UNIVERSITY OF SOUTHAMPTON

STRUCTURE AND SPECIFICITY STUDIES ON BATROXOBIN, A SNAKE VENOM DERIVED THROMBIN-LIKE ENZYME

by

Lorraine Earps

A thesis submitted for
the degree of
Doctor of Philosophy

Division of Biochemistry and Molecular Biology

August 1999


UNIVERSITY OF SOUTHAMPTON
LIBRARY

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SCIENCE

DIVISION OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Doctor of Philosophy

STRUCTURE AND SPECIFICITY STUDIES ON BATROXOBIN, A SNAKE VENOM DERIVED THROMBIN-LIKE ENZYME

by Lorraine Earps

Batroxobin is a thrombin-like enzyme derived from the venoms of the South American Pit Viper *Bothrops Atrox*. Batroxobin is termed thrombin-like due to its ability to cleave fibrinogen. However, unlike thrombin, which cleaves both the fibrinopeptides, A and B, batroxobin is only capable of cleaving fibrinopeptide A from fibrinogen.

Initially, batroxobin derived from two sources, *Bothrops atrox* and *Bothrops moojeni*, two subspecies of the same *Bothrops Atrox* species, were characterised. This was to assess any differences between batroxobin derived from the different species and also to look at their reactivity with fibrinogens derived from various species. During the course of the physical characterisation of the two batroxobins, it was not possible to identify any difference between them in terms of mass and N-terminal sequence analysis. However, on reaction with various fibrinogens it was evident that the batroxobin derived from *Bothrops atrox* was twice as active as the batroxobin derived from *Bothrops moojeni*. These specificity studies, on fibrinogens derived from eight species, showed that both batroxobins had a preference for fibrinogens derived from rat and mouse plasma, the natural prey of *Bothrops Atrox* pit vipers.

Sequence analysis of the fibrinogens indicated some differences in the sequences around the scissile bond of fibrinopeptide A, which could account for differences observed in specificity between the batroxobins and thrombin. To check the validity of this proposal a series of peptides were designed and synthesised using solid phase peptide chemistry. Reactions between these peptides and batroxobin and thrombin were carried out, with the specific intention of comparing the requirements of each enzyme for substrate recognition and cleavage. It was found that batroxobin has an absolute requirement for a proline residue at P2' position, unlike thrombin which can tolerate histidine. It was also evident that batroxobin has a preference to an aspartic acid at the P3 position, whereas thrombin prefers a glycine in this position. The synthesis of a novel inhibitor with a reduced peptide bond showed that both batroxobin and thrombin have a requirement for a carbonyl group at the scissile bond not only for cleavage, but also for recognition.

In an attempt to find a structural explanation for the observations, homology modelling was performed. A model produced from the structures of porcine glandular kallikrein and TSV-PA was finally decided upon. The sequences of various peptides were modelled into the active site cleft of the batroxobin model. This work showed that the preference of batroxobin for an aspartic acid at P3 could be due to its ability to form a stabilising interaction with a lysine in the active site cleft, which is not present in thrombin. It also indicated that batroxobin cannot tolerate histidine at the P2' position due to the highly positive nature of the enzyme at that point. This work also allowed the identification of a potential fibrinogen exosite on the batroxobin model that is different to those observed in thrombin. This region of high positive charge correlates very well to one identified recently on a closely related snake venom enzyme crotalase.

CONTENTS

CHAPTER ONE: General Introduction

1.1 Venomous snakes and their venoms 1

1.2 The pit vipers 2

1.3 Composition of snake venoms 4

1.4 Blood as a target for snake venom 6

1.5 Haemostasis 7

1.5.1 Intrinsic pathway 8

1.5.2 Extrinsic pathway 9

1.5.3 Fibrin formation 10

1.6 Fibrinolysis 14

1.7 The proteases 15

1.7.1 The serine proteases 15

1.8 Action of serine proteases 18

1.9 Catalytic action of serine proteases 22

1.9.1 General mechanism of serine proteases 22

1.9.2 Substrate specificity of serine proteases 24

1.10 Pit viper snake bites 26

1.11 Components within venom affecting haemostasis 26

1.11.1 Blood platelets and components affecting them 27

1.11.2 Venom components acting on the blood vessel wall 29

1.11.3 Venom components acting on endothelial cells 29

1.11.4 Venom components acting on the activation of prothrombin 29

1.11.5 Venom components acting on fibrinogen 30

1.11.6 Venom components that act upon protein C 32

1.12 Batroxobin the thrombin-like enzyme 32

1.13 Practical applications of batroxobin 35

1.13.1 Haemostyptic drugs 35

1.13.2 Defibrinogenating drugs 35

1.13.3 Laboratory reagents 36

1.13.4 Other uses 36

1.14 Comparison of thrombin and batroxobin 36

1.14.1 Catalytic triad 37
1.14.2 Cleavage of fibrinogen 37
1.14.3 Other actions of thrombin 39

CHAPTER TWO: Materials and Methods

2.1 Chemicals and reagents 41
2.2 Buffers and Solutions 41
2.2.1 Purification Buffers 41
2.2.2 S-2238 chromogenic substrate 41
2.2.3 HEPES buffered saline 41
2.2.4 Calcium chloride solution 42
2.2.5 Bovine fibrinogen stock solution 42
2.2.6 Bovine thrombin 42
2.2.7 Polyacrylamide gel electrophoresis 42
2.2.8 12% SDS-PAGE gel 43
2.2.9 1.6x Disruption buffer 44
2.2.10 5x Running buffer 44
2.2.11 Stain for SDS-PAGE gels 44
2.2.12 Destain for SDS-PAGE gels 45
2.3 Determination of protein concentration 45
2.4 Identification of thrombin-like enzymes 46
2.4.1 S-2238 chromogenic assay 46
2.4.2 Visual gelation clotting assay (fibrinogen) 47
2.4.3 Kinetic study of fibrin polymerisation 47
2.4.4 Coagulation of human plasma 48
2.4.5 Zymographic assay 48
2.4.6 Fibrinogen-agarose plate assay 49
2.5 Purification of the thrombin-like enzyme from crude venom 49
2.5.1 Purification using heparin sepharose affinity column 49
2.5.2 Purification using Poros 20HE Biosprint column 50
2.5.3 Purification using p-aminobenzamidine column 51
2.6 Homology modelling 52
2.7 Solid phase peptide synthesis 53

2.8 Reaction of thrombin and batroxobin on peptides 53
2.9 Crystal growth 54

CHAPTER THREE: Purification, Characterisation and Crystallisation of Batroxobin

3.1 Introduction 55
3.2 Batroxobin purification trials 56
3.2.1 Purification by heparin sepharose 56
3.2.2 Purification using Biosprint HE and HQ columns 56
3.2.3 Purification using p-aminobenzamidine column 59
3.3 Purification of batroxobin 61
3.4 Characterisation of batroxobin (*B. atrox*) supplied by LATOXAN 63
3.5 Characterisation of batroxobin (*B. atrox*) supplied by Pentapharm 67
3.6 Specificity studies of thrombin and batroxobin (*B. moojeni)* and (*B. atrox*) on various fibrinogens. 69
3.6.1 Reaction of thrombin and batroxobin with various fibrinogens. 70
3.6.2 Experimental procedure 71
3.6.2.1 End point clot time. 71
3.6.2.2 Kinetic study of fibrin formation 72
3.6.3 Comparison of the role of Ca^{2+} in clot formation between thrombin and batroxobin from *B. atrox* and *B. moojeni.* 72
3.6.4 Comparison of clot formation with various fibrinogens between thrombin and batroxobin (*B. moojeni*) and (*B. atrox*) 73
3,6,5 Kinetic study of clot formation for thrombin and batroxobin with various fibrinogens 80
3.7 Crystallographic trials. 83
3.7.1 Introduction. 83
3.7.2 Crystal growth. 84
3.7.3 Initial trial. 85
3.7.4 Second trial. 86
3.7.5 Refined trial 86

3.7.6 Initial diffraction. 89
3.7.7 Diffraction data. 90
3.7.7.1 Data collected at Grenoble. 90
3.7.8 Continuing crystallography work 93

CHAPTER FOUR: Synthesis and Reactions of Peptides and Inhibitors.

4.1 Introduction 94
4.2 Peptide design. 95
4.3 Fmoc peptide synthesis 99
4.3.1 Materials and methods. 99
4.3.2 Reaction vessel preparation. 99
4.3.3 General reaction protocol for Fmoc synthesis. 100
4.3.3.1 The coupling protocol. 100
4.3.3.2 Capping. 101
4.3.4. The ninhydrin assay. 101
4.3.4.1 The quantitative ninhydrin assay. 102
4.4 Cleavage of the peptide from the resin. 103
4.4.1 TFA cleavage from Wang resin. 103
4.4.2 HF Cleavage from the MBHA resin. 103
4.5 Purification and characterisation of peptides. 104
4.5.1 Gel filtration. 104
4.5.2 Analytical HPLC. 104
4.5.3 Preparative HPLC. 105
4.5.4 Electrospray mass spectrometry. 105
4.6 Reactions of thrombin and batroxobin with synthetic peptides. 107
4.7 Initial experimental screen 107
4.7.1 Experimental procedure. 107
4.7.2 Results of initial screen. 107
4.7.3 Production and reactions of decapeptide free acids. 110
4.7.4 Production and reactions of amide peptides. 110
4.8 Kinetic experimental analysis. 111
4.8.1 Experimental procedure. 111
4.8.2 Results of kinetic studies. 112

4.8.2.1 The role of Ca^{2+} in thrombin-like enzyme activity. 112
4.8.2.2 Comparison of the reactivity of batroxobin (*B. atrox*) and batroxobin (*B. moojeni*) with artificial peptide substrates 112
4.8.2.3 Reactions of peptides with each of the enzymes 112
4.9 Inhibitor studies. 118
4.9.1 Inhibitor synthesis. 118
4.9.1.1 Structure of the proposed inhibitor 118
4.9.1.2 The formation of the boc-arginine diol from the ornithine diol 119
4.9.1.3 The formation of the boc-valine-arginine aldehyde 120
4.9.1.4 Formation of the reduced peptide bond 120
4.9.2 Reaction of inhibitors. 123
4.9.2.1 Experimental procedure. 123
4.9.2.2 Results of inhibitor studies. 123
4.10 Discussion. 125

CHAPTER FIVE: Homology modelling of batroxobin.
5.1 Introduction 128
5.2 Sequence homology analysis. 128
5.3 Modelling of batroxobin from *B. moojeni* on porcine pancreatic kallikrein. 131
5.3.1 Validity of the modelled structure of batroxobin based upon porcine pancreatic kallikrein. 136
5.4 Batroxobin modelled on mouse salivary gland kallikrein 137
5.5 The structure of batroxobin based upon *Trimeresurus stenjnejeri* venom plasminogen activator (TSV-PA) 140
5.6 Comparison of batroxobin modelled on porcine kallikrein and TSV-PA with α-thrombin. 144
5.6.1 Comparison of the primary and overall structure of batroxobin and α-thrombin. 144
5.6.2 The β-sheeted barrels. 147
5.6.3 Disulphide bridges. 147

5.6.4 The external loops. 148
5.6.5 The active site region. 151
5.6.6 Specificity binding pockets. 151
5.7 Substrate modelling. 151
5.7.1 Batroxobin modelled with the decapeptide GGGVRGPRVV at the active site. 153
5.7.2 Comparison of the binding around the scissile bond in α-thrombin compared to the model of batroxobin. 157
5.7.3 Comparison of the substrate binding at the S3 site between α-thrombin and the batroxobin model. 159
5.7.4 Comparison of the substrate binding at S2' between α-thrombin and the batroxobin model. 161
5.7.5 Comparison of the fibrinogen binding sites in α-thrombin and batroxobin. 164
5.8 Discussion. 165

SUMMARY AND FUTURE WORK 168

REFERENCES 171

LIST OF FIGURES

CHAPTER ONE: General Introduction

Figure 1.1: Photograph of the South American pit-viper, Bothrops atrox. 3

Figure 1.2 : General overview of coagulation. 7

Figure 1.3: The intrinsic pathway of blood coagulation. 9

Figure 1.4: The extrinsic pathway of blood coagulation. 10

Figure 1.5: The structure of fibrinogen. 10

Figure 1.6: Structure of fibrin monomers, and the subsequent stacking of monomers to produce the fibrin polymer. 11

Figure 1.7: Cross-linking of fibrin by factor $XIII_a$ 11

Figure 1.8: Action of thrombin on fibrinogen. 12

Figure 1.9: Fibrinolysis 14

Figure 1.10: Positioning of substrate within the active site and formation of the tetrahedral intermediate by the action of the nucleophilic serine. 16

Figure 1.11: The collapse of the tetrahedral intermediate. 16

Figure 1.12: Attack by water molecules in the surrounding solution. 17

Figure 1.13: Mechanism for the collapse of the tetrahedral intermediate. 17

Figure 1.14: Chromogenic reaction of serine proteases. 19

Figure 1.15: Reaction of DFP with serine proteinase. 20

Figure 1.16: Reaction of His 57 with TFCK. 21

Figure 1.17: The active site residues of the serine proteases. 21

Figure 1.18: Charge relay mechanism of the active site of serine proteases. 22

Figure 1.19: Mechanism of serine protease, illustrated by chymotrypsin 23

Figure 1.20 Substrate binding specificity. 24

Figure 1.21: Hydrophobic specificity pocket of three different serine proteases. 25

Figure 1.22: The onset of circulatory shock as a result of envenomation. 27

Figure 1.23: Sequence homologies of various serine proteases. 33

Figure 1.24: Action of batroxobin on fibrinogen. 35

Figure 1.25: Catalytic triad of thrombin. 38

CHAPTER TWO: Materials and Methods

Figure 2.1: Reaction of thrombin or thrombin-like enzymes with S-2238. 47

CHAPTER THREE: Purification, Characterisation and Crystallisation of Batroxobin.

Figure 3.1: SDS-Gel comparison of the crude venom derived from B. Atrox and B. Moojeni. 55

Figure 3.2: Purification profile of batroxobin from B. atrox crude venom using a heparin sepharose affinity method. 57

Figure 3.3: Purification of crude B. atrox venom using HE Biosprint column. 58

Figure 3.4: Purification profile of partially purified B. Atrox venom on HQ Biosprint column. 59

Figure 3.5: Purification profile for crude venom (B. atrox) using the p-amino benzamidine affinity column. 60

Figure 3.6: SDS-Gel showing the purification of batroxobin from crude venom. 62

Figure 3.7: N-terminal sequence analysis of batroxobin (B. atrox and B. moojeni) 63

Figure 3.8: SDS-PAGE gel of purified batroxobin (B.atrox). 64

Figure 3.9: Graphical determination of the molecular mass of batroxobin (B. Atrox, LATOXAN) from a SDS-PAGE gel. 65

Figure 3.10: MALDI-TOF MS of batroxobin (B. moojeni). 66

Figure 3.11: MALDI-TOF MS of batroxobin (B. atrox, LATOXAN). 66

Figure 3.12: N-terminal sequence analysis of batroxobin (B. moojeni, Pentapharm) and Batroxobin (B. atrox, Pentapharm). 68

Figure 3.13: Electrospray MS analysis of batroxobin (B. moojeni, Pentapharm). 68

Figure 3.15: Effect of Ca^{2+} on clot formation with thrombin or batroxobin. 73

Figure 3.16: Comparison of end point clot formation of thrombin with various fibrinogens. 74

Figure 3.17: Comparison of end-point clot time for batroxobin

(B. moojeni, Pentapharm) with various fibrinogens. 74

Figure 3.18: Comparison of end-point clot time for batroxobin (B. atrox, Pentapharm) with various fibrinogens. 75

Figure 3.19: Comparison of end-point clot time for human fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm). 75

Figure 3.20: Comparison of end-point clot time for cow fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm). 76

Figure 3.21: Comparison of end-point clot time for pig fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm) 76

Figure 3.22: Comparison of end-point clot time for cat fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm). 77

Figure 3.23: Comparison of end-point clot time for dog fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm). 77

Figure 3.24: Comparison of end-point clot formation for rabbit fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm). 78

Figure 3.25: Comparison of end-point clot time for rat fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm) 78

Figure 3.26: Comparison of end-point clot time with mouse fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm). 79

Figure 3.27: Hofstee-Eadie Plot of thrombin with fibrinogens from various species. 80

Figure 3.28: Hofstee-Eadie plot of batroxobin (B. moojeni) with fibrinogens from various species. 81

Figure 3.29: Hofstee-Eadie plot for batroxobin (B. atrox) with fibrinogen from various species. 81

Figure 3.30: Hanging - drop method of crystal growth. 84

Figure 3.31: Photograph showing crystal formation in the second Hamilton crystal screen 87
Figure 3.32: An example of normal diffraction data. 91
Figure 3.33: Diffraction data obtained for batroxobin and example of mosaicity 91
Figure 3.34: An example of ideal unit cell packing 92
Figure 3.35: An example of irregular unit cell packing resulting in mosaicity 92

CHAPTER FOUR: Synthesis and Reactions of Peptides and Inhibitors

Figure 4.1a: The Fmoc protecting group 97
Figure 4.1b: The Boc protecting group 97
Figure 4.2: General scheme for solid phase peptide chemistry 98
Figure 4.3: Reaction vessel 99
Figure 4.4: HPLC trace for a purified peptide GVRGPR-COOH 106
Figure 4.5: Electrospray MS of the purified peptide GVRGPR-COOH 106
Figure 4.6: Representative HPLC trace of a hexapeptide free acid (GVRGPR-COOH) after a 0mins incubation with either thrombin or batroxobin 108
Figure 4.7: Representative HPLC trace of the same hexapeptide free acid after a 24-hour incubation with thrombin or batroxobin 108
Figure 4.8: Representative HPLC trace of the hexapeptide amide (GVRGPR-amide) after a 0mins incubation with thrombin or batroxobin 109
Figure 4.9: Representative HPLC trace of the hexapeptide amide (GVRGPR-amide) after a 24-hour incubation with thrombin or batroxobin 109
Figure 4.10: Eadie-Hoffstee plot of the reactions of synthetic peptides with thrombin 113
Figure 4.11: Eadie-Hoffstee plot of the reactions of synthetic peptides with batroxobin 113
Figure 4.12: Representative HPLC traces of peptides cleaved by thrombin or batroxobin, over a 30min period 117

Figure 4.13: Formation of the deprotected arginine-diol 119
Figure 4.14: Formation of the boc valine arginine aldehyde 120
Figure 4.15: Formation of the reduced peptide bond 121
*Figure 4.16: Analytical trace showing the inhibitor GGGVR-reduced-*GPRVV-amide 122
Figure 4.17: Electrospray MS of the inhibitor GGGVR-reduced-GPRVV-amide, showing the doubly charged species. 122
Figure 4.18: Eadie-Hoffstee plot of the effect of various inhibitors on the reaction of batroxobin with S-2238. 124
Figure 4.19: Eadie-Hoffstee plot of the effect of various inhibitors on the reaction of thrombin with S-2238. 125

CHAPTER FIVE: Homology modelling of batroxobin.

Figure 5.1 Alignment of the amino acid sequences of batroxobin with various serine proteases. 130
Figure 5.2: Alignment of the amino acid sequences of batroxobin with porcine pancreatic kallikrein. 132
Figure 5.3: Schematic representation of the backbone of the dimeric porcine pancreatic kallikrein. 134
Figure 5.4: Schematic representation of the backbone of porcine glandular kallikrein. 135
Figure 5.5: Schematic representation of the backbone of the model batroxobin based upon porcine glandular kallikrein 135
Figure 5.6: Alignment of the amino acid sequences of batroxobin with mouse salivary gland kallikrein 138
Figure 5.7a: Schematic representation of the backbone of mouse salivary gland kallikrein monomer 139
Figure 5.7b: Schematic representation of the backbone of the model of batroxobin based upon mouse salivary gland kallikrein. 139
Figure 5.8: Alignment of the amino acid sequences of batroxobin with porcine pancreatic kallikrein and TSV-PA. 141
Figure 5.9: Schematic representation of the backbone of the model of batroxobin based upon porcine glandular kallikrein and TSV-PA. 143

Figure 5.10: Alignment of amino acid sequences of batroxobin and α-thrombin. 145

Figure 5.11: Schematic representation showing the backbone of α-thrombin superimposed over the backbone of batroxobin modelled on porcine kallikrein and TSV-PA. 146

Figure 5.12: Active site region of the model batroxobin structure. 149

Figure 5.13: Active site region of α-thrombin. 150

Figure 5.14: Schematic representation of the backbone of α-thrombin with the thrombin specific inhibitor rhodonin 152

Figure 5.15: Schematic representation of the backbone of the batroxobin model, with the synthetic peptide GGGVRGPRVV bound to the active site cleft. 154

Figure 5.16: Schematic representation of the backbone of the batroxobin model with the synthetic peptide GGGVRGPRVV bound, rotated through 90° 155

Figure 5,17: Stereoimage of the backbone of batroxobin with the synthetic peptide GGGVRGPRVV bound to the active site cleft 156

Figure 5.18: The binding of the substrate to the batroxobin model around the scissile bond. 158

Figure 5.19: The binding of aspartic acid (P3) within the model batroxobin structure. 160

Figure 5.20: Surface potential representation of the batroxobin model. 162

Figure 5.21: Surface potential representation of α-thrombin. 163

LIST OF TABLES

CHAPTER ONE: General Introduction

Table 1.1: Plasma proteins involved in haemostasis and fibrinolysis 13

Table1.2: Snake venom derived platelet activators. 28

Table 1.3: Snake venom derived platelet inhibitors. 28

Table 1.4: Snake venom derived components affecting haemostasis. 31

Table 1.5: Snake venom derived components affecting fibrinolysis 31

CHAPTER THREE: Purification, Characterisation and Crystallisation of Batroxobin.

Table 3.1: Purification table of batroxobin from crude venom 62.

Table 3.2: In vivo effect of batroxobin on various species after differing modes of injection 71

Table 3.3: Kinetic data for thrombin with various fibrinogens. 82

Table 3.4: Kinetic data for batroxobin (B. atrox) with various fibrinogens. 82

Table 3.5: Kinetic data for batroxobin (B. moojeni) with various fibrinogens. 83

Table 3.6: Refined well conditions at pH 5, 5.5 and 6. 88

CHAPTER FOUR: Synthesis and Reactions of Peptides and Inhibitors.

Table 4.1: The reactions of thrombin and batroxobin with various amide peptides after 24 hours incubation. 110

Table 4.2: The kinetic data for all of the synthetic peptides with each enzyme capable of cleaving them. 114

Table 4.3: Inhibitory effects of each compound on thrombin and batroxobin. 123

Table 4.4: Table to show the relative Ki values of various inhibitors with thrombin and batroxobin. 125

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Professor P.M. Shoolingin-Jordan, for his help and support during the course of this study. I also wish to thank my industrial supervisor, Dr. S. Cederholm-Williams, for all his assistance, both financial and in deciphering the intrigue associated with the snake sub-species.

I am indebted to the following people for all their help:

Dr. Julian Marshall and Dr. Emma Akufo, for their help with thrombin and batroxobin assaying techniques.

Gyles Cozier and Shu'fen for their assistance with the molecular modelling work.

Gareth Lewis for his help with the crytallography work.

Dr. Ram. Sharma and Dr. Robert Broadbridge for help with peptide synthesis.

Will Reece-Blanchard for all his help.

All my friends and colleagues in the department, in particular the members of the Jordan lab.

I wish to thank the BBSRC and Oxford Bioresearch Laboratories for funding this work.

I wish to thank my parents for their continuing love and support, which has been of great importance over the years. Finally, I wish to thank my husband, Darren, whose ongoing help support, guidance and love have kept me going even throughout the most difficult of times.

ABBREVIATIONS

A	Absorption
Abs	Absorption
Analar H_2O	Analytical water
B. Atrox	*Bothrops Atrox (sp)*
B. atrox	*Bothrops Atrox Atrox* (maranhao type II locality)
B. moojeni	*Bothrops Atrox Moojeni* (moojeni type I locality)
Da	Dalton
EDTA	Ethylenediaminetetraacetic acid
Electrospray MS	Electrospray Mass Spectrometry
HPLC	High Pressure Liquid Chromatography
MALDI-TOF MS	Matrix Assisted Laser Desorption Ionisation Time of Flight Mass Spectrometer
M	Molar
mM	Millimolar
Mr	Relative molecular weight
nM	Nanomolar
OBL	Oxford Bioresearch Laboratories
Peg	Polyethylene glycol
Pip	Pipecolyl
SDS	Sodium dodecyl sulphate
TEMED	N,N,N',N'-tetrametylenediamine
Tris	Tris(hydroxymethyl)aminomethane
TSV-PA	*Trimeresurus stenejnejeri* venom plasminogen activator
µg	Microgram
µl	Microlitre
uv	Ultraviolet
v/v	Volume/Volume
w/v	Weight/Volume

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Venomous snakes and their venoms

In order to understand fully the role of enzymes contained within the venoms of snakes it is first essential to consider the significance of the biological role of venoms. Snakes belong to a suborder of the vertebrate class *Reptilia* called *Ophilia, Serpentes*. They are the most easily recognised members of this class by their obvious characteristics of elongation of the body and the absence of limbs. However, other less obvious characteristics have a significant role in the evolutionary development of venoms. Snakes are all strictly carnivorous and, due to their lack of cutting and grinding teeth, must consume their prey whole which is itself a potentially dangerous activity; consequently, they minimise the risk by feeding occasionally on relatively large animals [1].

To allow this unique method of feeding, special adaptations to the skull anatomy have occurred. Snake jaws, unusually for reptiles, are not rigidly attached to the brainpan but are instead connected kinetically to form a highly mobile unit. These adaptations allow the jaw to move independently at either side and consequently widen the mouth considerably. This means that the snake can swallow large, whole prey by using small movements of their skulls over the top of the prey. Since this skull movement has many potential risks the snakes need first to immobilise their prey before they attempt to consume it. Often snakes will choose to be opportunistic feeders, preying on animals that are weak and feeble. However, some method of immobilisation is needed to allow snakes to quieten their prey.

Two potential methods for the immobilisation of prey exist: the majority of snakes make use of constriction, whereby pressure is exerted on the victim by loops of their body encircling the prey leading to strangulation of the victim; the alternative makes use of the injection of a venom. These toxic secretions are produced in specialised exocrine dental glands that are found in the upper jaw of the snake. It is now believed that these venom glands are in fact modified salivary glands and that the venom proteins contained within them has evolved from digestive enzymes.

The classification of the various types of venomous snakes are made according to their various methods of injection of the venom. There are four possible modes of injection, two of which are fairly primitive: the ungrooved (or fangless) and the grooved (or backfanged). These two types of fangs are found in the largest family of

snake, the *Colubridae*. The other snake families, the *Elapidae* and the *Viperidae* have a much more sophisticated method of venom injection. The Elapidae have front-fanged venom apparatus in which there is a larger fang that is almost entirely closed. The Viperidae, including the pit vipers to which the snake venoms studied in this project belong, have the most sophisticated mode of injection. Their skulls are extremely mobile and their fangs, which have a completely enclosed venom canal, are folded back when the mouth is closed. The fangs of the *Viperidae* are therefore able to be much longer due to the structure of the maxillary bones. This group of venomous snakes belongs to the hollow fanged, or solenoglyphous, group. Due to this significant evolutionary development allowing efficient delivery of venom, these snakes can inject their venom and then remove themselves to a safe distance whilst waiting for the venom to take its effect [1,3].

1.2 The pit vipers

As has previously been mentioned, the pit vipers belong to the most sophisticated group of snakes, the hollow fanged or solenoglyphous. They are the group of interest during the course of this project, as the major areas of interest are in the pit vipers of South America and Asia. Principally, the area that has been looked at during the course of this work has been related to the *Bothrops* species of snakes from South America. This group of snakes was first classified using the Linnaeul system of classification as devised by Carl Linnaeus in 1753. Under this classification the species was only identified as existing in one form, referred to as *Bothrops atrox*, with the generic name written first and the specific name second. There were other snakes named under the generic name of *Bothrops*, which are of no interest in this project, except to say that all snakes of this generic name are commonly known as the South American lanceheads.

This classification of the species stood until 1985 when it was recognised by Hoge that the species ought to be further subdivided into five categories. These were decided upon due to differences that began to be recognised in the chemical composition of the snake venoms, the geography of their existence and their markings and scale pattern. This group has been divided as follows:

	B. atrox.atrox (Hoge)
	B. atrox.asper
B. atrox (Linneaus)	*B. atrox.marajoensis*
	B. atrox.moojeni
	B. atrox.pradoi

These names can be shortened *to B. atrox, B. asper B. marajoensis, B. moojeni* and *B. pradoi*. The two enzymes of most importance within this project are derived from the venoms of *B. moojeni* and *B. atrox*. These two are of interest as there is significant uncertainty surrounding the actual differences that exist between these species since it is very difficult to distinguish between them. Although they should come from differing geographical locations there is a possibility that there are some geographical regions were the two species co-exist. This also raises the possibility that some crossbreeds may exist, as the differences are not likely to be great enough to prevent interbreeding. The only real way to distinguish between these two snakes is to look at their scale numbering, as they both have the same markings. Therefore, classifying them whilst they are alive may prove to be impossible. This has obviously lead to some difficulties in verifying the source of the venoms that have been supplied during the course of this project. Due to the fact that prior to 1985 the species had not been subdivided, any work that was carried out before this time has been discredited as the samples used may have been mixed venoms [2].

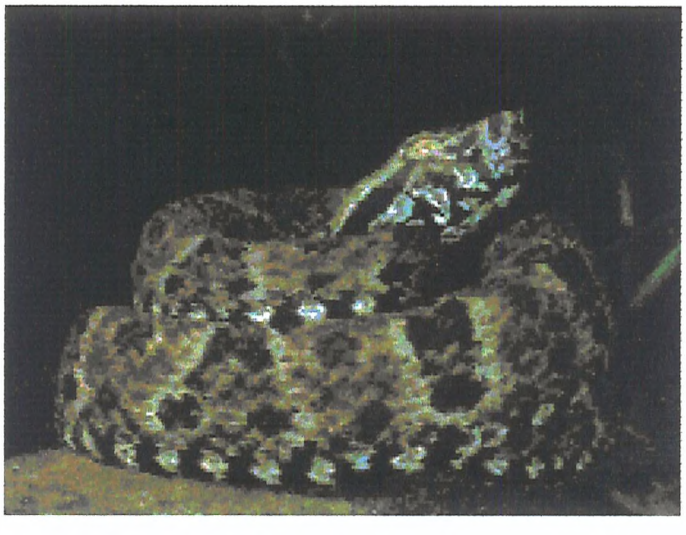


Figure 1.1: Photograph of the South American pit viper, Bothrops atrox

1.3 Composition of Snake Venoms

Snake venoms are highly concentrated secretory fluids containing between 15 and 50% solid matter dependent upon the species. The venoms of the *Bothrops* species of snakes seem to contain 24 - 40% solid matter. Over 90% of the solid snake venom components are proteins, which have a variety of the biological functions involved in the immobilisation of the prey. The remaining 10% of the solid matter are made up of non-protein components, including inorganic anions, cations, amino acids, small peptides, lipids, nucleosides, nucleotides, carbohydrates and amines. The protein components of the snake venoms are made up of both enzymatic and non-enzymatic factors. A very large variety of enzymes exist within snake venoms, many of which seem to target tissue proteins, polysaccharides, plasma proteins and cell membranes. These enzymes include phospholipases, phosphodiesterases, nucleosidases and elastases, as well as the thrombin-like enzymes implicated in coagulation [3].

As can be seen from this very wide ranging, diverse assortment of enzymes and enzymatic targets there are a whole variety of ways in which nature has evolved in order to bring about the immobilisation of the prey. In actual fact, no single method is relied upon, as a whole cocktail of enzymes with a whole variety of tissue targets exist in each venom which contribute to the downfall of the prey. Other non-enzymatic components, both protein and inorganic, also contribute to the action of the various venoms. These also vary according to the species of snake from which they are isolated, and there is also some evidence that there may even be some variation within the species.

Non-enzymatic protein components of snake venoms include:

Neurotoxins
- Low molecular weight proteins of ~ 7kDa
- Two groups: post-synaptic blocks and pre-synaptic blocks
- Result in cramps, convulsions and paralysis

Cytotoxins - 60-62 amino acids
- Result in haemolysis, cytolysis, cardiotoxicity and muscle membrane depolarisation

Myotoxins - 42-50 amino acids
- Result in skeletal muscle contracture

Cardiotoxins - 21 amino acids
- Vasoconstriction of coronary vessels leading to cardiac arrest

Other components are choline esterase inhibitors, phospholipase inhibitors, nerve growth factors, lectins, proteins affecting platelet function and proteins that act upon the complement system.

The non-proteins are composed of both organic and inorganic constituents. The organic components include lipids, carbohydrates, riboflavin, nucleosides, nucleotides, amino acids and biogenic amines. The inorganic components include cations and anions. The specific role of each of these components is as yet not fully understood. However, in actual fact, it would appear that in many cases the role might be in preservation of the venom and in establishing the correct environmental balance in which the enzymatic components can function most effectively.

The diversity of the composition of the snake venoms makes it virtually impossible to investigate the activity and functioning of the venom as a whole and therefore specific components are investigated in isolation. The areas for probing investigation are determined by the way in which the immobilisation occurs in the prey of the snakes that are being investigated. In the case of the vipers and pit vipers, the major cause of death is from circulatory shock. This is normally associated with many events, and can vary from snake to snake and species to species. This is due to the fact that there is now a significant feeling that the venom components interact in a species-specific manner. In other words, the venoms act differently on different prey and are evolved to be most efficient against the animals that are the natural prey of the snake. Many other factors, including the age of the snake, the geography in which the snakes exist, along with the potential individuality of each snakebite, may affect the outcome of a single snakebite. However, since it is known that attacking the

cardiovascular system of the prey is the most efficient and direct method of immobilising the animal by death, an investigation into the enzymes involved in this area and the way in which they interact with the prey is considered to be vital.

In this project, the area of interest is the enzymatic factors involved in targeting of the proteins in the haemostatic process, with specific emphasis on the thrombin-like enzymes. In order to investigate fully the role of thrombin-like enzymes derived from snake venoms, it is first important to understand the action of the haemostatic process that the enzyme is designed to disrupt and, most importantly, the mammalian thrombin that the activity is based upon [1,3].

1.4 Blood as a Target for Snake Venom

Blood is a target for the action of many venoms derived from both snakes and other animals. The role of blood is of vital importance for a number of reasons, most vitally in the transportation of oxygen from the lungs to the tissues for use in cell metabolism and the removal of carbon dioxide from these tissues for excretion at the lungs. Oxygenation of the tissues is of major importance and must be maintained undisrupted in order to ensure the healthy functioning of the animal [5,6]. Consequently any disruption in the process can be fatal. Therefore a sophisticated mechanism has been evolved by which the supply of oxygenated blood to the tissues can be maintained in the event of injury. This mechanism is part of the haemostatic and fibrinolytic system, where haemostasis is the term that describes any processes that are activated by vascular injury which lead to the occlusion of damage and the prevention of bleeding, and fibrinolysis is any process that leads to the breakdown of fibrin clots produced during the course of haemostasis. The two processes of haemostasis and fibrinolysis must be kept in a constant balance to ensure that a constant blood flow is maintained at all times. It is this property of blood flow that makes the blood a suitable target for the action of so many venoms [4,5,6].

1.5 Haemostasis

Haemostasis can be divided into two categories: primary haemostasis and blood coagulation. Primary haemostasis is the initial activity of the body to prevent blood loss by means of vasoconstriction of the damaged vessel, resulting in a decrease in the size of the injury, and subendothelial exposure to blood resulting in an accumulation of platelets at the site of injury [6,7,8]. The subsequent adhesion and aggregation of platelets and their release of biologically active components result in the further accumulation of platelets, the formation of a 'bung' and the activation of the second part of haemostasis, blood coagulation*(figure 1.2)*.

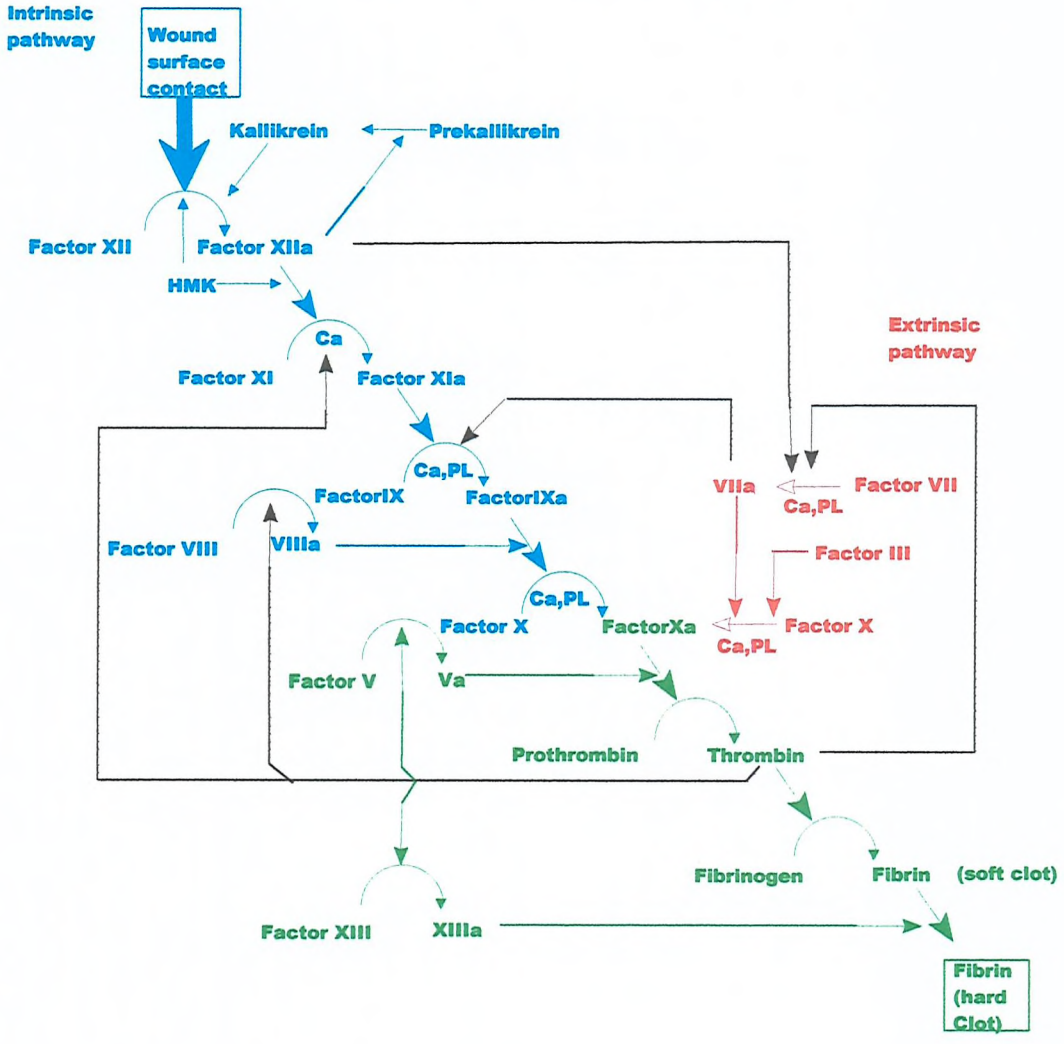


Figure 1.2: General overview of coagulation.

This can occur *via* two different pathways, the *intrinsic pathway (figure1.3)* and the *extrinsic pathway (figure1.4)*. The intrinsic pathway is a longer and slower method of bringing about coagulation in which all the factors needed to cause coagulation are circulating in the blood at all times[9,10,11]. Contact of the blood with a glass microscope slide, as well as contact with irregular surfaces within the blood vessels can activate this process. It is therefore this method that is thought to be responsible for the formation of thrombosis in patients with excessive fatty deposits in their arteries[12,13,14]. The extrinsic system is a much more rapid and direct method in which blood coagulation occurs. This system is triggered by vascular injury and is therefore more rapid as it requires an immediate prevention of blood loss from the site of injury [5,10]].

1.5.1 Intrinsic Pathway

As mentioned above, the intrinsic pathway of coagulation is a relatively slow process in which all the components necessary for coagulation are already circulating around in the blood at all times. This means that the coagulation does not need to be triggered by a vascular injury; thus it is this process of coagulation that is responsible for the formation of thrombosis due to interaction of the blood with fatty deposits in the arteries[15,16,17]. Factor XII, kininogen, prekallikrein, and factor XI bind to the surface of the subendothelium at the site of the vascular injury,[15,18,20,23,24,25] triggering the intrinsic pathway. On adsorption to the subendothelium by factor XII there is exposure of the active protease factor XII_a [17,18,19]. Factor XII_a cleaves prekallikrein and factor XI,[20] resulting in the formation of the active enzymes kallikrein and factor XI_a [21,22] . The active enzyme kallikrein is responsible for the activation of factor XII in a negative feedback mechanism,[23,24,25] and the active factor XI_a goes on to activate the change of factor IX into factor IX_a which is located on the surface of the aggregated platelets[26,27]. This interacts with factor VIII:C, [28,29,30]to form a complex which converts factor X into the active proteinase factor X_a [31,32]. This platelet bound factor binds to the cofactor V_a to form the complex that is responsible for the conversion of prothrombin into thrombin[37,38]. Thrombin is responsible for the conversion of fibrinogen into the insoluble coagulant fibrin[39,4042,43]. The entirety of this mechanism is carried out by the conversion of

inactive precursors, or zymogens, into active proteolytic enzymes through a cascade system [5,10].

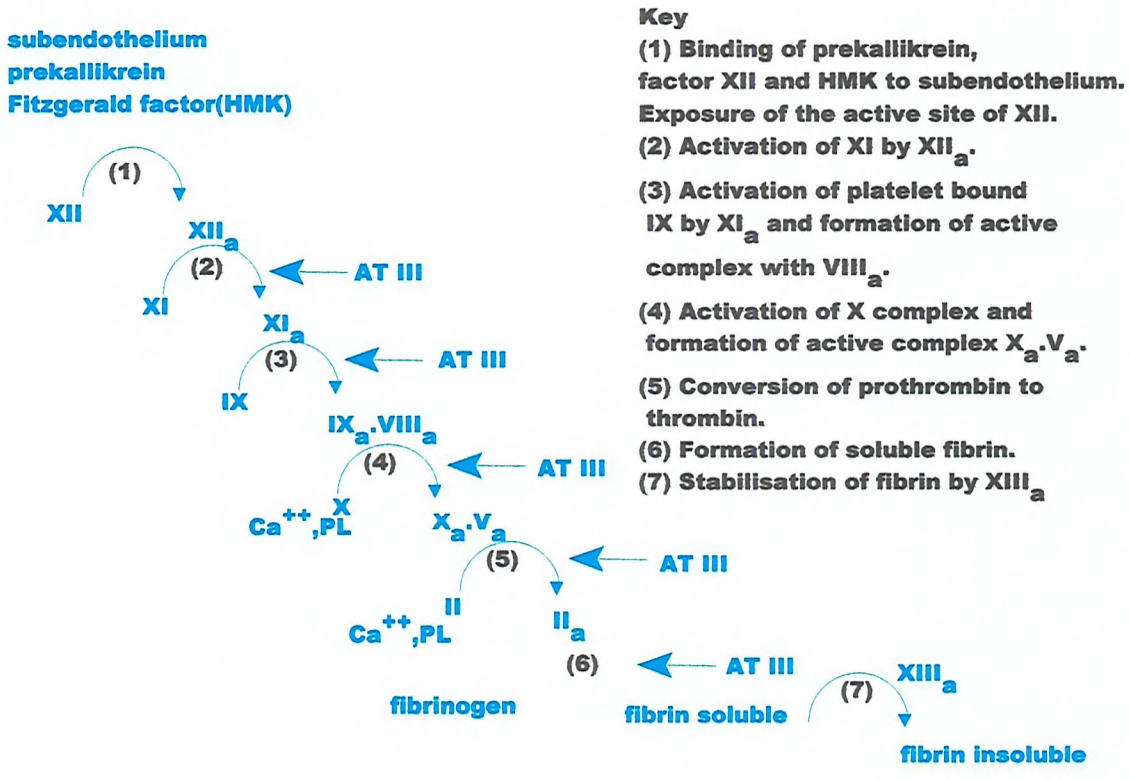


Figure 1.3: The intrinsic pathway of blood coagulation.

1.5.2 Extrinsic Pathway.

The alternative method by which thrombin can be generated is by the extrinsic pathway. This is activated by the generation of the tissue factor thromboplastin, which is a phospholipoprotein originating from disrupted tissue cells[33,34]. This then binds to factor VII, which undergoes a conformational change exposing the active serine proteinase factor VII_a[35]. This then goes on to activate the platelet bound factor X into X_a,[36] which interacts with factor V to form the complex that converts prothrombin into thrombin [5,37,38].

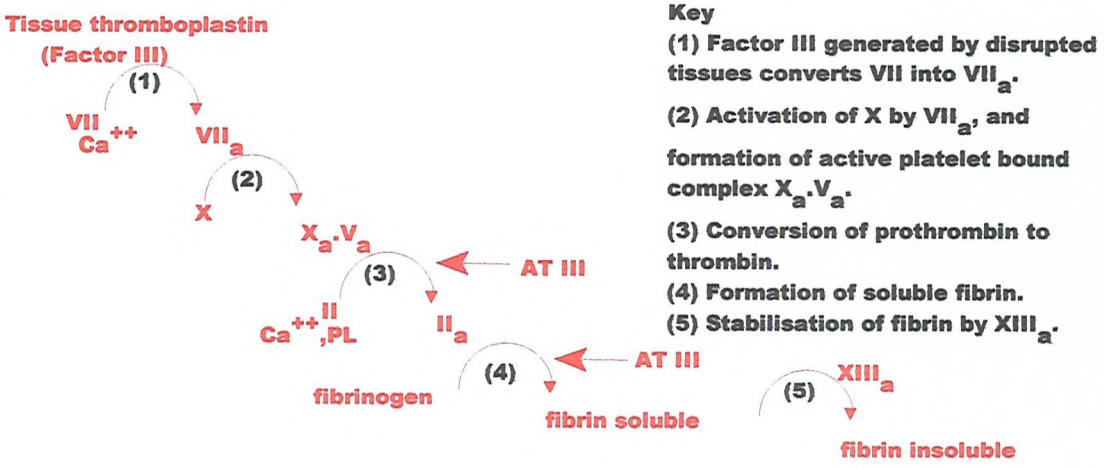


Figure 1.4: The extrinsic pathway of blood coagulation.

1.5.3 Fibrin Formation.

Fibrin is formed from fibrinogen by the action of thrombin [39,40,43,44]. Fibrinogen is a large molecule of 340kDa, which is made up of three pairs of differing chains [6,7,41,42]. There are two sets of Aα, Bβ and γ- chains that are linked together by disulphide bonds, as can be seen from *(figure 1.5)*[7].

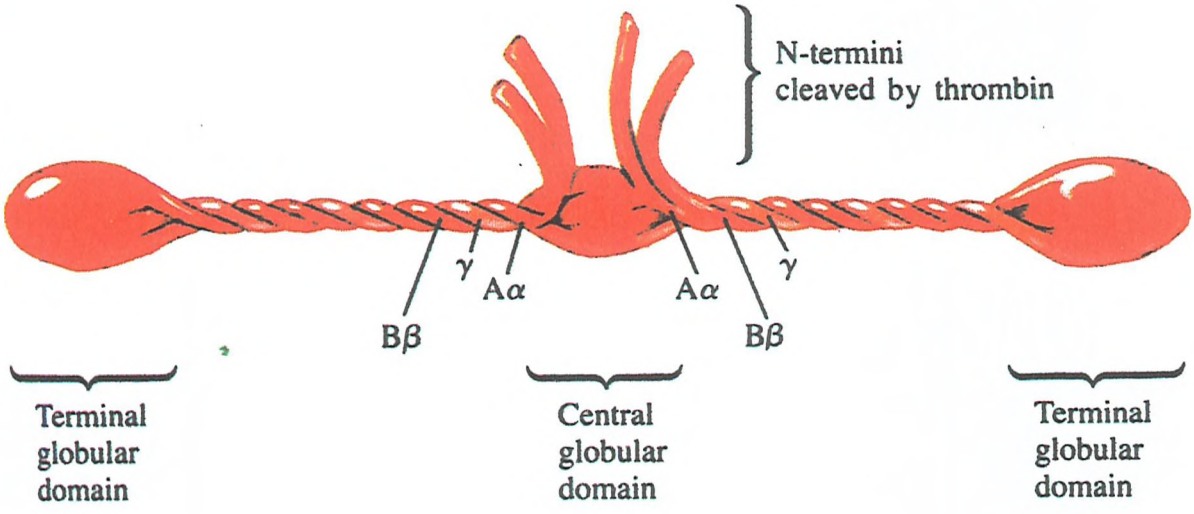


Figure 1.5: The structure of fibrinogen.

The thrombin cleaves the fibrinogen at the Aα-chain to yield fibrinopeptide A and the fibrin I monomer by breaking the Arg-16-Gly-17 bond[39,40,43,44]. The second cleavage breaks the Bβ-chains to form fibrinopeptide B and the fibrin II monomer, by the cleavage of the Arg-14-Gly-15 bond, as can be seen in *figure 1.6* [7].

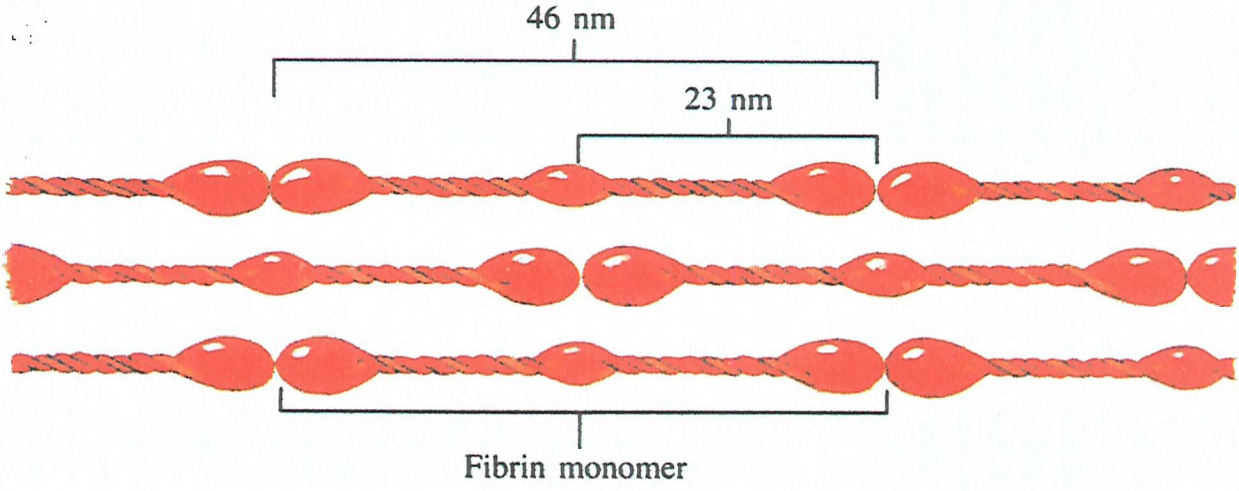


Figure 1.6: Structure of fibrin monomers, and the subsequent stacking of monomers to produce the fibrin polymer.

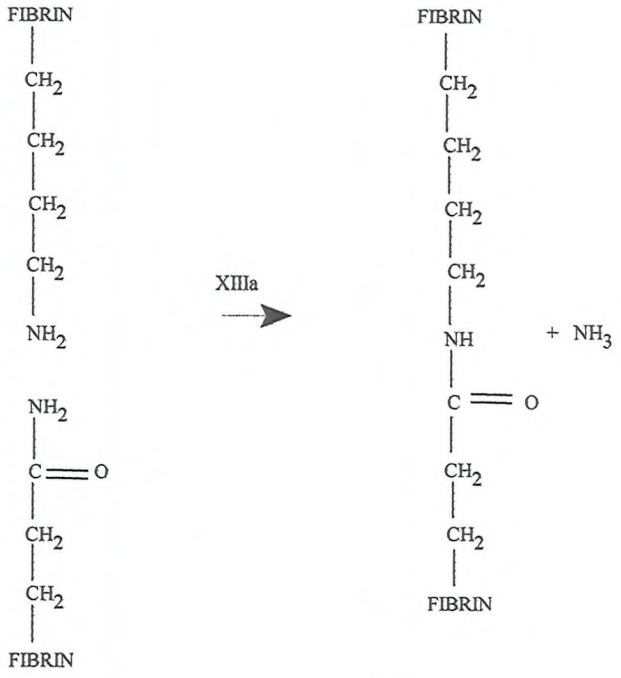


Figure 1.7: Cross-linking of fibrin by factor $XIII_a$.

The fibrin II monomers then stack up and through the specific transamidase action of factor $XIII_a$ Υ-glutamyl-ε-lysine covalent bridges are formed between adjacent fibrin II monomers, as is shown in *figure 1.7* above [7]. The covalent bonds produced by the fibrin stabilising factor result in the formation of the stable fibrin clot, as seen in *figure 1.8* [45,46].

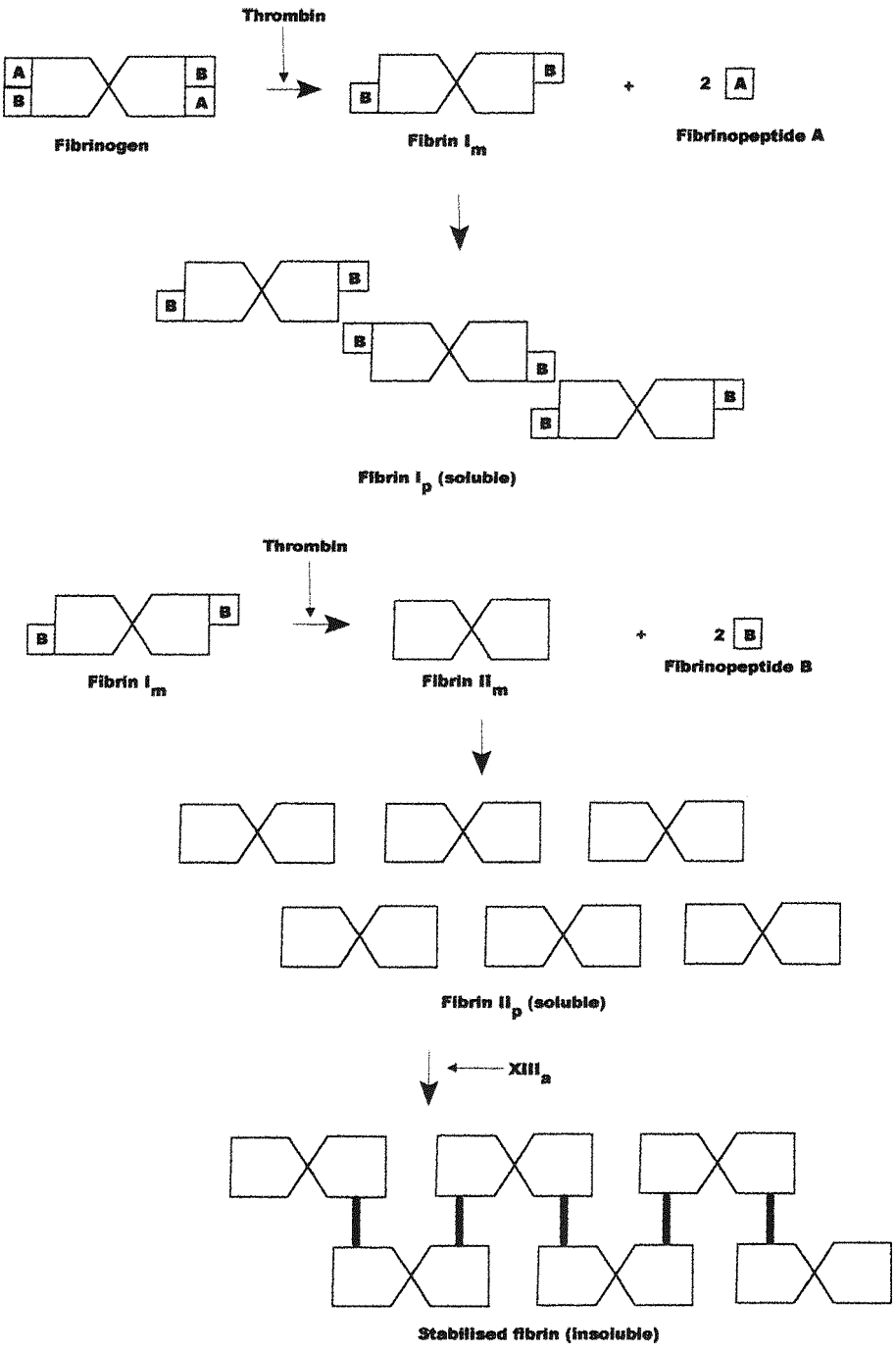


Figure 1.8: Action of thrombin on fibrinogen.

