NOVEL ANTICOAGULANTS FROM HEMATOPHAGOUS ANIMALS

IYER JANAKI KRISHNAMOORTHY

(M.Sc. Biotechnology, Madurai Kamaraj University)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES,

NATIONAL UNIVERSITY OF SINGAPORE

2015

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

_____

Iyer Janaki Krishnamoorthy

27th November 2015

ACKNOWLEDGEMENTS

I would like to thank my PhD supervisor, Prof. R. Manjunatha Kini, for his constant guidance and support. He was always available for discussions and encouraged me to independently think, plan and conduct experiments. His guidance and motivation has helped me develop as an independent and confident researcher. Endless discussions with him have helped me solve some of the most difficult problems that I encountered in my project and solve them systematically. It was with him that I realised, that sitting down and planning is the key to success! Thanks Prof.! Thanks for everything!!!

I would like to also thank A/ Prof. Jeroen Kool from the BioMolecular Analysis group, UV University, Amsterdam, who host me as my supervisor for the development of the online screening assay. Jeroen showed immense perseverance and constantly motivated me to troubleshoot the problems during the toughest of times. This assay development would not have been possible without the help and support of Reka Otvos. I also established many other long term friends and collaborators during this guest research program. A big thanks to Marija Mladic, for the discussions about the at-line assay development. I would also like to thank Prof. Wilfried Niessen, for his expert opinion on mass spectrometry that helped me understand many basic concepts.

I am grateful to our collaborators, RnDr. Maria Kazimirova and RnDr. Mirko Slovak, for providing me the salivary gland extracts of ticks as and when they became available. It was always a struggle obtaining sufficient ticks for my research, but both of you always took extra efforts just to make sure my project didn't stop. Another collaborator, Prof. Patricia Nuttall, from University of Oxford taught me a lot about the tiny ticks and helped me understand several basic biology ideas. It was always fun to have you in our lab in Singapore.

I would like to express gratitude to our collaborators, Prof. Jose Ribeiro and Dr. Ivo Francischetti for all the help for the transcriptomic sequencing and bioinformatics work of *Dermacentor reticulatus*. Thanks Ivo for all the additional career guidance that you gave me whenever we met. Thanks to Dr. Lin Qingsong and Teck Kwang from the Protein and Proteomics Centre for all the proteomics and iTRAQ help. Thanks to Dr. Jobichen Chako and Prof. K. Swaminathan for the help in solving the crystal structure. I would also like to thank the others who have being a great help along the way: Dr. Jun Mizuguchi, Dr. Takayuki Imamura, Dr. Chikateru Nozaki, and Dr. Sadaaki Iwanaga from KAKETSUKEN, Japan, for supplying the thrombin used for crystallization.

I would also like to thank A/Prof. Mark Chan from the Cardiovascular Research Instititute, NUS for all the clinical expertise about antithrombotics. Your insights gave us an idea about the type of experiments would help to get the anticoagulant in the market and take our drugs from the bench to bedside.

I am highly grateful to Angelina, the sister that I never had for answering all my research as well as non research related questions. Thanks for being my friend, philosopher and guide. Thank you so much for everything Angie. Things would have been such a mess without you! I have always missed you in my last year after you left.

I would like to thank the members of the variegin team- Cho Yeow, Norrapat and Fathiah for the constant brainstorming about our peptides. I am highly grateful to Cho Yeow who taught me how to solve crystal structures and has always been available to answer all my questions about thrombin inhibitors. Thanks Cho Yeow for all the additional career mentorship.

I would like to thank Bee Ling and Liyuan for always being so helpful in ordering all our consumables and taking good care of the lab. You always ensured we worked in a safe environment. I would also like to thank all the other members of the Protein Science Laboratory- Ryan, Girish, Bhaskar, Sheena, Sindhuja,

Chen Wan, Ritu, Bidhan, Summer, Varuna and all other members of the Structural Biology Laboratories. The lab was so much fun with you guys.

I would like to thank MOE for awarding me with a generous fellowship and for the grants that made infrastructure available for my projects. I would also like to thank the DBS staff members- Priscilla and Reena for patiently answering all my questions.

A special thanks to my family, thanks mom and dad for everything you have done for me. I love you and I miss you a lot. Thanks to my little brother for being so wonderful and taking care of parents when I have always been away.

The most special thanks to beloved Rahul, for all the help and support. This PhD has not been an easy ride and you have done your best to keep me motivated throughout my PhD. Thank you for being so understanding during all the difficult times, and persuading and pushing me so that I could complete my PhD on time. Thanks for being my best friend and giving me the most constructive critique.

TABLE OF CONTENTS

Summary

Cardiovascular diseases are the leading cause of death globally account for ~ 7.5 million deaths every year. Anticoagulants are a preferred choice of therapeutics for the prevention and control for a number of cardiovascular diseases. Because of the several limitations of the currently available anticoagulants, novel and superior anticoagulants with greater benefits are being sought.

Tick are haematophagous arthropods that rely exclusively on the host blood for their survival. In order to overcome the host defense response and to obtain a bloodmeal, ticks infuse a potent salivary cocktail at the site of feeding. This cocktail is a mixture of vasodilators, antiplatelets, anticoagulants, and immunomodulaors, that together disarm the host defense mechanisms. We have adapted different strategies for the identificaton of novel anticoagulants from the salivary glands of ticks. We have developed a a high throughput platform for the identification of molecules that target thrombin and FXa, the two major enzymes of the blood coagulation cascade. In this on-line post-separation bioassay complex mixtures can be separated, individual molecules can be separated, identified for their functional activity and their exact masses can be determined in tandem. Briefly, in this approach, a nano-HPLC is coupled to a microfluidic chip based bioassay system and a mass spectrometer. The instrumentation is designed in such a way that eluate from the nano-HPLC is split into two equal parts wherein one part is fed into a mass spectrometer (which identifies exact mass) and another part is fed into a microfluidic chip where the enzyme assay takes place (which

determines the functional activity of the molecule). This nano-HPLC coupled to a microfluidic bioassay system operates at submicrolitre flow rates, consumes very little starting material and can successfully potent functionally active compounds from mixtures that are available in limited amounts.

Hard ticks are long term feeders that remain attached to the host and feed on the host blood for periods as long as 9-12 days. Although both male and female ticks feed on blood, stark differences in the feeding behaviours between male and female ticks of the same species are observed. For example, the female ticks feed and increase in size by about 100 times of the unfed body weight, while the male ticks grow barely about 2 times their body weight after feeding.

Tick feeding triggers new protein synthesis in the saliva, and these newly synthesized proteins are the players which disarm the host defense mechanisms enabling prolonged tick feeding. In order to identify sex specific and feeding stage specific proteins, we have carried out quantitative transriptomic (Illumina) and proteomic (iTRAQ) profiling of the tick salivary gland extract of an ixodid tick, *Dermacentor reticulatus*. We have identified more than 30,000 transcripts in the transcriptome, and over 400 proteins in the proteome of the male and female ticks. We have proven that feeding stage specific expression takes place in tick saliva, and the alteration in the levels of these proteins is what mediates prolonged tick feeding. These proteins are mainly anticoagulants, antiplatelets and immune suppressors and represent a library of novel families of proteins with immense diversity, and could be developed further as potent therapeutics.

DTIs and direct FXa inhibitors have been the most preferred anticoagulants for the prevention and control of cardiovascular disorders. Despite being the most sought after options, these anticoagulants are fraught with limitations and do not present a good safety-efficacy balance. In our quest for better and safer anticoagulants, we have identified and characterized a novel thrombin inhibitor-avathrin from the salivary glands *Ambylomma variegatum*. We have solved the 3D-crystal structure of avathrin in complex with thrombin to study detailed structure activity relationships. We have successfully demonstrated the activities of avathrin using *in vitro* assays and *in vivo* animal models. Taking this molecule further, we have identified similar sequences from other tick species and demonstrated the presence of a family of thrombin inhibitors in ixodid ticks. These peptides are short simple sequences with a unique mechanism of inhibiting thrombin active site and exosite with affinities in the picomolar to femtomolar range. We have also conceived and synthesized structure based variants of these peptides with better potencies and different mode of inhibition than the native peptides. We are currently evaluating the safety-efficacy balance and pharmacokinetics-pharmacodynamics of some of these peptides in animal models to prove that these peptides hit the sweet spot and fulfil an unmet need in the current day anticoagulant market.

List of Tables

List of Figures

List of Abbreviations

Single and three letter abbreviations of amino acids were followed as per the IUPAC-IUBMB Joint Commission on Biochemical Nomenclature.

Chemicals and reagents

ACN	Acetonitrile
APS	Ammonium persulfate
BSA	Bovine serum albumin
$CaCl_2$	Calcium chloride
CNBr	Cyanogen bromide
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
FA	Formic acid
Fmoc	9-Fluorenylmethyloxycarbonyl
FII	Factor II
FIIa	Factor IIa
FVIIa	Factor VIIa
FIXa	Factor IXa
FXa	Factor Xa
FXIa	Factor XIa
FXIIa	Factor XIIa
FXIIIa	Factor XIIIa
Gla	Gamma-carboxyglutamic acid
HATU	O-(7-azabenzotriazol-1-yl)-1,1,3,-3-tetramethyluronium hexafluorophosphate

HCII	Heparin cofactor II
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HCl	Hydrochloric acid
LMWH	Low-molecular-weight heparin
NaCl	Sodium chloride
PBS	Phosphate buffered saline
PCI	Percutaneous coronary intervention
PEG	Poly ethylene glycol
S2222	Benzoyl-IIe-Glu (Glu-γ -methoxy)-Gly-Arg-p-nitroanilide (pNA) hydrochloride (HCl)
S2238	H-D-Phe-pipecolyl (Pip)-Arg-pNA•2HCl
S2251	H-D-Val-Leu-Lys-pNA•2HCl
S2288	H-D-IIe-Pro-Arg-pNA•2HCl
S2302	H-D-Pro-Phe-Arg-pNA•2HCl
S2366	PyroGlu-Pro-Arg-pNA•HCl
S2444	PyroGlu-Gly-Arg-pNA•HCl
S2586	Methoxysuccinyl-Arg-Pro-Tyr-pNA•HCl
S2765	Benzyloxycarbonyl-D-Arg-Gly-Arg-pNA•2HCl
TBS	Tris buffered saline
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	N,N,N’,N’-Tetramethylethylenediamine
TF	Tissue factor
TFPI	Tissue factor pathway inhibitor
TFA	Trifluoroacetic acid
tPA	Tissue plasminogen activator

Tris	Tris(hydroxymethyl)-aminomethane
UFH	Unfractionatd heparin

Units

Å	Angstrom
Da	Daltons
h	Hour
kbp	Kilo base-pair
kDa	Kilo daltons
M	Molar
mg	Milli-gram
min	Minute
ml	Milli-litre
mM	Milli-molar
mm	Milli-metre
ng	Nano-gram
nl	Nano-litre
nM	Nano-molar
nm	Nano-metre
µg	Micro-gram
µl	Micro-litre
µM	Micro-molar
µm	Micro-metre
^{0}C	Degree Celsius
%	Percent

Others

ADP	Adenosine diphosphate
BPTI	Bovine pancreatic trypsin inhibitor
CDD	Conserved domain database
CDS	Coding sequence
EST	Expressed sequence tags
GP	Glycoprotein
HIT	Heparin induce thrombocytopenia
IG	Immunoglobulin
iTRAQ	Isobaric tags for relative and absolute quantitation
MS	Mass spectrometry
NOACs	New oral anticoagulants
ORF	Open reading frame
PAR	Protease-activated receptors
PCR	Polymerase Chain Reaction
RMSD	Root mean square deviation
RPKM	Reads per kilobase per million
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TAP	Tick anticoagulant peptide
TE	Transposable elements
TIL	Trypsin inhibitor-like
TTO	Time to occlusion
UV	Ultraviolet
vWF	von Willebrand factor

CHAPTER 1

Introduction

1.1. Haemostasis

Blood circulates and transports nutrients, oxygen and other important nutrients to, and removes waste products from all parts of the body. In vertebrates, the blood circulatory system is a closed and a high-pressured system. Haemostasis enables organisms not only to maintain blood in a fluid state, but also seals any breach caused due to vascular injury and removes blood clots following restoration of vascular integrity (Versteeg et al., 2013). This crucial haemostatic system which is highly conserved from zebrafish to human involves a complex interplay of vasoconstriction, platelet aggregation, blood coagulation and fibrinolysis (Gonias and Pizzo, 1986; Jagadeeswaran, 2005). These systems involve specific proteinases, protein cofactors and highly specialized cell surfaces of platelets and endothelium.

1.1.1. Vasoconstriction

Vasoconstriction is an early response to a vascular injury and this response of the haemostatic system reduces blood flow through affected blood vessels and limits extravasation of the vascular components. The vascular endothelium, which, under normal conditions sustains blood vessels in a vasodilatory state by producing vasodilators, is also responsible for vasocontriction, in the event of a vascular injury (Wakefield et al., 2008). Any physical (trauma) or functional (sepsis) disturbance to the endothelium triggers the production and release of vasoconstrictors like endothelin, ADP, serotonin and thromboxane which act on the vascular smooth muscle cells to induce a localized vasoconstriction. In addition, the local concentration of vasodilators like nitric oxide, adenosine and

prostacyclins is reduced (Heindl and Kupatt, 2000). Vasoconstriction is usually followed by prothrombotic events which begin with the adherence of platelets to the subendothelial tissue (Gonias and Pizzo, 1986).

1.1.2. Platelet aggregation

Platelet activation and aggregation is an important part of the haemostatic system. In the event of a vascular injury, the platelets within circulation are exposed to the components of the subendothelial extracellular matrix like collagen, von Willebrand factor (vWF) and fibronectin, and this triggers the adherence of platelets to the site of injury (Kroll et al., 1991). Collagen binds to the glycoprotein (GP) VI and GPIa receptors on platelets, while the multivalent vWF bridges components of the subendothelium like collagen to the GPIb/V/IX receptor on the platelet surface (Clemetson, 1999; Jennings, 2009). This initial adhesion causes the platelets to roll, adhere and spread on the collagen matrix leading to the formation of an activated platelet monolayer at the site of the vascular injury. This is also termed as the collagen pathway of platelet activation. In addition, the activated endothelial cells resulting from a disrupted endothelium express elevated levels of cell surface adhesion molecules such as P-selectin and E-selectin promoting an increased adhesion of platelets to the site of injury (Roche et al., 1993; Wakefield et al., 2008). The initial platelet adhesion events bring resting platelets which are not normally in contact with each other, to develop local contacts between themselves. These cross-talks between adhered platelets trigger several signaling pathways and lead to the production and release of molecules that amplify platelet aggregation.

In the tissue factor-dependent pathway, platelet activation is mainly initiated by thrombin cleavage of protease activated receptors (PARs) on the surface of platelets.

Thrombin is generated from the tissue factor pathway through the classical blood coagulation cascade (described in 1.1.3.2). Thrombin also binds GPIb, and this binding strengthens its interaction with PARs. Both pathways finally result in platelet activation and aggregation at the site of injury, but which of the two pathways predominate to bring about the initiation of platelet aggregation depends on the injury. Activated platelets release molecules like ADP, serotonin, thromboxane, calcium, and various procoagulant molecules, stored within granules inside the platelets into the surrounding plasma (Reed, 2004). These bound platelet secretion products along with local prothrombotic factors (like tissue factor) lead to the generation of a platelet plug (Brass, 2003). Binding of fibrinogen, a plasma protein involved in the blood-coagulation cascade to the $\alpha_{IIb}\beta_3$ on the platelet surface further crosslinks activated platelets, strengthening the platelet plug (Mans and Neitz, 2004).

1.1.3. Blood coagulation

Blood coagulation is an important part of haemostasis in which circulating zymogens are activated by limited proteolysis in a sequential manner and these events result in the formation of a fibrin clot. Formation of the fibrin clot occurs concurrent with the platelet plug formation and both events together lead to the generation of a stable haemostatic clot (Furie and Furie, 2007). Blood coagulation which is also triggered in response to rupture of the endothelium comprises of

three main phases, initiation, amplification and propagation (Dahlbäck, 2000; Monroe and Hoffman, 2006).

1.1.3.1. Initiation

The initiation phase, also known as the extrinsic pathway of blood coagulation, is triggered on the surface of extravascular cells by the exposure of tissue factor (TF) to blood. Extravascular subendothelial cells like smooth muscle cells and fibroblasts constitutively express TF on their membranes. A rupture to the vascular endothelial cells exposes components of the bloodstream to the TF. TF binds both zymogen and activated form of the coagulation factor VII (FVII and FVIIa). A fraction of activated FVII (FVIIa) circulates in the blood and binding of this form to TF, forms the TF-FVIIa complex (the extrinsic tenase complex) which can activate blood coagulation FIX and FX to form FIXa and FXa, respectively (Dahlbäck, 2000). FIXa along with the cofactor FVIIIa, activates FX to FXa from FX. This FXa associates with cofactor FVa to form the prothrombinase complex on the surface of TF expressing cells, which then activates prothrombin to the active enzyme thrombin. The locally generated thrombin carries out further reactions leading to the amplification phase of blood coagulation (Maynard et al., 1975).

1.1.3.2. Amplification

Most of the thrombin generation takes place in the amplification phase, after the initial fibrin clot is generated.

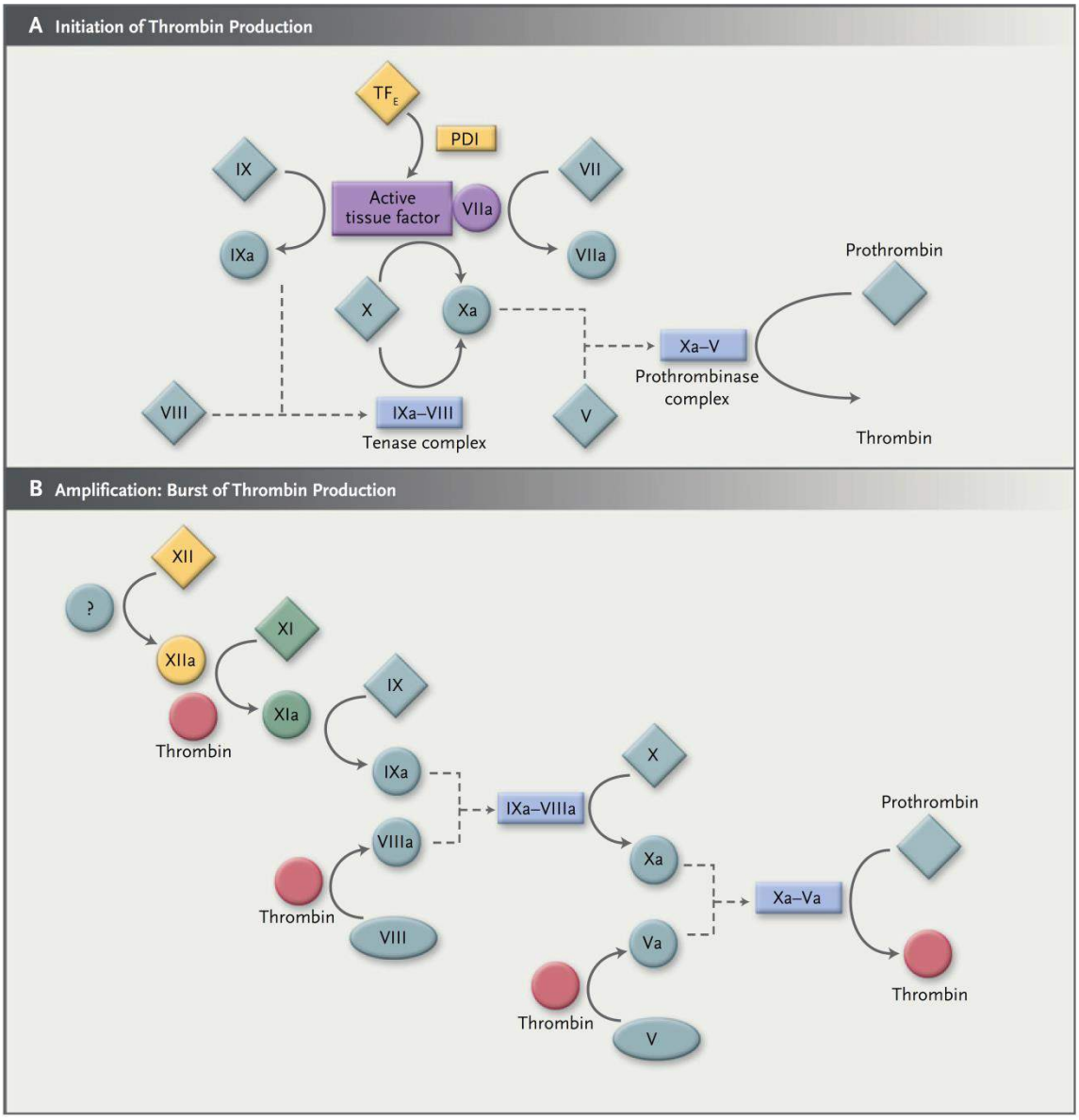


Figure 1.1. Blood coagulation cascade – initiation and amplification phase. FVIIa and tissue factor, form the extrinsic tenase complex and initiate the cascade. This complex activates FX to form FXa, which associates with FVa to produce thrombin. Thrombin then leads to the production of fibrin monomers. Generation of small amount of thrombin also results in a feedback which leads to a burst of thrombin in the amplification phase (Adapted from (Furie and Furie, 2008))

The small amount of thrombin generated on the surface of TF-expressing cells plays multiple roles which result in the amplification of its own production. A major procoagulant function of thrombin in the amplification phase is the cleavage and activation of the two cofactors, FV and FVIII to their activated forms, FVa and FVIIIa, respectively, both of which result in an amplified prothrombinase activity. Thrombin also activates FXI on the platelet surface to form FXIa. Alternatively, FXIa is also formed following the activation of FXII to FXIIa on a negatively charged surface by a process called contact activation (Gailani and Renné, 2007; Renné and Gailani, 2007). FXIa formed by these reactions converts FIX to FIXa, which together with the cofactor FVIIIa forms the intrinsic tenase complex, which subsequently activates FX to FXa, hence amplifying the thrombin generation in the next phase, the propagation phase (described in 1.1.3.3). This set of reactions that constitute the contact pathway, is also called the 'intrinsic pathway', because this property is intrinsic to the components of the plasma, where coagulation is initiated by constituents within the vascular system independent of the extravascular tissue factor.

Another crucial prothrombotic role that thrombin plays in the amplification phase is platelet activation by cleavage of the PARs, which are members of the superfamily, G-protein coupled receptors. Thrombin cleaves PAR1, PAR3 and PAR4, which are expressed on the platelet cell surface, at their extracellular N-termini and this cleavage exposes a new N-terminal domain (Di Cera, 2008). The peptide cleaved from the N-terminal of the PARs binds intramolecularly to the newly exposed N-terminal domain thereby signaling the

further downstream events of platelet aggregation (Jennings, 2009). By the end of the amplification phase, the stage is set for a full fledged propagation phase in which a burst of thrombin generation takes place.

1.1.3.3. Propagation

This final step of blood coagulation takes place on the surface of activated platelets which contains procoagulant phospholipids. FIXa produced by the intrinsic pathway quickly associates with the platelet surface bound FVIIIa, activating FX to FXa, which in turn associates with the platelet surface bound FVa, forming the prothrombinase complex and produces large amounts of thrombin required for the conversion of soluble fibrinogen to insoluble fibrin, which is crosslinked to form a fibrin mesh. Thrombin is also responsible for the activation of the transglutaminase, FXIII to FXIIIa, which covalently cross links the fibrin clot (Versteeg et al., 2013). In addition, thrombin activates thrombin activatable fibrinolysis inhibitor (TAFI), a carboxypeptidase that protects fibrin from the fibrinolytic attack (Dahlbäck, 2000; Monroe and Hoffman, 2006).

The increase in the amounts of phosphatidylserine on the outer leaflet of the activated platelets, makes the platelet cell surface a very highly specialized platform on which the tenase and prothrombinase complexes can be assembled, hence making the platelet surface an exclusive location for the propagation phase (Monroe et al., 2002). The phosphatidylserine on the platelet surface is important because it binds to the γ-carboxyglutamic acid of the coagulation proteins in a calcium dependent fashion, hence tethering the coagulation factors in close vicinity of each other to successfully assemble the tenase and prothrombinase

complexes (Lentz, 2003). Therefore, this stepwise activation of zymogens of the blood coagulation cascade culminates in the generation of sufficient amounts of fibrin which is crosslinked to efficiently seal the breached vascular barrier.

1.1.4. Fibrinolysis

The fibrinolytic system is responsible for clearing clots from the sites of injury following the regeneration and repair of damaged vascular structures. Plasmin is the most important enzyme of the fibrinolytic system, which degrades fibrin into soluble products and dissolves the fibrin clot. The zymogen - plasminogen is converted to the active enzyme, plasmin by the action of two enzymes, tissue type plasminogen activator (t-PA), produced by the endothelial cells, or the trypsin-like serine protease - urokinase (Draxler and Medcalf, 2014; Okafor and Gorog, 2015).

1.1.5. Regulation of coagulation by anticoagulant pathways

Haemostasis, if allowed to proceed in an uncontrolled manner can develop into pathologic thrombosis. Therefore blood coagulation is regulated at different stages by inhibition of coagulation factors or modulation of activity of the cofactors. TFPI almost always inhibits the interaction between TF and the small amounts of circulating FVIIa. Most coagulation factors in plasma including thrombin, FIXa, FXa and FXIa are inhibited by the serine protease inhibitor,- antithrombin (Lane et al., 2005). Circulating coagulation factors are almost instantaneously inhibited by antithrombin, while those coagulation factors that are produced and assembled on the procogulant cell surfaces such as components of the tenase or prothrombinase complexes are less accessible to antithrombin, and

hence remain active. Antithrombin itself is an inefficient serine protease inhibitor, but heparin, which is present on the cell surfaces of endothelial cells enhances antithrombin's inhibitory activity (Dahlbäck, 2000). This limits blood coagulation to occur only at sites of vascular injury.

The central enzyme thrombin plays paradoxical roles as both a procoagulant and an anticoagulant factor under different conditions. In contrast to its procoagulant roles (described in 1.1.3), thrombin when bound to thrombomodulin activates Protein C. The anticoagulant- Activated protein C, in presence of protein S proteolytically inactivates the cofactors, FVIIIa and FVa thus down regulating the activities of the tenase and prothrombinase complexes (Sadler et al., 1993). Thus, although at sites of vascular disruption, the procoagulant activities of thrombin are well pronounced, in an intact vascular system, thrombin acts as an anticoagulant preventing the development of unwanted clots.

1.2. Thrombosis

Subtle changes in the activities of the components of the delicate haemostatic balance can lead to life threatening complications. While the physiological formation of clots (thrombus) is critical for haemostasis and preservation of blood volume; abnormal thrombosis is related to pathologic conditions in which occlusive unwanted clots develop in circulation (Borissoff et al., 2009). Unwanted clots which can develop due to different reasons in the circulation, and lead to the development of dreaded disorders. Cardiovascular disorders are the single largest killer worldwide and account for about 25% of the deaths worldwide claiming more than 7.5 million lives each year and is estimated to reach about 25 million by 2030 (Chaudhari et al., 2014).

Typically, arterial thrombi are described to develop as a result of underlying plaques (atherosclerosis or vasculitis), a dysfunctional endothelium, and high shear stress; where platelet activation plays a crucial role forming the platelet rich 'white thrombi' which are characteristic of an arterial thrombosis event (Badimon and Vilahur, 2007). Occlusion of the coronary arteries leads to coronary syndromes or myocardial infarction, while obstructive thrombi that develop in cerebral arteries lead to thrombotic stroke, and occlusion of peripheral arteries results in peripheral arterial disease and gangrene (Kottke-marchant, 2010).

In contrast, venous thrombi are typically reported to develop at sites where the vein wall is undamaged, and shear stress is low, and thrombus formation occurs mainly due to the congenital dysfunction of coagulation proteins or

physiologic anticoagulants, giving rise to the red cell rich 'red thrombi' (Rosendaal, 1999). In the pathologic condition known as deep vein thrombosis (DVT), clots most often develop in the deep veins of the leg or pelvis. When these clots or small parts of the clot break off from the site of formation, and travel through the circulatory system, an embolism occurs. When this clot lodges in the lung, the condition then known as pulmonary embolism (PE), is a life-threatening disease (Lijfering et al., 2011).

Due to these differences in the underlying mechanisms of development of arterial and venous thrombosis different antithrombotic agents are used for different indications. Antiplatelets, and to a lesser extent anticoagulants, are used for the treatment of arterial thrombosis, whereas, anticoagulants are preferred for the control of venous thrombi (Wu and Matijevic-Aleksic, 2005).

1.3 Current antithrombotics

Antiplatelets, anticoagulants and fibrinolytics are the three major classes of currently available antithrombotic drugs in the market.

1.3.1. Antiplatelets

The interesting property of platelets to instantaneously adhere at sites of vascular injury implicates their crucial role in the development of highly cohesive, platelet rich arterial thrombi. This important role of platelets in arterial thrombosis has been extensively studied for indications such as unstable angina, as well as mycardial infarction and stroke (Vyasa et al., 2011). Platelet aggregation inhibitors prevent blood clots by blocking one of the receptors on the platelet surface. The currently available antiplatelet therapies include aspirin, clopidogrel, ticlopidine and dipyridamole. Out of these options, aspirin which is relatively safe and inexpensive has remained as the mainstay for treatment of several types of arterial thrombotic indications (Phillips et al., 2005). Due to the limitations of antiplatelet monotherapy, a combination therapy of antithrombotics (e.g. aspirin-plus-clopidogrel or aspirin-plus-anticoagulant) has been preferred for the management of certain indications (Curiale et al., 2011; Hong, 2014).

1.3.2. Fibrinolytics

Fibrinolytic agents allow reperfusion by converting plasminogen to plasmin, which can then dissolve the fibrin clot. Tissue plasminogen activator (tPA), streptokinase and urokinase are the three classes of fibrinolytic agents which, by different mechanisms, convert plasminogen to plasmin, thereby activating the fibrinolytic pathway. Since these type of drugs thrombolyse already

formed clots, they are most effective when they are administered within 4-6 hours following a cardiovascular event (Chaudhari et al., 2014; Piccolo et al., 2015).

1.3.3. Anticoagulants

Anticoagulants are used as both short- and long-term options for the management of arterial as well as venous thrombi. Heparins, Vitamin K-antagonists, direct thrombin inhibitors (DTIs) and direct FXa inhibitors comprise the four major classes of anticoagulants currently available in the market. Anticoagulants are the main focus of this thesis and are discussed in detail in this section.

1.3.3.1 Heparin

Heparin, the polysaccharide based anticoagulant, has been administered in several forms, for different indications such as haemodialysis, DVT, renal impairment, pulmonary embolism, venous thromboembolism (VTE), and angina blood vessel complications. There are three different forms of heparin: unfractionated heparin (UFH, a heterogeneous mixture of polysaccharide chains of molecular sizes in the range 3 to 50 kDa), low molecular weight heparin (LWMH, polysaccharides of the molecular size ~ 6 kDa) and ultra low molecular weight heparin (ULMWH, polysaccharides of molecular size < 2 kDa). Out of the three forms, UFH which is the most widely used form of heparin is a century old drug, and has remained as the preferred anticoagulant despite its various limitations (Linhardt and Liu, 2012).

The molecular mechanism of the anticoagulant action of heparin lies in its ability to bind to and enhance the inhibitory effect of the plasma serine protease inhibitor- antithrombin. Heparin bound antithrombin can effectively inhibit the enzymatic actions of thrombin and FXa, the two most crucial serine proteases of the blood coagulation cascade. In addition, heparin can also activate other serpins like heparin cofactor-II (HCII) and protein C inhibitor (PCI), which through different mechanisms bring about the anticoagulant action of heparin (Hoffman et al, 2012).

UFH shows a number of limitations, such as binding to other plasma proteins, which gives rise to variable pharmacokinetics, and it induces an immune response called heparin-induced thrombocytopenia (HIT). These limitations of UFH have triggered the introduction of LWMH products as agents for the management of different indications. LWMH also have the advantage of being subcutaneously bioavailable, and have a longer half-life than UFH. Despite these advantages, the lack of a readily available reversal agent for LWMH increases severe bleeding risks, and hence restricts the use of LWMH in clinical settings (Linhardt and Liu, 2012).

1.3.3.2. Vitamin K antagonists

Vitamin K antagonists (VKAs), of which warfarin is the most common example, exert their anticoagulant activity in an indirect manner. VKAs competitively inhibit vitamin K epoxide reductase, the enzyme which brings about the γ-carboxylation of prothrombin, FVII, FIX, FX (procoagulants), protein C and protein S (anticoagulants), impairing their activity. Despite being popular,

this class of anticoagulants suffers from serious side effects like food drug interactions, narrow therapeutic window, highly variable therapeutic dosages, and requires close monitoring of drug levels in the plasma (Marder et al., 2004). All these side effects increase the healthcare costs and require rigorous patient compliance, limiting the use of VKAs as anticoagulant drugs.

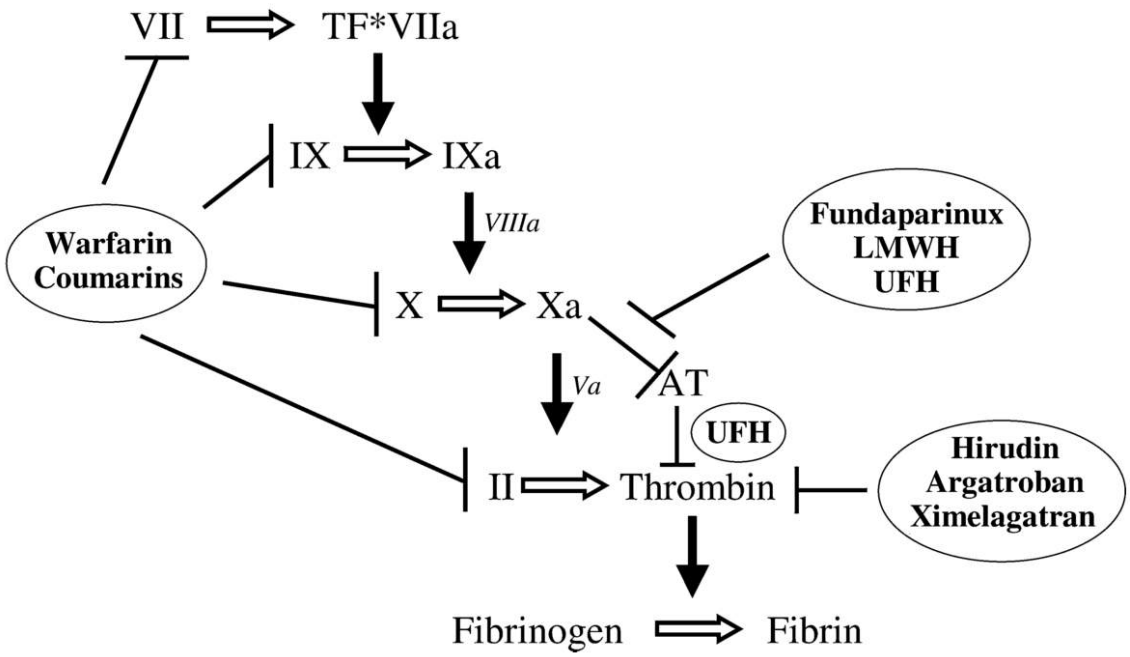


Figure 1.2. Targets of current anticoagulants. Warfarin and coumarins (vitamin K antagonists) inhibit the synthesis of blood coagulation zymogens (factors II, VII, IX and X). Antithrombin inhibits thrombin and FXa. LMWH, fondaparinux and UFH inhibit thrombin and FXa in an antithrombin dependent manner. Hirudin, argatroban and ximelagatran are DTIs. (Adapted from (Wu and Matijevic-Aleksic, 2005))

1.3.3.3. Direct thrombin inhibitors (DTIs) and direct FXa inhibitors

Designed to overcome the limitations of VKAs, DTIs and direct FXa inhibitors are becoming the most popular choice of anticoagulants for the prevention and treatment of venous and arterial thromboembolism. Unlike

warfarin, these agents directly and specifically bind to and inhibit thrombin and FXa - the most important serine proteases of the blood coagulation cascade. DTIs and direct FXa inhibitors have been studied extensively and have gained increasing popularity (Yeh et al., 2015). DTIs such as hirudin, bivalirudin, dabigatran and argatroban, and direct FXa inhibitors, such as rivaroxaban, apixaban and endoxaban are currently available in the market. These agents have a rapid onset of action and some of these direct acting inhibitors have excellent oral bioavailability (and hence are named as non-vitamin K antagonists oral anticoagulants (NOACs)).

Hirudin from the medicinal leech *Hirudo medicinalis* is a DTI. This 65-residue anticoagulant has been produced recombinantly and has been used for treatment in patients with HIT and for thrombosis prophylaxis after major orthopedic surgery (Greinacher et al., 1999). However, the clinical use of hirudin has been limited due to several drawbacks such as (a) high risk of bleeding, (b) pharmacokinetics that depend on renal function, (c) lack of antidote and (e) immunogenicity (Lee and Ansell, 2011).

Bivalirudin is a 20-residue synthetic oligopeptide, which has been developed by linking the exosite-I binding C-terminus of hirudin with the active site binding moiety, $_{D}$-Phe-Pro-Arg-Pro (Warkentin, 2004). The C-terminus of bivalirudin is based on the exosite-I binding sequence of hirudin. The N-terminal active site binding moiety is linked to the exosite-I binding C-terminus with 4 glycine residues as a linker. Bivalirudin selectively and reversibly binds to the active site of free and fibrin bound thrombin, has negligible immunogenicity,

significantly lower bleeding risks than hirudin, and is eliminated via proteolytic cleavage and renal routes. It has been used in the management of percutaneous coronary intervention (PCI) and is gaining increased attention for numerous clinical applications like acute coronary syndrome (ACS), myocardial infarction (MI), and venous HIT thrombotic events (Mavrakanas and Chatzizisis, 2015).

The small molecule DTIs and direct FXa inhibitors like argatroban, dabigatran, apixaban, endoxaban and rivaroxaban are the NOACs, which are used for several conditions such as ACS, atrial fibrillation, VTE, stroke, thromboprophylaxis after knee or hip arthroplasty. Because these NOACs directly bind to a single enzyme, they have well documented mechanism of action, and predicted outcomes and have shown to be safer than VKAs (Saito et al., 2015; Yeh et al., 2015). Despite all this progress, there are well documented side-effects of bleeding complications, as well as liver toxicities resulting from the use of NOACs (Santarpia et al., 2015).

Despite the presence of numerous antithrombotic options, these drugs are fraught with several limitations and do not present an ideal safety-efficacy balance. Therefore, in our quest for novel antithrombotics with superior benefits, we have identified and developed a novel family of anticoagulants from haematophagous animals and are developing them further in an attempt to fulfill an unmet need in the current day antithrombotic market.

