- Uhrin P, Dewerchin M, Hilpert M, Chrenek P, Schofer C, Zechmeister-Machhart M, Kronke G, Vales A, Carmeliet P, Binder BR, Geiger M (2000) Disruption of the protein C inhibitor gene results in impaired spermatogenesis and male infertility. *J Clin Invest* 106:1531-9
- Whitehead I, Kirk H, Kay R (1995) Expression cloning of oncogenes by retroviral transfer of cDNA libraries. *Mol Cell Biol* 15:704-710
- Wu G, Lilja H, Cockett A, Gershagen S (1998) Cloning and characterization of the alpha(1)-antichymotrypsin produced by human prostate tissue. *Prostate* 34:155-161
- Yamada M (2002) Risk factors for cerebral amyloid angiopathy in the elderly. *Ann N Y Acad Sci* 977:37-44

Table 1: Reactive centre loop sequences of the murine *serpina3* genes and tissue expression.

[illegible]

Table 2. Reactive center loop sequences of rat, cow and pig antichymotrypsin-like serpins.

The sequences shown below are derived from the publicly available rat genome database (<http://www.ncbi.nlm.nih.gov/genome/guide/rat/>) and published cow (Hwang et al. 1994; Hwang et al. 1995) and pig (Stratil et al. 1995) sequences. The predicted P₁/P₁' sites are indicated by the symbol ▼.

Species	Name	Accession no.	Sequence
Rat	SPI-2.4	XM_216782	GTEAAAATGANLVPR▼SGR-PPMIVWFNR
	CPI-26, SPI-3, SPI-2.2	XM_234494	GTEAEATTRVEYNFR▼PAKLNDTFVNFVR
	LOC299279	XM_234496	GTEAAAATGVKIIIPM▼CAKFYYVTMYFNR
	Spin2b, SPI2, Spin2, CPI-21	NM_012657	GTEGAAATAVTAALK▼SLPQTVPLLNFNR
	Spin2a, SPI1, Spin1, CPI-23	NM_182474	GTEATAATGVATVIR▼RQPRT---LNFNR
	Spin2c	NM_031531	GTEAAAATGVKFVPM▼SAKLDPLIIAFDR
Bovine	pHHK11	U13608	GTEGAAATGISME-R▼TISR--IIVRVNR
	pHHK12	U13609	GTEGAAVTAVVMA-T▼SSLLHTLTVSENR
	Pit ACT	S80570	GTEGAAATGIGIE-R▼TFLR--IIVRVNR
Porcine	SERPINA3-3	AJ293891	GTEAAAATGIEMMTS▼TL-QS-LTVIFSR
	SERPINA3-1	AJ293892	GTEGAASTGVVIERK▼SFEN--FIVRFDS
	SERPINA3-2	AJ293890	GTEAAAATGIDINVR▼SLER-IALHFNR

FIG. 1

	1180	P1	P1'	1221 1286	1314
Serpina3a	ACA ATA GCC CGA TAT AAC TTC CAA TCT --- GCA AFA ATT GTT AAG.....TAT GGG TAA GGT TAT CAA CCC TTT AAC AAB	ACA ATA GTC GGA TAC AAC TTC ATG TCT --- GCA AAA CTT AAA CCT.....CAT TTA ATG GGA AAG GTT ATC AAC CCC TCA	ACA GGG GTC AAT TTT CGC ATT CTT AGT --- ACA AGA ACC AGT CTG.....CTC TTT ATA GGC AAA ATC ACT CAC COT AAG	ACA AGA TTC AAA ATT GCC CCA CTG TCT --- GCA AAA TTT GAC ATT.....TTC GTG TTA GTG AAA GTC TTG AAT CCC AAG	ACA GGA TAT CAA AAT CTC CAA TGT TGT CAA GGT GTA ATC TAC TCT.....CTC TTT ATG GCA AAA GTT CCA AAT CCA GAG
Serpina3b	ACA AGA ATG GCA GGT GTC GGA TGT TGT --- GCA GTT TTT GAC TTT CTC TTT ATG GCA AAA GTT ACA AAT CCA GAG	ACA GGG GTC AAA TTA ATT CTA TGT TGT --- GAA AAA ATC TAT TCT.....CTC TTT ATG GGC AAA GTT ACA AAT CCC AAG	ACA GGA GTT AAA GTT AAT CTA CGT TGT --- GGA AAA ATC TAT TCT CTC TTT ATG GGC AAA GTT ACA AAT CCC AAL	ACA AGA GAC AAA TAT GAC TTC CTG TCC --- AAA TAT GAC TTC CTG.....ACT TTA TGG GCA AGA TCA ACA ACC CTG GAC	ACA GGG GTT ATT GGT GGC ATT CGT AAG GGC ATA TTA CCA GCT GTG.....CTC TTT ATG GGC AAA GTC ATT AAC CCC AAG
Serpina3c	TCG AGA GCC ATA TAT AAC TTC CAA TCT --- TCA AAA ATG TTC CTT CAT TCC TTT GGC AAG ATT GTG GAT COT TTA	ACA GGG TTC ATT TTT GGC TTT CGT TCT --- AGA AGA TTA CAA ACT.....CTC TTT ATG GGC AAA GTC ACT AAC CCC AAG	ACT GGA GTC AAA TTT GTC CCA ATG TCT --- GCG AAA CTG TAC CTT.....CCC TTT ATA GGC AAG ATA GCC AAC CCC AAA		
Serpina3d					
Serpina3e					
Serpina3f					
Serpina3g					
Serpina3h					
Serpina3i					
Serpina3j					
Serpina3k					
Serpina3l					
Serpina3m					
Serpina3n					

FIG. 2

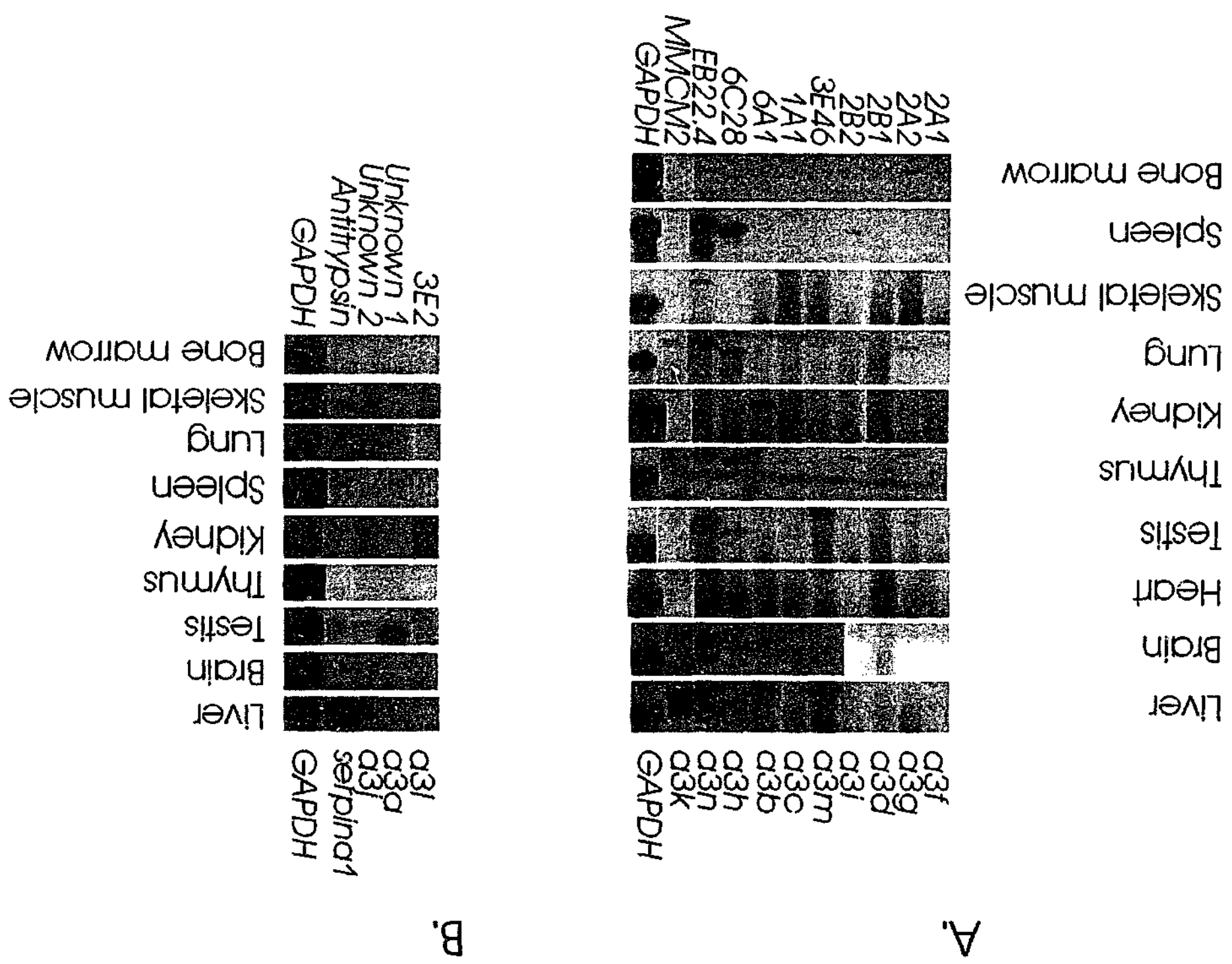
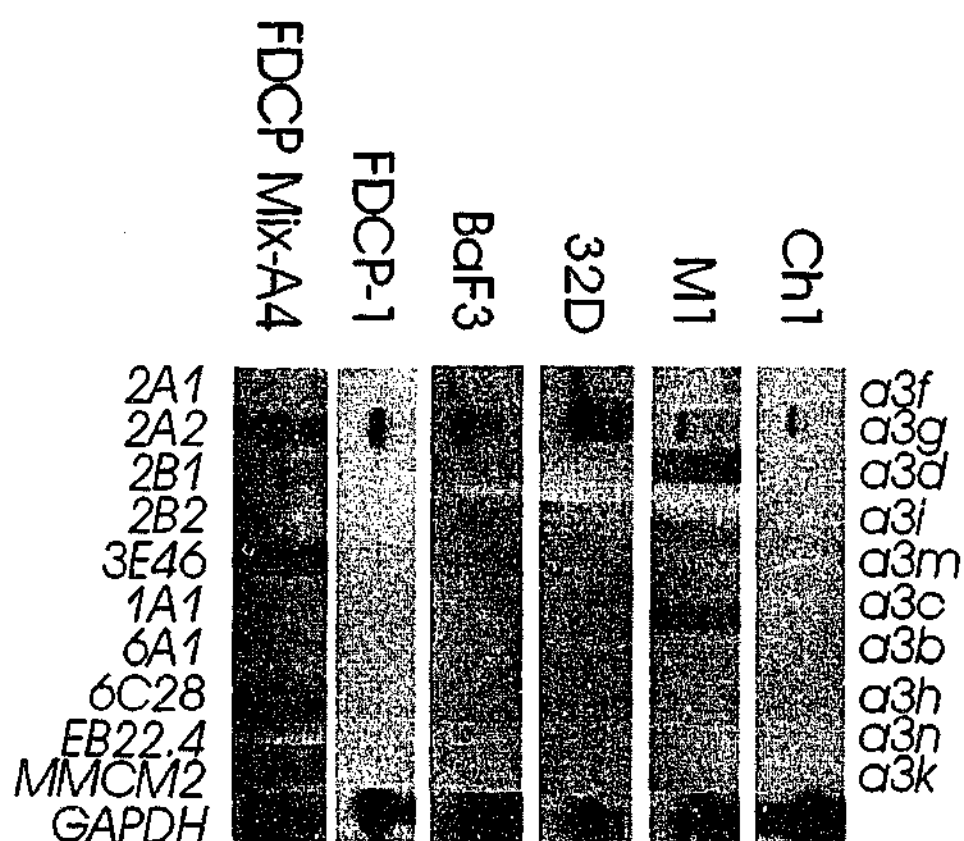
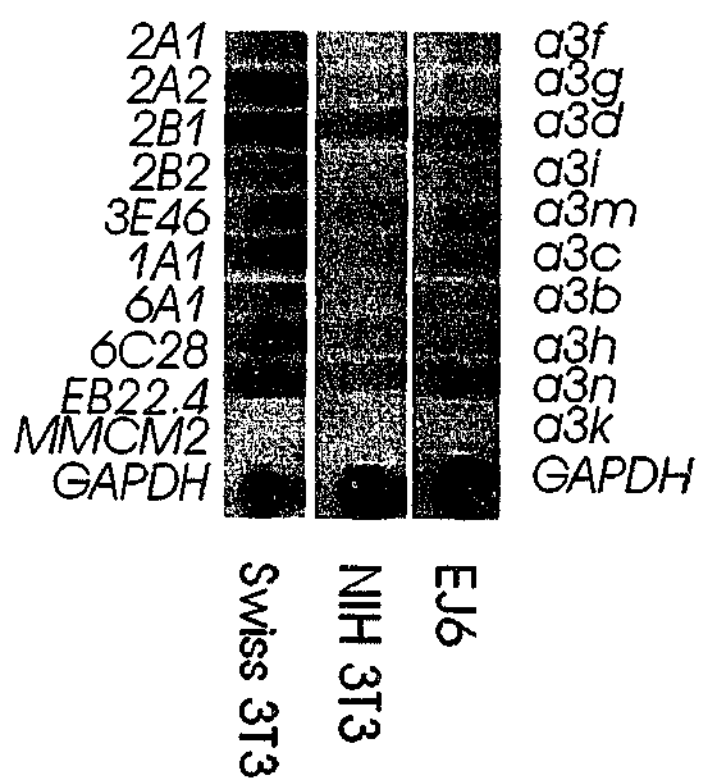


FIG. 3

A.



B.



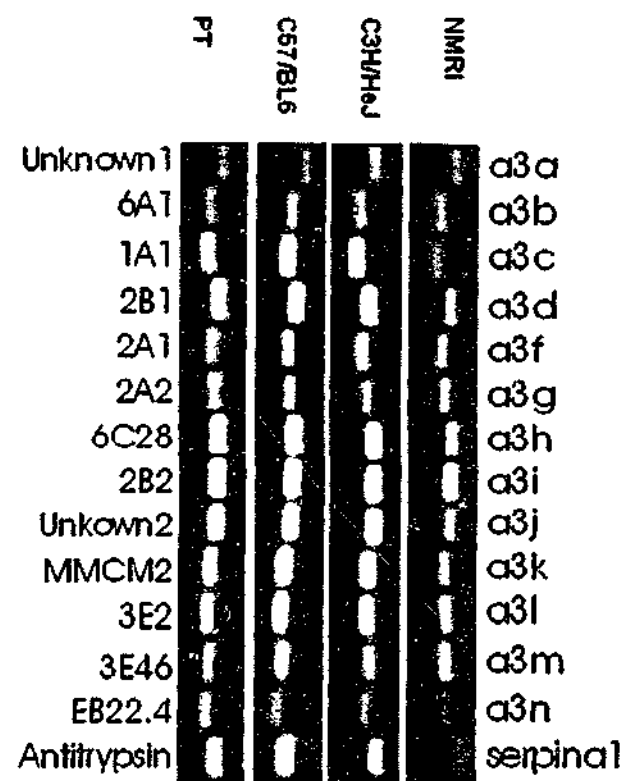
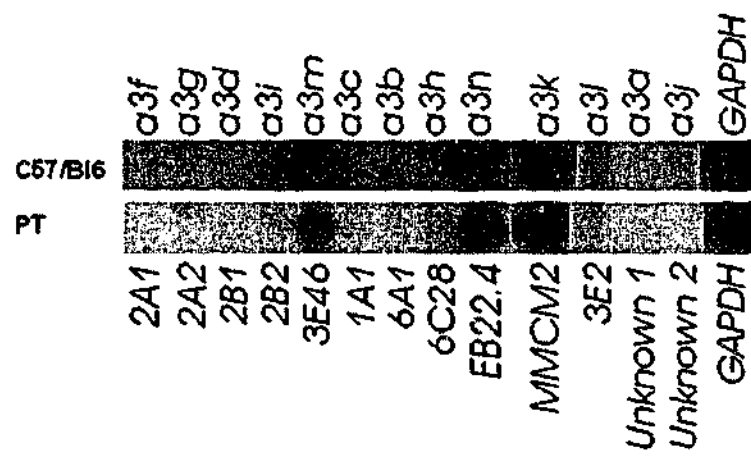


FIG. 4

FIG. 5



Mouse 12F1

a10 a6 HongrES1 a2 a1 a11 a9 a12 a4 A5ps a3 a3n

Spln26
 Spln28
 Spln29
 XLM_234496
 CP1-26
 XLM_26782

10 kb

Rat 6q32

a10 a6 HongrES1 a2 a1 a11 a9 a12 a4 a5 a3

Human 14q32.1

A10 A6 A2 A1 A11 A9 A12 A4 A5 A3



CHAPTER 4

Production and purification of bacterial recombinant mouse serpina3 proteins

4.1. INTRODUCTION

The studies in the previous chapter demonstrate that the most abundant *serpina3* genes expressed in mouse tissue are MMCM2, EB22.4, 3E46, and 6C28. The members MMCM2, EB22.4 and 3E46 share strong expression in liver and based on their predicted amino acid sequences, also have in common an N-terminal secretion peptide which suggests secretion into the circulation. 6C28 is expressed in a number of lymphoid tissues, however, the absence of a predicted N-terminal secretion peptide is suggestive of intracellular localisation. The predicted RCL sequences of all four serpins are significantly divergent and suggest distinct biological functions in mouse plasma and tissues. The possibility of functional diversity prompted an investigation into whether these four genes encode functional protease inhibitors. The aim of this chapter is to produce recombinant proteins for biochemical and biophysical analysis.

Recombinant proteins can be generated by a number of prokaryotic and eukaryotic expression systems. In our laboratory recombinant protein is routinely generated in *Escherichia coli*, *Pichia pastoris* and baculovirus. Eukaryotic expression has the advantage of yielding recombinant proteins with disulfide-bond formation and glycosylation. By contrast, bacterial systems are easy to manipulate, inexpensive and can generate a high yield of protein.

The studies undertaken in this chapter describe the production of the recombinant *serpina3* proteins; MMCM2, EB22.4, 3E46 and 6C28 in a bacterial expression system and their subsequent purification. Bacterial expression was chosen as a number of recombinant serpins, including human

α_1 -ACT, have been successfully generated in bacteria and yielded ample quantities of recombinant protein for structural and functional characterisation (Rubin *et al.* 1990; Hopkins *et al.* 1993; McCarthy and Worrall 1997; Irving *et al.* 2002 (b); Irving *et al.* 2003).

4.2. METHODS

4.2.1. RT-PCR of mouse tissue cDNA and construction of expression plasmids

Mouse liver, brain and spleen were isolated from 7 week NMRI mice and RNA extracted using the acid-guanidinium/phenol/chloroform method, as described in Chapter 2. 2 μ g of total RNA was treated with DNase prior to reverse transcription with Superscript II reverse transcriptase and 5 pmol/L oligo(dT), as described in Chapter 2. For RT-PCR, the sense primers were designed to amplify the *serpina3* genes EB22.4 (oligo 254), MMCM2 (oligo 254), and 3E46 (oligo 255), without the N-terminal secretion peptide. The 5' primer for 6C28 (oligo 265) was designed to the predicted start methionine, as shown in Figure 4.1. The antisense primers used were as follows: oligo 257 for EB22.4, oligo 256 for MMCM2, oligo 258 for 3E46 and oligo 260 for 6C28 (as described in Chapter 2 Table 2.2). Both sense and antisense primers included 5' tails encoding a *Bam*HI restriction site for cloning into the pETHis(3a) expression vector. The 5' primers were also designed to clone in frame with and immediately downstream of the His-tag sequence of the pETHis(3a) vector chosen for expression (Figure 4.1). The vector is modified from the parental pET3a with a poly-histidine tag inserted between the *Nde*I and *Bam*HI restriction site (constructed by Dr Paul Coughlin).

The open reading frames were amplified from their respective tissue cDNA using PCR, with either Vent polymerase, or Taq polymerase, as indicated in the results section. PCR cycles for the amplification of MMCM2 and EB22.4 were: 94°C 5min, [94°C 30 sec, 50°C 30 sec, 72°C 90 sec] x 30, 72°C 2 min.

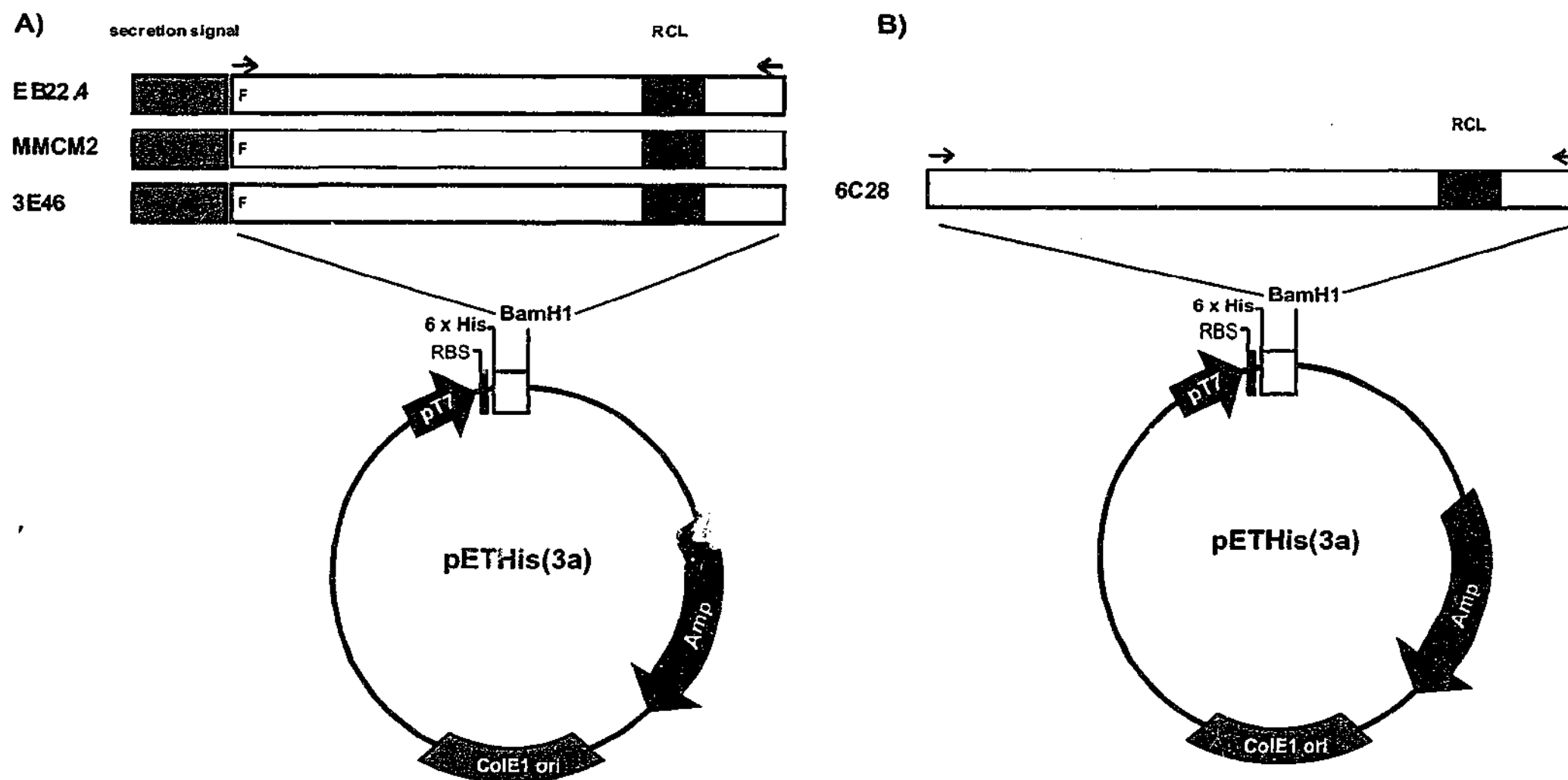


Figure 4.1: Cloning schematic for the introduction of the *serpina3* cDNA into the pETHis(3a) expression vector. A) cDNA for EB22.4, 3E46 and MMCM2 were PCR amplified after their predicted N-terminal secretion signal (red). **B)** cDNA for 6C28 was amplified from predicted start codon. Arrows indicate the location of primers. The *serpina3* cDNA were inserted in-frame with the polyhistidine tag (6x His) via the BamH1 restriction site.

For the amplification of 3E46 and 6C28 the PCR cycle conditions were: 94°C 5 min, [94°C 30 sec, 47°C 30 sec, 72°C 90 sec] x 30, 72°C 2 min. Products were identified on 1% (w/v) agarose/ethidium bromide gels as a band of approximately 1.2 kB. The amplified products were purified using the Qiagen PCR purification kit, according to manufacturer's instructions. The purified PCR products and the plasmid pETHis(3a) were digested with the restriction enzyme *BamH1* at 37°C overnight. 0.1U of calf intestine alkaline phosphatase was added to the restriction digest reaction of pETHis(3a) and incubated at 37°C for 1 hr. The digested products and pETHis(3a) vector were electrophoresed on 1% agarose gel and the bands excised and gel purified using the Qiagen gel extraction kit. The *serpina3* inserts and cut vector were ligated using T4 DNA ligase in a standard ligation reaction. The ligation reactions were transformed into chemically competent DH5 α cells and plated onto LB agar/ampicillin plates as described in Chapter 2.

Colony PCR reactions were performed to identify clones containing inserts in the correct orientation. The primers used were the T7 promoter primer and the insert specific oligonucleotides 256, 257, 285 and 260. Positive colonies were cultured overnight at 37°C, and miniprep DNA was isolated using the Qiagen miniprep kit. The presence of insert was confirmed by restriction enzyme digestion of the plasmids with *BamH1*. Correct orientation and sequence were confirmed by DNA sequencing with T7 primer and insert specific oligonucleotides 256, 257, 285, 260, 334, and 335 (Chapter 2: Table 2.2).

4.2.2. Trial inductions of *serpina3* proteins in BL21(DE3)pLysS

The four *serpina3* expression constructs and the pETHis(3a) empty vector were transformed into BL21(DE3)pLysS chemically competent cells. The BL21(DE3)pLysS cells are deficient in proteases that can degrade expressed proteins prior to purification. In addition, this host strain contains the pLysS plasmid which encodes the T7 lysozyme gene. Presence of the T7 lysozyme

lowers the basal expression of the target gene by binding to and inhibiting the T7 RNA polymerase.

Single colonies were used to inoculate an overnight culture of 2 ml 2 x TY, ampicillin (50 µg/ml). Cultures were incubated at 37°C overnight and shaking at 225 rpm. The following day cultures were diluted 1:10 into 2 ml 2 x TY/ampicillin (50 µg/ml) and incubated at 37°C in shaker incubator for 2 hr. Protein expression was induced by the addition of 0.2 mM IPTG. The addition of IPTG to bacterial cultures induces the lacUV5 promoter which allows the transcription/translation of T7 RNA polymerase and its subsequent binding to the T7 promoter drives target gene expression. Following addition of IPTG, the cultures were then incubated for an additional 2 hr at 37°C. For inductions at 30°C or room temperature, the cells were cooled to room temperature prior to the addition of 0.2 mM IPTG and were then incubated at either 30°C or room temperature for an additional 4 hr in a shaker incubator. Inductions in the presence of rifampicin were carried out as described above except that 30 min after induction with 0.2 mM IPTG, rifampicin was added to a final concentration of 100 µg/ml and cultures were grown for an additional 2 hr.

4.2.3. Trial inductions of His-MMCM2 and His-3E46 in AD494 cells

The expression vectors pETHis(3a)/MMCM2 and pETHis(3a)/3E46 were transformed into chemically competent AD494 cells and plated onto LB agar/kanamycin (30 µg/ml)/ampicillin (50 µg/ml) plates. The AD494 cells are an alternative bacterial strain for expression of recombinant proteins. These cells were a generous gift from Dr Noelene Quinsey (Department of Biochemistry and Molecular Biology, Monash University). Single colonies were grown overnight in 2ml cultures of 2 x TY/kanamycin (30 µg/ml)/ ampicillin (50 µg/ml) at 37°C in a shaker incubator at 225 rpm. The following day the cultures were diluted 1:10 in 2mL 2 x YT/kanamycin (30 µg/ml)/ampicillin (50 µg/ml) and incubated at 37°C in a shaker incubator for 2 hr. The cells were induced for expression (as with the BL21(DE3)pLysS) by the addition of

0.2 mM IPTG. The cultures were grown for an additional 2 hr after which the cells were harvested for analysis.

4.2.4. Analysis of trial inductions

The bacterial cells were harvested by centrifugation at 1,000 x g for 10 min at 4°C. The cells were resuspended in 500 µL ice-cold PBS, pH 7.4, lysozyme (1mg/ml), PMSF (20 µg/ml) and protease inhibitor cocktail (Sigma) 1/1000 and incubated on ice for 20 min. The cells were lysed by sonication on ice with 5 x 10 sec pulses at 30 sec intervals. Lysates were centrifuged at 16,000 x g for 10 min at 4°C and the supernatant, containing soluble protein, was transferred into a microfuge tube and set aside on ice. The pellet containing the insoluble protein fraction was resuspended in 500 µL ice-cold 1 PBS, pH 7.4, lysozyme (1 mg/ml), PMSF (20 µg/ml) and protease inhibitor cocktail (1/1000).

To determine the level of recombinant protein expression equal volumes of the soluble and insoluble fraction (10 µl) were combined with 2 x SDS-PAGE sample buffer (10 µl) and heated to 95°C for 5 min. The proteins were separated on 12.5% SDS-PAGE and then transferred to PVDF membrane by Western blotting. Membranes were probed with mouse-anti-His (1/2000 dilution) and goat-anti-mouse-HRP (1/5000). Proteins were detected with ECL reagent and visualised by autoradiography.

4.2.5. Large scale-production of *serpina3* proteins

A single colony of BL21(DE3)pLysS transformed with each of the pET(3a)His/*serpina3* plasmids was used to inoculate a 50 ml overnight culture of 2 x YT/ampicillin (50µg/ml). The overnight cultures were diluted 1/100 into 1L of 2 x YT/ampicillin (50µg/ml) and grown at 37°C until OD_{600nm} reached 0.7. They were then induced with 0.2 mM IPTG. Cultures of His-EB22.4 and His-MMCM2 were grown for a further 2-3 hrs at 37°C, at 225 rpm. His-3E46

and His-6C28 cultures were cooled to room temperature prior to induction and induced with 0.2 mM IPTG. The induced His-3E46 and His-6C28 cultures were incubated at 30°C and room temperature, respectively, for 4 hr with shaking at 225 rpm.

Cells were harvested by centrifugation (1,500 x g) and resuspended in 10 ml equilibration buffer (50 mM sodium phosphate, 100 mM NaCl, pH 7.8, 10 mM imidazole) to which lysozyme (1 mg/ml), 5 mM MgCl₂, 20 µg/ml PMSF, protease inhibitor cocktail (1/1000) and 5 mM β-mercaptoethanol were added. Lysates were incubated on ice for 30 min and DNase (25 µg/ml) added. The cells were freeze/thawed 3 times in liquid N₂ and at 37°C water bath respectively. The cell lysates were centrifuged at 27,000 x g for 15 min (Sorvall SS34 rotor) and the soluble supernatant transferred to a clean tube and stored on ice.

4.2.6. Purification of *serpina3* recombinants on Nickel-agarose

The first stage purification of recombinant proteins from soluble lysate was performed on Nickel-nitrilotriacetic acid (Ni-NTA) Agarose (Qiagen, Melbourne, Australia). The 50% (v/v) slurry of nickel-agarose (Ni-agarose) beads was washed with 10 column volumes of equilibration buffer and 2 ml of a 50% slurry was added to the soluble fraction and incubated overnight at 4°C with rotation. The lysate/beads were centrifuged at 1,500 x g for 10 min and supernatant removed and set aside as the unbound fraction. The beads were resuspended in 5 ml equilibration buffer and loaded onto a 10 cm X 1 cm chromatography column (Bio Rad) and attached to a Gilson peristaltic pump. The column was washed with 20 column volumes of equilibration buffer at 1 ml/min. Bound proteins were eluted with 20 column volumes of 125 mM imidazole/equilibration buffer in 1ml fractions. 10-50 µl of each fraction was diluted to a final volume of 800 µl with dH₂O and 200 µl of Bio-Rad Protein Assay reagent was added in order to measure protein concentration by absorbance at 595nm.

20 µl of each peak fraction was added to 20 µl of 2 x SDS-PAGE loading dye and samples heated to 95°C for 5 min. 20 µl of each sample was loaded onto 12.5% SDS-PAGE and proteins either visualised with Coomassie-stain or transferred to PVDF membranes by Western blotting. Membranes were probed with anti-His (1/2000) and anti-mouse-HRP (1/5000). Bound antibody was detected with ECL reagent and visualised by autoradiography.

4.2.7. Purification of recombinant His-tagged serpinA3 proteins by anion-exchange chromatography

To further purify the recombinant His-tagged serpinA3 proteins, anion exchange chromatography was performed using a Mono-Q column (Pharmacia,(Amersham) New Jersey, USA,) controlled by a Bio Logic Duo Flow HPLC (Bio Rad; California, USA; California, USA). The column was equilibrated with 15 column volumes of Buffer A [20 mM Tris, pH8.0, 5 mM β-mercaptoethanol, 5 mM EDTA], at a flow rate of 1 ml/min. The column was further washed with 10 column volumes of Buffer B, [20 mM Tris pH8.0, 0.5 M NaCl, 5 mM β-mercaptoethanol, 5 mM EDTA], and then with Buffer A until conductivity returned to baseline. The system was used at room temperature but all samples were kept on ice before and after chromatography.

The peak fractions from the Ni-agarose purification that contained recombinant protein, as assessed by 12.5% SDS-PAGE and Western blotting, were pooled and dialysed into 1L of Buffer A overnight. The dialysed recombinants were loaded onto the Mono-Q column by bulk loading. The column was then washed with 5 column volumes of Buffer A and bound protein was eluted with a 0-0.5 M NaCl gradient over 30 min. Peak fractions were determined by continuous recording of absorbance at 280nm and 1 mL fractions were collected by the automated fraction collector. The Mono-Q was washed in 20 column volumes of Buffer B and then re-equilibrated into Buffer A prior to commencing the next purification. Fractions containing the serpinA3 proteins

were identified by 12.5% SDS-PAGE and Western blotting (as described above in Section 4.2.6).

4.2.8. Purification of His-MMCM2 by Gel Filtration

To separate monomeric His-MMCM2 from the co-purifying immunoreactive 90kDa complex, gel filtration was performed. Superose12 (Bio-Rad) coupled with a Bio Logic Duo Flow HPLC (Bio-Rad) was utilised for this purpose. The column was equilibrated in Buffer C [20 mM Tris, pH 8.0, 0.15 M NaCl, 5 mM β -mercaptoethanol, 5 mM EDTA] for 1 hr at a flow rate of 1ml/min. 50 μ l of gel filtration standards (Bio Rad; California, USA) were loaded onto the column by injection via the 1 ml loop and separated over 60 min at a flow rate of 0.5 ml/min. The column was re-equilibrated with Buffer C for 1 hr at 1 ml/min. 500 μ l of a peak fraction eluted from the Mono-Q (identified on Coomassie stained SDS-PAGE and Western blotting) was loaded onto the Superose 12 column. The proteins were separated at 0.5 ml/min over 60 min and protein peaks detected by continuous measurement of absorbance at 280nm. 1 ml fractions were collected by the automated fraction collector and 10 μ l of each peak fraction was analysed on 12.5% SDS-PAGE and protein detected by Coomassie staining.

4.2.9. Mass spectroscopy

To determine the identity of the 90kDa complex, which co-purified with monomeric His-MMCM2, the complex was analysed by mass spectroscopy. Peak fractions of His-MMCM2 eluted from the anion-exchange column, were separated on 12.5% SDS-PAGE and proteins visualised by Coomassie staining. The band corresponding to high molecular weight immunoreactive complex, between 90-110 kDa was excised and the gel slice incubated with 500 μ l of 100 mM ammonium bicarbonate, pH 8.0 for 1 hr. The ammonium bicarbonate was removed and replaced with 500 μ l of 50% acetonitrile/100 mM ammonium bicarbonate, pH 8.9 and incubated for 1 hr

with shaking. This buffer was discarded and the gel slice was cut into 1 mm cubed gel slices and placed into three separate microfuge tubes. 50 μ l of acetonitrile was added to shrink the gel pieces and incubated for 10-15 min. The solvent was removed and the gel slices dried in a RC1010 Speedvac (Jouan: Easton, USA) followed by re-swelling in 40 μ l of 25 mM ammonium bicarbonate pH 8.0 containing 0.5 μ g of sequencing grade trypsin. The sample was incubated at 37°C overnight. The supernatant was transferred to a new tube and samples submitted for MALDI-mass spectroscopy (performed by Dr Simon Harris from Department of Biochemistry and Molecular Biology, Monash University).

4.2.10. Multiple sequence alignments and prediction of cleavage point, molecular weight, extinction co-efficient and pI.

The predicted amino acid sequences of the four cDNA isolated in this study were compared to sequences in GenBank as follows; EB22.4 (Accession No: M64086.1), 3E46 (Accession No: X55148), and MMCM2 (Accession No: X55147) and 6C28 (Mouse Genome), as well as the two mouse testis sequences isolated from EST databases, (EST1 GI: 28841458, EST2 GI: 29490274). Multiple amino acid sequence alignments were carried out using MultAlign (<http://prodes.toulouse.inra.fr/multalin/multalin>) (Corpet 1988). Signal P V1.1 was used to predict the location of the signal peptide cleavage site in the predicted amino acid sequences of MMCM2, EB22.4 and 3E46 (<http://www.cbs.dtu.dk/services/SignalP>). Molecular weight and extinction co-efficient of His-EB22.4, His-MMCM2, His-3E46 and His-6C28 were calculated based on the predicted amino acid sequence using the Peptide Properties Calculator (<http://www.basic.nwu.edu/biotools/proteincalc.html>). The pI for each recombinant protein was predicted using the Compute pI/Mw tool from Expasy (http://au.expasy.org/tools/pi_tool).

4.3. RESULTS

4.3.1. Isolation of *serpina3* cDNA transcripts

Figure 4.1 outlines the cloning strategy employed for the isolation of the *serpina3* transcripts and their subsequent insertion into the pETHis(3a) expression vector. The predicted amino acid sequences suggested that MMCM2, 3E46 and EB22.4 possessed an N-terminal secretion peptide thus to produce a recombinant resembling the mature extracellular protein the 5' primers for these were designed to amplify from the predicted N-terminal secretion signal cleavage point. The apparent absence of a secretion peptide for 6C28 suggested that the encoded protein would most likely localise intracellularly and accordingly the 5' primer was designed to amplify from the predicted start codon.

RT-PCR was employed to isolate the full length transcripts from the tissue cDNA in which the genes were dominantly expressed (Chapter 3) namely, brain for EB22.4, liver for MMCM2 and 3E46, and spleen and the cell line Swiss-3T3 for 6C28 (Figure 4.2). Using the proof reading Vent polymerase, cDNA transcripts were only amplified for MMCM2 and EB22.4 at the expected size of approximately 1.2kb (Figure 4.2(A)). The isolation of the full length cDNA for 3E46 and 6C28 using Vent polymerase was unsuccessful (Figure 4.2(A) and (B)). Therefore, Taq polymerase amplified products were isolated for these two genes and used for succeeding cloning steps. The four *serpina3* cDNA were inserted into the *Bam*HI site of the pET(3a)His vector as outlined schematically in Figure 4.1. Each full length *serpina3* cDNA was nucleotide sequenced with T7 primers and gene specific primers (Table 2.3) to verify the authenticity of the product (described in Chapter 2).

Figure 4.3 shows a multiple alignment of the predicted amino acid sequences of the EB22.4, 3E46 and MMCM2, as determined from the cDNA sequence of the amplified products. This is shown in comparison to corresponding

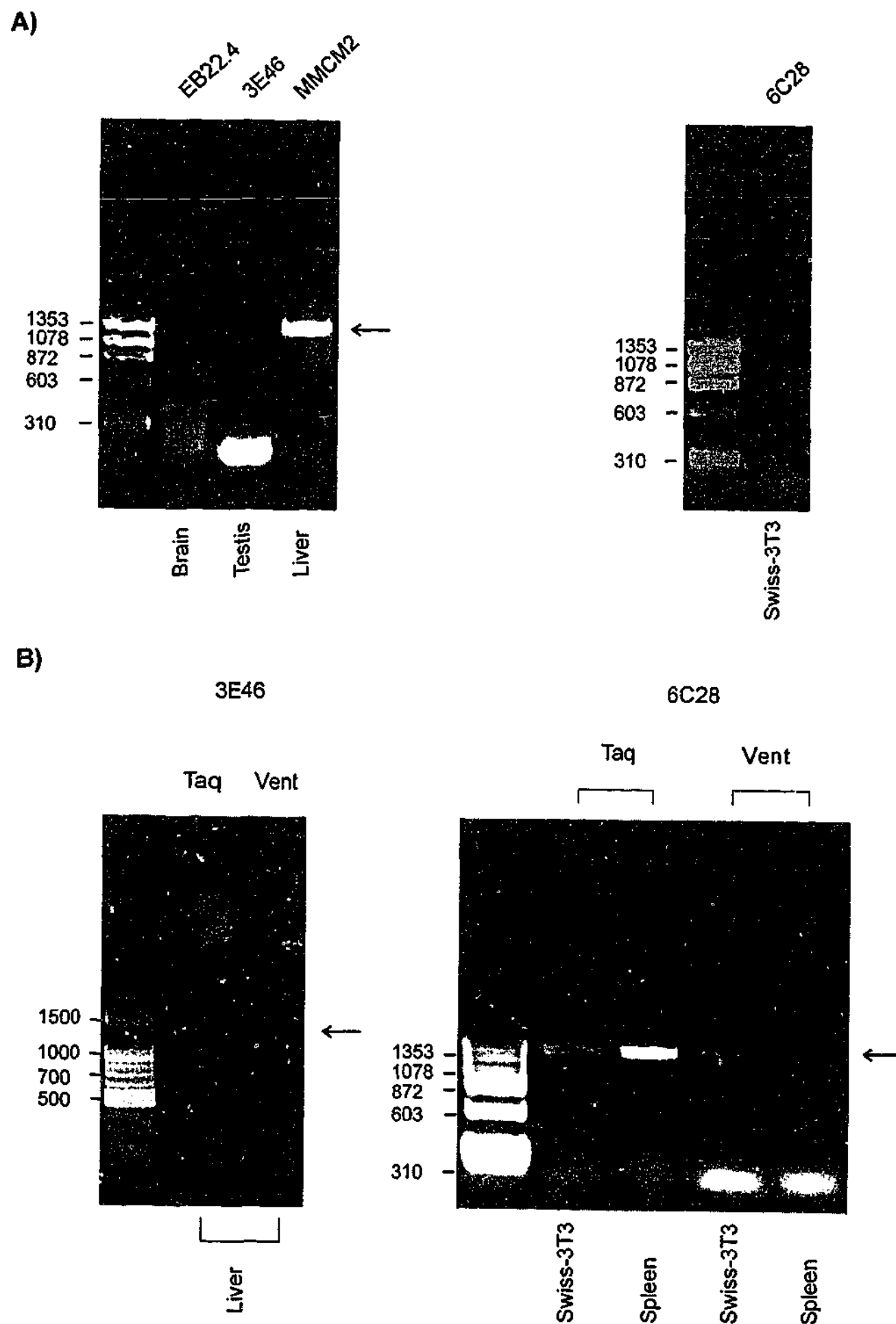


Figure 4.2: Isolation of full length transcripts for serpinA3 genes EB22.4, MMCM2, 3E46 and 6C28. **A)** RT-PCR using Vent polymerase and indicated mouse tissue cDNA. **B)** RT-PCR using Taq polymerase to isolate 3E46 and 6C28. Arrows indicate the presence of PCR amplified products of interest.

sequences from published sources present in GenBank. The predicted cleavage point is shown by the arrow, and the predicted N-terminal secretion peptide is indicated in green. Variations between the GenBank sequences and isolated sequences are highlighted in red (non-conserved) and blue (conserved). To determine whether these variations were to conserved amino acids, the highlighted residues were compared to the other members of the mouse *serpina3* family and to a multiple sequence alignment of the serpin superfamily reported by Irving *et al* (2000).

EB22.4 shows a Lys(70)→Thr substitution in the N-terminal region. Comparison of this substitution to amino acids which reside at this position in other mouse *serpina3* members revealed that it is conserved with Thr (70) in MMCM2. The Thr residue is also present in human and mouse α_1 -AT's. A further two substitutions at Val(85)→Leu and Met(86)→Val are also conserved with other mouse *serpina3* members. 3E46 also has an amino acid variation of Ile(210)→Val in strand3 of β -sheet A. Whilst all the mouse *serpina3* have an Ile residue at this position, comparison of this variation with other members of the serpin superfamily showed that Val(210) was present in human PI-6 and mouse SCCA-2. Therefore, it is unlikely that any of these variations would affect structural stability or functional activity. No predicted amino acid variations were observed for MMCM2.

Of particular interest is the 6C28 cDNA transcript amplified from spleen cDNA. Figure 4.4 is a comparison of the predicted amino acid sequence of the amplified 6C28 (spleen) to the sequence taken from the Mouse genome (6C28 (genome)). The figure demonstrates a total of 20 amino acid substitutions as highlighted in red (non-conserved) and blue (conserved). Table 4.1 shows an analysis of these variations to residues present in members of the mouse *serpina3* locus and the serpin superfamily. Six of the substitutions are located N-terminal of the first structural element helix A and are therefore unlikely to influence the structural stability of the serpin molecule. The variations that are located in structural regions are to residues which are conserved with either members of the mouse *serpina3* family or the serpin

	1					60
MMCM2 (G)	MAFIVAMGMI	LMAGVCPAVLC	FPDGTKEMD	IVFHEHQDNG	TQDDSLTLAS	VNTDFAFSLY
MMCM2 (R)			FPDGTKEMD	IVFHEHQDNG	TQDDSLTLAS	VNTDFAFSLY
3E46 (G)	MAFIAALG-I	LMAGICPTVLC	FSDDTWGID	ILLHKNQESG	TPDDSLTLAS	INTDFAFSLY
3E46 (R)			FSDDTWGID	ILLHKNQESG	TPDDSLTLAS	INTDFAFSLY
EB22.4 (G)	MAFIAALG-L	LMAGICPAVLC	FPDGTLGMD	AAVQEDHDNG	TQLDSLTLAS	INTDFAFSLY
EB22.4 (R)			FPDGTLGMD	AAVQEDHDNG	TQLDSLTLAS	INTDFAFSLY
	61					120
MMCM2 (G)	KKLALKNPDT	NIVFSPLSIS	AALALVSLGA	KGKTMEEILE	GLKFNLITETP	EADIHQGFNG
MMCM2 (R)	KKLALKNPDT	NIVFSPLSIS	AALALVSLGA	KGKTMEEILE	GLKFNLITETP	EADIHQGFNG
3E46 (G)	KELVLKNPDK	NIVFSPLSIS	AALALVSLGA	KGNTLEEILE	GLKFNLITETS	EADIHQGFNG
3E46 (R)	KELVLKNPDK	NIVFSPLSIS	AALALVSLGA	KGNTLEEILE	GLKFNLITETS	EADIHQGFNG
EB22.4 (G)	KELVLKNPDK	NIVFSPLSIS	AALAVMSLGA	KGNTLEEILE	GLKFNLITETS	EADIHQGFNG
EB22.4 (R)	KELVLKNPDT	NIVFSPLSIS	AALALVSLGA	KGNTLEEILE	GLKFNLITETS	EADIHQGFNG
	121					180
MMCM2 (G)	LLQSLSQPED	QDQINIGNAM	FIEKDLQILA	EFHEKTRALY	QTEAFTADFQ	QPTEAKNLIN
MMCM2 (R)	LLQSLSQPED	QDQINIGNAM	FIEKDLQILA	EFHEKTRALY	QTEAFTADFQ	QPTEAKNLIN
3E46 (G)	LLQRLSQPED	QDQINIGNAM	FIEKDLQILA	EFHEKARALY	QTEAFTADFQ	KPTEAKNLIN
3E46 (R)	LLQRLSQPED	QDQINIGNAM	FIEKDLQILA	EFHEKARALY	QTEAFTADFQ	KPTEAKNLIN
EB22.4 (G)	LLQRLNQPKD	QVQISTGSAL	FIEKRQOILT	EFQEKAKTLY	QAEFTADFQ	QPRQAKKLIN
EB22.4 (R)	LLQRLNQPKD	QVQISTGSAL	FIEKRQOILT	EFQEKAKTLY	QAEFTADFQ	QPRQAKKLIN
	181					240
MMCM2 (G)	DYVSNQTQGM	IKELISELDE	RTLMLVNYI	YFKGKWKISF	DPQDTFESEF	YLDEKRSVKV
MMCM2 (R)	DYVSNQTQGM	IKELISELDE	RTLMLVNYI	YFKGKWKISF	DPQDTFESEF	YLDEKRSVKV
3E46 (G)	DYVSNQTQGM	IKELISELDT	DTLMVLVNYI	YFKGKWKISF	DPQDTFESEF	YLDEKRSVKV
3E46 (R)	DYVSNQTQGM	IKELISELDT	DTLMVLVNYV	YFKGKWKISF	DPQDTFESEF	YLDEKRSVKV
EB22.4 (G)	DYVRKQTQGM	IKELVSDLDK	RTLMLVNYI	YFKAKWKVPF	DPLDTFKSEF	YCGKRRPVIV
EB22.4 (R)	DYVRKQTQGM	IKELVSDLDK	RTLMLVNYI	YFKAKWKVPF	DPLDTFKSEF	YCGKRRPVIV
	241					300
MMCM2 (G)	PMMKMKLLTT	RHFRDEELSC	SVLELKYTGN	ASALLILPDQ	GRMQQVEASL	QPETLRKWRK
MMCM2 (R)	PMMKMKLLTT	RHFRDEELSC	SVLELKYTGN	ASALLILPDQ	GRMQQVEASL	QPETLRKWRK
3E46 (G)	PMMKMKFLTT	RHFRDEELSC	SVLELKYTGN	ASALFILPDQ	GRMQQVEASL	QPETLRKWWK
3E46 (R)	PMMKMKFLTT	RHFRDEELSC	SVLELKYTGN	ASALFILPDQ	GRMQQVEASL	QPETLRKWWK
EB22.4 (G)	PMMSMEDLTT	PYFRDEELSC	TVVELKYTGN	ASALFILPDQ	GRMQQVEASL	QPETLRKWKV
EB22.4 (R)	PMMSMEDLTT	PYFRDEELSC	TVVELKYTGN	ASALFILPDQ	GRMQQVEASL	QPETLRKWKV
	301					360
MMCM2 (G)	TLFPSQIEEL	NLPKFSIASN	YRLEEDVLPE	MGIKEVFTEQ	ADLSGITETK	KLSVSQVVHK
MMCM2 (R)	TLFPSQIEEL	NLPKFSIASN	YRLEEDVLPE	MGIKEVFTEQ	ADLSGITETK	KLSVSQVVHK
3E46 (G)	SLKTRKIGEL	YLPKFSISTD	YNLK-DILPE	LGIKEIFSKQ	ADLSGITGTK	DLSVSQVVHK
3E46 (R)	SLKTRKIGEL	YLPKFSISTD	YNLK-DILPE	LGIKEIFSKQ	ADLSGITGTK	DLSVSQVVHK
EB22.4 (G)	SLKPRMIDEL	HLPKFSISTD	YSLE-DVLSK	LGIREVFSTQ	ADLSAITGTK	DLRVSQVVHK
EB22.4 (R)	SLKPRMIDEL	HLPKFSISTD	YSLE-DVLSK	LGIREVFSTQ	ADLSAITGTK	DLRVSQVVHK
	361		P₁-P₁'			420
MMCM2 (G)	AVLDVAETGT	EAAAATGVIG	GIR KAILPA-	-VHFNRPFLE	VIYHTSAQSI	LFAKVNNPK
MMCM2 (R)	AVLDVAETGT	EAAAATGVIG	GIR KAILPA-	-VHFNRPFLE	VIYHTSAQSI	LFAKVNNPK
3E46 (G)	AVLDVAETGT	EAAAATGFIF	GFR SRRLQTM	TVQFNRPFLM	VISHTGVQTT	LFAKVNTNPK
3E46 (R)	AVLDVAETGT	EAAAATGFIF	GFR SRRLQTM	TVQFNRPFLM	VISHTGVQTT	LFAKVNTNPK
EB22.4 (G)	AVLDVAETGT	EAAAATGVKF	VPM SAKLYPL	TVYFNRPFLI	MIFDTETEIA	PFIKIANPK
EB22.4 (R)	AVLDVAETGT	EAAAATGVKF	VPM SAKLYPL	TVYFNRPFLI	MIFDTETEIA	PFIKIANPK

Figure 4.3: Comparison of predicted amino acid sequences of EB22.4, MMCM2 and 3E46 isolated by RT-PCR (R) to reported sequences in Genbank(G). Alignment was analysed using MultAlign (<http://prodes.toulouse.inra.fr/multalin/multalin>). The arrow indicates the predicted cleavage site and P₁-P₁' are shown in bold (black). Conserved variations are shown in blue and non-conserved in red.

	1						60
6C28 (genomic)	MAGISPAVLG	CPCVTLDRDT	-AVHEVQENI	TSGDSFTLAS	SNTDFAFSLY	RKLVLKNPDE	
6C28 (spleen)	MAGISPAVFG	IPDVTLERNT	VAVHEVQENI	TSGDSLTVAS	SNTDFAFSLY	RKLALKNPDE	
Testis EST1	
Testis EST2	
3T3Spi2	
	61						120
6C28 (genomic)	NVVFSPFSIS	AALALLSLGA	KSNTLKEILE	GLKFNLITETP	EPDIHQGFY	LLDLLSQPGD	
6C28 (spleen)	NVVFSPFSIS	AALALLSLGA	KGNTLKEILE	GLKFNLITGAP	EPDIHQGFY	LLDLLSQPGD	
Testis EST1	
Testis EST2	
3T3Spi2	
	121						180
6C28 (genomic)	QVQISTGSAL	FIEKHLQILA	EFKEKARALY	QAEAFADFL	QPRQATKLIN	DYVSNQTQ GK	
6C28 (spleen)	QVQIRTGSAL	FIEKHLQILA	EFKEKARALY	QAEAFADFL	QPRQATKLIN	DYVSNQTQ GK	
Testis EST1	
Testis EST2	
3T3Spi2	
	181						240
6C28 (genomic)	IKELISDL DK	STLMVLVNYI	YFKGKWKMPF	DPHDTFNSVF	YLDEKRSVNV	SMMKIEELTT	
6C28 (spleen)	IKELISDL DK	STLMVLVNYI	YFKGKWKMPF	DPHDTFNSVF	YLDEKGSVNV	SMMKIEELTT	
Testis EST1	
Testis EST2	
3T3Spi2	
	241						300
6C28 (genomic)	PYFRDDELSC	TVVELKYTGN	ASAMFILPDQ	GRMQQVEASL	QPETLRKWK N	SLKPRRIDEL	
6C28 (spleen)	PYFRDDELSC	TVVELKYTGN	ASAMFILPDQ	GRMQQVEASL	QPETLRKWK N	SLKPRRIDEL	
Testis EST1	
Testis EST2	
3T3Spi2	
	301						360
6C28 (genomic)	HLPKFSISMD	NSLEHILPEL	GIRELFSTQA	DLSAITGTDK	LRVSQVVHKA	VLDVAETGTE	
6C28 (spleen)	HLPKFSISTD	NSLEHILPEL	GIREVFSTQA	DLSAITGTDK	LRVSQVVHKA	VLDVAETGTE	
Testis EST1QA	DLSAITGTDK	LRVSQVVHKA	VLDVAETGTE	
Testis EST2D	LRVSQVVHKA	VLDVAETGTE	VLDVAETGTE	
3T3Spi2KA	VLDVAETGTE	VLDVAETGTE	
	361						409
6C28 (genomic)	AAAATGVKLI	LC CEKIYSMT	IYFNRPF LMI	ISDINAHIAL	FMAKV TNP K		
6C28 (spleen)	AAAATGVKLI	LL CRKIYSMT	IYFNRPF LMI	ISDTKAHIAL	FMAKV TNP K		
Testis EST1	AAAATGVKLI	LL CRKIYSMT	IYFNRPF LMI	ISDTQAHIAL	FMAKV TNP K		
Testis EST2	AAAATGVKLI	LL CRKIYSMT	IYFNRPF LMI	ISDTKAHIAL	FMAKV TNP K		
3T3Spi2	AAAATGVKLI	LL CRKIYSMT	IYFNRPF LMI	ISDTKAHIAL	FMAKV TNP K		

Figure 4.4: Comparison of the predicted amino acid sequence of 6C28 (spleen) to 6C28 (genome), 3T3-Spi2 and to two sequences identified in the mouse EST database which are expressed in testis. Conserved amino acid variations are highlighted in blue and non-conserved variations are highlighted in red.

Table 4.1: Predicated amino acid substitutions in 6C28 (spleen) and comparison of these residues to those present in mouse *serpina3* and serpin superfamily.

Substitution	Location	Conserved
Leu(9)→Phe	N-terminus	
Cys(11)→Ile	N-terminus	
Cys(13)→Asp	N-terminus	
Asp(17)→Glu	N-terminus	
Asp(19)→Asn	N-terminus	
Insertion at P20	N-terminus	
Phe(35)→Leu	Helix A	Conserved: Leu residue conserved with serpin superfamily
Leu(37)→Val	Helix A	Conserved: Val residue present in avian MENT and human PAI-1
Val(53) →Ala	Helix A	Conserved: Ala residue conserved with serpin superfamily
Ser(81) →Gly	Helix C	Conserved: Glu residue conserved with serpin superfamily
Glu(97)→ Gly	Between helix C and D	Conserved: Gly residue present in human megsin and mouse antipiasmin
Thr(98)→ Ala	Between helix C and D	Conserved: Ala residue present in rat RASP-1
Arg(145)→Gly	Between s4C and s3C	Conserved: Gly residue present in viral serpins, Spi-3
Met(288)→Thr	strand6A	Conserved: Thr residue conserved with mouse <i>serpina3</i> members
Leu(304) → Val	Between helix I and s5A	Conserved: Val residue present in mouse <i>serpina3</i> members
Thr(307) →Met	Between helix I and s5A	Not conserved with any serpins
Cys(372) →Leu	RSL -P ₁ residue	Conserved: Leu residue present in 3T3-Spi2
Glu(374) → Arg	RSL- P ₂ residue	Conserved: Arg residue present in 3T3-spi2
Ile(393) →Thr	Between s4B and s5B	Conserved: Thr residue present in 3T3-Spi2 and <i>serpin2A</i>
Asn(394)→Lys	Between s4B and s5B	Conserved: Lys residue present in 3T3-Spi2 and <i>serpin2A</i>

superfamily (Table 4.1). However, the Thr(307)→Met substitution is not shared by any other members of the serpin superfamily. This substitution is located in a non-conserved region between helix I and strand 5 of β sheet A and is unlikely to influence serpin structure or function (Irving *et al.* 2000).

There are however, a number of functionally significant substitutions in the predicted amino acid sequence of 6C28 (spleen). Of particular importance is the Cys (372)→Lys at the predicted P₁ residue of the RCL, which would alter the target protease profile of the serpin. A further substitution at the predicted P₂' residue from Glu→Arg could also alter protease interactions. Figure 4.4 shows that the transcript isolated from spleen is identical to a partial transcript obtained from a fibroblast cell line in our laboratory by Dr Sharon Forsyth (unpublished results). 6C28 (spleen) similarly shows 100% amino acid identity with EST's isolated from mouse testis. The predicted amino acid sequence of 6C28 (spleen) and 6C28 (genome) share 96% nucleotide sequence homology and 95% overall amino acid identity. The highly homologous nature of the two genes suggests they are likely to be strain/species variants.

The predicted amino acid sequences of the four recombinant serpin3 proteins were used to calculate their molecular weight, extinction co-efficient and pI (Table 4.2).

Table 4.2: Predicted properties of recombinant serpin3 proteins

	His-EB22.4	His-MMCM2	His-3E46	His-6C28
Molecular Weight (Daltons)	45885	45837	46109	46329
Extinction co-efficient (u.ml.mg ⁻¹ .cm ⁻¹)	25700	23020	34400	24780
pI	5.92	5.36	5.69	6.40

4.3.2. *Small-scale induction of serpinA3 proteins*

The induction of the four serpinA3 proteins at 37°C with IPTG resulted in varied levels of expression. Analysis of protein expression was carried out by centrifugation of the total bacterial lysate into the soluble supernatant and insoluble pellet and both fractions were Western blotted and probed with the His-tag antibody. Figure 4.5 shows the Western blots of His-EB22.4 and His-MMCM2 inductions. Immunoreactive bands to the His-antibody were detected below the level of the 50kDa marker corresponding to the expected size of His-EB22.4 and His-MMCM2. The bands were present in both insoluble and soluble fractions, with significant quantities partitioning into the soluble fraction (Figure 4.5(A)). In both soluble and insoluble lysates of His-MMCM2 a high molecular weight band at approximately 90kDa was also immunoreactive with the His-antibody. This band was not observed in His-EB22.4 fractions. Figure 4.5(B) demonstrates that the proteins recognised by the His-antibody (at approximately 50kDa and 90kDa) in His-MMCM2 fractions, are also immunoreactive with the polyclonal serpin2A antibody. This was expected based on the high amino acid identity shared by the mouse serpinA3 family. No cross-reactivity with bacterial proteins was observed with either antibody in vector control samples.

Analysis of the soluble and insoluble fractions of His-3E46 induced at 37°C, revealed the presence of a prominent band below the level of the 50kDa marker (Figure 4.6(A)). Only small amounts of the recombinant protein partitioned into the soluble fraction and appeared to be smaller in molecular mass than His-EB22.4 and His-MMCM2. In the insoluble pellet fraction a second entity migrating at a slightly higher molecular weight than the dominant immunoreactive band is evident and is indicated by the upper arrow. These data raise the possibility that His-3E46 may be cleaved in bacteria during expression. Figure 4.6(B) shows the induction of His-6C28 at 37°C and demonstrates the presence of an immunoreactive inducible band below the 50kDa marker. However, all of the recombinant protein partitioned into the insoluble fraction, with no soluble protein detected.

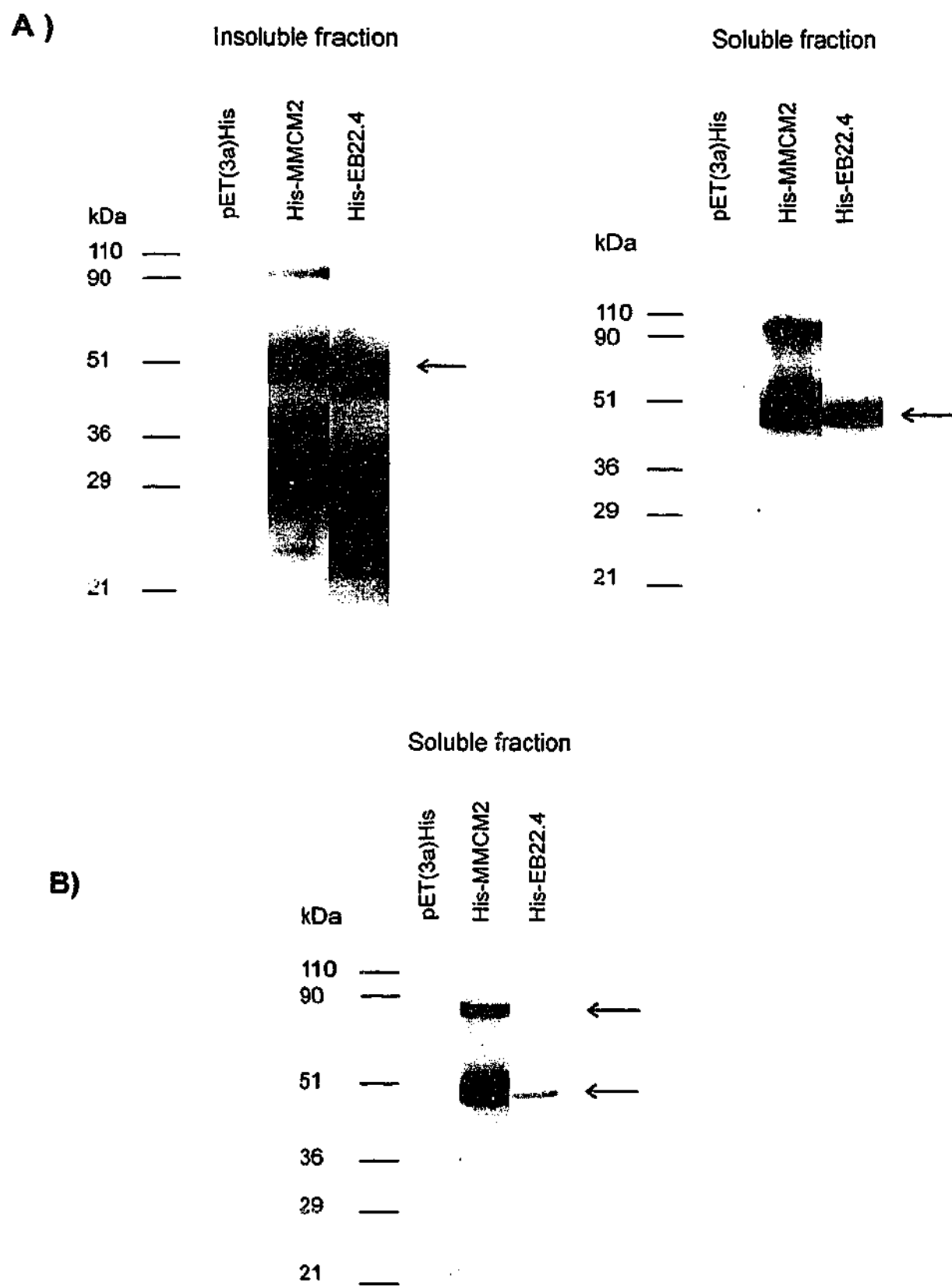


Figure 4.5: Trial induction of pET(3a)His constructs of His-MMCM2 and His-EB22.4. Cultures were induced by the addition of 0.2mM IPTG. Protein was then fractionated into either soluble or insoluble fractions and separated on 12.5% SDS-PAGE electrophoresis. The presence of recombinant proteins was detected by Western blotting with **A)** monoclonal His-tag antibody and **B)** polyclonal antibody to serpin 2A. Arrows indicate the presence of immunoreactive entities. pETHis(3a) lanes are lysates isolated from vector control inductions. Relative molecular mass is shown on the left.

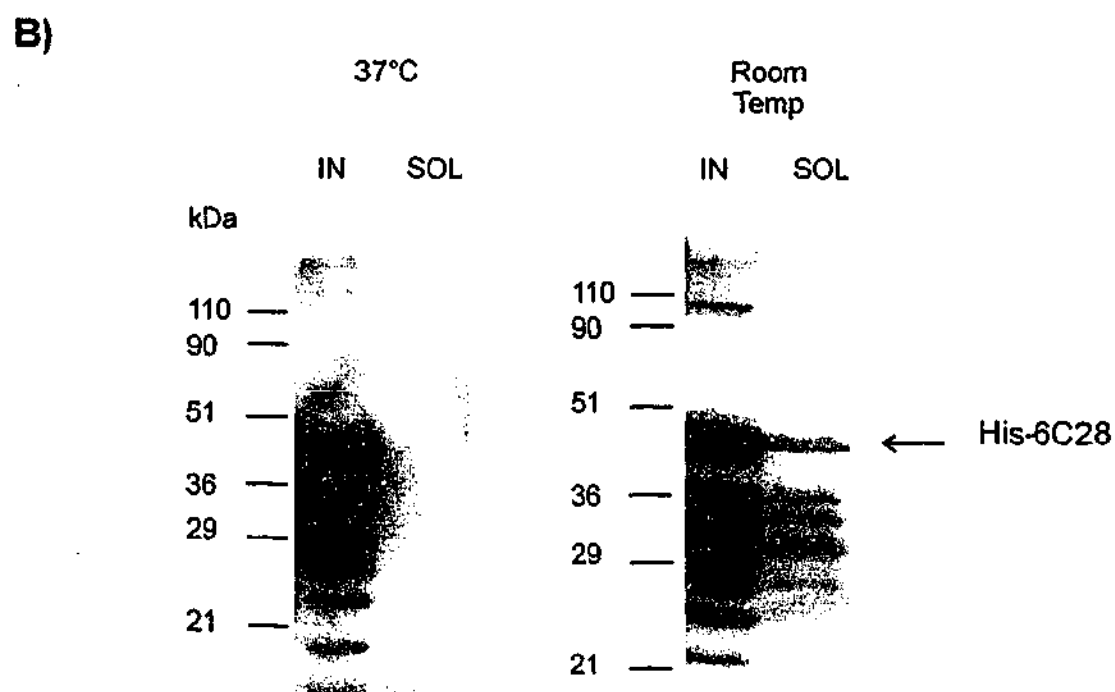
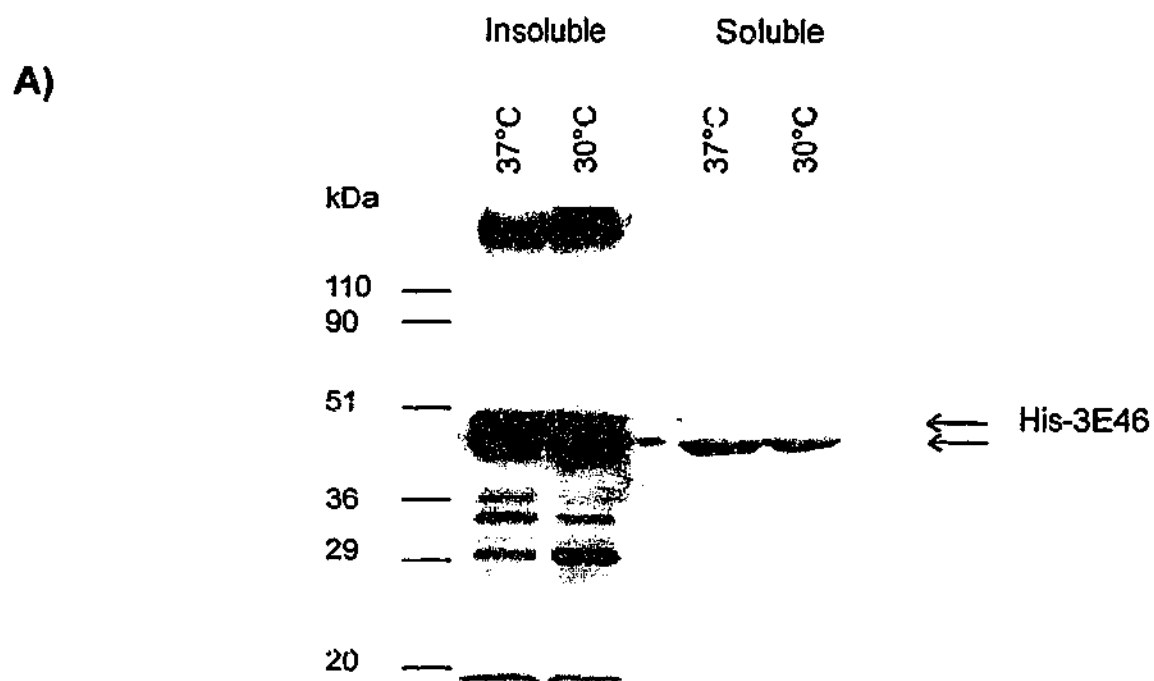


Figure 4.6: Trial inductions of recombinant His-3E46 and His-6C28. BL21 cells were induced for expression for either **A)** His-3E46 at 37°C or 30°C, or **B)** His-6C28 at 37°C or room temperature, with 0.2mM IPTG. Protein was then fractionated into insoluble (IN) and soluble (SOL) fractions, separated on 12.5% SDS-PAGE and analysed by Western blotting with a monoclonal His-antibody. Arrows indicate immunoreactive bands at the expected molecular weight for the recombinant proteins. Relative molecular mass is shown on the left.

