Biochemical characterization of the thrombin inhibitor of the tick, *Ornithodoros savignyi*, and investigation into the expression of its recombinant forms

by

Paul (Po-Hsun) Cheng

Submitted in partial fulfillment of the requirements for the degree of

Magister Scientiae

Department of Biochemistry

Faculty of Natural and Agricultural sciences

University of Pretoria

February 2005

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 3

LIST OF TABLES

LIST OF ABBREVIATIONS

a	Molecular radius
AEHPLC	Anion exchange high performance liquid chromatography
amu	Atomic mass unit
APC	Activated protein C
ATIII	Anti-thrombin III
bp	Base pair
BPTI	Bovine pancreatic trypsin inhibitor
BSA	Bovine serum albumin
CCMB80	Calcium chloride - manganese chloride - magnesium chloride - postassium acetate – glycerol broth
Csav	C-domain of savignin gene
Csav	C-domain of savignin
Da	Dalton
DEAE	Diethylaminoethyl cellulose
ddH_2O	Double distilled water
DLS	Dynamic light scattering
DNA	Deoxyribose nucleic acid
DTI	Direct thrombin inhibitor
DTT	dithiothreitol
EC	C-domain savignin containing enterokinase cleavage site
EDTA	Ethylene diamine tetra acetic acid
EF	Full length savignin containing enterokinase cleavage site
EK	Enterokinase
ELISA	Enzyme linked immunosorbent assay
EN	N-domain savignin containing enterokinase cleavage site
EPCR	Endothelial PC recetor
ESMS	Electrospray mass spectrometry
FDP	Fibrin degradation products
FI	Fibrinogen
FII	Prothrombin

FIIa	Thrombin
FIII	Tissue factor
FV	Proaccerlerin
FVII	Proconvertin
FVIII	Antihemophilic factor
FIX	Christmas factor
Fsav	Full length savignin gene
Fsav	Full length savignin
FX	Factor Xa (Stuart-power factor)
FXa	Activated factor Xa (activated stuart power factor)
FXa I	Factor Xa Inhibitor
FXI	Thromboplastin
FXII	Hageman factor
FXIII	Fibrin-stabilizing factor
H	Heparin binding site
HIHPLC	Hydrophobic interaction HPLC
HIT	Heparin-induced thrombocytopenia
HMWK	High molecular weight kininogen
HPLC	High performance liquid chromatography
I	Induction
IPTG	Isoporpyl-β-D thiogalactopyranoside
K	Kallikrein
K*av*	Partition coefficient
K_d	Dissociation constant
KDa	Kilodalton
K_i	Intrinsic dissociation constant
LB broth	Luria Bertani broth
LMM	Low molecular mass marker
MALDI-TOF MS	Matrix assisted laser desorption ionization time of flight mass spectrometry
MBP	Maltose binding protein
ME	Mercaptoethanol
MeCN	Acetonitrile

MM	Molecular mass
Mr	Relative molecular mass
MW	Molecular weight
Nsav	N-domain of savignin
P	Periplasmic protein
PAGE	Poly-acrylamide gel electrophoresis
PAI	Plasminogen activator inhibitor
PC	Protein C
PCI	Protein C inhibitor
PCR	Polymerase chain reaction
PDB	Protein databank
PK	Prekallikrein
PM	Peptide mass marker
R	Thrombin binding site
r	Radius of gel pore
rEK_L	Recombinant enterokinase (light chain)
Rh	Hydrodynamic radius
rNAPc2	Recombinant nematode anticoagulant protein
RPHPLC	Reversed phase HPLC
Rs	Stokes radius
RT	Retentions time
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SEHPLC	Size exclusion HPLC
SGE	Salivary gland extract
SOB	Tryptone-yeast extract- NaCl- KCl
SOC	SOB medium containing glucose
TAE	Tris-acetate-EDTA
TAP	Tick anticoagulant peptide
TEMED	N,N,N’,N’,-tetramethyl-ethylenediamine
TF	Tissue factor
TFA	Trifluoroacetic acid
TFPI	Tissue pathway inhibitor

TI	Trypsin inhibitor
TM	Thrombomodulin
Tricine	N-[Tris (hydroxymethyl) methyl] glycine
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultra violet
V	Volt
Vi	Proaccelerin inhibitor
VIIIi	Antihemophilic factor inhibitor
Vo	Void volume
Vt	Total volume
XC	C-domain savignin containing factor Xa cleavage site
XF	Full length savignin containing factor Xa cleavage site
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside
XN	N-domain savignin containing factor Xa cleavage site

ACKNOWLEDGEMENTS

It would not have been possible to complete this dissertation without the help of many people. I am particularly grateful to following people:

Dr A.R.M Gaspar, Department of Biochemistry, University of Pretoria, my supervisor for this dissertation for her guidance, support and valuable criticism.

Prof. A.W.H Neitz, Department of Biochemistry, University of Pretoria, my co-supervisor, for his interest, encouragement and financial support.

Dr B.J. Mans, Computational Biology branch, National Centre for Biotechnology and Information, National Institute of Health, my co-supervisor, for his valuable criticism and all the time he devoted to this study.

Prof. M.J. van der Merwe, Department of Biochemistry, University of Stellenbosch, for performing electrospray mass spectrometric analysis on savignin.

Prof. Summer, Department of Chemistry, University of South Africa, for allowing the usage of dynamic light scattering apparatus.

Mr. N.J. Taljaard, Department of Biochemistry (Before he retired), University of Pretoria, for performing amino acid analysis on savignin.

Mrs. S. van Wyngaardt, Department of Biochemistry, University of Pretoria, for her valuable suggestion and assisted in numerous experiments over the years.

All my fellow students, friends and my family for their interest, support, friendly advice and encouragement when things went wrong.

My God hear my prayer in my difficult time.

The National Research Foundation and University of Pretoria for their financial support.

Chapter 1

Literature Overview

1. Introduction

Thrombosis is tokhe most common cause of death in the industrialized world (Murray & Lopez, 1997). Heparin and warfarin have been used for the treatment of thromboembolic diseases for more than 50 years, but are associated with several limitations such as myocardial infarction, stroke and deep vein thrombosis (Hirsh, 2003). In the last 10 years, scientists in their quest to develop alternative agents have targeted almost every step in the blood coagulation pathway (Weitz & Crowther, 2002).

The search for novel anticoagulants has benefited from the isolation of inhibitors from hematophagous invertebrates (Urata, Shojo & Kaneko, 2003). Blood-sucking animals require a free flow of blood during feeding. The saliva of these organisms contains a rich cocktail of anticoagulants that inhibit the action of host blood coagulation factors (Law, Ribeiro & Wells, 1992, Urata, Shojo & Kaneko, 2003). The best characterized and the most advanced as a drug is hirudin. This 65-amino acid polypeptide, originally isolated from the medicinal leech, *Hirudo medicinalis*, inhibits thrombin (Markwardt, 1970). Recombinant hirudin has been successfully used in the treatment of heparin-induced thrombocytopenia (HIT) (Ortel & Chong, 1998). Tick anticoagulant peptide (TAP) (Waxman et al., 1990) and ornithodorin (van de Locht et al., 1996) isolated from the soft tick *Ornithodoros moubata*, are specific activated factor X (fXa) and thrombin inhibitors, respectively.

TAP has been evaluated as an antithrombotic in various animal models (Lyle et al., 1995), but have not undergone clinical evaluation. Two anticoagulants, a fXa inhibitor (fXaI) (Gaspar et al., 1996) and a thrombin inhibitor (savignin) (Nienaber, Gaspar & Neitz, 1999) have been isolated in our Department from the salivary glands of the tick, *Ornithodoros savignyi*. Both inhibitors have been cloned and sequenced. Structural and functional analyses of these natural molecules will provide valuable information that may be used for the design of novel anticoagulants.

1.1 Current treatment of thrombosis using warfarin and heparin

Warfarin (Figure 1.1A), a vitamin K antagonist, is the most widely used oral anticoagulant (Hirsh, 1991b). Frequent coagulation monitoring is necessary because of its unpredictable effect due in part to food and drug interactions and its narrow therapeutic window (Hirsh, 2003). Warfarin exerts its anticoagulant effect by inhibiting vitamin K epoxide reductase and possibly vitamin K reductase. This process leads to the depletion of vitamin K and limits the γ-carboxylation of the vitamin K-dependent coagulant proteins, prothrombin, factor VII, factor IX, and factor X (Whitlon, Sadowski & Suttie, 1978). In addition, the vitamin K antagonists limit the carboxylation of the regulatory anticoagulant proteins (protein C and protein S), and as a result impairs their function (Whitlon, Sadowski & Suttie, 1978).

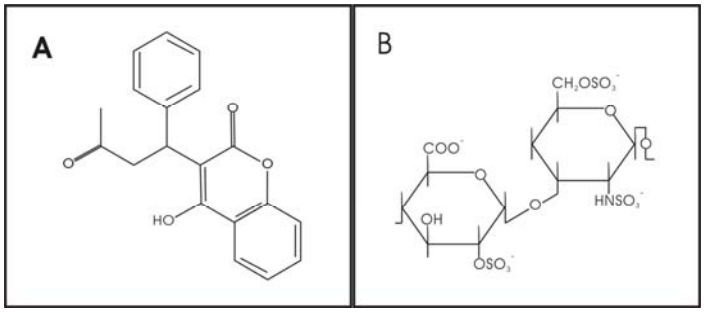


Figure 1.1 Structures of A) Warfarin and B) Heparin (repeating disaccharide unit) (Source: Adapted from www.pubmed.com).

Heparin is a glycosaminoglycan (Figure 1.1B) composed of chains of alternating residues of sulphated-D-glucosamine and D-glucuronic acid. It can be administered parenterally only and therefore coagulation monitoring is required (Choay & Petitou, 1986). The anticoagulant effect of heparin is mediated largely through its interaction with antithrombin (AT) (Figure 1.2); this produces a conformational change in AT and therefore markedly accelerates its ability to inactivate the coagulation enzymes thrombin, factor Xa, and factor IXa (Bjork & Lindahl, 1982). With heparin, the rate of inhibition is accelerated approximately 1 000 fold (Hirsh, 1991).

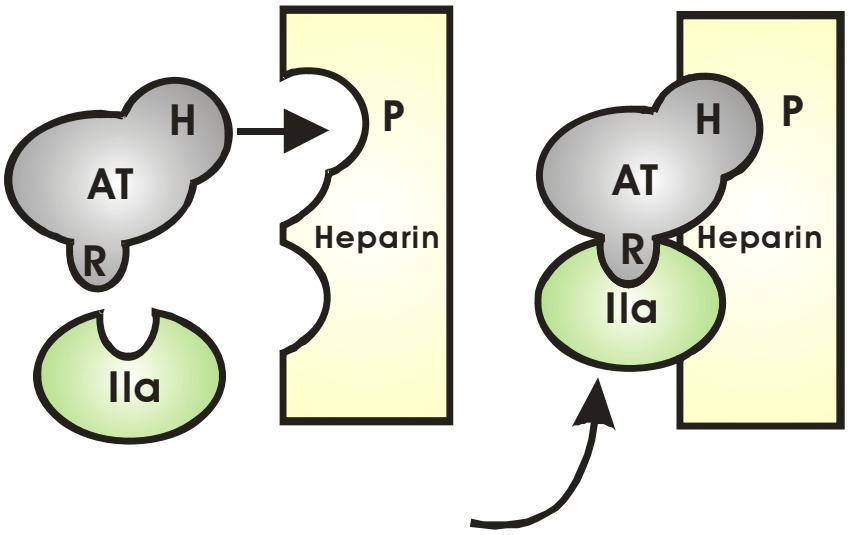


Figure 1.2 Simplified model describing how heparin catalyses the antithrombin-thrombin interaction. Symbols: AT = antithrombin, IIa = thrombin, H = heparin binding site, R = thrombin binding site, P = antithrombin binding segment (adapted from Weitz & Crowther, 2002).

Many anticoagulants have pharmacokinetic, biophysical and biological limitations. For example, the pharmacokinetic limitation of heparin is caused by its non-specific binding to proteins and cells (Hirsh, 1991). This limitation of heparin is manifested clinically by its poor bioavailability at low doses, the marked variability in its anticoagulant response among patients with thromboembolic disease, and its

relatively short plasma half-life (Hirsh et al., 1995). The biophysical limitation of heparin reflects the inability of the heparin-antithrombin complex to inactivate thrombin which is bound to fibrin and factor Xa. Fibrin and factor Xa are bound to phospholipid surfaces within the prothrombinase complex (Hirsh et al., 1995).

The biological limitation of heparin, namely internal bleeding, is the major adverse effect and occurs with both prophylactic and therapeutic use (Amerena, Mashfold & Wallace, 1990). Serious life-threatening haemorrhage may occur, although bleeding tends to be less of a problem with heparin than with oral anticoagulants, owing to its short half-life. Thus clearly, there exists a need for more specific, versatile and less toxic anticoagulants (Amerena, Mashfold & Wallace, 1990).

1.2 The ideal anticoagulant

The limitations described above have led to the development of new, improved agents to meet several currently unmet clinical needs. To permit this, the ideal agent should have: 1) no or minimal interactions with food or other drugs and low, non-specific plasma protein binding, 2) a wide therapeutic window, 3) an appropriate half-life, 4) a rapid onset of action and 5) a rapid offset of action (Hirsh, 2003).

1.3 The blood coagulation cascade

Whenever a vessel is ruptured, haemostasis is achieved by several mechanisms including 1) vasoconstriction 2) formation of a platelet plug and 3) blood coagulation. The physiological response to vascular damage culminates in the rapid generation of thrombin at the site of injury, subsequently leading to an increase in platelet deposition and the formation of the insoluble fibrin network. The response of the coagulation process is generally limited to the site of injury and is proportional in

magnitude to the extent of the vascular damage (Furie & Furie, 1992).

A blood clot (thrombus) forms through the action of a cascade of proteolytic reactions involving the participation of nearly 20 different substances, most of which are liver-synthesized plasma glycoproteins. The cascade is shown in Figure 1.3 below. All but two of these factors are designated by both a roman numeral and a common name.

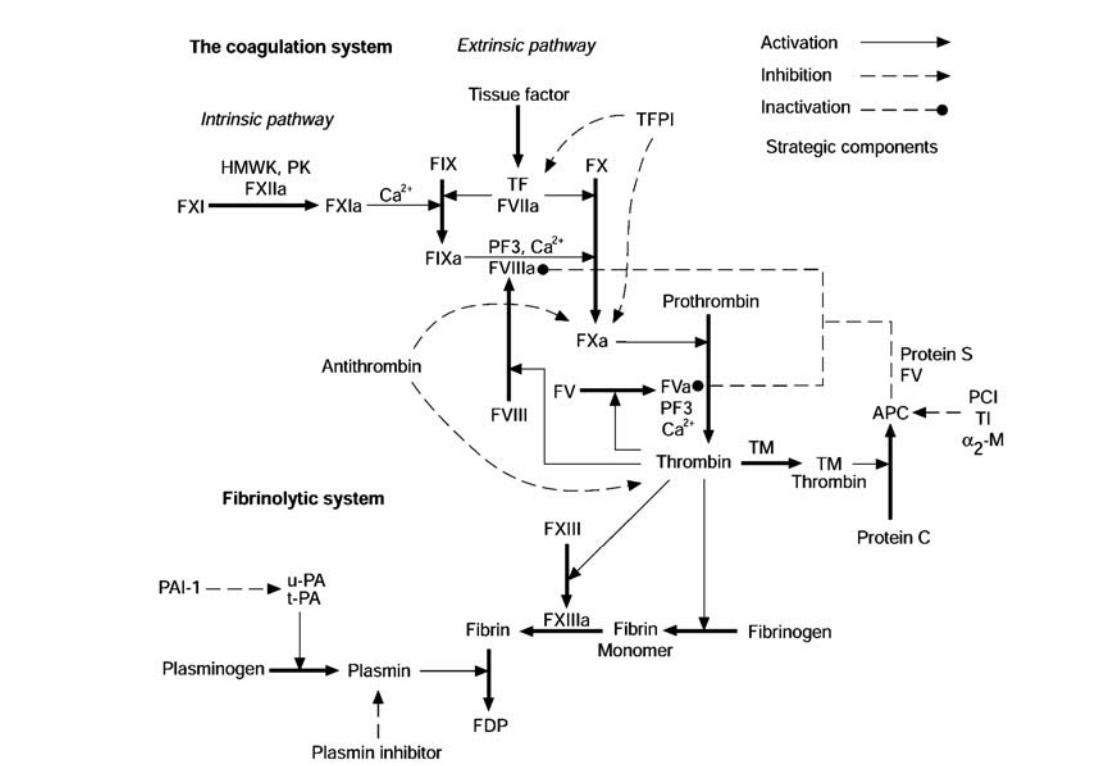


Figure 1.3 Schematic representation of the coagulation cascade. Blood coagulation is initiated by vascular injury and results in the explosive generation of thrombin which clots blood. Coagulation factors are represented by Roman numerals (a = activated). Abbreviations: HMWK = high molecular weight kininogen, PK = prekallikrein, K = kallikrein, TF = tissue factor, TFPI = tissue factor pathway inhibitor, PF3 = phospholipids, TM = thrombodulin, PC = protein C, APC = activated protein C, PCI = protein C inhibitor, TI = trypsin inhibitor, α_2-M = α_2-macroglobulin, FDP = fibrin degradation products**.** (Source: http:// www.chromogenix.com)

Seven of the clotting factors are zymogens of serine proteases that are proteolytically activated by serine proteases further up in the cascade (Table 1.1). Blood coagulation is the culmination of a series of proteolytic reactions that terminate in the thrombin catalyzed conversion of fibrinogen to fibrin. Activation of blood coagulation by either the intrinsic or extrinsic pathways results in the formation of activated factor X (fXa) pathway, which catalyzes the formation of thrombin from prothrombin. Tissue factor and tissue phospholipids activate the extrinsic pathway while contact of fXII and platelets with collagen in the vascular wall initiates the intrinsic pathway. The conversion of prothrombin to thrombin involves the formation of the prothrombinase complex. This complex consists of prothrombin and fXa bound to fVa on anionic phospholipids on the platelet membrane. Thrombin in turn converts soluble fibrinogen to insoluble fibrin and also activates factor XIII to its active form which stabilizes the clot by cross-linking reaction (Furie & Furie, 1992).

Table 1.1 Plasma coagulation components (Source:http://www.chromogenix.com)

Factor	Name	Size [KDa]	Concentration [µg/ml]	Factor	Name	Size [KDa]	Concentration [µg/ml]
I	Fibrinogen	340	3000	X	Stuart-Power factor	59	8
II	Prothrombin	69	100	XI	Thromboplastin	160	5
III	Tissue factor	47	-	XII	Hageman factor	80	30
IV	Calcium	-	-	XIII	Fibrin-stabilizing factor	320	10
V	Proaccelerin	330	10	-	Tissue factor	37	-
VI	-	-	-	-	Protein C	57	4
VII	Proconvertin	48	0.5	-	Protein S	75	25
VIII	Antihemophilic factor	330	0.1	-	Antithrombin	58	150
IX	Christmas factor	55	5	-	Heparin cofactor II	66	91

1.4 Control of blood coagulation

Blood coagulation is modulated by several mechanisms, including the heparin-antithrombin interaction (See section 1.1), the tissue factor pathway inhibitor (TFPI), activated protein C and the fibrinolytic system (Broze, 1995; Esmon et al., 1997; Collen, 1999).

1.4.1 Tissue factor pathway inhibitor (TFPI)

Inhibition of the factor VIIa-tissue factor complex is effected by TFPI, the majority of which is bound to endothelium (Broze, 1995). TFPI acts in a two-step manner: Firstly, it complexes and inactivates factor Xa. Secondly, the resulting complex inactivates factor VIIa within the factor VIIa-tissue factor complex. Because TFPI downregulates the initiation of coagulation by the factor VIIa-tissue factor complex, an alternate mechanism for propagating coagulation is necessary. This may be provided by factor XI, which is efficiently activated by thrombin in the presence of platelets (Gailani & Broze, 1993). By activating factor IX, a key component of the intrinsic tenase complex, factor XIa induces the generation of sufficient amounts of factor Xa to propagate coagulation (Gailani & Broze, 1993).

1.4.2 Protein C

Thrombin is inhibited when it binds to thrombomodulin (a 74 kDa glycoprotein), a thrombin receptor found on the endothelium (Figure 1.4). Once bound to thrombomodulin, thrombin undergoes a conformational change at its active site that converts it from a procoagulant enzyme into a potent activator of protein C. Thrombomodulin specifically binds thrombin so as to convert it to a form with decreased ability to catalyze clot formation and 1000-fold increased capacity to activate protein C (Esmon et al., 1997). Activated protein C (APC), together with its

cofactor, protein S (PS), acts as an anticoagulant by proteolytically degrading and inactivating activated factor V (Va) or factor VIII (VIIIa) on the platelet surface.

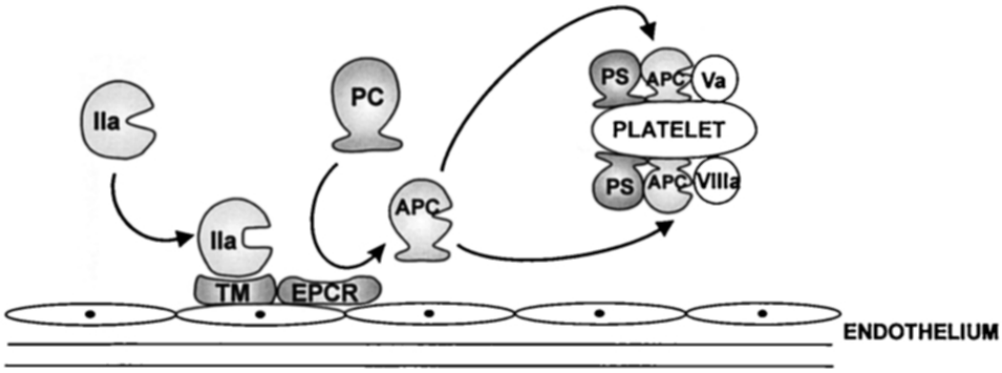


Figure 1.4 Protein C (PC) anticoagulant pathway. Thrombin (IIa) binds to thrombomodulin (TM), the endothelial cell thrombin receptor. Once bound, IIa undergoes a conformational change at its active site that converts it from a procoagulant to a potent activator of PC. Activation of PC occurs on the endothelial cell surface, where the zymogen binds to the endothelial PC receptor (EPCR) (Based on: Esmon et al., 1997).

1.4.3 The fibrinolytic system

Blood clots are only temporary patches; they are eliminated as wound repair progresses. Fibrin is dismantled in a process of fibrinolysis. This is particularly needed when a clot has inappropriately formed or has broken free into the general circulation (Collen, 1999). Plasminogen activators convert plasminogen to plasmin which is a plasma serine protease that specifically cleaves fibrin, a triple coiled protein. The protease cleaves a covalently linked α–chain which is not in the triple coiled region (Collen, 1999). This opens the mesh-like structures of the blood clot and gives plasmin free access to polymerized fibrin molecules thereby facilitating clot lysis. Fibrinolysis is regulated by the action of plasminogen activator inhibitor (PAI-1) and antiplasmin on the plasminogen activators and plasmin, repectively (Figure 1.5).

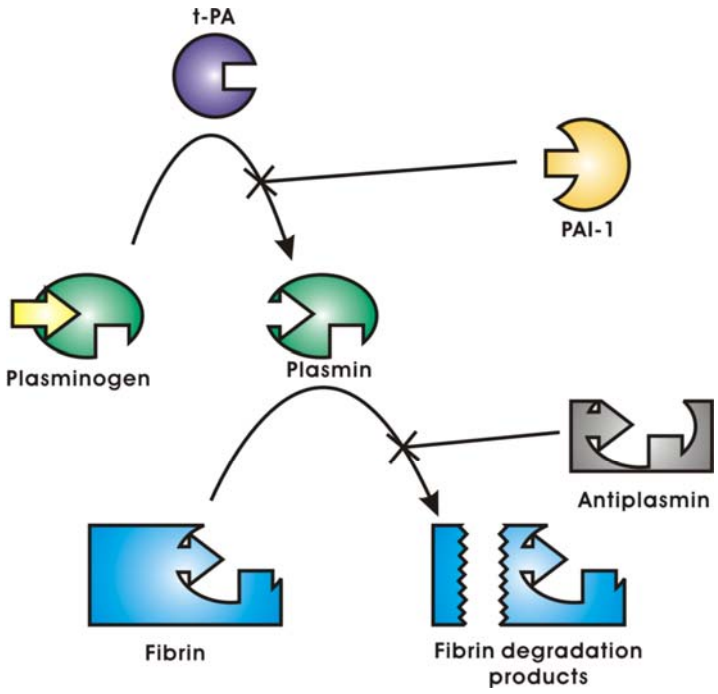


Figure 1.5 The fibrinolytic system. Plasminogen activators convert plasminogen to plasmin. Plasmin degrades fibrin to yield fibrin degradation products. The system is regulated at two levels; type 1 plasminogen activator inhibitor (PAI-1) inactivates the plasminogen activators, whereas antiplasmin inhibits plasmin (adapted from Collen, 1999).

1.5 Thrombin

Thrombin is a serine protease enzyme that is responsible for many aspects in the blood coagulation cascade. As a procoagulant, it cleaves numerous substrates such as fibrinogen, factors V, VIII and XIII, and it is also necessary for platelet aggregation. With its anticoagulant function it leads to activation of protein C (Mann & Lorand, 1993). In addition to its procoagulant and anticoagulant functions, it is an agonist for a number of cellular responses during wound repair. Its role in thrombotic disorders has made it the centre of focus for development of therapeutic agents and recent advances in the understanding of its molecular structure have made it possible to design and develop novel and more effective antithrombotic drugs to combat thrombosis (Narayanan & Thiagarajan, 2001).

Thrombin is composed of two polypeptide chains, A and B, which are covalently linked through a disulfide bond. The A chain has no documented functional role and runs opposite to the front hemisphere of the B chain, which hosts the entrance to the active site of the enzyme. The B chain is shaped like a sponge, with deep crevices and large protruberances on its water-accessible surface (Stubbs & Bode, 1992).

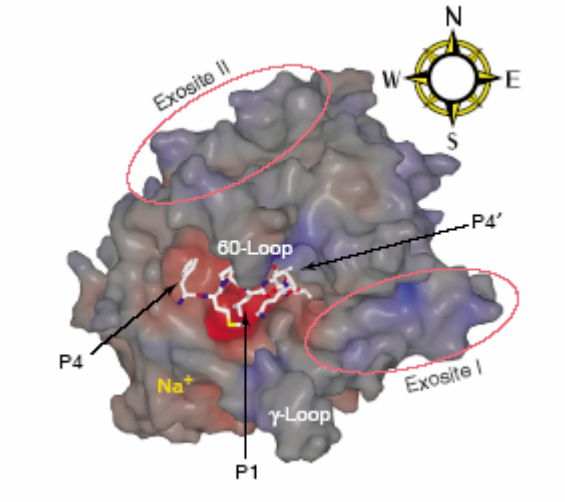


Figure 1.6. Surface representation of thrombin. The classic view of the active site cleft of thrombin. Traditionally, the structural features of thrombin are described relative to the active site with the substrate running from West to East (N-terminal to C-terminal). The surface of thrombin is coloured according to the electrostatic potential (blue for positive charge and red for negative charge). The active site possesses an overall negative potential and prefers an arginine in the P1 position. The non-primed side of the active site is a large hydrophobic cavity that prefers hydrophobic residues (P2, Pro; P3, Phe; and P4, Phe). The active site is buried deep in a canyon formed by the 60-loop and γ-insertion loop. Substrate and cofactor binding is often mediated by exosite interactions, and anion-binding exosites I and II are indicated. Thrombin activity is also affected by the binding of monovalent cations, and the Na^+ binding site is indicated (Source: Huntington & Baglin, 2003).

The active-site cleft appears at the centre of the molecule. Thrombin features a trypsin-like specificity and cuts preferentially at Arg residues. Unlike trypsin, however, thrombin cleaves selectively at specific Arg sites using alternative interactions from "exosites" distinct from the active site (Stubbs & Bode, 1992).

Exosite I (fibrinogen binding site) is located "east" of the active site (Stubbs & Bode 1992). It contains hydrophobic patches and numerous charged residues on its surfaces that provide electrostatic steering to fibrinogen on its approach to the active site of thrombin (Rose & Di Cera, 2002). Exosite I also provides the locus for the binding of thrombomodulin to thrombin (Pineda, Cantwell & Bush, 2002).

Exosite II (glycosamino-binding site) is positioned "west" of the active site (Stubbs & Bode, 1995), opposite to exosite I. It features a conspicuous number of charged residues, but unlike exosite I has no hydrophobic patches on its surface. Exosite II is the locale for interaction with polyanionic ligands like glycosaminoglycans and heparin (Sheehan & Sadler, 1994).

1.6 Novel anti-coagulants

1.6.1 Inhibitors of the TF-factor VII pathway

Inhibitors of the tissue factor (TF)-factor VIIa complex have been studied most extensively so far and some have been evaluated in patients (Gresele & Agnelli, 2002). These inhibitors include the naturally occurring anticoagulant TFPI and nematode-derived anticoagulant protein isolated from *Ancylastoma canicum*. Recombinant TFPI has shown promise in preclinical studies and is being evaluated *in vivo* (Narita et al., 1995). The recombinant nematode anticoagulant protein (rNAPc2), which binds to a noncatalytic site on factor X (or factor Xa) and inhibits factor VIIa within the TF-factor VIIa complex, is the most extensively clinically evaluated inhibitor of the TF-factor VIIa pathway (Stassens et al., 1996).

1.6.2 Factor Xa inhibitors

Both direct and indirect factor Xa inhibitors are under development. Direct inhibitors of factor Xa inactivate factor Xa bound to phospholipid surfaces, as well as free factor Xa. Direct factor Xa inhibitors include natural inhibitors, such as TAP (Waxman et al., 1990), antistasin (Tuszynski, Gasic & Gasic, 1987) and synthetic agents (Herbert et al., 1996). Synthetic factor Xa inhibitors are non-peptide, low molecular weight and reversible inhibitors of factor Xa (Herbert et al., 1996). TAP and antistasin are specific polypeptide inhibitors, originally isolated from the tick, *Ornithodoros moubata* (Vlasuk et al., 1993) and the leech, *Haementeria officinalis* (Tuszynski, Gasic & Gasic, 1987), respectively. Both native and recombinant forms of antistasin are tight-binding, slow reversible inhibitors of factor Xa. Like TAP, antistasin is highly selective for factor Xa.