1.4. Haematophagous animals

Different groups of animals in the arthropod, annelid and nematode taxa have evolved the ability to feed on vertebrate blood for their survival. Mosquitoes, ticks, bed bugs, leeches, tsetse flies and triatomines are a few examples of haematophagous animals. During feeding, these haematophagous animals puncture through the skin, lacerate the blood vessel and suck the host blood to obtain nutrients. This causes an injury to the vascular system and triggers the host defense mechanism which involves the haemostatic and the immune system. The host haemostatic system ensures minimal blood loss following a vascular injury, by sealing the breach to stop the blood flow; the host immune system causes pain and inflammation leading to a grooming response and aids in awareness in the host and subsequent removal of the parasite. To counter these defense mechanisms and to ensure a successful feeding, haematophagous animals produce an assortment of molecules in their saliva which they infuse into the host blood (Champagne, 2006; Mans, 2011; Mans and Neitz, 2004).

Evolutionary studies have shown that haematophagy (blood feeding behavior) has evolved at least 20 independent times in 15,000 species distributed over 400 genera of arthropods (Francischetti et al., 2009). At each of these stages of evolution, novel scaffolds, with new strategies, novel structures and mechanisms to overcome the host defense response have been evolved. In most cases, the saliva contains a complex mixture of components with vasodilatory, antiplatelet, anticoagulant and anti-inflammatory activities to disarm the host haemostatic and immune systems. Therefore, the saliva of these haematophagous

animals represent a rich pharmacoepia presenting a diverse set of molecules that target the host haemostatic and immune systems which can be developed into therapeutic agents for clinical use.

Since the sizes of these blood-feeding animals are significantly smaller than their vertebrate hosts, their mouthparts can be considered as tiny micro syringes that infuse minute amounts of saliva in the host blood, yet bring about the desired antihaemostatic response and help them obtain a successful blood meal. This indicates that the molecules isolated from the saliva of these animals would present as exceptionally potent antihaemostatics that could easily supersede the current day antithrombotics available in the market. Few examples of vasodilators, antiplatelets and anticoagulants studied from the saliva of haematophagous animals are discussed in detail below.

1.4.1. Vasodilators

The main role of vasodilators in the saliva of haematophagous arthropods injected into the skin is to help the arthropod enlarge the vessel and subsequently enhance the blood flow to the feeding site. Maxadilan, a 60-residue peptide, isolated from the saliva of the sand fly, *Lutzomyia longipalpis* (Lerner et al., 1991) and *Simulium vittatum* erythema protein (SVEP) from the saliva of the black fly, *Simulium vittatum* (Cupp et al., 1998) are the most potent known vasodilators, which bring about an erythema when injected at levels as low as 0.1 ng peptide per rabbit. The vasodilatory properties of maxadilan is mediated by alterations in the levels of intracellular cAMP in smooth muscle cells by binding to the pituitary adenylate cyclase-activating polypeptide type I receptor (PAC1)

(Lerner et al., 1991). Nitrovasodilators from *Rhodnius prolixus*, which were later identified as nitric oxide (NO) binding haeme proteins, have the lipocalin fold and were subsequently named as nitrophorins (Ribeiro et al., 1990a). These nitrophorins have been shown to bind and release NO in a pH dependent manner. In the case of hard ticks prostaglandins including prostacyclins, and prostaglandin E_2 (PGE_2) accomplish vasodilation of the host blood vessel (Champagne, 2006). Several other vasodilators from different classes of haematophagous arthropod saliva have been well studied, yet there remains immense potential to identify novel vasodilators from the largely unstudied groups of haematophagous animals.

1.4.2 Antiplatelets

Antiplatelets which inhibit platelet activation and aggregation are extensively produced in a vast diversity in the saliva of haematophagous animals. These antiplatelets primarily inhibit platelet activation or aggregation by blocking the activation of one of receptors on the platelet surface. Apyrases which degrade ADP, a crucial activator of platelet aggregation, have been isolated from the saliva of mosquitoes, bugs, sand flies, fleas, triatomines and ticks (Ribeiro and Francischetti, 2003). The lipocalin moubatin, isolated from the saliva of *Ornithodorus moubata* inhibits collagen-induced platelet aggregation by interfering with a pathway associated with a thromboxane A_2 (TXA_2) receptor (Francischetti, 2010). Other examples of platelet aggregation inhibitors include 'tick adhesion inhibitor' which inhibits integrin $\alpha_2\beta_1$-collagen adhesion and disagregin which inhibits integrin $\alpha II_b\beta_3$-fibrinogen mediated aggregation (Karczewski et al., 1994). The amine binding protein of the lipocalin superfamily

from the saliva of *Rodnius prolixus* inhibit serotonin and epinephrine mediated effects on platelets (Champagne, 2006). Since thrombin is an important activator of platelet aggregation, thrombin inhibitors have also been documented to exert antiplatelet activity (Tanaka et al., 2007).

1.4.3. Anticoagulants

The blood coagulation cascade is disarmed by anticoagulants injected by the haematophgous animal at the site of feeding. Numerous anticoagulants are produced by these blood sucking parasites, and these exogenous anticoagulants present an array of molecules with tremendous diversity. Studies by different groups have reported that in addition to disabling the host blood coagulation cascade at the site of feeding, anticoagulants produced in the saliva of these haematophagous arthropods play a significant role in the parasite gut, where they maintain blood in the fluid state (Bowman et al., 1997; Valenzuela, 2004). Evolutionary studies with soft ticks have revealed that anticoagulants preceded antiplatelets in the ancestral arthropod saliva, suggesting that adaptation to blood coagulation played an important role in successful evolution of haematophagy (Mans et al., 2002).

These exogenous anticoagulants from arthropod saliva are serine protease inhibitors that selectively inhibit specific coagulation factors or inhibit the activity of one of the complexes assembled on the procoagulant surfaces (described in 1.1.3.). Mechanistically, these exogenous anticoagulants are classified as extrinsic tenase complex inhibitors, intrinsic tenase complex inhibitors, contact system inhibitors, FXa inhibitors and thrombin inhibitors (Koh and Kini, 2008). Few

examples of anticoagulants which are present in the saliva of haematophagous animals are discussed in further detail below.

1.4.3.1. Extrinsic tenase complex inhibitors

The extrinsic tenase complex (TF-FVIIa complex) which is assembled on the surface of TF-expressing endothelial cells, is one of the early responses and is triggered during the initiation phase of the blood coagulation cascade (described in 1.1.3.1.). Extrinsic tenase complex inhibitors have evolved in the arthropod saliva to target this important step. Extrinsic tenase complex inhibitors from haematophagous animals have been isolated and characterized and examples from two classes are described below.

Ixolaris, a two domain Kunitz type inhibitor, with both domains homologous to TFPI has been characterized from the saliva of *Ixodes scapularis* (Francischetti et al., 2002). The 120 residue long, two domain Kunitz inhibitor has been shown to interact stoichiometrically with FX and FXa with affinities in the range of 0.5-10 nM, but not with FVIIa. Ixolaris is a fast and tight binding inhibitor that binds to the exosite of FX (FXa) in a Gla domain dependent manner. This binding impairs the interaction of FX or FXa with FVIIa and prothrombin, thereby blocking the assembly of the extrinsic tenase complex and the prothrombinase complex, respectively, on the surface of TF expressing endothelial cells. Penthalaris, an inhibitor with five tandem Kunitz domains also isolated from the saliva of *Ixodes scapularis* inhibits the assembly of the extrinsic tenase and the prothrombinase complex using FX or FXa as a scaffold (Francischetti et al., 2004).

The nematode anticoagulant protein 2 (NAPc2), from hookworm is a potent inhibitor of the TF-FVIIa complex. Similar to ixolaris, NAPc2 exerts its inhibition over the extrinsic tenase complex in a FX- or FXa-dependent fashion. But unlike ixolaris, NAPc2 bound to FX or FXa also binds to the TF-FVIIa complex and forms a tight ternary complex, with an equilibrium inhibitory constant of 10 pM (Lee and Vlasuk, 2003).

1.4.3.2. Intrinsic tenase complex inhibitors

The intrinsic tenase complex (FIXa-FVIIIa complex), which is assembled on the phospholipid membrane of the platelet surface plays an important role in FXa generation, which leads to a subsequent burst in thrombin generation in the amplification phase of blood coagulation (described in 1.1.3.2.). The lipocalin prolixin S, isolated and characterized from the kissing bug *Rhodnius prolixus* is the only identified intrinsic tenase complex inhibitor from haematophagous arthropods. This intrinsic tenase complex inhibitor inhibits the complex by an anti-FVIIIa activity, which makes it a unique anticoagulant inhibiting a blood coagulation cofactor unlike the usual anticoagulants that inhibit the blood coagulation enzymes. The anti-FVIIa activity of prolixin S lies in a short stretch of the molecule and this anti-FVIIa activity is independent of NO binding. This makes prolixin S a unique lipocalin because all the other lipocalins identified from *Rhodnius prolixus* are NO dependent proteins and act as vasodilators (Ribeiro et al., 1995).

1.4.3.3. Contact system protein inhibitors

The contact activation system of blood coagulation, which involves kallikriens, FXIIa, FXIa and other plasma serine proteases (the activities of some of these serine proteases in the blood coagulation is yet to be understood) is also targeted by components of the arthropod saliva.

BmTI-A, from the larvae of *Boophilus microplus* is a Kunitz type inhibitor and it inhibits trypsin, neutrophil elastase and plasma kallikrien. BmTI-A has two Kunitz domains and its sequence shows similarity to ornithodorin (a thrombin inhibitor) (van de Locht et al., 1996) and tick anticoagulant peptide (TAP, a FXa inhibitor) (Waxman et al., 1990), both of which have been identified from *Ornithodorus moubata* (Tanaka et al., 1999).

Triafestin-1 and triafestin-2 are the two kallikrien-kinin system inhibitors from the salivary glands of the kissing bug, *Triatoma infestans*. Both inhibitors are structurally related to lipocalins and inhibit the contact pathway of blood coagulation by inhibiting the reciprocal activation of factor XII and prekallikrein. They interact with factor XII and high molecular weight kininogen in a Zn^{2+}-dependent manner and also inhibit, the ability of factor XII and high molecular weight kininogen binding to negatively charged surfaces. This inhibition not only blocks the activation of FXI by FXIIa, but it also inhibits bradykinin release because of the inhibition of the kallikrien-kinin system. Since bradykinin is involved in numerous inflammatory responses around the injured site, these molecules play dual roles, both as anticoagulants as well as immune suppressors (Isawa et al., 2007).

Ixodes ricinus contact phase inhibitor (Ir-CPI), contains one Kunitz domain and inhibits the contact pathway and to a lesser extent, fibrinolysis *in vitro*. Ir-CPI exerts its inhibition on the contact pathway by inhibiting the reciprocal activation of FXII, prekallikrien and FXI in human plasma (Decrem et al., 2009).

1.4.3.4. Factor Xa inhibitors

FXa is an important enzyme of the blood coagulation cascade, which along with FVa forms the prothrombinase complex. Suppression of FXa activity by intrinsic inhibitors is essential for the physiologic regulation of the blood coagulation cascade, and this is achieved by the plasma inhibitors,- antithrombin and TFPI (Furie and Furie, 2008). The saliva of haematophagous animals have evolved numerous FXa inhibitors to control the activity of this crucial enzyme of the blood coagulation cascade (Gould et al., 2006).

1.4.3.4.1. Kunitz type Factor Xa inhibitors

The tick anticoagulant peptide (TAP, K_i 200 pM), is the most well studied FXa inhibitor from the haematophagous soft tick, *Ornithodorus moubata* (Waxman et al., 1990). TAP is a single domain Kunitz type, slow, tight binding and competitive FXa inhibitor that possesses the conserved cysteines of a Kunitz domain but differs significantly from other Kunitz type inhibitors in the overall amino acid sequence (Corral-Rodríguez et al., 2009; Waxman et al., 1990).

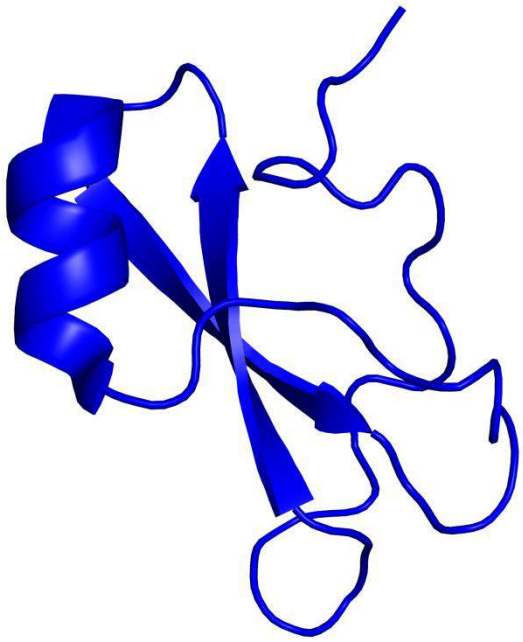


Figure 1.3. Structure of Tick anticoagulant peptide (1TAP). The single domain Kunitz type FXa inhibitor TAP contains one α-helix (residues 51-60) and an antiparallel β-sheet (residues 22-28 and 32-38). The N-terminal residues of the 3_{10} helix (residues 2-7) lock into the active site of FXa.

The three N-terminal residues of TAP bind to FXa active site in a non-canonical manner and exhibit a mechanism that is similar to hirudin inhibition of thrombin (Simone et al., 1998). During the formation of the TAP-FXa complex, an initial slow-binding occurs at the secondary binding site, which induces a rearrangement in the N-terminal residues of TAP to lock into the active site of factor Xa. In FXa, there are five basic residues at the secondary binding site that interact with the acidic residues of TAP. In thrombin and trypsin, the corresponding residues would not provide this kind of a charge interaction with TAP. The three N-terminal residues locked into the active site form extensive hydrogen bonds with the active site. (Dunwiddie et al., 1992).

1.4.3.4.2. Antistatin and ghilantens

Antistatin (ATS), isolated from the Mexican leech, *Haementaria officenalis*, is a slow binding FXa inhibitor (Tuszynski et al., 1987). The affinity

of ATS was 8300- to 12,600-times smaller than TAP, and unlike TAP, ATS was cleaved by FXa at the Arg^{35}-Val^{36} peptide bond (Mao, 1993). ATS contains two tandem antistasin-like domains, and each domain possesses 10 cysteines which form five intra-domain disulphide bonds. Several other FXa inhibitors homologous to antistatin (named ghilantens) have been isolated from the salivary glands of the South American leech *Haementaria ghilianii* (Blankenship et al., 1990).

1.4.3.4.3. Lufaxin

Lufaxin is a type member of a novel family of FXa inhibitors isolated from the saliva of the sand fly *Lutzomyia longipalpis*. The 32.4 kDa protein was identified as a slow, tight, noncompetitive, and reversible inhibitor of FXa. Lufaxin is a highly specific FXa inhibitor that does not interact with any other enzyme of the cascade. In addition to inhibiting the pro-coagulant properties of FXa, lufaxin also inhibits FXa-induced PAR2 activation. Lufaxin injected into the mice tails has been shown to prolong the time to occlusion in carotid artery thrombosis models. Because, the primary structure of lufaxin does not show any similarity to the other FXa inhibitor families, it has been classified as a novel family of FXa inhibitors (Collin et al., 2012).

1.4.3.4.4. Ascaris type FXa inhibitors

Anticoagulant peptide (AcAP) was isolated from the human hookworm, *Ancylostoma caninum* and was shown to inhibit FXa with an affinity of 323.5 pM. AcAP inhibits the active site of FXa and shows a 100- to 200-fold higher specificity towards FXa than to the other serine proteases of the coagulation

cascade (Cappello et al., 1995). Several other ascaris-type FXa-inhibiting proteins similar to AcAP have been reported from *Ancyclostoma caninum* (NAP5/6 and NAPc2/3/4) and from *Ancylostoma ceylanicum* (AcAP5 and AceAP1) (Mieszczanek et al., 2004; Stassens et al., 1996).

1.4.3.4.5. Serpin family of Factor Xa inhibitors

The ubiquitously found serpin family of serine protease inhibitors that inhibit blood coagulation proteases are also identified from the saliva of haematophagous animals and inhibit the key enzyme, FXa to control the host haemostatic system (Khan et al., 2011). Serpins contain 350-400 residues with molecular sizes in the range of 40-55 kDa (Khan et al., 2011). Alboserpin, from the salivary glands of the mosquito *Aedes albopictus* displays stoichiometric, competitive, reversible and tight binding to FXa. Like the other members of the serpin family, alboserpin undergoes huge conformational changes to bind to and inhibit FXa irreversibly (Calvo et al., 2011). Sequences of several other serpins from the salivary gland transcriptomes of other haematophagous animals have been identified and some of these sequences have been hypothesized to inhibit FXa in a manner similar to alboserpin (Karim et al., 2011; Ribeiro et al., 2006; Tan et al., 2015).

1.4.3.5. Thrombin inhibitors

The crucial role of thrombin is reflected by the presence of thrombin inhibitors in the saliva of blood-sucking animals (Huntington, 2014). The unique 3-dimensional organization of thrombin makes it a specialized enzyme restricting the access of most macromolecular substrates and inhibitors and only those which

exhibit specific properties that are able to fit this unique structure are able to enter the active site cleft.

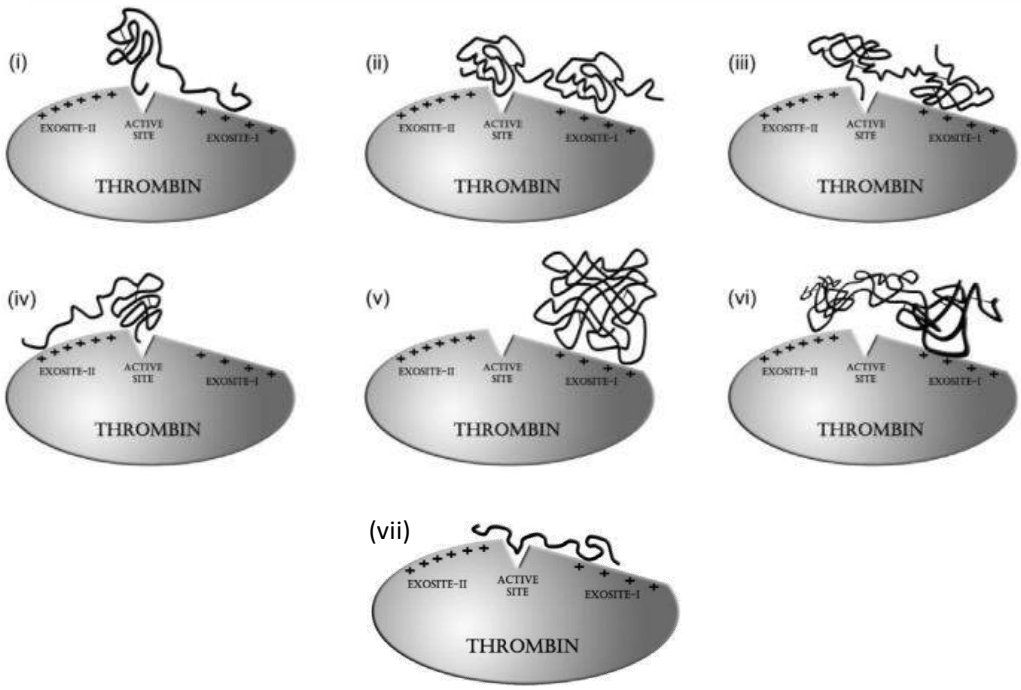


Figure 1.4. Schematic of different classes of thrombin inhibitors and their structural features. (i) Hirudin, N-terminus binds to active site, acidic and extended C-terminal domain binds to exosite-I; (ii) rhodniin: two Kazal-type domains, the N-terminus binds to active site and the C-terminal domain binds to exosite-I; (iii) ornithodorin: two tandem Kunitz-type domains, N-terminus domain binding to active site and the C-terminal domain to exosite-I; (iv) haemadin: N-terminus binds to active site, extended C terminus binds to exosite-II; (v) triabin: single lipocalin domain binds to exosite-I; (vi) bothrojaracin: two chains of the C-type lectin domain bind to exosite-I and exosite-II respectively; (vii) Variegin: N-terminal steers towards exosite-II, middle part of the molecule (MHTK) binds the active site and flexible acidic C-terminus binds to the exosite-I (Adapted and modified from (Koh et al., 2007)).

In addition to the two basic exosites that determine the high substrate and inhibitor specificity, the thrombin active site is placed in a narrow cleft which is lined by two loops (called the 60-loop and the 149-loop) (Polgár, 2005). Many

specialized thrombin inhibitors have been studied from the saliva of haematophagous animals and they fall into three general categories: ‘canonical’; ‘non-canonical (or ‘heretical’); and exosite binding inhibitors. When the inhibitor binds to the enzyme in a substrate-like fashion forming an antiparallel β-strand between the P3-P1 residues of the inhibitor with 214–216 residues of thrombin, the interaction is called 'canonical'. 'Non-canonical' refers to any inhibitor- active site interaction that is not canonical (Krowarsch et al., 2003).

1.4.3.5.1. Hirudin

Hirudin, the anticoagulant protein isolated from the salivary glands of the medicinal leech, *Hirudo medicinalis*, was the first thrombin inhibitor from a haematophagous animal that was used as a clinical antithrombotic agent (Stone and Hofsteenge, 1986). It is a polypeptide composed of 65 amino acids, and tightly and specifically binds to thrombin with a K_i of about 20 fM. It interacts with the thrombin active site as well as the exosite-I, preventing fibrinogen cleavage thus inhibiting the clot formation. Hirudin binds to the active site cleft of thrombin with its three N-terminal residues- Val, Val and Tyr. The amino group at the N-terminus of hirudin forms strong hydrogen bonds with the active-site His57 and with the main chain carbonyl group of Ser214. The P2 position is occupied by Vall and the P3 position is approximately occupied by Tyr3. Thus the hirudin polypeptide binds thrombin in a direction opposite to that expected for fibrinogen and is hence a non-canonical inhibitor. The carboxy tail of hirudin (residues 48-65) wraps around the exosite-I and forms several ionic interactions and salt bridges. This binding to both thrombin active site cleft as well as the

exosite-I gives hirudin its high selectivity (Grütter et al., 1990). Of special importance in the hirudin sequence is the sulfated Tyr64 residue, which plays a significant role in interactions with thrombin. Hirudin desulfated at this position inhibits thrombin with a 10-fold weaker affinity (Stone and Hofsteenge, 1986).

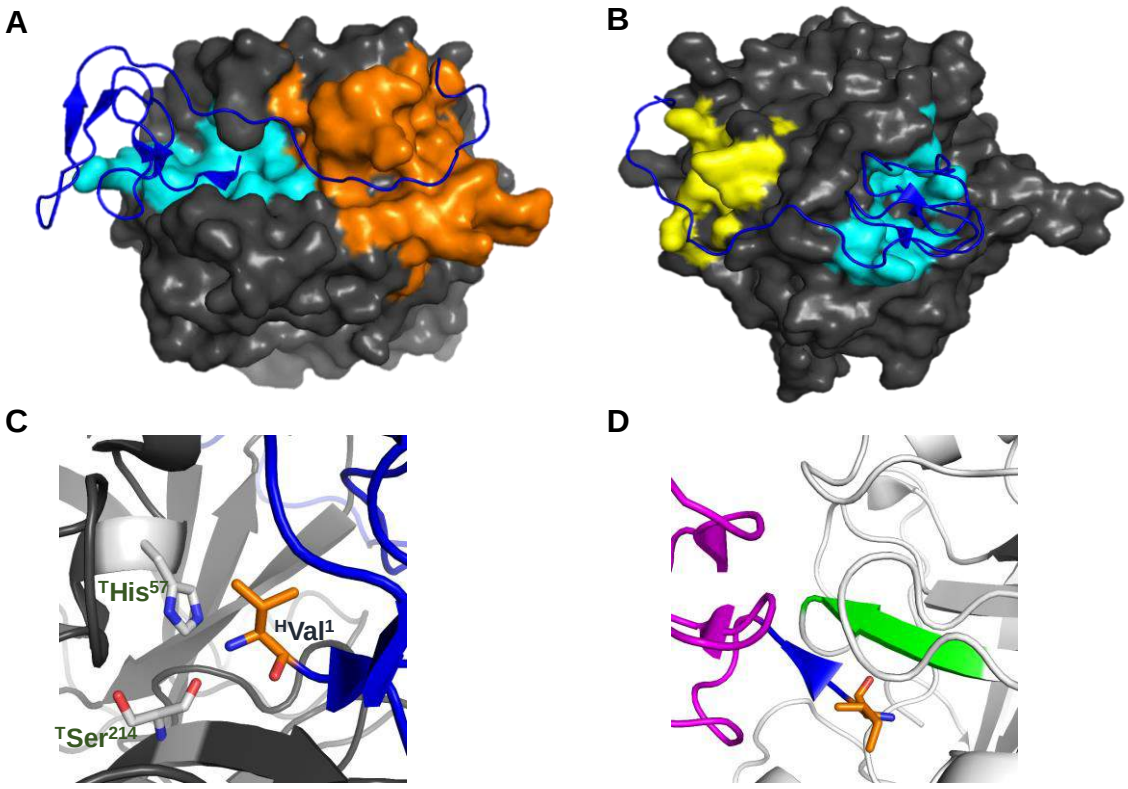


Figure 1.5. Structures of thrombin with inhibitors from leech. A. Hirudin binds thrombin active site with its globular N-terminal domain and exosite-I with its acidic C-terminal domain (4HTC) **B.** Haemadin binds thrombin active site with its globular N-terminal domain and shows overall similarity with hirudin, but unlike hirudin it binds exosite-II of thrombin with its acidic C-terminal domain (1E0F) (For panels A and B, thrombin surface view shown in gray, active site in cyan, exosite-I in orange, exosite-II in yellow; and hirudin and haemadin cartoons in blue. **C.** Main chain nitrogen of hirudin N-terminus fits in the active site cleft of thrombin forming hydrogen bond with main chain oxygen of Ser214 of thrombin. **D.** N-terminal residues of haemadin forming parallel β-sheet with Ser214-Gly216, a characteristic of the non-canonical type inhibitors. Isobutyl group of Ile1 occupies the S2 pocket of thrombin.

1.4.3.5.2. Haemadin

The slow, tight binding inhibitor, haemadin from the leech *Haemadipsa sylvestris* binds thrombin with an affinity of 200 fM (Strube et al., 1993). Haemadin is a 57 residue polypeptide with a globular N-terminal domain that is held compactly together with three disulphides bonds and an acidic C-terminal tail. Despite this similarity in the overall organization of hirudin and haemadin, haemadin binds to the thrombin active site and to the heparin binding exosite (exosite-II) (Richardson et al., 2000). The first three residues of the haemadin N-terminus bind to active site of thrombin in a non-canonical manner identical to hirudin binding. A truncation of the 17 residues of the C-terminus reduces the haemadin affinity towards thrombin by 20,000-fold, indicating that these residues contribute to the tight binding nature of haemadin (Richardson et al., 2002).

1.4.3.5.3. Kunitz type inhibitors

The Kunitz family of serine protease inhibitors is one of the most extensively studied class of thrombin inhibitors reported from ticks, and have been reported to be found in both soft and hard ticks (Corral-Rodríguez et al., 2009). Despite exhibiting a characteristic BPTI-like disulphide bond pattern, members of the Kunitz family of serine protease inhibitors display significant differences in the amino acid sequences and bind the target serine protease with different mechanisms.

Ornithodorin comprises of two tandem Kunitz domains and interacts with the thrombin active site by inserting its N-terminal residues inside the active site cleft in a manner similar to the thrombin-hirudin interactions (van de Locht et al.,

1996). Although ornithodorin contains two reactive site loops (RSLs) in the two Kunitz domains, neither of the two RSLs interacts with the enzyme active site. The two Kunitz domains of ornithodorin are oriented in a tail-to-tail manner and hence, the two RSLs are directed to point away from the enzyme surface. The first six residues of ornithodorin N-terminus are in direct contact with the thrombin active site and $^{O}Ser^{1}$, $^{O}Leu^{l}$, $^{O}Asn^{2}$ and $^{O}Val^{3}$ run towards $^{T}Ser^{l95}$ and form a parallel β-sheet arrangement with $^{T}(Ser^{214}\text{-}Gly^{219})$ of thrombin exhibiting a non-canonical mode of inhibition. The C-terminal domain of ornithodorin shows lesser contacts with thrombin in the form of electrostatic contacts and forms three salt bridges that fix it to the exosite-I of thrombin.

Similar to ornithodorin, boophilin inhibits thrombin in a non-canonical manner (Macedo-Ribeiro et al., 2008). Whereas BPTI, the type-member of Kunitz type inhibitor, binds thrombin active site in a canonical manner and causes an extensive rearrangement of the active site, the active site remains almost unaltered by boophilin binding (Soares et al., 2012). The guanidinum group of $^{B}Agr^{17}$ anchors to the S1 pocket of the enzyme and forms two hydrogen bonds with the carboxyl group of $^{T}Asp^{189}$ at the bottom of the S1 pocket. This feature of boophilin distinguishes it from ornithodorin, which does not possess an Arg at this position. The residues $^{B}Asn^{18}$, $^{B}Gly^{19}$, $^{B}Arg^{22}$, and $^{B}Phe^{39}$ are also involved in extensive interactions with different subsites of thrombin to facilitate boophilin binding (Macedo-Ribeiro et al., 2008). The C-terminal domain of boophilin forms ionic interactions with the exosite-I of thrombin. Ornithodorin and boophilin

differ greatly in the orientation of the two Kunitz domains with respect to their binding on the surface of the enzyme.

1.4.3.5.4. Kazal type inhibitors

Like the Kunitz family of serine protease inhibitors, the members of the Kazal family of inhibitors too have been extensively studied as thrombin inhibitors (Laskowski and Kato, 1980). Rhodniin, the double Kazal domain protein from the assassin bug *Rhodnius prolixus* inhibits thrombin with a K_i of 0.2 pM, with its N-terminal domain binding to thrombin's active site and its C-terminal domain binding to exosite I (van de Locht et al., 1995). A typical Kazal domain contains six cysteines forming three disulfide

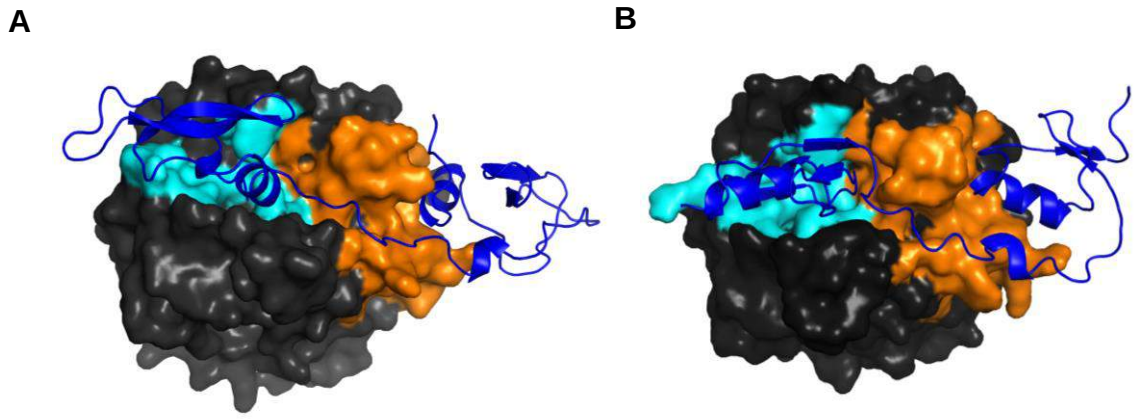


Figure 1.6. Structures of thrombin with kunitz and kazal inhibitors. A. Ornithodorin contains two kunitz type domains and binds to the active site with the N-terminal domain and to the exosite-I with the C-terminal domain (1TOC). **B.** Rhodniin is a double kazal domain inhibitor binding the active site with its N-terminal domain and exosite-I with the C-terminal domain (1TBQ). (For panels A and B, thrombin surface view shown in gray, active site in cyan, exosite-I in orange; and ornithodorin and rhodniin cartoons in blue.)

bonds, and the usual members of this family are slow, tight-binding competitive thrombin inhibitors (Huntington, 2014). And similar to the type member, rhodniin, the first domain of these Kazal proteins binds to the active site of thrombin canonically while the second domain, together with inter-domain linkers, binds to exosite-I. Other examples of Kazal-type thrombin inhibitors from other haematophagous animals include dipetalogastin from the blood-sucking bug, *Dipetalogaster maximus* and infestin from the assassin bug, *Triatoma infestans* (Campos et al., 2002; Mende et al., 1999).

1.4.3.5.5. Lipocalins

The members of lipocalin family of proteins which share a low sequence similarity possess an eight stranded, anti-parallel β-barrel structures and have a hydrophobic ligand binding pocket (Flower et al., 2000). Some members of this family have been identified as thrombin inhibitors. Triabin, the lipocalin identified from the saliva of the triatomine bug, *Triatoma palidipennis* inhibits thrombin exosite-I in a 1:1 molar ratio and prolongs thrombin clotting time and activated partial thromboplastin time. Triabin is the only known thrombin inhibitor from a haematophagous animal that does not block the active site. Triabin completely blocks fibrinogen cleavage by inhibiting the exosite-I with an affinity of 3 pM. Triabin also inhibits thrombin induced platelet aggregation, presumably due to the exosite I-dependence of PAR-1 cleavage (Noeske-Jungblut et al., 1995).

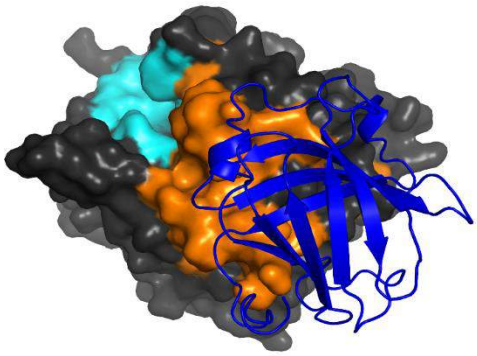


Figure 1.7. Structure of thrombin with triabin. The lipocalin fold containing triabin binds to thrombin exosite-I alone and not to active site (1AVG). (Thrombin surface view shown in gray, active site in cyan, exosite-I in orange; and triabin cartoon in blue.)

1.4.3.5.6. Variegin

The 32-residue long peptide from the salivary gland extracts of *Amblyomma variegatum* was characterized as a fast, tight-binding, competitive thrombin inhibitor. Structurally, variegin lacks secondary structures and exists as a random coil in solution. Variegin selectively binds thrombin with an affinity of 10.4 pM (Koh et al., 2007). Variegin binds to thrombin active site as well as the fibrinogen binding exosite and blocks the thrombus formation. The N-terminus of variegin has been described to be responsible for its fast binding kinetic properties. Variegin gets cleaved by thrombin and the cleaved product retains the ability to inhibit thrombin in a non-competitive way (Koh et al., 2009). The sequence of C-terminus of variegin resembles the sequence of hirulog which is a human-designed, bivalent thrombin inhibitor (Warkentin, 2004). It is therefore interesting to compare the development of a manmade rational antithrombotic drug with that of nature's strategy of developing an anticoagulant through evolution and natural selection.

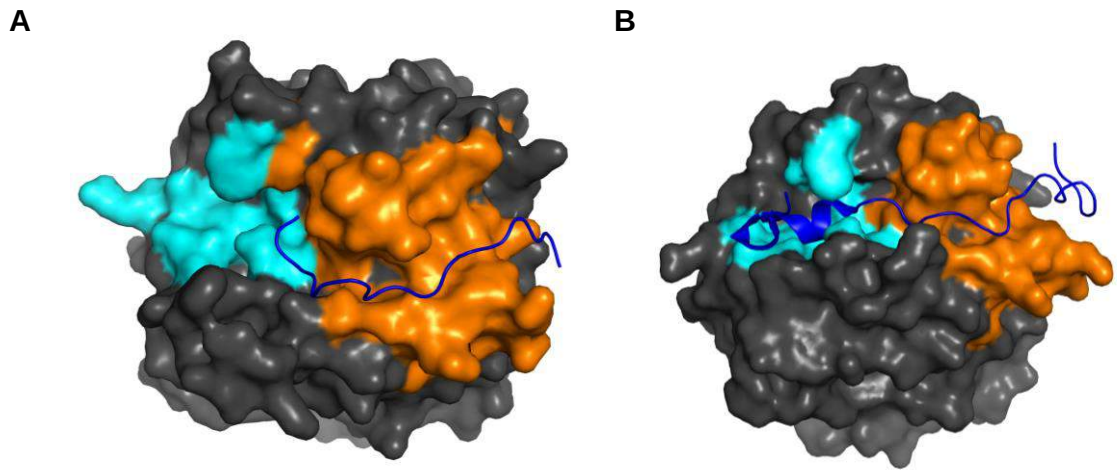


Figure 1.8. Structures of thrombin with peptide inhibitors. A. Variegin is a cleavable inhibitor and the C-terminal cleaved product remains bound to the thrombin active site and exosite-I (3B23). **B.** Anophelin binds thrombin in a unique fashion in which its C-terminus binds to thrombin active site and N-terminus binds to exosite-I (4E05). (For panels A and B, thrombin surface view shown in gray, active site in cyan, exosite-I in orange; and variegin and anophelin cartoons in blue.)

1.4.3.5.7. Anophelin

Anophelin, the 61-residue long peptide from the the mosquito is a unique thrombin inhibitor with a K_i value of 3.5 pM. Anophelin by itself lacks a secondary structure and exists in the form of a random coil in solution. It has an inverted hirudin like interactions with its C-terminus binding to the active site of thrombin and its N-terminus binding to the exosite-I of thrombin. $^{A}Asp^{50}$ of anophelin has been shown to be the most important residue of anophelin as it makes simultaneous hydrogen bonds with both $^{T}His^{57}$ and $^{T}Ser^{195}$ of the thrombin active site. The inhibitory mechnism of the active site of thrombin by anophelin is unique because it delocalizes the proton that is shared between His57 and Ser195, thus disrupting the catalytic triad. Although the N-terminus of anophelin binds to the exosite-I of thrombin, the interactions are sparse (Figueiredo et al., 2012).