Factor number	Name	Properties and functions
I	Fibrinogen	MW 340 000; six chains: $(A\alpha)_2(B\beta)_2(\gamma)_2$. Converted to fibrin by removal of 2A and 2B peptides.
II	Prothrombin Thrombin is II_a	MW 70 000; 582 amino acids with 12 disulfide bridges. Converted to thrombin (factor II_a) by factor X_a.
III	Tissue factor, thromboplastin	Acts with factor VII_a to activate factor X.
IV	Calcium ions	Activate many enzymes in the blood clotting cascade.
V	Proaccelerin, V_a is accelerin	Factor V_a accelerates conversion of factor II to factor II_a by factor X_a.
VII	Proconvertin	Activated by tissue trauma, part of extrinsic system.
VIII	Antihemophiliac factor	MW $> 10^6$. Acts with factor IX_a to activate factor X. Its absence causes hemophilia A.
IX	Christmas factor	MW 55 000. Its absence causes hemophilia B.
X	Stuart factor	MW 55 000; composed of two polypeptide chains. Activated by factor IX_a, $VIII_a$ and Ca^{2+} ions, or by factor VII_a, factor III_a and Ca^{2+} ions.
XI	Plasma thromboplastin antecedent	MW 124 000; composed of two similar chains linked by disulfide bonds. Activates factor IX.
XII	Hageman factor	MW 76 000. First factor in the intrinsic pathway. Activated by factor XII_a and by contact with abnormal surfaces.
XIII	Transglutaminase	MW 365 000. Cross-links precipitated fibrin monomers.

Table 1.1: Plasma proteins involved in haemostasis and fibrinolysis

1.6 Fibrinolysis

Once haemostasis has been completed, and the vascular injury it was initiated for has been repaired, it is imperative that the process of haemostasis is halted in order to prevent any additional damage[5]. Therefore the fibrinolytic system has evolved. The major role of this system is to lyse the fibrin polymers into fibrin degradation products (FDPs) which can then be excreted *via* the kidneys. By removing the excess fibrin from the system any risk of thrombosis as a result of clot formation is removed. It is believed that this system of regulation of haemostasis is in existence within the system at all times, and that there is an ongoing equilibrium between the two pathways.

The mechanism of fibrinolysis is activated by the adherence of plasminogen, the plasma zymogen, to the fibrin polymer where it is cleaved by either tissue-type plasminogen activator (tPA) or by urinary-type plasminogen activator (uPA) to form the active proteolytic enzyme plasmin[47]. tPA is released from the endothelial cells of the vascular wall and uPA is released by a mechanism involving proUK , factors XII, kininogen, prekallikrein and activated protein C[48,48]. It is the ensuing plasmin that is responsible for the breakdown of fibrin to FDPs[47,48,49].

Fibrinolysis is regulated by inhibitor α_2-macroglobulin. This is a broad-spectrum proteinase inhibitor which acts upon plasmin, along with the serpins, plasminogen activator inhibitor (PAI-1) which acts upon tPA and uPA, C_1-inhibitor which activates kallikrein and α_2-antiplasmin which acts upon plasmin[49]. Consequently, a highly effective method for the regulation of both haemostasis and fibrinolysis has developed in evolution in an attempt to protect this highly vulnerable cardiovascular system. *(figure 1.10).*

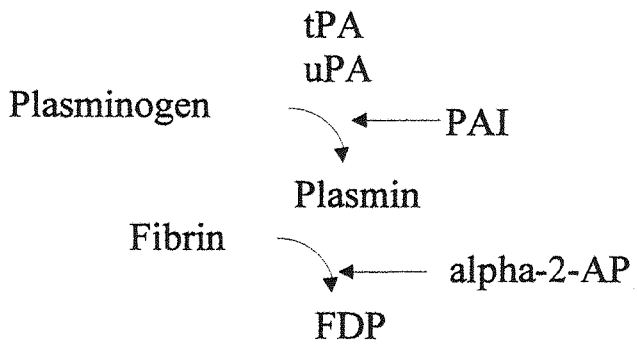


Figure 1.9: Fibrinolysis.

1.7 The proteases

The family of enzymes known as the proteases can be divided into four groups that are subdivided according to their activities and to their functional groups. The groups are the serine proteases, the carboxyl or aspartyl proteases, the thiol proteases and the zinc or metallo proteases. These enzymes, with the exception of some of the metalloproteases are endopeptidases each with a low molecular mass from 15 to 35kDa, with the exception of leucine aminopeptidase, which has a mass of 250kDa. All these enzymes catalyse fundamentally the same reaction, the hydrolysis of peptide bonds, but they all take advantage of different mechanisms in order to achieve this aim. The serine proteases have an active serine residue that has an optimum pH around neutrality or in the slightly basic range. The thiol proteases have a similar activity to the serine proteases except they have an active cysteine but equally they have a neutral pH optimum. The zinc proteases are metalloenzymes that have a lower pH optimum. The final group, the carboxyl or aspartyl proteases, differ in that they require a low pH for optimum activity due to their catalytically active carboxylates [51,52].

1.7.1 The serine proteases

The serine proteinases are perhaps the most extensively studied of all the protease family. Since this is the group of proteases of specific interest in this project, a more detailed description of the mechanism and the experiments leading to the elucidation of the mechanism will be discussed in the next section. There are three catalytically active residues within the structure of all serine proteinases: Ser-195, from which the name serine protease is derived, His-57 and Asp-102 (these refer to the chymotrypsin numbering). The substrate binds to the pocket within the enzyme between the two domains of the structure and binds from the N-acylamino chain to the carbonyl group of Ser-214. Other residues within the N-acylamino chain bind to sites within the pocket that are available. The reactive carbonyl of the polypeptide substrate is thus positioned between the backbone NH groups of Ser-195 and Gly-193 with its oxygen residue between the two.

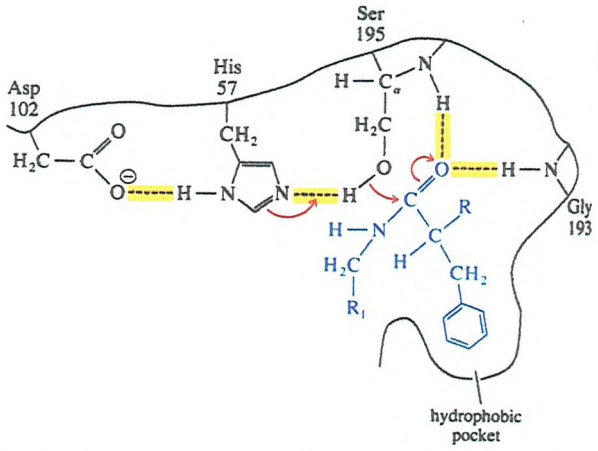


Figure 1.10: Positioning of substrate within the active site and formation of the tetrahedral intermediate by the action of the nucleophilic serine.

A tetrahedral intermediate is then formed *via* reaction of the hydroxyl group of the Ser-195 with the carbonyl carbon of the substrate. As this is occurring, the proton on the hydroxyl group is transferred to His-57, by a mechanism that will be discussed later. The bond between Ser-195 and the substrate formed results in the lengthening of the C=O to form a single bond, and thus this negative oxygen can move closer to the NH of Gly-193, forming a much shorter and stronger hydrogen bond. The tetrahedral transition state then collapses, resulting in the formation of an acylenzyme intermediate and the loss of the leaving group.

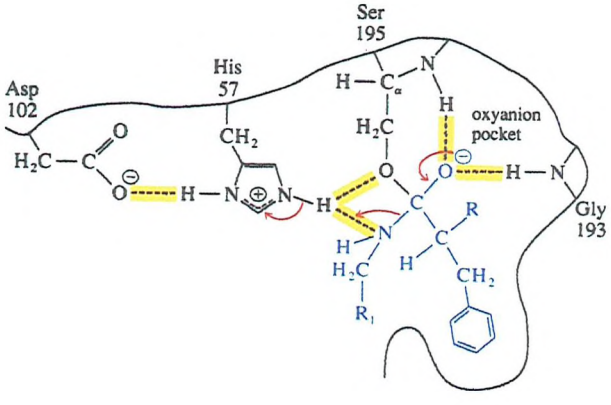


Figure 1.11: The collapse of the tetrahedral intermediate.

Deacylation then occurs, again by a mechanism that will be discussed in more detail later, *via* the activation of water molecules by the charge relay system within the active site.

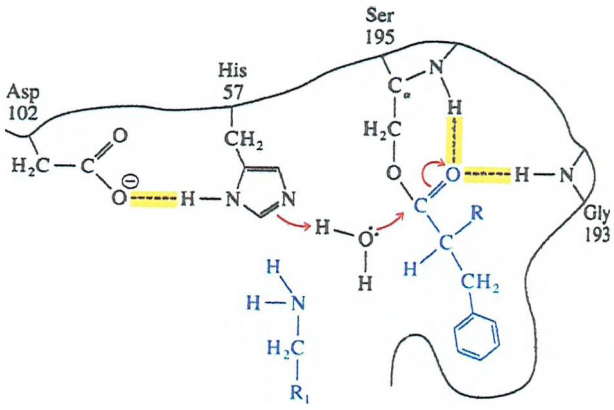


Figure 1.12: Attack by water molecules in the surrounding solution.

Once again a tetrahedral intermediate is formed; however, this collapses releasing Ser-195 and forming the enzyme-product complex.

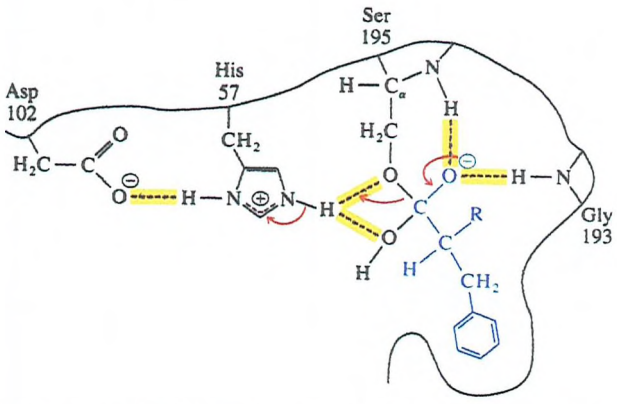


Figure 1.13: Mechanism for the collapse of the tetrahedral intermediate.

Serine proteinases are responsible for the hydrolysis of ester or amide substrates, with many bonds being the potential targets for cleavage. The significant

specificity for substrates that seems to exist for some of the enzymes within this group appears to be determined by the binding sites of the enzyme. Consequently enzymes like trypsin, which have evolved to have very little specificity for particular proteins due to its role in gastric digestion, show the same catalytic site as more specific enzymes but significant differences in binding and recognition. Like many other proteases serine proteases exist in an inactive zymogen precursor form. However, they are different in that they are enzymatically cleaved by proteins belonging to the same serine protease group and consequently these enzymes act biologically together, often in a cascade form, with one activating another [7,51,52].

1.8 Action of serine proteases

All the enzyme factors involved in the blood coagulation cascade belong to the group of enzymes known as serine proteinases. These are a family of very closely related enzymes, which are characterised by the presence of an active serine residue within the active site. Other enzymes that belong to this group of enzymes include pancreatic enzymes, for example trypsin and chymotrypsin. The original serine protease to be studied was chymotrypsin, although many of the family have been considered to date. However, it is still useful to understand the mechanism of action of chymotrypsin when considering the activities of thrombin and thrombin -like enzymes which belong to the serine proteinase family. The unusual activity of the serine residue within the active site was first demonstrated on chymotrypsin in 1954 by B.S Hartley and B.A Kilby,[53] by studying the action of the artificial substrate *p*-nitrophenyl acetate. The original results showed the existence of an acyl-enzyme intermediate. This was shown by monitoring the two distinct stages of the reaction by following the formation of p-nitrophenolate which is yellow in colour (*figure 1.14*)[7].

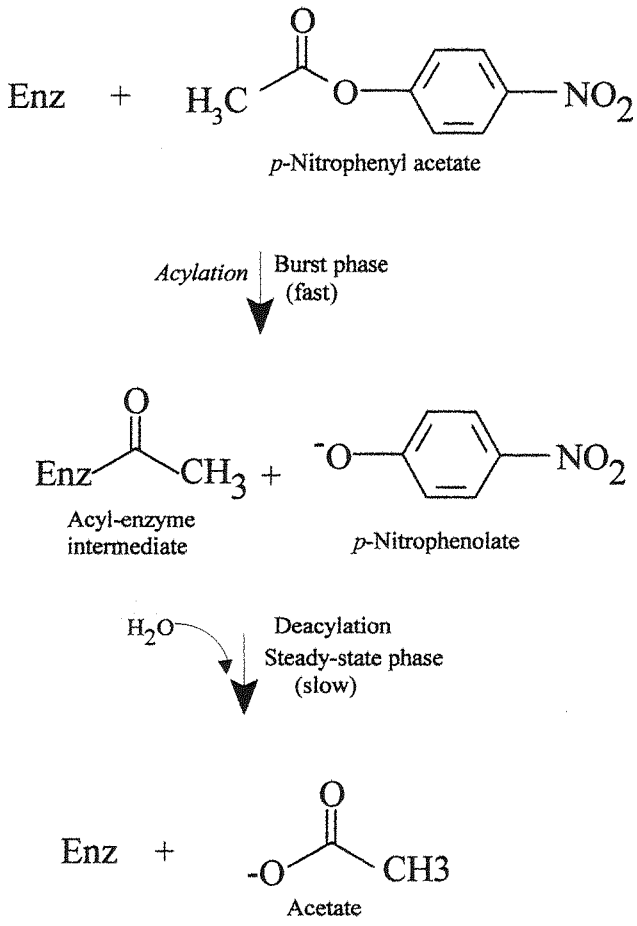


Figure 1.14: Chromogenic reaction of serine proteinase.

The highly active serine residue was identified when the acyl- enzyme intermediate was isolated. The amino acid analysis showed that the acyl group was attached to serine 195. Later it was shown that, despite the fact that there are 27 serine residues within the structure of chymotrypsin, only Ser-95 reacts chemically with the nerve toxin diisopropylfluorophosphate (DFP) showing the unique reactivity of this particular residue *(figure 1.15)*[7,54].

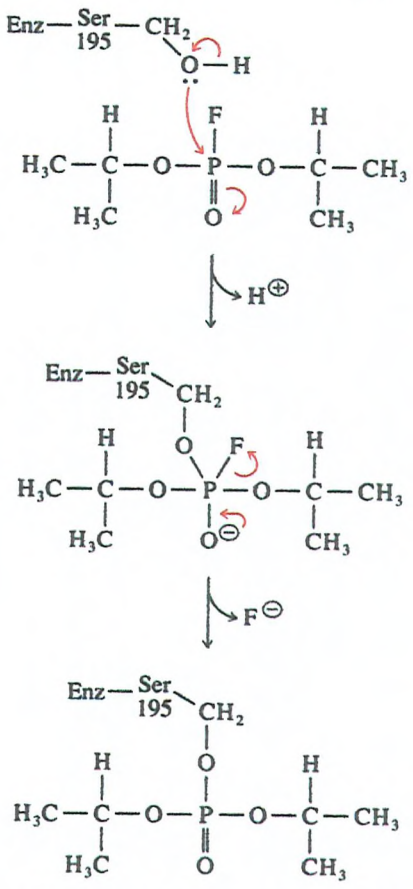


Figure 1.15: Reaction of DFP with serine proteinase.

Further studies showed that within the active site of the serine proteinases there are other active residues involved with the reaction, as was evident with the action of chymotrypsin with N-tosyl L-phenylalanyl chloromethyl ketone (TPCK)[56]. It was expected that this reagent would interact with the active serine residue. However, studies showed that it actually reacted with histidine-57, indicating that two nucleophilic residues existed within the active site *(figure 1.16)* [7].

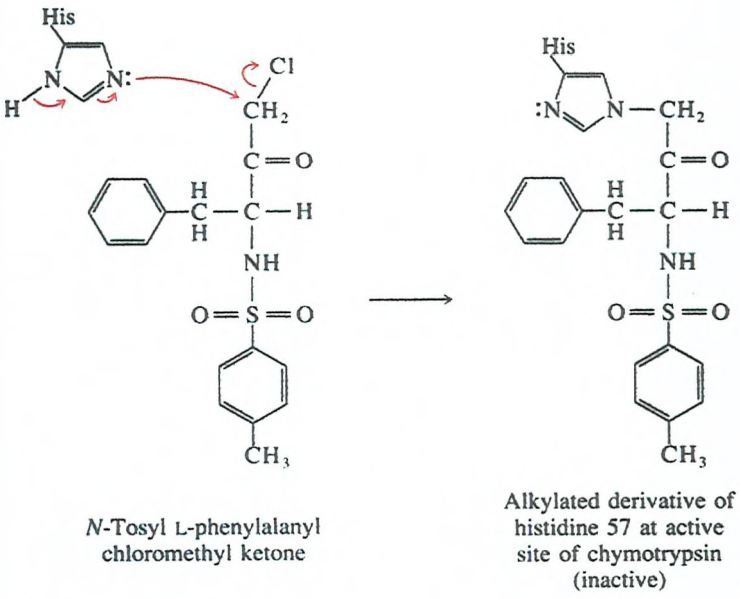


Figure 1.16: Reaction of His 57 with TFCK.

X-ray crystallography of the active site showed that the imidazole ring of His-57 lies 0.38nm from serine 195[56]. It was also identified that another amino acid residue, aspartate 102, was also found to lie close to the histidine residue *(figure 1.17)* [7]. It has since been shown that all three of these residues are of importance in the activity of this group of enzymes and that they are all found within the active site of all serine proteases *(figure 1.17)* [7].

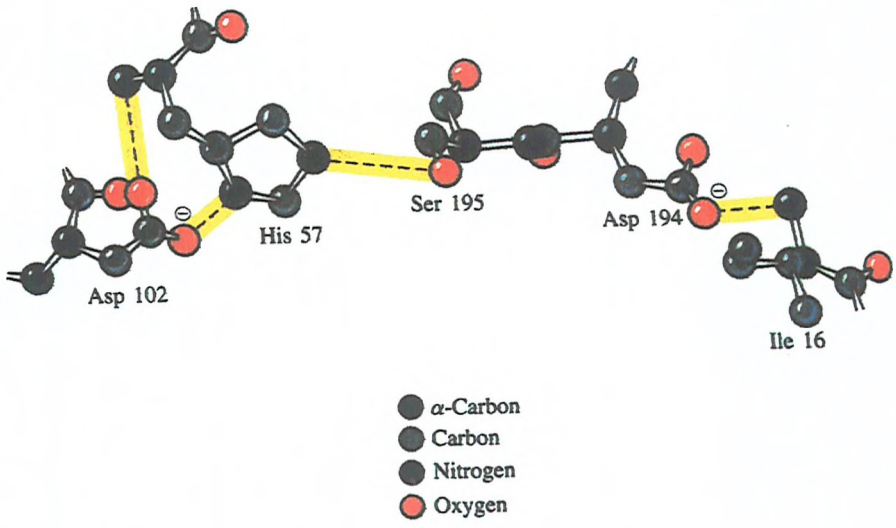


Figure 1.17: The active site residues of the serine protease chymotrypsin.

The three amino acids work together in a charge relay network in order to increase the nucleophilicity of the active serine residue[57]. This works by the passage of electrons around the three amino acids, effectively increasing the negative charge on the serine (*figure 1.18*) [7,51,52,57].

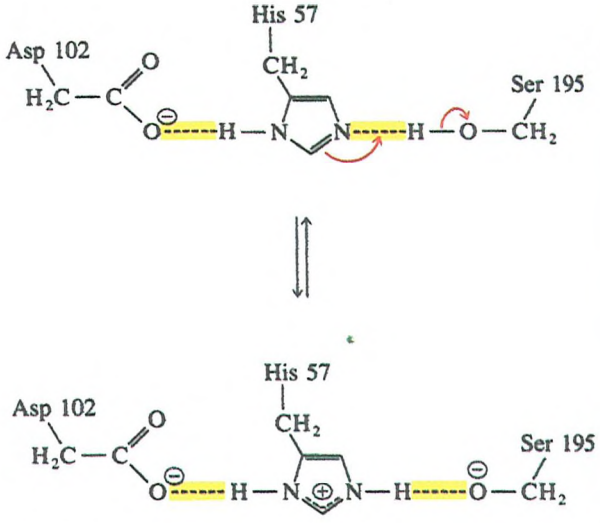
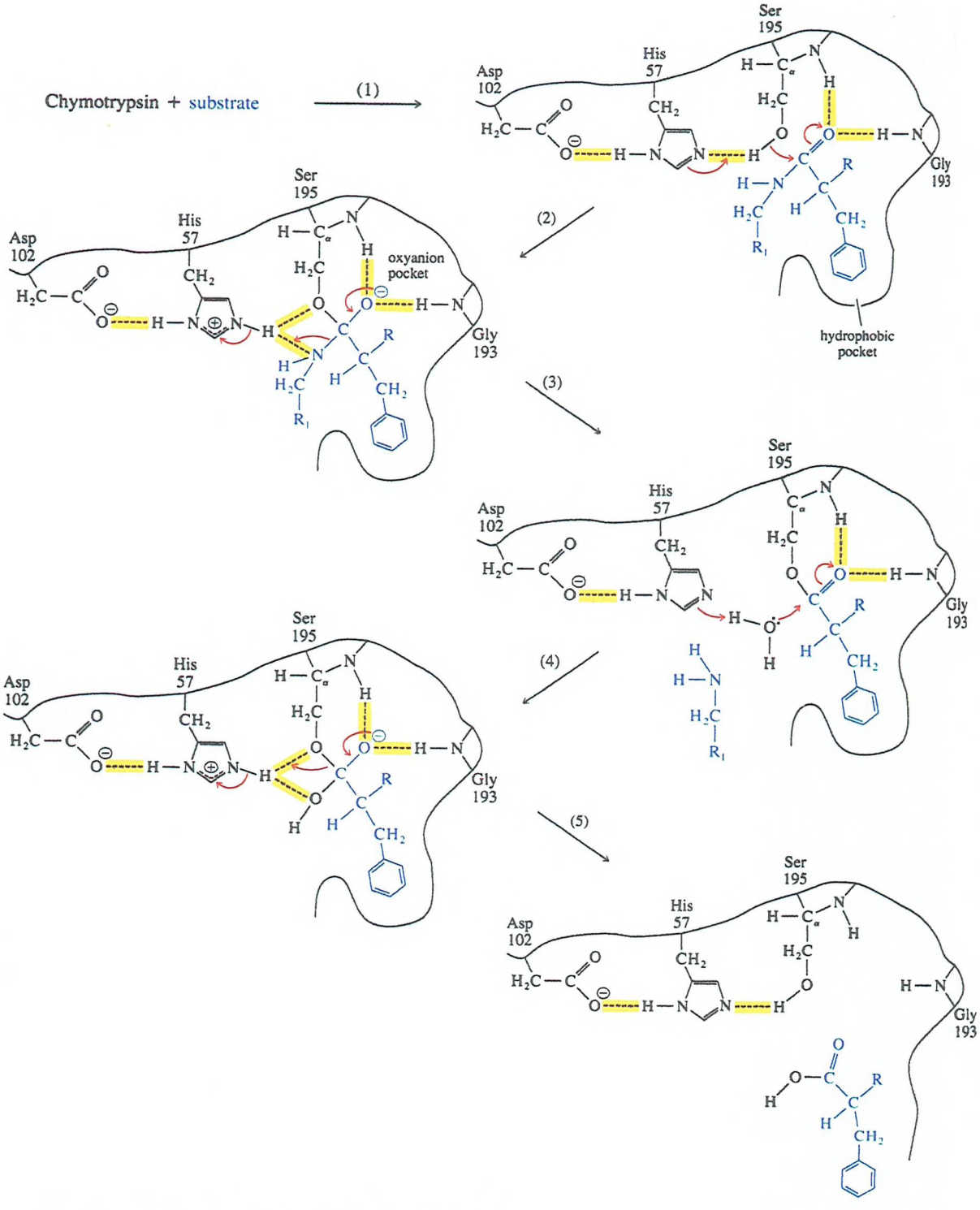


Figure 1.18: Charge relay mechanism of the active site of serine proteases.

1.9 Catalytic action of serine proteases

1.9.1 General mechanism of serine proteases

From the experimental evidence shown in the previous section, it is possible to determine the way in which the overall mechanism occurs. As has already been mentioned, chymotrypsin is the serine proteinase that has been most extensively studied, due to the fact that it has always been readily available in abundance, and that it was consequently relatively easy to purify large quantities of it [7]. Therefore, a general mechanism for this group of enzymes has been produced based largely on the mechanism of chymotrypsin, trypsin and kallikrein. It is this general mechanism which will be detailed in this section[58,59]. One point that is important when using this as a model for the action of specific enzymes of this general group, is the fact that it does not in any way look at differences in the specificity of bond cleavage resulting from variations in substrate specificity (*fig: 1.19*) [7,60].



1. Formation of enzyme-substrate complex.
2. Formation of tetrahedral intermediate by nucleophilic action of serine.
3. Cleavage of scissile bond to form acyl-enzyme intermediate.
4. Formation of tetrahedral intermediate by hydrolysis of acyl-enzyme intermediate.
5. Collapse of tetrahedral intermediate, reforming active enzyme.

Figure 1.19: Mechanism of serine protease, illustrated by chymotrypsin.

1.9.2 Substrate specificity of serine proteases

The individual specificity of serine proteases for particular bonds is due at least in part to the binding of residues around the scissile bond to the enzyme in question. This was first recognised by Schechter *et al* who through the investigation of the papain, an aspartate protease derived from the papaya, showed that amino acids on the substrate either side of the scissile bond were important in substrate recognition. The specificity of the enzyme was derived from the size and charge on binding sites within the active site; these binding sites, called subsites, are numbered S1, S2, etc. on the N-terminus of the catalytic site, and S'1, S'2,etc. on the C-terminus side. The respective sites on the substrate are termed P and P' sites, as shown in *figure 1.20* [61].

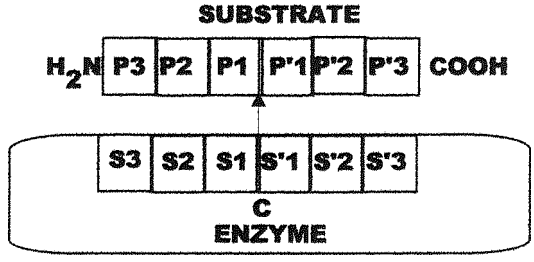


Figure 1.20: Subsite binding specificity.

Within the structure of each serine protease are specific differences in the subsite residues, which are responsible for the differences in recognition of the substrate. In particular, there are huge differences in the residues found at the S 1 site which correlate to differences in the bond that is cleaved, as shown in *figure 1.21* [7]. For example, chymotrypsin has a hydrophobic specificity pocket at the S 1 site which means that it will cleave bonds on the substrate that have an aromatic residue at the P 1 position. Trypsin, and consequently thrombin and batroxobin, have an aspartic acid residue at the S 1 site resulting in cleavage when a positively charged residue like lysine or arginine is in the P 1 site. On the other hand the S 1 binding site of elastase is blocked off by valine and threonine, resulting in elastase only being able to bind small groups at the P 1 site, for example glycine and alanine. It would be of great

interest to look at the specificity of batroxobin at these important binding sites. Therefore this will be considered in Chapter 4, where a series of peptides are designed and tested to probe this area.

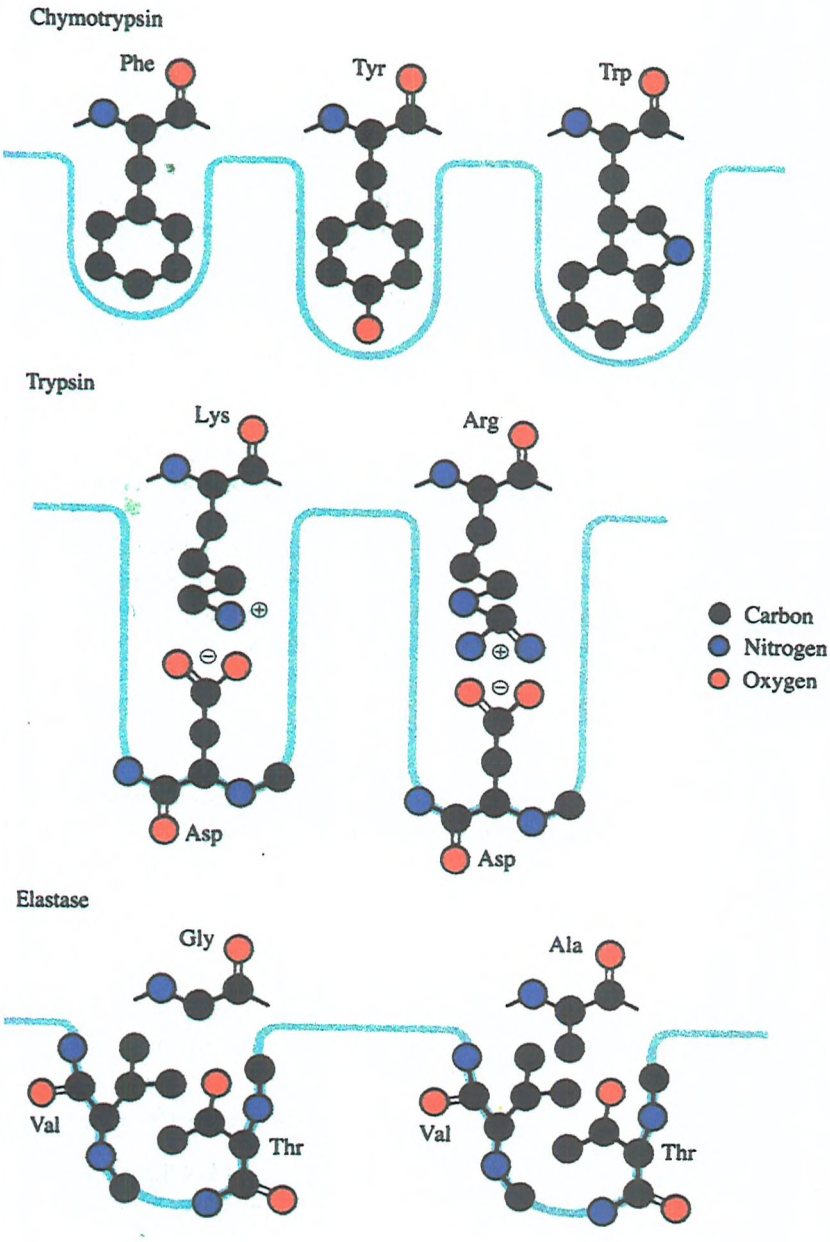


Figure 1.21: Hydrophobic specificity pocket of three different serine proteases.

1.10 Pit viper snake bites

The snake venoms considered in this project belong to the *Viperidae* group of snakes and, accordingly, an exploration of the bite produced by these snakes and the effect it has on the coagulation system is needed. As mentioned previously, circulatory shock is the major cause of death for victims of pit viper bites. One of the major causes of instantaneous death is by haemorrhaging into major organs. This is caused by direct damage of the capillary endothelium by haemorrhagic venom components.

However, this does not actually affect the coagulation system itself, and in fact when humans are bitten by pit vipers there is almost always a relatively small amount of venom injected into the system. In this case the major problems are actually concerned with the coagulation system, and a response known as the 'defibrination syndrome' occurs. This response is a disturbance of the coagulation system, which is brought about by the activation of the blood coagulation cascade. In this way, the fibrinogen is converted into fibrin and then broken down by the fibrinolytic system to form fibrin degradation products (FDP). These products are then excreted through the kidneys, resulting in severe hypofibrinogenaemia, which often leads to near-zero levels of fibrinogen in the blood.

Under these circumstances, although the low levels of fibrinogen will not in themselves cause haemorrhaging they can increase bleeding when the haemorrhagins within the venom cause epithelial cell damage. Many patients exhibit signs of renal failure due to fibrin deposits causing glomerular damage. In cases where there is an indication of such severe systemic envenomation the only treatment is with an antivenom (*figure: 1.22*)[1].

1.11 Components within venom affecting haemostasis

There are many points along the haemostatic system at which the components of the snake venom can act in order to affect the cardiovascular system. Therefore there are many ways in which the components of venoms may act in order to bring about the circulatory shock that was mentioned as the primary cause of death in pit viper bites.

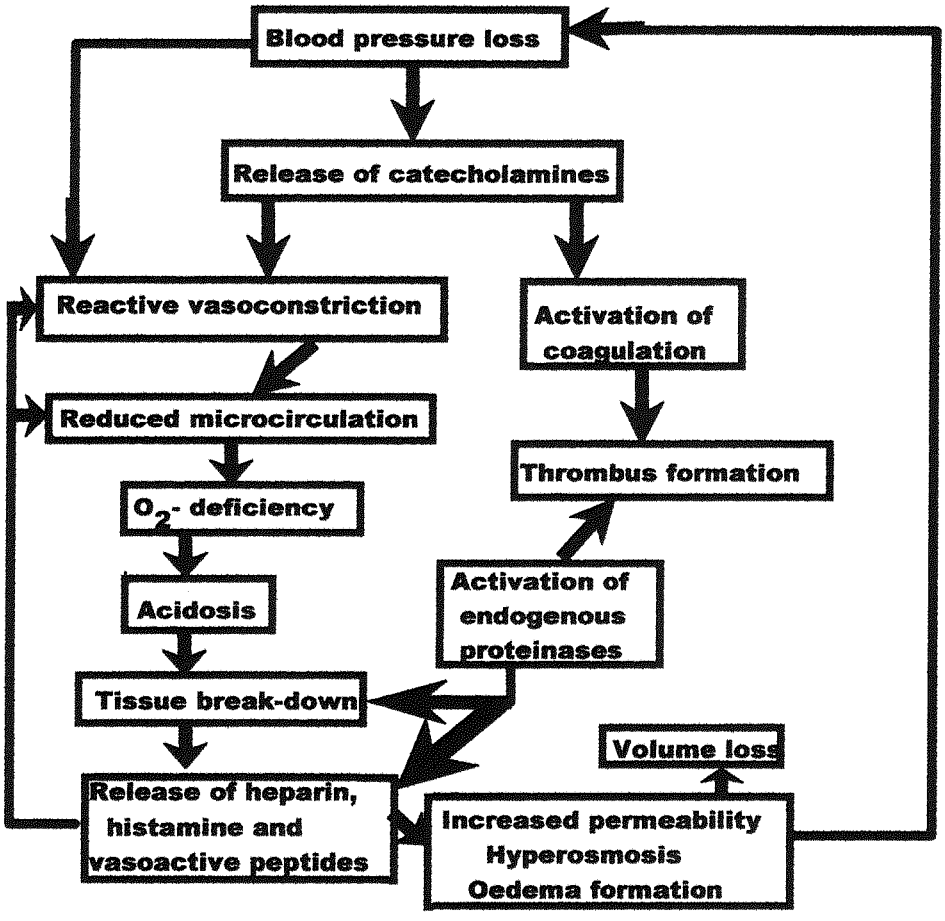


Figure 1.22: The onset of circulatory shock as a result of envenomation

1.11.1 Blood platelets and components affecting them

Platelets are of vital importance in the initiation of primary haemostasis. Therefore, any factors within the venom, which alter the way they act naturally, will have a disruptive effect upon the haemostatic system. There are components that act upon the aggregation of platelets, the release mechanisms within the haemostatic system for platelets, and the clot retraction. The major factors involved, the snakes from which they are derived, and the way in which they affect platelets can be seen in *Table: 1.2.*

Another mechanism that has been adopted by the snakes is for their venoms to contain various factors, both enzymatic and non-enzymatic that inhibit the

Name	Source (snake species)	Aggregation	Release reaction	Retraction
Enzymes				
Phospholipase A_2	*Trimeresurus mucrosquamatus*	+	—	—
Phospholipase A_2	*Agkistrodon c. contortrix*	+	—	—
Thombocytin	*Bothrops atrox*	+	+	+
Crotalocytin	*Crotalus horridus horridus*	+	+	—
Ecarin	*Echis carinatus*	+		
Nonenzymatic venom proteins				
Platelet aggregating factor	*Trimeresurus kinavensis*	+	+	—
Convulxin	*Crotalus durissus terrificus*	+	—	—
Convulxin	*C. durissus cascavella*	+	—	—
Aggregoserpentin	*T. mucrosquamatus*	+	—	—
Aggregoserpentin	*T. mucrosquamatus* (Hunan province)	+	—	—
Aggregoserpentin	*T. gramineus*	+	—	—
Aggregoserpentin	*Calloselasma rhodostoma*	+	—	—
Triwaglerin	*T. wagleri*	+	+	—

Table:1.2: Snake venom derived platelet activators.

Name	Source (snake species)
Enzymes	
Phospholipase A_2	*Agkistrodon halys*
Phospholipase A_2	*Vipera russelli*
Phospholipase A_2	*V. siamensis*
Phospholipase A_2	*Naja nigricollis crawshawii*
Fibrinogenase	*V. palaestinae*
Fibrinogenase	*V. aspis*
α-Fibrinogenase	*Trimeresurus mucrosquamatus*
α-Fibrinogenase	*Calloselasma rhodostoma*
F1-proteinase (α-Fibrinogenase)	*N. n. crawshawii*
ADPase/5'-nucleotidase	*V. aspis*
5'-Nucleotidase	*T. gramineus*
5'-Nucleotidase	*A. acutus*
ADPase	*A. acutus*
Amino acid oxidase	*Crotalus adamanteus*
Nonenzymatic venom proteins	
Carinatin	*Echis carinatus*
Platelet inhibitor	*C. rhodostoma*
Trigramin	*Trimeresurus gramineus*
Echistatin	*E. carinatus*
Applaggin	*A. p. piscivorus*
Bitistatin	*Bitis arietans*
Cardiotoxins CTX-1 to CTX-4	*N. n. crawshawii*

Table 1.3: Snake venom derived platelet inhibitors.

aggregation of the platelets, (*Table 1.3).* A further method that has been adopted, particularly in the venoms of the South American pit vipers, is the induction of agglutination.

Agglutination of platelets is a phenomenon in which protein or glycoprotein ligands that contain two or more binding domains bind to the platelets. These factors are the coagglutins and include Botrocetin, which is used clinically for the determination of certain blood coagulation disorders [5].

1.11.2 Venom components acting on the blood vessel wall

Snake venoms contain factors that affect the vessel wall by initiating cutaneous and subcutaneous bleeding at the bite site along with some generalised bleeding. These factors are called the haemorrhagic proteins and are normally zinc proteinases. These proteins appear to act upon the basal membrane of the blood vessel wall, without a proteolytic action [1,3].

1.11.3 Venom components acting on endothelial cells

There are a whole variety of components that act upon the endothelial cells of the blood vessels resulting in constriction or relaxation of the cells and, subsequently, the blood vessels. These factors include the sarafotoxins, which are highly toxic and lead to immediate death from cardiac arrest brought about by constriction of the coronary vessels [1,3,5].

1.11.4 Venom components acting on the activation of prothrombin

There are a whole range of factors within the venoms of various snakes that act upon the coagulation system in order to activate prothrombin. This results in an increase in the production of thrombin and consequently an increase in fibrin formation. These components have been divided into 3 groups in order to help characterise them [1].

Type 1: These activators are not affected by the non-enzymatic cofactors (factor V_a and phosholipid) of the prothrombinase complex. They include factors found in the South American lanceheads (*Bothrops sp).*

Type 2: Phospholipids plus Ca2+ and factors V affect these activators and they have many similarities to the clotting factor X_a. They are found in the venoms of the Australian elapid snakes.

Type 3: These activators are affected by phospholipids plus Ca^{2+} and are derived from the venoms of Australian taipans.

1.11.5 Venom components acting on fibrinogen

Since it is this area of attack of snake venom components that is of interest in this project, this will be covered in more detail than the above mentioned targets. A very large proportion of the *Viperidae* venoms contain enzymes belonging to the serine proteinase family which are involved in the splitting of fibrinogen into some form of fibrin. These enzymes have been termed thrombin-like due to the fact that their function relates most strongly to the known function of thrombin.

Certain of these enzymes have had primary structures determined and some even have a known cDNA. From the primary structure and from the organisation of the gene it has been decided that these enzymes belong to the trypsin-kallikrein group of serine proteinases. In actual fact some of the enzymes which are thrombin-like enzymes, for instance crotalase, have been shown to not only have an action on fibrinogen, but also to act upon kininogens which is the natural substrate of kallikrein.

These thrombin-like enzymes have many varying activities, in that some of them cleave both the fibrinopeptides, like thrombin, and that others cleave only one or the other fibrinopeptides. This suggests that the development of these enzymes has been very broad and highly flexible. They all show significant species specificity for various mammals. These enzymes include batroxobin (*B. atrox)* the enzyme of interest in this project; others are shown in *Table 1.4.* Others act by degrading fibrin into FDP these are shown in *Table 1.5* [1,3,5].

Source (snake species)	Trivial name	Hemorrhagic activity	Fibrinogenolytic activity	Fibrinolytic activity
α-Fibrinogenases				
Agkistrodon acutus		+	Aα	+
A. c. contortrix	Fibrolase	+	Aα	+
A. c. mokasen		–	Aα	+
A. piscivorus conanti	n.d.	Aα>Bβ	+	124
Bothrops asper		–	Aα>Bβ	n.d.
Calloselasma rhodostoma		n.d.	Aα	+
Cerastes cerastes	Cerastase F-4	–	Aα>Bβ	+
Crotalus atrox	Protease-I	n.d.	Aα>Bβ	n.d.
	Protease-IV	n.d.	Aα>Bβ	n.d.
		–	Aα>Bβ	–
	Atroxase	–	Aα>Bβ	+
Naja nigricollis	Proteinase F-1	n.d.	Aα	+
Trimeresurus gramineus	Hemorrhagic Proteinase 1	+	Aα	+
T. mucrosquamatus	+	Aα	+	88
	P-2	–	Aα>Bβ	n.d.
	P-3	–	Aα>Bβ	n.d.
	Hemorrhagic Factor b	+	aα	n.d.
β-Fibrinogenases				
A. p. piscivorus		n.d.	Bβ	n.d.
C. atrox	Protease II	n.d.	Bβ	n.d.
	Protease III	n.d.	Bβ	n.d.
		n.d.	Bβ>Aα	–
T. gramineus		–	Bβ>Aα	Weak
T. mucrosquamatus	–	Bβ>Aα	Weak	136
	P-1	–	Bβ>Aα	n.d.
	ME-1,2	–	Bβ>Aα	+
	ME-3,42	–	Bβ = Aα	+
	HF a	+	Bβ	n.d.

Table 1.4: Snake venom derived components affecting haemostasis.

Source (snake species)	Trivial name	Fibrino-peptide released
Agkistrodon acutus	Acutin	
A. caliginosus		
A. c. contortrix	Venzyme	B,(A)
A. halys blomhoffi		A,B
A. h. pallas		B > A
Bitis gabonica	Gabonase	A,B
Bothrops asper	Asperase	
B. atrox	Batroxobin	A,(B)
B. a. moojeni	Batroxobin	A,(B)
B. insularis		
B. jararaca		A,(B)
Calloselasma rhodostoma	Ancrod	A,(B)
Cerastes vipera	Cerastobin	
Crotalus adamanteus	Crotalase	A,(B)
C. atrox	Catroxobin	A,(B)
C. durissus terrificus	Gyroxin	
C. h. horridus	Defibrizyme	A,(B)
C. viridis helleri		
C. v. oreganus		
Dispholidus typus		
Lachesis muta muta		A,(B)
L. m. noctivaga	Clotase	
Trimeresurus flavoviridis	Flavoxobin	A
T. gramineus		

Table 1.5: Snake venom derived components affecting fibrinolysis

1.11.6 Venom components that act upon Protein C

Protein C activators have been isolated from the venoms of the copperhead (*Agkistrodon contortrix*) and other related vipers. They directly activate protein C, which is an inactive zymogen into the active enzyme protein C_a. This enzyme is a highly potent anti-coagulant. Other components exist that act at this point in the cascade but in a different manner. One is a slow activation of the protein C, by Russells viper venom X, along with a recent discovery that showed that batroxobin (*B. atrox. moojeni*) seemed to act as a thrombomodulin-dependent activator of protein C.

1.12 Batroxobin the thrombin-like enzyme

Batroxobin is the generic name given to any thrombin-like enzyme derived from the *B. atrox* venom. This is because the term was developed before the species was subdivided. Consequently, any published information concerning the structure and action of this enzyme pre-1985 is not wholly believable since the work may well have been carried out on a mixed venom batch of the enzyme. In addition, much of the work to date has been carried out on only partially purified protein and as such is not reliable. Since the term batroxobin refers to all members of this group, then it is vital that the name of the snake from which it is derived is specified to allow clear verification of the exact type of batroxobin that is being studied [5].

The most extensively studied batroxobin to date is the form derived from the venom of *B.moojeni*. This has a known primary structure, and the cDNA sequence was elucidated and has been expressed with partial success since 1991. The gene sequence of the batroxobin shows that the protein component of batroxobin (*B. moojeni*) has a molecular mass of 25,503 Da. However all the forms of batroxobin have been shown to be glycosylated and the mass with the carbohydrate attached is 32,312 Da. It is thought that this carbohydrate is vital for the effective functioning of this group of enzymes, as studies associated with the activity of the enzyme seem to indicate that different amounts of carbohydrate produce differing activities of the protein. It is perhaps worth noting that preliminary studies carried out indicate that

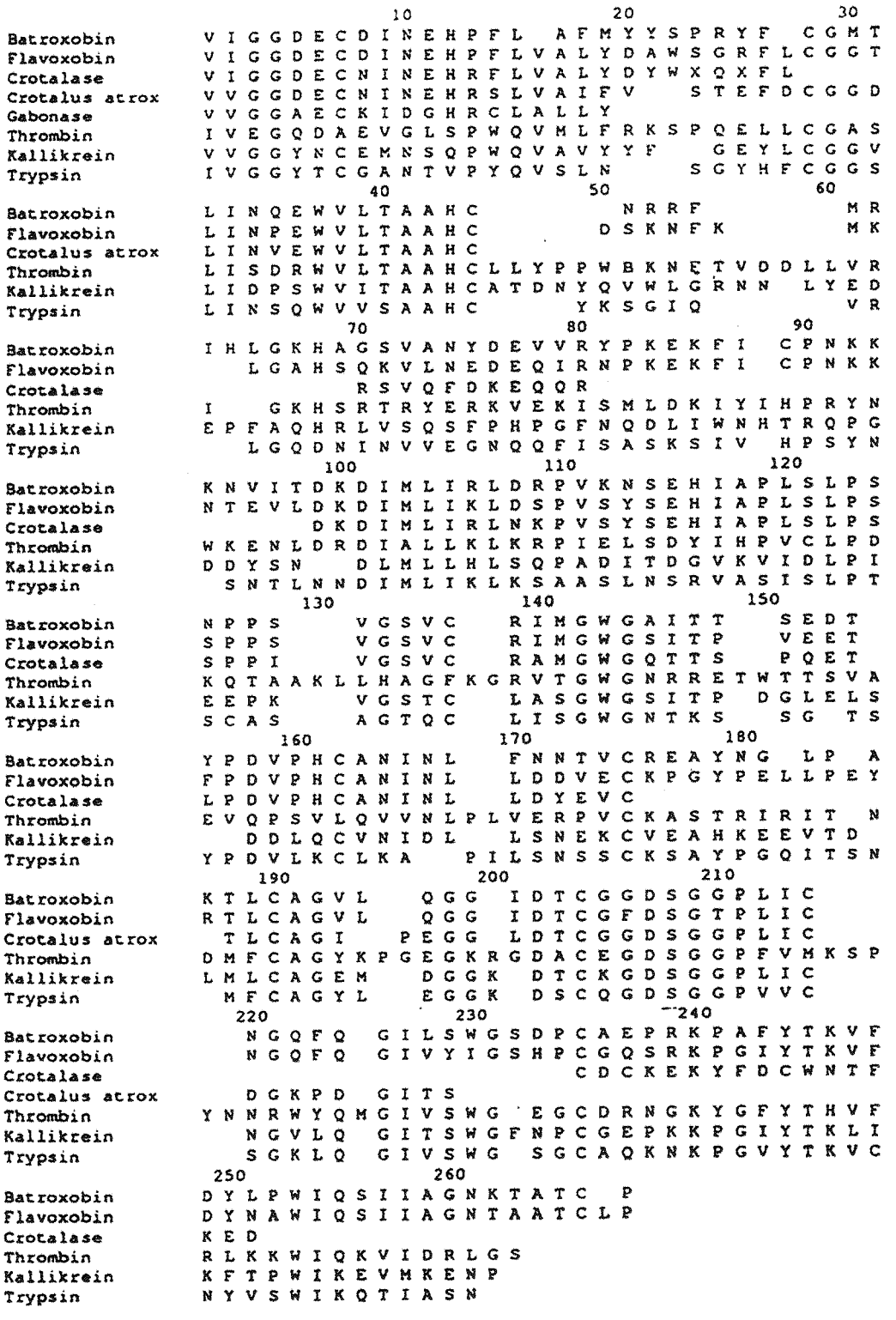


Figure 1.23: Sequence homologies of various serine proteases.

the carbohydrate component found occurring naturally does not correspond to the activity optimum for this enzyme, and that in fact a slightly reduced attachment seems to show higher activity. At present the exact role of this carbohydrate in the action of batroxobin is not known. However, it is clear from the variations within the activity with differing amounts of carbohydrate that it must indeed have some role to play. Batroxobin (*B. moojeni*) contains 6 disulphide bridges and it has a carbohydrate component of 27%, which consists of hexose, glucosamine and sialic acid.

Batroxobin shows significant sequence homology with the other enzymes belonging to the serine protease group, as can be seen from *figure 1.23* above [62]. However, it would seem that it shows the greatest homology ~40% with glandular kallikrein, and not thrombin. This leads to the idea that the thrombin-like enzymes are named according to their action as they cleave fibrinogen and not according to their structure [63,64].

As has previously been noted, the thrombin-like enzymes cleave fibrinogen in a variety of ways. Some like thrombin cleave both the fibrinopeptides A and B whereas others cleave one or another of them. Batroxobin (*B. atrox or B. moojeni*) cleave only the Arg-16-Gly-17 bond in the Aα-chain of fibrinogen. This results in the release of fibrinopeptide A and the subsequent formation of the fibrin I monomer. This spontaneously aggregates in an end-to-end manner in order to form the fibrin I clot. However, unlike thrombin, the fibrinopeptide B is not cleaved and therefore the stable insoluble fibrin cannot be formed as cross linking by factor XIII is not possible (see *figure 1.24*)[64,65].

Batroxobin has been shown to act in a species-dependent manner, showing a preference for the plasma on different mammalian species. This is an area that is of great interest and will be considered in detail in Chapter 3. However, it has also been shown to have an action on some very ancient proteins such as the coagulogen of the horseshoe crab. Some work that has been carried out on the specificity of batroxobin for fibrinogens shows its specificity for dog > man > gopher, and that it shows its lowest specificity to carp plasma [1].

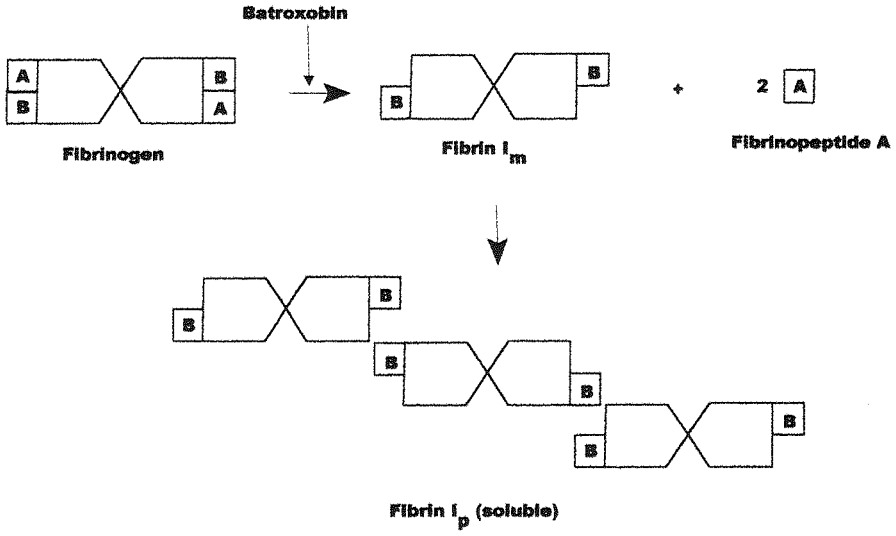


Figure 1.24: Action of batroxobin on fibrinogen.

1.13 Practical applications of batroxobin.

Considerable practical value has been found for batroxobins. These are as haemostatic agents (Reptilase®), defibrinogenating agents (Defibrase®) and for diagnostic purposes (Reptilase-reagent®)[1].

1.13.1 Haemostyptic drugs

Batroxobin has been shown upon local administration to the site of injury to exert coagulation. This acts in a dose-dependent manner and is not inhibited by any plasma thrombin inhibitor. Therefore batroxobin (Reptilase®) has found an application as a topical haemostatic agent. Small doses of batroxobin results in the formation of the fibrin I monomer. Therefore this drug has also found some value by enhancing the action of systemic haemostasis [1].

1.13.2 Defibrinogenating drugs

It has been found that intravenous or subcutaneous injection of large quantities of batroxobin (Defibrase®) results in the cleavage of the fibrinopeptide A and the

formation of fibrin I monomer. At the same time the batroxobin also induces the release of tPA from the endothelium, resulting in the activation of the fibrinolytic system. Therefore the fibrin I monomer is broken down into FDPs which are excreted *via* the kidneys. This results in dose-dependent decrease in the circulating fibrinogen and hence a reduction in the coagulability of the plasma. This drug has obvious and significant applications in the reduction of the risk of thrombosis and since the batroxobin does not affect other points in the coagulation cascade this treatment does not affect primary haemostasis and so reduces the risk associated with defibrinogenating treatments in the event of injury [1,65,66,67].

1.13.3 Laboratory reagents.

Since batroxobin can liberate fibrinopeptide A, without affecting the platelets and without being inhibited by the thrombin inhibitors, it has been found to be a highly useful tool in research and diagnosis. Therefore batroxobin (Reptilase®-Reagent) can be used to determine the concentration of fibrinogen in the plasma, as an investigation for dysfibrinogenemias and as a test for the contractile system of platelets [1].

1.13.4 Other uses

Batroxobin has also been found to be useful for certain preparative purposes, examples of these being the production of serum for patients that are heparinised and also for the production of the soluble fibrin I monomer which can be used as a stimulator of tPA for plasminogen activation [1].

1.14 Comparison of thrombin and batroxobin.

An area of particular interest within this project is a comparison of thrombin and batroxobin. As previously mentioned, batroxobin is termed a thrombin-like enzyme due to its catalytic cleavage of fibrinogen and not necessarily due to its structural similarity to thrombin. However, despite the fact that batroxobin like

thrombin cleaves fibrinogen, there are many ways in which the two proteins are enzymatically dissimilar. Thrombin unlike batroxobin is a multifunctional enzyme. This means that it has a wide variety of important physiological roles.

1.14.1 Catalytic triad.

As has previously been mentioned, the catalytic triad of the serine proteases is remarkably well conserved throughout the family. Thrombin, like all other members of this family has the conserved catalytic triad of His 57, Asp 102 and Ser 195, which is orientated in such a manner as to allow the formation of the highly nucleophilic serine by the action of the charge relay system. This is shown in *figure 1.25* . It is hoped that it will be possible during the course of this project to verify the position of these important residues in batroxobin during the course of the structural studies performed. A comparison of the position of these residues in thrombin and batroxobin is undertaken in Chapter 5.

1.14.2 Cleavage of fibrinogen.

Within fibrinogen there are four bonds potentially available for cleavage by thrombin-like enzymes. Two of these potential sites are within the Aα chain and the other two are within the Bβ chain of fibrinogen. Although there are four possible cleavage sites within fibrinogen there are only two recognition sequences. Within fibrinogen there are four arginine to glycine bonds, these are the potential cleavage sites. The sequence surrounding the two arg-gly bonds within the Aα chain of fibrinogen are identical, for human fibrinogen comprising of: -GGGV**RG**PRVV-. The sequences surrounding the two arg-gly bonds within the Bβ chain of fibrinogen are also identical, comprising in human fibrinogen of: -FFSA**RG**HRIP-. Thrombin has the ability to recognise both of these sequences and subsequently to cleave at both sites, although it has a higher affinity for the Aα sequences. Conversely, batroxobin will only recognise the sequence from the Aα chain.