4.3.3. Optimisation of expression

In order to optimise the expression of soluble His-3E46 and His-6C28, protein expression was induced at 30°C or room temperature and compared to inductions at 37°C. Figure 4.6(A) shows that there was no change in the level of soluble protein present by inducing His-3E46 expression at 30°C. In contrast, the induction of His-6C28 at room temperature did increase the production of soluble protein (Figure 4.6 (B)).

Previous studies with bacterially produced human α_1 -AT have shown that addition of rifampicin to bacterial cultures, following induction with IPTG, increased the level of soluble protein (Hopkins *et al.* 1993). Rifampicin inhibits *E.coli* RNA polymerases but not the T7 RNA polymerase, therefore, protein synthesis in bacteria is diverted towards transcripts produced by the T7 RNA polymerase. Figure 4.7 shows the effect of the addition of rifampicin to the His-3E46 and His-6C28 cultures following induction with 0.2 mM IPTG at 37°C and room temperature respectively. The addition of rifampicin to either His-3E46 or His-6C28 cultures did not alter the solubility of either protein.

The bacterial cell line AD494 was used to investigate whether expression in an alternative bacterial cell type could alter the size and solubility of His-3E46. Figure 4.7(A) shows that there was no difference observed in solubility or molecular weight of recombinant His-3E46 synthesised in AD494 cells in comparison to that generated in BL21(DE3)pLysS cells. Induction of His-MMCM2 was also trialled in the presence of rifampicin and in AD494 cells to see if the high molecular weight protein at approximately 90kDa could be removed. Figure 4.7(A) shows no significant difference in the amount of 90kDa protein present in the soluble fractions between the different expression trials. The small decrease observed in BL21(DE3)pLysS and rifampicin lysates is likely to be the result of a decrease in total protein expression. Expression of His-6C28 was not trialled in AD494 as no difference in solubility was observed with either His-3E46 or His-MMCM2 inductions in these cells. Figure 4.7(A) clearly demonstrates the lower molecular mass of His-3E46 to recombinant

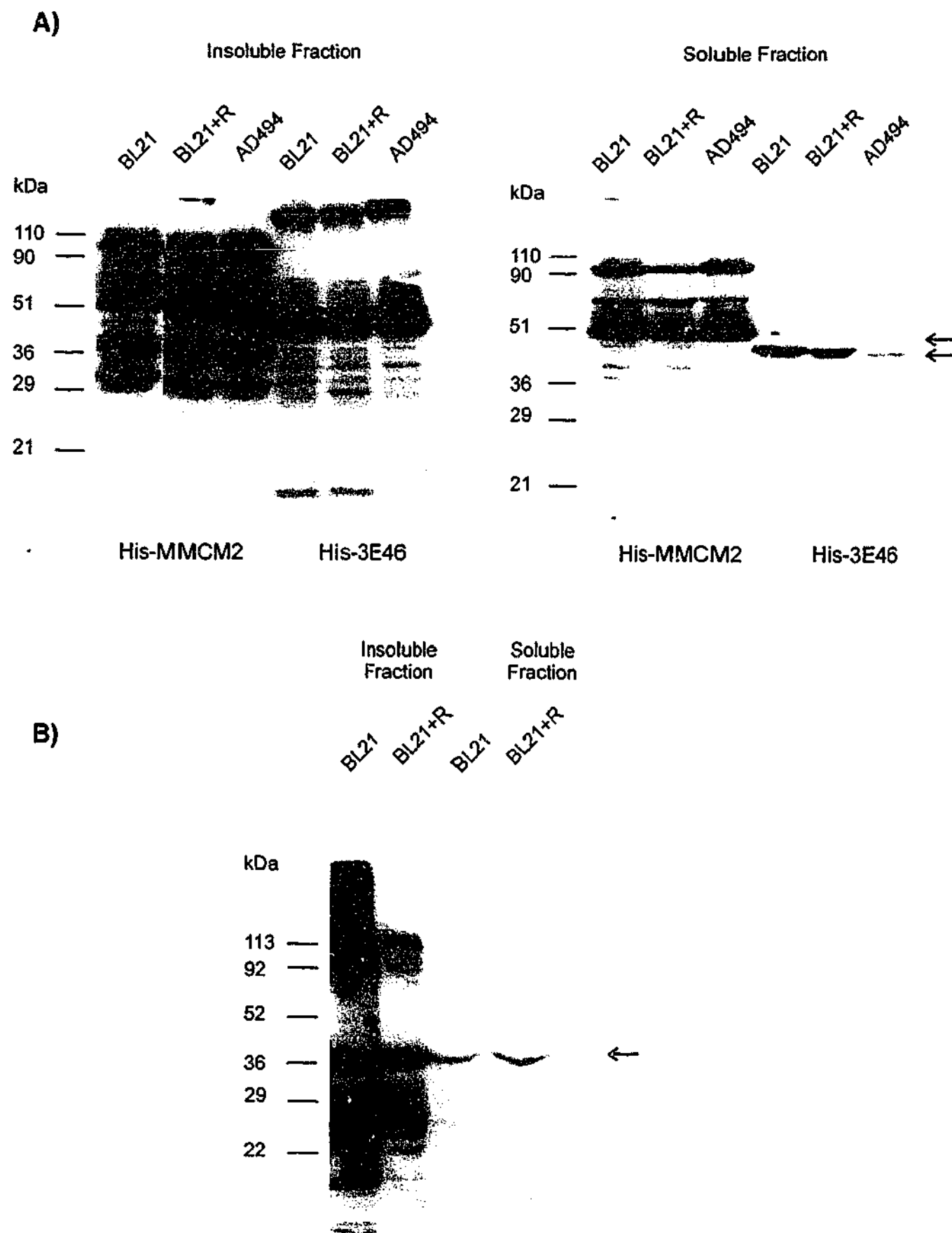


Figure 4.7: Optimisation of expression for His-3E46, His-MMCM2 and His-6C28. **(A)** Expression of His-3E46 and His-MMCM2 in the BL21 cells in the presence or absence of rifampicin (1 mg/ml) and in comparison to induction in AD494 cells. **(B)** Expression of His-6C28 in the BL21 cells in the presence or absence of rifampicin (1 mg/ml). Bacterial lysates were fractionated into soluble and insoluble fractions and proteins separated on 12.5% SDS-PAGE followed by Western blotting. Recombinant proteins were detected by the monoclonal His-antibody and anti-mouse-HRP. Arrows indicate immunoreactive bands of interest.

His-MMCM2, suggesting the protein may have been cleaved during expression.

Results of the trial inductions demonstrated that all four *serpina3* proteins could be generated as soluble proteins in BL21(DE3)pLysS. Induction at 37°C for 2 hr produced abundant soluble His-EB22.4 and His-MMCM2 and induction at room temperature for 4 hr was required to generate soluble His-6C28. Whilst trial inductions appeared to show no difference between the amount of soluble His-3E46 generated, large scale inductions at 37°C produced variable amounts of soluble protein between preparations. Therefore, His-3E46 inductions were carried out at 30°C.

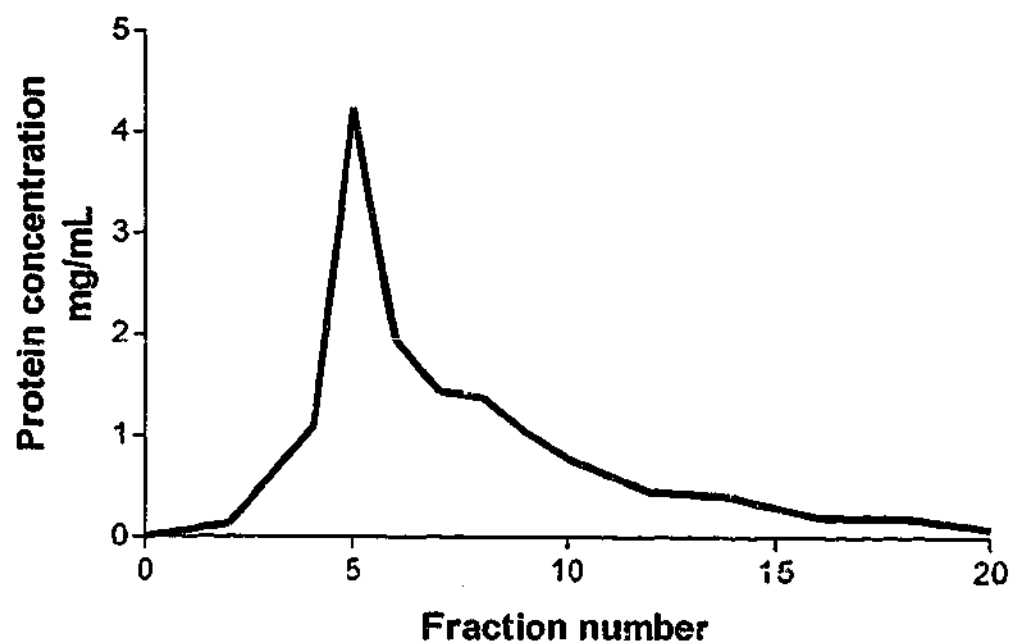
4.3.4. *Large scale expression and purification of *serpina3* proteins*

4.3.4.a. *Purification of His-EB22.4*

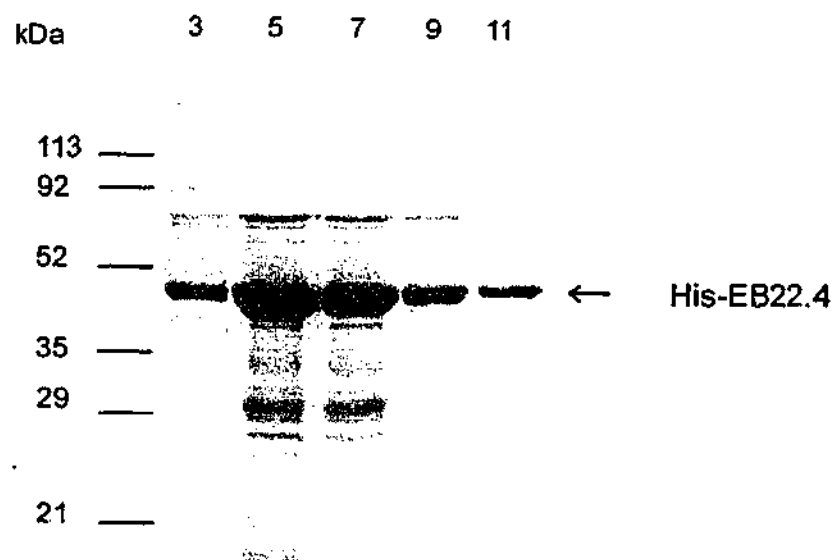
The soluble lysate of His-EB22.4 was bound to a 1 ml Ni-agarose affinity column and eluted with 125 mM imidazole (Figure 4.8(A)). Peak fractions were analysed on 12.5% SDS-PAGE, as shown in Figure 4.8(B). The Coomassie stained gel shows the elution of a dominant band near the 52kDa marker, which Western blotting with His-antibody confirmed to be the recombinant protein His-EB22.4. This procedure routinely generated approximately 10 mg of protein, with a small amount of contaminating entities observed on Coomassie stained SDS-PAGE.

Following dialysis into the anion exchange equilibration buffer A, recombinant His-EB22.4 was bound to the Mono-Q column and eluted using a linear 0-0.5M NaCl gradient. A typical elution profile is shown in Figure 4.9(A) and demonstrates the elution of a single dominant peak at a conductivity of 23.5 mS/cm. Figure 4.9(B) shows the analysis of fractions on Coomassie stained 12.5% SDS-PAGE and demonstrates that His-EB22.4 bound strongly to the Mono-Q and confirms that the dominant peak fraction contained

A)



B)



C)

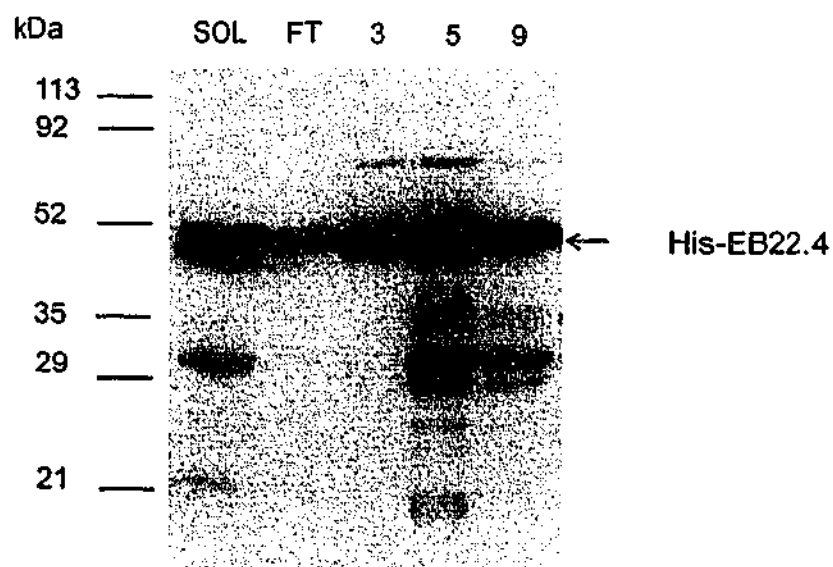
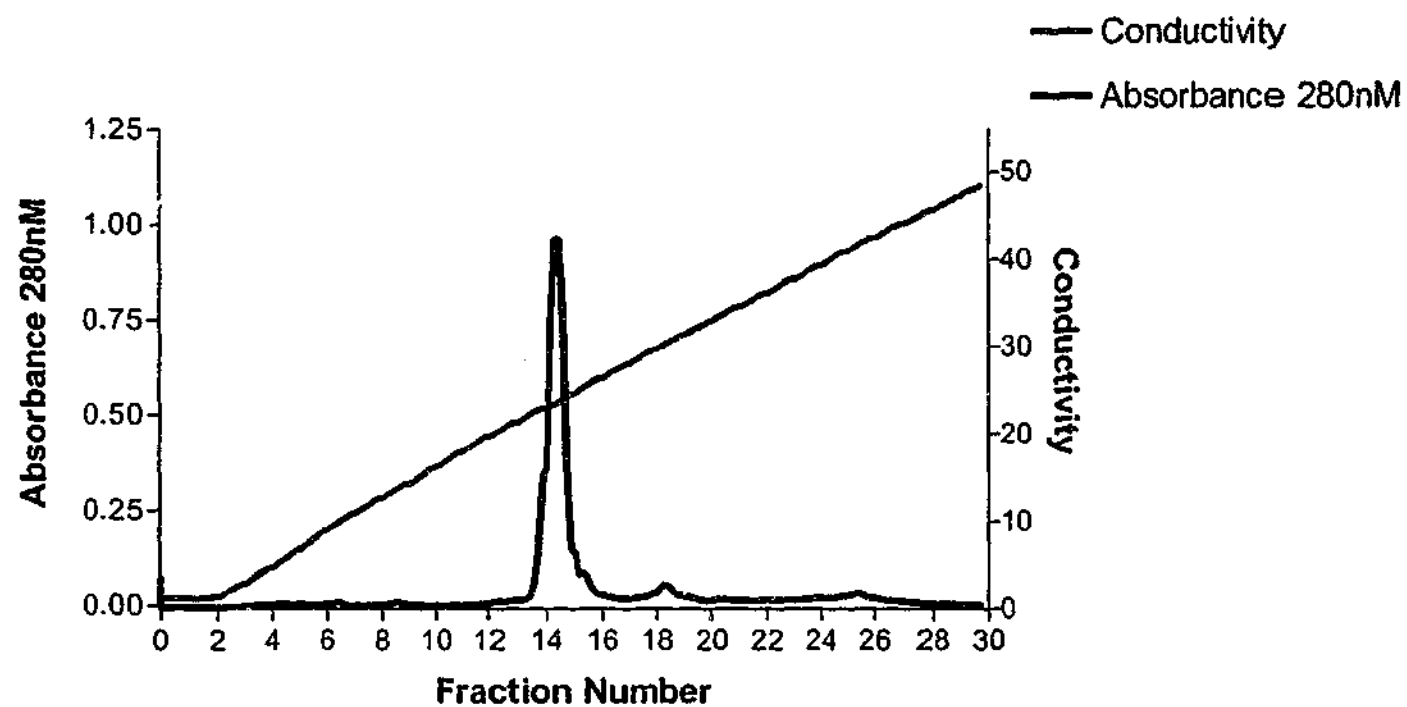


Figure 4.8: A) Elution profile of recombinant His-EB22.4 from Ni-agarose. Peak fractions were analysed on 12.5% SDS-PAGE by B) Coomassie staining and C) Western blotting with the His-antibody. The recombinant His-EB22.4 is indicated by the arrow. The soluble starting material (SOL) and flowthrough (FT) are also shown. Relative molecular mass is indicated on the left.

A)



B)

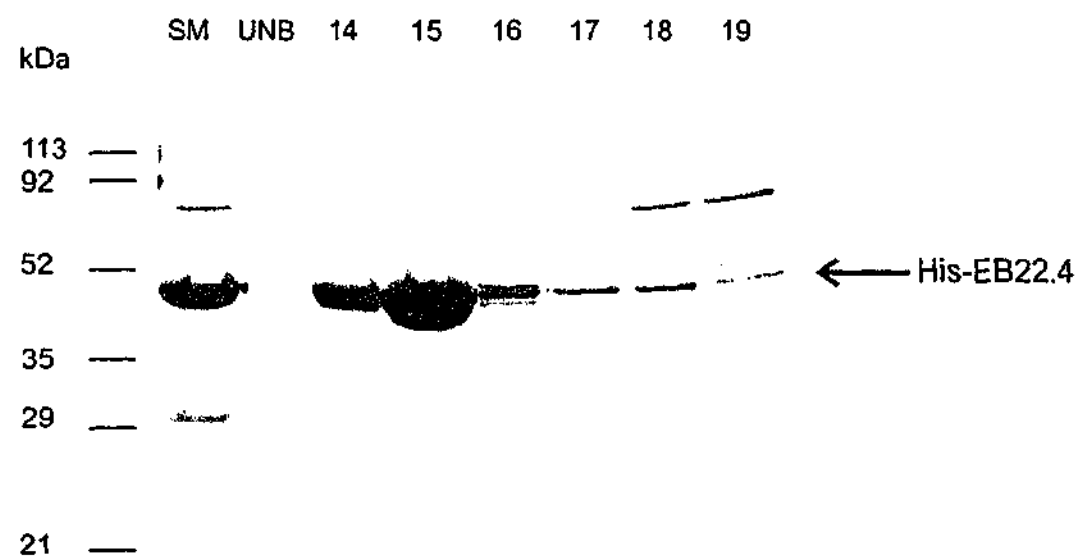


Figure 4.9: A) Elution profile of His-EB22.4 from Mono-Q. B) Peak fractions were analysed on Coomassie stained 12.5% SDS-PAGE gels. The arrow indicates the presence of His-EB22.4. Relative molecular mass is shown on the left.

recombinant His-EB22.4. Furthermore, no Coomassie-stainable contaminating bands were detected in the peak fraction, indicating that His-EB22.4 was purified to homogeneity. The total amount of protein present after this final purification stage was between 1-2 mg.

4.3.4.b. Purification of His-MMCM2

The purification of His-MMCM2 from the Ni-agarose column is shown in Figure 4.10(A). The elution profile shows a distinct peak, which correlates with the presence of recombinant protein as observed on Coomassie stained 12.5% SDS-PAGE (Figure 4.10(B)). Whilst the recombinant protein migrates at the predicted molecular weight of approximately 46kDa, a second dominant band at approximately 90kDa is also observed. The corresponding Western blot shows that the His-antibody reacts with both the 46 kDa and ~90kDa proteins (Figure 4.10 (C)). This suggests that during protein expression either a complex between His-MMCM2 and a bacterial protein, or a dimer of recombinant His-MMCM2 was generated. Ni-agarose purification yielded between 10-20 mg of protein, of which the monomeric recombinant His-MMCM2 appeared to be ~50% of the total protein present. The peak fractions were pooled and dialysed into Buffer A for anion-exchange chromatography.

His-MMCM2 displayed strong binding to the Mono-Q matrix and a single dominant peak was eluted at a conductivity of 22.8 mS/cm (Figure 4.11(A)). Peak fractions were shown to contain the recombinant His-MMCM2 by Coomassie stained 12.5% SDS-PAGE (Figure 4.11(B)). The peak fractions also contained the immunoreactive high molecular weight protein at 90kDa. The total amount of protein present after purification by anion-exchange chromatography was approximately 2-4 mg.

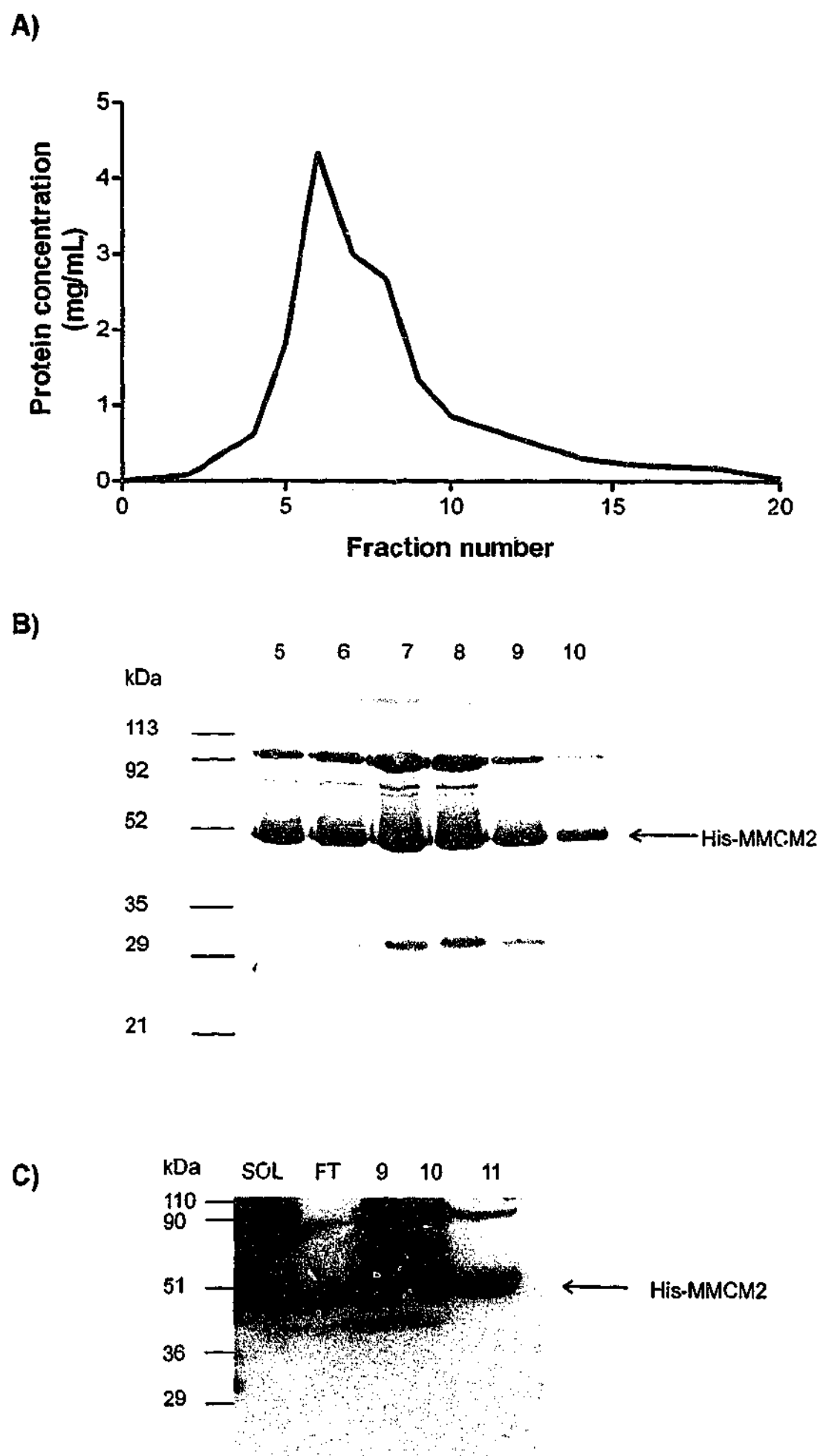
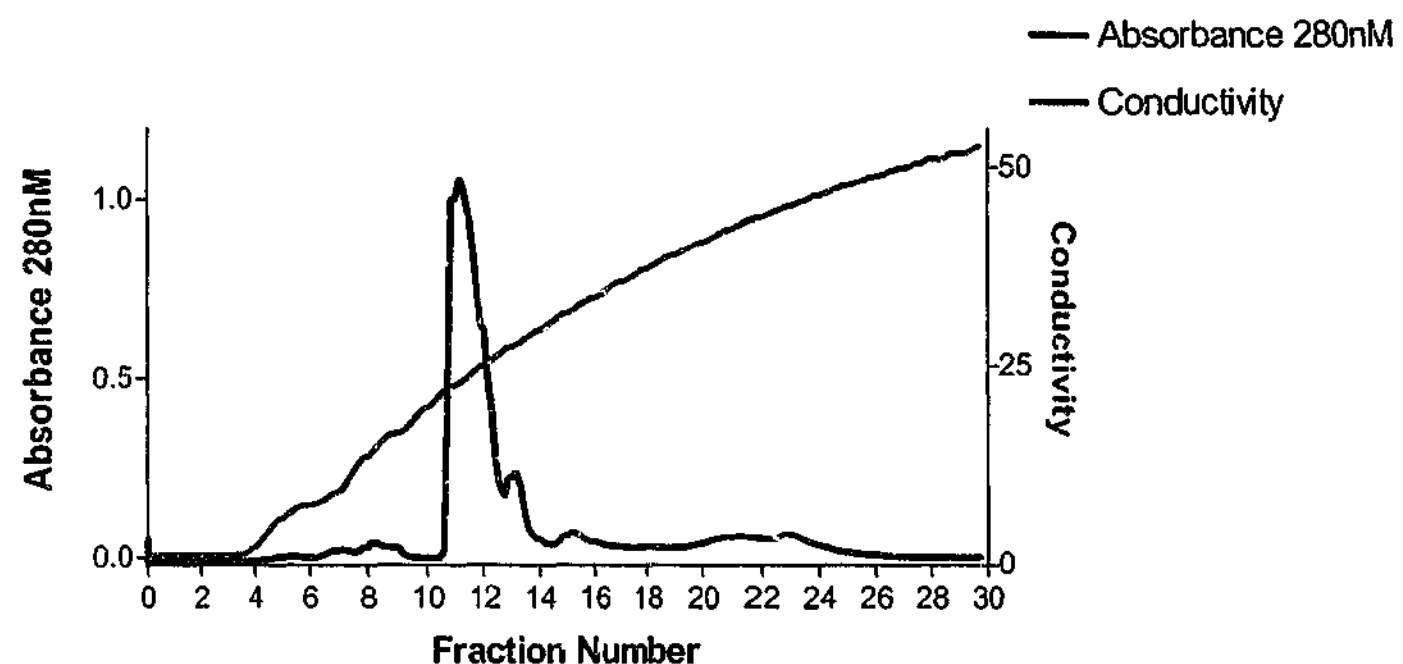


Figure 4.10: **A)** Elution profile of recombinant His-MMCM2 from Ni-agarose. Peak fractions were analysed on 12.5% SDS-PAGE by **B)** Coomassie staining and **C)** Western blotting with monoclonal His-antibody. The recombinant His-MMCM2 is indicated by the arrow. Soluble starting material (SOL) and the flowthrough (FT) are also shown. Relative molecular mass is indicated on the left.

A)



B)

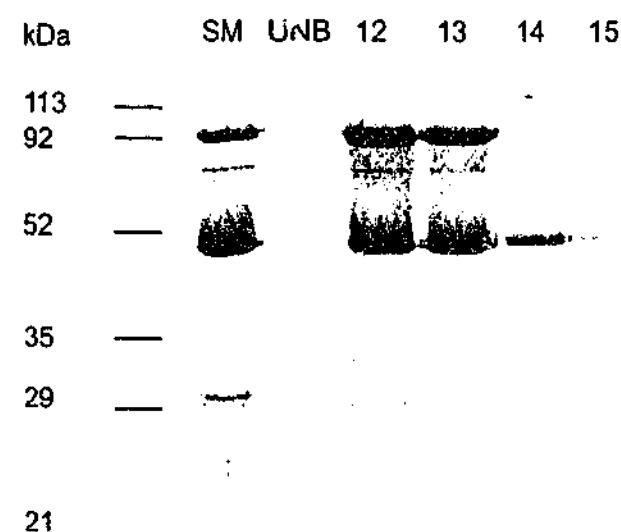


Figure 4.11: A) Elution profile of His-MMCM2 from Mono-Q. B) Peak fractions were analysed on Coomassie stained 12.5% SDS-PAGE gels. Relative molecular mass is indicated on the left.

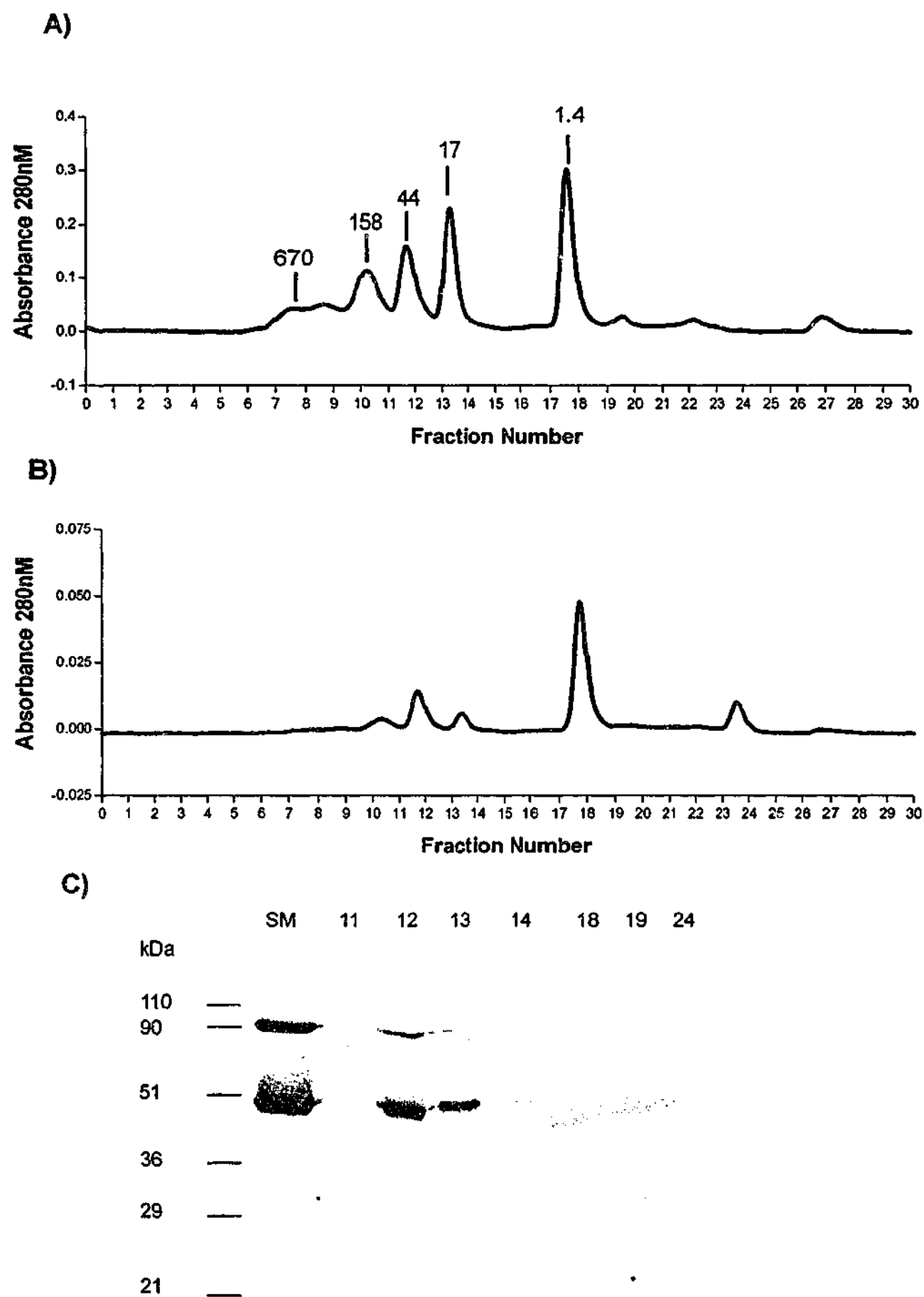


Figure 4.12: Gel filtration of His-MMCM2 on Superose 12. **(A)** Elution profile of Bio-Rad standards with relative molecular mass (kDa) shown above. **(B)** Elution profile of His-MMCM2 and **(C)** 12.5% SDS-PAGE of peak fractions eluted from Superose 12 after loading of His-MMCM2. Relative molecular mass is indicated on the left.

4.3.4.c. Gel filtration of His-MMCM2

The co-purification of a higher molecular weight protein with recombinant His-MMCM2 prompted an attempt to separate the two proteins by gel filtration. Figure 4.12(A&B) shows the separation profile on Superose 12 of the molecular weight standards and His-MMCM2. Analysis of the eluted fractions by 12.5% SDS-PAGE highlighted that gel filtration of His-MMCM2 on Superose 12 was unable to separate the 46kDa and 90kDa proteins (Figure 4.12(C)).

4.3.4.d. Purification of His-3E46

The purification of His-3E46 from the soluble bacterial lysate on Ni-agarose column is shown in Figure 4.13(A). The corresponding Coomassie stained gel indicated the dominant eluted protein to be recombinant His-3E46 (Figure 4.13(B)) and this was confirmed by Western blots probed with His-antibody (Figure 4.13(C)). Ni-agarose affinity purification of His-3E46 yielded between 5 mg of protein with only minor amounts of contaminating protein observed on Coomassie stained gels. The peak fractions were pooled and dialysed into Buffer A for anion-exchange chromatography.

Recombinant His-3E46 also eluted from the Mono-Q column as a single peak at conductivity of 25.8 mS/cm, as demonstrated by the elution profile recorded over the 0-0.5M NaCl gradient (Figure 4.14(A)). Coomassie stained 12.5% SDS-PAGE of the peak and adjacent fractions confirmed the presence of recombinant His-3E46 (Figure 4.14(B)). The gel also shows fraction 11 is free of any other Coomassie-stainable contaminating bands, indicating that His-3E46 has been purified to homogeneity. The total yield of protein after final purification by anion-exchange chromatography was approximately 500 µg-1 mg.

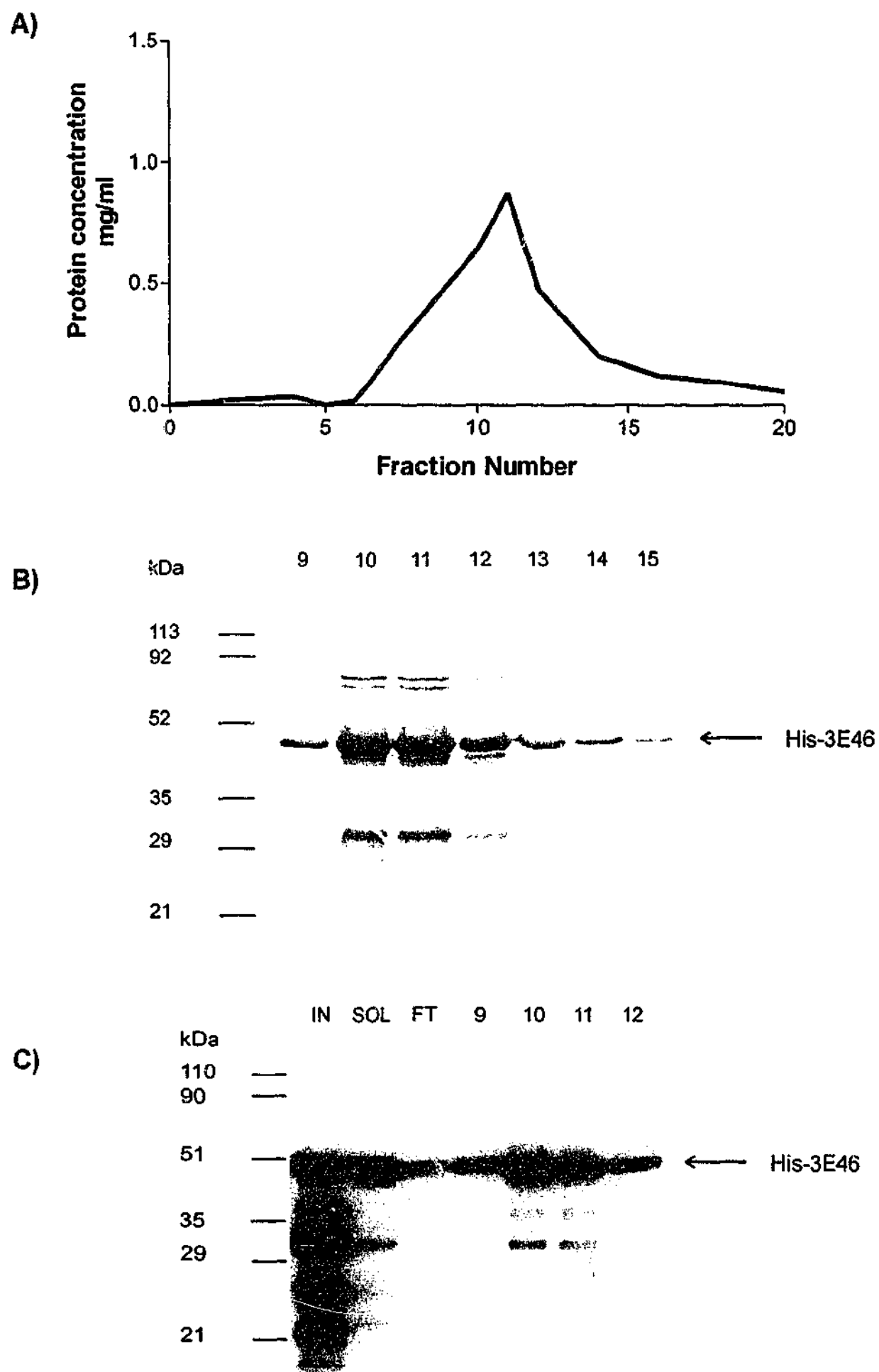


Figure 4.13: A) Elution profile of recombinant His-3E46 from Ni-agarose. Peak fractions were analysed on 12.5% SDS-PAGE by **B)** Coomassie staining and **C)** Western blotting with monoclonal His-antibody. The recombinant His-3E46 is indicated by the arrow. The insoluble material (IN), soluble starting material (SOL) and flowthrough (FT) are also shown. Relative molecular mass is indicated on the left.

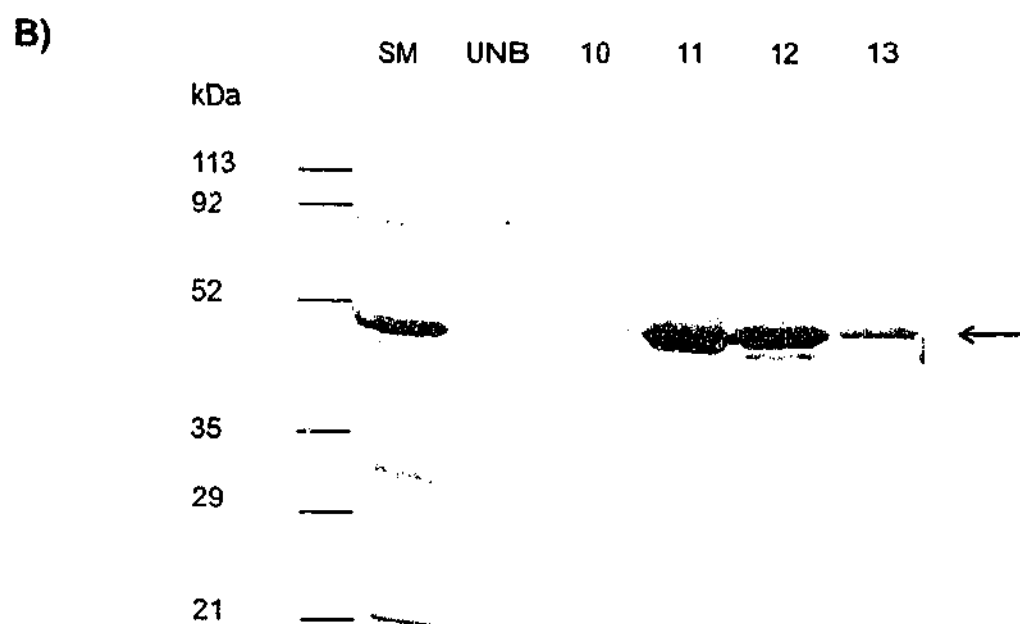
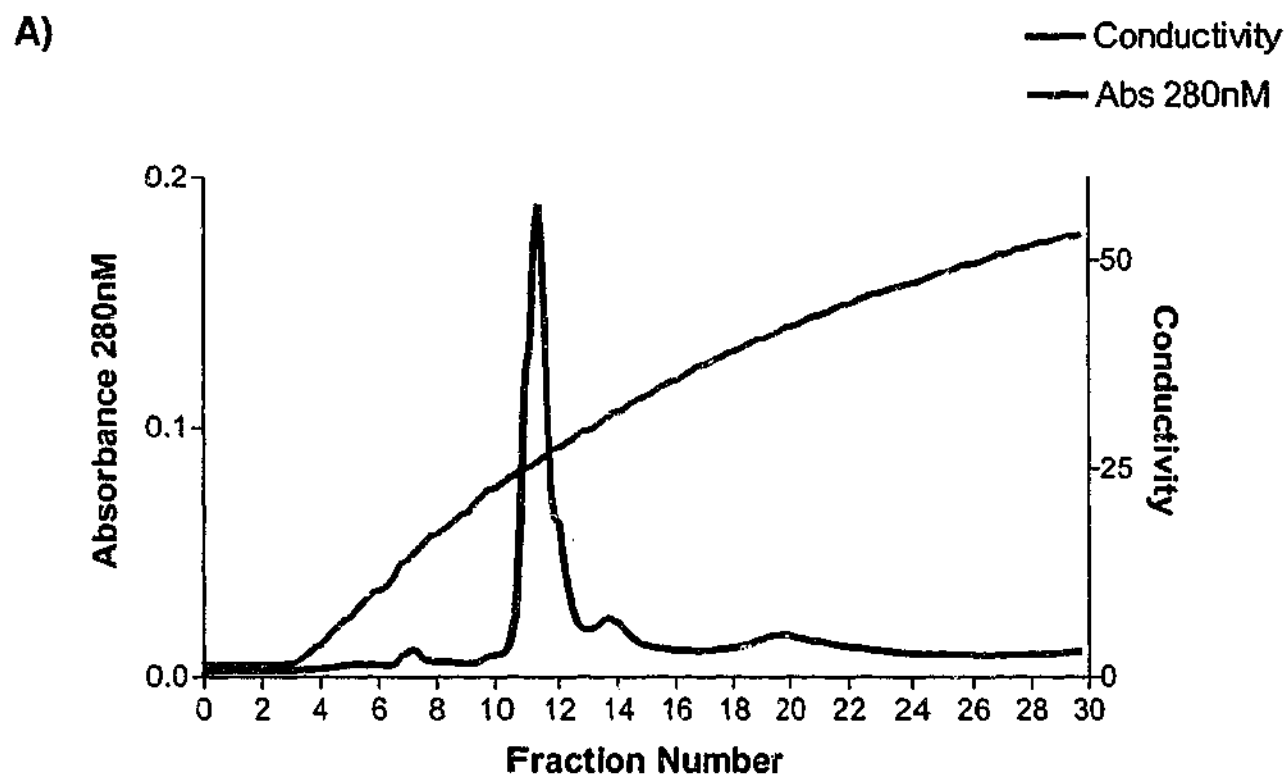


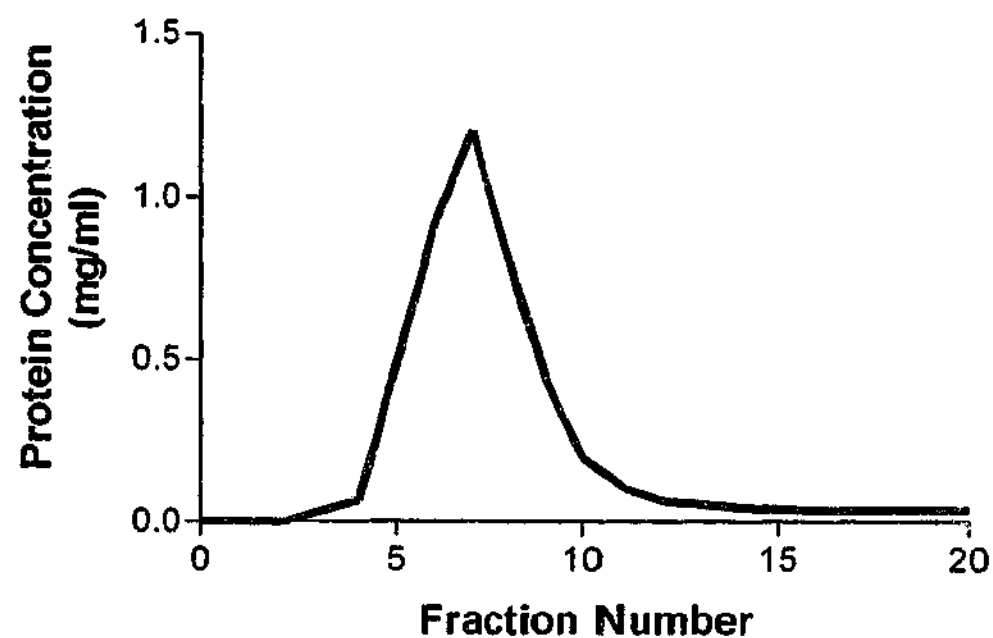
Figure 4.14: A) Elution profile of His-3E46 from Mono-Q. B) Peak fractions were analysed on Coomassie stained 12.5% SDS-PAGE. The arrow indicates the presence of His-3E46 and relative molecular mass is shown on the left.

4.3.4.e. Purification of His-6C28

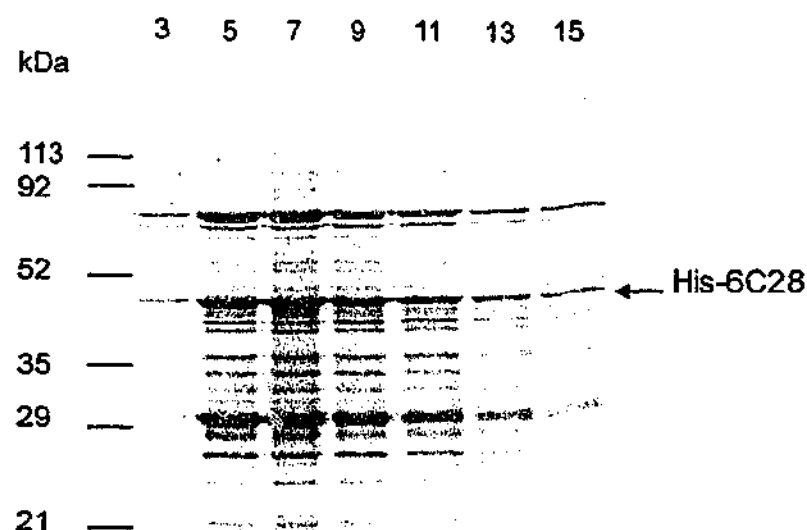
Figure 4.15(A) shows the purification of recombinant His-6C28 from the bacterial lysate by chromatography on Ni-agarose. The 12.5% Coomassie-stained SDS-PAGE gel demonstrates the presence of recombinant His-6C28 at approximately 46 kDa, but with a number of contaminating proteins (Figure 4.15(B)). The band observed at 46kDa corresponds with the immunoreactive band at this molecular weight observed on a Western blot (Figure 4.15(C)). In addition to recognising the recombinant His-6C28 at 46kDa, the His-antibody also reacted with a number of low molecular weight proteins suggesting serpin cleavage during expression. A prominent 80 kDa contaminant was not recognised by the His antibody and is therefore likely to be a bacterial protein.

Figure 4.16(A) shows the elution profile of recombinant His-6C28 from Mono-Q with a linear 0-0.5M NaCl gradient. The profile consists of multiple peaks, with the largest observed at a conductivity of 30 mS/cm, a broad elution peak between a conductivity of 38-42 mS/cm and two minor peaks at 7 and 10 mS/cm. A corresponding Coomassie stained 12.5% SDS-PAGE indicates that recombinant His-6C28 is eluted in the two minor peaks at a conductivity of 7 and 10 mS/cm corresponding to fractions 5 and 7 (Figure 4.16(B)). Western blotting also demonstrated the presence of His-6C28 in the flowthrough (Figure 4.16(C)). In addition, a small amount of recombinant protein was also detected in later elution fractions (Figure 4.16(C)). The presence of His-6C28 in the flowthrough together with its elution with a low NaCl concentration indicates weak binding to the Mono-Q matrix. The amount of purified material resulting from the bacterial expression and the two-stage purification was 50-100 µg/ml from a 1 litre culture and was pure as assessed by Coomassie stained gels. This provided enough material to commence biophysical and biochemical studies.

A)



B)



C)

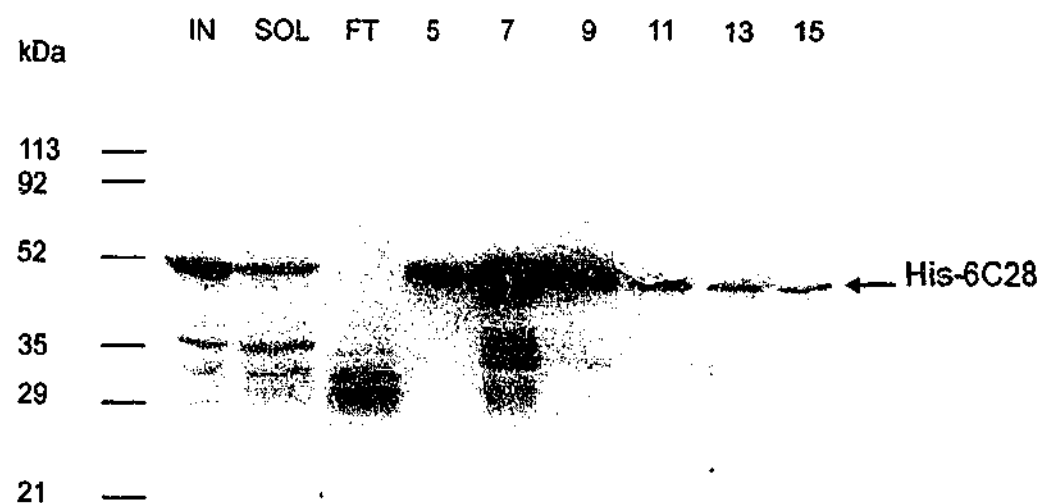


Figure 4.15: A) Elution of recombinant His-6C28 from Ni-agarose. Peak fractions were analysed on 12.5% SDS-PAGE by B) Coomassie staining and C) Western blotting with monoclonal His-antibody. The recombinant His-6C28 is indicated by the arrow. The insoluble material (IN), soluble starting material (SOL) and flowthrough (FT) are also shown.

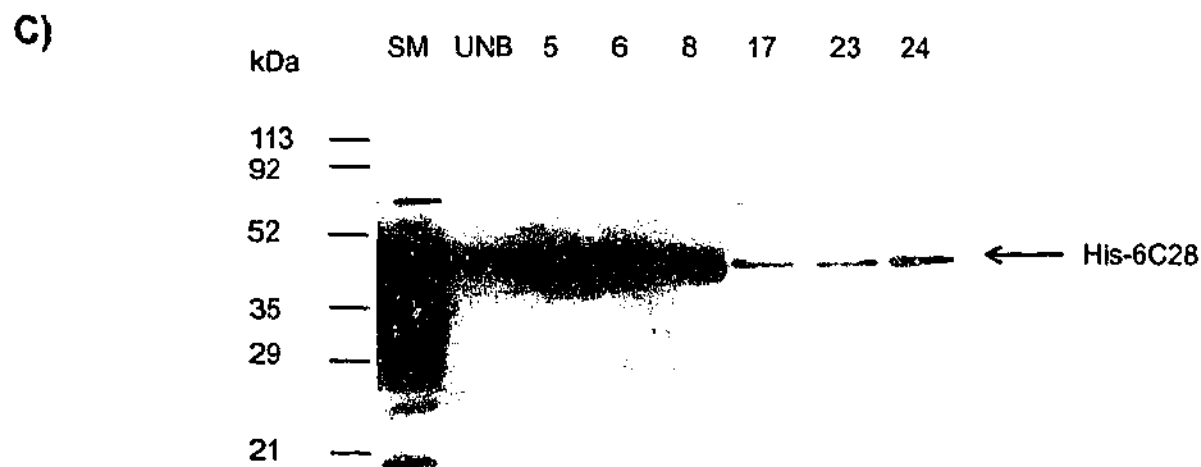
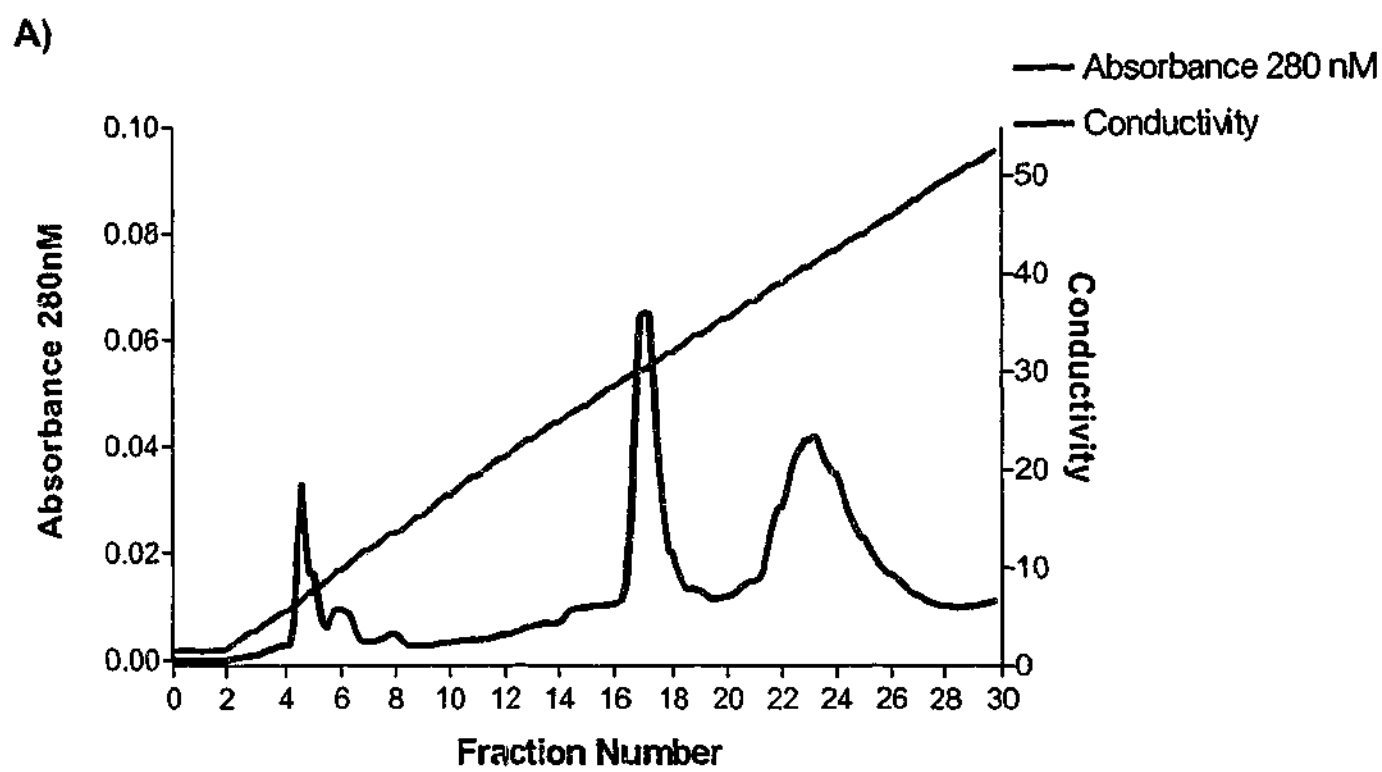


Figure 4.16: A) Elution profile of recombinant His-6C28 from Mono-Q. Peak fractions were analysed on 12.5% SDS-PAGE by B) Coomassie staining and C) Western blotting with monoclonal His-antibody. The recombinant His-6C28 is indicated by the arrow.

4.3.5. Mass Spectroscopy of His-MMCM2 90 kDa complex

The 90kDa complex which co-purified with His-MMCM2 was separated on 12.5% SDS-PAGE and the band excised and prepared for mass spectroscopy. Mass spectroscopy demonstrated that the only protein present in the complex was His-MMCM2.

4.3.6. Comparison of purified *serpina3* proteins

Figure 4.17 shows the appearance of the *serpina3* proteins after their final purification step. The four proteins migrate differently despite having similar predicted molecular weights. His-3E46 is clearly smaller than the other proteins raising the possibility that it is cleaved in bacteria. In addition, His-MMCM2 is slightly higher in molecular weight than His-EB22.4 and His-6C28.

4.4. DISCUSSION

The full length cDNA of the four *serpina3* genes, MMCM2, EB22.4, 3E46 and 6C28 were successfully isolated from mouse tissue using RT-PCR. Two of the transcripts, MMCM2 and EB22.4, were isolated using Vent polymerase however, 3E46 and 6C28 were isolated using Taq polymerase. DNA sequencing of the cDNA confirmed the nucleotide sequences for MMCM2 based on both published and GenBank sequences. Nucleotide substitutions in 3E46 and EB22.4 resulted in amino acid changes. The variations for EB22.4 were confirmed in two different EB22.4 transcripts and in both cases Vent polymerase was used. Therefore, it is unlikely that the changes were due to PCR errors and is more likely attributed to inter species variation or polymorphism. The amino acid substitutions for both EB22.4 and 3E46 are to conserved amino acids and are therefore unlikely to alter the serpin structure or function.

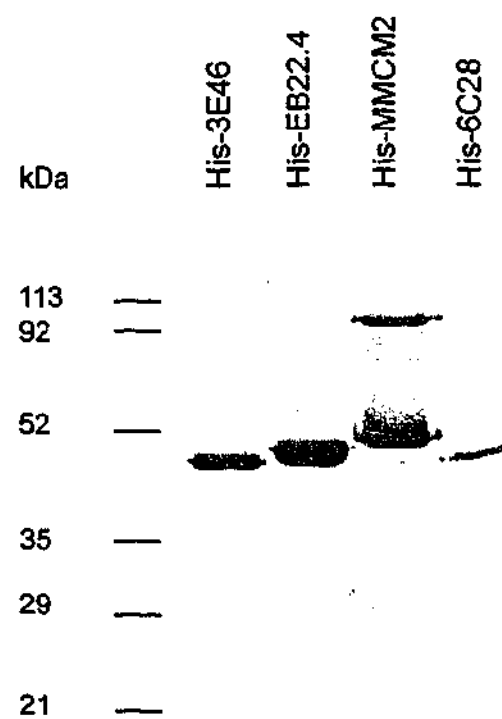


Figure 4.17: Coomassie stained 12.5% SDS-PAGE of purified recombinant serpinA3 proteins.

The full-length 6C28 transcript amplified from spleen cDNA is significantly different from the predicted sequence in the mouse genome database. All the amino acid substitutions (except Thr (307)→Met) are to conserved residues present in either members of the mouse serpin3 family or in the serpin superfamily. The 6C28 (spleen) predicted amino acid sequence is 100% homologous to a partial cDNA transcript from the NIH-3T3 cell line previously isolated in our laboratory by Dr Sharon Forsyth (unpublished results) and to two cDNA transcripts from mouse testis identified in EST databases. These sequences include the RCL region and support the existence of a 6C28 (spleen) gene in the mouse genome. This leads us to question the relationship between 6C28 (spleen) and 6C28 (genome). The sequences of these cDNA are highly homologous and share approximately 95% amino acid identity and 96% nucleotide sequence homology. The large number of amino acid differences between the two make it unlikely that 6C28 (spleen) is a polymorphism of 6C28 (genome). Database searches of the mouse genome using the RCL loop sequence of 6C28 (spleen) failed to identify a separate gene from 6C28 (genome). In addition, no mouse EST's were identified for 6C28 (genome) although the gene has been reported by Inglis *et al* (1991) and is present in the mouse genome database (Inglis and Hill 1991). A possible explanation is that 6C28 (spleen) is a variant of 6C28 (genome) resulting from strain divergence. A number of strain variations have been reported for the mouse α_1 -AT multigene locus (reviewed in Chapter 1) (Goodwin *et al*. 1997; Barbour *et al*. 2002).

There are two significant amino acid alterations in 6C28 (spleen) which affects the functionally important RCL. Of particular importance is the variation at the predicted P₁ of Cys→Leu which may alter target protease specificity. The Cys-Cys pairing at P₁-P₁' in 6C28 (genome) is unusual and its possible target protease specificity is still uncertain. Recently, Liu *et al* reported inhibitory activity towards cysteine proteases by serpin2A which also possess the Cys-Cys pairing at P₁-P₁' (Liu *et al*. 2003). The alteration of the P₁-P₁' pairing to Leu-Cys would make the serpin more like human α_1 -ACT which also possesses a Leu at P₁ and has target protease specificity towards

chymotrypsin-like proteases such as cathepsin G. A further variation at P₂', Glu→Arg, may also alter interactions with target proteases, although the exact effects of the C-terminal region of the RCL are yet to be fully elucidated and may depend on the target protease.

The two recombinant proteins, His-MMCM2 and His-EB22.4, were produced in significant quantities as soluble proteins in bacteria. The two-step purification resulted in His-EB22.4 purified to homogeneity, with Coomassie stained gels of peak fractions eluted from Mono Q consisting of a single band at approximately 46kDa. The amount of protein purified from this expression system routinely yielded between 1-2 mg of pure protein per litre of culture.

The purification of His-MMCM2 was complicated by the presence of a higher molecular weight moiety of approximately 90kDa which was immunoreactive with both the anti-His antibody and the polyclonal antibody to serpin2A (which recognises all mouse α_3 serpins). The 90 kDa complex co-purified with the 46 kDa recombinant His-MMCM2 on all chromatography media tested, including gel filtration using Superose 12. A gel filtration column which can better separate proteins between 40kDa and 100kDa is required to fully test whether the monomeric His-MMCM2 can be separated from the co-purifying complex. Mass spectroscopy of the immunoreactive 90kDa entity indicated that it consisted only of the recombinant protein His-MMCM2, leading to the conclusion that it was a serpin dimer. The nature of this dimer is obscure as it was not dissociable on SDS-PAGE excluding the possibilities of disulfide-linkage or non-covalent loop-sheet interactions. This phenomenon appeared specific for His-MMCM2 as no such complex formation was observed with any of the other recombinant proteins. The 90kDa moiety was also present when His-MMCM2 was expressed in AD494 cells, showing that the phenomenon was not bacterial strain specific. Transglutamination is a possible mechanism by which a covalent dimer could form. The transglutaminase catalyses the formation of an isopeptide bond between a glutamine and lysine residue in the presence of Ca²⁺ (Chen and Mehta 1999). However, at this stage this

mechanism is speculative. Further analysis of the 90kDa moiety is required and this can only be accomplished on its separation from the monomer.

The recombinant His-MMCM2 was purified to contain only the two proteins at 46kDa and 90kDa. The amount of the 90kDa entity varied between preparations, and appeared to be related to the quantity of recombinant protein expressed in bacteria. Longer induction times increased the relative amount of dimer to monomer therefore, induction times were limited to 2 hr. The protein yields obtained from the two-stage purification were approximately 1-2mg of recombinant His-MMCM2 per litre of culture. In comparison to His-6C28 and His-EB22.4, monomeric His-MMCM2 appeared slightly higher in molecular weight, although, no post translational modifications were expected for His-MMCM2. At this stage, the differences in molecular weight remain unexplained. Purified samples of His-MMCM2 with the least amount of high molecular weight contaminant were used for biochemical and biophysical analyses.

The third predicted extracellular serpin3 member, 3E46, was also purified by the two-stage purification process. Coomassie stained SDS-PAGE revealed that the peak fractions eluted from Mono-Q consisted of a single band of approximately 45kDa. The His-3E46 appeared to be slightly smaller than the other recombinant proteins, however, based on the varied migration of all four proteins on SDS-PAGE, it was difficult to ascertain whether this was due to cleavage in bacteria. The biophysical characteristics of recombinant His-3E46 are investigated in Chapter 5.

Abundant quantities of recombinant protein were produced for His-EB22.4, His-MMCM2 and His-3E46, although this was not the case for His-6C28. Whilst expression at room temperature did yield some soluble protein, the majority of the protein (between 95-98%) appeared in the insoluble pellet. His-6C28 differs from the other three serpin3 members under investigation as it is predicted to be an intracellular serpin (absence of a predicted N-terminal secretion peptide) and has an oxidisable residue at P₁. Other intracellular

serpins expressed in bacteria, namely serpin2A and PI-6, have also been found to be insoluble (Sun *et al.* 1995 (b); Morris *et al.* 2003).

The low levels of soluble His-6C28 in bacterial lysate may have promoted the binding of non-specific proteins to Ni-agarose resulting in significant amounts of contaminants. The pI of 6C28 is not significantly different from the other three recombinants (Table 4.2). Therefore, the unusual binding to and elution from the Mono-Q column, with some protein eluting at the start of the gradient and some at the end of the gradient, suggests that conformational differences may exist. The biophysical characteristics of recombinant His-6C28 are explored in subsequent chapters.

4.5. CONCLUSION

The bacterial expression system which utilises the pETHis(3a) expression vector and BL21(DE3)pLysS host cells was successful in producing the four serpinA3 proteins. The quantity and purity of the recombinant proteins was sufficient for biochemical and biophysical characterisation.

CHAPTER 5

BIOPHYSICAL CHARACTERISATION OF RECOMBINANT SERPINA3 PROTEINS

5.1. INTRODUCTION

One of the distinguishing features of the serpin superfamily is a dramatic conformational change which results in inhibition of the target protease (Huntington *et al.* 2000). Cleavage of the RCL by a protease results in a substantial structural rearrangement of the serpin molecule in order to open the A- β -sheet for insertion of the proximal loop (Stein and Chothia 1991). Loop insertion is accompanied by a gain in thermal stability, thus the cleaved serpin possesses increased stability over the native conformation (Carrell and Owen 1985). This change in molecular stability is termed the stressed to relaxed (S \rightarrow R) transition. Conformational change, in addition to being essential for protease inhibition, is also used by some non-inhibitory serpins to regulate activity. For example, the hormone binding serpin CBG requires cleavage of the RCL and conformational change to release cortisol at sites of inflammation (Pemberton *et al.* 1988). The metastable nature of the serpin molecule also renders it susceptible to adopting conformations which can abolish inhibitory activity. Inhibitory serpins are able to undergo polymerisation or form the latent state, both of which are relaxed conformations, and make the serpin extremely stable.

The aim of this chapter is to define the biophysical properties of the serpin3 recombinant proteins and, in particular, establish whether they undergo the conformational change necessary for protease inhibition. To investigate changes in stability the four proteins were assessed by chemical and thermal denaturation studies, namely transverse urea gradient gel (TUG) electrophoresis and circular dichroism (CD), respectively. When serpins are in

their native/stressed conformation a characteristic unfolding transition is observed by both methods. However, when the serpin is in the relaxed state, either due to loop insertion, polymerisation or latency, molecular stability is increased and therefore resists unfolding by both thermal and chemical denaturation. This chapter describes the biophysical characterisation of the four *serpina3* recombinants, His-EB22.4, His-MMCM2, His-3E46 and His-6C28.

5.2. METHODS

5.2.1. *Reactive centre loop cleavage studies*

Chymotrypsin and trypsin were prepared as described in Chapter 2. 2 µg of His-EB22.4 and His-MMCM2 were incubated at various serpin:protease ratios (w/w) (from 1:1 to 1:200), in a 20 µl reaction in 20 mM Tris pH 7.4, 2.5 mM CaCl₂ at 37°C for 2 hr. The reactions were terminated by the addition of an equal volume of 2 x reducing SDS-PAGE loading buffer and denatured at 95°C for 5 min. The cleavage reactions were analysed by 12.5% SDS-PAGE and protein detected by Coomassie staining.