1.5. Aim and scope of the thesis

Cardiovascular disorders are a single largest killer worldwide and account for more than 20% of the deaths in Singapore (Ministry of Health, 2012). These cardiovascular diseases involve the formation of unwanted clots within the bloodstream. Antiplatelets and anticoagulants, which are agents that inhibit different components of the haemostatic pathway have been used for the prevention and treatment of these unwanted clots. Heparin, coumarin, warfarin, DTIs and direct FXa inhibitors are some of the most extensively used options of anticoagulant therapy. Despite being the most popular options, these classes are fraught with a number of limitations, such as narrow therapeutic window, high bleeding risks, poor bioavailability and high food drug interactions. Therefore, to overcome these unwanted side effects, superior anticoagulants with greater benefits have been sought. Hematophagous animals produce a plethora of antihaemostatic compunds in their saliva, which they inject into the host at the site of feeding. These antihaemostatic compounds prevent the host haemostatic pathway from being triggered and hence facilitate a successful bloodmeal. Therefore, the saliva of hematophagous animals presents as a pharmacopeia of anticoagulants that can be developed into therapeutics. The aim of this thesis is to identify novel anticoagulants from the saliva of hematophagous animals using different startegies.

Towards this aim, the specific objectives which are within the scope of this thesis are outlined below

1. **Development of a microfluidic chip based online assay coupled to mass spectrometry for the identification of novel thrombin and FXa inhibitors**

Drug discovery from natural sources has been a herculean task due to the scarce availability of the starting material. To address this question, we have developed an online assay in which a nano-LC is coupled with a mass spectrometer and a microfluidic chip. Being operated in the submicrolitre flow rates, this system is a platform for the identification of novel thrombin and FXa inhibitors from sources which are available in limited quantities. The specific aims for this approach were as follows:

i. To standardize thrombin and FXa assays in microtiter plates
ii. To standardize thrombin and FXa assays in the online format and demonstrate the dose response of inhibition
iii. To demonstrate the thrombin and FXa online assay coupled with mass spectrometry
iv. To demonstrate the use of online assay for the identification of thrombin and FXa inhibitors from snake venom

2. **Quantitative sialome of salivary gland extracts of male and female *Dermacentor reticulatus* at different stages of feeding**

Tick salivary gland extracts have been proven to be a rich source of potent pharmacologically active molecules. Blood feeding is known to induce the expression of new anticoagulant proteins which aid the ticks to obtain a successful bloodmeal. We have exploited this phenomenon and tried to identify novel anticoagulants that are produced during different stages of

blood feeding. The specific aims that were adopted for this approach were as follows:

i. To generate the sialotranscriptome of male and female *D. reticulatus*
ii. To generate a quantitative proteome of salivary gland extracts of male and female *D. reticulatus* at different stages of feeding
iii. To identify the differences in the salivary composition between the male and female *D. reticulatus*
iv. To explore the Kunitz type protease inhibitors for the identification of novel subclasses

3. Structural and functional characterization of avathrin, a novel thrombin inhibitor from *Amblyomma variegatum*

Our collaborators, Dr. Maria Kazimirova and Dr. Ladislav Roller, identified variegin like sequences in the salivary glands of *Amblyomma variegatum*. Out of these sequences, we selected one representative peptide, avathrin, chemically synthesized it in our laboratory and tested its ability to inhibit thrombin. After successful confirmation of avathrin's ability to inhibit thrombin, we have studied its structure activity relationships and demonstrated its *in vivo* antithrombotic efficacy in a murine model. The specific aims for this approach were as follows:

i. To study the kinetics and mechanism of inhibition of avathrin
ii. To understand the molecular mechanistic details of avathrin's interactions with thrombin

iii. To design variants that allow better understanding of molecular interactions

iv. To demonstrate the *in vivo* antithrombotic effect of avathrin and establish the feasibility of developing it as an anticoagulant therapeutic

4. Novel family of thrombin inhibitors from ixodid ticks

After the characterization of avathrin, we carried out a detailed analysis of the transcriptomes of other tick species and identified similar sequences. We have selected few interesting sequences and studied their interactions with thrombin. The specific aims for this part of the thesis were as follows:

i. To identify variegin and avathrin like sequences from the sialomes of already studied species of hard ticks

ii. To demonstrate their abilities to inhibit thrombin

iii. To study their molecular interactiosn with thrombin

iv. To describe a pattern which is found in all members of this family

5. Identification of factor Xa inhibitors from the salivary gland extracts of *Rhipicephalus pulchellus*.

Since tick salivary gland extracts have proven to be excellent sources of anticoagulants, we have attempted to purify FXa inhibitors from the salivary gland extracts of female *R. pulchellus*. The specific aims for this part of the project are as follows:

i. To test the anti-FXa activity of male and female salivary gland extracts of *R. pulchellus*

ii. To partially purify the FXa inhibitor using size exclusion chromatography

iii. To purify the partially purified FXa inhibitor from the size exclusion chromatography using affinity chromatography

iv. To test the homogeneity of the purified FXa inhibitor

CHAPTER 2

Microfluidic chip-based online screening coupled to mass spectrometry for the identification of inhibitors of thrombin and factor Xa

Accepted for publication at Journal of Biomolecular Screening

(This work was performed by me at VU University, Amsterdam)

2.1. Introduction

Thrombin and FXa are two key enzymes of the blood coagulation cascade and inhibitors targeting these two enzymes have excellent potential as potent anticoagulant therapeutics (Chaudhari et al., 2014). In search for novel anticoagulants with superior benefits, numerous molecules that inhibit these two enzymes have been purified and characterized from natural sources (Koh and Kini, 2008). Natural sources like plant or animal extracts are usually complex mixtures which are available in limited amounts and are difficult to process.

High-resolution screening techniques in which separation methods are coupled to bioassays that eliminate traditional, labour-intensive purification and screening tasks, and facilitate rapid identification of active compounds from complex mixtures are becoming increasingly popular (Shi et al., 2009; Wigglesworth et al., 2015). These HRS techniques have also been adapted for the identification of enzyme inhibitors, where an LC-based separation of a complex mixture is coupled to an absorbance or fluorescence based bioassay (Entzeroth, 2003). Recent advancements have fuelled the hyphenation of separation techniques with mass spectrometry for accurate identification of active molecules (Graßmann et al., 2012; Kool et al., 2011). HRS techniques in which analyte separation coupled with biochemical detection in parallel with mass spectrometric identification, for ligand binding as well as enzyme inhibition have been developed in both at-line and on-line formats (Giera et al., 2009; Graßmann et al.,

2012). Such HRS techniques are being used for the identification of individual compounds as lead molecules from complex mixtures like biological extracts or combinatorial libraries (Shin and van Breemen, 2001). Scarce availability of starting material is a key limitation in this process for many types of biological extracts (Valenzuela, 2004). Hence, there is a need for new platforms that can use minute amounts of starting material for the identification of highly active lead molecules. Although microfluidic systems show lower sensitivities and thus higher limits of detection (approximately 4-6 times lower sensitivities) than macro scale systems, these sensitivities have been reported to be within the concentration range of bioactivity detection, and hence are implemented for screening of bioactives from limited sample amounts (Irth, 2007). Therefore, microfluidic systems provide detection platforms for identifying bioactives from limited sample amounts and consume lesser amounts of bioassay reagents without a compromise in assay sensitivity (de Boer et al., 2005). Moreover, when coupled to a sensitive nano-spray ESI-MS, these platforms also allow parallel mass determination and further peptide analysis from minute amounts of starting material which macro scale systems fail to handle.

We have described microfluidic high resolution screening (HRS) procedures for bioaffinity profiling of the acetylcholine binding protein (AChBP) along with accurate mass determination (Heus et al., 2014; Kool et al., 2010; Otvos et al., 2013). For this, a nano-HPLC is coupled online to a fluorescence enhancement assay in parallel with a mass spectrometer, to detect bioactive compounds and identify their masses in a single analysis. We have also developed

a microfluidic chip-based method for screening of cathepsin B inhibitors and an HRS LC-MS method for detection and identification of small molecule inhibitors of p38α mitogen-activated protein kinase (de Boer et al., 2005; Falck et al., 2010). Here, we describe a microfluidic post-column method for identifying thrombin and FXa inhibitors from a mixture of compounds. Since both enzymes prefer to hydrolyze an arginyl bond, we have used the same fluorogenic substrate, R22124 (Rhodamine 110, bis-(p-Tosyl-L-Glycyl-L-Prolyl-L-Arginine Amide)) for both enzymes, to keep the method simple.

2.2. Materials and Methods

2.2.1. Chemicals and Reagents

Thrombin and FXa were purchased from BV Bioconnect, Haematologic Technologies, Amsterdam, The Netherlands. Substrate R22124, Rhodamine 110, bis-(p-Tosyl-L-Glycyl-L-Prolyl-L-Arginine Amide) was obtained from Life technologies, Amsterdam, The Netherlands. NaCl, KCl, $Na_2HPO_4 \cdot 2H_2O$, KH_2PO_4, argatroban monohydrate, Tween 20 and PEG 6000 were purchased from Sigma-Aldrich (Zwijndrecht, The Netherlands). ELISA blocking reagent was purchased from Hoffman-La-Roche (Mannheim, Germany). ULC-MS grade methanol, ULC-MS grade acetonitrile (ACN) and trifluoroacetic acid (TFA) were obtained from Biosolve (Valkenswaard, The Netherlands). HPLC grade water was produced using Milli-Q purification system (Amsterdam, The Netherlands). Venom of *Dendroaspis polylepis* was provided by Dr. Ryan McCleary.

2.2.2. Thrombin and FXa assays in microtiter plates

Enzyme assays were performed in 96-well black microtiter plates (Griener Bio-One GmBH, Kremsmünster, Austria) to evaluate parameters such as K_M, V_{max}, and tolerance to additives. Briefly, 50 µL of PBS (pH 7.4) was mixed with 30 µL of R22124, and 20 µL of enzyme was then added to start the reaction. The rate of release of the fluorescent product R110 (excitation, 490 nm; emission, 520 nm) at room temperature was measured for 10 min with an Infinite 200Pro microtiter plate reader (Männedorf, Switzerland). For determination of K_M and V_{max}, two different enzyme concentrations (thrombin: 100 ng/mL and 20 ng/mL; FXa: 1000 ng/mL and 250 ng/mL) along with different substrate concentrations

(for thrombin assay: 60, 20, 6.6, 2.2, 0.74 and 0.24 µM; and for FXa assay: 50, 25, 12.5, 6.25, 3.12, 1.56 and 0.78 µM) were used. Initial reaction velocities were calculated by calculating the slopes of the first eight points of the reaction progression curves. K_M and V_{max} values were obtained by fitting the data to Michaelis-Menten equation using the GraphPad Prizm software. To determine tolerance of the assay to different additives, similar enzyme assays (with different concentrations of additives such as PEG 6000 (2, 1, 0.5, 0.25, 0.12, 0.62, 0.31 and 0.15, 0.007 mg/mL), ELISA blocking reagent (10, 5, 2.5, 1.25, 0.62, 0.31, 0.15, 0.07 and 0.03 mg/mL), Tween 20 (2.5, 1.25, 0.62, 0.31, 0.15, 0.0.08, 0.04, 0.02 and 0.01%), acetonitrile or methanol (10, 5, 2.5, 1.25, 0.62, 0.31, 0.15, 0.07 and 0.03%) and TFA (0.1, 0.05, 0.02, 0.01, 0.006, 0.003, 0.001, 0.0008 and 0.0004%)) were performed. The reaction rates were calculated as described above and percentage inhibition at different concentrations of additives was determined by comparing the reaction velocities in the presence and absence of additives. The parameters obtained from these studies were used for the development and optimization of the online assays.

2.2.3. Instrumentation

The instrumental setup for the online assay (Fig. 2..1 A) consisted of a nano-LC system, a miniaturized reaction coil in the form of a microfluidic chip, a microfluidic LED-based confocal detector and a mass spectrometer.

2.2.3.1. Nano-LC system

The Dionex Ultimate 3000 nano-LC system was from Thermo Fisher (Breda, The Netherlands).

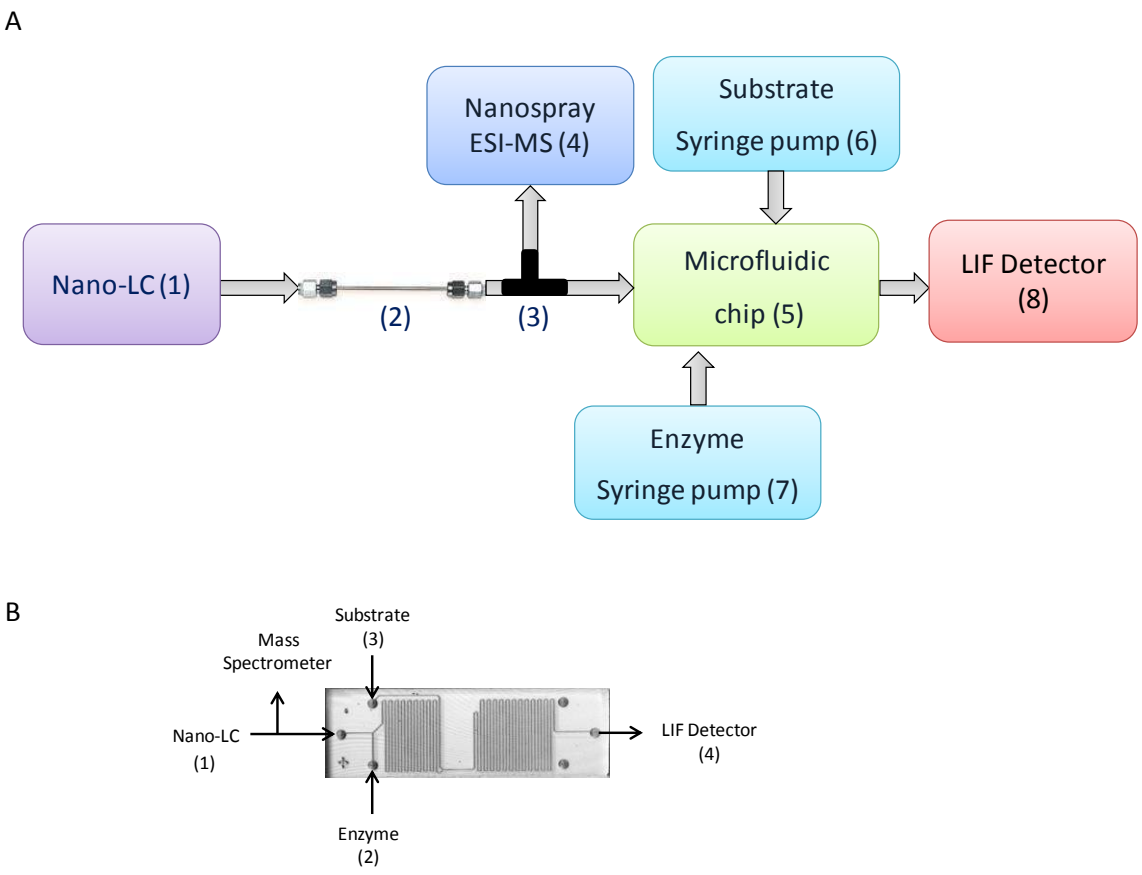


Figure 2.1. Online setup for identification of thrombin and FXa inhibitors. A. Instrumental set-up for online assay consists of three parts: a nano-LC, a microfluidic chip and a mass spectrometer. With the nano-LC injector (1), samples (500 nL- 1µL) were injected into the C18 capillary column (2). The eluate from the capillary column was split post-column in a 1:1 ratio using a T-splitter (3) in such a way that one part was fed into a nanospray ESI-MS (4), and the other part was infused into the microfluidic chip (5) in which the assay took place. The enzyme (thrombin/FXa) and substrate (R22124) were infused continually into the microfluidic chip using syringe pumps (6 & 7). The microfluidic chip outlet was channelled into the LIF detector (8), which measured the amount of fluorescent product formed. **B**. The microfluidic chip had three inlets connected to (1) nano-LC, (2) Enzyme, (3) substrate, and one outlet, (4) LIF detector. The first micro reactor (1.6 µL) acted as an in-flow incubation of enzyme with inhibitors eluting from nano-LC. The enzyme substrate reaction took place in the second micro reactor (2.4 µL). The amount of fluorescent products formed was measured by LIF detector.

Flow rate of the nano-LC was maintained at 400 nL min^{-1}. Samples (500 nL to 1 µL) were injected in the Flow Injection Analysis (FIA) mode or in the gradient analysis mode using nano-LC controlled autosampler. For FIA experiments, 100% Milli-Q water was used as the mobile phase. For gradient analysis, a capillary column (150 mm X 75 µm i.d.) packed in-house with Phenomenex Aqua C18 particles (5 µm, 200 Å pore diameter) was used with mobile phase eluent A as water/ACN 99:1 and 0.1% TFA; and eluent B as water/ACN 1:99 and 0.1% TFA. The column was first equilibrated for 5 min at 5% eluent B and then a linear gradient from 5% to 70% B in 60 min was applied. Eluate (400 nL min^{-1}) from the column was split via a T-splitter in a 1:1 ratio and one half was infused into the microfluidic chip and the other half was fed into a mass spectrometer. Two fused silica capillaries of identical length (150 cm) and diameter (25 µm i.d.) were used to connect the nano-LC with the mass spectrometer and with the microfluidic chip to obtain a 1:1 split ratio. After each run, the column was washed with 70% eluent B for 5 min and re-equilibrated with 5% eluent A for 3 min before the next run.

2.2.3.2. Microfluidic chip

The microfluidic chip and 4515 chip holder were from Micronit Microfluidics (Enschede, The Netherlands), and microfluidic chip nanoports were from Upchurch (Amsterdam, The Netherlands). The microfluidic chip (described previously) with dimensions 45 mm × 15 mm × 2.2 mm had an open tubular channel that was 125 µm wide × 70 µm deep (Irth, 2007). This open tubular channel had a total volume of 4 µL and was divided into two micro reactors of

volumes 1.6 µL and 2.4 µL by incorporating three inlets and one outlet (Fig. 1b). The enzyme (thrombin or FXa) was infused through the first inlet via a 2.5 mL syringe driven by syringe pumps (Harvard Apparatus, Holliston, USA) at a constant flow rate of 2 µL/min. Eluate from C18 column (containing potential inhibitors) was infused through the second inlet at a flow rate of 200 nL/min. This allowed an in-flow incubation of enzyme with the eluting inhibitor in the first micro reactor. Substrate was infused through the third inlet via a 2.5 mL syringe driven by a syringe pump at a flow rate of 2 µL/min. In the second micro reactor, enzyme hydrolyzes the substrate releasing the fluorophore. The outlet from the chip was channelled to a microfluidic confocal fluorescence detector (CFD) to measure fluorescence of the product formed. The chip was connected to syringe pumps, nano-LC and detector using fused silica capillaries, placed in the chip holder, and maintained at 37°C in a water bath.

2.2.3.4. LED-based microfluidic confocal fluorescence detector (CFD)

The LED-based microfluidic CFD as described previously consisted of a bubble cell capillary of 150 µm i.d. and connecting capillaries of 50 µm i.d. from Agilent Technologies (Amsterdam, The Netherlands) (Heus et al., 2010). Briefly, in the CFD setup, light emitted from the LED passes through the 465 nm single bandpass filter followed by collimation by the lens, and then reflected under a 90° angle using dichroic mirror. This filtered light is focused onto the center of the florescence optical cell. Emitted light is allowed to pass through the same dichroic mirror, the focussing lens and the 520 nm single bandpass filter, and subsequently detected by the photomultiplier tube.

2.2.4. Mass Spectrometry settings

The ion trap time-of-flight (IT-TOF) hybrid mass spectrometer from Shimadzu (Hertogenbosch, The Netherlands), with Picoview nano-ESI source from New Objective (Woburn, MA, USA), was operated in the positive ion mode. A 40 mm × 180 mm o.d. × 30 mm i.d. stainless-steel emitter from Thermo Scientific (Waltham, MA, USA) served as the spray needle. The temperature of the heating block was set to 200°C and interface voltage was set to 1.7 kV which resulted in a current of 32 µA. The scan range was *m/z* 200 to 1000.

2.2.5. Microfluidic online assay of snake venom

For the analysis of snake venom, the nano-LC was connected to MS and microfluidic chip (as described in *2.3*). After stabilization of the elevated baseline fluorescence, 5 µg of *Dendroaspis polylepis* venom dissolved in 500 nL and spiked with 500 ng of argatroban, dissolved in mobile phase eluent A was injected into the C18 capillary column. The column was equilibrated with 5% eluent B for 5 min and the snake venom was separated using an acetonitrile gradient (5% to 70% eluent B in 60 min).

2.3. Results and Discussion

Microfluidic chip-based enzymatic assays and receptor binding studies for the detection of bioactives from complex mixtures have been used as alternatives to traditional screening approaches because of their speed, robustness, reproducibility and sensitivity. They are most suitable for screening of novel drug leads from minute sample amounts. We have developed fluorescence based online screening methods coupled with mass spectrometry in miniaturized formats for bioaffinity profiling of AChBP ligands, cathepsin B inhibitors and inhibitors of p38α mitogen-activated protein kinase (de Boer et al., 2005; Falck et al., 2010; Heus et al., 2014; Kool et al., 2010; Otvos et al., 2013). In this study, we describe the development of an analytical platform for profiling of thrombin and FXa inhibitors.

2.3.1. Standardization of enzyme assays in microtiter plates

The assay is based on enzymatic cleavage of a fluorogenic substrate-R22124 by thrombin or FXa to release a fluorescent product. K_M and V_{max} values for R22124 with thrombin and FXa were determined as described in the methods. The K_M and V_{max} for R22124 with thrombin were 4.05±0.87 µM and 81±4.72 RFU min^{-1} ng^{-1} (thrombin) respectively, and K_M and V_{max} for R22124 with FXa were 4.29±0.55 µM and 10.01±0.38 RFU min^{-1} ng^{-1} (FXa) respectively (Fig. 2.2 A and B). As R22124 is a specific substrate for thrombin, it has 8-times higher V_{max} values, and thus, a 12-times higher concentration of FXa than thrombin was used to obtain a similar high fluorescence signal during the short (online assay)

incubation times. For all further assays in microtiter plate format, 20 ng/mL thrombin and 250 ng/mL and FXa respectively were used.

Non-specific adsorption of enzymes and substrates to glass surfaces of the microfluidic chip and fused silica capillaries leads to increasing baselines and peak broadening (Heus et al., 2010). To block this non-specific adsorption, blocking reagents like PEG 6000, Tween 20 and ELISA blocking reagent were used in the online assays.

A B

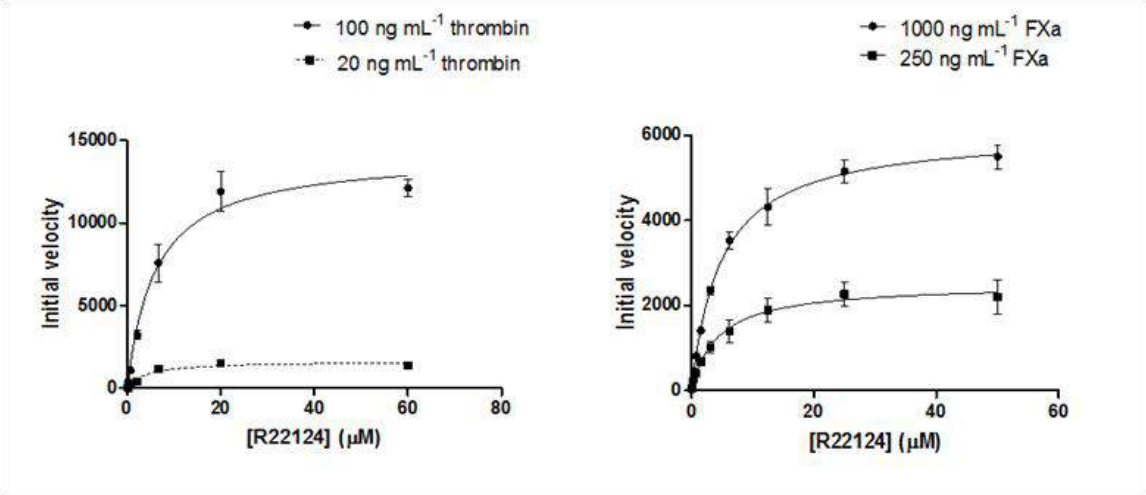


Figure 2.2. Effect of substrate concentrations microtiter assays. (A) Thrombin, (B) FXa. The K_M for R2214 were determined using 100 ng/mL (6.03±0.88 µM) and 20 ng mL^{-1} (4.05±0.87 µM) for thrombin, and 1000 ng/mL (4.87±0.25 µM) and 250 ng/mL for FXa (4.29±0.55 µM).

As these additives may inhibit the assay or lead to enzyme precipitation within the microfluidic chip, their effect on the assay was evaluated using microtiter plate assays prior to translation to online format. The enzyme activity of thrombin or FXa was not affected in the presence of PEG 6000 (up to 2

mg/mL), Tween 20 (up to 0.02%) or ELISA blocking reagent (up to 3 mg/mL) in the reaction mixture (data not shown).

In the online assay setup, organic mobile phase solvents like ACN and methanol used for gradient elution in the nano-LC system will be infused into the microfluidic chip and if the concentration of these solvents is significantly high, they might reduce thrombin and FXa enzymatic activities due to denaturation. Therefore, the effect of different concentrations of ACN and methanol (up to 10%) on activities of thrombin and FXa was evaluated using microtiter plate assays. At lower concentrations of these organic solvents (2% to 5%), there were slight increases in thrombin and FXa enzymatic activities. Overall, the activities of both thrombin and FXa were not significantly affected in the presence of ACN or methanol at concentrations that would be used in the online assay. In the online assay format, 200 nL min^{-1} of the nano-LC eluate will be infused with 4 μL min^{-1} of reagents of enzyme assay (2 μL min^{-1} each of enzyme and substrate). Since the online assay reagents are prepared in a PBS buffer (pH 7.4), the final concentration of organic solvents in the online assay taking place within the microfluidic chip will be diluted to a concentration that it will have little or no effect on the enzyme activities (organic solvent concentrations will remain well below 10% even when the mobile phase gradient of the nano-LC reaches up to 70% eluent B). The effect of different concentrations of ion pairing agents TFA and FA on enzyme activities was also evaluated. At the concentrations that would be used in the online assay (up to 0.01%), no inhibition was observed for both ion

paring agents. Since TFA is the most commonly used ion pairing agent and gives better resolution, TFA was used in the online assay.

2.3.2. Optimization of online assay parameters by flow-injection analysis (FIA)

2.3.2.1. Microfluidic setup

The microfluidic chip has two micro reactors, where the enzyme and inhibitor are infused in the first, and substrate is infused in the second. Thus the EI complex is formed in the former and the enzymatic cleavage of the substrate to release the fluorescent product takes place in the latter. The continuous infusion of enzyme and substrate, and thus constant release of fluorescent products in the second micro reactor, gives stable elevated baseline fluorescence. The in-flow incubation of inhibitors eluting from the nano-LC provides its efficient mixing with the assay buffer and enzyme in the first micro reactor. As these inhibitors bind to and inhibit the enzyme, the rate of substrate turnover is decreased which is observed as a negative peak.

2.3.2.2. Standardization of enzyme assays in the online format

For the online assays, enzymes and substrates were used at concentrations lower than those used for the microtiter plate assays. These lower concentrations were chosen to avoid the formation of saturating amounts of fluorescent products. Initially, different enzyme concentrations (1.5 ng/mL, 4.5 ng/mL and 13.5 ng/mL for thrombin, and 10 ng/mL, 25 ng/mL and 50 ng/mL for FXa) were used with a constant substrate concentration (3.3 µM of substrate for thrombin assays and 2.5

µM for FXa assays). Decreasing amounts of benzamidine (13.8 µg, 2.76 µg and 0.55 µg) were injected in triplicate into the system in the FIA mode (Fig. 2.3 A; Fig 2.4 A). For thrombin and FXa, 4.5 ng/mL and 25 ng/mL were chosen as optimal concentrations respectively. At lower enzyme concentrations the detection limits of inhibitors is decreased, while at higher enzyme concentrations, inhibition peaks were observed for blank injections.

The substrate concentration was optimized by using a constant enzyme concentration (thrombin, 4.5 ng/mL; FXa, 25 ng/mL) and different substrate concentrations (1.1 µM, 3.3 µM and 9.9 µM for thrombin; 1.25 µM, 2.5 µM, 4.5 µM for FXa). Decreasing amounts of benzamidine (13.8 µg, 2.76 µg and 0.55 µg) were injected in triplicate into the system in the FIA mode. For thrombin, and FXa, 3.3 µM and 2.5 µM were chosen as optimal R22124 concentrations, respectively (Fig. 2.3 B; Fig. 2.4 B). At lower substrate concentrations, the low baseline fluorescence obtained decreased the detection limit, whereas at higher substrate concentrations blank injections (without inhibitor) gave false inhibition peaks.

For all further experiments, enzymes were prepared in PBS buffer (PH 7.4) containing 0.5 mg/mL PEG 6000 and 0.5 mg/mL ELISA blocking reagent, and substrate R22124 was prepared in a buffer containing 0.01% Tween 20.

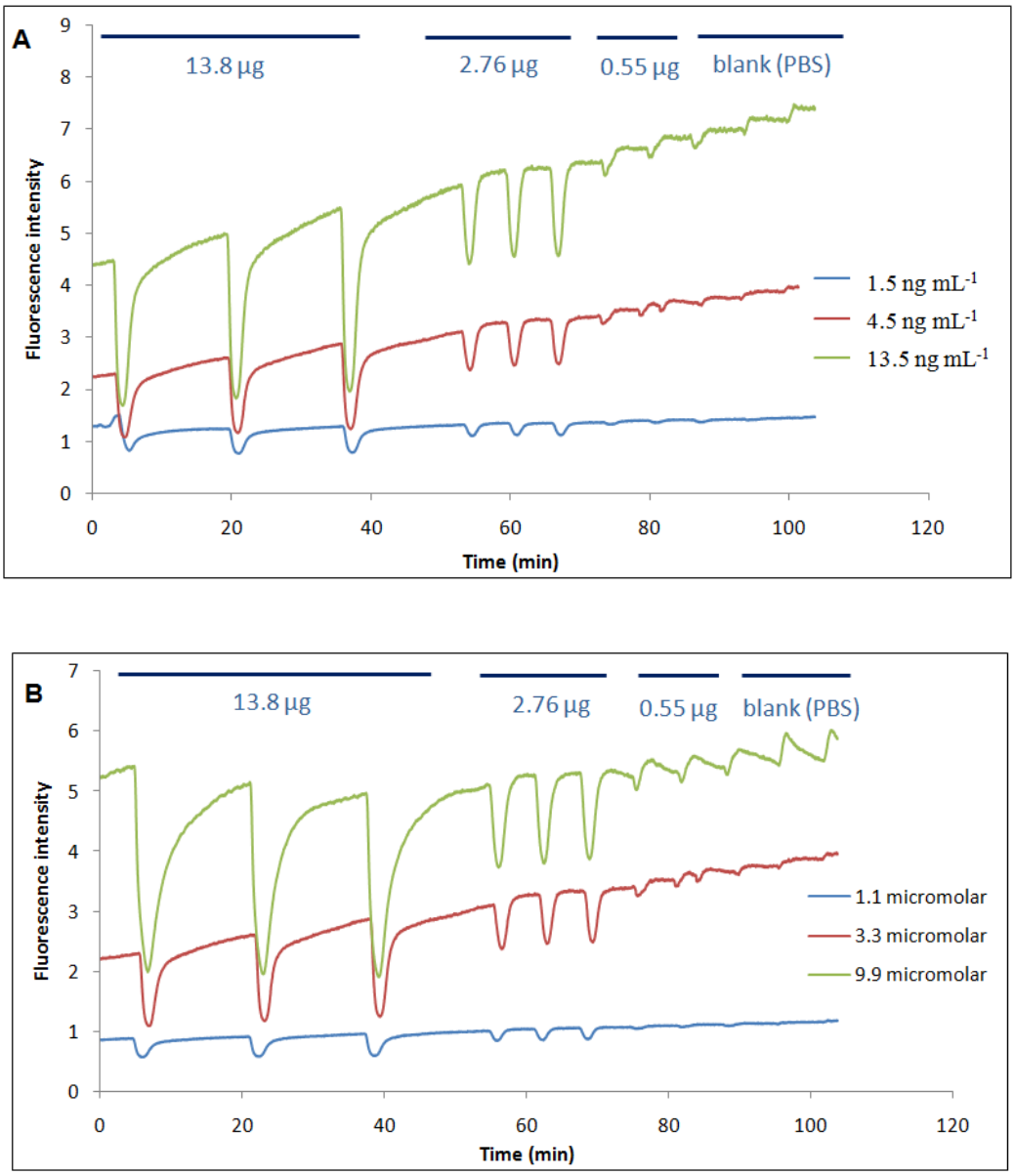


Figure 2.3. Optimization of enzyme and substrate concentration for online assays. A. Optimization of thrombin concentration. Three thrombin concentrations (1.5ng/mL, 4.5ng/mL and 13.5 ng/mL) were used. After stabilization of the elevated fluorescence baseline, triplicate injections of 13.8 µg, 2.76 µg and 0.55 µg of benzamidine were made. For further experiments, 4.5 ng/mL was selected as the optimum thrombin concentration. B. Optimization of R22124 concentration. Three R22124 concentrations (1.1 µM 3.3 µM and 9.9 µM) were used. After stabilization of the elevated fluorescence baseline, triplicate injections of 13.8 µg, 2.76 µg and 0.55 µg of benzamidine were made. For further experiments, 3.3 µM was selected as the optimum R22124 concentration.

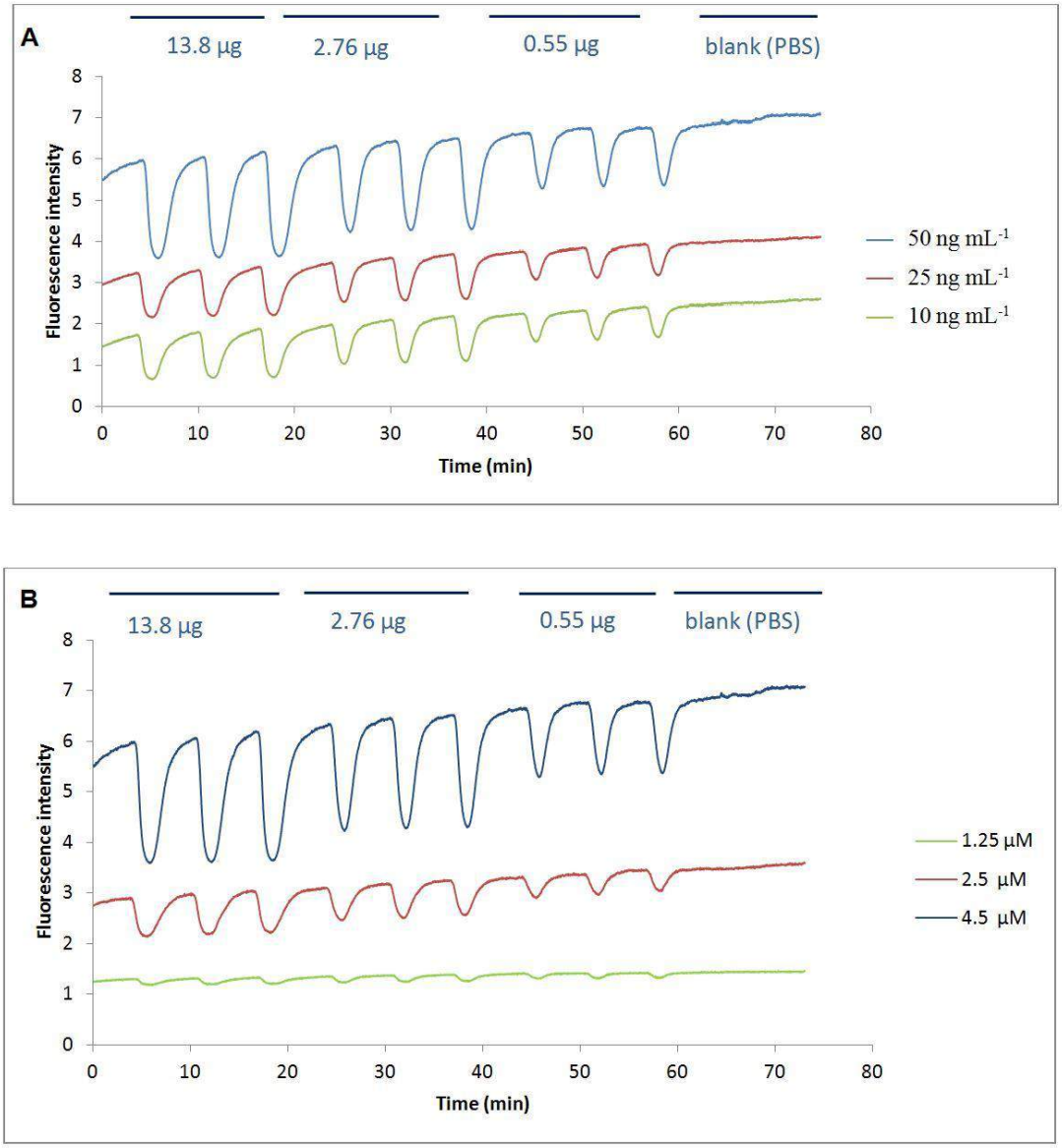
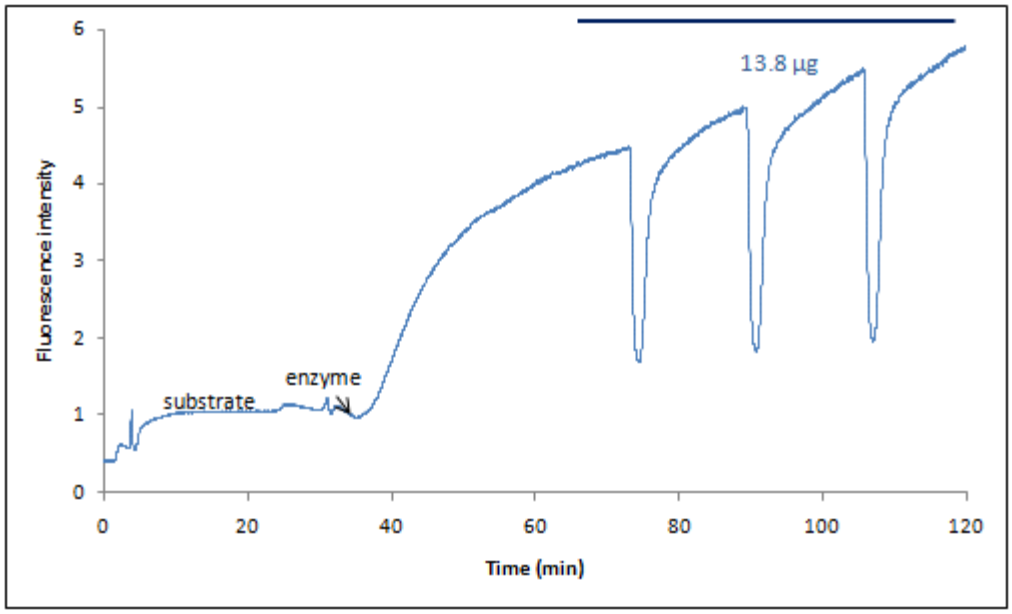


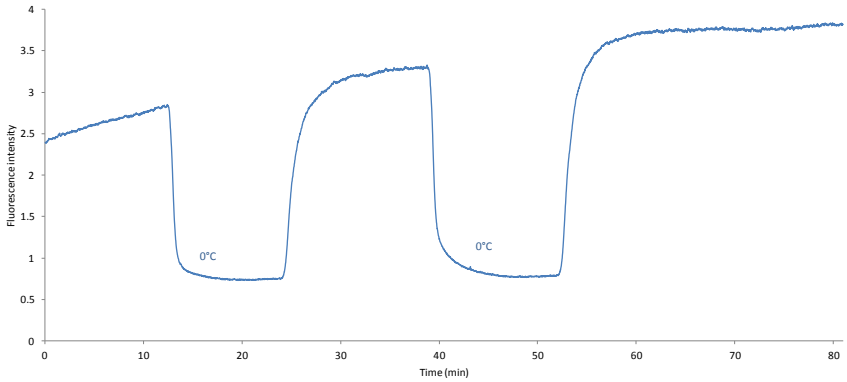
Figure 2.4. Optimization of the enzyme and substrate concentration for the online FXa assays. A. Optimization of FXa concentration. Three concentrations (50 ng/mL, 25 ng/mL and 10 ng/mL) of thrombin were used for optimization. After stabilization of the elevated fluorescence baseline, triplicate injections of 13.8 µg, 2.76 µg and 0.55 µg of inhibitor benzamidine were made. The negative peaks indicate the inhibition of thrombin by benzamidine. For further experiments, 25 ng/mL was selected as the optimum FXa concentration. B. Optimization of R22124 concentration. Three concentrations of R22124 (1.25 µM, 2.5 µM and 4.5 µM) were used. After the stabilization of the elevated fluorescence baseline, triplicate injections of 13.8 µg, 2.76 µg and 0.55 µg of benzamidine were made. For further experiments, 2.5 µM was selected as the optimum R22124 concentration.