Indirect factor Xa inhibitors require antithrombin for their action and only inhibit the activity of free factor Xa (Gresele & Agnelli, 2002).

1.6.3 Thrombin inhibitors

Thrombin inhibitors can inactivate thrombin either indirectly by activating antithrombin (eg. heparin) or directly by binding to thrombin and preventing its interaction with other substrates.

Direct thrombin inhibitors (DTIs) can be divided into those that bind bivalently to thrombin at exosite I or exosite II as well as the active site and those that bind univalently to the active site of thrombin (Hirsh, 2003). Examples of bivalent DTI's isolated from invertebrate species are listed in Table 1.2.

Table 1.2 Thrombin inhibitors isolated from invertebrate species (Source: Urata, Shojo & Kaneko, 2003)

Molecule	species	Amino acid residues	Ki	References
Hirudin	*Hirudo medicinalis* (European medicine leech)	65	22 fM	Stones & Hofsteenge, 1986
Haemadin	*Haemadipas sylvestris* (Land-living leech)	57	224 fM	Richardson et al., 2002 Strube et al., 1993
Theromin	*Theromyzon tessulatum* (Rhynchobdellid leech)	67	34 pM	Salzet et al., 2000
Dipetalogastin	*Dipetalogaster maximus* (Reduviid bug)	453	125fM	Lange et al., 1999
Dipetalogastin II	*Dipetalogaster maximus* (Reduviid bug)	344	49.3 fM	van de Locht et al., 1995
Rhodniin	*Rhodniin prolixus* (Triatomine bug)	103	203 fM	Friedrich et al., 1993
Triabin	*Triatoma pallidipennis* (Triatomine bug)	142	3.0 pM	Noeske-Jungblut et al., 1995
Ornithodorin	*Ornithodoros moubata* (Soft tick)	119	1.0 pM	van de Locht et al., 1996
Savignin	*Ornithodoros savignyi* (Soft tick)	134	4.89 pM	Nienaber, Gaspar & Neitz, 1999
TTI	*Glossina morsitans morsitans* (Tsetse fly)	32	584 fM	Capello et al., 1998 Li, Kwon & Aksoy, 2001
Anophelin	*Anopheles albimanus* (Mosquito)	83	5.87 pM	Valenzuela, Francischatti & Riberio 1999. Francischatti, Valenzuela & Riberio 1999.

The high inhibition potency and specificity of the thrombin inhibitors is probably due to binding of the inhibitors to two different sites on the thrombin molecule. The Ki values for these inhibitors are in the order of 10^{-4} ~ 10^{-11} M. The N- and C-terminal domains of hirudin (Stone & Hofsteenge, 1986), triabin (Noeske-Jungblunt et al., 1995), rhodniin (Friedrich et al., 1993) and ornithodorin (van de Locht et al., 1996) interact with the active site cleft and exosite I of thrombin, respectively.

Univalent DTIs are synthetic, low molecular weight inhibitors that bind non-covalently to the active site of thrombin and act as competitive inhibitors (Hilpert et al., 1994).

1.7 Anticoagulants from the soft tick, *O. savignyi*

The tick *O. savignyi* is a soft tick that occurs throughout the North-western regions of Southern Africa. This tick is also known as "the eyed sand tampan" and was used as the model in this study. During feeding the uniform integument which folds in on itself, allows the tick to obtain a large quantity of blood in a short period of time (Mans, 2002). When these ticks feed they secrete toxic substances which can be lethal to young animals (Howell, Neitz & Potgieter, 1975).

Figure 1.7 shows a dissected tick, the large salivary glands lie in an oblique position and resemble two bunches of white grapes.

Figure 1.7 Internal organ of the O.savignyi (Mans, 2002).

An electron micrograph of salivary gland is shown in Figure 1.8. Several bioactive compounds have been identified and characterized in the salivary glands of *O. savignyi* (Table 1.3). Kinetic data for these molecules is indicated in Table 1.3. The data for the substances isolated from the tick *O. moubata*, a close relative of *O. savignyi* is also shown in Table 1.3 for comparison.

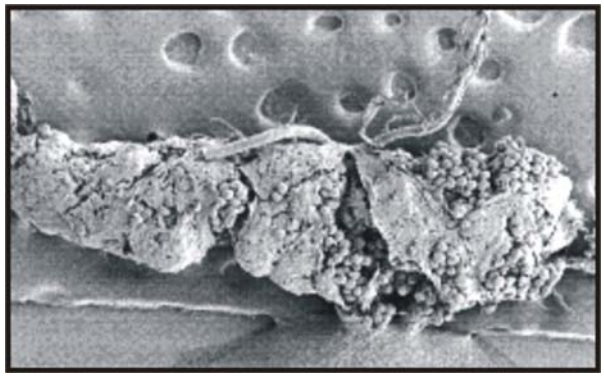


Figure 1.8 Electron micrograph of a salivary gland of O. savignyi (Mans, 2002).

Table 1.3: Inhibitors isolated from O. savignyi and O. moubata

Molecule	Species	Molecular mass	Ki	Reference
Thrombin inhibitors				
Savignin	*O. savignyi*	12 kDa	4.89 pM	Nienaber, Gaspar & Neitz, 1999
Ornithdorin	*O. moubata*	12 kDa	1.0 pM	van de Locht et al., 1996
Factor Xa inhibitors				
FXa I	*O. savignyi*	12 kDa	830 pM	Gaspar et al., 1996
TAP	*O. moubata*	7 kDa	142 pM	Waxman et al., 1990
Platelet aggregation inhibitors				
Savignygrin	*O. savignyi*	6.9 kDa	-	Mans, Louw & Neitz, 2002
Disagregrin	*O. moubata*	6 kDa	-	Karczewski et al., 1994

A number of natural proteins originally isolated from haematophagous organisms have been shown to be direct inhibitors of factor Xa. Factor Xa inhibitor (fXa I) was isolated from salivary gland extracts prepared from *O. savignyi*, a close relative of *O. moubata*. The N-terminal amino acid sequence (residues 1-12) was determined and found to share a 66% identity with the TAP (Gaspar et al., 1996). FXa I displays a slow, tight-binding inhibition of factor Xa. The interaction of the fXa-I was found to be competitive and dependent on ionic strength (Gaspar et al., 1996).

Savignin (12 kDa), a thrombin inhibitor, has been isolated from the salivary glands of the tick, *O. savignyi* (Nienaber, Gaspar & Neitz, 1999). The full-length gene encoding the inhibitor has been cloned and sequenced using both 5' and 3' RACE. Savignin is structurally similar to basic pancreatic trypsin inhibitor (BPTI). BPTI is a small, soluble and stable molecule. The easiest family characteristic to appreciate at a glance is the topological relationship between the disulphide bridges and the location of the reactive site. BPTI has 6 Cys residues that form 3-disulphide bonds (Figure 1.9).

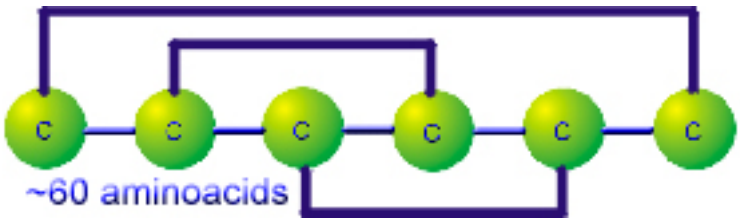


Figure 1.9: Topological structures of BPTI. There are 6 Cys residues indicated in green and three disulphide bridges in dark blue (Adapted from Mans, 2002).

Kinetic studies indicated that savignin is a competitive, slow-tight binding inhibitor of α–thombin (Ki = 4.89 ± 1.39 pM), whereas with γ–thrombin (Ki = 22.3 ± 5.9 nM), a thrombin derivative lacking exosite I, reduced inhibition was observed which suggests that savignin interacts with exosite I (Nienaber, 1999).

With the work that was done by Nienaber 1999, it was not possible to show the mechanism of savignin binding to thrombin. Structural modelling of savignin based on the crystal structure of ornithodorin, the inhibitor from *O. moubata*, and docking to thrombin indicated similarities to ornithodorin (Mans, Louw & Neitz, 2002a). The N-terminal domain of savignin binds to the active site, while the C-terminal domain interacts with the basic fibrinogen recognition exosite of thrombin (Mans, Louw & Neitz, 2002a). Mans (2002) hypothesized that the domains of savignin interact with each other, giving a globular form in the absence of thrombin. Binding of the C-terminal domain of savignin to the fibrinogen-binding site of thrombin leads to a dissociation of the savignin domains. This would yield an extended conformation that would allow the N-terminal residues from the N-terminal domain of savignin, to bind within the thrombin's active site.

1.8 Aims

This dissertation describes the further biochemical characterization of native savignin as well as the investigation into its recombinant production in order to obtain adequate amounts of inhibitor to study the binding mechanism and its structure and kinetics. Characterization of the structural conformation of uncomplexed savignin will aid in the elucidation of its mechanism of inhibition.

The aims of the present study are:

1. To investigate the structural conformation of uncomplexed native savignin.
2. To recombinantly express full-length savignin and the individual N- and C-domains.

Chapter 2

Investigation into the proposed association of the N- and C-terminal globular domains of savignin

2.1 Introduction

This chapter describes the investigations into the hypothesis proposed by Mans (2002) concerning the mechanism of inhibition of thrombin by savignin. The hypothesis proposes that the N- and C-terminal domains of savignin interact with each other in the uncomplexed state to give a compact globular form (Figure 2.1). Binding of the C-terminal domain to thrombin's fibrinogen-binding exosite leads to a dissociation of the domains to give an extended conformation. This allows the N-terminal residues from the N-terminal domain of savignin to bind within thrombin's active site.

The hypothesis proposed by Mans thus contains two ideas – the association of the two globular domains resulting in a compact, globular structure and the targeting of the exosite of thrombin by the C-terminal domain, resulting in a conformational change to an extended form (Figure 2.1).

The suggested mechanism regarding the conformational change and binding of savignin is based on the crystal structure of the ornithodorin-thrombin complex (van de Locht et al., 1996), the docking model of savignin with thrombin (Mans, Louw & Neitz, 2002) and kinetic studies of TAP and savignin (Jordan et al., 1992; Wei et al., 1998; Nienaber, Gaspar & Neitz, 1999).

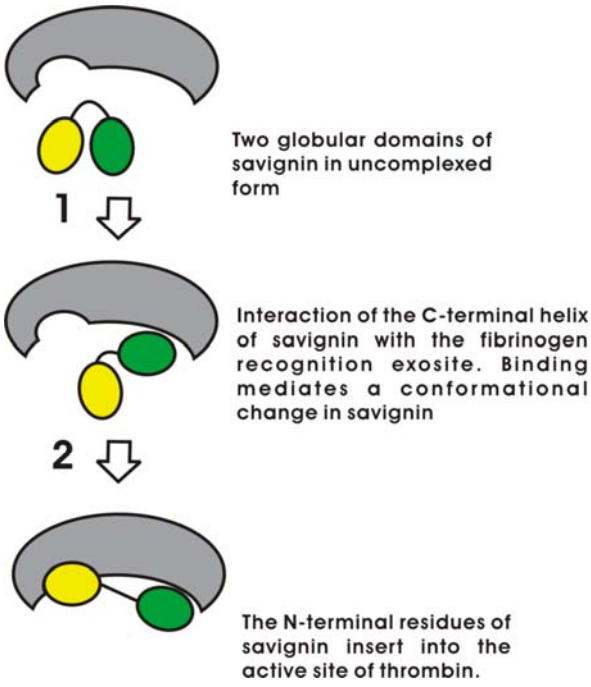


Figure 2.1 A schematic representation of the two-step mechanism for savignin binding to thrombin as proposed by Mans (2002). Green: C-terminal, Yellow: N- terminal.

Furthermore, savignin shows an unexpected substantially higher molecular mass after SDS-PAGE electrophoresis in comparison to the true molecular mass determined by electrospray mass spectrometry (ESMS). This may indicate an extensive unfolding of the structure or low binding of SDS (Nienaber, 1999). These observations upon which the hypotheses are based are described in more detail below.

Ornithodorin, a thrombin inhibitor from the tick, *O. moubata,* has been crystallized in complex with thrombin (van de Locht et al., 1996). It consists of two BPTI-like domains connected by a linker of 7 amino acid residues. The N-terminal domain is involved in the interaction with the active site of thrombin, while the C-terminal domain interacts with the basic fibrinogen recognition site via the C-terminal helix and probably the overall negative electrostatic nature of this domain (Figure 2.2B).

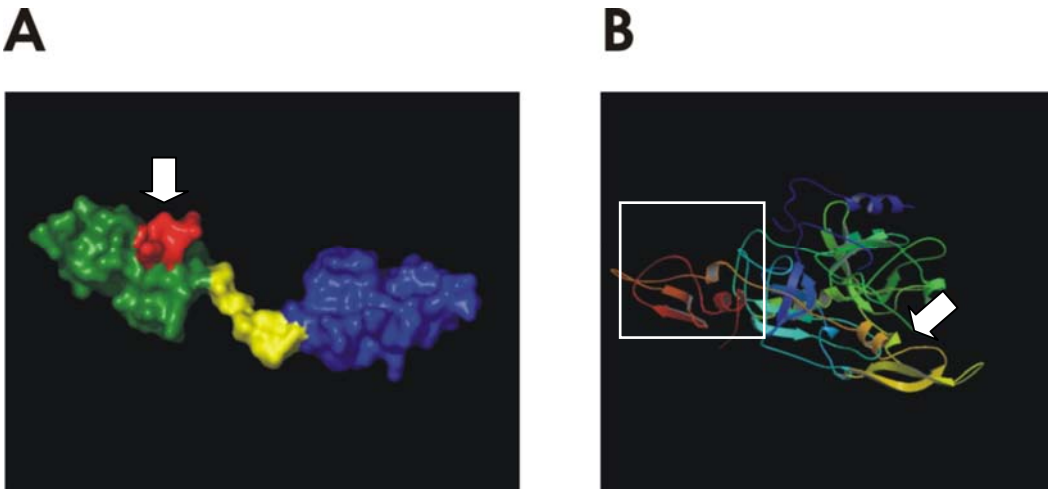


Figure 2.2* A) *A surface model of extended savignin. The N-terminal domain is indicated in green. Yellow indicates the linker region. Blue indicates the C-terminal domain of savignin. Red indicates amino acids that are involved in the active site binding. **B) *A surface model of the ornithodorin-thrombin complex (Accession Code: 1TOC).*** The N-terminal domain of ornithodorin fits into the active site cleft of thrombin (indicated with white arrow), while the C-terminal α-helix of ornithodorin shows close proximity to the fibrinogen-binding exosite of thrombin (indicated with box).

Savignin, the thrombin inhibitor described for the tick, *O. savignyi*, shows 83% sequence identity with ornithodorin (Mans, Louw & Neitz 2002). Structural modelling of savignin and docking to thrombin indicate that its inhibitory mechanism is similar to that of ornithodorin (Figure 2.2B).

TAP from *O. moubata* is an inhibitor of fXa and consists of 60 amino acids with a molecular mass of 6 850 Da (Waxman et al., 1990). It has limited homology to Kunitz-type inhibitors although it possesses a BPTI-like fold. It has been shown that TAP is a slow, tight binding inhibitor of fXa and binds to fXa in a two-step fashion, involving a secondary binding site (Wei et al., 1998).

The crystal structure of the TAP-fXa complex shows that three N-terminal amino acid residues bind inside the active site of fXa whereas the C-terminal end interacts with a

secondary binding site close to the active site (Wei et al., 1998). To explain the observed two-step kinetic mechanism observed for TAP it was proposed that an initial slow-binding step occurs at the secondary binding site resulting in a conformational change of TAP with concomitant binding into the active site (Jordan et al., 1992).

The crystal structure of ornithodorin and the modelled structure of savignin have unusual conformations. Both consist of two globular domains with a flexible linker region (Figure 2.2). The likelihood of a compact form of uncomplexed savignin is supported by considering the following observations. In proteins in general, the number of thermodynamically unfavourable solute-solvent interactions are minimized by concealing hydrophobic residues in the interior structure. Hereby the surface area of a protein exposed to an aqueous solvent is reduced by folding into a globular conformation (Jones & Thornton, 1995). The extended conformation thus seems unusual with the linker region exposed to solvent. To establish if this differs from the general behaviour, Mans (2002), investigated the relationship between molecular mass and volume of several globular proteins. A clear relationship between these parameters was found (Figure 2.3).

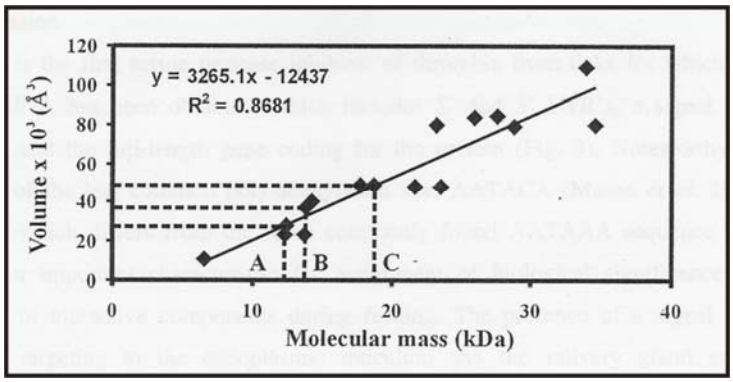


Figure 2.3 Relationship between molecular mass and volume (Å^3) of various proteins. Hydrodynamic data of the various proteins were obtained (Creighton 1992). (A) Volume of savignin derived from its molecular mass, (B) volume for bikunin derived from its molecular mass (C) molecular mass of savignin derived from measured volume (Rasmol) (Mans 2002).

For savignin a volume of ~ 28 000 Å^3 is indicated for a molecular mass of 12 430 Da (Figure 2.3). However, the volume measured with the Rasmol package for the complexed extended form of ornithodorin and for the modelled structure of savignin is ~ 49 000 Å^3 that indicates a molecular mass of ~19 kDa, which is clearly outside normal deviation. Bikunin, a plasma serine protease inhibitor which also contains two BPTI-like domains, falls neatly into the expected mass to volume relationship. In uncomplexed bikunin the two globular domains are packed upon each other with a turn in the linker region resulting in an overall globular structure (Xu et al., 1998)

To characterize the structural conformation of savignin, one approach is to determine the Stokes radius (Rh or Rs). This parameter can be obtained by different analytical techniques such as size-exclusion chromatography, electrophoresis, sedimentation analysis, light scattering, electron microscopy, (Rowe, 1978) etc. For the determination of Rh, calibration is required using proteins of known molecular mass.

The earliest applications of SEC were aimed at preparative procedures; further development of chromatography allowed the study and characterization of molecular systems. These include procedures for the determination of molecular size and molecular mass, studies of heterogeneous distributions of biopolymers, conformational transitions and subunit interactions (Ackers, 1970).

Ackers (1964) defined the fundamental principle that underlies the use of SEC systems as analytical tool in protein chemistry, such as partition coefficients (K*av*, a parameter that can be used to characterize the interaction between a solute molecule and a porous gel network). K*av* depends on the molecular size and shape of the protein in contrast to surface or charge properties.

$K_{av} = (1 - a/r)^3$	**K_{av}: Partition coefficients** **a: Molecular radius** **r: Radius of gel pore**

Size exclusion chromatography (SEC) is a viable technique for the determination of the hydrodynamic or Stokes radius (Cabré, Canela & Canela 1989). Comparison of the elution volumes of unknown proteins with that of globular proteins with known molecular masses and hydrodynamic radii allows the determination of the hydrodynamic radius of the unknown protein (Cabré, Canela & Canela 1989).

One of the aims of this chapter was to apply SEC as a method for investigating the proposed association of the N- and C-terminal globular domains of savignin by estimating the hydrodynamic radii (Rh) of the different forms of savignin.

2.2 Materials and methods

2.2.1 Calculations of the hydrodynamic radius of savignin

As mentioned earlier the hypothesis of Mans (2002), proposed that savignin might exist in either a compact or extended form. It was previously indicated that a compact conformation exists for bikunin which is also a double-BPTI domain protein (Xu et al., 1998). Because the molecular mass of bikunin (12 048 Da) is so close to that of savignin (12 430 Da) we used its structure (Protein Databank Accession Code: 1BIK) as a model for the compact form of savignin. To determine whether it will be feasible to investigate the proposed conformational differences by measurement of the hydrodynamic radius (i.e. are the Rh values different enough to be measured?), two approaches were used. HYDROPRO programme and the formula shown in Figure 2.4 ($Rh = 0.6452X^{0.4115}$). These calculations also allow choice of standard proteins with known hydrodynamic data that can be used for calibration during SEC.

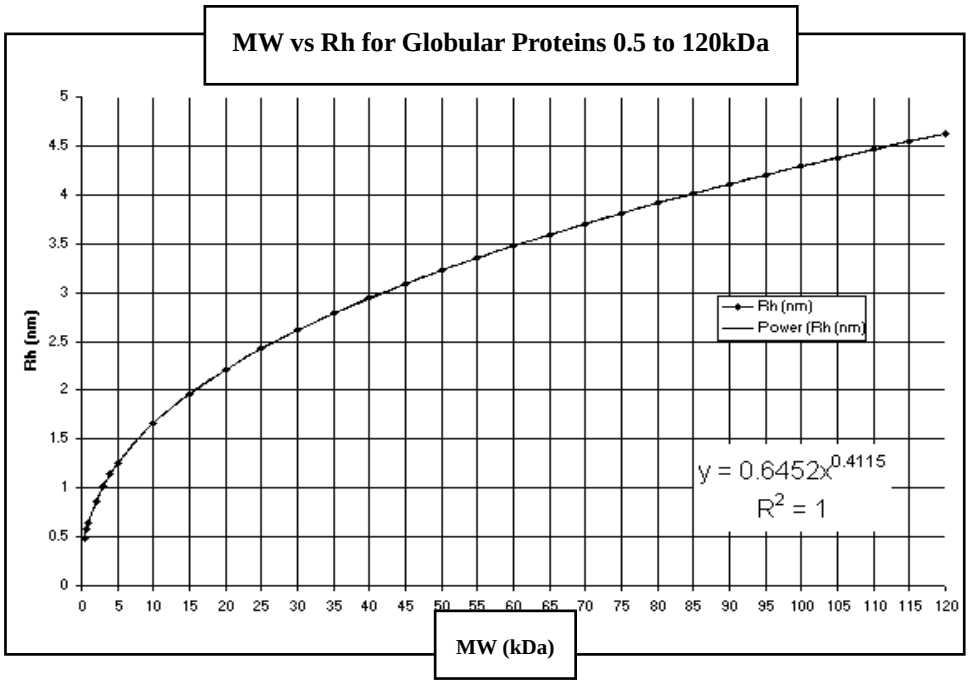


Figure 2.4: Standard hydrodynamic curve. Plots are included to help visualize the mathematical relationship. Y axis: indicates the hydrodynamic radius. X axis: indicates the molecular weight of globular proteins (obtained from Precision Detectors Inc.).

The program HYDROPRO allows the calculation of hydrodynamic properties from crystal structure data (Carasco & Garcia de la Torre, 1999; Garcia de la Torre, Huertas & Carrasco, 2000). Non-hydrogen atoms are replaced by spheres of a fixed radius and the resulting structure with overlapping spheres is represented by a shell model that allows accurate calculation of hydrodynamic parameters. To run the programme a structure file such as a PDB file must be supplied, as well as the known molecular mass of the protein.

Default values for other variables were used and include the effective radius of the atomic elements at 3.1 Å, the minimum and maximum radius of the beads in the shell (1 Å and 2 Å, respectively), the temperature (293 Kelvin), the viscosity of the solvent in poise (0.01) and the solvent density (1 g/cm^3). The program calculates the partial specific volume of the molecule from the sequence extracted from the structure file. The Rh values calculated for ornithodorin and bikunin correspond with that of chymotrypsinogen and lysozyme respectively (Mans personal communication, Garcia de la Torre, Huertas & Carrasco, 2000). In the present study the structural files (chymotrypsinogen, Protein Databank Accession Code: 2CGA; lysozyme, Protein Databank Accession Code: 6LYS) were used for recalculation of their Rh values.

2.2.2 Preparation of tick salivary gland extracts (SGE)

Ticks were collected in the Upington region by sifting of sand. The salivary glands were obtained by dissection, frozen in liquid nitrogen and stored at -70 ^{o}C (Mans et al., 1998). Salivary gland extracts were prepared by sonication with a Branson sonifier (Branson Sonic Power Co.), using 3 x 6 pulses at 30 % duty cycle. Glands were kept on ice throughout the procedure. Sonified products were centrifuged in a microfuge (10 000 g for 5 minutes) and the supernatant was used for further studies.

2.2.3 Thrombin chromogenic substrate assay

Thrombin activity was determined by monitoring the increase in the absorbance at 405 nm caused by the release of *p*-nitroaniline when Chromozym TH (tosyl-glycyl-propyl-arginine-4-nitroanilide acetate, Roche Mannheim) is cleaved. The reactions were performed in 96 well Sterilin microtiter plates (Bibby Sterilin, UK), and the absorbance was monitored over 10 minutes at room temperature using an ELISA microplate reader (Thermo Labsystem, Multiskan Ascent). Human thrombin (8×10^{-4} U in assay buffer, Roche Mannheim) was pre-incubated with 10 – 20 µl of salivary gland fractions obtained from HPLC or with buffer (20 mM Tris-HCl, 1 M NaCl, pH 7.6) as control in a final volume of 200 µl for 30 minutes at room temperature. Then 20 µl Chromozym TH (1.9 mM) prepared in deionized double distilled water was added.

2.2.4 Purification of savignin by HPLC

It was initially planned to employ dynamic light scattering (DLS) in this study but was later found not to be possible - please refer to the results section for an explanation. To determine an accurate hydrodynamic radius of savignin, purification of savignin was done by HPLC. Since relatively large quantities of highly purified samples are required for DLS, new purification strategies were used to obtain a better yield compared to previous purification methods (Nienaber, Gaspar & Neitz, 1999). A Beckman HPLC aparatus, consisting of two pumps, an analogue interface for control of solvent delivery, a UV detector and a Beckman 340 organiser was used. This system was controlled with Beckman System Gold software. All buffers and samples were filtered through a 0.22 µm membrane (potassium acetate matrix, Millex GV4, Millipore Corporation, USA) prior to injection.

2.2.4a Anion exchange HPLC (AEHPLC)

Salivary gland extract (100 salivary glands), prepared in buffer A [20 mM Tris-HCl, pH 7.6] was filtered through a 0.22 µm filter and applied to an anion exchange column (DEAE-5PW, 7.5 mm x 7.5 cm, TosoHaas), pre-equilibirated with buffer A. The proteins were fractionated using a linear gradient of 0-1 M NaCl with buffer B [20 mM Tris-HCl, 1 M NaCl, pH 7.6] as set out in Table 2.1a. The absorbance was monitored at 280 nm and the flow rate was 1 ml/min. Fractions (1 ml) were collected and assayed for anti-thrombin activity.

Table 2.1a: Flow conditions used during AEHPLC

Time (min.)	Flow (ml/min.)	% Buffer A	% Buffer B	Duration (min.)
0	1	100	0	5
5	1	40	60	21
26	1	0	100	10
36	1	100	0	0.5
36.5	1	100	0	9
45.5	End			

Buffer A – 20 mM Tris-HCl, pH 7.6

Buffer B – 20 mM Tris-HCl, 1 M NaCl, pH 7.6

2.2.4b Hydrophobic interaction HPLC (HIHPLC)

Fractions from AEHPLC were adjusted to 2 M using 4 M ammonium sulphate and applied to a hydrophobic interaction column (TSK-Phenyl-5PW, 7.5 mm x 7.5 cm, Bio-Rad, Richmond, California). The column was pre-equilibrated with buffer A (20 mM Tris-HCl, 1.7 M $(NH_4)_2SO_4$, pH 7.4). Fractions were eluted with a linear gradient (1.7 – 0 M $(NH_4)_2SO_4$) with buffer B (20 mM Tris-HCl, pH 7.4) over 15 minutes, as set out in Table 2.1b. The flow rate was 1 ml/min and the absorbance was followed at 280 nm. Fractions (1 ml) were collected and assayed for inhibitory activity.

Table 2.1b: Flow conditions used during HIHPLC

Time (min.)	Flow (ml/min.)	% Buffer A	% Buffer B	Duration (min.)
0	1	100	0	5
5	1	0	100	10
15	1	0	100	10
25	End			

Buffer A – 20 mM Tris-HCl, 1.7 M $(NH_4)_2SO_4$, pH 7.4

Buffer B - 20mM Tris-HCl, pH 7.4

2.2.4c Reversed phase chromatography (RPHPLC)

Fractions from HIHPLC were desalted by dialysis (room temperature, overnight) and applied to a C_{18} column (Jupiter 5.6 mm x 25 cm, Phenomenex, USA), pre-equilibrated with mobile phase A (0.1 % TFA, 0.1 % acetonitrile). Elution was achieved with a linear gradient of acetonitrile (0 – 60 %) with mobile phase B (0.1 % TFA, 60 % acetonitrile) over 60 minutes, as set out in Table 2.1c. The flow rate was 1ml/min and the absorbance was monitored at 230 nm. Fractions (1 ml) were collected dried in a vacuum concentrator (Bachoffer) and assayed for anti-thrombin activity.

Table 2.1c: Flow conditions used during RPHPLC

Time (min.)	Flow (ml/min.)	% Buffer A	% Buffer B	Duration (min.)
0	1	100	0	5
5	1	0	100	60
65	1	0	100	10
75	1	100	0	0.5
75.5	1	100	0	10
85.5	End			

Buffer A – 0.1 % TFA, 0.1 % acetonitrile

Buffer B – 0.1 % TFA, 60 % acetonitrile

2.2.5 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

2.2.5.1 Glycine SDS-PAGE analysis

Electrophoresis was performed according to the method described by Laemmli (1970). A 12 % separating gel (0.375 M Tris-HCl, 0.1 % SDS, pH 8.8) and 4 % stacking gel (0.125 M Tris-HCl, 0.1 % SDS, pH 6.8) were prepared from a 30 % acrylamide/0.1 % bisacrylamide stock solution. These solutions were degassed for 30 minutes and polymerised by addition of 50 µl of 10 % ammonium persulphate and 5 µl of TEMED.