5.2.2. *TUG electrophoresis of native recombinant *serpina3* proteins*

TUG gels were prepared as described in Chapter 2. To determine the conformational stability of native *serpina3* recombinant proteins, 75 µg of recombinant His-EB22.4, His-MMCM2 and His-3E46 and 10 µg of His-6C28 were electrophoresed on TUG gels. TUG gels of His-EB22.4, His-MMCM2 and His-3E46 were Coomassie stained, while His-6C28 was detected by electroblotting to a PVDF membrane and probed with monoclonal His-antibody.

5.2.3. TUG electrophoresis of cleaved recombinant *serpina3* proteins

Cleavage of the reactive site loop was achieved by incubating 75 μ g of His-EB22.4 and His-MMCM2 with trypsin at 1:20 and 1:5 protease:serpin (w/w), respectively, at 37°C for 2 hr. Control reactions in which native His-EB22.4 and His-MMCM2 were incubated in the absence of protease were set up in parallel. The reactions were stopped by the addition of PMSF to a final concentration of 20 μ g/mL and samples placed on ice. Reaction products were electrophoresed on 7.5% TUG gels and proteins visualised by Coomassie staining.

5.2.4. Circular dichroism (CD) analysis of thermal stability

CD analysis was performed on a Jasco 820s spectropolarimeter (Jasco: Easton, USA). Changes in protein secondary structure were monitored by measuring the change in ellipticity at 222nm using a 0.05cm path-length cuvette. Native and cleaved *serpina3* protein concentration was 0.2 mg/ml in PBS, pH 7.4. Thermal unfolding experiments were performed by heating at a rate of 1°C/min from 25-90°C. To generate RCL cleaved serpin, 1 mg of His-EB22.4 and His-MMCM2 were incubated with trypsin at 20:1 and 5:1 (w:w) respectively, at 37°C for 2 hr. The reactions were stopped by the addition of PMSF to a final concentration of 20 μ g/mL and samples placed on ice. Cleaved proteins were purified on a Mono-Q (as described in Chapter 4 for native proteins), to remove traces of protease.

To determine the melting temperature (T_m), the CD data was modified to the first derivative and represented as mdeg as a function of temperature. The T_m was determined as the minimum point of the curve. The reported melting temperature of His-MMCM2 and His-EB22.4 represents the average of three scans. The limiting amount of protein generated from the expression of recombinant His-3E46 and His-6C28 only allowed for a single scan.

5.2.5. Polymerisation Studies

To determine if His-MMCM2 and His-EB22.4 formed polymers, 2 µg of each recombinant protein was incubated at temperatures between 37°C and 95°C for 10 min. Samples were placed immediately on ice and 5 x native sample buffer added. The reactions were analysed on 7% non-denaturing polyacrylamide gels and protein detected by Coomassie staining.

5.2.6. In Vitro Transcription/Translation

³⁵S-labelled His-3E46 and His-6C28 were synthesised by *in vitro* transcription/translation for 90 min at 30°C from the pETHis(3a)-3E46 and pETHis(3a)-6C28 constructs, using the TNT T7 Coupled Reticulocyte Lysate System, as described in Chapter 2. Loop cleavage of the serpins was accomplished by incubating 45 µl of *in vitro* transcribed/translated protein with 1.8 µg of trypsin (His-6C28) or 1.8 µg of chymotrypsin (His-3E46). Reactions were made up to 200 µl in 20 mM Tris pH 7.4, 10 mM CaCl₂ and incubated at 37°C for 15 min. Equal amounts of *in vitro* transcribed/translated serpins (5 µl of native and 22 µl of cleaved), were analysed by 12.5% SDS-PAGE and the remainder loaded onto 7.5% TUG gels. Proteins were detected by autoradiography.

5.3. RESULTS

5.3.1. Conformational analysis of native recombinant serpin^{a3} proteins

The initial task undertaken was an analysis of the conformational state of the four native recombinant serpin^{a3} proteins. The four proteins were electrophoresed on TUG gels which are composed of a 0M-8M gradient of the chemical denaturant urea. As the concentration of urea increases across the

gel, the protein begins to unfold and its passage through the gel becomes retarded. Consequently an unfolding curve can be observed.

Figure 5.1(A) shows the appearance of His-EB22.4 on a TUG gel. The observed unfolding pattern is similar to the that seen with the prototypical serpin α_1 -AT (Lomas *et al.* 1995 (a); Elliott *et al.* 1998 (b)). At low concentrations of urea the serpin is able to traverse through the gel with minimal impediment, however, between 1-2 M urea, His-EB22.4 begins to unfold. As the concentration of urea continues to increase the ability of His-EB22.4 to migrate through the gel becomes completely retarded. A similar unfolding transition was observed with His-MMCM2 on TUG gels, with the protein beginning to unfold between 1-2 M urea, and completely retarded by 8 M urea (Figure 5.1 (B)). These data suggest that His-EB22.4 and His-MMCM2 are in a native conformation.

Figure 5.1 (C) shows the appearance of recombinant His-3E46 on TUG gels. Unlike His-EB22.4 and His-MMCM2, His-3E46 is resistant to unfolding, even at 8M urea, suggesting that the protein is extremely stable. This appearance is similar to that observed with cleaved or latent serpins. As already suggested in Chapter 4, His-3E46 has a lower molecular weight on reducing SDS-PAGE to the other three serpin3 recombinants suggesting its cleavage in bacteria. The presence of minor additional bands, that are resistant to high concentrations of urea, may represent a small amount of polymerised protein as is similarly observed with polymerised human α_1 -AT (Elliott *et al.* 1998 (b)).

The migration of recombinant His-6C28 on TUG gels is atypical of native serpins. Figure 5.1(D) displays a late unfolding transition at 5-6 M urea. This is in marked contrast to that seen with His-MMCM2 and His-EB22.4 which display an unfolding transition characteristic of native serpins. The unusual appearance of His-6C28 on TUG gels raises the possibility that the protein may be misfolded in bacteria.

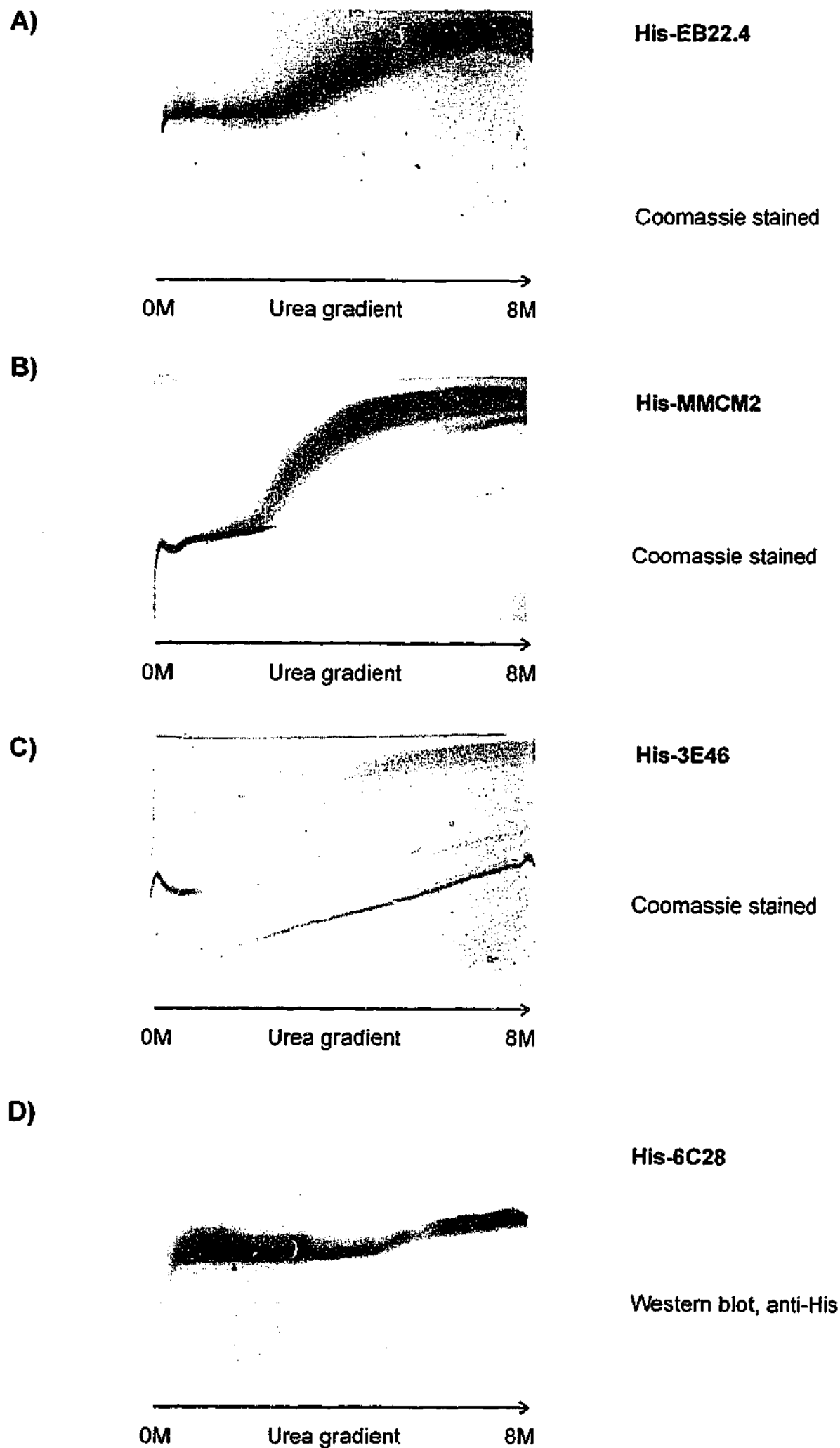


Figure 5.1: Transverse urea gradient gel electrophoresis of native recombinant serpinA3 proteins **A)** His-EB22.4, **B)** His-MMCM2, **C)** His-3E46 and **D)** His-6C28. Recombinant His-MMCM2, His-EB22.4 and His-3E46 were detected by Coomassie staining. The TUG gel of His-6C28 was Western blotted and detected with monoclonal antibody to the 6xHis-tag. The concentration of urea increases from left to right as shown.

5.3.2. Cleavage studies of His-MMCM2 and His-EB22.4

Cleavage studies were undertaken to define the conditions under which proteases could specifically cleave the RCL of His-MMCM2 and His-EB22.4. Incubations with trypsin and chymotrypsin were set up using protease:serpin ratios of 1:1 to 1:200 (w/w) and incubated at 37°C for 2 hr. The reactions were then analysed by SDS-PAGE to determine the conditions which resulted in loop-cleavage, as demonstrated by a reduction in serpin molecular mass of approximately 5kDa.

Figure 5.2 shows the effect of incubating His-EB22.4 with either trypsin or chymotrypsin at the indicated ratios. Incubation with trypsin revealed that His-EB22.4 is cleaved by the protease, with a number of low molecular weight species present at a 1:1 ratio, suggesting a number of cleavage points in the serpin molecule (Figure 5.2(A)). As the trypsin:His-EB22.4 ratio increased to 1:100, the lower molecular weight bands were no longer evident and the majority of the protein appeared as a single band at a molecular weight 5kDa lower than uncleaved His-EB22.4. A small amount of protein is also present corresponding to native His-EB22.4. By the protease:serpin ratio of 1:200 there is more native serpin than cleaved.

Figure 5.2(B) shows that His-EB22.4 is both cleaved by and able to form complex with chymotrypsin. At the protease:serpin ratio of 1:1, a small amount of low molecular weight species are present, consistent with loop-cleavage. In addition, high molecular weight entities are also observed, suggestive of complex formation. As the ratio between the protease and serpin increases more native serpin is seen, accompanied by a decrease in intensity of the His-EB22.4/chymotrypsin complex at approximately 70kDa.

As described in Chapter 4, the purified fractions of recombinant His-MMCM2 often contained a contaminating 90kDa immunoreactive protein. For cleavage studies, fractions of His-MMCM2 with minimal 90kDa contaminating complex were chosen in order to simplify interpretations. Figure 5.3(A) shows that the

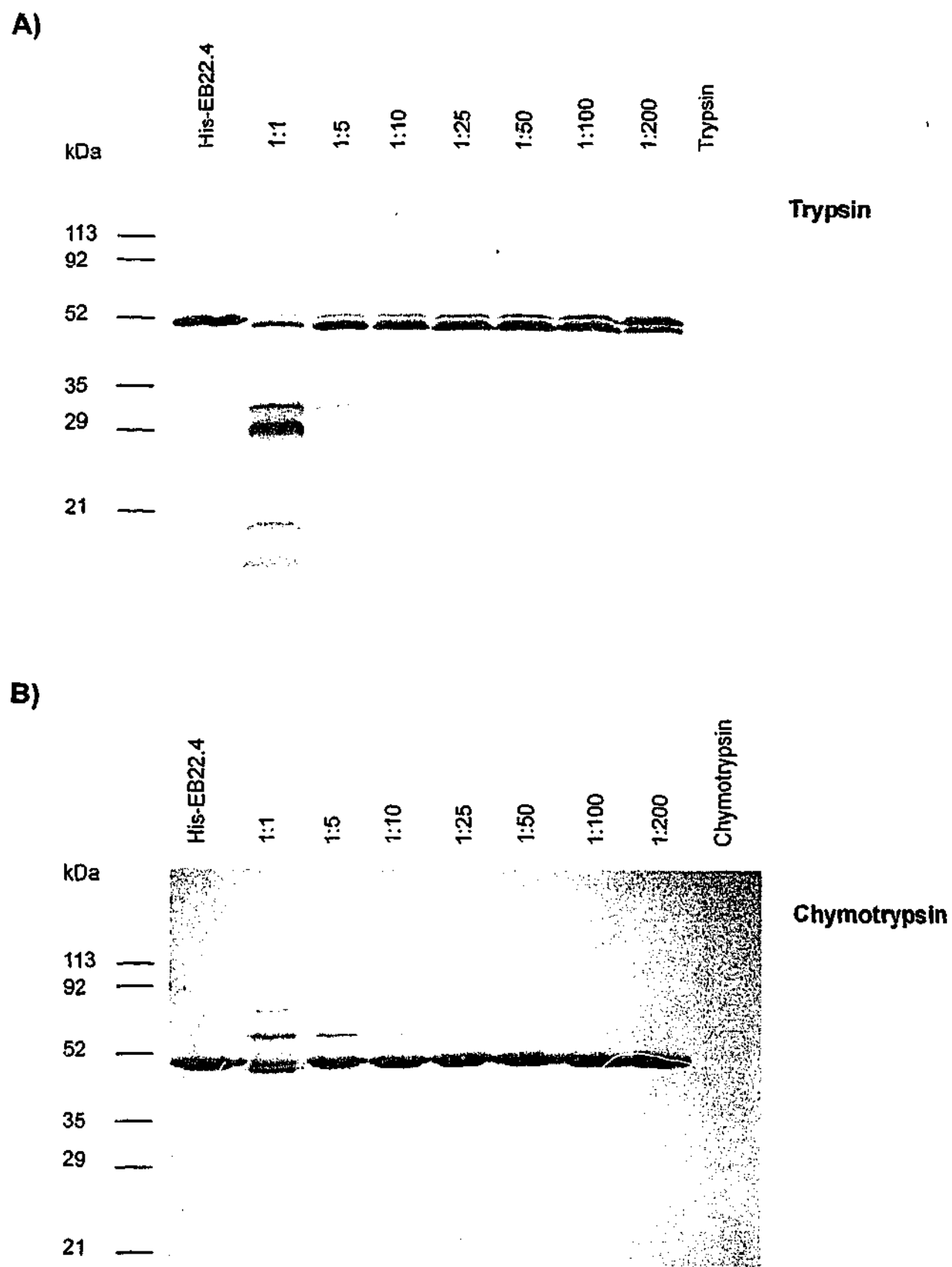


Figure 5.2: Protease cleavage studies of recombinant His-EB22.4. His-EB22.4 was incubated with (A) trypsin or (B) chymotrypsin at indicated protease:serpin ratios (w:w) for 2 hr at 37°C. Serpin cleavage was analysed on Coomassie stained 12.5% SDS-PAGE. The relative molecular mass is shown on the left.

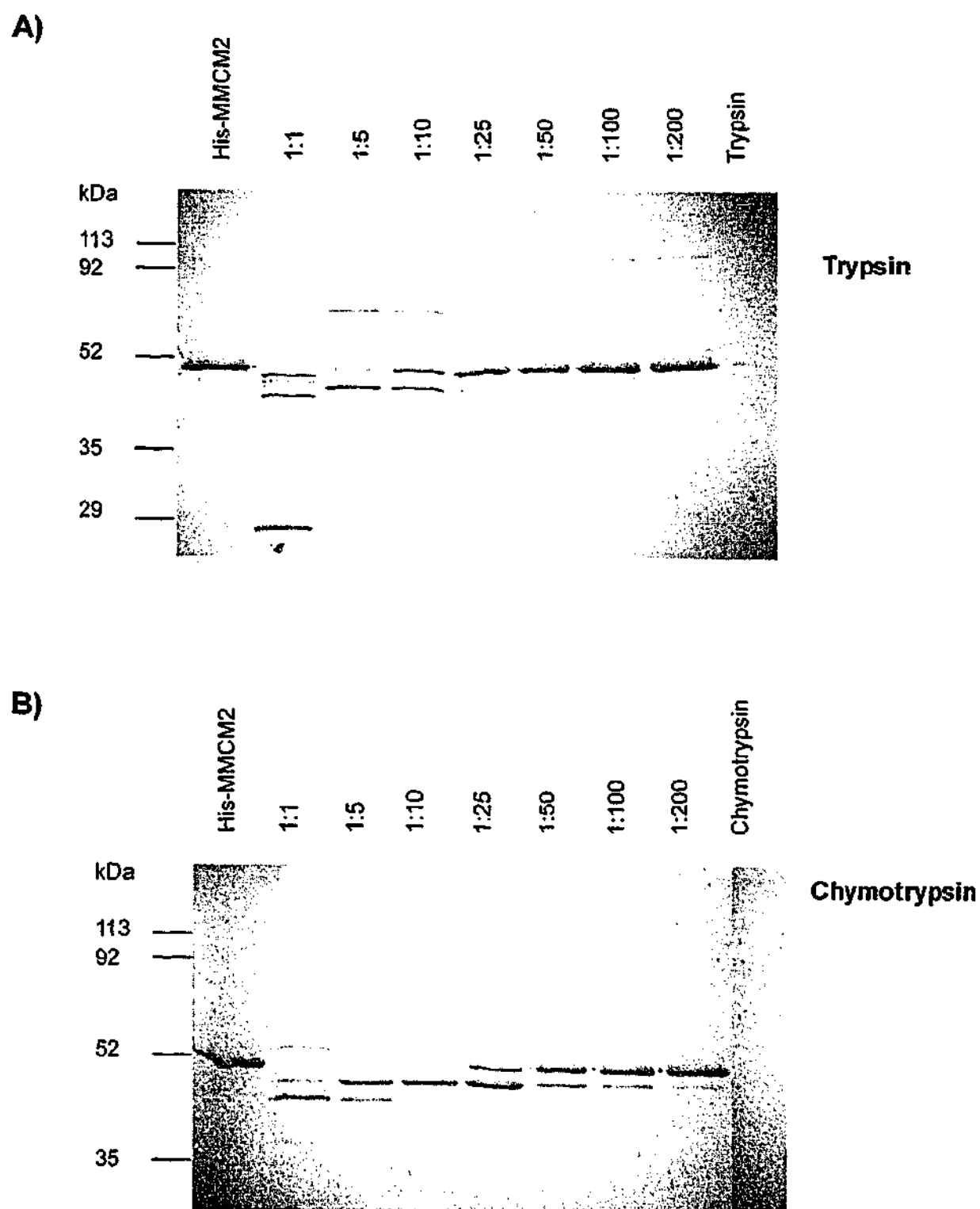


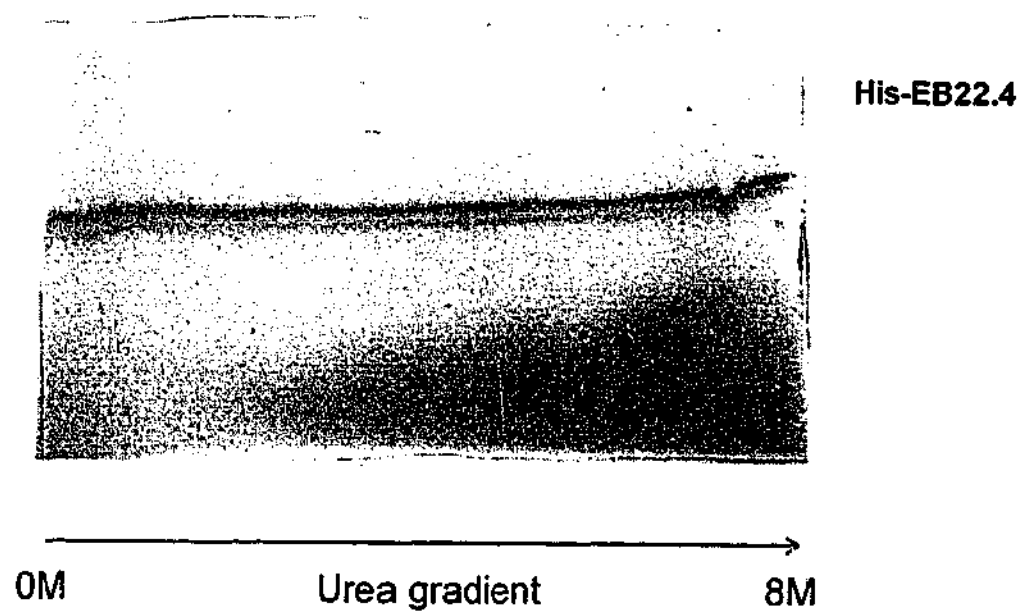
Figure 5.3: Protease cleavage studies of recombinant His-MMCM2. His-MMCM2 was incubated with **A)** trypsin or **B)** chymotrypsin at indicated protease:serpin ratios (w:w) for 2hr at 37°C. Reactions were analysed on Coomassie stained 12.5% SDS-PAGE. The relative molecular mass is shown on the left.

incubation of His-MMCM2 with trypsin produced cleaved serpin at protease:serpin ratios of 1:1 and 1:5. At a 1:1 ratio two dominant entities below the native His-MMCM2 are present at approximately 42 and 45kDa. This suggests the presence of two different cleavage points within the serpin molecule. A further cleavage product is present at 27kDa. Incubation at a 1:5 ratio resulted in the majority of the protein cleaved to a single species at around 42kDa. In addition, the presence of high molecular weight entities at approximately 70kDa was consistent with complex formation. Figure 5.3(B) shows the results of incubating His-MMCM2 with chymotrypsin. His-MMCM2 was digested by the protease at lower ratios of protease:serpin with a number of low molecular weight fragments detected. Incubation of His-MMCM2 with lower concentrations of chymotrypsin decreased the degradation of the serpin, but no serpin-protease complex was observed.

The data presented above demonstrates that, under specified conditions, trypsin was able to digest His-EB22.4 and His-MMCM2 in a manner suggestive of specific RCL cleavage. Therefore, TUG electrophoresis and CD analysis were performed on the cleaved species. Figure 5.4(A) shows trypsin cleaved His-EB22.4 on a TUG gel. Unlike the unfolding curve observed with the native protein, the cleaved form runs as a straight band across the urea gradient. This suggests that trypsin cleavage of His-EB22.4 resulted in a stable serpin, resistant to unfolding at high concentrations of urea and is indicative of RCL cleavage and loop insertion into the A- β -sheet.

The effect of treatment of His-MMCM2 with trypsin, as observed by TUG electrophoresis, is shown in Figure 5.4(B). This shows a highly stable serpin, with no observed unfolding, at high concentrations of urea, indicative of loop insertion into A- β -sheet. The presence of minor species, which are also retarded within the gradient, is consistent with the small amount of serpin/protease complex seen in Figure 5.3(A).

A)



B)

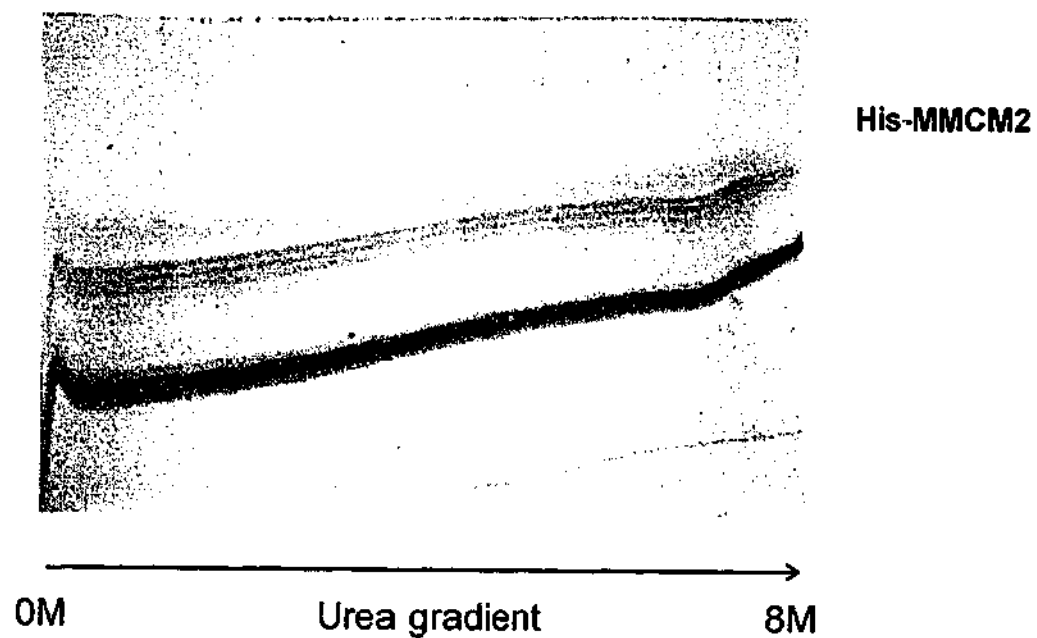


Figure 5.4: TUG electrophoresis of cleaved His-EB22.4 and His-MMCM2. The cleaved recombinant **A)** His-EB22.4 and **B)** His-MMCM2 were generated by incubation with trypsin at protease:serpin ratio (w/w) of 1:20 and 1:5 respectively, for 2hr at 37°C. Proteins on TUG gels were detected by Coomassie staining. The concentration of urea increases from left to right as shown.

Preliminary cleavage studies of recombinant His-6C28 demonstrated extensive degradation of the protein. Therefore, subsequent investigations of His-3E46 and His-6C28 were carried out using *in vitro* transcribed/translated protein.

5.3.3. Circular dichroism analysis of recombinant *serpina3* proteins

To confirm the observations from TUG electrophoresis, the four native recombinant *serpina3* proteins and the trypsin cleaved His-MMCM2 and His-EB22.4 were analysed by circular dichroism. The structural stability of the proteins was determined by measuring changes in ellipticity at 222nm whilst increasing temperature from 25-90°C at 1°C/min. The wavelength 222nm was chosen as this reflects the α -helical content of the protein.

Figure 5.5(A) shows the appearance of native His-EB22.4 and trypsin cleaved His-EB22.4 prior to CD analysis. Whilst most of the protein is cleaved there is minimal residual uncleaved/native material present. Figure 5.5(B) shows the CD data obtained for native and cleaved forms of His-EB22.4. In its native conformation, His-EB22.4 undergoes a sigmoidal transition with a loss in secondary structure. This is similar to that observed for other native serpins, such as human PAI-1 (Lawrence *et al.* 1994 (b)). In the cleaved state this transition was absent. This was consistent with the TUG analysis, and indicates that RCL cleavage of the His-EB22.4 results in a stable conformation.

Figure 5.6(A) is a 12.5% SDS-PAGE illustrating the appearance of native and trypsin cleaved His-MMCM2 prior to CD analysis. Incubation with trypsin resulted in the majority of the serpin being cleaved although some residual uncleaved/native material is present (Figure 5.6(B)). The native His-MMCM2 underwent a sigmoidal transition, with a loss in secondary structure observed at 55°C. The CD signal increases between approximately 60°C and 65°C indicating an increase in secondary structure. Whilst this unfolding pattern is different from that observed with His-EB22.4, the apparent increase in secondary structure suggests the formation of polymers at higher

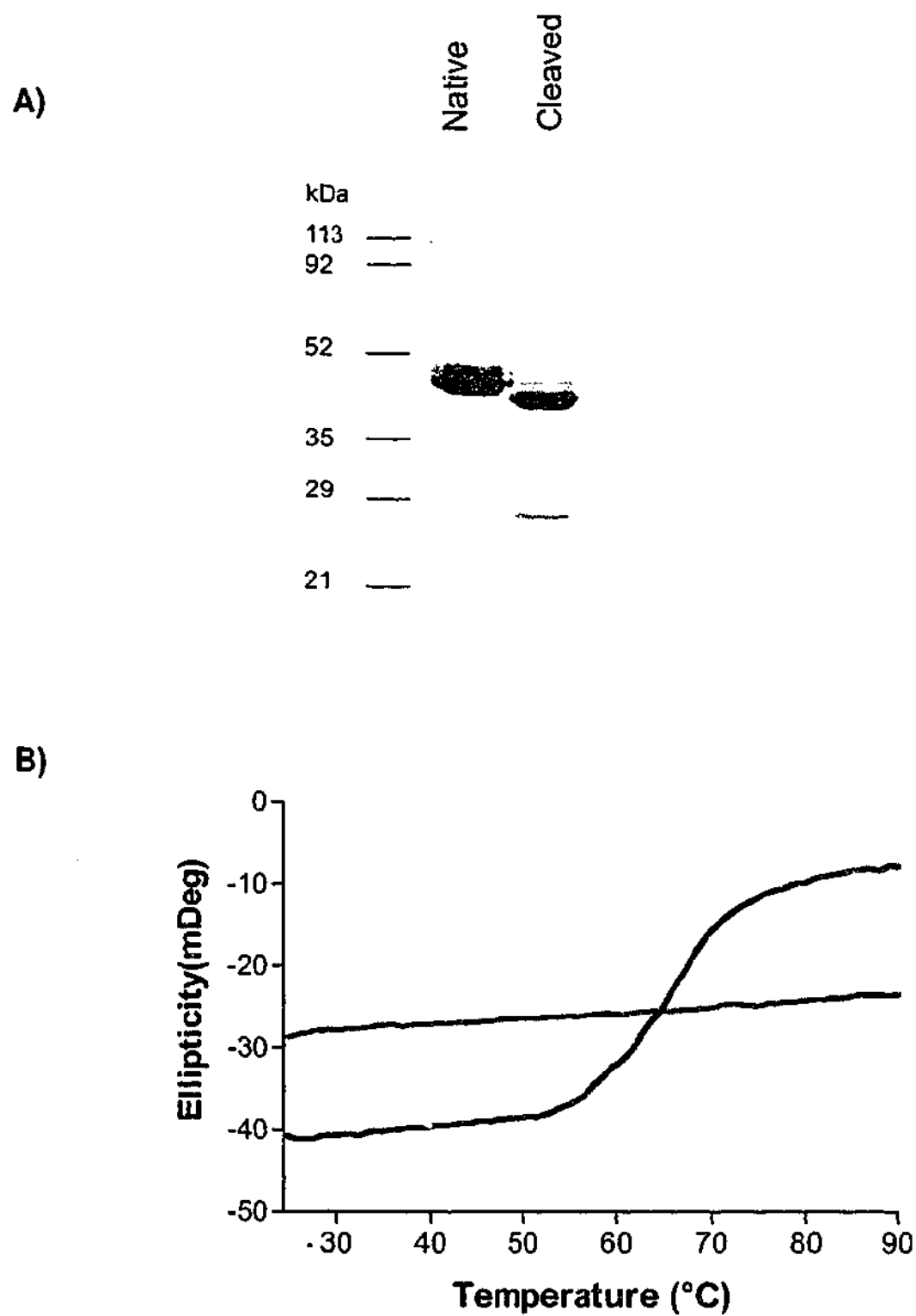
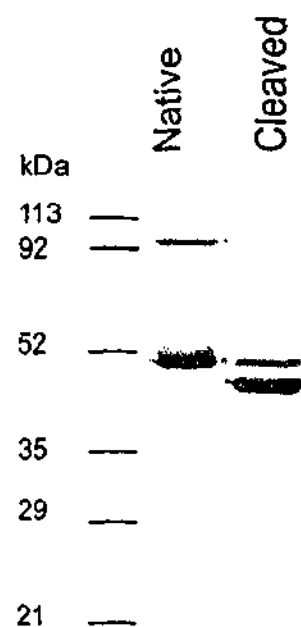


Figure 5.5: Thermal stability of native and trypsin cleaved recombinant His-EB22.4. A) 12.5% SDS-PAGE of native and cleaved His-EB22.4 prior to CD analysis. B) Thermal stability of native (black) and trypsin cleaved (red) His-EB22.4, measured as changes in ellipticity at 222nm.

A)



B)

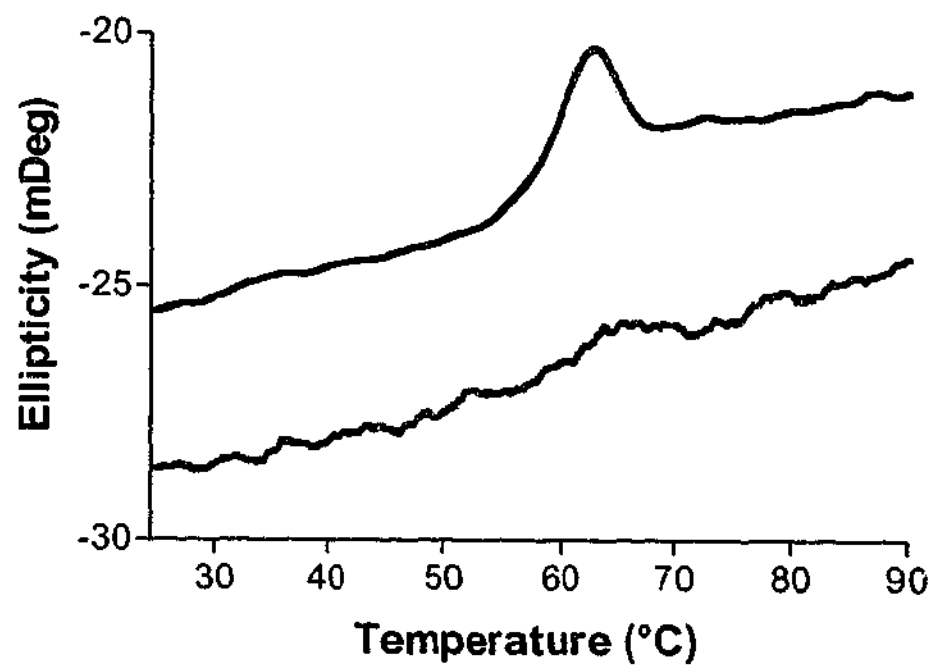


Figure 5.6: Thermal stability of native and trypsin cleaved recombinant His-MMCM2. **A)** 12.5% SDS-PAGE of native and cleaved His-MMCM2 prior to CD analysis. **B)** Thermal stability of native (black) and trypsin cleaved (red) His-MMCM2, measured as changes in ellipticity at 222nm.

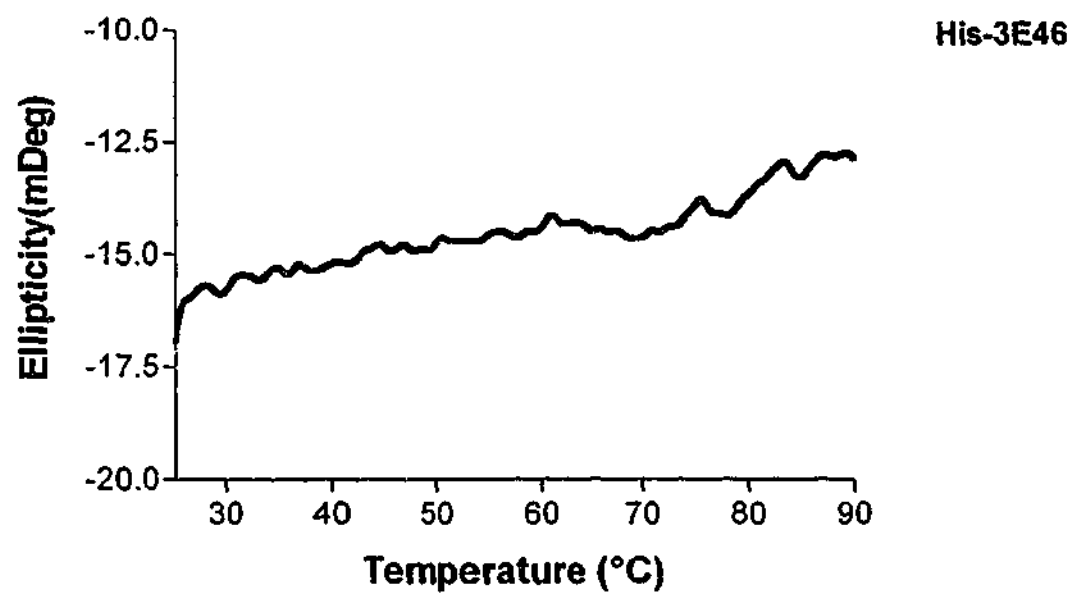
temperatures. This type of transition is also observed on thermal denaturation of human α_1 -AT and antithrombin (Dafforn *et al.* 1999; Zhou *et al.* 2003). Cleaved His-MMCM2 does not demonstrate an unfolding transition suggesting that it has been stabilised by RCL insertion. The thermal denaturation of recombinant His-MMCM2 confirms the results obtained by TUG analysis.

Figure 5.7 shows the thermal unfolding of native recombinant His-3E46 and His-6C28 as detected by changes in ellipticity by CD. No change in secondary structure was observed by thermal denaturation of His-3E46 (Figure 5.7(A)). This is consistent with bacterial recombinant His-3E46 being produced as either a RCL cleaved or latent conformation and supports the TUG analysis.

The native His-6C28 underwent a loss in secondary structure, as shown by a decrease in ellipticity between 70°C and 90°C. This agreed with the analysis by TUG electrophoresis of His-6C28 and indicates that the native serpin is unusually stable at high temperatures and requires up to 70°C to commence unfolding (Figure 5.7 (B)).

Thermal denaturation studies of native His-EB22.4, His-MMCM2 and His-6C28 allowed for the calculation of the melting temperature for each protein and the values obtained are presented in Table 5.1. The melting temperatures of His-EB22.4 and His-MMCM2, of 61.5°C and 60.2°C respectively, are comparable to human α_1 -ACT (67.7°C) and α_1 -AT (61.4°C) (Dafforn *et al.* 1999; Gooptu *et al.* 2000). In comparison, the melting temperature of His-6C28 is significantly higher than both His-EB22.4 and His-MMCM2 and other native serpins.

A)



B)

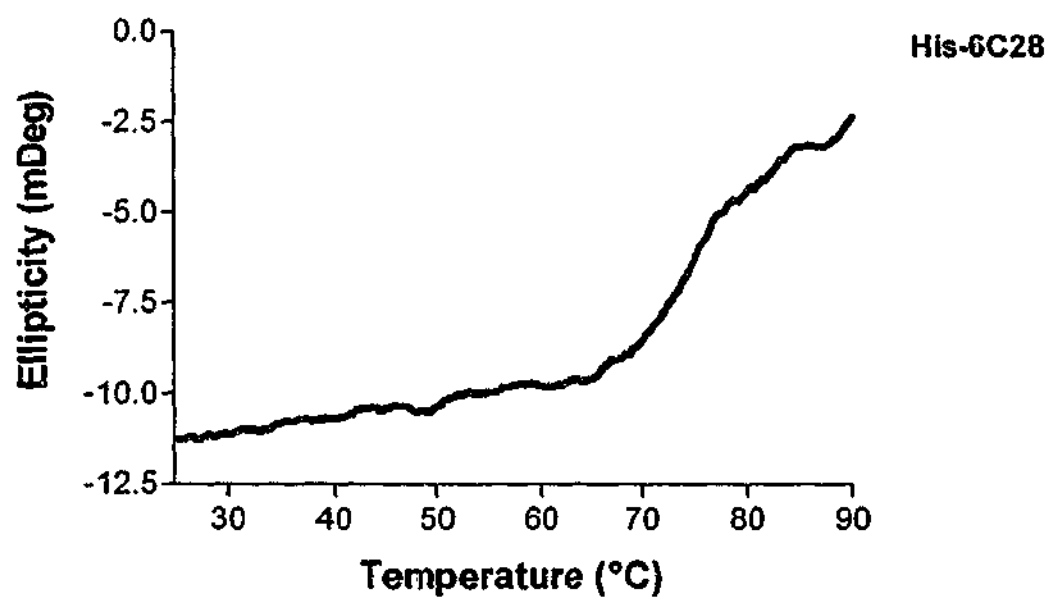


Figure 5.7: Thermal stability of A) His-3E46 and B) His-6C28. The effect on secondary structure was measured as change in ellipticity at 222nm.

Table 5.1: Melting temperatures of recombinant serpinA3 proteins

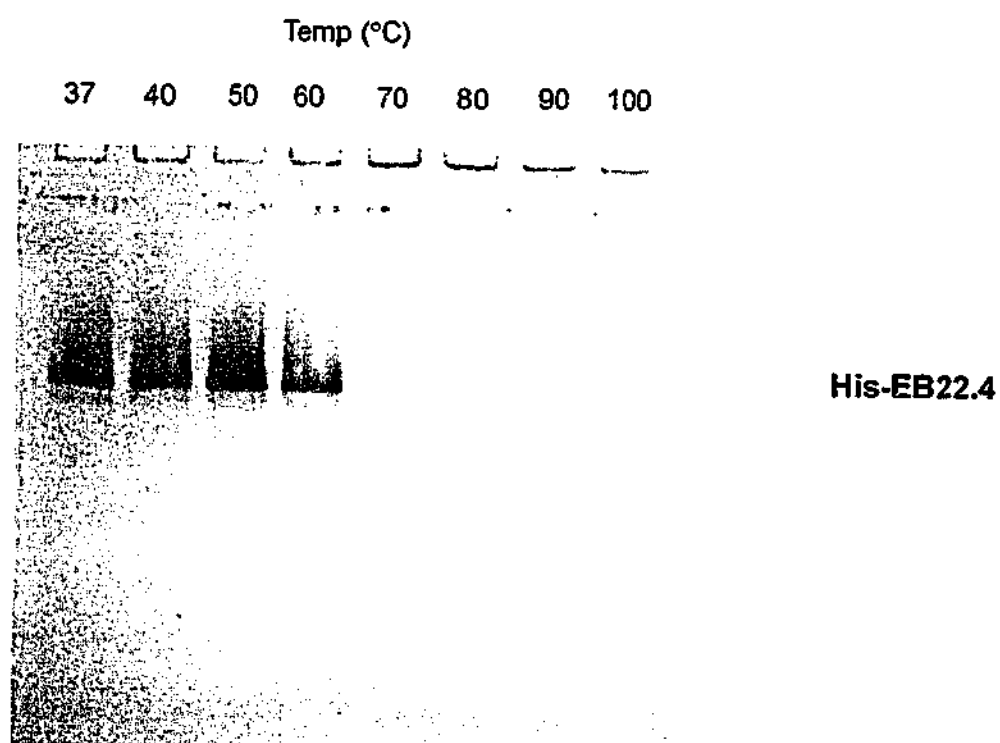
	His-EB22.4	His-MMCM2	His-6C28
$T_m(^{\circ}\text{C})$	61.5 ± 0.74 (n=3)	60.2 ± 0.25 (n=3)	75.9

5.3.4. Polymerisation studies

To determine if the recombinant proteins formed heat induced polymers, His-EB22.4 and His-MMCM2 were incubated at temperatures between 37°C and 100°C, and analysed on 7% native PAGE. Figure 5.8(A) shows the effect of heat-treatment on His-EB22.4. At temperatures between 37°C and 60°C His-EB22.4 runs as a single band. The band corresponding to the native serpin is lost at temperatures higher than 60°C and no high molecular weight polymers are evident. There is little evidence of retarded protein in the well or at the interface between the stacking and separating gels. A small amount of insoluble material is evident in the wells at 70°C and 80°C suggesting that His-EB22.4 has aggregated which may explain the absence of polymeric bands.

His-MMCM2 runs as a set of three bands on non-denaturing PAGE at temperatures between 37°C-50°C. This is similar to the appearance of recombinant human α_1 -AT on non-denaturing PAGE which also separates into several bands at these temperatures (Bottomley and Chang 1997). Figure 5.8(B) shows the presence of higher molecular weight entities between 60°C-100°C which is suggestive of serpin polymerisation. The high molecular weight species are unlike human plasma α_1 -AT polymers which separate into several bands on non-denaturing PAGE (Lomas *et al.* 1995 (a)). The His-MMCM2 high molecular weight entities are extremely large and are unable to migrate into the gel. They appear in the wells and at the interface between the stacking and separating gel. The presence of these high molecular weight entities agrees with the CD data which indicate that, at temperatures higher

A)



B)

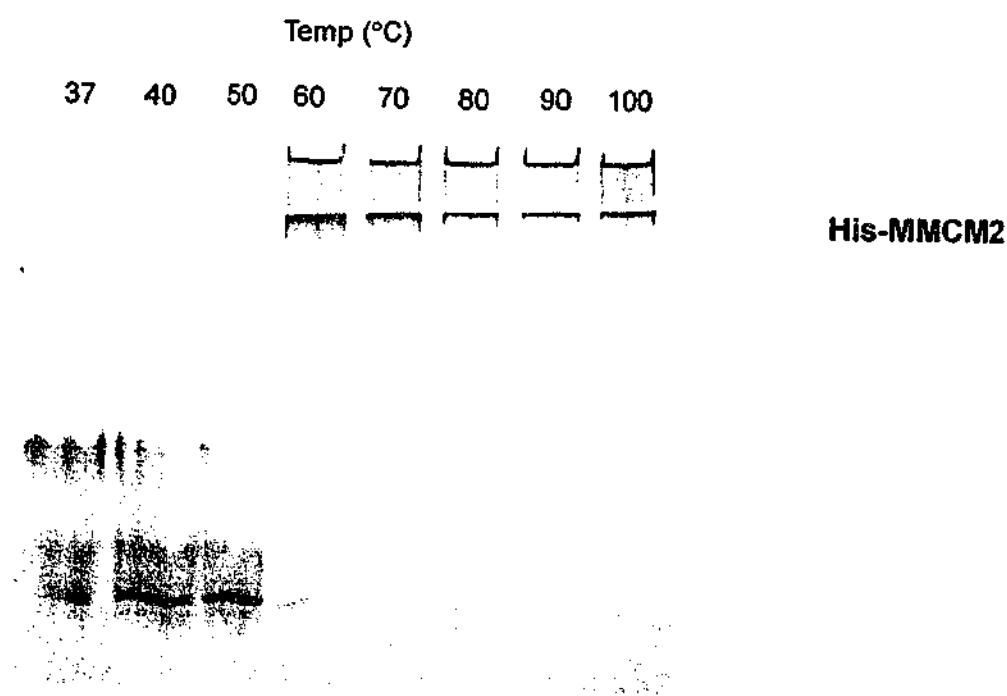


Figure 5.8: Polymerisation studies of recombinant **A)** His-EB22.4 and **B)** His-MMCM2. 2 µg of recombinant protein was incubated at the indicated temperatures for 10 minutes and then subjected to 7% native acrylamide gel electrophoresis. Protein was detected by Ccomassie staining.

than 60°C, there is a gain in secondary structure possibly due to polymerisation.

5.3.5. Conformational studies of *in vitro* transcribed/translated His-3E46 and His-6C28

Both recombinant His-3E46 and His-6C28 display atypical behaviour on conformational analysis and a possible explanation for this is that, during bacterial expression, His-3E46 is cleaved by a bacterial protease and His-6C28 is misfolded. Therefore, an alternative expression system, namely *in vitro* transcription/translation was employed to generate native, intact protein. Serpins produced by *in vitro* transcription/translation have previously been shown to form functional protease inhibitors and examples include PI-6 and MNEI (Coughlin *et al.* 1993(a); Benarafa *et al.* 2002). This method utilises the eukaryotic translation apparatus derived from rabbit reticulocyte lysate and generates between nanogram and microgram amounts of protein. In order to analyse these products the protein is labelled with ^{35}S at methionine/cysteine residues.

Figure 5.9(A) shows a 12.5% SDS-PAGE gel of *in vitro* transcribed/translated His-3E46. A dominant band at the expected size for native His-3E46, of approximately 48kDa, is evident. A number of minor low molecular weight bands are also present and possibly result from the translation of a truncated protein caused by the ribosome recognizing a methionine residue C-terminal of the start codon. Preliminary investigations with a panel of proteases showed that His-3E46 is loop cleaved by chymotrypsin (presented in detail in Chapter 6). Figure 5.9(A) shows the results of incubating His-3E46 with chymotrypsin, demonstrating a loss in molecular weight of approximately 5kDa which is indicative of serpin RCL loop cleavage.

The native and chymotrypsin treated His-3E46 were subjected to TUG gel electrophoresis and the results are presented in Figure 5.9(B & C). The *in vitro* transcribed/translated native His-3E46 shows a sigmoidal unfolding transition

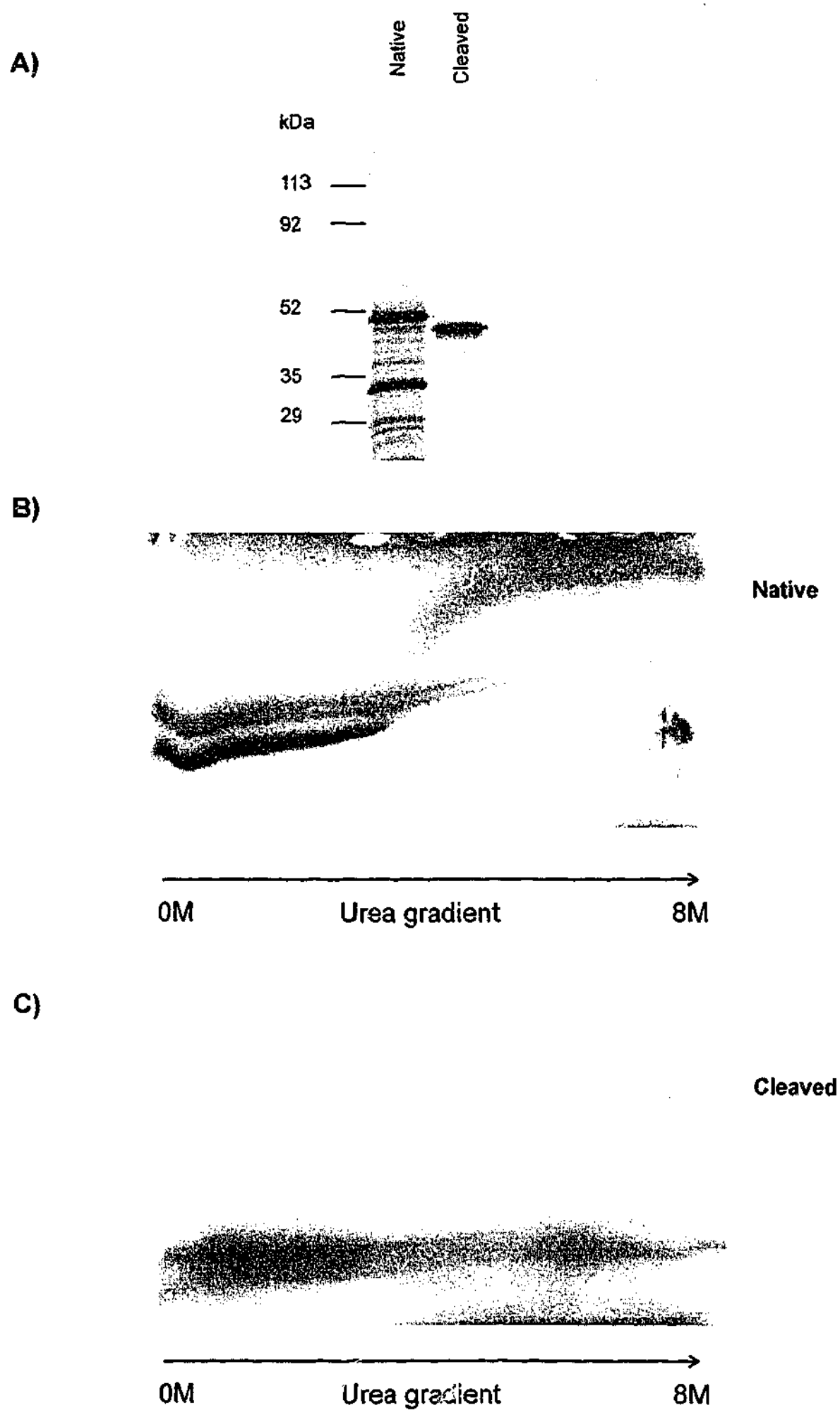


Figure 5.9: A) 12.5% SDS-PAGE of *in vitro* transcribed/translated ³⁵S-labelled His-3E46 in the absence (native) or presence (cleaved) of chymotrypsin. Post incubation of **(B)** native and **(C)** chymotrypsin-treated *in vitro* transcribed/translated His-3E46 were analysed by TUG gel electrophoresis. Proteins were detected by autoradiography.

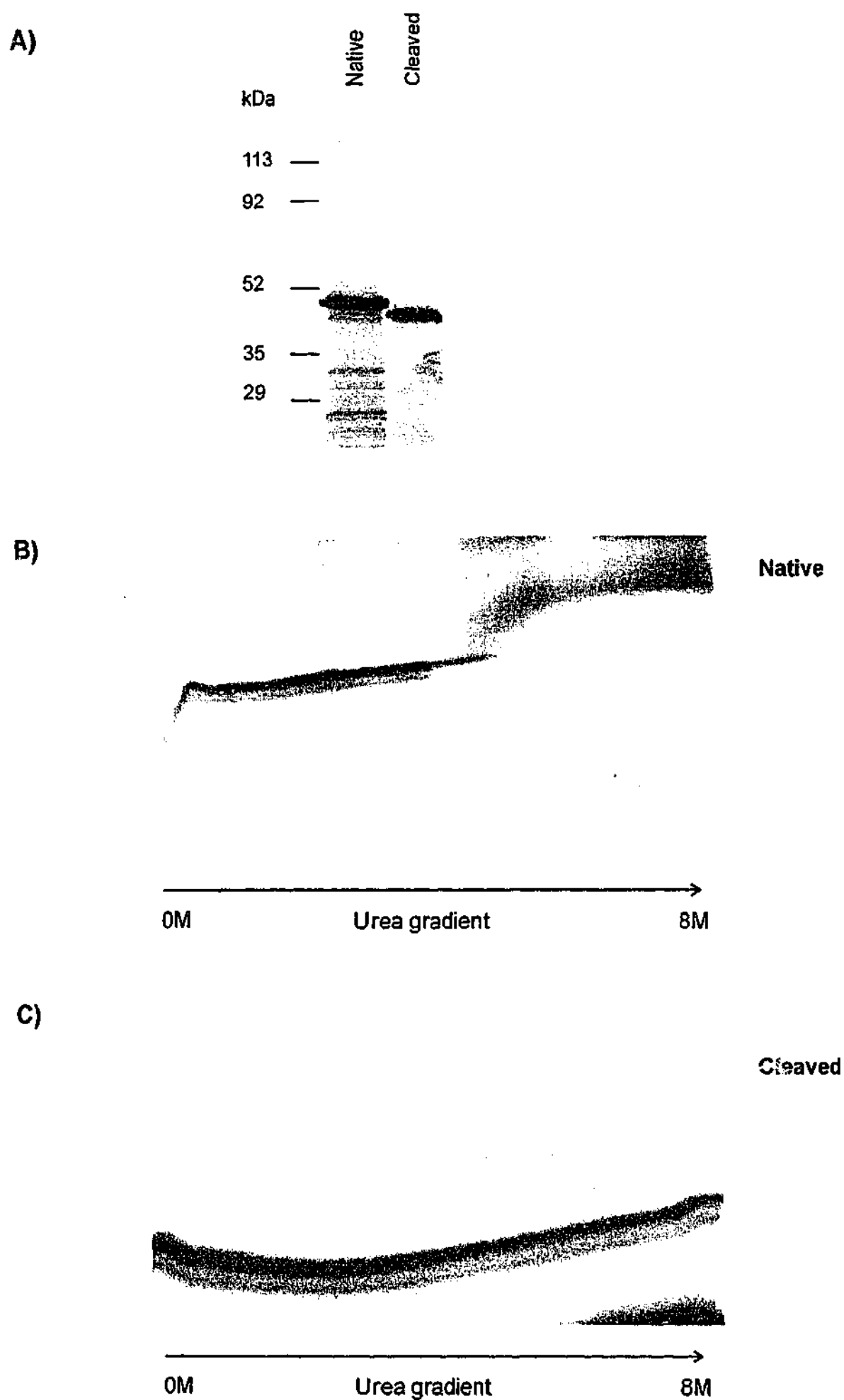


Figure 5.10: (A) 12.5 % SDS-PAGE of *in vitro* transcribed/translated, ³⁵S-labelled native and trypsin cleaved His-6C28. Post incubation, *in vitro* transcribed/translated (B) native and (C) trypsin treated His-6C28 were electrophorised on TUG gels and proteins detected by autoradiography.

as it migrates through increasing concentrations of urea. This unfolding pattern is similar to that observed for the recombinant proteins His-EB22.4 and His-MMCM2. A small amount of protein resistant to unfolding suggests the presence of some latent or loop-cleaved His-3E46 or, alternatively, may relate to the minor truncated bands observed on SDS-PAGE. Chymotrypsin treated His-3E46 migrates as a horizontal line indicating a gain in molecular stability due to RCL insertion into the A- β -sheet. These results suggest that *in vitro* transcribed/translated His-3E46 is in a native and potentially functional conformation.

Figure 5.10(A) shows a 12.5% SDS-PAGE of the product of *in vitro* transcribed/translated His-6C28. The presence of a prominent band at approximately 48kDa indicates its successful production in this system. In protease interaction studies (described in detail in chapter 6) *in vitro* transcribed/translated His-6C28 could be cleaved by trypsin. Figure 5.10(A) further shows that treatment with trypsin resulted in loop cleavage of His-6C28 with an observed loss in molecular weight of approximately 5kDa. Figure 5.10(B) illustrates that *in vitro* transcribed/translated native His-6C28 unfolds with increasing concentrations of urea. However, similar to the bacterially produced recombinant, the *in vitro* transcribed/translated His-6C28 requires a high concentration of urea (approximately 4-5M urea) to unfold the protein. This resistance to unfolding is considerably greater than other native serpins. Surprisingly, RCL cleavage of *in vitro* transcribed/translated His-6C28 results in the formation of a stable protein, as evidenced by the absence of an unfolding transition at high urea concentrations on TUG electrophoresis (Figure 5.10(C)). This suggests that His-6C28 is able to undergo the conformational change associated with loop insertion into the A- β -sheet.

5.4. DISCUSSION

Expression in the bacterial system produced four recombinant mouse serpinA3 proteins at approximately the correct molecular weight, as observed on reduced SDS-PAGE. However, it was unknown whether the four recombinants

have the structural requirements for protease inhibition. The stressed→relaxed conformational change, crucial for serpin function, provides a means to assess the structural characteristics of the four recombinants. The two techniques applied, chemical denaturation by TUG gel electrophoresis and thermal denaturation by CD, gave complimentary evidence on the structural stability of the proteins. Analysis by TUG gel electrophoresis gives qualitative data on protein unfolding in increasing concentrations of urea. CD analysis provides both qualitative and quantitative information regarding the stability of a protein, both in the appearance of the unfolding curve and the calculation of melting temperature.

Analyses of the four native recombinant proteins by both chemical and thermal denaturation studies revealed that only two of the recombinants, His-MMCM2 and His-EB22.4, behaved as inhibitory serpins. Both techniques demonstrated that the two recombinants underwent a characteristic sigmoidal unfolding pattern in the native metastable/stressed conformation. TUG gel electrophoresis showed an unfolding of both native His-MMCM2 and His-EB22.4 between 1 and 3 M urea. The thermal denaturation of native His-EB22.4 and His-MMCM2 demonstrated a loss in α -helical secondary structure at $61.5 \pm 0.74^\circ\text{C}$ and $60.2 \pm 0.25^\circ\text{C}$, respectively. These values are similar to the published melting temperatures for both recombinant human α_1 -ACT and α_1 -AT (Dafforn *et al.* 1999; Gooptu *et al.* 2000).

The thermal denaturation unfolding curve of native His-MMCM2 indicated an increase in secondary structure between temperatures 60°C and 65°C . This is reminiscent of the thermal unfolding curves of native human serpins, α_1 -AT and antithrombin, which also display a gain in secondary structure at temperatures higher than 60°C (Dafforn *et al.* 1999; Zhou *et al.* 2003). This gain in secondary structure has been attributed to the formation of heat-induced serpin polymers. Polymerisation studies of recombinant His-MMCM2 demonstrated the formation of high molecular weight species at 60°C and higher, however no polymer ladder was observed. This is different to that observed with human plasma α_1 -AT and α_1 -ACT which form a ladder of

slow migrating, high molecular polymers (Dafforn *et al.* 1999; Crowther *et al.* 2003). Previous studies have demonstrated that heat-induced recombinant polymers form higher order polymers than those derived from plasma serpins and is associated with poor resolution on non-denaturing PAGE (Bottomley and Chang 1997; Chang *et al.* 1997).

The cleavage of the RCL of His-EB22.4 and His-MMCM2 with trypsin resulted in the generation of a stable species, resistant to chemical denaturation up to 8M urea and thermal denaturation up to 90°C. The increased stability observed on cleavage of the reactive loop is characteristic of a serpin which has undergone conformational change due to insertion of the RCL. Therefore, taken together with the analysis of the native recombinants, His-MMCM2 and His-EB22.4 appear to be in their native conformation and are able to undergo the stressed→relaxed conformational change characteristic of inhibitory serpins.

Neither recombinant His-3E46 nor His-6C28 behaved as typical native serpins on TUG gels or in thermal denaturation studies. The recombinant His-3E46 appeared to be in a stable conformation, resisting chemical denaturation even at 8M urea and thermal unfolding up to 90°C. The unexpectedly low molecular weight of His-3E46 on SDS-PAGE indicated that the protein may have been cleaved in bacteria. In some preparations of recombinant His-3E46, a minor entity migrating at a slightly higher molecular weight than the dominant band was observed and may represent a small amount of non-cleaved serpin. Based on its lower molecular weight recombinant His-3E46 is likely to be RCL cleaved as opposed to other stable (relaxed) conformational states such as latent or polymerised serpin.

In contrast to the recombinant, *in vitro* transcribed/translated His-3E46 underwent a loss in secondary structure at approximately 2M urea as shown by TUG electrophoresis. This is characteristic of a serpin in the native metastable/stressed conformation. Furthermore, His-3E46 underwent a stressed→relaxed conformational change on treatment with chymotrypsin,

indicating the potential for protease inhibitory activity. This confirms the potential cleavage of His-3E46 in the bacterial expression system.

Native recombinant His-6C28 demonstrated atypical unfolding behaviour when analysed by TUG electrophoresis and CD analysis. TUG analysis showed a delayed unfolding at 5-6M urea and these observations were paralleled by a late loss in α -helical secondary structure at 74°C, as measured by circular dichroism. The unusual results observed with bacterial recombinant His-6C28 were corroborated by the *in vitro* transcribed/translated protein. This material also demonstrated delayed unfolding on TUG gels between 5-6M urea. Whilst recombinant His-6C28 is more stable than other native serpins, its delayed unfolding is inconsistent with loop cleavage. This is supported by the surprising observation that *in vitro* transcribed/translated His-6C28 is converted to a stable serpin by trypsin cleavage. Therefore, despite an atypical increased stability, His-6C28 is able to undergo the stressed→relaxed conformational change characteristic of inhibitory serpins. This also rules out the possibility that His-6C28 is in a conventional latent conformation, as latent serpins do not undergo conformational change (Mast *et al.* 1992; Lawrence *et al.* 1994 (c); Lomas *et al.* 1995 (a)).

Whilst *in vitro* transcribed/translated His-3E46 and His-6C28 produced functional serpins, the system only yields limiting quantities of protein. Further detailed analysis of the biophysical properties of both proteins requires an alternative recombinant expression system. If time had permitted, 3E46 and 6C28 could have been produced in baculovirus, yeast or mammalian cells. These approaches have the advantage over their prokaryotic counterparts of more authentic post-translational modifications and a greater likelihood of producing correctly folded proteins. Both baculovirus and *Pichia pastoris* have successfully produced several other serpins in a soluble and native form (Pemberton *et al.* 1995; Sprecher *et al.* 1995; Sun *et al.* 1995 (b); Cooley *et al.* 1998; Jayakumar *et al.* 2003) and this includes another member of the *serpina3* locus, serpin2A (Morris *et al.* 2003).

5.5. CONCLUSION

The two bacterial recombinants His-MMCM2 and His-EB22.4 possess the structural requirements for protease inhibition, as demonstrated by their ability to undergo the stressed→relaxed conformational change. The bacterial recombinant of His-3E46 is in a relaxed conformation although analysis of *in vitro* transcribed/translated His-3E46 demonstrates that the serpin can undergo the conformational change required for protease inhibition. The conformational nature of His-6C28 remains in doubt and will require further biophysical analysis.

CHAPTER 6

Protease interactions of murine serpin3 serpins

6.1. INTRODUCTION

Elucidating the cognate protease of a serpin is central to determining its inhibitory function *in vivo*. As described in the previous chapter, four mouse serpin3 proteins were shown to possess the structural requirements which would predict protease inhibitory function. The aim of this chapter is to characterise the interactions of His-EB22.4, His-MMCM2, His-3E46 and His-6C28 against a panel of serine and cysteine proteases *in vitro* to develop a profile of possible *in vivo* activity.

The amino acid sequence of the RCL determines target protease specificity with the P₁ residue being the most critical as it interacts with the S1 subsite of the target protease (Bode and Huber 2000). The regions surrounding the reactive site P₁ can influence the kinetics of interaction and have been shown to influence the division between hydrolysis and inhibition pathways (ChaillanHuntington *et al.* 1997; Plotnick *et al.* 1997).

The initial interaction between a serpin and protease is a reversible Michaelis-Menten reaction which proceeds to the formation of an acyl-intermediate. Subsequent reactions may proceed down two different pathways; 1) the completion of the inhibitory pathway with loop insertion and translocation of the protease or 2) cleavage of the serpin without protease inactivation. The balance between these two pathways is reflected in the stoichiometry of inhibition (SI). If the inhibitory pathway proceeds faster than the substrate pathway then the SI approaches 1. However, if serpin proteolysis and the substrate pathway prevail over the inhibitory pathway then the SI is greater than 1 (Gettins 2002(b)). A typical inhibitory serpin forms a covalent complex with its cognate protease which is resistant to SDS and thermal

denaturation, has a stoichiometry of inhibition close to 1 and an association constant (k_a) of $\geq 10^5 \text{ M}^{-1}\text{sec}^{-1}$ (Askew *et al.* 2001).

The diversity of residues within the RCL of the serpin3 members suggest that they may interact with a diverse range of cognate proteases. This chapter describes the protease interactions of the four mouse serpin3 proteins, His-EB22.4, His-MMCM2, His-3E46 and His-6C28.

6.2. METHODS

6.2.1. Analysis of complex formation of recombinant His-MMCM2 and His-EB22.4 using SDS-PAGE

6.2.1.a. Serine proteases

Serine proteases used in this investigation include: bovine chymotrypsin, bovine trypsin, human neutrophil cathepsin G (CatG), human leukocyte elastase (HLE), human plasma thrombin and factor Xa. The serine proteases were prepared as outlined in Chapter 2.

10 μl of each serine protease (1 μM) was incubated with 10 μl of either His-EB22.4 or His-MMCM2 (1 μM or 5 μM) in 20 μl of PBS, pH 7.4. Reactions were incubated at 37°C for 30 min and immediately placed on ice. 5 μl of 5 x SDS-PAGE reducing buffer was added. Samples were denatured at 95°C for 5 min, separated on either 10% or 12.5% reducing SDS-PAGE (as indicated in figure legends) and transferred to PVDF by Western blotting. Membranes were probed with mouse-His-antibody (1/2000) and goat-anti-mouse-HRP (1/5000). Proteins were detected with ECL reagent and visualised by autoradiography.

6.2.1.b. Cysteine proteases

The cysteine proteases used in this investigation included human recombinant cathepsin L and cathepsin V. Complex formation was investigated on SDS-PAGE by a method developed by Poh Chee Ong from the laboratory of Ass Prof Robert Pike. 2 μ M of either His-EB22.4 or His-MMCM2 was incubated with either 0.4 μ M or 2 μ M of each cysteine protease in cysteine protease buffer (0.1 M acetate, pH 5.5, 1 mM EDTA, 0.1% (w/v) Brij-35, 10 mM cysteine) in a 25 μ L reaction at 37°C for 1 hr. A control reaction of recombinant His-MENT (kindly provided by Poh Chee Ong) and cathepsin L at 1:8 molar ratio was set-up in parallel. Samples were immediately placed on ice. To resolve the complexes the method utilises a low pH 2x SDS-PAGE loading buffer (0.125 M Tris-HCl, pH 4.3, 4% (w/v) SDS, 20% (v/v) glycerol, 0.2 M DTT, 0.02% (w/v) bromophenol blue). The low pH loading buffer was added to the samples and heated at 80°C for 2 min. Proteins were separated on 12.5% SDS-PAGE and transferred to PVDF membranes for Western blotting. Membranes were probed with His-antibody (1/2000) and goat-anti-mouse-HRP (1/5000). Proteins were detected with ECL reagent and visualised by autoradiography.

6.2.2. Complex formation of *in vitro* transcribed/translated *serpina3* proteins

³⁵S-labelled His-3E46 and His-6C28 were synthesised by *in vitro* transcription/translation for 90 min at 30°C from the pETHis(3a)-3E46 and pETHis(3a)-6C28 constructs, using the TNT T7 Coupled Reticulocyte Lysate System (as outlined in Chapter 2). 2.5 μ L of *in vitro* transcribed/translated protein was incubated with either 20 ng or 100 ng of each serine protease in a 20 μ L reaction in PBS for 10 min. Samples were placed on ice, 5 x SDS-PAGE sample buffer dye added, heated at 95°C for 5 min and electrophoresed on 10% SDS-PAGE. The gels were dried and ³⁵S-labelled proteins detected by autoradiography. Incubations with cysteine proteases were carried out as for

serine proteases, except that cysteine protease buffer was used in reactions and samples were prepared for SDS-PAGE as described in section 6.2.1.b.