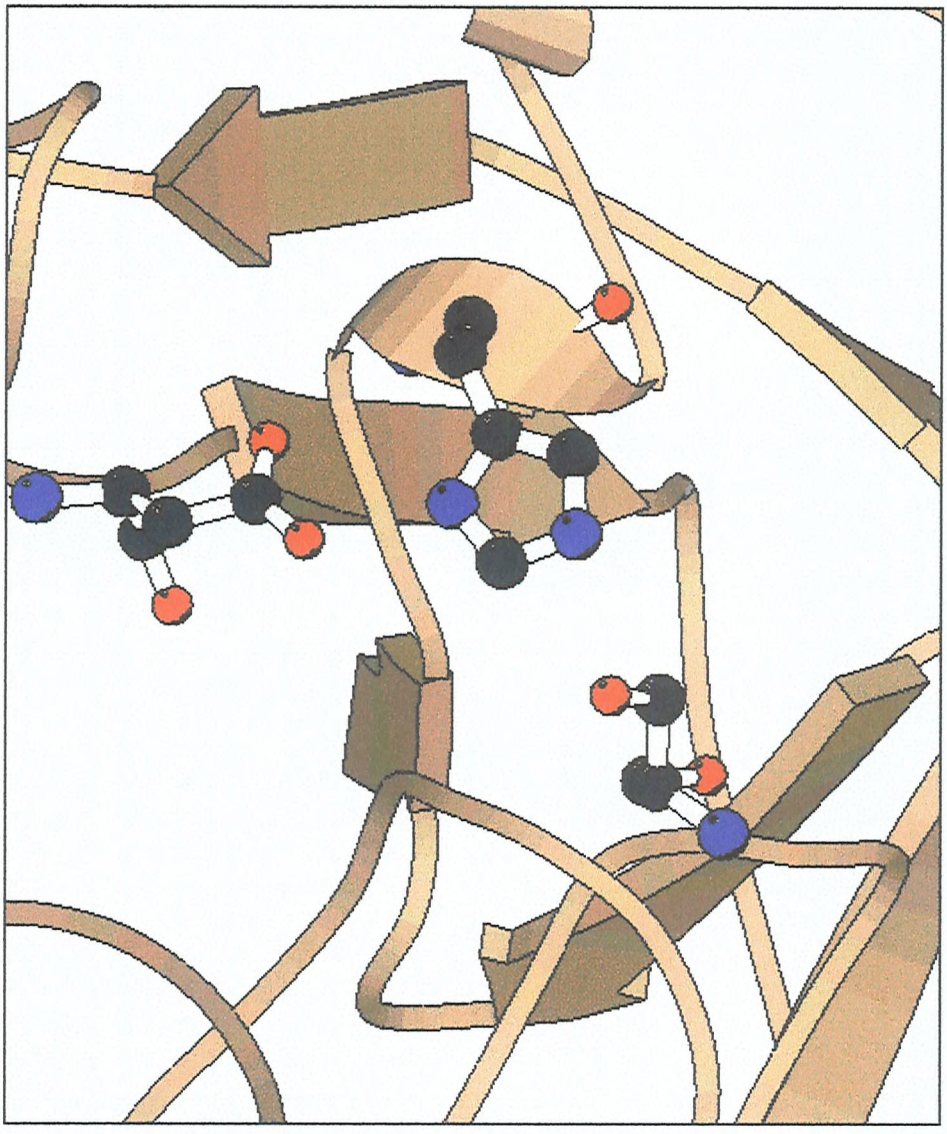


Figure 1.25: The catalytic triad of thrombin. From left to right: Asp102, His 57 and Ser 195.

One important aim of this project was to attempt to account for the differences in the specificity of thrombin and batroxobin for these two sequences. It was hoped that through a combination of structural studies and specificity studies on specially designed peptides that some insight could be gained for the unique action of batroxobin. Consideration of the differences in the specificity for the natural fibrinogen substrates between thrombin and batroxobin is shown in Chapter3. Differences in the specificity of each enzyme at the sites around the scissile bond are shown by their reactions with synthetic peptides in Chapter 4.

1.14.3 Other actions of thrombin

As previously mentioned, thrombin is a multifunctional enzyme, which exhibits many important activities in addition to its most well documented action on fibrinogen. It induces local platelet aggregation, is involved in regulation of the coagulation pathways by induction of the plasma factors V, VIII, IX and XIII for the initiation of coagulation, as well as its role in the negative feed back of coagulation by the induction of the anticoagulant protein C. In addition to its role in coagulation, it also has a vital part to play in cell proliferation and wound healing, affecting fibroblasts, stimulating secretion in platelets and endothelial cells, regulating prostaglandin synthesis in platelets, endothelial cells and fibroblasts, mediating chemotaxis of leucocytes to fight infection, increasing smooth muscle contractility and in stemming blood flow and activating components of the complement system. In fact, to date over 50 proteins have been identified that are susceptible to attack by thrombin.

Interestingly of these only two, fibrinogen and a potential binding site on endothelial cells resulting in the release of tissue plasminogen activator (tPA) are also substrates for batroxobin[5]. Therefore it is vital that an explanation for the increased specificity of thrombin over batroxobin can be found. It is hoped that in achieving this it might be possible in the future to produce a mutant thrombin which exhibits the same action as batroxobin. This protein would potentially be of value in the treatment of thrombosis but with less of the antigenic problems associated with snake venom derived proteins.

Recent findings suggest that there are two distinct actions of thrombin, the fast action, responsible for fibrin formation, and the slow action responsible for all the cellular interactions. The conversion of thrombin from the fast to the slow form seems to be mediated by binding of sodium to Tyr. Only proteins that contain tyrosine at this position have the ability to bind sodium, which is needed for the conversion of the enzyme to the fast form. Interestingly, batroxobin has a proline at this position and is therefore not able to change between two different conformations, however, it would appear from its activity that it must be in the conformation allowing the fast thrombin action as it is only able to cleave fibrinogen[68,69,70,71,72].

Consideration of the differences between the structures of thrombin and batroxobin is performed in Chapter 5. It is hoped that some differences may be identified that will provide potential explanations for the differences in action shown between the two enzymes, batroxobin and thrombin. This includes the specificity for fibrinogen and also the ability of batroxobin to recognise only fibrinogen as a substrate. It is hoped that a novel recognition site for fibrinogen could be identified, which could help to explain the unique specificity exhibited by batroxobin for fibrinogen. In these ways, it will hopefully prove possible to more fully characterise batroxobin, which is to date poorly understood.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Chemicals and reagents

All materials used in this study were supplied by SIGMA chemicals (Poole, UK), with the following exceptions: the S-2238 chromogenic substrate, which was supplied by Kabi Diagnostica (Stockholm, Sweden); Gel Bond® the agarose gel support medium, supplied by Flowgen Instruments (Sittingbourne, UK); the *Bothrops atrox* crude venom which was supplied by LATOXAN (France); and the *Bothrops atrox* (Type II locality, Maranhao) and *Bothrops moojeni* (Type I locality, MinasGerais) crude venoms, which were supplied by PENTAPHARM (Sweden).

2.2 Buffers and Solutions

2.2.1 Purification Buffers

These solutions were made up as outlined in section 2.6

2.2.2 S-2238 chromogenic substrate

The S-2238 chromogenic substrate was supplied in 25mg quantities, which were dissolved in 12.5mls of distilled water. This gave a stock solution of 1mmol/l, which could then be diluted down for use in the assay. The stock was then divided into 100µl aliquots, and stored at -20°C.

2.2.3 HEPES buffered saline (HBS)

1.9g - HEPES
2.34g - NaCl

This was made up to 200ml with distilled H_2O and titrated carefully with concentrated NaOH solution to pH 7.4±0.1.

2.2.4 Calcium chloride solution

88.8 mg - dry $CaCl_2$

This was made up to $20cm^3$ with distilled H_2O and stored at 4°C.

2.2.5 Bovine fibrinogen stock solution - 10-15mg/ml

30mg- bovine fibrinogen (dry powder)

This was dissolved in 2ml of HBS buffer, centrifuged for 5min at 2,500g. The concentration of the fibrinogen was determined spectrophotometrically. This solution was stored at -20°C in 0.2ml aliquots

2.2.6 Bovine thrombin - 10U/ml

1 vial bovine thrombin was dissolved in distilled H_2O up to 10 NIH units per ml. The solution was stored at -20°C in 1ml aliquots

2.2.7 Polyacrylamide gel electrophoresis - acrylamide stock

38g - acrylamide
2g - N,N-methylenebisacrylamide

This was dissolved in distilled H_2O to a final volume of 100ml and stirred for 30min. The resulting solution was then filtered through a Whatman 0.2µl nylon membrane filter and stored at 4°C.

2.2.8 12% SDS - PAGE gel

Main Gel

3.2ml - acrylamide stock

1.0ml - 1.5M Tris/HCl pH8.6

2.8ml – distilled H_2O

80µl - 10%(w/v) SDS (Sodium Dodecyl Sulphate)

100µl 10% (w/v) APS (ammonium persulphate)

20µl - TEMED (N,N,N',N' tetramethylethylenediamine)

Stacking gel

1.0ml - acrylamide stock

1.0ml - 0.5M Tris/ HCl pH6.8

3.0ml – distilled H_2O

50µl - 10%(w/v) SDS

100µl - 10%(w/v) APS

20µl - TEMED

To construct the gel the gel plates were set up and all the solutions for the main gel were combined, with the exception of 10%(w/v) APS and TEMED. When ready to pour the main gel, the APS and TEMED were quickly mixed to the other ingredients and the mix was cast into the gel plates. Whilst waiting for the main gel to set all the components, once again with the exception of the APS and TEMED, for the stacking gel were mixed[73]. Once the main gel had solidified any excess water lying on the top of the main gel was removed, the APS and TEMED were added and the stacking gel was cast with a comb to form the sample wells. Once set, the comb was removed and the gel was placed in the tank which was then filled with running buffer. Samples were prepared by the addition of 3/5 volume of 1.6x disruption buffer and placed in boiling water for 1 minute. The sample (15µl) was then loaded onto the gel and electrophoresis was carried out at 30mA until the bromophenol blue dye front had reached the bottom. The gel was then stained, destained and visualised[74].

2.2.9.6x Disruption Buffer

2ml - 10%(w/v) SDS
0.5ml - 1M Tris pH6.8
0.6ml - glycerol
0.5ml - β-mercaptoethanol
0.01g - bromophenol blue
6.4ml – distilled H_2O

This was stored at -20°C in 500µl aliquots.

2.2.10 5x Running buffer

9g - Tris
43.2g - glycine
3g - SDS

This was dissolved in 600ml of distilled H_2O and stored at 4°C. This concentrated stock solution was diluted for use, with 60ml of 5x stock made up to 300ml with distilled H_2O.

2.2.11 Stain for SDS gels

70ml - glacial acetic acid
400ml - methanol
530ml - dH_2O
0.25% brilliant blue R

2.2.12 Destain for SDS gels

70ml - glacial acetic acid

400ml - methanol

530ml - dH_2O

2.3 Determination of protein concentration

Protein standards were made up using bovine serum albumin (BSA) along with the Bradford dye reagent (Bio-rad) used to calculate protein concentration, as shown below:

	1	2	3	4	5
Volume BSA (µl)	—	2	4	6	8
Volume H_2O (µl)	800	798	796	794	792
Bradford Dye (µl)	200	200	200	200	200

The unknown samples were made up at two differing concentrations as shown below:

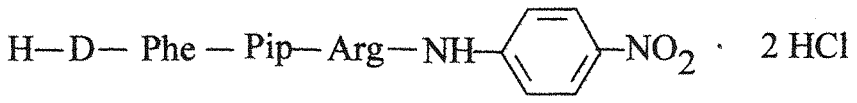
	1	2
Volume protein (µl)	5	10
Volume H_2O (µl)	795	790
Bradford dye (µl)	200	200

The samples and the standards were allowed to rest for 5-10mins at RT in order to develop fully. The absorbance was measured at 595nm and the concentration of protein was calculated from the standard curve obtained from the BSA standards[75].

2.4 Identification of thrombin-like enzymes

2.4.1 S-2238 chromogenic assay

S-2238 is a chromogenic substrate designed for thrombin or thrombin-like enzymes. It is chemically known as H-D-phenylalanyl-L-piecolyl-L-arginine-p-nitroanilide dihydrochloride, which is cleaved by serine proteinases in order to produce para-nitroaniline (*p*NA), which is yellow in colour and can therefore be identified spectrophotometrically at 316nm in H_2O. This has a formula of:



It has a molecular weight of 625.6 and an E_{max} of $1.3 \times 10^4 mol^{-1}\ l\ cm^{-1}$;the concentration of thrombin or thrombin-like enzyme can therefore be determined, along with the rate at which it turns over the substrate. However, due to the fact that it may be cleaved by any serine proteinase it is imperative that it only used on the already purified protein, since the crude venom contains many serine proteinases that will activate the substrate along with the batroxobin.

Standards were made up containing known concentrations of batroxobin in order that a dose response curve for this enzyme could be produced. This was produced as shown below:

Activity Required (BU/ml)	µl of 20BU/ml Batroxobin	µl of 0.1M TRIS + 1M NaCl
5.0	45	135
7.5	67.5	112.5
10.0	90	90
15.0	135	45
20.0	180	–

This was plotted in order to make a standard curve, and the unknown sample's activity was plotted against the standards in order to determine the activity of the batroxobin in batroxobin units. One batroxobin unit is equivalent to 0.17NIH units; the international standard used for thrombin [76].

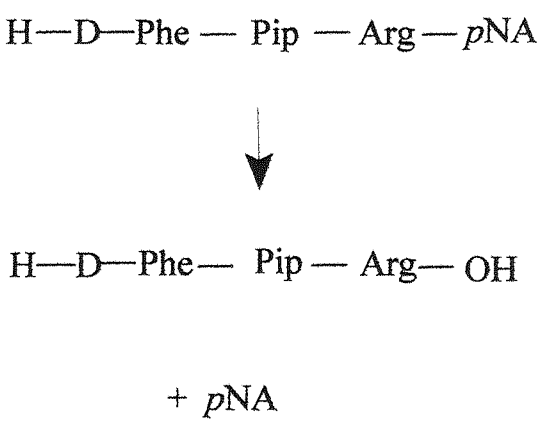


Figure 2.1: Reaction of thrombin or thrombin-like enzymes with S - 2238.

2.4.2 Visual gelation clotting assay (purified fibrinogen)

The bovine fibrinogen stock was diluted to a 4mg/ml solution with HBS pH7.4 and 0.1ml volumes were preincubated at 37°C for 10min, and seperately 0.2ml volumes of thrombin or thrombin-like enzyme. After incubation an equal volume of 0.04M $CaCl_2$ was added to the fibrinogen and carefully mixed. Immediately, the preincubated thrombin was added to the fibrinogen $CaCl_2$ mixture to initiate clotting. Upon addition, the stopwatch was started and the tube was gently tilted in order to allow the very first signs of gelation to be noted. The time taken for the clot to form was recorded. This is the clot time and should be between 10 seconds and 1 minute for the time to be accurate. If the time is quicker than 10 seconds the solution must be diluted in order that the time can be calculated accurately. Any longer than this and the sample must be concentrated [76].

2.4.3 Kinetic study of fibrin polymerisation

The fibrinogen stock was made up to a concentration of 4mg/ml with analytical H_2O. A cuvette was set up as shown:

0.75 ml – HEPES buffered saline

0.10 ml - fibrinogen

0.10 ml - $CaCl_2$ solution

The mixture was then incubated at 37°C for 3 min. The cuvette was placed in the UV-visible spectrophotometer with the wavelength set to 350nm, start time 10 seconds, cycle time 5 seconds, and the running time to 300 seconds. Then 0.05ml of thrombin or thrombin-like enzyme was added to the cuvette immediately after setting up the 10-second delay and the contents were mixed well. A turbidity curve was calculated and the lag period was calculated by extrapolation to the X-axis. The rate of the reaction was calculated by measuring the slope at 100 seconds [76].

2.4.4 Coagulation of human plasma

This method of assay is of major significance in the calculation of the activity of the enzyme once it has been purified, as this is the standard to which the thrombin-like enzymes are compared. In this case, the above method for gelation of fibrinogen was followed, with the exception that 0.2ml of human plasma was used instead of the fibrinogen and $CaCl_2$.

During the course of the purification process the purified fibrinogen must be used to follow the progress of the purification, since there are other components within the crude venom mix, which act upon the blood coagulation cascade at other points. Therefore, if human plasma was used to assess the purity of the mixture other venom components acting higher up in coagulation may trigger the coagulation cascade of the plasma, resulting in an impression that there is a thrombin-like enzyme present. In the same way, the chromogenic assay cannot be used to evaluate the purity of the process as this assay is active for all serine proteases, and since the enzymes acting on the coagulation cascade are all serine proteases any one of them may be responsible for the reaction observed [76].

2.4.5 Zymographic assay

A non-reducing SDS-PAGE gel of the thrombin or thrombin-like enzyme was run, with 4mg/ml fibrinogen solution substituted in place of the distilled H_2O, to

produce a fibrinogen impregnated slab. Once the gel had run it was washed with 2.5% TRITON X-100 for about 2 hrs to enable the denatured enzyme to be reformed. The gel was then extensively washed with dH_2O in order to remove all the TRITON from the gel. The gel was then stained in 0.1% amido black solution for 10 mins. before washing. This process works upon the basis that the thrombin - like enzyme will convert the fibrinogen in the SDS gel at the point where the enzyme has run. Thus fibrin formation can then be visualised by staining with the amido black [76].

2.4.6 Fibrinogen - agarose plate assay

The fibrinogen - agarose film was produced by making up the agarose in 4mg/ml fibrinogen solution and pouring it onto the hydrophilic polyester support film. 3mm diameter wells were cut into the gel and 20µl of thrombin-like enzymes either sample or standards were placed in the wells. The film was incubated at 37°C overnight and then washed in 1l of 0.9% NaCl at RT for 2hrs. The gel was dried on a piece of filter paper between paper towels with a 10kg weight on top of the film covered by a glass plate, so that it was dried under high pressure for 20mins. The film was then re-washed and re-dried under pressure a further two times. The film was then allowed to dry in an incubator and stained with 0.1% amido black stain [76].

2.5 Purification of thrombin-like enzyme from crude venom

2.5.1 Purification using heparin sepharose affinity column

The heparin Sepharose affinity gel was made up according to the manufacturers instructions, washing the gel extensively prior to dissolving into the equilibrium buffer (0.01M Tris/HCl (pH 8.5) + 0.1M Glycine – base). As some of the packing powder mixed in with the gel may not have fully dissolved, the gel was centrifuged to remove any undissolved particles. The pH was checked and if the gel was found to be too acidic it was neutralised. This process sometimes resulted in the re-precipitation of some of the affinity gel; under these circumstances it would be necessary to re-centrifuge it. The gel was then kept on ice.

The gel (10ml) was placed into a 1cm-diameter column and equilibrated with a minimum of 5 column volumes with the equilibration buffer. The venom was dissolved in analytical H_2O and the protein concentration was measured spectrophotometrically at 280nm. The venom mixture was then centrifuged at 1000-x g for 5-10 minutes to clarify before putting onto the column. The venom was then placed on the column slowly under gravity, allowing it to fully load prior to any washing. Following this the column was washed with 0.01M Tris/HCl + 0.1M Glycine - buffer + 0.1M NaCl readjusted to pH 8.5. Each fraction was monitored at 280nm until it reduced to less than 0.05 absorption units, showing that little protein was present in the fractions.

The column was then washed with 0.01M Tris/HCl + 0.1M Glycine - base + 0.25M NaCl readjusted to pH8.5. Once again the absorption at 280nm was checked until below 0.05 units before washing with 0.01 M Tris/HCl + 0.1 M Glycine - base + 1.0M NaCl, readjusted to pH 8.5. The absorption was once again monitored until adsorption at 280nm was below 0.05 absorption units before finally washing with 0.01M Tris/HCl + 0.1M Glycine - base + 4.0M NaCl readjusted to pH 8.5 with HCl.

This was again checked that no more protein was eluting from the column. The fractions were then all screened using the visual coagulation gelation of purified fibrinogen in order to determine which fractions contained the thrombin - like enzyme batroxobin [76].

2.5.2 Purification using Poros 20HE Biosprint column

This purification was carried out using the same basic buffer system. However, the instrument set-up enabled one low salt and one high salt buffer to be mixed to produce a continuous gradient of 0M NaCl to 2M NaCl. The column was packed using the self-pack device that was supplied with the instrument and using the equilibrium buffer described in 2.5.1.

The sample was loaded onto the column and the proteins were eluted using a salt gradient from 0-2M over a 30-column volume period. The fractions of 0.5ml were collected in a fraction collector.

2.5.3 Purification using *p*-aminobenzamidine column

The *p*-aminobenzamidine gel (10mls) was packed into a column with a diameter of 1cm using the equilibrium solution (0.05M Tris/HCl + 0.4M NaCl, readjusted to pH 9.0) to facilitate the packing. All solutions were equilibrated to 4°C; the column was packed and run at the same temperature. The venom was dissolved in analytical H_2O and the protein concentration was measured at 280nm and using the Bradford assay. The venom was centrifuged at 1000-x g for 5-10mins in order to clarify the venom before loading onto the column. The gel was equilibrated using at least 5 bed volumes of the equilibrium buffer, before slowly loading the clarified venom onto the column.

At 4°C the thrombin-like enzyme binds to the benzamidine attached to the agarose *via* a *para*-amino link in the form of an enzyme inhibitor complex, as the benzamidine is a highly effective competitive inhibitor for the batroxobin. Therefore, since the desired protein is bound *via* hydrophobic interactions to the column, the gel is washed in high salt using the equilibrium buffer to elute any contaminating proteins from the crude venom, which may have bound *via* weak hydrophilic interactions to the gel. The column was washed until the fractions show readings of less than 0.2 ABS units at 280nm.

The bound active enzyme was then eluted from the column using the elution buffer 0.05M Tris/HCl + 0.1M NaCl + 0.15M Benzamidine readjusted to pH 9.0. This elution buffer acts by competing with the benzamidine in the gel for the bound thrombin-like enzyme. As it elutes it forms an enzyme- inhibitor complex with the batroxobin removing it from the gel. When the absorption of the elutant was monitored spectrophotometrically at 280nm the high reading observed was due to the benzamidine and not to any protein content. However, the protein was eluted about 2 fractions behind the benzamidine solvent front. This was identified by some activity that could be identified using the fibrinogen visual gelation assay, which was shown to have some effect even when the enzyme was still bound to the benzamidine inhibitor.

In order to remove the enzyme from the enzyme-inhibitor complex it was dialysed against 0.02M Citrate + 0.2M NaCl pH 6.0 for 24hrs with three changes of the dialysis solution. The extensive dialysis was sufficient to remove all the

benzamidine leaving the free enzyme. This solution was then freeze-dried to leave totally purified freeze-dried batroxobin [77].

2.6 Homology modeling.

All analysis and modelling of batroxobin was performed using the program package Quanta/CHARMm. The sequence alignment was initially determined using a Needleman Wunsch algorithm[78], and a protein score matrix[79]. This alignment was then refined by hand assuming the conservation of residues known to be significant within the active site of serine proteases. Any insertions or deletions (indels) in the aligned sequences, were whenever possible introduced into loop regions of the structure rather than the more conserved β-pleated sheets or α helicies of the structure. The backbone co-ordinates of the aligned residues in kallikrein were mapped to the equivalent amino acids in the model. At the places were indels occurred within the sequence that was being modelled, a fragment database was used to search for short sequences that overlapped with the structure either side of the unknown region and which contained the same number of residues[80]. The best fit fragment was then chosen and used in the model. After the backbone co-ordinates had been defined, the validity of the emerging structure was assessed using the Protein Health module in Ouanta. Any irregularities that were identified, were remodelled, either manually or by re-using the fragment database. This process was repeated until the backbone co-ordinates had been properly defined, and then the side chain co-ordinates were considered. The side chain co-ordinates from identical amino acids were copied directly from the kallikrein structure to the model structure. Quanta/CHARMm was then used to generate the side chains for any non identical residues, in their most rotameric form. Any major side chain clashes were removed manually and then CHARMm was used to perform an energy minimisation. This was carried out whilst applying constraints in order to preserve the basic backbone structure of kallikrein. Since the energy minimisation program is based upon the theoretical structure at absolute zero, then this iterative process was continued until the the final energy of the molecule was of the same order of magnitude as that of the kallikrein structure. This final structure was then checked for any abnormalities and

inconsistancies using the protein health option in Quanta. A complete description of this process can be found in Chapter 5.

2.7 Solid phase peptide synthesis

In the synthesis carried out, a solid phase peptide synthesis was undertaken using Fmoc protection, with both the formation of the free acids and the amides and a combination of Fmoc and tBoc protection for the inhibitors. This was achieved manually, using a Griffin flask shaker. All the solvents and chemicals used were of peptide synthesis grade. Arginine-Wang resin, 0.9mmole/g was used for the synthesis of all the peptide free acids, with the exception of one which required the synthesis of the R-resin from the Wang resin, 1mmole/g. *p*-Methylbenzhydralamine (MBHA) resin, 0.9 mmole/g, 100-200 mesh (Calbiochem-Novabichem Ltd., UK) was used for the synthesis of the peptide amides. The side chain protections that were used were for Fmoc chemistry in the production of all the peptides, but a combination of Fmoc and tBoc chemistry for the inhibitors. All the filtrations were carried out using positive nitrogen pressure. A complete account of this process can be found in Chapter 4[81,82].

2.8 Reaction of thrombin and batroxobin on peptides.

0.02 M solutions of each peptide were made up in analytical water, and a $1x10^{-5}$M solutions of thrombin or batroxobin were made up in 50mM TRIS/HCL, buffer pH7.5. Each enzyme solution (50µl) was added to 200µl of the peptide solution and incubated at 37°C. Aliquots (40µl) were taken from the mixture after 0min, 30min, 1hr, 2hrs, and 4hrs, and the reaction was stopped with 40µl of a 50:50 solution of acetonitrile/TFA[83,84,85].

A 1/100 dilution of each of the reaction mixtures was loaded onto a C18 reverse phase analytical column, and eluted with a 0-20% linear gradient of solution B, where solution A is 0.1% TFA in H_20 and solution B is 0.1% TFA in acetonitrile. Elution of the peptides alone resulted in the production of the hexapeptide peaks at

~15% acetonitrile, with any tripeptide cleavage products expected at some point before. This standard procedure was then refined, as detailed in chapter 5.

2.9 Crystal growth.

The hanging-drop method was used to grow crystals of batroxobin. The lip of each of the wells on the 4x6 well plate was greased using high vacuum grease, and each well was filled with 1ml of buffer. Initially these buffers were obtained from the Hamilton Crystal Screens I and II and later, during refinement, using buffers made up on site. A droplet of the buffer, between 1-4µl dependant on the volume of the protein to be used, was placed on a coverslip presiliconated with dimethyldichlorosilane[86]. An equal volume of the protein solution at a concentration of between 10-35mg/ml was added to the droplet and mixed thoroughly. The droplet was then suspended over the corresponding well of buffer. The plates were then incubated at a constant temperature of 19°C to enable crystal growth. This standard procedure was modified as detailed in Chapter 3.

CHAPTER THREE

PURIFICATION, CHARACTERISATION AND CRYSTALLISATION OF BATROXOBIN

3.1 Introduction

In order that any of the work to be carried out in this project could be satisfactorily completed a better method of purification for batroxobin needed to be found, since much of the previous work carried out on this snake venom protease had been discredited due to the presence of impurities. Therefore, a large amount of time was spent attempting to find the best possible way in which a reliable purification could be achieved.

Once this had been carried out, the purified protein could be used in crystallography studies, as well as in attempts to characterise the differences between batroxobin from *B. moojeni* (Type I locality, Minas Gerais) and batroxobin from *B. atrox* (Type II locality, Maranhao). Batroxobin derived from the venom of *B. moojeni* has been extensively characterised, however, the enzyme derived from its very close relative, *B. atrox,* had yet to be considered in any detail and, in fact, it was unclear whether any differences actually existed.

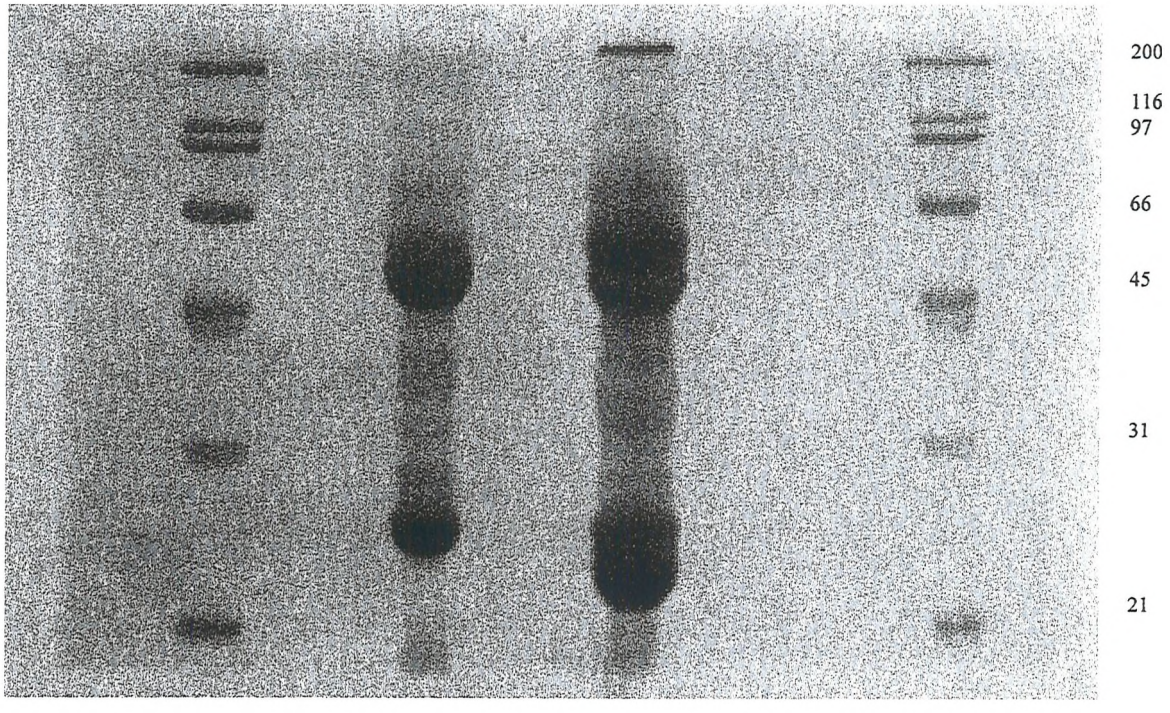


Figure 3.1: SDS-Gel comparison of the crude venom derived from B. Atrox and B. Moojeni. Lane 1 and 4 are MW standards, lane 2 B. moojeni crude venom and lane 3 is B. atrox crude venom.

3.2 Batroxobin purification trials

3.2.1 Purification by heparin Sepharose

The previously published purification protocol, as shown in Chapter 2, was followed in order to attempt to isolate the batroxobin from the crude venom mixture. It soon became evident, however, that there were certain problems associated with this method of purification. Heparin affinity columns have been traditionally used for the purification of many blood coagulation proteins, or proteins that act upon the coagulation cascade. As snake venoms are largely made up of proteins that do act upon this cascade then it is entirely plausible that many of the proteins within the venom mixture may bind to the heparin. In fact this was a problem that was encountered time and again whilst following this purification technique and it was never possible to separate fully the proteins from each other upon elution. Attempts were made to elute batroxobin specifically from other bound components by making the salt gradient as shallow as possible for the elution at this point, however, contaminating proteins were always evident and some proteolysis occurred also

A more rapid reliable method of purifying the batroxobin therefore needed to be found in order to produce a homogeneous enzyme and to avoid the problems associated with degradation. Therefore, it was decided that the same protocol should be carried out, but that this time it should be attempted on the Biosprint instrument, which would allow the purification to take place at a much higher speed. It was also hoped that since the process was so rapid, it may prove possible to run the sample through the column many times in order to try and fully separate the proteins that bound to the heparin and, consequently, produce a highly pure batroxobin sample [76].

3.2.2 Purification using Biosprint He and HQ columns.

Once again, the protocol as shown in Chapter 2 was followed. This time the heparin was used on the Biosprint instrument which enables the process to be carried out at high flow rates in only 3 minutes. Unfortunately the batroxobin isolated by this

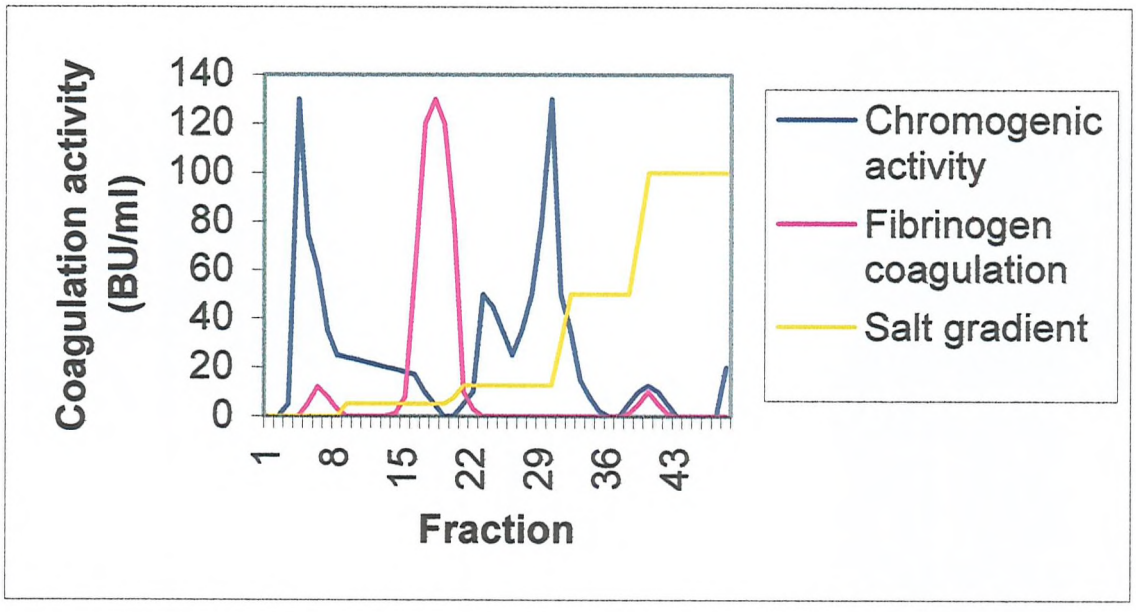


Figure 3.2: Purification profile of batroxobin from B. atrox crude venom using a heparin sepharose affinity method. Samples were eluted in 2ml fractions using the salt gradient shown in yellow. Each fraction collected was assayed for serine protease activity using s-2238 a chromogenic substrate as shown in blue, and coagulation activity using purified fibrinogen as shown in magenta.

method seemed to be spread quite widely through the fractions from the Biosprint. It was originally thought that this must be due to the fact that the process was carried out at too high a speed to allow complete binding to the column or alternatively, that there was not enough capacity on the very small column used on the Biosprint instrument. It was also possible that the characteristics of the heparin binding to many protein components within the venom were also partly responsible for the poor separation. Attempts were made to run the sample through the column several times in an attempt to allow the complete separation of all the bound proteins. Despite numerous attempts to use heparin sepharose it became clear that batroxobin could not be resolved from other heparin binding proteins sufficiently well to allow this method to be used reliably. It was therefore decided that a second stage of purification should be attempted, pooling the fractions that showed thrombin-like activity, concentrating and desalting them, and then running them through another separation gel in an attempt to separate them under different conditions.

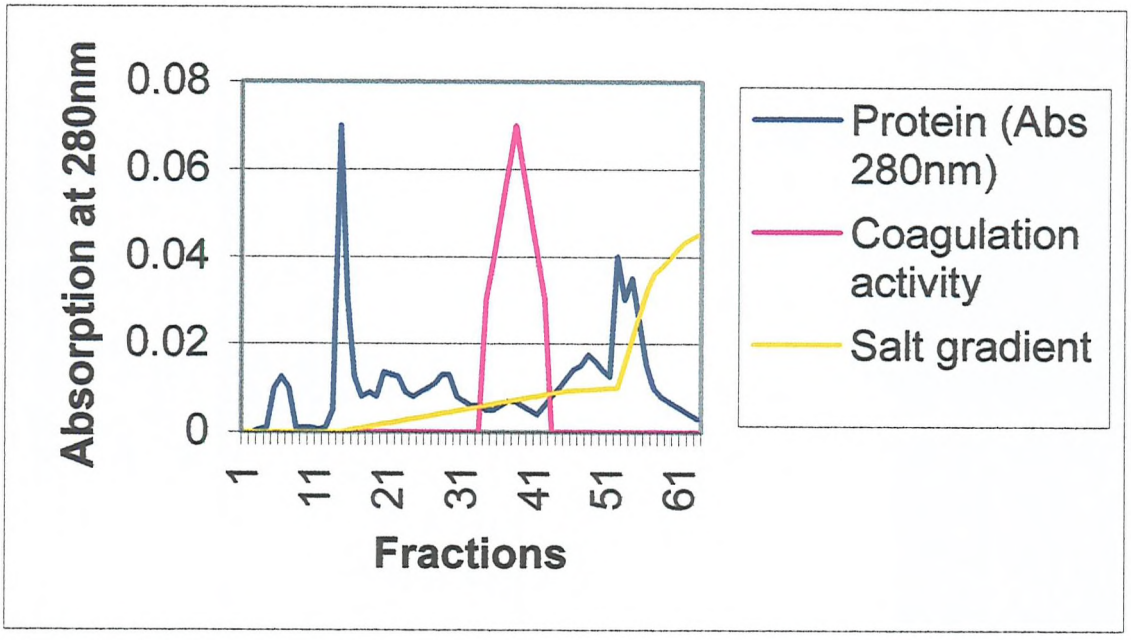


Figure 3.3: Purification of crude B. atrox venom using HE Biosprint column. Samples were eluted in 1ml fractions using the salt gradient shown in yellow. Each fraction was then assayed at 280nm protein, shown in blue, and fibrinogen coagulation activity shown in magenta.

It was decided that an anion exchange column, the POROS 20HQ column should be used for this additional step. It was hoped that it would allow batroxobin to be completely purified from the other components. Unfortunately, two other venom components also eluted at the same point as batroxobin from the anion exchange column. It appeared that only one of these proteins could ever be removed from the batroxobin, irrespective of the salt gradient used. Although it may have proven possible to run this semi-purified sample through another separation gel, by this point the amount of sample left was extremely small, with the vast majority of the protein having been lost along the way. Therefore, it was decided that rather than try and further purify from this point in the protocol it would be more effective to attempt a completely different approach to this purification.

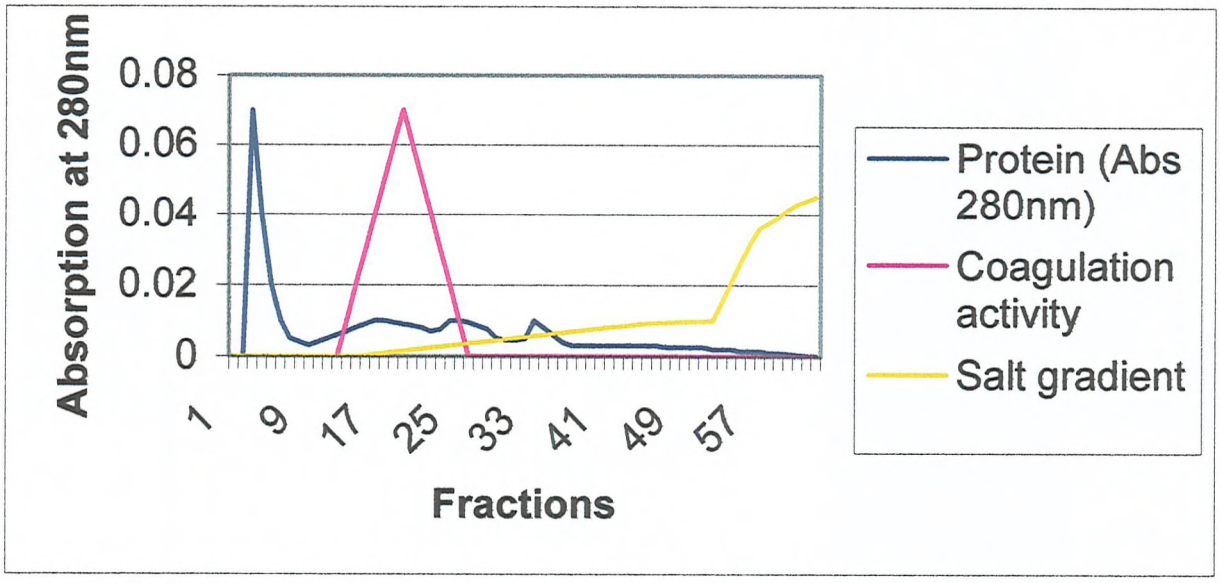


Figure 3.4: Purification profile of partially purified B. Atrox venom on HQ Biosprint column. Samples were eluted in 1ml fractions using the salt gradient shown in yellow. Each fraction was then monitored at 280nm for protein, shown in blue, and assayed for fibrinogen coagulation activity, shown in magenta.

3.2.3 Purification using *p*-amino benzamidine column

Having decided that an alternative approach was necessary, it was decided that a method of purification that had previously been described by Holleman *et al* for the thrombin-like enzymes belonging to this family of snakes should to be tried. This method was based on the observation that benzamidine acts as a good competitive inhibitor of thrombin-like enzymes. The benzamidine was attached to the agarose by a *p*-amino linkage to produce an affinity chromatography gel. This affinity gel has since become commercially. Elution of the bound enzyme from the enzyme-inhibitor complex is easily facilitated by the addition of free benzamidine to the buffers. In this case, when the protocol as shown in Chapter 2 was followed, it was found that a purer sample, with only minor contaminants, was obtained after only one run through the benzamidine column. The benzamidine could easily be removed by dialysis against citrate and a fully purified active thrombin-like enzyme could then be obtained. It was hoped that this method of purification would bind batroxobin specifically since only

thrombin-like enzymes, rather than all enzymes acting on the coagulation cascade, would actually bind [77].

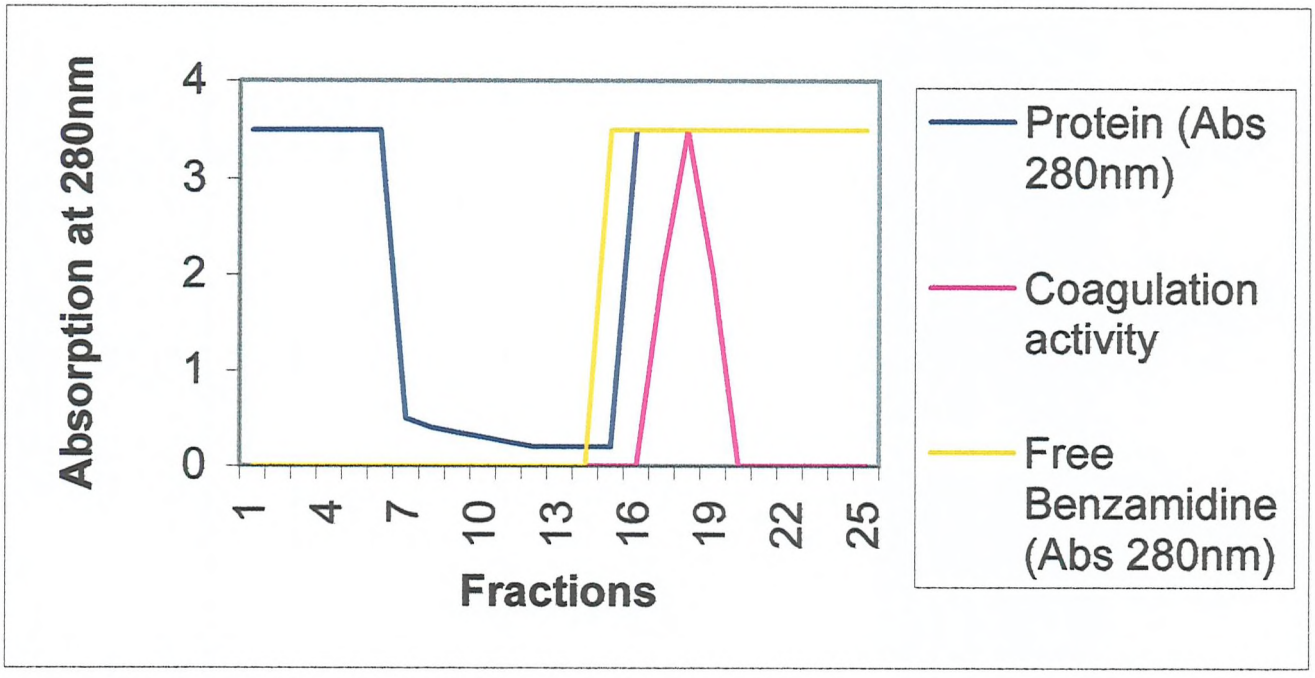


Figure 3.5: Purification profile for crude venom (B. atrox) using the p-amino benzamidine affinity column. Absorption at 280nm is shown in blue, however, since benzamidine absorbs at 280nm then once this has been added the protein cannot be seen. Therefore, fibrinogen coagulation activity is used to identify the batroxobin, as shown in cerise.

This method of purification of batroxobin was found to be much more effective than the heparin sepharose at purifying the batroxobin, however, even this method resulted in some minor contamination by other proteins from the venom mixture. It was therefore decided that, since the contaminating proteins from the two affinity methods were different, it might be of value to attempt the purification by coupling both methods. In addition to this, it was found that some of the contaminant proteins within the venom mixture were not as soluble as the batroxobin, consequently this could be used as a further purification step.

3.3 Purification of batroxobin.

Having carried out sufficient pilot experiments using 10mg amounts of crude venom for the purification a large scale isolation was possible. Typically 500mg of crude venom was placed in 5mls of analytical water and inverted to mix. It was vital not to agitate too violently in order to prevent frothing and protein denaturation. The solution was centrifuged at 1000-x g for 5-10min. and the supernatant was decanted. Any venom residue left in the pellet was collected and assayed for thrombin activity. This process was repeated until all the coagulation activity was removed from the residue. This residue was then run on a SDS-PAGE gel to confirm that no batroxobin was left undissolved. Therefore, if a band at the right size for batroxobin was found on the SDS-PAGE gel, then the process was repeated until the band was no longer visible.

The dissolved venom was passed through a PD10 Pharmacia column to remove salt and other small molecules from the highly complex mixture of the venom in order to enhance binding to the affinity column. This is due to the fact that the high levels of salt found in the crude venom were found to affect complete binding of the batroxobin to the affinity column. The desalted sample was then loaded slowly onto the *p*-aminobenzamidine column and the method outlined in Chapter 2 was followed. Fractions containing coagulation activity were pooled and dialysed against citrate to remove the benzamidine, desalted in a PD_{10} column and the resulting solution was then freeze-dried. The fractions collected from the high salt washes were assayed for coagulation activity to ensure the column had not been overloaded. Since any bound batroxobin can only be eluted with the specific addition of free benzamidine, then any batroxobin found in the high salt washes could not have bound to the column and was therefore present due to overloading of the column. The lyophilised powder was redissolved in 5mls of analytical water before loading onto the heparin column. This was eluted following the method outlined in Chapter 2. The fractions were screened for coagulation activity and any fractions showing activity were pooled and freeze dried. This process was found to produce sufficiently highly purified batroxobin for all of the biochemical testing, but a further step of gel filtration was undertaken prior to crystallography trials.

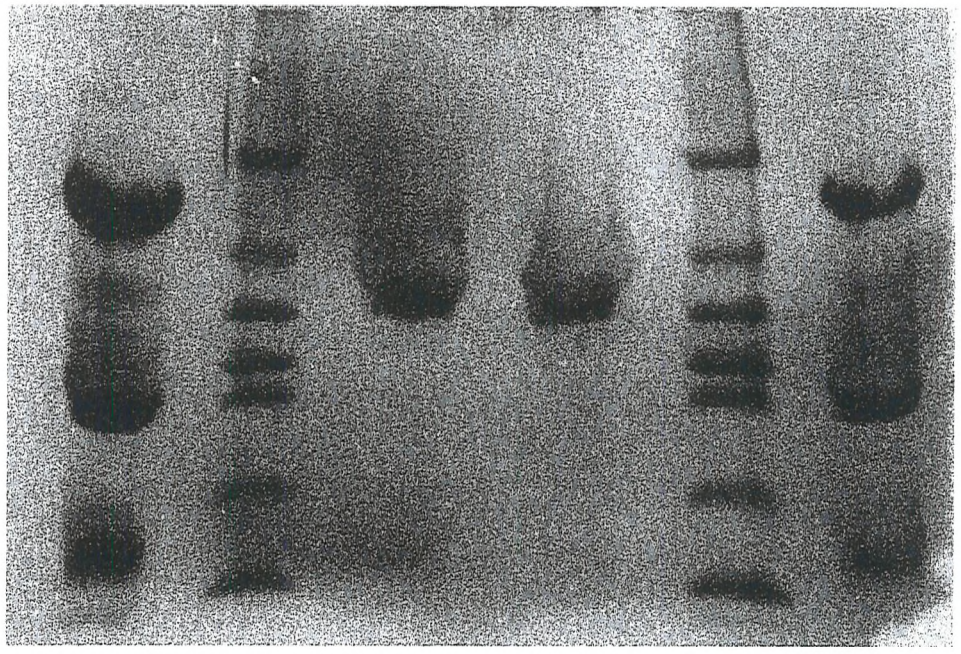


Figure 3.6: SDS-Gel showing the purification of batroxobin from crude venom. Lanes 1 is B. atrox crude venom, Lanes 2 and 5 are MW markers, Lane 3 is purified batroxobin (B. atrox), Lane 4 is purified batroxobin (B. moojeni) and Lane 6 is B. moojeni crude venom.

Purification	Conc. by Bradford assay (mg/ml)	Total protein (mg)	Activity (BU)	Specific activity (BU/mg)	% Yield
Crude venom	50	500	6086	12	-
After *p*-aminobenz-amidine column	2.5	20	4716	236	~77%
After heparin column	2	10	4211	421	~70%

Table 3.1: Purification table of batroxobin from crude venom.

3.4 Characterisation of batroxobin (*B. atrox*) supplied by LATOXAN.

Once the enzyme had been purified, it was immediately assessed in an attempt to determine its similarity to batroxobin (*B. moojeni*) the better characterised enzyme. At present one of the vital questions being considered is whether or not there is any difference between batroxobin derived from different snake sub-species. Many of them have never been fully characterised and, accordingly, it is very difficult to see if the differences in the enzymes are real protein primary sequence variations or whether they relate to different carbohydrate content. Since variations in the glycosylation of batroxobin from within the same sub-species have been documented, it would be hard to specify any great differences between batroxobin from other sub-species if the only variation was in this part of the structure[5]. Therefore it was important to determine if the primary sequence between the sample of batroxobin (*B. atrox*) was the same as that published for batroxobin (*B. moojeni*). It had already been suggested from the literature that the two enzymes had different masses but it was unclear as to the reason for this difference.

Therefore an N-terminal sequence analysis of the proteins was carried out. This was possible on the first 20 amino acids from the N-terminal of the proteins. If the first 20 amino acids of a protein are found to be identical to those of a known standard the statistical likelihood is that they are the same protein.

BATROXOBIN **VI***DE*DINEHPFLAFMY-**
(*B. atrox*)

BATROXOBIN **VIGGDECDINEHPFLAFMY-**
(B. moojeni)

Figure 3.7: N-terminal sequence analysis of batroxobin (B. atrox and B. moojeni)

As can be seen from these two sequences the primary structure of the batroxobin isolated from the venom of *B. atrox,* as supplied by LATOXAN, was similar to the published sequence for the enzyme from *B. moojeni.* However, there were a few points on the sequence that did not produce any definitive answers. These

regions can be explained: as cysteine residues are notoriously difficult to sequence and the SDS-PAGE was carried out in a glycine buffer and carry-over could have occurred, which swamped the signals coming from the glycine residues. Consequently, the two sequences can be seen as identical. The next step was to finally determine the molecular mass of batroxobin from *B. atrox* and it to the known value for batroxobin from *B. moojeni*. This was attempted initially using standard SDS-PAGE techniques, as shown in *figures 3.8-3.9.*

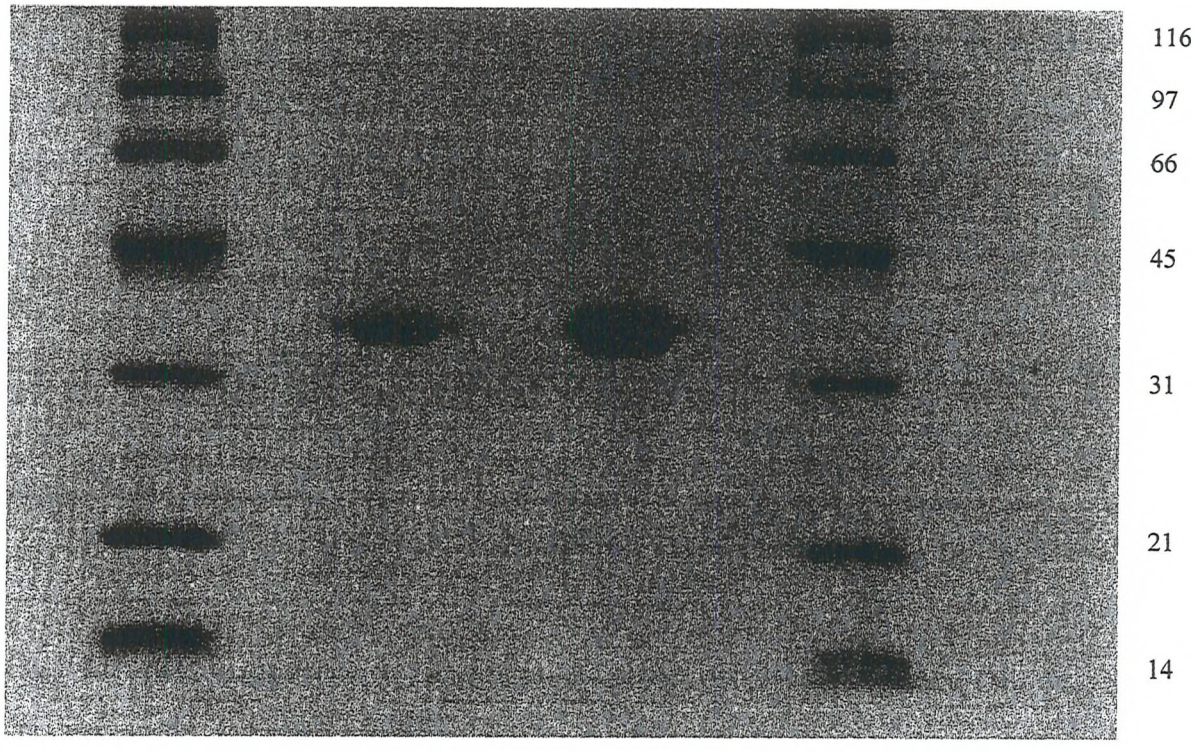


Figure 3.8: SDS-PAGE gel of purified batroxobin (B.atrox). Lanes 1 and 4 are MW markers. Lanes 2 and 3 are batroxobin (B. atrox).

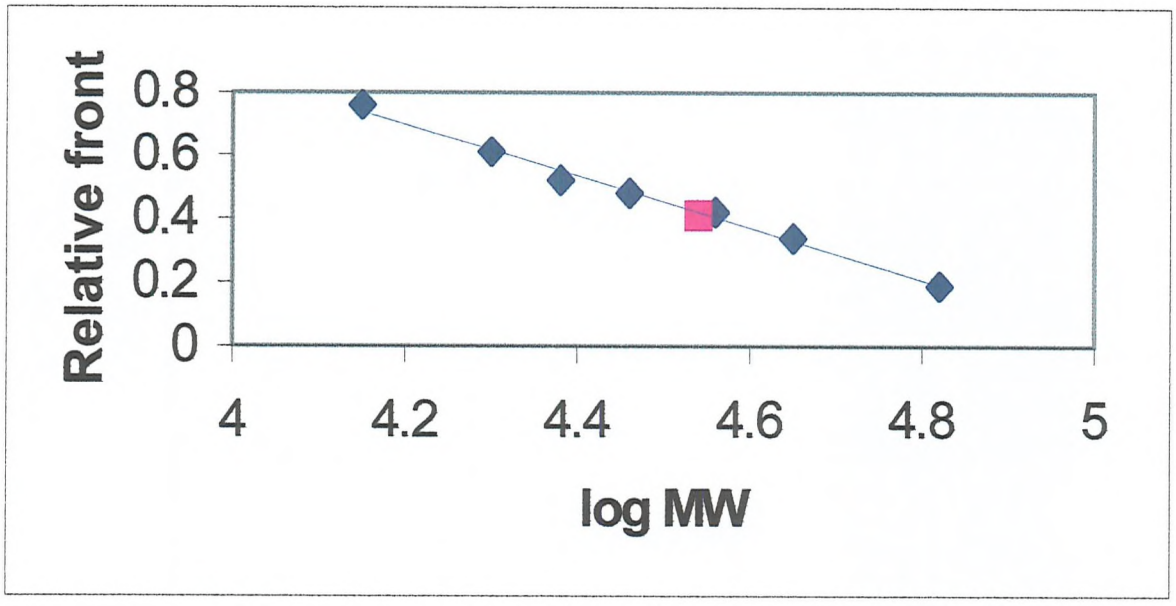


Figure 3.9: Graphical determination of the molecular mass of batroxobin (B. Atrox, LATOXAN) from a SDS-PAGE gel. The blue line shows the standard curve produced from calculation of the Rf value for each MW marker. The point in magenta shows the position of batroxobin (B. atrox). The MW of this species is calculated at ~34kDa.