2.3.3. Validation of the online assay by flow-injection analysis (FIA)

The setup for the online enzymatic assays was validated in the FIA mode. The substrate, R22124 was continually infused into the microfluidic chip until a stable substrate baseline was obtained (Fig. 2.5 A). When thrombin infusion was started, there was a large increase in the fluorescence because of the formation of the fluorescent product and this fluorescence stabilized as an elevated baseline indicating that product formation had reached a steady state. Subsequently, 13.8 µg of benzamidine was injected in triplicate through the nano-LC. Each injection resulted in a reproducible negative peak (Fig. 2.5 A). To demonstrate that the increase in the elevated fluorescence baseline is due to the enzyme-substrate reaction, the chip was incubated at different temperatures (at 0°C and at 37°C). An increase in the baseline fluorescence was observed when the chip was placed from 0°C to 37°C and a sudden dip in fluorescence when the chip was again incubated at 0°C. This behaviour shows that the elevated baseline fluorescence is due to the enzymatic formation of fluorescent product because the enzyme-substrate reaction is optimal at 37°C, while low or no product formation occurs at 0°C (Fig. 2.5 B). The difference in the height of the baseline is also a good indication of the total assay window in the online assay. This assay window was confirmed by injecting a high concentration of inhibitor (20 µg) that fully inhibited enzymatic activity (data not shown).

A





B

Figure 2.5. A. Validation of the online assay. Substrate was continually infused into the microfluidic chip until a stable baseline was obtained. Thrombin infusion led to an increase in fluorescence. After stabilization of the elevated fluorescence baseline, triplicate injections of 13.8 µg of benzamidine were made. The negative peaks indicated inhibition of thrombin by benzamidine. **B. Temperature dependence of the enzymatic reaction.** The microfluidic chip was incubated at different temperatures (0°C and 37°C). The temperature dependence was demonstrated by the sudden dip in the fluorescence intensity when the chip is incubated at 0°C.

In order to demonstrate the robustness of the assay, statistical parameters like screening window coefficient (Z'-factor), dynamic range, signal-to-background ratio (S/B), signal-to-noise ratio (S/N) and signal window (SW) were calculated for the thrombin assay (Table 2.1) and for the FXa assay (Table 2.2) using different substrate (R22124) concentrations according to previously described methods (Zhang et al., 1999).

For the thrombin assay, Z'-factors in the range of 0.5-1 were obtained for all R22124 concentrations of 3 µM and above, indicating an excellent assay and the highest dynamic range (2.9) was obtained for 20 µM R22124. For the FXa assay, Z'-factors in the range of 0.5-1 were obtained for all R22124 concentrations of 2.5 µM R22124 and above, and the highest dynamic range (2.33) was obtained for 2.5 µM R22124.

Table 2.1. Calculated statistical parameters for thrombin assay using different R22124 concentrations.

[R22124] (µM)	Z'-factor	Dynamic range	S/B	S/N	SW
20	0.85	2.90	1.47	71.83	22.72
10	0.55	2.69	2.12	8.98	14.25
3	0.55	1.00	1.71	21.23	5.25
1	-0.02	0.30	1.49	4.88	-0.17

Table 2.2. Calculated statistical parameters for FXa assay using different substrate (R22124) concentrations.

[R22124] (µM)	Z'-factor	Dynamic range	S/B	S/N	SW
4.5	0.84	2.33	1.62	42.36	27.86
2.5	0.67	0.73	1.32	20.24	11.26
1.25	0.14	0.13	1.11	7.50	0.90

2.3.4. Dose-response of inhibition in the online assay

To demonstrate the dose-dependent inhibition, the online assay was set up using optimized thrombin and substrate concentrations, and after stabilization of the elevated fluorescence baseline, different amounts of the inhibitor benzamidine (13.8 µg, 4.6 µg, 1.53 µg, 0.52 µg, 0.17 µg, 0.05 µg and 0.016 µg) followed by a blank were injected in triplicates into the online assay system through the nano-LC using the FIA mode. After each injection, a negative peak in fluorescence was observed (Fig. 2.6 A). The peak heights directly correlated with the concentrations of inhibitor. One important note in this regard is that the final inhibitor concentration in the assay after injection is diminished due to two factors: (i) due to the mixing of nano-LC eluate with reagents of the online assay (mixing dilution D_M) and (ii) due to chromatographic dilution (D_C). Detailed descriptions of calculating the final inhibitor concentrations considering these two dilutions are given elsewhere (Falck et al., 2010). The peak heights were plotted against log final inhibitor concentrations to determine the IC_{50} values. The IC_{50}

for benzamidine was determined to be 1.13±0.02 µM for thrombin and 4.96±0.12 µM for FXa (Fig. 2.6 B).

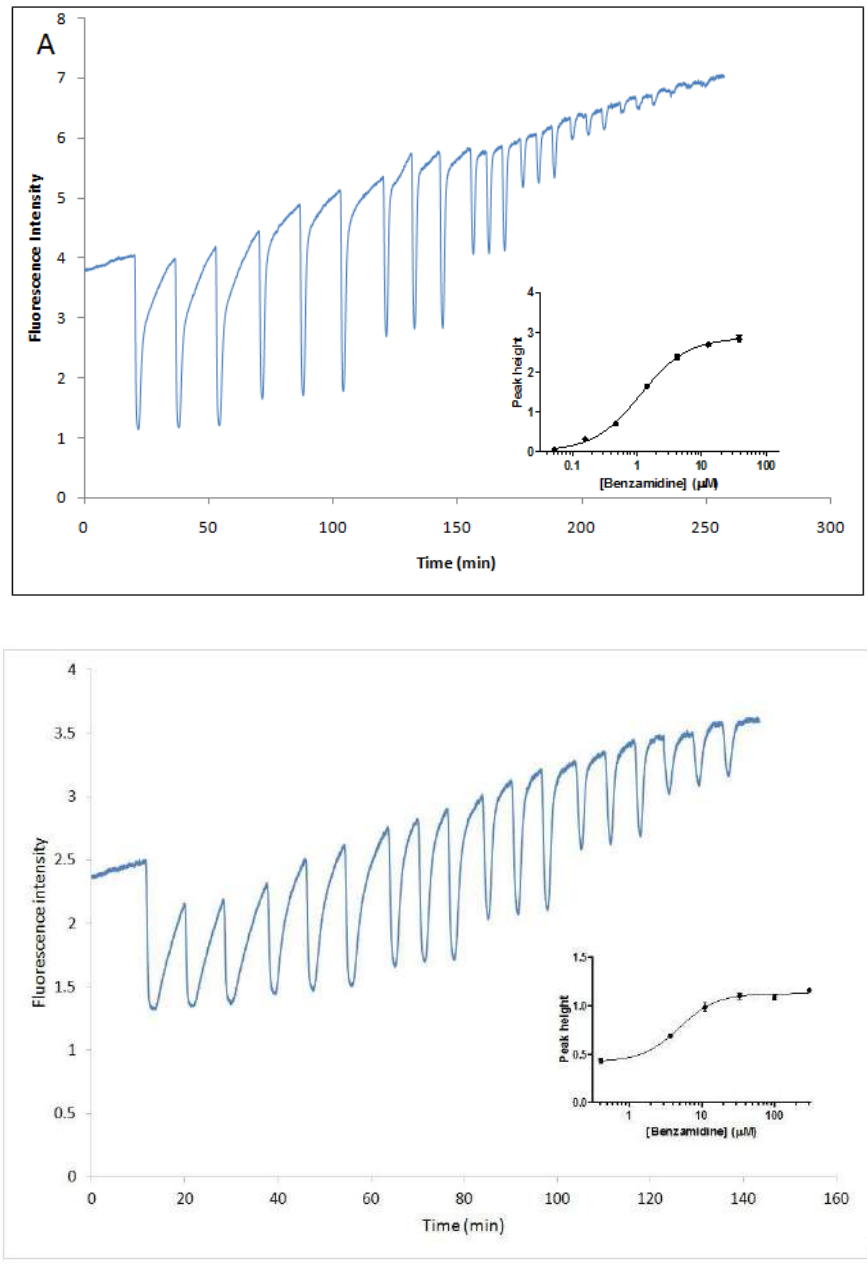


Figure 2.6. Dose-response of inhibition in the online (A) thrombin assay, (B) FXa assay. Decreasing amounts of benzamidine were injected in triplicates. The final inhibitor concentration after injection is diminished due to two factors, (1) due to the mixing of nano-LC eluate with reagents of online assay and (2) due to chromatographic dilution and peak broadening. The final concentration of

inhibitor in the online assay was calculated considering these two dilutions as described previously. Inset, the plot of peak heights obtained in the online assay versus benzamidine concentrations. IC_{50} for thrombin assay was 1.13±0.02 µM. The IC_{50} for FXa assay was determined to be 4.96±0.12 µM.

2.3.5. Online thrombin assay coupled with mass spectrometry

To demonstrate the online thrombin assay coupled to the MS, the nano-LC was connected to the MS and the microfluidic chip such that one half of the eluate from the nano-LC was fed into the assay and the other half into the MS. Briefly, once the elevated baseline fluorescence was obtained for the enzyme-substrate reaction, 1 µM argatroban was injected into the C18 capillary column and eluted using an acetonitrile gradient (5% to 70% in 60 min). The eluted argatroban (507.24 Da) detected by the MS at 48 min could be correlated as a bioactive negative peak in the online assay at 54 min. This delay is due to the additional volume that the inhibitor has to flow through the microfluidic chip before it is detected in the fluorescence detector. Similarly, when a final concentration of 1 mM benzamidine was injected into the C18 column, dimer masses (277.11 and 241.14 Da) detected by the MS at 10 min with a bioactive negative peak at 17 min. Subsequently, a mixture of benzamidine (1 mM) and argatroban (1 µM) was injected into the capillary C18 column and eluted as described above (Fig. 2.7 A). From this mixture, benzamidine dimers (277.11 and the 241.14 Da) were detected at 10 min in the MS with a negative bioactive peak at 17 min. Argatroban (507.24 Da) was detected at 48 min in the MS with the negative peak at 54 min in the assay. The second bioactive negative peak eluting at 46 min was not observed in MS. This bioactive was a low molecular weight impurity from benzamidine and in such a case, with the scan range used (*m/z* 200-1000), the mass of this eluting

compound could not be detected. Benzamidine is a low affinity, mechanism based inhibitor. This type of inhibition where inactivation of the active site occurs slowly, has a clear time dependency on the efficacy of inhibition. Sensitive measurement of mechanism based inhibitors in the online assay, which has a very short post-column incubation time, is a drawback of this assay format. For analyzing benzamidine, a high concentration was injected. With the scan range used (*m/z* 200 to 1000) only the early eluting benzamidine dimers (277.11 and the 241.14 Da) resulting from ($[2M+HCl]^+$ and $[2M+H]^+$) could be observed at 10 min in the MS. However additional experiments with a scan range of *m/z* 100-1000 showed that monomer benzamidine eluted at 11 min in the UV and in the MS. But severe ion suppression from tailing benzamidine at a scan range *m/z* 100-1000 masked the compounds eluting at lower concentration to be observed in MS and hence scan ranges lower than *m/z* of 200 were not used. When additional experiments with the same mixture using a scan range of *m/z* 100 to 1000 were performed, due to severe ion suppression, only the benzamidine monomer (121.07 Da) could be observed, but benzamidine dimers (277.11 and 241.14 Da) were observed at much lower intensities and argatroban (509.27 Da) could not be observed anymore. Measurement with a scan range of *m/z* 200 to 1000 on the contrary, did allow the observation of the early eluting benzamidine dimers (277.11 and 241.14 Da). When the two analytical runs were compared, an elution time difference of 1 min between the different forms of benzamidine was observed with the monomer benzamidine (121.07 Da) observed with *m/z* 100 to 1000 scan range eluting one minute later than the benzamidine dimers (277.11

and 241.14 Da) observed with *m/z* 200 to 1000 scan range. UV data can in this case help by providing additional information to be used for correlation purposes.

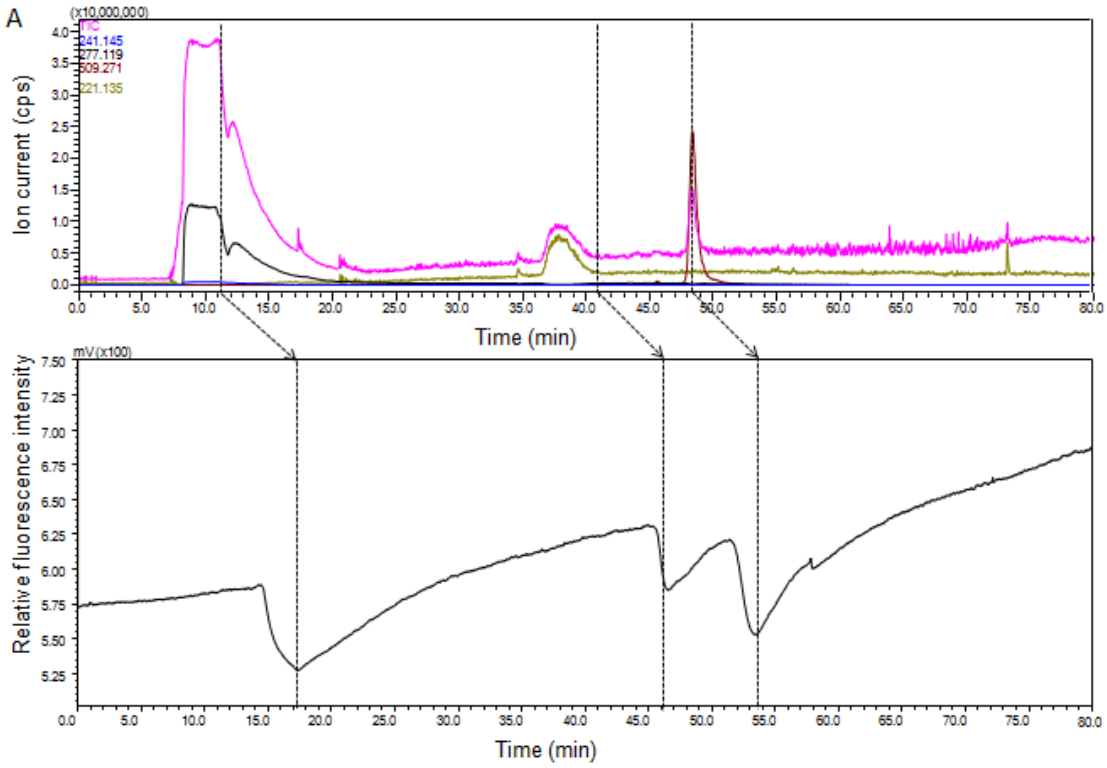


Figure 2.7 A. Online thrombin assay with parallel mass spectrometry after nano-LC for analysis of a mixture of benzamidine and argatroban. Benzamidine eluting from the nano-LC was detected by the MS as its dimer masses (277.11 and 241.14 Da) at 10 min in the MS and 17 min in the assay. The scan range used of 200 -1000 did not allow the observation of benzamidine (121.07 Da). Argatroban (507.24 Da) detected by the MS at 48 min correlated with the bioactive negative at 54 min. The second bioactive negative peak observed at 46 min could be a low molecular weight impurity and was not observed in the MS.

2.3.6. Online FXa assay coupled with mass spectrometry

Similar to the thrombin assay, the online FXa assay coupled to mass spectrometer was also demonstrated. The FXa assay was set up in the microfluidic chip with the optimized FXa and R22124 concentrations and one half of the flow from the nano-LC column was fed into the assay and the other

half into the mass spectrometer. After the stabilization of the elevated fluorescence baseline, the inhibitors- argatroban (1 µM) or benzamidine (1 mM) were injected into the C18 capillary column and were eluted using an acetonitrile gradient from 5% to 70% in 60 min.

B

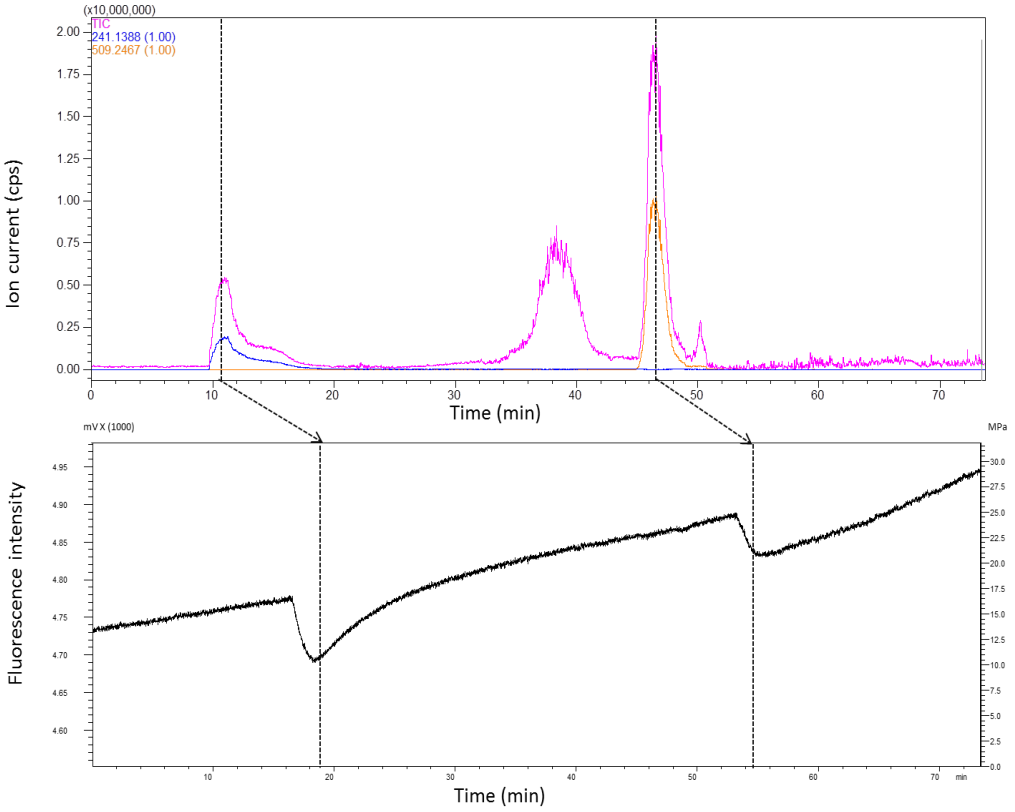


Figure 2.7. B. Online FXa assay coupled with mass spectrometry using a mixture of benzamidine and argatroban. Benzamidine that eluted from the nano-LC resulted in a negative peak at 18 min and was detected by the mass spectrometer at 10 min with an XIC of 277.1145 at 10 min. Argatroban that eluted from the nano-LC resulted in a negative peak at 54 min and was detected by the mass spectrometer at 47 min with an XIC of 507.2431.

Their respective *m/z* peaks at 10 min and 48 min in the mass spectrometer and bioactive negative peaks at 17 min and 54 min in the assay were observed.

Further, a mixture of argatroban (1 µM) and benzamidine (1 mM) was injected into the nano-LC and was eluted using the same acetonitrile gradient and their corresponding peaks in the mass spectrometer and in the online assay were observed (Fig. 2.7 B).

2.3.7. Identification of inhibitors in snake venoms with online assay

Previously, we have screened snake venoms for acetylcholine receptors antagonists using online methods. To evaluate the suitability of this method to identify FXa inhibitors from complex mixtures, we used the venom of the snake *Dendroaspis polylepis* spiked with argatroban. In addition to argatroban (50 min; 509.25 Da), three inhibitors from the snake venom were detected at 36, 38 and 42 min with masses of 6798.18 Da (*m/z* 1134.03; charge +6), 3492.76 Da (*m/z* 874.19; charge +4;) and 7208.58 Da (*m/z* 1202.43; charge +6) corresponding to inhibitor peaks in the online assay at 42, 44 and 48 min respectively (Fig. 2.8). These proteins can be purified based on their masses, and further characterized for structure, function and selectivity.

In conclusion, we have developed a platform for the identification of inhibitors of thrombin and FXa from a mixture of compounds. The robustness of the system was evaluated and optimized to deliver reliable data comparable to the traditional microtiter plate reader assays. This setup was used to identify FXa inhibitors from mixtures like snake venom.

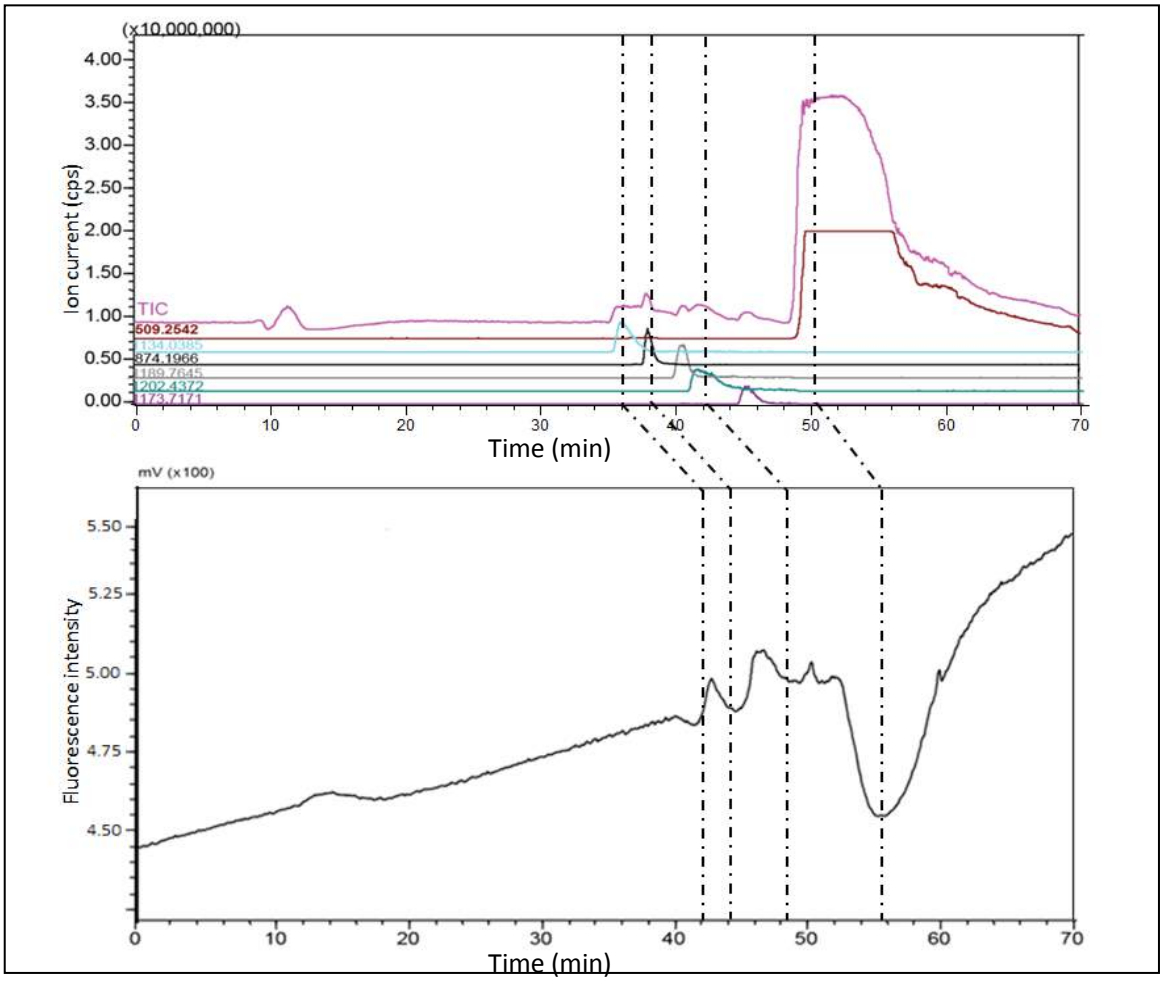


Figure 2.8. Identification of FXa inhibitors in *Dendroaspis polylepis* venom. Crude venom spiked with argatroban was injected to nano-LC. Argatroban was detected as a bioactive peak at 56 min with a mass of 509.25 Da at 50 min in the MS. In addition three other FXa inhibitors from the venom were detected at 42, 44 and 48 min in the online assay with masses of 6798.18 Da (*m/z* 1134.03; charge +6), 3492.76 Da (*m/z* 874.19; charge +4) and 7208.58 Da (*m/z* 1202.43; charge +6) at 36, 38 and 42 min respectively in the MS.

CHAPTER 3

Quantitative sialome of male and female *Dermacentor reticulatus*

3.1. Introduction

Ticks are a specialized group of obligate haematophagous ectoparasites which have evolved for millions of years and acquired potent pharmacologically active molecules in their saliva to disarm and counteract the host haemostatic system (Bowman et al., 2008). As a result, several novel compounds present in tick saliva and their roles as antihaemostatic compounds facilitating tick feeding have been studied. Therefore, the knowledge about the molecules present in tick saliva and their functions has accrued significantly over the past few years (Hovius et al., 2008). Hard ticks (Ixodidae) remain attached to their host for several days and throughout the feeding process, they infuse a cocktail of pharmacologically active molecules into the host blood. Due to their long term feeding patterns, ixodid ticks have also proven to act as excellent vectors pathogen transmission (Richter et al., 2013). Like the saliva from other hematophagous animals, tick saliva contains a wide range of physiologically active molecules that aid in the tick attachment to the host and facilitate blood feeding. Pathogens transmitted by ticks also utlilize the components of the tick saliva to facilitate their own transmission to other hosts (Francischetti et al., 2009). Although both male and female ixodid ticks are blood feeders, stark differences in the blood feeding patterns of the two sexes have been observed and well reported (Bowman and Sauer, 2004; Sauer et al., 1995). For example, the female ticks feed on larger volumes of blood, and grow to about 100 times of their unfed body weights, whereas the male ticks barely grow to more than 2-times of their unfed sizes. During this blood feeding phase, various morphological and

physiological changes occur in the salivary glands of the male and female sexes and their salivary components reflect these differences in the feeding habits of the two sexes. With the salivary glands increasing in mass and protein content by more than 25 times, the granular type II and type III acinar cells exhibit remarkable morphological changes. The type III acini are transformed into a tissue that mainly performs fluid transporting functions. The male-specific type-IV acinar cells have shown to increase in protein content selectively only if the female ticks are present on the host. These type-IV acinar cells are therefore hypothesized to produce proteins that transfer the male spermatophore into the females which involves copious amounts of saliva production by the males (Bowman and Sauer, 2004; Tan et al., 2015).

Blood feeding induces more gene expression of already expressed genes or expression of new genes that were not expressed earlier in the salivary glands. Salivary glands of female ixodid ticks have been shown to increase in mass and protein content by more than 25 times that enable the female tick to feed on the large volumes of blood (Šimo et al., 2004). The mass and total protein content of the ixodid male ticks also changes and new genes are expressed during feeding, but the magnitude of these changes are not to the same extent as those seen in the female ticks (Oaks et al., 1991). The small amounts of protein obtained from the salivary glands of these tiny ticks have largely hampered the identification of specific proteins that are up- or down-regulated during this feeding process (Bowman et al., 2008; Valenzuela, 2004).

Advancements in modern molecular techniques and the introduction of highly sensitive instruments has allowed the use small amounts of biological starting material for the complete profiling of molecules present in these complex mixtures. Transcriptomics and proteomics have become the two most common approaches that have facilitated the identification and isolation of a vast repertoire of transcripts and proteins from the tick salivary glands (Bowman et al., 2008). A quantitative profiling for the identification of specific proteins expressed at specific stages of tick feeding has not been done yet. The identification of these pharmacologically active molecules can lead to their development as potent antithrombotic agents, immune suppressors or as potential vaccine candidates for the control tick-borne diseases (Kazimírová and Štibrániová, 2013).

Dermacentor reticulatus is the second most important hard tick species in Europe after the common tick, *Ixodes ricinus* (Karbowiak, 2014). Commonly known as the ornate dog tick, *Dermacentor reticulatus* has a widespread occurrence across Western Europe, and feeds on the blood of vertebrate hosts like dog, cattle, sheep, horse, pig and human. *D. reticulatus* is the primary vector that transmits *Babesia canis* which is the pathogen for canine babesiosis (Halos et al., 2014). It is also a vector for tick-borne encephalitis virus, which is the most frequent tick borne viral disease in Europe (Wójcik-Fatla et al., 2011). The protein content and body weight of the completely engorged female *D. reticulatus* is reported to increase drastically compared to the unfed ticks (Vančová et al., 2010). Therefore, to identify novel, sex specific and feeding stage specific proteins in the male and female salivary gland extracts of *D. reticulatus,* we have

carried out a detailed the quantitative transcriptomic and proteomic profiling of the salivary gland extracts of males and females of this tick species.

3.2. Materials and Methods

3.2.1. Salivary gland extracts

Salivary gland and salivary gland extracts of adult male and female *D. reticulatus* were obtained from collaborator RNDr. Mirko Slovak from the Slovak Academy of Sciences, Slovakia. The ticks were reared under laboratory conditions (Slovak et al., 2002) in the Institute of Zoology, SAS, Slovakia. Ticks from the fourth breeding generation were used in the experiments (Slovak et al., 2002). Briefly, the ticks were maintained at 24 °C in desiccators filled with concentrated KCl solution, with 85-90% relative humidity and a photoperiod of 16:8 h (L:D). White New Zealand rabbits were used as hosts for both male and female ticks at all stages of feeding. The usage of animals in these experiments was approved by the State Veterinary and Food Administration of the Slovak Republic (928/10-221 and 1335/12-221). Ticks were fed on rabbits for pre-determined number of days, after which they were detatched from the rabbits and their salivary glands were dissected phosphate buffered saline (PBS, pH 7.2).

For the preparation of salivary gland extracts, the glands were homogenized in 150 µl of PBS, using a handheld homogenizer in a microcentrifuge tube. The homogenized mixture was then centrifuged at 15,000 rpm, and the supernatant was harvested as the salivary gland extracts. This was repeated a second time with the pellet obtained from the centrifuged sample, and the pooled with the supernatant from the first extraction. The salivary gland extracts were then freeze dried until further use.

For isolating RNA for transcriptomics, the salivary glands were immersed in RNAlater (Qiagen, Hilden, Germany) immediately after dissection. Salivary glands were kept in 4^0C for a minimum of 48 h to ensure penetration of RNAlater solution into the tissue. Thereafter, they were kept frozen in -30^0C until further processing.

3.2.2. Transcriptomics

3.2.2.1. cDNA library construction

Salivary glands from male and female ticks fed for the following time points were pooled as follows: Unfed, 1, 3 and 7 h, 1, 2, 3, 4, 5, 6, 7 and 8 days fed. Six male and six female ticks were used for each time point, except for 1 day where 12 ticks were used and separate libraries for male and female tick salivary gland extracts were generated. Salivary glands were dissected and the tissues were washed three times in the same buffer before being immersed into RNAlater (Qiagen) and stored at -30^0C. mRNA was isolated from the salivary glands using a Micro-FastTrack 2.0 mRNA isolation kit (Invitrogen, San Diego, CA) according to the manufacturer's protocol and fragmented using a Covaris E210 (Covaris, Woburn, MA) sonicator. Library amplification was performed using eight cycles to minimize the risk of over-amplification. Unique barcode adapters were applied to male and female libraries for their identification separately. Individual libraries were quantified using qPCR and pooled in an equimolar ratio before sequencing on a HiSeq 2000 (Illumina) with ver. 3 flow cells. Two lanes of the HiSeq 2000 were used and both libraries were run together in both lanes to avoid lane bias. A total of 20,144,555 paired ended sequences from adult females and 25,193,086

sequences from adult male ticks (100 nt long) were obtained. Raw reads were processed using RTA 1.12.4.2 and CASAVA 1.8.2.

3.2.2.2. Transcriptome assembly

Reads from the Illumina sequencing were assembled with the Abyss software (Garber et al., 2011; Simpson et al., 2009) with different k values (every even number from 50 to 96). In addition to Abyss, which sometimes tends to miss highly expressed contigs (Zhao et al., 2011), the reads were assembled using the Trinity assembler (Grabherr et al., 2011). The resulting assemblies were joined by iterative BLAST and cap3 assembler (Karim et al., 2011). Coding sequences (CDS) were extracted using an in-house automated pipeline, based on similarities to known proteins, or by obtaining coding sequences from the larger open reading frame (ORF) of the contigs containing a signal peptide. A non-redundant set of the coding sequences and their translated protein sequences were mapped to a hyperlinked excel spreadsheet. Signal peptide, transmembrane domains, furin cleavage sites and mucin-type glycosylation were determined by software analysis available from the Center for Biological Sequence Analysis, Denmark (Duckert et al., 2004; Nielsen et al., 1999). To map raw reads to the coding sequences and determine their sex bias, raw reads from each library were blasted to the coding sequences using BLASTN with a word size of 25 and allowing recoveries of up to three matches and matches with less than two gaps were used. The resulting BLAST file was compiled and the number of reads each CDS received from each library was determined, and the number of hits at each base of the CDS was counted. This allowed the calculation of the average CDS coverage,

maximum coverage, minimum coverage and the reads per kilobase per million (RPKM) values. These results were statistically tested by X^2 test (using the number of reads per CDS), and the results were reported as significant if the 'p' value was < 0.05. Phylogenetic analyses were done with ClustalX and Mega softwares. For quantitative analysis of the transcriptome, only CDS with RPKM values > 4 were considered (Garber et al., 2011; Wagner et al., 2012).

3.2.3. Proteomics

Mass spectrometric based quantitation of proteins expressed in the male and female salivary glands at different stages of feeding was carried out by using the isobaric tag for relative and absolute quantification (iTRAQ^{TM}) technique. iTRAQ^{TM} is a chemical labeling technique that allows multiplexing of protein samples and quantification of individually identified proteins from each of the multiplexed sample (Wiese et al., 2007). Four time points from both sexes (unfed, day 0; early stage of feeding, day 2; late stage of feeding, day 4; and fully fed, day 5 for males and day 6 for females) were used for this analysis.

3.2.3.1. Protein quantification

Freeze dried salivary gland extracts of male (80 ticks at day 0, 60 at day 2, 60 at day 4 and 50 at day 5) and female (70 ticks at day 0, 50 at day 2, 50 at day 3 and 30 at day 6) *D. reticulatus* at different stages of feeding were reconstituted in MilliQ water and filtered with a 0.22 µm Ministart filter (Sartorius Stedium Biotect, Goettingen, Germany). Protein quantification was performed by the Bradford assay. The diluted Bradford reagent concentrate (Bio-rad Laboratories, California, USA) (200 µl) of was mixed with the salivary gland extract protein

sample (10 µl) in a 96-well microtitre plate, and incubated at room temperature for 10 min. Absorbance was read at 595 nm with a microplate reader (Tecan M200, Männedorf, Switzerland). The protein concentration was estimated with a standard curve plotted using BSA. For all further proteomics experiments, samples were processed as technical replicates.

3.2.3.2. Tryptic digestion and iTRAQ labeling

30 µg (1 mg/ml) of salivary gland extracts from each feeding stage from male and female ticks was used as the starting material. All reagents used for digestion and labeling (including trypsin) were from 8-plex iTRAQ™ labeling kit supplied by AB SCIEX. Briefly, each sample was denatured and reduced by adding 1 µl denaturant (2% SDS), followed by the addition of 2 µl reducing agent (50 mM tris(2-carboxyethyl)phosphine (TCEP)), and incubated at 60 °C for 1 h. This was followed by alkylation of the cysteines by the addition of 1 µl cysteine blocking reagent (200 mM methyl methane thiosulfonate (MMTS)), and incubated at room temperature for 10 min. Trypsin was added to the mixture in a ratio of 6 µg trypsin : 30 µg protein and the mixture was incubated at 37^0C for 16 h, with light shaking.

Complete digestion of the sample was confirmed by running 0.5 µl of pre- and post-digested samples on a 12% (w/v) Mini-PROTEAN® TGX™ precast polyacrylamide gel (Biorad Laboratories, California, USA). 2X SDS loading dye (62.5 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% glycerol, 50 mM DTT, 0.01% (w/v) bromophenol blue) was mixed with the protein sample in a 1:1 ratio and heated at 95 ^{0}C for 10 min. After allowing the mixture to cool down to room

temperature, samples were loaded into the SDS-PAGE gel. Electrophoresis was carried out 160 V in a mini-PROTEAN® cell powered by the PowerPac™ Basic Power Supply (Bio-rad Laboratories, California, USA). Following electrophoresis, gels were washed briefly in MilliQ water and stained with a silver-stain kit (Thermo Fisher Scientific Inc., Massachusetts, USA) according to the manufacturer's protocol. These tryptic digested samples were used for further iTRAQ labeling.