6.2.3. *Enzymes, assay buffers and substrates*

Chymotrypsin was assayed in 25 mM Tris, pH 8.0, 10 mM CaCl_2 and cathepsin G in 20 mM Tris-HCl, pH 7.4, 500 mM NaCl, 0.1% polyethylene glycol (PEG) 8000. The activity of chymotrypsin and cathepsin G was measured with the substrate N-succinyl-Ala-Ala-Pro-Phe-pNA at a final concentration of 100 μM and 1 mM respectively. The release of the chromogen pNA was monitored by absorbance at 405 nm. Elastase was assayed in 0.1 M HEPES, 0.5 M NaCl, pH 8.0 and activity measured using the substrate N-methoxy-Ala-Ala-Pro-Val-pNA at a final concentration of 200 μM . Release of pNA was measured at an absorbance of 405 nm. Trypsin activity was assayed in 25 mM Tris, pH 8.0, 10 mM CaCl_2 with the substrate Z-Phe-Arg-AMC. The release of the fluorogenic leaving group, AMC, was measured using excitation/emission wavelengths of 370/460 nm. Enzymatic measurements were performed on a Polarstar Optima (BMG: Offenburg, Germany) for trypsin and chymotrypsin and elastase and cathepsin G were measured on a Thermomax microplate reader (Molecular Devices: California, USA).

6.2.4. *Active site titration of proteases*

The active concentration of human neutrophil elastase, cathepsin G and trypsin was determined by titrating the proteases against a known active amount of purified human plasma α_1 -AT, kindly provided by Dr. Steve Bottomley (see Section 6.2.6). Plasma α_1 -AT activity was titrated against active site-titrated bovine chymotrypsin.

6.2.5. BSA coated-microtitre plates

A solution consisting of 0.2% BSA (w/v), 0.05% PEG-8000 (w/v) and 0.01% Triton X-200 (v/v) was added to each well of a 96-well microtitre plate. The plates were incubated overnight at room temperature and washed three times in ddH₂O. The plates were placed face-down and left to air dry.

6.2.6. Determination of the Stoichiometry of Inhibition

Prior to determining the stoichiometry of inhibition values with His-EB22.4 and His-MMCM2 the proteases were titrated against a known active amount of human α_1 -AT. These reactions with α_1 -AT were performed as described below for His-MMCM2 and His-EB22.4.

Stoichiometry of inhibition (SI) values for His-EB22.4 were determined with chymotrypsin, cathepsin G and elastase with His-EB22.4 in 60 μ L reactions. Different concentrations of His-EB22.4 (0.5nM-4nM) were incubated with a constant concentration of chymotrypsin (2nM) at 37°C for 4hrs and residual activity assayed with the substrate N-succinyl-Ala-Ala-Pro-Phe-pNA. Inhibition of cathepsin G (30nM) was titrated with increasing concentrations of His-EB22.4 (7.5nM-60nM) at room temperature for 2hr and residual activity assayed with the substrate N-succinyl-Ala-Ala-Pro-Phe-pNA on BSA coated microtitre plates. HLE (2nM) was incubated with increasing concentrations of His-EB22.4 (0.5nM-4nM) at 37°C for 2hr and assayed with 200 μ M N-methoxy-Ala-Ala-Pro-Val-pNA on BSA coated microtitre plates. The SI of trypsin inhibition by His-MMCM2 was determined by incubating 10nM trypsin with increasing concentrations of His-MMCM2 (2.5nM-50nM) and assaying for residual trypsin activity with Z-Phe-Arg-AMC. Linear regression analysis of the loss in protease activity with increasing serpin concentration was performed and the SI was determined by extrapolating to the protease:serpin ratio where protease activity is zero. The SI values represent the average of three separate experiments.

6.2.7. Kinetics of inhibitory reactions

6.2.7.a. Continuous method

Rate constants were measured under pseudo-first order conditions using the progress-curve method for the interaction of His-EB22.4 with HLE and catG. For HLE, assays were carried out using 0.4 nM HLE, 0.4-2.8 nM His-EB22.4 and 200 μ M N-methoxy-Ala-Ala-Pro-Val-pNA at 37°C in BSA coated microtitre plates. For catG, assays were performed with 20 nM catG, 4-40 nM His-EB22.4 and 1 mM N-succinyl-Ala-Ala-Pro-Phe-pNA at room temperature in BSA coated microtitre plates. A constant amount of protease was mixed with varying amounts of His-EB22.4 and substrate. The rate of product formation was measured at OD_{405nm}. Product formation is described by Equation 1 where P is the concentration of product at time t, k_{obs} is the apparent first-order rate constant and v_0 is the initial velocity.

$$P = \frac{v_0}{k_{obs}} \times [1 - e^{(-k_{obs}t)}]$$

Equation 1

For each combination of protease and His-EB22.4, a k_{obs} value was calculated by non-linear regression analysis of the data using Equation 1. The k_{obs} values were plotted against the respective serpin concentration [I] and linear regression was used to provide a line to the points. The slope of this line gave the second-order rate constant k' . As the rate of inhibition is dependent on the SI and the inhibitor is in competition with the substrate, the second order rate constant k' was corrected for substrate concentration, the K_M of the protease for substrate and SI to calculate by equation 2 to calculate k_a .

$$k_a = k' \times \left(1 + \frac{[S]}{K_M}\right) \times SI$$

Equation 2

The k_a values for the inhibition of cathepsin G and elastase represent the average of two and three separate experiments, respectively. Linear and non-linear regression analysis was fitted using Graphpad Prism 3.

6.2.7.b. Discontinuous method

The discontinuous method was used to determine the rate of inhibition (k_a) of chymotrypsin by His-EB22.4 and of trypsin with His-MMCM2. The pseudo-first order rate constant with chymotrypsin (2 nM) and His-EB22.4 (10-100 nM) was determined by incubation for different periods of time (0-5mins) followed by measurement of residual chymotrypsin activity. For MMCM2 and trypsin, 30 nM trypsin was incubated with various concentrations of His-MMCM2 (150 nM-240 nM) and assayed for residual activity at various time points (0-5mins). The pseudo first order constant, k_{obs} , was determined by the slope of a semi-log plot of the residual protease activity against time. The k_{obs} values were then plotted against serpin concentration and linear regression was used to fit the points to a line. The slope of this line gave an estimate of the second-order rate constant k_a . The k_a values represent the average of three separate experiments.

6.3. RESULTS

6.3.1. The RCL of the mouse *serpina3* proteins predict a varied protease inhibitory spectrum

Figure 6.1 shows the reactive centre loop residues of the four *serpina3* members in comparison to their human counterpart, α_1 -ACT, as well as the human serpins α_1 -AT, protein C inhibitor (PCI) and MNEI. All four mouse *serpina3* proteins have proximal hinge regions typical of inhibitory serpins, with the conserved Gly-Thr-Glu motif between P₁₅-P₁₃ followed by an Ala-rich region. Sequence through the protease specificity determining region (P₄-P₄')

shows marked divergence between the mouse *serpina3* members and their closest human counterpart, α_1 -ACT. What is immediately apparent is that none of the *serpina3* members have RCL residues identical to α_1 -ACT. Only one of the four members, 6C28, shares a Leu in the major specificity P_1 position. The presence of Leu at P_1 suggests that 6C28 may have inhibitory activity towards chymotrypsin-like proteases resembling human α_1 -ACT. However the P_1' residue of 6C28 is Cys.

The reactive site pairing of Met-Ser in EB22.4 is identical to that of human α_1 -AT. Furthermore, EB22.4 shares a Pro residue at P_2 with human α_1 -AT. This predicts that EB22.4 is most likely to behave like α_1 -AT, with inhibitory activity towards the serine proteases elastase and cathepsin G. The P_5 - P_3' RCL residues of 3E46 show striking similarity to human PCI and importantly, they share Arg-Ser at P_1 - P_1' . This suggests that 3E46 is likely to have inhibitory properties in common with PCI and inhibit trypsin-like serine proteases. MMCM2 has an unusual reactive centre, a pairing of Arg-Lys, and a large number of Gly residues between P_6 - P_1 . This P_1 - P_1' pairing is not present in any other known serpins although the presence of Arg at P_1 suggests inhibitory activity towards trypsin-like proteases.

A number of serpins have been reported to use dual reactive sites within the RCL to inhibit different types of target proteases. Serpins with dual specificity include MNEI, which has separate sites for elastase/proteinase 3 and cathepsin G/chymase/chymotrypsin (Cooley *et al.* 2001), PI-6 and PI-8 have sites for thrombin and chymotrypsin (Riewald and Schleef 1996; Dahlen *et al.* 1998 (b)) and PI-9 is able to inhibit granzymes B and neutrophil elastase (Sun *et al.* 1996; Dahlen *et al.* 1999). This suggests that the *serpina3* members may also inhibit proteases through the use of alternate reactive sites. Serpin inhibition of cysteine proteases demonstrates that some serpins can also function as cross-class inhibitors. Both crmA and PI-9 have been shown to inhibit caspases, whilst SCCA1, MENT and hurpin are all cathepsin inhibitors (Schick *et al.* 1998; Annand *et al.* 1999; Young *et al.* 2000;

	P_1 - P_1'	Protease Specificity
α_1 -ACT	GTEASAATAVKITLL LS ALVETRTIVRFNRPF	Chymotrypsin-like, cathepsin G, chymase
α_1 -AT	GTEAAGAMFLEAIP MS IPPEVK----FNKPF	Elastase, chymotrypsin- and trypsin- like
EB22.4	GTEAAAATGVKFVP MS AKLYPLTVY-FNRPF	Elastase and chymotrypsin-like
MMCM2	GTEAAAATGVIGGIR K AVLPAVC---FNRPF	Trypsin-like
3E46	GTEAAAATGFIFGFR S RRLQTMTVQ-FNRPF	Trypsin-like
6C28	GTEAAAATGVKLILL LC RKIYSMTIY-FNRPF	Chymotrypsin-like
PCI	GTRAAAATGTIFTFR S ARLNSQRLV-FNRPF	Trypsin-like, kallikrein, acrosin
MNEI	GTEAAAATAGIATF C MLMPEENFTA--DHPF	Neutrophil elastase, cathepsin G

Figure 6.1: Comparison of the predicted RCL residues of the four mouse serpin3 members EB22.4, MMCM2, 3E46 and 6C28 with the human homologue α_1 -ACT, and the human serpins α_1 -AT, PCI and MNEI. RCL residues are shown from P_{15} - P_{15}' and P_1 and P_1' residues are shown in bold. P_1 residues conserved with: α_1 -ACT are shown in red, α_1 -AT are shown in blue, and with PCI are shown in green. Predicted protease specificity is indicated on right.

Irving *et al.* 2002 (b); Welss *et al.* 2003). Therefore, it is also possible that the serpin3 members may inhibit cysteine proteases.

To establish the target protease specificity of the four serpin3 proteins, a panel of proteases consisting of six serine proteases and two cysteine proteases were analysed. Each protease was incubated with the serpin at two different protease:serpin ratios and the formation of a covalent complex between the serpin and protease analysed on SDS-PAGE. Conventional serpin-serine protease covalent complexes migrate at approximately the molecular weight of the sum of the proteins. Cleavage of the serpin within the RCL, without subsequent complex formation, causes it to migrate at approximately 5kDa smaller than the native protein.

6.3.2. Protease interactions of His-EB22.4

Figure 6.2(A&B) shows an SDS-PAGE of His-EB22.4 incubated with the panel of serine proteases. Recombinant His-EB22.4 forms SDS-stable complexes with elastase and cathepsin G and minimal complex formation is observed with chymotrypsin at a ratio of 1:1. Increasing the serpin concentration, whilst keeping the protease constant, results in increased complex formation with elastase and chymotrypsin.

His-EB22.4 is cleaved by trypsin as shown by the presence of a 42kDa predominant band at both protease:serpin concentrations. A higher molecular weight species at approximately 55kDa is also observed. This most likely represents complex degradation by excess uninhibited protease. His-EB22.4 did not interact with factor Xa, although some complex formation was observed with thrombin, at the higher protease:serpin ratio of 1:5 (Figure 6.2(B)).

Incubations with cysteine proteases were carried out by maintaining a constant serpin concentration and diluting the protease 1:5. Figure 6.2(C) shows the effect of incubating His-EB22.4 with cathepsins L and V. The serpin is completely digested by cathepsin V at a 1:1 ratio and cleaved by cathepsin L

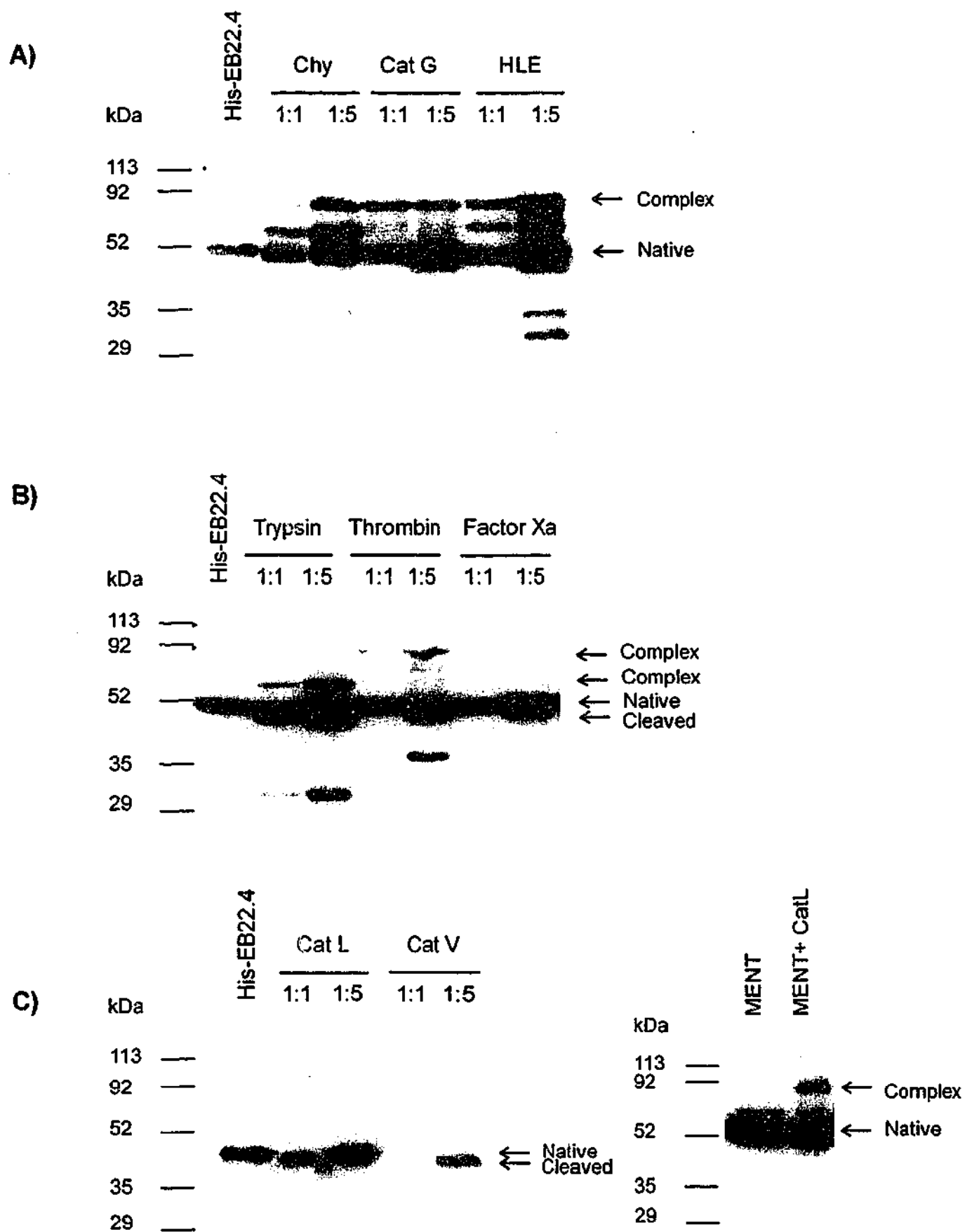


Figure 6.2: (A&B) Protease interactions of recombinant His-EB22.4 with panel of serine proteases and analysed for complex formation on 10% SDS-PAGE. Lanes marked 1:1 and 1:5 signify different protease:serpin ratios. **C)** 12.5% SDS-PAGE of showing the interaction of recombinant His-EB22.4 with two cysteine proteases, cathepsin L and V. A positive control reaction showing the interaction of MENT with cathepsinL is included. Gels were Western blotted and probed with anti-His antibody. Chy: chymotrypsin, HLE: human leukocyte elastase, CatG: cathepsin G, CatL: cathepsin L, CatV: cathepsin V.

to a low molecular weight entity in the absence of inhibitory complex formation. To demonstrate that a cysteine protease-serpin interaction can be resolved by this method, a positive control of the serpin, His-MENT, incubated with cathepsin L is included. Figure 6.2(C) demonstrates that complex formation between His-MENT and cathepsin L is able to be resolved on 12.5% SDS-PAGE.

6.3.3. *Protease interactions of His-MMCM2*

Incubation of His-MMCM2 with a panel of serine proteases was carried out using fractions of His-MMCM2 which contained minimal contaminating 90kDa complex (Figure 6.3(A&B)). His-MMCM2 was cleaved by the serine proteases chymotrypsin, cathepsin G and elastase (Figure 6.3(A)). Incubation of His-MMCM2 with chymotrypsin was associated with low molecular weight digestion products at both protease:serpin ratios. Incubation with cathepsin G resulted in the complete digestion of His-MMCM2 with no residual material apparent on the gel. A similar effect was observed by incubating His-MMCM2 with elastase, where complete digestion was observed at a ratio of 1:1 and only a small amount of low molecular protein detected at the ratio of 1:5.

No complex formation was observed with trypsin at a ratio of 1:1, however increasing the protease:serpin ratio to 1:5 resulted in complex being observed at approximately 70kDa. No interaction was observed with the other trypsin-like proteases, thrombin or factor Xa, either in complex formation or cleavage of the serpin. Incubation of His-MMCM2 with the cysteine proteases cathepsin L and cathepsin V was carried out by maintaining a constant serpin concentration and diluting the protease 1:5. Figure 6.3(C) demonstrates that incubation of His-MMCM2 with cathepsin L and V resulted in cleavage of the serpin by both proteases. A band smaller than native His-MMCM2, of approximately 40kDa, is evident at a ratio of 1:5. His-MMCM2 was also cleaved by cathepsin L at 1:1 (w/w) ratio but not a 1:5 ratio.

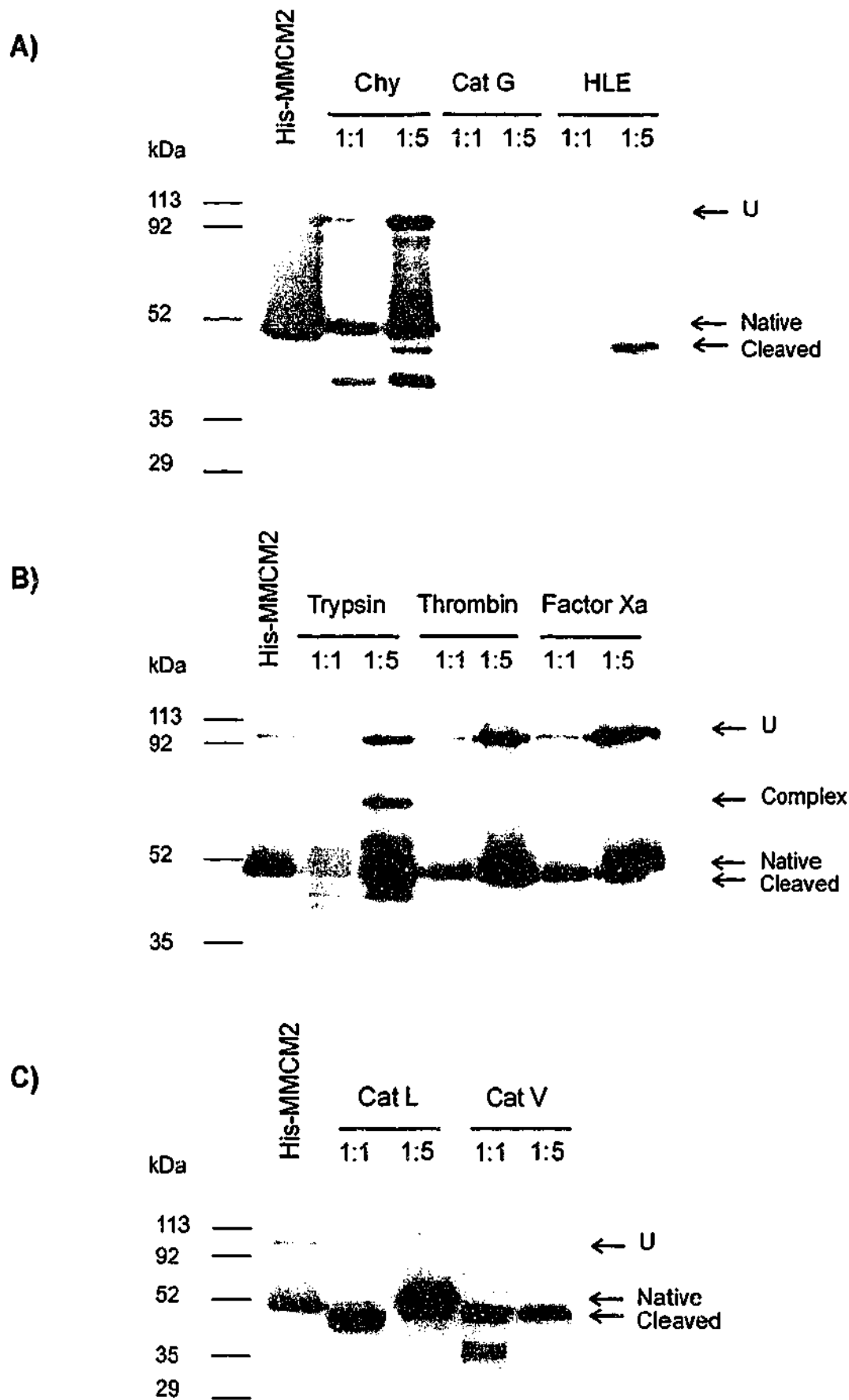


Figure 6.3: A&B) Protease interactions of recombinant His-MMCM2 with the serine protease panel and analysed for complex formation on 10% SDS-PAGE. Lanes marked 1:1 and 1:5 indicate protease:serpin ratios. **C)** 12.5% SDS-PAGE of showing the interaction of recombinant His-MMCM2 with two cysteine proteases. U indicates the presence of the unrelated His-MMCM2 dimer. Gels were Western blotted and probed with anti-His. Chy: chymotrypsin, HLE: human leukocyte elastase, CatG: cathepsin G, CatL: cathepsin L and CatV: cathepsin V.

6.3.4. *Protease interactions of His-3E46*

As the recombinant His-3E46 was shown to be in a relaxed conformation, *in vitro* transcribed/translated protein was used to investigate for protease interactions. His-3E46 was incubated with either 20 ng or 100 ng of each protease from the panel for 10 min at 37°C. Inhibitory complex formation was analysed on 10% SDS-PAGE and the results are presented in Figure 6.4. His-3E46 was cleaved in the presence of 100 ng of HLE, indicated by the reduction in molecular weight of approximately 5 kDa with respect to native serpin. It was not cleaved at the lower HLE concentration of 20 ng. His-3E46 was also cleaved by chymotrypsin at both protease concentrations. The presence of high molecular weight entities at a concentration of 20 ng indicates a small amount of complex formation between His-3E46 and chymotrypsin. Neither complex formation nor RCL cleavage was observed with cathepsin G.

Figure 6.4(B) demonstrates that high molecular weight complexes were formed when His-3E46 was incubated with trypsin, thrombin and factor Xa. The greatest amount of complex formation occurred with thrombin, with prominent complexes at both concentrations of protease. Less inhibitory complex was formed with factor Xa. Incubation with trypsin at both concentrations resulted in the formation of an SDS-stable complex of approximately 60kDa and an entity of approximately 40kDa, consistent with loop-cleavage of His-3E46.

Figure 6.4(C) demonstrates that incubation with the cysteine proteases cathepsin V and cathepsin L resulted in the cleavage of His-3E46. Cathepsin V appeared to cleave the serpin at two points due to the presence of a doublet at approximately 40kDa. Incubations with cathepsin L also resulted in serpin cleavage as demonstrated by the presence of a number of low molecular weight entities around the 40 kDa.

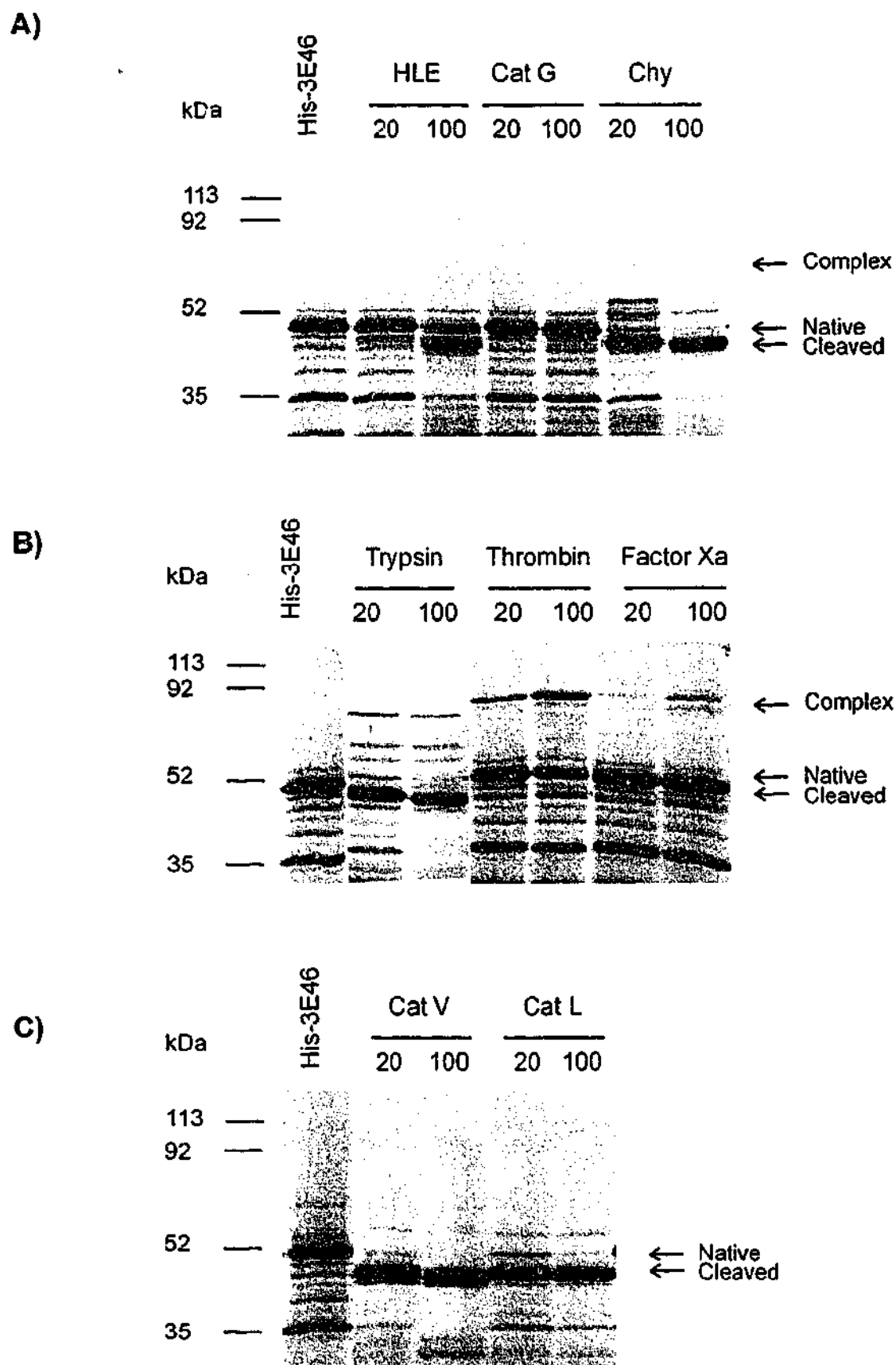


Figure 6.4: Protease interactions of ^{35}S -Met-labelled His-3E46 with protease panel. His-3E46 was incubated with either 20 or 100 ng of protease and complex formation analysed on 10% SDS-PAGE gels. **A)** Interactions with Chy: chymotrypsin, CatG: cathepsin G and HLE: human leukocyte elastase. **B)** Interactions with trypsin-like serine proteases. **C)** Interactions with cysteine proteases, CatL: cathepsin L and CatV: cathepsin V.

6.3.5. Protease interactions of His-6C28

The presence of Leu-Cys at the predicted reactive site of 6C28 suggests that the protein would inhibit chymotrypsin-like proteases which cleave after hydrophobic residues. Preliminary experiments with recombinant His-6C28 demonstrated that bacterial recombinant His-6C28 was completely degraded in complex formation assays with chymotrypsin-like proteases. Therefore, to assess if this was a true feature of the 6C28, it was generated by *in vitro* transcription and translation.

In vitro transcribed/translated His-6C28 was incubated with 20 ng and 100 ng of the protease panel for 10 min at 37°C. Surprisingly, inhibitory activity was observed despite the atypical behaviour of His-6C28 seen on TUG gels (Chapter 5). Figure 6.5(A) shows the effect of incubating His-6C28 with the serine proteases HLE, cathepsin G and chymotrypsin. The ³⁵S-labelled His-6C28 formed complex with HLE at both protease concentrations, as indicated by the presence of inhibitory complex of approximately 70kDa. At the higher elastase concentration cleavage is also observed with an entity present of approximately 42kDa. His-6C28 was able to form inhibitory complexes with cathepsin G at both protease concentrations and more complex was observed in the presence of 100 ng of cathepsin G than 20 ng. Incubation of His-6C28 with 20 ng of chymotrypsin yielded an inhibitory complex at approximately 70kDa. Increasing the concentration of chymotrypsin to 100ng resulted in a series of high molecular weight entities between 50-60kDa which most likely represent proteolysis of the His-6C28-chymotrypsin complex.

Figure 6.5(B) shows the effect of incubating His-6C28 with trypsin-like proteases. His-6C28 appeared to be loop-cleaved by trypsin as shown by the loss in molecular weight of approximately 5kDa. There was no evidence of complex formation with trypsin, thrombin or factor Xa. His-6C28 is cleaved by 20 and 100 ng cathepsin V, producing a product at approximately 40kDa (Figure 6.5(C)). However, incubation with 20 ng of cathepsin L resulted in the formation of a complex between 75-80 kDa. In the presence of larger amounts

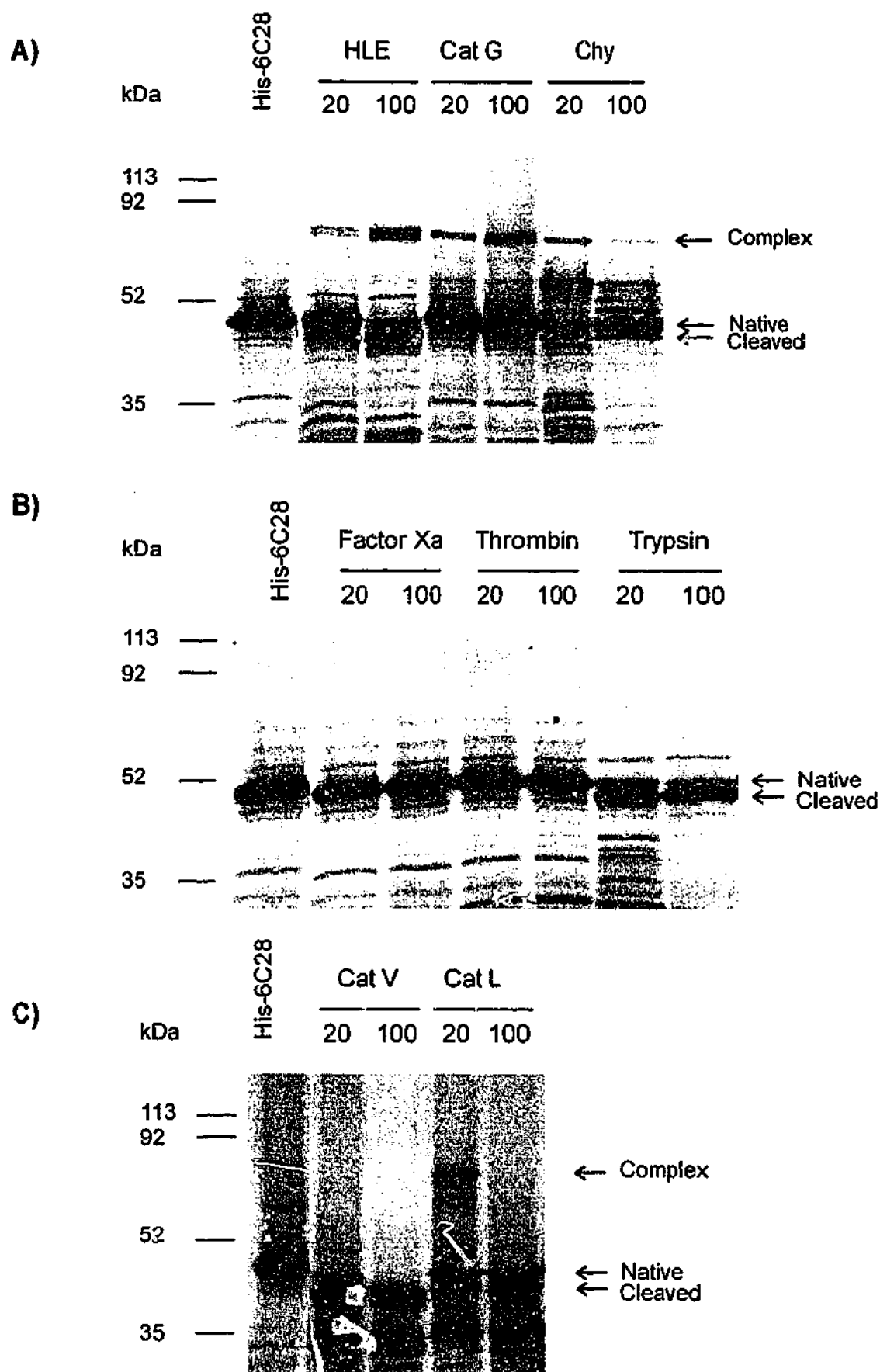


Figure 6.5: Protease interactions of ³⁵S-Met-labelled His-6C28 with protease panel. His-6C28 was incubated with either 20 or 100 ng of protease and complex formation analysed on 10% SDS-PAGE gels. **A)** Interactions with serine proteases, Chy: chymotrypsin, CatG: cathepsin G, and HLE: human leukocyte elastase. **B)** Interactions with trypsin-like serine proteases. **C)** Interactions with cysteine proteases, CatL: cathepsin L and CatV: cathepsin V.

of cathepsin L the amount of complex present was reduced which is likely due to proteolysis of the complex by excess protease. Repeated experiments showed that complex formation with cathepsin L was reproducible. This suggested that, in addition to being an active inhibitor of cathepsin G and elastase, this serpin is also a cross class inhibitor of the cysteine proteases.

6.3.6. Kinetics of interactions

6.3.7. Inhibitory kinetics of recombinant His-EB22.4 with chymotrypsin, cathepsin G and elastase

From the panel of proteases tested, His-EB22.4 appeared to inhibit the serine proteases chymotrypsin, cathepsin G and elastase. In order to establish whether these protease interactions are physiologically relevant, the kinetics of inhibition were determined.

To determine the SI of His-EB22.4 and the proteases chymotrypsin, cathepsin G and elastase, increasing concentrations of serpin were incubated with a constant amount of protease. The residual protease activity was measured and plotted against the respective serpin[I]:protease[E] concentration. Figure 6.6 shows a linear decrease in protease activity as serpin concentration increased. The SI was determined by extrapolating to the serpin[I]:protease[E] concentration which resulted in total loss of protease activity. The SI of His-EB22.4 with chymotrypsin was 2.1 which is higher than that reported for human α_1 -ACT and α_1 -AT. The SI for interactions with elastase and cathepsin G were determined as 1.1 and 1.0 respectively, which agree with SI values previously reported for α_1 -ACT and α_1 -AT (Table 6.1).

The association rate constant (k_a) between His-EB22.4 and chymotrypsin was measured under pseudo-first order conditions using the discontinuous method. Different concentrations of His-EB22.4 and chymotrypsin were pre-incubated for various lengths of time (0-5 min), after which substrate was added and

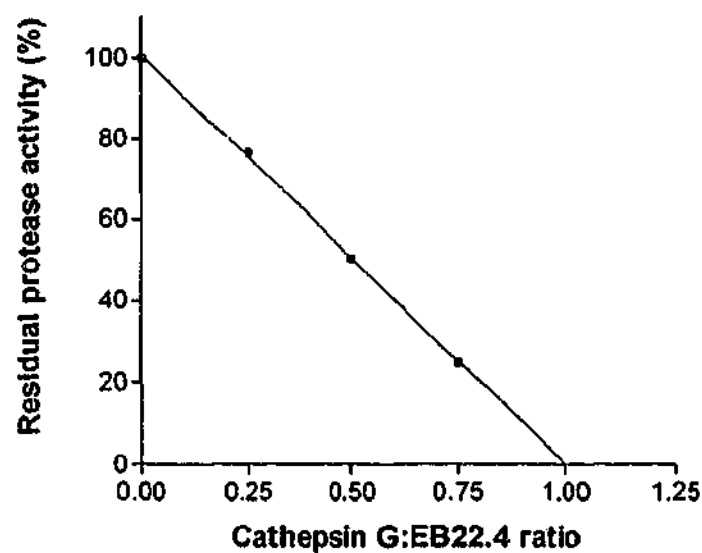
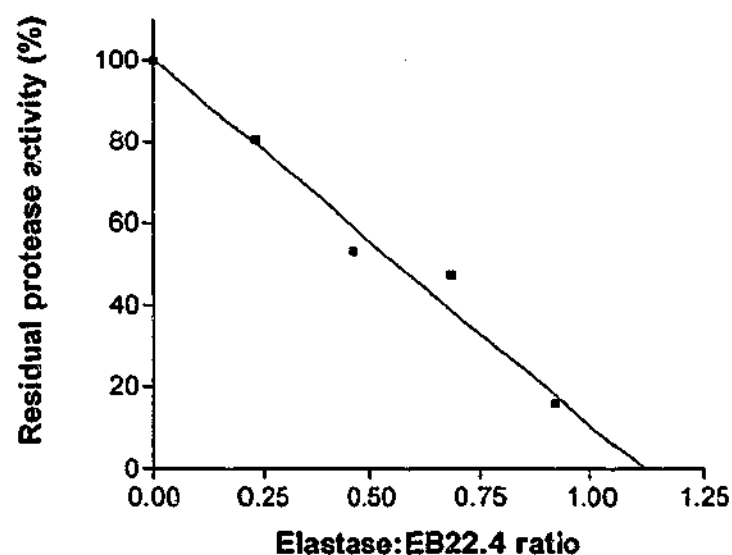
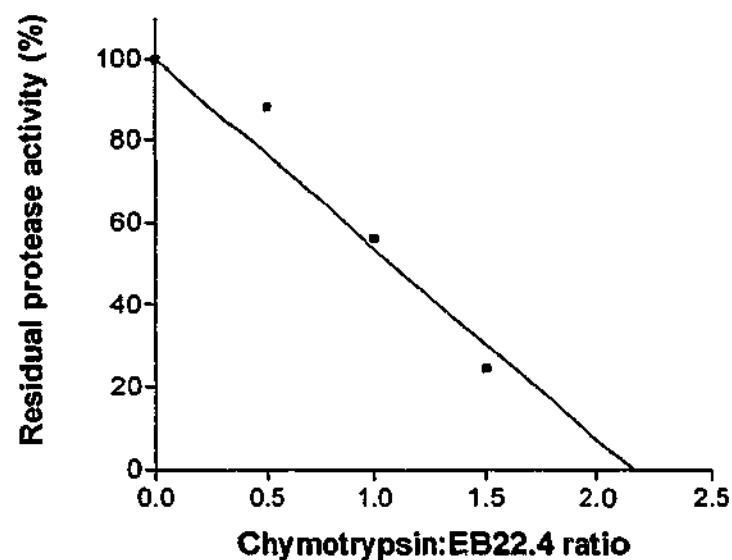


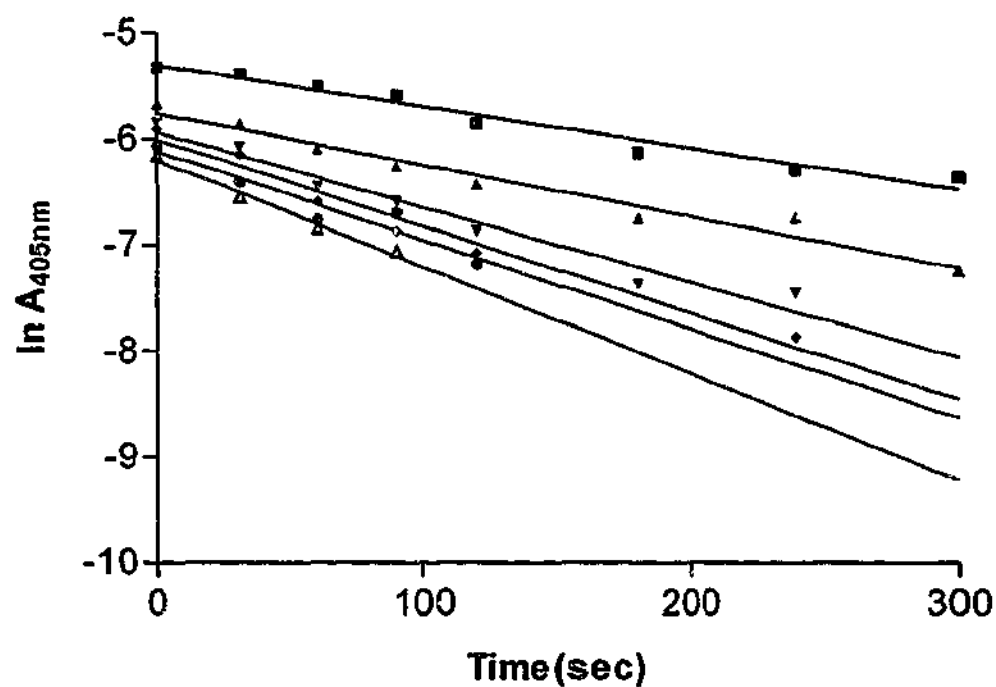
Figure 6.6: Inhibition of chymotrypsin, elastase and cathepsin G by His-EB22.4. The proteases were incubated with varying amounts of His-EB22.4 at the indicated serpin:protease ratios (mol:mol) and residual protease activity assayed. The stoichiometry of inhibition was determined by extrapolation to serpin:protease ratio where residual protease activity was zero.

residual enzyme activity measured. When the log of enzyme activity, (represented by $\ln A_{405\text{nm}}$) was plotted against time, a linear relationship was observed (Figure 6.7(A)). Linear regression analysis of the slope of each line provided the different association rates, k_{obs} , at the respective serpin concentrations. The k_{obs} values were plotted against serpin concentration and the slope of this line (Figure 6.7(B)), gave the k_a value of $5.8 \pm 0.7 \times 10^4 \text{ M}^{-1} \text{ sec}^{-1}$.

In preliminary investigations, the inhibition of cathepsin G and HLE was tested using the discontinuous method. However, inhibition occurred too rapidly and we were unable to measure residual protease activity post incubation. Therefore, the association rate constant between His-EB22.4 with cathepsin G and elastase was determined using the progress curve method. This method has the advantage of substrate being present throughout the incubation and inhibition is measured by the continuous loss of protease activity.

Figure 6.8(A) is a representative example of the studies undertaken to determine the association constant of elastase with His-EB22.4. The progress of inhibition was monitored as change in absorbance at 405nm as a function of time and the rate of loss of enzyme activity was calculated as k_{obs} (Equation 1). The k_{obs} values obtained with different concentrations of His-EB22.4 were then plotted against serpin concentration (Figure 6.8(B)). The slope of this line, the K_M of the substrate and the SI were then used to calculate the overall k_a as per Equation 2. The results show that His-EB22.4 is a fast inhibitor of elastase with a k_a of $4 \pm 0.9 \times 10^6 \text{ M}^{-1} \text{ sec}^{-1}$, which is 5-10 fold slower than human α_1 -AT (Beatty *et al.* 1980). The interaction with cathepsin G gave a k_a value of $7.9 \pm 0.9 \times 10^5 \text{ M}^{-1} \text{ sec}^{-1}$, which makes it a faster inhibitor than human α_1 -AT. Although the k_a is approximately 60-fold less than human α_1 -ACT reported by Beatty *et al* (1980), it does agree with the inhibitory constants reported by Lomas *et al* (1995 (b)) and Duranton *et al* (1998) (Table 6.1).

A)



B)

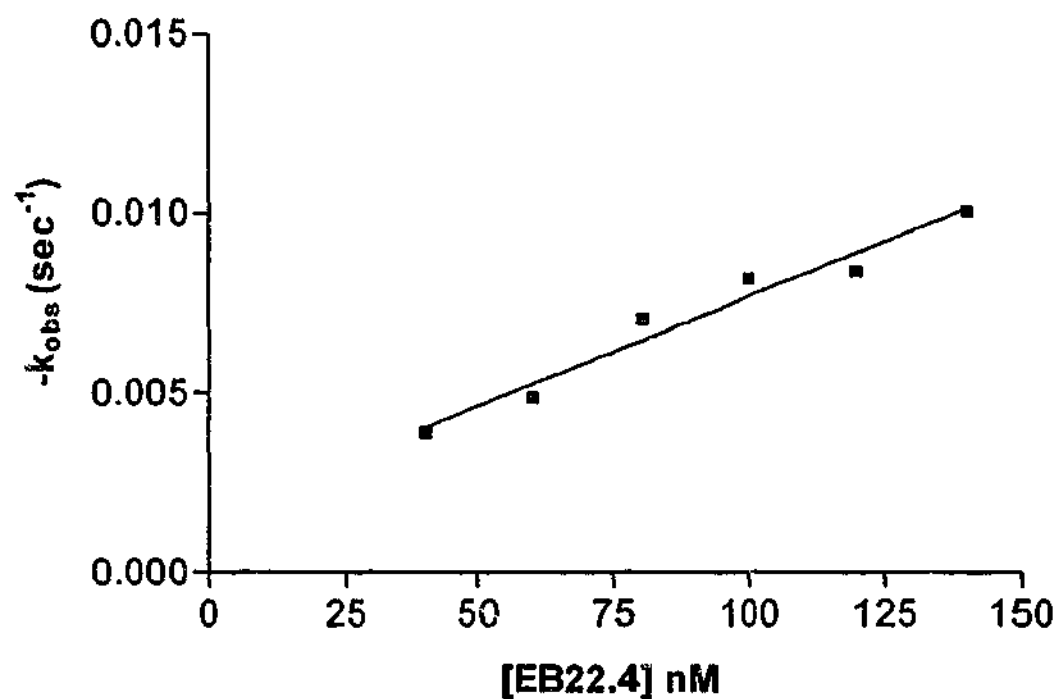
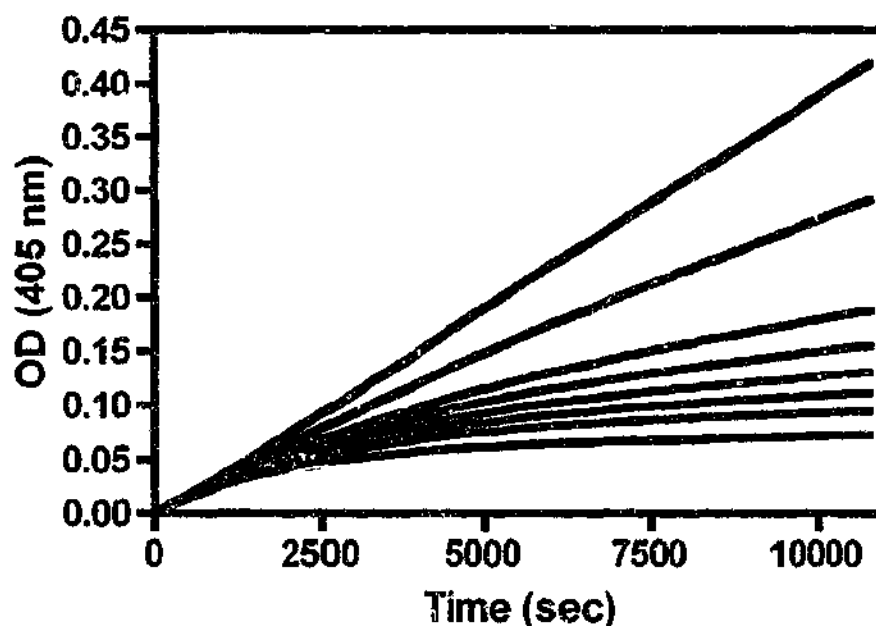


Figure 6.7: Determination of the inhibitory constant k_i for the inhibition of chymotrypsin by His-EB22.4. **A)** Semilog plots of residual chymotrypsin activity vs time for reactions at various concentrations of His-EB22.4. **B)** Plot of k_{obs} as a function of His-EB22.4 concentration. Linear regression of the slope gave the second-order rate constant k_i for the inhibition of chymotrypsin by His-EB22.4.

A)



B)

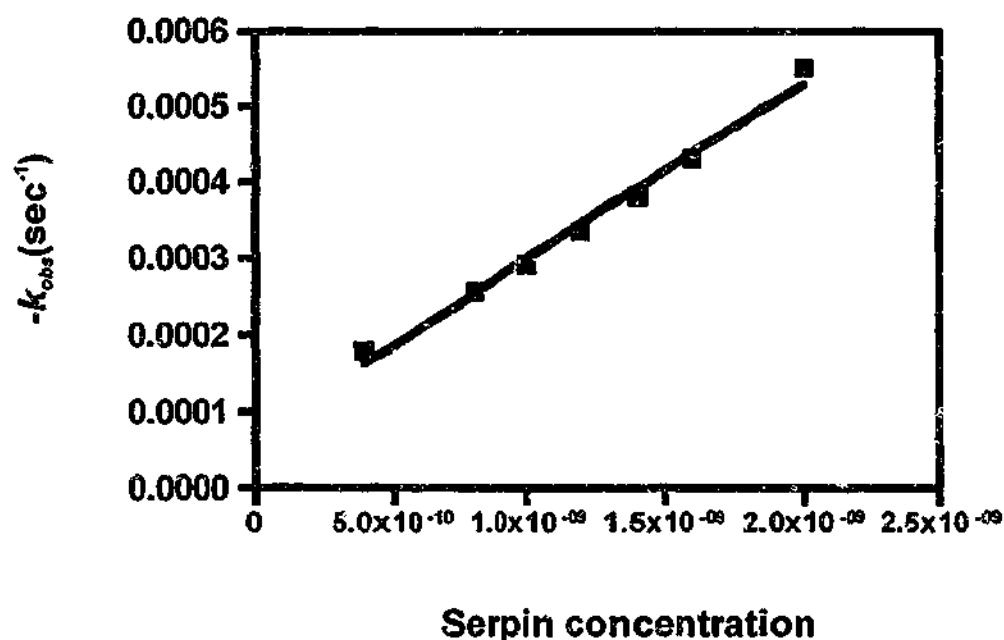


Figure 6.8: Kinetic analysis of the inhibition of human neutrophil elastase by His-EB22.4. **A)** Progress curves of the interaction between elastase (0.4nM) and His-EB22.4 (0.4-2nM), monitored by continuous measurement of the change in A_{405nm} . k_{obs} at each serpin concentration was determined by non-linear regression analysis of each curve using Equation 1. **B)** k_{obs} were plotted against His-EB22.4 concentration, and linear regression analysis was used to determine second-order rate constant (k). k_s was determined by accounting for the K_m of the protease for substrate (Equation 2).

Table 6.1: Kinetic constants for recombinant His-EB22.4 and His-MMCM2 in comparison with human α_1 -ACT and α_1 -AT.

PROTEASE	SERPIN	SI	k_a ($M^{-1}sec^{-1}$)
Chymotrypsin (bovine)	His-EB22.4	2	$5.8 \pm 0.7 \times 10^4$
	His-MMCM2	Substrate	-
	α_1 -ACT	1	6×10^4 (Beatty <i>et al.</i> 1980) 5.5×10^5 (Lomas <i>et al.</i> 1995 (b))
	α_1 -AT	1	5.9×10^5 (Beatty <i>et al.</i> 1980) 3.6×10^6 (Lomas <i>et al.</i> 1995 (b))
Cathepsin G (human)	His-EB22.4	1	$7.9 \pm 0.9 \times 10^5$
	His-MMCM2	Substrate	-
	α_1 -ACT	1	5.1×10^7 (Beatty <i>et al.</i> 1980) 8.1×10^5 (Lomas <i>et al.</i> 1995 (b)) 7.0×10^5 (Duranton <i>et al.</i> 1998)
	α_1 -AT	1.1	4.1×10^5 (Beatty <i>et al.</i> 1980)
Neutrophil Elastase (human)	His-EB22.4	1.1	$4.0 \pm 0.9 \times 10^6$
	His-MMCM2	Substrate	-
	α_1 -ACT	Substrate	- (Lomas <i>et al.</i> 1995 (b))
	α_1 -AT	1.1	5×10^7 (Beatty <i>et al.</i> 1980) 1.9×10^7 (Lomas <i>et al.</i> 1995 (b))
Trypsin (bovine)	His-EB22.4	Substrate	-
	His-MMCM2	10	$2.7 \pm 0.9 \times 10^4$
	α_1 -ACT	Substrate	-
	α_1 -AT	1.1	1.69×10^5 (ChaillanHuntington <i>et al.</i> 1997)

6.3.8. Inhibitory kinetics of recombinant His-MMCM2 with trypsin

Figure 6.9 is a representative example of the studies used to determine the inhibitory kinetic parameters for the interaction of His-MMCM2 with trypsin. Fractions of His-MMCM2 were used that contained the least amount of contaminating 90kDa complex. The stoichiometry of inhibition was determined as described for His-EB22.4 (Figure 6.9(A)) and three separate calculations gave an SI of 10.0. The k_a was determined using the discontinuous method, as previously outlined for the interaction of His-EB22.4 and chymotrypsin.

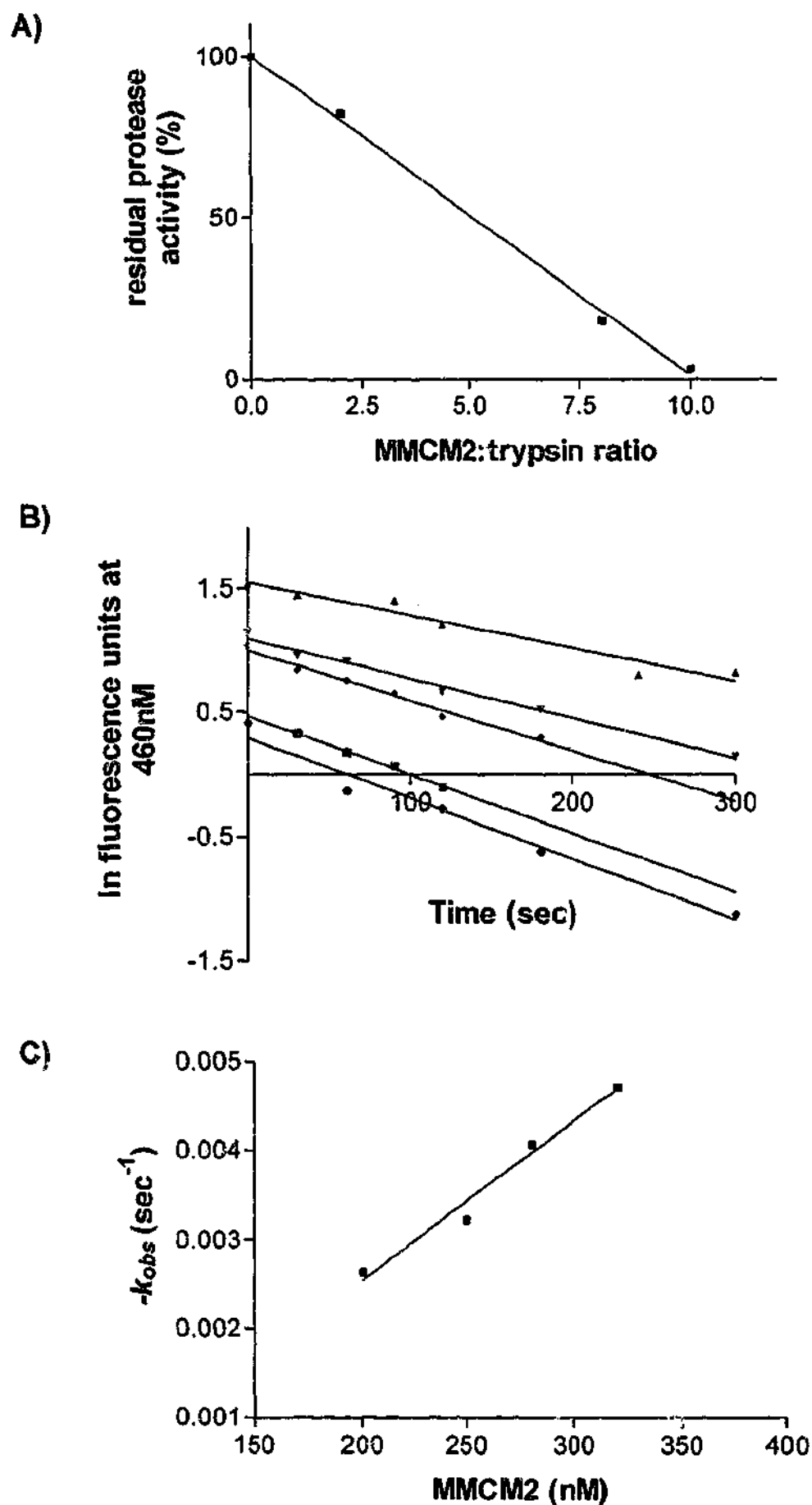


Figure 6.9: Determination of the kinetic parameters of inhibition of trypsin by His-MMCM2. **A)** His-MMCM2 and trypsin were incubated with varying amounts at the indicated serpin:protease ratios and residual protease activity assayed. The stoichiometry of inhibition was determined by extrapolation to serpin:protease ratio where residual protease activity was zero. **B)** Semilog plots of residual trypsin activity vs time for reactions with various concentrations of His-MMCM2. **C)** Plot of k_{obs} as a function of His-MMCM2 concentration. Linear regression of the slope gave the second-order rate constant k_i for the inhibition of trypsin by His-MMCM2.

His-MMCM2 was found to be a moderate inhibitor of trypsin with a k_i value of $2.7 \pm 0.9 \times 10^4 \text{ M}^{-1}\text{sec}^{-1}$ (Table 6.1).

6.4. DISCUSSION

Target protease specificity is determined by the amino acid composition of the reactive centre loop. Comparison of the RCL residues of human α_1 -ACT and the mouse *serpina3* paralogues EB22.4, MMCM2, 3E46 and EB22.4 reveal that only 6C28 shares the Leu of α_1 -ACT at the critical P_1 position. Furthermore, the highly divergent RCL residues suggest that the *serpina3* mouse serpins possess highly variable protease specificity. As predicted, this study demonstrates that inhibitory activity is mediated towards a wide range of serine proteases, including members from the chymotrypsin, trypsin and elastase subfamilies. Cross class inhibitory activity towards the cysteine protease family is also observed.

In addition to residue composition, the effectiveness of protease inhibition is also modulated by RCL length (Lawrence *et al.* 2000). The human α_1 -ACT RCL length is four residues longer than the archetypal serpin human α_1 -AT (Table 6.1). In comparison to α_1 -AT, the MMCM2 loop is one residue longer, whilst the other three mouse serpin RCL's have three extra residues. The longer loop lengths did not appear to interfere with protease inhibitory activity, as complex formation was observed with all four serpins. However, the longer loop-lengths may have implications for RCL conformation.

The mouse serpin EB22.4 shares many features with human α_1 -ACT and α_1 -AT and functionally appears to be an intermediate between the two serpins. Whilst EB22.4 has greatest sequence homology to human α_1 -ACT outside the RCL, the P_1 - P_1' residues (Met-Ser) are identical to human α_1 -AT. Complex formation assays show that His-EB22.4 shares the target proteases chymotrypsin, cathepsin G and neutrophil elastase, with α_1 -ACT and α_1 -AT. Although there is no direct evidence, the inhibitory activity is likely to be

mediated through the Met residue at P₁, resembling α_1 -AT, where P₁ Met functions in the inhibition of both elastase and cathepsin G (Beatty *et al.* 1980).

Kinetic analyses show that His-EB22.4 is a fast and efficient inhibitor of cathepsin G and elastase with k_a values of approximately $10^6 \text{ M}^{-1}\text{sec}^{-1}$ and stoichiometry of inhibition of ~ 1 with both proteases. The ability of His-EB22.4 to inhibit elastase is a feature in common with human α_1 -AT but not human α_1 -ACT which has no inhibitory activity towards elastase. Previous attempts to convert human α_1 -ACT into an elastase inhibitor by producing a series of α_1 -ACT- α_1 -AT chimeras, in which sections of the α_1 -AT RCL (between P₆-P_{3'}) were replaced in the corresponding region of α_1 -ACT, were unable to yield a fast inhibitor of elastase (Plotnick *et al.* 1997). The most effective α_1 -ACT-chimera (P₃-P_{3'}) possessed at least 60-fold slower rate of inhibition than α_1 -AT. Furthermore, the complexes of α_1 -ACT (P₃-P_{3'})-chimera-HLE broke down more rapidly than α_1 -AT-HLE complexes (Rubin *et al.* 1994). The inability to convert α_1 -ACT into an HLE inhibitor was attributed to differences between RCL conformations of α_1 -ACT and α_1 -AT (Plotnick *et al.* 1997). His-EB22.4 is at least 30-fold faster than any of the α_1 -ACT-chimeras which suggests that the RCL of His-EB22.4 is in a more favourable conformation to interact with HLE. It would be interesting to determine the conformation of the His-EB22.4 RCL and compare it to those of human α_1 -ACT and α_1 -AT.

Despite some complex formation observed between His-EB22.4 and trypsin on SDS-PAGE, kinetic assays show no inhibitory activity towards trypsin. It is possible that the SI is so high that we are unable to observe an inhibitory effect in kinetic assays. The presence of basic Lys residues at P₅ and P_{3'} are ideal cleavage points for trypsin, which is consistent with the loop-cleavage observed on SDS-PAGE. Cleavage of the RCL away from the P₁ residue is more likely to result in substrate-like behaviour of the serpin due to change in loop length (Zhou *et al.* 2001).

The prominent mRNA expression of EB22.4 in liver and the presence of a predicted N-terminal secretion peptide suggest plasma localisation *in vivo*.

Subsequently, it is likely to parallel human α_1 -ACT and α_1 -AT in the regulation of elastase and cathepsin G at sites of inflammation. The fast inhibition of cathepsin G and elastase by EB22.4 further supports that the two proteases are likely to be physiological targets in the mouse. Elastase and cathepsin G are serine proteases produced by myelomonocytic cells and stored in the azurophil granules of neutrophils (Borregaard and Cowland 1997). The proteases are released from the granules into the phagosome to eliminate phagocytosed particles including pathogens, parasites or apoptotic cells. Stimulation of cells during an inflammatory response signals the release of cathepsin G and elastase into the extracellular milieu (Borregaard and Cowland 1997). These neutrophil proteases are associated with several destructive tissue disease states including emphysema, chronic obstructive pulmonary disease (COPD) and cystic fibrosis. Human α_1 -ACT and α_1 -AT are the primary extracellular regulators of cathepsin G and elastase activity and this is corroborated by patients with α_1 -antitrypsin or α_1 -antichymotrypsin deficiency who develop emphysema (Faber *et al.* 1993; Poller *et al.* 1993).

The efficiency with which His-EB22.4 is able to inhibit cathepsin G also supports EB22.4 to be the closest functional homologue of human α_1 -ACT. It is important to note that the proteases used in this study were of human origin. A better understanding of the physiological role of EB22.4 would come from the use of mouse proteases as it is likely that association rates with mouse cathepsin G and elastase would be greater than the values determined with their human counterparts. Furthermore, one of the consequences of using bacterial recombinants is the absence of glycosylation. Glycosylation rates may alter function and using native plasma EB22.4 may result in a further increase in the kinetics of interaction may also be observed.

Analysis on TUG gels suggested that His-6C28 was extremely stable and resisted unfolding to fairly high concentrations of urea. Thus the inhibitory complex formation observed with chymotrypsin, cathepsin G and elastase was unexpected. This complex formation appeared to increase with greater amounts of elastase and cathepsin G, although in the case of HLE, this also

resulted in more cleaved serpin. Both α_1 -ACT and 6C28 have a Leu residue at P₁, which supports the inhibitory activity observed towards cathepsin G. It is interesting that inhibition of neutrophil elastase is observed when human α_1 -ACT does not interact. Unlike α_1 -ACT, 6C28 possesses a Cys residue at P1'. Other serpins previously shown to inhibit elastase include the drosophila protein Necrotic (Nec) which also has a Leu at P₁ and MNEI which has a Cys at P₁. (Cooley *et al.* 2001; Robertson *et al.* 2003). Therefore, the inhibitory activity towards elastase could be mediated through either P₁ or P₁' of 6C28.

6C28 is a predicted intracellular serpin with mRNA expression detected in lymphoid tissue and fibroblast cells but not in haemopoietic cells. Given that both cathepsin G and elastase are exclusively expressed in myelomonocytic cell lineages, it is unlikely that 6C28 and the two proteases could physically interact, given their compartmental division. It is more likely that 6C28 interacts with other chymases present in the intracellular compartment.

His-6C28 did not show any inhibitory activity towards trypsin-like proteases factor Xa and thrombin however, it was loop cleaved by trypsin. This is consistent with the presence of the basic residues Lys at P₅ and P₃' and Arg at P₂', which are ideal cleavage sites for trypsin.

The ³⁵S-Met-labelled His-6C28 also formed a stable inhibitory complex with cathepsin L and was loop cleaved by cathepsin V in the absence of complex formation. Substrate specificity for cathepsin L is primarily determined by the S2 subsite which has a preference for hydrophobic or branched amino acids (Chapman *et al.* 1997). The presence of Ile at P₃ makes the 6C28 RCL a good pseudo-substrate for cathepsin L and V. Cathepsin L is a ubiquitously expressed lysosomal cysteine protease with predominant intracellular localisation (Turk *et al.* 2000). Therefore, it is possible that 6C28 interacts with either the murine homologue of cathepsin L or a cysteine protease which possesses cathepsin L-like activity *in vivo*. Further work with a functional recombinant would be required to fully characterise this interaction.

The presence of an Arg at P₁ in His-MMCM2 suggested inhibitory activity towards trypsin-like proteases. His-MMCM2 was shown to inhibit trypsin with a k_a of $1 \times 10^{-5} \text{ M}^{-1}\text{sec}^{-1}$. The observed SI of 10.0 is consistent with SDS-PAGE analysis which demonstrates that His-MMCM2 is also loop cleaved. The SI also suggests that trypsin is unlikely to be its physiological target *in vivo* and therefore, MMCM2 is likely to interact with another trypsin-like serine protease. There are a number of trypsin-like proteases involved in the regulation of coagulation, fibrinolysis and the complement cascade towards which MMCM2 may function as a specific inhibitor. However, His-MMCM2 showed neither inhibitory or substrate activity towards the trypsin-like proteases factor Xa or thrombin. Cathepsin G, elastase and chymotrypsin digested MMCM2 with no evidence of complex formation. Incubation of His-MMCM2 with the cysteine proteases also resulted in apparent RCL cleavage in the absence of complex formation which may have occurred via the hydrophobic Ile residue at P₃ in the RCL.

The *in vitro* transcribed/translated His-3E46 formed SDS-PAGE stable complexes with all three trypsin-like proteases investigated. The predicted P₁ Arg of 3E46 is consistent with trypsin-like inhibitory activity. It is interesting that complex formation with thrombin is observed, whilst incubation of trypsin results in the majority of the serpin being cleaved. The 3E46 RCL has a further two Arg residues at P₂' and P₃' which can also act as substrates for trypsin, and may explain the presence of loop-cleavage in conjunction with complex formation. Thrombin has an almost exclusive cleavage preference for P₁ Arg, and can accommodate P₂ Pro/Gly and P₁' Ser/Thr/Gly plus has a requirement for no acidic residues between P₃ and P₃' (Le Bonniec *et al.* 1991). Therefore, thrombin may only cleave the RCL of 3E46 at P₁-P₁', as the Arg residues located at P₂' and P₃' may not meet its specific substrate requirements.

The RCL residues of 3E46 are very similar to human PCI (Figure 6.1). PCI is a heparin-binding serpin and inhibits several trypsin-like proteases including activated protein C, thrombin and factor Xa (Suzuki *et al.* 1984; Espana and Griffin 1989). The binding of heparin to PCI accelerates the inhibition of

thrombin and activated protein C approximately 30-fold (Suzuki *et al.* 1984; Pratt and Church 1992). PCI binds heparin primarily through four basic residues of helix H (Lys-266, Arg-269, Lys-270 and Lys-273), and these residues are also present in 3E46. However, 3E46 does not share the additional residues on helix D, which also contribute to heparin binding of PCI (Kuhn *et al.* 1990; Neese *et al.* 1998). Furthermore, the residues important for antithrombin-heparin interactions are also absent in 3E46 (Jin *et al.* 1997). In the absence of a functional recombinant it is difficult to ascertain the effectiveness of the interactions between 3E46 and target proteases and to assess its heparin binding ability. It will be of interest to determine whether 3E46 activity is also modulated by co-factors.