Samples were diluted 1:2 in reducing buffer [(0.06 M Tris-HCl, pH 6.8, 2 % SDS (w/v), 0.1 % glycerol (v/v), 0.05 % β-mercaptoethanol (v/v) and 0.025 % bromophenol blue (w/v)] and boiled at 94 ^{o}C for 10 minutes. Low MW markers (boiled in 100 µl reducing buffer) were applied per well (10 µl) as well as samples (30 µl/well). The low MW markers were phosphorylase b (94 kDa), BSA (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20 kDa) and α-lactoalbumin (14.5 kDa). Electrophoresis was carried out in running buffer (0.02 M Tris-HCl, 0.1 M glycine and 0.06 % SDS, pH 8.3) using a Biometra electrophoresis system with an initial voltage of 60 V for 45 minutes and then a voltage of 100 V for a further two hours.

2.2.5.1a) Coomassie staining

Proteins were visualised by staining with 0.1 % coomassie brilliant blue (40 % methanol, 10 % acetic acid) and destained in an excess of destaining solution (40 % methanol, 10 % acetic acid).

2.2.5.1b) Silver staining

The silver staining procedure of Morrisey (1981) was used. The gel was fixed in 100 ml of fixing solution (40 % methanol, 10 % acetic acid) for 30 minutes. The fixing solution was poured off and the gel immersed in 50 % methanol and 10 % acetic acid in water and agitated slowly for 30 minutes. The solution was poured off and the gel was covered with 50 ml of a 10 % glutaraldehyde solution (Sigma) and agitated slowly for 10 minutes in a fume hood. After the glutaraldehyde was poured off, the gel was washed thoroughly with several changes of water for 2 hours to ensure low background levels. The gel was then soaked in 100 ml of a 5 µg/ml DTT (Sigma) solution for 30 minutes.

The DTT was discarded and without rinsing the gel, the gel was immersed in 100 ml of a 0.1 % silver nitrate solution (Merck, Darmstadt Germany) and agitated for 30 minutes. After the silver nitrate was decanted, the gel was thoroughly rinsed with water. The gel was then soaked in 100 ml of the carbonate developing solution (0.5 ml of a 37 % formaldehyde solution/litre, 3 % sodium carbonate) and agitated slowly until the desired level of staining was achieved. The staining reaction was stopped by adding 2.3 M citric acid solution per 100 ml of carbonate developing solution. The gel was then rinsed in water for another 30 minutes.

2.2.5.2 Tricine SDS-PAGE analysis

Purified proteins were analyzed using a tricine SDS-PAGE system (Schägger & von Jagow, 1987) that is suitable for resolution of proteins in the range of 1 – 100 kDa. A 16.5 % T, 3 % C separating gel (1 M Tris-HCl, 0.1 % SDS, pH 8.45) and a 4 % T, 3 % C stacking gel (0.75 M Tris-HCl, 0.075 % SDS, pH 8.45) was prepared from acrylamide (48 % acrylamide/ 1.5 % N', N'-methylene bisacrylamide) and

electrophoresis buffer (3 M Tris-HCl, pH 8.45, 0.3 % SDS) stock solutions.

These solutions were degassed for 10 minutes and polymerized by addition of 50 µl 10 % ammonium persulphate and 5 µl of TEMED. Protein was diluted 1:2 in reducing buffer [0.06 M Tris-HCl, pH 6.8, 2 % SDS (w/v), 0.1 % glycerol (v/v), 0.05 % β-mercaptoethanol (v/v) and 0.025 % bromophenol blue (w/v)] and boiled at 94 ^{o}C for 10 minutes. Low molecular mass peptide markers were used for molecular mass determination. Unfolding of inhibitors in the presence of urea was investigated by including 8 M urea in the tricine separating gel.

The gel system used was from Biometra, with an initial voltage of 30 V for 1 hour and then a voltage of 100 V for a further 2 hours. The anode buffer was 0.2 M Tris, pH 8.9 and cathode buffer, 0.1 M Tris, 0.1 M Tricine, 0.1 % SDS, pH 8.25. The tricine gel was visualised by silver staining.

2.2.6 Matrix assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI –TOF MS)

MALDI-TOF MS was performed to determine the purity of savignin. Savignin (5 pmol) was crystallised in an equal volume of α-cyano-4-hydroxycinnamic acid matrix before analysis with a MALDI-TOF mass spectrometer (Micromass) with a laser intensity of 1 200 V, accelerating voltage (2 000 V), grid voltage (91 - 92 %), guide wire voltage (0.1 %) using a linear flight path. The spectrometer was calibrated using a calibration mixture of external standards (ubiquitin I, cytochrome C, myoglobin and insulin).

2.2.7 Electrospray mass spectrometry (ESMS)

Electrospray mass spectrometry was performed by Dr. M. J. van der Merwe (Department of Biochemistry, University of Stellenbosch). A VG Micromass Quattro triple quadrupole mass spectrometer equipped with an electrospray ionisation source was used (Micromass, UK). The capillary voltages applied in the positive mode was 3.5 kV, the source temperature 80 ^{o}C and cone voltage 70 V with the skimmer lens set at 5 V. All other lenses were turned for maximum sensitivity. About 300 pmol of protein was dissolved in 100 µl of 50 % acetonitrile containing 0.1 % formic acid. Application of samples was by injection of 10 µl into a stream of 50 % aceonitrile at a flow speed of 10 µl/minute supplied from a LKB/Pharmacia 2249 gradient pump.

Data acquisition was in the continuum mode, scanning from m/z 500 to 2 000 atomic mass units per second (amu/s). The instrument was calibrated with the multiple charged spectrum of horse heart myoglobin, acquired at the same resolution as unknown proteins.

2.2.8 Amino acid analysis

Dried fractions of savignin were dissolved in 25 µl ddH_2O and hydrolyzed with 200 µl of (6 N HCl, 1 % phenol) in a closed vessel. The vessel was sealed and evacuated after flushing with N_2 gas and placed in an oven at 110 ^{o}C for 24 hours. After hydrolysis, 10 µl of a mixture containing methanol:water:triethylamine (2:2:1) was added to the sample and dried under vacuum. This was followed by derivatization using derivatizing agent (methanol:water:triethylamine:phenylisothiocyanate, 7:1:1:1) at room temperature for 20 minutes and vacuum dried. Dried derivatized sample was dissolved in 100 µl of 5 mM Na_2HPO_4, 10 % H_3PO_3, 5 % MeCN (pH 7.4) and filtered through a 0.45 µm membrane (Microsep, USA).

Samples (20 µl) were subsequently subjected to the reversed phase RP-HPLC column (PICO.TAG, 3.9 mm x 15 cm), and eluted with a gradient from buffer A (0.14 M NaOAc.$3H_2O$, pH 5.7) and buffer B (60 % acetonitrile). The data was analyzed by System Gold software (Beckman).

2.2.9 Temperature stability assay

For the temperature stability assay, savignin (10 ng of protein in 50 µl of 20 mM Tris-HCl, 1 M NaCl buffer, pH 7.6) was incubated at 94 °C for different time periods and after each incubation placed on ice for 10 minutes. Samples were centrifuged (14 000 *g* for 15 minutes at room temperature) and supernatants were tested for residual activity with the chromogenic substrate assay as described in 2.2.3.

2.2.10 Size exclusion chromatography (SEC) for determination of hydrodynamic radii

Sephadex G-50 matrix (Pharmacia) was pre-swollen in water by boiling at 94 °C for one hour. The column (120 cm x 35 mm) was then manually packed to give a column height of 116 cm. This resulted in a void volume (Vo) of ~ 40 ml and a total volume (Vt) of 120 ml. The column was pre-equilibrated with buffer (20 mM Tris-HCl, 0.15 M NaCl, pH 7.6) and purified savignin (25 µg) or salivary gland extract (50 salivary glands, ~ 2000 µg protein) were applied to the column. The column was calibrated using 2 mg of both chymotrypsinogen (25 kDa) and lysozyme (14 kDa), both obtained from Pharmacia. The column was eluted isocratically at a flow rate of 2 ml/min and 2 ml fractions were collected and assayed for anti-thrombin activity. The absorbance was monitored at 280 nm. The AKTA prime system (Amersham) was used for solvent delivery, spectrophotometric monitoring of fractions and fraction collection.

2.3 Results

2.3.1 Calculation of hydrodynamic radii of the extended and compact forms of savignin

Hydrodynamic radii using the HYDROPRO program were calculated for ornithodorin and bikunin which were used as models, of the extended and compact forms of savignin respectively (Table 2.2; Figure 2.4). Values calculated for chymotrypsinogen and lysozyme are also shown. The hydrodynamic radii for bikunin and lysozyme are similar (2 nm), ornithodorin and chymotrypsinogen are also similar (2.3 nm). This confirms that lysozyme and chymotrypsinogen will be suitable references to use for measurement of the hydrodynamic radius of uncomplexed savignin. It also indicates that ornithodorin will, due to its extended conformation, behave like a protein twice its molecular mass (25 kDa) even though it has a molecular mass of 12 kDa (Figure 2.5).

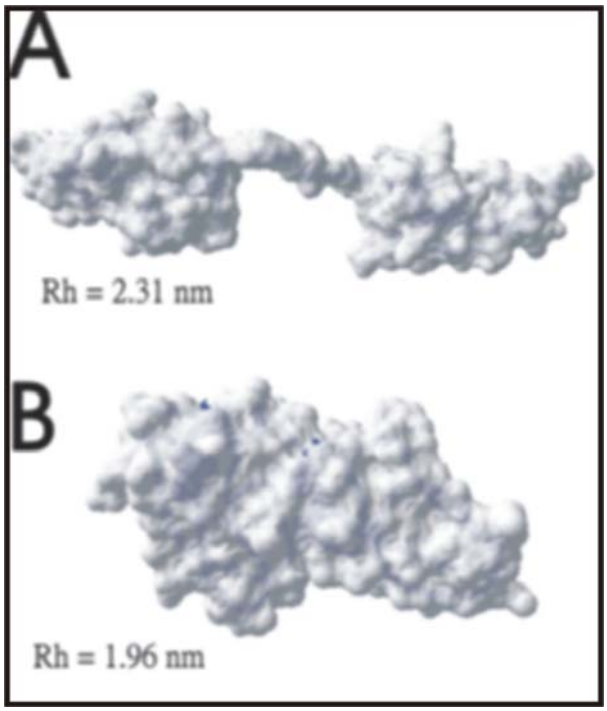


Figure 2.5 Rh values of savignin determined by the HYDROPRO programme. A): extended form. B): compact or globular form.

Table 2.2 Hydrodynamic radii calculated for bikunin, ornithodorin, lysozyme and chymotrypsinogen. Also indicated are their molecular masses and Protein Databank Accession Codes (PDB).

Protein	**PDB**	**MW (Da)**	**Rh (nm)**
Bikunin	1BIK	12048	1.965
Ornithodorin	1TOC	12430	2.319
Lysozyme	6LYS	14320	1.988
Chymotrypsinogen	2CGA	25660	2.316

The relationship between the volume and molecular mass of a molecule can be more formally explained as a dependence of its molecular mass and its hydrodynamic radius (Figure 2.4) (Cabré, Canela & Canela, 1989). Using the equation that relates hydrodynamic radius and molecular mass ($Rh = 0.6452X^{0.4115}$), a hydrodynamic radius for savignin can be calculated as 1.84 nm from its molecular mass of 12 430 Da. The hydrodynamic radius calculated for uncomplexed savignin from its modelled structure is 2.31 nm, which corresponds to that of a molecule with molecular mass of 25 kDa.

The difference in the calculated hydrodynamic radius (~ 0.35 nm) between ornithodorin and bikunin would be sufficient to be distinguished by dynamic light scattering. We were thus confident that we should be able to determine the conformation of savignin in its uncomplexed form by measuring its hydrodynamic radius. Because dynamic light scattering needs high pure samples, the next step was to purify savignin.

2.3.2 Purification of savignin

Savignin was previously purified by size exclusion, anion exchange and reversed phase HPLC (Nienaber, Gaspar & Neitz, 1999). This strategy gave yields of ~ 0.148 % of the total protein or 5 µg from 35 salivary glands (Nienaber, personal communication). Low yields could have been due to adsorption of proteins to the gel matrix during SEC that was used to purify savignin from a contaminating 7 kDa protein (Nienaber, personal communication). At the time of this study, no size exclusion column was available to repeat the procedure previously described. As an alternative, a strategy was employed that included AEC, HIC and RPHPLC. Ion exchange HPLC was used as a first step in the purification process (Figure. 2.6). A DEAE-5PW anion exchange column was used for fractionation based on charge differences. Based on Nienaber 1999, the anti-thrombin activity eluted at 16 - 17 minutes. Fractions in this range (14 - 20 minutes) were collected and assayed for inhibitory activity. Fractions corresponding to a retention time (RT) of 16 - 17 minutes were retained for further analysis.

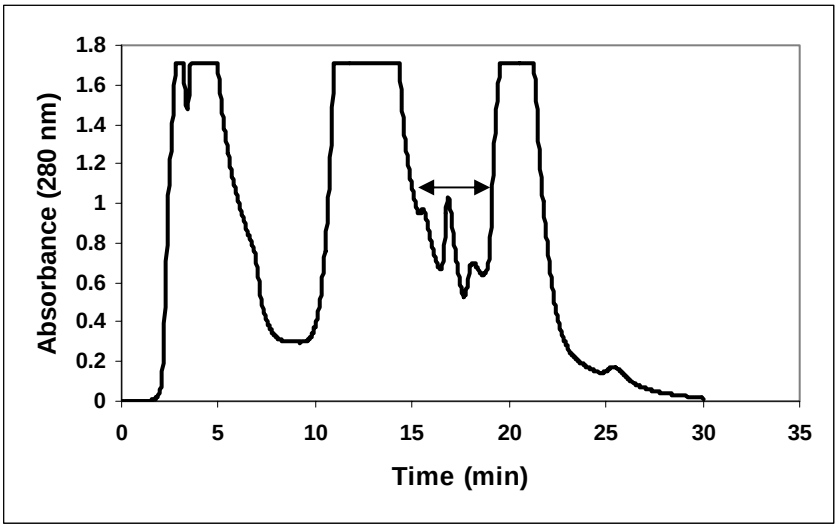


Figure 2.6 Anion exchange chromatography of salivary gland extracts. A linear gradient of 0 - 1 M NaCl was applied over 35 minutes starting at 5 minutes. Protein was monitored at 280 nm. One millilitre fractions were collected and 10 µl samples were assayed for inhibition of anti-thrombin activity. ↔ indicates fractions with anti-thrombin activity.

In an attempt to separate the 7 kDa contaminant previously observed, the pooled fractions were subjected to hydrophobic interaction HPLC (Figure 2.7). This was used as a second step in the purification process. Analysis of savignin revealed that fractions collected during a retention time between 23.5 - 27.1 minutes, exhibited the anti-thrombin activity.

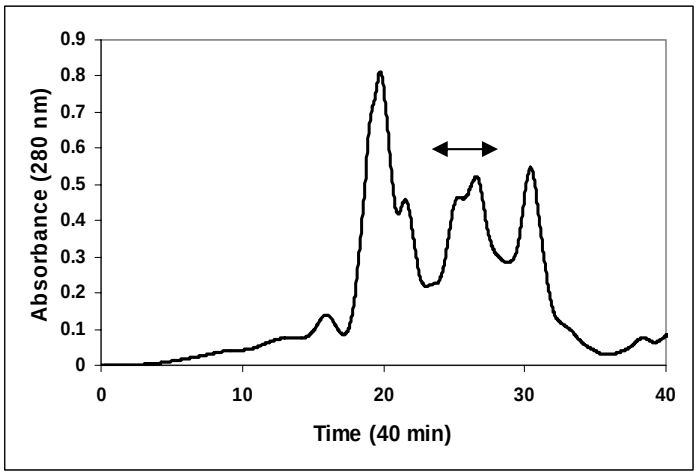


Figure 2.7 Hydrophobic interaction chromatography of the anti-thrombin activity obtained after anion exchange chromatography. Fractions (1ml) were collected and assayed for thrombin inhibition. ↔ indicates anti-thrombin activities.

The pooled fractions from HIC were subjected to reversed phase C18 chromatography. The chromatogram in Figure 2.8 indicates a distinctive peak at a retention time of 31 - 32 minutes.

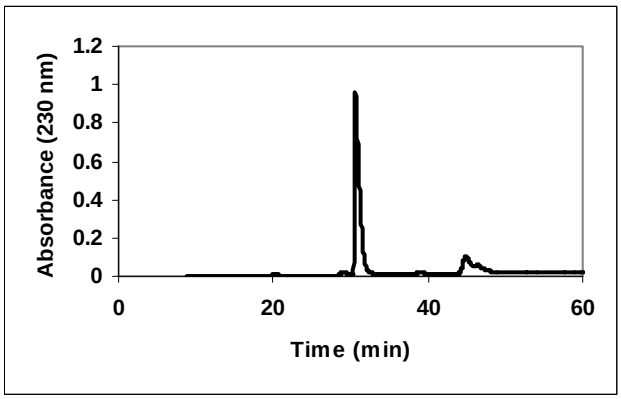


Figure 2.8 Reversed phase chromatography of the anti-thrombin activity obtained after hydrophobic interaction chromatography. A linear gradient of 0.1 – 60 % acetonitrile was applied after 5 minutes over period of 60 minutes.

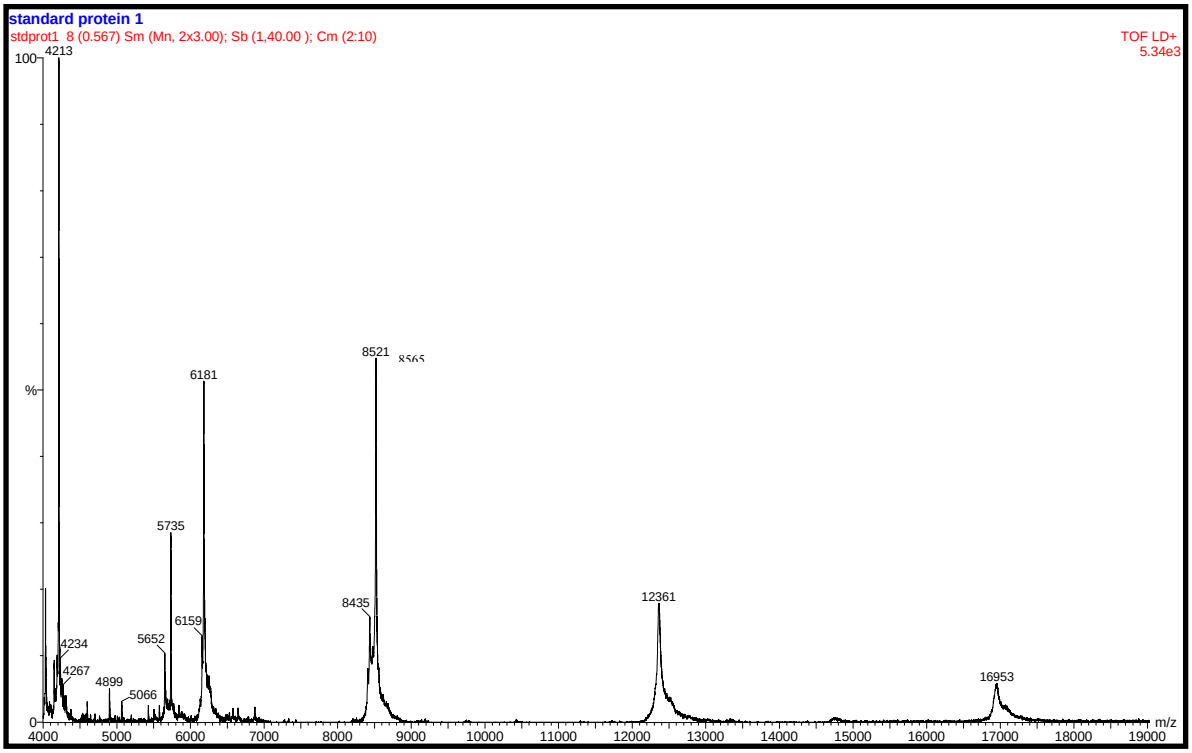
Amino acid analysis of this peak indicated a yield of 138 µg from 100 salivary glands. This corresponds to a yield of ~ 1.38 % of the total salivary gland extract protein and is almost 10 fold higher than previously obtained. In order to confirm that this protein was savignin, amino acid analysis, glycine SDS-PAGE, tricine SDS-PAGE, MALDI-TOF MS and ESMS were performed.

2.3.3 MALDI-TOF MS of purified savignin

The isolated protein was subjected to MALDI-TOF MS for the determination of its purity and molecular mass. To calibrate the MALDI-TOF MS, standard proteins included insulin $[M+H]^+$: 5 734.56 Da; ubiquitin I $[M+H]^+$: 8 565.89 Da; cytochrome C $[M+H]^+$: 12 361.09 Da; myoglobin $[M+H]^+$: 16 952.55 Da; cytochrome C $[M+2H]^{2+}$: 6 181.05 Da and myoglobin $[M+2H]^{2+}$: 8 476.77 Da.

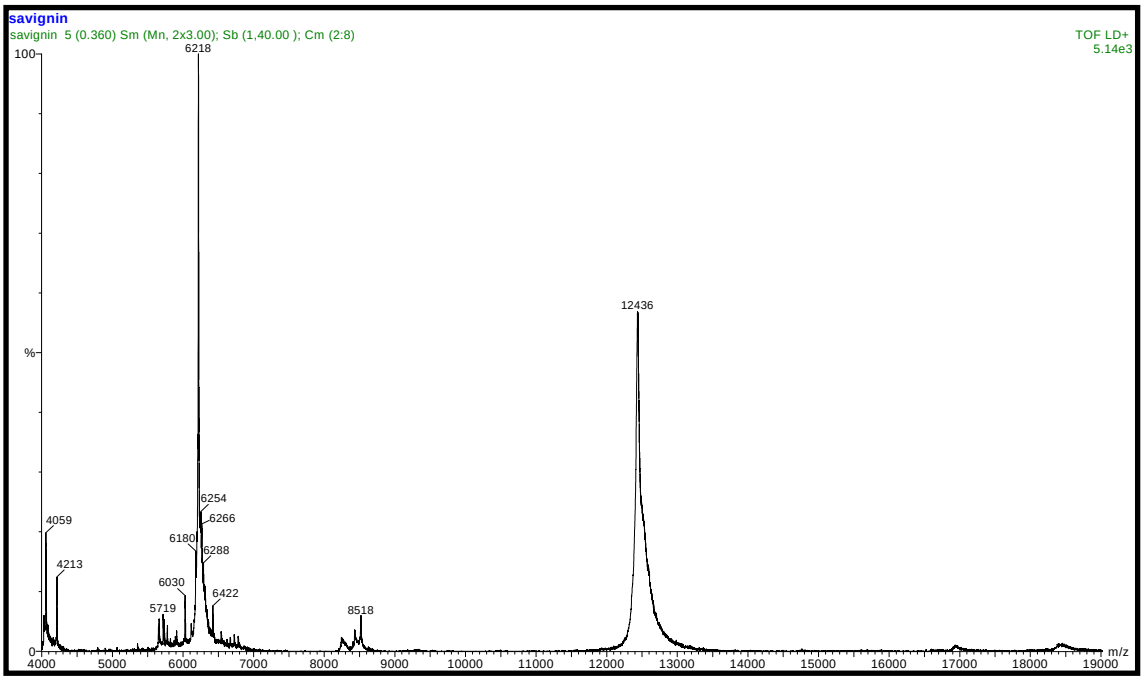
The standard protein profile is indicated in Figure 2.9A and shows that the MALDI-TOF MS was adequately calibrated. The mass obtained for purified savignin is for the single protonated form $[M+H]^+$: 12 436 Da (Figure 2.9B). Also indicated is the diprotonated form $[M+2H]^{2+}$: 6 218 Da.

A



m/z

B



m/z

Figure 2.9 MALDI-TOF MS analysis of calibration standard proteins and savignin. A) The M^{+} and M^{2+} protonated species of standard proteins. B) The M^{+} and M^{2+} protonated species of savignin

2.3.4 Electrospray mass spectrometry of savignin

To determine an accurate and precise molecular mass, savignin was subjected to ESMS. The mass spectrum in Figure 2.10 reveals the $(M+H)^+$ ion of a protein with a molecular mass of 12 440.6 Da, compared to 12 430.4 Da observed by Nienaber 1999. This result confirms the purity and identity of savignin.

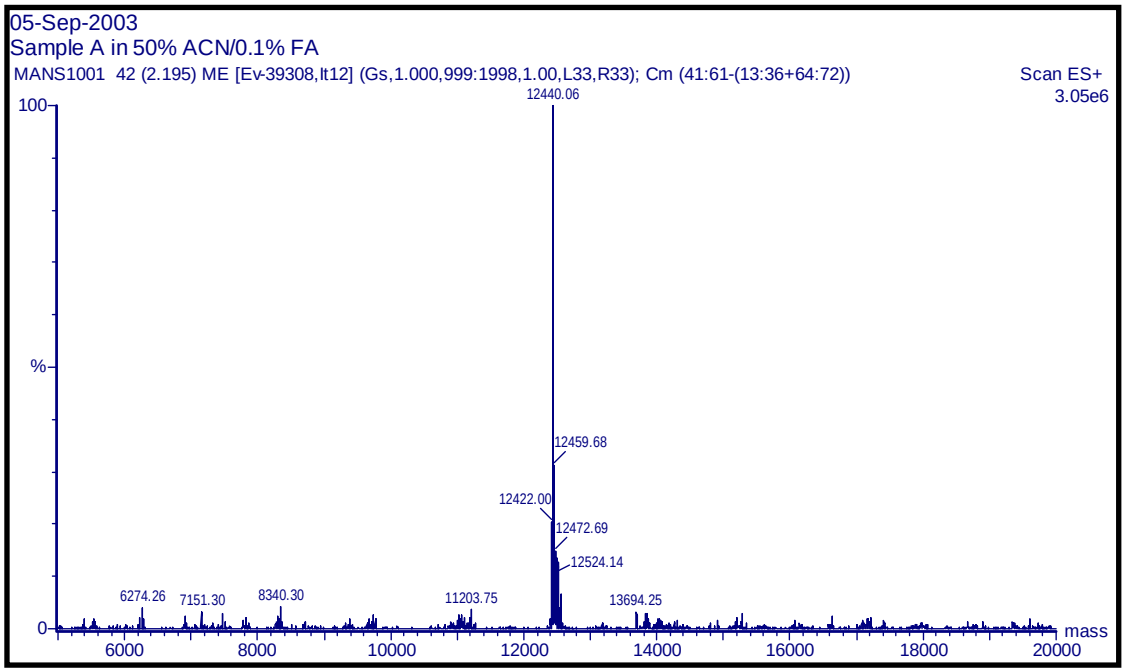


Figure 2.10 Electrospray mass spectrometry analysis of savignin. Indicated is the deconvoluted mass spectrum.

2.3.5 Amino acid analysis

Further confirmation of the identity of savignin was obtained by a comparison of the amino acid composition reported previously for savignin and the composition obtained in this study (Figure 2.11). This shows that the amino acid compositions for both purifications compare fairly well.

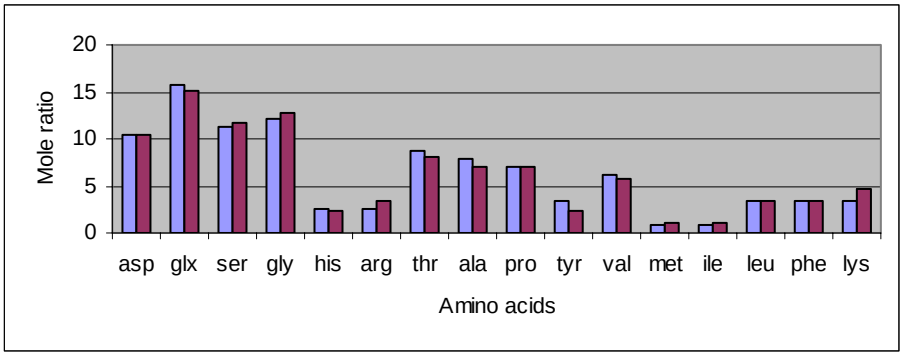


Figure 2.11 Amino acid analyses of savignin. Y axis: Indicates mole ratios of amino acids. The blue bars indicate the composition obtained by Nienaber, Gaspar & Neitz, 1999 and the purple bars indicate results of the present study.

2.3.6 Electrophoretic mobility of savignin

Previously, savignin showed anomalous behaviour during SDS-PAGE analysis by showing a lower electrophoretic mobility than its expected size and hence a higher apparent molecular mass (Nienaber, 1999). This anomalous behaviour could be explained by a compact conformation such as that observed for bikunin. In this case savignin would presumably not unfold completely during boiling with SDS, which would cause less SDS binding and hence a higher apparent molecular mass.

The behaviour of savignin during tricine SDS-PAGE was thus investigated further in this study. Savignin that was treated in the usual manner (heated for 5 minutes at 94°C in sample buffer containing reducing agent) and applied to a normal tricine SDS-PAGE gel, migrated at an apparent molecular mass of ~ 17 kDa (Figure. 2.12).

If savignin is truly in a compact form that binds less SDS, then incubation in the presence of urea should unfold it resulting in the expected amount of SDS being bound. Inclusion of 8 M urea in the tricine SDS-PAGE gel shifted the apparent molecular mass from ~ 17 kDa to ~ 30 kDa (Figure. 2.12).

In the absence of boiling, the band migrated faster; hence less unfolding occurred resulting in higher mobility due to smaller surface area.