Three separate iTRAQ™ runs for the analysis of male and female tick salivary gland extracts were performed (Male+female, female alone, male alone). 8-plex iTRAQ™ labels (113, 114, 115, 116, 117, 118, 119, and 121) were prepared freshly by adding 50 µl of isopropanol to each vial of the label. The contents of one iTRAQ™ reagent label vial were transferred to one tryptic digest sample tube. The different stages of feeding and iTRAQ™ labels with which they were labeled are summarized in Table 3.1. The pH of the solution was tested using a pH paper and was adjusted within the range of 7.5-8.5 by addition of the dissolution buffer. The organic concentration was maintained at a minimum of 60% for efficient iTRAQ™ labeling. The tubes were incubated for two hours at room temperature before the clean-up.

Table 3.1. Details of the three different iTRAQ™ runs along with the labels used.

iTRAQ label	Male+female	Female alone	Male alone
113	MD 0 (1)	FD 2 (1)	MD 0 (1)
114	MD 5 (1)	FD 4 (1)	MD 5 (1)
115	FD 0 (1)	FD 0 (1)	MD 2 (1)
116	FD 6 (1)	FD 6 (1)	MD 4 (1)
117	MD 0 (2)	FD 2 (2)	MD 0 (2)
118	MD 5 (2)	FD 4 (2)	MD 5 (2)
119	FD 0 (2)	FD 0 (2)	MD 2 (2)
121	FD 6 (2)	FD 6 (2)	MD 4 (2)

All experiments were carried out in duplicates and numbers 1 or 2 within the parantheses represent duplicate 1 and 2 respectively. (M: salivary gland extracts of male ticks; F: salivary gland extracts of female; D followed by the number represents the number of days for which the ticks were fed on the host before detatching and extraction of the salivary glands.)

3.2.3.3. Sample clean-up

Digested samples were desalted using a C-18 Sep-Pak plus cartridge (Waters, Massachusetts, USA). The Sep-Pak cartridge was conditioned with 10 ml 100% ACN, and then equilibrated with 10 ml 2% ACN. iTRAQ™ labeled sample mixture was diluted with 2% ACN to more than 10 times the volume (approximately 10 ml) and then loaded into the cartridge. A wash step was performed with 10 ml 2% ACN and the peptides were eluted with 10 ml 80% ACN. Sample was freeze dried prior to mass spectrometric analysis.

3.2.3.4. Mass spectrometry

The dried sample mixture was reconstituted with 30 µl of 5% ACN and 0.05% formic acid (FA) for LC MS/MS analysis. For each sample, 3 µg was injected into a microbore reverse phase column (75 µm × 150 mm, ReproSil-Pur C18-AQ, 3 µm, 120 Å (Eksigent, California, USA)) and eluted using mobile phase A as 2% ACN, 0.1% FA and mobile phase B as 98% ACN, 0.1% FA. The gradient (in terms of mobile phase B percentage) for the LC was as follows: 5% for 1 min, 5-12% for 2 min, 12-30% for 120 min, 30-90% for 1 min and held at 90% for 11 min, with a flow rate of 300 nl/min. MS analysis was performed with the TripleTOF™ 5600 system (AB SCIEX).

Data acquisition was performed with a mass range set at *m/z* 400-1250 and an accumulation time of 250 ms per spectrum. 20 precursor ions were subjected to MS/MS analysis, with dynamic exclusion enabled for a duration of 15 s. Proteins from the three iTRAQ^{TM} runs were identified and the labels were quantified using the ProteinPilot™ Software v.4.5 (AB SCIEX) using the Paragon™ Algorithm (Hultin-Rosenberg et al., 2013). ID focus was set to biological modification, with cysteine alkylation by iodoacetamide, digestion by trypsin, and search effort set to thorough. The database used for the search was the set of proteins derived from the transcriptome of *D. reticulatus*, containing 37,036 protein sequences. The false discovery rate (FDR) analysis was performed with the Proteomics System Performance Evaluation Pipeline (PSPEP) feature in the ProteinPilot™ software using a decoy database search strategy in which the protein database sequences were reversed and searched against. The reported FDR for both male and female

proteomes at different stages of feeding with less than 1% based on the global FDR fit. Only proteins with at least two peptides with >95% confidence were taken for further analysis. For iTRAQ^{TM} quantification analysis the total intensity of spectra for each label from each run was calculated and a manual normalization was carried out to ensure equal loading. The integration of spectra from one label was considered as the reference label and the total intensity of spectra from other labels were divided by the total intensity of the reference label.

3.3. Results

3.3.1. Transcriptome of *Dermacentor reticulatus*

In an attempt to identify novel proteins, the transcriptomes of male and female salivary gland extracts were generated. The transcriptomes of male and female salivary gland extracts represented the entire repertoire of proteins expressed at different stages of feeding in the two sexes in the highly specialized salivary glands of *D. reticulatus*. A statistical analysis of the assembled and annotated transcriptome data also helped us to compare the levels of expression of different proteins in the two sexes of this tick.

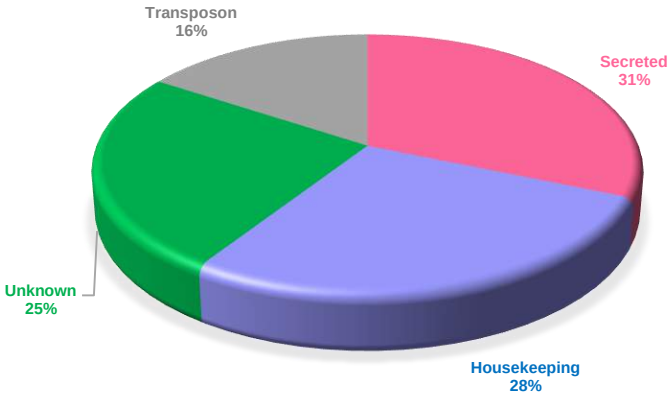
A total of 49,177 coding sequences (CDS) were extracted from the combined assembly of 35,982,329 and 40,274,259 paired-end sequences (100 nt long) from male and female samples, respectively, out of which 26 CDS resulting from a total of 30,766,591 had matches to the rRNA database. These reads were separated from the main transcripts to avoid rRNA induced bias in the fold change and RPKM calculation. The remaining 49,152 were sorted based on their RPKM values and the 37,036 CDS that had RPKM values >4 and were used for quantitative analysis. The CDS were classified into four main classes: housekeeping proteins (H), secreted and putative secreted proteins (S), transposable elements (TE), and proteins with unknown function (U) (Fig. 3.1 and Table 3.2.).

In summary, a total of 6,145 (27.94% reads) CDS were classified into the housekeeping (H) class; 10,328 CDS (31.3%) to the secretory (S) class and, 4,688

(15.80% reads) CDS including fragments, belonged to the transposon (TE) class and a total of 15,875 CDS (24.94% reads) were classified as unknowns (U) class CDS with unknown functions and these CDS were not analyzed further. All CDS, their similarity hits to different databases used for classification are presented as a spreadsheet in the supplementary files DVD, with the file named "DR transcriptome".

Table 3.2. Functional classification of coding sequences from the transcriptome of *D. reticulatus*

Class	Number of CDS	Number of reads	Percentage of total reads (%)
Housekeeping	6,145	12,667,655	27.94
Secreted	10,328	14,194,752	31.30
Unknown	15,875	11,308,847	24.94
Transposon	4,688	7,166,387	15.80



Major families of secreted proteins

- Enzymes
- IgG binding proteins
- Lipocalins
- Proteinase inhibitor domain
- DA-p36 family
- Mucins
- Ixodegrins
- Antigen-5 binding proteins
- Tick specific proteinase inhibitors
- Glycine rich superfamily

Figure 3.1. Components of the sialotranscriptome of *D. reticulatus*. The sialotranscriptome was classified into four main classes - housekeeping (H), secreted (S), transposable (T) and unknown (U). Major families of secreted products are listed in the box.

For each CDS, the numbers of reads derived from the male and female libraries, along with a chi-squared test, and RPKM values for the significance of the fold differences in expression in their read counts are tabulated.

3.3.1.1 Housekeeping proteins

A total of 6,145 CDS were classified into the H class and were further subclassified into 23 families based on their function. Among the 23 subclasses, proteins associated with signal transduction, nuclear regulation, transporters and channels, transcription and translation were proteins which were most highly expressed. Since the salivary gland is a specialized organ associated mainly with glandular functions, polypeptide secretion and osmotic regulation, proteins with transcription, translation and transport functions are expected to be expressed in high numbers (Alarcon-Chaidez et al., 2007; Karim et al., 2011; Nene et al., 2002). The high quality and number of reads generated by a deep sequencing technique allowed the detection and annotation of transcripts not usually detected by usual sequencing techniques. These included proteins belonging to minor subclasses like extracellular matrix, nuclear transport, storage and several others.

Table 3.3. Functional classification of extracted CDS of housekeeping proteins from the *D. reticulatus* sialotranscriptome

Class	CDS	Associated Reads	% Total reads
Transcription machinery	372	406100	0.895724
Transcription factors	207	207881	0.458517
Storage	13	22361	0.049321
Soluble carrier protein	43	45132	0.099546
Translational machinery	351	1837438	4.052787
Oxidant metabolism/detoxification	102	102586	0.226271
Protein export	327	575762	1.269943
Protein folding/ chaperone	88	217984	0.480801
Transporters and channels	426	564620	1.245367
Proteasomal machinery	259	408963	0.902039
Signal transduction	961	1071215	2.36275
Extracellular matrix	185	2881704	6.356096
Lipid metabolism	359	439659	0.969744
Energy metabolism	176	448035	0.988219
Intermediary metabolism	34	22377	0.049356
Cytoskeletal elements	346	588616	1.298294
Nuclear regulation	801	1162468	2.564024
Immune related	94	246061	0.54273
Protein modifying	301	407648	0.899138
Amino acid metabolism	103	179062	0.394952
Nuclear transport	42	58453	0.128928
Carbohydrate metabolism	253	459226	1.012902
Nucleotide metabolism	302	314304	0.693252

The transcription machinery proteins mainly constituted of RNA polymerases, transcription initiation and elongation factors, zinc finger proteins and while most of these were expressed in similar numbers in the male and female salivary glands, certain transcripts were expressed with very high RPKM values (>20) denoting that these functions were carried out mainly by these proteins, and for certain transcripts, the proteins were exclusive to males of females. This indicated that although the overall number of CDS were approximately equal in male and female transcriptomes, certain transcripts were expressed in a sex-specific manner (transcripts SigP-620365 and SigP-602832 had high male RPKM values of 199 and 219 respectively which were more than 2 times higher than the RPKM values for the female transcripts, whereas transcripts DrIxod-559348 and DrIxod-220182 had female RPKM values of 25 and 48 respectively which were almost 3 times higher than the RPKM values for the corresponding male transcripts). Similar observations were made for specific transcripts from other classes such as those with translation and export functions, where a sex-specific expression of particular transcripts was noted.

Transcripts encoding for proteins associated with carbohydrate metabolism, lipid metabolism, nucleotide metabolism, protein modification and cytoskeletal elements were the next most abundant subclasses of transcripts within the H class. The lipase precursor transcripts (DrIxod-562420 and SigP-606455) with male RPKM values >200 were overexpressed in males (>150 times) compared to the female transcripts. The glycolate oxidase (SigP-589640) was overexpressed in females (>100 times). This overexpression of proteins involved

in carbohydrate synthesis was observed in females which corroborated with the increased gluconeogenesis functions of the female salivary glands owing to the increased volumes of blood ingested by the females. The protein modifying fucosyltransferases were observed to be overexpressed in males (>15 fold) compared to females. This could be attributed to the higher amounts of glycoproteins and lubricating proteins produced in the male saliva which assist them in sexual reproduction (Feldman-Muhsam et al., 1970). These examples hence suggested that in addition to their direct roles in housekeeping functions, these proteins could be also be associated with the functions of the salivary glands and play sex-specific roles.

The other subclasses of protein families which were detected in low numbers, such as those associated with oxidant metabolism and detoxification may also be associated with direct salivary gland functions and assist the tick in feeding and osmoregulation. Enzymes and detoxification proteins such as glutathione-s-peroxidases, sulfotransferases, dehydrogenases and cytochrome-P450 which were members of these classes could detoxify host oxidants and participate in prostaglandin synthesis which are important functions associated with tick saliva (Szabó et al., 2007). Of special interest were storage proteins, such as vitellogenins and transferrins, which were specifically overexpressed in females (>3 times overexpression in the total number of reads). These storage proteins are directly associated with female salivary functions such as salivary gland degeneration and ovarian development (Friesen and Kaufman, 2009). The extracellular matrix proteins like cell wall mannoproteins and collagen binding

proteins (>200 times overexpressed in males) were characteristic of the male salivary transcriptome. These proteins and glycoproteins play important roles in male tick, feeding, immunity and sexual reproduction (Donohue et al., 2008).

3.3.1.2. Secreted proteins

The 10,329 reads which were classified as secreted proteins were further classified into 63 subclasses based on their function (Karim et al., 2011; Ribeiro and Francischetti, 2003; Ribeiro et al., 2011; Tan et al., 2015). These 63 families could be further subdivided into ubiquitous and tick specific proteins.

The ubiquitous families of proteins constituted of proteins which similar to proteins that were found in the salivary transcriptomes of other ticks. These families included enzymes such as metalloproteases, dipeptidyl peptidases, ADAMs (or ADAMTs), M13 peptidases, lipases, serine carboxypeptidases, zinc carboxypeptidases, serine proteases, apyrases and ribonucleases. The non-enzymatic proteins included in ubiquitous families of proteins were the proteinase inhibitor families like serine protease inhibitors, cystatins, madanins, kunitz, kazal and trypsin inhibitor like (TIL) domains. In addition, the immunomodulatory proteins like lipocalins, lysozymes, defensins, ixoderins and other similar proteins which were also a part of the ubiquitous family of proteins were found in the *D. reticulatus* transcriptome similar to the salivary trancriptomes of other ixodid ticks.

Tick specific protein families included those with highly specialized functions associated with the tick salivary glands and tick feeding. These families

included glycine-rich superfamily proteins (cement protein), mucins (lubrication and sexual reproduction), evasins (bind and neutralize host cytokines), DA-p36 (immunosuppressor), japanins (dendritic cell modulators), ixodegrins (potential antiplatelet agents) and tick-specific protease inhibitors such as carboxypeptidase inhibitors (Hajnická et al., 2001; Vančová et al., 2010; Wang and Nuttall, 1995). Some of these protein families are discussed in detail further.

Table 3.4. Functional classification of extracted CDS of housekeeping proteins from the *D. reticulatus* sialotranscriptome

Class	CDS	Associated Reads	% secreted
Metalloprotease	103	174024	1.225974
ADAMs	6	3269	0.02303
M13	42	68779	0.484538
Dipeptidyl peptidase	9	10063	0.070892
Lipocalin	366	1311467	9.239098
Putative basic tail	1	156	0.001099
18.3 kDa protein	8	32195	0.226809
Carboxypeptidase inhibitor	15	15415	0.108596
Serpins	6	3183	0.022424
Putative metastriate	2	2452	0.017274
8.9 kDa protein	131	231426	1.630363
28 kDa protein	10	2330	0.016415
IgG binding protein	8	363536	2.561059
23 kDa protein	6	12115	0.085348
Serine protease	15	17884	0.12599
Male specific salivary serine protease	1	938	0.006608

Serine carboxypeptidase	34	60835	0.428574
Zinc carboxypeptidase	5	6403	0.045108
Catalytically inactive chitinase like lectins	2	695	0.004896
Apyrase/ ribonuclease	7	4283	0.030173
Antigen 5 family/CAP family	7	41654	0.293446
Vasodilators	2	474	0.003339
alpha macroglobulin	4	4984	0.035112
Cystatins	7	14332	0.100967
Tick til	51	127974	0.901559
Putative metastriate one of each	1	46	0.000324
Defensins	24	57826	0.407376
Lysozyme	3	1292	0.009102
Peptidoglycan recognition protein	3	1408	0.009919
Salivary lipase	2	1450	0.010215
Galectin	3	4858	0.034224
Ixoderin/ficolin	5	14237	0.100298
Ixostatin	2	8840	0.062277
Thyropin	1	395	0.002783
Putative 8 kDa	2	995	0.00701
TGF beta induced	1	9526	0.067109
Ixodegrins	6	8452	0.059543
Dap36	31	63477	0.447186
Evasin	21	18792	0.132387
Kunitz	167	363038	2.557551
Kazal	5	663	0.004671
Similar to chymotrypsin/elastase inhibitor	19	70855	0.499163
Serine protease inhibitors	20	7220	0.050864

Glycine rich superfamily	124	4243216	29.89285
Mucins	68	119697	0.843248
Madanin	2	1774	0.012498
Microplusin	5	9599	0.067624
Insulin like growth factor binding proteins	5	1039	0.00732
22.5 secreted histamine binding protein	1	3645	0.025679
Fed tick salivary protein	1	260	0.001832
Salivary gland expressed bhlh	1	557	0.003924
Ribonuclease	1	490	0.003452
Salivary selenoprotein precursor	1	2052	0.014456
Phospholipase A2	2	1871	0.013181
Der and secreted protein	10	11732	0.08265
Hypothetical acid tail secreted protein	4	52443	0.369453
Conserved secreted	105	393667	2.773328
Japanin	8	3621	0.025509
Secreted peptide precursor	479	1228358	8.653607
Hypothetical conserved secreted	37	156236	1.10066
Hypothetical secreted	8301	4808961	33.87844
Other secreted proteins that could not be classified	10	11298	0.079593

3.3.1.2.1. Enzymes

A total of 229 CDS were classified as enzymes in the S class. 103 transcripts out of these were metalloproteases and more than half of these 103

metalloproteases were identified to be similar to the Reprolysin_2 PFAM domain. This family of metalloproteases found in snake venoms and salivary transcriptomes from other ixodid ticks are a highly specialized group of enzymes (4 of which contained a thrombospondin like motif, hence called ADAMTs) and are involved in tissue remodelling, processing of procollagen and von Willebrand factors, imflammation, angiogenesis and cell migration (Apte, 2004). Distinct sex-specific expression of certain transcripts (Transcripts SigP-585676 and SigP-560950 were overexpressed in females while DrIxod-590135 was overexpressed in males).

M13 domain/neprilysin proteases, are transmembrane metalloendopeptidases and are zinc dependent enzymes involved in the inactivation of peptide hormones. In tick saliva, these metalloproteases could function by degrating inflammatory peptide mediators like cytokines, bradykinin or anaphylatoxins from the hosts. A total of 42 CDS were associated with these metalloproteases, and one transcript (SigP-620954) was >13 times over expressed in the male transcriptome compared to the females.

Apyrases are enzymes which hydrolyze ATP and ADP to AMP and hence suppress inflammation and platelet aggregation. Apyrases derived from the saliva of hard tick such as *I. scapularis* and the soft ticks *O. moubata* and *O. savignyi* have been previously shown to inhibit ADP and collagen induced platelet aggregation (Mans et al., 1998; Ribeiro et al., 1991). Apyrases usually possess a signature signal peptide which causes the protein to be transported to the extracellular environment, and a carboxy terminus in which a

phosphoglycoinositol (GPI) anchor fixes the protein to the extracellular side of the cell membrane (Ogata et al., 1990). Apyrases from mosquitoes and ticks belong to the 5′ nucleotidase family and lack the carboxyterminal membrane spanning domain that links them to a phosphoglycoinositol membrane anchor. In addition to antihemostatic functions, apyrases from the mosquito *Anopheles gambiae* have also shown to be involved in the pathogen transmission of several pathogens (Reno and Novak, 2005). Seven transcripts encoding for apyrases were found in the *D. reticulatus* transcriptome and all these 7 transcripts were over-expressed in the males indicating the sex bias.

Several serine proteases from the sialomes of other ixodid ticks like *A. maculatum R. pulchellus, A. variegatum* have been reported, but their detailed structure activity relationships are not yet known (Karim et al., 2011; Nene et al., 2002; Tan et al., 2015). In the tabanids (horsefly), serine proteases have shown to possess fibrinolytic activities (Xu et al., 2008). Although only 15 transcripts corresponding to serine proteases were identified in the *D. reticulatus* transcriptome, the sex-specific expression of transcripts is well pronounced in these 15 transcripts. The overall number of reads for male transcripts was about 5 times more as compared to the female transcripts, and two transcripts in particular, DrIxod-592189 and DrIxod-587803 were 80 and 161 times respectively overexpressed in males compared to the females. This difference in the amounts of serine proteases expressed the male and female salivary glands has also been observed for other animals, like *Drosophila*, in which they are a part of the male specific seminal fluid and these serine proteases are activated during

mating (LaFlamme et al., 2012). The transfer of these male specific serine proteases to females during mating brings about important changes in the mated females like egg production and reduced receptivity to remating (Avila et al., 2011). Serine proteases have also been shown to be present in the seminal fluid of other organisms such as human and *C. elegans* where they are known to affect male fertility (Smith and Stanfield, 2012; Veveris-Lowe et al., 2007). Therefore these male specific serine proteases in ticks could also be hypothesized to play significant roles in sexual reproduction.

Serine carboxypeptidases are a family of enzymes within the S class which are distinctly overexpressed in males than in females (>5 times) in the *D. reticulatus* transcriptome. A similar sex-specific pattern of expression of serine carboxypeptidases has also been noted in the salivary gland transcriptomes of other hard ticks. Despite this known sex-specific expression pattern, the role of this family of enzymes in tick feeding or tick reproduction is not yet known.

Carboxypeptidases were identified in both males and females of the *D. reticulatus* transcriptome and these carboxypeptidases could function in immune suppression of the host as observed in the saliva of *I. dammini* (Ribeiro et al., 1990b).

3.3.1.2.2. Proteinase inhibitor domains

To counter the host blood coagulation proteases, ixodid ticks produce enormous amounts of protease inhibitors in their saliva. Some of the proteins families with protease inhibitor domains that were abundantly expressed in the

saliva of *D. reticulatus* were kunitz, kazal, serpins, cystatins, thyropins and trypsin inhibitor like (TIL) proteins.

3.3.1.2.2.1. Kunitz type serine protease inhibitors

The Kunitz domains are small (50-60 residues), disulfide-bonded molecules found in a variety of species. The typical kunitz domain is restrained by three disulfide bonds, but certain proteins containing unique pattern of disulfide bonds in the Kunitz domain have also been studied (Krowarsch et al., 2003). Endogenous serine protease inhibitors (e.g. TFPI) which possess the Kunitz domain target a large number of blood coagulation enzymes for efficient control of haemostasis (Wesselschmidt et al., 1993). Kunitz domain of protease inhibitors have shown to be evolved in large numbers in the saliva of haematophagous animals (Corral-Rodríguez et al., 2009). Whereas most proteins with the Kunitz domains inhibit serine proteases, some proteins have also been identified with functions that block ion channels (Schwarz et al., 2014). Proteins containing multiple Kunitz domains have also been characterized and have been named as monolaris, bilaris, trilaris and so on corresponding to the number of Kunitz domains. The single Kunitz domain containing ornithorodin and tick anticoagulant peptide are well studied examples of high affinity thrombin and FXa inhibitors respectively (Lim-Wilby et al., 1995; van de Locht et al., 1996) from the soft tick genera *Ornithodorus*. The two Kunitz domain containing protease inhibitor family is typified by ixolaris from *I. scapularis* which typically inhibits the extrinsic tenase complex (Francischetti et al., 2002).

A total of 167 CDS corresponding to Kunitz domain containing proteins were identified in the *D. reticulatus* transcriptome. These transcripts were further classified and grouped into monolaris, bilaris and trilaris subclasses based on the number of cysteines and the observed pattern of disulphide bonds. Several other transcripts that could not be classified into one of these subclasses were grouped separately and analyzed. Among these sequences, several transcripts contained multiple Kunitz domains and their sequence lengths were more than 500 residues indicating they might possess over 10 tandem Kunitz domains. A sex specific expression pattern in a majority of these Kunitz domain containing proteins was observed. Some of these observations are discussed below.

Monolaris

The monolaris (one Kunitz domain) group of proteins was studied in detail. A total of 62 CDS were classified as monolaris proteins. In a typical Kunitz domain, the protein structure is held by three conserved disulphide bonds: C1-C6, C2-C4 and C3-C5. Interestingly, none of the monolaris identified in the *D. reticulatus* transcriptome contained this typical disulphide pattern of the usual Kunitz domain. Twenty two CDS in the monolaris subclass contained 2 cysteines in addition to the typical 6 cysteines (Fig 3.2 A and B). Since the two extra cysteines in this group of monolaris may form an extra disulphide bond at the N-terminus of the molecule, this monolaris family of proteins may have a fold different than the other studied monolarises and it could target a unique serine protease with a novel mechanism. Three CDS were >10 times overexpressed in the male ticks while 8 CDS were > 10 times overexpressed in the female ticks.

Four of these CDS were almost exclusive to the females with their expression levels >1000 times greater in the females compared to the males. This observation supports the hypothesis that different sets of unique inhibitors may be expressed in the males and females. (There is a chance that the N-terminal cysteines might form a disulphide. In addition to the extra cysteines, these transcripts also contain a long (7 residues insertion) between C2 and C3 of the typical Kunitz. this longer sequence might position the second disulphide in a different way allowing the formation of an extra disulphide at the N-terminus.)

Bilaris

Bilarises are proteins which contain two Kunitz domains in tandem. In the *D. reticulatus* transcriptome, 54 CDS were classified as the bilaris subclass of proteins. Eight out of these CDS contained 10 cysteines in a specific pattern and were overexpressed in females. A group of 12 CDS which contained 14 cysteines were overexpressed (5-20 times) in the male ticks (Fig. 3.3.).

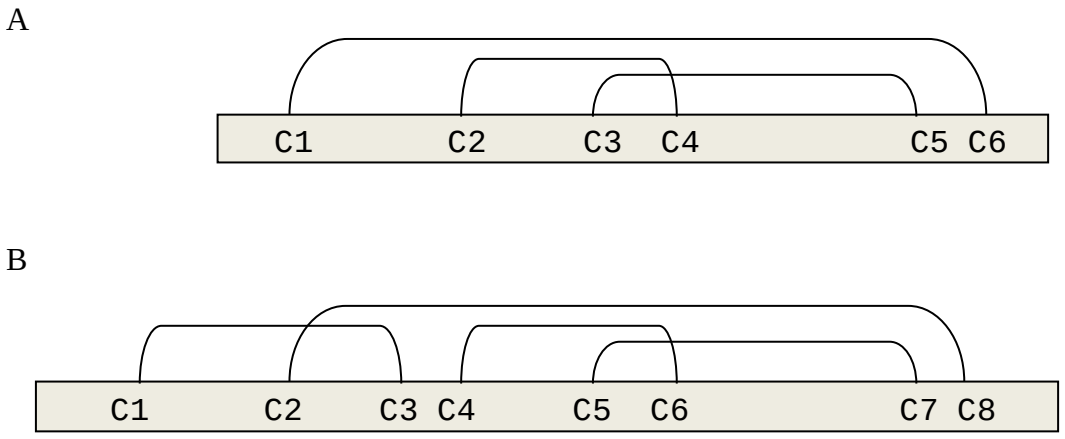


Figure 3.2. Comparison of disulphide bonding pattern of typical Kunitz domain with novel monolaris subclass found in *D. reticulatus* sialotranscriptome. A. 6 cysteines of the typical Kunitz domain form 3 disulphide bonds (C1-C6, C2-C4, C3-C5). **B**. 8 cysteines of the novel monolaris subclass form 4 disulphide bonds (C1-C3, C2-C8, C4-C6, C5-C7).

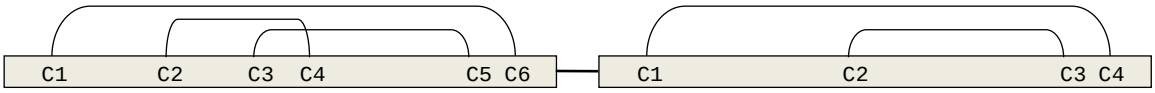


Figure 3.3. Disulphide bonding pattern of novel bilaris subclass found in *D. reticulatus* sialotranscriptome. 6 cysteines of the first typical Kunitz domain form 3 disulphide bonds (C1-C6, C2-C4, C3-C5). 4 cysteines of the second domain for 2 disulphide bonds (C1-C4, C2-C3).

Trilaris

Proteins containing 18 to 22 cysteines were classified as trilaris proteins. A total of 19 CDS were classified as trilaris proteins. Five CDS which had 18 cysteines with three typical Kunitz domains were separately grouped as they were slightly overexpressed in the males.

3.3.1.2.2.2. Kazal

The two kazal domain containing rhodniin from the assassin bug *Rhodnius prolixus* is a well studied example of a kazal inhibitor from the saliva of a haematophagous animal (van de Locht et al., 1995). A total of 5 transcripts encoding for proteins containing kazal domain were detected in the transcriptome of *D. reticulatus*. these transcripts did not show significant sequence identity with rhodniin. Four transcripts of these contained 2, 3, 6 or 8 cysteines, while one transcript contained 20 cysteines indicating the possibility of more than one tandem kazal domains

3.3.1.2.2.3. Serpins

The serpin (for Serine Protease Inhibitor) class of proteins are a common class of protein inhibitor family found in many species (*Ixodes scapularis* and *Rhipicephalus microplus*) and are physiological inhibitors of blood coagulation enzymes that control haemostasis. Serpins from saliva of other haematophagous arthropods such as *R. microplus* and *I. ricinus* inhibiting trypsin and other coagulation cascade enzymes have been previously reported (Chmelar et al., 2011; Rodriguez-Valle et al., 2012). In our studies, 26 transcripts encoding for

serpins from the salivary gland extracts of *D. reticulatus* were identified. A clear sex bias in the expression pattern of serpins was observed where female reads were about two times more abundant than male reads. Five out of the 26 transcripts were >10 times overexpressed in the females compared to males.

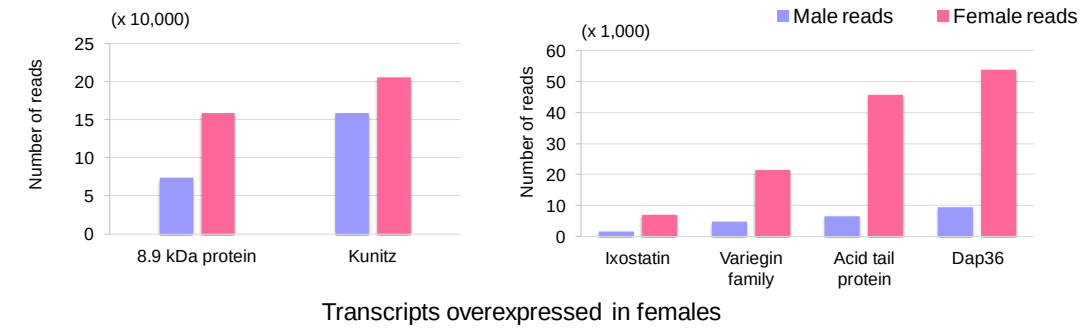
3.3.1.2.2.4. Cystatins

Cystatins are cysteine protease inhibitors, inhibit peptidases belonging to the legumain and papain family (Yadav et al., 2013). Salivary cystatins (sialostatin L and sialostatin L2) from *I. scapularis* have shown to inhibit cathepsin L and cathepsin S and have shown to have anti-inflammatory and immunosuppresive functions (Kotsyfakis et al., 2007). Seven transcripts encoding for cystatins were identified from the transcriptome of *D. reticulatus*. The cystatins were clearly >3 times overexpressed in males as compared to the females. This could be correlated to the known observations that cystatins play an important role in sexual reproduction in the male ticks, as high levels of cystatins have also been found in seminal fluids of other organisms (Yadav et al., 2013).

3.3.1.2.2.5. TIL domain containing proteins

The trypsin inhibitor-like (TIL) domain containing proteins are protease inhibitors found in different organisms. TIL containing proteins found in ixodid ticks have found to possess anti-trypsin and anti-elastase like activities (Fogaça et al., 2006). In the *D. reticulatus* transcriptome, 51 TIL domain containing proteins were identified, and the expression patterns of several of these were sex biased, and one transcript in particular (DrIxod-625451) was 11 times overexpressed in males.

A



B

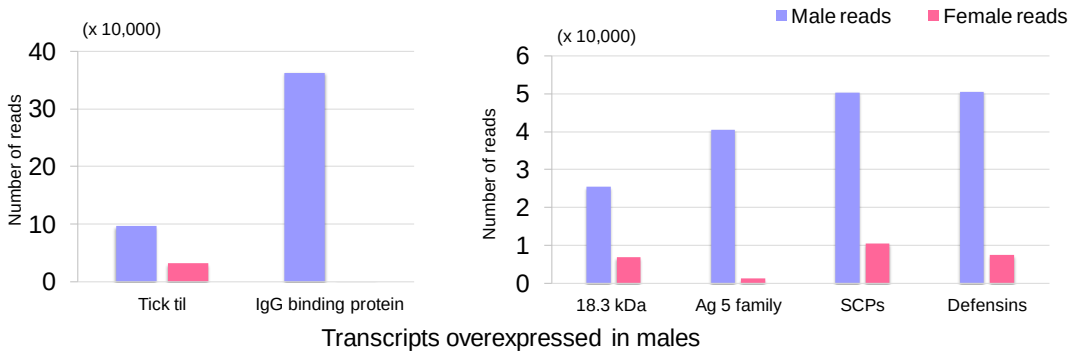


Figure 3.4. Transcripts selectively overexpressed in A. female and B. male sialotranscriptome of *D. reticulatus*. 8.9 kDa protein, Kunitz type inhibitors, ixostatins, variegin family proteins, acid tail proteins and DA-p36 were the proteins that were mainly overexpressed in the female salivary gland extracts. Tick til, IgG binding proteins, 18.3 kda, Ag5, SCPs and defensins were the families that were overexpressed in male salivary gland extracts.

3.3.1.2.3. Immunity associated proteins

Lysozymes, peptidoglycan recognition proteins, galectins, ficolins and ML domain containing proteins were some of the immunity associated proteins found in the salivary gland transcriptomes of *D. reticulatus*. These protein families which are produced in the salivary glands during tick feeding are infused into the host blood to modulate the host immune system to ensure successful tick feeding (Johns et al., 2001).

3.3.1.2.4. Lipocalins

Lipocalins were one of the largest groups of secreted proteins found in the salivary gland extracts of *D. reticulatus*. A total of 365 transcripts encoding for lipocalins were found out of which 57 were female specific and 46 were male specific in their expression pattern. Lipocalins are a unique group of structurally similar proteins with barrel-like structures that are known to carry hydrophobic ligands and are associated with transport functions across the membranes (Flower et al., 2000). Despite their similarity in structures, tick lipocalins are known to be associated with a diverse set of functions. Tick lipocalins have been reported to be associated with functions like scavenging biogenic amines, anti-complement, anti-clotting and immunoglobulin binding activities, and mediators of inflammation (Mans et al., 2008; Sangamnatdej et al., 2002).

Japanin

Japanins are a group of glycosylated lipocalins which were first identified in the salivary glands of *Rhipicephalus appendiculatus* . Japanins modulate the

dendritic cell (DC) activity by altering their expression of co-stimulatory and co-inflammatory transmembrane molecules and their secretion of pro-inflammatory, anti-inflammatory and T cell polarising cytokines (Preston et al., 2013). Japanin like sequences have been identified in the transcriptomes of *D. andersoni* and *R. microplus*. In the *D. reticulatus* transcriptome, 8 CDS (which did not fall in the lipocalin group) were identified as japanin sequences. A neighbor joining tree of these sequences with the native japanin sequence from *R. appendiculatus* and japanin like sequences from *D. andersoni* is shown in Fig 3.5.

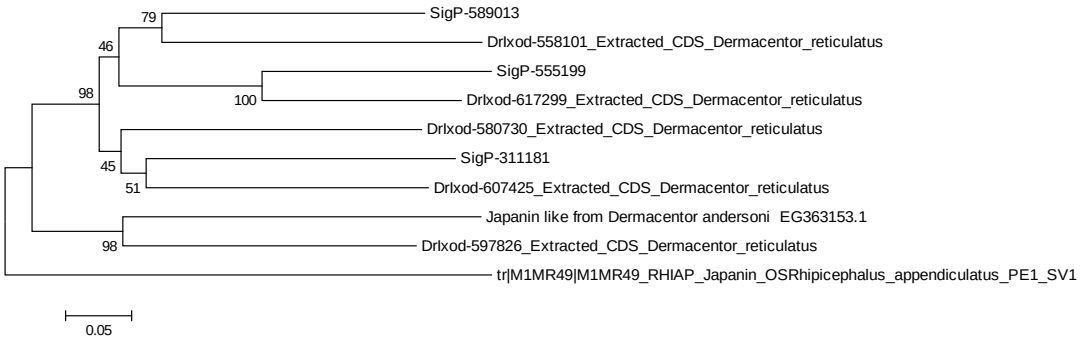


Figure 3.5. Neighbor joining tree of japanin like sequences from different ticks. NJ tree of japanin with japanin like sequences from the sialotranscriptome of *D. reticulatus* and *D. andersoni*.

3.3.1.2.5. DA-p36 family

The 36 kDa immunosuppressor protein from *D. andersoni* was shown to possess Ig-G binding properties and hence modulate host immune responses (Bergman et al., 2000). This family of proteins has also been reported to be found in transcriptomes of other ticks and 31 transcripts were identified in *D. reticulatus*. Out of the 31 transcripts, 13 transcripts were exclusively

overexpressed in females (some of these transcripts were >800 times overexpressed in females than in males).

3.3.1.2.6. Immunoglobulin G-binding proteins

This specialized family of proteins found in ticks bind to host IgG to suppress an immune reaction. When the tick feeds on host blood, host antibodies pass through the tick midgut into the haemolymph, and antibodies specific for tick antigens and cause damage to the ticks (Wang and Nuttall, 1995). The IgG binding proteins produced in tick saliva bind to the host IgG and recycle them back to the host thus preventing IgG mediated immune response. Male specific IgG binding proteins have been previously reported in *Rhipicephalus appendiculatus* and these have been shown to help the gregarious feeding of the female ticks (Wang and Nuttall, 1995). Previous observations have suggested that like other tick species, IgG binding proteins might be produced exclusively in the male *D. reticulatus* (Šimo et al., 2004). In our analysis we identified 8 transcripts encoding for IgG binding proteins, however these 8 transcripts were mapped from a total of 363536 reads, which were almost exclusively from the male ticks. The RPKM values for male reads for three of these transcripts > 1000 (which is well above the RPKM average for a moderately expressed proteins that have RPKM values ~10). This indicated that these three transcripts were expressed in very high numbers in the male salivary glands. These transcripts were >200 times overexpresssed in the male salivary glands compared to the females.

3.3.1.2.7. Glycine rich superfamily

This family of proteins which includes cuticle proteins, collagen, small GGY peptides, and large GGY peptides function as tick cement proteins and are associated with immunity. The metastriate ticks which have smaller mouthparts produce copius amounts of cement proteins that forms a cement cone and attached the tick to the host (Francischetti et al., 2009). This family of proteins was found to be about 3 times overexpressed in the males compared to the females. This could result from the male ticks producing large amounts of cement proteins because of their intermittent feeding nature (and they have to attach to the host multiple times) whereas the female ticks which feed continuously produce smaller amounts to produce the cement cone only once that helps them in attaching to the host.