Once an estimate of the mass was known, using SDS-PAGE then a more accurate mass could be found. This was attempted using the electrospray mass spectrometer but, unfortunately, neither of the two forms of batroxobin (*B. atrox* or *B. moojeni*) showed any signs of flying within the electrospray instrument. Despite the fact that the conditions were altered in order to achieve a more successful outcome no Mr was obtained. It is still unclear as to the reason for the lack of flight, although the fact that the enzymes have a high degree of glycosylation (up to 30%) has been proposed as a possible explanation.

Since the electrospray mass spectrometer experiments had been unsuccessful a new approach to mass determination needed to be taken. Matrix Assisted Laser Desorption Ionisation Time of Flight Mass Spectrometer (MALDI TOF MS) was therefore used. This was not such a desirable method for mass determination, as the electrospray instrument can give masses to a1a.m.u accuracy, whereas the MALDI TOF is a much less accurate instrument with errors of more than 100 a.m.u.

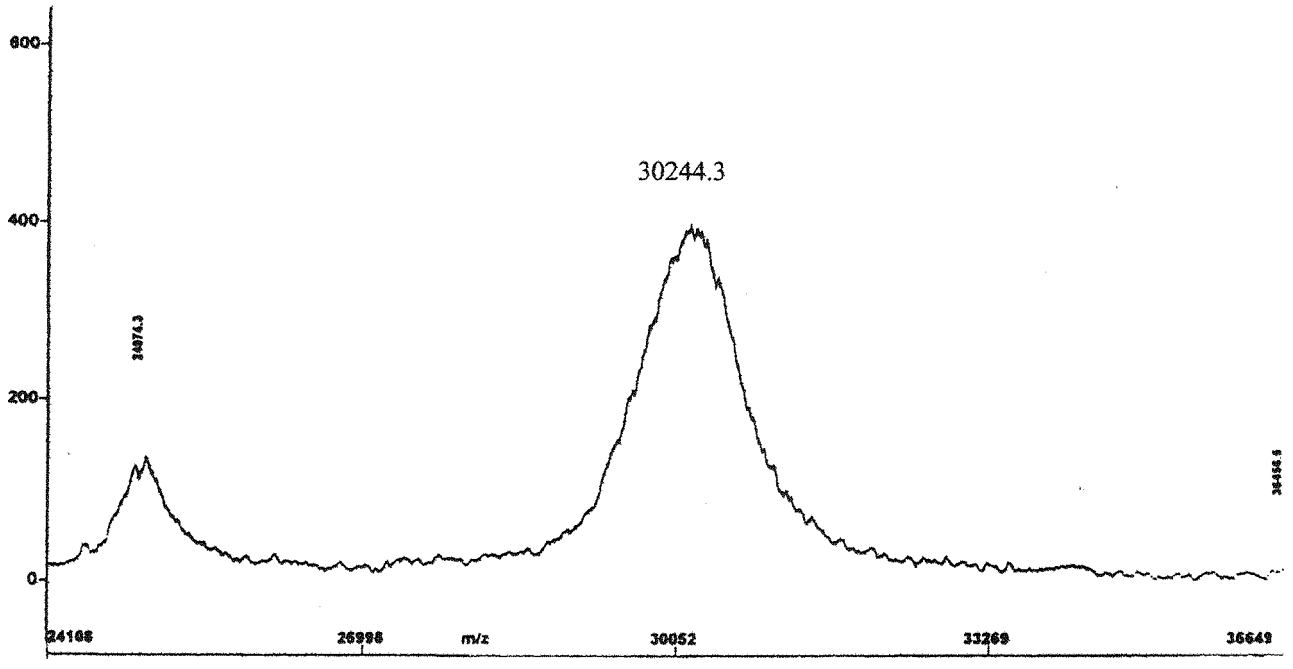


Figure 3.10: MALDI-TOF MS of batroxobin (B. moojeni).

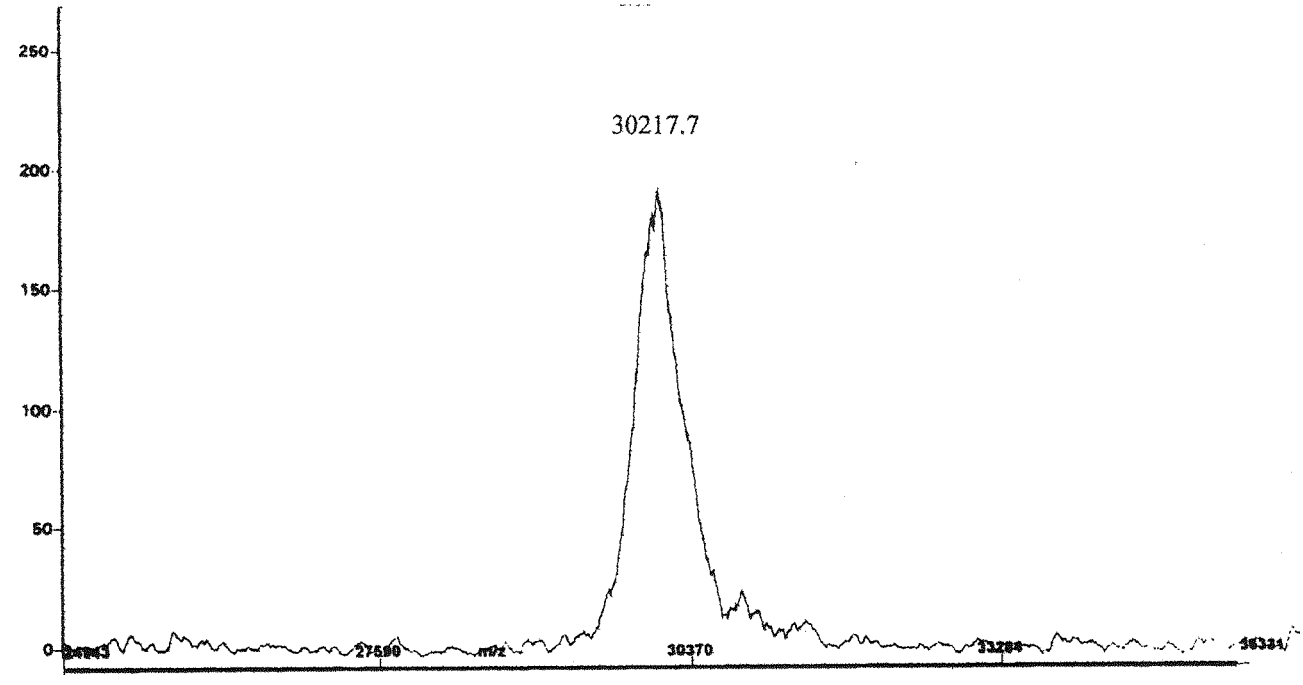


Figure 3.11: MALDI-TOF MS of batroxobin (B. atrox, LATOXAN).

As can be seen from these two profiles, within the error of the instrument, the proteins appear to have identical masses. This was highly unexpected, as it was believed from the literature that these two enzymes should have vastly differing masses[5]. Since there appeared to be a lower molecular mass protein in the sample of batroxobin (*B. moojeni*) obtained as a gift from OBL, it was initially believed that perhaps the problem lay in the highly invasive procedure of the MALDI-TOF cleaving away some of the glycosylation. The mass of the batroxobin (*B. atrox*) could have been less than expected because it could have been partially glycosylated.

However, on further questioning of the supplier (OBL), it was revealed also found the same protein in all their samples and that it was simply an impurity due to the problems outlined with the heparin column. It would therefore appear that the masses of these two enzymes are identical. Since literature suggests that there have been considerable problems with the identification of these snakes at source and there have been many instances of researchers being supplied with the wrong venom, it is possible that the reason why these proteins are showing identical masses is because they come from the same sub-species (*B. moojeni*).

However, on further inspection it would appear that the documented masses for the batroxobin (*B. atrox*) range from 26kDa to 42kDa. This variation is dependent on either the technique used to determine the mass, or the purity of the sample. A recent review suggests that much of the previous work was carried out on impure samples. Thus it is also possible that the premise that these enzymes should have differing masses may not be correct. It is also true to say that the masses of none of these enzymes have ever been determined using any method of mass spectrometry, and that most values are derived from SDS-PAGE analysis, which is a notoriously inaccurate method particularly for glycoproteins. Therefore, it is possible that the obvious conclusion is accurate, and that the thrombin-like enzyme derived from these two subspecies are identical.

3.5 Characterisation of batroxobin (*B. atrox*) supplied by Pentapharm.

Since it was unclear as to the authenticity of the pedigree of the venom supplied by Latoxan, it was decided that this work should be repeated using venoms supplied by Pentapharm, a guaranteed source. The proteins were purified as outlined previously

and the samples were N-terminally sequenced and their masses determined using electrospray MS.

BATROXOBIN **VIGGDECDINEHPFLAFMY-**
(B. atrox)

BATROXOBIN **VIGGDECDINEHPFLAFMY-**
(B. moojeni)

Figure 3.12: N-terminal sequence analysis of batroxobin (B. moojeni, Pentapharm) and Batroxobin (B. atrox, Pentapharm).

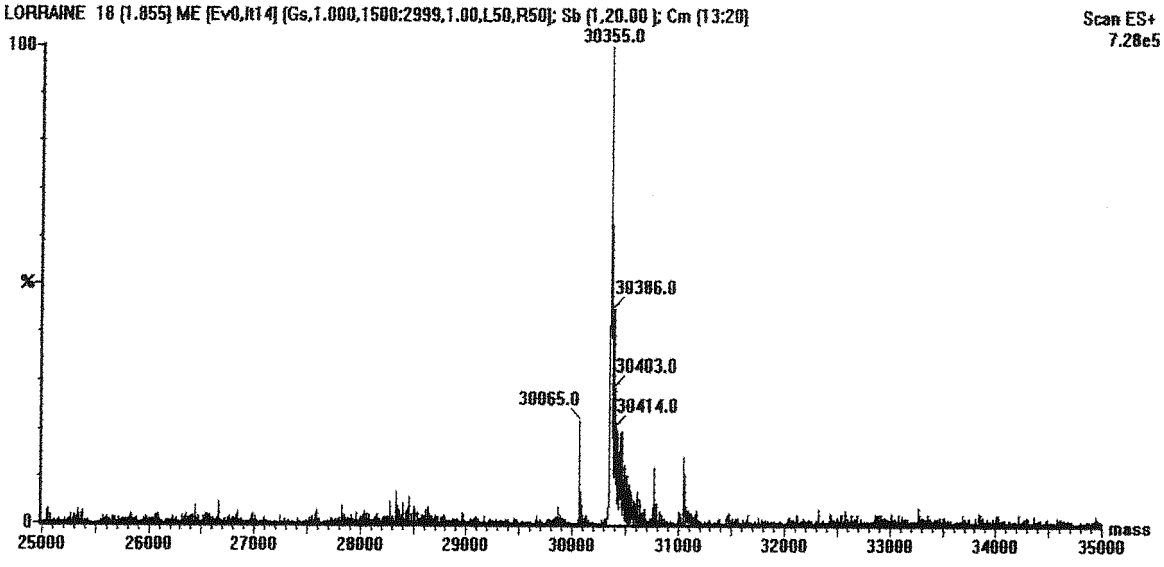


Figure 3.13: Electrospray MS analysis of batroxobin (B. moojeni, Pentapharm).

Unfortunately, it was never possible to produce any electrospray MS data for the batroxobin derived from *B. atrox*, despite successive attempts using extensive conditions including those equivalent to the ones used for batroxobin (*B. moojeni*). This may be due to differences in the carbohydrate content of the two enzymes, since batroxobin derived from *B. atrox* is believed to have a much higher carbohydrate content to that derived from *B. moojeni.*

3.6 Specificity studies of thrombin and batroxobin (*B. moojeni*) and (*B. atrox*) on various fibrinogens.

One of the major aims of this project was to probe the specificity of batroxobin. This has been considered to be an area of interest, due to the fact that the enzyme acts in a manner similar, but not identical, to thrombin. As has been previously mentioned in Chapter 1, batroxobin is a much more highly selective enzyme than thrombin and is only capable of cleaving the arg16-gly17 bond within Aα chain of fibrinogen, resulting in the cleavage of fibrinopeptide A and the formation of the fibrin I monomer. Thrombin on the other hand is capable of the additional cleavage of an arg-gly bond within the Bβ chain of fibrinogen to produce the fibrinopeptide B and the fibrin II monomer. Thrombin is also able to interact with the coagulation process at a variety of other positions, allowing it to be involved in the regulation of the overall process. Therefore any information that can be obtained which may help to explain the increased specificity of batroxobin compared to thrombin would be useful in explaining the activities of the two enzymes[88,89,90,91].

It was decided that two differing methods should be attempted to probe this problem. These include both the use of the natural substrate of batroxobin fibrinogen, and also artificial peptide substrates that have been specifically designed to address the problem posed, as discussed in Chapter 4. Batroxobin, like many other enzymes, acts in a species-specific manner. This means that its activity is different with fibrinogen obtained from different species. Consequently, it may be possible to look at the activity of the enzyme with differing fibrinogens and then relate the activity to the known sequences of those fibrinogens.

3.6.1 Reaction of thrombin and batroxobin with various fibrinogens

Close consideration of the effect of batroxobin *in vivo* on a variety of different species was used as an initial guideline to the types of fibrinogen that should be probed in this series of experiments. Certain unpublished data supplied by OBL suggested that the venous, subcutaneous and intramuscular injection of batroxobin into different animals had vastly differing effects on the animals. These differences were in terms of the daily dose in BUs per kg body mass needed to have a defibrinogenating effect on the animal.

These values were used as an indication as to the species that should be considered, whilst considering the possibility that the values obtained from the *in vivo* studies may not be confirmed when comparing activity with purified fibrinogens. This is due to the fact that there are other factors to take into consideration when looking at the effect of the batroxobin on the complete coagulation cascade, as compared to the purified substrate. This is of particular importance when considering the fact that the *in vivo* studies show low activity of batroxobin with the rat, which is the natural prey of the pit-viper family (including *Bothrops sp.*) from which batroxobin is collected.

In addition to this information, the other criteria that were used when deciding which fibrinogens to study initially were those that already had a published primary structure, as these would be of much more value when attempting to decipher which aspects of the sequence had an effect on the specificity. Therefore, although all the above fibrinogens were used in this study, it was necessary in some cases to resolve the amino acid sequence in the vicinity of the scissile bond in order to clarify the sequence specificity for each enzyme.

Species	Mode of application	Daily defibrinogenating dose* (BU) required per kg body weight
Human	subcutaneous	0.2-1.0
	intravenous	0.5-2.0
Dog	subcutaneous	0.4-2.0
	intravenous	1.0-2.0
	intramuscular	~5
Cat	intravenous	0.2
Pig	intravenous	~1000
Rabbit	intravenous	~1000
Rat	intravenous	~200
Mouse	intravenous	~40
	subcutaneous	~20

*Plasma fibrinogen: <100mg% during 24hrs

Table 3.2: In vivo effect of batroxobin on various species after differing modes of injection

3.6.2 Experimental procedure

The fibrinogen specificity assays were carried out using the coagulation of the fibrinogen to fibrin as the easiest means of assessing the progress of the reaction. This was carried out using a straightforward clot time as the end point of the reaction, as well as a kinetic study. The clot time was monitored using the coagulometer at Oxford Bioresearch Laboratories.

3.6.2.1 End point clot time

A 1mg/ml solution of each fibrinogen was made up in phosphate buffered saline (PBS) pH 7.4. This fibrinogen concentration refers to the clottable fibrinogen

present in the sample. This was calculated as described in Chapter Two. Solutions of batroxobin and thrombin at 0.1, 0.2 and 0.3 mg/ml w/v concentrations were made up in PBS. The method was followed as described in the clot time assay of Chapter 2, with a final concentration of 0.5mg/ml clottable fibrinogen used for each assay in a final volume of 0.5ml. The reactions were carried out using each concentration of thrombin and batroxobin in triplicate, both with and without 2mM $CaCl_2$ solution. All buffers were saturated with EGTA, a specific Ca^{2+} chelator for the reactions where calcium was to be omitted. The calcium ions were included in this set of experiments as there is some contradictory evidence as to the requirement for Ca^{2+} in the reactions with batroxobin, although the need for calcium ions with thrombin is well documented.

3.6.2.2 Kinetic study of fibrin formation.

A series of 0.8, 1.6, 2.4, 3.2 and 4mg/ml concentrations of each fibrinogen were made up in PBS. The method was followed as described in Chapter 2, with final concentrations of clottable protein in each reaction of 0.2, 0.4, 0.6, 0.8 and 1mg/ml respectively. The reactions were carried out using each concentration of thrombin and batroxobin, both with and without 2mM $CaCl_2$. All buffers were saturated with EGTA, a specific Ca^{2+} chelator for those reactions where calcium was to be omitted.

3.6.3 Comparison of the role of Ca^{2+} in clot formation between thrombin and batroxobin from *B. atrox* and *B. moojeni.*

The standard end point clotting times were performed, as outlined in Chapter 2, with the buffers saturated with EGTA, a specific Ca^{2+} chelator in order to remove all Ca^{2+} present. No clot formation was detected for any of the reactions at any concentration of fibrinogen or enzyme for thrombin or either of the batroxobins. This finding was confirmed by kinetic assays of thrombin and the two batroxobins (*B. atrox* and *B. moojeni*), with various concentrations of fibrinogen from all species, as outlined in Chapter 2, once again with all buffers saturated with EGTA. No change in turbidity was detected without the presence of Ca^{2+}, showing that no clot formation took place.

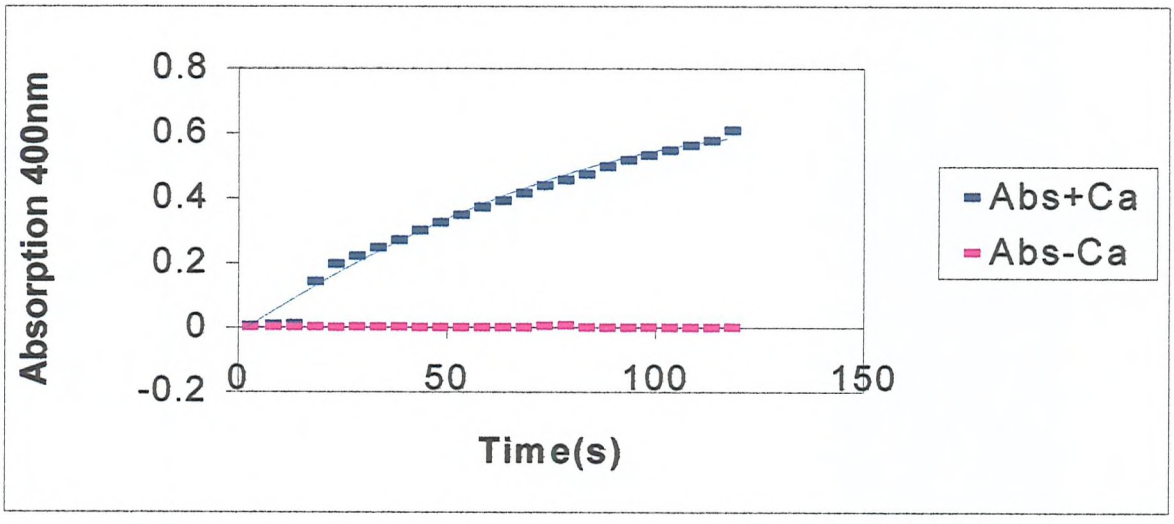


Figure 3.15: Effect of Ca^{2+} on clot formation with thrombin or batroxobin.

3.6.4 Comparison of clot formation with various fibrinogens between thrombin and batroxobin (*B. moojeni)* and (*B. atrox*).

The end point clot formation assays were carried out as outlined in Chapter 2 in the presence of Ca^{2+} ions. Seven fibrinogens from various species were chosen: human, cow, pig, cat, dog, rabbit, rat and mouse, and their clottable fibrinogen concentration were standardised as outlined in Chapter 2. The samples were then desalted in a PD10 column, lyophilised and re-dissolved in PBS, pH 7.4, to a standard clottable protein concentration. In this way it was possible to ensure that any differences in clot formation were due to species differences and not due to altering clottable fibrinogen concentrations or ionic strength of the solutions since batroxobin is very sensitive to changes in ionic strength.

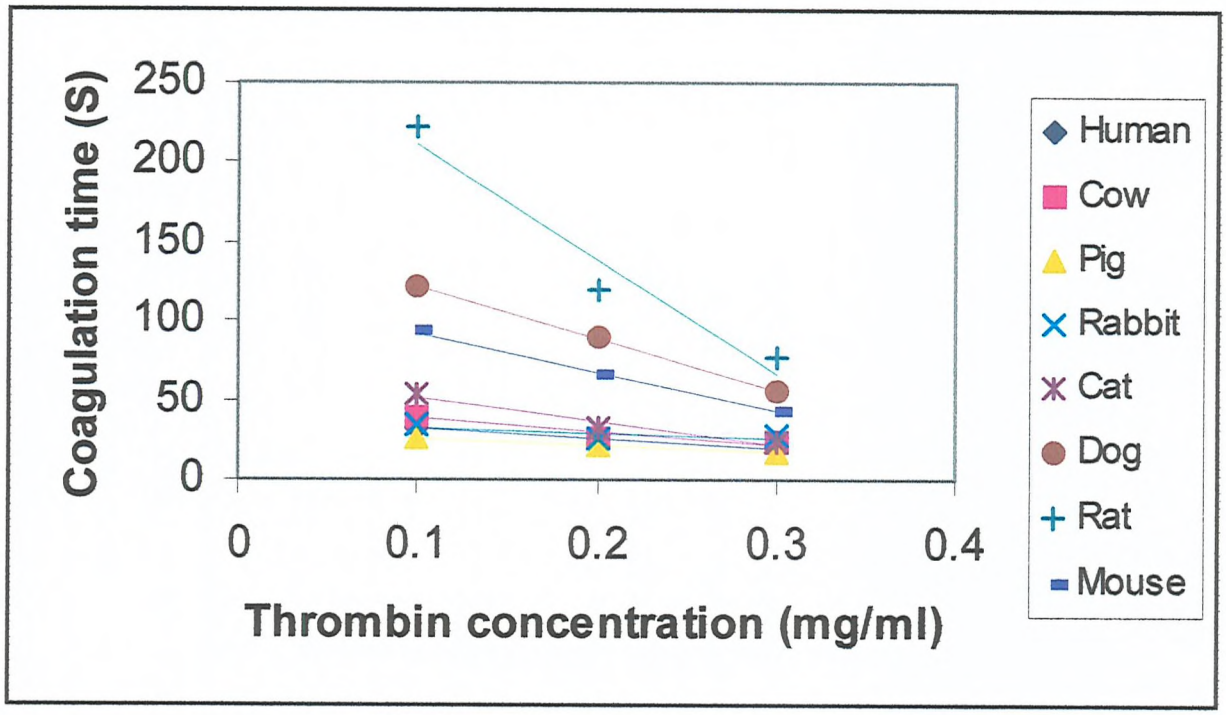


Figure 3.16: Comparison of end point clot formation of thrombin with various fibrinogens.

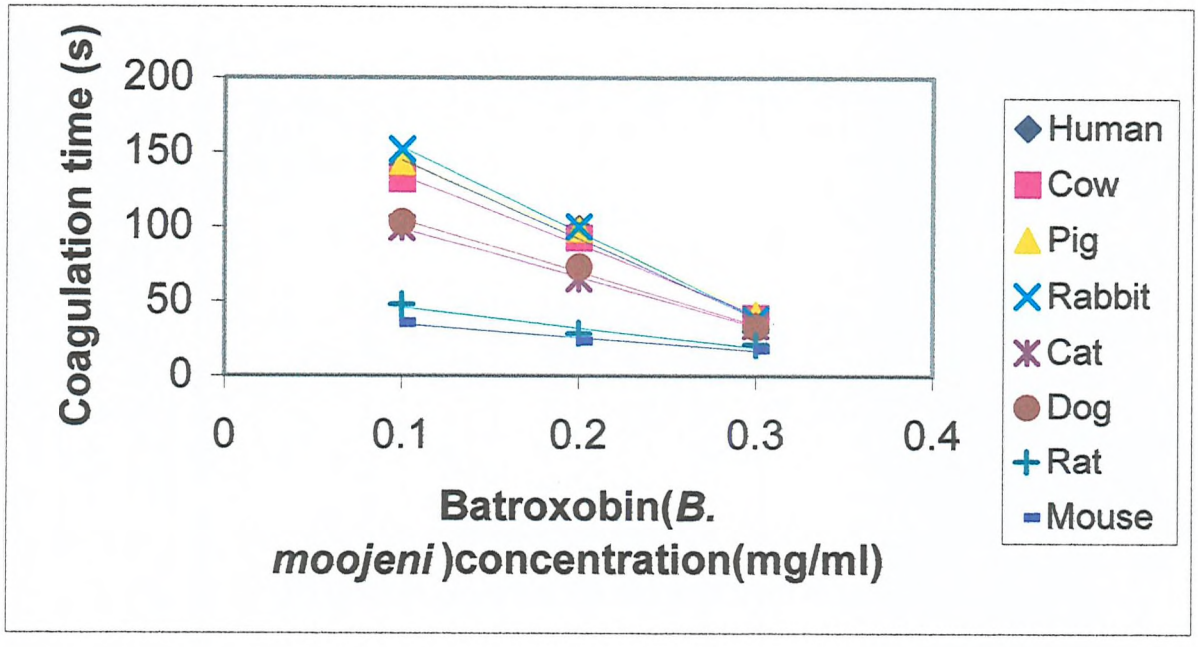


Figure 3.17: Comparison of end-point clot time for batroxobin (B. moojeni, Pentapharm) with various fibrinogens.

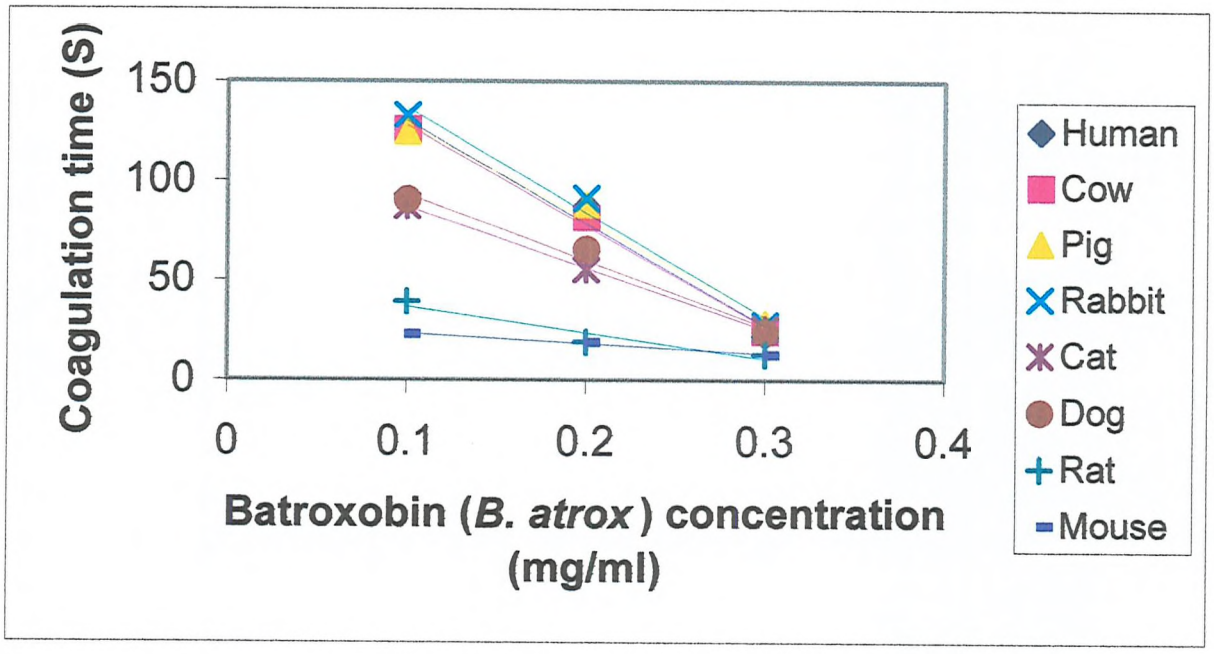


Figure 3.18: Comparison of end-point clot time for batroxobin (B. atrox, Pentapharm) with various fibrinogens.

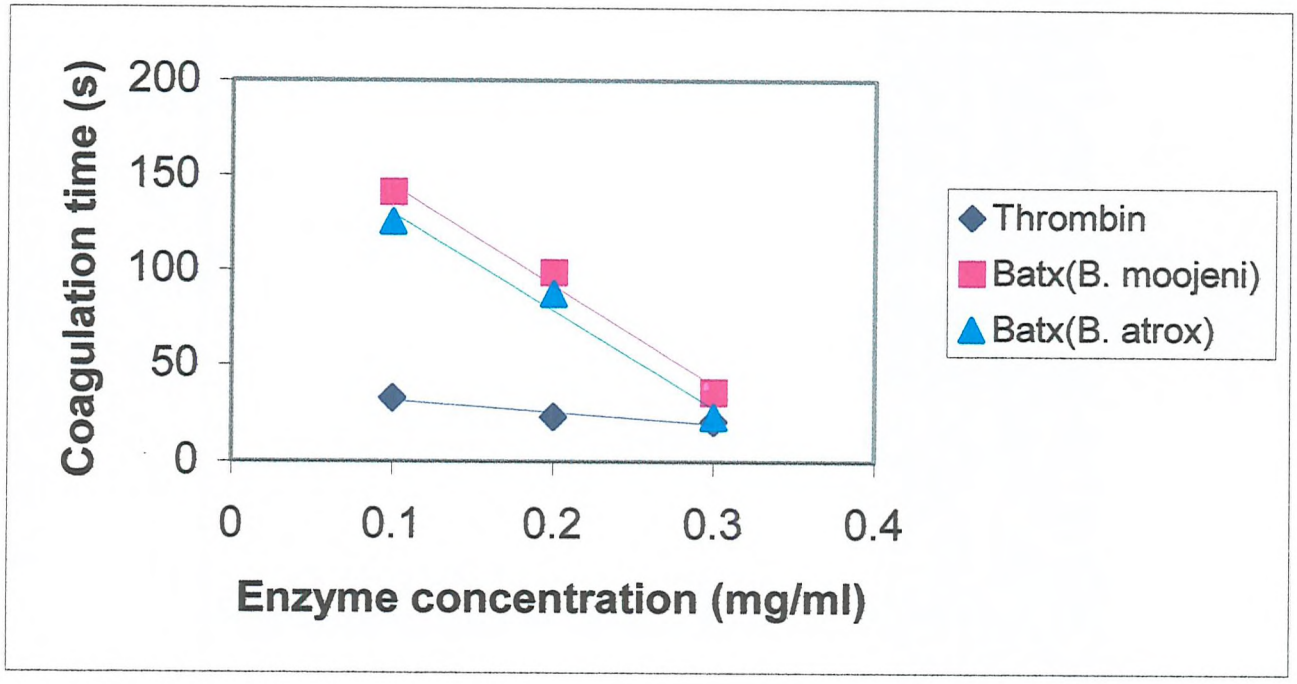


Figure 3.19: Comparison of end-point clot time for human fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm).

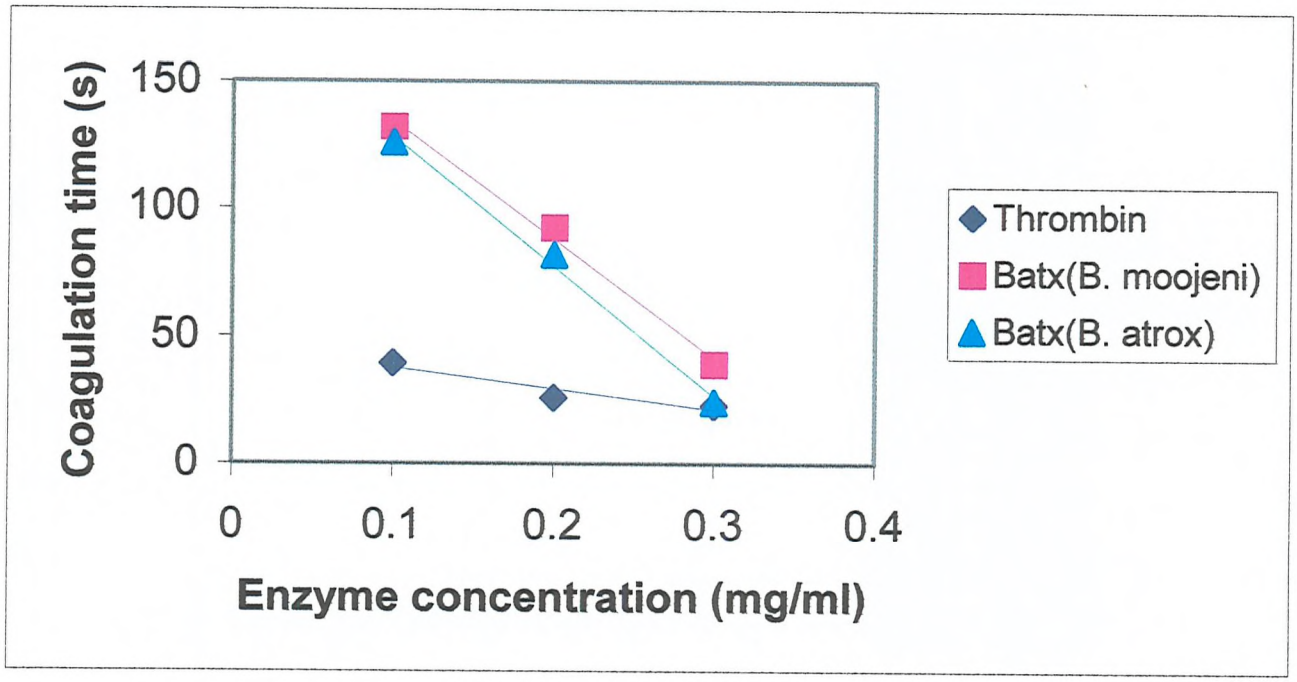


Figure 3.20: Comparison of end-point clot time for cow fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm).

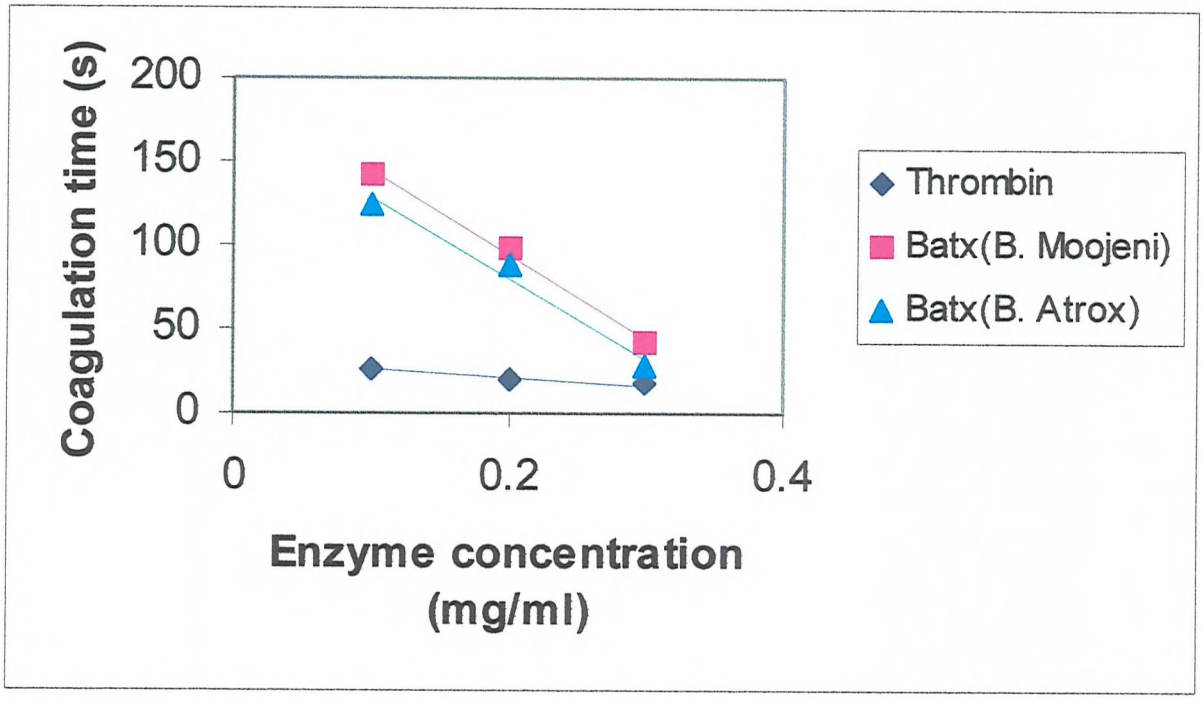


Figure 3.21: Comparison of end-point clot time for pig fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm)

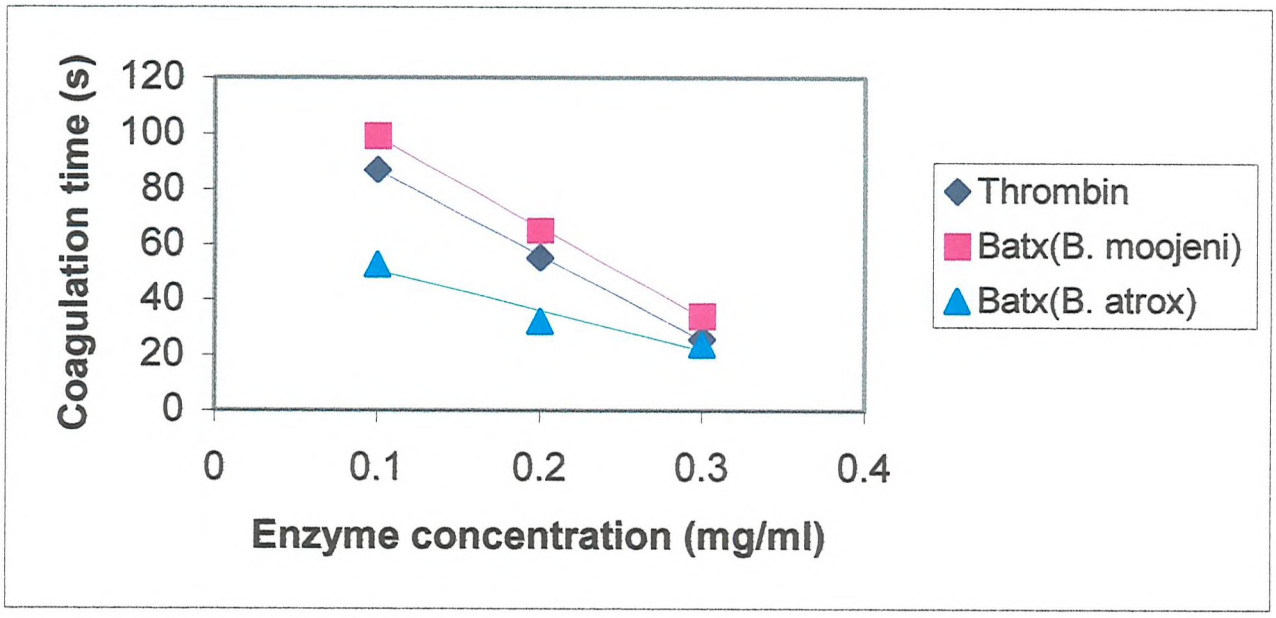


Figure 3.22: Comparison of end-point clot time for cat fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm).

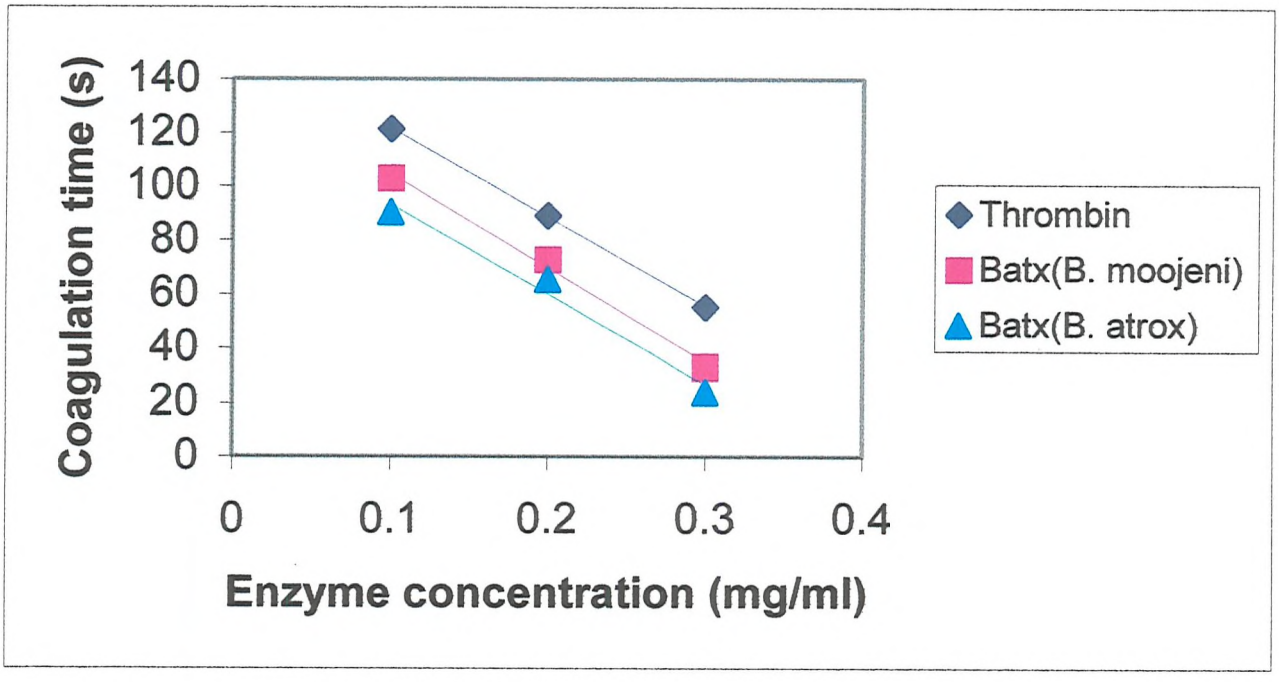


Figure 3.23: Comparison of end-point clot time for dog fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm).

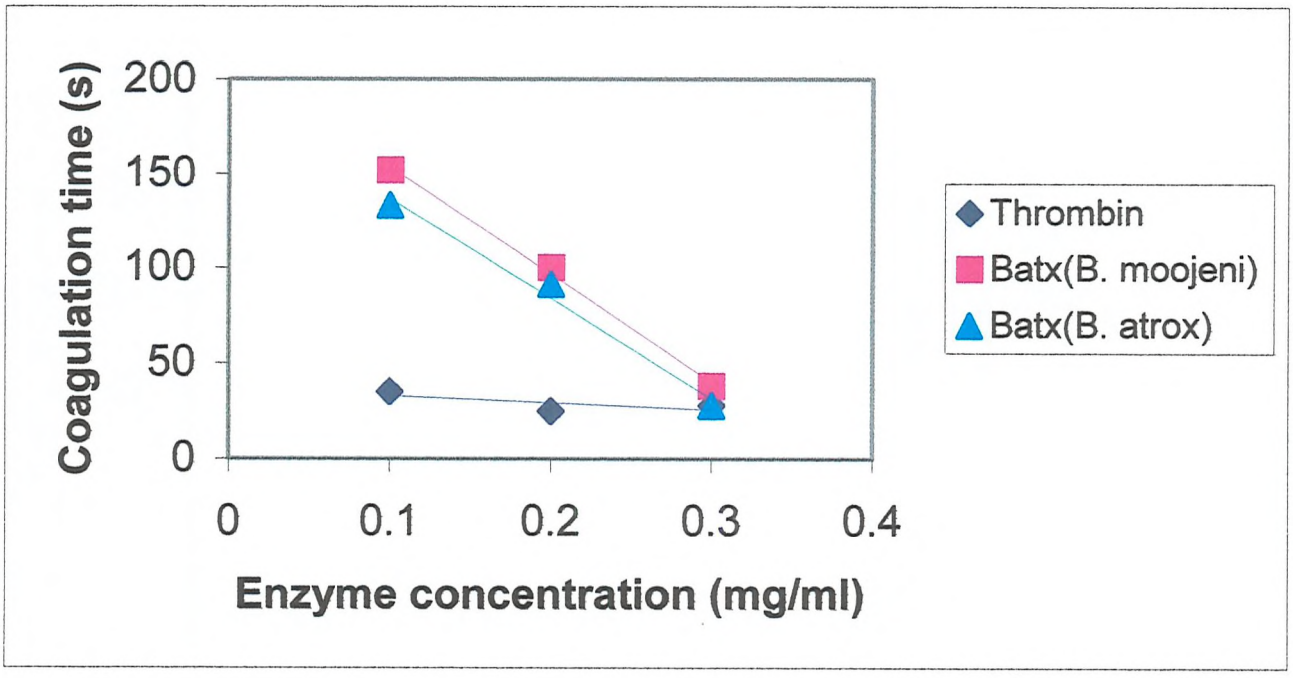


Figure 3.24: Comparison of end-point clot formation for rabbit fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm).

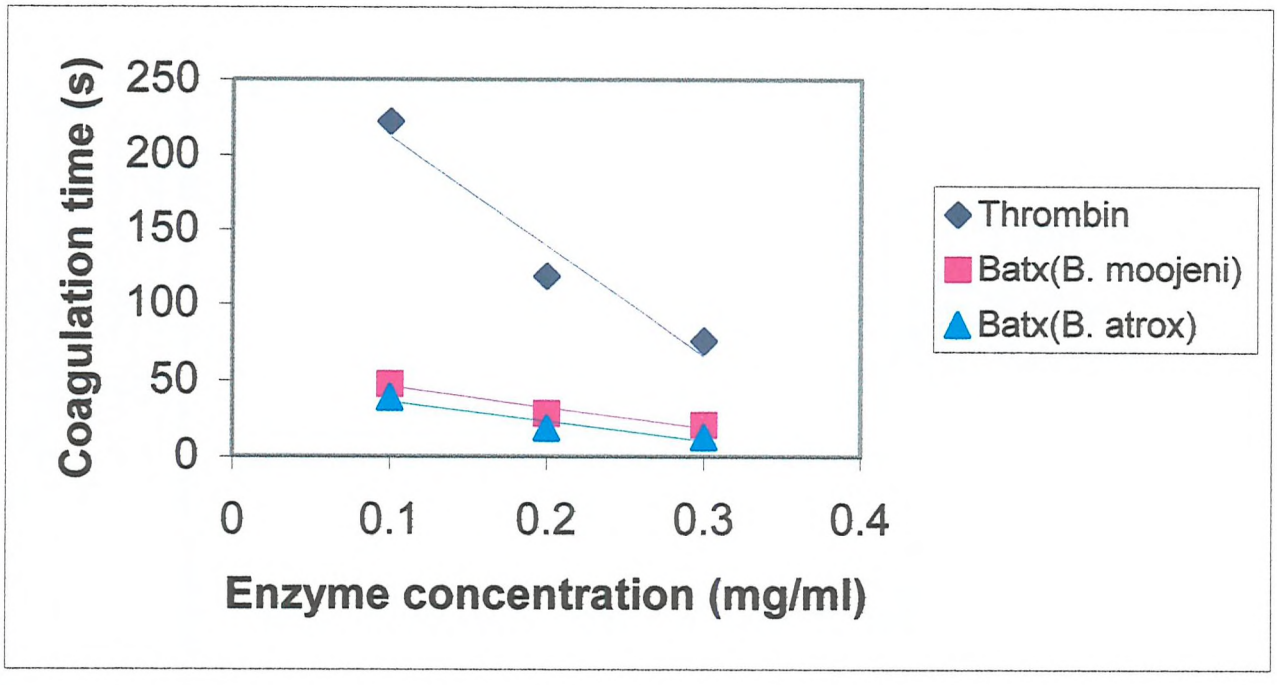


Figure 3.25: Comparison of end-point clot time for rat fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm)

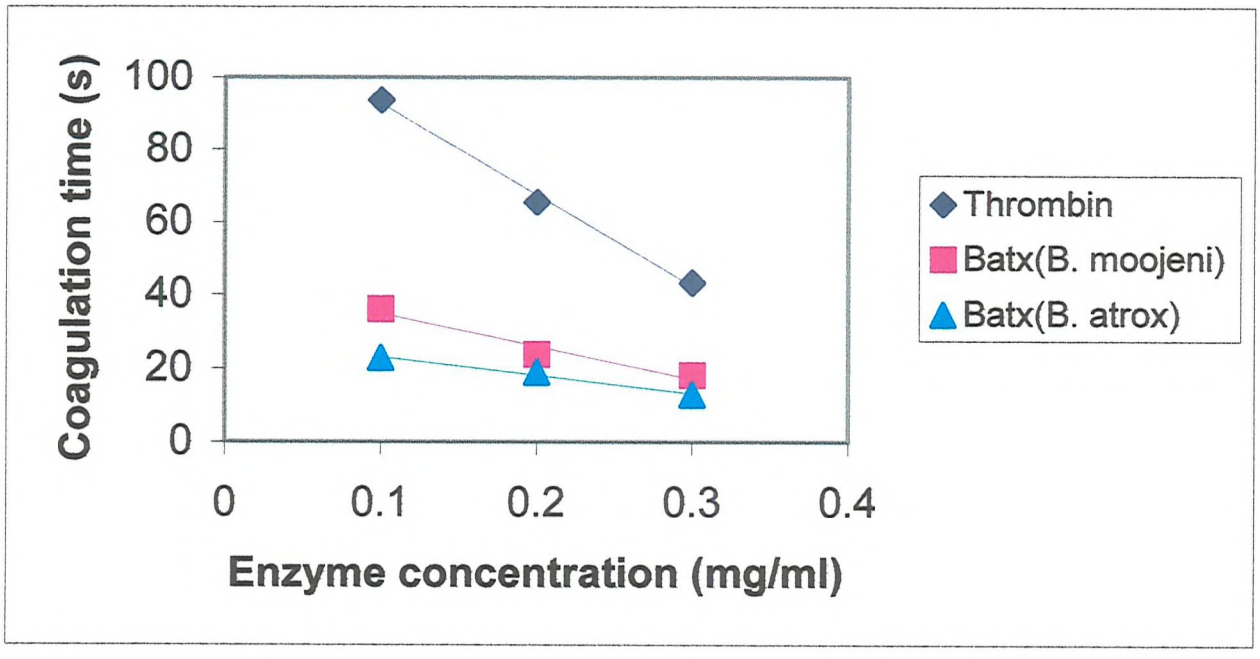


Figure 3.26: Comparison of end-point clot time with mouse fibrinogen with thrombin and batroxobin (B. moojeni and B. atrox, Pentapharm).

The end-point clot time assays indicate that there are some very distinct differences between the natural substrate specificity of thrombin and batroxobin. The thrombin used in these studies was obtained from human plasma and it is therefore perhaps not surprising that this enzyme has a high degree of specificity for human fibrinogen. However it also shows a high degree of specificity for the fibrinogens derived from porcine and bovine blood. Conversely, batroxobin derived from both species have much faster clot time formation with the fibrinogens derived from rat and mouse blood, which is interesting as these are the natural prey of the two snakes. However, one point of considerable interest is that the batroxobin from *B. atrox* is almost twice as active to all the fibrinogens compared with the enzyme derived from the closely related snake, *B. moojeni*.

3.6.5 Kinetic study of clot formation for thrombin and batroxobin with various fibrinogens.

In order to clarify the natural substrate specificity of each enzyme for the various fibrinogens, the work was repeated as a kinetic study as outlined in Chapter 2 and section 3.6.2.2.

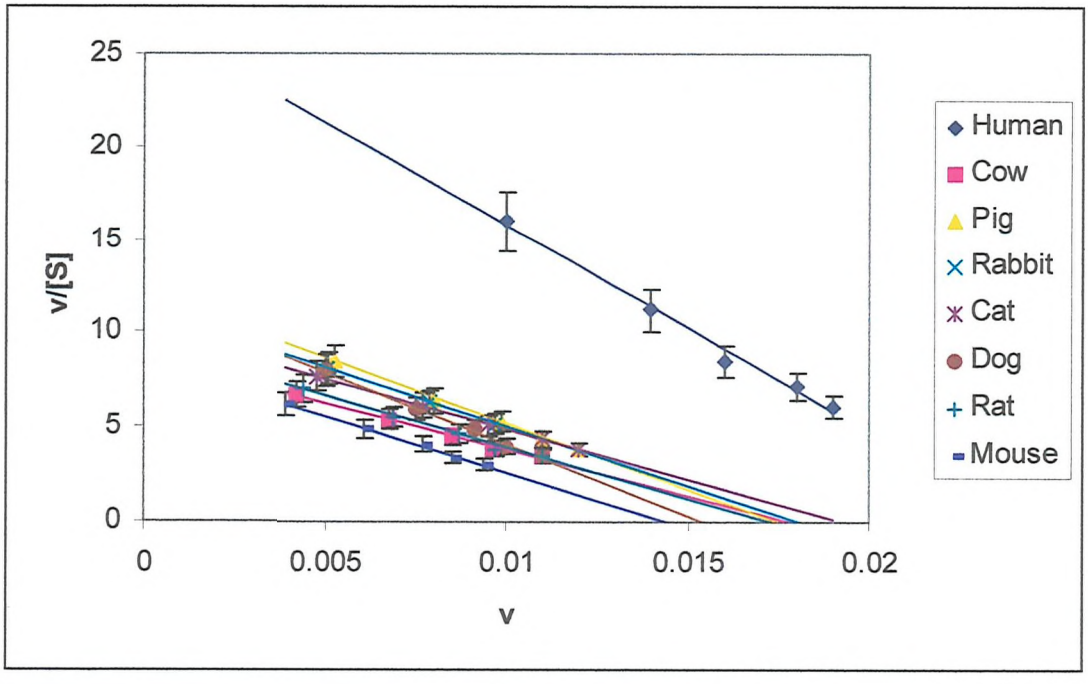


Figure 3.27: Eadie-Hoffstee Plot of thrombin with fibrinogens from various species. Each fibrinogen is represented by a different colour as described in the legend.

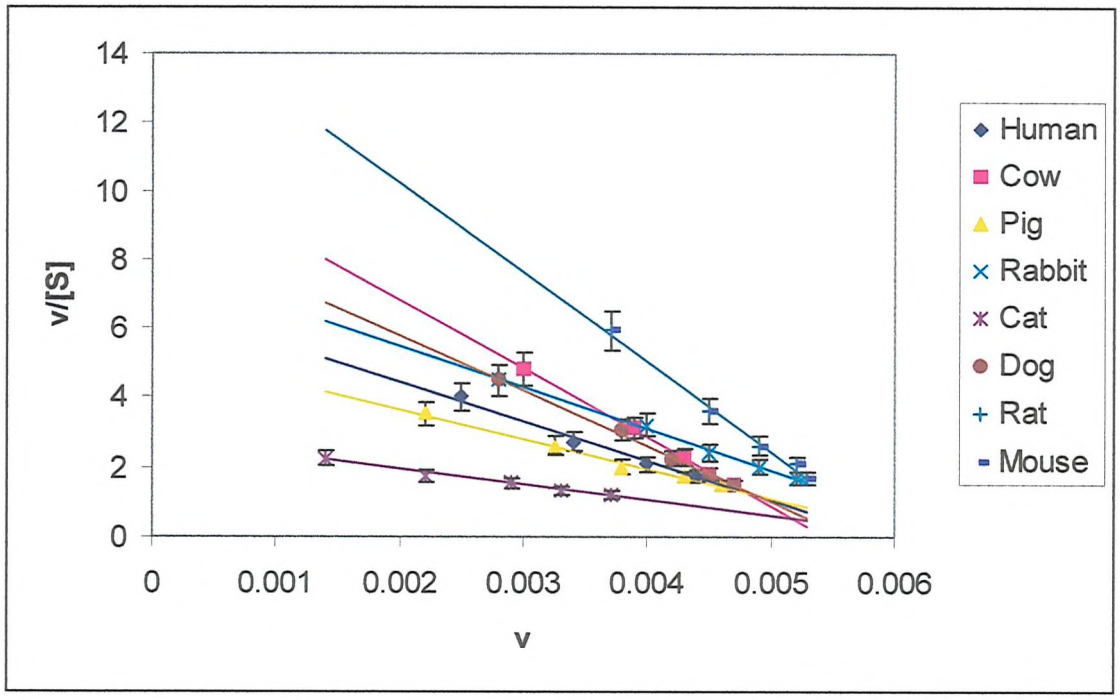


Figure 3.28: Eadie-Hoffstee plot of batroxobin (B. moojeni) with fibrinogens from various species. Each fibrinogen is represented by a different colour as described in the legend.