6.5. CONCLUSION

The serpin3 proteins investigated have varied target protease specificity as predicted by their hypervariable RCL. The four serpins, EB22.4, MMCM2, 3E46 and 6C28 were demonstrated to inhibit proteases of different classes, and from different families. This suggests that they possess distinct inhibitory and biological functions *in vivo*.

CHAPTER 7

Establishment of a retroviral-mediated gene transfer system to introduce serpin2A into murine haemopoietic progenitors

7.1. INTRODUCTION

A number of human intracellular serpins are expressed in haemopoietic cells at different stages of differentiation where they are involved in a diverse range of biological functions including cell growth, differentiation, and apoptosis (as described in Section 1.5.1). Serpin2A is an intracellular member of the mouse *serpina3* locus with high level expression in haemopoietic progenitor cells (HPC) and is down-regulated on differentiation (Hampson *et al.* 1997; Terskikh *et al.* 2001). Constitutive expression of serpin2A in the primitive multipotential cell line FDCP-MixA4 resulted in a higher clonogenic potential and a delay in differentiation in comparison to control cells (Hampson *et al.* 1997). These studies suggested that serpin2A may play a regulatory role in haemopoietic differentiation. However, the effect of serpin2A expression in primitive haemopoietic progenitors *in vivo* is unknown.

Primitive haemopoietic progenitor cells have the capacity to self-renew and to differentiate into all blood cell lineages. Integration of a gene into the genome of HPC's is an effective way to investigate its biological function in haemopoiesis as the introduced gene will be transferred to all HPC-derived progeny. The most common approach of gene delivery to HPC's is through the use of retroviruses that carry the gene of interest. The retrovirally-transduced HPCs can then be assessed for their ability to repopulate the marrow and contribute to haemopoiesis in ablated host mice.

The aim of this chapter was to establish a retroviral system to introduce constitutive serpin2A expression in HPC's and to investigate the effects on

normal haemopoietic cell proliferation and differentiation. Based on the work reported by Hampson *et al*, the enforced expression of serpin2A is likely to lead to delayed maturation with a consequent reduced contribution to mature peripheral blood cells.

7.2. METHODS

A schematic of the approach undertaken to introduce the serpin2A gene into the haemopoietic compartment is presented in Figure 7.1. Retroviral gene transfer requires cells to be cycling. Donor mice were treated with 5-fluorouracil (5-FU) which kills the cycling progeny normally present in the marrow driving the normally quiescent primitive stem cells into cycle (Szilvassy and Cory 1993). To further enrich for primitive progenitors, bone marrow cells were sorted and cultured *ex vivo* for 2-3 days in the presence of cell cycle promoting cytokines. The progenitors were then transduced with retroviral vector encoding either green fluorescent protein (GFP) alone or serpin2A/GFP. Post-infection the transduced cells were sorted by fluorescence activated cell sorting (FACS), based on GFP positivity (GFP⁺) and introduced into lethally irradiated host mice or plated into HPP-CFC assays (described in 7.2.11). Isolation of blood and bone marrow in the recovering mice was then examined to assess the contribution of transduced donor progenitor cells to haemopoiesis in host mice. These experiments were performed in collaboration with and in the laboratory of Dr Ivan Bertoncello from the Peter MacCallum Cancer Centre, Australia.

7.2.1. Cell lines and vector construction

The ecotropic packaging cell line GP⁺E86 was used for the generation of the ecotropic retroviruses. GP⁺E86 and NIH-3T3 cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum, 2 mmol/L glutamine, 100U/ml penicillin and 100 µg/ml streptomycin at 37°C in 5% CO₂.

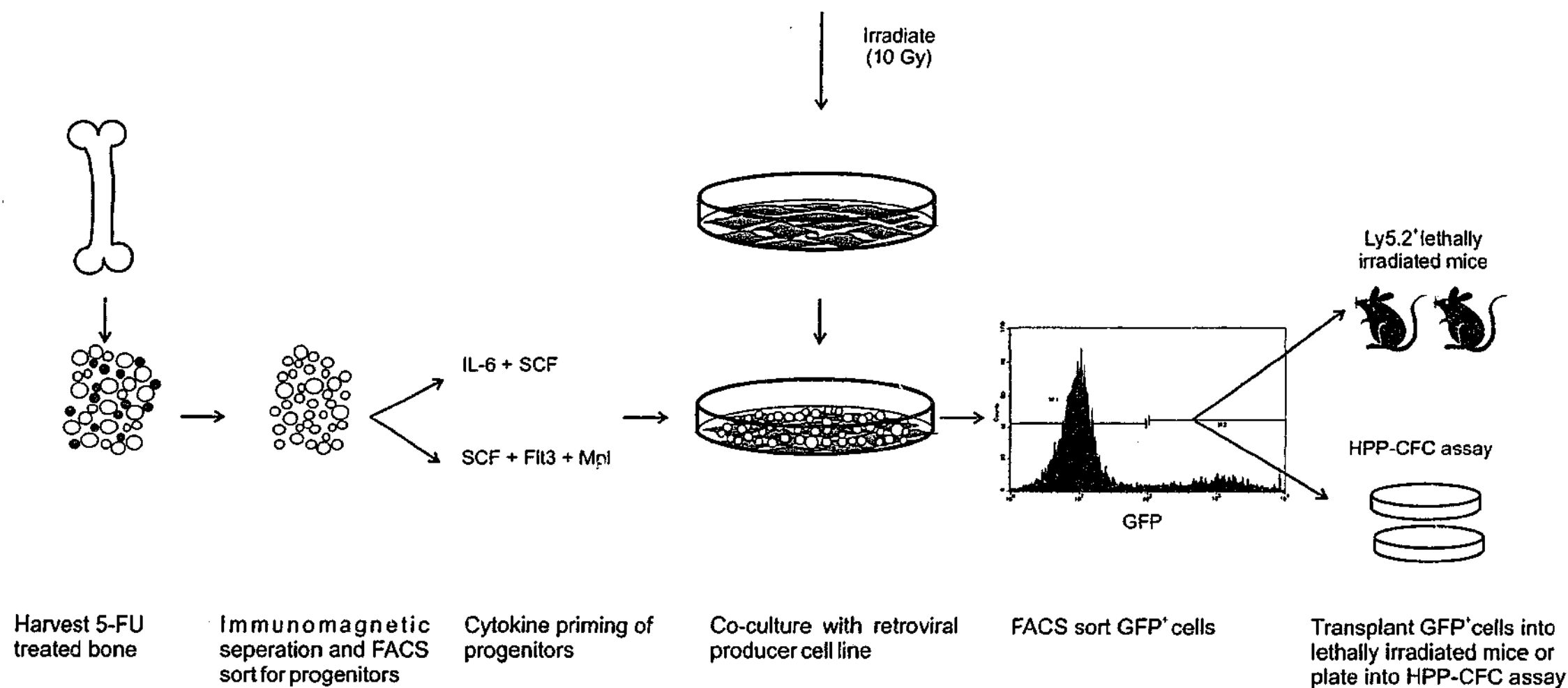


Figure 7.1: Schematic representation of progenitor enrichment and retroviral transduction. Bone marrow from 5-FU_{3d} treated Ly5.1⁺ mice was isolated and enriched for progenitors. After priming in cytokines for 2 days, the progenitor cells were co-cultured for 2 days with irradiated retroviral producer cell lines. Post-retroviral transduction the cells were sorted based on GFP expression and either transplanted into lethally irradiated Ly5.2⁺ mice or plated into *in vitro* HPP-CFC assay.

The retroviral expression vector pMSCV-IRES-GFP was obtained from Dr Paul Simmons, (Peter MacCallum Cancer Centre, Melbourne, Australia). The serpin2A cDNA was subcloned into MSCV-GFP vector at the first translation start site, 5' of the IRES.

7.2.2. *Generation of ecotropic producer cells.*

The pMSCV-IRES-GFP and pMSCV-IRES-GFP/2A were transfected into the ecotropic retroviral cell line GP+E86 by the calcium phosphate transfection procedure according to manufacturer's instructions (Invitrogen: Melbourne, Australia). Transduced cells were sorted by FACS based on GFP expression one week post-transfection and then either expanded in culture (mass culture) or single cell plated into 96-well plates to produce clonal producers of each vector.

A high titre viral stock from each producer population was generated by growing each producer cell line in T₇₅-flasks until they were 80% confluent. 7 ml of fresh culture media was added and harvested once the pH indicator in the media turned yellow. It was then centrifuged at 3,000 x g for 10 min to remove debris, aliquoted and stored at -80°C.

The viral titre of conditioned media from each of these producer cell populations was determined by infection of NIH-3T3 cells. The NIH-3T3's were established in 6-well dishes at 5×10^4 cells per well. The following day either 5 µl, 50 µl or 500 µl of viral supernatant was added to the cells in the presence of 8 µg/ml polybrene and made up to 1.5 ml with DMEM/10% FCS. The cells were incubated for 48 hr at 37°C, 5%CO₂. After incubation the NIH-3T3 cells were washed in PBS, trypsinised and centrifuged at 500 x g. Cells were resuspended in 500 µl PBS and analysed by flow cytometry for GFP expression. Viral titre was expressed as number of cells infected per ml and calculated as follows:

Viral titre = (% GFP⁺ cells x 5x10⁴(NIH-3T3 cells) x dilution factor of supernatant)
(cells/ml)

Each viral titre is the average \pm SEM calculated from the volume of viral supernatant added to NIH-3T3 cells.

7.2.3. Western blot analysis

5 x 10⁵ NIH-3T3 cells were plated into 6cm tissue culture dishes. The following day 1 ml retroviral supernatant from each clone, 8 μ g/ml polybrene and 0.5 ml DMEM/10%FCS was added to separate dishes of NIH-3T3 cells. Cells were incubated for 48 hr at 37°C, 5%CO₂. The cells were harvested and washed in PBS. Transduced cells were lysed in PBS, pH 7.4, 0.2% Triton-X100, 20 μ g/ml PMSF at a concentration of 10⁷ cells/ml for 20 min on ice. The lysate was spun at 16,000 x g for 10 min at 4°C. The supernatant was removed and an equal volume of 2 x SDS-PAGE loading dye added. Purified recombinant His-serpin2A was included as a positive control. Samples were heated at 95°C for 5 min, loaded onto 12.5% SDS-PAGE gels and electrophoresed. The separated proteins were transferred to PVDF membranes by Western blotting. To identify cells expressing the serpin2A gene the membrane was probed with the serpin2A-polyclonal antibody and anti-rabbit-HRP conjugated antibody. Proteins were detected with ECL reagent and visualised by autoradiography.

7.2.4. Mice

Congenic C57Bl/6J (Ly 5.2) and B6.SJL-Ptprc^a Pep3^b/BoyJ (Ly 5.1) were purchased from the Animal Resources Centre (Perth, WA, Australia). The revised terminology for the Ly 5 alleles of the leukocyte common antigen (CD45) as proposed by Morse *et al.*, (1987) has been used in these studies. The Ly5.1 and Ly5.2 congenic markers are cell surface markers present on the two different congenic strains of C57Bl/6 mice (Morse *et al.* 1987). This

provided a means to differentiate between donor cells (Ly5.1⁺) and host cells (Ly5.2⁺) in transplant models.

7.2.5. Bone marrow harvesting and sorting

Regenerating bone marrow was harvested from mice intravenously injected with 5-fluorouracil (FU; David Bull Laboratories, Mulgrave, Australia) via the lateral tail vein at a dose of 200mg/kg body weight and sampled at day 3 post-FU injection (FU_{3d}). Mice were anaesthetised by CO₂ inhalation and killed by cervical dislocation. Bone marrow (BM) cell suspensions were prepared by excising femurs, tibiae and iliac crests then removing any excess tissue from the bones by scraping with a sterile #11 surgical blade. The whole bones were crushed with a sterile mortar and pestle in cold PBS-2% heat-inactivated FCS. The cells were filtered through a 40µm nylon cell strainer (Falcon, Becton Dickinson, NJ, USA) to remove bone debris and washed twice by centrifugation at 500 x g and cells resuspended in 2% FCS/PBS.

Low-density EM (LD-BM) cells were isolated by discontinuous density gradient separation using Nycoprep Animal (density 1.077g/cm³; osmolarity 265mOsm; Nyegaard, Oslo, Norway). The washed BM cells were layered onto a cushion of Nycoprep and centrifuged at 500 x g for 25 min at 4°C, with the brake off to prevent disruption of the gradient. Gradients were set up in 50ml tubes with a maximum of 15 femur equivalents of BM cells in 9ml of 2% FCS/PBS and layered onto 10 ml of Nycoprep. The LD-BM cells (<1.077g/cm³) at the interface of the Nycoprep and PBS were collected into 2% FCS/PBS using a Pasteur pipette and then washed once in 2% FCS/PBS and resuspended at approximately 10⁶ cells/ml.

7.2.6. Negative Immunomagnetic Cell Separation

LD-BM cells were depleted of cells expressing mature haemopoietic cell antigens by immunomagnetic selection using the MACS system (Miltenyi Biotech, Bergisch Gladbach, Germany). The lineage antibody cocktail comprised pre-titred concentrations of antibodies directed against B cells (anti-B220), T cells (anti-CD3, anti-CD4, anti-CD5, anti-CD8), neutrophils (anti-Gr-1), and macrophages (anti-Mac-1) either in their purified or biotin-conjugated forms.

LD-BM cells were centrifuged at 500 x g for 5 min at 4°C and the supernatant decanted. 50 µl of the antibody cocktail described above was added per 5.0×10^6 cells and the suspension incubated on ice for 20 min. The cells were washed and incubated with goat anti-rat microbeads (Miltenyi Biotech, Bergisch Gladbach, Germany) at 13 µl beads plus 87 µl 0.5% BSA/PBS, 5 mM EDTA per 10^7 cells, with regular agitation for 15 mins on ice. The cells were then added to the column (D column, maximum capacity 10^9 cells), run into the mesh and left to magnetise for 5 minutes. The non-magnetic fraction (mature cell lineage antigen-negative cells; Lin⁻) was collected by eluting the cells through a 20-G needle with a solution of 100 ml 0.5% BSA/PBS, 5 mM EDTA. The cells were then washed. The needle was changed to a larger diameter (18-G) and the column demagnetised for 5 min. The bound cells were then flushed into the upper reservoir and the column magnetised for 5 min. The wash fraction was then collected as described above and cell washed in 2% FCS/PBS.

7.2.7. Isolation of Lin⁻ Sca-1⁺ HPC

Lin⁻ cells were isolated as described above and then resuspended at 5.0×10^6 cells per 100 µl. These were then stained with 1 µl of FITC-conjugated stem cell antigen (Sca-1). Residual biotinylated Lin⁺ cells were identified by labelling with 50 µl of Streptavidin-RED670 per 5.0×10^6 cells for 20 min on ice. The

cells were then washed by centrifugation and resuspended at 15.0×10^6 cells/ml for sorting. Non-viable cells were identified by labelling with 7-AAD. Cells for sorting were kept on ice and sorted at a rate of approximately 10,000 cells/sec at 4°C, and collected into serum-coated tubes.

7.2.8. Fluorescence activated cell sorting (FACS)

FACS was performed with a FACStar^{Plus} cell sorter (Becton Dickinson), equipped with a 5-watt argon ion laser running at 200mW power, and a Spectra-Physics ultraviolet (UV) laser (Mountain View, CA), running at 50mW power. Low angle forward scatter (FSC) and 90° light scatter (SSC) were collected through a 488nm band pass 10 filter, and a 1 decade neutral density filter in the forward light scatter path. Green fluorescence (FITC and GFP) was collected through a FITC 530nm filter, with a bandwidth of ± 15 nm. Red fluorescence (PE) was collected through a 575DF26 filter, whilst 7-AAD emission was collected through a red glass filter. Pulses emitted following the excitation of Streptavidin-RED670 were collected through a long-pass RG655 filter. A 610 DMSP (610nm short pass dichroic mirror) was used to separate the emission wavelengths. The PMT voltages were adjusted to compensate for the overlap of FITC, PE and 7AAD emission spectra using cells labelled with 7AAD, FITC or PE alone.

7.2.9. Retroviral transduction of haemopoietic progenitors

Haemopoietic progenitor cells (either low density BM or Lin⁻Sca⁺ cells) were stimulated with either 100ng/ml of stem cell factor (SCF) and IL-6, or 100ng/ml SCF, 100ng/ml Flt-3-ligand and 100ng/ml Mpl-ligand, for 3 days in alpha modification of Eagles Minimal Essential Medium (α -MEM) supplemented with 20% FCS. GP+E86 cell lines stably expressing MSCV/GFP (clone H3) or MSCV/GFP/serpin2A (clone C12) were plated into T₂₅-flasks at 7×10^5 cells. The following day the retroviral producer cell lines were irradiated (10 Gy). The stimulated progenitor cells were then added to irradiated retroviral producer

cells in DMEM/10% FCS, the above cytokines, and 8 µg/ml polybrene. Progenitors were co-cultured for 48 hr at 37°C, 5%CO₂. For transduction of progenitors using retroviral supernatant, the cytokine stimulated progenitors were centrifuged at 500 x g, to remove media and then resuspended in 0.5 mL DMEM/10% FCS, 8 µg polybrene, above cytokines and 1 ml retroviral supernatant prepared as per Section 7.2.2. Cells were incubated for 48 hr at 37°C, 5%CO₂.

7.2.10. Flow cytometry of retrovirally transduced progenitor cells

Progenitor cells were removed from the co-culture and washed twice in PBS. The cells were resuspended in 500 µl of PBS, 7-AAD added for viability and analysed for GFP expression (as described 7.2.8). Cells were then sorted as (as described in Section 7.2.8) and either plated into HPP-CFC assays or used for transplantation. After the completion of the viral transduction procedure, progenitor cells were washed twice in phosphate buffered saline (PBS), and resuspended in PBS for FACS sorting. The cells were treated with 7-AAD 30 mins on ice prior to cell sorting. The GFP⁺ and GFP⁻ were gated as described in Results section and FACS sorted as described in Section 7.2.8. Dead cells were excluded from the sorted population based on their uptake of the viability marker 7-AAD and low forward scatter. Sorted populations were reanalysed for GFP expression and routinely showed purities at approximately 99%.

7.2.11. High Proliferative Potential Colony Forming (HPP-CFC) Assay

The *in vitro* HPP-CFC assay provides a short-term surrogate analysis of the primitive haemopoietic compartment and allows for the identification of cell phenotype based on cytokine requirements (Bertoncello and Kriegler 1997). Two types of colony forming cells are identified by the assay, cells of low or high proliferative potential (LPP-CFC and HPP-CFC). The HPP-CFC cells share many features of primitive haemopoietic stem cells, namely the ability to repopulate, they are multipotential and require multiple cytokines to proliferate.

LPP-CFC cells are more differentiated, lineage restricted and require less cytokines to proliferate (Bertoncello and Kriegler 1997).

Cells were assayed for colony-forming ability in 35mm Petri dishes using a double layer nutrient agar culture system (Bradley and Hodgson 1979). Double strength α -MEM, supplemented with Eagle's MEM vitamins, 1% L-glutamine, and 40% FCS was used, and diluted with double strength agar to give the final working concentrations. Growth factors were diluted in saline/1% BSA and included in the 1 ml underlay, which contained α -MEM medium and agar at a final concentration of 0.5%. The progenitor cells were plated at 1000 cells/dish for Lin⁻Sca⁺ progenitors and low density 5-FU_{3d} BM at 1,250 cells/dish. Target cells were plated in triplicate in a 0.5 ml overlay, which contained above α -MEM medium and agar at a final concentration of 0.3%. The dishes were gassed in plastic boxes at low oxygen tension (Bradley *et al.* 1978), using a gas mixture of 5% O₂, 10% CO₂ and 85% N₂ at a flow rate of 2.0 L/minute for 20 minutes, prior to sealing for incubation at 37°C for 14 days.

Macrophage lineage restricted progenitor cells of low proliferative potential (LPP-CFC) grown in the presence of CSF-1 alone typically formed colonies less than 0.5mm in diameter containing at least 50 and less than 50,000 cells. High proliferative potential colony forming cells (HPP-CFC) grown in the presence of multiple growth factor combinations typically formed colonies greater than 0.5mm in diameter containing 50,000 cells or more, and could be distinguished on the basis of their growth factor requirements (Bartelmez *et al.* 1989). Colonies were counted at times 20 magnification, using a binocular dissecting microscope and indirect light.

7.2.12. Competitive repopulation transplant assay

The relative haemopoietic regenerative capacity of sorted, cultured and manipulated HPC derived from Ly5.1 donor mice was assessed by transplantation into congenic Ly5.2 myeloablated recipients which were lethally irradiated (10 Gy irradiation, 1.3Gy/min, split dose, 4 hours apart) using a ^{137}Cs source (Gamma cell 40, Atomic Energy, Ottawa, Canada), 24 hours prior to the transplant. Each transplant recipient received 1000-5000 sorted or sorted/transduced Ly5.1 donor cells plus 50,000 supporting Ly5.2 bone marrow cells in 200 μl PBS, by intravenous injection via the lateral tail vein. Transplanted mice were maintained in microisolators and analysed for donor contribution to blood cell production at intervals up to 1 year post-transplant.

7.2.13. Transplant analysis

Donor cell engraftment was determined by immunofluorescence and FACS analysis using a PE-conjugated anti-mouse Ly5.1 monoclonal antibody and GFP expression. At 5-6 weeks post-transplantation, 200 μl of peripheral blood from the transplant recipients and three unmanipulated Ly5.1 and Ly5.2 mice was collected via the retro-orbital sinus using heparinized micro-hematocrit tubes. The percentage of donor cells in individual mice was assessed by measuring the percentage of donor-positive cells in individual mice, correcting for background fluorescence utilising a standard curve constructed of mixtures of Ly5.1 and 5.2 peripheral blood cells comprising 0%, 10%, 15%, 25%, 50%, 75% and 100% Ly5.1 cells (Figure 7.2). Excessive red blood cells were lysed by adding 3 ml of 1 x NH_4Cl_2 (freshly made) per 200 μl of whole blood solution, for 10 min on ice. The cell suspension was washed three times in PBS-2%FCS and the remaining nucleated blood cells incubated on ice for 20 min with 50 μl of the Ly5.1-PE antibody.

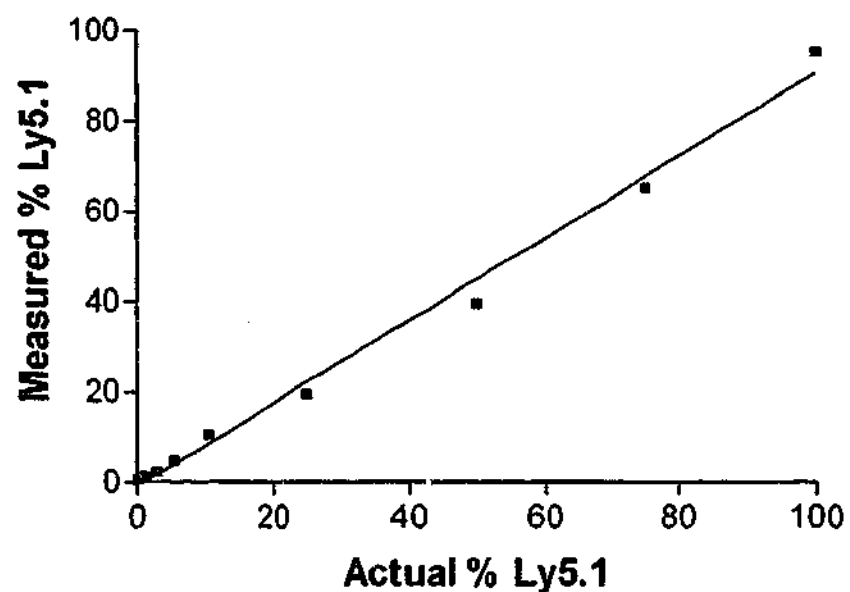


Figure 7.2: Standard curve for the measurement of Ly5.1 expression. In each transplant experiment, peripheral blood cells from Ly5.1 and Ly5.2 mice were titrated and labeled with PE-conjugated mouse anti-mouse monoclonal antibody against the Ly 5.1 epitope to generate a standard curve in order to demonstrate a linear relationship between the percentage of Ly 5.1 measured by immunofluorescence and the percentage of Ly 5.1-positive cells in the sample. Data is representative of one experiment.

7.3. RESULTS

7.3.1. *Establishment of a retroviral producer cell line*

Figure 7.3(A) shows the bicistronic, murine stem cell virus (MSCV)-based retroviral vectors used to enforce GFP or GFP/serpin2A expression in haemopoietic progenitors. The bicistronic vector allows for the co-expression of GFP and serpin2A. The presence of an internal ribosome entry site (IRES) transcriptionally links the co-translation of GFP and serpin2A. The control MSCV-GFP vector contains the GFP cDNA at the second translational position. In the MSCV-GFP/serpin2A vector, the serpin2A cDNA is positioned at the first translational position linked to GFP cDNA through the IRES. The co-expression of the serpin2A gene with GFP facilitates the identification and quantification of expressing cells.

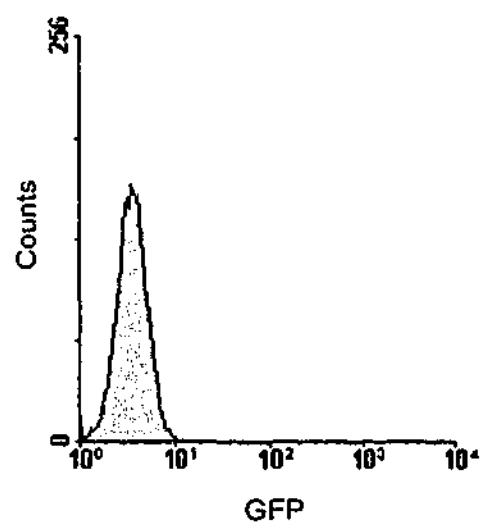
The two vectors were transfected into the GP+E86 ecotropic packaging cell line and one week later analysed for GFP⁺ expression. Flow cytometric analysis shown in Figure 7.3(B) demonstrates intense residual GFP fluorescence in >1% of the population. The brightest 10% of the cells expressing GFP were FACS sorted and either expanded in culture or plated into clonal assays.

Table 7.1 shows the viral titre of the conditioned media obtained from either the expanded culture of GFP⁺ sorted cells or the clonally expanded cell lines. Viral titre was determined by flow cytometric analysis of NIH-3T3 cells infected by retrovirus from either GFP or GFP/serpin2A cell lines and calculated based on the percentage of GFP expression. All viral titres had infection rates, between 1×10^4 – 3×10^5 cells/ml of viral supernatant. The best viral producer clones for each construct were GFP(cH3) and GFP/serpin2A(cC12) which had titres of 2.7×10^5 cells/ml and 1.4×10^5 cells/ml of viral supernatant, respectively. To assess the expression of serpin2A protein, whole cell lysates from NIH-3T3 cells transduced with GFP and GFP/serpin2A were analysed by

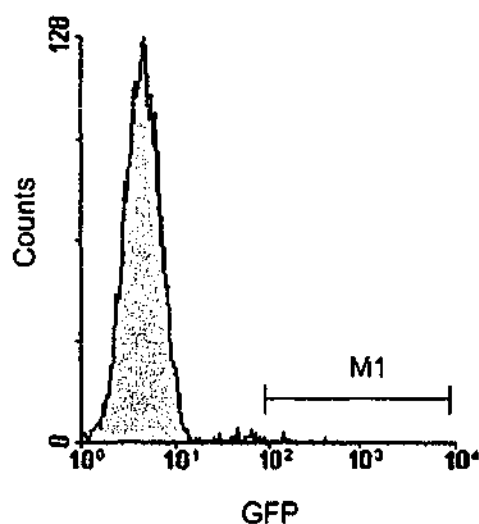
A)



B)



C)



D)

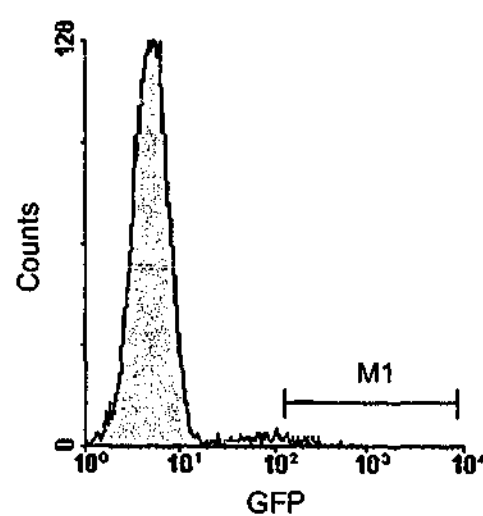


Figure 7.3: A) Schematic representation of MSCV-GFP retroviral vectors. The internal ribosome entry site (IRES) allows bicistronic translation of GFP and serpin2A. Flow cytometry analysis of GFP expression 1 week post transfection of B) GP+E-86 parental cells, C) GP+E-86 transfected with MSCV-GFP, and D) GP+E-86 transfected with MSCV-GFP/serpin2A. M1 represent the cells sorted for the establishment of the stable retroviral cell line.

immunoblotting with the serpin2A polyclonal antibody. Figure 7.4 demonstrates the expression of an immunoreactive band of the correct molecular size (~48kDa) serpin2A in all NIH-3T3 cells infected by the viral producer cells GP+E86-MSCV-serpin2A/GFP. No immunoreactive band was observed in NIH-3T3 cell lysates transduced with control vector MSCV-GFP. The retroviral producer clones chosen for subsequent infections were the ones with the highest calculated titre, the GFP/serpin2A(cC12) and GFP(cH3).

Table 7.1: Viral titre of GP+E86 retroviral producer clones

	GP+E86 viral producer clone				
	GFP (MC)	GFP (cH3)	GFP/serpin2A (MC)	GFP/serpin2A (cA7)	GFP/serpin2A (cC12)
Viral Titre (cells/ml)	$1.1 \pm 2.4 \times 10^5$	$3.0 \pm 1.3 \times 10^5$	$7.6 \pm 2.9 \times 10^4$	$7.2 \pm 3.0 \times 10^4$	$1.4 \pm 0.4 \times 10^5$

7.3.2. Retroviral transduction of murine progenitors by co-culture

Low density BM cells were isolated from bone marrow of 3 day 5-FU treated mice and stimulated with IL-6 and SCF for 48hrs. The cells were co-cultured on either GP+/MSCV-GFP(cH3) or GP+/MSCV-GFP/serpin2A(cC12) retroviral producer cells for 48hrs. Figure 7.5(A) shows the flow cytometric analysis of murine BM progenitor cells after co-culture. Non-viable cells were identified by labelling with 7-AAD and viable progenitor cells were analysed for GFP expression. Approximately 10% and 14% of viable progenitor cells were positive for GFP expression after co-culture with the retroviral producer cell lines MSCV-GFP and MSCV-GFP/serpin2A, respectively.

A significant reduction in viable cells was observed between the number of cells added to the producer cells on day one of the co-culture (1×10^6 cells per/ per flask) and the number of viable cells present 48hrs post incubation,

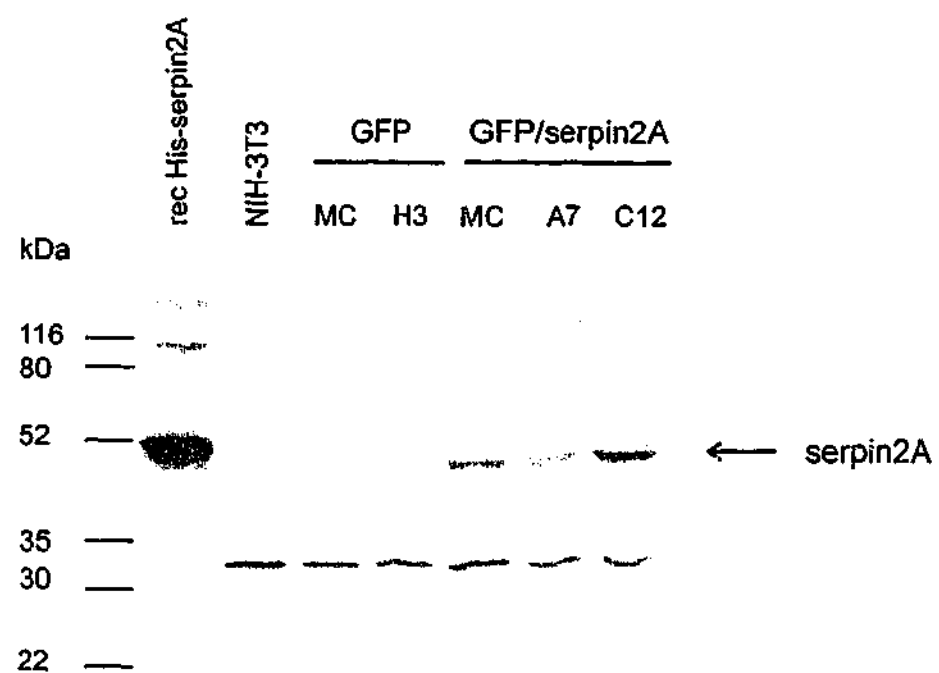


Figure 7.4: Western blot analysis of cell lysates of NIH-3T3 cells transduced with GFP/control or GFP/serpin2A probed with the polyclonal anti-serpin2A. rec His-serpin2A represents recombinant His-serpin2A purified from *Pichia Pastoris*.

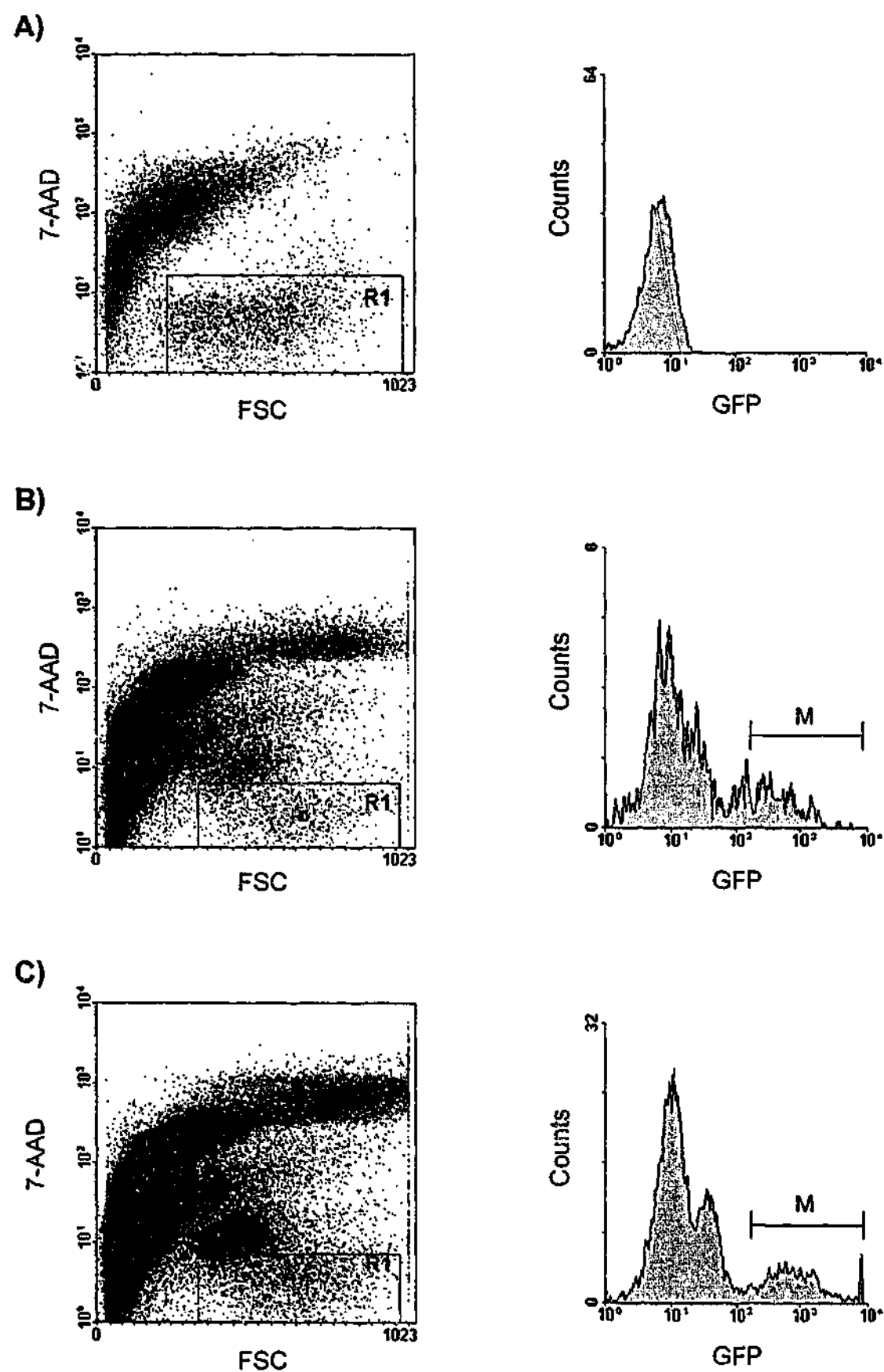


Figure 7.5: Flow cytometric analysis of low density BM cells post retroviral transduction. Cells were gated for viability (R1: 7-AAD⁻) and analysed for GFP expression. **A)** Whole BM from C57Bl/6 (control) **B)** low density BM transduced with GFP/control **C)** low density BM transduced with GFP/serpin2A. M represents GFP⁺ cells sorted for transplant into donor mice.

$\sim 5 \times 10^4$ total for each vector. Figure 7.5(B&C) demonstrates that a significant number of cells post co-culture took up the viability marker 7-AAD in comparison to control cells. These could be either retroviral producers (which have detached from the base of the flask) or dead transduced BM cells. No difference in the percentage of viable cells was observed between GFP and GFP/serpin2A transduced cells. This suggested that loss in BM cell number could be attributed to cell death during co-culture and was independent of serpin2A expression.

The GFP⁺ cells were FACS sorted and yielded approximately 1.5×10^3 and 4×10^3 cells GFP and GFP/serpin2A, respectively. To evaluate the ability of the GFP⁺ sorted progenitor cells to reconstitute haemopoiesis, the cells were combined with 50,000 BM cells of host origin (C57Blk/6-Ly5.2⁺), and injected into lethally irradiated host mice.

Whilst haemopoiesis was effectively reconstituted in all transplanted mice, only two mice showed GFP expression in peripheral blood (PB). Figure 7.6 shows the flow cytometric analysis of peripheral blood from mice 5-weeks post engraftment. Host peripheral blood cells were analysed for cells of donor origin via the expression of the Ly5.1⁺ congenic marker and GFP expression. Figure 7.6 (C&D) show the GFP and Ly5.1(donor) analysis of peripheral blood from the two positive mice. Only a small percentage of cells expressed GFP (less than 0.5%) in both GFP/control and GFP/serpin2A transplanted mice. All PB cells which expressed the Ly5.1 congenic marker also expressed GFP and this was demonstrated for both PB from GFP/control and GFP/serpin2A transplanted mice. This suggested that the low level of GFP expression was not due to gene silencing in donor progenitors.

These results were reproducible over a number of experiments, and the transduced GFP⁺ cells made either no or low levels of contribution to haemopoietic reconstitution in host mice. These data suggested that insufficient numbers of progenitor cells had been transduced and transplanted to achieve a robust contribution in host mice.

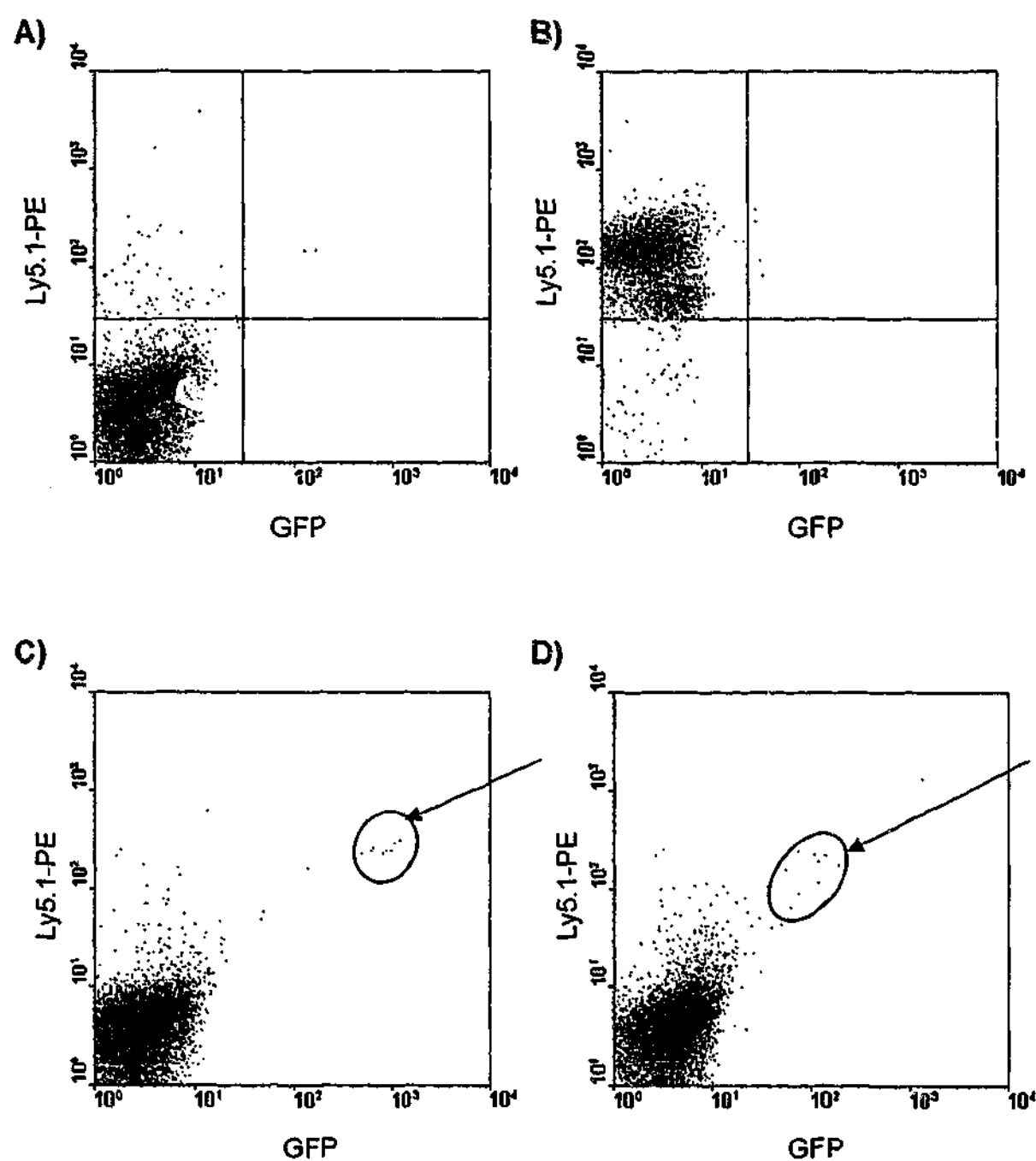


Figure 7.6: Flow cytometric analysis of peripheral blood (PB) from host mice transplanted with retrovirally transduced progenitor cells. Cells were labelled with Ly5.1-PE and analysed for Ly5.1 (donor) and GFP expression. **A)** C57Blk6-Ly5.2(control-host) **B)** C57Bl/6-Ly5.1(donor) **C)** C57Bl/6(host) transplanted with GFP/control cells and **D)** C57Bl/6(host) transplanted with GFP/serpin2A donor cells. Arrows indicate Ly5.1⁺/GFP⁺ cells.

Analysis of the mice one year post transplant demonstrated that a single host mouse transplanted with GFP/serpin2A donor cells possessed a small amount of GFP⁺ cells in peripheral blood (Figure 7.7(A)). These cells were also Ly5.1 positive demonstrating that the cells were of donor origin (Figure 7.7(B)). This suggested that retroviral transduction was able to infect primitive progenitors capable of long-term repopulation.

7.3.3. Optimisation of retroviral transduction protocol

Different cytokine combinations have been shown to promote cycling and proliferation of progenitor cells (Luskey *et al.* 1992; Szilvassy and Cory 1994). To increase the number of primitive progenitors for retroviral transduction an investigation of cytokine conditions was undertaken. A comparison was made between the two cytokine combination IL-6 and SCF, and the three cytokine combination SCF, Mpl ligand and Flt-3 ligand. Figure 7.8 shows the forward scatter (FSC)/7-AAD profile of progenitor cells stimulated by the two different cytokine cocktails and subsequently transduced with GFP/serpin2A. In comparison to untreated BM cells (Figure 7.8(A)), a significant proportion of treated cells took up the viability marker 7-AAD (Figure 7.8(B&C)). Live cells (7-AAD⁻) were gated and subsequently analysed for GFP expression. No significant difference in GFP expression of cells cultured in cells was observed between cells cultured in IL-6 and SCF (22% GFP⁺) and SCF/Mpl-L/Flt-3-L (16% GFP⁺). The viable cells were sorted for GFP⁺ based on the position of the markers shown in Figure 7.8(B&C).

The HPP-CFC assay was used as an *in vitro* surrogate assay of the stem cell potential of GFP⁺ sorted cells. Table 7.2 shows that the three-cytokine combination of SCF, Mpl-L and Flt-3-L gave a 10-fold increase in total HPP-CFC number in comparison to the two-cytokine cocktail. When HPP-CFC numbers were adjusted for total CFC number, the three cytokine cocktail still gave a 5-fold increase in HPP-CFC. This suggested that lower numbers of progenitors were stimulated into cycle and to proliferate with IL-6 and SCF.

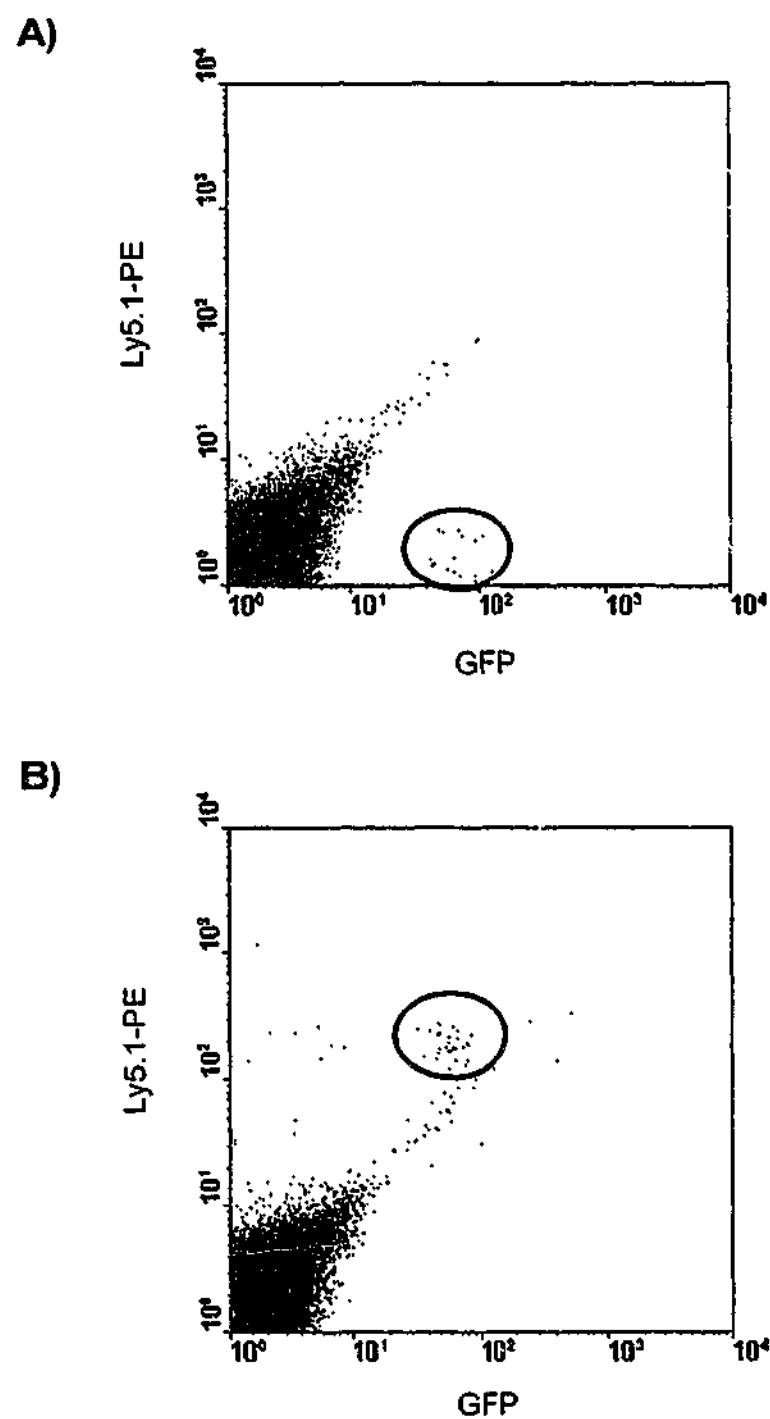


Figure 7.7: Peripheral blood analysis 1 year post transplant of a single host mouse transplanted with progenitors transduced with GFP/serpin2A. **A)** Unlabelled peripheral blood analysed for GFP (x-axis) vs. Ly5.1-PE auto-fluorescence (y-axis). **B)** Peripheral blood labelled with Ly5.1-PE and analysed for GFP (x-axis) and Ly5.1-PE (y-axis). Circled population indicates Ly5.1⁺/GFP⁺ cells of donor origin.

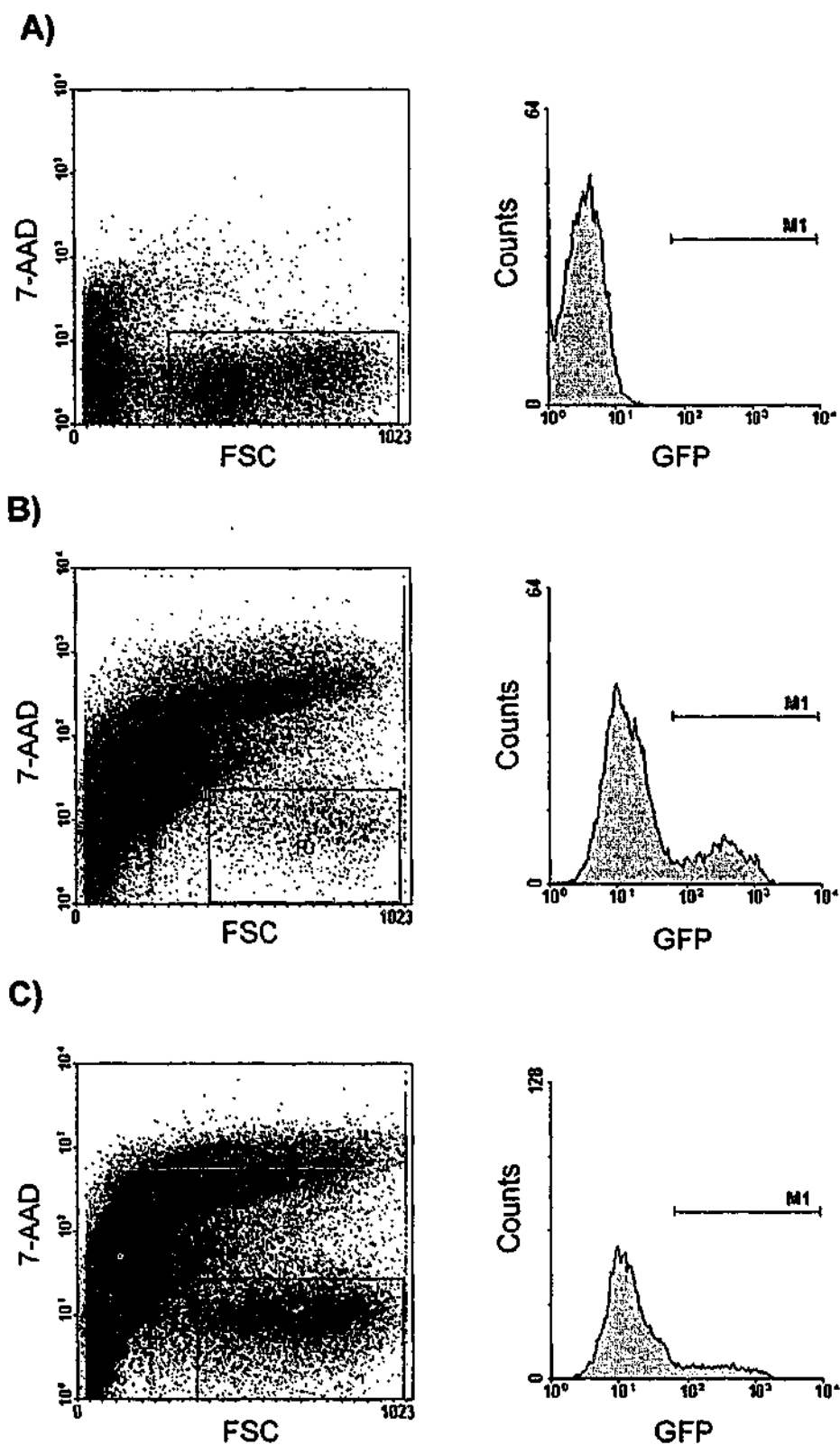


Figure 7.8: Flow cytometric analysis of low density BM post cytokine pre-stimulation and retroviral transduction with GFP/serpin2A. Cells were stained with 7-AAD prior to analysis. **A)** Whole BM analysis of 7-AAD and GFP expression. **B)** Low density 5-FU BM pre-stimulated with IL-6 and SCF and **C)** low density 5-FU BM pre-stimulated with Mpl-L, Flt-3-L and SCF. M1 represents GFP⁺ cells sorted for HPP-CFC assays.

Table 7.2: Colony forming assay of GFP⁺ post priming in either two cytokine or three cytokine cocktails.

		LPP-CFC	HPP-CFC
		Colonies per 25,000 cells	
SCF + IL-6	GFP ⁺	962	25
SCF, Flt-3-L + Mpl-L	GFP ⁺	1475	250

Values are the Mean of duplicate plates from the single experiment.

To further enrich for primitive progenitors, the low density 5-FU BM derived cells were enriched for cells lacking lineage markers (Lin⁻) and sorted on the basis of Sca antigen expression (Sca⁺). The gating strategy employed for this isolation is shown in Figure 7.9 (A&B). The sorted Lin⁻Sca⁺ cells were stimulated with the three cytokine combination and exposed to GFP/control and GFP/serpin2A irradiated retroviral producer cell lines. Figure 7.9 (C&D) shows the flow cytometry analysis of progenitors post co-culture. Analysis based on 7-AAD uptake showed that a significant proportion of cells were 7-AAD⁺, demonstrating again that many cells died during co-culture. GFP expression analysis of 7-AAD⁻ cells (Quadrant R1) showed that 12% of GFP/control and 22% of GFP/serpin2A cells were transduced with the retroviral vector. The GFP⁺ cells were then FACS sorted and either transplanted into lethally irradiated mice or analysed by HPP-CFC assays.

Table 7.3 shows a 67-fold increase in HPP-CFC number after enrichment of low density 5-FU BM for Lin⁻Sca⁺ cells. After cytokine treatment and retroviral transduction of Lin⁻Sca⁺ progenitor cells a 5-fold increase in total HPP-CFC number is observed. Comparisons between GFP⁺ and GFP⁻ fractions suggested no significant difference between HPP-CFC number, although progenitors transduced with GFP/control appeared to have higher HPP-CFC in the GFP⁺ progenitor cells. Given the high number of HPP-CFC present post

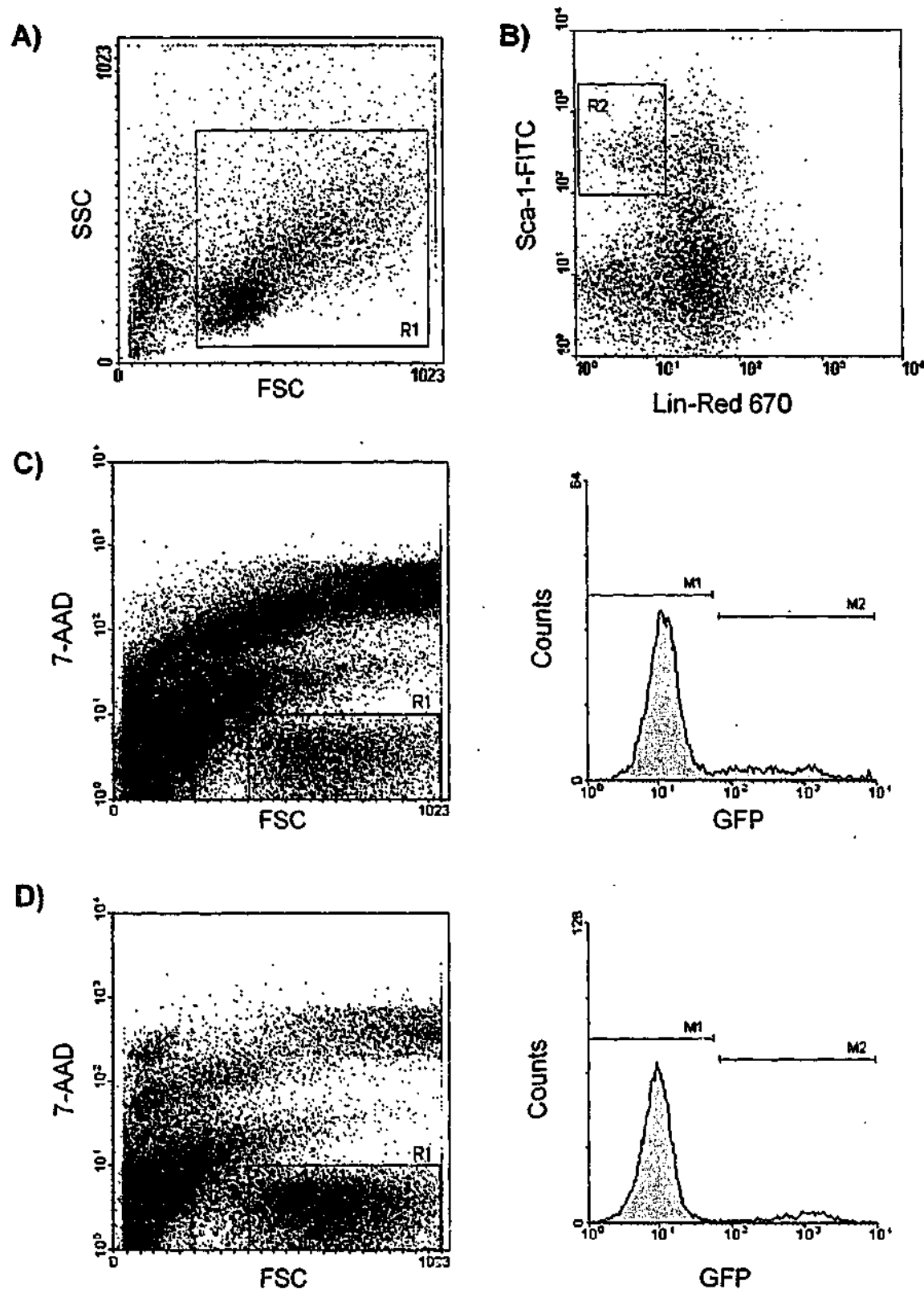


Figure 7.9: A&B) FACS sort of low density 5-FU BM cells. Cells gated in quadrant R1 were sorted based on Lin⁻ and Sca⁺ expression (quadrant R2). Lin⁻Sca⁺ sorted cells were stimulated with Flt-3-L, Mpl-L and SCF and then co-cultured with either GFP/control or GFP/serpin2A retroviral producer cells. **C)** Lin⁻Sca⁺ progenitors transduced with GFP/control. **D)** Lin⁻Sca⁺ progenitors transduced with GFP/erpin2A. 7-AAD⁺ cells were gated (R1) and then analysed for GFP expression. Cells were then sorted based on GFP⁻(M1) or GFP⁺(M2).

transduction we would expect to see a robust contribution by donor cells to engraftment in host mice.

Table 7.3: HPP-CFC assay of sorted progenitor cells

		LPP-CFC	HPP-CFC
		Colonies per 25,000 cells	
Primary explanted and sorted bone marrow cells	Unfractionated FU _{3d} bone Marrow	66.3 ± 2.9	7.0 ± 1.2
	Lin ⁻ Sca-1 ⁺ cells	892 ± 106	467 ± 33
Following priming and co-culture with producer cell lines	MSCV GFP ⁻	633 ± 156	300 ± 25
	GFP ⁺	1692 ± 358	908 ± 183
	Serpin2A GFP ⁻	1875 ± 200	488 ± 63
	GFP ⁺	1540 ± 175	408 ± 71

Values are the Mean ± SEM of one experiment plated in triplicate in the presence of SCF + IL-1 + IL-3 + CSF-1

When transduced Lin⁻Sca⁺ cells were transplanted into host mice they made no contribution to reconstituted haemopoiesis. This was demonstrated by the absence of GFP and Ly5.1 expression in peripheral blood 6-7 weeks post-transplant. This was in direct contrast to the data obtained from the HPP-CFC assays which demonstrated a significant number of progenitor cells present with high proliferative potential. Based on this, we would have expected to see a more robust engraftment in host mice.

7.3.4. Retroviral transduction of Lin⁻ Sca⁺ progenitors using retroviral supernatant

One of the concerns raised during the infection of the progenitors was the significant amount of cells to stain positive with the viability marker 7-AAD. In addition, a large amount of debris within the infection media was also apparent during and post infection. The presence of debris was attributed to the retroviral feeder layer detaching from the flask and it was postulated that this may adversely influence the progenitors. To overcome these issues it was decided to incubate the enriched progenitors in the absence of the viral producer cell line using the retroviral supernatant.

Incubation of progenitor cells with retroviral supernatant resulted in a dramatic increase in progenitor numbers present post transduction with both GFP/control and GFP/serpin2A vectors. The cells doubled every 12 hrs in contrast to the dramatic loss in cell number observed with retroviral co-culture. Flow cytometry analysis for GFP expression revealed a low level of retroviral transduction (2%), with both GFP/control and GFP/serpin2A supernatants. The low level of transduction was not surprising as the retroviral titre was determined as 1×10^4 cells/ml and 2.9×10^4 cells/ml for the GFP/control and GFP/serpin2A supernatants respectively. Given the low level of transduction the cells were not sorted and the complete population of *in vitro* cultured cells were transplanted into lethally irradiated mice.

Figure 7.10(A) shows the results of flow cytometry analysis of GFP expression and Ly5.1 labelling of PB isolated from host mice. Only a single mouse transplanted with GFP/serpin2A transduced progenitor cells had low level GFP expression (1.3%). No GFP expression in PB was detected in the other transplanted mice despite haemopoiesis being restored. As host mice were transplanted with unsorted donor cells, progenitors that were GFP⁻ but Ly5.1⁺ could also contribute to host reconstitution. Flow cytometry analysis using the congenic marker Ly5.1 showed that transduced GFP/control and GFP/serpin2A donor cells contributed approximately 50.5% and 63.3%,

A)

	% Ly5.1 ⁺	% GFP ⁺ / Ly5.1 ⁺
GFP/serpin2A mouse #1	65.4	1.3
GFP/serpin2A mouse #2	59.7	0
GFP/serpin2A mouse #3	64.9	0
GFP/control mouse #1	44	0
GFP/control mouse #2	53.4	0
GFP/control mouse #3	54.2	0

B)

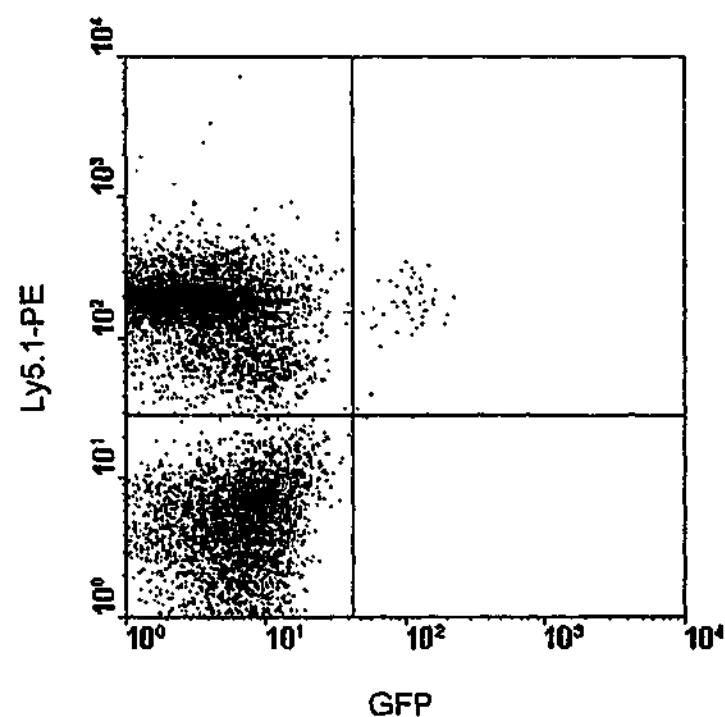


Figure 7.10: A) Summary of the flow cytometric analysis of peripheral blood from host mice transplanted with GFP/serpin2A and GFP/control. Lin⁻Sca⁺ cells were transduced with retroviral supernatant and the complete population introduced into host mice. Peripheral blood cells were labeled with Ly5.1-PE and analysed for GFP and Ly5.1 expression to identify cells of donor origin. **B)** Dot plot of the analysis of the single host mouse with positive for donor cells (Ly5.1⁺/GFP⁺) transduced with GFP/serpin2A.

respectively (Figure 7.10(A)). This demonstrated that Lin⁻Sca⁺ donor cells could engraft and contribute to haemopoiesis in host mice.

7.4. DISCUSSION

To investigate the biological role of serpin2A in haemopoiesis, a retroviral mediated gene transfer system was developed to introduce constitutive serpin2A expression into haemopoietic progenitors. This system successfully yielded clonal viral producer cell lines capable of infecting target mouse cells with GFP and GFP/serpin2A. The presence of the IRES in the retroviral vector translationally linked the expression of GFP and serpin2A and facilitated the identification of vector expressing cells and their progeny. Western blot analysis confirmed the expression of serpin2A in cells infected with virus from the producer line GP⁺E86-MSCV/GFP/serpin2A.

Progenitor cells transduced with either GFP/control or GFP/serpin2A resulted in varied levels of transduction efficiency between experiments, with 10-50% GFP expression as assessed by flow cytometry. There was no significant difference in transduction efficiency between the producer cell lines, GFP/control and GFP/serpin2A. A consistent feature of the retroviral transduction protocol was the high level of cell death observed after co-culture. The high uptake of the viability marker 7-AAD could be attributed to either shedding of the retroviral producer layer or death of transduced progenitors. Given the considerable reduction in progenitor numbers post-transduction, it is likely that many of the 7-AAD⁺ cells are progenitors.

The use of retroviral supernatant to infect progenitors did not result in high levels of cell death as demonstrated by a 20-fold increase in cell number post retroviral transduction. However, by using the retroviral supernatant only 2% of the cells were transduced. Therefore, it is difficult to ascertain whether the difference in cell death observed was due to vector expression in progenitor cells, or due to culture conditions. In order to distinguish between these

possibilities, a higher retroviral titre supernatant would need to be produced and examined for the effect on progenitors.

The HPP-CFC assay was used as a surrogate assay to identify the phenotype of the haemopoietic progenitors present pre- and post-retroviral transduction and to determine reconstitution potential. HPP-CFC's share many features with primitive haemopoietic stem cells; the ability for long-term haemopoietic reconstitution, they are able to proliferate in the presence of multiple cytokines, and are multipotential (Bertoncello and Kriegler 1997). As expected, the HPP-CFC assays revealed that HPP number was enriched 67-fold from 5-FU treated low density bone marrow to cytokine treated Lin⁻Sca⁺. Retroviral transduction of progenitor cells with either GFP/control or GFP/serpin2A did not affect HPP-CFC numbers, and there was no difference observed in HPP-CFC number between GFP⁺ and GFP⁻ fractions. Furthermore, the numbers of HPP-CFC present after co-culture would predict a robust engraftment in ablated mice.

Previous studies by Hampson *et al* (1997) demonstrated that constitutive expression of serpin2A in the multipotential cell line FDCP-MixA4 resulted in an increase in clonogenicity. We observed no difference in colony numbers between serpin2A and control transduced cells. Although these results are apparently contradictory, the two studies are not directly comparable. The current study utilised primary isolated and enriched progenitors whilst Hampson *et al* used FDCP-MixA4 cells which require IL-3 and serum for the maintenance of a primitive, self-renewing phenotype (Spooncer *et al.* 1984). This would suggest that the cell line is more differentiated than the multi-cytokine dependent progenitors used in this investigation.

The ability of donor cells to reconstitute haemopoiesis in host mice was assessed by competitive repopulation studies. The retrovirally transduced progenitor cells failed to contribute more than 2% to haemopoietic repopulation in host mice. This was observed with both GFP and GFP/serpin2A transduced cells. These observations were in direct contrast to *in vitro* HPP-CFC assays which predicted that the number of HPP-CFC present in these populations

would produce a robust engraftment in host mice. Furthermore, untransduced Lin⁻Sca⁺ donor cells contributed between 40-65% to haemopoiesis in host mice. These results validated that the enrichment process and cytokine incubations did not affect the ability of the donor cells to reconstitute haemopoiesis in host mice. This suggests that either co-culture with the retroviral producer cell lines or retroviral transduction had altered their engraftment potential.

The MSCV retroviral vectors have previously been shown to efficiently infect haemopoietic progenitors, and these transduced cells were able to reconstitute haemopoiesis in host mice. (Persons *et al.* 1997; Persons *et al.* 1999; Leimig *et al.* 2002) Therefore, it is unlikely that the presence of the vector altered the engraftment potential of the progenitors. Furthermore, this vector system has recently been shown to correct the lysosomal disorder, galactosialidosis, in cathepsin A knockout mice by retroviral gene transfer of cathepsin A into haemopoietic stem cells. Leimig *et al.* (2002) demonstrated that progenitor cells transduced with the MSCV vector resulted in 20-50% GFP expression in erythrocytes, lymphocytes and platelets from peripheral blood of transplanted mice.