3.3.1.2.8. Mucins

Mucins are serine- and/or threonine-rich proteins which are glycosylated at N-acetyl -galactosamine residues. These proteins function in tick feeding by coating and lubricating the chitinous feeding mouthparts. Proteins from this family have a shown to possess a distinct chitin binding domain in addition to the Ser/Thr-rich domain. In the *D. reticulatus* transcriptome, 68 transcripts encoding for mucins were identified and one transcript in particular was 136 times overexpressed in males compared to the female ticks.

3.3.2. Proteome of *Dermacentor reticulatus*

Isobaric tag for relative and absolute quantification (iTRAQ^{TM}), the chemical labeling technique based on stable isotopes that allows multiplexing was

used for the detection of sex-specific and feeding stage specific proteins from the salivary proteomes of *D. reticulatus*. Briefly, in the iTRAQ™ technique, peptides are labeled at their primary amine groups with a stable isotope label using the N-hydroxysuccinimide (NHS) chemistry (Fuller and Morris, 2012; Hultin-Rosenberg et al., 2013). In an 8-plex iTRAQ™ (AB SCIEX), up to 8 different protein samples can be labeled with the 8 isobaric tags. Following labeling, identical peptides from different protein samples with different labels have the same physicochemical properties in an LC-MS and are selected for fragmentation as a single precursor ion. During the fragmentation, the isobaric labels from the peptides are released and identified with different masses in the MS spectrum (113, 114,...,119, 121) (Karp and Lilley, 2007; Pütz et al., 2005). Therefore a relative quantification of different amounts of proteins present in different samples can be done by comparison of intensities of the different isobaric labels.

Since tick size increases drastically during feeding, it has been hypothesized that the composition of the salivary gland proteins undergoes a feeding induced change and different sets of proteins are synthesized during different stages of feeding (Xu et al., 2008). The increasing body weight is reported to be brought about by an increase in total protein content of the salivary glands in different ixodid ticks (Kubes et al., 1994; Tan et al., 2015). To identify specific proteins that increase or decrease during different stages of feeding, we have carried out a quantitative proteomics of the salivary gland extracts of male and female *D. reticulatus* at different stages of feeding. When the total protein content from the male and female salivary gland extracts was quantified at

different stages of feeding, it was observed that the protein content of the female salivary gland extracts increases by about 70 times as the feeding progresses (Fig. 3.6.).

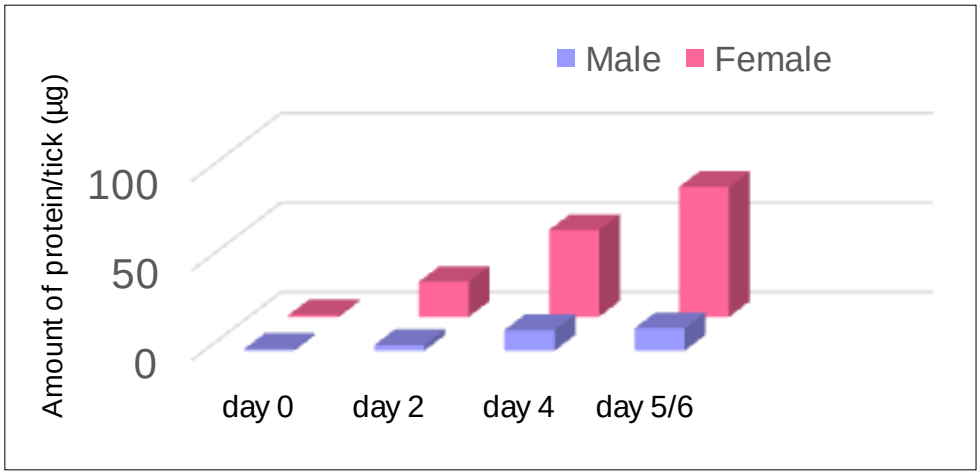


Figure 3.6. Protein quantification in the male and female salivary gland extracts at different stages of feeding. The amount of protein in the female ticks increased from 1.05 µg/ tick at day 0 to 72.6 µg/ tick at day 6 (more than 70 times), while the amount of protein in the male ticks increased from 0.92 µg/ tick at day 0 to 12.57 µg/ tick at day 5 (more than 10 times).

Table 3.5. Setup of three iTRAQ runs

Run 1	Male + Female	Male unfed, Male completely fed, Female unfed, Female completely fed
Run 2	Female alone	Female unfed, Female early fed, Female late fed, Female completely fed
Run 3	Male alone	Male unfed, Male early fed, Male late fed, Male completely fed

All experiments were carried out in technical replicates (with 8 iTRAQ labels as described in methods) and data from the three iTRAQ runs were analyzed separately. For equal loading, data was normalized with 113 as the denominator label for all three runs.

A total of 552 proteins were identified in the male+female run, 439 proteins were identified in the male alone run and 636 proteins were identified in the female alone run. These proteins have been mapped to the transcriptome and we are currently carrying out the quantification of different proteins at different stages of feeding.

3.4. Discussion

We have generated the sialomes of male and female D. reticulatus. From the transcriptome we have reported a total of 37,036 CDS, classified them into various classes and compared their expression patterns in male and female ticks. We have also generated the proteomes of both male and female tick salivary gland extracts at different stages of feeding. We have identified over 400 proteins in each of the three iTRAQ^{TM} run and we are currently analyzing this data for detailed expression profiling of the proteins expressed at different stages of feeding in the male and female salivary glands.

As indicated in the transcriptome, many transcripts were differentially expressed in the male or the female salivary glands, and in some cases, the fold differences were well over 1000 times. Interestingly, the number of proteins overexpressed in females was more than twice that of males. This supports the observation made in the total protein content where the protein content of the female increases drastically during feeding (>10 times) while the protein content of the males shows only a marginal increase (~2 times). This may also be attributed to the observed differences in the feeding patterns of the male and female ticks. Female ticks, which feed on larger volumes of blood, produce more proteins for the necessary metabolic processes.

Interesting classes of proteins which were selectively overexpressed in females were 8.9 kDa proteins, Kunitz type inhibitors, lipocalins and DA-p36 family of proteins. On the other hand male specific serine proteases, glycine rich superfamily, mucins IgG binding proteins and cystatins were some families which

were overexpressed in the male ticks. With these observations we could conclude that the female ticks produce abundant amounts of anti-haemostatic and immunomodulatory proteins, while the male ticks produce copius amounts of proteins required for sexual reproduction. Members of these families of proteins may represent new molecular structures with the same basic scaffold but novel mechanisms to disarm the host haemostatic and immune system. We have therefore generated a library of molecules which could be drug leads as antithrombotics or immunesuppressors and can be developed further for clinical use. We have in fact inititated the recombinant expression of three Kunitz type proteins of the monolaris subclass and intent to study their structure activity relationship and develop them further.

The importance of anti-haemostatic and immunomodulatory proteins produced by the females has been discussed in the previous sections and the importance of proteins involved sexual reproduction produced by the male ticks is discussed below. The mouthparts of male ticks play an important role in the sexual reproduction. The male ticks salivate enormously on the spermatophore, and push it inside the female's genital pore using their mouthparts (Feldman-Muhsam et al., 1970). Glycine rich proteins and mucins are the classes of proteins which aid this process (Kubes et al., 1994). These proteins coated on the spermatophore avoid stickiness of the spermatophore while it is being transferred inside the female genital pore. An interesting class of proteins exclusively expressed by the male ticks (>200 times overexpressed in males compared to females) was the IgG binding proteins. These IgG binding proteins have been

reported to combat the injested host IgGs. The IgG binding proteins produced and secreted at the site of feeding assist the female ticks in the vicinity to engorge at faster rates (Wang and Nuttall, 1995).

CHAPTER 4

Structure-activity relationship of avathrin, a novel thrombin inhibitor

4.1. Introduction

The action of thrombin (FIIa) is pivotal in both haemostasis and thrombosis (Stubbs and Bode, 1993). This crucial enzymes plays roles as both a procoagulant and an anticoagulant under different conditions (Di Cera, 2008). In its pro-coagulant role: (a) it cleaves soluble fibrinogen to fibrin monomers, which polymerize to form a nascent clot; (b) it activates the transglutaminase (FXIII) that covalently crosslinks fibrin monomers to stabilize the clot; (c) it activates non-enzymatic cofactors (FV and FVIII) required for its own amplification; (d) it activates FXI which in turn activates the intrinsic pathway; and (e) it activates platelets by cleaving protease-activated receptors leading to their shape change, degranulation and aggregation (Monroe et al., 2002; Versteeg et al., 2013). Conversely, thrombin also plays a significant role as an anticoagulant by down regulating the progression and amplification of the blood coagulation process: upon binding to thrombomodulin, it activates protein C, which in turn inactivates both cofactors FVa and FVIIIa to mitigate further thrombin generation (Di Cera, 2008). These paradoxical procoagulant and anticoagulant roles of thrombin maintain a balance between uncontrolled bleeding and formation of obstructive thrombi, with sufficient thrombus formation when desired.

Cardiovascular disease is the single largest killer worldwide and is a hefty contributor to the burden of non-communicable diseases (Chaudhari et al., 2014). Ischemic heart disease and stroke, both of which are pathological manifestations of thrombosis are the most common examples of cardiovascular disease and account for up to one in four deaths worldwide (Raskob, 2014). Anticoagulants

like direct thrombin inhibitors (DTIs), direct factor Xa (FXa) inhibitors and vitamin K antagonists (VKAs) comprise a significant fraction of the current therapeutic options as antithrombotic drugs. Some examples of DTIs used as therapeutic options are bivalirudin, the synthetic analogue of hirudin which is a bivalent inhibitor binding to the thrombin active site and exosite-I; argatroban and dabigatran, small molecule univalent DTIs that bind to the active site alone; and low-molecular-weight-heparins (LMWHs) which inhibit thrombin in an antithrombin-dependent manner (Coppens et al., 2012). Despite being popular options of anticoagulant therapy, these classes are fraught with limitations like narrow therapeutic window, individual dosing, high bleeding risks, poor bioavailability and high food-drug interactions (Bauer, 2013). Therefore, novel, superior anticoagulants with greater benefits are being sought (described in detail in 1.3).

Haematophagous animals have adapted a blood feeding diet and have evolved an assortment of molecules that control host haemostasis to ensure a continuous blood flow for successful feeding (Koh and Kini, 2008). Due to thrombin's crucial role in haemostasis, it is not surprising that ixodid ticks have evolved numerous thrombin inhibitors, and among anticoagulants, thrombin inhibitors take a central stage in these blood sucking parasites. Hirudin, haemadin, triabin, ornithodorin and rhodniin are some of the most extensively studied examples of specific families of thrombin inhibitors from haematophagous animals (Huntington, 2014) (described in detail in 1.4.3.5).

We have previously described a novel class of thrombin inhibitor variegin, a 32-residue long peptide which is a fast, tight binding, and competitive thrombin inhibitor from the salivary gland extracts of the hard tick- *Amblyomma variegatum* (Koh et al., 2007). Here we describe the structural and functional characterization of 'Avathrin', a 30-residue long peptide, the transcripts of which were detected in the form of a polypeptide in the salivary glands of *Amblyomma variegatum*. We have synthesized avathrin using solid phase peptide synthesis, studied its structure-activity relationship and studied its in vivo efficacy in a murine thrombosis model. Although avathrin shows only 40% sequence identity with variegin, it selectively inhibits thrombin in a similar fast, tight binding competitive mode with a Ki of 545 pM.

4.2. Materials and Methods

4.2.1. Materials

9-Fluorenylmethyloxycarbonyl (Fmoc)-L-amino acids, Wang resin, N, N-Dimethylformamide, Piperidine, O-(7-azabenzotriazol-1-yl)- 1,1,3,3-tetramethyluronium hexafluorophosphate (HATU), and N,N-diisopropylethylamine (DIPEA) kallikrein, human fibrinogen, and bovine trypsin were from Merck Chemicals Ltd. (Nottingham, UK). Trifluoroacetic acid, acetonitrile, 1, 2- ethanedithiol, thioanisole, bovine chymotrypsin, 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), HEPES sodium salt and polyethylene glycol (PEG) 8000, ferric chloride hexahydrate and bovine serum albumin were purchased from Sigma- Adrich (St. Louis, Missouri, USA). Human alpha thrombin, human factor VIIa (FVIIa), factor XIIa (FXIIa), tissue plasminogen activator (TPA), urokinase, human factor IXa (FIXa), factor Xa (FXa), factor XIa (FXIa), APC, and plasmin were from Hematologic Technologies, Inc. (Essex Junction, VT). Recombinant thrombin was a gift from the Chemo-Sero-Therapeutic Research Institute (KAKETSUKEN). Chromogenic substrates S2222, S2238, S2251, S2288, S2302, S2366, S2444, S2586, and S2765 were purchased from Chromogenix (Milano, Italy). Spectrozyme FIXa was from American Diagnostica Inc. (Stamford, CT, USA). Crystallization greased plates and coverslips were purchased from Hampton Research (Aliso Viejo, California, USA). All other chemicals and reagents used were of analytical grade.

4.2.2. Peptide synthesis and purification

All peptides (avathrin, ultravariegin, QT26, IS20, GL16, avathrinS12A, avathrinS12H, avathrinK10R, avathrinL16P,G17P and β-avathrin) were synthesized using solid-phase peptide synthesis on an Intavis MultiPep RSi peptide synthesizer. Fmoc groups of the amino acids were removed using 20% v/v piperidine in N, N-dimethylformamide and coupled using HATU/DIPEA. All peptides were synthesized on standard Wang resin which had a loading capacity of 0.54 mmol/g. The peptides were cleaved from the resin using a cleavage cocktail of trifluoroacetic acid/1, 2-ethanedithiol/thioanisole/water (92.5%/2.5%/2.5%/2.5%). Crude peptides were purified using reverse-phase HPLC on an AKTA basic purifier from GE Healthcare (Uppsala, Sweden) with a Jupiter proteo (5 µm, 250 mm X 10 mm) column. The purity and mass of all peptides were determined by electrospray ionization mass spectrometry using an LCQ Fleet Ion Trap mass spectrometer from Thermo Fisher Scientific (Waltham, MA, USA).

4.2.3. CD Spectroscopy

Far-UV CD spectra (260–190 nm) of avathrin, QT26 and IS20 dissolved in 10 mM sodium phosphate buffer (pH 7.4) were measured using a Jasco J-810 spectropolarimeter (Easton, MD, USA). All measurements were carried out at room temperature using a 0.1-cm path length cuvette with a scan speed of 50 nm/min, a bandwidth of 2 nm and a resolution of 0.2 nm.

4.2.4. Inhibition of thrombin amidolytic activity and determination of inhibitory constants

All thrombin amidolytic activity assays for avathrin and all variants were performed in 96-wells microtiter plates in 50 mM Tris buffer (pH 7.4) containing 100 mM NaCl and 1 mg/ml bovine serum albumin at room temperature. Typically, 100 µl of peptide and 100 µl of thrombin were preincubated for different durations before 100 µl of the S2238 was added to the reaction wells. The rates of formation of coloured product p-nitroaniline were followed by measuring the absorbance at 405 nm for 10 min with a Tecan InfinitePro Microplate reader. Percentage inhibition was calculated by taking the rate of increase in absorbance in the absence of inhibitor as 0%. Dose-response curves were fitted using the Prizm Software to calculate IC_{50} values and Hill coefficients. Assays to measure the inhibitory constants were also carried out in a similar way as described above. Typically, 100 µl of thrombin (0.81 nM) was added to wells containing 100 µl of different concentrations of the peptide and 100 µl of different concentrations of S2238 (100 µM). Product formations was measured and residual velocities were determined. The resulting equation was fitted to the Morrison's tight binding equation (Copeland, 2000) and the inhibitory constants were determined.

4.2.5. Inhibition of thrombin fibrinogenolytic activity

The abilities of avathrin, QT26, IS20 and GL16 to prolong fibrinogen clotting time were tested using a BBL fibrometer from BD Biosciences (Franklin Lakes, NJ, USA). Typically, 100 µl of fibrinogen (final concentration, 3 mg/ml)

was added to wells conatining 100 µl of peptides (different concentrations) at 37 °C. Clotting of fibrinogen was initiated by the addition of 100 µl of thrombin (final concentration, 20 nM). All reagents and samples were dissolved in 50 mM Tris buffer (pH 7.4) containing 100 mM NaCl.

4.2.6. Serine protease selectivity

The selectivity profile of avathrin was examined against 13 serine proteases: anticoagulant serine protease APC, procoagulant serine proteases (FXIIa, FXIa, FXa, FIXa, FVIIa, kallikrein, and thrombin), fibrinolytic serine proteases (plasmin, TPA, and urokinase), and classic serine proteases (chymotrypsin and trypsin). Effects of avathrin on these serine proteases were determined by inhibition of their amidolytic activities assayed using particular chromogenic substrates. All assays were carried out in 96-well microtiter plates, in a 50 mM Tris buffer (pH 7.4) containing 100 mM NaCl and 1 mg/ml BSA at room temperature (5 mM of CaCl2 were also present in the buffer for FVIIa, FIXa, FXa, chymotrypsin and activated protein C assays). Typically, 100 µL of avathrin (concentrations for thrombin assay: 1000 nM, 100 nM and 10 nM; concentrations for all other serine proteases: 100 µM, 10 µM and 1 µM) was incubated with 100 µL of the respective serine protease for 5 minutes followed by the addition of 100 µL of the respective substrate. The reaction rates were followed for 10 minutes using a Tecan InfinitePro microplate reader and percentage inhibition at different concentrations of avathrin were calculated using the rate of increase in absorbance in the absence of inhibitor as 0%.

4.2.7. Cleavage of avathrin by thrombin

Avathrin (150 μM) was incubated with human α-thrombin (5 μM), in a 50 mM Tris buffer (pH 7.4) containing 150 mM NaCl and 1 mg/ml BSA. After different incubation times, reactions were quenched with 1% TFA (pH 1.8) and loaded onto a Jupiter Proteo (4μm, 90Å) microbore column (100 x 1.0 mm) attached to a Dionex nano-HPLC system and eluted using an acetonitrile gradient with 0.05% TFA and 99.95% milliQ water as eluent A and 0.05% TFA, 19.95% milliQ water and 80% ACN as eluent B. New peaks other than those present in the control chromatogram were taken as the peaks of cleaved products. The masses of the cleaved products were confirmed by analyzing their masses using electrospray ionization mass spectrometry. Cleavage products were quantified by integrating the peaks and calculating the area under the curves.

Effect of different pre-incubation times (hence cleavage), on the thrombin inhibitory activity of avathrin was measured. Briefly, amidolytic assays were performed as described in 4.2.4. where different concentrations of avathrin (1000 nM, 100 nM and 10 nM were incubated with thrombin for different pre-incubation times (up to 36h), followed by which the substrate was added and inhibitory activities were measured by monitoring the rates of reaction.

4.2.8. Crystallization of avathrin in complex with thrombin

Recombinant α-thrombin (in 150 mM NaCl) as a generous gift from Chemo-Sero-Therapeutic Research Institute (KAKETSUKEN, Japan) was desalted using 3000 MWCO spin filters in 20 mM ammonium bicarbonate (NH_4HCO_3), and lyophilized before using it for crystallization. The reported

crystallization conditions for thrombin in complex with other inhibitors, like variegin, hirugen and hirulog were used and optimized further (Koh et al., 2011; Skrzypczak-Jankun et al., 1991). Lyophilized avathrin was dissolved in 50 mM HEPES buffer (pH 7.4) containing 375 mM NaCl to a concentration of 81.73 µM (256 µg/mL). Desalted, lyophilized recombinant α-thrombin was subsequently dissolved in the avathrin solution to a final concentration of 54.49 µM (2 mg/ml). The amount of avathrin in the mixture was 1.5 fold in molar excess of thrombin. Crystallization was achieved using hanging drop method vapor diffusion method. Typically, 1 µl of mixture containing avathrin and thrombin was mixed with 1 µl of precipitant buffer (100 mM HEPES buffer pH 7.4), containing 20 to 25% (w/v) PEG 8000 and were equilibrated against 1 ml of precipitant buffer. The crystallization plates were left at 4 °C and crystals appeared after approximately sixr weeks and were harvested for data collection. The entire process for setting up, growing and harvesting of crystals were performed in 4 °C as the crystals are unstable at room temperature.

Prior to data collection, crystals were soaked in a cryoprotectant solution containing the mother liquor, supplemented with 25% (v/v) glycerol, and flash cooled at 100 K in cold nitrogen gas stream (Cryostream cooler, Oxford Cryosystem, Oxford, United Kingdom). A data set of 180 frames was collected using a CCD detector. The data set was processed and scaled using mosflm and aimless respectively (Leslie and Powell, 2007). The structure of the complex was determined by molecular replacement using Phaser (McCoy, 2006) using light and heavy chains of thrombin from the thrombin-variegin crystal structure (3B23)

as the template. The crystal belonged to the monoclinic space group C2 and diffracted up to 2.25 Å resolution with unit cell dimensions as a = 70.2, b =71.2, c=72; and β = 100°. This unit cell dimensions and space group were the same as that of the thrombin-hirulog-1 complex.

Model building and refinement was performed using COOT (Emsley and Cowtan, 2004) and R-work/R-free for the refined structure was 0.185/0.226. The completely refined structure, had no Ramachandran outliers (0%) and had a high MolProbity score of 1.35. Details of crystallographic parameters are given in Table 4.1.

Table 4.1. Crystallographic data and refinement statistics.

Data collection & processing	
Wavelength (Å)	1.54
Space group	C2
Unit cell parameters [a, b, c (Å); β (°)]	70.2, 71.2, 72.0; 100.0
Resolution (Å)	38.46 – 2.25
Unique reflections	14872 (859)
Completeness (%)	89.4 (56.3)
R_{merge}	0.091 (0.362)
R_{pim}	0.055 (0.224)
$CC_{1/2}$	1.00 (0.87)
Mean $I/\sigma(I)$	10.1 (3.0)
Multiplicity	4.6 (3.4)
Refinement	
Resolution (Å)	38.46 – 2.25
R_{work}	0.185
R_{free}	0.226
RMSD bonds (Å)	0.007
RMSD angles (°)	1.19
No. atoms (thrombin/avathrin/water)	2229/92/129
Residues in favoured regions (%)[#]	96.4
Residues in allowed regions (%)	3.6
Residues in disallowed regions (%)	0
Average B factors for atoms (thrombin/avathrin/water) (Å^2)	34.3/68.8/29.6

4.2.9. Ferric chloride carotid artery thrombosis model

All animal experiments were carried out under protocol 041/12 approved by Institutional Animal Care and Use Committee, National University of Singapore. The FeCl3- induced carotid artery thrombosis model was performed to demonstrate the in vivo efficacy of avathrin in a murine model using previously described protocols and compared with the efficacy of hirulog-1 using the same model (Eckly et al., 2011). Briefly, C57BL/6 male mice (9–11 weeks old, 24.5–27.5 g) were anesthetized with an intraperitoneal injection of ketamine (75 mg/kg) and medetomidine (1 mg/kg). In each set of experiment (n = 6), 0.1 mL of different doses of avathrin or hirulog were injected via tail vein into the mice. The right carotid artery was exposed using blunt dissection, and vascular injuries were caused by applying filter paper (2 mm x 2 mm) saturated with 10% FeCl3 on top of the carotid artery. Following 3-min FeCl3 exposure, the filter paper was removed and the vessel was washed with sterile normal saline. To determine the time to occlusion (TTO), a miniature Doppler flow probe was placed around the carotid artery, and blood flow was recorded using a Transonic flowmeter, and data was acquired using ADInstruments and recorded wing the Labcharts software. The maximal monitoring time after injury was 30 min. Mice were killed via cervical dislocation immediately after conclusion of the experiment and prior to recovery from anesthesia.

4.3. Results

4.3.1. Detection of avathrin transcripts in salivary glands of *Amblyomma variegatum*

Transcripts encoding variegin-like peptides were detected in the large basal granular cells of salivary gland acini type II of nymphs, adult males and females by in situ hybridization with a single stranded DNA probe labelled with digoxigenin. There were differences in the expression pattern of the transcript at different stages of feeding (Dr M. Kazimirova, unpublished observations). Strongest expression was observed at 2-4 days of feeding in the nymphs, 12 days of feeding in males and 5 days of feeding in females. The transcript encodes for a 219-residue protein that contains a putative secretion signal and five highly similar repetitive sequences, each around 30-residue-long. Each of the five repeats showed ~40% sequence identity with the 32-mer variegin. Initially, this transcript was named 'AvaHIRU' (BAD29729), based on the structural similarity of its C-terminus with that of hirudin. Subsequently, one representative sequence out of the five repeats (30 residues) was synthesized and characterized to be a potent thrombin inhibitor as reported here. As a result, we renamed this AvaHIRU derived peptide as 'Avathrin' (*Amblyomma variegatum* derived thrombin inhibitor) for a better reflection of its origin and function.

Variegin:SDQGDVAEPKMHKTAPPFDFEAIPEEYLDDES

Avahiru: SGGHQTAVPKISKQGLGGDFEEIPSDEIIE

Figure 4.1. Comparison of variegin and avathrin sequences. Variegin and avathrin had the typical sequence for thrombin active site interactions (P2 Pro, P1 Lys). Both variegin and avathrin had an acidic C-terminus that could interact with the exosite-I. His12 in variegin which disrupted the catalytic triad was replaced by Ser12 in avathrin. Identical residues are underlined.

4.3.2. Synthesis and purification of avathrin and its variants

For the structural and functional characterization, avathrin was synthesized using solid phase peptide synthesis, purified by reversed-phase high performance liquid chromatography (RP-HPLC) and its purity and mass were determined by ESI-MS (Fig. S1-S3, Table S1). ESI-MS spectrum of avathrin revealed 3 *m/z* peaks with +4, +3 and +2 charged states and a mass of 3139.5 ± 0.2 Da. Thus, the estimated masses of the avathrin peptides matched the calculated masses. The CD spectra of avathrin and two of its variants were typical of random coil (Fig. 4.2).

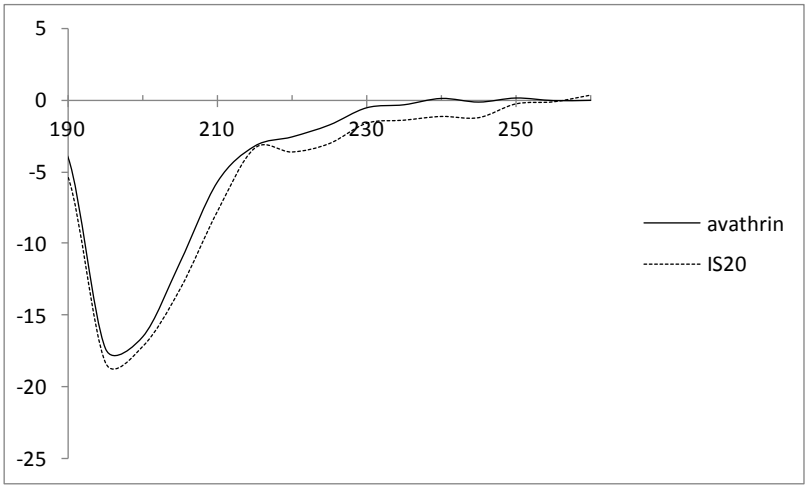


Figure 4.2. Far UV CD spectra of avathrin and IS20. Avathrin and its truncated variant, IS20 showed CD spectra characteristic of random coil.

4.3.3. Inhibition of thrombin amidolytic activity

The ability of avathrin to inhibit the amidolytic activity of thrombin was assayed using S2238 as the substrate. Avathrin inhibited thrombin amidolytic activity in a dose-dependent manner with IC_{50} and Hill coefficient of 6.95 ± 0.42 nM and 0.92 ± 0.01 respectively. Significant inhibition was observed at equimolar concentrations of avathrin and thrombin (~25% inhibition at 0.81 nM avathrin) indicating that avathrin is a tight binding thrombin inhibitor (Fig. 4.3A). Reaction progress curves showed that a steady-state equilibrium was achieved upon mixing indicating a fast binding mode of avathrin (Fig 4.3B). Residual reaction velocities of thrombin amidolytic activity in presence of avathrin at different S2238 concentrations were determined to calculate K_{iapp}. A straight line plot of the linearly increasing K_{iapp} with increasing S2238 concentrations indicated that avathrin is a competitive inhibitor and these K_{iapp} were fitted to the Morrison equation using the GraphPad Prizm software and a K_i of 545.78 ± 0.29 pM of avathrin towards thrombin was obtained (Fig. 4.3C and 4.3D). Therefore, avathrin is a fast, tight binding, competitive inhibitor of thrombin.

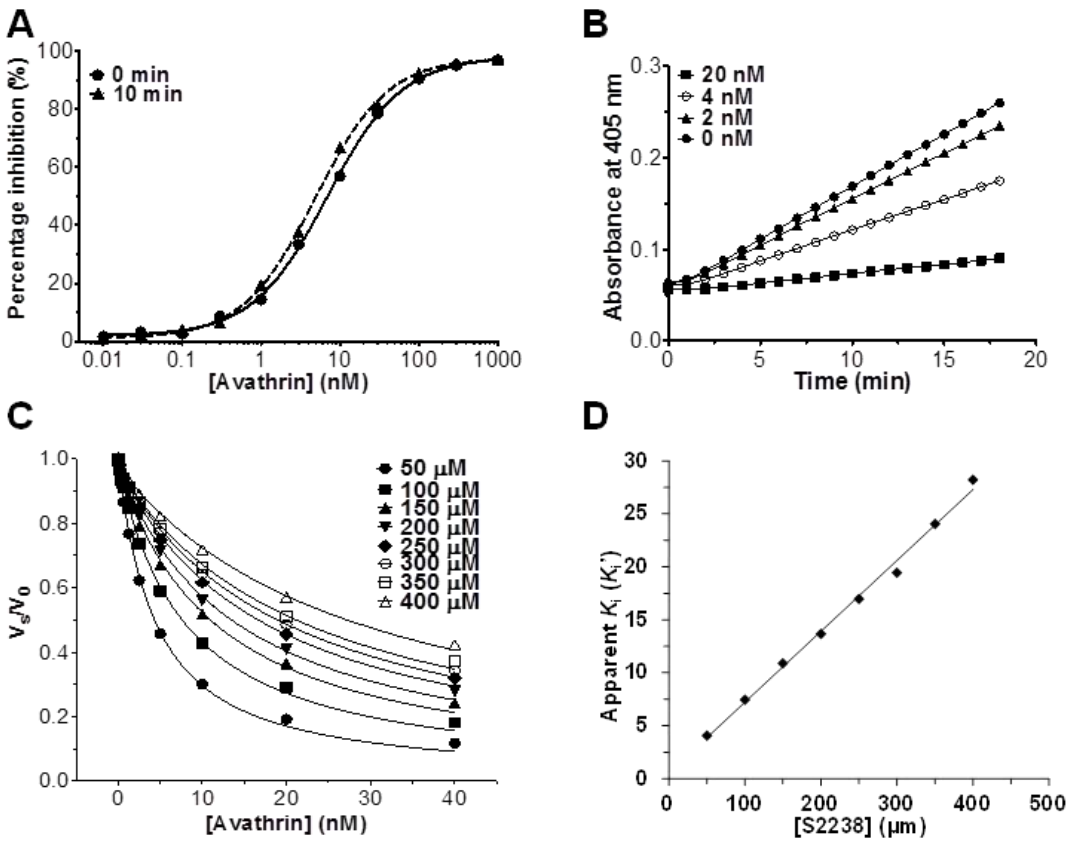


Figure 4.3. Kinetics of thrombin inhibition by avathrin using the chromogenic substrate S2238. A. Avathrin inhibits thrombin amidolytic assay in a dose-dependent manner. The effect of various concentrations of avathrin on the amidolytic activity of thrombin (0.81 nM) was measured using the substrate S2238 (100 µM) with (10 min) or without pre-incubation. IC_{50} and Hill coefficient of the inhibition are 6.95 ± 0.42 nM and 0.92 ± 0.01 at 0 min and 4.86 ± 0.36 nM and 0.94 ± 0.02 nM at 10 min, respectively. Each data point is the mean ± S.D. of at least three experiments. **B. Avathrin is a fast binding inhibitor**. Thrombin (0.81 nM) amidolytic assay using S2238 (100 µM) in presence of various concentrations of avathrin was carried out and linear progression curves of thrombin inhibition in presence avathrin were achieved – a characteristic of fast binding inhibitor. **C. Avathrin is a tight binding inhibitor.** The residual thrombin amidolytic activity in presence of various concentrations of avathrin was measured at different concentrations of S2238 and the K_i' (apparent K_i) was determined. Reactions were started with the addition of thrombin (0.81 nM). Data were fitted to the Morrison tight binding equation using GraphPad Prizm software. Each data point is the mean ± S.D. of at least three experiments. **D. Avathrin is a competitive thrombin inhibitor**. Plot of K_i' against S2238 concentration increased linearly, indicating avathrin is a competitive inhibitor. The inhibitory constant K_i was determined to be 545.3 ± 3.1 pM.

4.3.4. Inhibition of thrombin fibrinogenolytic activity

Fibrinogen is an important physiological substrate for thrombin which binds to both active site and exosite-I of thrombin. Therefore, the ability of avathrin to inhibit thrombin's fibrinogenolytic activity was evaluated.

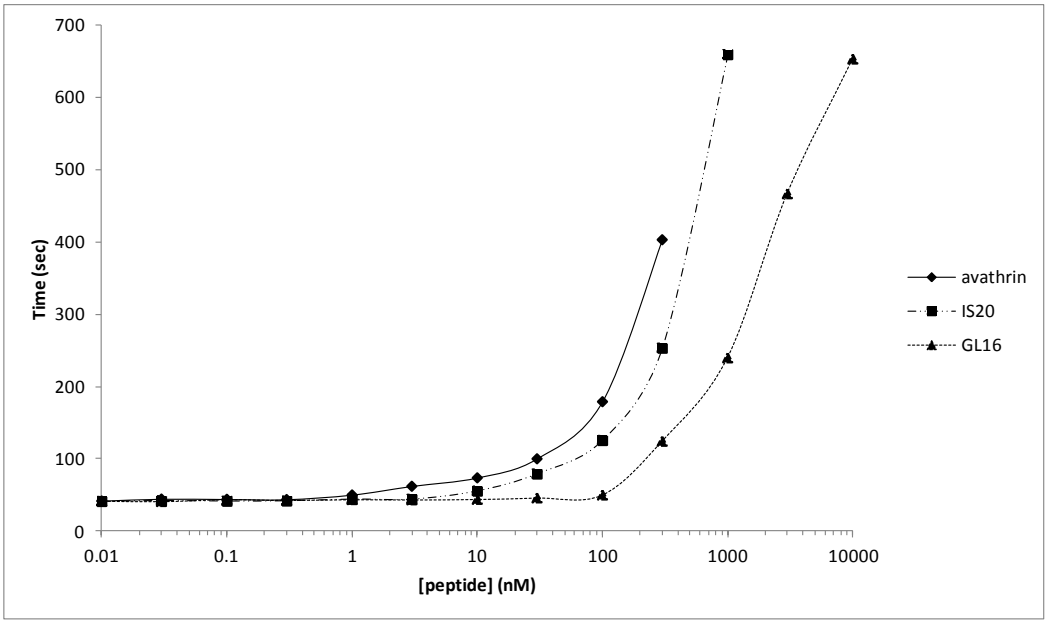


Figure 4.4. Inhibition of fibrinogenolytic actvity of thrombin. Avathrin and its truncated variants (IS20 and GL16) prolonged fibrinogen clotting times (n = 3, error bars represents S.D.). GL16 inhibited fibrinogenolytic activity but not amidolytic activity of thrombin, suggesting the C-terminal binding to exosite-I.

Avathrin prolonged fibrinogen clotting time in a dose-dependent manner (Fig. 4.4). This observation is consistent with observations for variegin and the C-terminus of hirudin which also inhibit thrombin's fibrinogenolytic activity (Koh et al., 2007; Stone and Hofsteenge, 1986).

4.3.5. Serine protease selectivity

The serine protease selectivity of avathrin was examined by screening against 13 serine proteases which included pro- and anti-coagulant serine

proteases of the blood coagulation cascade and the two classic serine proteases trypsin and chymotrypsin. At 10 nM, avathrin inhibited 64.95% of thrombin's activity. However, even at 1 μM, it showed no significant inhibition (<5%) of any of the serine protease other than thrombin. Trypsin and fXIa activities were inhibited ~30% at 100 μM avathrin (Fig. 4.5). Thus, avathrin is a highly potent and selective (at least 4 orders of magnitude) thrombin inhibitor.

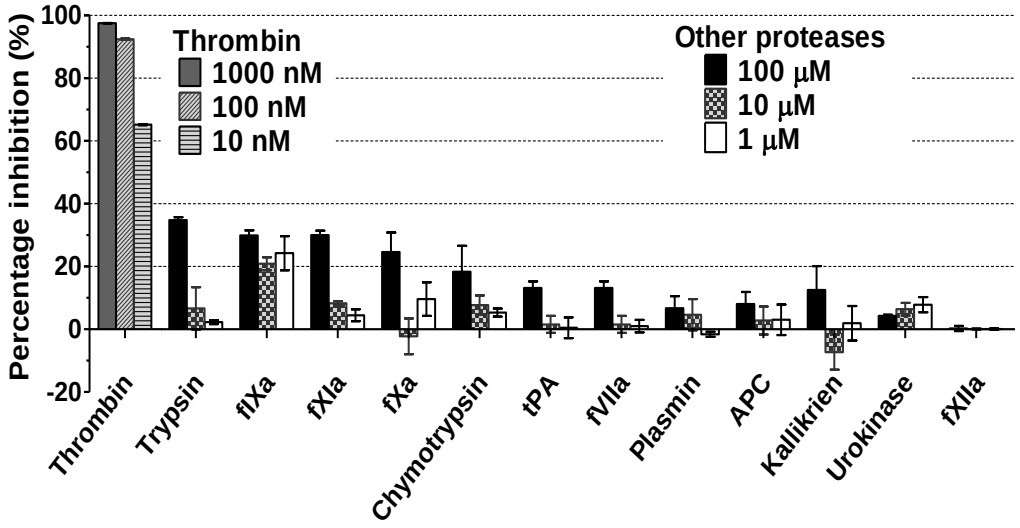


Figure 4.5. Serine protease selectivity of avathrin. Avathrin was screened for its selectivity against 13 serine proteases: procoagulant serine proteases (fXIIa, fXIa, fXa, fIXa, fVIIa, kallikrein and thrombin); anticoagulant serine protease (APC); fibrinolytic serine proteases (plasmin, tPA and urokinase) and classical serine proteases (chymotrypsin and trypsin). The final concentrations of proteases and substrates used for the amidolytic assays are given in parentheses in nM and μM, respectively, unless mentioned otherwise: α-thrombin (0.81)/S2238 (0.1), trypsin (0.87)/S2222, fIXa (333)/Spectrozyme® fIXa (0.4), fXIa (0.125)/S2366 (1000), fXa (0.24)/S2765 (650), chymotrypsin (1.2)/S2586 (0.67), tPA (36.9)/S2288 (1000), fVIIa (460)/S2288 (1200), plasmin (3.61)/S2251 (1200), APC (2.74)/S2366 (600), kallikrein (0.93)/S2302 (1100), urokinase (32 U/ml)/S2444 (650), fXIIa (20)/S2302 (1000). Activity of thrombin was tested at lower concentrations of avathrin (1000 nM, 100 nM and 10 nM) while the other proteases were tested at much higher concentrations of avathrin (100 μM, 10 μM and 1 μM). Each data point is the mean ± S.D. of at least three experiments.