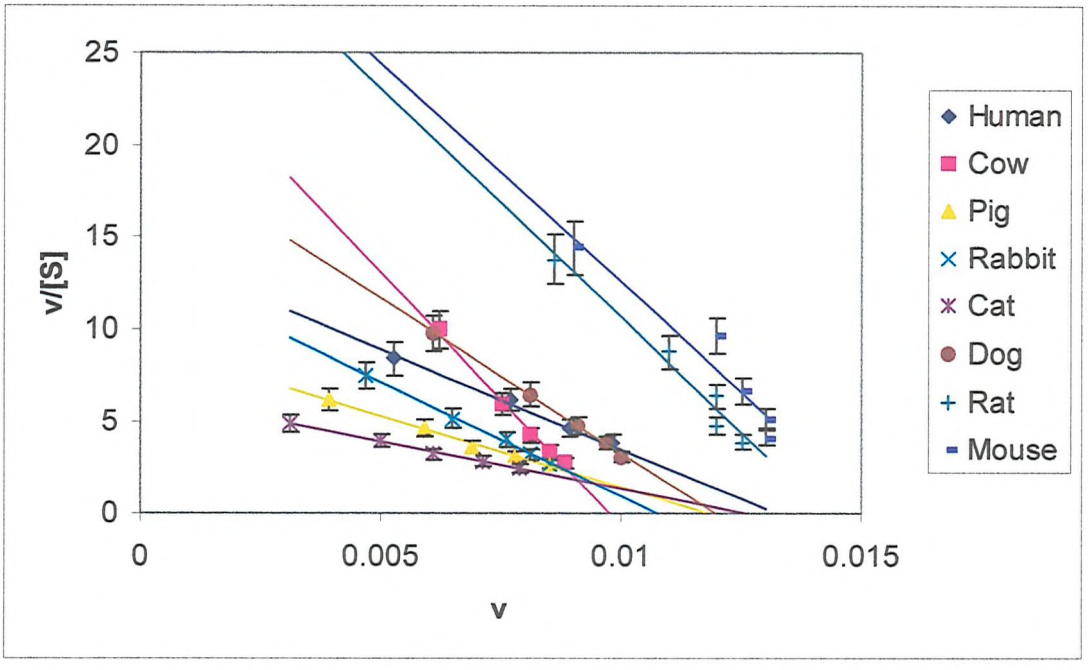


Figure 3.29: Eadie-Hofstee plot for batroxobin (B. atrox) with fibrinogen from various species. Each fibrinogen is represented by a different colour as described in the legend.

The data obtained from the kinetic study of thrombin and batroxobin on various fibrinogens was analysed by plotting Hofstee-Eadie plots as shown in *Figures 3.27-3.29*. This was used in order to determine K_M and V_{max} values for each enzyme with each substrate. It has been possible, from this information, to calculate k_{cat}/K_M for each enzyme with each substrate, which is an expression that considers the specificity of a substrate for an enzyme.

$$V_{max} = k_{cat}[E]_t$$

where $[E]_t$ = total enzyme conc.

k_{cat} = turnover number

FIBRINOGEN	V_{max} (s^{-1})	K_M (mM)	k_{cat}	k_{cat}/K_M
Human	0.0242	1	2.42×10^5	2.4×10^8
Bovine	0.0173	2	1.7×10^5	0.8×10^8
Porcine	0.0172	1.5	1.7×10^5	1.1×10^8
Rabbit	0.0178	1.6	1.8×10^5	1.1×10^8
Cat	0.0184	1.8	1.8×10^5	1.0×10^8
Dog	0.0151	1.3	1.5×10^5	1.2×10^8
Rat	0.0158	1.7	1.6×10^5	0.9×10^8
Mouse	0.0164	1.6	1.6×10^5	1.0×10^8

Table 3.3: Kinetic data for thrombin with various fibrinogens.

FIBRINOGEN	V_{max} (s^{-1})	K_M (mM)	k_{cat}	k_{cat}/K_M
Human	0.0138	1	1.4×10^5	1.4×10^8
Bovine	0.01	0.4	1.0×10^5	2.5×10^8
Porcine	0.0119	1.2	1.2×10^5	1.0×10^8
Rabbit	0.0108	0.8	1.1×10^5	1.4×10^8
Cat	0.013	2	1.3×10^5	0.6×10^8
Dog	0.012	0.6	1.2×10^5	2.0×10^8
Rat	0.014	0.4	1.4×10^5	3.5×10^8
Mouse	0.0145	0.4	1.45×10^5	3.53×10^8

Table 3.4: Kinetic data for batroxobin (B. atrox) with various fibrinogens.

FIBRINOGEN	V_{max} (s^{-1})	K_M (mM)	k_{cat}	k_{cat}/K_M
Human	0.0061	1	0.6×10^5	0.6×10^8
Bovine	0.0054	0.5	0.5×10^5	1.0×10^8
Porcine	0.0063	1.2	0.6×10^5	0.5×10^8
Rabbit	0.0064	0.8	0.6×10^5	0.8×10^8
Cat	0.0059	2	0.6×10^5	0.3×10^8
Dog	0.0055	0.6	0.5×10^5	0.9×10^8
Rat	0.0060	0.4	0.6×10^5	1.5×10^8
1	0.0060	0.4	0.6×10^5	1.5×10^8

Table 3.5: Kinetic data for batroxobin (B. moojeni) with various fibrinogens.

As can be seen from the kinetic data shown in *Tables 3.3-3.5*, thrombin and batroxobin show very different specificity for various substrates to each other. Thrombin shows its greatest specificity for human fibrinogen, which is perhaps not surprising as the thrombin is derived from a human blood source. Both batroxobins on the other hand, show much greater specificity for rat and mouse fibrinogens than for any of the others. The results obtained from this kinetic study support the results obtained from the end-point clot times, both in terms of the order of specificity for each enzyme and substrate, but also in terms of the higher activity of batroxobin (*B. atrox*) compared to batroxobin *(B. moojeni)*.

The k_{cat}/K_M for each of these enzymes with every substrate indicates that they are all highly evolved towards reacting with any of the fibrinogen substrates. This indicates that they have good specificity for all fibrinogens, but the increased values for specific substrates suggests that the enzymes have evolved to catalyse the reactions with the fibrinogen they need to act on most effectively.

3.7 Crystallographic trials.

3.7.1 Introduction.

X-ray crystallography is the single best method for determination of the molecular structure of large biological molecules. It is a technique that makes use of

the physical phenomenon of interference of waves when they come into contact with an object. This interference can either be constructive, when the wave is enhanced, or destructive, when it is diminished. Therefore, when an object, for example an atom obstructs a wave, it produces a characteristic diffraction pattern made up of regions of enhanced and diminished intensities.

X-rays possess wavelengths that are very similar to bond lengths found in molecules and hence the spacing of atoms in crystals (~0.1nm); therefore these X-rays can be diffracted by atoms. Since every structure has a characteristic diffraction pattern, due to the unique way in which the atoms are arranged in space, it is possible to use the diffraction pattern as a blueprint for the structure of the molecule. By analysing the pattern obtained, it is possible to produce a highly detailed picture of each of the atoms within the structure and their corresponding positioning within the overall molecule. X-rays are not alone in their ability to be diffracted by atoms, for example, electrons that have been accelerated through 4kV to ~20 000km/s have wavelengths of 40pm, and neutrons, that have been slowed to thermal velocities, have similar wavelengths and can also be used[92].

3.7.2 Crystal growth.

This work has made use of a technique for growing crystals called the hanging-drop method. This technique allows protein crystals to be grown at high concentrations of protein, whilst using only relatively small volumes. Thus it is of great value when attempting crystallography using relatively limited amounts of protein.

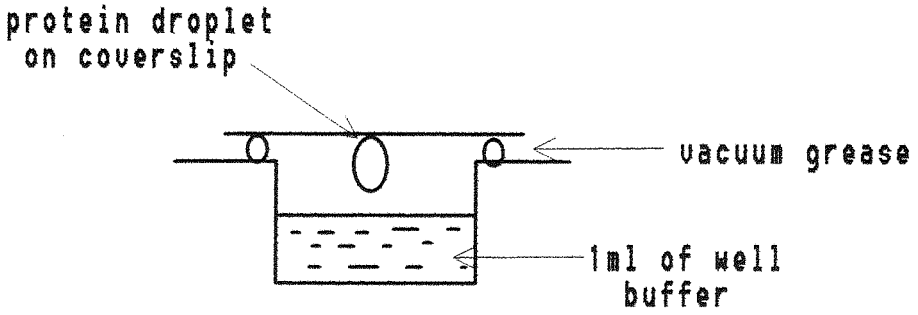


Figure 3.30: Hanging - drop method of crystal growth.

In the method used the lip of each of the wells was greased with high vacuum grease, and the well was filled with 1ml of the buffer. A 1-4µl droplet of the buffer, dependent on the volume of protein solution being used, was placed on a cover slip and presiliconised with dimethyldichlorosilane. An equal volume of the protein solution was added to the droplet and mixed thoroughly. This droplet was then suspended over the corresponding 'well' buffer. The plates were incubated for up to two months at constant temperatures waiting for crystal formation.

The process of crystal growth by the hanging-drop method works by slow diffusion. Due to the closed environment of each well and droplet, a diffusion gradient is established between the well and the droplet. Gradually, as the water is removed from the droplet into the environment between the drop and the buffer, the concentration of the protein in the droplet increases very slowly, along with the concentration of salt. Due to the natural preference of molecules to aggregate at higher concentrations the protein molecules start to collect together. If this process occurs too rapidly disordered aggregates will result and the protein will precipitate out of solution. However, if all the conditions are correct, the aggregation of protein molecules will be uniform and controlled, resulting in the formation of crystals.

The initial screens for crystallography are normally carried out using the Hamilton screen. These are a set of one hundred pre-made combinations of buffers, precipitants and salts at set pH values. These make use of those combinations of reagents that have been shown to be successful at aiding the crystallisation of many proteins[93]. Once the type of conditions that may result in the formation of crystals has been found the conditions can be further refined, making use of the Hamilton screen solution as a starting point. Many attempts at this refinement may be necessary in order to ascertain the ideal conditions for crystallisation of any given protein to produce usable crystals.

3.7.3 Initial trial

Initially a Hamilton screen trial using 1µl of protein solution at a concentration of 10mg/ml in each well was carried out. This made use of protein that had been isolated from the *B. atrox* venom supplied by LATOXAN. The plates were laid down for up to 2 months (in total) at 19°C in the dark. Although there was some limited,

transient crystal formation, it was decided that this concentration of protein was insufficient to enable any high quality crystal growth. Therefore the procedure was repeated using an increased concentration.

3.7.4 Second trial

The Hamilton screen trial was repeated as above, using a protein concentration of 35mg/ml. Once again the trays were laid down at 19°C, in the dark, for three weeks. After this time it became apparent that some crystal growth had occurred, but it was decided that the plates should be left for a further three weeks to see if any change in the crystal formation resulted.

After this time, all the wells that had contained crystals, with the exception of one, had lost their crystals by dissolving back into solution, a phenomenon that was also seen in the initial trial. However, one well (well 40 of Screen I) had significant crystal development (*Figure 3.31*). These crystals were not of a high enough quality to enable their use for the collection of diffraction data, but they did pinpoint the conditions around which further trials should be made.

3.7.5 Refined trial

It was considered appropriate to use the conditions of well 40 Screen I as a starting point for the refinement of the conditions necessary for crystallisation of the batroxobin from *B. atrox*.

The solution contained:

Buffer - 0.1 M Sodium Citrate, pH 5.6

Precipitant - 20% v/v 2-propanol and

-20% w/v PEG 4000

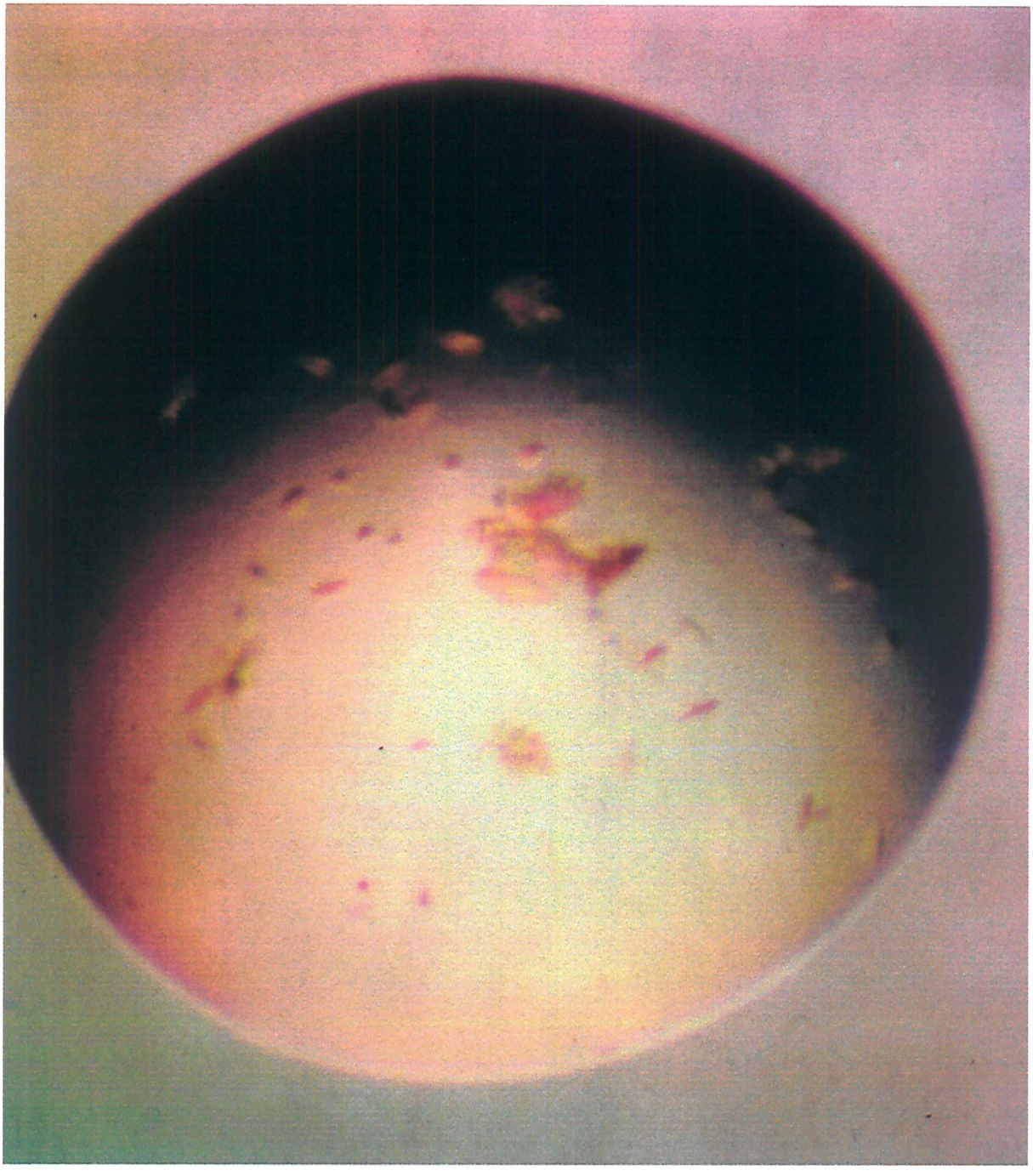


Figure 3.31: Photograph showing crystal formation in the second Hamilton crystal screen.

It was decided that this should be attempted immediately, using protein that had been supplied by Oxford Bioresearch Laboratories. Therefore this protein was plated out at 25mg/ml using the well conditions outlined above. The plates were left for a total of 2 months at 19°C in the dark. However, crystals did not form at any point during this time.

1	2	3	4	5	6
100mM Na Citrate 10% PEG 4000 10% 2-Propanol	100mM Na Citrate 15% PEG 4000 15% 2-Propanol	100mM Na Citrate 20% PEG 4000 20% 2-Propanol	100mM Na Citrate 25% PEG 4000 25% 2-Propanol	100mM Na Citrate 30% PEG 4000 30% 2-Propanol	
100mM Na Citrate 40% PEG 4000 - % 2-Propanol	100mM Na Citrate 30% PEG 4000 - % 2-Propanol	100mM Na Citrate 20% PEG 4000 -% 2-Propanol	100mM Na Citrate -% PEG 4000 40% 2-Propanol	100mM Na Citrate -% PEG 4000 30% 2-Propanol	100mM Na Citrate -% PEG 4000 30% 2-Propanol
100mM Na Citrate 30% PEG 4000 10% 2-Propanol	100mM Na Citrate 10% PEG 4000 30% 2-Propanol				

Table 3.6: Refined well conditions at pH 5, 5.5 and 6.

A possible reason for this could have been due to the excessive manipulation of the protein that was necessary before use for crystallisation. The lack of information with regard to the protein solution from which this sample of enzyme had originally been freeze dried was thought to be important, since it was possible that the protein had already denatured before being plated out. Another possible explanation was that, since there is some evidence that the batroxobin supplied by OBL was derived from a different sub-species of *B. atrox*, there may be slight differences in the carbohydrate content which may have resulted in a differing response during crystallography. Alternatively, there is also a possibility that the protein that was purified on site at Southampton was stabilised in some way by the presence of small amounts of benzamidine that remained bound to the protein after the purification process. This possibility was somewhat supported by the fact that the reported structures for porcine kallikrein and bovine trypsin were both crystallised in the presence of benzamidine.

Consequently, it was decided that the next trial should be repeated using protein that had been reclaimed from the crystal plates containing the batroxobin purified on site. This protein was then laid down in the way outlined above, at 19°C in the dark for 3 weeks. On inspection it was apparent that two of the wells contained crystals at positions A2 and A3 of the plate.

Conditions A2 100mM Sodium citrate, pH5
15% PEG 4000
15% 2-Propanol

Conditions A3 100mM Sodium citrate, pH 5
20% PEG 4000
20% 2-Propanol

These are final concentrations in an overall volume of 1ml

The conditions in A3 showed the best crystal formation. Some further work is required, to make more refinements to these conditions. It appears obvious that the conditions associated with pH and sodium citrate are likely to be correct, however, the actual concentrations of 2-propanol and PEG 4000 need to be looked at more closely in order to achieve usable crystals.

Since the protein that has been used in this case has been reclaimed several times and, since freshly purified proteins seem to crystallise more efficiently than older ones, it seemed prudent to wait to carry out this next trial until after some more protein had been purified. This must wait until some more of the crude venom is available.

3.7.6 Initial diffraction

Since beam time at Grenoble was available, it was decided that an attempt should be made to obtain diffraction from some of the crystals that had developed during the refined trial. Four of the better crystals were chosen and an attempt to find a suitable cryoprotectant was undertaken. It was decided that the conditions in the

wells that crystals were found should be used as a basic starting point, with the addition of glycerol to aid in the protection against ice.

100mM Na citrate pH 5
22.5% PEG 4000
20% 2-propanol
10% glycerol

The crystals were soaked, mounted and then frozen before packaging in a dewier for transportation.

3.7.7 Diffraction data

3.7.7.1 Data produced at Grenoble

The crystals that were taken to Grenoble allowed diffraction data to be collected up to 2 angstroms in resolution. However, this data has certain problems associated with it, as it only appears to give information on two of the three dimensions. This was thought to be due to the fact that the crystals were growing in a manner that was too plate-like, thus only allowing diffraction through the two planes. However, further consideration of the problem, leads to the belief that in fact this is not the case and that the crystal may be mosaic in nature[94].

Under normal circumstances, when diffraction data is collected, there are three components to the diffraction obtained. There are diffraction spots in straight lines in a uniform lattice type that occur both vertically and horizontally and there are spots that occur in curved patterns around the outer regions of the diffraction pattern, as can be seen in *Figure 3.32*. It is these curved diffraction spots that allow the computer based analysis packages to determine the orientation of the crystal with respect to the X-ray beam. Without these curves, the determination of the third dimension is not possible. In the case of the data produced by the batroxobin crystals, these curved diffraction points do not actually exist and consequently programs the analysis are only able to calculate two of the dimensions, as can be seen from *Figure 3.33*.

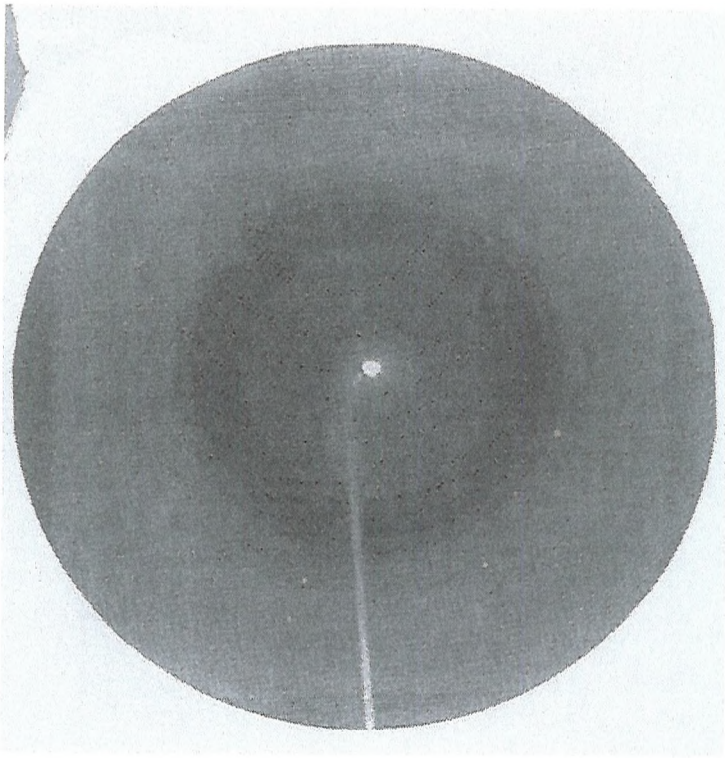


Figure 3.32: An example of normal diffraction data.

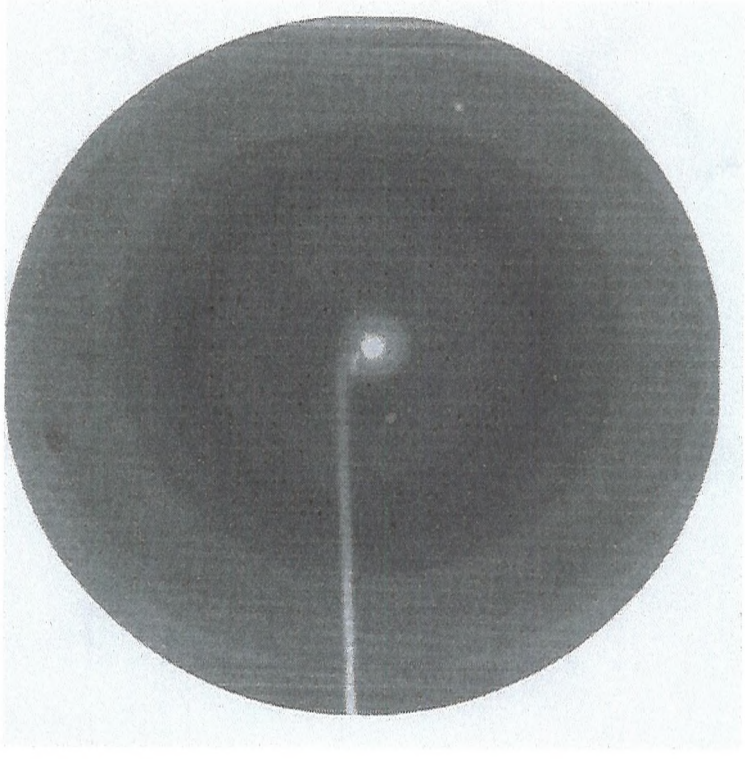


Figure 3.33: Diffraction data obtained for batroxobin, and example of mosaicity.

The major cause of these strange diffraction data obtained from the crystals is likely to be due to them being mosaic in nature. The ideal packing of the unit cells within the crystals is in a uniform manner *Figure 3.34*. The more the packing deviates from this ideal, and the less uniform it becomes, and it is said to possess mosaicity. *Figure 3.35*. This mosaicity can be brought about by external influences, for example the cryoprotectant, or it can be a natural phenomena associated with those particular crystals.

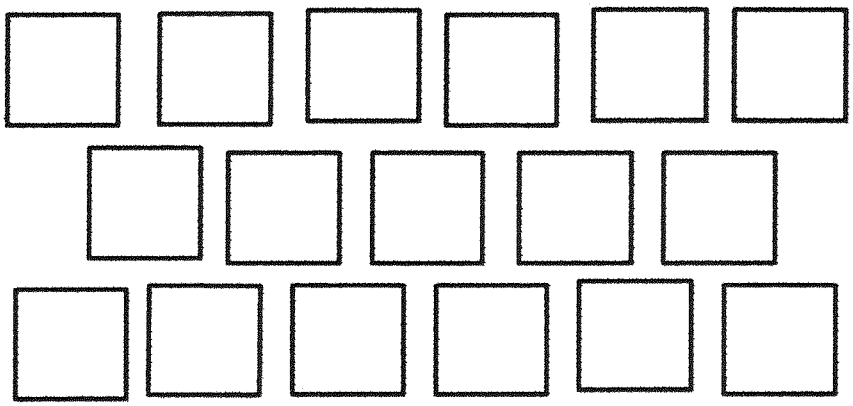


Figure 3.34: An example of ideal unit cell packing.

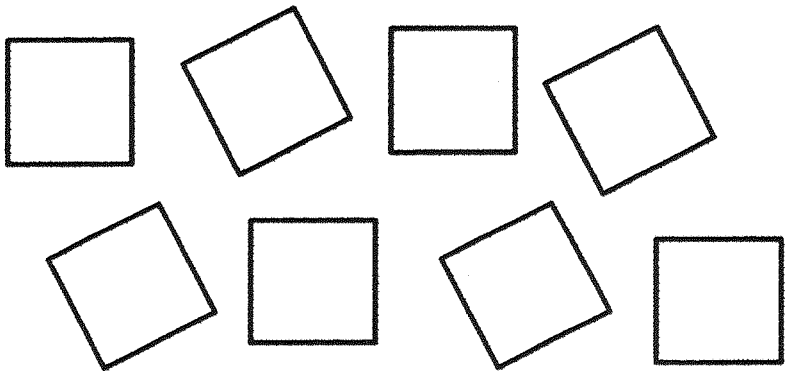


Figure 3.35: An example of irregular unit cell packing resulting in mosaicity.

3.7.8 Continuing crystallography work

It was decided that the problem of mosaicity might be as a result of there being partial binding of some benzamidine inhibitor within the active site of the batroxobin. This could help to explain why some crystal formation was occurring with batroxobin, which is notoriously soluble and difficult to crystallise. It is believed that some of the free benzamidine from the purification is still bound within the structure of some of the some of the molecules. Therefore some of the molecules are locked in differing conformations to the other, aiding the crystallisation process but causing the resultant mosaicity. Therefore, it was decided that the best way to overcome this problem was to add free benzamidine to all the crystallisation buffers and to the protein in order to attempt to lock all the molecules in the same conformation.

This was attempted at various inhibitor concentrations and in all cases the protein precipitated out immediately. It is believed that this is due to the fact that the increased inhibitor has further stabilised the protein, resulting in the 25mg/ml concentration being far too high. Attempts are ongoing to find a concentration of inhibitor and protein that with facilitate good crystal growth.

CHAPTER FOUR

SYNTHESIS AND REACTIONS OF PEPTIDES AND INHIBITORS

4.1 Introduction

Solid-phase peptide synthesis was first devised by Bruce Merrifield in 1963, at Rockerfeller University [95]. This synthesis was pioneering as it allowed the reaction to take place in solution whilst still attached to an insoluble support medium. This medium takes the form of a resin that holds the C-terminal amino acid by the carboxyl group, thus resulting in the formation of the peptide from C-N rather than N-C as found in nature. The resin is a polymeric structure, usually polystyrene, containing approximately 1% *p*-(chloromethyl)styrene units. The amino group of the first amino acid, and then subsequent residues, is initially blocked before removal to allow reaction to the next residue.

Two possible types of blocking agents can be used. The original agent is the *t*-butoxycarbonyl (tboc) protecting group, which is acid labile, requiring treatment with anhydrous acid to remove it from the amino group of the residue, and HF for the final cleavage of the peptide from the resin. The other approach, devised in 1983 by Sheppard uses 9-fluorenylmethoxycarbonyl (Fmoc) as the protecting group [82], *figure 4.1a.* This is base labile, and requires the addition of pyridine to remove the protection from the amino group of the residue, with the final cleavage from the resin carried out in TFA.

The advantages to solid phase peptide synthesis are numerous. In particular it is a highly efficient process allowing a much easier and quicker separation of the peptides from the reaction mixture than could be hoped for in solution phase chemistry and without the losses of material which are also common to reactions in solution. This high efficiency arises from the fact that the peptides remain bound to the resin during the course of the reaction means that the soluble reaction components can be used in large excess as the resin can easily be removed by filtration.

Solid phase peptide synthesis makes use of the sequential addition of amino acids. These residues have their α-amino side chains protected by either Fmoc or tboc to a solid, insoluble support. This protection is removed by a deprotection step, before the addition of the next protected amino acid residue by the use of either a coupling agent or a pre-activated residue. A linker is added to the synthesis in both cases, between the resin and the first amino acid residue, in order to bring about an easier cleavage of the peptide from the resin. With the linker present, the cleavage

can be more specific, allowing the chemist to choose the type of C-terminus required in the peptide. In this case, it is possible to produce not only free acids, but also amides, hydrazides and side chain protected peptide fragments. It is often convenient to choose side chain protection that can be simultaneously cleaved with the final cleavage of the peptide from the resin, *figure 4.2.*

4.2 Peptide Design

The peptides were defined to allow the specificity of batroxobin to be probed. It was therefore important that the literature of similar experiments on related enzymes should be considered. It became apparent that peptide studies on thrombin and kallikrein indicated that only the three residues either side of the scissile bond showed any significant impact on the rate of cleavage [96,97]. Therefore, it was decided that for these reactions only this type of hexapeptide should be considered.

It was then vital that a complete search of the sequences of fibrinopeptides A and B should be considered in an attempt to identify areas that may explain why thrombin cleaves at both sites and batroxobin at only one.

Sequence surrounding the scissile bond in fibrinopeptide A

HUMAN	-GGGVRGPRVV-
BOVINE	-GGGVRGPRLV-
CANINE	-GGGVRGPRIV-
RAT	-GGDIRGPRIV-

Sequence surrounding the scissile bond in fibrinopeptide B

HUMAN	-FFSARGHRPL-
BOVINE	-GLGARGHRPY-
CANINE	-TVDARGHRPL-
RAT	-LSIARGHRPV-

As can be seen, there is significant sequence homology between all the sequences in fibrinopeptide A, but much less homology between the sequences in fibrinopeptide B. It was therefore decided that the peptides produced should be based upon the sequences shown for the human fibrinogen, as this is the most significant in terms of the role of batroxobin as a possible therapeutic agent.

Peptides produced based on human fibrinopeptides A and B.

GVRGPR-COOH*

GVRGHR-COOH

SVRGHR-COOH

GARGHR-COOH

SARGHR-COOH#

SARGPR-COOH

GVRGPR-NH_2*

GVRGHR-NH_2

SVRGHR-NH_2

GARGHR-NH_2

SARGHR-NH_2#

SARGPR-NH_2

GGGVRGPRVV-COOH*

FFSARGHRIV-COOH#

GGGVRGPRVV-NH_2*

FFSARGHRIV-NH_2#

NB. * sequence as found in fibrinopeptide A. # sequence as found in fibrinopeptide B

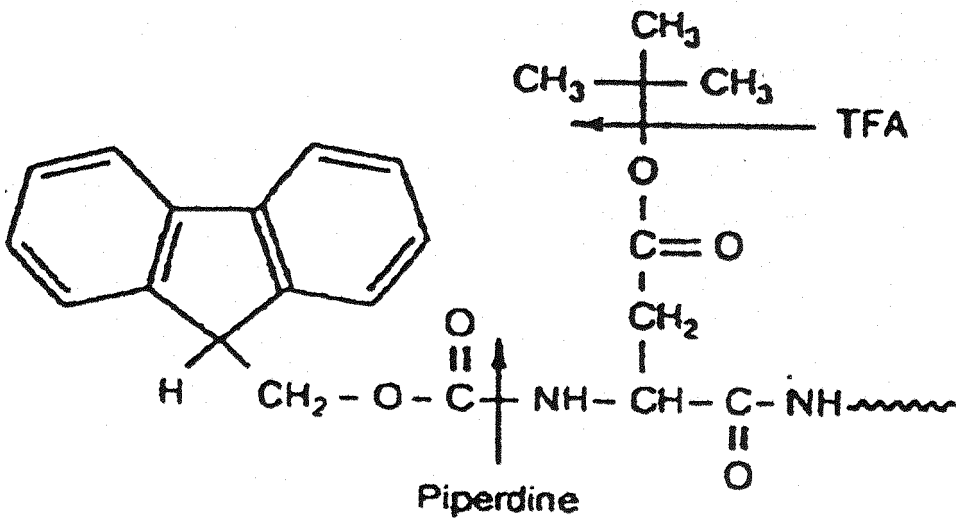


Figure 4.1a: The Fmoc protecting group.

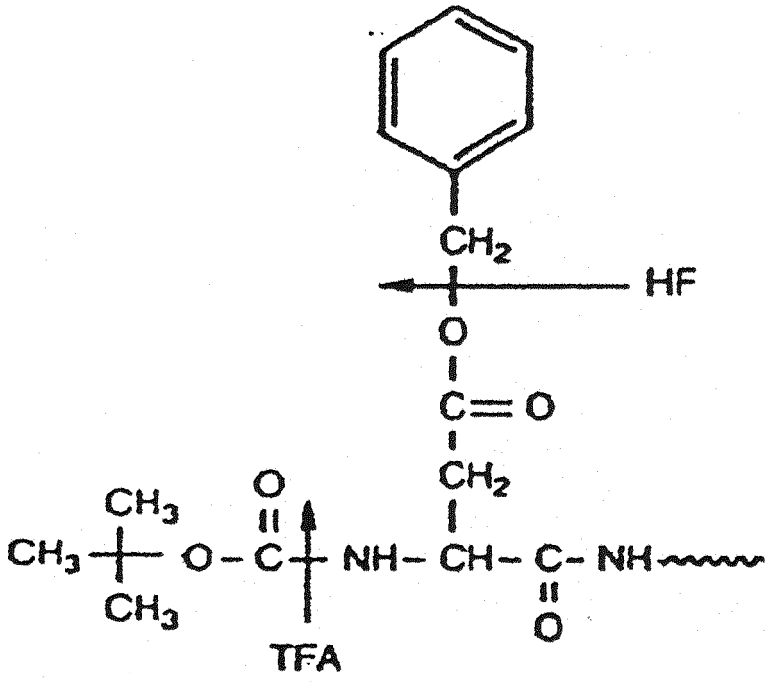


Figure 4.1b: The Boc protecting group.

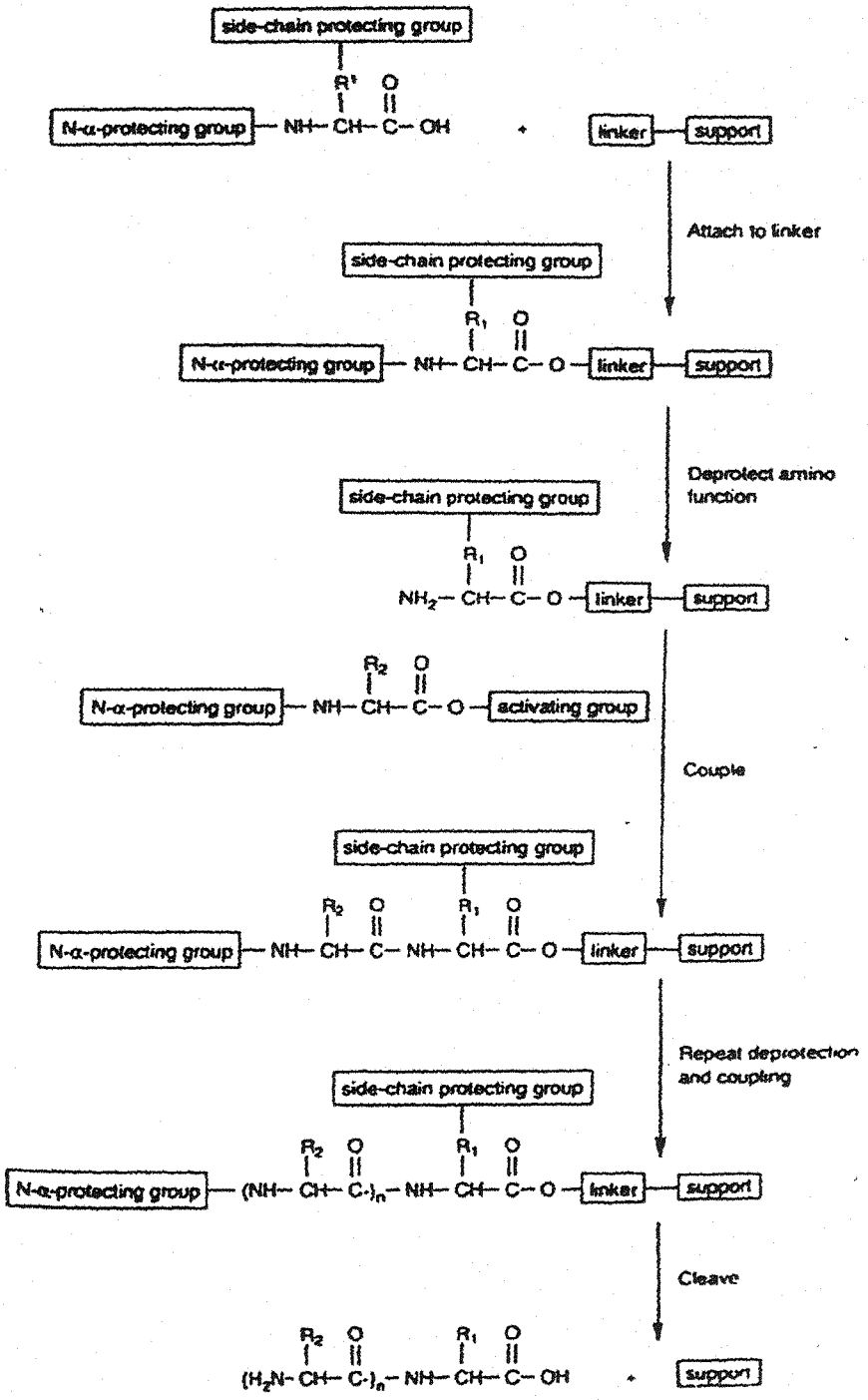


Figure 4.2: General scheme for solid phase peptide chemistry.

Therefore a positive and negative control were produced with the exact sequences of the two fibrinopeptides and the remainder of the peptides showed substitutions of one or more residue into each of the sequences to look at which of the sites were responsible for the differences in reactivity. One amide was produced of the fibrinopeptide A sequence to ensure that the highly negative charge of the free acid was not affecting the reactivity of the batroxobin.

4.3 Fmoc peptide synthesis

4.3.1 Materials and methods

In the synthesis carried out, a solid phase peptide synthesis was undertaken using Fmoc protection, with both the formation of the free acids and the amides. This was achieved manually, using a Griffin flask shaker. All the solvents and chemicals used were of peptide synthesis grade. Arginine-Wang resin, 0.9mmole/g was used for the synthesis of all the peptide free acids, with the exception of one which required the synthesis of the R-resin from Wang resin, 1mmole/g. *p*-Methylbenzhydrylamine (MBHA) resin, 0.9 mmole/g, 100-200 mesh (Calbiochem-Novabiochem Ltd., UK) was used for the synthesis of the peptide amides. The side chain protections that were used were for Fmoc chemistry. All the filtrations were carried out using positive nitrogen pressure.

4.3.2 Reaction vessel preparation

Dichloromethylsilane (15%) in toluene (20ml) was added to a dry sintered reaction vessel and allowed to stand for 30 mins. The solution was then removed, and the flask was washed with acetone before allowing to dry.

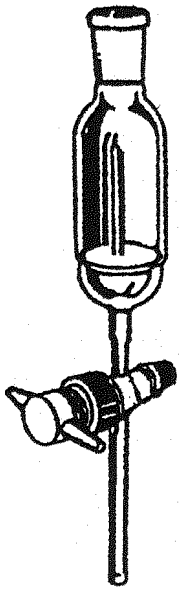


Figure 4.3: Reaction vessel

4.3.3 General reaction protocol for fmoc synthesis

Peptides were synthesised in a C to N direction on either Wang or MBHA resin following the procedure outlined in *figure 4.2*.

4.3.3.1 The coupling protocol.

The Wang or MBHA resin (200mg, sub. 0.9mmole/g, 0.2 mmole scale) was added to the reaction vessel with 10ml of DMF and allowed to swell. BOP: HOBt: α-N-Fmoc protected amino acids: DIPEA in a 1:1:1:3 ratio with the resin, with the addition of 266mg of BOP, 82mg of HOBt, 200µl of DIPEA and 0.6mmole x M.W. of Fmoc-amino acid were added to each coupling. Each Fmoc-amino acid coupling cycle involved the following steps:

1. DMF wash (3 x 10ml x 1 min)
2. 20% (V/V) piperidine/DMF deprotection (1 x 10ml x 5 mins, 1 x 10ml x 15 mins).
3. DMF wash (5 x 10ml x 1 min).

4. Ninhydrin assay to confirm deprotection.
5. Coupling with BOP, HOBt, DIPEA, and α-N-Fmoc protected amino acid (in DMF for one hour).
6. DMF wash (4 x 10ml x 1 mins).
7. Ninhydrin assay to monitor the coupling efficiency. A positive result for the addition of the amino acid confirmed that deprotection could take place. A negative result led to a repeat of the coupling step.

This cycle was repeated for each amino acid required for the production of the peptide.

4.3.3.2 Capping

Acetylation, or capping of free amines, was performed after any incomplete coupling. Incomplete coupling was monitored using the ninhydrin assay, as outlined in section 6.3.4. Capping was achieved by incubation of the resin with 10ml of 50% v/v of acetic anhydride/pyridine for 25mins and then washing the resin with DMF and DCM. A successful capping reaction was identified as a negative ninhydrin assay.

4.3.4 The ninhydrin assay

Wet peptidyl-resin (2-5mg) was mixed with 100μl of solution A and 25μl of solution B and heated at 100°C for 5-10 min. The colour development was monitored qualitatively by eye or quantitatively as described later.

Solution A - Phenol (40g) was added to 10ml of ethanol and warmed until completely dissolved. 65mg of potassium cyanide was dissolved in 100ml of distilled water. 2ml of this solution were added to 100ml of pyridine and 4g of Amberlite MB-3 resin was added and stirred for 1hr. The solution was then filtered to remove the resin and collected.

Solution B - Ninhydrin (2.5 g) was dissolved in 50 ml of ethanol.

4.3.4.1 The quantitative ninhydrin assay

The assay was carried out using a known weight (2-5mg) of dry peptide-resin.

1. The sample was placed in a small Pyrex tube.
2. Solution A (100μl) was added to the tube and to an empty tube to act as a blank.
3. Solution B (25μl) of solution B was added to each tube, and heated at 100°C for 10 min.
4. After cooling, 2ml of 60% (v/v) aqueous ethanol were added to each tube, and the adsorption, A_{570nm}, was monitored.
5. If resulting colour was too intense, any necessary dilutions were made in 60% ethanol.

The following formula was used to calculate the amount of free amine present following the quantitative assay:

$$\mu mole / g = (A_{570nm} \times Vol \times 10^6) / (E \times Wt)$$

Where:

μmole / g = μmole of free amine per gram of resin

A_{570nm} + optical density at 570 nm corrected for dilution

Vol = volume of the assay reaction (2.125ml)

E = the average extinction coefficient ($1.5 \times 10^4 M^{-1}cm^{-1}$)

Wt = accurate mass (mg) of the peptide-resin being assayed

Quantitative assay example

The procedure as described above was followed, with 2.8mg of dry resin used. This was diluted 17 times and the A_{570nm} was measured.

$A_{570} = 0.9981$

A_{570nm} corrected for dilution = $0.9981 \times 17 = 16.9677$

μmole of free amine / g = $(16.9677 \times 2.125 \times 10^6) / (1.5 \times 10^4 \times 2.8) = 896.4$

= 0.8964 mmole of free amine / g

Since the substitution of the MBHA resin is known to be 0.9mmole / g, the coupling efficiency of the first reaction = 0.8964 / 0.9 = 99.6%

4.4 Cleavage of the peptide from the resin

4.4.1 TFA cleavage from the Wang resin.

In the synthesis of the free acids, a standard TFA cleavage for Fmoc chemistry was carried out. This was achieved by the addition of 95% TFA to the resin that had been dried by nitrogen gas following the final washes. In addition 100µl of anisole was added as a scavenger to help clean up any tertiary butyl protecting groups that might have been removed on final cleavage.

The mixture was then shaken for ½ hr to allow the cleavage from the resin to occur. The solution was filtered from the resin into a clean round bottomed flask by nitrogen pressure and the resin was re-washed in a small volume of TFA, before this was also collected. Excess TFA was removed by rotary evaporation of the mixture until an oily substance was found collecting around the side of the flask. This oil was washed extensively with diethyl ether, and a white solid was produced. This solid peptide was then ready for analysis and verification.

4.4.2 HF Cleavage from the MBHA resin

During the production of all the amide peptides produced in this work, a standard HF cleavage [82] was used to cleave the peptide from the resin. In addition this cleavage resulted in the removal of all the fmoc side chain protection. Dr. R. P Sharma or Dr. R. Broadbridge kindly performed this.

Peptidyl resin (350mg) was placed in a Teflon reaction vessel with 200mg of *p*-cresol. The reaction vessel was attached to the HF line on the HF apparatus (Peptide Institute, Osaka, Japan), and cooled under liquid nitrogen. Evacuation of the reaction vessel was then achieved by the use of a Teflon coated pump. Liquid HF (10ml) was distilled into the reaction vessel, and the mixture was stirred at 0°C in an ice bath for one hour.

After the reaction had finished, the HF was evaporated from the reaction vessel with a stream of nitrogen. The cleaved product was washed with diethyl ether, and filtered through a sintered glass funnel. The resulting powder was dissolved into 10ml of TFA, and re-filtered. The excess solvent was removed by rotary evaporation, and the crude peptide was obtained by a final precipitation using diethyl ether.

4.5 Purification and characterisation of peptides

4.5.1 Gel filtration

Gel filtration was used in order to desalt the peptide samples prior to purification. This was carried out on a PD10 column (Sephadex® G-25M, Pharmacia Biotech). The crude peptide, approximately 60mg, was dissolved in 2ml of 0.1% TFA and loaded onto the column. The fractions were collected from the column in a total final volume of 3ml. These were then screened for peptide content using the ninhydrin assay as previously described. The pooled fractions containing the peptide were then freeze dried, resulting in the formation of a white powder.

4.5.2 Analytical HPLC

The purity of the peptides produced was examined by the use of HPLC. A Gilson 715 HPLC system with two slave 306 pumps and a reverse-phase analytical Vydac C-18 column was used. The absorbance at 216nm wavelength was monitored, and the following conditions were used:

Solution A: 0.1% (V/V) TFA in Analar H_2O
Solution B: 0.1% (V/V) TFA in acetonitrile

Gradient: Solution B, 0% over 2 min
0-80% over 20min
80% over 1 min
80-0% over 8 min
0% over 1 min

At a flow rate of 1ml/min

This approach allowed the identification of the desired peptide prior to purification using a preparative HPLC technique (as described below). The purified peptides were then rechromatographed on the same C-18 column, under the same conditions to assess the success of the purification procedure.

4.5.3 Preparative HPLC

The peptides were purified using preparative HPLC. The same system was used as described previously for the analytical HPLC, with the exception that a Vydac C-18 reverse phase preparative column was used. The purification was carried out using the conditions outlined below:

Solution A: 0.1% TFA in Analar H_2O
Solution B: 0.1% TFA in acetonitrile

Gradient: Solution B

0% over 2 min
0-80% over 32 min
80% over 1 min
80-0% over 8 min
0% over 1min

The peptide peak was collected and its identity was confirmed using electrospray MS.

4.5.4 Electrospray mass spectrometry

The characterisation of the peptides produced was carried out using electrospray mass spectrometry of the peptide in a 50:50 solution of acetonitrile/methanol. There was no necessity for protonation using formic acid, as the presence of the TFA from purification was sufficient. The identity of the peptide could be accurately confirmed since the instrument was able to give accurate mass measurements to within 1 atomic mass unit.

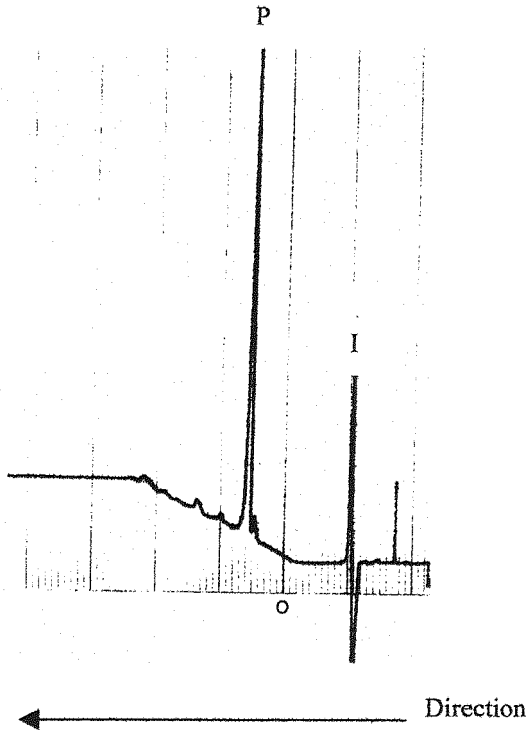


Figure 4.4: HPLC trace for a purified peptide GVRGPR-COOH, representative of all other peptide purifications. The injection artefact is represented as I and the peptide as P.

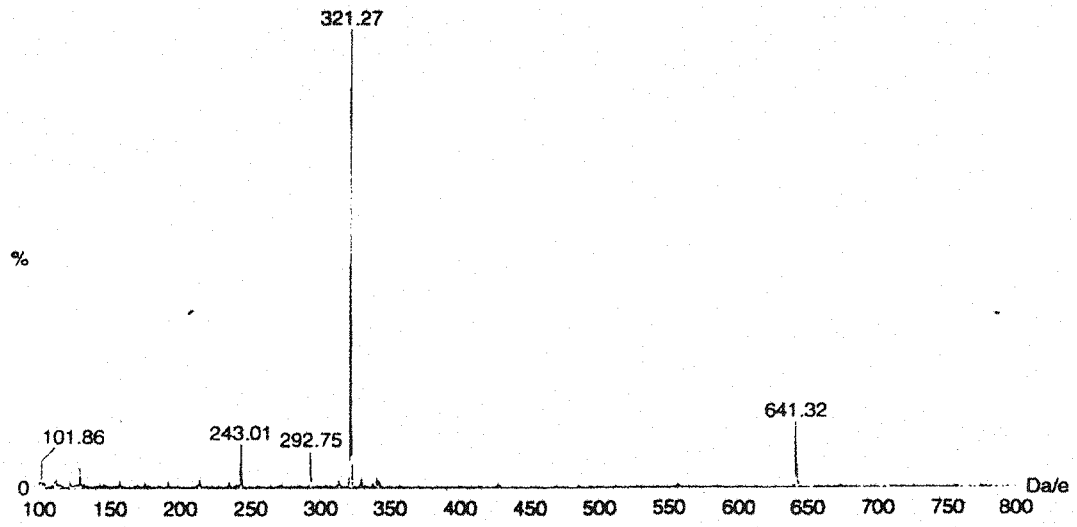


Figure 4.5: The electrospray MS of the purified peptide GVRGPR-COOH. This is representative of all the others purified. The predicted mass is 641Da Both the singly charged species at 641.32Da and the doubly charged at 321.27Da can be seen.

4.6 Reactions of thrombin and batroxobin with synthetic peptides.

A series of experiments were performed to look at the reactions of thrombin and batroxobin, (*B. moojeni* and *B. atrox*), with each of the synthetic peptides. This was in order to ascertain the significance of each of the residues surrounding the scissile bond in both fibrinopeptide A and B in the relative specificity of each of the enzymes. Initially this was carried out over a 24-hour period in order to rule out immediately any peptides where there was no cleavage. The peptides that showed some reactivity after 24 hrs were then considered in more detail for specificity studies with each enzyme.

4.7 Initial experimental screen.

4.7.1 Experimental procedure.

Solutions (0.02 M) of each peptide were made up in analytical water, and $1x10^{-5}$ M solutions of thrombin or batroxobin were made up in 50mM TRIS.HCl buffer pH7.5. Enzyme solution (50µl) was added to 200µl of the peptide solution and incubated at 37°C. Aliquots (40µl) were taken from the mixture after 0min, 30min, 1hr, 2hrs, and 4hrs, and the reaction was stopped with 40µl of a 50:50 solution of acetonitrile/TFA.

A 1/100 dilution of each of the reaction mixtures was loaded onto a C18 reverse phase analytical HPLC column, and eluted with a 0-20% linear gradient of solution B, where solution A is 0.1% TFA in H_20 and solution B is 100% acetonitrile. Elution of the peptides alone resulted in the production of the hexapeptide peaks at ~15% acetonitrile, with any tripeptide cleavage products expected at some point before.

4.7.2 Results of initial screen.

Initially, all the hexapeptide free acids were incubated for 24 hours in the presence of both thrombin and batroxobin (*B. moojeni*). This initial set of experiments resulted in no cleavage for any of the free acids with any of the enzymes.

It is therefore evident that all these enzymes are incapable of tolerating a highly electronegative group at their C-terminal end in the proximity of the active site. The hexapeptide based on the structure of fibrinopeptide A, which was synthesised with an amide group attached to the C-terminus, was found to cleave after 24 hours with both batroxobin and thrombin.

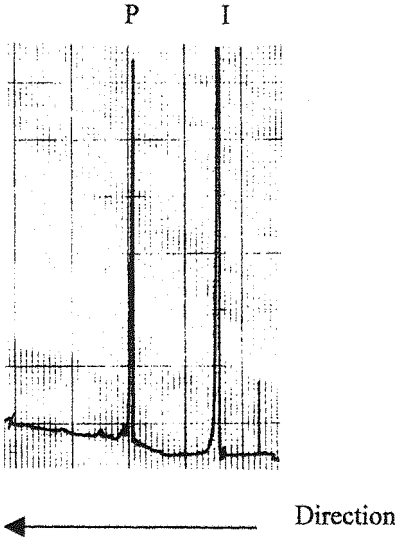


Figure 4.6: A representative HPLC trace of a hexapeptide free acid (GVRGPR-COOH) after a 0mins incubation with either thrombin or batroxobin.

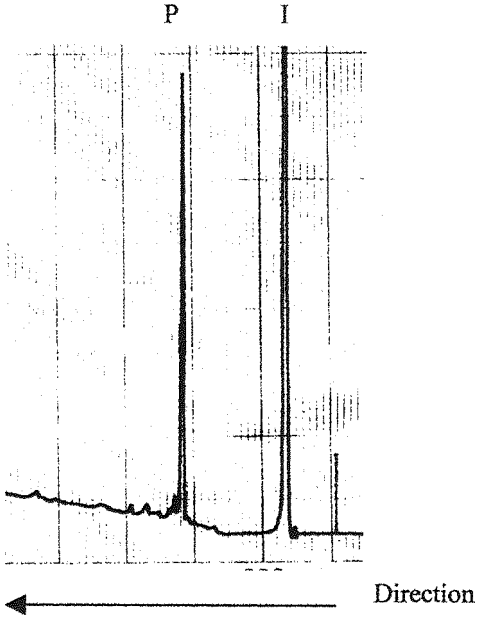


Figure 4.7: A representative HPLC trace of the same hexapeptide free acid after a 24-hour incubation with thrombin or batroxobin. The injection artefact is represented by I and the peptide by P.