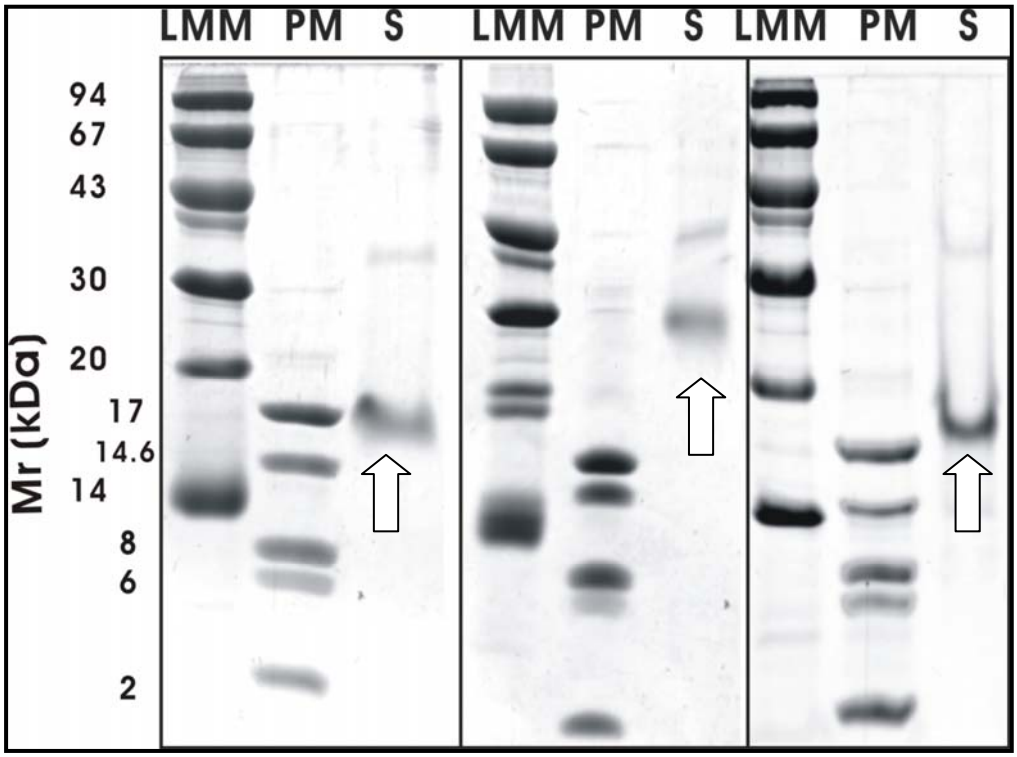


Figure 2.12 Tricine SDS-PAGE analysis of savignin. The panel on the left shows savignin under normal reducing conditions with mercaptoethanol (ME), in the middle panel 8M urea has been included in the separating gel (with ME) and the panel on the right shows savignin incubated with sample buffer that contains 8M urea and ME without boiling. LMM: low molecular mass marker. PM: peptide marker. S: savignin indicated with arrows.

2.3.7 Heat stability of savignin

To investigate the heat stability of savignin it was incubated at 94 °C for various time intervals (Figure 2.13). Savignin retained 100 % anti-thrombin activity after 2 hours at 94 ^{o}C. This supports the hypothesis that savignin might not unfold completely during sample preparation for SDS-PAGE analysis.

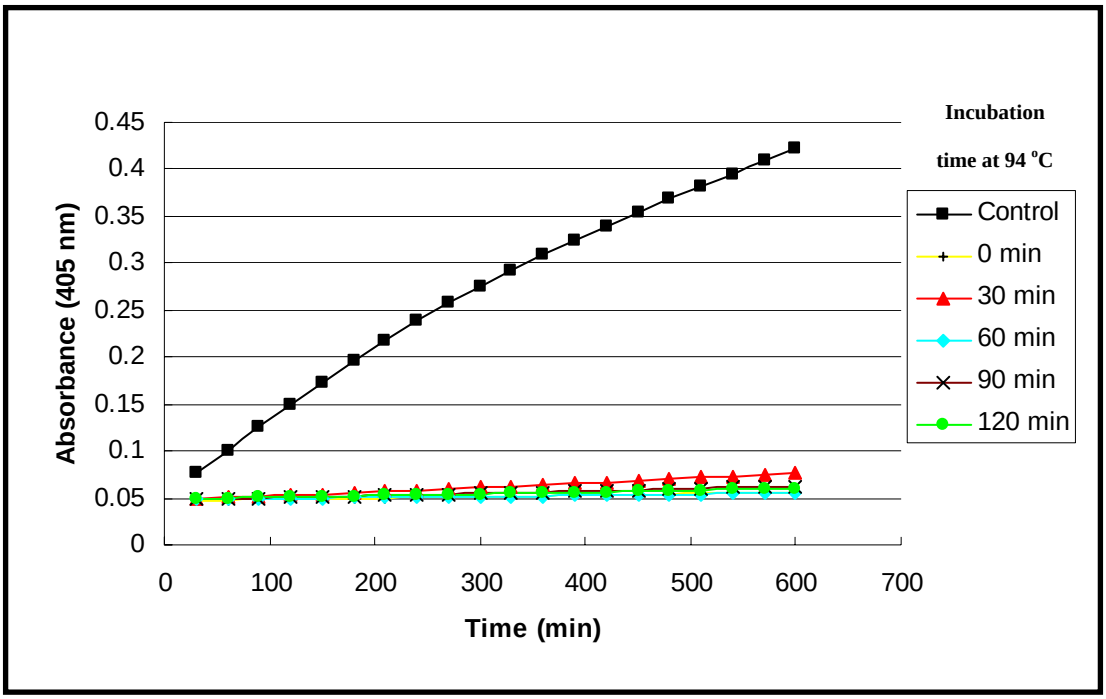


Figure 2.13 The anti-thrombin activity of heat-treated savignin. The progression curves of thrombin activity (0.2 U) in the presence of savignin that was incubated at 94 °C for various time periods as indicated. The control indicates thrombin activity in the absence of savignin.

2.3.8 Measurement of the hydrodynamic radius of savignin by dynamic light scattering

Light scattering equipment is currently available in South Africa at the Chemistry Department of the University of South Africa. According to initial enquiries with both the Head of the Chemistry Department and a representative of the company that supplies the equipment (Microsep), the equipment (Precision Detector Model 2000 Dual angle light-scattering detector) was adequate for the measurement of hydrodynamic properties of proteins. This had an important influence on the experimental design of this study and its subsequent execution, which included the purification of savignin in high yield and to a high level of purity.

Only later when we performed measurements on-site with the light-scattering equipment, did we discover that the equipment did not have all the necessary software

for the measurement of the hydrodynamic radius. Measurements of the radius of gyration were possible, but only for molecules with a radius above 12 nm. Therefore an alternative approach was used to measure hydrodynamic radii (section 2.3.9).

2.3.9 Measurement of the hydrodynamic radius of savignin by size exclusion chromatography

The inability of the light scattering equipment to measure Rh values limited the way in which we could test the hypothesis that uncomplexed savignin had a compact globular form. An alternative way was to measure the Rh of savignin using size exclusion chromatography. We confirmed that size exclusion would be suitable to distinguish between the two proposed forms of savignin, by separating lysozyme and chymotrypsinogen to almost base-line separation using a Sephadex G-50 column (Figure 2.14A). Chymotrypsinogen eluted at 30 minutes and lysozyme at 45 minutes.

We could thus predict that savignin will elute close to 30 minutes if it exists in an extended conformation or else close to 45 minutes if it exists in a compact conformation (Mans, 2002). Purified savignin (25 µg, which was all that was left from the purification) was applied to the size exclusion column, but no detectable peak could be observed, even when the detector was set to its most sensitive setting. This was not surprising, considering the fact that the peak heights observed for lysozyme and chymotrypsinogen are for a total concentration of 2 mg applied to the column. Therefore, and because savignin is the predominant inhibitor of thrombin in salivary gland extracts (SGE), SGE was used in further experiments (Figure 2.14B).

A large SEC column was used in this study because a SEC HPLC column was not functional at the time of study.

The major inhibitory activity eluted at 30 minutes which was similar to that of chymotrypsinogen (Figure 2.14 C). This indicated that the major conformation of active savignin was present as the extended conformation. Another region, between 42 - 47 minutes, displayed minor inhibitory activity and was found to be in the same region where lysozyme (compact form) had been detected. High background signals in this region were however observed. A final region, between 52 – 56 minutes, also displayed inhibitory activity which may due to the presence of second unidentified thrombin inhibitor.

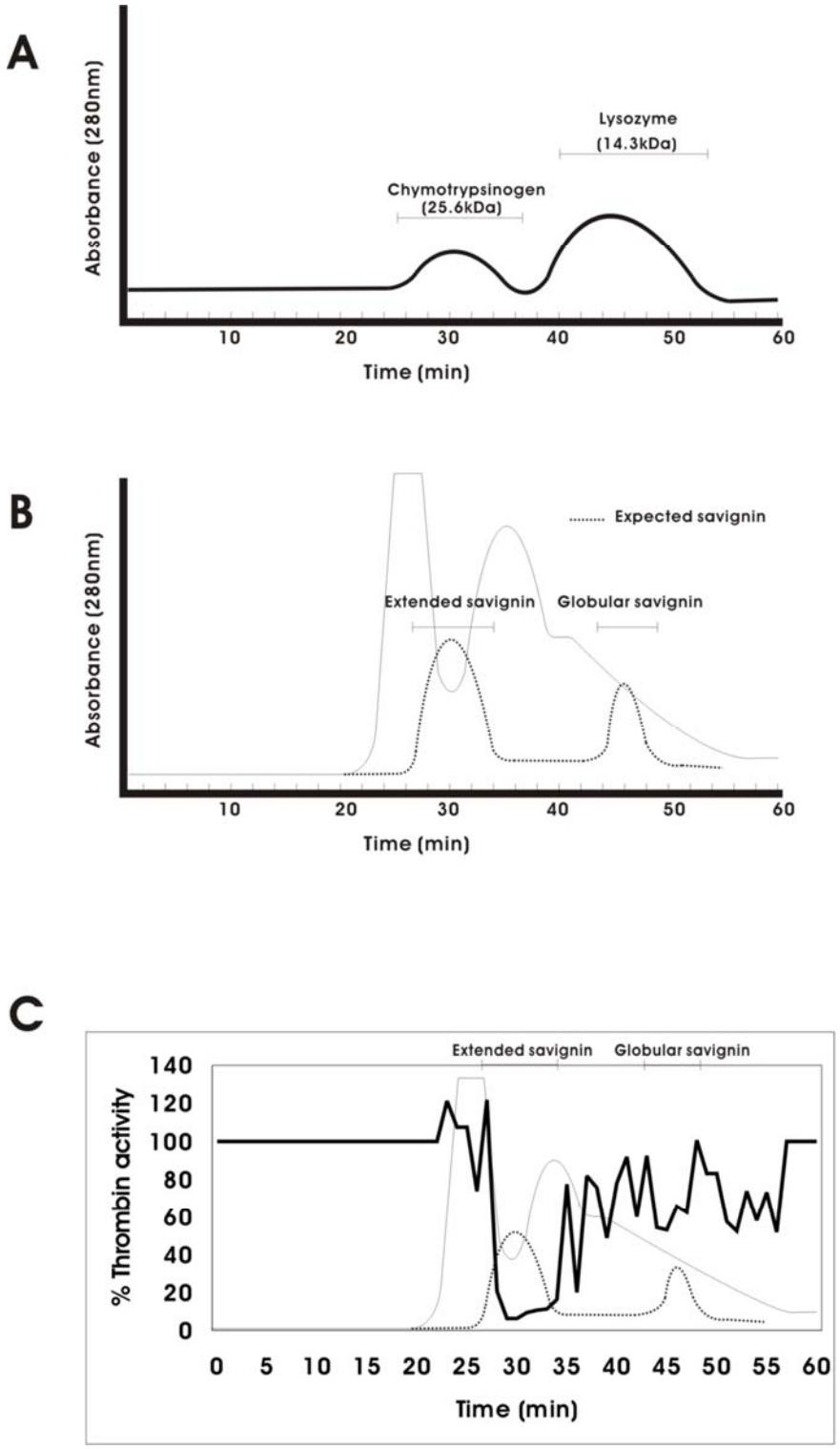


Figure 2.14 Measurement of Rh of savignin by using SEC. A) Size exclusion chromatography of chymotrypsinogen and lysozyme. B) Size exclusion chromatography of crude salivary gland extracts. Solid line represents the elution profile of the salivary gland extract. Dashed line represents the expected extended and globular form of savignin. C) The percentage inhibition of thrombin of collected fractions from the size exclusion column.

2.4 Discussion

Savignin is classified as a slow, tight binding competitive inhibitor of thrombin. It has also been indicated that the fibrinogen-binding exosite of thrombin is important for savignin's inhibitory activity (Nienaber, Gaspar & Neitz, 1999). A molecular model of savignin was constructed which was based on the crystal structure of ornithodorin (Mans, Louw & Neitz, 2002). This model was docked with the structure of thrombin and suggests that the C-terminal domain of savignin does indeed interact with thrombin's fibrinogen-binding exosite (Mans, Louw & Neitz, 2002). It also suggests that the N-terminal domain of savignin binds with its N-terminal residues into the active site of thrombin. This model of interaction is based on the crystal structure of the ornithodorin-thrombin complex (van de Locht et al., 1996). No kinetic data has been reported for ornithodorin (van de Locht et al., 1996). The savignin-thrombin complex indicates that there are two binding sites for savignin (Mans, Louw & Neitz, 2002). This could explain the tight-binding kinetics observed for savignin (Nienaber, Gaspar & Neitz, 1999). It does not however, explain the slow-binding kinetics. Slow-binding kinetics indicates some form of structural rearrangement (Sculley, Morrison & Cleland, 1996). If two binding sites exist, it is probable that the inhibitor first binds to one site and then to the other. The hypothesis of Mans (2002) tried to give a mechanistic description of this, which would also explain some of the anomalous electrophoretic behaviour observed for savignin.

The hypothesis of Mans (2002) was tested by determining the hydrodynamic radius of uncomplexed savignin. It was found that the hydrodynamic radius of uncomplexed savignin is similar to that of chymotrypsinogen, as inferred from its similar elution behaviour during SEC. The hydrodynamic radius of chymotrypsinogen and

complexed ornithodorin as calculated from their crystal structures is similar as well. In contrast, the calculated hydrodynamic radius of bikunin and lysozyme are similar and differ significantly from ornithodorin. It is concluded that uncomplexed savignin exists in an extended conformation that correlates with that of the crystal structure of complexed ornithodorin.

If savignin exists in an extended conformation, then the only satisfying explanation for its anomalous behaviour during electrophoresis is that the N- and C-terminal domains are very stable on their own and do not unfold under usual denaturing conditions. This would also explain the thermal stability of savignin (Figure 2.15).

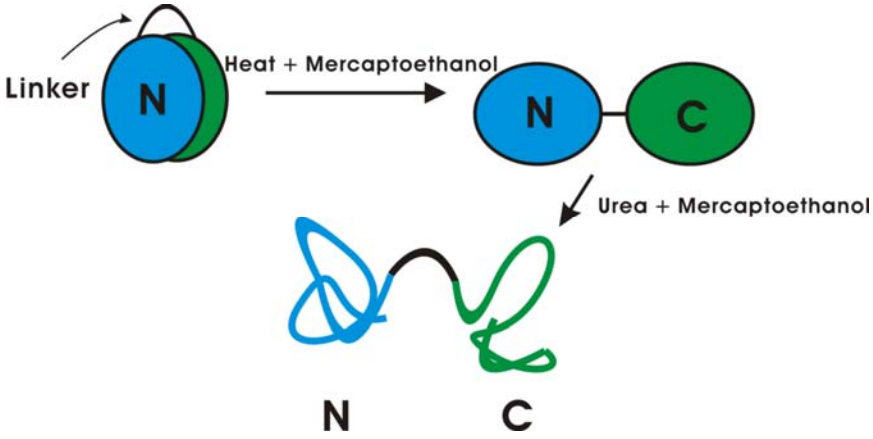


Figure 2.15 Unfolding of savignin. Blue indicates N-domain, green indicates C-domain, black line indicates the linker region. Two-steps of unfolding are depicted: unfolding to an extended form and subsequently unfolding of the two globular domains.

One of the possible explanations of the electrophoretic behaviour that was observed, was that the compact form was opened up and formed the extended form when it was exposed to heat and a reductive environment and further reduced into a linearized protein chain.

During the present study a method to purify savignin in higher yields was developed. This purification method might be used in the future to purify more native inhibitor for the purpose of studying the kinetics of savignin in greater detail.

While some support could be found for a compact conformation of savignin, the present study did not show whether savignin binds to thrombin in a two step manner as proposed by Mans (2002). Targeting of savignin to the fibrinogen-binding site of thrombin by its C-terminal BPTI-domain and subsequent binding to thrombin's active site with its N-terminal domain is still a viable hypothesis that can explain the slow-binding kinetics observed for savignin. In order to investigate this, binding data for savignin with thrombin and the individual domains with their binding sites is needed. Site-directed mutagenesis of savignin could also help to test this hypothesis.

Results from SEC suggested the presence of both globular and extended conformations. Inhibitory activity was found in both the extended form (major inhibition) and globular form (minor inhibition). There was a third region (retention time 52 - 56 minutes) which also displayed partial inhibition, which may be due to an even more compact form, or to the presence of a unidentified second thrombin inhibitor.

Generation of adequate concentrations of savignin could be used for further structural studies, for example to determine the structure of savignin in the uncomplexed form and in complex with thrombin. These studies would require molecular biological techniques. The rest of the thesis thus deals with an investigation into the recombinant expression of the N- and C-terminal domains of the inhibitor as well as the full length savignin as this will facilitate further kinetic analysis.

Chapter 3

Investigation into the recombinant expression of savignin and its N- and C- domains

3.1 Introduction

An in-depth characterization into the interaction of savignin with its target enzyme requires large quantities of material that cannot be isolated from its natural source. Recombinant production of savignin and its N- and C-domains will enable the elucidation of the mechanism of inhibition of savignin since this was not possible in previous investigations using the native inhibitor (Nienaber, Gaspar & Neitz, 1999). A comparison between the kinetics of the full-length savignin as well as the N- and the C-terminal domain will also be possible and allow the determination of the dissociation constants (Ki) for savignin and for the separate domains.

The production of recombinant savignin will also allow the *in vitro* evaluation of the inhibitor's anticoagulant potential. These studies were initially performed on crude SGE and could not be repeated on the purified native inhibitor due to insufficient material.

There are various expression systems, such as yeast and bacteria, which have been used in the production of eukaryotic proteins. Various points need to be taken into consideration when deciding which expression system to use such as, cellular localization and solubility of the target protein.

Ornithodorin was successfully expressed in the pMAL-p2 expression system (van de Locht, patent 1996). Savignin has similar characteristics to the ornithodorin protein and therefore the pMAL-p2 expression system was used to express savignin although other factors were taken into consideration.

The full length gene of savignin (Figure 3.1A) was previously obtained using 3'– and 5'–rapid amplification of cDNA ends (RACE) and subsequently sequenced and cloned into pGEM-T easy plasmid (Mans, Louw & Neitz, 2002). In this study the polymerase chain reaction (PCR) was used to amplify the cDNAs encoding the individual N-terminal and C-terminal domains of savignin as well as the full length inhibitor. The gene for full-length savignin (*Fsav*), cloned into the pGEM-T easy vector (Figure 3.2A) was used as a template for PCR synthesis of savignin and its C-domain (*Csav*). The gene for the N-domain was cloned into the same vector. This was then used as a template for PCR amplification of the N-domain (*Nsav*).

The forward (Nsavkpn-, Csavkpn-, NfXa-, and CfXa) primers containing the sequence of the AvaI restriction enzyme site (5'-CTC GGG-3') and either enterokinase (kpn) or factor Xa (fXa) recognition sites as 5' overhangs were used for amplification at the N-terminals of these constructs (Figure. 3.1B). The sequence of the reverse primer (Sp6) is indicated in Figure 3.1B. cDNA molecules were subcloned into the pGEM-T easy vector for subsequent sequencing and this was followed by cloning genes into the pMAL-p expression plasmid to generate recombinant products. Proteins were then expressed either in TB1 or *Sure E.coli* cells.

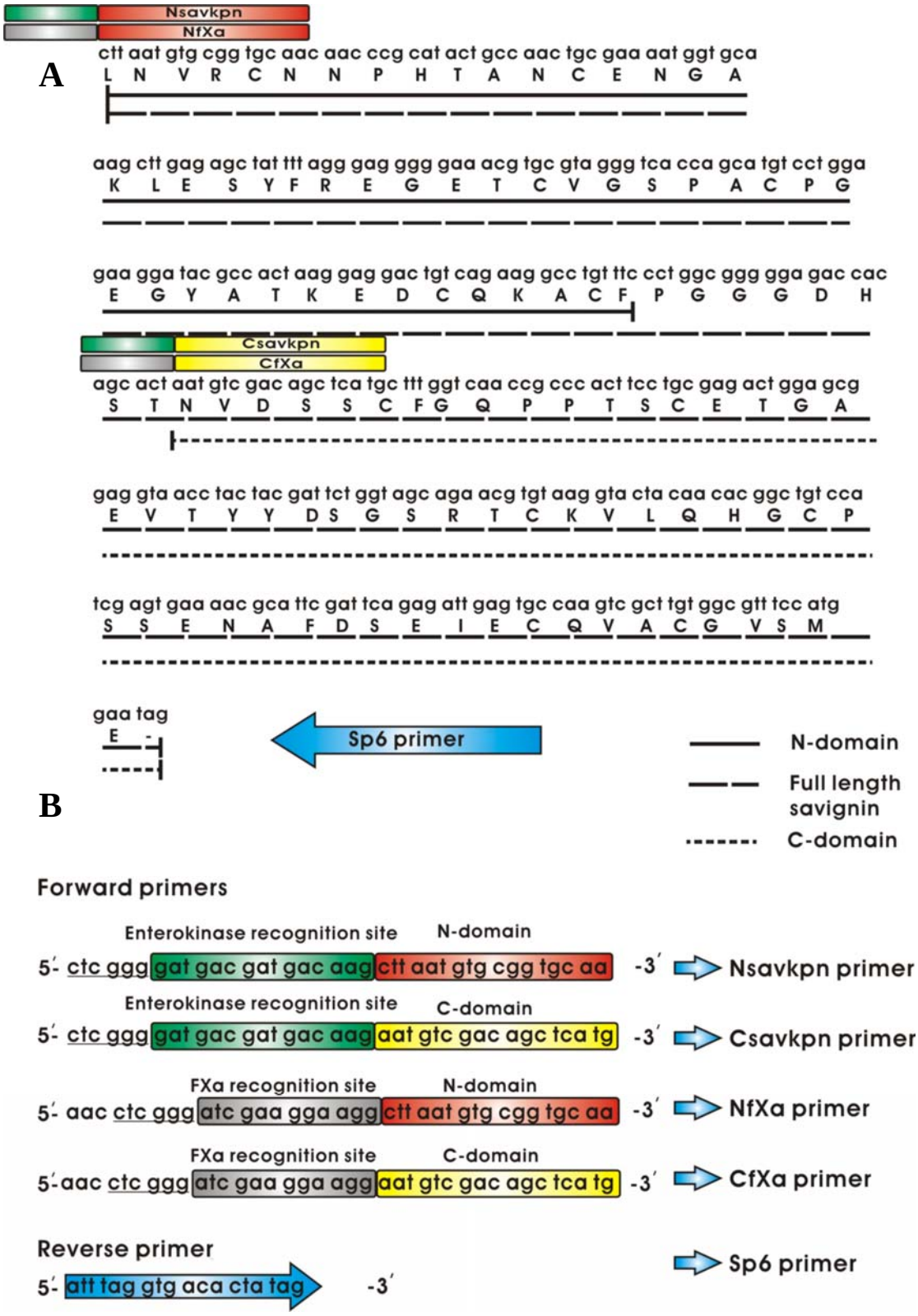


Figure 3.1 A): Sequence of full-length savignin. B): Primers used in the production of full length, N- and C-domains of savignin. Underlined: AvaI restriction enzyme site. Green: enterokinase protease recognition site. Grey: factor Xa protease recognition site. Red: N-terminal sequence of N-domain savignin. Yellow: N-terminal sequence of C-domain savignin.

Another reason as to why the pMAL-p2 expression system was used was to produce active and properly folded savignin and its truncated N- and C-domains. This method uses the strong "tac" promoter and the malE translation initiation signal to give high-level expression of the cloned sequences (Duplay et al., 1984).

With this system, the cloned gene is inserted downstream from the malE gene of pMAL-p2 plasmid which encodes the maltose-binding protein (MBP) (Figure 3.2B), resulting in the expression of a MBP fusion protein (Guan et al., 1987; Maina et al., 1988).

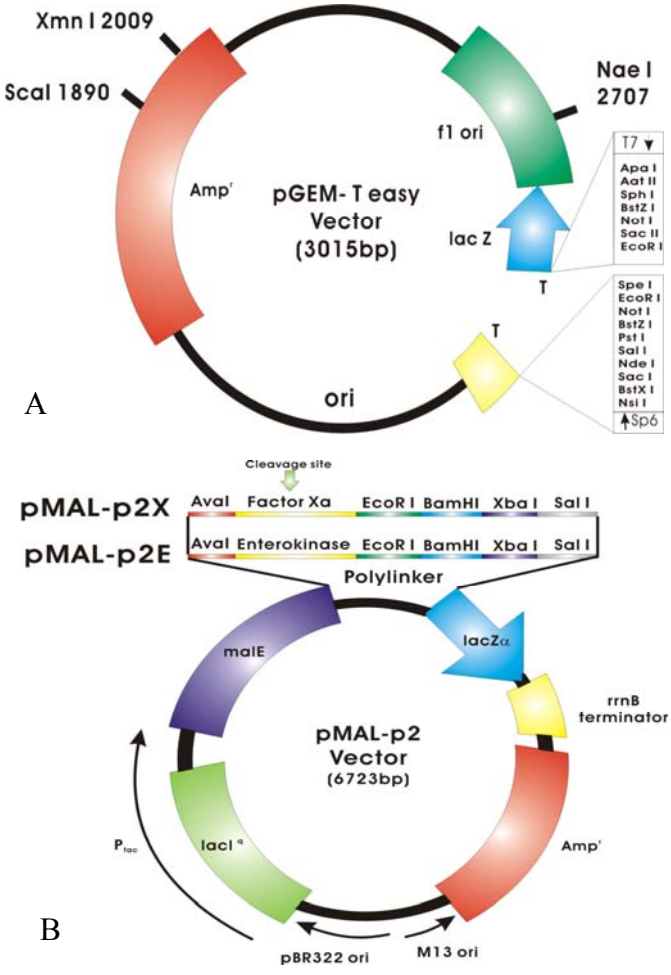


Figure 3.2 A): pGEM-T easy plasmid cloning vector used for both cloning and sequencing of the N-domain, C-domain and full length savignin. B): pMAL-p2 plasmid is an expression vector used for the expression of the N-domain, C-domain and full length savignin sequences.

When present, the signal peptide on MBP directs fusion proteins to the periplasm (Duplay et al. 1984). This allows for successful folding and disulfide bond formation in the periplasm of *E. coli* and is especially suited for eukaryotic secretory proteins. Translocation into the periplasm also serves as an enrichment step during purification of the recombinant proteins. The pMAL vectors carry the *lacI*^{q} gene, which code for the Lac repressor. This keeps expression of P_{tac} low in the absence of isopropyl-β-D-thiogalactopyranoside (IPTG) induction. The pMAL-p2 plasmids also contain the sequence coding for the recognition site of a specific protease, located just 5´ to the polylinker insertion sites. This allows MBP to be cleaved from the proteins of interest after purification. The pMAL-p2E and pMAL-p2X vectors encode the recognition-sites for enterokinase (Asp-Asp-Asp-Asp-Lys) and factor Xa (Ile-Glu-Gly-Arg), respectively (Nagai & Thøgersen, 1987). The proteases cleave the C-terminal end of their amino acid recognition sequences, so that few or no vector-derived residues are attached to the proteins of interest depending on the site used for cloning.

The enterokinase- and the factor Xa-protease sites are located between the carboxyl terminus of MBP and the amino terminus of both full length and truncated savignin. The MBP and full length or truncated savignin are linked by a potentially flexible and solvent-exposed linker of 10 Asn residues.

Light-chain enterokinase (rEKL, 26.3 kDa), displays major differences in the efficiency of digestion of the purified fusion protein in comparison to native enterokinase (51 kDa). As little as 1:100 000 ratio (w/w) of light chain enterokinase to fusion protein is required for cleavage (Collins-Racie et al., 1995).

Recombinant proteins expressed in *E.coli* are often produced as aggregates called inclusion bodies (Wilkinson & Harrison 1991). When inclusion bodies are formed, some of the target proteins are semi-soluble within the cell. For the production of soluble and functional savignin, the pMAL-p2 system was used. As mentioned previously, the presence of the signal peptide directs the savignin into the periplasmic space resulting in disulphide bond formation (New England BioLabs) and making the target protein more soluble. Target proteins exported to the periplasm containing the signal sequence may be purified using simplified experimental procedures. Although the pMAL-p2 is a good expression system, some target proteins will not be good candidates for periplasmic localization.

Typically, a MBP fusion protein expressed from pMAL-p2 is first purified by amylose affinity chromatography (Figure 3.3) (Kellerman & Ferenci, 1982). The fusion protein is eluted from the column with a buffer containing maltose (Figure 3.3)

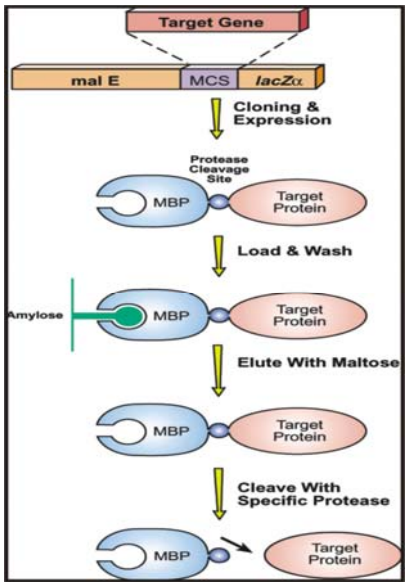


Figure 3.3: pMAL protein fusion and purification system. In the protein fusion and purification system, the cloned gene is inserted into a pMAL vector downstream from the malE gene, which encodes MBP. This results in the expression of an MBP-fusion protein. The fusion protein is then purified by one-step amylose affinity chromatography (Source: www.neb.com).

The fusion protein is then cleaved either with factor Xa or enterokinase and target protein purified from the MBP using additional chromatographic steps.

This chapter describes attempts to clone, express and purify full length savignin (Fsav) and its N- and C-truncated domains (Nsav and Csav).