4.3.6. Cleavage of avathrin by thrombin

Because avathrin binds to the thrombin active site, we hypothesized that like variegin and the other thrombin inhibitors (Krowarsch et al., 2003), avathrin may be prone to proteolytic cleavage by thrombin. A time course analysis of avathrin incubated with thrombin indicated that avathrin was indeed cleaved of by thrombin. A RP-HPLC separation of the reaction mixture incubated for different time points was carried out to separate and quantify the cleaved products. At 0 min, a single peak corresponding to full length avathrin and a peak corresponding to thrombin were observed (Fig. 4.6A upper panel). With increasing periods of incubation, two new peaks corresponding to the cleaved products (SGGHQTAVPK with a mass of 981.3 Da and ISKQGLGGDFEEIPSDEIIE with a mass of 2176.3 Da indicating the cleavage of Lys10-Ile11 peptide bond) started to appear (Fig. 4.6A middle panel). With increasing incubation times, the amounts of these two peaks increased, and the amount of the full length avathrin decreased. After 10 h, only three peaks, two peaks corresponding to the cleavage products of avathrin, and one peak corresponding to thrombin were detected. The peak corresponding to avathrin was not detected indicating complete cleavage (Fig. 4.6A lower panel). These cleavage products wer quantified by integrating the areas under the curve (Fig. 4.6B). In order to verify the effect of thrombin cleavage on the

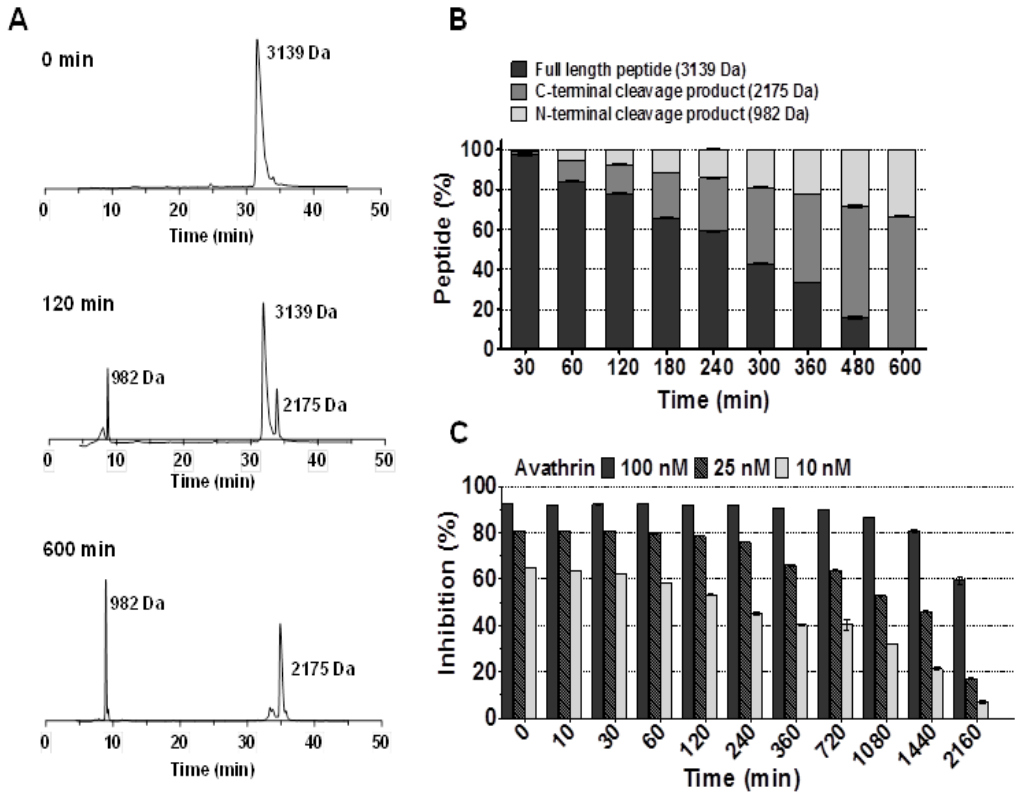


Figure 4.6. Cleavage of avathrin by thrombin. A. Typical HPLC chromatograms of avathrin cleavage by thrombin. Avathrin (150 μM) was incubated with thrombin (5 μM) for different lengths of time and the reaction mixtures were separated using RP-HPLC; and the masses of the cleavage products were analyzed with ESI-MS. At 0 min (upper panel), a single peak corresponding to full length avathrin (mass 3139 Da) was identified. At 120 min (middle panel), two new peaks corresponding to N-terminal cleavage product (SGGHQTAVPK; mass 982Da) and C-terminal cleavage product (ISKQGLGGDFEEIPSDEIIE; mass 2175 Da) were identified in addition to avathrin peak. At 600 min (lower panel), two peaks corresponding to the N- and C-terminal cleaved products were observed while avathrin peak was not observed, indicating complete cleavage. **B. Quantification of cleavage products.** Relative percentages of avathrin, its N- and C-terminal cleavage products were quantified by calculating areas under the curve. Each data point is the mean ± S.D. of at least three experiments. **C. Prolonged inhibitory effect of avathrin.** Avathrin was incubated with thrombin (0.81 nM) for up to 36 h, and assayed at different time points for its ability to inhibit thrombin amidolytic activity on the chromogenic substrate S2238. At 25 nM of avathrin, the inhibitor was present in ~30-fold excess of thrombin (0.81 nm), and these ratios are similar to that used in HPLC analysis of cleavage products. After 24 h, cleavage products retained >50% of the

original inhibitory activity, although full length avathrin was completely cleaved around 10 h. Thus, cleavage products, particularly the C-terminal cleavage product of avathrin appears to remain bound to thrombin and continues to inhibit thrombin. Each data point is the mean ± S.D. of at least three experiments.

inhibitory activity, avathrin was incubated with thrombin for up to 36 h and the inhibitory effect of avathrin at different time points was examined by testing the amidolytic activity of thrombin (Fig. 4.6C). At 25 nM concentration (30 times excess peptide), avathrin inhibited more than 30% of thrombin activity for 18 h. Thus, avathrin exhibited prolonged inhibition of thrombin and we concluded its cleaved products retained strong binding to thrombin.

4.3.7. Truncated versions of avathrin

In order to optimize the chain length, we synthesized three truncated versions of avathrin. The three truncated versions, QT26, IS20 and GL16 were synthesized using solid phase peptide synthesis, purified by reversed-phase high performance liquid chromatography (RP-HPLC) and its purity and mass were determined by ESI-MS (Fig. 7). QT26 and IS20 inhibited thrombin amidolytic activity with IC_{50}s and Hill coefficients of 8.94 ± 0.64 nM and 0.89 ± 0.03 and 12.17 ± 0.32 nM and 0.87 ± 0.06 respectively. Inhibitory kinetics of QT26 and IS20 showed that both peptides were tight binding inhibitors with K_i of 760.32 ± 0.91 pM and 5760.00 ± 0.23 pM respectively. A plot of the K_{iapp} versus increasing concentrations of S2238, increased linearly for QT26, but remained constant for IS20, indicating that QT26, like full length avathrin was a competitive inhibitor, while IS20, the cleavage product of avathrin inhibited thrombin with a non-competitive mode (Fig. 4.8A and B). Both QT26 and IS20 inhibited thrombin

fibrinogenolytic activity in a dose-dependent manner. The third and the shortest truncated variant, GL16 did not inhibit thrombin amidolytic activity even at 300 µM, but a slight activation (5-10%) of thrombin amidolytic activity was seen with GL16. However, GL16 did inhibit the fibrinogenolytic activity of thrombin in a dose dependent manner (Fig. 6).

```
Avathrin: SGGHQTAVPKISKQGLGGDFEEIPSDEIIE
QT26:          QTAVPKISKQGLGGDFEEIPSDEIIE
IS20:                ISKQGLGGDFEEIPSDEIIE
GL16:                    GLGGDFEEIPSDEIIE
```

Figure 4.7. Truncated versions of avathrin. Three truncated versions of avathrin - QT26, IS20 and GL16 were synthesized and their abilities to inhibit thrombin active site and exosite were tested.

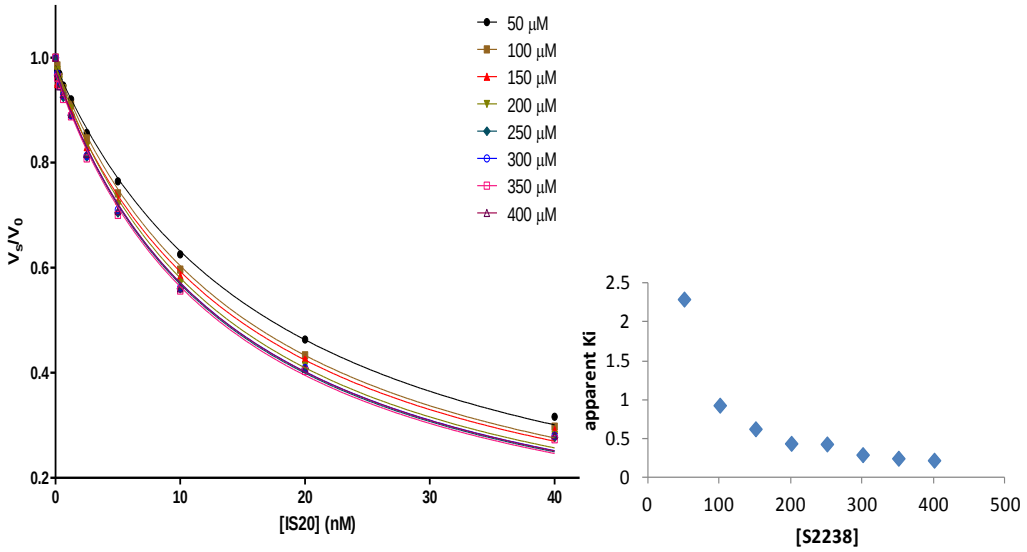


Figure 4.8. Affinity of IS20, the cleavage product of avathrin. A. IS20 at different concentrations (0.078 nM, 0.156 nM, 0.312 nM, 0.625 nM, 1.25 nM, 2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM) was mixed with different concentrations of S2238 (50 µM, 100 µM, 150 µM, 200 µM, 250 µM, 300 µM, 350 µM and 400 µM and the Ki' was determined. Reactions were started by the addition of thrombin (0.81 nM). Data were fitted to the Morrison equation equation using GraphPad

prizm software (n = 3, error bars represent S.D.). **B.** Plot of K_i' against substrate concentration decreased gradually, indicating avathrin competitively inhibited thrombin amidolytic activity and the inhibitory constant Ki was shown to be 5760.3 ± 0.23 pM.

4.3.8. Crystal structure

The crystal structure of thrombin-avathrin complex was determined at 2.25 Å. The electron density for residues of thrombin was well defined except for termini residues of chain A [^{T}(^{1}TFGSG^{5}, ^{13}DGR^{15} and E^{259})]. Only 14 out of the 30 residues of avathrin had well defined densities [^{A}(^{9}PKI^{11} and ^{18}GDFEEIPSDEI^{28})]. Since avathrin gets cleaved by thrombin, a single crystal may represent a heterogeneous population of molecules in which thrombin may either be bound to the full-length peptide or the C-terminal fragment after cleavage. The structure that we have modelled however represents thrombin in complex with the full length avathrin.

Active site interactions - Side chain density of avathrin residues interacting with thrombin active site is not distinctive enough to unambiguously identify them – but based on our sequence, cleavage and kinetic data, we conclude these residues to be ^{A}(^{9}PKI^{11}). The N of P1' (^{A}I^{11}) is in close proximity of the γO of ^{T}S^{195} (distance of 3.04 Å). This proximity would facilitate the nucleophilic attack during the catalytic cleavage of the peptide bond between P1 and P1' residues. In the region where the peptide bound to the active site of thrombin, one hydrogen bond between main chain carbonyl O of ^{A}K^{10} and backbone amide N of ^{T}G^{193} (2.88 Å) was observed. The P1 (^{A}K^{10}) of avathrin is coordinated with S1 (^{T}D^{189}) of thrombin. The N ζ of ^{A}K^{10} is in proximity to the ^{T}189^{D} and possibly interacts

coordinated by a water molecule (distances: $^{T}D^{189}$ to H_2O: 3.58 Å and $^{A}K^{10}$ to H_2O is 1.89 Å respectively). This interaction is similar to that observed for thrombin's interactions with the PAR3 receptor which also possesses a Lys at P1 position and the

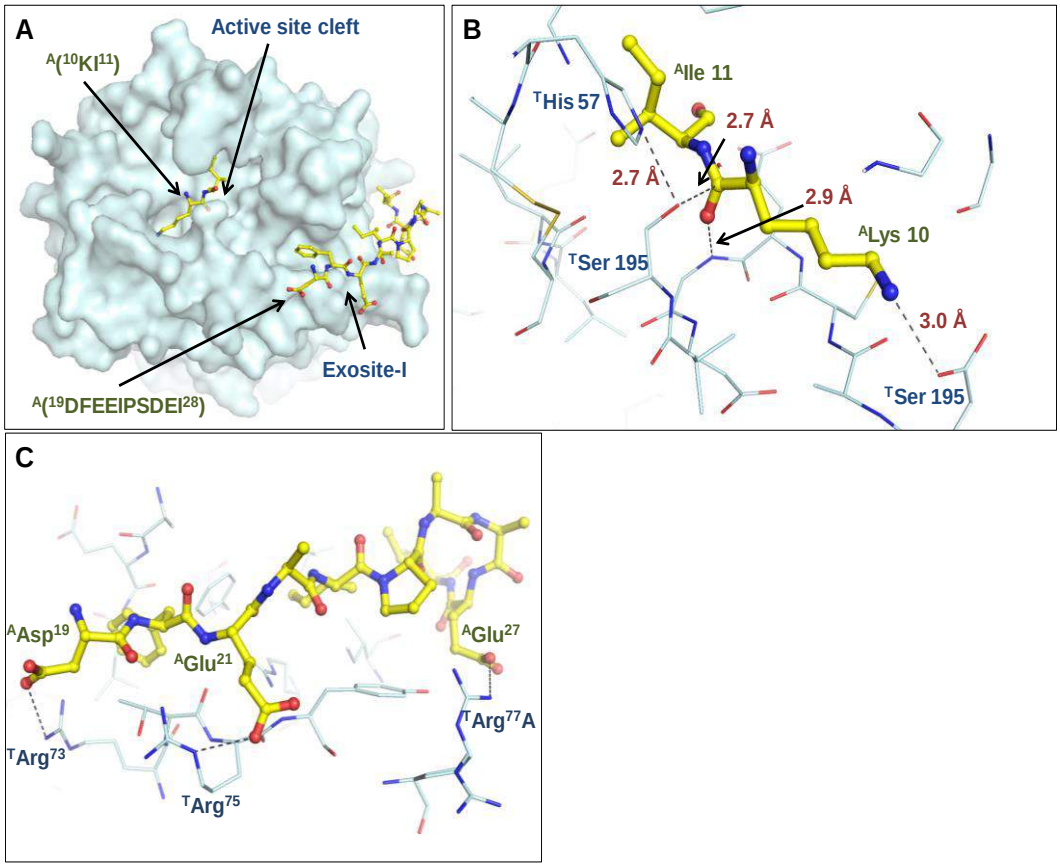


Figure 4.9. Crystal structure of thrombin-avathrin complex. A. Overall thrombin-avathrin complex crystal structure. Surface representation of thrombin is shown in cyan; and ball and stick representation of avathrin is shown (carbon: yellow; nitrogen: blue; oxygen: red). Two avathrin residues (^{A}Lys10 - ^{A}Ile11) were identified in the proximity of the thrombin active site cleft. **B. Avathrin interactions at thrombin active site.** Oγ of ^{T}Ser195 is 2.7 Å away from Nε of ^{T}His57, forming the nucleophile poised to attack the carbon atom of ^{A}Lys10 (P1) 2.7 Å away. Carbonyl oxygen of ^{A}Lys10 is stabilized in the oxyanion hole at 2.9 Å from backbone nitrogen of ^{T}Gly193. ^{A}Lys10 fits into the S1

specificity pocket of thrombin and stabilized by an electrostatic interaction (3.0 Å) with Oδ of ^{T}Asp189 at the bottom of the pocket. **C. Avathrin interactions with thrombin exosite-I.** Three electrostatic interactions (^{A}Asp19-^{T}Arg73; ^{A}Glu21-^{T}Arg75; and ^{A}Glu27-^{T}Arg77A) are observed between avathrin C-terminus and exosite-I. Distances of all three electrostatic interactions lie in between 2.9 to 4.1 Å.

interaction is coordinated by an intermediate water molecule (Bern et al., 1991). The corresponding residue P1 residue in hirulog-1 is an Arg. The side chain nitrogens of this Arg interact directly with its side chain oxygens of $^{T}D^{189}$, but the backbone carbonyl oxygen of the backbone amide N of $^{T}G^{193}$ (Perona et al., 1995) and this interaction is also seen in avathrin-thrombin complex. The basic P1 residue of fibrinogen occupies the S1 specificity pocket of thrombin and forms an antiparallel beta sheet with ^{T}(^{214}S-G^{216}) (Perona et al., 1995). The main chain N of $^{A}K^{10}$ comes close to main chain O of $^{T}S^{214}$ (3.43 Å). And $^{A}P^{9}$ points in the direction of (or is close to) $^{T}G^{216}$ (distance between main chain amide N of $^{A}P^{9}$ and main chain carbonyl O of $^{T}G^{216}$ is 4.51 Å). Therefore, we concluded that avathrin is a canonical inhibitor as it binds to thrombin in a substrate-like fashion.

Interactions with the prime subsite - The apolar binding site of thrombin is a hydrophobic pocket formed by the 60 loop (Page et al., 2005) which is lined by $^{T}H^{57}$ $^{T}Y^{60A}$, $^{T}W^{60D}$, aliphatic (amphipathic) chain of $^{T}K^{60F}$ and the disulphide bond between $^{T}C^{42}$ and $^{T}C^{58}$ (Skrzypczak-Jankun et al., 1991). The P1' side chain of $^{A}I^{11}$ is buried inside this hydrophobic cavity of S1' subsite of thrombin.

Interactions with thrombin exosite - Clear electron density for the main chain residues of ^{A}(^{18}GDFEEIPSDEI^{28}) was observed out of which clear side chain densities for only $^{A}E^{21}$ and ^{A}(^{23}IPSDEI^{28}) could be observed. Although the side

chain of $^{A}D^{19}$ points towards the exterior, the main chain carbonyl O is in close proximity to the side chain η^2 N of $^{T}R^{75}$ (2.81 Å). The side chain oxygens of the cognate Asp corresponding to $^{A}D^{19}$ in hirulog-1/-3 and hirugen structures form ion pairs with the side chain nitrogens of $^{T}R^{73}$. Side chain density for the aromatic ring of $^{A}F^{20}$ is not observed but the C β of the $^{A}F^{20}$ points in a direction towards $^{T}F^{34}$. The cognate Phe in variegin, hirulog-1/3 and hirugen all of which have their aromatic rings pointing towards TF34 and show a T shaped π-Stacking interactions with this Phe of thrombin (Chang, 1983; Skrzypczak-Jankun et al., 1991). $^{A}E^{21}$ side chain points to the surface of thrombin and lies at the interface between the structure. The O ε^1 of $^{A}E^{21}$ forms a salt bridge with the amide N of $^{T}Y^{76}$ (3.75 Å) and the O ε^1 forms a salt bridge with the ε N of $^{T}R^{75}$ (2.88 Å). This is different from what is observed for variegin, hirulog, hirugen. Density for side chain of $^{A}E^{22}$ is not observed but for the C β which points away from the interface towards the solvent. Side chain of $^{A}I^{23}$ is buried in a hydrophobic cavity lined by $^{T}I^{28}$, $^{T}F^{34}$, $^{T}L^{65}$, $^{T}Y^{76}$and $^{T}I^{82}$. The Ile in this position in variegin, hirugen, hirulog all of them appear to be buried within this cavity. $^{A}P^{24}$ comes close to the aromatic ring of $^{T}Y^{76}$ (within 5 Å). Proline can interact aromatic residues favorably, due to both the hydrophobic effect and the interaction between the π aromatic face and the polarized C-H bonds, called a CH/π interaction. But this proline occupies a similar position in variegin and sulfo-hirudin structures. $^{A}S^{25}$ and $^{A}D^{26}$ side chains points away from thrombin into the solvent. C β of $^{A}E^{27}$ is seen. Side chain oxygens are not seen. But the C β points towards $^{T}R^{77}$ (distance between C β and N η of TR77 is 6.54 and 6.75 Å). This may be due to the

oxygens of the glutamate side chain forming ionic interactions with the $^{T}R^{77}$. $^{A}I^{28}$ side chain goes into a cavity formed by $^{T}L^{68}$ and $^{T}I^{82}$.

4.3.9. Structure based variants

For further characterization of the structure-activity relationship and to identify the functionally important residues responsible for inhibitory activity, we synthesized a series of avathrin substitution mutants. From the crystal structure of variegin-thrombin complex, we have shown that $^{V}H^{12}$ binds to the prime subsite and its side chain nitrogen forms a hydrogen bond with $^{T}S^{195}$, hence disrupting the charge relay system of the thrombin catalytic triad. However the electron density of the cognate $^{A}S^{12}$ in the avathrin-thrombin crystal structure could not be observed. Therefore, to identify if the $^{A}S^{12}$ in avathrin plays a role similar to $^{V}His^{12}$ in variegin, we synthesized two mutant peptides replacing this $^{A}S^{12}$ with Ala and His and named these mutants as avathrinS12A and avathrinS12H respectively. AvathrinS12A and was a competitive inhibitor with an IC_{50} and Ki of 101.20 ± 1.32 nM and 6075 ± 1.82 pM (Suppl Fig. 4C and 4D). Since the replacement of $^{A}S^{12}$ resulted in a more than 10-fold loss in the inhibitory activity, we concluded that $^{A}S^{12}$ of avathrin was indeed important for disrupting the catalytic triad similar to the $^{V}H^{12}$ of variegin. AvathrinS12H was a tight binding inhibitor with IC_{50} and K_i of 18.51 ± 0.32 nM and 1237.91 ± 2.46 pM.

From the crystal structure of variegin in complex with thrombin, we have identified the contribution of two proline residues in variegin at position 16 and 17. $^{V}P^{16}$ and $^{V}P^{17}$ induce a kink in the backbone of variegin, which causes a slight upward bend in the main chain of the peptide and this kink has been shown to be

important for the binding of the P' residues of the peptide with the prime subsite of thrombin. In the sequence of avathrin, position 16 and 17 are occupied by $^{A}Leu^{16}$ and $^{A}Gly^{17}$. Therefore, to identify if the two prolines at these two positions could introduce a similar kink in the avathrin backbone, hence making the interactions of the peptide stronger with the thrombin prime subsite, we replaced $^{A}Leu^{16}$ and $^{A}Gly^{17}$ by two proline residues. The activity of this mutant, avathrinL16P,G17P dropped severely and it had an IC_{50} of 181.32 ± 3.76 nM and its binding mode changed from tight binding to competitive mode. The N-terminal acidic residues of variegin are hypothesized to steer towards the basic surface of the exosite-II. The N-terminus of avathrin lacked any acidic residues, and therefore, we introduced acidic residues at two positions in the N-terminus of avathrin (avathrinG2D,Q5D). This mutant did not gain a significant increase in activity and in fact its inhibitory capacity was a little weaker than avathrin with an IC_{50} of 13 ± 1.23 nM and K_i of 981.32 ± 0.45 pM. The Lys at the P1 site of avathrin interacts with $^{T}Asp^{189}$ mediated by a water molecule. The P1 residue of most thrombin substrates is Arg, that directly interacts with the side chain oxygen of $^{T}Asp^{189}$. Therefore we substituted the Lys at this position by Arg, and observed a 5 fold gain in activity with avathrinK10R showing an IC_{50} and Ki of 1.32 ± 0.45 nM and 172.92 ± 2.91 pM . The results of all the variants are summarized in Table 4.2.

Table 4.2. IC_{50} and K_i values of avathrin and all its variants

Peptide	Sequence	IC_{50} (nM)	Hill Slope	K_i (nM)
Avathrin	SGGHQTAVPKISKQGLGGDFEEIPSDEIIE	6.95 ± 0.42	0.92 ± 0.02	0.545 ± 0.003
QT26	QTAVPKISKQGLGGDFEEIPSDEIIE	8.94 ± 0.64	1.22 ± 0.03	0.760 ± 0.009
IS20	ISKQGLGGDFEEIPSDEIIE	12.17 ± 0.32	0.86 ± 0.02	5.760 ± 0.230
GL16	GLGGDFEEIPSDEIIE	N. I.*	N. I.	N. I.
AvathrinS12A	SGGHQTAVPKIAKQGLGGDFEEIPSDEIIE	101.20 ± 1.32	0.62 ± 0.01	6.075 ± 0.180
AvathrinS12H	SGGHQTAVPKIHKQGLGGDFEEIPSDEIIE	18.51 ± 0.32	0.88 ± 0.02	1.230 ± 0.046
AvathrinL16P,G17P	SGGHQTAVPKISKQGPPGDFEEIPSDEIIE	181.32 ± 3.76	0.54 ± 0.02	-
AvathrinG2D,Q5D	SDGHDTAVPKISKQGLGGDFEEIPSDEIIE	12.98 ± 1.23	0.71 ± 0.03	0.932 ± 0.015
AvathrinK10R	SGGHQTAVPRISKQGLGGDFEEIPSDEIIE	1.32 ± 0.45	1.22 ± 0.01	0.172 ± 0.002
β-avathrin	SGGHQTAVPβISKQGLGGDFEEIPSDEIIE	332.16 ± 1.32	0.62 ± 0.01	32.04 ± 0.36

4.3.10. Uncleavable avathrin

Avathrin possesses the typical sequence like thrombin substrates which makes it a favourable candidate to be cleaved while it is bound to the active site. Because of the cleavage of avathrin by thrombin, and owing to its flexibility, the electron densities for avathrin residues interacting with the thrombin active site could not be clearly seen. Therefore, we synthesized a variant of avathrin, called β-avathrin in which the P1 Lys was substituted by a β-homoArg. The scissile peptide bond of avathrin (Lys10-Ile11) was substituted by a proteolytically stable bond (β-homoArg10-Ile11) in β-avathrin. Owing to the proteolytically stable bond, β-avathrin could not be cleaved by thrombin for upto 24 hours and it retained the inhibition of thrombin's amidolytic activity but its potency was significantly reduced. β-avathrin inhibited thrombin with an IC_{50} of 332 ± 1.32

nM and a hill coefficient of β-avathrin was a competitive thrombin inhibitor with a K_i of 32 nM. Crystallization of β-avathrin in complex with thrombin is in process.

4.3.11. Clot bound thrombin inhibition

The blood coagulation cascade culminates with the formation of a haemostatic clot and this clot mainly traps active α-thrombin within the clot limiting its circulation in the blood. However, the α-thrombin trapped within the clot also acts as a reservoir of active thrombin and can become instrumental in causing re-thrombosis. Therefore, inhibition of clot bound thrombin may prevent re-thrombosis from occurring. Therefore, we evaluated the ability of avathrin to inhibit clot bound thrombin. Avathrin was able to inhibit to inhibit clot bound thrombin in a dose-dependent manner with an IC_{50} of 1.12 ± 0.35 μM.

4.3.12. Ferric chloride carotid artery thrombosis model

We used an established $FeCl_3$-induced carotid artery thrombosis model to evaluate the in vivo antithrombotic efficacy of avathrin in mice. The average time to occlusion in control animals was 7.24 min, while the time to occlusion in animals injected with avathrin increased in a dose dependent way.

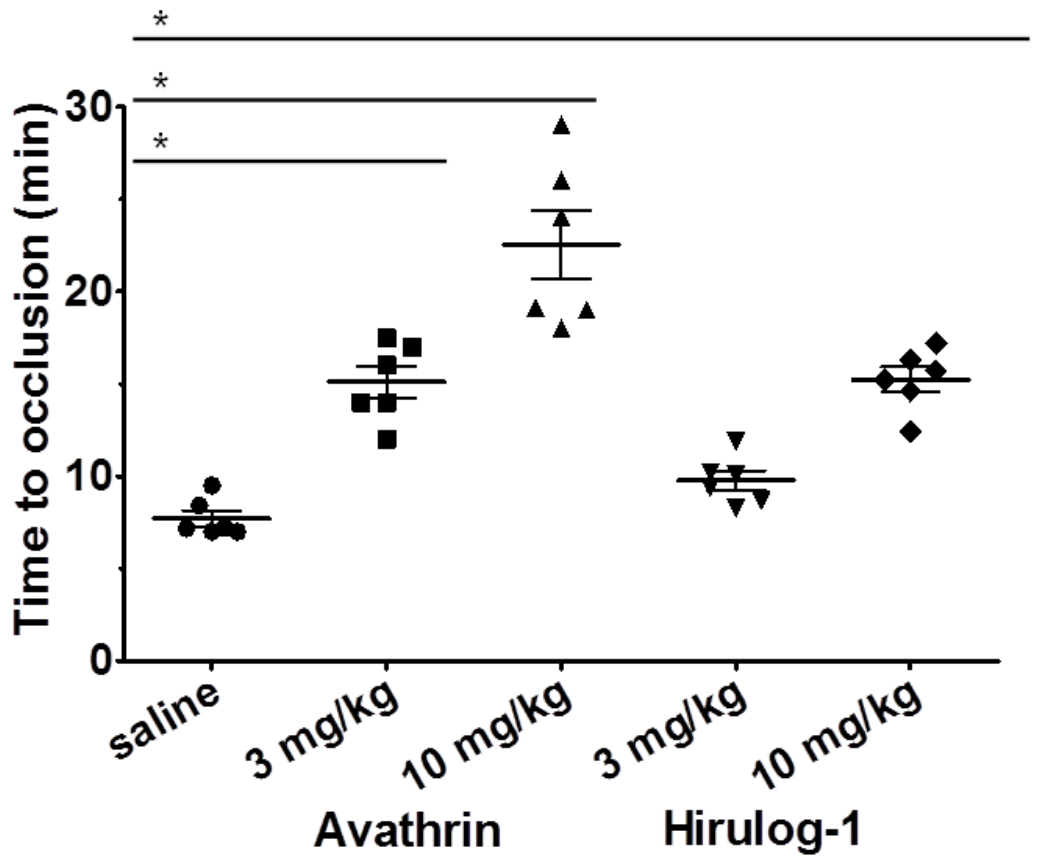


Figure 4.10. in vivo antithrombotic effect of avathrin in murine model. $FeCl_3$-induced carotid artery thrombosis model was used to evaluate the in vivo antithrombotic efficacy of avathrin in mice. Two concentrations (10mg/kg and 3 mg/kg) of avathrin or bivalirudin were injected in mice and the time to occlusion after a $FeCl_3$-induced thrombosis was tested and compared with control animals.

Time to occlusion in animals injected with 10mg/kg and 3 mg/kg were 24 min and 16 min respectively. To compare this efficacy, same doses of the commercially available drug, hirulog were also injected in another set of animals. Time to occlusion in animals injected with 10mg/kg and 3 mg/kg were 15 min and 9 min respectively (Fig. 15.).

4.4. Discussion

Several thrombin inhibitors have been identified from the saliva of hematophagous animals. To control the host haemostatic system, anticoagulants targeting thrombin are produced in the saliva of *A. variegatum*. We have previously described variegin, which is a potent fast binding thrombin inhibitor from the salivary gland extracts of female *A. variegatum*. By using in situ hybridization, we identified variegin like sequences from the salivary gland extracts of *A. variegatum* and synthesized one representative peptide- avathrin, and have studied it structure-activity relationship and demonstrated its *in vivo* efficacy in a murine model.

Although avathrin was identified as a variegin like peptide, it showed certain striking differences in its sequence and interactions with thrombin. Avathrin showed only 40% sequence similarity with variegin. The crystal structure of s-variegin relevaled that His12 of variegin was the residue which interacted with Ser195 of the thrombin active site catalytic triad and disrupted the charge relay system. This residue is replaced by Ser in avathrin sequence, and this Ser12 also interact with the catalytic triad of thrombin in a manner similar to His12 of variegin. Ser12 of avathrin most likely interacts with His57 of thrombin active site hence disrupting the charge relay system of the catalytic triad. Unfortunately this was not directly evident from the crystal structure of avathrin in complex with thrombin, but we have shown the importance of Ser12 using mutagenesis. When Ser12 was replaced with Ala, the potency of the peptide dropped > 15 times, which indicated the importance of Ser at that position for the

inhibitory effect of avathrin. The native variegin (n-variegin) which was the peptide originally isolated from the salivary gland extracts of *A. variegatum*, was glycosylated at Thr14. S-variegin which was synthesized using solid phase peptide synthesis which lacked the glycosylation at Thr14 (due to difficulties in synthesizing the glycosylated peptide) was 14 times less potent than n-variegin. Avathrin sequence however lacked a glycosylation site and instead possessed a Gln at that position. And we hypothesized that this Gln14 could be extensively involved in electrostatic interactions and compensate for the missing glycosyl moiety. Unfortunately, this hypothesis could not be confirmed, since the electron density of Gln14 was not observed in the crystal structure. The variegin sequence contained more acidic residues (10 residues) as compared to the avathrin sequence (6 residues). These acidic residues are important because, the surface of thrombin has exosite-I and exosite-II which are highly basic in nature, and the acidic residues in the C-terminus of variegin were shown to interact with the thrombin exosite-I, while the acidic residues towards the N-terminus of vareigin were hypothesized to steer the molecule towards the exosite-II. Avathrin contained acidic residues only in its C-terminus, which we have shown to interact with exosite-I of thrombin. However, the addition of acidic residues in the N-terminus of avathrin did not improve its potency, indicated that despite possessing a flexible molecular structure, the way avthrin interacts with thrombin is different than variegin-thrombin interaction.

Both variegin and avathrin possess a basic residue at the P1 site which makes them susceptible to thrombin cleavage. Therefore, similar to variegin,

avathrin was also cleaved by thrombin in a time dependent manner. However, avathrin was cleaved at a much slower rate (8-10 hr), compared to vareigin (3 hr). This could be attributed to the presence of more number of glycines in the avathrin sequence. These glycines could make avathrin more flexible in its binding with thrombin, and hence slightly less susceptible to cleavage. This flexibility could also be responsible for the slightly lower potency of avathrin compared to variegin. Similar to MH22, the cleaved product of variegin, IS20, the cleaved product of avathrin, inhibited thrombin amidolytic activity in a non-competitve manner, and hence both variegin and avathrin, retained the ability to exert a prolonged inhibitory effect over thrombin for as long as 36 hr.

Variegin and avathrin are distinct from the other described thrombin inhibitors (discussed in 1.4.3.) and can be classified to form a new family of thrombin inhibitors. Two thrombin inhibitors, tsetse thrombin inhibitor (TTI), isolated from tsetse fly *Glossinia morsitans morsitans*, and NTI-1, isolated from the camel tick *Hyalomma dromedarii* (Barker and Murrell, 2004) have molecular sizes in a similar to variegin and avathrin. Despite, similar molecular sizes, they appear to be unrelated to variegin and avathrin.

In fact, the best comparison of variegin and avathrin could be made with the sequence of bivalirudin. Bivalirudin is a synthetic thrombin inhibitor which has been developed by linking the exosite-I binding C-terminus of hirudin with the active site binding moiety, $_{D}$-Phe-Pro-Arg-Pro (Bourdon, 1991). Bivalirudin is the current anticoagulant of choice for prevention of thrombosis during certain cardiovascular procedures (Mavrakanas and Chatzizisis, 2015). Bivalirudin

competitively inhibits thrombin with an affinity of 3.23 nM, and is cleaved by thrombin. However, unlike variegin and avathrin, which exert prolonged inhibitory effect over thrombin, bivalirudin loses its ability to inhibit thrombin after it is cleaved. An *in vitro* comparison of avathrin with bivalirudin therefore demonstrated that avathrin, which has an affinity of 545 pM is more potent than bivalirudin, and retains the ability to inhibit thrombin for periods longer than bivalirudin. Avathrin may therefore perform as a better anticoagulant than bivalirudin for the control of cardiovascular disorders.

We have shown that avathrin is able to inhibit free as well as clot bound thrombin and have evaluated the efficacy of avathrin to inhibit thrombosis in an established ferric chloride carotid artery thrombosis model. With observations from these experiments and inferences from other DTIs, we are certain that avathrin can be used for several clinical indications and cardiovascular procedures such as (1) prevention of arterial thrombosis and reocclusion during invasive procedures such as percutaneous coronary intervention (PCI); (2) venous thrombosis prophylaxis after an orthopaedic surgery such as hip or knee replacement; (3) management of myocardial infarction (MI) (Bates and Weitz, 2005). We have also compared the efficacy of avathrin with bivalirudin using in murine models, and demonstrated that avathrin may indeed perform as an anticoagulant with better efficacy. Safety and efficacy are the two most basic clinical parameters which are used for comparison of performance of anticoagulants. Safety and efficacy of different anticoagulants vary due to their inherent differences in pharmacokinetics, pharmacodynamics, and interactions

with other plasma proteins. Therefore, although initial experiments with small animals may be a starting point for the development of antithrombotics, details about pharmacokinetics, pharmacodynamics, and other side effects should be considered before accurate clinical comparisons could be made. We plan to conduct pharmacokinetics/ pharmacodynamics of variegin, avathrin and few other peptides (ultravariegin, described in 5.3.) to identify the peptide which would have the best safety-effficacy balance and develop that peptide as a therapeutic.