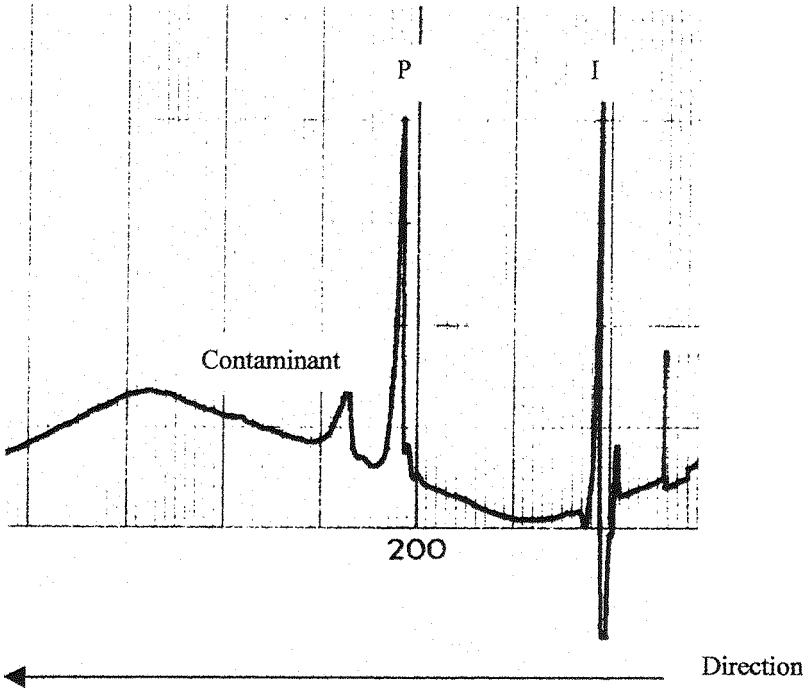


Figure 4.8: A representative HPLC trace of the hexapeptide amide (GVRGPR-amide) after a 0 min incubation with thrombin or batroxobin.

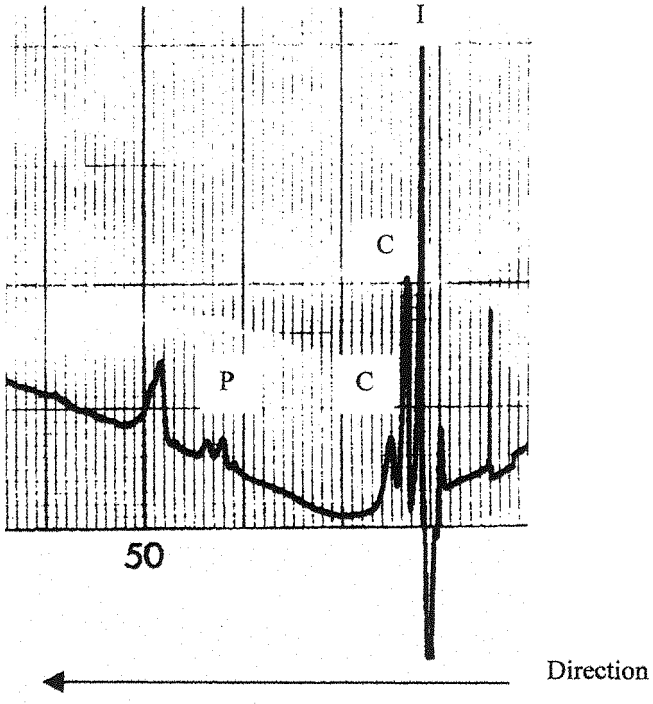


Figure 4.9: A representative HPLC trace of the hexapeptide amide (GVRGPR-amide) after 24 hours incubation with either thrombin or batroxobin. I represents the injection artefact, P the peptide and C the cleavage products.

4.7.3 Production and reactions of decapeptide free acids.

It was decided that since none of the hexapeptide free acids were cleaved by either of the enzymes that two decapeptide free acids should be produced in order to assess whether this was due to the proximity of the carbonyl group to the active site. Therefore the peptides based upon the sequences of fibrinopeptides A and B were produced: GGGVRGPRVV-COOH and FFSARGHRPL-COOH. These peptides were once again incubated for 24hrs with thrombin and batroxobin.

Once again, there was no cleavage of either of these peptides with either of the enzymes. This indicates very strongly that this group of enzymes simply cannot tolerate a highly negative charge close to its active site.

4.7.4 Production and reactions of amide peptides.

Since the only peptide that had thus far cleaved was the one with the C-terminal amide group attached, it was decided that a whole series of these should be produced as outlined previously. These peptides were once again incubated with thrombin and batroxobin for 24 hours to see if any reaction occurred. Any peptides that showed no cleavage after 24 hours were assumed not to be substrates for the particular enzyme.

SUBSTRATE	THROMBIN	BATROXOBIN
GVRGPR-NH_2	✓	✓
GVRGHR-NH_2	✓	✗
SVRGHR-NH_2	✓	✗
GARGHR-NH_2	✓	✗
SARGHR-NH_2	✓	✗
SARGPR-NH_2	✓	✗
GGGVRGPRVV-NH_2	✓	✓
FFSARGHRPL-NH_2	✓	✗

Table 4.1: The reactions of thrombin and batroxobin with various amide peptides after 24 hours incubation.

As can be seen from *Table 4.1*, thrombin is capable of cleaving all of the amide peptides produced, including those derived from fibrinopeptide B. However, batroxobin was incapable of cleaving any of the peptides that did not correlate directly with the sequence found in fibrinopeptide A. It was therefore decided that another hexapeptide should be produced which correlated to the sequence found around the scissile bond of fibrinopeptide A found in rat fibrinogen (GGDVRGPRVV-amide). This was in order to assess the significance of this sequence with particular reference to the aspartic acid residue, as previous studies carried out on the natural substrate fibrinogen suggested that rat fibrinogen was the best substrate for batroxobin (see Chapter3).

This was again set up as a 24-hour incubation following the protocol previously outlined with both thrombin and batroxobin. This preliminary study again showed that both enzymes cleaved this sequence.

4.8 Kinetic experimental analysis

The results obtained from the initial screens were used as the starting point for a more in depth analysis. Those combinations of peptides and enzymes, which had shown activity, were considered in a complete kinetic study were samples were taken every 5 minutes over the course of 30 minutes. This analysis also took into consideration the differences between batroxobin derived from both sub-species being studied *B. moojeni* and *B. atrox* as well as assessing the possible role of Ca^{2+} in these reactions. This part of the work was undertaken to confirm that the requirement for Ca^{2+} ions as shown in Chapter 3 was due to and interaction with the fibrin monomers and not with the enzyme. If the enzyme required the Ca2+ then its omission in these experiments would result in a decrease in the rate of the cleavage of the peptides.

4.8.1. Experimental procedure.

These experiments were performed as detailed in Chapter 2. It was necessary to refine the conditions gradually in an attempt to produce cleavage of the substrate within a 30 minute period. This resulted in differing enzyme concentrations, see *Table 4.2.*

4.8.2 Results of kinetic studies.

4.8.2.1 The role of Ca^{2+} in thrombin-like enzyme activity.

The reactions of GVRGPR-amide with thrombin and batroxobin (*B. moojeni*) were carried out both with and without Ca^{2+} present. No difference was found between the course of the reaction in the presence of Ca^{2+} as compared to the reactions in its absence, as seen in *Table 4.2*. This supports the established view that Ca^{2+} is only significant in facilitating stacking of the fibrin monomers produced as a result of the enzyme activity on fibrinogen, rather than being required for the enzyme action itself.

4.8.2.2 Comparison of the reactivity of batroxobin (*B. atrox*) and batroxobin (*B. moojeni*) with artificial peptide substrates.

This set of reactions was carried out in order to compare the actions of the two forms of batroxobin. This was in an attempt to explain the differences in their specificity to fibrinogens derived from different species as shown in Chapter 2.

As can be seen from *Table 4.2* there appears to be no difference in the action of either form of batroxobin to the artificial peptide. This suggests that the differences shown in Chapter 3, with the increased rate of cleavage of natural fibrinogen substrates by batroxobin (*B. atrox*), are due to external binding and recognition sites and not due to the sequences around the scissile bond.

4.8.2.3 Reactions of peptides with each of the enzymes.

The kinetic data obtained from this study were used to calculate K_M, V_{max} and k_{cat} for each of the synthetic peptides with each enzyme. These values were subsequently used to calculate the K_{cat}/K_M for each of the substrates, thus providing an indication of the specificity of the enzyme for each substrate. *Figure 4.12* shows a representative HPLC trace for a peptide that has been cleaved. The information presented in *Figure 4.10, 4.11* and *Table 4.2* is the kinetic data obtained from the analysis of all the reactions of thrombin and batroxobin with the synthetic peptides.

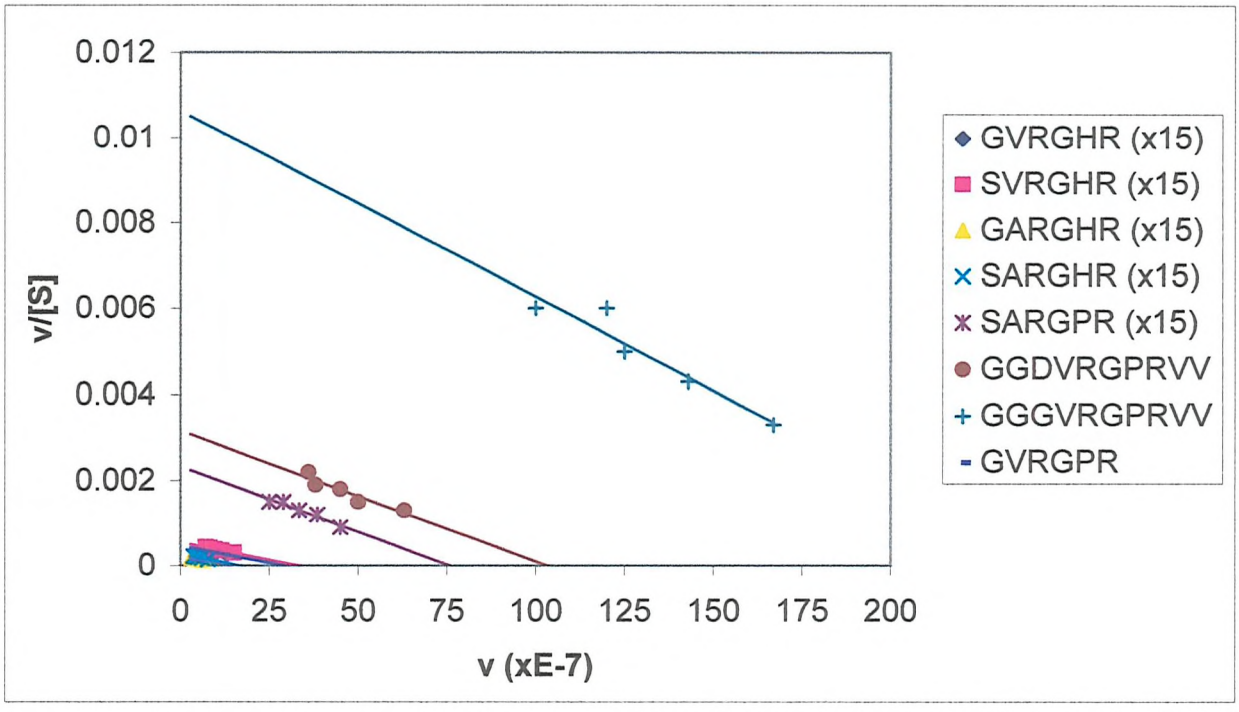


Figure 4.10: Eadie-Hoffstee plot of the reactions of synthetic peptides with thrombin. Each peptide is represented as a colour, as indicated in the legend..

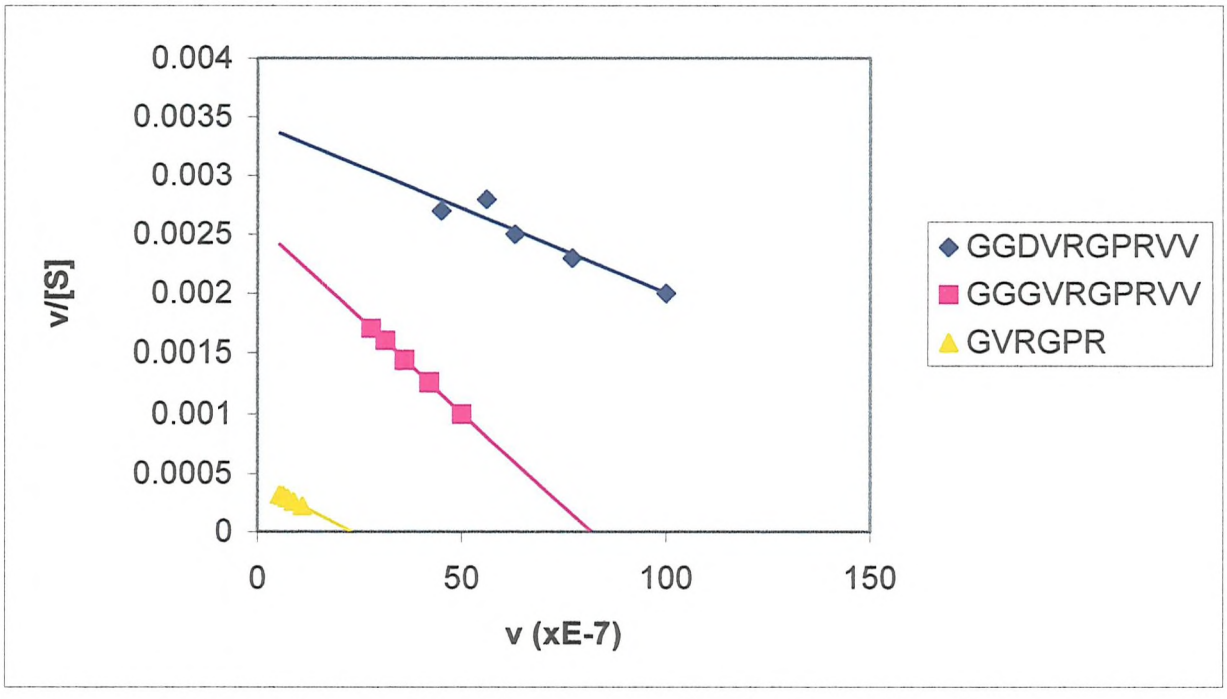
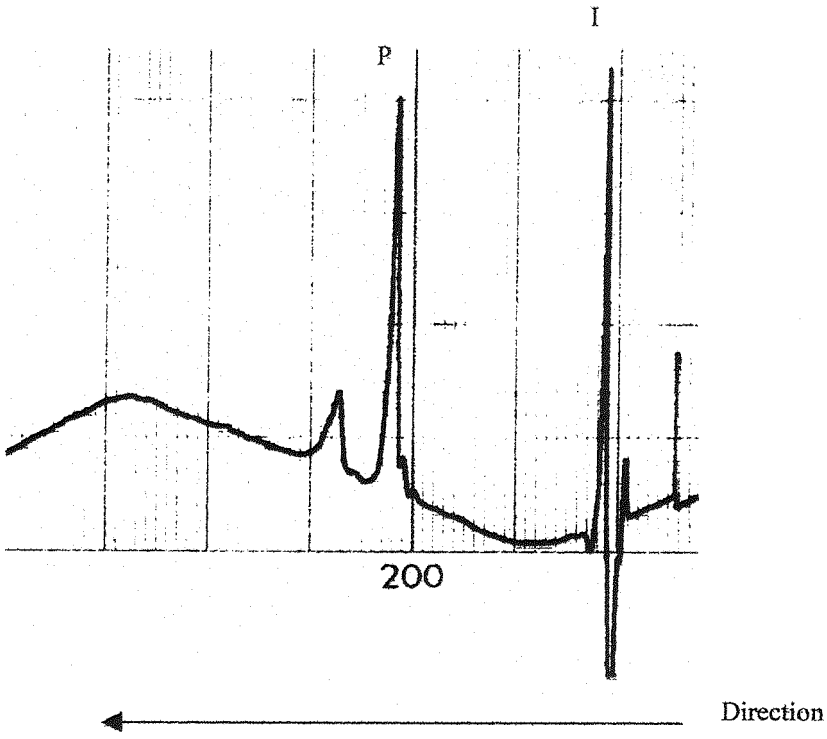


Figure 4.11: Eadie-Hoffstee plot of the reactions of synthetic peptides with batroxobin. Each peptide is represented as a colour, as indicated in the legend..

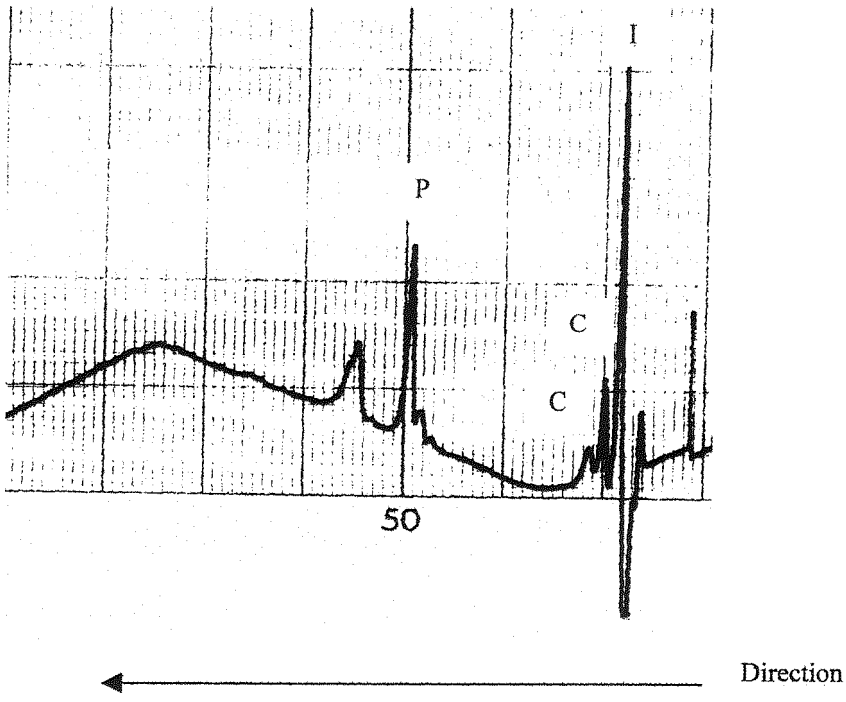
Enzyme	Peptide	[E] (M)	K_M	V_{max} (s^{-1})	k_{cat}/K_M
Batx(*atrox*)	GVRGPR	$1x10^{-6}$	$5.94x10^{-3}$	$2.43x10^{-6}$	$0.4x10^{3}$
Batx(*mooj*)	GVRGPR	$1x10^{-6}$	$5.88x10^{-3}$	$2.51x10^{-6}$	$0.42x10^{3}$
Batx-Ca^{2+}	GVRGPR	$1x10^{-6}$	$5.96x10^{-3}$	$2.53x10^{-6}$	$0.42x10^{3}$
Thrombin	GVRGPR	$1x10^{-6}$	$7.5x10^{-3}$	$1.6x10^{-6}$	$0.21x10^{3}$
Batx(*mooj*)	GGGVRGPRVV	$1x10^{-6}$	$3.6x10^{-3}$	$8.6x10^{-6}$	$2.4x10^{3}$
Thrombin	GGGVRGPRVV	$1x10^{-6}$	$2.2x10^{-3}$	$2.4x10^{-5}$	$10.9x10^{3}$
Batx(*mooj*)	GGDVRGPRVV	$1x10^{-6}$	$2.5x10^{-3}$	$2.5x10^{-5}$	$10.0x10^{3}$
Thrombin	GGDVRGPRVV	$1x10^{-6}$	$3.1x10^{-3}$	$1x10^{-5}$	$3.2x10^{3}$
Thrombin	GVRGHR	$1.5x10^{-5}$	$5.72x10^{-3}$	$2.71x10^{-6}$	$0.03x10^{3}$
Thrombin	SVRGHR	$1.5x10^{-5}$	$5.34x10^{-3}$	$2.9x10^{-6}$	$0.04x10^{3}$
Thrombin	GARGHR	$1.5x10^{-5}$	$6.49x10^{-3}$	$1.4x10^{-6}$	$0.01x10^{3}$
Thrombin	SARGHR	$1.5x10^{-5}$	$6.78x10^{-3}$	$1.68x10^{-6}$	$0.02x10^{3}$
Thrombin	SARGPR	$1.5x10^{-5}$	$3.54x10^{-3}$	$8.32x10^{-6}$	$0.1x10^{3}$
Thrombin	FFSARGHRPLV	$1.5x10^{-5}$	$3.2x10^{-3}$	$2.92x10^{-5}$	$0.6x10^{3}$

Table 4.2: The kinetic data for all of the synthetic peptides with each enzyme capable of cleaving them.

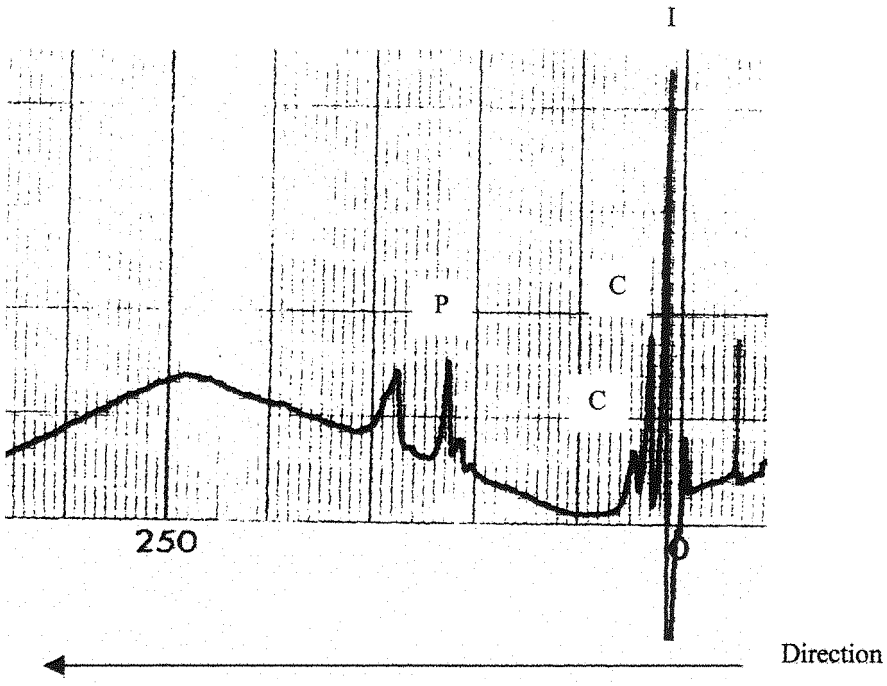
As can be seen from the table of kinetic data, for both thrombin and batroxobin, the sequences surrounding the scissile bond of fibrinopeptide A are much better substrates than any of the others. In the case of batroxobin, no cleavage is seen for any of the other sequences. On the other hand, thrombin which is capable of cleaving all the peptides produced, with the exception of the free acids, is seen to be much more reactive with the fibrinopeptide A sequence than any of the others.
It is also apparent from the data table that both enzymes are more reactive with the decapeptide sequences than they are with the hexapeptides, which is not surprising as these sequences provide a larger number of recognition sites. The data also seems to indicate that batroxobin has a higher specificity for the sequences found in rat fibrinopeptide A than in human, which correlates to the evidence gathered in



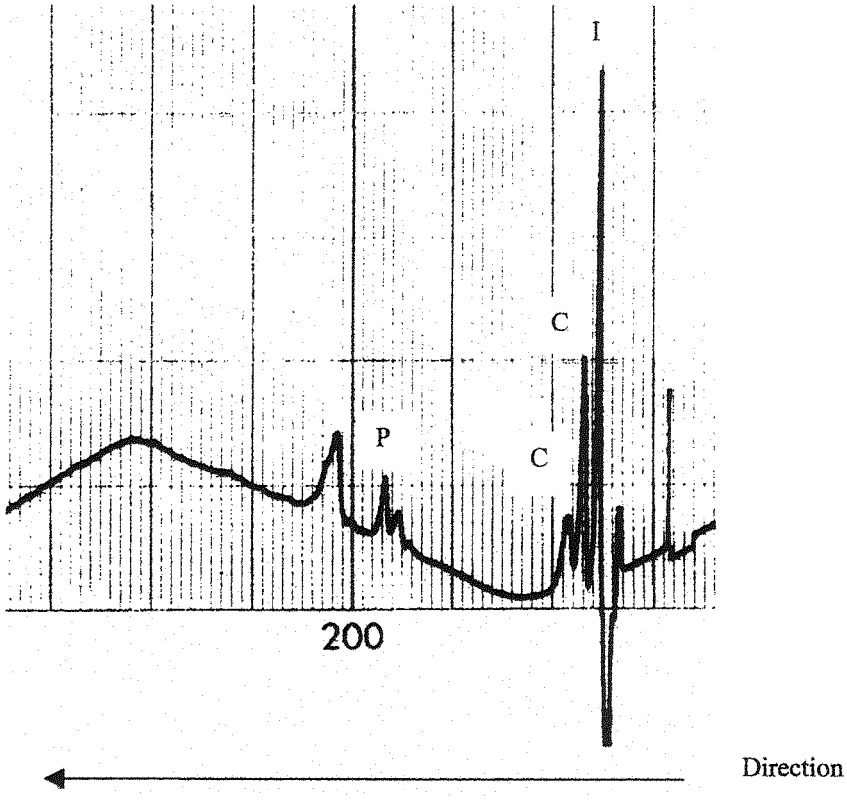
A. HPLC trace of GVRGPR-amide incubated with batroxobin taken at t=0mins. I is the injection artefact, P is the peptide and C are the cleavage products.



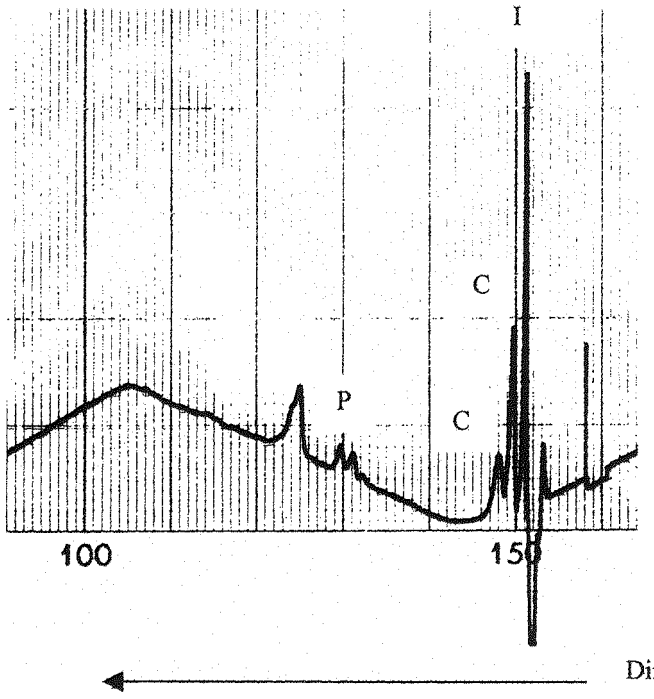
B. HPLC trace of GVRGPR-amide incubated with batroxobin taken at t=5mins. I is the injection artefact, P is the peptide and C are the cleavage products



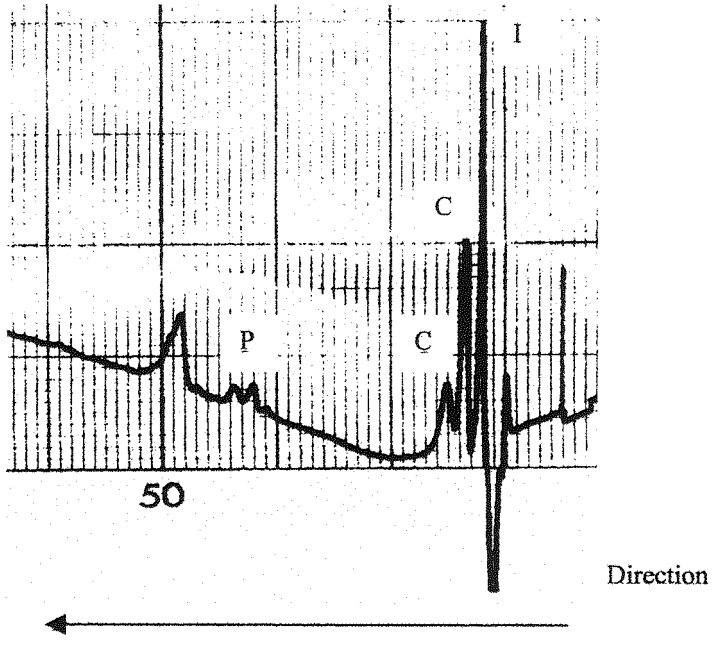
C. HPLC trace of GVRGPR-amide incubated with batroxobin taken at t=10mins. I is the injection artefact, P is the peptide and C are the cleavage products.



D. HPLC trace of GVRGPR-amide incubated with batroxobin taken at t=15mins. I is the injection artefact, P is the peptide and C are the cleavage products.



E. HPLC trace of GVRGPR-amide incubated with batroxobin taken at t=20mins.



F. HPLC trace of GVRGPR-amide incubated with batroxobin taken at t=25mins.

Figure 4.12: Representative HPLC traces of peptides cleaved by thrombin or batroxobin, with samples taken at 5 minute intervals for 30 minutes. I is the injection artifact, P is the peptide and C are the cleavage products.

Chapter 2. Thrombin on the other hand shows its greatest specificity to the sequence found in human fibrinopeptide A. Of the additional sequences, it appears that the proline residue found in the fibrinopeptide A sequence is important for recognition by thrombin since when this is introduced into the sequence found in fibrinopeptide B (SARGHR compared to SARGPR) there is an increase in the rate of cleavage.

4.9 Inhibitor studies.

Since the best substrate for the two enzymes was found to be the sequence from fibrinopeptide A, it was decided that this sequence should be used in the synthesis of an inhibitor. It was hoped that this inhibitor could serve two purposes: Firstly, to aid in the characterisation of the enzymes in terms of their substrate specificity, and secondly in the stabilisation of crystal formation.

It was decided that the best inhibitor to produce was the decapeptide sequence GGGVRGPRVV-amide, with a non- reducible peptide bond at the scissile bond. The process to bring about the production of this inhibitor was a complex chemical procedure, which started with the production of the ornithine diol. This residue was kindly supplied by Dr.R.Sharma and Dr. R.Broadbridge.

4.9.1 Inhibitor synthesis.

4.9.1.1 Structure of the proposed inhibitor.

Gly-Gly-Gly-Val-(Arg)-CH_2-NH-(Gly)-Pro-Arg-Val-Val-NH_2

It was believed that this structure should be a good inhibitor of batroxobin, as it comprised of the identical sequence found in fibrinopeptide A. Therefore it should be recognised as a substrate by the enzyme. However, as it does not contain a conventional peptide bond, but instead has no carbonyl group just a CH_2, then it cannot undergo cleavage. This is due to the fact that the carbonyl group is necessary for the nucleophilic attack of the catalytic serine.

4.9.1.2 The formation of the boc-arginine diol from the ornithine diol.

The protecting group on the nitrogen of the ornithine was removed via hydrogenation using H_2 and a platinum catalyst as shown in *Figure 4.13.* This allowed the exposure of the amine group and its subsequent guanylation to arginine via the reaction with the classical guanylating agent 1-Guanyl 3,5-dimethyl pyrazole nitrate. This reaction resulted in the formation of the boc arg diol, as can be seen in *Figure 4.13.* The boc group was then removed by bubbling HCl gas through a solution of the amino acid in ether.

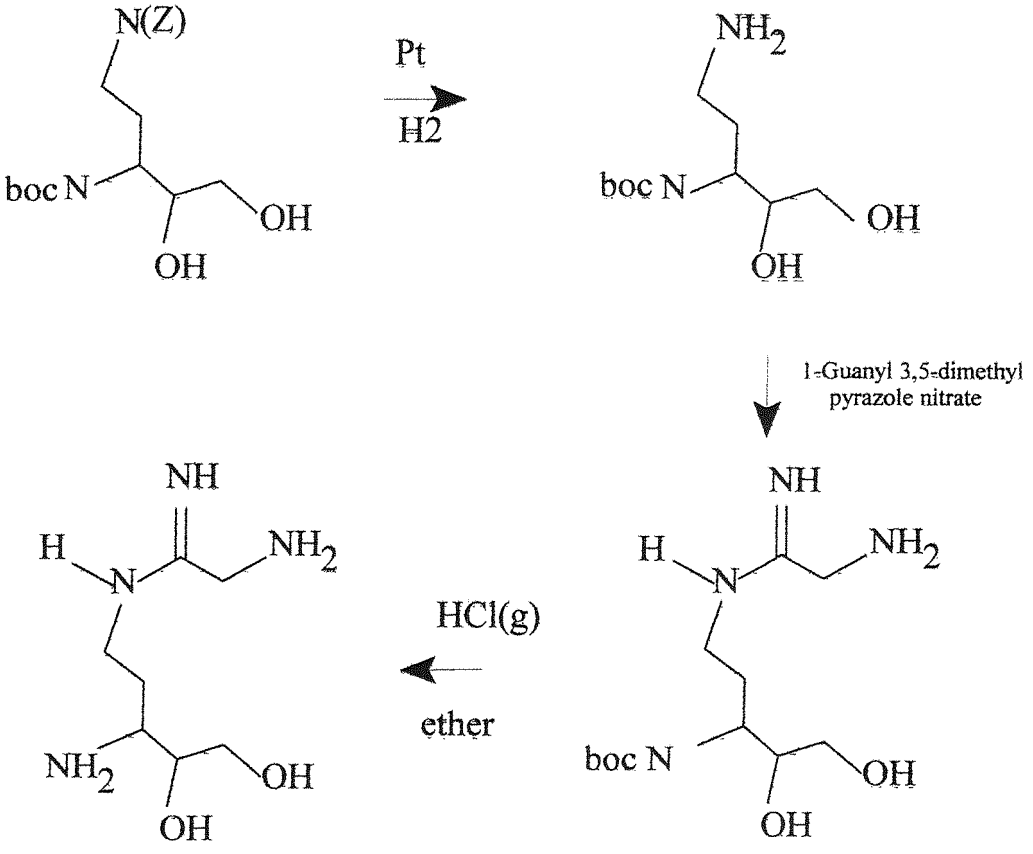


Figure 4.13:Formation of the deprotected arginine-diol.

4.9.1.3 The formation of the boc-valine-arginine aldehyde.

This arginine diol was then coupled in solution to a boc protected valine residue, which was pre-activated by its coupling to O-succinate. This reaction can be seen in *Figure 4.14.* The boc val arg diol was then oxidised to form the aldehyde by the action of metaperiodate (IO_4^-) ions, as seen in *Figure 4.14.*

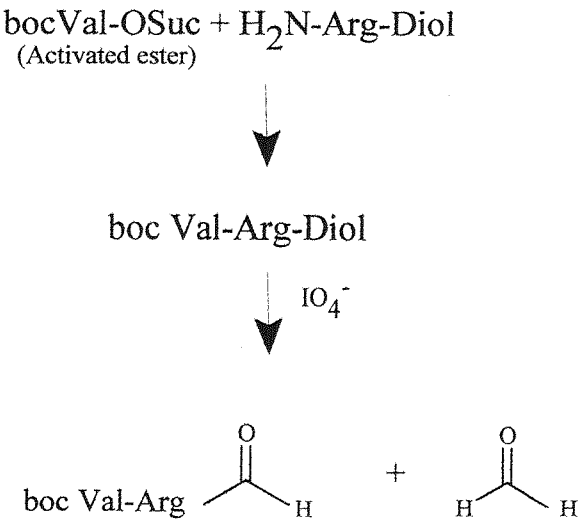


Figure 4.14: Formation of the boc valine arginine aldehyde.

4.9.1.4 Formation of the reduced peptide bond.

The first 5 residues from the c-terminus end of the peptide were coupled using conventional Fmoc solid phase peptide synthesis to the resin as outlined in section 4.3. The aldehyde was then coupled to the glycine attached to the resin under reducing conditions. This was achieved by the reaction with the solid phase bound glycine residue, the aldehyde and an excess of sodium cyanoborobydrate ($NaCNBH_3$) in DMF. This was carried out in the presence of enough acetic acid to maintain the pH throughout at 5, since this provided a balance between a low enough pH to facilitate the coupling, but a high enough pH to allow the reduction to be performed. This reaction can be seen in *Figure 4.15.* The remainder of the synthesis followed the

conventional Fmoc couplings as previously described. The cleavage and purification were followed as described for all other amide peptides.

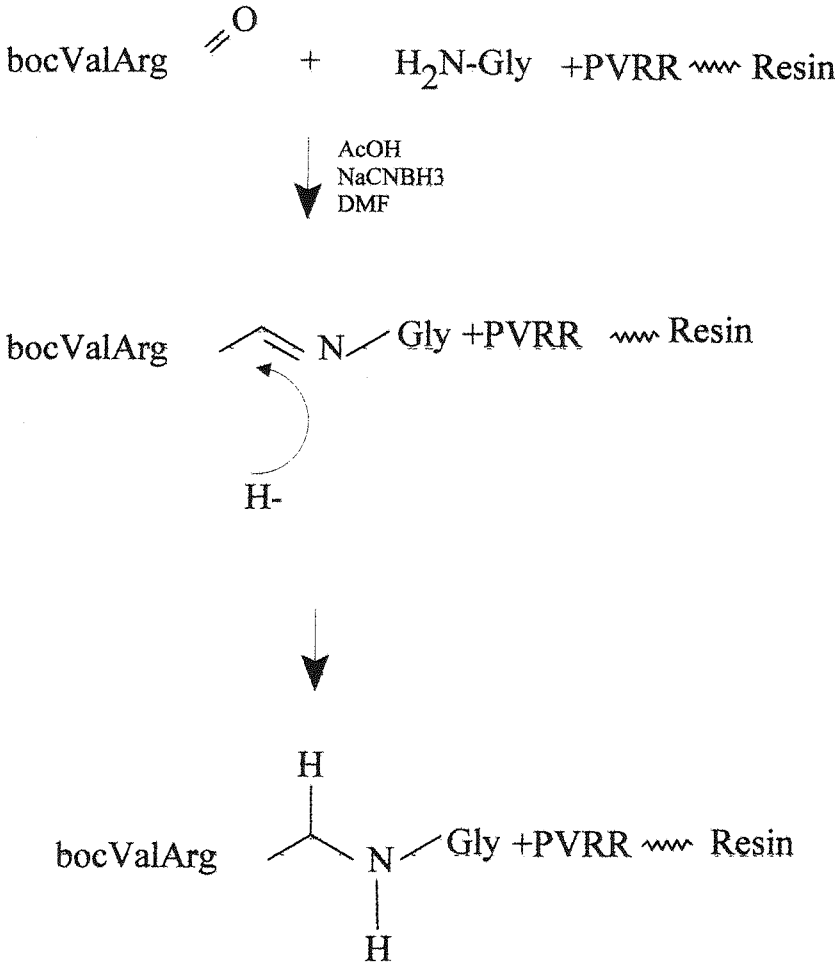


Figure 4.15: Formation of the reduced peptide bond.

The predicted mass for the inhibitor was 937Da. However, only the doubly charged species with a mass of 469Da could be identified using electrospray MS, as can be seen in *Figure 4.17.* This is not unexpected, as this peptide is an amide and they are notoriously difficult to identify as the singly charged parent ion using electrospray.

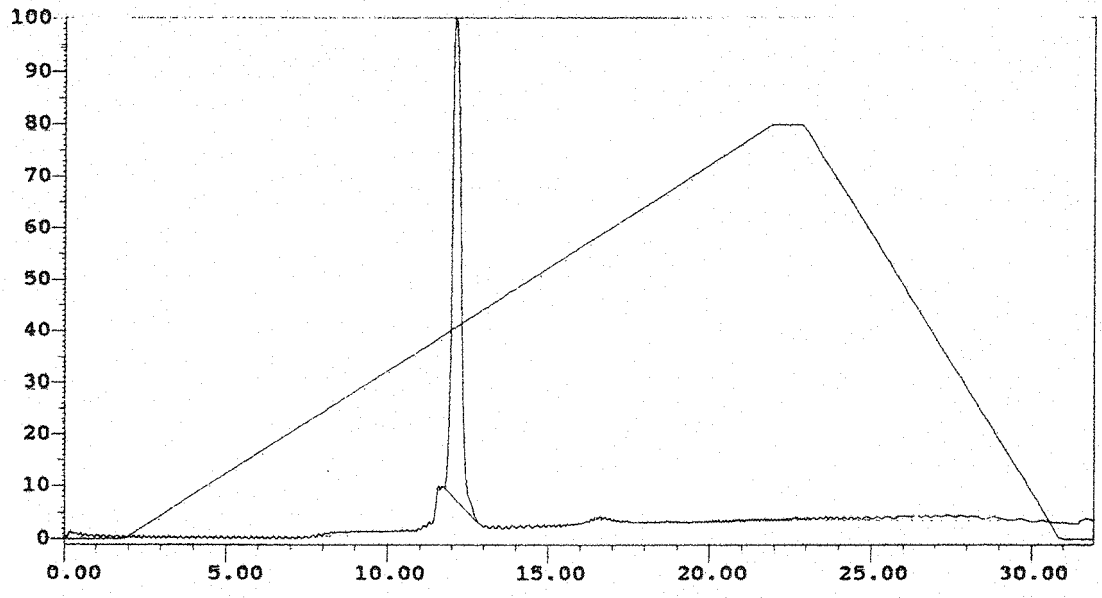


Figure 4.16: Analytical trace showing the inhibitor GGGVR-reduced-GPRVVamide.

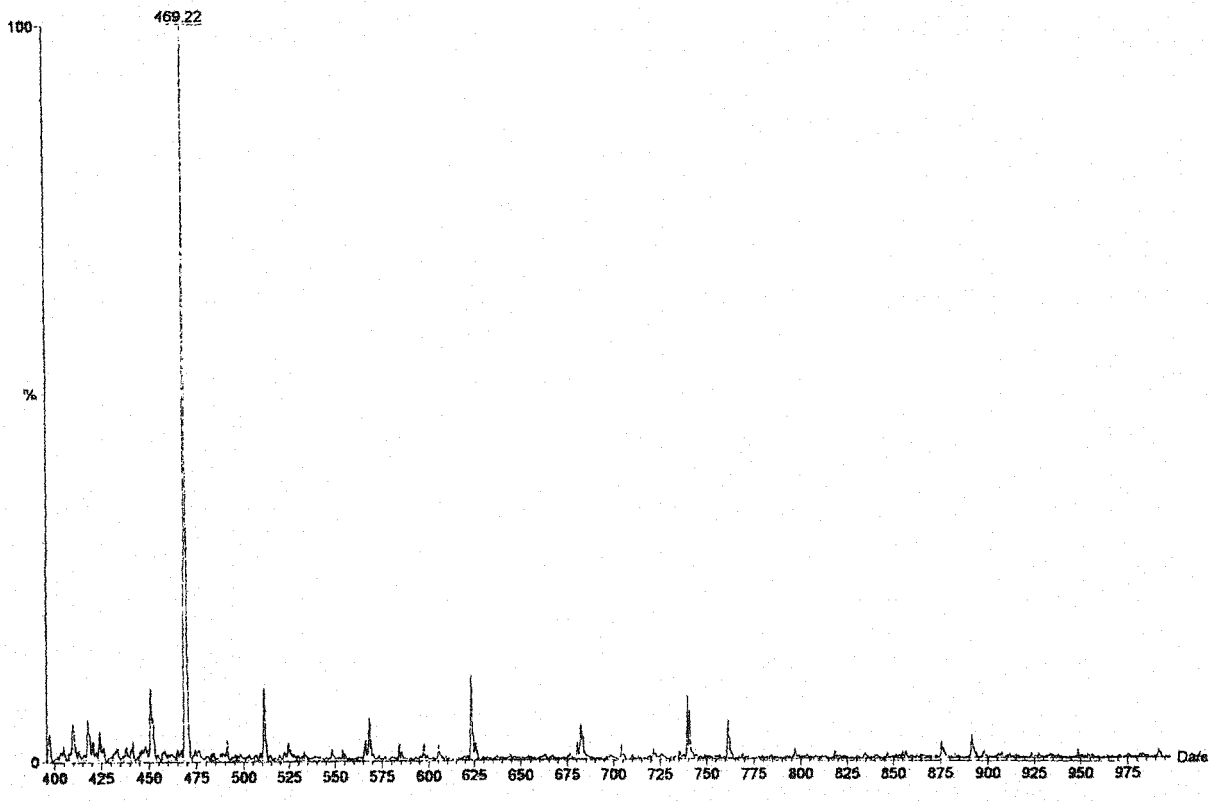


Figure 4.17: Electrospray MS of the inhibitor GGGVR-reduced-GPRVV-amide, showing the doubly charged species.

4.9.2 Reaction of inhibitors.

During the synthesis of the one major inhibitor detailed in section 4.9.2 a number of by-products were produced which could potentially have exhibited inhibition properties with one or all of the enzymes. All of these inhibitors plus some other more conventional inhibitors were tested

4.9.2.1 Experimental procedure

The action of each substance as an inhibitor was monitored by observing the effect that they had on the standard reaction of each enzyme with the chromogenic substrate S-2238. The protocol for these reactions is detailed in Chapter 2.

4.9.2.2 Results of inhibitor studies

Each of the potential inhibitors was considered to assess whether they had any inhibitory effect on each of the enzymes. The results of this preliminary study are shown in *Table 4.3.* Any compound that showed some effect was then considered in more detail as a kinetic study and the data was used to calculate the Ki for each compound. The results from this study are shown in *Table 4.4.*

Inhibitor(AMIDES)	Thrombin	Batroxobin
GGGVR^{RED}GPRVV	✗	✗
FmocGGGVR^{RED}GPRVV	✗	✗
GGGVO^{RED}GPRVV	✗	✗
FmocGGGVO^{RED}GPRVV	✗	✗
GGGVRGPRVV	✓	✓
GGDVRGPRVV	✓	✓
GVRGPR	✓	✓
V-R-aldehyde	✓	✓
FmocV-R-aldehyde	✓	✗
Leu-peptin	✓	✓

Table 4.3: Inhibitory effects of each compound on thrombin and batroxobin.

As can be seen from *Table 4.3*, none of the potential inhibitors containing the reduced peptide bond showed any inhibitory effect on either thrombin or batroxobin. This is surprising since identical decapeptide sequences that contained a normal peptide bond had been shown to act as good substrates, see section 4.8. In addition to this was the fact that these same peptides were shown to have some inhibitory effect on both of enzymes. As the only major structural differences between these compounds is either the lack of a carbonyl group, or in the fact that the reduced bond is planar rather than tetrahedral in shape, it was decided that the arginine aldehyde should be investigated as a potential inhibitor. It was thought that this would clarify whether the carbonyl was necessary for the recognition of the sequence.

As can be seen from *Table 4.3,* the aldehyde groups did in fact act as inhibitors to both of the enzymes, although they were considerably better inhibitors of thrombin than batroxobin, see *Table 4.4*. However, this can be explained by the earlier observation that batroxobin has a requirement for quite a long region of amino acid recognition for it to react well with a substrate, and that any change in this sequence will result in no reaction, see section 4.8. It is interesting to note that the batroxobin obtained from both species show identical reactivities with S-2238 and all the inhibitors.

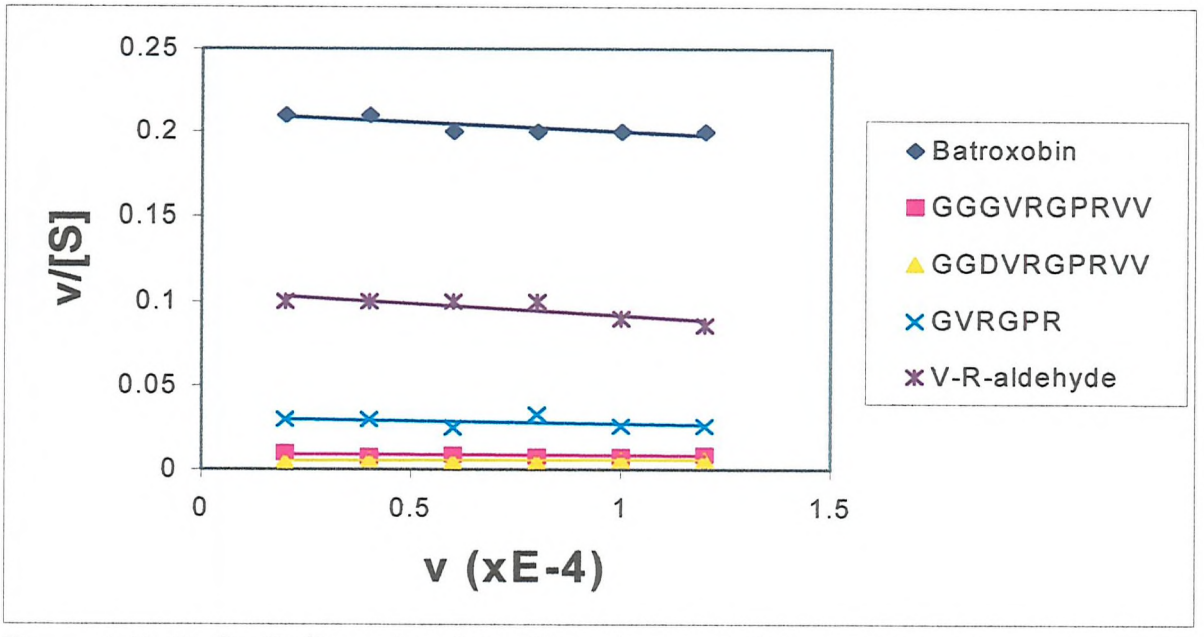


Figure 4.18: Eadie-Hoffstee plot of the effect of various inhibitors (shown as described in the legend) on the reaction of batroxobin with S-2238.

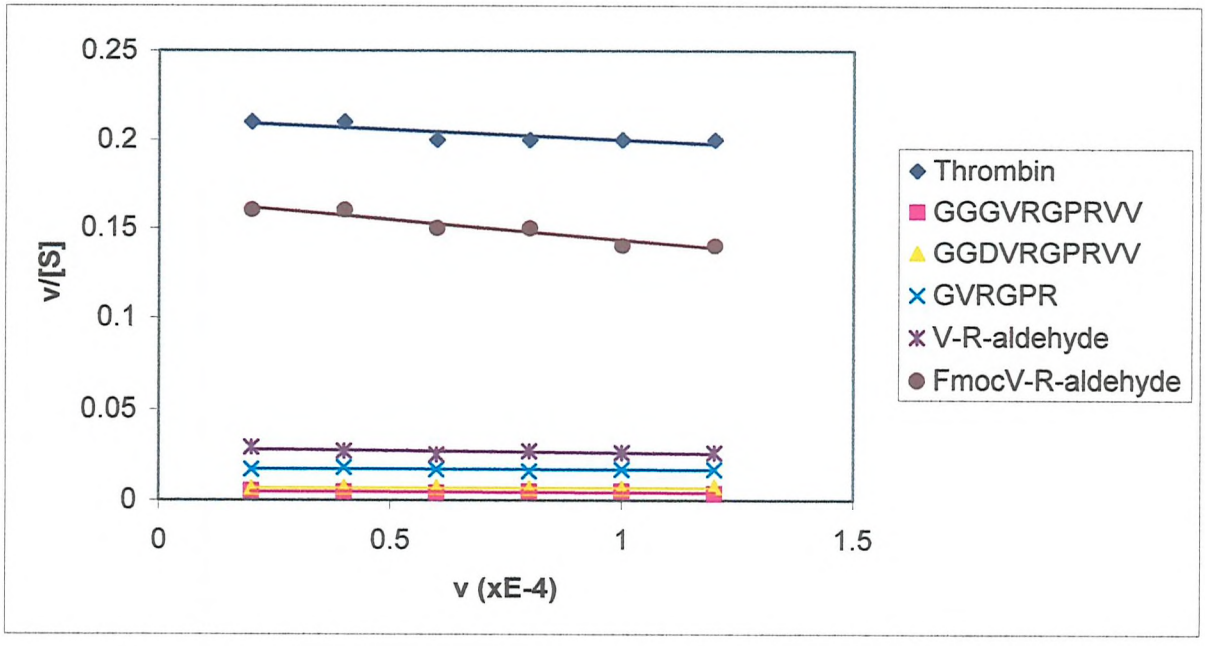


Figure 4.19: Eadie-Hoffstee plot of the effect of various inhibitors (shown as described in the legend) on the reaction of thrombin with S-2238.

Inhibitor	Ki for Thrombin	Ki for Batroxobin
GGGVRGPRVV	$2.5x10^{-6}$	$4.0x10^{-6}$
GGDVRGPRVV	$3.2x10^{-6}$	$2.4x10^{-6}$
GVRGPR	$1.0x10^{-5}$	$1.4x10^{-5}$
V-R-aldehyde	$1.5x10^{-5}$	$1.2x10^{-4}$
FmocV-R-aldehyde	$2.4x10^{-4}$	-

Table 4.4: Table to show the relative Ki values of various inhibitors with thrombin and batroxobin

4.10 Discussion.

The results obtained from this study of various synthetic peptide substrates and inhibitors have helped to elucidate the importance of certain residues and chemical groups around the scissile bond of the substrate. It has long been thought that the differences in specificity exhibited by batroxobin compared to thrombin was due to external binding regions on the surface of the enzyme[102]. However, the

results obtained in this study for thrombin as compared to batroxobin recognition seem to indicate that the residues and binding in the vicinity of the active site are of great importance.

Thrombin appears to be capable of recognition, binding and cleavage of hexa- and decapeptides containing the R-G scissile bond that are both identical to fibrinopeptide A & B and are combinations of the two. Although it does show a marked preference to the sequences obtained in fibrinopeptide A, it is nevertheless still capable of cleavage of the other sequences. Batroxobin, on the other hand, shows a very great preference for the sequences identical to fibrinopeptide A. In fact, batroxobin shows a much higher specificity when an aspartic acid residue is introduced into the sequences GGGVRGPRVV-amide of human fibrinopeptide A at the P3 position to produce the peptide GGDVRGPRVV-amide. This is observed by an increase in the rate of the reaction with this peptide. This is very interesting as this sequence is derived from rat and mouse fibrinogens, which are the natural prey for *Bothrops Atrox* pit vipers. This seems to suggest that there has been some evolution of batroxobin for the fibrinogens of its natural prey. However, this peptide was found to be a much poorer substrate for thrombin than the sequence containing glycine at the P3 position. This is again interesting, as the thrombin used in this study was derived from human plasma and human fibrinogen contained a glycine at the P3 position. It would be interesting to determine if batroxobin is behaving in a similar way to rat and mouse fibrinogens, which might exhibit sequence specificity for substrates with an aspartic acid at the P3 position. This would be in keeping with the known sequence of their own fibrinogen. In fact, even when the proline residue, which was considered to be important for recognition, was re-introduced into the sequence there appeared to be no reaction.

However, it was eventually possible during the course of these experiments to use a sufficiently large concentration of batroxobin relative to the peptide substrate concentration to force some reaction with the peptide SARGPR-amide. This sequence is identical to that found in fibrinopeptide B, with the notable difference that this histidine at the P2' position SARGHR-amide has been replaced by a proline. When a 100x excess of batroxobin was used in these experiments and incubated at 37°C for 24hrs it was possible to detect some cleavage of the peptide. This was never seen with any of the other sequences and as such was taken to be valid. This suggests

that the residue at the P2' position is extremely important. A proline in the P2' position as found in fibrinopeptide A is not only neutrally charged, compared to the histidine found in fibrinopeptide B, but it also results in the formation of a bend in the substrate. This conformational change between the two substrates could be sufficient to prevent batroxobin from cleaving the scissile bonds of sequences that do not contain proline. The fact that some small activity can be forced with the reintroduction of the proline into the sequence, albeit only a fraction of that originally exhibited, shows that this is an important residue recognition.