The migration of HSCs through the circulation and back to organs that support their proliferation and differentiation, such as the bone marrow, is referred to as homing. To successfully reconstitute and contribute to haemopoiesis in ablated hosts requires the maintenance of self-renewal capacity in the donor cell population, and the ability of these cells to home to the bone marrow (Szilvassy *et al.* 1999). The HPP-CFC assays indicated that progenitor cells post transduction maintained self-renewal capacity due to their ability to form HPP-CFC *in vitro*. One possible explanation for the absence of GFP⁺ cells in the periphery of host mice is that the homing capacity of the progenitors may have been altered by co-culture or retroviral transduction. Another possibility was that the donor cells and their progeny were retained within the bone marrow. However, no difference in GFP expression was observed between peripheral blood and bone marrow from reconstituted mice.

Successful reconstitution is also dependent on survival of these primitive cells during the period of *in vitro* manipulation. In the current investigation, the co-culture of progenitor cells with retroviral producer cell lines consistently resulted in a loss in progenitor number associated with cell death. This suggests that necrotic or apoptotic pathways may have been stimulated in the donor cells. The activation of apoptosis has been implicated to target donor cells for destruction upon transplantation, subsequently compromising their ability to engraft in host mice (Young *et al.* 2001; Moore 2002).

7.5. CONCLUSION

In this investigation, we set out to analyse the effect of constitutive serpin2A expression on haemopoietic progenitor differentiation *in vivo* by establishing a retroviral-mediated transduction protocol. Using this method we were unable to obtain successful engraftment in ablated mice but this effect was not related to serpin2A expression. The role of serpin2A in primitive haemopoietic progenitors *in vivo* therefore remains an open question.

CHAPTER 8

GENERAL DISCUSSION

8.1. *The mouse *serpina3* locus*

Members of the mouse *serpina3* locus have undergone remarkable expansion and diversification within the RCL. The multiplication of the murine *serpina3* genes was most likely driven by positive selection to have conserved the serpin structure whilst expanding the range of potential protease targets through the hypervariable RCL. This is the first study to comprehensively demonstrate that the individual members of the *serpina3* locus are functional and the hypervariable RCL mediates diverse activity.

Nature has adopted two methods to increase the functional diversity of serpin loci. In higher organisms, gene duplication and gene conversion events have resulted in serpin multigene families with hypervariable RCL. These include the mouse *serpina3*, α_1 -AT, PI-6, PI-8 and MNEI which have evolved to encode fourteen members for the *serpina3* loci, five paralogues for both mouse α_1 -AT and PI-6, and four and seven paralogues for MNEI and PI9 respectively (Borriello and Krauter 1991; Kaiserman *et al.* 2002; Forsyth *et al.* 2003). Instead of duplication, the variable genes of *Manduca Sexta*, *Drosophila* and *C.Elegans* arise from the multiplication of the exon encoding the RCL. Sequence divergence within these exons has led to the ability to produce different RCL from a single gene by alternate exon splicing. *Manduca sexta* contains twelve alternatively spliced variants (Jiang and Kanost 1997) whilst the *Drosophila* gene *sp4* encodes ten different mRNA's (Kruger *et al.* 2002) and the *srp-7* gene in *C.Elegans* has three variants (Jiang and Kanost 1997; Pak *et al.* 2004). The mouse *serpina3* locus is remarkable in being the largest multiplication observed in any species to date.

In the absence of an evolutionary advantage, it would be expected that the *serpina3* genes would fall into disuse and degenerate into pseudogenes. However, only one member of the family, *serpina3e* (2B2(b)), is predicted to be pseudogene due its truncated 5' end. The other thirteen members encode serpins with full open reading frames. Furthermore, the variable tissue specific expression of the *serpina3* genes, as demonstrated in this study, supports the idea that individual gene products carry out unique functions.

Human α_1 -ACT is prominently expressed in liver and is secreted into the circulation (Travis and Salvesen 1983). The current study demonstrated three mouse *serpina3* genes; EB22.4, MMCM2 and 3E46, with high liver expression. All possess predicted N-terminal secretion peptides, consistent with localisation into the extracellular compartment and plasma. However, the relative concentrations of α_1 -ACT orthologues in mouse plasma are unknown. An antichymotrypsin-like protein, contrapsin, with inhibitory activity towards trypsin, was previously purified from mouse plasma (Takahara and Sinohara 1982). The relationship between the plasma contrapsin activity described by Takahara *et al* (1982) and the contrapsin gene identified in liver by Hill *et al* (1984) is unclear. Based on the studies presented in this thesis, the liver expressed serpins 3E46 and MMCM2 have inhibitory activity towards trypsin and therefore, either could be responsible for the reported contrapsin activity. Four additional *serpina3* genes, with predicted N-terminal secretion peptides, are expressed in other mouse tissue. It is more likely that these serpins act locally in the tissues of their expression. Clues to the *in vivo* function of these genes may be provided by elucidation of the cells that express them.

The basis for tissue specific gene expression remains unknown as the 5' sequences upstream of the transcription start are yet to be fully investigated, except for *serpin2A*. The promoter region for *serpin2A* was previously shown to contain a NF- κ B and a Moloney murine leukaemia enhancer factor Lva binding sites (Hampson *et al*. 2001). Subsequent studies demonstrated that *serpin2A* mRNA was strongly induced by NF- κ B in murine embryonic fibroblasts (Liu *et al*. 2003). Consistent with the presence of a NF- κ B binding

site *serpin2A* expression was reported to be induced in macrophages by the inflammatory stimuli LPS and bacterial products (Hamerman *et al.* 2002). In addition, a putative STAT binding site was reported for its stimulation by IFN- γ (Hamerman *et al.* 2002). Inflammatory stimuli have also been reported to up-regulate the expression of *serpina3* genes in the liver but it is unclear which members of the cluster were induced (Hill *et al.* 1985; Inglis *et al.* 1991). No locus control regions have been identified nor the role of chromatin structure analysed. Further work is required to fully understand the mechanisms involved in gene regulation of the locus.

Analysis of the predicted amino acid sequences of the *serpina3* members demonstrates that the *serpina3* locus is distinct from other serpin multigene families. It does not encode an obvious murine orthologue of human α_1 -ACT, as none have RCL composition like their human counterpart. This is unlike the other murine expanded serpin loci which all have an obvious murine orthologue (Goodwin *et al.* 1997; Benarafa *et al.* 2002; Kaiserman *et al.* 2002).

8.2. Molecular characterisation of extracellular *serpina3* members

From the work presented in this thesis, EB22.4 is likely to be the closest functional murine orthologue of human α_1 -ACT. His-EB22.4 is a fast inhibitor of cathepsin G with an SI of 1 and k_a of $7.9 \times 10^5 \text{ M}^{-1}\text{sec}^{-1}$ which is equivalent to the k_a reported for human α_1 -ACT (Lomas *et al.* 1995 (b); Duranton *et al.* 1998). However, the P_1 - P_1' of EB22.4 is Met-Ser and this contrasts with Leu-Ser for human α_1 -ACT. The presence of Met at P_1 explains the fast inhibitory activity observed towards elastase, with an SI of 1 and k_a of $4 \times 10^6 \text{ M}^{-1}\text{sec}^{-1}$. The elastase inhibitory activity is unlike human α_1 -ACT which is a substrate of elastase. The fast inhibition of cathepsin G and elastase makes EB22.4 a good candidate to be the physiological inhibitor of both proteases.

Biophysical and biochemical studies demonstrated that recombinant His-MMCM2 is an inhibitory serpin. Analysis of interactions with a panel of proteases revealed His-MMCM2 to be an inhibitor of trypsin but not thrombin or factor Xa. Kinetic assays showed a moderate inhibition of trypsin, with an SI of 10 and a k_a of $2.7 \times 10^4 \text{ M}^{-1}\text{sec}^{-1}$. The interactions of MMCM2 and its cognate protease in the physiological context may be more favourable than those observed in this study. It is likely that *in vivo* MMCM2 in mouse plasma would interact with other trypsin-like proteases, for example, those involved in the fibrinolytic pathway and complement activation.

The RCL sequence of 3E46 is strikingly similar to both human and mouse Protein C Inhibitor (PCI) and the serpins share expression in liver and testis. This suggests that they may have a similar target protease profile and overlap in function. The current investigation demonstrated that 3E46 is able to form inhibitory complexes with trypsin-like proteases, including trypsin, thrombin and factor Xa. As expected, this inhibitory profile is similar to PCI, which is also able to inhibit these clotting factors (Espana and Griffin 1989). However, gene knockout studies demonstrated that PCI^{-/-} mice were viable but were infertile as a result of impaired spermatogenesis (Uhrin *et al.* 2000). This suggests that, at least in mouse testis, 3E46 was not able to compensate for loss of PCI inhibitory activity. Further experiments with a functional recombinant are required to fully characterise interactions with cognate protease/s to give pointers to the *in vivo* function of 3E46.

8.3. Molecular characterisation of intracellular *serpina3* members

The current study investigated the biochemical features of the predicted intracellular member of the *serpina3* locus, 6C28. The bacterial recombinant His-6C28 demonstrated atypical unfolding in biophysical studies with a late transition at 75.9°C. The increased stability is unlikely to be attributable to misfolding of the serpin in bacteria as this was also observed when *in vitro* transcribed/translated protein was analysed.

Other native inhibitory serpins have melting temperatures between 45°C and 65°C. The question arises whether recombinant His-6C28 is latent. It is known that latent PAI-1 and the δ -conformation of α_1 -ACT (Leu55→Pro) variant have high melting temperatures of 67.5°C and 70.7°C respectively (Lawrence *et al.* 1994 (b); Gooptu *et al.* 2000). Both these conformations render the serpin molecule functionally inactive (Lawrence *et al.* 1994 (c); Gooptu *et al.* 2000). However, in the case of His-6C28, loop cleavage of *in vitro* transcribed/translated protein resulted in a stable species completely resistant to unfolding and complex formation assays suggested inhibitory activity towards a number of proteases. There were several predicted amino acid differences between the 6C28 cDNA isolated from spleen and the sequence present in the mouse genomic database. Although the majority of these variations were to conserved residues present in other members of the mouse *serpina3* locus or in the serpin superfamily, we can not rule out that these variations did not alter the stability of the protein. Isolation of an independent cDNA clone from an alternative mouse strain would clarify whether these amino acid substitutions contribute to the unusual biophysical characteristics of the protein. Formal demonstration of 6C28 hyperstability awaits a functional recombinant protein in which biophysical and inhibition studies can be performed in detail.

What is the *in vivo* function of 6C28? From the studies presented in this thesis, 6C28 is a cross-class inhibitor, able to form inhibitory complexes with serine and cysteine proteases. Gene expression studies demonstrated prominent expression in a number of lymphoid tissue and fibroblasts however its expression was absent in haemopoietic cells. Therefore, 6C28 is unlikely to interact with cathepsin G and elastase intracellularly as these proteases are exclusively expressed in the haemopoietic compartment. The presence of an inhibitory complex with cathepsin L also suggests that 6C28 may inhibit lysosomal cysteine proteases. This is similar to the observations with *serpin2A* which has also been demonstrated to inhibit cysteine proteases (Liu *et al.* 2003). The identification of cell types that express 6C28 may provide clues towards *in vivo* protease targets.

Previous studies have shown serpin2A to be expressed in primitive haemopoietic progenitors and there was some evidence to suggest modulation of differentiation (Hampson *et al.* 1997). We therefore developed a system in which the biological role of serpin2A *in vivo* could be examined. This method employed retroviral mediated gene transfer of serpin2A into primary isolated progenitors. We showed that the primary isolated progenitors were multi-potential in colony forming assays and allowed haemopoietic re-constitution in radio-ablated mice. Furthermore, we could transduce these HPC with control and serpin2A constructs and colony forming assays demonstrated that transduced cells were multipotent, demonstrating their potential to give rise to multi-lineage engraftment. However, both control and serpin2A transduced progenitors contributed minimally to reconstituted haemopoiesis in competitive repopulation assays. Clarification of the function of serpin2A in murine haemopoietic progenitors and its role in hematopoiesis will require further studies.

In order to overcome the minimal engraftment in host mice, new viral producer cell lines need to be generated to give efficient transduction of progenitors. An alternative to retroviral vectors are those derived from lentiviruses. These vectors can transduce non-dividing cells and are therefore more efficient in transducing quiescent haemopoietic progenitors than retroviral vectors (Logan *et al.* 2002). Changing the vector used may help to overcome the low viral titre and extensive cell death observed in the current system. An alternative approach would be to produce transgenic mice with expression of serpin2A targeted to the haemopoietic lineages. For example, either the CD45 or *vav* promoters could be used to target serpin2A to the stem cell compartment and for multilineage expression (Ogilvy *et al.* 1999; Radomska *et al.* 2002).

8.4. FUTURE DIRECTIONS

The main conclusion drawn from this thesis is that the gene expansion observed at the mouse *serpina3* locus gives rise to functionally diverse serpins. They are expressed in a tissue specific manner and are likely to undertake specialised inhibitory activities consistent with diverse protease inhibitory profiles. The challenge is to establish the roles of these serpins *in vivo* and to ascertain their relationship to human α_1 -ACT.

For two members investigated in this study, namely 3E46 and 6C28, functional recombinant proteins are required to characterise their inhibitory interactions with proteases. This would require expression in an alternative system such as, *Pichia pastoris* or baculovirus. The recombinant proteins generated by these expression systems could then be used for biophysical and biochemical analysis. In particular, a functional recombinant of 6C28 would elucidate whether the unusual biophysical characteristics observed in this investigation are truly a feature of the 6C28 serpin structure or are a reflection of expression in bacteria or *in vitro* transcription/translation.

The absence of a predicted N-terminal secretion peptide for 6C28 suggests an intracellular localisation. Further work is required to investigate the subcellular localisation of 6C28 which could be accomplished by insertion of the cDNA into a mammalian expression vector and the transfection into mammalian cell lines. Subcellular localisation could be determined by indirect immunofluorescence.

Human α_1 -ACT has been shown to bind the amyloid β peptide and is implicated in the progression of Alzheimer's disease (Abraham 2001). The crossing of plaque-producing PDGF-APP mice with transgenic mice expressing human α_1 -ACT in astrocytes resulted in a 2-fold increase in plaque formation compared to control mice (Mucke *et al.* 2000). EB22.4 is the closest orthologue of human α_1 -ACT and is highly expressed in mouse brain. We would expect that EB22.4 is also expressed in astrocytes however, RT-PCR of

cDNA from primary isolated astrocytes would need to be undertaken to confirm its expression. To determine whether EB22.4 has the same effect on plaque formation as human α_1 -ACT, transgenic mice producing EB22.4 in astrocytes driven by the glial fibrillary acidic protein (GFAP) promoter could be produced (Mucke *et al.* 2000). These mice could then be crossed with the plaque-producing PDGF-APP transgenic mice and assessed for amyloid plaque formation in EB22.4/APP brain. These studies could further be extended in EB22.4 knockout mice. Crossing of PDGF-APP transgenics with EB22.4^{-/-} would predictably have a reduced rate of amyloid plaque formation.

The studies in this thesis present a biophysical and functional characterisation of EB22.4. Human α_1 -ACT structure provides a poor comparator because of its relatively low resolution and results from a chimeric recombinant protein. We have a recombinant protein which has been purified to homogeneity and is suitable for crystallisation and structure determination. The data from this could be compared with not only α_1 -ACT but also α_1 -AT.

This study investigated only five members of the *serpina3* locus. A further four members were shown to be expressed in mouse tissue and cell lines. The generation of recombinant proteins and characterisation of their biophysical and biochemical properties is required to give a more complete understanding of the diversity of functions carried out by the member of the mouse *serpina3* locus.

BIBLIOGRAPHY

Abraham, C. R. (2001). "Reactive astrocytes and alpha1-antichymotrypsin in Alzheimer's disease." Neurobiol Aging 22(6): 931-6.

Abts, H. F., Welss, T., Mirmohammadsadegh, A., Kohrer, K., Michel, G. and Ruzicka, T. (1999). "Cloning and characterization of hurpin (protease inhibitor 13): A new skin-specific, UV-repressible serine proteinase inhibitor of the ovalbumin serpin family." J Mol Biol 293(1): 29-39.

Aertgeerts, K., De Bondt, H. L., De Ranter, C. J. and Declerck, P. J. (1995). "Mechanisms contributing to the conformational and functional flexibility of plasminogen activator inhibitor-1." Nat Struct Biol 2(10): 891-7.

Annand, R. R., Dahlen, J. R., Sprecher, C. A., De Dreu, P., Foster, D. C., Mankovich, J. A., Talanian, R. V., Kiesel, W. and Giegel, D. A. (1999). "Caspase-1 (interleukin-1beta-converting enzyme) is inhibited by the human serpin analogue proteinase inhibitor 9." Biochem J 342 Pt 3: 655-65.

Aronsen, K. F., Ekelund, G., Kindmark, C. O. and Laurell, C. B. (1972). "Sequential changes of plasma proteins after surgical trauma." Scand. J. Clin. Lab. Invest 29: 127-136.

Askew, Y. S., Pak, S. C., Luke, C. J., Askew, D. J., Cataltepe, S., Mills, D. R., Kato, H., Lehoczy, J., Dewar, K., Birren, B. and Silverman, G. A. (2001). "SERPINB12 is a novel member of the human ov-serpin family that is widely expressed and inhibits trypsin-like serine proteinases." J Biol Chem 276(52): 49320-30.

Atchley, W. R., Lokot, T., Wollenberg, K., Dress, A. and Ragg, H. (2001). "Phylogenetic analyses of amino acid variation in the serpin proteins." Mol Biol Evol 18(8): 1502-11.

Baglin, T. P., Carrell, R. W., Church, F. C., Esmon, C. T. and Huntington, J. A. (2002). "Crystal structures of native and thrombin-complexed heparin cofactor II reveal a multistep allosteric mechanism." Proc. Natl. Acad. Sci. 99(17): 11079-11084.

Bank, U., Reinhold, D., Schneemilch, C., Kunz, D., Synowitz, H. J. and Ansorge, S. (1999). "Selective proteolytic cleavage of IL-2 receptor and IL-6 receptor ligand binding chains by neutrophil-derived serine proteases at foci of inflammation." J Interferon Cytokine Res 19(11): 1277-87.

Barbour, K. W., Goodwin, R. L., Guillonneau, F., Wang, Y., Baumann, H. and Berger, F. G. (2002). "Functional diversification during evolution of the murine alpha(1)-proteinase inhibitor family: Role of the hypervariable reactive center loop." Mol Biol Evol 19(5): 718-27.

Barbour, K. W., Wei, F., Brannan, C., Flotte, T. R., Baumann, H. and Berger, F. G. (2002). "The murine alpha₁-proteinase inhibitor gene family: Polymorphism, chromosomal location, and structure." Genomics 80(5): 515-522.

Bartelmez, S. H., Bradley, T. R., Bertoncello, I., Mochizuki, D. Y., Tushinski, R. J., Stanley, E. R., Hapel, A. J., Young, I. G., Kriegler, A. B. and Hodgson, G. S. (1989). "Interleukin 1 plus interleukin 3 plus colony-stimulating factor 1 are essential for clonal proliferation of primitive myeloid bone marrow cells." Exp Hem 17(3): 240-5.

Baumann, U., Huber, R., Bode, W., Grosse, D., Lesjak, M. and Laurell, C. B. (1991). "Crystal structure of cleaved human alpha₁-antichymotrypsin at 2.7 Å resolution and its comparison with other serpins." J. Mol. Biol. 218: 595-606.

Beatty, K., Bieth, J. and Travis, J. (1980). "Kinetics of association of serine proteinases with native and oxidized alpha-1-proteinase inhibitor and alpha-1-antichymotrypsin." J Biol Chem 255(9): 3931-4.

Becerra, S. P., Sagasti, A., Spinella, P. and Notario, V. (1995). "Pigment Epithelium-derived Factor Behaves Like a Noninhibitory Serpin." J Biol Chem 270(43): 25992-25999.

Belaouaj, A., McCarthy, R., Baumann, M., Gao, Z., Ley, T. J., Abraham, S. N. and Shapiro, S. D. (1998). "Mice lacking neutrophil elastase reveal impaired host defense against gram negative bacterial sepsis." Nat Med 4(5): 615-8.

Belin, D., Wohlwend, A., Schleuning, W. D., Kruithof, E. K. and Vassalli, J. D. (1989). "Facultative polypeptide translocation allows a single mRNA to encode the secreted and cytosolic forms of plasminogen activators inhibitor 2."

EMBO J 8(11): 3287-94.

Belorgey, D., Crowther, D. C., Mahadeva, R. and Lomas, D. A. (2002). "Mutant Neuroserpin (S49P) That Causes Familial Encephalopathy with Neuroserpin Inclusion Bodies Is a Poor Proteinase Inhibitor and Readily Forms Polymers in Vitro." J Biol Chem 277(19): 17367-17373.

Benarafa, C., Cooley, J., Zeng, W., Bird, P. I. and Remold-O'Donnell, E. (2002). "Characterization of four murine homologs of the human ov-serpin monocyte neutrophil elastase inhibitor MNEI (SERPINB1)." J Biol Chem. 277(44):42028-33

Bertoncello, I. and Kriegler, A. B. (1997). "High proliferative potential colony-forming cells." Methods Mol Biol 75: 265-72.

Bird, C. H., Sutton, V. R., Sun, J., Hirst, C. E., Novak, A., Kumar, S., Trapani, J. A. and Bird, P. I. (1998). "Selective regulation of apoptosis: the cytotoxic lymphocyte serpin proteinase inhibitor 9 protects against granzyme B-mediated apoptosis without perturbing the Fas cell death pathway." Mol Cell Biol 18(11): 6387-98.

Bode, W. and Huber, R. (2000). "Structural basis of the endoproteinase-protein inhibitor interaction." Biochimica et Biophysica Acta 1477(1-2): 241-52.

Bories, D., Raynal, M. C., Solomon, D. H., Darzynkiewicz, Z. and Cayre, Y. E. (1989). "Down-regulation of a serine protease, myeloblastin, causes growth arrest and differentiation of promyelocytic leukemia cells." Cell 59(6): 959-68.

Ørregaard, N. and Cowland, J. B. (1997). "Granules of the human neutrophilic polymorphonuclear leukocyte." Blood 89(10): 3503-21.

Borriello, F. and Krauter, K. S. (1991). "Multiple murine alpha 1-protease inhibitor genes show unusual evolutionary divergence." Proc Natl Acad Sci 88(21): 9417-21.

Bottomley, S. P. and Chang, W. (1997). "The effects of reactive centre loop length upon serpin polymerisation." Biochem Biophys Res Comm 241: 264-269.

Bottomley, S. P. and Stone, S. R. (1998). "Protein engineering of chimeric serpins: an investigation into effects of the serpin scaffold and reactive centre loop length." Protein Eng 11(12): 1243-7.

Bradley, T. R., Hodgson, G. S. and Rosendaal, M. (1978). "The effect of oxygen tension on haemopoietic and fibroblast cell proliferation in vitro." J Cell Physiol 97(3 Pt 2 Suppl 1): 517-22.

Bradley, T. R. and Hodgson, G. S. (1979). "Detection of primitive macrophage progenitor cells in mouse bone marrow." Blood 54(6): 1446-50.

Brem, R., Certa, U., Neeb, M., Nair, A. and Moroni, C. (2001). "Global analysis of differential gene expression after transformation with the v-H-ras oncogene in a murine tumour model." Oncogene 20: 2854-2858.

Bruce, D., Perry, D. J., Borg, J.-Y., Carrell, R. W. and Wardell, M. R. (1994). "Thromboembolic Disease Due to Thermolabile Conformational Changes of Antithrombin Rouen-VI (187 Asn-->Asp)" J Clin Invest 94(6): 2265-2274.

Calugaru, S. V., Swanson, R. and Olson, S. T. (2001). "The pH dependence of serpin-proteinase complex dissociation reveals a mechanism of complex stabilization involving inactive and active conformational states of the proteinase which are perturbable by calcium." J Biol Chem 276(35): 32446-55.

Carrell, R. W. and Owen, M. C. (1985). "Plakalbumin, α 1-antitrypsin, antithrombin and the mechanism of inflammatory thrombosis." Nature 317: 730-732.

Carrell, R. W. and Travis, J. (1985). "alpha(1)-antitrypsin and the serpins: variation and countervariation." Trends Biochem Sci 10: 20-24.

Carrell, R. W., Stein, P. E., Wardell, M. R. and Fermi, G. (1994). "Biological implications of a 3-angstrom structure of dimeric antithrombin." Structure 2: 257-270.

Chaillan-Huntington, C. E., Gettins, P., Huntington, J. A. and Patston, P. A. (1997). "The P6-P2 region of serpins is critical for proteinase inhibition and complex stability." FASEB J 11: 2149-2149.

Chang, W., Whisstock, J., Hopkins, P., Lesk, A. M., Carrell, R. W. and Wardell, M. R. (1997). "Importance of the release of strand 1C to the polymerization mechanism of inhibitory serpins." Protein Sci 6: 89-98.

Chapman, H. A., Riese, R. J. and Shi, G. P. (1997). "Emerging roles for cysteine proteases in human biology." Annual Review of Physiology 59: 63-88.

Chen, J. S. K. and Mehta, K. (1999). "Tissue transglutaminase: an enzyme with a split personality." Int J Biochem Cell Biol 31(8): 817-836.

Chiang, C. S., Stalder, A., Samimi, A. and Campbell, I. L. (1994). "Reactive gliosis as a consequence of interleukin-6 expression in the brain: studies in transgenic mice." Dev Neurosci 16(3-4): 212-21.

Chomczynski, P. and Sacchi, N. (1987). "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction." Anal Biochem 162: 156-159.

Chuang, T. L. and Schleef, R. R. (1999). "Identification of a nuclear targeting domain in the insertion between helices C and D in protease inhibitor-10." J Biol Chem 274(16): 11194-8.

Cichy, J., Potempa, J., Chawala, R. K. and Travis, J. (1995). "Stimulatory effect of inflammatory cytokines on alpha1-antichymotrypsin expression in human lung-derived epithelial cells." J Clin Invest 95: 2729-2733.

Confort, C., Rochefort, H. and Vignon, F. (1995). "Insulin-like growth factors (IGFs) stimulate the release of alpha 1-antichymotrypsin and soluble IGF-II/mannose 6-phosphate receptor from MCF7 breast cancer cells." Endocrinology 136(9): 3759-66.

Cooley, J., Mathieu, B., Remold-O'Donnell, E. and Mandle, R. J. (1998). "Production of recombinant human monocyte/neutrophil elastase inhibitor (rM/NEI)." Protein Expr Purif 14(1): 38-44.

Cooley, J., Takayama, T. K., Shapiro, S. D., Schechter, N. M. and Remold-O'Donnell, E. (2001). "The serpin MNEI inhibits elastase-like and chymotrypsin-like serine proteases through efficient reactions at two active sites." Biochemistry 40: 15762-15770.

Cooper, D. D. and Spangrude, G. J. (1999). "(R)Evolutionary Considerations in Hematopoietic Development." Ann NY Acad Sci **872**(1): 83-93.

Copeland, N. G., Jenkins, N. A. and Gilbert, D. J. *et al* (1993). A genetic linkage map of the mouse: current applications and future prospects Science **262**(5130):57-66.

Corpet, F. (1988). "Multiple sequence alignment with hierarchical clustering." Nucl Acids Res **16**(22): 10881-10890.

Coughlin, P., Sun, J. R., Cerruti, L., Salem, H. H. and Bird, P. (1993 (a)). "Cloning and molecular characterization of a human intracellular serine proteinase-inhibitor." Proc Natl Acad Sci **90**: 9417-21.

Coughlin, P. B., Tetaz, T. and Salem, H. H. (1993 (a)). "Identification and purification of a novel serine proteinase- inhibitor." J Biol Chem **268**: 9541-47.

Croall, D. E. and DeMartino, G. N. (1991). "Calcium-activated neutral protease (calpain) system: structure, function, and regulation." Physiol Rev **71**(3): 813-47.

Crowther, D. C., Serpell, L. C., Dafforn, T. R., Gooptu, B. and Lomas, D. A. (2003). "Nucleation of alpha 1-antichymotrypsin polymerization." Biochemistry **42**(8): 2355-63.

Cunningham, D. D., Wagner, S. L. and Farrell, D. H. (1992). "Regulation of protease nexin-1 activity by heparin and heparan sulfate." Adv Exp Med Biol **313**: 297-306.

Dafforn, T. R., Mahadeva, R., Elliott, P. R., Sivasothy, P. and Lomas, D. A. (1999). "A kinetic mechanism for the polymerization of alpha1-antitrypsin." J Biol Chem **274**(14): 9548-55.

Dafforn, T. R., Della, M. and Miller, A. D. (2001). "The Molecular Interactions of Heat Shock Protein 47 (Hsp47) and Their Implications for Collagen Biosynthesis." J Biol Chem **276**(52): 49310-49319.

Dahlen, J. R., Foster, D. C. and Kisiel, W. (1997). "Human proteinase inhibitor 9 (P19) is a potent inhibitor of subtilisin A." Biochem Biophys Res Comm **238**: 329-333.

Dahlen, J. R., Jean, F., Thomas, G., Foster, D. C. and Kisiel, W. (1998 (a)). "Inhibition of soluble recombinant furin by human proteinase inhibitor 8." J Biol Chem 273: 1851-1854.

Dahlen, J. R., Foster, D. C. and Kisiel, W. (1998 (b)). "The inhibitory specificity of human proteinase inhibitor 8 is expanded through the use of multiple reactive site residues." Biochem Biophys Res Com 244: 172-177.

Dahlen, J. R., Foster, D. C. and Kisiel, W. (1999). "Inhibition of neutrophil elastase by recombinant human proteinase inhibitor 9." Biochim Biophys Acta 1451(2-3): 233-41.

Dale, D. C., Person, R. E., Bolyard, A. A., Aprikyan, A. G., Bos, C., Bonilla, M. A., Boxer, L. A., Kannourakis, G., Zeidler, C., Welte, K., Benson, K. F. and Horwitz, M. (2000). "Mutations in the gene encoding neutrophil elastase in congenital and cyclic neutropenia." Blood 96(7): 2317-22.

Darnell, G. A., Antalis, T. M., Johnstone, R. W., Stringer, B. W., Ogbourne, S. M., Harrich, D. and Suhrbier, A. (2003). "Inhibition of retinoblastoma protein degradation by interaction with the serpin plasminogen activator inhibitor 2 via a novel consensus motif." Mol Cell Biol 23(18): 6520-32.

Davis, R. L., Holohan, P. D., Shrimpton, A. E., Tatum, A. H., Daucher, J., Collins, G. H., Todd, R., Bradshaw, C., Kent, P., Feiglin, D., Rosenbaum, A., Yerby, M. S., Shaw, C. M., Lacbawan, F. and Lawrence, D. A. (1999). "Familial encephalopathy with neuroserpin inclusion bodies." Am J Pathol 155(6): 1901-13.

Derechin, V. M., Blinder, M. A. and Tollefsen, D. M. (1990). "Substitution of arginine for Leu444 in the reactive site of heparin cofactor II enhances the rate of thrombin inhibition." J Biol Chem 265(10): 5623-8.

Deussing, J., Kouadio, M., Rehman, S., Werber, I., Schwinde, A. and Peters, C. (2002). "Identification and characterisation of a dense cluster of placenta-specific cysteine peptidase genes and related genes on mouse chromosome 13." Genomics 79(2): 225-240.

Dichtl, W., Moraga, F., Ares, M. P. S., Crisby, M., Nilsson, J., Lindgren, S. and Janciauskiene, S. (2000). "The Carboxyl-Terminal Fragment of [alpha]1-Antitrypsin Is Present in Atherosclerotic Plaques and Regulates Inflammatory Transcription Factors in Primary Human Monocytes*1." Mol Cell Biol Res Comm 4(1): 50-61.

Dickinson, J. L., Bates, E. J., Ferrante, A. and Antalis, T. M. (1995). "Plasminogen activator inhibitor type 2 inhibits tumor necrosis factor alpha-induced apoptosis. Evidence for an alternate biological function." J Biol Chem 270(46): 27894-904.

Dickinson, J. L., Norris, B. J., Jensen, P. H. and Antalis, T. M. (1998). "The C-D interhelical domain of the serpin plasminogen activator inhibitor-type 2 is required for protection from TNF-alpha induced apoptosis." Cell Death Differ 5(2): 163-71.

Djie, M. Z., Le Bonniec, B. F., Hopkins, P. C., Hipler, K. and Stone, S. R. (1996). "Role of the P2 residue in determining the specificity of serpins." Biochemistry 35(35): 11461-9.

Djie, M. Z., Stone, S. R. and LeBonniec, B. F. (1997). "Intrinsic specificity of the reactive site loop of alpha(1)- antitrypsin, alpha(1)-antichymotrypsin, antithrombin III, and protease nexin I." J Biol Chem 272: 16268-16273.

Drouet, C., Bouillet, L., Csopaki, F. and Colomb, M. G. (1999). "Hepatitis C virus NS3 serine protease interacts with the serpin C1 inhibitor." FEBS Lett 458(3): 415-8.

Du, X., Saido, T. C., Tsubuki, S., Indig, F. E., Williams, M. J. and Ginsberg, M. H. (1995). "Calpain cleavage of the cytoplasmic domain of the integrin beta 3 subunit." J Biol Chem 270(44): 26146-51.

Dunstone, M. A., Dai, W., Whisstock, J. C., Rossjohn, J., Pike, R. N., Feil, S. C., Le Bonniec, B. F., Parker, M. W. and Bottomley, S. P. (2000). "Cleaved antitrypsin polymers at atomic resolution." Protein Sci 9(2): 417-20.

Duranton, J., Adam, C. and Bieth, J. G. (1998). "Kinetic mechanism of the inhibition of cathepsin G by alpha 1-antichymotrypsin and alpha 1-proteinase inhibitor." Biochemistry 37(32): 11239-45.

Elliott, P. R., Abrahams, J. P. and Lomas, D. A. (1998 (a)). "Wild-type alpha(1)-antitrypsin is in the canonical inhibitory conformation." J Mol Biol 275: 419-425.

Elliott, P. R., Bilton, D. and Lomas, D. A. (1998 (b)). "Lung polymers in Z alpha(1)-antitrypsin deficiency-related emphysema." Am Resp Cell And Mol Biol 18: 670-674.

Eriksson, S., Carlson, J. and Velez, R. (1986). "Risk of cirrhosis and primary liver cancer in alpha 1-antitrypsin deficiency." N. Engl. J. Med. 314(12): 736-9.

Espana, F. and Griffin, J. H. (1989). "Determination of functional and antigenic protein C inhibitor and its complexes with activated protein C in plasma by ELISA's." Thromb Res 55(6): 671-82.

Faber, J. P., Poller, W., Olek, K., Baumann, U., Carlson, J., Lindmark, B. and Eriksson, S. (1993). "The molecular basis of alpha 1-antichymotrypsin deficiency in a heterozygote with liver and lung disease." J Hepatol 18(3): 313-21.

Flink, i. L., Bailey, T. J., Gustafson, T. A., Markham, B. E. and Morkin, E. (1986). "Complete amino acid sequence of human thyroxine-binding globulin deduced from cloned DNA: close homology to the serine antiproteases." Proc Natl Acad Sci 83(20): 7708-12.

Forsyth, S., Horvath, A. and Coughlin, P. (2003). "A review and comparison of the murine alpha(1)-antitrypsin and alpha(1)-antichymotrypsin multigene clusters with the human clade A serpins." Genomics 81(3): 336-45.

Frank, P. S., Douglas, J. T., Locher, M., Llinas, M. and Schaller, J. (2003). "Structural/functional characterization of the alpha 2-plasmin inhibitor C-terminal peptide." Biochemistry 42(4): 1078-85.

Frazer, J. K., Jackson, D. G., Gaillard, J. P., Lutter, M., Liu, Y. J., Banchereau, J., Capra, J. D. and Pascua, V. (2000). "Identification of centerin: a novel human germinal center B cell-restricted serpin" Eur J Immunol 30(10): 3039-48.

Geboes, K., Ray, M. B., Rutgeerts, P., Callea, F., Desmet, V. J. and Vantrappen, G. (1982). "Morphological identification of alpha-1-antitrypsin in the human small intestine." Histopathology 6(1): 55-60.

Gettins, P. G. (2002 (a)). "Serpin structure, mechanism, and function." Chem Rev 102(12): 4751-804.

Gettins, P. G. (2002 (b)). "The F-helix of serpins plays an essential, active role in the proteinase inhibition mechanism." FEBS Letters 523(1-3): 2-6.

Goldenberg, D. P. (1989 (a)). Non-denaturing Gel Electrophoresis of Proteins. Protein Structure a Practical Approach. T. E. Creighton, IRL Press: 228-238.

Goldenberg, D. P. (1989 (b)). Transverse Gradient Gel Electrophoresis. Protein Structure a Practical Approach. T. E. Creighton, IRL Press: 239-250.

Goodwin, R. L., Barbour, K. W. and Berger, F. G. (1997). "Expression of the alpha 1-proteinase inhibitor gene family during evolution of the genus Mus." Mol Biol Evol 14(4): 420-7.

Gooptu, B., Hazes, B., Chang, W. S., Dafforn, T. R., Carrell, R. W., Read, R. J. and Lomas, D. A. (2000). "Inactive conformation of the serpin alpha(1)-antichymotrypsin indicates two-stage insertion of the reactive loop: implications for inhibitory function and conformational disease." Proc Natl Acad Sci 97(1): 67-72.

Goselink, H., van Damme, J., Hiemstra, P., Wuyts, A., Stolk, J., Fibbe, W., Willemze, R. and Falkenburg, J. (1996). "Colony growth of human hematopoietic progenitor cells in the absence of serum is supported by a proteinase inhibitor identified as antileukoproteinase." J. Exp. Med. 184(4): 1305-1312.

Grigoryev, S. A. and Woodcock, C. L. (1998). "Chromatin structure in granulocytes. A link between tight compaction and accumulation of a heterochromatin-associated protein (MENT)." J Biol Chem 273(5): 3082-9.

Grigoryev, S. A., Bednar, J. and Woodcock, C. L. (1999). "MENT, a heterochromatin protein that mediates higher order chromatin folding, is a new serpin family member." J Biol Chem 274(9): 5626-36.

Hamerman, J. A., Hayashi, F., Schroeder, L. A., Gygi, S. P., Haas, A. L., Hampson, L., Coughlin, P., Aebersold, R. and Aderem, A. (2002). "Serpin 2a is induced in activated macrophages and conjugates to a ubiquitin homolog." J Immunol **168**(5): 2415-23.

Hammond, G. L., Smith, C. L., Goping, I. S., Underhill, D. A., Harley, M. J., Reventos, J., Musto, N. A., Gunsalus, G. L. and Bardin, C. W. (1987). "Primary structure of human corticosteroid binding globulin, deduced from hepatic and pulmonary cDNAs, exhibits homology with serine protease inhibitors." Proc Natl Acad Sci **84**(15): 5153-7.

Hampson, I. N., Hampson, L., Pinkoski, M., Cross, M., Heyworth, C. M., Bleackley, R. C., Atkinson, E. and Dexter, T. M. (1997). "Identification of a serpin specifically expressed in multipotent and bipotent hematopoietic progenitor cells and in activated T cells." Blood **89**: 108-118.

Hampson, L., Hampson, I. N., Babichuk, C. K., Cotter, L., Bleackley, R. C., Dexter, T. M. and Cross, M. (2001). "A minimal serpin promoter with high activity in haemopoietic progenitors and activated T cells." Hematol J **2**: 150-160.

Harrop, S. J., Jankova, L., Coles, M., Jardine, D., Whittaker, J. S., Gould, A. R., Meister, A., King, G. C., Mabbutt, B. C. and Curmi, P. M. (1999). "The crystal structure of plasminogen activator inhibitor 2 at 2.0 Å resolution: implications for serpin function." Structure Fold Des **7**(1): 43-54.

Hastings, G. A., Coleman, T. A., Haudenschield, C. C., Stefansson, S., Smith, E. P., Barthlow, R., Cherry, S., Sandkvist, M. and Lawrence, D. A. (1997). "Neuroserpin, a brain-associated inhibitor of tissue plasminogen activator is localized primarily in neurons - Implications for the regulation of motor learning and neuronal survival." J Biol Chem **272**: 33062-33067.

Hedstrom, L. (2002). "Serine protease mechanism and specificity." Chem Rev **102**(12): 4501-24.

Hill, R. E., Shaw, P. H., Boyd, P. A., Baumann, H. and Hastie, N. D. (1984). "Plasma protease inhibitors in mouse and man: divergence within the reactive centre regions." Nature **311**(5982): 175-7.

Hill, R. E., Shaw, P. H., Barth, R. K. and Hastie, N. D. (1985). "A genetic locus closely linked to a protease inhibitor gene complex controls the level of multiple RNA transcripts." Mol Cell Biol 5(8): 2114-22.

Hill, R. E. and Hastie, N. D. (1987). "Accelerated evolution in the reactive centre regions of serine protease inhibitors." Nature 326(6108): 96-9.

Hirayoshi, K., Kudo, H., Takechi, H., Nakai, A., Iwamatsu, A., Yamada, K. M. and Nagata, K. (1991). "HSP47: a tissue-specific, transformation-sensitive, collagen-binding heat shock protein of chicken embryo fibroblasts." Mol Cell Biol 11(8): 4036-44.

Holmes, W. E., Lijnen, H. R., Nelles, L., Kluft, C., Nieuwenhuis, H. K., Rijken, D. C. and Collen, D. (1987). "Alpha 2-antiplasmin Enschede: alanine insertion and abolition of plasmin inhibitory activity." Science 238(4824): 209-11.

Hopkins, P., Carrell, R. W. and Stone, S. R. (1993). "Effects of mutations in the hinge region of serpins." Biochemistry 32: 7650-7657.

Horwitz, M., Benson, K. F., Person, R. E., Aprikyan, A. G. and Dale, D. C. (1999). "Mutations in ELA2, encoding neutrophil elastase, define a 21-day biological clock in cyclic haemopoiesis." Nat Genet 23(4): 433-6.

Huber, R. and Carrell, R. W. (1989). "Implications of the three-dimensional structure of α 1-antitrypsin for structure and function of serpins." Biochemistry 28(23): 8951-8966.

Huntington, J. A. (2003). "Mechanisms of glycosaminoglycan activation of the serpins in hemostasis." J Thromb Haemost 1(7): 1535-49.

Huntington, J. A., Fan, B. Q., Karlsson, K. E., Deinum, J., Lawrence, D. A. and Gettins, P. (1997). "Serpine conformational change in ovalbumin. Enhanced reactive center loop insertion through hinge region mutations." Biochemistry 36: 5432-5440.

Huntington, J. A., Pannu, N. S., Hazes, B., Read, R. J., Lomas, D. A. and Carrell, R. W. (1999). "A 2.6 Å structure of a serpin polymer and implications for conformational disease." J Mol Biol 293(3): 449-55.

Huntington, J. A., Read, R. J. and Carrell, R. W. (2000). "Structure of a serpin-protease complex shows inhibition by deformation." Nature 407(6806): 923-6.

Inglis, J. D. and Hill, R. E. (1991). "The murine Spi-2 proteinase inhibitor locus: a multigene family with a hypervariable reactive site domain." EMBO J 10(2): 255-261.

Inglis, J. D., Lee, M., Davidson, D. R. and Hill, R. E. (1991). "Isolation of two cDNAs encoding novel antichymotrypsin-like proteins in a murine chondrocytic cell line." Gene 106: 213-220.

Irving, J. A., Pike, R. N., Lesk, A. M. and Whisstock, J. C. (2000). "Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function." Genome Res 10(12): 1845-64.

Irving, J. A., Steenbakkens, P. J. M., Lesk, A. M., Op den Camp, H. J. M., Pike, R. N. and Whisstock, J. C. (2002 (a)). "Serpins in prokaryotes." Mol Biol Evol 19(11):1881-1890.

Irving, J. A., Shushanov, S. S., Pike, R. N., Popova, E. Y., Bromme, D., Coetzer, T. H., Bottomley, S. P., Boulyenko, I. A., Grigoryev, S. A. and Whisstock, J. C. (2002 (b)). "Inhibitory activity of a heterochromatin-associated serpin (MENT) against papain-like cysteine proteinases affects chromatin structure and blocks cell proliferation." J Biol Chem 277(15): 13192-201.

Irving, J. A., Pike, R. N., Dai, W., Bromme, D., Worrall, D. M., Silverman, G. A., Coetzer, T. H., Dennison, C., Bottomley, S. P. and Whisstock, J. C. (2002 (c)). "Evidence that serpin architecture intrinsically supports papain-like cysteine protease inhibition: engineering alpha(1)-antitrypsin to inhibit cathepsin proteases." Biochemistry 41(15): 4998-5004.

Irving, J. A., Cabrita, L. D., Rossjohn, J., Pike, R. N., Bottomley, S. P. and Whisstock, J. C. (2003). "The 1.5 Å Crystal Structure of a Prokaryote Serpin: Controlling Conformational Change in a Heated Environment." Structure 11(4): 387-397.

Isa, M. N., Boyd, E., Morrison, N., Harrap, S., Clauser, E. and Connor, J. M. (1990). "Assignment of the human angiotensinogen gene to chromosome 1q42-q43 by nonisotopic in situ hybridization" Genomics 8(3): 598-600.

Jallat, S., Tessier, L. H., Benavente, A., Crystal, R. G. and Courtney, M. (1986). "Antiprotease targeting: altered specificity of alpha 1-antitrypsin by amino acid replacement at the reactive cent.e." Rev Fran de Transfus Immunohematol 29(4): 287-98.

Janciauskiene, S. and Lindgren, S. (1999). "Human monocyte activation by cleaved form of alpha-1-antitrypsin. Involvement of the phagocytic pathway." Eur J Biochem 265(3): 875-882.

Jayakumar, A., Kang, Y., Frederick, M. J., Pak, S. C., Henderson, Y., Holton, P. R., Mitsudo, K., Silverman, G. A., AK, E. L.-N., Bromme, D. and Clayman, G. L. (2003). "Inhibition of the cysteine proteinases cathepsins K and L by the serpin headpin (SERPINB13): a kinetic analysis." Arch Biochem Biophys 409(2): 367-74.

Jensen, P. H., Schuler, E., Woodrow, G., Richardson, M., Goss, N., Hojrup, P., Petersen, T. E. and Rasmussen, L. K. (1994). "A unique interhelical insertion in plasminogen activator inhibitor-2 contains three glutamines, Gln83, Gln84, Gln86, essential for transglutaminase-mediated cross-linking." J Biol Chem 269(21): 15394-8.

Jiang, H. B. and Kanost, M. R. (1997). "Characterization and functional analysis of 12 naturally occurring reactive site variants of serpin-1 from *Manduca sexta*." J Biol Chem 272: 1082-1087.

Jin, L., Abrahams, J. P., Skinner, R., Petitou, M., Pike, R. N. and Carrell, R. W. (1997). "The anticoagulant activation of antithrombin by heparin." Proc Natl Acad Sci 94: 14683-14688.

Johnson, D. and Travis, J. (1978). "Structural evidence for methionine at the reactive site of human alpha-1-proteinase inhibitor." J Biol Chem 253(20): 7142-4.

Kaiserman, D., Knaggs, S., Scarff, K. L., Gillard, A., Mirza, G., Cadman, M., McKeone, R., Denny, P., Cooley, J., Benarafa, C., Remold-O'Donnell, E., Ragoussis, J. and Bird, P. I. (2002). "Comparison of human chromosome 6p25 with mouse chromosome 13 reveals a greatly expanded ov-serpin gene repertoire in the mouse." Genomics 79(3): 349-62.

Kaslik, G., Patthy, A., Balint, M. and Graf, L. (1995). "Trypsin complexed with [alpha]1-proteinase inhibitor has an increased structural flexibility." FEBS Letters 370(3): 179-183.

Keijer, J., Linders, M., Wegman, J. J., Ehrlich, H. J., Mertens, K. and Pannekoek, H. (1991). "On the target specificity of plasminogen activator inhibitor 1: the role of heparin, vitronectin, and the reactive site." Blood 78(5): 1254-61.

Komiyama, T., Gron, H., Pemberton, P. A. and Salvesen, G. S. (1996). "Interaction of subtilisins with serpins." Protein Sci 5(5): 874-882.

Korbling, M. and Anderlini, P. (2001). "Peripheral blood stem cell versus bone marrow allotransplantation: does the source of hematopoietic stem cells matter?" Blood 98(10): 2900-8.

Kordula, T., Rydel, R. E., Brigham, E. F., Horn, F., Heinrich, P. C. and Travis, J. (1998). "Oncostatin M and the Interleukin-6 and Soluble Interleukin-6 Receptor Complex Regulate alpha 1-Antichymotrypsin Expression in Human Cortical Astrocytes." J Biol Chem 273(7): 4112-4118.

Krebs, M., Uhrin, P., Vales, A., Prendes-Garcia, M. J., Wojta, J., Geiger, M. and Binder, B. R. (1999). "Protein C inhibitor is expressed in keratinocytes of human skin." J Invest Dermatol 113(1): 32-7.

Kroczyńska, B., Evangelista, C. M., Samant, S. S., Elguindi, E. C. and Blond, S. Y. (2003). "The SANT2 domain of murine tumor cell DnaJ-like protein 1 human homologue interacts with alpha 1-antichymotrypsin and kinetically interferes with its serpin inhibitory activity." J. Biol. Chem 279(12):11432-43.

Kroll, J. and Briand, P. (1988). "Estrogen-dependent release of serum proteins from MCF-7 breast carcinoma cells in vitro." Anticancer Res 8(1): 89-91.

Kruger, O., Ladewig, J., Koster, K. and Ragg, H. (2002). "Widespread occurrence of serpin genes with multiple reactive centre-containing exon cassettes in insects and nematodes." Gene 293(1-2): 97-105.

Krugliak, L., Meyer, P. R. and Taylor, C. R. (1986). "The distribution of lysozyme, alpha-1-antitrypsin, and alpha-1-antichymotrypsin in normal hematopoietic cells and in myeloid leukemias: an immunoperoxidase study on cytocentrifuge preparations, smears, and paraffin sections." Am J Hematol 21(1): 99-109.

Kuhn, L. A., Griffin, J. H., Fisher, C. L., Greengard, J. S., Bouma, B. N., Espana, F. and Tainer, J. A. (1990). "Elucidating the structural chemistry of glycosaminoglycan recognition by protein C inhibitor." Proc Natl Acad Sci 87(21): 8506-10.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature 227: 680-685.

Laursen, I. and Lykkesfeldt, A. E. (1992). "Purification and characterization of an alpha 1-antichymotrypsin-like 66 kDa protein from the human breast cancer cell line, MCF-7." Biochim Biophys Acta 1121(1-2): 119-29.

Lawrence, D. A., Berkenpas, M. B., Palaniappan, S. and Ginsburg, D. (1994 (a)). "Localization of vitronectin binding domain in plasminogen activator inhibitor-1." J Biol Chem 269(21): 15223-8.

Lawrence, D. A., Olson, S. T., Palaniappan, S. and Ginsburg, D. (1994 (b)). "Engineering plasminogen activator inhibitor 1 mutants with increased functional stability." Biochemistry 33(12): 3643-8.

Lawrence, D. A., Olson, S. T., Palaniappan, S. and Ginsburg, D. (1994 (c)). "Serpine reactive center loop mobility is required for inhibitor function but not for enzyme recognition." J Biol Chem 269(44): 27657-62.

Lawrence, D. A., Ginsburg, D., Day, D. E., Berkenpas, M. B., Verhamme, I. M., Kvassman, J. O. and Shore, J. D. (1995). "Serpine-protease complexes are trapped as stable acyl-enzyme intermediates." J Biol Chem 270(43): 25309-12.

Lawrence, D. A., Palaniappan, S., Stefansson, S., Olson, S. T., Francis-Chmura, A. M., Shore, J. D. and Ginsburg, D. (1997). "Characterization of the binding of different conformational forms of plasminogen activator inhibitor-1 to vitronectin. Implications for the regulation of pericellular proteolysis." J Biol Chem 272(12): 7676-80.

Lawrence, D. A., Olson, S. T., Muhammad, S., Day, D. E., Kvassman, J. O., Ginsburg, D. and Shore, J. D. (2000). "Partitioning of serpin-proteinase reactions between stable inhibition and substrate cleavage is regulated by the rate of serpin reactive center loop insertion into beta-sheet A." J Biol Chem **275**(8): 5839-44.

Le Bonniec, B., MacGillivray, R. and Esmon, C. (1991). "Thrombin Glu-39 restricts the P'3 specificity to nonacidic residues." J Biol Chem **266**(21): 13796-13803.

Lee, C., Piazza, F., Brutsaert, S., Valens, J., Strehlow, I., Jarosinski, M., Saris, C. and Schindler, C. (1999). "Characterization of the Stat5 protease." J Biol Chem **274**(38): 26767-75.

Leimig, T., Mann, L., Martin, M. d. P., Bonten, E., Persons, D., Knowles, J., Allay, J. A., Cunningham, J., Nienhuis, A. W., Smeyne, R. and d'Azzo, A. (2002). "Functional amelioration of murine galactosialidosis by genetically modified bone marrow hematopoietic progenitor cells." Blood **99**(9): 3169-3178.

Levesque, J. P., Takamatsu, Y., Nilsson, S. K., Haylock, D. N. and Simmons, P. J. (2001). "Vascular cell adhesion molecule-1 (CD106) is cleaved by neutrophil proteases in the bone marrow following hematopoietic progenitor cell mobilization by granulocyte colony-stimulating factor." Blood **98**(5): 1289-97.

Levesque, J. P., Hendy, J., Takamatsu, Y., Williams, B., Winkler, I. G. and Simmons, P. J. (2002). "Mobilization by either cyclophosphamide or granulocyte colony-stimulating factor transforms the bone marrow into a highly proteolytic environment." Exp Hematol **30**(5): 440-9.

Levesque, J. P., Hendy, J., Takamatsu, Y., Simmons, P. J. and Bendall, L. J. (2003). "Disruption of the CXCR4/CXCL12 chemotactic interaction during hematopoietic stem cell mobilization induced by GCSF or cyclophosphamide." J Clin Invest **111**(2): 187-96.

Licastro, F., Campbell, I. L., Kincaid, C., Veinbergs, I., Van Uden, E., Rockenstein, E., Mallory, M., Gilbert, J. R. and Masliah, E. (1999). "A role for apoE in regulating the levels of alpha-1-antichymotrypsin in the aging mouse brain and in Alzheimer's disease." Am J Pathol **155**(3): 869-75.

Lijnen, H. R. and Collen, D. (1985). "Protease inhibitors of human plasma. Alpha-2-antiplasmin." J Med 16(1-3): 225-84.

Liu, N., Raja, S. M., Zazzeroni, F., Metkar, S. S., Shah, R., Zhang, M., Wang, Y., Bromme, D., Russin, W. A., Lee, J. C., Peter, M. E., Froelich, C. J., Franzoso, G. and Ashton-Rickardt, P. G. (2003). "NF-kappaB protects from the lysosomal pathway of cell death." EMBO J 22(19): 5313-22.

Loebermann, H., Tokuoka, R., Deisenhofer, J. and Huber, R. (1984). "Human α_1 -proteinase inhibitor." J. Mol. Biol. 177: 531-556.

Logan, A. C., Lutzko, C. and Kohn, D. B. (2002). "Advances in lentiviral vector design for gene-modification of hematopoietic stem cells." Curr Opin Biotech 13(5): 429-436.

Lomas, D. A., Evans, D. L., Finch, J. T. and Carrell, R. W. (1992). "The mechanism of α_1 -antitrypsin accumulation in the liver." Nature 357: 605-607.

Lomas, D. A., Elliott, P. R., Chang, W., Wardell, M. R. and Carrell, R. W. (1995 (a)). "Preparation and characterization of latent α_1 -antitrypsin." J Biol Chem 270: 5282-5288.

Lomas, D. A., Stone, S. R., Llewellyn-Jones, C., Keogan, M.-T., Wang, Z.-m., Rubin, H., Carrell, R. W. and Stockley, R. A. (1995 (b)). "The Control of Neutrophil Chemotaxis by Inhibitors of Cathepsin G and Chymotrypsin." J Biol Chem 270(40): 23437-23443.

Lomas, D. A. and Carrell, R. W. (2002). "Serpinoopathies and the conformational dementias." Nat Rev Genet 3(10): 759-68.

Lorenz, E., Uphoff, D., Reid, T. R. and Shelton, E. (1951). "Modification of irradiation injury in mice and guinea pigs by bone marrow injections." J Natl Cancer Inst 12(1): 197-201.

Luskey, B. D., Rosenblatt, M., Zsebo, K. and Williams, D. A. (1992). "Stem cell factor, interleukin-3, and interleukin-6 promote retroviral-mediated gene transfer into murine hematopoietic stem cells." Blood 80(2): 396-402.

- Lutz, P. G., Moog-Lutz, C., Coumou-Gatbois, E., Kobari, L., Di Gioia, Y. and Cayre, Y. E. (2000). "Myeloblastin is a granulocyte colony-stimulating factor-responsive gene conferring factor-independent growth to hematopoietic cells." Proc Natl Acad Sci **97**(4): 1601-6.
- Ma, J., Yee, A., Brewer, H. B. J., Das, S. and Potter, H. (1994). "Amyloid-associated proteins alpha-1 antichymotrypsin and apolipoprotein E promote assembly of Alzheimer beta-protein into filaments." Nature **372**: 92-94.
- MacIvor, D., Shapiro, S., Pham, C., Belaaouaj, A., Abraham, S. and Ley, T. (1999). "Normal neutrophil function in cathepsin-G deficient mice." Blood **94**(12): 4282-4293.
- Marsden, M. D. and Fournier, R. E. K. (2003). "Chromosomal Elements Regulate Gene Activity and Chromatin Structure of the Human Serpin Gene Cluster at 14q32.1." Mol. Cell. Biol. **23**(10): 3516-3526.
- Mast, A. E., Enghild, J. J., Pizzo, S. V. and Salvesen, G. (1991). "Analysis of the plasma elimination kinetics and conformational stabilities of native, proteinase-complexed, and reactive site cleaved serpins: comparison of a1-proteinase inhibitor, a1-antichymotrypsin, antithrombin III, a2-antiplasmin, angiotensinogen, and ovalbumin." Biochemistry **30**: 1723-1730.
- Mast, A. E., Enghild, J. J. and Salvesen, G. (1992). "Conformation of the reactive site loop of a1-proteinase inhibitor probed limited proteolysis." Biochem. **31**: 2720-2728.
- McCarthy, B. J. and Worrall, D. M. (1997). "Analysis of serpin inhibitory function by mutagenesis of ovalbumin and generation of chimeric ovalbumin PAI-2 fusion proteins." J Mol Biol **267**: 561-569.
- McCoy, A. J., Pei, X. Y., Skinner, R., Abrahams, J.-P. and Carrell, R. W. (2003). "Structure of [beta]-Antithrombin and the Effect of Glycosylation on Antithrombin's Heparin Affinity and Activity." J Mol Biol **326**(3): 823-833.
- Meyer, J., Jucker, M., Ostertag, W. and Stocking, C. (1998). "Carboxyl-truncated STAT5beta is generated by a nucleus-associated serine protease in early hematopoietic progenitors." Blood **91**(6): 1901-8.

Mikolajczyk, S. D., Marks, L. S., Partin, A. W. and Rittenhouse, H. G. (2002). "Free prostate-specific antigen in serum is becoming more complex." Urology 59(6): 797-802.

Missen, M. A., Haylock, D. N., Simmons, P. J. and Coughlin, P. B. (2004). "Serpin expression in human haemopoietic and peripheral blood cells." (manuscript in preparation).

Moore, M. A. (2002). "Cytokine and chemokine networks influencing stem cell proliferation, differentiation, and marrow homing." J Cell Biochem (Sup) 38: 29-38.

Moore, M. A. S. (1996). "Overview of hemoiesis and hemopoietic reconstitution." Cell Therapy: Stem Cell Transplantation, Gene Therapy and Cellular Immunotherapy. Cambridge University Press New York: 3-17.

Mori, Y., Miura, Y., Oiso, Y., Hisao, S. and Takazumi, K. (1995). "Precise localization of the human thyroxine-binding globulin gene to chromosome Xq22.2 by fluorescence in situ hybridization." Human Genetics 96(4): 481-2.

Morris, E. C., Dafforn, T. R., Forsyth, S. L., Missen, M. A., Horvath, A. J., Hampson, L., Hampson, I. N., Currie, G., Carrell, R. W. and Coughlin, P. B. (2003). "Murine serpin 2A is a redox-sensitive intracellular protein." Biochem J 371(Pt 1): 165-73.

Morse, H. C., 3rd, Shen, F. W. and Hammerling, U. (1987). "Genetic nomenclature for loci controlling mouse lymphocyte antigens." Immunogenetics 25(2): 71-8.

Mottonen, J., Strand, A., Symersky, J., Sweet, R. M., Danley, D. E., Geoghegan, K. F., Gerard, R. D. and Goldsmith, E. J. (1992). "Structural basis of latency in plasminogen activator inhibitor-1." Nature 355: 270-273.

Mourey, L., Samama, J. P., Delarue, M., Petitou, M., Choay, J. and Moras, D. (1993). "Crystal structure of cleaved bovine antithrombin III at 3.2 Å resolution." J Mol Biol 232(1): 223-41.

Mucke, L., Yu, G. Q., McConlogue, L., Rockenstein, E. M., Abraham, C. R. and Masliah, E. (2000). "Astroglial expression of human alpha(1)-antichymotrypsin enhances alzheimer-like pathology in amyloid protein precursor transgenic mice." Am J Pathol 157(6): 2003-10.

Naidoo, N., Cooperman, B. S., Wang, Z. M., Liu, X. Z. and Rubin, H. (1995). "Identification of lysines within alpha 1-antichymotrypsin important for DNA binding. An unusual combination of DNA-binding elements." J Biol Chem 270(24): 14548-55.

Neese, L. L., Wolfe, C. A. and Church, F. C. (1998). "Contribution of basic residues of the D and H helices in heparin binding to protein C inhibitor." Arch Biochem Biophys 355(1): 101-8.

Nilsson, L. N. G., Das, S. and Potter, H. (2001). "Effect of cytokines, dexamethazone and the A/T-signal peptide polymorphism on the expression of alpha1-antichymotrypsin in astrocytes: significance for Alzheimer's disease." Neurochem Int 39(5-6): 361-370.

Ogawa, M. (1993). "Differentiation and proliferation of hematopoietic stem cells." Blood 81(11): 2844-53.

Ogilvy, S., Metcalf, D., Gibson, L., Bath, M. L., Harris, A. W. and Adams, J. M. (1999). "Promoter elements of vav drive transgene expression in vivo throughout the hematopoietic compartment." Blood 94(6): 1855-63.

Olson, S. T. and Björk, I. (1991). "Predominant contribution of surface approximation to the mechanism of heparin acceleration of the antithrombin-thrombin reaction. Elucidation from salt concentration effects." J Biol Chem 266(10): 6353-6364.

Olson, S. T. and Björk, I. (1992). Regulation of thrombin by antithrombin and heparin cofactor II. Thrombin: Structure and function. L. J. Berliner. New York, Plenum Press: 159-217.

Olson, S. T., Björk, I., Sheffer, R., Craig, P. A., Shore, J. D. and Choay, J. (1992). "Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. Resolution of the antithrombin conformational change contribution to heparin rate enhancement." J Biol Chem 267(18): 12528-12538.

Owen, C., Campbell, M., Sannes, P., Boukedes, S. and Campbell, E. (1995). "Cell surface-bound elastase and cathepsin G on human neutrophils: a novel, non-oxidative mechanism by which neutrophils focus and preserve catalytic activity of serine proteinases." J. Cell Biol. **131**(3): 775-789.

Owen, M. C., Brennan, S. O., Lewis, J. H. and Carrell, R. W. (1983). "Mutation of antitrypsin to antithrombin. alpha 1-antitrypsin Pittsburgh (358 Met leads to Arg), a fatal bleeding disorder." N. Engl. J. Med **309**(12): 694-8.

Pak, S. C., Kumar, V., Tsu, C., Luke, C. J., Askew, Y. S., Askew, D. J., Mills, D. R., Bromme, D. and Silverman, G. A. (2004). "Srp-2 is a cross-class inhibitor that participates in post-embryonic development of the nematode *Caenorhabditis elegans*: Initial characterization of the Clade L serpins." J. Biol. Chem **279**(15):15448-59

Patel, S., Wang, F., Whiteside, T. L. and Kasid, U. (1997). "Constitutive modulation of Raf-1 protein kinase is associated with differential gene expression of several known and unknown genes." Mol Med **3**: 674-685.

Patrone, L., Damore, M. A., Lee, M. B., Malone, C. and Wall, R. (2001). "Genes expressed during the IFN-gamma-induced maturation of pre-B cells." Mol Immunol **38**: 597-606.

Patston, P. A., Hauert, J., Michaud, M. and Schapira, M. (1995). "Formation and properties of C1-inhibitor polymers." FEBS Letters **368**(3): 401-4.

Pavlovic, D., Chen, M., Gysemans, C. A., Mathieu, C. and Eizirik, D. L. (1999). "The role of interferon regulatory factor-1 in cytokine-induced mRNA expression and cell death in murine pancreatic beta-cells." Euro Cytokine Netw **10**(3): 403-12.

Pemberton, P. A., Stein, P. E., Pepys, M. B., Potter, J. M. and Carrell, R. W. (1988). "Hormone binding globulins undergo serpin conformational change in inflammation." Nature **336**(6196): 257-8.

Pemberton, P. A., Wong, D. T., Gibson, H. L., Kiefer, M. C., Fitzpatrick, P. A., Sager, R. and Barr, P. J. (1995). "The tumor suppressor maspin does not undergo the stressed to relaxed transition or inhibit trypsin-like serine proteases. Evidence that maspin is not a protease inhibitory serpin." J Biol Chem **270**(26): 15832-7.

Perkins, S. J., Smith, K. F., Amatayakul, S., Ashford, D., Rademacher, T. W., Dwek, R. A., Lachmann, P. J. and Harrison, R. A. (1990). "Two-domain structure of the native and reactive centre cleaved forms of C1 inhibitor of human complement by neutron scattering." J Mol Biol 214(3): 751-63.

Perlmutter, D. H., Cole, F. S., Kilbridge, P., Rossing, T. H. and Colten, H. R. (1985). "Expression of the alpha 1-proteinase inhibitor gene in human monocytes and macrophages." Proc Natl Acad Sci 82(3): 795-9.

Perona, J. J. and Craik, C. S. (1995). "Structural basis of substrate specificity in the serine proteases." Protein Sci 4(3): 337-60.

Perry, D. J., Harper, P. L., Fairham, S., Daly, M. and Carrell, R. W. (1989). "Antithrombin Cambridge, 384 Ala to Pro: a new variant identified using the polymerase chain reaction." FEBS Lett 254: 174-176.

Persons, D. A., Allay, J. A., Allay, E. R., Smeyne, R. J., Ashmun, R. A., Sorrentino, B. P. and Nienhuis, A. W. (1997). "Retroviral-mediated transfer of the green fluorescent protein gene into murine hematopoietic cells facilitates scoring and selection of transduced progenitors in vitro and identification of genetically modified cells in vivo." Blood 90(5): 1777-86.

Persons, D. A., Allay, J. A., Allay, E. R., Ashmun, R. A., Orlic, D., Jane, S. M., Cunningham, J. M. and Nienhuis, A. W. (1999). "Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis." Blood 93(2): 488-99.

Plotnick, M. I., Schechter, N. M., Wang, Z. M., Liu, X. Z. and Rubin, H. (1997). "Role of the P6-P3' region of the serpin reactive loop in the formation and breakdown of the inhibitory complex." Biochemistry 36: 14601-14608.

Plotnick, M. I., Rubin, H. and Schechter, N. M. (2002). "The effects of reactive site location on the inhibitory properties of the serpin alpha(1)-antichymotrypsin." J Biol Chem 277(33): 29927-35.

Poller, W., Faber, J. P., Weidinger, S., Tief, K., Scholz, S., Fischer, M., Olek, K., Kirchgesser, M. and Heidtmann, H. H. (1993). "A leucine-to-proline substitution causes a defective alpha 1-antichymotrypsin allele associated with familial obstructive lung disease." Genomics **17**(3): 740-3.

Potempa, J., Korzus, E. and Travis, J. (1994). "The serpin superfamily of proteinase inhibitors: structure, function, and regulation." J Biol Chem **269**(23): 15957-60.

Pratt, C. W. and Church, F. C. (1992). "Heparin binding to protein C inhibitor." J Biol Chem **267**(13): 8789-94.

Radomska, H. S., Gonzalez, D. A., Okuno, Y., Iwasaki, H., Nagy, A., Akashi, K., Tenen, D. G. and Huettnr, C. S. (2002). "Transgenic targeting with regulatory elements of the human CD34 gene." Blood **100**(13): 4410-9.

Ragg, H., Lokot, T., Kamp, P. B., Atchley, W. R. and Dress, A. (2001). "Vertebrate serpins: construction of a conflict-free phylogeny by combining exon-intron and diagnostic site analyses." Mol Biol Evol **18**(4): 577-84.

Remold-O'Donnell, E. (1985). "A fast-acting elastase inhibitor in human monocytes." J Exp Med **162**(6): 2142-55.

Remold-O'Donnell, E., Chin, J. and Alberts, M. (1992). "Sequence and molecular characterization of human monocyte/neutrophil elastase inhibitor." Proc Natl Acad Sci **15**(89): 5635-9.

Remold-O'Donnell, E. (1993). "The ovalbumin family of serpins." FEBS Letters **315**(2): 105-108.

Riewald, M. and Schleef, R. R. (1995). "Molecular cloning of bornapin (protease inhibitor 10), a novel human serpin that is expressed specifically in the bone marrow." J Biol Chem **270**(45): 26754-7.

Riewald, M. and Schleef, R. R. (1996). "Human cytoplasmic antiproteinase neutralizes rapidly and efficiently chymotrypsin and trypsin-like proteases utilising distinct reactive site residues." J Biol Chem **271**(24): 14526-32.