3.2 Materials and Methods

Part 1: CONSTRUCTION OF RECOMBINANT PMAL FUSION PLASMIDS (See flow diagram, Part 1)

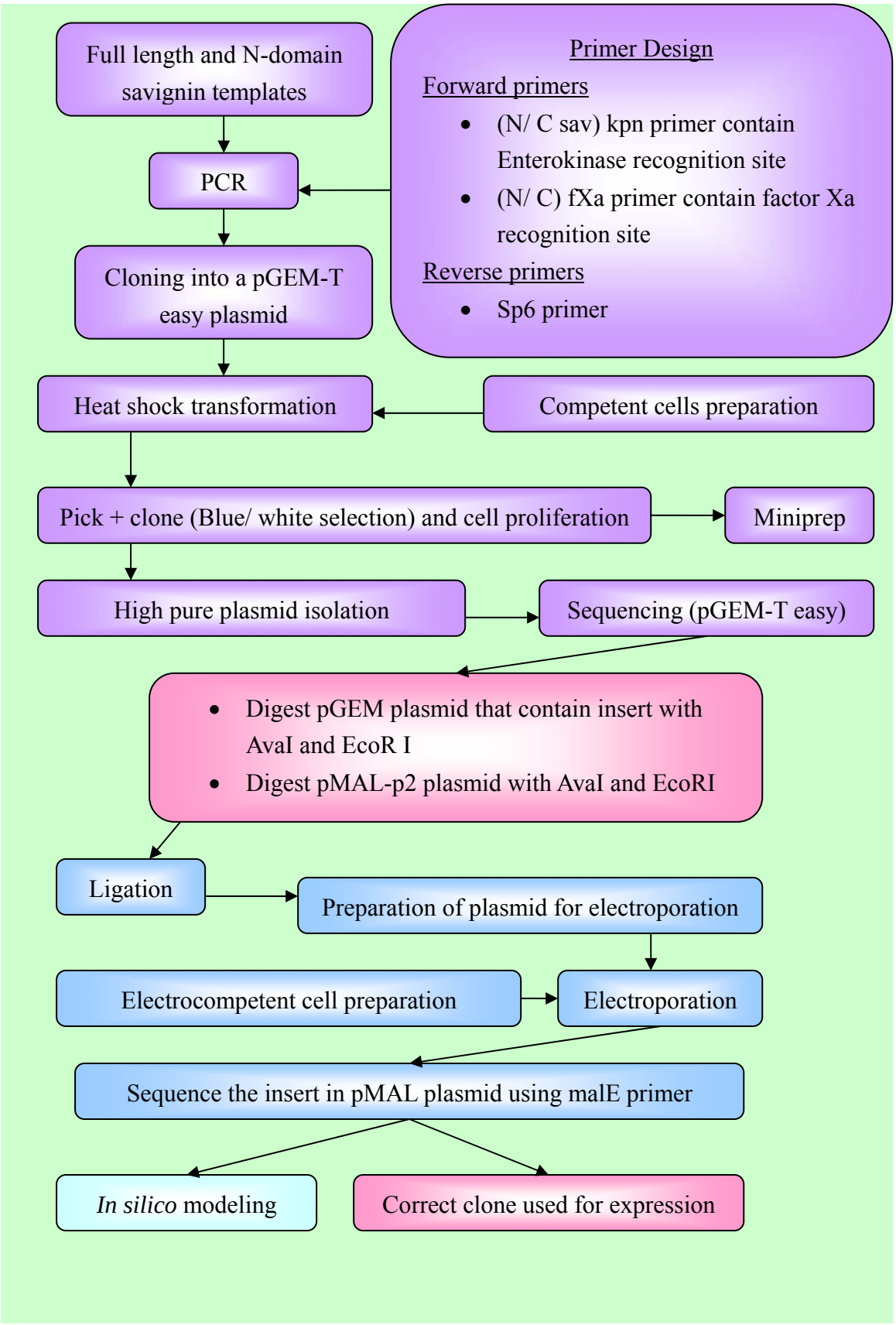
Primers used were synthesized by IDT (Integrated DNA Technologies, Coralville, USA) and re-suspended at 100 pmole/µl in 20 % acetonitrile and stored at -20 ^{o}C. Final concentrations were confirmed by spectrophotometric measurements at 260 nm. In the forward primers, protease recognition sites were engineered just before the recombinant genes, because precise cleavage of fusion protein is needed to gain the functional and active recombinant protein. Double-distilled, deionised and sterilised water was used in all experiments. The N-domain and full-length savignin cloned into pGEM-T easy vectors, respectively, were obtained from B. J. Mans (Department of Biochemistry, University of Pretoria).

3.2.1 PCR mediated synthesis of the N-, C-domains and full length savignin

The synthetic oligonucleotides (Nsavkpn, Csavkpn, NfXa, CfXa and Sp6 upstream primers) were used for amplification for N-, C-domains and full-length savignin (Figure 3.1 A and B).

PCR synthesis of the N-domain and full length savignin uses a forward (Nsavkpn/ NfXa, 5 pmole/µl) primer and a reverse primer (Sp6, 5 pmole/µl). For amplification of the C-domain the forward primer was either Csavkpn or CfXa and the reverse Sp6 primer. The PCR mixture (16 µl) containing [1 µl of forward primer (Nsavkpn/NfXa/Csavkpn/CfXa 5 pmole/µl), 1 µl of reverse primer (Sp6, 5 pmole/µl), 10 x PCR buffer (100 mM Tris-HCl, 500 mM KCl, pH 8.3), $MgCl_2$ (2.5 mM final concentration) and 4 µl dNTP's (200 µM final concentration)] was adjusted to a final

PART 1: CONSTRUCTION OF RECOMBINANT pMAL FUSION PLASMIDS



volume of 49.5 µl with H_2O. The DNA template (200 ng pGEM-T easy) of full length savignin (used for amplification of savignin and the C-domain) and the N-domain of savignin (used for amplification of the N-domain), were denatured at 94 ^oC for 3 minutes and then cooled to 80 ^oC after which 0.5 µl TaKaRa Taq^{TM} (5 U/µl, TAKARA Biotechnology) was added. Amplification consisted of 30 cycles of DNA denaturation (94 ^oC, 30 s); annealing (in the case of the NsavKpn / CsavKpn primers at 55 ^oC, the NfXa/CfXa primer s at 45 ^oC, 30 s); extension (72 ^oC, 2 minutes) and a final extension (72 ^oC, 7 minutes). All amplification procedures were conducted in a Gene Amp® PCR system 9 700 (Applied Biosystems).

3.2.2 Recloning into pGEM-T easy plasmid

After PCR amplification, duplicate samples of the respective genes were purified using the PCR purification kit (Roche Molecular Biochemicals). A volume of 3 µl insert (~ 6.7 ng), 5 µl 2 x rapid ligation buffer, 1 µl T4 ligase (3 U) and 1 µl EcoRI digested pGEM-T easy vector (~ 50 ng/µl) was added in a reaction tube and incubated at 4 ^oC on ice for 3 days. The concentration of insert was determined by using the following formula:

[(ng of pGEM vector x kb size of insert) / (kb size of pGEM vector (3.0 kb))] **X** 3 (3:1 insert: template) = ng of insert

3.2.3 Preparation of *E.coli* competent cells (*Sure*)

Sure cells were used in both cloning and expression. Competent *SURE* (Stratagene, La Jolla, CA, USA) *E.coli* cells were prepared using the calcium/manganese-based method (Hanahan, Jesse & Bloomer, 1991). Bacteria were plated onto Luria Bertani (LB) plates [1 % tryptone, 0.5 % yeast extract, 1 % NaCl, pH 7.5, 1.5 % (w/v) agar

with 12.5 µg/ml tetracycline] and allowed to proliferate overnight at 30 °C. Several colonies were suspended in SOB medium (2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, pH 7.0) by vortexing and inoculated into 50 ml SOB medium. Cells were grown at 30 °C with shaking until the OD_{600} was approximately 0.5, after which cells were pelleted by centrifugation (10 000 g, 15 minutes at 4 °C). The supernatant was discarded and the pellet was resuspended in 16.6 ml CCMB 80 medium (80 mM $CaCl_2.2H_2O$, 20 mM $MnCl_2.4H_2O$, 10 mM $MgCl_2.6H_2O$, 10 mM K-acetate, 10 % glycerol, pH 6.4) and incubated on ice for 20 minutes. Cells were pelleted by centrifugation (10 000 g, 10 minutes at 4 °C) and the supernatant discarded. The pellet was suspended in 4.1 ml CCMB 80 medium, aliquotted and stored at -70 °C.

3.2.4 Heat shock transformation

Competent cells were thawed on ice for 30 minutes before transformation. A volume of 100 µl competent cells was added to 10 µl ligation mixture [3 µl insert (~ 6.7 ng), 5 µl 2 x rapid ligation buffer, 1 µl T4 ligase (3 U) and 1 µl EcoR I digested pGEM-T easy vector (~ 50 ng/µl)]. These competent cells with plasmid were incubated on ice for 30 minutes. The mixtures were heat-shocked at 42 °C for 60 s and placed immediately on ice for 3 minutes. A volume of 900 µl SOC (2 % tryptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl and 10 µl 2 M D-glucose) medium was added to the sample and mixed by inversion. The mixture was then incubated at 30 °C for 1 hour in a shaker. For transformation, LB-ampicillin plates were prepared as follows: 200 ml LB medium (1 % tryptone, 0.5 % yeast extract, 1 % NaCl, pH 7.5) was added to 3.0 g agar and autoclaved. Then 1/500 of 100 mg/ml ampicillin was added and mixed thoroughly before it was poured into a petri dish. Each plate was coated with 16 µl of X-gal [5-bromo-4-chloro-3-indolyl D-galactopyranoside (20 mg/ml in DMF)]

and 4 µl of IPTG (200 mM), and allowed to dry. Transformed cells (100 µl) were plated out onto a solidified LB-plate and incubated overnight at 30 ^{o}C for a maximum of 16 hours.

3.2.5 Plasmid purification using miniprep

Cells from white colonies (1 positive colony in 1 ml LB Broth) were collected by centrifugation (13 500 g, 1 minute) in an eppendorf tube and re-suspended in 100 µl miniprep solution I (25 mM Tris-HCl, 50 mM glucose, 10 mM EDTA, pH 8.0). Miniprep solution II (150 µl, 0.2 M NaOH, 1 % SDS) and the cell suspension were mixed by gentle inversion before incubation on ice (5 minutes) to lyse the cells. Pre-chilled miniprep solution III (250 µl, 3 M KOAc, pH 4.2) was added and mixed by gentle inversion and incubation on ice (5 minutes), before centrifugation (12 000 g, 5 minutes), to precipitate chromosomal DNA. Absolute ethanol (1 ml) was added to the supernatant and incubated at room temperature (10 minutes), before centrifugation (12 000 g, 5 minutes) to precipitate plasmids. The supernatant was discarded and the pellet was dried using the Bachoffer vacuum concentrator.

Pellets were suspended in 30 µl water (contain 20 µg RNAse) and 8 µl of this mixture was digested with 1 µl Ava I (12 U) in digestion buffer (60 mM Tris-HCl pH 7.5, 500 mM NaCl, 60 mM $MgCl_2$ and 10 mM DTT at 37 ^{o}C) for the first 2 hours. Thereafter, 1 µl Eco RI (10 U) in digestion buffer (90 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, pH 7.5 at 37 ^{o}C) was added and incubated for a further 2 hours.

3.2.6 Agarose gel electrophoresis

Agarose (0.6 g) was dissolved in 30 ml of TAE buffer [40 mM Tris-acetate, 1 M methylene diamine tetra-acetic acid (EDTA)] and cooled at 55 - 60 ^{o}C, followed by

addition of 5 µl ethidium bromide (0.5 µg/ml). The agarose solution was poured onto the gel tray to a thickness of 5 mm and the comb inserted immediately after pouring. The gel was left to set for an hour.

Digested plasmids obtained from miniprep were mixed with loading buffer (50 % glycerol, 0.4 % bromophenol blue, 1 mM EDTA) and loaded in the wells of the gel. The gel was run at 80 V for an hour. After separation of the bands, the desired fragment was excised from the gel for further silica purification.

3.2.7 Plasmid purification using high pure plasmid isolation kit

The purification method was followed as described in the 'High Pure Plasmid Isolation Kit' instruction manual 1999 (Roche Molecular Biochemicals). Cells from 3 ml culture were pelleted at 12 000 g for 30 s at room temperature. The supernatant was discarded and the pellet re-suspended in 250 µl suspension buffer (50 mM Tris-HCl, 10 mM EDTA, pH 8.0 at room temperature containing RNase). This was followed by addition of 250 µl lysis buffer (0.2 mM NaOH and 1 % SDS), mixed gently and incubated for 5 minutes at room temperature. Pre-chilled binding buffer (350 µl) (4 M guanidine hydrochloride, 0.5 M potassium acetate, pH 4.2) was then added. The sample was incubated for further 5 minutes on ice. After incubation, the samples were centrifuged at 12 000 g for 10 minutes. The supernatant was transferred to high pure filter tube and centrifuged at 12 000 g for 1 minute. After centrifugation, the flow-through solution was discarded and 700 µl wash Buffer II [20 mM NaCl, 2 mM Tris-HCl, 20 % ethanol (v/v), pH 7.5 at room temperature] was added and centrifuged at 12 000 g for 1 minute.

Following the centrifugation, the flow-through solution was discarded and the

remaining pellet was re-centrifuged at 12 000 g for 1 minute to remove residual wash buffer. The filter tube was inserted into a clean 1.5 ml microcentrifuge tube and 100 µl of double distilled deionized water was added, followed by centrifugation at 12 000 g for 1 minute to elute the plasmid.

3.2.8 Elution of inserts using silica purification

The procedure as described by Boyle & Lew (1995) was followed. The agarose-containing selected inserts were dissolved in three times the fragment volume of fresh 6 M sodium iodide (NaI) and melted at 60 ^{o}C for 5 minutes. Thereafter, 10 µl silica solution (Sigma, USA) in 6 M NaI was added. The eppendorf tubes were inverted gently and placed on ice for 30 minutes. These were then centrifuged at 12 000 g for 1 minute and the supernatant removed. The pellet was re-suspended with 1 ml wash buffer (10 mM Tris-HCl, 50 mM NaCl, 2.5 mM EDTA, pH 7.6 and 50 % ethanol) and centrifuged as before. The wash buffer was removed and the silica was then suspended in 10 µl elution buffer (1 mM Tris-HCl, pH 8.0 at 60 ^{o}C) and centrifuged for another minute. The supernatant was transferred into a second eppendorf tube and the elution procedure was repeated for each sample (Boyle & Lew, 1995). Yield was determined spectrophotometrically.

3.2.9 pGEM sequencing of inserts

Silica purified plasmids were used for sequencing. Plasmids (400 ng) were sequenced using 3.2 pmoles of the upstream T7 primer with the Big Dye Kit (Perkin Elmer, Foster City) on an ABI 377 sequencer (Applied Biosystems) according to the manufacturer's instructions. Different clones were sequenced for each insert and each clone was sequenced with upstream T7 primer. Sequences obtained were analyzed using the Bioedit programme (Hall, 2001). The DNA and deduced protein sequences

were analyzed using BLAST (Altschul et al., 1990) and alignments were performed with Clustal X (Jeanmougin et al., 1998). Protein mass and amino acid composition of the deduced sequences were determined with the PAWS programme.

3.2.10 Cloning of full length savignin and the N- and C-domains into pMAL-p2E fusion plasmids

For cloning, pMAL-p2E and pGEM-T easy plasmids (1 µg) (containing full length and truncated forms of savignin) were digested with Ava I and EcoR I restriction enzymes to generate cohesive 5'- and 3'-sticky ends (refer to figure 3.2B). Digested products were subjected to agarose gel electrophoresis (section 3.2.6).

The desired inserts were excised from agarose gels and purified using High Pure PCR Purification Kit (Roche Molecular Biochemicals). Volumes of 300 µl Binding buffer [3 M guanidine-thiocyanate, 10 mM Tris-HCl, 5 % ethanol (v/v), pH 6.6] was added to dissolve the agarose gel (300 µl Binding buffer:100 mg agarose gel) for 10 minutes at 56 ^{o}C, followed by the addition of isopropanol (150 µl isopropanol:100 mg agarose gel) before addition to High Pure filter tubes. Tubes were centrifuged for 1 minute at 12 000 g and the flow-through discarded. Filters were washed with 500 µl wash buffer [2 mM Tris-HCl, 20 mM NaCl, 20 % ethanol (v/v), pH 7.5] by centrifugation (12 000 g, 1 minute), followed by another wash with 200 µl wash buffer. The desired fragment was eluted with 100 µl elution buffer (1 mM Tris-HCl, pH 8.5). The yield was determined using *GENEQUANT*. The inserts (15 ng ~ 25 ng) were ligated into the pMAL-p2E and pMAL-p2X vectors (50 ng) at 4 ^{o}C (for 16 hours) with 1 µl T4 ligase (1 U, Promega) and 1 µl 10 x ligase buffer [300 mM Tris-HCl (pH 7.8), 100 mM $MgCl_2$, 100 mM DTT and 10 mM ATP, Promega].

3.2.11 Electroporation

3.2.11a Preparation of *E. coli* competent cells (TB1)

Fresh colonies of *E. coli* cells were used to inoculate 50 ml of SOB medium in a 500 ml flask. Cells were proliferated with vigorous aeration overnight at 37 ^{o}C. Cells (2.5 ml) were diluted into 250 ml of SOB in a 1 L flask. They were grown for 2 to 3 hours with vigorous aeration at 37 ^{o}C until the cells reached an OD_{600} ~ 0.5. Cells were harvested by centrifugation at 2 600 g for 10 minutes. The cell pellet was washed by re-suspension in 250 ml of sterile ice-cold wash buffer [10 % redistilled glycerol, 90 % distilled water, (v/v)]. The cell suspension was centrifuged at 2 600 g for 15 minutes. Cells were washed again in 250 ml of sterile ice-cold double distilled water and the same wash process was repeated. The cell suspension was centrifuged at 2 600 g for 15 minutes and the supernatant was poured off. The cell pellet was resuspended in wash buffer to a final volume of 1 ml. Cells were frozen in 0.1 ml aliquots in microcentrifuge tubes using liquid nitrogen. Frozen cells were stored at -70 ^{o}C.

3.2.11b Preparation of plasmid for electroporation

The pMAL-p2 plasmids containing the desired insert (full length or truncated forms of savignin) as prepared in 3.2.10 was purified by precipitation. These mixtures were precipitated by adding 1/10 of both tRNA and 5 M ammonium acetate to the ligation reaction. The precipitates were incubated with 3 volumes of absolute ethanol -70 ^{o}C for 1 hour. The supernatant was poured off after centrifugation at 13 000 g for 15 minutes at 4^{o}C. The pellet was further washed with 60 µl of 70 % ethanol and centrifuged at 13 000 g for 15 minutes at room temperature. The plasmids were dried at room temperature, followed by reconstitution with water to a concentration of 10 ng/µl of DNA.

3.2.11c Transformation by electroporation

An aliquot of cells, that had been prepared as described in 3.2.11a, was thawed and 100 µl transferred to microfuge tubes on ice. This was followed by the addition of 5 µl plasmid containing the savignin gene (10 ng/µl per 20 µl of cell suspension) prepared as in 3.2.11b. This mixture (20 µl) was pipetted into the channels of an electroporation cuvette (placed in ice before use). The cuvette was placed in a slot in the chamber rack and pulsed at 2 000 mV for 5ms.

For the selection of cell transformation, cells were inoculated into 2 ml of SOC medium and shaken for 60 minutes to allow cell proliferation. Cells were then plated onto LB ampicillin plates and allowed to grow for a further 16 hours at 37 ^{o}C. A single cell colony was picked out from the plates and inoculated in 5 ml LB broth containing 5 µl glucose.

To screen if the savignin gene was present in the plasmid, the overnight culture was subjected to a miniprep (refer to section 3.2.5), agarose gel electrophoresis (refer to section 3.2.6) and plasmid purification (refer to section 3.2.7).

3.2.12 Sequencing of inserts cloned into pMAL plasmids

Plasmids to be sequenced were purified using the high pure plasmid isolation kit (section 3.2.7, Roche Molecular Biochemicals) and sequenced using the upstream malE primer, 5'-GGTCGTCAGACTGTCGATGAAGCC-3' (New England Biolabs), with the Big Dye Kit (Perkin Elmer, Foster City), using an ABI 377 sequencer (Applied Biosystems).

3.2.13 Amino acid composition of full length and truncated forms of savignin

The Bioedit programme (Hall, 2001) was used to calculate the theoretical molecular mass and the amino acid composition of the recombinant forms.

3.2.14 *in Silico* modelling of savignin and its truncated forms

To model the structures of truncated and full length recombinant savignin, the crystal structure of ornithodorin was used as the template. The *de novo* amino acid sequences of cloned sequences were submitted to the SWISS-MODEL Automated Comparative Protein Modelling Server (Guex et al., 1999). All worm figures with secondary structures were constructed with the graphical representation and analysis of surface properties program (Pymol).

PART 2: EXPRESSION AND PURIFICATION OF RECOMBINANTS (See flow diagram, Part 2)

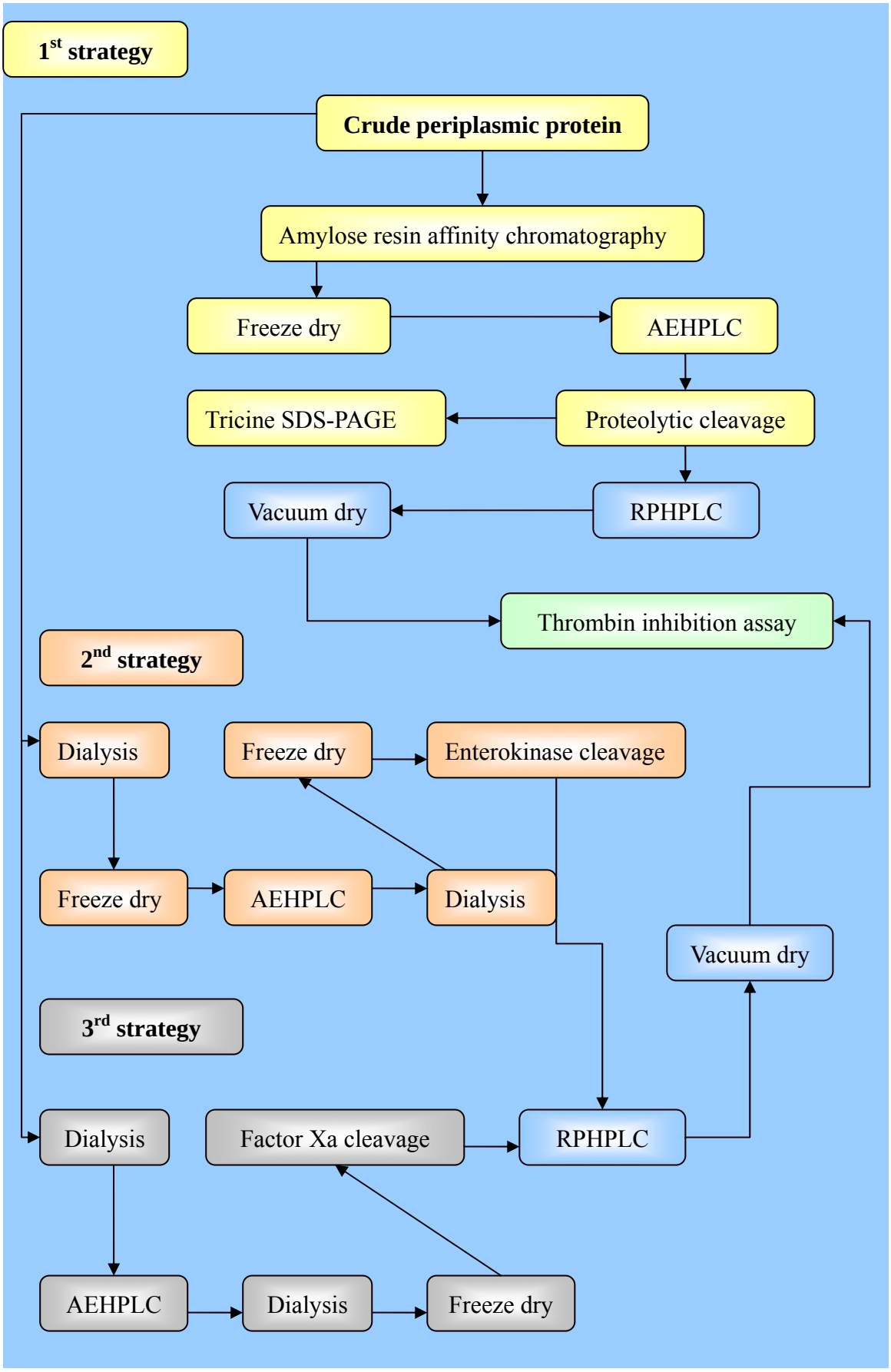
With the exception of the modifications outlined below, all reagents used and methods employed were carried out as described by the respective suppliers. Human thrombin was purchased from Sigma and the chromogenic substrate, Chromozym-TH, was purchased from Roche Molecular Biochemicals.

3.2.15 Optimization of protein induction

An overnight culture (0.8 ml) of both TB1 and *Sure* cells containing the fusion plasmid (pMAL-p2E) was added to 80ml of rich broth (10 g tryptone, 5 g yeast extract, 5 g NaCl, 2 g glucose, 100 µg ampicillin). Cells were grown at 37 °C with good aeration to 2 x 10^8 cells/ml (A_{600} of ~ 0.5). A sample of 1 ml was taken and centrifuged for 2 minutes (sample 1:uninduced cells). Supernatant was discarded and the cells re-suspended in 50 µl of SDS-PAGE sample buffer and stored at –20 °C. IPTG was added to the remaining culture to a final concentration of 0.3 mM.

Incubation was continued at 37 °C for 2 hours. A 0.5 ml sample was withdrawn and centrifuged for 2 minutes (sample 2:induced cells) for SDS-PAGE electrophoresis (see section 2.2.5 for detail). The supernatant was discarded and the cells re-suspended in 100 µl of SDS-PAGE sample buffer and then frozen at –20 °C. The remaining culture was divided into two aliquots for protein isolation. This was repeated for time intervals of 3, 6, 12 & 24 hours. The cell pellets obtained from different induction time periods were re-suspended in 10 ml suspension buffer [(30 mM Tris-HCl, 20% sucrose, pH 8.0), (8 ml / 0.1 g cells wet weight)].

PART 2: EXPRESSION AND PURIFICATION OF RECOMBINANTS



To prepare the periplasmic protein extracts, EDTA (1 mM final concentration) was added to the cell suspension and incubated for 5 – 10 minutes at room temperature with shaking before centrifugation at 8 000 g at 4 °C for 10 minutes. The supernatants were discarded and pellets re-suspended in 10 ml chilled magnesium sulphate (5 mM $MgSO_4$). The suspended cells were then shaken for 10 minutes in an ice-water bath followed by centrifugation as above. Supernatant (10 µl of cold osmotic shock fluid) was removed, and reconstituted with 10 µl of 2 x SDS-PAGE sample buffer for further SDS-PAGE analysis.

3.2.16 Quantitative protein assay

The total amount of protein from the crude periplasmic extract was determined by the Bradford method (1976) using the Pierce Coomassie Protein Assay Kit (Pierce, USA). Bovine serum albumin (BSA) was used as a standard reference protein. Standard protein or samples (150 µl) were pipetted into the wells of a microtitre plate followed by 150 µl of the Coomassie Protein Assay Reagent. The microplate was shaken for 30 minutes before reading the absorbance at 595 nm with an ELISA microplate reader (Thermo Labsystem, Multiskan Ascent). Determinations were performed in triplicate.

3.2.17 Preparative amylose resin affinity chromatography

The amylose resin, pre-swollen in 20 % ethanol (New England Biolab) was poured into a 25 ml syringe (2.5 cm diameter) plugged with silanized glass wool. The column was washed with 12 column volumnes of 20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA and adjusted to have a flow rate of 1 ml/minute. The crude periplasmic protein extract was loaded. The flow-through was reloaded to ensure maximum binding and this was followed by washing with 12 column volumes of column buffer. After each washing step, fractions were collected to monitor for unbound protein by SDS-PAGE.

To elute the fusion protein, column buffer with 10 mM maltose was used. Between 10 - 20 fractions of 3 ml each were collected. Elution was monitored by measuring UV absorbance at 280 nm, determining the protein concentration by the Bradford method and by SDS-PAGE analysis. All fractions, including fractions collected during the washing steps, were dialyzed overnight at room temperature. The dialyzed fusion proteins were immersed in liquid nitrogen with swirling motion until frozen and then freeze dried.

3.2.18 Anion exchange HPLC (AEHPLC)

The freeze dried fusion proteins were reconstituted with 1 ml of Tris-HCl (20 mM, pH 7.6) and centrifuged for 10 minutes. The supernatant was filtered through a 0.22 µm filter (Millipore Corporation, Bedford, Massachusetts) before applying to the anion exchange column as described in the Materials and Methods 2.2.4a.

3.2.19 Cleavage of fusion proteins

3.2.19 a) Enterokinase Cleavage

Enterokinase digestions with *E.coli*-derived recombinant light chain enterokinase (rEK_L) (New England Biolabs) were performed by mixing 100 µl (100 µg) aliquots of purified MBP-fusion protein with 1 ng (1:100 000), 2.5 ng (1:40 000) or 5 ng (1:20 000) of EK_L in buffer (20 mM Tris-HCl pH 7.6), each in a final volume of 120 µl (Collins-Racie et al., 1995). Digestions were carried out at 37 ^{o}C for 12 and 24 hours, and 10 µl of each reaction was analyzed with tricine SDS-PAGE (refer to section 2.2.5.2).

3.2.19 b) Factor Xa Cleavage

An alternative enzyme used to cleave the fusion protein was factor Xa. To do so, pMAL-p2E plasmid was modified to pMAL-p2X using appropriate primers. The contents of one vial of protease factor Xa (Roche Molecular Biochemicals) was dissolved in buffer containing 100 mM NaCl, 50 mM Tris-HCl, 1 mM $CaCl_2$, pH 8.0 to a final concentration of 1 mg/ml. The purified MBP-fusion proteins to be cleaved were dissolved in buffer containing 100 mM NaCl, 50 mM Tris-HCl, 1 mM $CaCl_2$, pH 8.0.