Another important observation was the lack of inhibition observed when the reduced bond inhibitor of the fibrinopeptide A sequence was used in the reactions. This is again an interesting point, since it has always been assumed that the carbonyl group of the peptide bond was important only in the mechanism of cleavage of the substrate by the enzyme, and not in the recognition and binding of the substrate. This suggests that in this group of enzymes at least the first step of the catalytic mechanism is also part of the recognition and binding of the substrate. Therefore the reaction proceeds *via* a continuous recognition, binding and catalysis step, and two distinct steps of binding followed by catalysis. The fact that the arginine aldehyde acts as an inhibitor reinforces the significance of the carbonyl group.

These observations together indicate that this region around the active site of the enzyme must be important for substrate recognition. This is very interesting, as previously held views suggested that the differences exhibited by batroxobin compared to thrombin were due only to exosites on the enzymes and that no differences would be exhibited by residue changes around the scissile bond. Although it still appears evident that fibrinogen recognition exosites are very important in the specificity of batroxobin, it now also appears evident that more subtle sequence differences may also be important.

CHAPTER FIVE

HOMOLOGY MODELLING OF BATROXOBIN

5.1 Introduction

Structural determination of the snake venom proteases by X-ray crystallography has so far proven to be extremely difficult. In fact, none of these proteins, including batroxobin, have had their structures resolved to date. This is a problem that is exacerbated by the fact that recombinant forms of the enzymes have been produced only in a very few cases, resulting in a difficulty in isolating sufficiently large quantities of the proteins. It was therefore decide that homology modelling could be a useful tool in producing a model structure that could help to explain some of the more pressing structural questions, whilst at the same time producing a structure upon which to base any crystallographic refinement.

The proteins belonging to the serine proteinase super family have all been shown to have significant sequence homology,[5] particularly in the regions surrounding the catalytic site, the 2 β-pleated sheets either side of the catalytic cleft and the C-terminal α-helix. However, various external loop regions are not conserved from one structure to another and it is believed that these areas may be significant in substrate recognition. Since it is the substrate specificity of batroxobin that is of major interest in this study, a detailed knowledge of the structure is essential.

5.2 Sequence homology analysis

The sequences of the different serine proteases (*figure 5.1*) were aligned with the sequence obtained for batroxobin (*B.moojeni*). Due to the fact that the sequences and structures of various kallikreins became available at different times, the initial alignments were between:

CHYMOTRYPSIN B (RAT)
CHYMOTRYPSIN B (HUMAN)
GRANZYME F (MOUSE)
GRANZYME F (HUMAN)
ELASTASE (HUMAN)
DUODENASE I (BOVINE)
PROTHROMBIN (RAT)
PROTHROMBIN (HUMAN)

VENOMIN A (LACHERIS MUTA MUTA)
GLANDULAR KALLIKREIN II (HUMAN)
GLANDULAR KALLIKREIN (PIG)
GLANDULAR KALLIKREIN (GUINEA PIG)
GLANDULAR KALLIKREIN I (HUMAN)

Other sequences were added to these alignments, as they became available. The sequence alignment was initially determined using a Needleman-Wunsch algorithm[78] and a protein sequence score matrix [79]. The alignment was then refined by eye where necessary and hydropathy plots were produced for each sequence, in order to look at the overall similarity in the electronic environments between each sequence and batroxobin. This was achieved by assuming the conservation of particular residues known to be structurally significant. In particular, the regions that correlated to the central β-pleated sheet structure, which appears to be conserved throughout the whole family, were used to refine the initial alignment. This allowed the arrangement of batroxobin in the same way as all the other members of the family. This was an important step as it made the location of the other residues known to be of structural significance much easier to locate accurately, particularly the residues involved in the catalytic triad His 57, Asp 102 and Ser 195 (by chymotrypsin numbering). Once the highly conserved areas in the central β-pleated sheet regions had been set in place, not only did the catalytic triad align extremely well but also the major insertions and deletions and alignments aligned with loop regions in the other serine proteases. The batroxobin aligned with the kallikrein structures particularly well. These facts add credibility to the fact that batroxobin may have the same basic structure as the kallikrein group of the serine proteases, and not the thrombin group.

All the serine proteases considered showed a reasonably high identity to the batroxobin sequence. All the identities within the serine protease group are around 20%, including in all cases the structurally significant β-pleated sheet and the catalytic triad. The initial alignments with porcine pancreatic kallikrein showed ~40% identity with batroxobin and as much as 70% homology for similar residues. Later work, once the sequences became available, showed a similar degree of identity

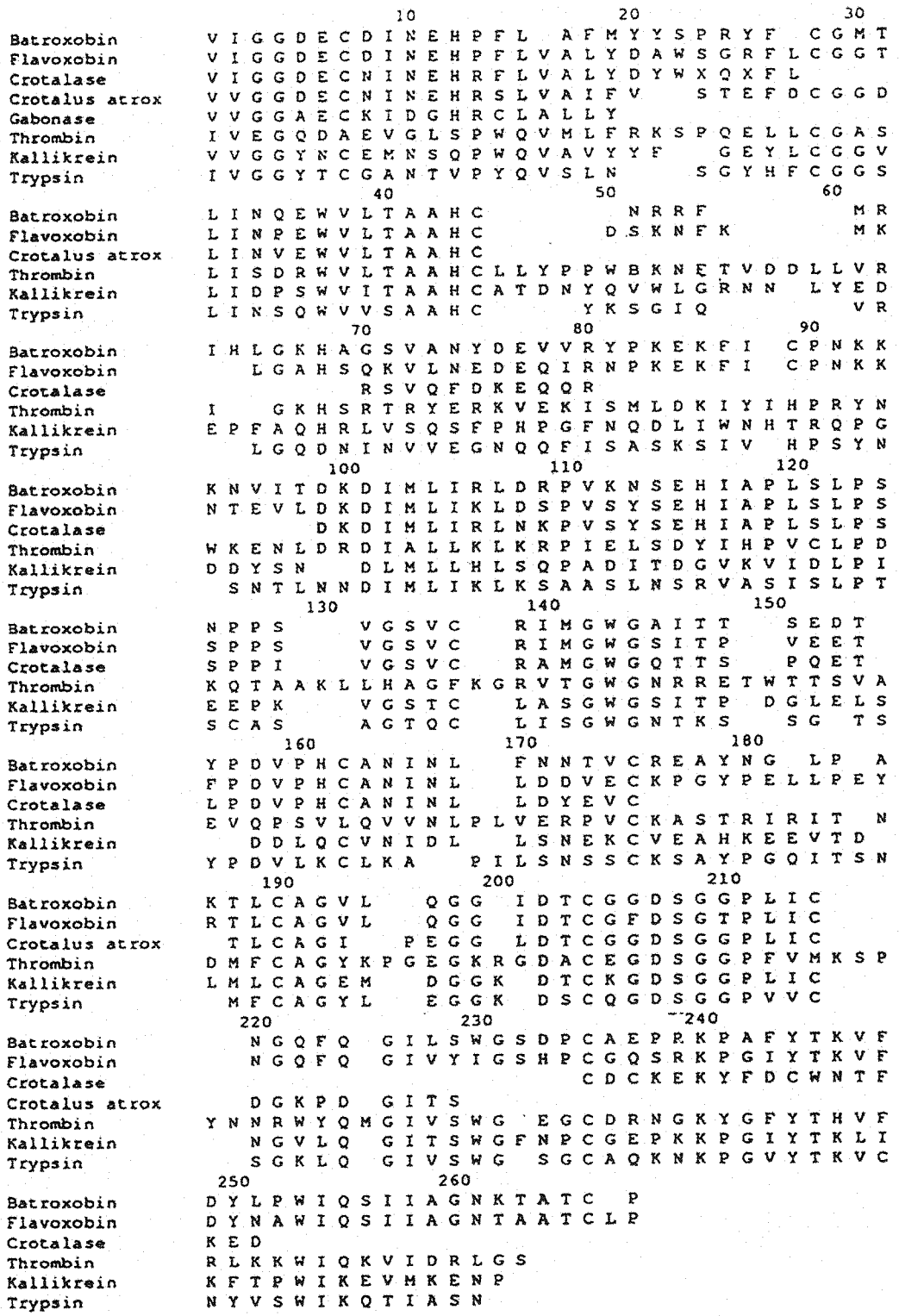


Figure 5.1 Alignment of the amino acid sequences of batroxobin with various serine proteases.

and homology for batroxobin with mouse salivary gland kallikrein and around 70% identity with the snake venom derived plasminogen activator. Unfortunately, the molecular co-ordinates for this structure are not due to be released until late 1999, meaning that a full model based upon this structure could not be produced. However, it has been possible to incorporate some of the structural findings from that sequence into the models produced from the kallikreins.

5.3 Modelling of batroxobin from *B. moojeni* on porcine pancreatic kallikrein.

The sequence homology between batroxobin and kallikrein derived from the pancreas of *Sus scrofa,* the wild boar, (40%) was the highest out of all the serine proteases initially considered, but not as high as the sequence homology between pancreatic kallikreins from different species (>60%). This fact was in many ways suprising as it was thought that as a thrombin-like enzyme it might well exhibit high degrees of structural similarity with serine proteases from the coagulation protease family. The differences were largely due to the fact that batroxobin does not contain any N or C-terminal extensions as found in the majority of coagulation enzymes, including thrombin, but rather has the smaller, simpler structure as exhibited by the kallikreins[98].

Once this alignment was satisfactorily completed, the Quanta/CHARMm program was used to perform homology modelling. The sequence homology found was used to model the backbone of the batroxobin to the porcine glandular kallikrein with certain key residues, for example the catalytic triad, being totally conserved. Any insertions or deletions in the aligned sequences were, where possible, incorporated into the loop regions of the structure rather than the more highly conserved β-sheet or α-helix regions. The backbone co-ordinates of aligned residues in porcine glandular kallikrein were mapped to the equivalent amino acids in the batroxobin model. Where any insertions or deletions occurred in the sequence, a fragment database [80] was used to search for short sequences that overlapped with the residues on either side of the unknown region and which contained the same number of residues[99].

```
Porcine
Kallikrein  IIGGRECEKN  SHPWQVAIYH  YSSFQCGGVL  VNPKWVLTAA
Batroxobin  VIGGDECDIN  EHPFLAFMYY  SPRYFCGMTL  INQEWVLTAA
                           -b1-       ----b2----  --b3--

Porcine
Kallikrein  HCKNDNYEVW  LGRHNLFENE  NTAQFFGVTA  DFPHPGFNLS
Batroxobin  HCNRRFMRIH  LGKHA*****  GSVANYDEVV  RYPKEKFICP
              ---b4---                  -----b5----


Porcine
Kallikrein  **ADGKDYSH  **DLMLLRLQ  SPAKITDAVK  VLELPTQEPE
Batroxobin  NKK*KNVVIT  DKDIMLIRLD  RPVKNSEHIA  PLSLPSNPPS
                           --b6--


Porcine
Kallikrein  LGSTCEASGW  GSIEPGPDBF  EFPDEIQCVQ  LTLLQNTFCA
Batroxobin  VGSVCRIMGW  GAITTSEDTY  **RDVPHCAN  INLFNNTVCR
              ---b7--                  ---b8---    ----a1


Porcine
Kallikrein  BAHPBKVTES  MLCAGYLPGG  KDTCMGDSGG  PLICNGMWQG
Batroxobin  EAYN*GLPAK  TLCAGVLQGG  IDTCGGDSGG  PLICNGQFQG
            ----         --b9--                 -b10- ----


Porcine
Kallikrein  ITSWGHTPCG  SANKPSIYTK  LI*FYLDWIN  BTITENP***
Batroxobin  ILSWGSDPCA  EPRKPAFYTK  *VFDYLPWIQ  SIIAGNKTAT
            b11--           -b12-    -------a2-------


Porcine
Kallikrein  **
Batroxobin  CP
```

Figure 5.2: Alignment of the amino acid sequences of batroxobin with porcine pancreatic kallikrein. The stars represent any gaps in the sequences. Residues shown in bold are part of the catalytic triad. Regions of the sequence that are α-helices are represented by a1-a2. Regions of the sequence that are β sheets are shown by b1-b12.

The fragment with the best fit was then chosen and used in the model structure. Once the backbone co-ordinates had been defined, the validity of the modelled structure was assessed using the protein health facility within Quanta. Any irregularities in the structure identified from this process were either remodelled manually, or by re-using the fragment database to find a more appropriate fragment. Once the backbone residues had been satisfactorily defined, the side chains of any identical residues were copied directly from the porcine glandular kallikrein to the batroxobin model. Quanta/CHARMm was then used to generate side chains in appropriate conformations for any non-identical residues. Throughout this process, any major side chain clashes were moved manually and then CHARMm was used to perform energy minimisation. During this process, constraints were applied to preserve the basic backbone structure. The final structure produced was then checked for any inconsistencies using the protein health module of Quanta.

The alignment between batroxobin and porcine pancreatic kallikrein shows a few small insertions and deletions, but the only major difference appeared to occur in the C-terminal region. Within the batroxobin sequence there was a sequence of seven amino acid residues that does not have a corresponding sequence within the porcine pancreatic kallikrein. This sequence also contained a cysteine residue (cys-230) that could form an additional disulphide bridge with another unpaired cysteine residue (cys-) within the body of the structure. This would result in six disulphide bridges within the structure of batroxobin, one more than the five found in kallikrein. Initially this s-s bond was not made, as there was insufficient evidence to support its existence. However, information that came to light later in the process, resulted in this area of sequence being reconsidered, for a possible s-s bond, as will be discussed in section 5.5.

As a result of the sequences aligning so well, the basic structure from the porcine pancreatic kallikrein (*figure 5.4*) was retained in the model of batroxobin (*figure 5.5*). That is, the two regions of β-pleated sheet either side of the catalytic cleft were conserved, along with the short region of α-helix at the C-terminal end, with the remainder of the structure being made up of loop regions between these three conserved areas, which are also well conserved. It is the tail at the end of the C-terminal α-helix that shows the major difference between the structures.

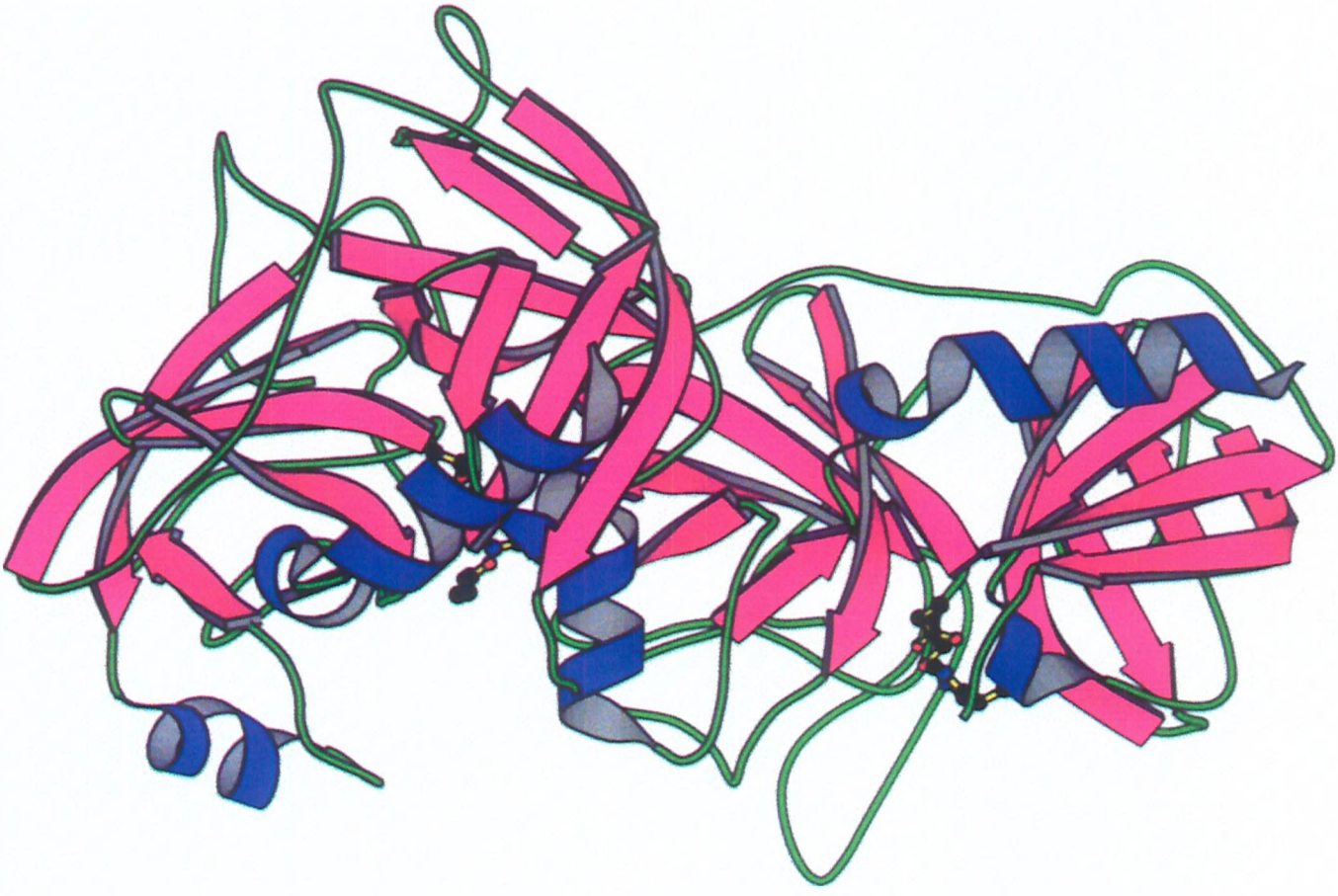


Figure 5.3: Schematic representation of the backbone of the dimeric porcine pancreatic kallikrein. Catalytic residues are shown as ball and stick.

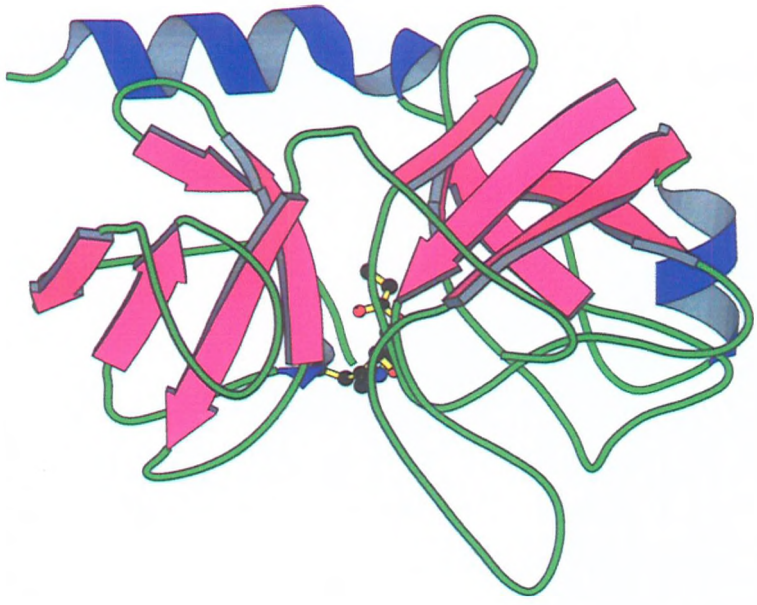


Figure 5.4: Schematic representation of the backbone of porcine glandular kallikrein. Catalytic residues are shown as ball and stick.

Figure 5.5: Schematic representation of the backbone of the model batroxobin based upon porcine glandular kallikrein

Interestingly, pig glandular kallikrein appears to exist in a dimeric form under crystallography conditions. An open heterologous dimer was shown that had one accessible binding site, whilst the other was inaccessible due to the manner in which it was bound to the other molecule of enzyme. It remains unclear as to the actual validity of this structure *in vivo* as it is possible that the presence of the dimeric form is merely an artefact of the crystallisation process.

One of the ways in which the molecular modelling was hoped to assist in the project was in its potential to provide structural evidence to explain the highly selective way in which batroxobin acts. Interestingly enough, the work on solving the structure of the porcine kallikrein carried out by Bode *et al* attempted to do just this with kallikrein in relation to the structure of trypsin. Kallikrein is a more specific enzyme than trypsin and, unlike trypsin, it is unable to use large proteins as substrates. It had been suggested from the work carried out by Bode *et al* that this might be explained by the differences that exist at the external surface loops between the two structures. Therefore, despite the fact that the active sites between the two structures are very similar, the elongation of the peptide segment from 217 to 220, chymotrypsin numbering, mean that the specificity pocket is bigger in kallikrein than in trypsin, thus allowing a more bulky group to bind in the P1 (S1) site. It also seems apparent the side chain of Ser226 in kallikrein partially covers the bottom of the specificity pocket and a mobile phenolic group, on Tyr599, which extends into the binding site, these result in the partial occlusion of the site and its subsequent specificity for small residues.. These factors have been postulated in explaining the binding of substrates to kallikrein.

5.3.1 Validity of the modelled structure of batroxobin based upon porcine pancreatic kallikrein.

The model that was produced of batroxobin, based upon the structure of porcine kallikrein was in many ways highly satisfactory, as it was able to show the positioning of the active site catalytic triad very clearly. The homology between the primary sequences of the two enzymes was calculated to be ~40%; however, in many ways the areas of the structures that were not identical did indeed show great similarities in terms of the types of residues present. However, a highly extended

region at the C-terminus at the end of the final helix was difficult to model, as it would have required a large amount of 'guesswork', and therefore it was left largely unassigned. This area may indicate that the structure upon which the modelling was carried out was not ideal – however, as it was the best structure available.

5.4 Batroxobin modelled on mouse salivary gland kallikrein

When the structure of mouse salivary gland kallikrein became available later in the project, it was decided to use this as a possible modelling template. Mouse salivary gland kallikrein was chosen due to the possible evolutionary significance, along with a possible structural difference that may, if it was present in the batroxobin, help to explain the uniquely specific nature of the enzyme. As has previously been mentioned (Chapter 1), there is much evidence that the snake venom glands are directly evolved from salivary glands. It therefore follows that the snake venom derived serine proteases may also be derived from salivary gland serine proteases. The other major factor that led to the decision that this structure should also be used for modelling purposes was the fact that it is the only kallikrein structure that has a significantly pronounced "kallikrein loop". This loop may allow the additional residues that exist at the C-terminal end of the helix to be assigned[100].

The alignment of the sequences and the modelling of batroxobin on mouse salivary gland kallikrein were performed as previously detailed. As with the porcine kallikrein, the underlying homology was found to be about 70% with identity at 40%. The area surrounding the active site cleft, including the two β sheeted barrels, were well conserved between all three structures, as were the two α-helices. The major differences between the structures were exhibited in the loop regions and the extended C-terminal, as anticipated from the sequence of the salivary gland kallikrein.

```
Mouse
Kallikrein  VVGGFNCKKN  SQPWQVAVYY  QKEHICGGVL  LDRNWVLTAA
Batroxobin  VIGGDECDIN  EHPFLAFMYY  SPRYFCGMTL  INQEWVLTAA
                              -b1-       ----b2----  --b3--

Mouse
Kallikrein  HCYVDQYEVW  LGKNKLFQEE  PSAQHRLVSK  SFPHPGFNMS
Batroxobin  HCNRRFMRIH  LGKHA*****  GSVANYDEVV  RYPKEKFICP
              ---b4---                      -----b5----


Mouse
Kallikrein  SLLMLQTIPP  GADFSN**DL  MLLRLSKPAD  ITDVVKPIAL
Batroxobin  ******NKK*  KNVVITDKDI  MLIRLDRPVK  NSEHIAPLSL
                                    --b6--


Mouse
Kallikrein  PTKEPKPGSK  CLASGWGSIT  PTRWQK**PD  DLQCVFITLL
Batroxobin  PSNPPSVGSV  CRIMGWGAIT  TSEDTY**RD  VPHCANINLF
                    ---b7--                     ---b8---

Mouse
Kallikrein  PNENCAKVYL  QKVTDVMLCA  GEMGGGKDTC  RDDSGGPLIC
Batroxobin  NNTVCREAYN  *GLPAKTLCA  GVLQGGIDTC  GGDSGGPLIC
             ----a1----          --b9--                -b10-

Mouse
Kallikrein  DGILQGTTSY  GPVPCGKPGV  *AIYTNLIKF  NS*WI**KDT
Batroxobin  NGQFQGILSW  GSDPCAEPRK  PAFYTK*VFD  YLPWIQSIIA
              ---b11---             -b12-  -------a2------

Mouse
Kallikrein  MMKNA
Batroxobin  GNKTATCP
            -
```

Figure 5.6: Alignment of the amino acid sequences of batroxobin with mouse salivary gland kallikrein. All stars indicate gaps between the sequences. Residues shown in bold print are part of the catalytic triad. Regions of α-helix are shown as a1-a2, and areas of β-sheet are represented by b1-b12.

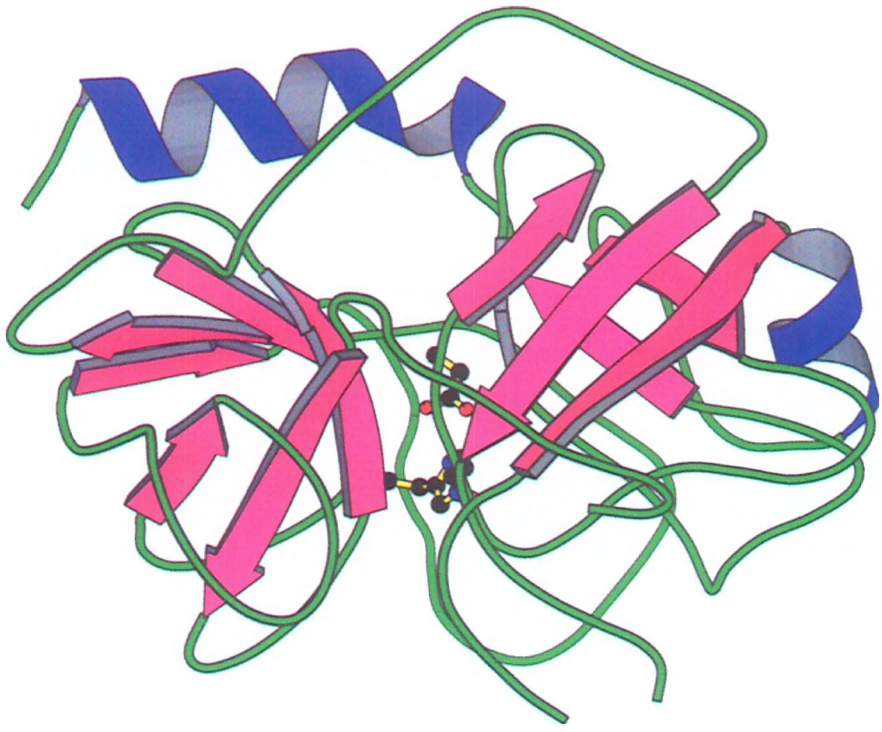


Figure 5.7a: Schematic representation of the backbone of mouse salivary gland kallikrein monomer. Catalytic residues are shown as ball and stick.

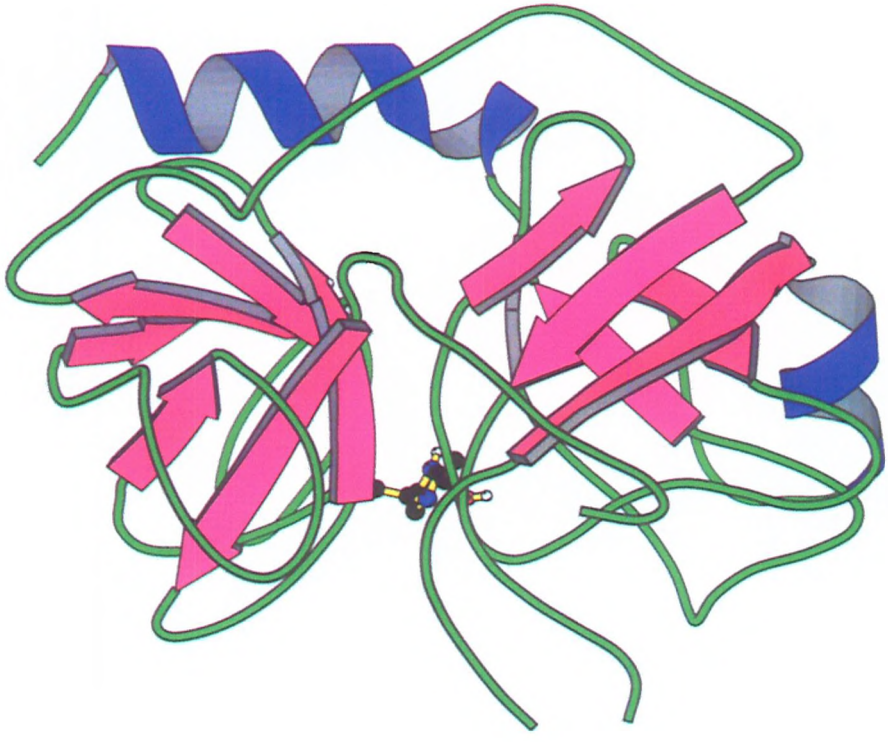


Figure 5.7b: Schematic representation of the backbone of the model of batroxobin based upon mouse salivary gland kallikrein.

As can be seen from *figure 5.7*, the extended C-terminal region obtained in the porcine kallikrein batroxobin model has been incorporated into the kallikrein loop within the batroxobin modelled on salivary gland kallikrein

5.5 The structure of batroxobin based upon *Trimeresurus stenjnejeri* venom plasminogen activator (TSV-PA)

Parry *et al* has recently solved the structure of a snake venom derived serine proteinase to 2.5 angstroms [101]. This amino acid sequence for this structure shows an extremely high homology to batroxobin, with almost 70% identity. Upon closer inspection of this structure it became apparent that there was, in fact, a C-terminal tail region, which was virtually identical to that suggested from the modelling of batroxobin on porcine pancreatic kallikrein. It would therefore appear that the first model of batroxobin, based upon porcine pancreatic kallikrein was more likely to be correct than that based upon salivary gland kallikrein.

Unfortunately, the co-ordinates for TSV-PA will not become available until August 1999 and consequently cannot be used for the purposes of this project. However, in light of the fact that the structure is obviously of immense importance and due to the fact that there are significant similarities with the model derived from porcine kallikrein, it was decided that the original model should be modified to incorporate any important features introduced by the TSV-PA structure.

The most striking structural difference between porcine pancreatic kallikrein and TSV-PA, is the existence of the C-terminal tail, which runs through a depression between the C-terminal α-helix and the 99 loop (a loop region 90 and 102). This is disulphide bridged between Cys91-Cys-245e) with an additional stabilising salt bridge between the C-terminal Pro245g and Lys101. The numbering used in this section corresponds to chymotrypsin, consequently any additional residues found in the TSV-PA and batroxobin structures are given the number corresponding to the last residue found in chymotrypsin and a letter which indicates the position of the residue in the extended loop. For example TSV-PA has 7 residues more than chymotrypsin at the c-terminal, therefore these residues are given the number 245 which is the last residue in chymotrypsin and each of the 7 residues are called a-g respectively. On

```
Porcine
Kallikrein   IIGGRECEKN   SHPWQVAIYH   YSSFQCGGVL   VNPKWVLTAA
Batroxobin   VIGGDECDIN   EHPFLAFMYY   SPRYFCGMTL   INQEWVLTAA
TSV-PA       VFGGDECNIN   EHRSLVVLFN   SNGFLCGGTL   INQDWVVTAA
                                -b1-        ----b2----   --b3--

Porcine
Kallikrein   HCKNDNYEVW   LGRHNLFENE   NTAQFFGVTA   DFPHPGFNLS
Batroxobin   HCNRRFMRIH   LGKHA*****   GSVANYDEVV   RYPKEKFICP
TSV-PA       HCDSNNFQLL   FGVHS*****   KKILNEDEQT   RDPKEKFFCP
                  ---b4---                    =====b5====


Porcine
Kallikrein   **ADGKDYSH   **DLMLLRLQ   SPAKITDAVK   VLELPTQEPE
Batroxobin   NKK*KNVVIT   DKDIMLIRLD   RPVKNSEHIA   PLSLPSNPPS
TSV-PA       NRK*KDDIEV   DKDIMLIKLD   SSVSNSEHIA   PLSLPSSPPS
                             --b6--


Porcine
Kallikrein   LGSTCEASGW   GSIEPGPDBF   EFPDEIQCVQ   LTLLQNTFCA
Batroxobin   VGSVCRIMGW   GAITTSEDTY   **RDVPHCAN   INLFNNTVCR
TSV-PA       VGSVCRIMGW   GKTIPTKEIY   **PDVPHCAN   INILDHAVCR
                ---b7--                      ---b8---     ----


Porcine
Kallikrein   BAHPBKVTES   MLCAGYLPGG   KDTCMGDSGG   PLICNGMWQG
Batroxobin   EAY*NGLPAK   TLCAGVLQGG   IDTCGGDSGG   PLICNGQFQG
TSV-PA       TAYSWRQVAN   TLCAGILQGG   RDTCHFDSGG   PLICNGIFQG
             a1---        --b9--                --b10--   ---


Porcine
Kallikrein   ITSWGHTPCG   SANKPSIYTK   LI*FYLDWIN   BTITENP***
Batroxobin   ILSWGSDPCA   EPRKPAFYTK   *VFDYLPWIQ   SIIAGNKTAT
TSV-PA       IVSWGGHPCG   QPGEPGVYTK   *VFDYLDWIK   SIIAGNKDAT
             b11-            -b12--    -------a2-------


Porcine
Kallikrein   **
Batroxobin   CP
TSV-PA       CPP
```

Figure 5.8: Alignment of the amino acid sequences of batroxobin with porcine pancreatic kallikrein and TSV-PA. Any stars indicate gaps between the sequences. Residues shown in bold represent the catalytic triad. Regions of α-helix are shown by a1-a2 and areas of β-sheet are shown by b1-b12.

consideration of the sequence homologies at these regions between TSV-PA and batroxobin, it is apparent that the disulphide bridge formation is definitely possible and that the additional stabilisation salt bridge could also be present (*figure 5.8).* This additional bond formation results in the movement of the 99 loop in TSV-PA from its position above the helix in both kallikreins (*figure 5.3 and 5.7)* to a lower region closer to the catalytic site to allow Cys91 to be close enough to Cys245e for bond formation to occur. It has been postulated by the authors that this loop region my result in the occlusion of the active site and may be responsible for the increased specificity demonstrated by TSV-PA.

The additional disulphide bridge was modelled into the batroxobin structure based upon porcine kallikrein as previously described. This required a great deal of manipulation in order to manoeuvre the Cys74 and Cys230 close enough to form a bond without causing any unreasonable contacts between the rest of the residues. In fact, in order to ensure that this was possible the 99 loop needed to be moved to occupy a similar position to that found in TSV-PA. This is an interesting observation as it may be that this loop is also significant in the increased specificity of batroxobin and its relative resistance to the effects of inhibitors. It remains unclear whether or not the additional stabilising effect of the Pro245g to Lys101 salt bridge seen in TSV-PA occurs in batroxobin. As can be seen from the sequence alignments, batroxobin has a six residue C-terminal extension and not the seven-residue extension exhibited by TSV-PA. This means that batroxobin has no residue at 245g, but it does have a proline at 245f. It is therefore possible that a salt bridge could be formed between Lys101 and Pro245f if these two residues could get close enough to form it. Modelling studies suggest that these residues could be in close enough proximity to form this interaction but it is unclear as to the validity of this observation.

The rest of the structure of the model of batroxobin based upon porcine kallikrein remains largely unchanged by the effect of the disulphide bond formation, with the two β sheeted barrels and the two α helices maintaining the same position. Closer inspection shows that the core of the structure around the active site cleft is also unchanged from the porcine kallikrein model, with the only changes displayed in the external loops. Interestingly the most affected loop bordering the active site, the

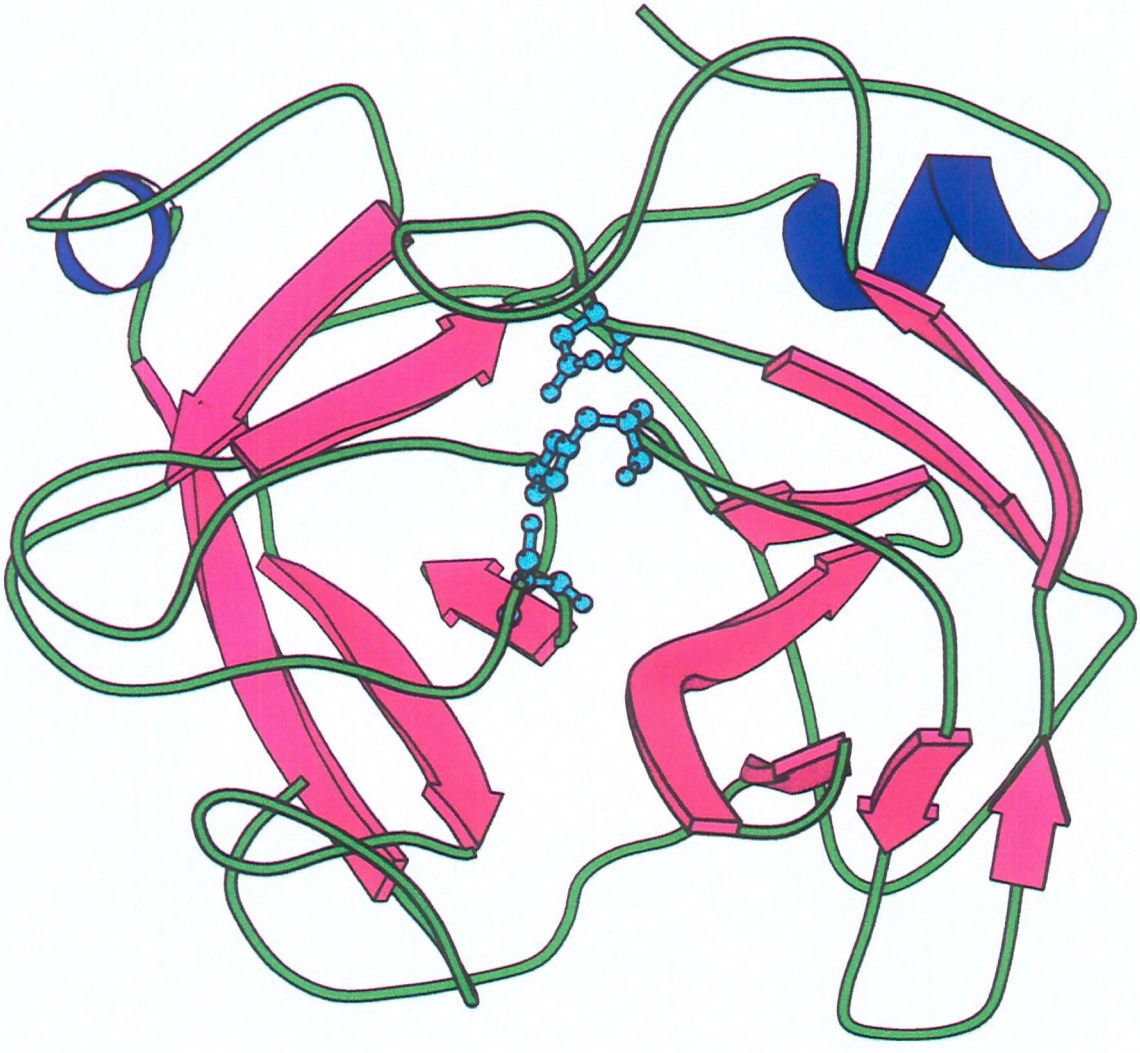


Figure 5.9: Schematic representation of the backbone of the model of batroxobin based upon porcine glandular kallikrein and TSV-PA. Active site residues are shown in cyan. This is shown to visualise a rotation through 180° compared to previous figures .

99 loop, appears to have residues that impinge into the top of the active site cleft. These residues seem to be involved in the formation of the roof of the active site, some falling very close to His57 of the catalytic triad, as also shown in the TSV-PA structure. It seems likely that the C-terminal tail and subsequent disulphide bridge formation, which is conserved throughout the snake venom serine proteinase family, results in the movement of the 99-loop to a region that may be close enough to the active site to explain the increased specificity displayed by these enzymes. However, since this area has been modelled onto a structure without the availability of the co-ordinates, it is important not to assume too much.

5.6 Comparison of batroxobin modelled on porcine kallikrein and TSV-PA with α-thrombin.

One of the major reasons for producing a good model of batroxobin was to attempt to explain the increased specificity exhibited by batroxobin compared to thrombin both in terms of the published specificities and in terms of the observed specificities for synthetic peptides, as seen in Chapter 4. It was therefore decided that once a good model was produced, a comparison should be made of its structure, particularly surrounding the active site with thrombin.

5.6.1 Comparison of the primary and overall structure of batroxobin and α-thrombin.

The sequence homology between α-thrombin and batroxobin is not as good as those exhibited between batroxobin and kallikrein, lending support to the belief that batroxobin is derived from kallikrein and is thrombin-like in activity, not structure. Comparisons show that there is about 40% homology but only ~25% identity between the two enzymes, as can be seen from the sequence alignment in *figure 5.10*. There are certain areas of major differences between the two structures, in particular there is no C-terminal extension in thrombin as displayed in batroxobin.

However, these large differences in sequence homology do not translate to differences in the overall appearance of the structure. When the two structures are compared directly by superimposing the batroxobin structure onto the α-thrombin

```
Human
Thrombin      IVEGQDAEVG  LSPWQVMLFR  KSPQELLCGA  SLISDRWVLTA
Batroxobin    VIGGDECDIN  EHPFL*AFMY  YSPRYF*CGM  TLINQEWVLTA
                                -b1-        ----b2----  --b3--

Human
Thrombin      AHCLLYPPWV  KNETVDDLLV  RI**GKHSRT  RYERKVEKISM
Batroxobin    AHC******N  DF*******M  RIHLGKHAGS  VANYDEVVRYP
                                  ---b4---              -----

Human
Thrombin      LDKIYIHPRY  NWKENLDRDI  ALLKLKRPIE  LSDYIHPVCLP
Batroxobin    KEKFI*CPNK  KKNVITDKDI  MLIRLDRPVK  NSEHIAPLSLP
              b5----                  --b6--

Human
Thrombin      DKQTAAKLLH  AGFKGRVTGWG  NRRETWTTS  VAEVQPSVLQ
Batroxobin    SNPPS***VG  SVC**RIMGWG  AITT**SED  T*YRDVPHCA
                           ---b7----                    ---

Human
Thrombin      VVNLVERPVC  KASTRIRITN  DMFCAGYKPGE  GKRGDACEGD
Batroxobin    NINLFNNTVC  REAYNG*LPA  KTLCAGVL**Q  GG*IDTCGGD
              b8-   ----a1-----        --b9--

Human
Thrombin      SGGPFVMKSP  YNNRWYQMGI  VSWGFNPCGEP  KKPGIYTKLIR
Batroxobin    SGGPLIC***  **NGQFQ*GI  LSWGSDPCAEP  RKPAFYTKVFD
                 ----b10----- ----b11----            -b12-- --

Human
Thrombin      LKKWIQKVID  RLGS
Batroxobin    YLPWIQSIIA  GNKTATCP
              -----a2-------
```

Figure 5.10: Alignment of amino acid sequences of batroxobin and α-thrombin. Any stars represent gaps between the sequences. Residues shown in bold are part of the catalytic triad. Regions of α-helix are shown as a1-a2 and β-sheets are represented by b1-b12.

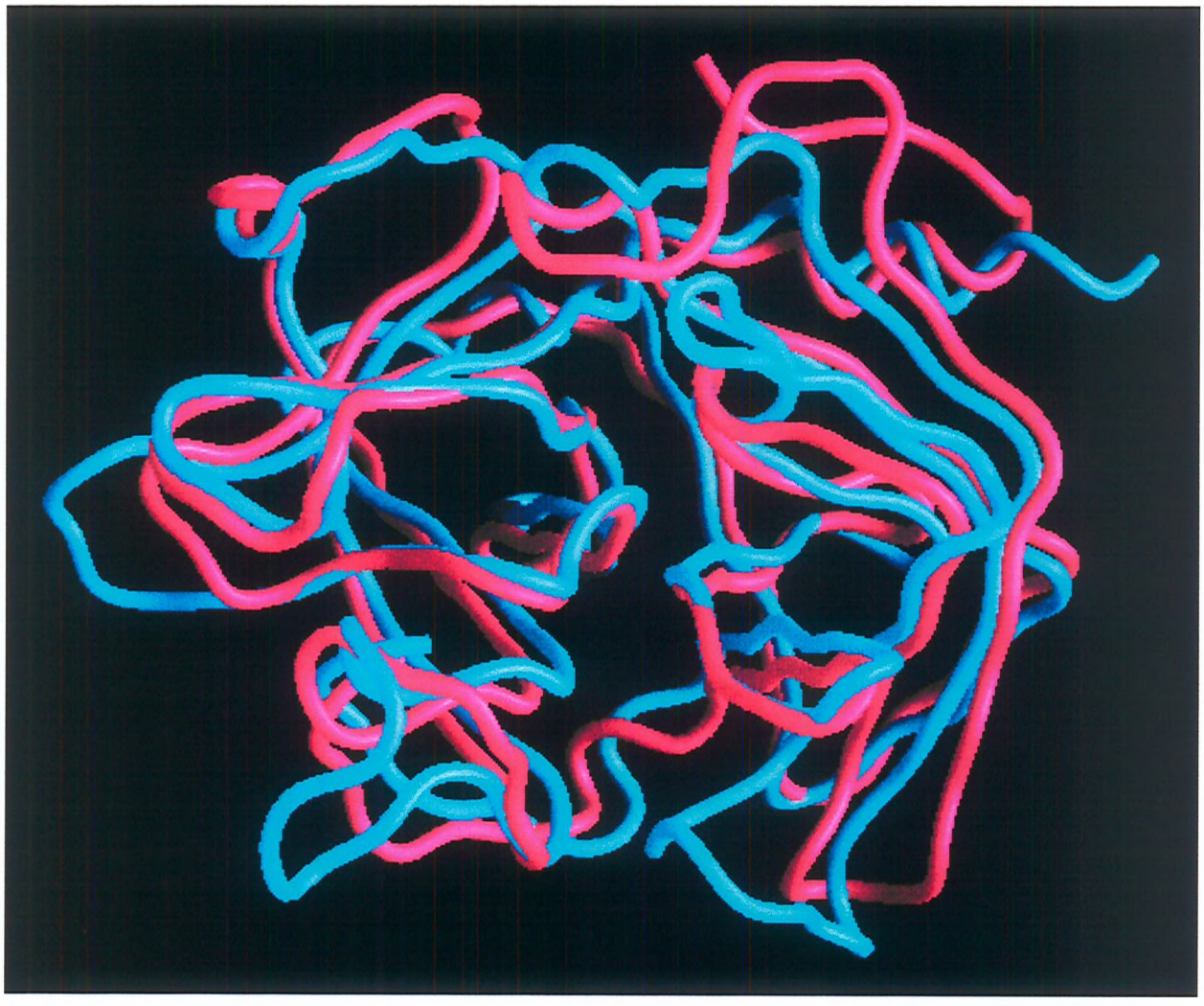


Figure 5.11: Schematic representation showing the backbone of α-thrombin superimposed over the backbone of batroxobin modelled on porcine kallikrein and TSV-PA.

structure they appear to be virtually identical, as seen in *figure 5.11*. All the major areas of the structure are highly conserved, even within areas where the sequences are not. The two α- helices are conserved, as are the two β sheeted barrels along with the whole of the core structure surrounding the active site cleft. This lends support to the view that the serine proteinases are a group of enzymes that show an extremely high degree of conservation of the structural integrity of the protein, whilst only needing to retain the most important catalytic and substrate-binding residues to maintain specificity.

5.6.2 The β-sheeted barrels.

The results of modelling batroxobin on porcine kallikrein and adapting the structure to incorporate the features thought to be conserved in snake venom proteases from the TSV-PA structure produced very similar β sheeted barrel regions to those found in α-thrombin. These regions, shown in blue on the sequence alignment *figure 5.9*, appear to be highly conserved amongst all the serine proteases. It is interesting to note the similarity between the batroxobin model and α-thrombin at this point since the model was not produced from α-thrombin. This provides additional evidence that this area is highly conserved throughout the group of enzymes.

5.6.3 Disulphide bridges.

α-Thrombin has three disulphide bridges Cys28-Cys44; Cys173-187 and Cys201-Cys-231 compared to batroxobin, which has six. The six disulphide bridges correlate to the five found in glandular kallikrein Cys7-Cys139, Cys26-Cys42, Cys118-Cys184, Cys150-Cys163 and Cys174-Cys199 along with one extra which is between the C-terminal extension from Cys74-Cys230 as found in TSV-PA. It is this extra disulphide bridge that results in the connection of the C-terminal tail to the core of the enzyme, as predicted earlier, that appears to be characteristic of snake venom proteases. Despite the fact that there are only three disulphide bridges within the α-thrombin structure compared to the six formed on the batroxobin model, the overall structure and shape between the two enzymes is well conserved.

5.6.4 The external loops.

The major differences between the structures of α-thrombin and the batroxobin model lie in external loop regions rather than the core areas of the enzymes that seem to be highly conserved throughout the group. The 99 loop of the batroxobin is about five residues longer than that found in thrombin. This extension means that the loop bordering the active site lies closer to His57, part of the catalytic triad, and it may be that this area is significant in the increased specificity demonstrated by batroxobin over thrombin. It also appears possible that this region may be responsible for the relative resistance of the snake venom proteases to inhibition by larger molecular weight inhibitors like bovine pancreatic trypsin inhibitor (BPTI). This finding is consistent with the crystallographic evidence from the structure of the closely related TVS-PA [101]. Batroxobin, unlike thrombin, does not appear to have the extended loop region found at the entrance to the active site cleft 60a-60f, this numbering is as described in section 5.5. This loop is of significance in thrombin, as it is believed to constitute part of the fibrinogen-binding site. In the batroxobin model this extension is not present. This is also consistent with evidence from structural and modelling studies on crotalase, a snake venom thrombin-like enzyme derived from *Crotalus adamanteus* [102]. Where the equivalent site is not present but an alternative-binding region of Arg60f, Lys85, Lys87 and Arg107, for fibrinogen binding within this enzyme is suggested. Closer inspection of the batroxobin structure suggests that this region is also present in the batroxobin structure, and that this additional binding site may be significant in the increased specificity for thrombin demonstrated by batroxobin over thrombin. It is interesting to note that the exosite binding domain in thrombin is not specific for fibrinogen, but is also thought to recognise sequences within thrombin receptors and protein C. It appears likely that the exosite suggested in crotalase, and which may also exist in batroxobin, is specific for fibrinogen and may therefore account for the lack of reactivity of these enzymes to other substrates than fibrinogen. It does not, however, account for the increased specificity of batroxobin within the fibrinogen structure.

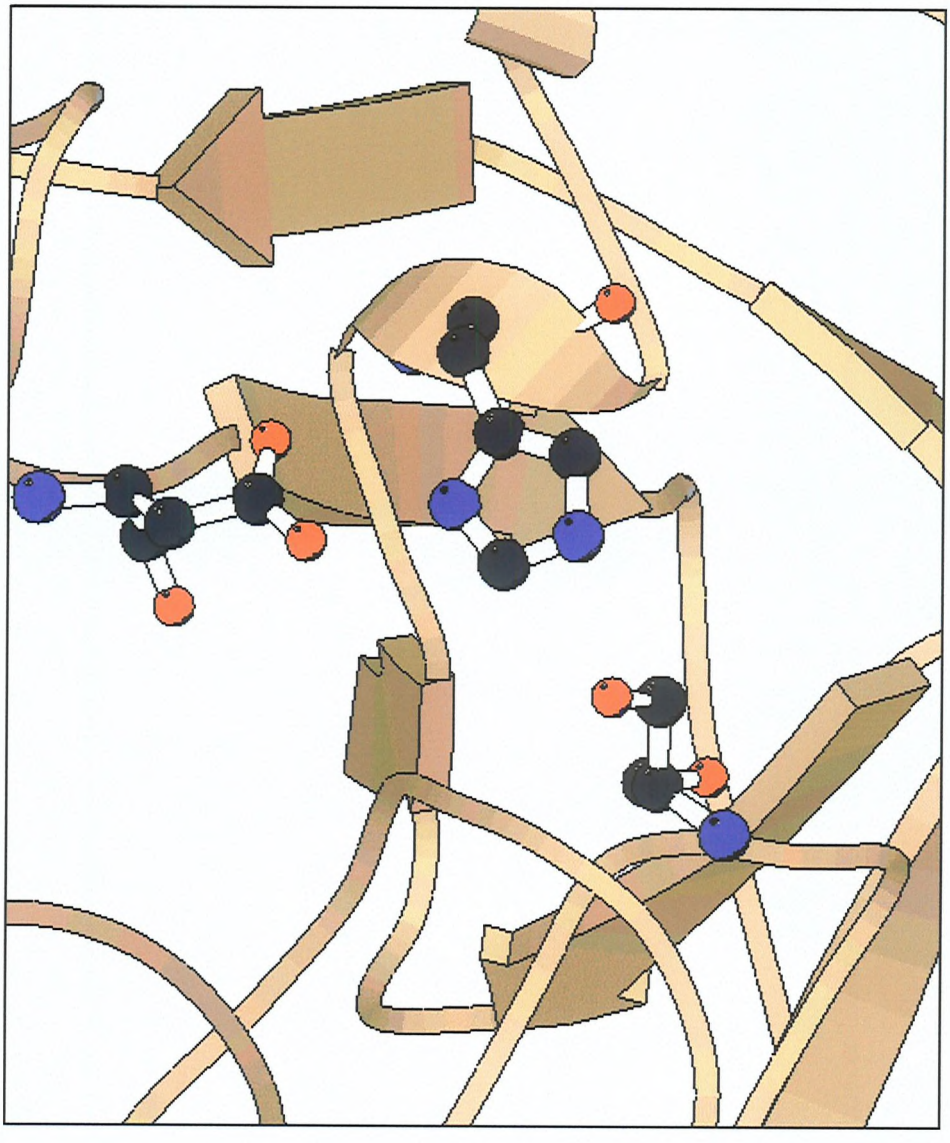


Figure 5.12: Active site region of the model batroxobin structure. From left to right: Asp102(86), His57(41) and Ser195(178). Numbering is based on chymotrypsin (numbering in brackets is the batroxobin number).

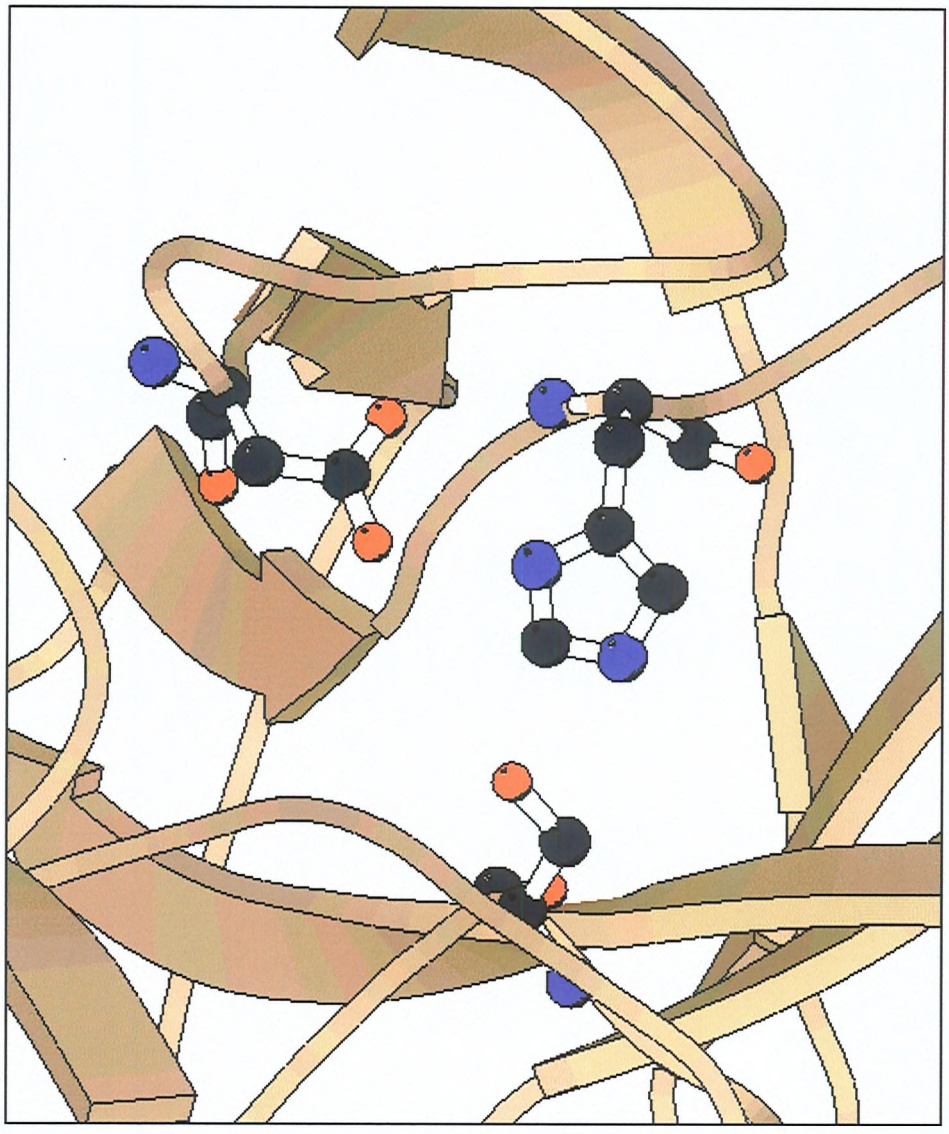


Figure 5.13: Active site region of α-thrombin. The top residues are from left to right Asp102 and His57. The bottom residue is Ser 195. All numbering is based on chymotrypsin.