- Riewald, M., Chuang, T., Neubauer, A., Riess, H. and Schleef, R. R. (1998). "Expression of bomapin, a novel human serpin, in normal/malignant hematopoiesis and in the monocytic cell lines THP-1 and AML-193." Blood **91**(4): 1256-62.
- Rijken, D. C., Groeneveld, E., Kluft, C. and Nieuwenhuis, H. K. (1988). "Alpha 2-antiplasmin Enschede is not an inhibitor, but a substrate, of plasmin." Biochem J **255**(2): 609-15.
- Robertson, A. S., Belorgey, D., Lilley, K. S., Lomas, D. A., Gubb, D. and Dafforn, T. R. (2003). "Characterization of the necrotic protein that regulates the Toll-mediated immune response in *Drosophila*." J Biol Chem **278**(8): 6175-80.
- Rollini, P. and Fournier, R. E. K. (1997). "A 370-kb Cosmid Contig of the Serpin Gene Cluster on Human Chromosome 14q32.1: Molecular Linkage of the Genes Encoding [alpha]1-Antichymotrypsin, Protein C Inhibitor, Kallistatin, [alpha]1-Antitrypsin, and Corticosteroid-Binding Globulin." Genomics **46**(3): 409-415.
- Rovelli, G., Stone, S. R., Guidolin, A., Sommer, J. and Monard, D. (1992). "Characterization of the heparin-binding site of glia-derived nexin/protease nexin-1." Biochemistry **31**(13): 3542-3549.
- Rubin, H., Wang, Z. M., Nickbarg, E. B., McLarney, S., Naidoo, N., Schoenberger, O. L., Johnson, J. L. and Cooperman, B. S. (1990). "Cloning, expression, purification, and biological activity of recombinant native and variant human alpha 1-antichymotrypsins." J Biol Chem **265**(2): 1199-207.
- Rubin, H., Plotnick, M., Wang, Z. M., Liu, X., Zhong, Q., Schechter, N. M. and Cooperman, B. S. (1994). "Conversion of alpha 1-antichymotrypsin into a human neutrophil elastase inhibitor: demonstration of variants with different association rate constants, stoichiometries of inhibition, and complex stabilities." Biochemistry **33**(24): 7627-33.
- Salvesen, G. S. and Dixit, V. M. (1997). "Caspases: Intracellular signalling by proteolysis." Cell **91**: 443-446.
- Schechter, I. and Berger, A. (1967). "On the size of the active site in proteases. I. Papain." Biochem Biophys Res Commun **27**(2): 157-62.

Schick, C., Pemberton, P. A., Shi, G. P., Kamachi, Y., Cataltepe, S., Bartuski, A. J., Gornstein, E. R., Bromme, D., Chapman, H. A. and Silverman, G. A. (1998). "Cross-class inhibition of the cysteine proteinases cathepsins K, L, and S by the serpin squamous cell carcinoma antigen 1: A kinetic analysis." Biochemistry **37**: 5258-5266.

Schleef, R. R. and Chuang, T. L. (2000). "Protease inhibitor 10 inhibits tumor necrosis factor alpha -induced cell death. Evidence for the formation of intracellular high M(r) protease inhibitor 10-containing complexes." J Biol Chem **275**(34): 26385-9.

Schreuder, H. A., de Boer, B., Dijkema, R., Mulders, J., Theunissen, H. J., Grootenhuis, P. D. and Hol, W. G. (1994). "The intact and cleaved human antithrombin III complex as a model for serpin-proteinase interactions." Nat Struct Biol **1**(1): 48-54.

Schulze, A. J., Baumann, U., Knof, S., Jaeger, E., Huber, R. and Laurell, C.-B. (1990). "Structural transition of α 1-antitrypsin by a peptide sequentially similar to β -strand s4A." Eur. J. Biochem. **194**: 51-56.

Schuster, M., Enriquez, P., Curran, P., Cooperman, B. and Rubin, H. (1992). "Regulation of neutrophil superoxide by antichymotrypsin-chymotrypsin complexes." J Biol Chem **267**(8): 5056-5059.

Scott, F. L., Eyre, H. J., Lioumi, M., Ragoussis, J., Irving, J. A., Sutherland, G. A. and Bird, P. I. (1999 (a)). "Human ovalbumin serpin evolution: phylogenic analysis, gene organization, and identification of new PI8-related genes suggest that two interchromosomal and several intrachromosomal duplications generated the gene clusters at 18q21-q23 and 6p25." Genomics **62**(3): 490-9.

Scott, F. L., Hirst, C. E., Sun, J., Bird, C. H., Bottomley, S. P. and Bird, P. I. (1999 (b)). "The intracellular serpin proteinase inhibitor 6 is expressed in monocytes and granulocytes and is a potent inhibitor of the azurophilic granule protease, cathepsin G." Blood **93**(6): 2089-97.

Sejima, H., Hayashi, T., Deyashiki, Y., Nishioka, J. and Suzuki, K. (1990). "Primary structure of vitamin K-dependent human protein Z." Bioch Biophys Res Comm **171**(2): 661-8.

Shapiro, S. D. (2002). "Neutrophil Elastase . Path Clearer, Pathogen Killer, or Just Pathologic?" Am. J. Respir. Cell Mol. Biol. 26(3): 266-268.

Sheehan, J. P., Tollefsen, D. M. and Sadler, J. E. (1994). "Heparin cofactor II is regulated allosterically and not primarily by template effects. Studies with mutant thrombins and glycosaminoglycans." J Biol Chem 269(52): 32747-51.

Silverman, G. A., Bird, P. I., Carrell, R. W., Coughlin, P. B., Gettins, P. G., Irving, J. I., Lomas, D. A., Luke, C. J., Moyer, R. W., Pemberton, P. A., Remold-O'Donnell, E., Salvesen, G. S., Travis, J. and Whisstock, J. C. (2001). "The serpins are an expanding superfamily of structurally similar but functionally diverse proteins: Evolution, mechanism of inhibition, novel functions, and a revised nomenclature." J Biol Chem 276(36):33293-6

Simonovic, M., Gettins, P. G. W. and Volz, K. (2000). "Crystal structure of viral serpin crmA provides insights into its mechanism of cysteine proteinase inhibition" Protein Sci 9(8): 1423-7.

Sitnicka, E., Lin, N., Priestley, G. V., Fox, N., Broudy, V. C., Wolf, N. S. and Kaushansky, K. (1996). "The effect of thrombopoietin on the proliferation and differentiation of murine hematopoietic stem cells." Blood 87(12): 4998-5005.

Sivasothy, P., Dafforn, T. R., Gettins, P. G. and Lomas, D. A. (2000). "Pathogenic alpha 1-antitrypsin polymers are formed by reactive loop-beta -sheet A linkage [In Process Citation]." J Biol Chem 275(43): 33663-8.

Smyth, M. J. and Trapani, J. A. (1995). "Granzymes - exogenous proteinases that induce target-cell apoptosis." Immunol Today 16(4): 202-206.

Spooner, E., Boettiger, D. and Dexter, T. M. (1984). "Continuous *in vitro* generation of multipotential stem cell clones from src-infected cultures." Nature 310(5974): 228-30.

Sprecher, C. A., Morgenstern, K. A., Mathewes, S., Dahlen, J. R., Schrader, S. K., Foster, D. C. and Kisiel, W. (1995). "Molecular cloning, expression, and partial characterization of two novel members of the ovalbumin family of serine proteinase inhibitors." J Biol Chem 270(50): 29854-61.

Stavridi, E. S., O'Malley, K., Lukacs, C. M., Moore, W. T., Lambris, J. D., Christianson, D. W., Rubin, H. and Cooperman, B. S. (1996). "Structural change in alpha-chymotrypsin induced by complexation with alpha 1-antichymotrypsin as seen by enhanced sensitivity to proteolysis." Biochemistry 35(33): 10608-15.

Stein, P. E., Tewkesbury, D. A. and Carrell, R. W. (1989). "Ovalbumin and angiotensinogen lack serpin S-R conformational change." Biochem J 262: 103-107.

Stein, P. E., Leslie, A. G. W., Finch, J. T., Turnell, W. G., McLaughlin, P. J. and Carrell, R. W. (1990). "Crystal structure of ovalbumin as a model for the reactive centre of serpins." Nature 347(6288): 99-102.

Stein, P. and Chothia, C. (1991). "Serpins tertiary structure transformation." J. Mol. Biol. 221: 615-621.

Stein, P. E., Leslie, A., Finch, J. T. and Carrell, R. W. (1991). "Crystal-structure of uncleaved ovalbumin at 1.95 Å resolution." J Mol Biol 221: 941-959.

Stein, P. E. and Carrell, R. W. (1995). "What do dysfunctional serpins tell us about molecular mobility and disease." Nat Struct Biol 2: 96-113.

Stephan, C., Jung, K., Diamandis, E. P., Rittenhouse, H. G., Lein, M. and Loening, S. A. (2002). "Prostate-specific antigen, its molecular forms, and other kallikrein markers for detection of prostate cancer." Urology 59(1): 2-8.

Stone, S. R. and LeBonniec, B. F. (1997). "Inhibitory mechanism of serpins. Identification of steps involving the active-site serine residue of the protease." J Mol Biol 265: 344-362.

Sun, J., Rose, J. B. and Bird, P. (1995 (a)). "Gene structure, chromosomal localisation, and expression of the murine homologue of human proteinase inhibitor6 (PI6) suggests divergence of PI6 from the ovalbumin serpins." J Biol Chem 270: 16089-16096.

Sun, J. R., Coughlin, P., Salem, H. H. and Bird, P. (1995 (b)). "Production and characterization of recombinant human proteinase- inhibitor-6 expressed in pichia-pastoris." Biochim Biophys Acta 1252: 28-34.

Sun, J. R., Bird, C. H., Sutton, V., McDonald, L., Coughlin, P. B., Dejong, T. A., Trapani, J. A. and Bird, P. I. (1996). "A cytosolic granzyme-B inhibitor related to the viral apoptotic regulator cytokine response modifier A is present in cytotoxic lymphocytes." J Biol Chem **271**: 27802-27809.

Sun, J., Stephens, R., Mirza, G., Kanai, H., Ragoussis, J. and Bird, P. (1998). "A serpin gene cluster on human chromosome 6p25 contains PI6, PI9 and ELANH2 which have a common structure almost identical to the 18q21 ovalbumin serpin genes." Cytogenet Cell Genet **82**(3-4): 273-7.

Sun, J. R., Ooms, L., Bird, C. H., Sutton, V. R., Trapani, J. A. and Bird, P. I. (1997). "A new family of 10 murine ovalbumin serpins includes two homologs of proteinase inhibitor 8 and two homologs of the granzyme B inhibitor (proteinase inhibitor 9)." J Biol Chem **272**: 15434-15441.

Suzuki, D., Miyata, T., Nangaku, M., Takano, H., Saotome, N., Toyoda, M., Mori, Y., Zhang, S. Y., Inagi, R., Endoh, M., Kurokawa, K. and Sakai, H. (1999). "Expression of megsin mRNA, a novel mesangium-predominant gene, in the renal tissues of various glomerular diseases." J Am Soc Nephrol **10**(12): 2606-13.

Suzuki, K., Nishioka, J., Kusumoto, H. and Hashimoto, S. (1984). "Mechanism of inhibition of activated protein C by protein C inhibitor." J Biochem **95**(1): 187-95.

Sveger, T. (1976). "Liver disease in alpha1-antitrypsin deficiency detected by screening of 200,000 infants." N. Engl. J. Med **294**(24): 1316-21.

Szilvassy, S. J. and Cory, S. (1993). "Phenotypic and functional characterization of competitive long-term repopulating hematopoietic stem cells enriched from 5-fluorouracil-treated murine marrow." Blood **81**(9): 2310-20.

Szilvassy, S. J. and Cory, S. (1994). "Efficient retroviral gene transfer to purified long-term repopulating hematopoietic stem cells." Blood **84**(1): 74-83.

Szilvassy, S. J., Bass, M. J., Van Zant, G. and Grimes, B. (1999). "Organ-Selective Homing Defines Engraftment Kinetics of Murine Hematopoietic Stem Cells and Is Compromised by Ex Vivo Expansion." Blood **93**(5): 1557-1566.

Szilvassy, S. J. (2003). "The biology of hematopoietic stem cells." Arch Med Res 34(6): 446-460.

Tabatabai, A., Fiehler, R. and Broze, G. J., Jr. (2001). "Protein Z circulates in plasma in a complex with protein Z-dependent protease inhibitor." Thromb Haemost 85(4): 655-60.

Taggart, C., Cervantes-Laurean, D., Kim, G., McElvaney, N. G., Wehr, N., Moss, J. and Levine, R. L. (2000). "Oxidation of either methionine 351 or methionine 358 in alpha 1-antitrypsin causes loss of anti-neutrophil elastase activity." J Biol Chem 275(35): 27258-65.

Takada, S., Tsuda, M., Fujinami, S., Yamamura, M., Mitomi, T. and Katsunuma, T. (1986). "Incorporation of alpha-1-antichymotrypsin into carcinoma cell nuclei of human stomach adenocarcinoma transplanted into nude mice." Cancer Res 46(7): 3688-91.

Takada, S., Tsuda, M., Matsumoto, M., Fujinami, S., Yamamura, M. and Katsunuma, T. (1988). "Incorporation of alpha-1-antichymotrypsin into human stomach adenocarcinoma cell nuclei and inhibition of DNA primase activity." Tokai J Exp Clin Med 13(6): 321-7.

Takahara, H. and Sinohara, H. (1982). "Mouse plasma trypsin inhibitors: Isolation and characterization of alpha-1-antitrypsin and contrapsin, a novel trypsin inhibitor." J Biol Chem 257: 2438-2446.

Takahara, H. and Sinohara, H. (1983). "Inhibitory spectrum of mouse contrapsin and alpha-1-antitrypsin against mouse serine proteases." J Biochem 93: 1411-1419.

Teglund, S., McKay, C., Schuetz, E., van Deursen, J. M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G. and Ihle, J. N. (1998). "Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses." Cell 93(5): 841-50.

Terskikh, A. V., Easterday, M. C., Li, L., Hood, L., Kornblum, H. I., Geschwind, D. H. and Weissman, I. L. (2001). "From hematopoiesis to neuropoiesis: evidence of overlapping genetic programs." Proc Natl Acad Sci 98(14): 7934-9.

To, L. B., Haylock, D. N., Simmons, P. J. and Juttner, C. A. (1997). "The Biology and Clinical Uses of Blood Stem Cells." Blood 89(7): 2233-2258.

Travis, J., Owen, M., George, P., Carrell, R., Rosenberg, S., Hallewell, R. A. and Barr, P. J. (1985). "Isolation and properties of recombinant DNA produced variants of human alpha 1-proteinase inhibitor." J Biol Chem 260(7): 4384-9.

Travis, J. and Salvesen, G. S. (1983). "Human plasma proteinase inhibitors." Ann Rev Biochem 52: 655-709.

Tsuda, M., Masuyama, M. and Katsunuma, T. (1986). "Inhibition of human DNA polymerase alpha by alpha 1-antichymotrypsin." Cancer Res 46(12 Pt 1): 6139-42.

Tsujimoto, M., Tsuruoka, N., Ishida, N., Kurihara, T., Iwasa, F., Yamashiro, K., Rogi, T., Kodama, S., Katsuragi, N., Adachi, M., Katayama, T., Nakao, M., Yamaichi, K., Hashino, J., Haruyama, M., Miura, K., Nakanishi, T., Nakazato, H., Teramura, M., Mizoguchi, H. and Yamaguchi, N. (1997). "Purification, cDNA cloning, and characterization of a new serpin with megakaryocyte maturation activity." J Biol Chem 272: 15373-15380.

Turk, B., Turk, D. and Turk, V. (2000). "Lysosomal cysteine proteases: more than scavengers." Biochim Biophys Acta 1477(1-2): 98-111.

Turk, V., Turk, B. and Turk, D. (2001). "Lysosomal cysteine proteases: facts and opportunities." EMBO J 20(17): 4629-33.

Uhrin, P., Dewerchin, M., Hilpert, M., Chrenek, P., Schofer, C., Zechmeister-Machhart, M., Kronke, G., Vales, A., Carmeliet, P., Binder, B. R. and Geiger, M. (2000). "Disruption of the protein C inhibitor gene results in impaired spermatogenesis and male infertility." J Clin Invest 106(12): 1531-1539.

Van Deerlin, V. M. and Tollefsen, D. M. (1991). "The N-terminal acidic domain of heparin cofactor II mediates the inhibition of alpha-thrombin in the presence of glycosaminoglycans." J Biol Chem 266(30): 20223-31.

van Gent, D., Sharp, P., Morgan, K. and Kalsheker, N. (2003). "Serpins: structure, function and molecular evolution." Int J Biochem Cell Biol 35(11): 1536-1547.

von Heijne, G., Liljestrom, P., Mikus, P., Andersson, H. and Ny, T. (1991). "The efficiency of the uncleaved secretion signal in the plasminogen activator inhibitor type 2 protein can be enhanced by point mutations that increase its hydrophobicity." J Biol Chem 266(23): 15240-3.

Wei, A., Rubin, H., Cooperman, B. S. and Christianson, D. W. (1994). "Crystal structure of an uncleaved serpin reveals the conformation of an inhibitory reactive loop." Nat Struct Biol 1(4): 251-8.

Weissman, I. L., Anderson, D. J. and Gage, F. (2001). "Stem and progenitor cells: Origins, Phenotypes, Lineage Commitments, and Transdifferentiations." Ann Rev Cell Devel Biol 17(1): 387-403.

Welss, T., Sun, J., Irving, J. A., Blum, R., Smith, A. I., Whisstock, J. C., Pike, R. N., von Mikecz, A., Ruzicka, T., Bird, P. I. and Abts, H. F. (2003). "Hurpin is a selective inhibitor of lysosomal cathepsin L and protects keratinocytes from ultraviolet-induced apoptosis." Biochemistry 42(24): 7381-9.

Whisstock, J. C., Skinner, R., Carrell, R. W. and Lesk, A. M. (2000). "Conformational changes in serpins: I. The native and cleaved conformations of alpha(1)-antitrypsin." J Mol Biol 295(3): 651-65.

Whitehead, I., Kirk, H. and Kay, R. (1995). "Expression cloning of oncogenes by retroviral transfer of cDNA libraries." Mol Cell Biol 15(2): 704-710.

Wilczynska, M., Fa, M., Ohlsson, P. I. and Ny, T. (1995). "The inhibition mechanism of serpins. Evidence that the mobile reactive center loop is cleaved in the native protease-inhibitor complex." J Biol Chem 270(50): 29652-5.

Winkler, I. G., Hendy, J., Coughlin, P., Horvath, A. and Levesque, J. P. (2004). "Down-Regulation of the Serine-Protease Inhibitors, Serpina1 and Serpina3, in the Bone Marrow During Hematopoietic Progenitor Cell Mobilization Results in Increased Neutrophil Serine-Protease Activity." (Submitted).

Wright, H. T., Qian, H. X. and Huber, R. (1990). "Crystal structure of plakalbumin, a proteolytically nicked form of ovalbumin Its relationship to the structure of cleaved alpha-1-proteinase inhibitor." J. Mol. Biol. 213: 513-528.

Ye, R. D., Wun, T. C. and Sadler, J. E. (1988). "Mammalian protein secretion without signal peptide removal. Biosynthesis of plasminogen activator inhibitor-2 in U-937 cells." J Biol Chem 263(10): 4869-75.

Young, J. C., Lin, K., Travis, M., Hansteen, G., Abitorabi, A., Sirenko, O., Murray, L. and Hill, B. (2001). "Investigation into an engraftment defect induced by culturing primitive hematopoietic cells with cytokines." Cytotherapy 3(4): 307-20.

Young, J. L., Sukhova, G. K., Foster, D., Kiesel, W., Libby, P. and Schonbeck, U. (2000). "The serpin proteinase inhibitor 9 is an endogenous inhibitor of interleukin 1beta-converting enzyme (caspase-1) activity in human vascular smooth muscle cells." J Exp Med 191(9): 1535-44.

Yu, J., Li, Y., Ishizuka, T., Guenther, M. G. and Lazar, M. A. (2003). "A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation." EMBO J, 22(13): 3403-3410.

Zhang, M., Volpert, O., Shi, Y. H. and Bouck, N. (2000). "Maspin is an angiogenesis inhibitor." Nat Med 6(2): 196-9.

Zhou, A., Carrell, R. W. and Huntington, J. A. (2001). "The serpin inhibitory mechanism is critically dependent on the length of the reactive center loop." J Biol Chem 276(29): 27541-7.

Zhou, A., Stein, P. E., Huntington, J. A. and Carrell, R. W. (2003). "Serpin polymerization is prevented by a hydrogen bond network that is centered on his-334 and stabilized by glycerol." J Biol Chem 278(17): 15116-22.

Zou, Z. Q., Anisowicz, A., Hendrix, M., Thor, A., Neveu, M., Sheng, S. J., Rafidi, K., Seftor, E. and Sager, R. (1994). "Maspin, a serpin with tumor-suppressing activity in human mammary epithelial-cells." Science 263: 526-529.



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

GENOMICS

Genomics 81 (2003) 336–345

www.elsevier.com/locate/ygeno

Review

A review and comparison of the murine α_1 -antitrypsin and α_1 -antichymotrypsin multigene clusters with the human clade A serpins

Sharon Forsyth, Anita Horvath, and Paul Coughlin*

Department of Medicine, Monash University, Melbourne 3128, Australia

Received 24 September 2002; accepted 6 December 2002

Abstract

The major human plasma protease inhibitors, α_1 -antitrypsin and α_1 -antichymotrypsin, are each encoded by a single gene, whereas in the mouse they are represented by clusters of 5 and 14 genes, respectively. Although there is a high degree of overall sequence similarity within these groupings, the reactive-center loop (RCL) domain, which determines target protease specificity, is markedly divergent. The literature dealing with members of these mouse serine protease inhibitor (serpin) clusters has been complicated by inconsistent nomenclature. Furthermore, some investigators, unaware of the complexity of the family, have failed to distinguish between closely related genes when measuring expression levels or functional activity. We have reviewed the literature dealing with the mouse equivalents of human α_1 -antitrypsin and α_1 -antichymotrypsin and made use of the recently completed mouse genome sequence to propose a systematic nomenclature. We have also examined the extended mouse clade "a" serpin cluster at chromosome 12F1 and compared it with the syntenic region at human chromosome 14q32. In summarizing the literature and suggesting a standardized nomenclature, we aim to provide a logical structure on which future research may be based.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Serpin; α_1 -antitrypsin; α_1 -antichymotrypsin; Clade A; Murine; Human; Evolution; Gene duplication

Introduction

Human disease has been the inspiration for the study of genetics. The conventional approach to inherited disorders has been to track phenotypes through family studies and, more recently, through positional cloning, to identify the responsible gene. Mass sequencing of genomes has made available large volumes of genetic data, and there is a need to correlate this information with human disease. To address this problem, researchers have relied increasingly on animals for the purpose of modifying or deleting specific genes of interest and assessing the consequent phenotype. Implicit in this approach is the assumption that the molecular mechanisms subserving major biological processes are conserved among species. In some cases this seems accurate, and valuable lessons can be learned from genetically modified organisms as far removed from humans as nematodes or fruitflies. A more relevant model organism for mammalian

biology is the genetically tractable mouse and, in general, there is a high degree of conservation between human and mouse genomes, not only within the coding regions but also in the intergenic, noncoding DNA [1,2]. This is reflected overall in strong functional conservation, making the mouse an instructive model organism for genetic manipulation, although there are clear exceptions to this general rule.

Our particular interest is the serpin gene family and the function of its various members in human biology and disease. This review deals with the cluster of genes at mouse chromosome 12F1 corresponding to human chromosome 14q32.1 in which occur the clade A serpins, antitrypsin and antichymotrypsin [3]. Both of these genes have undergone considerable expansion in the mouse, and as a consequence the number, function, and nomenclature of all of these genes are confusing and inconsistent. This review summarizes the literature dealing with the mouse antitrypsin and antichymotrypsin serpins and, using the resources available from the recently completed mouse genome, describes their genomic organization and presents an updated, systematic nomenclature. In addition the genomic organization and

* Corresponding author. Fax: +61-3-98950332.

E-mail address: [redacted] (P. Coughlin).

homologies of genes present in the human and mouse serpin A clades on Hsa14 and Mmu12 are compared.

The mouse α_1 -protease inhibitor (antitrypsin) and con-trapsin (antichymotrypsin) clusters comprise 5 and 14 members, respectively. Both clusters represent the results of multiple gene duplication events, with a high degree of conservation in secondary structural elements surrounding substantially divergent reactive centers giving rise to functional inhibitors with variable specificities [4,5]. The complexity of the mouse *12F1* locus, and the forces driving its diversity, may well hold some of the keys to understanding the biology of human α_1 -protease inhibitor and antichymotrypsin. Furthermore, the mechanisms underlying the evolution of the serpins have implications for other gene families. The finding that the serpin scaffold is under strong conservative pressure, whereas the specificity-determining reactive-site loop is relatively free to diverge, provides an unusual evolutionary paradigm. Some appreciation of the intricacies of this locus is essential to understand and interpret the biochemical and biological literature dealing with these serpins.

The serpin superfamily—structure, function, and phylogenetic history

The serine protease inhibitors (serpins) compose a large family of functionally diverse proteins (>700 to date). Most serpins are inhibitors of either serine or cysteine proteases involved in numerous intracellular and extracellular processes including blood coagulation, fibrinolysis, cell migration, proliferation, embryo implantation, complement activation, and tumor suppression (for reviews see [3,6]). Some serpins have noninhibitory roles such as blood pressure regulation by angiotensinogen, chaperoning of collagen folding by the heat shock protein Hsp47, and hormone binding by corticosteroid-binding globulin (for reviews see [3,6]). Representatives of the serpin family have been found in all eukaryotes as well as pox viruses, and have recently been described in prokaryotes as well [7,8]. However, none have been identified in unicellular yeast, and several prokaryotic genomes have been fully sequenced without finding any members of the serpin family. In general, the evidence suggests that serpins may have arisen before animal and plant divergence (~2.7 billion years ago) and even before eukaryote development (>3.9 billion years ago). Another possibility is that the serpins are in fact a younger gene family and that horizontal transfer has occurred through viruses. Although there is no direct evidence for this, it would explain the apparently patchy representation of serpins in prokaryotes and yeast.

Despite their diversity of function, serpins demonstrate a highly conserved structure consisting of three β -sheets (A, B, and C) and at least seven α -helices (Fig. 1) (for review see [3]). The critical feature of the serpin mechanism is that the active, native conformation is metastable and interaction with a target protease triggers the transition to a stable

configuration. In the native conformation the RCL, a short flexible strand composing the protease recognition region, is exposed for interaction with its ligand. After formation of an initial noncovalent complex, a covalent link is established during cleavage of the peptide bond between P_1 and P'_1 . The free RCL inserts into the middle of β -sheet A, distorting the tethered proteinase, which impedes deacylation and traps the protease in an effectively irreversible complex (Fig. 1).

Numerous chromosomal localization studies and, more recently, whole-genome analyses reveal that serpin genes are often clustered, presumably as a result of gene duplications [9]. Many serpin gene subfamilies demonstrate conserved intron-exon patterns providing clues to their phylogenetic history. Using these patterns, together with diagnostic amino acid sites and rare indels (insertion or deletion of three nucleotides corresponding to a codon), Ragg et al. have assigned the serpins to six gene families [10]. Subsequently they assigned the serpins to various clades by means of tree reconstruction analyses incorporating amino acid sequences, exon-intron structure, and family-specific diagnostic amino acid sites [9]. Irving et al., in contrast, assigned the serpins to clades by means of consensus analysis based on examination of patterns of sequence conservation to assess phylogenetic relationships within the superfamily [7]. Genomic structure was not taken into account. With some exceptions, the phylogenetic trees suggested by Ragg et al. and Irving et al. are similar. However, from a practical perspective, phylogenetic analysis based on amino acid sequence alone provides the most accessible method for serpin classification. Therefore, subsequent to the Second International Symposium on the Structure and Biology of Serpins (1999), a nomenclature for the entire serpin family was proposed based on the clade structure suggested by Irving et al. and was presented as supplementary data in the review by Silverman et al. [3].

The mouse α_1 -antitrypsin, α_1 -protease inhibitor family

Human α_1 -protease inhibitor (α_1 PI) or α_1 -antitrypsin (A1AT) (SERPINA1) is a 50–55-kDa glycosylated protein that is synthesized primarily in the liver. It is expressed at high levels and is an efficient inhibitor of neutrophil elastase. Deficiency of A1AT results in unregulated leukocyte elastase activity, leading to destruction of lung elastic tissue (emphysema; for review see [11]). In striking contrast to the human and bovine genomes, wherein α_1 PI is represented by a single gene, four gene members are found in the guinea pig and rabbit, and individual mouse species possess as many as five α_1 PI isoforms (Table 1) [5,12–14]. However, the number of different α_1 PI genes is species-specific, with *Mus caroli* and *Mus cookii* expressing only one form of the α_1 PI gene, whereas *Mus saxicola* express four α_1 PI isoforms and *Mus domesticus* five [5]. Surprisingly, intraspecies variation of expression is also observed, with *M. domesticus* C57BL/6J, A/J, and C3H/HeJ strains expressing five variants whereas AKR/J and DBA/2J express only three



Fig. 1. Schematic representation of a serpin and a serpin-protease complex. The structure of native human α_1 -antitrypsin is shown on the left with the reactive center loop at the top of the molecule. The complex of α_1 -antitrypsin with trypsin is shown on the right. The cleaved reactive center loop (shown in black) is inserted into the A sheet and the tethered, inactivated protease (dark gray) is at the base of the α_1 -antitrypsin molecule. The figures were produced using Protein Explorer (<http://molvis.sdsc.edu/protexpl/frntdoor.htm>) using PDB Database entries 1QLP and 1EZK from Elliot et al., 2000 [48] and Huntington et al., 2000 [49].

[5]. The simplest explanation for this observation would be variability at the transcriptional level, yet evidence based on gene-specific PCR indicates that intraspecies and interspecies variability is caused by different α_1 PI gene numbers [5] (see Table 3 for revised nomenclature of the *M. domesticus* α_1 PI genes). All species studied so far express at least one orthodox α_1 PI (with methionine at the P_1 position), and differences in expression between strains involve only the

unorthodox members (variation at the P_1 amino acid) (Table 1). The presence of methionine at the P_1 position may be favored because it produces efficient inhibition of elastase while conferring sensitivity to oxidation; this property of α_1 PI allows relatively unrestrained proteolytic activity within the privileged microenvironment of an inflammatory site [15–17].

The α_1 PI gene family is notable for its strikingly rapid accumulation of amino acid substitutions in the RCL resulting from a higher than average rate of nonsynonymous substitutions [5,18]. The evolutionary forces driving diversity within the α_1 PI RCL are not understood. It is likely that the requirement for maximum control of elastase activity acts as a selective constraint on the orthodox α_1 PI genes. It is noteworthy, however, that although the P_1 amino acid remains unchanged in the orthodox α_1 PI, there is considerable variation in the remaining reactive site, especially between P_2 and P_6 , which may confer a broader inhibitory spectrum or resistance to cleavage by nontarget proteases. The duplication of α_1 PI genes combined with rapid divergent evolution within the RCL implies that positive darwinian selection is operating. In contrast, the 5' domain of the serpin genes, which encode most of the secondary structural elements, is highly conserved [13]. Calculations based on

Table 1

	RCL
Human SERPINA1	AMFLEAIPMSIPPEVKFNK
Murine DOM1	VTVLQMVPMSMPPIILRFDH
Murine DOM2	ATVFEAVPMSMPPIILRFDH
Murine DOM3	ATVLLAVPYSMPPIVRFDH
Murine DOM4	ATVLQVATYSMPPIIVRFDH
Murine DOM5	ATVLQAGFLSMPPILHFNH
Murine DOM6	ATVLLAVPYSMPPILRFDH
Murine SAX1	TTIVEAVFMSLPPILHFNH
Murine SAX2	TTIVEGVFMSLPPILHFNH
Murine SAX3	TTVLGASYMSAPPILNFNC
Murine SAX4	TTVLGSTLYSAPAILHFNH
Murine SAX5	TTVLACTFTSWPPILNFNR

(Adapted from Goodwin et al., 1997.)

Table 2

Common name	Revised	RCL
Unknown (1) or 4933406L18RIK	<i>Serpina3a</i>	IARYNFQSA-KIKAKIVK
6A1	<i>Serpina3b</i>	IVGYNFMSA-KLKPVFVK
1A1 or KLKBP	<i>Serpina3c</i>	GVNFRILSR-RTS---LW
2B1	<i>Serpina3d</i>	RFKIAPLSA-KFDIVNVN
2B2 (b)	<i>Serpina3e</i>	GVKVNLRG-KIYSMTIY
2A1	<i>Serpina3f</i>	GYQNLQCCQGVIIYSMKIY
2A2 or Spi2/eh1 or Spi2-1	<i>Serpina3g</i>	GMAGVGCCA-VDFLEIF
6C28	<i>Serpina3h</i>	GVKLILCCE-KIYSMTIY
2B2	<i>Serpina3i</i>	GVKVNLRG-KIYSMTIY
Unknown2	<i>Serpina3j</i>	RDKYDFLST-KSNPTVVN
MMCM2 or MMSpi2	<i>Serpina3k</i>	GVIGGIRKA-VLP--AVC
3E2	<i>Serpina3l</i>	RAIYNFQSS-KMYPMLLR
3e46 or MMCM7 or MMSpi2.4	<i>Serpina3m</i>	GFIPGFRSR-RLQTMTVQ
Spi2/eb22.4 or Spi2.2	<i>Serpina3n</i>	GVKFPMSA-KLYPLTVY
Human antichymotrypsin	<i>SERPINA3</i>	AVKITLLSALVETRTIVR

the rate of synonymous substitutions (which do not result in an amino acid change and accumulate at a constant rate) suggest that amplification events within the α_1 PI gene occurred before *Mus* speciation, resulting in three to five genes in several species and only one in others [13].

Positive darwinian selection within the RCL would imply altered functional activity between the various gene products. Of the five gene members expressed in *M. domestica*, the two orthodox genes, with methionine at the P₁ position, are efficient inhibitors of elastase, as expected [19]. These two, together with two unorthodox members, were also found to be efficient inhibitors of chymotrypsin, though not trypsin. No cognate protease was found for the fifth unorthodox member. A recent study by Barbour et al. confirmed that different *Serpina1* gene products within the same species differ in their ability to bind various serine proteases indicating distinct biochemical functions [20]. RCL-swapped chimeras showed that the variation in binding was specific to the reactive-site sequence. This group similarly examined the binding of three *M. saxicola* α_1 PI variants to crude snake venom proteases and again found marked differences in their recognition of different unidentified ligands. These results are consistent with the theory that the array of pathogen- or predator-introduced proteases may drive accelerated evolution and the resultant isoforms of the α_1 PI family. Another interesting consequence of this study was the observation that two of the α_1 PI isoforms expressed in *M. domestica* are sexually dimorphic, with a greater expression in males than in females.

The mouse α_1 -antichymotrypsin, *Spi2* family

Human α_1 -antichymotrypsin (SERPINA3) is a 66-kDa glycoprotein synthesized primarily in the liver, although expression in other tissues has been observed [21,22]. It is an efficient inhibitor of the chymotrypsin-like protease cathepsin G as well as chymase, and deficiency is associated

with emphysema, presumably by a mechanism similar to the one that occurs in antitrypsin deficiency [23,24]. α_1 -antichymotrypsin has also been reported to inhibit purified DNA polymerase [25] as well as DNA primase [26]; this observation may be physiologically relevant, in that the serpin contains a nuclear localization signal and nuclear distribution has been shown in some cells [25–28]. The complex of SERPINA3 with the prostate-specific antigen is commonly used as a clinical marker of prostate cancer (for review see [29] and [30]). More recently the binding of SERPINA3 to the amyloid- β peptide and subsequent deposition of this complex in amyloid plaque deposits has been linked to the progression of Alzheimer's disease (for review see [31]).

A mouse gene with sequence similarity to human antichymotrypsin (SERPINA3) was originally given the name contrapsin [4]. It displayed 70% nucleotide and 60% amino acid identity with human SERPINA3; however, the RCL was substantially different. Further investigation showed that contrapsin was a member of a multigene cluster on chromosome 12 that has been termed the *Spi-2* locus [32] and, until recently, was thought to encompass 11 members. A high degree of sequence homology suggests these genes were all derived from a common ancestor represented in humans by the single-copy SERPINA3 gene on chromosome 14q32.1 [33]. Recent availability of the mouse genome from the Mouse Genome Sequencing Consortium (http://www.ensembl.org/Mus_musculus/) indicates 14 members of this gene family at chromosomal location 12F1. To simplify analysis of this gene cluster, we have assigned systematic names consistent with the recently suggested serpin nomenclature [3] and that have been approved by the Mouse Gene Nomenclature Committee (MGNC) (<http://www.informatics.jax.org/mgihome/nomen/>) (Table 2). The gene structure of the family is illustrated in Figure 2 and shows no ordered orientation of the genes. The direction of 3E2 (*Serpina3l*), 3E46 (*Serpina3m*), and 1A1 (*Serpina3c*)

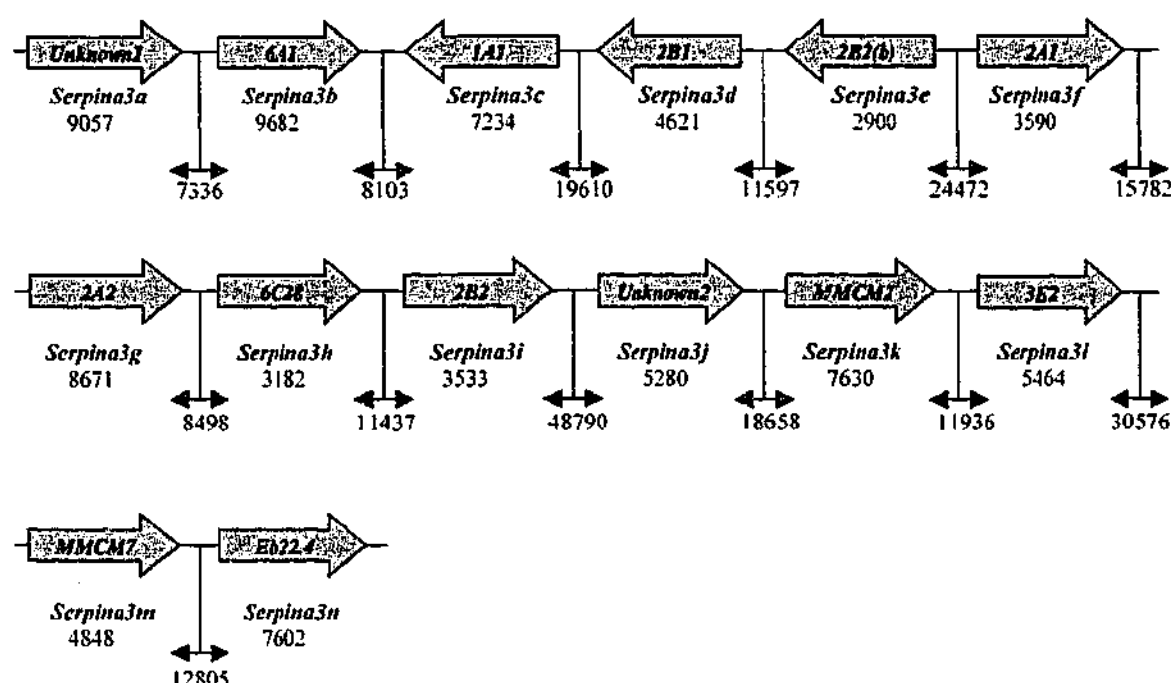


Fig. 2. The structure of the *Serpina3* (Spi2) locus on murine chromosome 12. All 14 genes are represented and the direction of transcription is indicated by the block arrow. The genes are labelled according to the classification in Table 2. The 'common' names are indicated within the arrow and the revised name is indicated below and in bold. The length of each gene as well as the intergenic sequence is shown. Information for this figure was acquired from the publicly available mouse genome sequence database (http://www.ensembl.org/Mus_musculus/).

differs from that predicted by Inglis and Hill [33]. Of the 14 genes, 2 (*Serpina3a* and *Serpina3j*) have not been identified before and another is a shortened form of 2B2 (*Serpina3i*). Although 1A1 (*Serpina3c*) has been identified in the public databases as the mouse homolog of human *SERPINA4* (kallistatin) [34,35], it actually shows a higher degree of homology with *SERPINA3* or antichymotrypsin (59% amino acid identity) than with *SERPINA4* (45% amino acid identity). We have therefore given it a name consistent with its genomic position within the *Serpina3* cluster. Intron/exon boundaries between members are conserved, except that four of the genes (1A1/*Serpina3c*, 2A2/*Serpina3g*, MMSpi2/*Serpina3k*, and eb22.4/*Serpina3n*) have a fifth exon at the 5' end. All intron/exon boundaries correspond with those in the human *SERPINA* clade [3]. It is interesting to note that five members of this family (1A1/*Serpina3c*, Unknown (2)/*Serpina3j*, MMSpi2/*Serpina3k*, 3E46/*Serpina3m*, and eb22.4/*Serpina3n*) possess an amino-terminal secretion signal, whereas another five (6A1/*Serpina3b*, 2A1/*Serpina3f*, 2A2/*Serpina3g*, 2B2/*Serpina3i*, 3E2/*Serpina3l*) would be predicted to have an intracellular localization. The 5' end of the remaining four genes is absent from the predicted cDNA sequences in the mouse genome database; therefore, their targeting to intracellular or extracellular sites cannot be predicted at this stage. The remaining sequence of these genes indicates that, despite substantial variation in the reactive center, most should be able to act as functional protease inhibitors, some with a predictable inhibitory spectrum, others with novel inhibitory activity. In particular the unusual Cys-Cys motif at the P₁-P_{1'} appears in more than one of the gene members (Table 2).

The evolution of this gene family is not as well documented as the *Serpina1* cluster but similarly involves accelerated substitution and divergence at the reactive site and conservation of the remaining sequence [33,36]. This was suggested to be the result of at least three localized gene conversion events that have preserved the structurally important framework of the protein outside of the reactive center. No gene conversion events were detected that encompass the reactive site, thus allowing variation within this region [33]. As with the "a1" cluster, the expansion of the "a3" group appears to be the result of a comparatively recent series of duplication events.

Since the identification of the multiple *Spi-2* genes [32], there has been little progress toward characterizing members of the family. However, one of the genes, namely 2A2 or, as it has been since described, serpin 2A [37], and now *Serpina3g*, has demonstrated a number of interesting characteristics. One of these was its expression at high levels in the primitive, pluripotent, hematopoietic cell line, FDCP-mix A4, and more importantly, its marked downregulation upon differentiation of these cells [37]. Similarly, downregulation of serpin 2A (*Serpina3g*) expression was observed in granulocyte/macrophage colony-forming cells isolated from mouse bone marrow. Constitutive expression of serpin 2A by retroviral transduction into FDCP-mix A4 cells also resulted in a delayed differentiation and increased clonogenicity [37]. In addition, serpin 2A expression was shown to be substantially upregulated upon T-cell activation in splenocyte culture [37]. An independent DNA subtraction study conducted to identify genes selectively expressed in mouse adult hematopoietic stem cells found that the two

most abundant transcripts encoded the protease inhibitors, serpin 2A and CTLA-2 [38]. Interestingly CTLA-2 is also selectively expressed in activated T cells. The importance of these observations remains to be elucidated, and as yet no target protease for serpin 2A has been identified. Examination of the amino acid sequence of this serpin, however, indicates a number of unique features. Despite homology with the secreted human SERPINA3, serpin 2A has no N-terminal secretion signal. In addition serpin 2A has the unusual Cys-Cys at the P₁-P₁' of the reactive center and possesses a unique carboxy-terminal extension of ~30 amino acids containing an additional two cysteine residues. These residues may be involved in some form of redox control mechanism for the activity of the inhibitor. A recent study by Hampson et al. identified the promoter region of serpin 2A and found it contained both an NF- κ B as well as the Moloney murine leukemia enhancer factor Lva binding sites [39]. In the FDCP-mix A4 cells they demonstrated the binding of a unique and unidentified protein to the NF- κ B site and flanking residues. NF- κ B was, however, the major DNA-binding protein in the T cells. The promoter was shown to act in a tissue- and differentiation stage-specific manner. The combined evidence suggests that serpin 2A contributes to the differentiation of primitive hematopoietic cells as well as to the activation of T cells, and that this activity may be regulated by an intracellular redox environment. It will be of great interest to find the physiological target of this inhibitor and to elucidate its mechanism of action.

Mouse *Serpina3* genes have been implicated in the pathogenesis and progression of Alzheimer's disease in studies conducted by Chiang et al. and Licastro et al. [40,41]. They measured the expression of mouse *Serpina3* by RNA protection assay and in situ hybridization. The probes used were derived from *Serpina3g* (eb22.3 or Serpin 2A) [42], but were not located in a unique area of sequence and were thus unlikely to be able to distinguish among the multiple *Spi-2* genes. Results may thus reflect expression of any one or more of the 14 paralogs. The data showing a correlation between expression of one of the *Serpina3* genes, and the progression of Alzheimer's disease-like pathology in apoE4-deficient mice is reminiscent of the up-regulation of human SERPINA3 in Alzheimer's disease.

The substantial expansion of the *Serpina3* gene cluster gives rise to some obvious questions. Why are there 14 genes in the mouse and only one in humans? Why have they diverged to such an extent, are all of the genes expressed, and what functions are they carrying out?

Comparison of the clade A serpin cluster on human chromosome 14 with the syntenic region on mouse chromosome 12

Considering that multiple gene duplications have occurred at the mouse *Serpina1* and *Serpina3* loci, we wondered about the status of the other clade "a" serpins on

mouse chromosome 12F1 and the syntenic region on human chromosome 14q32.1 [3]. By examining the recently available *M. musculus* and *H. sapiens* genomes, we were able to make a direct comparison of this A clade cluster between the two species. An assessment of amino acid identities (Table 3) confirms the high homology of five mouse genes with human SERPINA1 and 14 mouse genes with human SERPINA3. It also shows that the mouse equivalents of SERPINA10 (PZI; protein Z-dependent protease inhibitor), SERPINA6 (CBG; corticosteroid-binding globulin), and SERPINA5 (PCI; protein C inhibitor) are all represented by single homologous genes. The organization of these genes in the chromosome is depicted in Figure 3. What is immediately clear from this figure is that the organization of the A clade serpin genes is largely conserved between the two species, and SERPINA1 and SERPINA3 are the only genes within this locus that have undergone expansion in the mouse. Therefore, the overall organization of the clade A locus predates the evolutionary divergence of mouse and human.

It is odd that there is no obvious homolog for the kallikrein inhibitor (SERPINA4) in the mouse. There is a mouse gene in the syntenic chromosomal location whose predicted transcript shows 57% amino acid identity with SERPINA4 (Table 3); however, it contains stop codons and therefore appears to be a pseudogene. Although the mouse gene 1A1 (*Serpina3c*) has previously been nominated as the homolog of SERPINA4 [34,35], it can be seen from Table 3 that the mouse gene shows a far greater overall similarity with SERPINA3 (59% as opposed to 45% identity with SERPINA4). Nevertheless, the RCLs of SERPINA4 and *Serpina3c* are very similar, and it is predictable that both proteins inhibit kallikrein efficiently. Comparison of the tissue-specific expression of the two genes and *in vitro* biochemical examination of their protease specificities would be useful to demonstrate their functional similarity. SERPINA9 (centerin) [43] is conserved in the mouse, but the gene appears deficient at the 3' end. Whether this represents incomplete sequencing or rather that the gene has fallen into disuse awaits further clarification.

Within the clade A locus there are two novel, uncharacterized serpin genes in the human and one in the mouse. In keeping with the revised serpin nomenclature, the human genes have been named SERPINA11 and SERPINA12, respectively. There is no mouse homolog for SERPINA11, but there exists a mouse gene syntenic with SERPINA12 with 60% predicted amino acid identity that we have designated *Serpina12* (Fig. 3). These names have been approved by the HUGO Gene Nomenclature Committee (HGNC) (<http://www.gene.ucl.ac.uk/nomenclature/>) and the MGNC (<http://www.informatics.jax.org/mgihome/nomen/>), respectively. SERPINA12 contains a Met-Glu at the P₁-P₁' position, whereas SERPINA11 contains a Gln-Pro at the P₁-P₁' position, similar to human SERPINA7 (TBG; thyroxine-binding globulin).

In conclusion, the phenomenon evident within the mouse

Table 3
Mouse clade "a" serpins versus the human clade A serpins^a

Human clade A serpins													
Common name	AIAT	ATR	AACT	KAL	PCI	CBG	Centerin	PZI	Novel	Novel			
Revised name	SERPINA1	SERPINA2	SERPINA3	SERPINA4	SERPINA5	SERPINA6	SERPINA9	SERPINA10	SERPINA11	SERPINA12			
Ensembl gene location on chromosome 14q32.1	92351933-362346	92337866-340248	92585960-597606	92535005-543340	92560866-566668	92277799-296886	92436655-443399	92792832-802543	924161680-426404	92460862-471956			
Mouse clade 'a' serpins													
Revised name	Common name	Ensembl gene location on chromosome 12F1	Ensembl gene ID (ENSMUSGXXXXXX#)										
Serpina1a	Spit-1/AIT1		21077	61%	53%	44%	45%	42%	41%	41%	33%	42%	40%
Serpina1b	Spit-2/AIT2		21077	59%	52%	42%	43%	39%	40%	41%	31%	42%	39%
Serpina1c	Spit-3/AIT3		21077	59%	53%	42%	42%	40%	40%	41%	32%	43%	40%
Serpina1d	Spit-4/AIT4		21077	61%	53%	43%	45%	41%	42%	40%	32%	42%	40%
Serpina1e	Spit-5/AIT5	98142801-147191	21083	60%	53%	45%	46%	43%	41%	42%	33%	44%	38%
	"ATR"	97902032-906318	21081	36%	38%	32%	28%	32%	30%	29%	25%	33%	28%
Serpina3a	Unknown (1)	98316084-325141	41536	37%	33%	51%	34%	39%	33%	35%	31%	38%	29%
Serpina3b	6A1	98332477-342159	21085	40%	39%	51%	39%	40%	41%	41%	28%	45%	32%
Serpina3c	KALBP/1A1	98350262-357496	41513	45%	40%	59%	45%	43%	42%	45%	31%	45%	38%
Serpina3d	2B1	98377106-381727	41504	35%	37%	53%	36%	40%	38%	40%	29%	39%	32%
Serpina3e	2B2 (b)	98393324-396224	41497	38%	35%	58%	40%	40%	41%	42%	29%	41%	32%
Serpina3f	2A1	98420696-424286	41489	45%	43%	59%	44%	41%	43%	45%	33%	45%	38%
Serpina3g	2A2	98440068-448739	41481	41%	41%	53%	40%	39%	39%	40%	28%	41%	37%
Serpina3h	6C28	98457237-460419	41474	41%	38%	60%	41%	44%	40%	42%	33%	40%	35%
Serpina3i	2B2	98471856-475389	21073	46%	42%	59%	43%	45%	43%	47%	33%	44%	36%
Serpina3j	Unknown (2)	98524359-529639	41458	39%	37%	53%	37%	38%	38%	37%	29%	42%	34%
Serpina3k	MMSpi2	98548297-555927	41449	43%	39%	58%	43%	44%	44%	40%	34%	46%	38%
Serpina3l	3E2	98567863-573327	41443	39%	40%	51%	40%	38%	38%	39%	27%	40%	35%
Serpina3m	3E46	98603903-608751	21092	43%	41%	61%	45%	45%	43%	43%	34%	46%	38%
Serpina3n	Spit2/Eb.4	98621556-629158	21091	43%	40%	61%	44%	45%	44%	43%	31%	43%	38%
	"KALps"	98281878-290337	41558	39%	36%	41%	37%	41%	48%	47%	29%	40%	35%
Serpina5	PCI	98305037-309494	41550	43%	41%	46%	38%	61%	38%	43%	31%	41%	36%
Serpina6	CBG	97859489-869385	41595	39%	38%	41%	36%	34%	57%	36%	24%	35%	33%
Serpina9	Centerin	98201091-217066	21084	38%	33%	42%	44%	43%	42%	64%	31%	45%	34%
Serpina10	PZI	97830594-844279	41603	28%	28%	26%	23%	25%	24%	25%	65%	23%	25%
Serpina12	Novel	98232183-248067	41567	39%	30%	37%	35%	38%	35%	32%	31%	37%	61%

^a The mouse serpins are compared with the human clade A serpins by percentage amino acid identity, which was obtained using pairwise BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>). Highest homology is indicated by a shaded box. Where known, the genomic location and Ensembl gene ID of the mouse serpin is indicated (as derived from http://www.ensembl.org/Mus_musculus/). The genomic locations of the human serpins are also indicated (as derived from http://www.ensembl.org/Homo_sapiens/). PCI, Protein C inhibitor; ATR, α_1 -antitrypsin-related protein; AIAT, α_1 -antitrypsin; AACT, α_1 -antichymotrypsin; CBG, corticosteroid-binding globulin; KAL, kallikrein inhibitor; PZI, protein Z-dependent protease inhibitor. The mouse "ATR" is in quotation marks because, although it is described as the human ATR homolog in the genome, it has very little homology with it. "KALps" appears to be a pseudogene with homology to human *SERPINA4* (KAL) gene. The new name of each mouse serpin, as approved by the MGNC, is shown. New serpins identified in the human genome are shown and their names, as approved by the HGNC, are underlined. See Fig. 3 for organization of the genes.

clade "a" serpin cluster, namely hypervariability of a defined domain with conservation of surrounding regions encoding structural elements, has been observed elsewhere. In the highly polymorphic MHC genes, variability allows for recognition of a vast array of antigenic peptides derived from pathogens. Variation is largely restricted to the peptide-binding region and results from a high ratio of nonsynonymous to synonymous substitutions [44].

Additional examples of gene expansion may come to light with the recent completion of the mouse genome. Mouse chromosome 16 (Mmu16), for example, contains 731 genes; 509 of these represent human homologs whereas 222 (30%) do not. Of these, 44 (8%) are gene paralogs having arisen from duplication events [1,2]. A recent example of gene expansion involves the C2H2-type zinc-finger genes, which in the human are dispersed among 11 clusters on chromosome 19 and encode putative transcription factors, olfactory receptors, and putative pheromone receptors

[45]. Many of the related mouse clusters, however, contain a very different complement of genes, and these appear to have arisen from duplication and independent selection from the original gene. The resulting genes appear to be functional and have nonredundant adaptive functions in addition, and complementary to, the ancestral gene. In one extreme example of these, a single zinc-finger gene on the human chromosome is represented by a cluster of 12 genes in the mouse. In the case of the putative pheromone receptors only one functional gene has been identified on human chromosome 19 as opposed to at least 17 genes at the syntenic locus in the mouse. Differences observed in the mouse olfactory and putative pheromone receptor genes, when compared with those in the human, are suggested to account for differences in the way mice and humans taste food and attract sex partners [1,2,45].

Besides the antitrypsin and antichymotrypsin genes described earlier, the intracellular (clade B) serpins have also

Mm

Hsa

Fig. 3
with
and S
geno
(base
mark
four
inver
in the
pseud
relate
appro
(http:

unde
chro
cellu
mosc
cons
and c
are c
exter
thus
acter
indee
paral
of the
Sp
gence
expos
prote

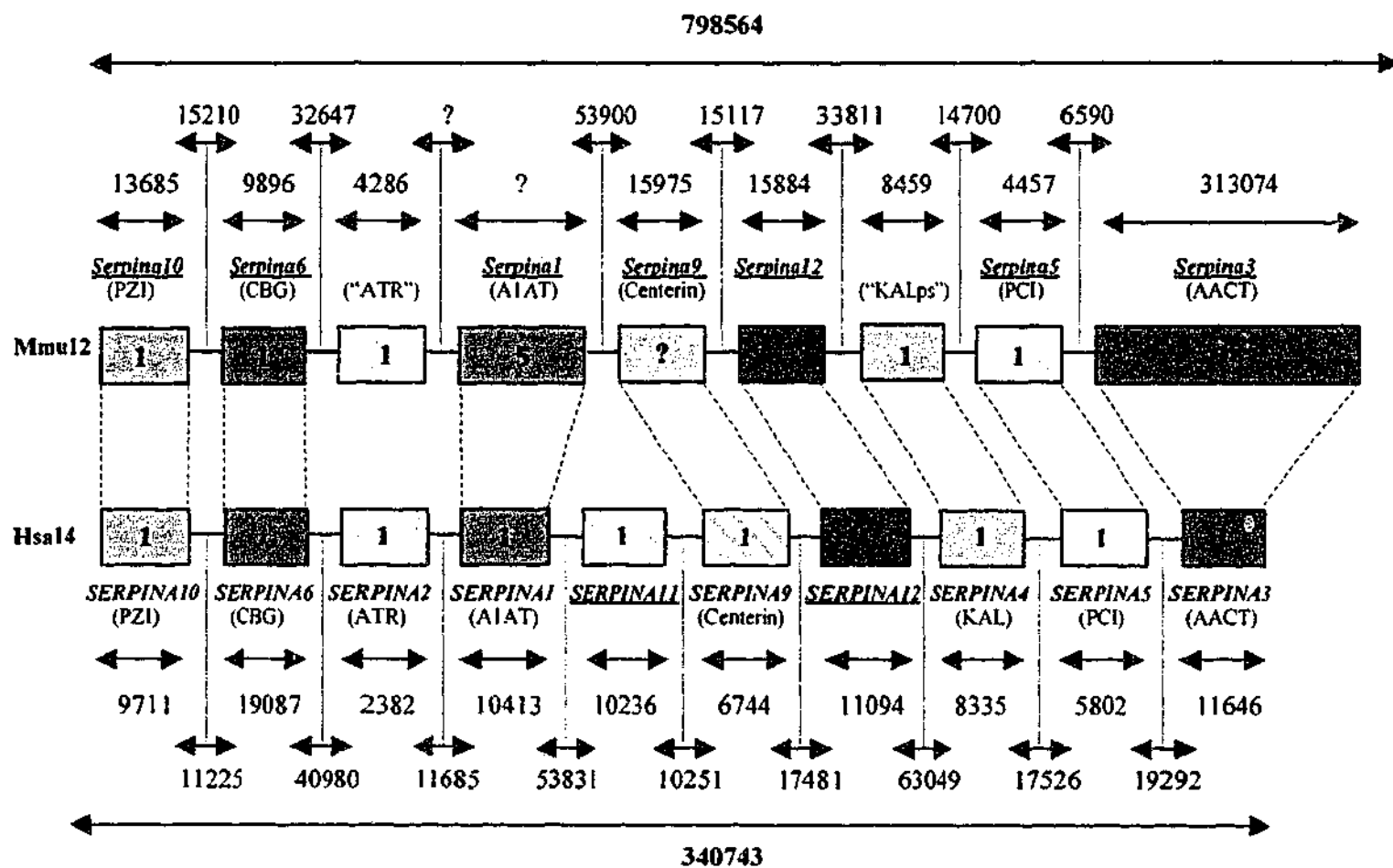


Fig. 3. Organization of mouse chromosome 12F1 containing the clade 'a' serpins versus the syntenic region on human chromosome 14q32.1. The number within each box indicates the number of genes within the locus. A question mark within a box indicates a gene fragment. The hatched boxes (*SERPINA11* and *SERPINA9*) are represented in the ENSEMBL database as one gene (ENSG00000170054) with two alternative transcripts but closer examination of the genomic sequence using the gene finding NIX tool (<http://www.hgmp.mrc.ac.uk/NIX/>) indicates two separate genes (as shown here). The size of the gene/s (base pairs) as well as the intergenic sequence is indicated below (Hsa) and above (Mmu). The length of the entire cluster of genes is also shown. The question mark for the *Serpina1* locus in the mouse indicates an area with poor sequence and only *Spi1-5* (*Serpina1e*) has been confirmed here. However, the remaining four *Spi1* genes (see text) have been identified within an unassigned chromosome fragment which is likely to belong in this locus. Murine "ATR" is in inverted commas as, although it has been identified in the databases as the homolog, it shows little identity with the human gene (Table 3). The murine gene in the corresponding position to human *SERPINA4* (Kallikrein inhibitor) has 57% amino acid identity with *SERPINA4* (Table 3) but appears to be a pseudogene and is described as "KALps". PZI = protein-Z dependent protease inhibitor, CBG = corticosteroid binding globulin, ATR = α_1 -antitrypsin related protein, AIAT = α_1 -antitrypsin, KAL = Kallikrein inhibitor, PCI = protein C inhibitor and AACT = α_1 -antichymotrypsin. New names, as approved by the MGNC and the HGNC, are underlined. Data for this figure was derived from the publicly available mouse and human genomes (<http://www.ensembl.org/>).

undergone gene duplication in the mouse [46,47]. Mouse chromosome 13 encompasses at least 15 genes of the intracellular gene family compared to 3 genes on human chromosome 6p25. Once again the mouse genes demonstrate considerable variation in the RCL, whereas the proximal and distal hinge motifs, critical for anti-proteinase function, are conserved. The RCLs of these genes also vary to the extent that they are likely to inhibit different proteases and thus demonstrate very different functional activities. Characterization of the mouse homologs of MNE1 (*SERPINB1*) indeed show altered binding specificity of at least one of the paralogs as well as alternate expression patterns of all four of the gene members [47].

Speculation on the driving force behind the rapid divergence of the mouse serpins has focused on the potential exposure of mice to a large range of pathogen-associated proteases [33,36]. The interspecies and intraspecies varia-

tion in the α_1 PI gene repertoire argues against this theory. It seems unrealistic to propose that *M. caroli*, which expresses only one form of α_1 PI, has been exposed to substantially fewer pathogens than *M. domesticus*, for example, which express up to five members. The evolutionary pressures and ultimate benefits of these gene duplications remain to be seen but, quite apart from the interesting evolutionary issues, they reinforce the proposition that mice may not always be an ideal model for human biology.

Abbreviations

RCL; reactive center loop, α_1 PI; α_1 -protease inhibitor, PCI; protein C inhibitor, ATR; α_1 -antitrypsin related protein, AIAT; α_1 -antitrypsin, AACT; α_1 -antichymotrypsin, CBG; corticosteroid binding globulin, KAL; kallikrein in-

hibitor and PZI; protein-Z dependent protease inhibitor, MGNC; Mouse Gene Nomenclature Committee, HGNC; HUGO Gene Nomenclature Committee here were derived from C57Bl/6J mice, and Inglis and Hill obtained their sequences from the inbred mouse strain 129/St. The gene described as 2B2 (b) is novel but is a truncated form of 2B2.

Acknowledgments

We thank Phillip Bird, James Whisstock, and James Irving (Department of Molecular Biology and Biochemistry, Monash University) for helpful discussions. P. C. is a Wellcome Senior Research Fellow in Medical Science.

References

- [1] N.G. Copeland, N.A. Jenkins, S.J. O'Brien, Genomics, Mmu 16—comparative genomic highlights, *Science* 296 (2002) 1617–1618.
- [2] R.J. Mural, et al., A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome, *Science* 296 (2002) 1661–1671.
- [3] G.A. Silverman, et al., The serpins are an expanding superfamily of structurally similar but functionally diverse proteins: evolution, mechanism of inhibition, novel functions, and a revised nomenclature, *J. Biol. Chem.* 276 (2001) 33293–33296.
- [4] R.E. Hill, P.H. Shaw, P.A. Boyd, H. Baumann, N.D. Hastie, Plasma protease inhibitors in mouse and man: divergence within the reactive centre regions, *Nature* 311 (1984) 175–177.
- [5] R.L. Goodwin, K.W. Barbour, F.G. Berger, Expression of the α 1-proteinase inhibitor gene family during evolution of the genus, *Mus*, *Mol. Biol. Evol.* 14 (1997) 420–427.
- [6] J. Potempa, E. Korzus, J. Travis, The serpin superfamily of proteinase inhibitors structure, function, and regulation, *J. Biol. Chem.* 269 (1994) 15957–15960.
- [7] J.A. Irving, R.N. Pike, A.M. Lesk, J.C. Whisstock, Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function, *Genome. Res.* 10 (2000) 1845–1864.
- [8] J.A. Irving, et al., Serpins in prokaryotes, *Mol. Biol. Evol.* 19 (2002) 1881–1890.
- [9] W.R. Atchley, T. Lokot, K. Wollenberg, A. Dress, H. Ragg, Phylogenetic analyses of amino acid variation in the serpin proteins, *Mol. Biol. Evol.* 18 (2001) 1502–1511.
- [10] H. Ragg, T. Lokot, P.B. Kamp, W.R. Atchley, A. Dress, Vertebrate serpins: construction of a conflict-free phylogeny by combining exon-intron and diagnostic site analyses, *Mol. Biol. Evol.* 18 (2001) 577–584.
- [11] R.W. Carrell, D.A. Lomas, α 1-antitrypsin deficiency—a model for conformational diseases, *N. Engl. J. Med.* 346 (2002) 45–53.
- [12] F. Borriello, K.S. Krauter, Multiple murine α 1-protease inhibitor genes show unusual evolutionary divergence, *Proc. Natl. Acad. Sci. USA* 88 (1991) 9417–9421.
- [13] C. Rheaume, R.L. Goodwin, J.J. Latimer, H. Baumann, F.G. Berger, Evolution of murine α 1-proteinase inhibitors: gene amplification and reactive center divergence, *J. Mol. Evol.* 38 (1994) 121–131.
- [14] R.L. Goodwin, H. Baumann, F.G. Berger, Patterns of divergence during evolution of α 1-proteinase inhibitors in mammals, *Mol. Biol. Evol.* 13 (1996) 346–358.
- [15] N.R. Matheson, H.L. Gibson, R.A. Hallewell, P.J. Barr, J. Travis, Recombinant DNA-derived forms of human α 1-proteinase inhibitor. Studies on the alanine 358 and cysteine 358 substituted mutants, *J. Biol. Chem.* 261 (1986) 10404–10409.
- [16] T.W. Stief, N. Heimburger, Inactivation of serine proteinase inhibitors (serpins) in human plasma by reactive oxidants, *Biol. Chem. Hoppe. Seyler* 369 (1988) 1337–1342.
- [17] C. Taggart, et al., Oxidation of either methionine 351 or methionine 358 in α 1-antitrypsin causes loss of anti-neutrophil elastase activity, *J. Biol. Chem.* 275 (2000) 27258–27265.
- [18] T. Ohta, On hypervariability at the reactive center of proteolytic enzymes and their inhibitors, *J. Mol. Evol.* 39 (1994) 614–619.
- [19] T. Paterson, S. Moore, The expression and characterization of five recombinant murine α 1-protease inhibitor proteins, *Biochem. Biophys. Res. Commun.* 219 (1996) 64–69.
- [20] K.W. Barbour, et al., Functional diversification during evolution of the murine α 1-proteinase inhibitor family: role of the hypervariable reactive center loop, *Mol. Biol. Evol.* 19 (2002) 718–727.
- [21] I. Laursen, A.E. Lykkesfeldt, Purification and characterization of an α 1-antichymotrypsin-like 66 kDa protein from the human breast cancer cell line, MCF-7., *Biochim. Biophys. Acta* 1121 (1992) 119–129.
- [22] M. Higashiyama, et al., α 1-Antichymotrypsin expression in lung adenocarcinoma and its possible association with tumor progression, *Cancer* 76 (1995) 1368–1376.
- [23] W. Poller, et al., A leucine-to-proline substitution causes a defective α 1-antichymotrypsin allele associated with familial obstructive lung disease, *Genomics* 17 (1993) 740–743.
- [24] J.P. Faber, et al., The molecular basis of α 1-antichymotrypsin deficiency in a heterozygote with liver and lung disease, *J. Hepatol.* 18 (1993) 313–321.
- [25] S. Takada, et al., Incorporation of α 1-antichymotrypsin into human stomach adenocarcinoma cell nuclei and inhibition of DNA primase activity, *Tokai. J. Exp. Clin. Med.* 13 (1988) 321–327.
- [26] M. Tsuda, M. Masuyama, T. Katsunuma, Inhibition of human DNA polymerase α by α 1-antichymotrypsin, *Cancer Res.* 46 (1986) 6139–6142.
- [27] S. Takada, et al., Incorporation of α 1-antichymotrypsin into carcinoma cell nuclei of human stomach adenocarcinoma transplanted into nude mice, *Cancer Res.* 46 (1986) 3688–3691.
- [28] N. Naidoo, B.S. Cooperman, Z.M. Wang, X.Z. Liu, H. Rubin, Identification of lysines within α 1-antichymotrypsin important for DNA binding. An unusual combination of DNA-binding elements, *J. Biol. Chem.* 270 (1995) 14548–14555.
- [29] S.D. Mikolajczyk, L.S. Marks, A.W. Partin, H.G. Rittenhouse, Free prostate-specific antigen in serum is becoming more complex, *Urology* 59 (2002) 797–802.
- [30] C. Stephan, et al., Prostate-specific antigen, its molecular forms, and other kallikrein markers for detection of prostate cancer, *Urology* 59 (2002) 2–8.
- [31] C.R. Abraham, Reactive astrocytes and α 1-antichymotrypsin in Alzheimer's disease, *Neurobiol. Aging* 22 (2001) 931–936.
- [32] R.E. Hill, P.H. Shaw, R.K. Barth, N.D. Hastie, A genetic locus closely linked to a protease inhibitor gene complex controls the level of multiple RNA transcripts, *Mol. Cell. Biol.* 5 (1985) 2114–2122.
- [33] J.D. Inglis, R.E. Hill, The murine Spi-2 proteinase inhibitor locus: a multigene family with a hypervariable reactive site domain, *EMBO J.* 10 (1991) 255–261.
- [34] K.X. Chai, J. Chao, L. Chao, Molecular cloning and sequence analysis of the mouse kallikrein-binding protein gene, *Biochim. Biophys. Acta* 1129 (1991) 127–130.
- [35] K.X. Chai, et al., Molecular cloning and expression of rat kallistatin gene, *Biochim. Biophys. Acta Gene Struct. Expr.* 1353 (1997) 277–286.
- [36] R.E. Hill, N.D. Hastie, Accelerated evolution in the reactive centre regions of serine protease inhibitors, *Nature* 326 (1987) 96–99.
- [37] I.N. Hampson, et al., Identification of a serpin specifically expressed in multipotent and bipotent hematopoietic progenitor cells and in activated T cells, *Blood* 89 (1997) 108–118.