The enzyme:substrate ratio used for cleavage was from 1:10 to 1:100 (w/w). Incubation was carried out at 25 ^{o}C for a total of 16 hours with a sample collection every 4 hours (Nagai & Thøgersen, 1987). The cleavage of the fusion proteins were analyzed with tricine SDS-PAGE (refer to section 2.2.5.2).

3.2.20 Reversed-phase HPLC (RPHPLC)

The cleaved proteins were centrifuged (10 minutes, 12 000 g) and the supernatant was filtered through a 0.22 μm filter (Millipore Corporation, Bedford, Massachusetts) before application to the reversed phase column (refer to section 2.2.4c). Fractions were collected and dried in a Bachoffer vacuum concentrator.

3.3 Results

Part 1: CONSTRUCTION OF RECOMBINANT pMAL FUSION PLASMIDS

3.3.1 PCR amplification of savignin

PCR amplified cDNA's encoding savignin and its truncated forms were analyzed using agarose gel electrophoresis (2 %). Amplified N- and C-domains as well as full length savignin fragments corresponded to bands of sizes 300 bp, 300 bp and 500 bp bands, respectively (Figure 3.4).

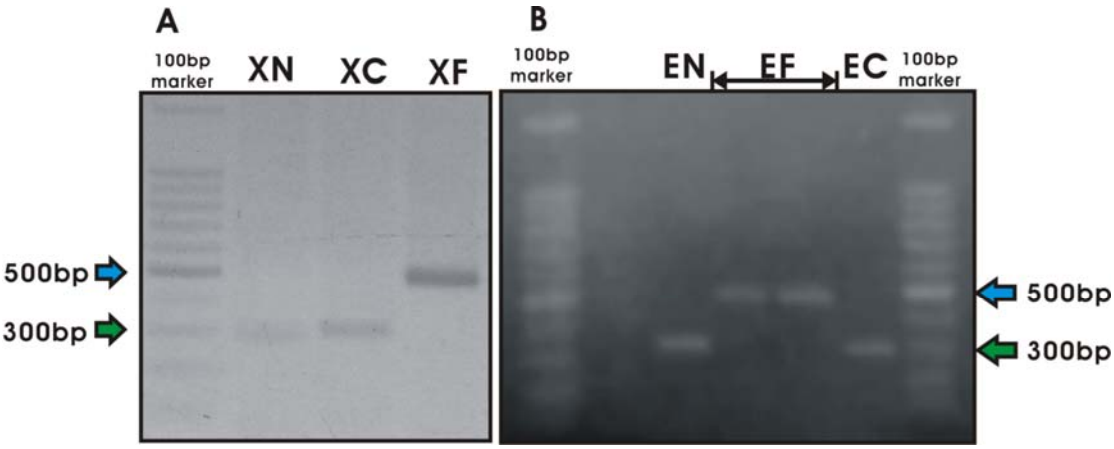


Figure 3.4 Agarose gel electrophoresis analysis of PCR products. A). PCR synthesis of savignin and its truncated forms from pGEM-T easy templates containing either the N-domain or full length savignin genes respectively using 5' overhang primers containing the factor Xa recognition sequence and the 3' Sp6 primer. 100 bp DNA ladder. XN = N-domain of savignin containing factor Xa cleavage site. XC = C-domain of savignin containing factor Xa cleavage site. XF = full-length savignin containing factor Xa cleavage site. B). PCR synthesis using 5' overhang primers containing the enterokinase recognition sequence and the 3' Sp6 primer. EN, EC, EF = N-, C-domains and full length savignin containing enterokinase cleavage site.

3.3.2 Cloning of cDNA's into pMAL plasmids and transformation into TB1 cells

After PCR products containing either fXa or enterokinase recognition sites were purified, they were re-ligated to the pGEM-T easy plasmid. To confirm the presence of the protease recognition sites, the pGEM-T easy plasmids were sequenced. Once the sequences were confirmed, restriction enzymes AvaI and EcoRI were used to

isolate the inserts (full length, N- and C-domains of savignin). These inserts (pre-digested with AvaI and EcoRI) were then ligated into pMAL-p2E expression plasmids, and transformed into TB1 cells. The cells were then plated and a single colony was proliferated. After proliferation the cloned plasmids were harvested from the cells which contained the inserts. These were digested with EcoRI and AvaI and elecrophoresed (Figure 3.5). Two bands corresponding to ~ 300 bp and ~ 400 bp were observed which are in the expected range of truncated and full-length inserts, respectively.

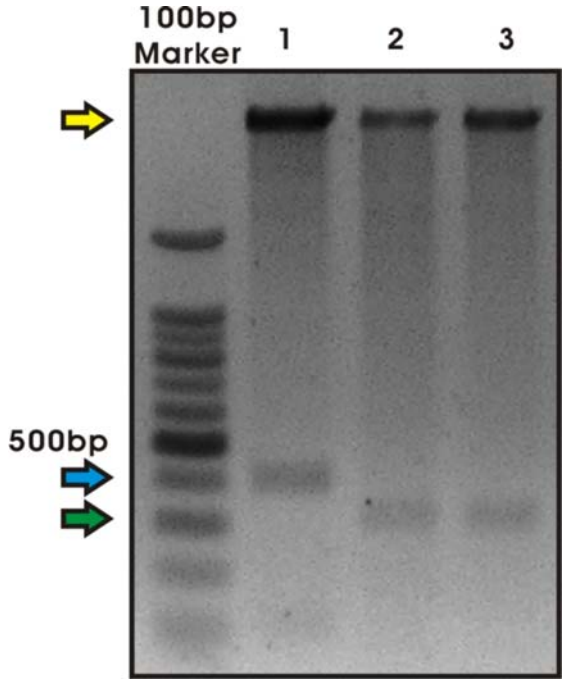


Figure 3.5 Agarose gel electrophoretic analysis of pMAL-p2 plasmid (containing inserts) after AvaI and EcoRI restriction enzyme digestions. Lane 1: full length savignin gene, Lane 2: N-domain of savignin. Lane 3: C-domain of savignin.

3.3.3 Nucleotide and amino acid sequence of savignin and its truncated genes.

The nucleotide sequences of pMAL plasmid inserts were determined and found to contain 153 bp, 180 bp and 357 bp. These sequences correspond to the N-domain savignin (*Nsav*), C-domain savignin (*Csav*) and full length savignin (*Fsav*), respectively (Figure. 3.6). A homology search of the PDB database revealed that the

deduced amino acid sequences of the cloned *Fsav* gene showed 100 % and 83 % identity with native savignin and with ornithodorin (van de Locht et al., 1996), respectively. The deduced amino acid sequences of the cloned genes for the N- and C-domains showed 100 % identity with savignin.

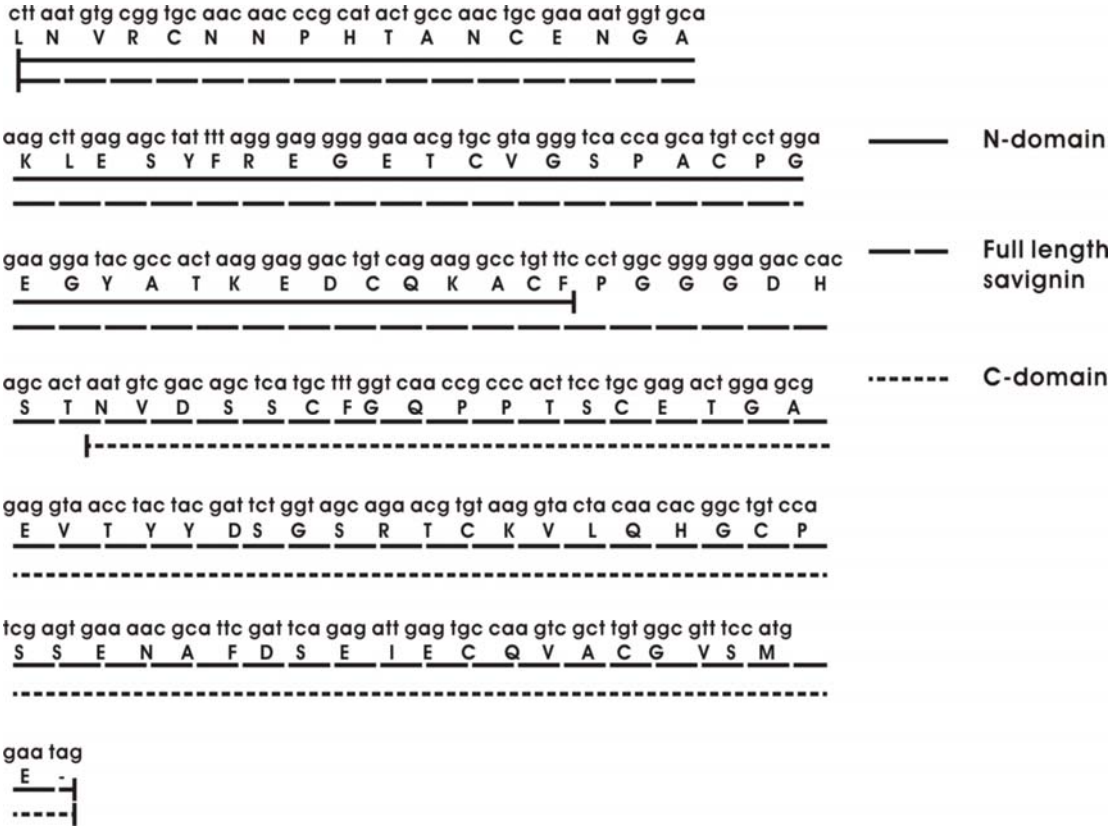
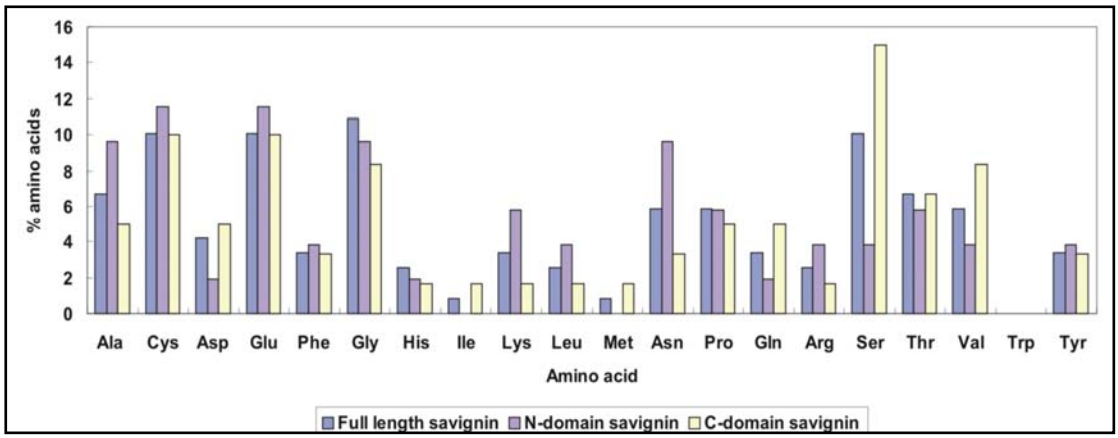


Figure 3.6 cDNA sequence and deduced amino acid sequence of savignin and its truncated forms.

3.3.4 Amino acid composition and molecular mass of savignin and its truncated forms

The deduced amino acid sequence revealed a protein of 118 amino acids residues for Fsav. The predicted amino acid composition of Fsav and its truncated forms Nsav and Csav correlated well with native savignin. The molecular mass of native savignin is 12 430.4 Da (determined by electrospray mass spectrometry, Nienaber et al. 1999). This result correlates well with the native molecular mass obtained for the full length savignin (12 446.96 Da). The predicted molecular masses for the truncated N- and C-



domains of savignins were 5 513.85 Da and 6 242.48 Da, respectively.

Figure 3.7 Amino acid compositions of savignin and its truncated forms. Comparison of the amino acid composition of savignin and its truncated forms deduced from the corresponding cDNA sequences (Bioedit Programme).

3.3.5 *in silico* modelling of savignin and the truncated forms

A full length structural model of savignin based on the crystal structure of ornithodorin (A23191) was published by Mans (2002). In this study, the N- and C-domains of savignin were also modelled and were based on the previous modelled savignin.

Superposition of the α-carbon backbone structures of savignin and its truncated forms onto ornithodorin (A23191) gave an RSMD value of 0.252 Å for full length savignin (Figure 3.8 A). The N- and C-terminal domains gave values of 0.108 Å (Figure 3.8B) and 0.105 Å (Figure 3.8C), respectively. The linker region showed the largest deviation of 0.177 Å (Figure 3.8). The modelled ribbon structure of full-length savignin shared two distinctive domains each containing one hairpin and one helix region joined by a linker sequence. The modelled ribbon structure of the N- and C-domains each contained one hairpin and one helix.

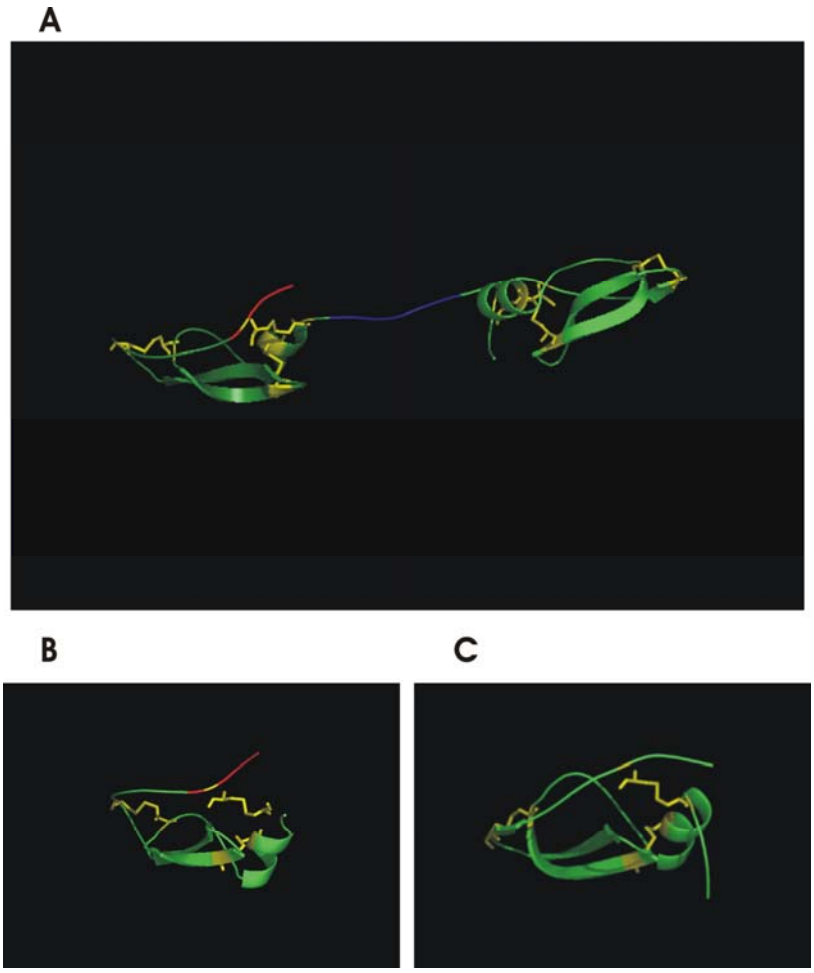


Figure 3.8 Modelled structures of savignin and its truncated forms. A): Model of the full length savignin. Red: shows the N-terminal residues that are accommodated in the active site cleft of thrombin. Yellow: indicates the presence of disulphide bonds. Blue: indicates the hinge region that joins the two globular domains of savignin. Green: shows the globular domains of savignin. B): Modelled structure of the N-terminal domain of savignin. C): Modelled structure of the C-terminal domain of savignin.

PART 2: EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS

3.3.6 Expression of MBP – fusion proteins

The gene for full length savignin was inserted downstream from the *malE* gene of pMAL-p2 plasmid, resulting in expression of the target gene product as a protein fused to MBP. MBP is a water-soluble protein with a molecular weight of 42.7 kDa. The expected molecular weight of MBP-savignin (MBP-Fsav), MBP-Csav, and MBP-Nsav are 54.7 kDa, 48.7 kDa and 48 kDa, respectively. Fusion proteins were stable in both TB1 and *Sure* strains during expression at 37 ^{o}C. Upon induction with IPTG the cells stopped proliferating, while production of the fusion proteins continued between 3 to 6 hours. A number of variables (such as different cell lines, induction times, concentration of IPTG) were investigated in order to maximize the cell and membrane protein yields. Each variable was monitored using SDS-PAGE analysis to determine the overall periplasmic protein production.

3.3.7a) Comparative studies between TB1 cells and *Sure* cells

Fusion protein (MBP-Fsav) yields varied between 0.2 and 0.5 mg/l as determined from the Bradford standard curve (Figure 3.9) for TB1 and *Sure* cells, respectively.

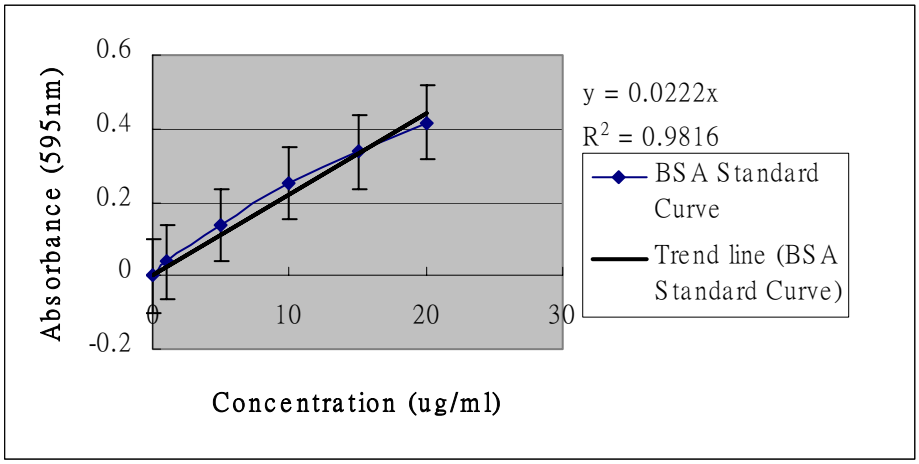


Figure 3.9: BSA standard curve generated using the Bradford method.

An important consideration that had to be taken into account when choosing the host strain was the lethality of expression. The TB1 strain is recommended by the manufacturer for cloning and expression with pMAL-p2 systems. TB1 strains had lower levels of protein expression than *Sure* strains. For *Sure* cells the yields were 2.4 mg/l and 3.2 mg/l fusion protein after 3 and 6 hours, respectively. TB1 strains require 6 hours incubation to yield a visable periplasmic band compared to 3 hours for *Sure* cells (Figure 3.10).

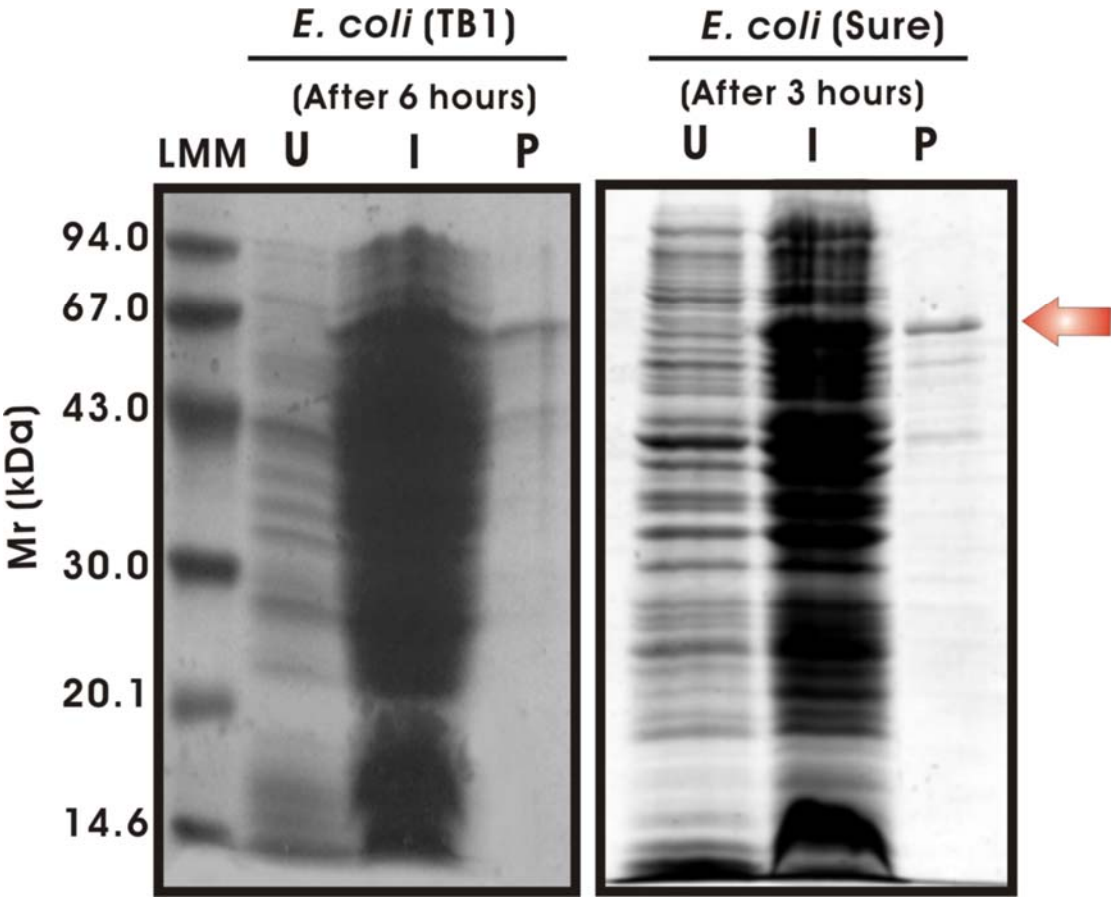


Figure 3.10 SDS-PAGE analyses of MBP-Fsav (indicated by red arrows) using TB1 and Sure strains. LMM = low molecular mass marker. U = uninduced protein. I = total induced protein. P = periplasmic protein. A) Left panel: TB1 cells. B) Right panel: Sure cells.

3.3.7b) Optimization of the induction time for *Sure* cells

The next variable to be investigated was the induction time in order to maximize the cell and fusion protein yields.

To monitor the release of protein in *SURE* cells, both the Bradford protein assay and SDS-PAGE analysis were used. The samples of periplasmic protein were harvested at various time points in order to determine the optimum production of fusion proteins (Figure 3.11). Lane 8 (5.5 hours) and lane 9 (6 hours) indicate the optimum production of the protein.

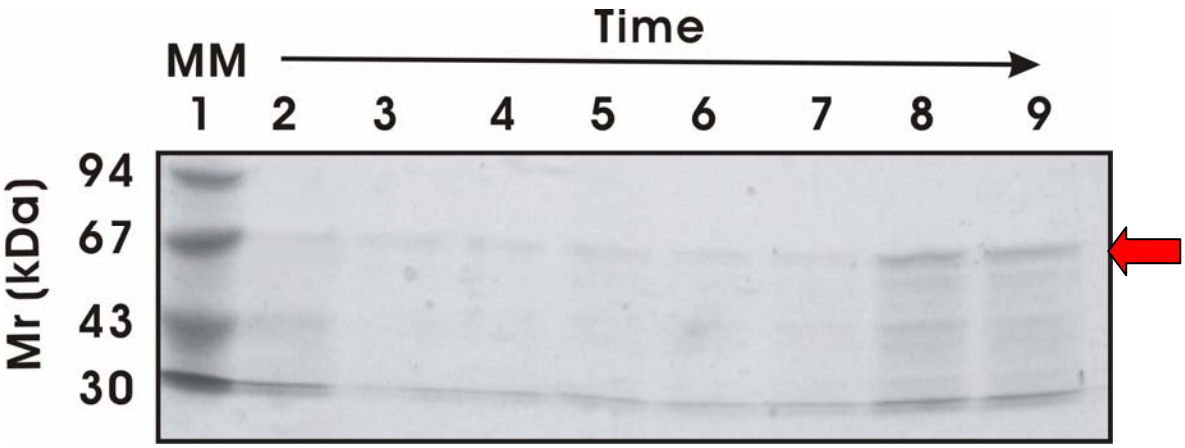


Figure 3.11 SDS-PAGE analysis of periplasmic protein at various time points of induction of Sure cells. Lane 1: low molecular mass marker. Lane 2: 2.5 hours. Lane 3: 3 hours. Lane 4: 3.5 hours. Lane 5: 4 hours. Lane 6: 4.5 hours. Lane 7: 5 hours. Lane 8: 5.5 hours. Lane 9: 6 hours. Arrows indicate fusion protein.

3.3.7c) Quantity of IPTG used for induction

Other variables investigated included the concentration of the inducer (IPTG) as well as culture batch size used for growth. It was also found that increased IPTG concentrations reduced cell population significantly (data not shown). When the concentration of IPTG increased from 0.3 mM (Lane 2: control) to 1 mM (Lane 4) protein expression remained the same (Figure 3.12) determined by Bradford method (data not shown).

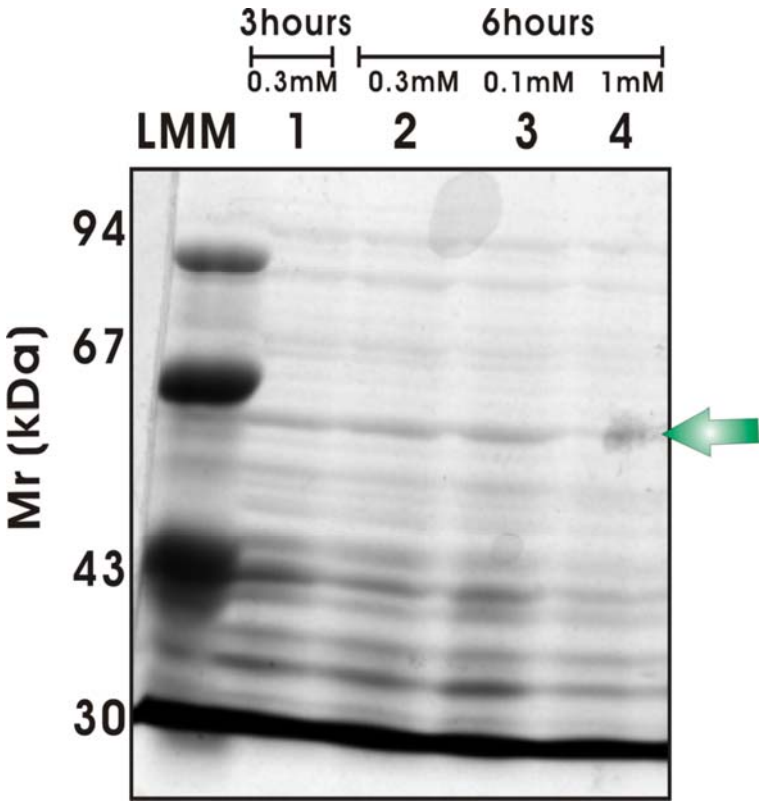


Figure 3.12 SDS-PAGE analysis of total protein using various concentrations of IPTG for induction. LMM: low molecular mass marker. Lane 1: 0.3 mM IPTG induced for 3 hours. Lane 2: 0.3 mM IPTG induced for 6 hours. Lane 3: 0.1 mM IPTG induced for 6 hours. Lane 3: 1 mM IPTG induced for 6 hours. Arrow indicates fusion protein.

When concentrations of IPTG decreased from 0.3 mM (Lane 2) to 0.1 mM (Lane 3) no changes in total protein concentration were observed, whereas MBP-Fsav expression was reduced close to 10-fold (determined by Bradford assay). Proliferating cells in smaller batch sizes (i.e., 80 and 250 ml) have insignificant effect on the protein yields per ml (less than 20 %). One litre batch growth was used in order to maximize overall protein production.

3.3.7 d) Optimization of temperature

The next variable to be investigated was the temperature in order to maximize the cell protein and fusion protein yields.

To monitor the expression levels of protein, the periplasmic protein produced at 25 ^{o}C and 37 ^{o}C, respectively, were analysed using SDS-PAGE (Figure 3.13). The expression levels of MBP-Fsav (indicated in blue arrow) displayed no differences between 25 ^{o}C and 37 ^{o}C.

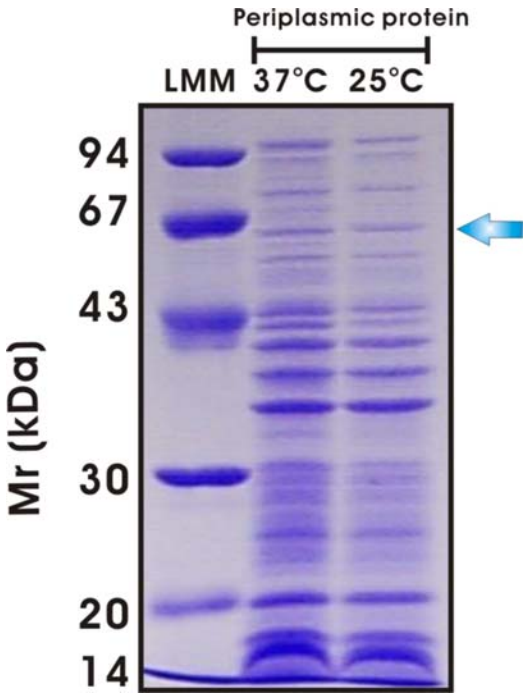


Figure 3.13 SDS-PAGE analysis of periplasmic protein at 25 ^{o}C and 37 ^{o}C of induction for Sure cells.
LMM: low molecular mass marker. Lane 2: 37 ^{o}C Lane 3: 25 ^{o}C. Arrow indicates MBP-Fsav.

3.3.8 Production of recombinant proteins

To monitor the production of fusion protein, MBP (42.7 kDa) was employed as positive control. Both MBP-Fsav (54.7 kDa) and MBP-Csav (48.7 kDa) protein were expressed, whereas MBP–Nsav (48 kDa) was not expressed (Figure 3.14).