5.6.5 The active site region.

As with all the serine proteases, the catalytic triad as identified in chymotrypsin as containing His57, Asp102 and Ser195 and is conserved in both thrombin and the batroxobin model. As can be seen from the two active site structures, *figure 5.12* and *figure 5.13*, batroxobin and α-thrombin are virtually identical in structure at this region. In both cases the charge relay mechanism can be set up between the Asp102 and His57 and the His57 and the Ser195, although in the batroxobin model the distances between Asp102, His57 and Ser195 are longer than those found in thrombin.

5.6.6 Specificity binding pockets.

The majority of the substrate residue binding sites surrounding the catalytic serine of thrombin are remarkably well conserved in batroxobin. The S1 and S1' sites are totally conserved between the two enzymes accounting for their similar specificity for arginine at P1 and glycine at P1' site. The other areas where substrate binding occurs are very similar between thrombin and batroxobin, with the exception of the S3 site and the S2' site which shows differences which will be discussed further later.

5.7 Substrate modelling.

One of the major aims of this modelling work was as an attempt to explain the specificity of batroxobin for the synthetic peptides used in the studies in Chapter 4 and the natural substrates used in Chapter 3. It was therefore decided that each of the peptide substrates should be modelled into both thrombin and batroxobin in order to look for any differences between the binding sites of the two enzymes. It was hoped that these could possibly account for the differences in specificity exhibited by thrombin and batroxobin for both the natural and synthetic substrates. Initially, it was necessary to find a template which could be used to allow the peptides to be modelled as accurately as possible into the active site cleft of each of the enzymes. Two potential templates were found; α-thrombin co-crystallised with the short peptide sequence [103] and α-thrombin co-crystallised with rhodonin[104]. Rhodonin is a

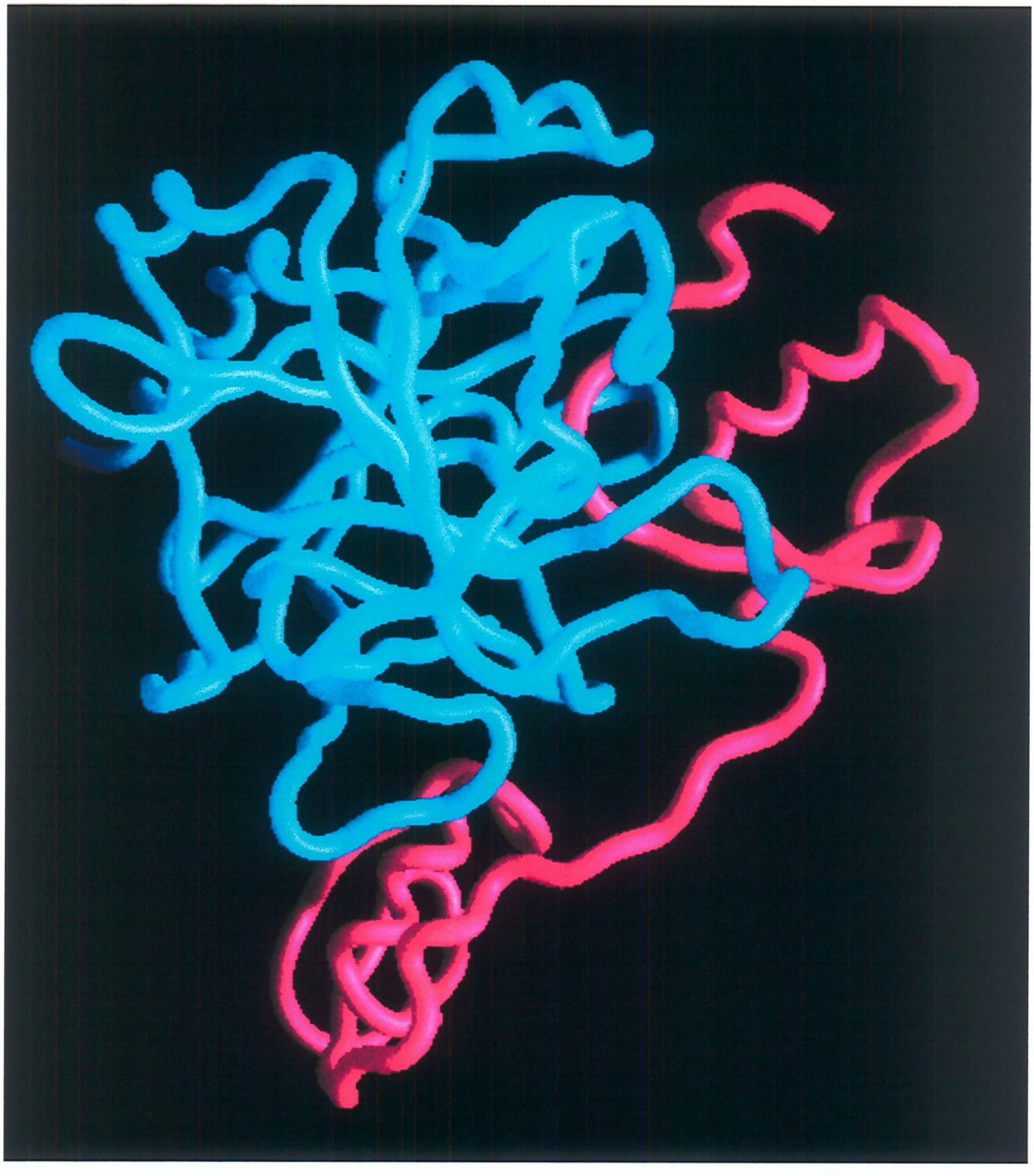


Figure 5.14: Schematic representation of the backbone of α-thombin with the thrombin specific inhibitor rhodonin[102], used as a template for the modelling of thrombin and batroxobin with the synthetic peptides.

specific thrombin inhibitor derived from the mosquito *Rhodnius Prolixus*. The insect uses this inhibitor to prevent blood coagulation at the site of insertion of the proboscis, enabling it to remove blood from its target animal.

Both of these structures were used. The former allowed the first five residues in the decapeptide sequence to be assigned unambiguously as they were also present in the template structure. This enabled the arginine in the S1 site to be positioned correctly. However, since the arginine was the "last" residue co-crystallised, then the rhodonin structure was needed to model remaining amino acids from the P1' to the P5' position.

5.7.1 Batroxobin modelled with the decapeptide GGGVRGPRVV at the active site.

Initially, the decapeptide GGGVRGPRVV-amide was modelled into the thrombin and batroxobin structures. This was chosen as it corresponds to the sequence found around the scissile bond in fibrinopeptide A, as described in Chapter 4. The modelling was carried out as previously outlined with continuous refinement of the structure to ensure a favourable model. This process prove to be difficult as it was important that the peptide was positioned in the correct place compared to the two templates but without any side chain clashes. Care needed to be taken to ensure that any allowed interactions in the form of hydrogen bonds between the peptide and the enzyme were formed as these may be involved in stabilisation of the substrate, but that no unfavourable interactions were present.

Once the peptide had been satisfactorily modelled into the enzyme, the whole structure was again energy minimised to refine the structure. As can be seen from *figures 5.15, 5.16* and *5.17*, the peptide enters the enzyme from the bottom of the active site cleft and continues up to the arginine at the P1 site where it is bound into a deep cleft. At the P1' glycine, the peptide bends back on itself, before continuing out of the bottom of the active site cleft. *Figures 5.16 and 5.17* show how the peptide protrudes out of the front of batroxobin, as the active site cleft is not completely surrounded by the external loops.

Figure 5.15: Schematic representation of the backbone of the batroxobin model, with the synthetic peptide GGGVRGPRVV bound to the active site cleft shown in yellow.

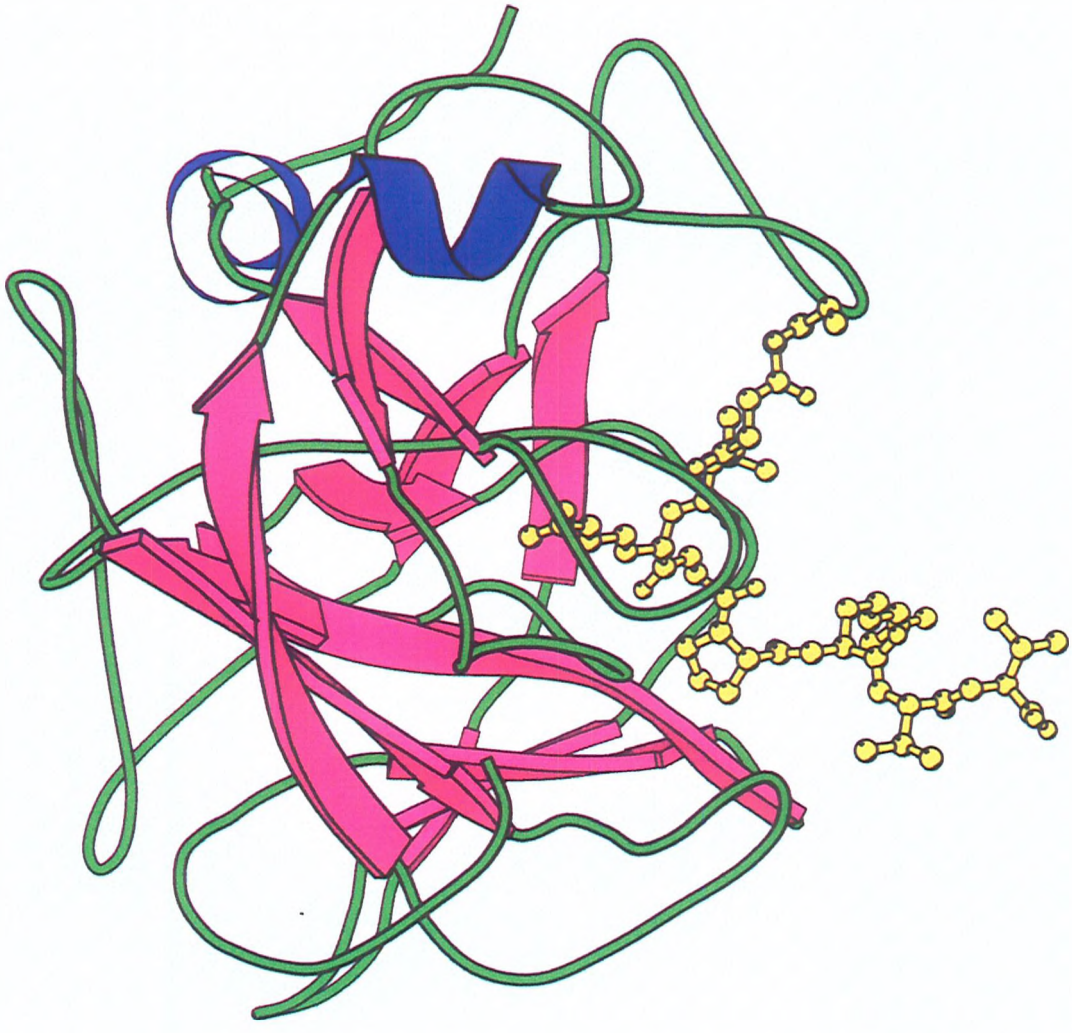


Figure 5.16: Schematic representation of the backbone of the batroxobin model with the synthetic peptide GGGVRGPRVV bound shown in yellow, rotated through 90°

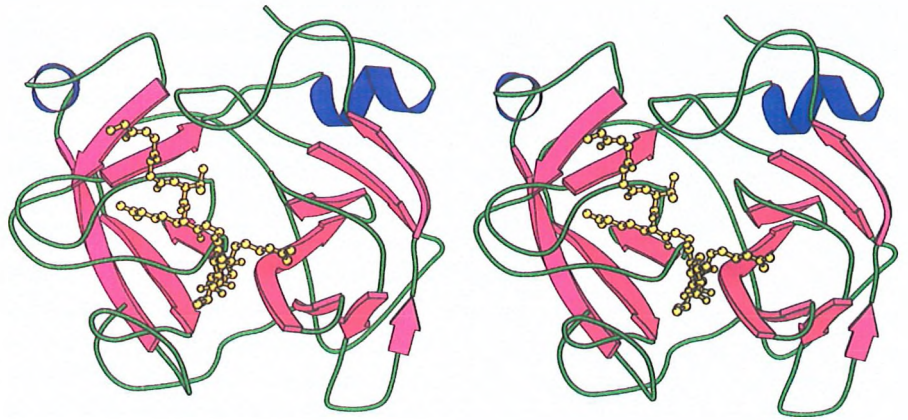


Figure 5,17: Stereoimage of the backbone of batroxobin with the synthetic peptide GGGVRGPRVV bound to the active site cleft shown in yellow

The binding of this substrate (GGGVRGPRVV) to batroxobin is very similar to that with thrombin. Like thrombin, the external residues glycine(P4 and P5) and valine(P4' and P5') do not appear to be bound to the enzymes at all. It is therefore unlikely that these areas are of any great importance in terms of recognition of the substrate, as it would appear that any residue that is small enough to fit into these sites would be acceptable to the enzyme. This probably helps to explain why the increase in the rate of cleavage of the scissile bond is not significantly increased by the extension from the hexapeptide GVRGPR to the decapeptide GGGVRGPRVV. In addition, since the kinetic data obtained from the experiments on the natural substrates in Chapter 3 and the synthetic substrates in Chapter 4 show vastly increased rates of reaction for fibrinogens compared to the peptides, then it would appear evident that other factors were involved in substrate recognition than the binding in the active site cleft alone. Consequently it is important that this modelling work should not only look for an explanation within the active site but should also attempt to find a fibrinogen-binding site for batroxobin that could help to explain some of the specificity.

5.7.2 Comparison of the binding around the scissile bond in α-thrombin compared to the model of batroxobin.

As previously mentioned the catalytic triad within the active site of α-thrombin and the model of batroxobin are almost identical. This is also true of the general binding and positioning of the substrate to the active site cleft around the scissile bond. The catalytic triad of His57, Asp 102 and Ser195 remain set up in close proximity to each other in order to allow the formation of the highly nucleophilic Ser195. This serine lies close to Arg (P1) of the substrate, thus enabling the nucleophilic attack by serine on the carbonyl carbon of the Arg (P1) to Gly (P1') peptide bond. In both enzymes the correct positioning of the substrate in enabled by the stabilising effect of two conserved residues, Gly196 forms a hydrogen bond between the main chain NH group of the glycine and the main chain carbonyl of Arg (P1) of the substrate. The other conserved environment brings about the preference for arginine at this position in both enzymes; an aspartic acid (Asp192) at the bottom

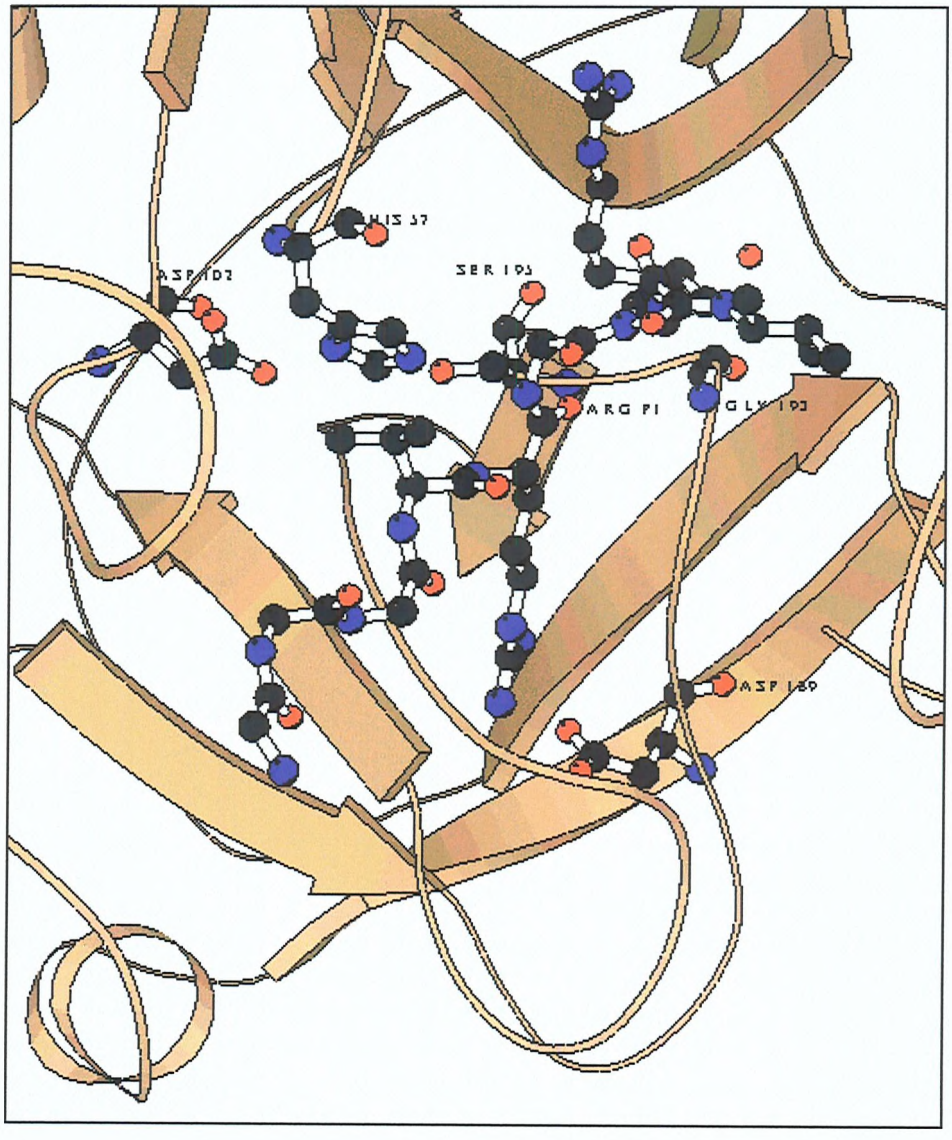


Figure 5.18: The binding of the substrate to the batroxobin model around the scissile bond.

of the S1 specificity holds the side chain of the arginine in place. These observations can be seen from *figure 5.18.* At this area, as with most of the binding sites, there are very small differences between α-thrombin and batroxobin. However, there are two areas that appear to show differences between the two enzymes, which may help to explain the specificity shown between the two. These include the regions at the S3 and S2' binding sites.

5.7.3 Comparison of substrate binding at the S3 site between α-thrombin and the batroxobin model.

Experimental evidence obtained from the reactions of the natural fibrinogen substrates and the synthetic substrates with both enzymes indicate that batroxobin has a preference for an aspartic acid at the P3 site. This is an interesting observation since most fibrinogens have a glycine at this residue, with rat and mouse fibrinogens the notable exceptions. This is of interest since these animals are the natural prey of the pit vipers and hence it would appear that there could be some evolution of specificity for this sequence within batroxobin.

It was therefore of value to try and explain this observation on a molecular level from a comparison of the thrombin and batroxobin binding at this point. The model of batroxobin indicates that there are a large number of positively charged residues surrounding this area. It is therefore not suprising that batroxobin might favour a negatively charged residue like aspartic acid in this position. In addition to this general environmental feature, modelling studies show that batroxobin has a lysine residue (lys95) in very close proximity to the P3 residue of the substrate. This lysine, which is positively charged is potentially capable of forming a stabilising hydrogen bond with an aspartic acid in the P3 position. These observations may provide an explanation for the preference of this enzyme for a negatively charged residue at this position. This can be seen in *figure 5.19.*

On the other hand, α-thrombin is not as positively charged generally around this area, and the overall environment is much more neutral. α-Thrombin is also not

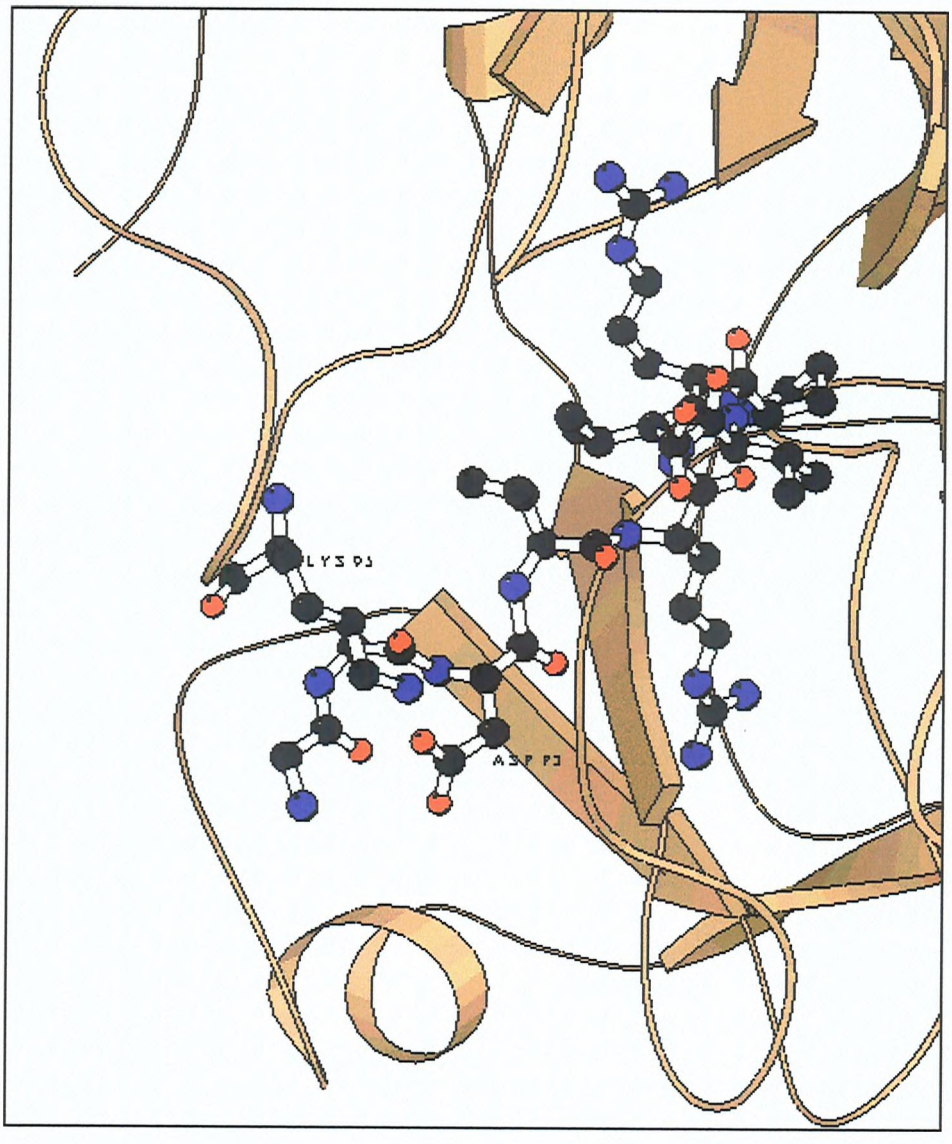


Figure 5.19: The binding of aspartic acid (P3) within the model batroxobin structure.

capable of forming a stabilising hydrogen bond at this position as it does not have the positively charged Lys95, but rather has a tryptophan at this position. Interestingly, both enzymes have enough space for either glycine or aspartic acid at this position but batroxobin with its positive charged environment has a preference for aspartic acid and thrombin with its neutral environment has a preference for glycine. This may help to explain why both enzymes can tolerate either glycine or aspartic acid in this position as substrates. It also provides a reasonable explanation for the preference of batroxobin for aspartic acid and α-thrombin for glycine.

5.7.4 Comparison of substrate binding at S2' between α-thrombin and the batroxobin model.

Experimental evidence obtained from the synthetic peptide experiments in Chapter 4 and the documented observation that batroxobin is not capable of cleaving fibrinopeptide B, seems to indicate that a proline residue at P2' is vital for recognition by batroxobin. Previous work suggests that if proline is substituted into the fibrinopeptide B sequence SARGHR, in place of the histidine, then some of the activity can be recovered, whereas thrombin is capable of cleaving any of the sequences to some extent. It therefore appears likely that although both enzymes have a very strong preference for proline at this position, a difference must exist between the two enzymes at this point, which will account for the requirement of this residue for batroxobin recognition.

On first consideration of the S2' binding site of both enzymes, it appeared that the only real reason why proline is required at this position was that in both enzymes a change to histidine slightly alters the conformation of the main chain of the substrate, resulting in a very slight change in the position of the scissile bond of the substrate. This seemingly minor alteration in structure may well help to explain why substrates with a proline at this position are so much better substrates for both enzymes than substrates with histidine present. This is due to the fact that this area is of major importance as the exact angle is vital for catalysis. However, this does not help to explain the differences seen between batroxobin and α-thrombin.

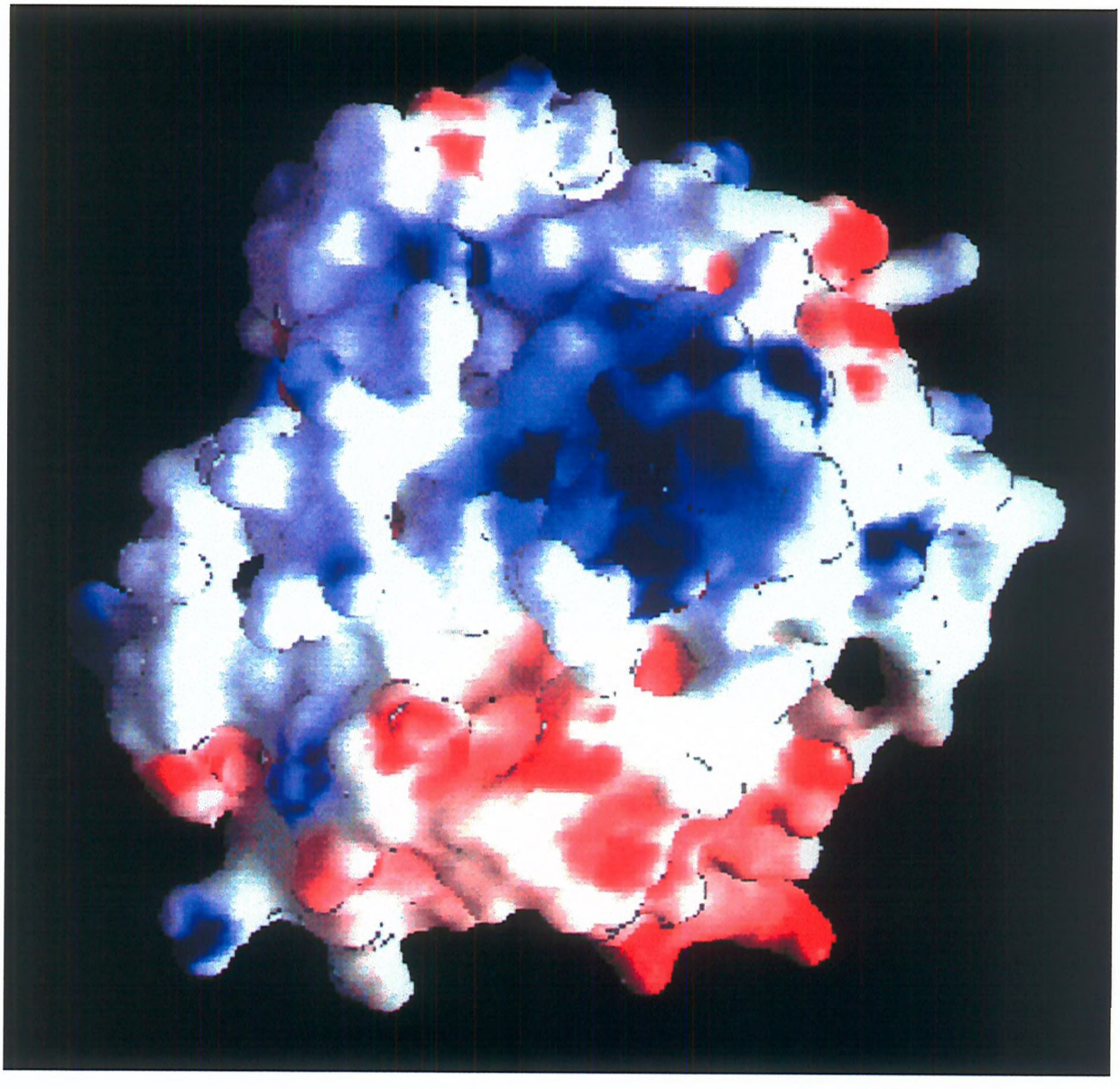


Figure 5.20: Surface potential representation of the batroxobin model. Areas in blue show positive charge and areas of red show negative charge. The main area of blue contains Arg60F, Lys85 and Lys87 and is also the point of binding of P2' residues.

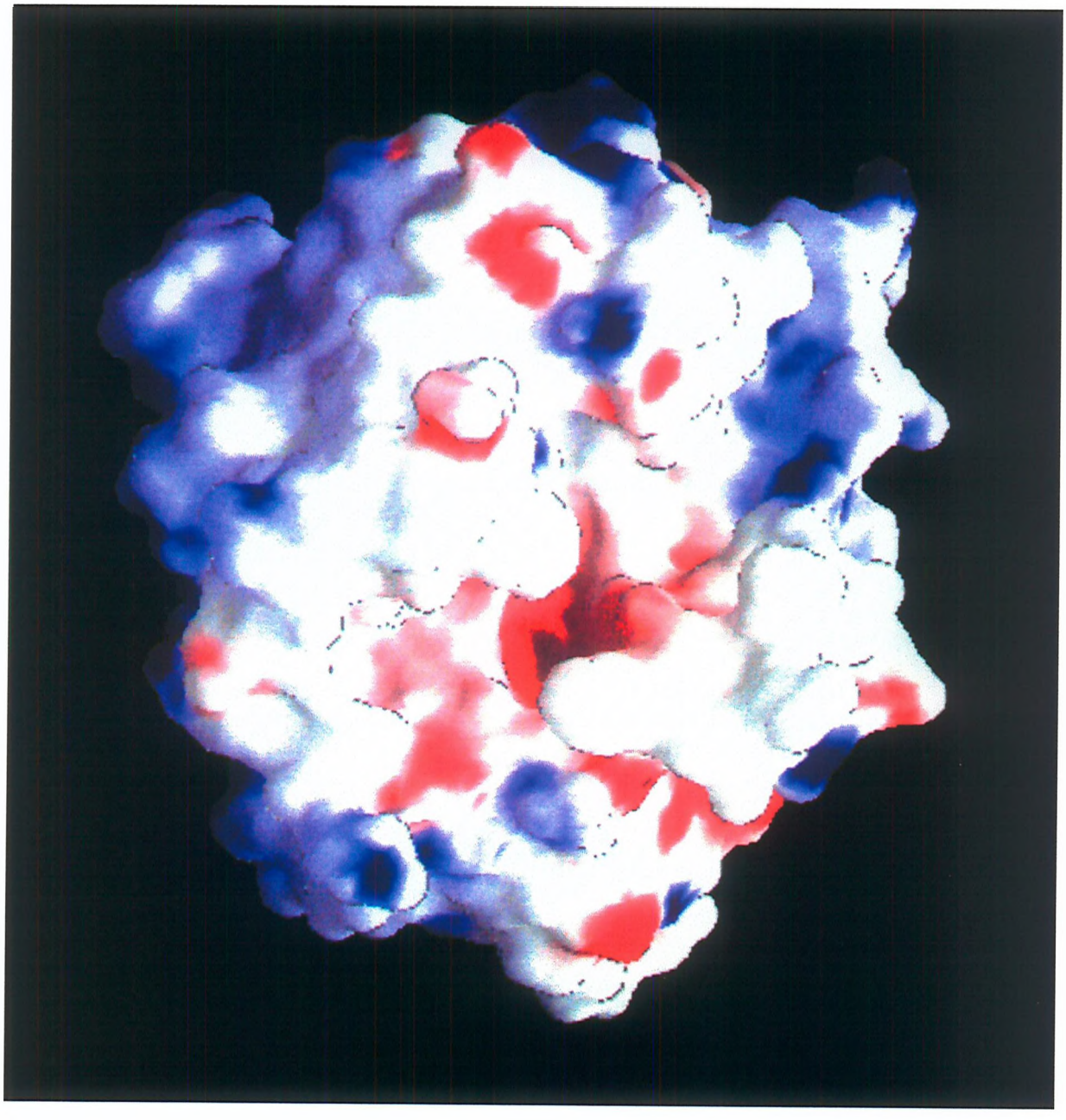


Figure 5.21: Surface potential representation of α-thrombin. Blue regions are areas of positive charge and red are negative. The large area of white in the centre of the structure is the area of binding for P2' residues, it also contains the 60A to 60F thrombin loop.

On closer examination it appeared that the environment around this position in batroxobin is positively charged. Therefore it was decided to perform a calculation of the surface potential of the two enzymes to try and find an explanation for the experimental observations. As can be seen from figures *5.20* and *5.21* there is a vast difference between the surface potentials of the two enzymes. Interestingly, batroxobin has one very large area of positive charge, compared to α-thrombin, which appears to have charge more equally distributed across the whole structure. The most interesting observation is that the area on the batroxobin which shows a very high positive charge is also the area to which the proline in the P2' position binds. It appears very likely therefore that batroxobin is incapable of recognising histidine at this position as histidine is positively charged and the massive charge on the enzyme surface is likely to repel any histidine containing peptide and reduce binding capacity.

5.7.5 Comparison of the fibrinogen binding sites in α-thrombin and batroxobin.

Although these areas of substrate binding within the active site cleft are of great interest in helping to explain the increased specificity of batroxobin for fibrinogen, other potentially important areas have come to light during the course of this modelling work. On close comparison of the batroxobin model to α-thrombin it became apparent that batroxobin did not have the extended "thrombin loop" from Leu60a to Lys60f. This region of thrombin is implicated in substrate recognition, [105] including the recognition of fibrinogen, and is completely missing in batroxobin. This observation is supported by the structure of TSV-PA which is also missing this area [101], and sequence and modelling studies carried out on crotalase [102].

Upon completion of the surface potential map of thrombin and batroxobin an area of high positive charge was found on batroxobin, as seen in *figure 5.20*. Areas of high charge on the surface of proteins are often found to be involved in recognition and binding of substrates. The region on batroxobin contains residues Arg60f, Lys85 and Lys87 which have previously been put forward as part of a fibrinogen recognition site in crotalase, another snake venom derived thrombin-like enzyme [102]. It appears likely that this region on batroxobin, which is conserved only in thrombin-like snake venom proteases and which has the characteristic highly charged region is a specific

fibrinogen recognition site. It is entirely possible that this region, unlike the thrombin loop in thrombin, is specific only for fibrinogen and helps to account for the increased specificity of batroxobin over thrombin.

5.8 Discussion.

The batroxobin model finally arrived at was based on a combination of porcine pancreatic kallikrein and TSV-PA and had a structure that helps to explain many of the characteristics associated with snake venom-derived serine proteases. For this reason, it is believed that, although only a model, the structure is a reasonably accurate representation of batroxobin. These structure exhibits the characteristic C-terminal tail associated with snake venom derived enzymes. The tail is bridged back into the core of the protein as a disulphide bond, characteristic of this group of enzymes. This feature is thought to be involved in contributing to the great specificity demonstrated by snake venom derived proteins for their substrates. This characteristic of batroxobin is shared by crotalase, ancrod and TSV-PA, to name but a few.

The new model structure of batroxobin was used to explain the unique specificity exhibited by batroxobin. As previously discussed in this chapter, the extended tail region when linked through the C-terminal disulphide bridge to the body of the enzyme results in some alterations in the external loop regions of the structure. It seems entirely plausible that this change from the structures found in both thrombin and kallikrein may provide evidence of some occlusion of the active site as demonstrated with TSV-PA[101], which may account for some of the increased specificity of batroxobin for fibrinogen. Upon modelling of the relevant peptides into the active site of the model structure, it was possible to find satisfactory explanations for the experimental observations, discussed in Chapters 3 and 4. These explanations are not definitive as they are derived only from a model structure; however, they do provide a basis for further investigation of batroxobin specificity.

The work carried out showed that the batroxobin model is remarkably similar to thrombin in the region of the catalytic cleft and substrate binding pockets. This is particularly interesting, since batroxobin exhibits a much greater specificity than thrombin, capable of interacting with many different substrates. During the course of this work, it was possible to identify a region of very high positive charge around the

catalytic cleft of batroxobin. As previously discussed, this region is very different to that found in the equivalent location of thrombin and is thought to provide a possible explanation for the loss of recognition of substrates containing histidine at the P2' site instead of proline, (GVRGHR and GVRGPR), as demonstrated experimentally in Chapter 4. This lack of recognition is due to the fact that the highly positive charge is unable to interact with the positively charged histidine residue in the substrate.

Another region in the active site cleft in the batroxobin model, that was found to differ from the thrombin structure, was around the S3 binding pocket of the enzymes. This area in thrombin was shown to be largely neutral, compared to the positively charged region in batroxobin. In particular, one residue Lys 95 was found to be capable of forming a stabilising interaction when aspartic acid was substituted for the glycine in the P3 position (GGDVRGPRVV compared to GGGVRGPRVV). The stabilising bond formation was not possible in thrombin, which contains a tryptophan at the same position. This provides a possible explanation for the preference exhibited by batroxobin for substrates with aspartic acid in the P3 position compared to thrombin which has a preference for the neutral glycine residue at this position. This helps to explain many of the experimental findings arising from specificity studies carried out on the natural substrate fibrinogen (Chapter 3) and the synthetic peptides (Chapter 4).

Finally, it was also possible during the course of the modelling work to identify a region on the surface of the batroxobin that may be responsible for the recognition of the fibrinogen substrate. This region is of very high positive charge, as shown in *Figure 5.20*, and correlates well with a region also identified in crotalase as a possible fibrinogen-binding site. This region, at the opening to the active site cleft is extremely interesting and important, as the fibrinogen binding site found in thrombin does not exist in batroxobin model and therefore another region is likely to be responsible for recognition.

In summary, these regions identified as being of potential importance within batroxobin, both in terms of the overall recognition of the fibrinogen substrate and sequence preferences around the scissile bond, provide a useful starting point for future study. The batroxobin model has suggested a number of future experiments with thrombin mutants at key residues, specificity studies with the synthetic peptides and crystallographic studies of the wild type and mutated forms, that the actual role of

these areas can be examined more fully. Hopefully, in this way the definitive reasons for the unique specificity exhibited by batroxobin can be found.

SUMMARY AND FUTURE WORK

Summary and Future Work.

The work that has been carried out as part of this project has allowed some of the questions surrounding the relative specificities of thrombin and batroxobin to be addressed. It has also given some insight into the structure of batroxobin and how variations in structure may be implicit in the differences between the two enzymes.

The work carried out in Chapter 3 showed that the coagulation process with batroxobin has an absolute requirement for Ca^{2+} as previously described for thrombin. However, it was only when the requirement for Ca^{2+} was examined with the synthetic peptides in Chapter 4 that it could be conclusively ascertained that this requirement was due to its action as an accelerant of fibrin polymerisation, and not due to some interaction with the batroxobin itself. This was demonstrated by its lack of effect on the cleavage of synthetic peptides and supports totally its role in coagulation with thrombin.

The comparison of the actions of thrombin and batroxobin on various natural substrates in Chapter 3 showed a marked difference between the two enzymes. The human thrombin used in the study, perhaps not surprisingly, exhibited higher activity for human fibrinogen than for any of the others. However it was also very active with bovine and porcine fibrinogen, showed good specificity for rabbit, cat and dog fibrinogens, but had very poor recognition and activity towards rat and mouse fibrinogens. Conversely, batroxobin had much higher activity towards rat and mouse fibrinogens with all other species exhibiting a similar but much lower activity. On closer inspection of the fibrinogens it became apparent that rat and mouse fibrinogens had sequence variations from the others around the scissile bond, with an aspartic acid residue instead of the more normal glycine at the P3 position. It therefore appeared possible that batroxobin has a preference for aspartic acid at this site, compared to the preference for glycine exhibited by thrombin.

In order to assess the validity of this proposal and to attempt to map the recognition sites in the active site cleft of batroxobin, a series of peptides were designed and synthesised using solid phase peptide chemistry. This work, shown in Chapter 4, indicates that as suggested the aspartic acid residue at the P3 position is indeed significant. Although it is not required for batroxobin recognition it does appear to be preferred. This work also showed that batroxobin has an absolute

requirement for proline at the P2' site, which helps to explain why batroxobin is incapable of cleaving fibrinopeptide B, which has a histidine at this position. Thrombin is capable of recognising sequences with histidine at the P2' position, although it does have a marked preference for proline at this position. This correlates well with the physiological ability of thrombin to cleave both fibrinopeptides A and B, but with a much slower cleavage of fibrinopeptide B.

In an attempt to find plausible explanations for the actions observed homology modelling was undertaken. This work, in Chapter 5, shows a model structure of batroxobin based on a combination of porcine pancreatic kallikrein and TSV-PA, a snake venom derived serine proteinase. Modelling studies of relevant peptides within the active site show reasons why batroxobin has different actions to thrombin. The preference for proline (P2') could be accounted for by either the very highly positive charge of batroxobin around this area repelling the positively charged histidine or by the slight change in conformation of the scissile bond brought about by a change to histidine. The preference exhibited by batroxobin for aspartic acid at P3 is likely to be explained by the ability of Lys 95 to form a stabilising bond with aspartic acid, which not possible with glycine. This interaction is not possible in thrombin.

In addition to these findings, both batroxobin and thrombin were found to require the carbonyl group of the peptide bond for recognition of the substrate. This can be seen in the inhibitor studies of Chapter 4, suggesting that this position is not only required for catalysis but also for recognition. It was also possible to identify from the modelling studies of Chapter 5 a possible exosite for the recognition of fibrinogen by batroxobin. This is of huge interest as its existence has often been postulated, but it has never to date been identified.

It is therefore evident that this project has opened up the work on batroxobin and other snake venom derived thrombin-like enzymes for much more investigation. It would be interesting to look at mutagenesis studies of thrombin and batroxobin at key residues to assess the overall validity of the proposals for specificity put forward in this work. However, this will not be possible until a fully functioning recombinant batroxobin can be produced. It would also be useful to look at the specificity of rat and mouse thrombin for various fibrinogens and synthetic substrates to see if the observations found with batroxobin hold true for them also, with specific reference to the preference for aspartic acid at the P3 position.

A continuation of the crystallisation studies for batroxobin as detailed in Chapter 3 is desirable to provide a definitive structure upon which the specificity of the enzyme can be more fully investigated. It is only in this way that the accuracy of the proposed structure in this project can be ascertained. In addition to this work on batroxobin, it would be advantageous to consider some of the other snake venom derived thrombin-like enzymes in this way. This would allow a greater overall understanding of this group of enzymes, particularly with reference to their unique specificity. Finally, it is hoped that the further probing of these enzymes could lead in the future to the development of a novel thrombin, with some of characteristics of increased specificity shown by batroxobin to be used in the treatment of thrombosis. It is hoped that this may help to reduce some of the antigenic effects observed in patients treated with snake venom proteins.

References

[1] Meier, J. and Stocker, K. *Critical Reviews in Toxicology* (19), **21(3)**, 171-182

[2] Pirkle, H. and Stocker, K. *Thrombosis and Haemostasis* (1991), **65(4)**, 444-450

[3] Stocker, K. In: *Medical Use of Snake Venom Proteins*, Chapter 3, CRC Press

[4] Burne, R. M. and Levy, M. N. *Physiology*, Second Ed., (1988), C. V. Mosby Co., St. Louis

[5] Stocker, K. In: *Medical Use of Snake Venom Proteins*, Chapter 5, CRC Press

[6] Ogston, D. In: *The Physiology of Haemostasis*, (1983), Croom Helm, London and Canberra

[7] Rawn, J. D. *Biochemistry* (1989), Neil Patterson publishers

[8] Burne, R. M. and Levt, M. N. *Physiology*, Second Edition (1988) C. V. Mosby Co., St. Louis

[9] Jackson, C. M. *Ann. Rev. Biochem.* (1980) **49**, 765-811

[10] Sanders, A. G. In: *The Haemostatic Mechanism in Man and Other Animals* Ed. R. G. Macfarlane (1970) 109-120, Ac. Press, London

[11] Sixma, J. J. In: *Haemostasis and Thrombosis* Eds. A. L. Bloom and D. P. Thomas (1981) 253-257, Churchill Livingstone, Edinburgh

[12] Ratnof, O. D. and Rosenblum, J. M. *Am. J. Med.* (1958) **25**, 160-168

[13] Griffin, J. H. and Cochrane, C. G. *Proc. Natl. Acad. Sci. US* (1976) **73**, 2554

[14] Ratnof, O. D. *Science* (1968) **162**, 1007-1009

[15] Macfarlane, R. G. *Nature* (1964) **202**, 498-499

[16] Davie, E. W. and Ratnoff, O. D. *Science* (1964) **145**, 1310-1312

[17] Macfarlane, R. G. *Thromb, Diath. Haemorrh.* (1965) **15**, 591-602

[18] Ratnof, O. D. and Saito, H. *Proc. Nat. Aced. Sci.* (1979a) **796**, 958-961

[19] Schiffmann, S. and Lee, P. *Brit. J. Haem.* (1974) **27**, 101-114

[20] Griffin, T. H. and Cochrane, C. G. *Proc. Nat. Acad. Sci.* (1976)

[21] Kurachi, K. and Davie, E. W. *Biochemistry* (1977) **16**, 5831-5839

[22] Mandle, R and Kaplan, A. P. *J. Biol. Chem.* (1977) **252**, 6097-6104

[23] Wuepper, K. D. and Cochrane, C. G. *J. Exptl. Med,* (1972) **135**, 1-20

[24] Thompson, R. E., Mandle, R. and Kaplan, A. P. *J. Exptl. Med.* (1978) **147**, 488-499

[25] Kerbiriou, D. M. and Griffin, J. H. *J. Biol. Chem.* (1979) **254,** 12020-12027

[26] Fujikawa, K., Thompson, A. R., Legaz, M.E., Meyer, R. G. and Davie, E. W. *Biochemistry* (1973) **13**, 5290-5299

[27] Fujikawa, K., Legaz, M.E., Kato, H. and Davie, E. W. *Biochemistry* (1974b) **13**, 4508-4516

[28] Sakariassen, K. S., Bolhuis, P. A., Sixma, J. J. *Nature* (1979) **279**, 638

[29] Doucet-de Bruine, M. H. M., Sixma, J. J., Over, J. and Beeser-Visser, N. H. *J. Lab. Clin. Med.* (1978) **92**, 96-107

[30] Vehar, G. A. and Davie, E. W. *Biochemistry* (1980) **19**, 401-410

[31] Fujikawa, K., Legaz, M. E. and Davie, E. W. *Biochemistry* (1972) **11**, 4882-4891

[32] Radcliffe, R. D. and Barton, P. G. *J. Biol. Chem.* (1973) **248**, 6788-6795

[33] Chargraff, E., Bendich, A. and Cohen, S. S. *J. Biol. Chem.* (1944) **156**, 161-178

[34] Nemerson, Y and Pitlick, F. A. ***Biochemistry*** (1970) **9**, 5100-5105

[35] Radcliffe, R. and Nemerson, Y. *J. Biol. Chem.* (1975) **250**, 388-395

[36] Radcliffe, R. and Nemerson, Y. *J. Biol. Chem.* (1976) **251**, 4797-4802

[37] Suttie, J. W. and Jackson, C. M. *Physiol. Rev.* (1977) **57**, 1-70

[38] Owen, W. G, Esmon, C. T. and Jackson, C. M. *J. Biol. Chem.* (1974) **249**, 594-605

[39] Fenton, J. W., Fasco, M. J., Strackrow, A. B., Aronson, D. L. , Young, A. M. and Finlayson, J. S. *J. Biol. Chem.* (1977) **252**, 3587-3598

[40] Lollar, P, Hoak, T. C. and Owen, W. G. *J. Biol. Chem.* (1980) **255**, 10279-10283

[41] Doolittle, R. F., Watt, K. W. K., Cottrell, B. A., Strong, D. D. and Riley, M. *Nature* (1979) **280**, 464-468

[42] Hall, C and Slater, H. *J. Biophys. Biochem. Cytology* (1959) **5**, 11-15

[43] Bilezkian, S. B., Nossel, H. L., Butler, V. P. and Canfield, R. E. *J. Clin. Inv.* (1975) **56**, 438-445

[44] Blomback, B., Hessel, B., Hogg, D. and Therkildsen, L. *Nature* (1978) **275**, 501-505

[45] Schwartz, M. L., Pizzo, S. V., Hill, R. L. and McKee, P. A. *J. Biol. Chem.* (1973) **248**, 1395-1407

[46] Pisano, J. J., Finlayson, J. S., Peyton, M. P. and Nagai, Y. *Proc. Nat. Acad. Sci.* (1971) **68**, 770-772

[47] Pizzo, S. V., Taylor, L. M., Schwartz, M. L., Hill, R. L. and McKee, P. A. *J. Biol. Chem.* (1973) **248**, 4584-4590

[48] Crawford, G. P. M. and Ogston, D. *Biochim. Biophys. Acta* (1975) **391** 189-192

[49] Gaffney, P. J. *Thrombosis Research* (1973) **2**, 201-218

[50] Hoylaerts, M., Rijken, D. C., Lijnin, H. R. and Collen, D. *J. Biol. Chem.* (1982) **257**, 2912-2919

[51] Fersht, A. In: *Enzyme Structure and Mechanisms*, Second Ed., (1985), W. H. Freeman and Co., New York

[52] Walsh, C. In: *Enzyme Reaction Mechanisms* (1977), W. H. Freeman and Co., New York

[53] Hartley, B. S. and Kilby, B. A.

[54] Jencks, W. P. *Catalysis in Chemistry and Enzymology* (1969) McGraw-Hill, New York

[55] Kossiakoff, A, A. and Spencer, S. A. *Biochemistry* (1981) **20**, 6462-6474

[56] Shulman, R. G. *Ann. Rev. Biophys. Bioeng.* (1982) **11**, 419-444

[57] Neurath, H. *Science* (1984) **224**, 350-357

[58] Lipscomb, W. N. *Ann. Rev. Biochem.* (1983) **52**, 17-34

[59] Boyer, P. D. *The Enzymes*, Third Edition (1970-1983) **1-15**, Ac. Press, New York

[60] Warshel, A. *Acc. Chem. Res.* (1981) **14**, 284-290

[61] Schechter, I and Berger, A. *Biochem. Biophys. Res. Comm.* (1967) **27**, 157

[62] DiBella, E. E. and Scheraga, H. A. *J. Protein Chem.* (1998) **17(3)**, 197-208

[63] Lochnit, G. and Geyer, R. *Eur. J. Biochem.* (1995) **228**, 805-816

[64] Nahas, L., Kamiguti, A. S. and Alice, M. *Thrombosis and Haemostasis* (1979) **41**, 2

[65] Meh, D. A., Siebenlist, K. R., Bergtrom, G. and Mosesson, M. W. *Blood Coag. Fibrin.* (1993) **4(1)**, 107-112

[66] Meh, D. A., Siebenlist, K. R., Bergtrom, G. and Mosesson, M. W. *Thrombosis Research* (1993) **70(6)**, 437-449

[67] Janahashi, N., Fukuuchi, Y, Tomita, M., Kobari, M., Takeda, H., Yokoyama, M., and Itoh, D. *Clinical Hemorheology* (1995) **15(1)**, 89-96

[68] Marsh, N. A. *Blood Coag. Fibrin.* (1994) **15(1)**, 89-96

[69] DiCera, E., Dang, Q. D., And Ayala, Y. M. *Cell. Mol. Life Sci.* (1997) **53**, 701-730

[70] Mann, K. G. *Am. J. Clin. Nutr.* (1997) **65**, 16575-16645

[71] DiBella, E. E. and Scheraga, H. A. *Biochemistry* (1996) **35**, 4427-4433

[72] DiCera, E., Guinto, E. R., Vindigni, A., Dang, Q. D., Ayala, Y. M., Wuyi, M. and Tulinsky, A. *J. Biol. Chem.* (1995) **270**, 22089-22092

[73] Laemmli, U. K. *Nature* (1970) **227**, 680-685

[74] Webber, K. and Osborn, M. *The Proteins Third Edition* (1975) **1**, 179-223

[75] Bradford, M. *Analyt. Biochem.* (1976) **72**, 248-254

[76] Personal communication from Oxford Bioresearch Labs.

[77] Hollemann, W. H. and Weiss, L. J. ***J. Biol. Chem.*** (1976) **251(6)**, 1663-1669

[78] Needleman and Wunsch (1970)

[79] Dayhoff (1978)

[80] Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meye, E. F., Jnr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasmuri, M. *J. Mol. Biol.* (1977) **112**, 535-542

[81] Steward, J. M. and Young, J. D. *Solid Phase Peptide Synthesis*, Second Edition (1984), Pierce Chemical Co., Rockford, Illinois

[82] Atherton, E. and Sheppard, R. C. *Solid Phase Peptide Synthesis, a Practical Approach* (1989), Oxford University Press

[83] Hageman, T. C. and Scheraga, H. A. *Arc. Biochem. Biophys.* (1974) **164**, 707-715

[84] Liem, R. K. H. and Scheraga, H. A. *Arc. Biochem. Biophys.* (1974) **160**, 333-339

[85] Liem, R. K. H., Andreatta, R. H. and Scheraga, H. A. *Arc. Biochem,. Biophys.* (1971) **68(2)**, 253-256

[86] Lewis, G. *PhD Thesis* University of Southampton (1997)

[87] Rawn, J. D. *Biochemistry* (1989) Neil Patterson Publishers

[88] Dibella, E. E. and Scheraga, H. A. *Biochemistry* (1996) **35**, 4427-4433

[89] DiCera, E., Guinto, E. R., Vindigini, A., Dang, Q. D, Ayala, Y. M., Wuyi, M and Tulinsky, A. *J. Biol. Chem.* (1995) **270(38)**, 22089-22092

[90] Dang, Q. D. And DiCera, E. *Proc. Nat. Acad. Sci.* (1996) **93**, 10653-10656

[91] Dang, Q. D., Guinto, E. R. and DiCera, E. *Nature Biotech.* (1997) **15**, 146-149

[92] Atkins, P. W. *Physical Chemistry* Second Edition (1988) Oxford University Press

[93] Crystal Screen Reagent Formulation, Hampton Research, CA, USA

[94] Personal communication from Prof. S. Wood

[95] Merrifield, R. B. *J. Amer. Chem. Soc.* (1963) **85**, 2149

[96] Liem, R. K. H., Andreatta, R. H. and Scheraga, H. A. *Arc. Biochem. Biophys.* (1971) **147**, 201-213

[97] Powers, J. C., McRae, B. J., Tanaka, T., Cho, K. and Cook, R. R. *Biochem. J.* (1984) **220**, 569-573

[98] Bode, W. Chen, Z. and Bartels, K. *J. Mol. Biol.* (1983) **164**, 237-282

[99] Chen, Z. and Bode, W. *J. Mol. Biol.* (1983) **164**, 283

[100] Timm, D. E. *Protein Science* (1997) **6**, 1418

[101] Parry, M. A. A., Jacob, U., Huber, R., Wisner, A, Bon, C. and Bode, W. *Structure* (1998) **6(9)**, 1195-1206

[102] Fuentes-Prior, P., Noeske-Jungblut, W., Donner, P., Schkuning, W. D., Huber, R. and Bode, W. *Proc. Natl. Acad. Sci.* (1997) **94**, 11845-11850

[103] Henschen-Edman, A. H. *Thrombosis and Haemostasis* (1999) **81(1)**, 81-86

[104] Van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kroeger, B., Hoeffken, H. W. and Bode, W. *Embo. J.* (1995) **14**, 5149-5157

[105] Meh, D. A., Siebenlist, K. R. and Mosesson, N. W. *J. Biol. Chem.* (1996) **271**, 23121-23125