- [38] A.V. Terskikh, et al., From hematopoiesis to neuropoiesis: evidence of overlapping genetic programs, *Proc. Natl. Acad. Sci. USA* 98 (2001) 7934–7939.
- [39] L. Hampson, et al., A minimal serpin promoter with high activity in haemopoietic progenitors and activated T cells, *Hematol. J.* 2 (2001) 150–160.
- [40] C.S. Chiang, A. Stalder, A. Samimi, I.L. Campbell, Reactive gliosis as a consequence of interleukin-6 expression in the brain: studies in transgenic mice, *Dev. Neurosci.* 16 (1994) 212–221.
- [41] F. Licastro, et al., A role for apoE in regulating the levels of α 1-antichymotrypsin in the aging mouse brain and in Alzheimer's disease, *Am. J. Pathol.* 155 (1999) 869–875.
- [42] J.D. Inglis, M. Lee, D.R. Davidson, R.E. Hill, Isolation of two cDNAs encoding novel antichymotrypsin-like proteins in a murine chondrocytic cell line, *Gene* 106 (1991) 213–220.
- [43] J.K. Frazer, et al., Identification of centerin: a novel human germinal center B cell-restricted serpin, *Eur. J. Immunol.* 30 (2000) 3039–3048.
- [44] J. Trowsdale, The gentle art of gene arrangement: the meaning of gene clusters, *Genome Biol.* 3 (2002).
- [45] P. Dehal, et al., Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution, *Science* 293 (2001) 104–111.
- [46] D. Kaiserman, et al., Comparison of human chromosome 6p25 with mouse chromosome 13 reveals a greatly expanded ov-serpin gene repertoire in the mouse, *Genomics* 79 (2002) 349–362.
- [47] C. Benarafa, J. Cooley, W. Zeng, P.I. Bird, E. Remold-O'Donnell, Characterization of four murine homologs of the human ov-serpin monocyte neutrophil elastase inhibitor MNEI (SERPINB1), *J. Biol. Chem.* 277 (2002) 42028–42033.
- [48] P.R. Elliott, X.Y. Pei, T.R. Dafforn, D.A. Lomas, Topography of a 2.0 Å structure of α 1-antitrypsin reveals targets for rational drug design to prevent conformational disease, *Protein Sci.* 9 (2000) 1274–1281.
- [49] J.A. Huntington, R.J. Read, R.W. Carrell, Structure of a serpin-protease complex shows inhibition by deformation, *Nature* 407 (2000) 923–926.

Down-Regulation of the Serine-Protease Inhibitors, Serpina1 and Serpina3, in the Bone Marrow During Hematopoietic Progenitor Cell Mobilization Results in Increased Neutrophil Serine-Protease Activity

Ingrid G. Winkler,¹ Jean Hendy,¹ Paul Coughlin,² Anita Horvath,² and Jean-Pierre Lévesque¹

¹Adhesive Interactions and Cell Trafficking laboratory, Peter MacCallum Cancer Centre, East Melbourne, Victoria 3002, Australia, and ²Department of Medicine, Monash University, Box Hill Hospital, Box Hill, Victoria 3128, Australia.

Character count: Text and figures 69,584 (Text alone: 54,584)

Corresponding authors:

Dr Jean-Pierre Lévesque or Dr Ingrid G Winkler

Adhesive Interactions and Cell Trafficking laboratory

Peter MacCallum Cancer Centre

Locked Bag 1

A'Beckett Street

Melbourne, VIC 8006

Australia

Tel: +

Fax: +61 3 9656 3738

E-mail:

Keywords: α 1-antitrypsin, α 1-antichymotrypsin, granulocyte colony-stimulating factor, hematopoiesis, neutrophil elastase

Abstract

We report that the activity of neutrophil serine-proteases within the bone marrow is regulated by the expression of naturally-occurring inhibitors (serpina1 and serpina3) produced locally within this tissue. We found that serpina1 and serpina3 are transcribed in the bone marrow by many different hematopoietic cell populations and that a strong reduction in expression occurs both at the protein and mRNA levels during mobilization induced by granulocyte colony-stimulating factor or chemotherapy, whereas expression of neutrophil elastase and cathepsin G remains unaltered. This decreased expression was restricted to the bone marrow as serpina1 expression was maintained in the liver, leading to no change in plasma concentrations during mobilization. The down-regulation of serpina1 and serpina3 during mobilization, shifts the balance between serine-proteases and their inhibitors, allowing the accumulation of active neutrophil serine-proteases in bone marrow extravascular fluids that cleave and inactivate molecules essential to the retention of hematopoietic progenitor cells within the bone marrow. These data reveal an unexpected role for serpina1 and serpina3 in regulating the bone marrow hematopoietic microenvironment as well as influencing the migratory behavior of hematopoietic precursors.

Introduction

Hemopoietic stem and progenitor cells (HPC) are responsible for the renewal of all mature blood cells. In adult mammals, the majority of HPC reside in the bone marrow (BM) whereas their differentiated progeny migrate to the circulation. Transient increases in the number of HPC circulating in the peripheral blood (mobilization) occur in response to a wide variety of stimuli including strenuous physical exercise, myelosuppressive chemotherapy, polyanions, chemokines, and hematopoietic growth factors (1).

Mobilized HPC are now the favored source of transplantable cells to reconstitute hematopoiesis following high-dose chemotherapy. Currently, the agent most commonly used to elicit HPC mobilization is granulocyte-colony stimulating factor (G-CSF) used alone or in combination with myelosuppressive chemotherapy (2, 3). The administration of G-CSF induces a 10- to 100- fold increase in the level of circulating HPCs in both humans and mice. G-CSF-induced mobilization is time- and dose-dependent, involving a rapid neutrophilia (evident within hours) and a gradual increase in HPC numbers in the blood peaking between 4-7 days of G-CSF administration. Mobilization with chemotherapeutic agents such as cyclophosphamide (CY) occurs during the recovery phase following the chemotherapy-induced neutropenia, that is, days 6-8 in mice, and days 10-14 in humans. Although mobilized HPC collected from the peripheral blood are extensively used to rescue hematopoiesis in patients undergo high-dose myeloablative chemotherapy, the exact molecular mechanisms responsible for the mobilization of HPC from the BM into the peripheral blood remain unclear.

Essential for the retention of HPC in the BM are adhesive and chemotactic interactions. Particularly important are 1) the adhesive interaction between the vascular cell adhesion molecule VCAM-1 (CD106) expressed by the BM stroma with its counter receptor integrin $\alpha 4 \beta 1$ (VLA-4) expressed by HPC, and 2) HPC chemotaxis due to binding of the chemokine CXCL12 (SDF-1) produced by the BM stroma, to its cognate receptor CXCR4 (CD184) expressed at the surface of HPC. Blocking either of these interactions; the VCAM-1/ $\alpha 4 \beta 1$ adhesive interaction (4-7) or the CXCL12/CXCR4 chemotactic interaction (8-10), by means of antibodies, antagonists or tissue-specific gene-targeted deletion has been shown to result in mobilization of HPC *in vivo*.

Our group has shown that HPC mobilization induced by G-CSF or CY coincides with the accumulation of high concentrations of active neutrophil serine-proteases within the BM. The predominant proteases were identified as neutrophil elastase (NE) and cathepsin G (CG),

both released by neutrophils upon activation (11, 12). Once released into the BM extracellular fluid, these proteases can disrupt locally the two important mechanisms by which HPC home to and remain within the BM: 1) adhesion to the BM stroma (by cleaving the extracellular domain of VCAM-1) (11, 12), and 2) chemotaxis of HPC by degrading and inactivating the chemokine CXCL12 and cleaving the 1st extracellular domain of CXCR4 expressed by human HPC (13). As previous studies have shown, inactivation of either is sufficient to induce mobilization. However the mechanism to explain how high concentrations of active neutrophil serine-proteases can accumulate in the BM during mobilization remains unclear.

Neutrophil serine-protease activity is controlled *in vivo* by naturally occurring serine-protease inhibitors or serpins (14, 15). Serpins function like mousetraps with a protruding reactive center loop in a metastable conformation. Cleavage of this reactive center loop by a specific protease triggers a spring-like mechanism, which entraps the protease within the serpin structure blocking access of potential substrates to the proteolytic center of the protease (16-18). Once formed, the inactive serpin-protease complex is rapidly cleared from the body. Although serpin structure is very conserved, the reactive center loop sequence is highly variable between different serpins allowing the targeting of specific proteases. Two groups of serpins located on chromosome 14 in humans and chromosome 12 in mice (referred to as 'Clade A' serpins) can inactivate neutrophil serine-proteases in blood and tissues. These are SERPINA1 (also known as α 1-antitrypsin: AAT) and SERPINA3 (α 1-antichymotrypsin: ACT) (19, 20). Although humans express a single copy of each *SERPINA1* and *SERPINA3* gene, in the house mouse *Mus musculus* the *serpina1* gene has replicated 5 times (*serpina1a-e*) (21-23) whereas the *serpina3* gene has replicated 14 times (*serpina3a-n*) (23, 24).

In this study we report that in the mouse, *serpina1* and *serpina3* are produced in steady-state BM in sufficient concentrations to inhibit local serine-proteases. During mobilization induced by either G-CSF or CY, the levels of *serpina1* and *serpina3* in the BM dramatically drop boosting the levels of active neutrophil serine-proteases with concomitant cleavage and inactivation of molecules essential for the retention of HPC within the BM.

Materials and Methods

Mobilization of mice, tissue collection and clonogenic assay. Male 8-10 week old BALB/c mice (4-6 mice per group) were injected twice daily (days 0-5) with 250 μ g/kg/day recombinant human G-CSF (Neupogen Filgastrim; Amgen, Thousand Oaks, CA) or an equal volume of saline (control mice). Mice were harvested at days 2, 4, 6 following G-CSF injection or allowed to rest and harvested at days 8 and 10. Cyclophosphamide treated mice received a single intra-peritoneal injection of 200mg/kg CY (Cycloblastin, Pharmacia and Upjohn, Rydalmere, Australia) on day 0.

Peripheral blood was collected in EDTA. Plasma was separated and red cells lysed with 10mM NaHCO₃ pH 7.4, 150mM NH₄Cl. Nucleated cells were plated in double-layer nutrient agar clonogenic cultures (1,500 cells/dish) in the presence human IL-1 α , mouse IL-3, mouse CSF-1 and rat KIT ligand as previously described (11). Colonies were scored after a two week culture at 37°C in the presence of 5% O₂, 10% CO₂ and 85% N₂.

The BM content of 1 femur from each mouse was flushed into 1 mL PBS on ice. Following centrifugation for 5 min at 400xg, the supernatant fraction (BM extracellular fluid) was removed and used directly in western blots or for the digestion of recombinant human VCAM-1. The BM cell pellet, as well as the liver and spleen were used for RNA extraction, immunostaining and cell sorting, or to generate frozen sections.

Flow cytometry and cell sorting. Bone marrow and spleen cells were stained using 1 μ g/mL of directly conjugated mAb (Becton Dickinson, San Jose, CA) for myeloid cells (Gr-1-FITC (cloneRB6-8C5) and CD11b-PE (clone M1/70)), for B-lymphocytes (B220-PE (clone RA3-6B2)), for T-lymphocytes (mix of CD3-FITC (clone 17A2), CD4-FITC (clone H129.19), CD5-FITC (clone 53-7.3) and CD8a-FITC (clone 53-6.7)), and for total hematopoietic cells (CD45-PE (clone 30-F11)). Cells positive for each category were sorted and RNA extracted as described below.

RNA extraction and quantitative real-time RT-PCR. Total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA). Following DNase treatment and reverse transcription using random hexamers, quantitative real-time PCR with SYBR green (ABI systems, Foster City, CA) was performed following manufacturer's instructions using the oligonucleotide combinations shown in Table 1. RNA levels were standardized by parallel RT-PCRs using primers to two different house-keeping genes, vimentin (a cytoskeletal protein, forward primer 5'-CACCCTGCAGTCATTGAGACA-3' and reverse 5'-

GATTCCACTTTCCGTTCAAGGT-3') and β_2 -microglobulin (forward primer 5'-TTCACCCCCACTGAGACTGAT and reverse 5'-GTCTTGGGCTCGGCCATA-3'). Additional RT-PCRs for the serine proteases NE (forward primer 5'-ACCCTCATTGCCAGGAACTTC-3' and reverse 5'-CCTGCACTGACCGGAAATTT-3'), CG (forward primer 5'-AGGCAGGGAAGATCATTGGA-3' and reverse 5'-TGGATCAGAAGAAATGCCATGT-3') as well as VCAM-1 (forward primer 5'-CTGGGAAGCTGGAACGAAGTA-3' and reverse 5'-GCCACTGAATTGAATCTCTGGAT-3') and CXCR4 (forward primer 5'-GGCTGACTGGTACTTTGGGAAA-3' and reverse 5'-CCGGTCCAGGCTGATGAA-3') were also performed. A PCR from each sample prior to reverse transcription was also performed to confirm the absence of contaminating genomic DNA.

VCAM-1 cleavage and immunoblotting. BM extracellular fluids collected from three mice at day 4 of G-CSF-induced mobilization were pooled and mixed with an equal volume (10 μ L) of either pooled BM fluids from three non-mobilized mice or with PBS. Where indicated, protease inhibitors were added to the following concentrations: 1mg/mL human α_1 -antitrypsin (Sigma Chemicals, St Louis, MO), 1 μ M BB-94/Batimastat (British Biotech Pharmaceuticals, Oxford, UK), 1mM O-phenantroline. Following a 20 min incubation at room temperature, 10ng recombinant human VCAM-1 (98 kDa glycosylated extracellular domain, R&D Systems, Minneapolis, MN) was added and the mixture re-incubated at 37°C for 20 min. An equal volume of loading buffer (125mM Tris-HCl pH6.8, 20% glycerol, 2% SDS) was then added and samples boiled for 3 min, before separation by electrophoresis on 10% SDS-PAGE and transfer onto nitrocellulose membrane. Membranes were immunoblotted using a goat anti-human VCAM-1 antibody as previously described (11, 12).

Immunoblotting for serpin1 and α_2 -macroglobulin. 10 μ L aliquots of pooled BM extracellular fluids from 3 mice at indicated time-points of G-CSF or saline administration were mixed with an equal volume of loading buffer with 10mM dithiothreitol and boiled for 3 min. Following separation by electrophoresis on SDS-PAGE (10% for serpin1; 8% for α_2 -macroglobulin) and transfer onto a nitrocellulose membrane, membranes were blocked with PBS containing 0.05% Tween-20 (PBST) and 5% dry skim milk and incubated for 1 hour with 1/3,000 dilution of either a purified rabbit anti-human α_1 -antitrypsin Ab (Sigma Chemicals) cross-reacting with mouse serpin1 or a rabbit anti-human α_2 -macroglobulin Ab cross-reacting with mouse α_2 -macroglobulin (Sigma Chemicals). Following extensive washes with PBST and incubation with a 1/10,000 dilution of horseradish peroxidase-

conjugated donkey F(ab)₂ anti-rabbit IgG (Jackson ImmunoResearch, West Grove, PA), blots were revealed by enhanced chemoluminescence.

Immunohistochemistry. Frozen sections of mouse liver and spleen were air dried, then fixed for 30 min in methanol containing 0.3% H₂O₂ to destroy endogenous peroxidases. Slides were then blocked in 4X SSC containing 5% BSA, 5% skim milk powder and 0.05% Triton X-100 for 2 hours, before incubation with either 25 µg/ml purified rabbit anti-human α 1-antitrypsin or non-immune purified rabbit IgG in PBST with 5% skim milk and 5% BSA overnight at 4°C. Following extensive washes, slides were incubated with a 1/500 dilution of horseradish peroxidase-conjugated donkey F(ab)₂ anti-rabbit IgG with minimal cross-reactivity to mouse proteins for 2 hours at room temperature, then washed. Staining was revealed using 0.5 mg/mL 3,3' diaminobenzidine in 50 mM Tris-HCl pH 7.4, 0.3% H₂O₂ before counterstaining with hematoxylin and mounting with Aquamount (BDH).

For BM wax-embedded sections, femoral BM was fixed by perfusing 0.05% glutaraldehyde, 2% paraformaldehyde into the descending femoral aorta and femurs processed exactly as described by Nilsson et al (25). Longitudinal 3.5 µm sections were cut, dewaxed in xylene then rehydrated using progressive baths from ethanol to water. Following antigen retrieval (12 min boiling in 10mM sodium citrate pH 6.0), endogenous peroxidase elimination and antigen labelling were performed as described above.

Statistics. Levels of significance were calculated using the non-parametric Mann-Whitney test.

Results

Serpins are present in steady-state BM extracellular fluids. We have previously reported a dramatic increase in active serine-protease levels in the BM at the peak of mobilization induced either by G-CSF or CY (11-13). This could be due to i) a corresponding increase in protease production/release, or ii) a drop in the level of inhibitors which control the activity of these proteases. To determine whether naturally-occurring protease inhibitors are involved in the regulation of protease activity in the BM, recombinant VCAM-1 was incubated with pooled BM fluids from three saline-injected mice (Figure 1, lane 1), three G-CSF day 4 mice (lane 3), or a mixture of the two (lane 2). As expected from our previous work, no cleavage of VCAM-1 was observed when incubated with BM fluids from saline-injected non-mobilized mice, while cleavage of VCAM-1 did occur when incubated with BM

fluids from G-CSF day 4 mice (11, 12). Interestingly, when both BM fluids were mixed together (lane 2) cleavage was also prevented. This result demonstrates that naturally-occurring protease inhibitors are present in the BM fluids from steady-state (saline-injected) mice in sufficient quantities to completely inhibit the proteolytic activity of BM fluids from mice mobilized with G-CSF.

When these same pooled BM fluids from G-CSF day 4 mice were incubated with purified human SERPINA1 (α 1-antitrypsin), complete inhibition of VCAM-1 cleavage was observed, suggesting that murine steady-state BM may contain similar serine protease inhibitors (Figure 1, compare lanes 1, 3 and 4). Inhibitors of matrix metalloproteinases (MMPs) such as BB-94 (lane 5) or O-phenanthroline (lane 6), did not prevent VCAM-1 cleavage. Together these results indicate that the proteases in G-CSF-mobilized murine BM fluids able to cleave VCAM-1 are predominately serine proteases (inhibited by SERPINA1) and not MMPs. Thus the inhibition of protease activity by steady-state BM fluids is essentially due to endogenous serpins and not to tissue inhibitors of metalloproteinases (TIMPs).

Serpinal protein concentration decreases in the BM during G-CSF-induced or CY-induced mobilization. A decrease in serpin expression at the peak of mobilization was confirmed by immunoblotting of BM fluids using a polyclonal anti-human SERPINA1 Ab that cross-reacts with murine serpin1. Immunoblotting revealed that serpin1 was present in pooled BM fluids before and after mobilization but not at the peak of mobilization on days 4–6 following G-CSF administration (Figure 2A). This decrease was restricted to the BM as no corresponding change in serpin1 concentration was observed in the blood plasma from these animals (Figure 2B). Similarly, serpin1 levels were decreased in the BM fluids from CY-treated mice between days 6–10 again corresponding with the peak of mobilization in these mice whereas blood plasma concentrations remained unchanged.

In contrast, on day 3 following CY injection, an increase in serpin1 proteins was observed. As the BM is very damaged with high red cell content at this time-point, we hypothesized that this peak in serpin1 in the BM could be due to blood infiltration across the endothelial barrier between the blood and the bone marrow extravascular compartment. This was confirmed by immunoblotting the same BM fluids as above for α 2-macroglobulin, a major plasma protein. Because of its large size (four 160 kDa subunits linked by disulfide bridges), α 2-macroglobulin cannot normally diffuse through vascular walls from the blood into tissues. While α 2-macroglobulin was readily detected in all blood samples, it was

detected in the BM extravascular compartment of mice only at day 3 following CY administration, but was absent from BM fluids collected from saline-injected or G-CSF injected mice, or from mice harvested 6 or 8 days following CY injection (Figure 2E and 2F). In summary, the spike in serpin1 protein concentration in BM extracellular fluids 3 days following CY injection is most likely due to blood infiltration rather than up-regulation of serpin production in the BM. Together, these results suggest that 1) serpin presence in the extravascular compartment of the BM is due to local secretion by BM cells and not to diffusion of blood across the endothelial barrier, and 2) serpin expression in the BM drops during HPC mobilization induced by G-CSF or CY.

Reduction of serpin1a-e mRNA levels in the BM during mobilization. To determine whether transcription of *serpin1* and *serpin3* genes was decreased in the BM at the peak of mobilization, total RNA was extracted from BM cells of 6 individual mice injected with either saline for 4 days, G-CSF for 4 days, or 8 days following a single injection of CY, and quantitative real-time RT-PCR performed. Results were standardized using primers specific for murine vimentin, a cytoskeleton protein ubiquitously expressed.

Using a primer pair amplifying transcripts from the five murine *serpin1* genes (*serpin1a-e*), we observed a dramatic decrease in *serpin1a-e* mRNA levels in the BM of mice mobilized with G-CSF (8,000 fold decrease) or CY (790 fold decrease) (Figure 3A). When primers to β 2-microglobulin (another ubiquitously expressed gene) were substituted for the pan-serpin1 primers, there was no difference between saline-, G-CSF- or CY-injected mice. These results confirm that the decrease in concentration of mouse *serpin1a-e* transcripts is specific (no changes are found with other RNAs such as β 2-microglobulin) when compared to the vimentin and that a sharp decrease in *serpin1a-e* transcripts is associated with the drop in serpin1 protein concentration in the BM (as shown by immunoblotting in Figure 2).

To confirm real-time RT-PCR results, a repeat PCR was stopped at 30 cycles for *serpin1* (mid-log phase of reaction) or 25 cycles for β 2-microglobulin, run on 8% PAGE and bands visualized with ethidium bromide (Figure 3B). After 30 cycles, *serpin1* product (140bp) was detectable in RNA from the BM of three individual saline injected mice but not in three G-CSF mobilized mice.

Real-time RT-PCR was performed from BM at different time-points of G-CSF-induced mobilization. The decrease in *serpin1* mRNA on days 2 and 4 (Figure 3C) correspond to the timing of decreased serpin1 protein in BM fluids (Figure 2) as well as the rise in colony-

forming cells mobilized into the peripheral blood (Figure 3D). Interestingly, this comparison of *serpina1* mRNA and protein levels with the timing of G-CSF-induced mobilization reveals a time lag at day 6. At this point in time, *serpina1* transcript level has rebounded and serpin protein levels are beginning to be detectable again; however high levels of active proteases remain in the BM with large numbers of CFC still in the PB.

Similar decreases in *serpina1a-e* mRNA levels were observed in the BM at the peak of CY-induced mobilization (days 6-8). Together these data indicate that a significant 2-3 log decrease in *serpina1a-e* mRNA concentration occurs in the BM at the peak of mobilization even when two completely different mobilization protocols (either G-CSF, a cytokine, or CY, a cytotoxic agent) are used.

Transcription of the neutrophil proteases NE, CG as well as VCAM-1 and CXCR4 in the BM remains unchanged during mobilization. In sharp contrast to *serpina1*, no significant change in NE or CG transcripts was observed during mobilization induced by G-CSF or CY (Figures 3D and 4). Similarly, the transcription of other molecules reported to be important for the retention of HPC in the BM such as VCAM-1 or CXCR4 remained unaltered (Figure 4). Taken together, these results support the notion that a dramatic drop in serpin expression, not an increase in protease expression, is responsible for the enhanced levels of active serine-proteases observed in mouse BM at the peak of G-CSF- and CY-induced mobilization.

*Decrease of *serpina3* mRNA in the BM at the peak of mobilization.* Using the same approach, we followed the mRNA levels of six murine *serpina3* genes which encode serpins with an N-terminal secretion sequence and thus are potentially secreted by cells (*serpina3b*, *serpina3c*, *serpina3d*, *serpina3k*, *serpina3m* and *serpina3n*) (23). A similar decrease in transcripts (between 8 and 1,400 fold) was found in the BM at the peak of mobilization induced by either G-CSF or CY (Figure 4, top panel) with the exception of *serpina3b* which was not transcribed in the BM of BALB/c mice. We also analyzed by real-time RT-PCR, transcripts of another five murine *serpina3* genes that have no secretion sequences and are thus presumed to remain cytoplasmic (*serpina3e*, *serpina3f*, *serpina3g*, *serpina3h*, *serpina3i*), the majority of which were detected in the BM of steady-state BALB/c mice (Table I). Similar to the secreted *serpina3*, the mRNA levels of these other *serpina3* genes were significantly decreased at the peak of mobilization (Fold reduction in BM transcription at day 4 of G-CSF was 85 with $p < 0.001$ for *serpina3e* and *serpina3i*, 621 with $p < 0.01$ for *serpina3g*, and 1,592 with $p < 0.01$ for *serpina3h*). *Serpina3f* was not expressed in steady-state or mobilized BM.

Decreased expression of serpin1 and serpin3 in the BM during mobilization is due to reduced mRNA levels in serpin producing cells. The exact identity of cell types expressing serpin1 and serpin3 in mouse BM has not been thoroughly investigated. We sorted BM cells from saline- and G-CSF-treated mice at day 4 based on lineage markers to determine which cell types were expressing serpin1 and serpin3 and whether serpin transcription in these cells was down-regulated during mobilization.

We first examined myeloid cells as they are the predominant cell type accumulating in the BM at the peak of mobilization (Figure 5A). Bone marrow myeloid cells were sorted on the basis of positive CD11b expression and relative maturity on the basis of Gr-1 expression (Figure 5B and 5C) with primitive-mono/myeloid cells being Gr-1^{dim} (sort gate R3), immature myeloid intermediate for Gr-1 staining (sort gate R4), and mature BM neutrophils Gr-1^{bright} (sort gate R2). With G-CSF-induced mobilization a 2-fold increase in both the proportion and total cell number of primitive Mac-1⁺Gr-1^{dim/intermediate} myeloid cells (R3 plus R4 regions) was found in the BM.

By real-time RT-PCR, serpin1a-e mRNA transcripts were greatest in BM neutrophils (Mac-1⁺ Gr-1^{bright}) with one or two log less mRNA in more primitive Gr-1^{intermediate} and Gr-1^{dim} cells respectively (Figure 5D). Interestingly, at the peak of G-CSF mobilization, a dramatic and significant transcriptional down-regulation of *serpin1a-e* genes was observed when compared to saline-injected BM in all myeloid subsets (Figure 5D). This transcriptional down-regulation was not just specific to *serpin1a-e* genes, but was also found with the secreted α_1 -anti-chymotrypsin serpin3c during mobilization (mRNA concentration was 5-fold lower in BM neutrophils and 10-fold lower in more primitive myeloid cells). However, during mobilization no change was found for other mRNAs transcribed by these cells such as the protease CG (Figure 5E). Although CG transcription is higher in more primitive myeloid cells than in mature BM neutrophils (consistent with packaging in primary granules), administration of G-CSF was not found to alter CG transcript levels in any of the three myeloid cell populations sorted from the BM.

A similar 100-1,000 fold decrease in *serpin1a-e* mRNA levels was also observed during mobilization for other sorted BM cell populations such as B220⁺ B-lymphocytes, CD4⁺ and CD8⁺ T-lymphocytes and total CD45⁺ hematopoietic cells (data not shown).

Decreases in serpin1 transcripts is restricted to the BM of mobilized mice. During mobilization, a large number of the mobilized HPC home to the spleen. We therefore queried whether the drop in serpin1 and serpin3 transcripts during mobilization was a general

phenomenon or restricted to the BM. For this purpose, we analyzed *serpina1* mRNA levels in liver, BM and spleen by real-time RT-PCR and confirmed these results at the protein level by immunohistochemistry of tissue sections.

SERPINA1 and SERPINA3 proteins are present in the blood at high concentrations (2mg/mL and 0.5 mg/mL respectively in humans). They are mainly produced by the liver, however the regulation of expression has not been exhaustively studied in the mouse. As expected, we found high levels of *serpina1a-e* mRNAs (Figure 6A, left panel) and *serpina1* protein expression (Figure 6B, left panel) in liver cells, with no change in mRNA or protein expression during G-CSF-induced mobilization. Conversely, *serpina1a-e* transcripts dropped dramatically (1,000 fold) in mouse BM during G-CSF-induced mobilization (Figure 6A) with a corresponding sharp drop in protein expression (Figure 6B, middle panel).

Interestingly, while *serpina1a-e* transcription in the spleen of mice was minimal, it increased during mobilization (Figure 6A). This finding was confirmed at the protein level by immunohistochemistry showing accumulation of cell clusters producing *serpina1* proteins within the myeloid areas of the spleen from mobilized mice but barely detectable in the spleens from non-mobilized mice. However, no increase in *serpina1a-e* mRNA was found in Gr-1⁺ cells sorted from spleens of mobilized versus non-mobilized mice (data not shown) suggesting that the increase in *serpina1* protein and mRNA levels in the spleen following the administration of G-CSF is most likely due to the migration of *serpin*-producing cells (such as myeloid cells) from the BM into the spleen. This is supported by the findings that the spleen of saline-injected mice contained $6 \pm 2\%$ Gr-1⁺ neutrophils increasing to $25 \pm 5\%$ Gr-1⁺ neutrophils following 4 days of G-CSF. When absolute numbers are considered, this corresponds to $5 \pm 2 \times 10^6$ neutrophils/spleen in non-mobilized animals compared to $38 \pm 7 \times 10^6$ neutrophils/spleen with mobilization.

RT-PCRs using primers for *serpina3c* revealed a similar phenomenon, that is a drop in mRNA levels in the BM and increased mRNA concentrations in spleen during mobilization (Figure 6A, right panel). In the liver, mobilization resulted in a 10-fold decrease in *serpina3c* transcripts, but interestingly no decrease in *serpina1* mRNA levels.

Discussion

In this report we demonstrate that the mRNA levels of five murine *serpina1* and nine *serpina3* genes are decreased specifically in the BM during HPC mobilization induced by either G-CSF or chemotherapy. Conversely, we found that the increased proteolytic activity

of NE and CG in mobilized BM (11, 12) was not associated with changes in the transcription of these proteases. Similar results were found in other inbred mouse strains such as C57BL6J, C3H/HEJ and 129Svj confirming that this phenomenon is not restricted to a particular mouse strain (data not shown). Collectively, these data support the notion that the sharp drop in concentration of *serpina1* and *serpina3* protein and mRNA levels (either due to transcriptional down-regulation and/or mRNA destabilization) is responsible for the accumulation of active NE and CG in the BM during mobilization.

Interestingly, hepatic transcription of *serpina1* is not altered, resulting in constant blood plasma levels during HPC mobilization. This would explain why there is no active NE or CG in the blood and no cleavage of VCAM-1 on the lumen of BM capillaries (11). We have previously shown that the expression of VCAM-1 in the spleen does not change during mobilization (11). In this study using immunohistochemistry as well as real-time RT-PCR we show that cells producing *serpina1* proteins accumulate in the spleen during mobilization, possibly protecting VCAM-1 from proteolytic cleavage by neutrophil serine-proteases in this tissue.

The human genome has only one copy of the *SERPINA1* and *SERPINA3* gene and their target protease specificity is well established (26). Both human *SERPINA1* and *SERPINA3* proteins can inhibit NE and CG *in vitro*, however human *SERPINA1* preferentially targets NE (with the reactive center loop P₁P₁' amino acids 'MS') (19) while *SERPINA3* preferentially targets CG (with reactive center loop P₁P₁' amino acids 'LS') (20). The reactive center loop of many of the murine *serpina1* and *serpina3* proteins have diverged (27), although several still exhibit the canonical α 1-antitrypsin reactive center (such as murine *serpina1a*, *serpina1b*, *serpina3b* and *serpina3n*) while others have the canonical α 1-antichymotrypsin reactive center (such as murine *serpina3c*, *serpina3d*, *serpina1e*, and *serpina3j*) (see Table 1). Therefore members of both the murine *serpina1* and *serpina3* families are likely to be involved in the inhibition of serine protease activity in steady-state BM as shown in Figure 1.

An interesting finding of this study, is that although the level of *serpina1a-e* transcripts dropped dramatically in the BM of mice during mobilization, hepatic mRNA levels did not alter resulting in no changes in plasma *serpina1* concentrations. Earlier studies on human *SERPINA1* transcription/expression reveal five alternative transcription starts can be used depending on the tissue in which expression occurs (with three transcription start sites in leukocytes, one in cornea and one in hepatocytes) (28) resulting in alternative splicing of

regions upstream of the translational 'ATG' start codon (28). Hence human SERPINA1 expression is controlled by both tissue specific transcriptional and pre-translational mechanisms (28). In this study we found a sharp reduction in *serpina1* transcripts specifically in BM leukocytes during mobilization while hepatic transcription remained unchanged suggesting that different tissue-specific *serpina1* transcriptional start sites are also employed in murine leukocyte and hepatic cells, although we cannot exclude mechanisms involving mRNA destabilization. In contrast to *SERPINA1*, the human *SERPINA3* gene appears to use a single promoter regardless of tissue-type in which it is expressed (28).

Both human *SERPINA1* and *SERPINA3* genes are reported to be 'acute phase proteins' as expression increases 2 to 5-fold in the liver during inflammation leading to increased levels of SERPINA1 and SERPINA3 proteins in plasma (29, 30). Whether a similar increase in the transcription of hepatic murine *serpina1* occurs during inflammation remains controversial (31, 32). A few studies have reported that several cytokines (such as IL-6, IL-1 β or TNF- α) increase *SERPINA1* transcription in human monocytes (33, 34). However, to our knowledge, no down-regulation of human or mouse SERPINA1 or SERPINA3 in response to cytokines/growth factors has previously been reported in leukocytes or even hepatocytes. There is one report that the glucocorticoid analogue dexamethasone inhibits monocytic production of human SERPINA1 (35). In summary, the mechanisms by which murine *serpina1* and *serpina3* expression is down-regulated in the BM during mobilization remain to be determined.

An intriguing finding of this study is that the same murine BM Gr-1⁺ myeloid cells contain large amounts of serine proteases as well as large amounts of serine protease inhibitors which would suggest that the serpins and proteases, although present in the same cells, are packaged in different granules. This is likely to be the case as *serpina1* transcription is greatest in BM neutrophils, while serine protease (NE and CG) transcription was highest in primitive mono/myeloid cells (Figure 5). It is known that the serine proteases synthesized by primitive myeloid cells are packaged predominantly in primary granules, while proteins synthesized by more mature BM neutrophils are more likely packaged in secondary or tertiary granules (36). The release of these distinct granules by a neutrophil occurs at different stages during the inflammatory process (37), so although contained within the one cell, both the serine proteases and their specific inhibitors may be released separately. Alternatively, serpins may be constitutively secreted with the proteases remaining sequestered in granules.

How are these findings relevant to mobilization in humans? We have previously reported several observations that strongly imply that active neutrophil serine-proteases accumulate in human BM during mobilization. Firstly, CXCR4 is cleaved from the surface of human CD34⁺ cells in the BM and blood during mobilization, resulting in its inactivation (13). Furthermore this cleavage and inactivation of CXCR4 can be recapitulated *in vitro* using NE or CG (13). Secondly, the concentration of cleaved VCAM-1 fragments is increased in the blood of mobilized patients and donors and this increase is significantly correlated to number of clonogenic cells mobilized in the peripheral blood (11). Finally, RT-PCR experiments on sorted populations of human BM cells indicate that transcription of human *SERPINA1* gene is also greater in mature granulocytes, with lower transcription in immature myeloid cells (Coughlin P, manuscript in preparation). The similarity in *SERPINA1* expression profiles between human and mouse myeloid cells, suggests that a decrease in *SERPINA1* and *SERPINA3* transcripts with mobilization may also occur in human BM.

In conclusion, our study is the first to identify a novel role of *serpinal* and *serpina3* proteins in regulating the homeostasis of the hematopoietic microenvironment which ultimately determines the fate of hematopoietic stem cells. This conclusion is strengthened by other studies using genetic linkage analysis showing that the *Aat* locus containing the *serpinal* genes, is significantly linked to BM cellularity in inbred mouse strains (38). As previously reported, many proteoglycans (39, 40), extracellular matrix proteins (41-44), cytokines and chemokines (45, 46) expressed in the BM hematopoietic microenvironment and playing a major role in determining the fate of hematopoietic stem cells, are substrates of the neutrophil serine-proteases that these serpins inhibit. Similarly, an earlier report has shown that the addition of either recombinant human *SERPINA1* or antileukoproteinase to serum-free culture media boosts *in vitro* proliferation and survival of HPC, probably by protecting cytokines from the proteases released by differentiated leukocytes (47). Therefore we propose that 1) *serpinal* and *serpina3* proteins play a major role in maintaining the molecular and functional integrity of the BM hematopoietic microenvironment, 2) the stoichiometric balance between these serpins and neutrophil serine-proteases is disrupted locally in the BM during systemic administration of G-CSF or chemotherapy due to a reduction in *serpinal* and *serpina3* transcripts, and 3) the disruption of the serpin/protease balance allows proteases to evade their naturally-occurring inhibitors, to accumulate in active form and to contribute to the mobilization of HSC.

Acknowledgements

The authors are grateful to Dr Susan K Nilsson for providing whole femur sections of mobilized and non-mobilized mice and to Ms Kamilla Supplit for her excellent technical assistance with immunohistochemistry. This work was supported by grants 080193 and 288701 (JPL) from the National Health and Medical Research Council of Australia. Paul Coughlin is a Wellcome Trust Research Fellow.

References

1. To, L.B., D.N. Haylock, P.J. Simmons, and C.A. Juttner. 1997. The biology and clinical uses of blood stem cells. *Blood* 89:2233-2258.
2. Korbiling, M., and P. Anderlini. 2001. Peripheral blood stem cell versus bone marrow allotransplantation: does the source of hematopoietic stem cells matter? *Blood* 98:2900-2908.
3. To, L.B., D.N. Haylock, T. Dowse, P.J. Simmons, S. Trimboli, L.K. Ashman, and C.A. Juttner. 1994. A comparative study of the phenotype and proliferative capacity of peripheral blood (PB) CD34⁺ cells mobilized by four different protocols and those of steady-phase PB and bone marrow CD34⁺ cells. *Blood* 84:2930-2939.
4. Papayannopoulou, T., and B. Nakamoto. 1993. Peripheralization of hemopoietic progenitors in primates treated with anti-VLA4 integrin. *Proc. Natl. Acad. Sci. USA* 90:9374-9378.
5. Kikuta, T., C. Shimazaki, E. Ashihara, Y. Sudo, H. Hirai, T. Sumikuma, N. Yamagata, T. Inaba, N. Fujita, T. Kina, and M. Nakagawa. 2000. Mobilization of hematopoietic primitive and committed progenitor cells into blood in mice by anti-vascular adhesion molecule-1 antibody alone or in combination with granulocyte colony-stimulating factor. *Exp. Hematol.* 28:311-317.
6. Vermeulen, M., F. Le Pesteur, M.C. Gagnerault, J.Y. Mary, F. Sainteny, and F. Lepault. 1998. Role of adhesion molecules in the homing and mobilization of murine hematopoietic stem and progenitor cells. *Blood* 92:894-900.
7. Scott, L.M., V. Zafiropoulos, G.V. Priestley, and T. Papayannopoulou. 2003. Conditional genomic ablation of $\alpha 4$ integrin and its impact on homeostasis and perturbations of adult hematopoiesis. *Mol. Cell. Biol.* 23:9349-9360.
8. Liles, W.C., H.E. Broxmeyer, E. Rodger, B. Wood, K. Hubel, S. Cooper, G. Hangoc, G.J. Bridger, G.W. Henson, G. Calandra, and D.C. Dale. 2003. Mobilization of hematopoietic progenitor cells in healthy volunteers by AMD3100, a CXCR4 antagonist. *Blood* 102:2728-2730.
9. Hattori, K., B. Heissig, K. Tashiro, T. Honjo, M. Tateno, J.H. Shieh, N.R. Hackett, M.S. Quitoriano, R.G. Crystal, S. Rafii, and M.A. Moore. 2001. Plasma elevation of stromal cell-derived factor-1 induces mobilization of mature and immature hematopoietic progenitor and stem cells. *Blood* 97:3354-3360.

10. Shen, H., T. Cheng, I. Olszak, E. Garcia-Zepeda, Z. Lu, S. Herrmann, R. Fallon, A.D. Luster, and D.T. Scadden. 2001. CXCR-4 desensitization is associated with tissue localization of hemopoietic progenitor cells. *J. Immunol.* 166:5027-5033.
11. Lévesque, J.P., Y. Takamatsu, S.K. Nilsson, D.N. Haylock, and P.J. Simmons. 2001. Vascular cell adhesion molecule-1 (CD106) is cleaved by neutrophil proteases in the bone marrow following hematopoietic progenitor cell mobilization by granulocyte colony-stimulating factor. *Blood* 98:1289-1297.
12. Lévesque, J.P., J. Hendy, Y. Takamatsu, B. Williams, I.G. Winkler, and P.J. Simmons. 2002. Mobilization by either cyclophosphamide or granulocyte colony-stimulating factor transforms the bone marrow into a highly proteolytic environment. *Exp. Hematol.* 30:430-439.
13. Lévesque, J.P., J. Hendy, Y. Takamatsu, P.J. Simmons, and L.J. Bendall. 2003. Disruption of the CXCR4/CXCL12 chemotactic interaction during hematopoietic stem cell mobilization induced by GCSF or cyclophosphamide. *J. Clin. Invest.* 111:187-196.
14. Potempa, J., E. Korzus, and J. Travis. 1994. The serpin superfamily of proteinase inhibitors: structure, function, and regulation. *J. Biol. Chem.* 269:15957-15960.
15. Silverman, G.A., P.I. Bird, R.W. Carrell, F.C. Church, P.B. Coughlin, P.G. Gettins, J.A. Irving, D.A. Lomas, C.J. Luke, R.W. Moyer, P.A. Pemberton, E. Remold-O'Donnell, G.S. Salvesen, J. Travis, and J.C. Whisstock. 2001. The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *J. Biol. Chem.* 276:33293-33296.
16. Huntington, J.A., R.J. Read, and R.W. Carrell. 2000. Structure of a serpin-protease complex shows inhibition by deformation. *Nature* 407:923-926.
17. Lomas, D.A., A. Loubakos, S.A. Cumming, and D. Belorgey. 2002. Hypersensitive mousetraps, alpha1-antitrypsin deficiency and dementia. *Biochem. Soc. Trans.* 30:89-92.
18. Huntington, J.A., and R.W. Carrell. 2001. The serpins: nature's molecular mousetraps. *Sci. Prog.* 84:125-136.
19. Beatty, K., J. Bieth, and J. Travis. 1980. Kinetics of association of serine proteinases with native and oxidized alpha-1-proteinase inhibitor and alpha-1-antichymotrypsin. *J. Biol. Chem.* 255:3931-3934.

20. Duranton, J., C. Adam, and J.G. Bieth. 1998. Kinetic mechanism of the inhibition of cathepsin G by alpha 1-antichymotrypsin and alpha 1-proteinase inhibitor. *Biochemistry* 37:11239-11245.
21. Paterson, T., and S. Moore. 1996. The expression and characterization of five recombinant murine alpha 1-protease inhibitor proteins. *Biochem. Biophys. Res. Commun.* 219:64-69.
22. Barbour, K.W., R.L. Goodwin, F. Guillonneau, Y. Wang, H. Baumann, and F.G. Berger. 2002. Functional diversification during evolution of the murine alpha(1)-proteinase inhibitor family: role of the hypervariable reactive center loop. *Mol. Biol. Evol.* 19:718-727.
23. Forsyth, S., A. Horvath, and P. Coughlin. 2003. A review and comparison of the murine alpha(1)-antitrypsin and alpha(1)-antichymotrypsin multigene clusters with the human clade A serpins. *Genomics* 81:336-345.
24. Inglis, J.D., and R.E. Hill. 1991. The murine Spi-2 proteinase inhibitor locus: a multigene family with a hypervariable reactive site domain. *Embo J.* 10:255-261.
25. Nilsson, S.K., R. Hulspas, H.U. Weier, and P.J. Quesenberry. 1996. In situ detection of individual transplanted bone marrow cells using FISH on sections of paraffin-embedded whole murine femurs. *J. Histochem. Cytochem.* 44:1069-1074.
26. Lomas, D.A., and R. Mahadeva. 2002. Alpha1-antitrypsin polymerization and the serpinopathies: pathobiology and prospects for therapy. *J. Clin. Invest.* 110:1585-1590.
27. Borriello, F., and K.S. Krauter. 1991. Multiple murine alpha 1-protease inhibitor genes show unusual evolutionary divergence. *Proc. Natl. Acad. Sci. USA* 88:9417-9421.
28. Kalsheker, N., S. Morley, and K. Morgan. 2002. Gene regulation of the serine proteinase inhibitors alpha1-antitrypsin and alpha1-antichymotrypsin. *Biochem. Soc. Trans.* 30:93-98.
29. Aronsen, K.F., G. Ekelund, C.O. Kindmark, and C.B. Laurell. 1972. Sequential changes of plasma proteins after surgical trauma. *Scand. J. Clin. Lab. Invest. Suppl* 124:127-136.
30. Perlmutter, D.H. 2002. Liver injury in alpha1-antitrypsin deficiency: an aggregated protein induces mitochondrial injury. *J. Clin. Invest.* 110:1579-1583.
31. Baumann, H., J.J. Latimer, and M.D. Glibetic. 1986. Mouse alpha 1-protease inhibitor is not an acute phase reactant. *Arch. Biochem. Biophys.* 246:488-493.

32. Frazer, J.M., S.A. Nathoo, J. Katz, T.L. Genetta, and T.H. Finlay. 1985. Plasma protein and liver mRNA levels of two closely related murine alpha 1-protease inhibitors during the acute phase reaction. *Arch. Biochem. Biophys.* 239:112-119.
33. Knoell, D.L., D.R. Ralston, K.R. Coulter, and M.D. Wewers. 1998. Alpha 1-antitrypsin and protease complexation is induced by lipopolysaccharide, interleukin-1beta, and tumor necrosis factor-alpha in monocytes. *Am. J. Respir. Crit. Care Med.* 157:246-255.
34. Perlmutter, D.H., L.T. May, and P.B. Sehgal. 1989. Interferon beta 2/interleukin 6 modulates synthesis of alpha 1-antitrypsin in human mononuclear phagocytes and in human hepatoma cells. *J. Clin. Invest.* 84:138-144.
35. Cichy, J., J. Potempa, and J. Travis. 1995. Effect of dexamethasone on alpha 1-proteinase inhibitor synthesis in human cells of monocytic origin. *Biochem. Biophys. Res. Commun.* 208:216-222.
36. Borregaard, N., and J.B. Cowland. 1997. Granules of the human neutrophilic polymorphonuclear leukocyte. *Blood* 89:3503-3521.
37. Borregaard, N. 1997. Development of neutrophil granule diversity. *Ann. N. Y. Acad. Sci.* 832:62-68.
38. de Haan, G., A. Ausema, M. Wilkens, G. Molineux, and B. Dontje. 2000. Efficient mobilization of haematopoietic progenitors after a single injection of pegylated recombinant human granulocyte colony-stimulating factor in mouse strains with distinct marrow-cell pool sizes. *Br. J. Haematol.* 110:638-646.
39. Klebanoff, S.J., M.G. Kinsella, and T.N. Wight. 1993. Degradation of endothelial cell matrix heparan sulfate proteoglycan by elastase and the myeloperoxidase-H₂O₂-chloride system. *Am. J. Pathol.* 143:907-917.
40. McDonnell, J., J.M. Lobner, W.B. Knight, M.W. Lark, B. Green, M. Poe, and V.L. Moore. 1993. Comparison of the proteoglycanolytic activities of human leukocyte elastase and human cathepsin G in vitro and in vivo. *Connect. Tissue. Res.* 30:1-9.
41. Kubes, P., R. Smith, M.D. Grisham, and D.N. Granger. 1993. Neutrophil-mediated proteolysis. Differential roles for cathepsin G and elastase. *Inflammation* 17:321-332.
42. Steadman, R., M.H. Irwin, P.L. St John, W.D. Blackburn, L.W. Heck, and D.R. Abrahamson. 1993. Laminin cleavage by activated human neutrophils yields proteolytic fragments with selective migratory properties. *J. Leukoc. Biol.* 53:354-365.
43. Pipoly, D.J., and E.C. Crouch. 1987. Degradation of native type IV procollagen by human neutrophil elastase. Implications for leukocyte-mediated degradation of basement membranes. *Biochemistry* 26:5748-5754.

44. Kafienah, W., D.J. Buttle, D. Burnett, and A.P. Hollander. 1998. Cleavage of native type I collagen by human neutrophil elastase. *Biochem. J.* 330 (Pt 2):897-902.
45. El Ouriaghli, F., H. Fujiwara, J.J. Melenhorst, G. Sconocchia, N. Hensel, and A.J. Barrett. 2003. Neutrophil elastase enzymatically antagonizes the in vitro action of G-CSF: implications for the regulation of granulopoiesis. *Blood* 101:1752-1758.
46. Lévesque, J.P., J. Hendy, I.G. Winkler, Y. Takamatsu, and P.J. Simmons. 2003. Granulocyte colony-stimulating factor induces the release in the bone marrow of proteases that cleave c-KIT receptor (CD117) from the surface of hematopoietic progenitor cells. *Exp. Hematol.* 31:109-117.
47. Goselink, H.M., J. van Damme, P.S. Hiemstra, A. Wuyts, J. Stolk, W.E. Fibbe, R. Willemze, and J.H. Falkenburg. 1996. Colony growth of human hematopoietic progenitor cells in the absence of serum is supported by a proteinase inhibitor identified as antileukoproteinase. *J. Exp. Med.* 184:1305-1312.

Figure legends

Figure 1. Extracellular fluids from steady-state BM contain serine-protease inhibitors while those from mobilized BM do not. Extracellular fluids from the BM of G-CSF-mobilized mice were mixed with an equal volume of PBS or extracellular fluids from saline-injected mice or with purified human SERPINA1, BB-94 or O-phenantroline as indicated. The proteolytic activity of each sample was estimated by cleavage of recombinant human VCAM-1 visualized following immunoblotting with a goat anti-human VCAM-1 antibody.

Figure 2. Concentration of *serpina1* and $\alpha 2$ -macroglobulin proteins in BM extracellular fluids drops during mobilization induced by G-CSF or CY. Aliquots BM extracellular fluids and blood plasma were taken at indicated time-points of mobilization induced by G-CSF (A, B) or by CY (C, D) and analyzed by immunoblotting with a rabbit anti-human SERPINA1 cross-reacting with mouse *serpina1*. Panels E and F are BM extracellular fluids and blood plasma at different time-points following a single CY injection and analyzed by immunoblotting with a rabbit anti-human $\alpha 2$ -macroglobulin.

Figure 3. Reduced *serpina1* mRNA levels in the BM during mobilization induced by G-CSF or CY. (A) Total RNA was isolated from BM cells of mice injected for 4 days with saline (Sal), G-CSF (G) or 8 days following a single cyclophosphamide injection (Cy). Murine *serpinala-e* (left panel) and β_2 -microglobulin (β_2m) mRNA levels were measured by quantitative real-time RT-PCR. Results are expressed as mRNA amount relative to mRNA for the cellular cytoskeleton protein vimentin (on a log scale). Each symbol represents the mRNA level from a different mouse. Black bars represent average of each group. (B) RNA was extracted from the BM cells of six mice injected with saline or G-CSF for 4 days. Products following RT-PCR for *serpinala-e* (30 cycles) or β_2m (25 cycles) were loaded on 8% PAGE and visualized by ethidium bromide staining. (C) Kinetics of *serpina1* mRNA levels during the course of G-CSF-induced (left panel) and CY-induced (right panel) mobilization. Concentrations of *serpina1* mRNA were measured by real-time RT-PCR from whole BM cells and are relative to vimentin mRNA. Data are mean \pm SD of 6-9 mice per time-point. (D) Kinetics of NE mRNA levels during the course of G-CSF-induced (left panel) and CY-induced (right panel) mobilization. Concentrations of mRNA were measured as

described in (C). (E) Kinetics of CFC mobilization into the peripheral blood induced by G-CSF (left panel) and CY (right panel). Data are mean \pm SD of 6-9 mice per time-point.

Figure 4. mRNA levels of secreted murine *serpina3*, *NE*, *CG*, *VCAM-1*, *CXCR4* at the peak of HPC mobilization induced by G-CSF or CY. mRNA levels were measured by real-time RT-PCR from BM cells isolated from mice injected for 4 days with saline (Sal) or G-CSF (G) or at day 6 following a single CY injection (Cy). Results are expressed as mRNA level relative to mRNA for the cytoskeleton protein vimentin (on a log scale). Each symbol represents the mRNA level from a different mouse. Black bars represent average of each group. Fold decreases between saline-treated groups and G-CSF- or CY-treated groups, together with corresponding significance levels (* for $p < 0.05$, ** for $p < 0.01$ and NS for non significant) are indicated underneath.

Figure 5. *Serpina1* transcripts are decreased in all BM mono/myeloid populations during HPC mobilization. (A) Number of nucleated cells and Mac-1^+ (CD11b^+) mono/myeloid cells in the femoral BM during the course of G-CSF administration. Data represent mean \pm SD from 6-9 mice per group. (B) Phenotypic expression and sorting of primitive mono/myeloid Mac-1^+ Gr-1^{dim} cells (gate R3), immature myeloid Mac-1^+ $\text{Gr-1}^{\text{intermediate}}$ cells (gate R4) and mature Mac-1^+ $\text{Gr-1}^{\text{bright}}$ neutrophils (gate R2) from mouse BM following 4 days of saline (left panel) or G-CSF (right panel) injection. Percentages in sorting gates indicate the proportion of these cells among total BM nucleated cells. (C) Giemsa staining of the R3, R4 and R2 populations sorted as above. Top row corresponds to cells sorted from BM of saline-injected mice whereas the bottom row corresponds to cells sorted from the BM of G-CSF-injected mice. (D) *Serpina1* mRNA expression levels in R3, R4 and R2 populations sorted from the BM of mice injected for 4 days with saline (gray columns) or G-CSF (white columns). Results are expressed as mRNA concentration relative to vimentin mRNA on a log scale. (E) Cathepsin G (CG) mRNA expression levels in R3, R4 and R2 populations sorted from the BM of mice injected for 4 days with saline (gray columns) or G-CSF (white columns). Results are expressed as mRNA concentration relative to vimentin mRNA on a log scale.

Figure 6. Down-regulation of *serpina1* is restricted to the BM during HPC mobilization. (A) *Serpina1a-e* (left panel) and *serpina3c* (right panel) mRNA levels were measured in liver, whole BM and spleen cells from mice injected for 4 days with saline (gray columns) or G-

CSF (white columns). Results are expressed as mRNA concentration relative to vimentin mRNA on a log scale. (B) Immunohistochemical staining of liver, BM and spleen sections with a rabbit anti-human SERPINA1 antibody cross-reacting with mouse serpin1. The top row are sections from mice injected for 4 days with saline while the bottom row are sections from mice injected for 4 days with G-CSF. The small inserts shown in the liver sections are negative controls performed with non-immune control rabbit IgG used at the same concentration as the rabbit anti-human SERPINA1. All sections are magnified 200 times.

Table 1

Murine serpin genes and primer sequences.

	Old nomenclature	Sequence of reactive loop, P ₁ P ₁ '	Potentially secreted from cells	Oligonucleotide sequences for realtime RT-PCR	RT-PCR BALBc BM
<i>Serpina3a</i>	Unknown1	QS	yes		ND
<i>Serpina3b</i>	6A1	MS ^a	yes	f 5' -TCATGTCTGCAAACTTAAAC-3' b 5' -TGTGGGGTTGATAACCTTTCC-3'	- ^c
<i>Serpina3c</i>	1A1	LS ^b	yes	f 5' -CTTAGTAGAAGAACCAGTCTG-3' b 5' -CTTAGGGTGAGTGATTTTGGC-3'	+
<i>Serpina3d</i>	2B1	LS ^b	yes	f 5' -AGATTCAAAATTGCCCACTG-3' b 5' -CTTGGGATTCAAGACTTTCAC-3'	+
<i>Serpina3e</i>	2B2(i-)	RC	?	f 5' -AGGAGTTAAAGTTAATCTACG-3' b 5' -CTTGGGATTGTAACTTTGGC-3'	+ ^d
<i>Serpina3f</i>	2A1	CC	no	f 5' -ATCTCCAATGTTGTCAAGGTG-3' b 5' -TG (T/G) AACTTTTGCCATAAAGAG-3'	- ^c
<i>Serpina3g</i>	2A2	CC	no	f 5' -CAGGAATGGCAGGTGTGG-3' b 5' -TG (T/G) AACTTTTGCCATAAAGAG-3'	+ ^c
<i>Serpina3h</i>	6C28	CC	no	f 5' -CCACAGGGGTCAAATTAATTC-3' b 5' -CTTGGGATTGTAACTTTGGC-3'	+
<i>Serpina3i</i>	2B2	RC	no	f 5' -AGGAGTTAAAGTTAATCTACG-3' b 5' -CTTGGGATTGTAACTTTGGC-3'	+ ^d
<i>Serpina3j</i>	Unknown2	LS ^b	yes		ND
<i>Serpina3k</i>	MMCM	RK	yes	f 5' -GTTATTGGTGGCATTCTGAAG-3' b 5' -CTTGGGGTTA (T/G) TGACTTTGGC-3'	+
<i>Serpina3l</i>	3E2	QS	yes		ND
<i>Serpina3m</i>	3e46	RS	yes	f 5' -GCTTTCGTTCTAGAAGATTAC-3' b 5' -CTTGGGGTTA (T/G) TGACTTTGGC-3'	+ ^c
<i>Serpina3n</i>	EB22.4	MS ^a	yes	f 5' -CAATGTCTGCGAACTGTACC-3' b 5' -TTTGGGGTTGGCTATCTTGGC-3'	+
<i>Serpina1a-e</i>			yes	f 5' -GAAGCTGCAGCAGCTACAGTC-3' b 5' -TGTGGGATCTACCAC'TTTCC-3'	+

a. 'MS' is the canonical human α_1 -antitrypsin reactive center loop P₁P₁'b. 'LS' is the canonical human α_1 -antichymotrypsin reactive center loop P₁P₁'

c. denotes that PCR was performed with a lower annealing temperature of 55°C

d. oligonucleotides for *serpina3e* and *serpina3i* do not distinguish between these two transcripts.

Figure 1.

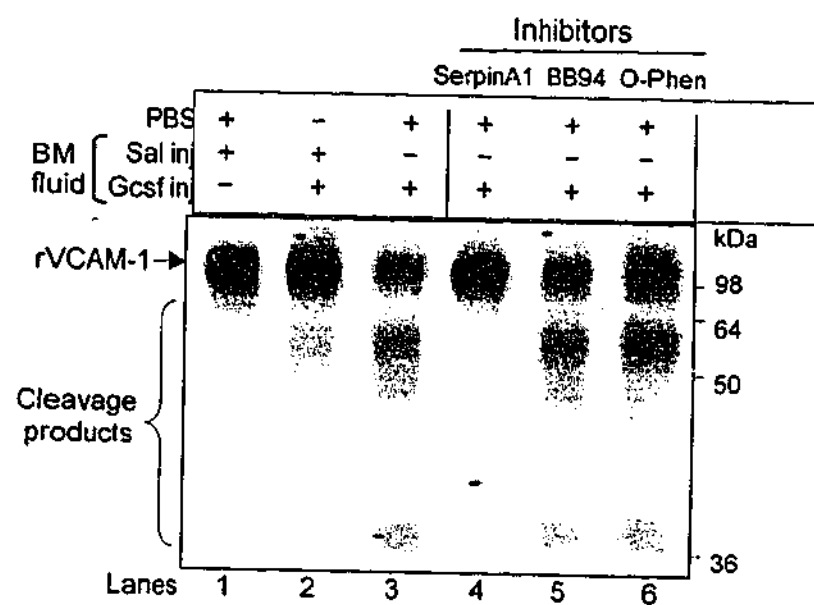


Figure 2.

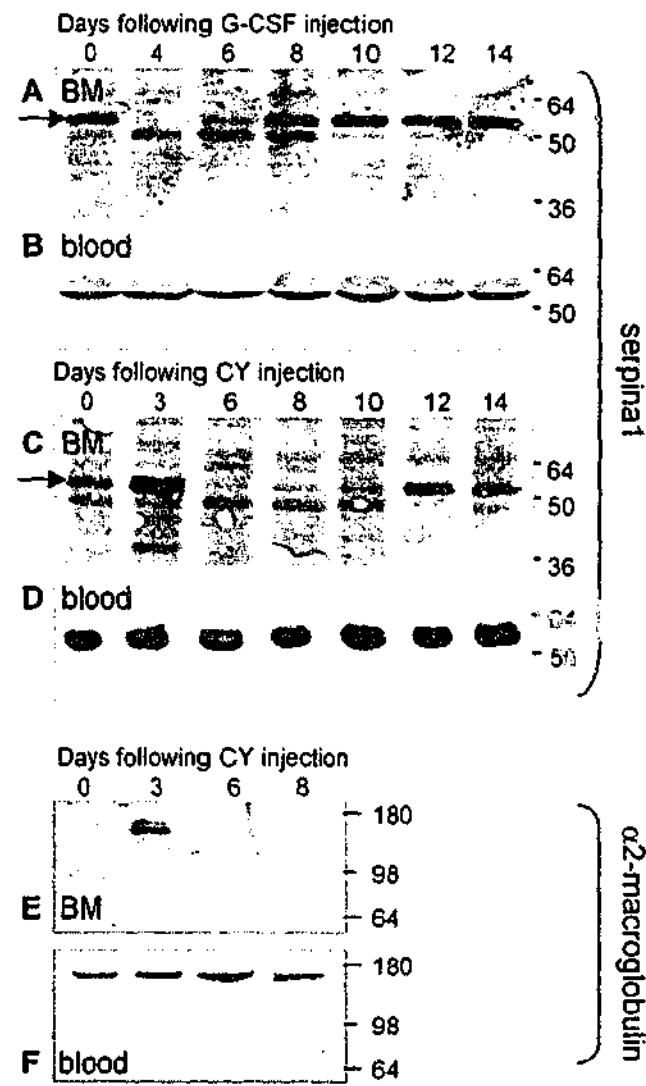


Figure 3.

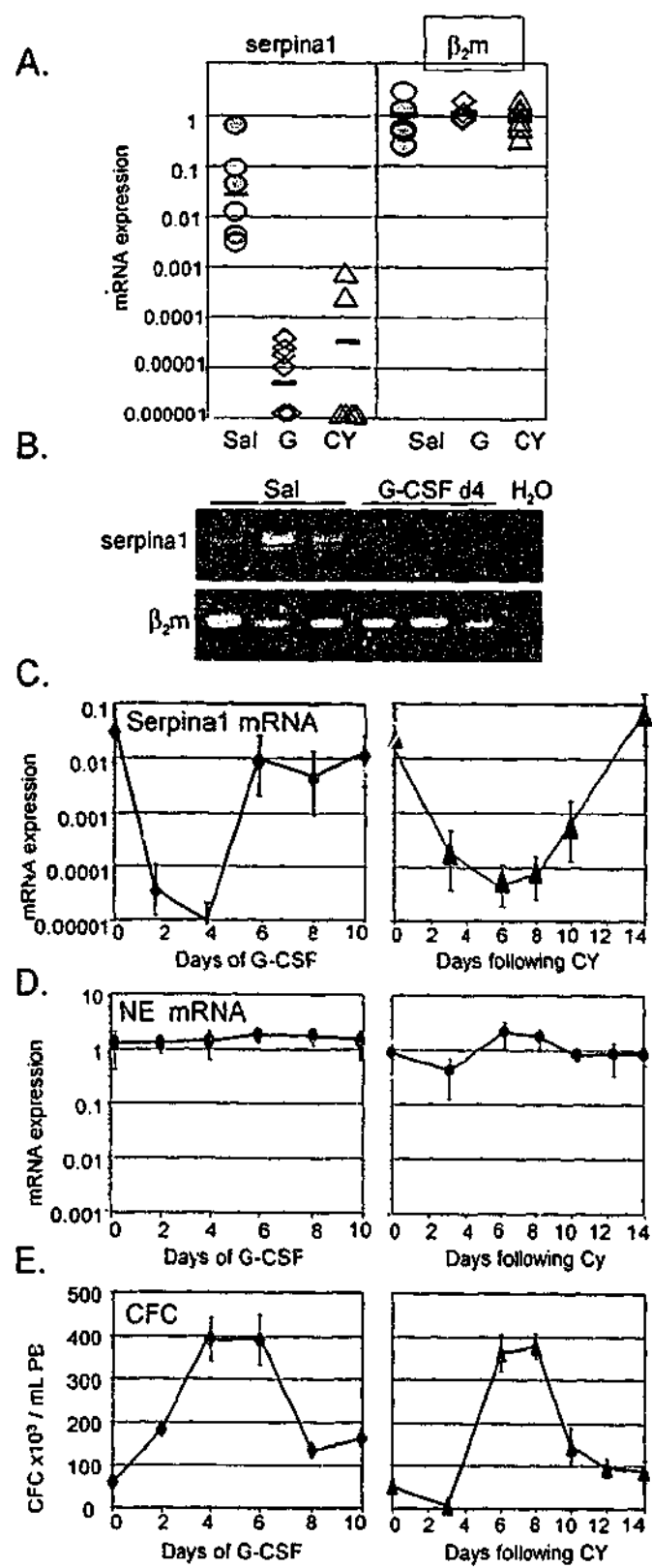


Figure 4.

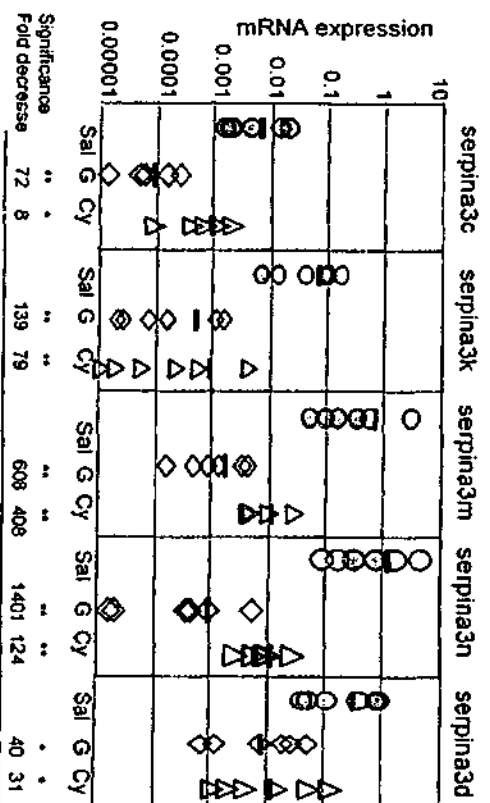
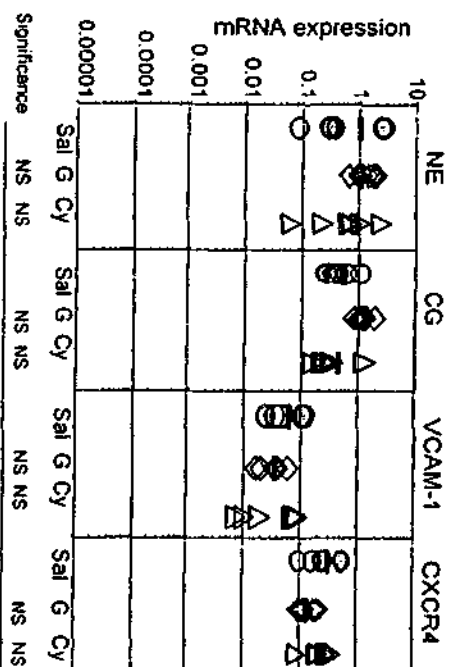


Figure 5.

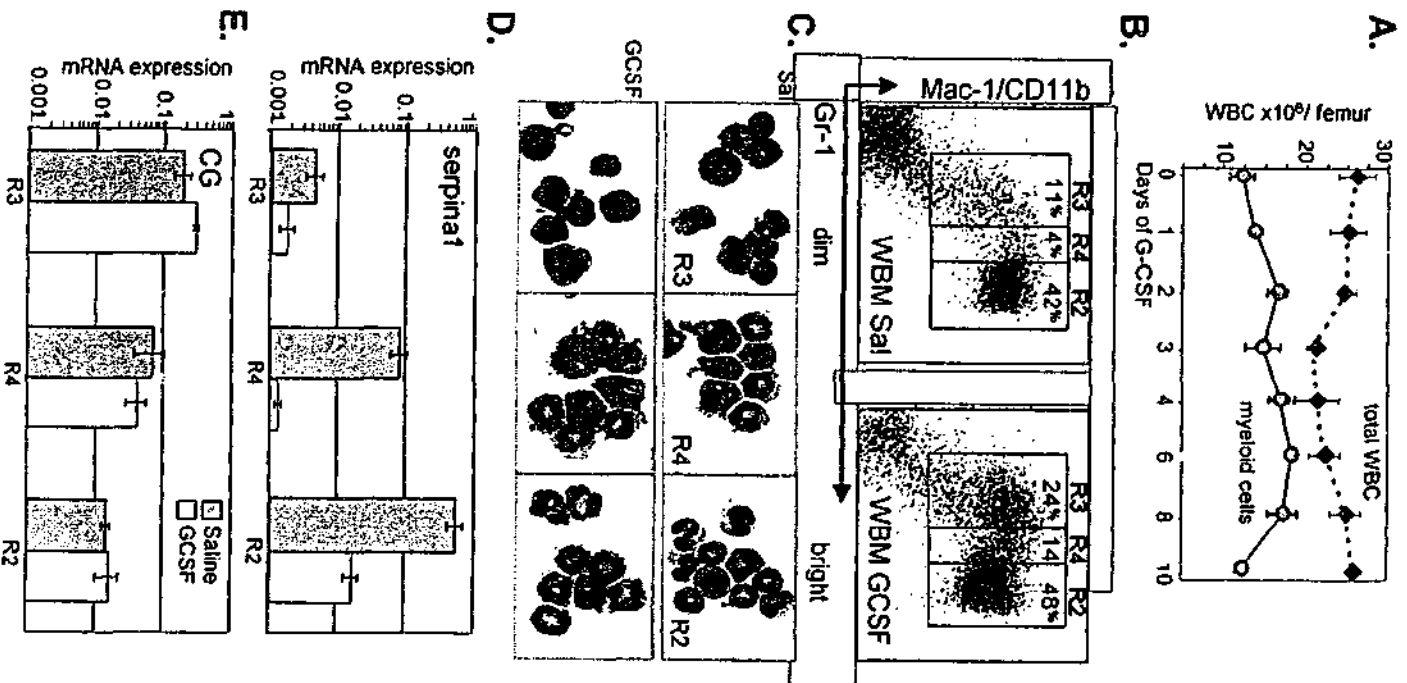


Figure 6.