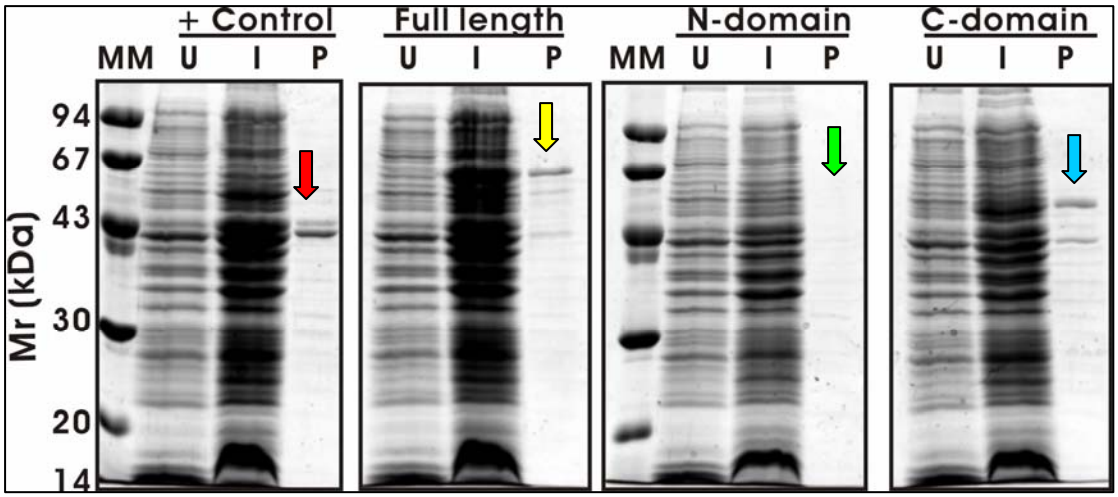


Figure 3.14 SDS-PAGE analysis of various recombinant forms of savignin. MM: low molecular mass marker. U: uninduced cells. I: induced cells. P: periplasmic space protein. First SDS gel: Production of MBP. Second SDS gel: Production of MBP-Fsav. Third SDS gel: Production of MBP-Nsav. Fourth SDS gel: Production of MBP-Csav. Red arrow indicates MBP. Yellow arrow indicates periplasmic MBP-Fsav. Green arrow indicates absence of MBP-Nsav. Blue arrow indicates periplasmic MBP-Csav.

3.3.9 Proteolytic cleavage efficiency

Figure 3.15 shows the reaction conditions of the proteolytic processing. The amount of enterokinase required to cleave a fusion protein in 12 and 24 hours at room temperature, varies from 1:20 000 to 1:100 000 (w/w) of enzyme to fusion protein as indicated in Figure 3.15. Optimum cleavage of the MBP-Fsav under these conditions requires 1:20 000 (w/w) of enzyme to substrate for 12 hours. The expected Fsav band was not observed on the tricine SDS-PAGE gel with either coomassie or silver staining.

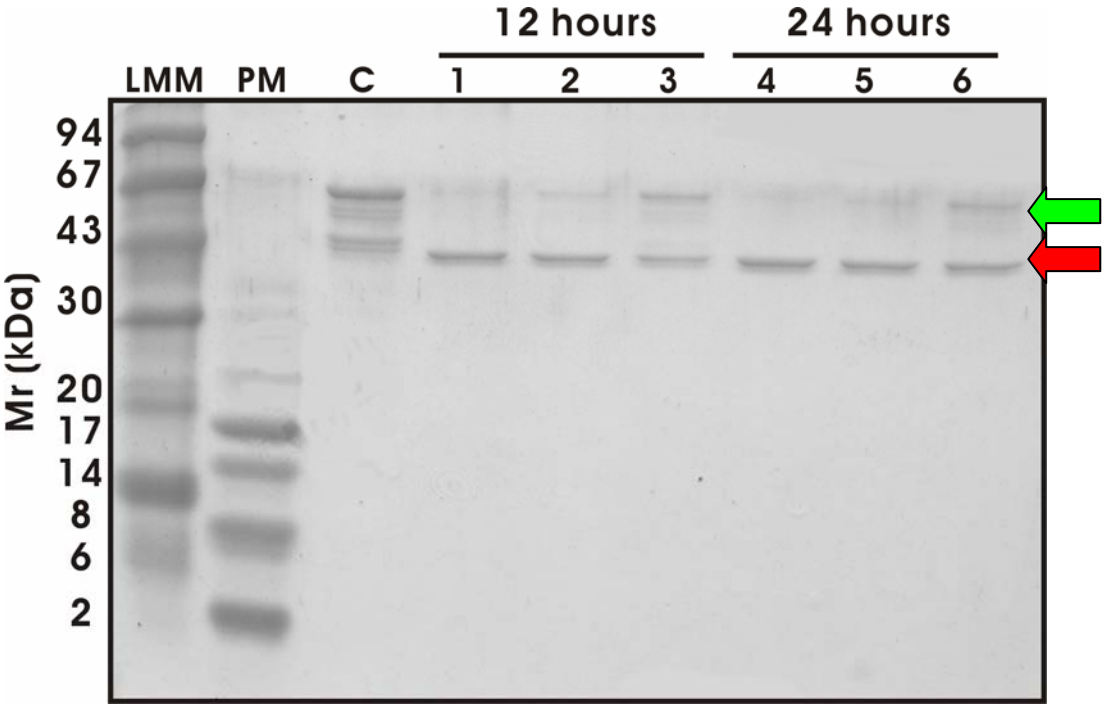


Figure 3.15 Analysis of the light-chain enterokinase cleavage of MBP-Fsav by tricine SDS-PAGE. LMM: low molecular mass marker. PM: peptide marker. C: MBP-Fsav fusion protein (produced from *SURE* cells). Lane 1- 3: 1:20 000, 1:50 000, 1:100 000 (w/w) a 12 hour cleavage period. Similarly Lane 4- 6: a cleavage period of 24 hours. Green arrow indicates MBP-Fsav. Red arrow indicates MBP.

In the case of factor Xa, the cleavage is after its preferred cleavage site Ile-Glu-Gly-Arg, so that no residues are attached to the protein of interest after cleavage. In a pilot experiment it was shown that a 1:50 (w/w) of factor Xa: fusion protein was optimal for cleavage. The factor Xa was employed to cleave both purified MBP-Fsav and MBP-Csav molecules. Analysis using tricine SDS-PAGE (Figure 3.16) indicates the mobility shift between digested and undigested fusion protein. In the case of MBP-Fsav, the mobility shift was from 54.7 kDa to 42.7 kDa, whereas in the case of MBP-Csav a mobility shift from 48.7 kDa to 42.7 kDa was observed. The expected full length savignin (~ 20 kDa) and C-domain savignin (~ 10 kDa) bands were however not observed.

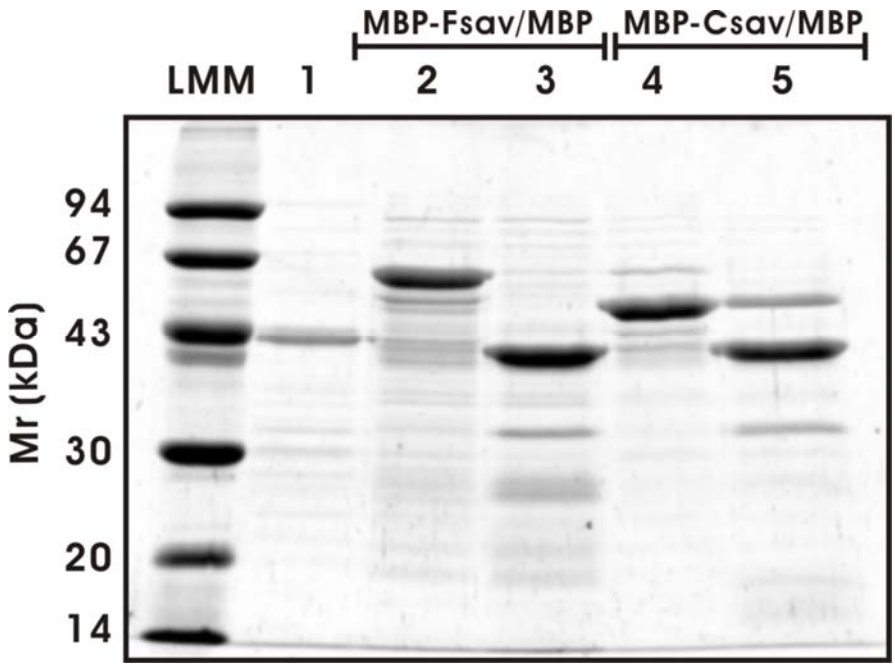


Figure 3.16 Analysis of factor Xa cleavage of MBP-Fsav and MBP-Csav by tricine SDS-PAGE. LMM: low molecular mass marker. Lane 1: MBP. Lane 2: undigested MBP-Fsav. Lane 3: digested MBP-Fsav. Lane 4: undigested MBP-Csav. Lane 5: digested MBP-Csav

3.3.10 Protein purification strategies

3.3.10 a) Strategy 1

The first attempt at purifying recombinant proteins (OD_{280} yield of ~ 430.63 µg /ml) included amylose affinity chromatography (Figure 3.17A) and AEHPLC (Figure 3.17D). Analysis of eluted fractions from the amylose column is presented in Figures 3.17B and C. As shown in Figure 3.17C, the fusion protein did not bind as expected to the affinity column. The fraction obtained from AEHPLC containing MBP-Fsav was then subjected to enterokinase cleavage (~ 100 µg purified fusion protein: ~ 2 ng light chain enterokinase). The cleaved fraction was then chromatographed using RPHPLC. The result of this experiment is shown in Figure 3.17E. Three peaks were observed with retention times of 26 - 28, 45 – 47 and 51 - 53 minutes. Thrombin inhibition assays for fractions obtained from RPHPLC is shown in Figure 3.17F. No inhibitory activity was observed for any of three peaks.

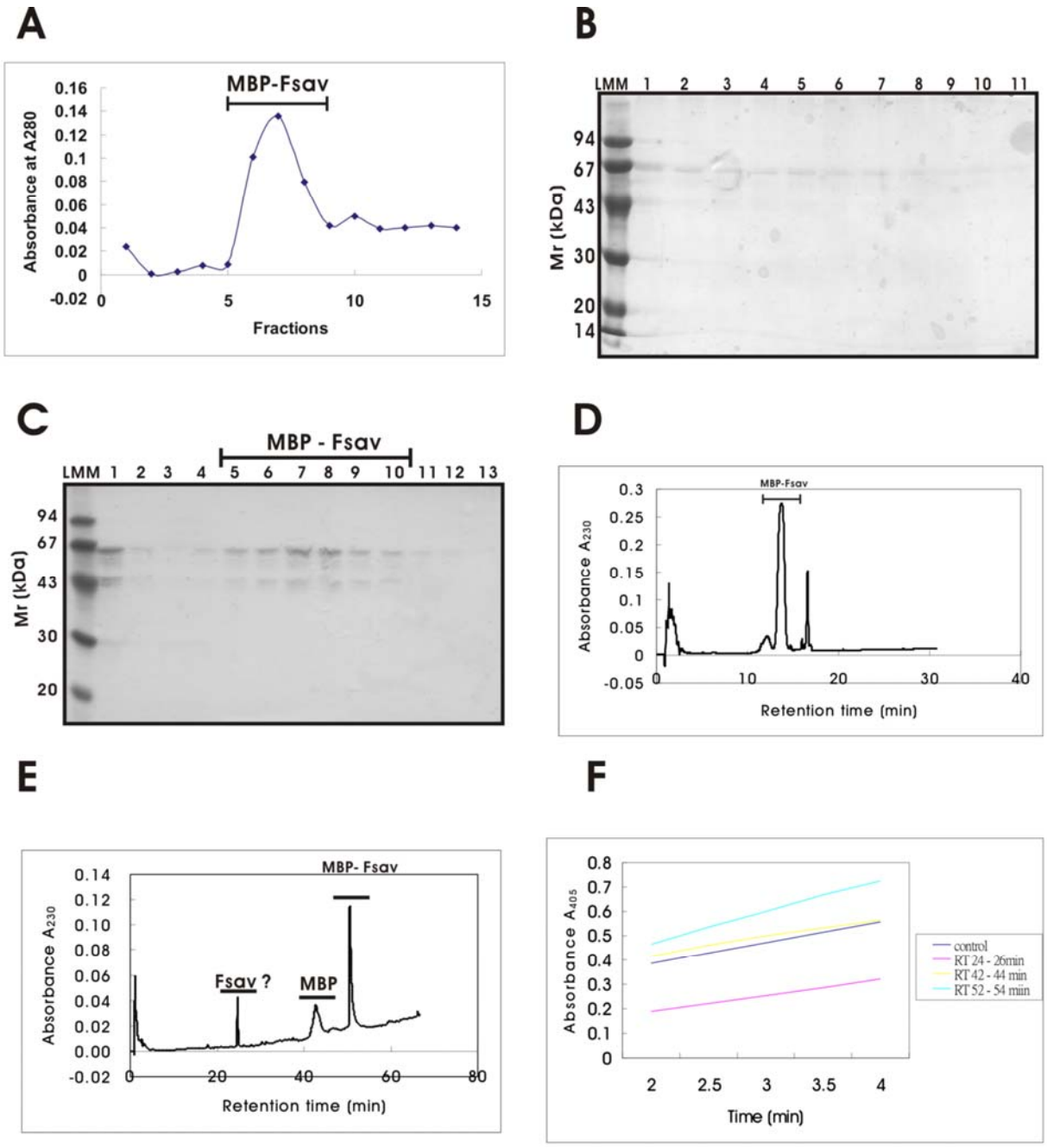


Figure 3.17 Results from strategy 1. A): Amylose affinity chromatography after the wash step. B): SDS-PAGE analysis of washes with 1 ml column buffer. Lane 1 - 11: indicates the fraction collected after each wash. C): SDS-PAGE analysis of fractions eluted from the amylose resin. LMM: low molecular mass marker, Lane 1 - 13: the fraction collected from the elution. D): Anion exchange chromatography of MBP-Fsav. Fraction (1 ml) was collected (retention time 12 ~ 13 minutes), dialyzed, and freeze dried for further enterokinase cleavage. E): Reversed phase chromatography of cleaved sample. Three peaks were detected (Retention time of 24 - 26 minutes, Retention time of 45 - 47 minutes, Retention time of 55 - 57 minutes). F): Anti-thrombin activity for each peak.

3.3.10 b) Strategy 2

The crude periplasmic fusion protein as monitored at OD_{280} gave a yield of ~ 430 μg/ml. In the second strategy, the affinity column was omitted to minimize protein loss. The periplasmic protein extract was applied to an anion exchange column. AEHPLC chromatogram showed a peak which contained the MBP-Fsav that was eluted at 12 – 14 minutes (Figure 3.18A). The presence of MBP-Fsav in this peak was confirmed by SDS-PAGE analysis (results not shown).

After purification by AEHPLC, fusion protein (~ 100 μg/ml) was subjected to enterokinase cleavage (refer to section 3.2.19b). The cleaved sample was subjected to a C18 reversed-phase column. Only one peak was observed at a retention time of 45 – 47 minutes (MBP). The expected savignin peak (from previous observations by Nienaber, Gaspar & Neitz, 1999, at retention between 24 – 26 minutes) was absent from the chromatogram (Figure 3.18B). The fractions between retention times 1 - 2, 24 - 26, 45 - 47 minutes were collected and vacuum dried and assayed for anti-thrombin activity (Figure 3.18C). Anti-thrombin activity was detected in the fraction collected at retention time 24 – 26 minutes, even though no peak was observed in chromatogram at this retention time.

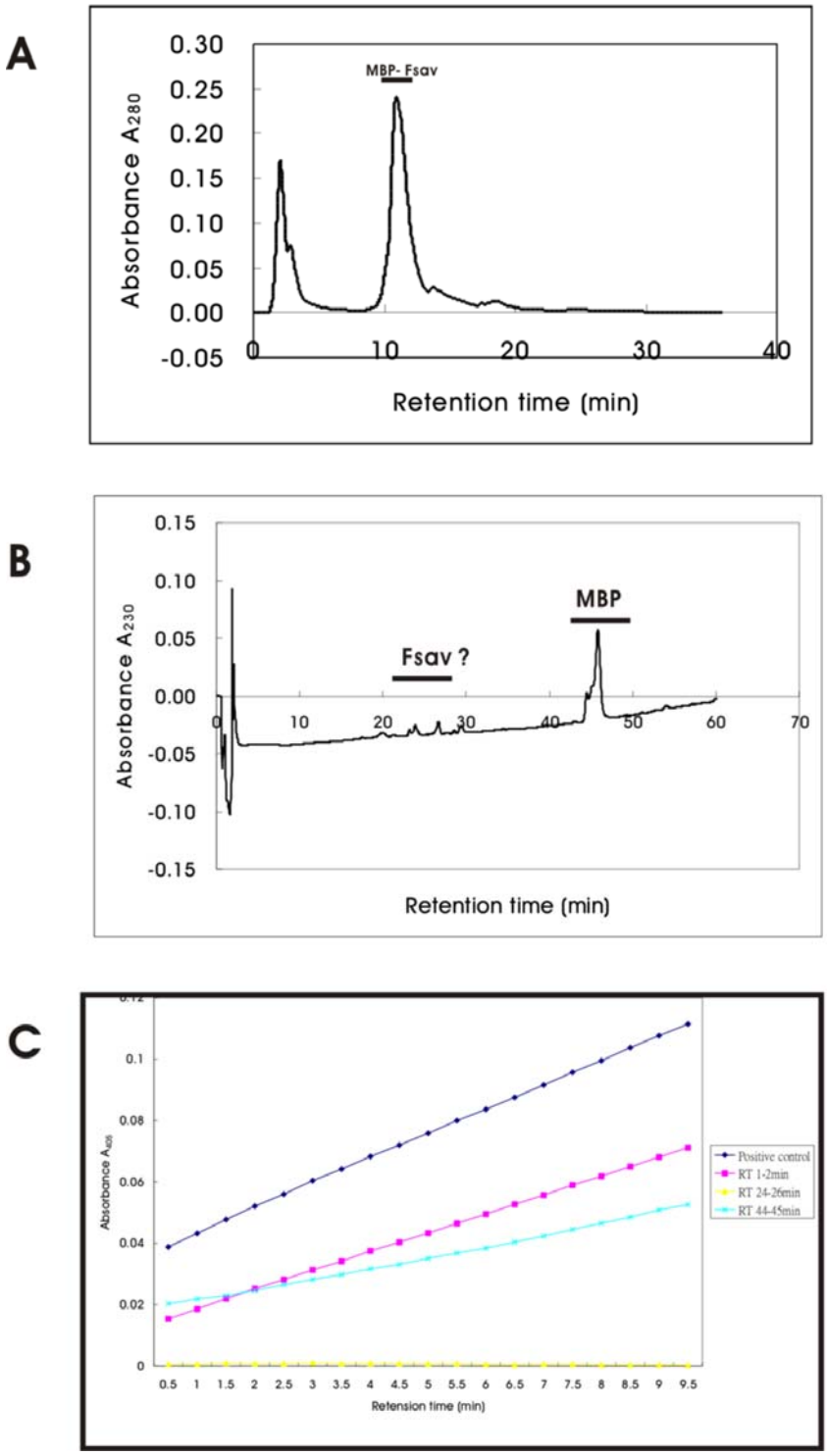


Figure 3.18 Results from purification strategy 2. A): Anion exchange chromatography of a MBP–Fsav. Fractions (1 ml) at RT 12 – 13 minutes were collected, dialyzed and freeze dried for enterokinase cleavage. B): Reversed phase chromatography of cleaved sample. Two distinctive peaks were detected (RT 1 - 2minutes, RT 45 - 47minutes). C): Fractions from reversed phase were vacuum dried and the anti-thrombin activity monitored. RT 1 - 2 minutes indicated in pink. RT 24 - 26 minutes indicated in yellow. RT 45 - 47 minutes indicated in light blue. Positive control indicated in dark blue.

3.3.10 c) Strategy 3

As in strategy 2 the affinity column was omitted, but fXa was used instead of rEK_L. As in strategy 2, crude periplasmic protein extracts containing MBP-Fsav (~ 485.3 µg/ml) and MBP-Csav (~ 651 µg/ml), respectively, were to an anion exchange column (Figure 3.19A and B). To confirm the identity of these peaks SDS-PAGE analysis was done. The results indicated that MBP-Fsav eluted at retention time of 26 - 29 minutes, whereas MBP-Csav eluted at retention time of 26 - 30 minutes.

The purified MBP-Fsav and MBP-Csav were subjected to factor Xa cleavage (refer to section 3.2.19b). After factor Xa cleavage, digested samples were applied to a pre-equilibrated reversed phase column. Results of the RPHPLC profile for the MBP control is shown in Figure 3.19C. MBP elutes at a with retention time of 22 - 25 minutes. This retention time is similar to the expected retention time of native savignin (Nienaber, Gaspar & Neitz, 1999). In previous results (strategy 1 and 2) the retention time for MBP was from 45 - 47 minutes. The Fsav peak obtained in the chromatogram (Figure 3.19D) cannot therefore discerned from MBP, since the retention times were similar (22 - 25 minutes). In Figure 3.19E (chromatogram of Csav) only the peak corresponding to MBP was observed. No peak was for Csav observed.

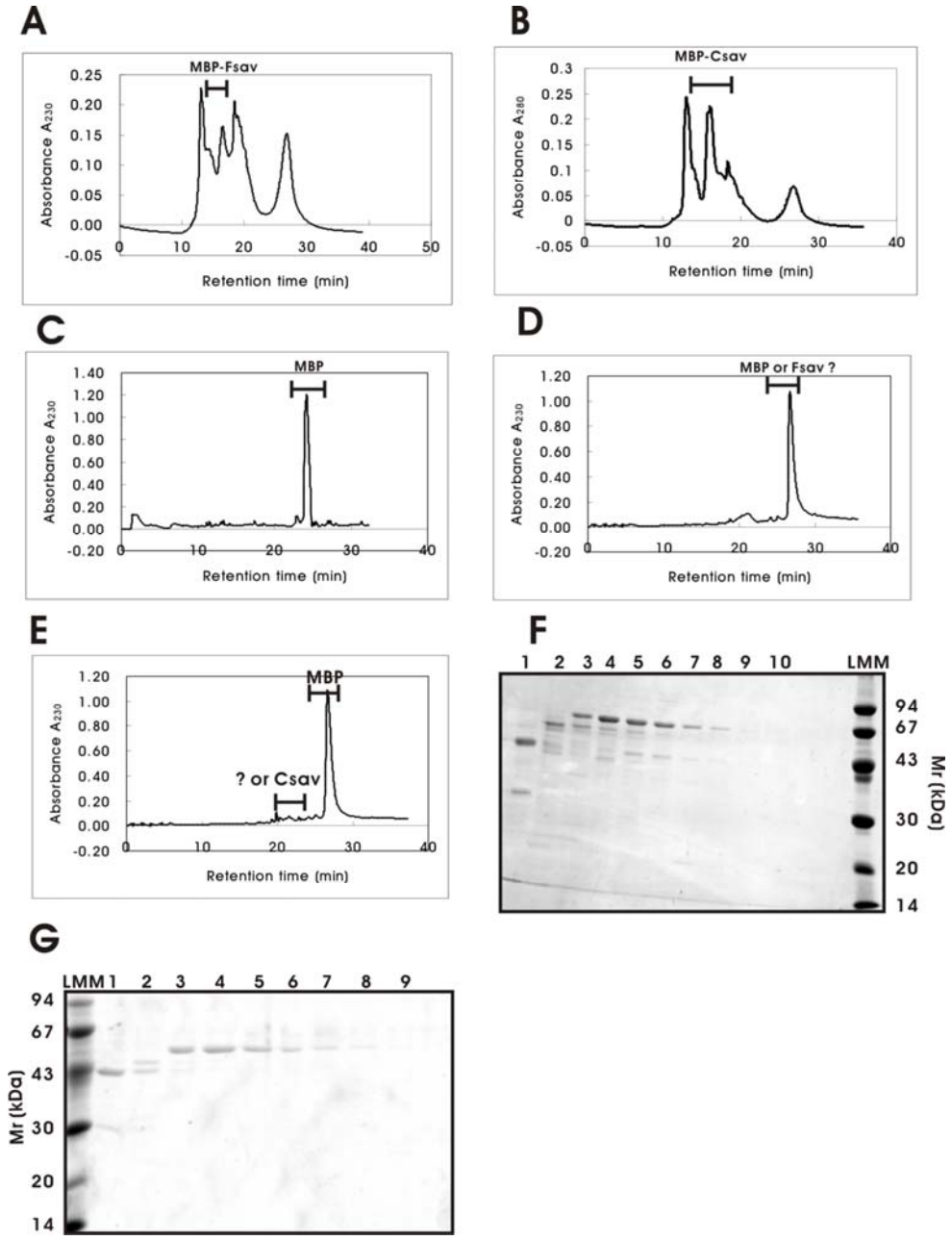


Figure 3.19 Results from purification strategy 3. A): Chromatogram of periplasmic protein contains MBP-Fsav from AEHPLC. B): Chromatogram of periplasmic protein contains MBP-Csav from AEHPLC. C): Chromatogram of MBP (positive control) from RPHPLC. D): Reversed phase chromatography of recombinant full length savignin after factor Xa cleavage. E): Reversed phase chromatography of recombinant C-domain savignin after factor Xa cleavage. F): Tricine SDS-PAGE analysis of Fsav fractions obtained from reversed phase chromatography. LMM: low molecular mass marker, Lane 1: RT 22 - 24 minutes. Lane 2: RT 24 - 26 minutes. Lane 3 - 10: RT 26 - 32 minutes. G): Tricine SDS-PAGE analysis of Csav fractions obtained from reversed phase chromatography. LMM: low molecular mass marker, Lane 1: RT 22 - 24 minutes. Lane 2: RT 24 - 26 minutes. Lane 3 - 9: RT 26 - 32 minutes.

Since MBP has a similar retention time to native savignin, the peaks were analyzed or the presence of Fsav and Csav. The tricine SDS-PAGE results indicate that more than one protein band is present in peaks containing Fsav and Csav, respectively (Figure 3.16F and G). Samples were also examined for anti-thrombin activity. Slight inhibitory activity was detected in the peak containing Fsav (Figure 3.20). No inhibitory activity was observed in any of the fractions obtained after RP chromatography of cleaved MBP-Csav.

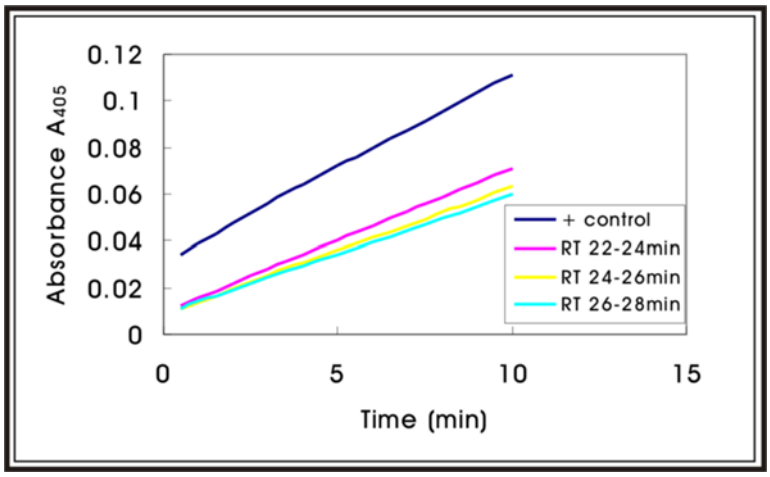


Figure 3.20 Anti-thrombin activity analyses for Fsav. Anti-thrombin activity analysis of Fsav fractions obtained from the reversed-phase chromatography.

3.4 Discussion

In order to investigate the structural features of savignin in the uncomplexed form and in complex with thrombin, the individual domains with their binding sites as well as adequate concentrations of savignin are needed. These investigation and kinetic studies require molecular biological techniques in order to obtain sufficient amounts of materials.

pMAL vectors have commonly been used for expression of soluble proteins or soluble domains of membrane proteins (Ko et al., 1993; Hiraoka et al., 1994). The vector pMAL-p2, which possesses a periplasmic localization signal, significantly increases the expression level of intact fusion proteins. The expression of MBP-Fsav and MBP-Csav were successful, but the expression of MBP-Nsav was not. This was most unexpected. Possible explanations for this result could be that 1) there was a loss of selection pressure, resulting in rejection of plasmid, 2) the protein was not targeted into the periplasmic space, 3) the proteins taking advantage of the larger volume of cytoplasm compared to the periplasmic space, 4) the cell tries to minimize toxicity by reducing the accumulation of foreign proteins in its periplasmic space, and 5) protease degradation occurs during the formation of the fusion protein.

No blue/white selection process was done with pMAL-p2 plasmid even though the β-galactosidase fragment is the same on both pGEM and pMAL vectors. The tac promoter on the pMAL expression vectors is much stronger than the lac promoter on the pGEM cloning vector. Cells bearing a pMAL vector when induced with IPTG die eventually. Therefore, alternative selection processes such as sequencing and electrophoresis were used.

Cell transformation was further tested using TB1 and *Sure* strains of *E. coli*. The best results were obtained using *Sure* cells with respect to efficiency of transformation and protein yield for both savignin and its truncated forms. *Sure* cells were transformed more efficiently by both electroporation and heat shock procedures.

During the efforts to overexpress the recombinant forms of savignin, expression of recombinant pMAL-p2E vector in TB1 and *Sure* strains of *E. coli* were tested. The expression levels in TB1 cells as monitored by the Bradford method and SDS-PAGE were either undetectable or very low (Figure 3.10). The expression level of recombinant forms of savignin from TB1 strains required a longer time period for induction. In striking contrast, the expression levels of recombinant forms of savignin from *Sure* strains appeared to have higher yields and required shorter induction time periods

To monitor the MBP-Fsav/Csav production (indicated by the expression of the crude fusion proteins) quantitatively, measurements were taken using the Bradford assay and SDS-PAGE analysis. These approaches were used to determine the optimal IPTG concentrations and the time of induction (Figure 3.12) for protein expression. Inducing the expression of a MBP fusion protein with either 0.1 mM IPTG or 0.3 mM IPTG gave the same results, whereas induction with 1 mM IPTG decreased in the fusion protein level. This was possibly due to toxic side effects by protein overexpression. Induction at 25 °C and 37 °C resulted in similar expression yields of fusion proteins and prolonged expression time resulted in a significant increase in expression levels (Figure 3.13 and Figure 3.11, respectively).

Because the present *E. coli* expression and purification strategies for recombinants did not give enough protein yields and purity of purified savignin, it was difficult to make a direct, quantitative comparison between purification strategies. In the purification strategies only minute quantities of recombinants were recovered. Even though MBP-Fsav and MBP-Csav yielded large quantities of material, separation of the target protein from the fusion partner by enzymatic means presented a major hurdle.

When considering proteolytic methods, poor results may occur due to unfavourable conformation at or near the scissile bond or the presence of additional, undesired cleavage sites. For example, enterokinase recognizes and cleaves after lysine at the amino acid sequence (Asp-Asp-Asp-Asp-Lys) site, but may sometimes cleave at other basic residues depending on the conformation of the protein substrate and the adjacent sequence (Pacher, Jürgens & Guttenberg, 2004).

Unsatisfactory results obtained using enterokinase to cleave target proteins from fusion proteins combined with its high cost led to exploration of factor Xa instead. A factor Xa cleavage site was introduced into pMAL-p2E plasmid via primers.

The normal recognition for factor Xa is Ile-Glu-Gly-Arg (Nagai & Thøgersen, 1987). Wearne (1990) found that a number of secondary sites showed factor Xa cleavage after Gly-Arg amino acid residues. Although this sequence is not found in savignin, non-specific cleavage may have occurred and explains the absence of the recombinant form after cleavage from the fusion protein.

It is noteworthy to mention, that even though the expected Fsav peak was not detected after RPHPLC, in the second purification strategy anti-thrombin activity was observed in the expected region. Peaks in the reversed-phase chromatogram shown in strategy 1 displayed no thrombin inhibition.

In general, the pMAL-p2 expression system gave the highest yield for MBP-Fsav (485.3 μg/ml) and MBP-Csav (651 μg/ml). No expression of MBP-Nsav was obtained. MBP-Nsav may have been degraded because the passage through the membranes may have been prevented resulting that the fusion proteins trapped in the membrane would be hindered from folding in the correct way (Gentz et al., 1988). Many proteins, especially those that have multiple disulphide bonds, only fold properly when exported. The *E. coli* cytoplasm is a reducing environment, and the proteins that catalyze disulphide bond formations are present in the periplasm (Bardwell et al., 1991). This probably explains the reason why the recombinant MBP-Nsav proteins are either degraded or inactive.

The retention time of MBP in the reversed-phase chromatogram (Figure 3.19D) differs between strategy 3 and strategies 1 and 2. One of the possible reasons to explain this behaviour is that partial truncation of MBP occurs during the fXa cleavage resulting in a surface area alteration which resulted in changes in the retention time.

In conclusion, the savignin recombinant may be unstable, misfolded or toxic to *E. coli* cells even though they were fused to a stable MBP. To obtain sufficient recombinant proteins alternative expression systems will have to be explored.

Chapter 4

Concluding Discussion

Savignin was characterized as a competitive, slow-binding inhibitor; and kinetic investigations indicate some form of structural rearrangement (Nienaber, Gaspar & Neitz, 1999). It was shown that savignin binds to γ-thrombin, a derivative of α-thrombin, which lacks exosite 1 (fibrinogen binding site). The Ki values of savignin for α- and γ-thrombin are 5 pM and 20 nM, respectively. The reduced affinity of savignin for γ-thrombin suggests that savignin may interact with exosite I, but this remains to be confirmed. Structural modelling of savignin based on the crystal structure of ornithodorin, the inhibitor from *O. moubata*, and docking to thrombin indicated similarities to ornithodorin (Mans, Louw & Neitz, 2002). Like ornithodorin, the N-terminal domain of savignin binds to the active site, while its C-terminal domain interacts with the basic fibrinogen recognition exosite of thrombin (Mans, Louw & Neitz, 2002).

Mans (2002) hypothesized that in the absence of thrombin, the domains of savignin interact with each other, giving a globular form. Binding of the C-terminal domain of savignin to the fibrinogen-binding site of thrombin leads to the dissociation of the savignin domains. This would yield an extended conformation that would allow the N-terminal residues of the N-terminal domain of savignin to bind within the thrombin's active site. If two binding sites exist, it is probable that the inhibitor first binds to one site and then to the other. This dissertation describes investigations that will contribute to a better understanding of the inhibition mechanism and the structural conformation of uncomplexed savignin.

In the first experimental chapter the hypothesis of Mans was tested by trying to determine the Rh of uncomplexed savignin. DLS was proposed to aid the study. Unfortunately, we found out that the specific equipment required does not have the dynamic detector needed for Rh measurement (the plan was to use the Precision Detector at UNISA) and no other DLS equipment is available in SA at present. Therefore, the Rh of uncomplexed-savignin could not be determined. This may require international collaboration.

The SEC results suggest the presence of both globular and extended conformations. Inhibitory activity was found in both the extended form (major inhibition) and globular form (minor inhibition). There was a third region (retention time 52 - 56 minutes) which also displayed some inhibition, which may be due to the presence of a second unidentified thrombin inhibitor. From these investigations it was concluded that uncomplexed savignin exists primarily in an extended conformation that correlates with that of the crystal structure of complexed ornithodorin.

If savignin truly exists in an extended conformation, then the only satisfying explanation for its anomalous electrophoretic mobility during electrophoresis is that the N- and C-terminal domains are very stable on their own and do not unfold under usual denaturing conditions, explaining the thermal stability of savignin. Another explanation for the observed electrophoretic behaviour could be that the compact, globular form opened up into an extended conformation when the compact, globular form was exposed to both the heated and reductive environment, which could then be further reduced into a linearized protein chain.

During this study a method to purify savignin in higher yields was developed (yield of of 138 µg from 100 salivary glands) when compared to a yield of 14 µg obtained from the same number of glands with the previous strategy. This purification method might be used in the future to purify more inhibitor for the purpose of studying the kinetics of native savignin in greater detail.

The pMAL-p2 expression system was used for recombinant expression. This vector, which possesses a periplasmic localization signal, significantly increases the expression level of intact fusion proteins (Ko et al., 1993; Hiraoka et al., 1994), but has shortcomings due to its selection process. No blue/white selection could be done, because the lacZ promoter region was truncated in the pMAL-p2 plasmid.

Cell transformation was further tested using TB1 and *Sure* strains of *E.coli*. The best results were obtained using *Sure* cells with respect to the efficiency of transformation and protein yield for both savignin (MBP-Fsav ~485.3 µg/ml) and its truncated form (MBP-Csav ~651 µg/ml). *Sure* cells were transformed more efficiently by both electroporation and heat shock procedures. No MBP-Nsav was detected. This may be due to the specific folding of the recombinant which may prevent its passage through the membranes. The trapped fusion proteins are thus hindered from folding in the correct way, which could lead to their degradation (Gentz et al., 1988).

MBP-Fsav and MBP-Csav yielded large quantities of material but separation of the target protein from the fusion partner by enzymatic means presented a major hurdle, and as a result only minute quantities of cleaved recombinants were recovered. This could be due to unfavourable conformations at or near the scissile bond and/or the presence of additional, undesired cleavage sites (Pacher, Jürgens & Guttenberg, 2004).

Unsatisfactory results obtained using enterokinase to cleave target proteins from fusion proteins combined with its high cost led to exploration of factor Xa as cleavage enzyme instead. A factor Xa cleavage site was introduced into the pMAL-p2E plasmid via primers used to generate pMAL-p2X containing inserts. Although fXa is less costly, it was found that cleavage of either MBP-Fsav or MBP-Csav resulted in proteolytic degradation of the recombinants. The normal recognition site for factor Xa is Ile-Glu-Gly-Arg (Nagai & Thøgersen, 1987). Wearne (1990) found that a number of secondary sites showed that factor Xa cleaves after Gly-Arg amino acid residues. Although this Gly-Arg amino acid residue sequence is not found in savignin, factor Xa probably cleaved savignin non-specifically.

It is noteworthy to mention, that even though the expected Fsav peak was not detected after RPHPLC in strategy 1, in the second purification strategy anti-thrombin activity was observed in the expected region. None of the peaks obtained during the reversed-phase chromatography in strategy 1 displayed thrombin inhibition.

The retention time of MBP in the reversed-phase chromatogram (Figure 3.19D) of strategy 3 differs from strategies 1 and 2. One of the possible reasons to explain this behaviour is that partial truncation of MBP occurs during fXa cleavage, resulting in a surface area alteration.

The goal for initiation of this project was overexpression of recombinant forms of savignin to allow the high throughput generation of clones producing active and soluble protein for purification. This low yield obtained after purification of the recombinant forms after cleavage from the fusion protein is likely attributable to proteolytic degradation as well as the number of purification steps involved. An

expression system that will yield protein in a form that does not involve proteolytic digestion or many purification steps should herefore be investigated for the production of recombinant savignin.

There are various expression strategies which have been used to express inhibitors from hematophagous organisms. With the recent advances in plant biotechnology, transgenic plants have been targeted as inexpensive means for the production of hirudin (Parmenter, 1995). The advantages of this approach are the fast, efficient, and economical isolation of the target protein under mild conditions, thus preserving the activity of the target protein. However, the current plant purification techniques lack a generally applicable, economic, large-scale strategy. Expression in a yeast (*Saccharomyces cerevisiae*) system was used to express tick anticoagulant peptide, a factor Xa inhibitor from *O. moubata* (Cook et al., 1998). The baculovirus-insect cell expression system was used to express triabin, a thrombin inhibitor from triaomine bugs (Prior et al., 1997).

Alternatively, pET vectors may be employed. These pET vectors contain different sequences adjacent to the cloning sites that encode a number of peptide “tags”, which perform localization, detection or purification functions when fused with the target protein. The method of cloning will determine whether or not these “tags” or any additional amino acids from the vector are expressed as a fusion to the protein of interest. There are several cloning options to produce target proteins with or without fusions (pET instruction manual, Novagen 2002). This sytem may have disadvantages because of the tags, since the tags attach to the savignin and might interfere with the inhibitory activity and the tags are not easily removed.

A cell-free *in vitro* expression system (T7 sample system) is another option. It is a unique translation system for one-step gene expression from circular and linear plasmid DNA templates, as well as non-purified PCR products. These expression systems offer significant time savings over *in vivo* systems and are easy to use, quantitative and suitable for high-throughput systems. The use of DNA in a coupled transcription/translation format allows for the development of many *in vitro* expression applications (T7 sample system technical bulletin, Promega).

Expression of savignin should provide the means for further structural characterization in terms of mechanism. An intriguing possibility that emerges from this study is the design of a chimeric protein that attacks fXa, thrombin and platelet aggregation simultaneously. Such a protein might be useful as a multi-functional agent to control thrombosis in a regulated manner. It could also be used as a possible vaccine agent, to generate an immune response that could knock out more than one function necessary for tick feeding.

Summary

Mans (2002) hypothesized that the two domains of savignin interact with each other, giving a globular form in the absence of thrombin. Binding of the C-terminal domain of the inhibitor to the fibrinogen-binding site of thrombin leads to the dissociation of the domains. This would yield an extended conformation that would allow binding of the N-terminal residues of the N-terminal domain to thrombin's active site.

To test this hypothesis, both theoretical and experimental approaches were employed to determine the molecular dimensions of uncomplexed savignin. In the theoretical approach, the hydrodynamic radius (Rh) of the extended form of savignin was calculated from the crystal structure data of the thrombin-ornithodorin complex, and found to be 2.319 nm. With the same programme, based on the crystal structure data for bikunin, a protein in which both domains are closely associated, the Rh value for the compact form of savignin was estimated as 1.96 nm. Using the equation that relates Rh to molecular mass, a value of 1.84 nm was calculated for savignin (12 430 Da).

In the experimental approach, the SEC of salivary gland extracts, using lysozyme (Rh = 1.99 nm) as standards and chrymotrypsinogen (Rh = 2.31nm) indicated that uncomplexed savignin exists in both the globular and extended conformations. However, the majority of inhibitory activity was associcated with the extended form. Heat stability assays as well as SDS-PAGE experiments indicated the possible existence of the compact form of savignin.

Generation of adequate amounts of savignin will allow further structural studies, to determine the structure of savignin in the uncomplexed form and in complex with thrombin.

Expression of full length savignin and the separate N- and C-domains will facilitate further kinetic analysis. In the recombinant production of savignin, various factors were investigated: cell-lines, transformation efficiency, induction times, purification strategy and protease cleavage of expressed fusion protein. Even though large quantities of expressed fusion protein were obtained, cleavage of the target protein and its separation from the fusion partner by enzymatic means presented a major hurdle. Expression of Nsav was not observed and is most likely as a result of misfolding of the recombinant form.

Due to the probable non-specific cleavage of the fusion protein, switching the prokaryotic expression system to other expression systems, like yeast- or baculovirus-insect cell-expression systems, is warranted.

References

Ackers GK. Molecular exclusion and restricted diffusion processes in molecular-sieve chromatography. *Biochemistry.* 1964; **72**: 723-730

Ackers GK. Analytical gel chromatography of proteins. *Adv Protein Chem.* 1970; **24**: 343-446

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic locale alignment search tool. *J Mol Biol.* 1990; **215**: 403-410

Bardwell JC, McGovern K, Beckwith J. Identification of a protein required for disulfide bond formation in vivo. *Cell.* 1991; **67**:581-589

Bjork I, Lindahl U. Mechanism of the anticoagulant action of heparin. *Mol Cell Biochem.* 1982; **48**:161–182

Boyle JS, Lew AM. An inexpensive alternative to glassmilk for DNA purification. *Trends Gen.* 1995; **11**: 8

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities od protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976; **72**: 248-254

Broze GJ Jr. Tissue factor pathway inhibitor. *Thromb Haemost.* 1995; **74**:90–93

Cabré F, Canela EI, Canela MA. Accuracy and precision in the determination of Stokes radii and molecular masses of proteins by gel filtration chromatography. *J Chrom. 1989;* **472**: 347-356

Cappello M, Li S, Chen X, Li CB, Harrison L, Narashimhan S, Beard CB, Aksoy S. Tsetse thrombin inhibitor: bloodmeal-induced expression of an anticoagulant in salivary glands and gut tissue of *Glossina morsitans morsitans*. *Proc Natl Acad Sci USA.* 1998; **95**: 14290–14295.

Carrasco B, Garcia de la Torre J. Hydrodynamic properties of rigid particles. Comparison of different modelling and computational strategies. *Biophys J.* 1999; **76**: 3044-3057

Choay J, Petitou M. The chemistry of heparin: a way to understand its mode and action. *Med J Aust.* 1986; **144**: 7-10

Creighton TE. Proteins. Structures and Molecular Properties. Second Edition. W.H. Freeman and Company, NewYork. 1992.

Collen D. The plasminogen (fibrinolytic) system. *Thromb Haemost.* 1999; **82**:259–270

Collins-Racie LA, McColgan JM, Grant KL, Diblasio-Smith EA, McCoy JM, LaVallie ER. Production of recombinant bovine enterokinase catalytic subunit in *Escherichia coli* using the novel secretory fusion partner DsbA. *Biotech.* 1995; **13**: 982-987

Cook JC, Schultz LD, Huang J, George HA, Herber WK. Ip C, Joyce JG, Mao SS, Markus HZ, Miller WJ, Sardana MK, Lehman ED. Expression and purification of recombinant tick anticoagulant peptide (Y1W/D10R) double mutant secreted by *Saccharomyces cerevisiae*. *Protein Exp and Puri.* 1998; **13**: 291-300

Duplay P, Bedouelle H, Fowler A, Zabin I, Saurin W, and Hofnung M. Sequences of the malE gene and of its product, the maltose binding protein of *Escherichia coli.* K12. *J Biol Chem.* 1984; 259: 10606-10613

Esmon CT, Ding W, Yasuhiro K, Gu JM, Ferrell G, Regan LM, Stearns-Kurosawa DJ, Kurosawa S, Mather T, Laszik Z, Esmon NL. The protein C pathway: new insights. *Thromb Haemost.* 1997; **78**:70–74

Francischetti IMB, Valenzuela JG, Ribeiro JMC. Anophelin: kinetics and mechanism of thrombin inhibition. *Biochemistry.* 1999; **38**: 16678–16685.

Friedrich T, Kröger K, Bialojan S, Lemaire HG, Höffken HW, Reuschenbach P, Otte M, Dodt J. A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus. J Biol Chem.* 1993; **268**: 16216-16222.

Furie B, Furie BC. Molecular and cellular biology of blood coagulation. *N Engl J Med.* 1992; **326**:800–806

Gailani D, Broze GJ Jr. Factor XI activation by thrombin and factor XIa. *Semin Thromb Hemost.* 1993; **19**:396–404

Garcia de la Torre J, Huertas ML, Carrasco B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J.* 2000; **78**: 719-730

Gaspar AR, Joubert AM, Crause JC, Neitz AW. Isolation and characterization of an anticoagulant from the salivary glands of the tick, *Ornithodoros savignyi* (Acari: Argasidae). *Exp Appl Acarol.* 1996; **20**:583-98.

Gentz R, Kuys Y, Zwieb C, Taatjes D, Taatjes H, Bannwarth W, Stueber D, Ibrahimi I. Association of degradation and secretion of three chimeric polypeptides in *Escherichia coli*. *J Bacteriol.* 1988; **170**: 2212-2220.

Gresele P, Agnelli G. Novel approaches to the treatment of thrombosis. *Trends Pharm Sci.* 2002; **23**: 25-32

Guan C, Li P, Riggs PD, Inouye H. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene.* 1987; **67**: 21-30

Guex N, Diemand A, Peitsch MC. Protein modelling for all. *Trends Biochem Sci.* 1999; **24**: 364-367

Hall T. BioEdit programme. Department of Microbiology. North Carolina State University. 2001; Version 5.0.6

Hanahan D, Jesse J. Bloomer FR. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol.* 1991; **204**: 63-113

Herbert JM, Bernat A, Dol F, Herault JP, Crepon B, Lormeau JC. DX-9065a, a novel, synthetic, selective and orally active inhibitor of factor Xa: *in vitro* and *in vivo* studies. *J Pharmacol Exp Ther.* 1996; **276**:1030–1038

Hilpert K, Ackerman J, Banner DW, Gast A, Gubernator K, Hadvary P, Labler L, Muller K, Schmid G, Tschopp TB. Design and synthesis of potent and highly selective thrombin inhibitors. *J Med Chem.* 1994; **37**: 3889–3901

Hiraoka O, Anaguchi H, Yamasaki K, Fukunaga R, Nagata S, Ota Y. Ligand binding domain of granulocyte colony-stimulating factor receptor. *J Biol Chem.* 1994; **269**: 22412-22419

Hirsh J. Heparin. *N Engl J Med.* 1991a; **324**: 1565–1574

Hirsh J. Oral anticoagulant drugs. *N Engl J Med.* 1991b; **324**: 1865–1875

Hirsh J, Dalen JE, Anderson DR, Poller L, Bussey H, Ansell J. Heparin: mechanism of action, pharmacokinetic, dosing considerations, monitoring, efficacy and safety. *Chest.* 1995; **108**: 258S-275S

Hirsh J. Current anticoagulant therapy-unmet clinical needs. *Thromb Res.* 2003; **109**: S1-S8

Huntington JA, Baglin TP. Targeting thrombin – rational drug design from natural mechanism. *Trends Pharm Sci.* 2003; **24**: 589-595

Howell CJ, Neitz AWH, Potgieter DJJ. Some toxic and chemical properties of the oral secretion of the sand tampan, *Ornithodoros savignyi* Audouin (1825). *Onderstepoort J Vet Res.* 1975; **43**: 99-102

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Fibson TJ. Multiple sequence alignment with Clustal X. *Trends Biochem Sci.* 1998; **23**: 403-405

Jones S, Thornton JM. Protein-protein interactions: a review of protein dimer structures. *Prog Biophys Mol Biol.* 1995; **63**:31-59

Jordan SP, Mao SS, Lewis SD, Shafer JA. Reaction pathways for inhibition of blood coagulation factor Xa by tick anticoagulant peptide. *Biochemistry.* 1992; **31**: 5374-5480

Karczewski J, Endris R, Connolly TM. Disagregin is a fibrinogen receptor antagonist lacking the Arg-Gly-Asp sequence from the tick, *Ornithodoros moubata. J Biol Chem.* 1994; **269**:6702-6708.

Kellerman OK, Ferenci T. Maltose binding protein from *E.coli*. *Methods Enzymol.* 1982; **90**: 459-463

Ko YH, Thomas PJ, Delannoy MR, Pedersen PL. The cystic fibrosis transmembrane conductance regulator. Overexpression, purification, and characterization of wild type and delta F508 mutant forms of the frist nucleotide binding fold in fusion with the maltose-binding protein. *J Biol Chem.* 1993; **268**: 24330-24338

Laemmli U. Cleavage of sturctural proteins during the asssembly of the head of bacteriophage T4. *Nature.* 1970; **227**: 680-685

Lange U, Keilholz W, Schaub GA, Landmann H, Markwardt F, Nowak G, Biochemical characterization of a thrombin inhibitor from the bloodsucking bug Dipetalogaster maximus, *Haemostasis.* 1999; **29**: 204–211.

Law JH, Ribeiro JMC, Wells MA. Biochemical insights derived from insect diversity. *Annu Rev Biochem.* 1992; **61**: 87–111.

Li S, Kwon J, Aksoy S, Characterization of genes expressed in the salivary glands of the tsetse fly, Glossina morsitans morsitans, *Insect Mol Biol.* 2001; **10**: 69–76.

Lyle EM, Fujita T, Conner NW, Connolly TM, Vlasuk GP, Lych JL Jr. Effect of inhibitors of factor Xa or platelet, adhesion, heparin, and aspirin on platelet deposition in an atherosclerotic rabbit model of angioplasty. *J Pharmacol Toxicol Methods.* 1995; **33**: 53-61

Maina CV, Rigss PD, Grandea AG. III. Slatko BE. Moran LS. Tagliasmonte JA. McREynolds LA. And Guan C. A vector to express and purify foreign proteins in *Escherichia coli* by fusion to, and separation from, maltose binding protein. *Gene*. 1988; **74**: 365-373

Mann KG, Lorand L. Introduction: blood coagulation. *Methods Enzymol.* 1993; **222**: 1-10

Mans BJ. PhD Thesis. Functional perspectives on the evolution of argasid tick salivary gland protein superfamilies. 2002. University of Pretoria. South Africa

Mans BJ, Louw AI, Gaspar ARMD, Neitz AW. Apyrase activity and platelet aggregation inhibitors in the ticks *Ornithodoros savgnyi. Exp Appl Acarol.* 1998; **22**: 353-366

Mans BJ, Louw AI, Neitz AW. Amino acid sequence and structure modeling of savignin, a thrombin inhibitor from the tick, *Ornithodoros savignyi. Insect Biochem Mol Biol.* 2002a; **32**:821-828.

Mans BJ, Louw AI, Neitz AW. Savignygrin, a platelet aggregation inhibitor from the soft tick *Ornithodoros savignyi*, presents the RGD integrin recognition motif on the kunitz-BPTI fold. J Biol Chem. 2002b; **277**: 21371-21378

Markwardt F. Hirudin as an inhibitor of thrombin. *Methods Enzymol.* 1970; **19**: 924-932

Morrisey JH. Silver staining for proteins in polyacrylamide gels: A modified procedure with enhanced uniformed sensitivity. *Anal Biochem.* 1981; **117**: 307-310.

Murray CJ, Lopez AD. Mortality by cause for eight regions of the world; global burden of disease study. *Lancet.* 1997; **349**: 1269-1276

Nagai K, Thøgersen HC. Synthesis and sequence-specific proteolysis of hybrid proteins produced in

Escherichia coli. Methods in Enzymol. 1987; **153**: 461-481

Narayanan AS, Thiagarajan P. Thrombin. *Encyclopedia of Life Science Nature* 2001; 1-7

Narita M, Bu G, Olins GM, Higuchi DA, Herz J, Broze GJ. Schwartz AL. Two receptor systems are involved in the plasma clearance of tissue factor pathway inhibitor *in vivo*. *J Biol Chem.* 1995; **270**: 24800-24804.

New England BioLabs. pMAL™ Protein Fusion and Purification System (Expression and Purification of Proteins from Cloned Genes). Instruction Manual version 5.01

Nienaber J. MSc. Dissertation. Isolation and characterization of a thrombin inhibitor form the salivary glands of the tick, *Ornithodoros savignyi*. 1999. University of Pretoria. South Africa.

Nienaber J, Gaspar ARM, Neitz AWH. Savignin, a potent thrombin inhibitor isolated from the salivary glands of the tick, *Ornithodoros savignyi* (Acari: Argasidae). *Exp Parasitology.* 1999; **93:** 82–91.

Noeske-Jungblunt C, Haendler B, Donner P, Alagon A, Possani L, Schleuning WD. Triabin, a high potent exosite inhibitor of thrombin. *J Biol Chem.* 1995; **270**: 28629-28634

Novagen. pET system manual. TB055 10th edition 2002

Ortel TL, Chong BH. New treatment for heparin-induced thrombocytopenia and deep venous thrombosis. *Semin Hematol.* 1998: **35**, 26-34

Pacher TB, Jürgens G, Guttenberger M. Casein provides protection against proteolytic artifacts. *Anal Biochem.* 2004: **329**;148-150

Parmenter DL, Boothe JG, van Rooijen GJ, Yeung EC, Moloney MM. Production of biologically active hirudin in plant seeds using oleosin partitioning. *Plant Mol Biol.* 1995; **29**:1167-1180.

Pineda AO, Cantwell AM, Bush LA. The thrombin epitope recognizing thrombomodulin is a highly cooperative hot spot in exosite I. *J Biol Chem.* 2002; **277**:32015–32019

Prior PF, Noeske-Jungblut C, Donner P, Sclenuning WD, Huber R, Bode W. Structure of the thrombin complex with triabin, a lipocalin-like exosite binding inhibitor derived from a triamomine bug. *Proc Natl Acad Sci USA.* 1997; **94**: 11845-11850

Promega. T7 sample system technical bulletin, TB293.

Richardson JL, Fuentes-Prior P, Sadler JE, Huber R, Bode W, Characterization of the residues involved in the human α-thrombin haemadin complex: an exosite II-binding inhibitor. *Biochemistry.* 2002; **41**: 2535–2542.

Roche Instruction manual. High Pure Plasmid Isolation Kit. 1999 Version 2

Rose T, Di Cera E. Three-dimensional modeling of thrombin-fibrinogen interaction. *J Biol Chem.* 2002; **277**:18875-80.

Rowe AJ. Techniques in protein and enzyme biochemistry. Part 1. Elsevier, North-Holland 1978; 1-31

Salzet M, Chopin V, Baert J, Matias I, Malecha J. Theromin, a novel leech thrombin inhibitor, *J Biol Chem.* 2000; **275**: 30774–30780.

Schägger H, von Jagow G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem.* 1987; **166**: 368-379.

Sculley MJ, Morrison JF, Cleland WW. Slow-binding inhibition: the general case. *Biochem Biophys Acta.* 1996; **1298**: 78-86

Sheehan JP, Sadler JE. Molecular mapping of the heparin-binding exosite of thrombin. *Proc Natl Acad Sci USA.* 1994; **91**:5518-5522

Stassens P, Bergum PW, Gansemans Y, Jespers L, Laroche Y, Huang S, Marki S, Messens J, Lauwereys M, Cappello M, Hotez P, Lasters I, Vlauk GP. Anticoagulant repertoire of the hookworm *Ancylostoma caninum*. *Proc Natl Acad Sci USA.* 1996; **93**: 2149-2154

Stone SR, Hofsteenge J. Kinetics of the inhibition of thrombin by hirudin. *Biochemistry.* 1986; **25**: 4622-4628

Strube KH, Kröger B, Bialojan S, Otte M, Dodt J. Isolation, sequence analysis, and cloning of haemadin: an anticoagulant peptide from the Indian leech. *J Biol Chem.* 1993; **268**: 8590–8595.

Stubbs MT, Bode W. A player of many parts: The spotlight falls on thrombin's structure. *Thromb Res.* 1992; **69**: 1-58

Stubbs MT, Bode W. The clot thickens: clues provided by thrombin structure. *Trends Biochem Sci.* 1995; **20**:23-28.

Tuszynski G, Gasic TB, Gasic GJ. Isolation and characterization of antistasin. *J Biol Chem.* 1987; **262**: 9718–9723

Urata J, Shojo H, Kaneko Y. Inhibition mechanisms of hematophagous invertebrate compounds acting on the host blood coagulation and platelet aggregation pathways. *Biochimie.* 2003; **85**: 493-500

van de Locht A, Lamba D, Bauer M, Huber R, Friedrich T, Kröger B, Höffken W, Bode W. Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. *EMBO J.* 1995; **14**:5149–5157.

van de Locht A, Stubbs MT, Bode W, Friedrich T, Bollschweiler C, Hoffken W, Huber R. The ornithodorin-thrombin crystal structure, a key to the TAP enigma? *EMBO J.* 1996; **15**:6011-6017.

Valenzuela JG, Francischetti IMB, Ribeiro JMC. Purification, cloning, and synthesis of a novel salivary anti-thrombin from the mosquito *Anopheles albimanus*, *Biochemistry.* 1999; **38**: 11209–11215.

Vlasuk GP. Structural and functional characterization of tick anti-coagulant peptide (TAP): a potent and selective inhibitor of blood coagulation factor Xa. *Thromb Haemost.* 1993; **70**:212–216

Waxman L, Smith DE, Arcuri KE, Vlasuk GP. Tick anti-coagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. *Science.* 1990; **248**: 593–596.

Wearne SJ. Factor Xa cleavage of fusion proteins. Elimination of non-specific cleavage by reversible acylation. *FEBS lett.* 1990; **263**; 23-26

Wei A, Alexander RS, Duke J, Ross H, Rosenfeld SA, Chang CH. Unexpected binding mode of tick anticoagulant peptide complexed to bovin factor Xa. *J Mol Biol.* 1998; **283**: 147-154

Weitz JI, Crowther M. Direct thrombin inhibitor. *Thromb Res.* 2002; **106**:275-284

Whitlon DS, Sadowski JA, Suttie JW. Mechanisms of coumarin action: significance of vitamin K epoxide reductase inhibition. *Biochemistry* 1978; **17**:1371–1377

Wilkinson DL, Harrison RG. Predicting the solubility of recombinant proteins in *E.coli Biotech.* 1991; **9**: 443-448

Xu Y, Carr PD, Guss J.M, Ollis DL. The crystal structure of bikunin from the inter-alpha-inhibitor complex: a serine protease inhibitor with two Kunitz domains. *J Mol Biol*. 1998. **276**, 955–966.