CHAPTER 5

Novel family of thrombin inhibitors from ixodid ticks

5.1. Introduction

The crucial role of thrombin in haemostasis (discussed in section 1.1.3 and 4.1) is reflected in the prevalence of the variety of thrombin inhibitors in the saliva of haematophagous animals (Dodt et al., 1996). These thrombin inhibitors, which are mostly proteinaceous molecules, fall into three categories: canonical, non-canonical and exosite binding inhibitors and are grouped into distinct families based on similarities in their structure, mode of binding and mechanism of inhibition (discussed in detail in 1.4.3.5). Hirudin and hirudin like, kunitz, kazal, lipocalin, madanins are some of the examples of different families of thrombin inhibitors from which type members have been characterized for their structure-activity relationship (Huntington, 2014). These families usually possess unique secondary and tertiary structures, with several disulfide bonds that give the inhibitor a typical fold that confers the unique mechanism of inhibition. These inhibitors that bind and inhibit one of the three functional sites of thrombin (thrombin active site, exosite-I, or exosite-II) are called univalent inhibitors. Inhibitors binding to two of these sites on the surface of thrombin (a combination of active site and exosite-I/exosite-II, or exosite-I and exosite-II) have also been studied and are named bivalent inhibitors (Warkentin, 2004).

The unique structure of thrombin in comparison to the other serine proteases of the blood coagulation cascade, allow only the binding of highly selective substrates and inhibitors to the thrombin active site (Pechik et al., 2004). The thrombin active site contains the classical catalytic triad formed by His57, Asp102 and Ser195. The active site is placed in a canyon like cleft which is

buried in between two insertion loops (60- and autolysis loops). Thrombin active site of has an acidic S1 subsite, hence it preferentially cleaves substrates with a basic side chain at the P1 position of the scissile bond (Fuentes-Prior et al., 2000; Polgár, 2005). Furthermore, exosite-I and exosite-II which are two highly basic surfaces on two opposing sides of the active site form specialized patches which allow only specific substrates to enter the active site cleft (Myles et al., 2001). Owing to this specialized structure, thrombin inhibitors possess complimentary surfaces which can bind to these specific sites of thrombin and selectively inhibit it with very high potencies (Malovichko et al., 2013).

Out of the described families of thrombin inhibitors, madanins, which bind in a substrate like fashion are quickly processed by thrombin resulting in fragments that are devoid of thrombin inhibitory activities (Figueiredo et al., 2013). Variegin and avathrin are the two novel thrombin inhibitors from the salivary gland extracts of *Amblyomma variegatum*, which we have described as highly selective thrombin inhibitors. These inhibitors, like madanins are processed by thrombin. However, but unlike madanins, they retain the ability to inhibit thrombin post processing (Koh et al., 2009). In this chapter, we describe a novel family of thrombin inhibitors from ixodid ticks, which bear certain key similarities with variegin and avathrin, bind to thrombin with a fast, tight binding mode, and inhibit it with high affinities. These inhibitors are processed by thrombin, and retain their ability to inhibit thrombin after they are cleaved. We hypothesize that some of these inhibitors possess properties that would allow

them to bind thrombin in a trivalent fashion, by binding to the active site as well as to thrombin exosite-I and exosite-II.

5.2 Materials and methods

5.2.1. Identification of peptide sequences from Ixodid tick transcriptomes

Peptide sequences which were similar to variegin and avathrin were identified by performing a standalone BLAST analysis of published transcriptomes of *Amblyomma variegatum, Rhipicephalus pulchellus, Amblyomma americanum, Amblyomma cajenesse, Amblyomma maculatus* and *Dermacentor marjinatum rupifies* (Cavassani et al., 2005; Nene et al., 2002). These sequences were manually aligned with variegin and avathrin and one peptide from each tick was selected for further analysis.

5.2.2. Peptide synthesis and purification

All selected peptides (ultravariegin and peptides from *R. sanguineus, A. americanum, A. cajenesse, A. maculatum* and *D. marjinatum rupifies*) were synthesized using solid-phase peptide synthesis and purified using reversed phase HPLC as described in 4.2.2.. The purity and mass of all peptides were determined by electrospray ionization mass spectrometry using an LCQ Fleet Ion Trap mass spectrometer from Thermo Fisher Scientific (Waltham, MA, USA). These peptide sequences and molecular masses of the selected peptides are given in (Table 5.1).

5.2.3. Inhibition of thrombin amidolytic activity

Thrombin amidolytic activity assays for all peptides were performed in 96-wells microtiter plates as described in Section 4.2.3. All reactions were carried out at room temperature in 50 mM Tris buffer (pH 7.4) containing 150 mM NaCl. The resulting reaction rates were used to calculate the IC_{50} and Hill slope.

Table 5.1. Sequences and molecular weights of variegin and avathrin like sequences from other hard ticks.

Sequence	Tick	Molecular weight (Da)
SDEAVRAIPKMYSTAPPGDFEEIPDDAIEE	*A. variegatum*	3293.5
SGEDHTAVPKMSRKGLGGDFEDIPPEAYERALEAR	*A. americanum*	3830.4
ELESGDEDSEGGDSQSSPTESAAPRLHQREGGGGDFENVEYDQDQK	*A. maculatum*	4940.6
SDVAPADYESDEGDNDGGHDGSEVAKPKMPRGNGGGGDFEEIPEVE	*R. sanguineus*	4773
TGSDDDDEYDMYESDGDSNEGNDNDEFETAVPRLPNPNSGRDSEHIPMPVN	*D. marginatum rupifies*	5660.73

Assays for determining the inhibitory constants were also carried out in a similar way. Typically, 100 µl of thrombin (0.81 nM) was added to wells containing 100 µl of different concentrations of the peptide and 100 µl of different concentrations of S2238 (100 µM). The resulting reaction rates and residual velocities were used to calculate the inhibitory constants (K_i).

5.2.4. Inhibition of thrombin fibrinogenolytic activity

The abilities of all peptides to prolong fibrinogen clotting time were tested as described in 4.2.4.

5.2.5. Serine protease selectivity

The selectivity profile of all peptides was examined against 13 serine proteases: anticoagulant serine protease APC, procoagulant serine proteases (FXIIa, FXIa, FXa, FIXa, FVIIa, kallikrein, and thrombin), fibrinolytic serine proteases (plasmin, TPA, and urokinase), and classic serine proteases (chymotrypsin and trypsin). All assays were carried out with respective substrates as described in Section 4.2.5.

5.2.6. Cleavage of peptides by thrombin

The ability of thrombin to cleave ultravariegin and peptide from *R. sanguineus* was tested by incubating 150 µM ultravariegin (or *R. sanguineus* peptide) with 5 µM thrombin. The reaction mixtures were incubated in a 50 mM Tris buffer (pH 7.4) containing 150 mM NaCl and 1 mg/ml BSA. After different incubation times, reactions were quenched with 1% TFA (pH 1.8) and loaded onto a Dionex Acclaim100 PepMap microbore column (100 x 1.0 mm) attached

to a Dionex nano-HPLC system and eluted using an acetonitrile gradient with 0.1% TFA and 99.9% milliQ water as eluent A and 0.1% TFA, 19.9% milliQ water and 80% ACN as eluent B. New peaks other than those present in the control chromatogram were taken as the peaks of cleaved products. The masses of the cleaved products were confirmed by analyzing their masses using electrospray ionization mass spectrometry. Cleavage products were quantified by integrating the peaks and calculating the area under the curves.

Effect of different pre-incubation times (hence cleavage), on the thrombin inhibitory activity of ultravariegin was measured in a way similar to that measured for avathrin as described in 4.2.6.

5.3. Results

5.3.1. Inhibition of thrombin amidolytic activity

The ability of different thrombin inhibitor peptides to inhibit the amidolytic activity of thrombin was assayed using S2238 as the substrate. The selected peptides from different ticks inhibited thrombin amidolytic activity in a dose-dependent manner. IC_{50} and Hill coefficient of all these peptides are outlined in Table 5.2. Of special interest was one peptide from the sialome of *A. variegatum* which had an IC_{50} of 0.5 nM. Since this peptide showed 50% sequence identity with variegin and was more potent to inhibit thrombin, this peptide was named as ultravariegin. Significant inhibition was observed at equimolar concentrations of ultravariegin and thrombin (~80% inhibition at 0.81 nM ultravariegin) indicating that ultravariegin is a tight binding thrombin inhibitor (Fig. 5.2.). Reaction progress curves showed that steady-state equilibrium was achieved upon mixing indicating a fast binding mode of ultravariegin. Residual reaction velocities of thrombin amidolytic activity in presence of ultravariegin at different S2238 concentrations were determined to calculate K_{iapp}. A straight line plot of the linearly increasing K_{iapp} with increasing S2238 concentrations indicated that ultravariegin is a competitive inhibitor and these K_{iapp} were fitted to the Morrison equation using the GraphPad Prizm software and a Ki of 1.5 ± 0.19 pM of ultravariegin towards thrombin was obtained. Therefore, ultravariegin is a fast, tight binding, competitive inhibitor of thrombin.

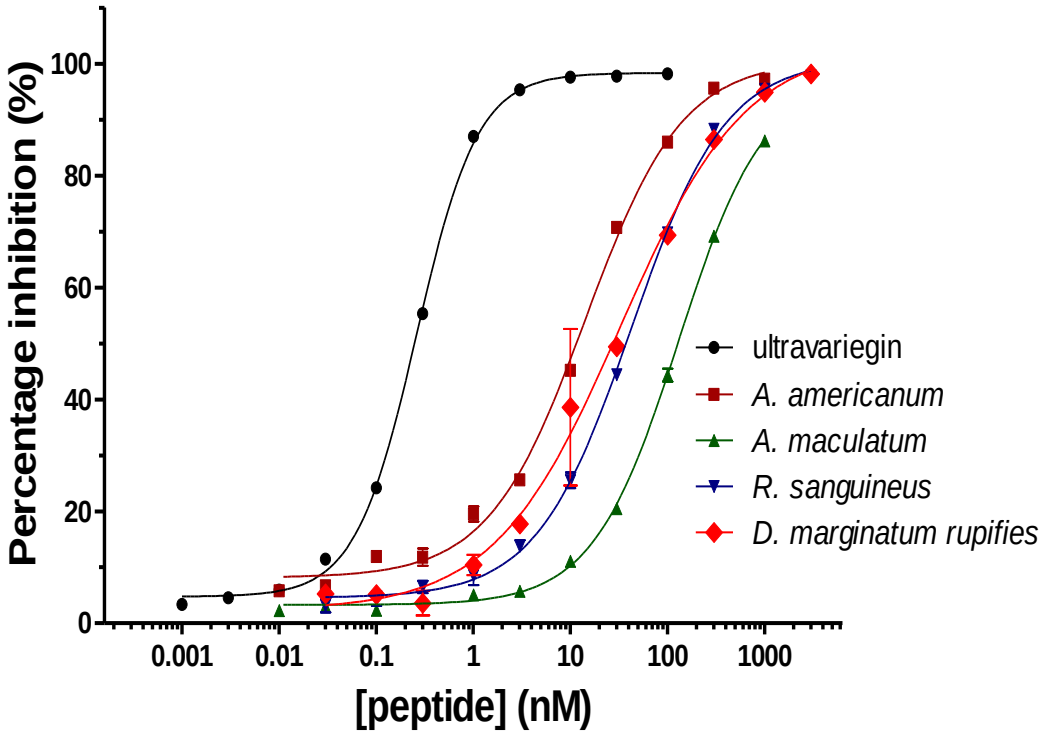


Figure 5.1. Inhibition of thrombin amidolytic assay by peptides of variegin family. Selected peptides at different concentrations (3000- 0.0003 nM) were tested for their abilities to inhibit thrombin. Ultravariegin, the peptide identified from the sialome of *A. variegatum* was the most potent peptide with an IC_{50} of 0.26 ± 0.008 nM. All other peptides had IC_{50}s in the range of 14 to 130 nM.

Table 5.2. IC_{50} and K_i values of members of variegin family

Peptide	IC_{50} (nM)	Affinity (nM)
Avathrin	6.95 ± 0.42	0.545 ± 0.002
Variegin	4.17 ± 0.93	0.283 ± 0.01
Ultravariegin	0.26 ± 0.008	0.001
A. americanum	14.29 ± 0.12	-
A. maculatum	130.20 ± 1.74	-
R. sanguineus	42.38 ± 0.62	8.79 ± 0.61
D. marginatum rupifies	32.48 ± 3.94	6.135 ± 0.39

(Experiments for determining K_i values for *A. americanum* and *A. maculatum* are not yet completed)

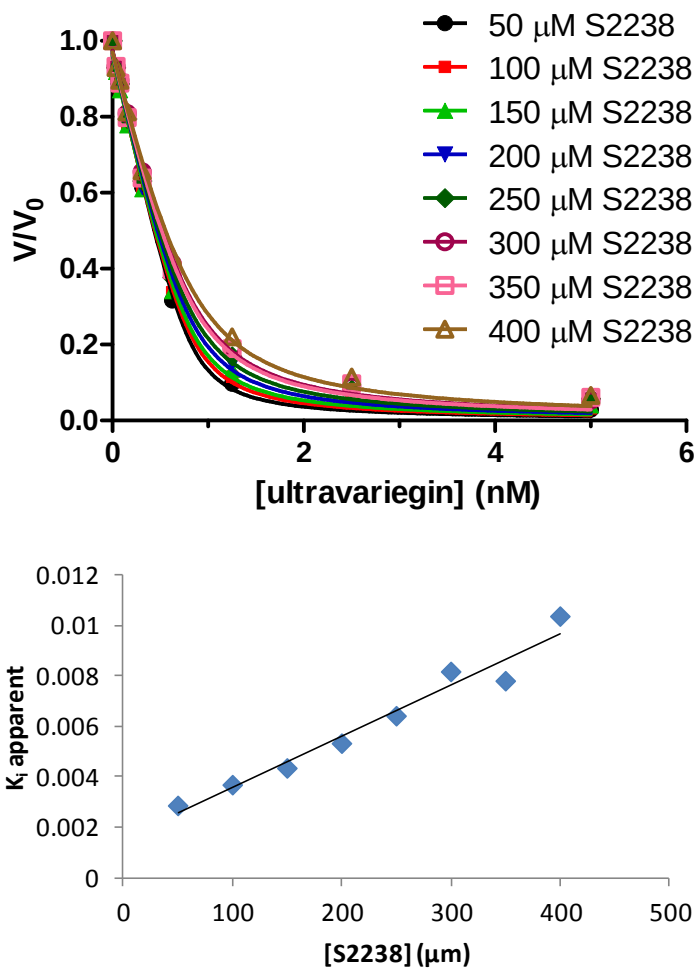


Figure 5.2 Inhibitory constant K_i of ultravariegin. K_i of ultravariegin, the most potent member of the vareigin family is shown as a representative. Ultravariegin is a tight binding inhibitor of thrombin. Different concentrations of ultravariegin were mixed with different concentrations of S2238 (50 μM, 100 μM, 150 μM, 200 μM, 250 μM, 300 μM, 350 μM and 400 μM and the Ki' was determined. Reactions were started with the addition of thrombin (0.81 nM). Data were fitted to the Morrison equation equation using GraphPad prizm software (n = 3, error bars represent S.D.). (B) Plot of K_i' against substrate concentration increased linearly, indicating ultravariegin competitively inhibited thrombin amidolytic activity and the inhibitory constant K_i was determined to be 1.5 ± 0.03 pM (error bars represent S.D.).

5.3.2. Inhibition of thrombin fibrinogenolytic activity

The ability of all thrombin inhibitor peptides from the ixodid ticks to inhibit thrombin's fibrinogenolytic activity was tested. Ultravariegin and all other peptides prolonged fibrinogen clotting time in a dose-dependent manner (data not shown). This observation is consistent with observations for variegin, avathrin and the C-terminus of hirudin which also inhibit thrombin's fibrinogenolytic activity.

5.3.3. Serine protease selectivity

The serine protease selectivity of ultravariegin and the other peptides was examined by screening against 13 serine proteases which included pro- and anti-coagulant serine proteases of the blood coagulation cascade and the two classic serine proteases trypsin and chymotrypsin. At 10 nM, ultravariegin inhibited 94.56% of thrombin's activity. However, even at 1 μM, ultravariegin showed no significant inhibition (<5%) of any of the serine protease other than thrombin (Fig. 5.3). The other thrombin inhibitor peptides also showed high selectivities for thrombin. Thus, these thrombin inhibitors from ixodid ticks are highly potent and selective (more than 4 orders of magnitude) thrombin inhibitors.

5.3.4. Cleavage of ultravariegin by thrombin

Because ultravariegin binds to the thrombin active site, we hypothesized that like variegin, avathrin and the other described thrombin inhibitors, ultravariegin (and peptide from *R. sanguineus*) may be prone to proteolytic cleavage by thrombin. Preliminary experiments with ultravariegin and the peptide from *R. sanguineus* have shown that these peptides are cleaved by thrombin.

Preliminary experiments have also shown that ultravariegin is cleaved completely after 16-18 h and further experiments to confirm this observation are currently being carried out.

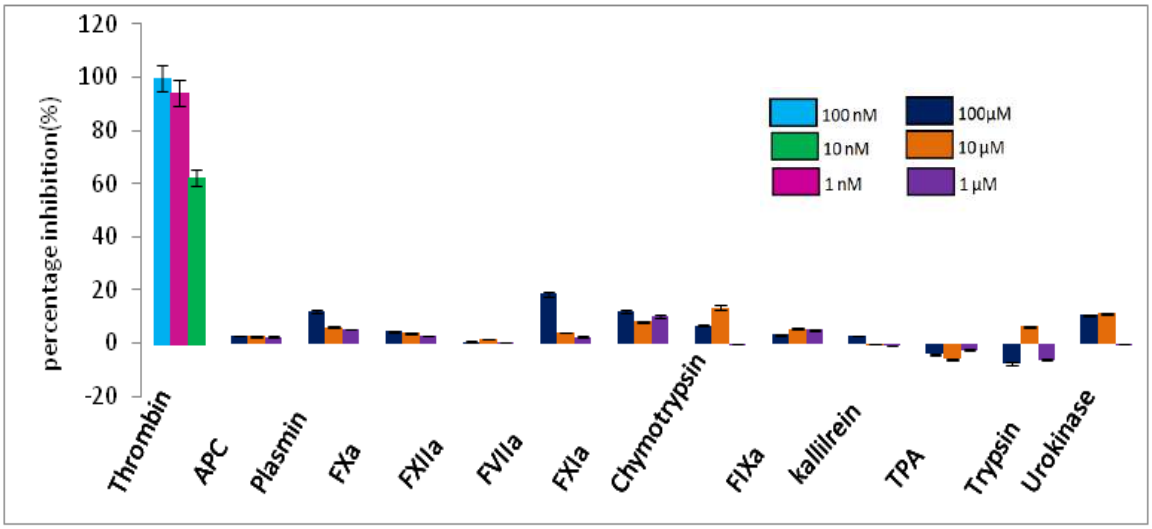


Figure 5.3. Serine protease selectivity of ultravariegin. The serine protease selctivity of all members of variegin family was determined by screening them against 13 serine proteases: procoagulant serine proteases (FXIIa, FXIa, FXa, FIXa, FVIIa, kallikrein and thrombin), anticoagulant serine protease APC, fibrinolytic serine proteases (plasmin, tPA and urokinase), and classical serine proteases (chymotrypsin and trypsin). The selectivity profile of the most potent peptide, ultravariegin is shown in this figure. All members of this family are more than 3-5 orders of magnitude more selective towards thrombin than the other serine proteases of the blood coagulation cascade.

5.4. Discussion

We have successfully demonstrated the presence of a novel families of thrombin inhibitors in hard ticks and named this family as the 'variegin family'. These peptides which are 25-50 residues long selectively inhibit thrombin with affinities in the nanomolar to picomolar range. The absence of cysteines in these peptide sequences differentiates them from the other prototypic thrombin inhibitors such as hirudin, haemadin, triabin and bothrojaracin which are discussed in earlier chapters.

These peptides possess certain striking similarities because of which they are classified as a separate family of thrombin inhibitors. All these peptides contain a basic residue at the P1 site and are cleaved by thrombin. The C-termini of these peptides contain acidic residues with which they are able to bind to the thrombin exosite-I. An interesting observation about the peptides from *R. sanguineus* and *D. marginatum rupifies* is that in addition to the acidic residues at the C-terminus which binds to exosite-I, these peptides have an extended N-terminus which contains several acidic residues that could bind to thrombin exosite-II. Thrombin exosite-II, which is also referred to as the heparin binding exosite hosts a number of positively charged residues (She et al., 2014). Heparin enhances the inhibition of thrombin by antithrombin by mediating its interaction with the thrombin exosite-II. Heparin first forms a complex with antithrombin and this complex then docks into the exosite-II by electrostatic coupling. Additionally, exosite-II is also involved in the interaction with the platelet receptor, glycoprotein Ibα (GPIbα). Cleavage of the PAR1 by thrombin proceeds in a

GPIbα dependent manner (Wallace and Smyth, 2013). Several inhibitors binding to the thrombin active site and to exosite-I have been studied and are discussed in 1.4.3.

Bothrojaracin and haemadin are two inhibitors which are known to bind to thrombin exosite-II (Strube et al., 1993). Bothrojaracin, a C-type lectin like protein binds to both thrombin exosite-I and exosite-II without binding to the active site (Zingali, 1993). Haemadin, the slow tight-binding thrombin inhibitor derived from the land living leech docks to the non-prime subsite of thrombin with its N-terminal segment in a manner similar to hirudin. In contrast to hirudin, the extended acidic C-terminus of haemadin does not interact with the exosite-I, but binds to the exosite-II forming direct intermolecular salt bridges (Richardson et al., 2000).

Trivalent thrombin inhibitors inhibitors which bind to the thrombin active site and both the thrombin exosite-I and exosite-II simultaneously have not been described yet. We hypothesize that the long peptides of the variegin family may form a group of trivalent thrombin inhibitors. We have shown that these peptides bind the active site and exosite-I of thrombin using standard assays, and are currently carrying out experiments to prove that the extended acidic N-terminus may bind to the exosite-II. Because of the presence of the basic P1 residue, these peptides are cleaved by thrombin, and similar to variegin, avathrin and ultravariegin, the cleaved products of these peptides may retain their inhibitory property against thrombin.

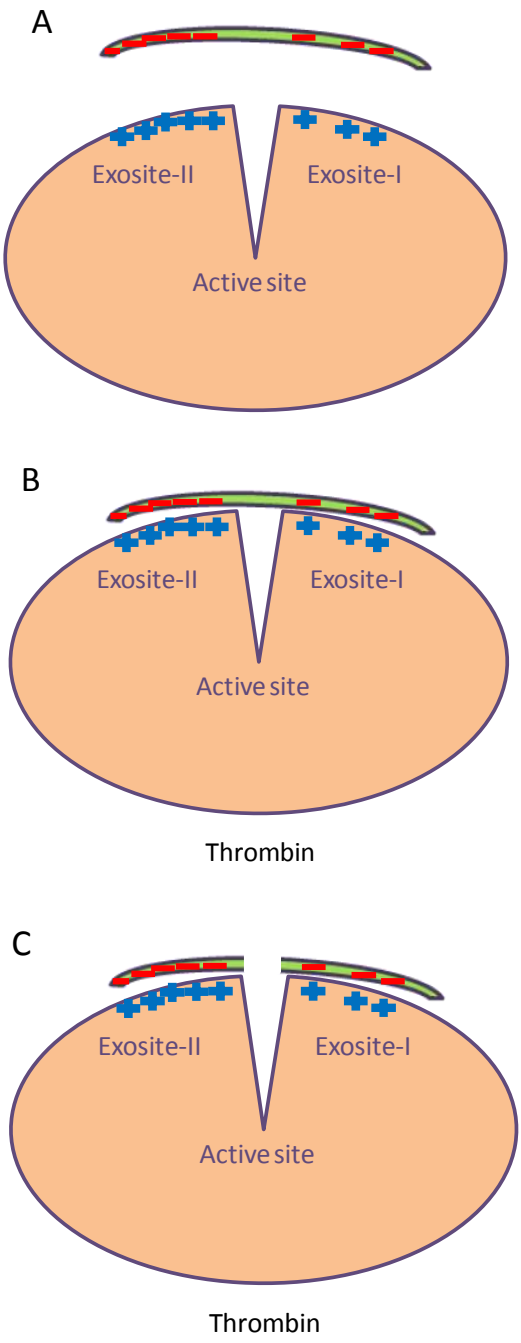


Figure 5.4. Hypothesized mode of binding of trivalent inhibitors of the vareigin family. A. Initial recognition of the thrombin active site, exosite-I and exosite-II by the C- and N-termini. **B.** C-terminus would bind to the exosite-I and N-terminus would bind to exosite-II. **C.** Due to the presence of a basic P1 residue,

the peptide would be cleaved, and N-terminus would remain bound to the exosite-II and C-terminus would remain bound to exosite-I.

Therefore, we propose that these trivalent inhibitors first bind to thrombin active site and to exosite-I and exosite-II and they get cleaved by thrombin at the P1 residue. Subsequently, the cleaved fragments may remain bound to the two thrombin exosites (Fig 5.4.). The variegin family of thrombin inhibitors could be described as a novel molecular scaffold in hard ticks. Generally different sets of scaffold are evolved by hematophagous to target and inhibit various stages of the blood coagulation cascade.

All molecules within one set possess a basic scaffold, which are further tailored to target the same or a different serine protease with a unique mechanism. For example, exogenous serine protease inhibitors from ticks are largely of the Kunitz type, whereas those from leeches are of the hirudin type. This is due to the evolution of different molecules of these sets from a common ancestral molecule. Despite having a common molecular scaffold, in come cases, different members within a group may show striking structural and mechanistic differences. These differences may be due to the adaptation to different hosts, or their evolution to target the same or different serine protease under a different condition. Hirudin, hirullin and bufrudin are examples of thrombin inhibitors within a single class. Yet hirullin and bufrudin show deviations from the features that confer its inhibitory property on hirudin. While *Hirudo medicinalis,* the leech from which hirudin is obtained is an amphibian parasite, *Hirudinaria manillensis,* from which both hirullin and bufrudin have been isolated is primarily a mammalian parasite. This suggests that hirullin and bufrudin may be tailored to inhibit mammalian

thrombin in contrast to hirudin which may be most efficient in inhibiting the amphibian enzyme.

Similarly, we have observed that although members of the variegin family possess a basic scaffold, they show striking differences in certain key properties like their affinities to inhibit thrombin, times at which they are cleaved by thrombin, and the differences in their molecular interactions with thrombin as revealed by the crystal structures. Variegin and avathrin which show 40% sequence similarity were able to inhibit thrombin amidolytic activity with similar affinities while (300-500 pM), ultravariegin which shows 50% sequence similarity with variegin inhibited thrombin amidolytic activity with >200 fold higher potency compared to variegin (~ 1.5 pM). The key residue of variegin, His12, which was shown to disrupt the catalytic triad of thrombin, was replaced by Ser in avathrin and Tyr in ultravariegin. The Ser12 in avathrin plays a role similar to His12 of variegin in disrupting the catalytic triad of thrombin and is confirmed by mutations (described in detail in 4.3.). Tyr12 in ultravariegin may interact with the thrombin catalytic triad using a novel mechanism which could confer on it a higher inhibitory potency compared to variegin and avathrin. We are currently crystallizing ultravariegin in complex with thrombin, to study the details of its structure-activity relationship.

We are currently also carrying out safety and efficacy experiments in animal models to study and compare their *in vivo* effects. At the clinical level, these three molecules may be developed to target different indications depending to their potencies. Unfractionated heparin (UFH), the most widely used

anticoagulant for cardiovascular disorders shows severe bleeding side effects and causes HIT (Coppens et al., 2012). Bivalirudin which is currently the anticoagulant of choice for percutaneous coronary intervention (PCI) has recently been shown to cause severe bleeding and intracranial haemorrhaege (Yeh et al., 2015). These limitations indicate the need for novel anticoagulants with better safety-efficacy balance that may prove as superior therapeutics. Variegin and avathrin, which have >5 times higher potency than bivalirudin may be developed for use in PCI, depending on their safety profiles. Ultravariegin, which is 1500 times more potent than bivalirudin, may be developed for use in cardiopulmonary bypass (CPB) and extracorporeal membranous oxygenation (ECMO). CPB and ECMO, in which blood is exposed to highly thrombogenic surfaces, require anticoagulants with extremely high potencies, and currently heparin is the only anticoagulant used for this purpose. Heparin causes unmanageable side-effects like severe bleeding and HIT, and no other drug has proven to be safe and efficacious enough for effectively managing these patients undergoing these two procedures. We are currently evaluating the safety-efficacy balance of ultravareigin in animal models and plan to develop it to as a therapeutic for these procedures.

CHAPTER 6

Factor Xa inhibitors from salivary gland extracts of female *Rhipicephalus pulchellus*

6.1. Introduction

Based on the theory of interaction of coagulation factors, and on the results of clinical outcomes, there is accumulating evidence that FXa may be a better target for inhibition than thrombin (Mousa, 2008). FXa occupies a crucial juncture in the blood coagulation cascade and produces thrombin. The activation of one molecule of FX to FXa can result in the generation of 1000 molecules of thrombin (Davie et al., 1991). On a molar basis, FXa is more thrombogenic than thrombin and it has also been shown that less heparin is required to inhibit thrombosis prior to thrombin formation than afterwards (Lee and Ansell, 2011). Additionally, FXa has limited roles other than producing thrombin.

Thrombin has several important activities within and outside of the haemostatic system and therefore the inhibition of thrombin may hamper other physiological processes (Borissoff et al., 2009; Myles and Leung, 2008). Some of the important prothrombotic roles of thrombin within the haemostatic system are its positive feedback on the coagulation factors to amplify its own production and its ability to activate platelets. In addition thrombin also has antithrombotic roles by activating protein C and thrombin activatable fibrinolysis inhibitor. On the other hand, FXa does not possess any role other than thrombin geneneration. FXa inhibitors have been shown to be more efficacious than DTIs in decreasing the endogenous thrombin potential and in prolonging the lag phase in thrombin generation time assays (Szlam, 2007). Therefore, FXa inhibitors may be interesting targets, especially to overcome the limitations posed by thrombin inhibitors. Whether or not FXa inhibitors are indeed better than thrombin

inhibitors with superior safety and efficacy profiles will remain controversial until direct clinical comparisons are made.

The role of antihaemostatic molecules in the saliva of ixodid ticks has already been discussed in the previous chapters. In order to identify novel FXa inhibitors from tick saliva, we analyzed the salivary gland extracts of the zebra tick, *Rhipicephalus pulchellus*. We have studied the anti-FXa properties of the salivary gland extracts of *R. pulchellus* and partially purified the FXa inhibitor. We are currently carrying out further purification of this protein and we will be studying the detailed structure-activity relationship of this protein to develop it as an antithrombotic agent.

6.2. Materials and Methods

Salivary gland extracts of male and female *R. pulchellus* which were fed for 6 days on the host were prepared and the amount of protein was estimated as explained in 3.2. Thrombin and FXa inhibitory actitivities of the crude extracts and purified fractions was tested in 384 well plates as described in 4.2.

For the purification of the FXa inhibitor, female salivary gland extracts were separated using size exclusion chromatography with a Superdex 75 10/300 column (GE Healthcare, Uppasala, Sweden). The column was first equilibrated with 25 ml of 50 mM Tris buffer (pH 7.4), and 100 salivary gland pairs reconstituted in 1 ml MilliQ water were injected into the column. Proteins were eluted using the same buffer at a flow rate of 0.8 ml/min and elution was monitored at UV wavelengths 215 nm and 280 nm and 0.8 ml fractions were collected.

Fractions from the size exclusion chromatography which showed FXa inhibition were further purified using FXa affinity chromatography. FXa was first coupled to CNBr-Activated Sepharose beads (GE Healthcare, Uppasala, Sweden) according to manufacturer's protocol. 1 ml of FXa coupled beads were used in a 10 ml poly-prep column (Bio-rad, California, USA). Protein samples were incubated with the beads for 30 min, with gentle rocking. The column was then washed with 40 ml Tris buffered saline and proteins were eluted using increasing concentrations of NaCl (0.5 M, 1.0 M, 1.5 M and 2.0 M) in 20 mM HCl. Proteins were eluted in tubes containing 1M Tris buffer, pH 8.0.

Eluted fractions from the FXa affinity chromatography which showed FXa inhibition were further purified using reverse phase chromatography using mobile phase eluent A as 84.9% H_20/15% ACN/0.1% TFA and eluent B as 99.9% ACN/0.1% TFA. An Agilent Zorbax 300SB-C18 column was first equilibrated with eluent A. 20 µl protein samples were then loaded into the column and eluted using a linear gradient of 5% eluent B to 70% eluent B in 70 min at a flow rate of 40 µl/min. The elution was monitored at UV wavelengths 215 nm and 280 nm and 60 µl fractions were collected, lyophilized and tested for anti-FXa activities.

6.3. Results

6.3.1. Protein quantification of *Rhipicephalus pulchellus* salivary gland extracts

Similar to *D. reticulatus*, there was a huge difference in the protein content in the salivary gland extracts of male and female *R. pulchellus*. The total protein content from the female ticks (43.72 µg/tick) was 5-times > than the total protein content of the male ticks (8.34 µg/tick).

6.3.2. Activity of crude salivary gland extracts

Equivalent amounts of male and female crude salivary gland extracts were assayed for their anti-thrombin and anti-FXa activities. Fig. 6.1. shows the percentage inhibition the male and female salivary glands had towards FXa. While the female salivary gland extracts were able to inhibit thrombin and FXa (92%), the male salivary glands showed little inhibition of thrombin (18%) and FXa (4%) amidolytic activity. We proceeded with the purification the FXa-inhibitor protein.

6.3.3. Purification of factor Xa inhibitor salivary gland extracts

Female salivary gland extracts of *R. pulchellus* were fractionated using size exclusion chromatography and their anti-FXa activities were measured as described in 6.2. FXa inhibitory activity was observed in two protein peaks (FXaI-1 and FXaI-2) (Fig. 6.2).

For further purification of the FXa inhibitor, an affinity chromatography was carried out using a FXa affinity column. Fractions containing FXaI-1 from

the size exclusion chromatography were pooled and incubated in the FXa affinity column for 30 min before washing away the unbound sample. The sample was eluted with increasing salt concentration and all fractions were tested for their FXa inhibitory activity. Most of the FXa had eluted in the 1.5 M NaCl fraction.

The active fraction from the affinity chromatography was further purified using a reverse phase microbore column as described in 6.2. There were two active inhibitor peaks in the reverse phase chromatogram which were well separated (data not shown). However due to limited amount of protein obtained from the tick salivary gland extracts, FXa inhibitor could not be sequenced and studied further.

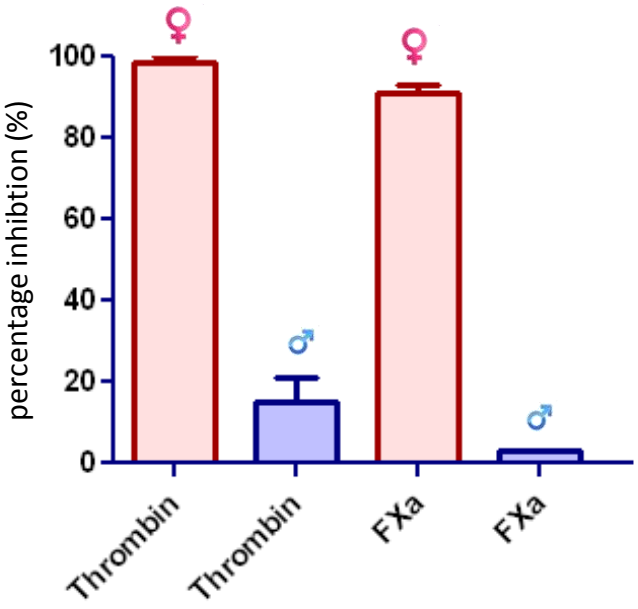


Figure 6.1. Anti-thrombin and anti-FXa activity of female *R. pulchellus* salivary gland extracts. Equivalent amounts of male and female crude salivary gland extracts were assayed for their anti-thrombin and anti-FXa activities. Female salivary gland extracts were able to inhibit thrombin and FXa (92%), male salivary glands showed little inhibition of thrombin and FXa amidolytic activities.

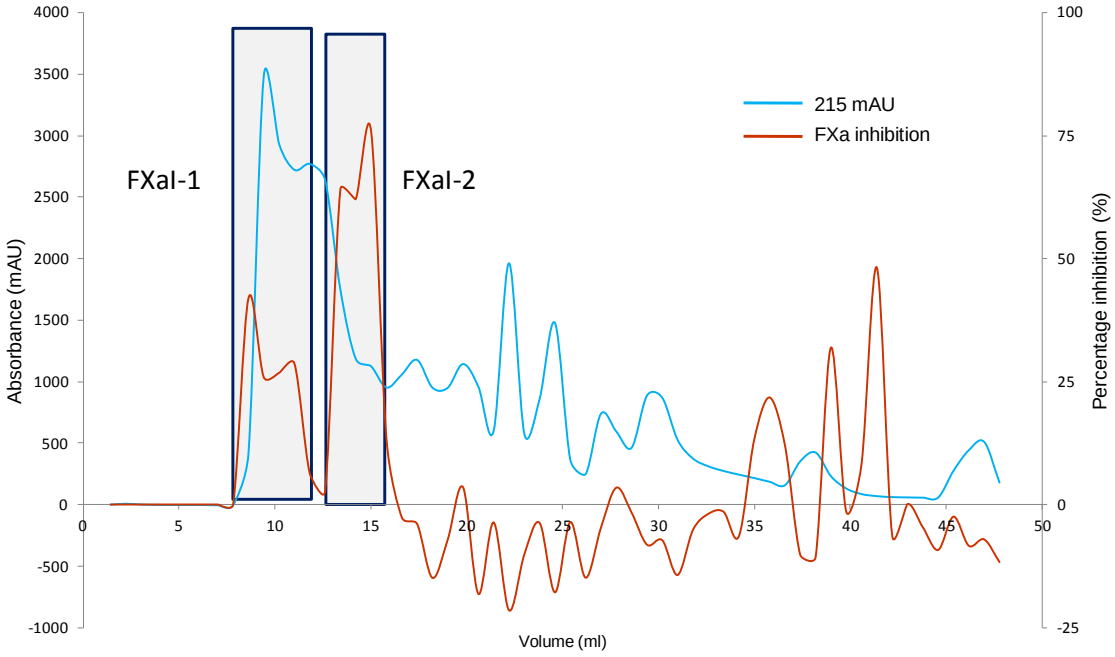


Figure 6.2. Gel filtration chromatogram of female *R. pulchellus* salivary gland extracts. Fractions collected from size exclusion chromatograhy were tested for their anti-FXa activities. An overlay of the protein elution chromatogram with the anti-FXa activity was plotted.

6.4. Discussion

We have successfully demonstrated that salivary gland extracts of female *R. pulchellus* contain anti-FXa activity and we have developed a strategy to purify this inhibitor. Using three types of chromatographies, the FXa inhibitor could be purified. However, the minute amounts of FXa inhibitor in the salivary gland extracts made the sequencing of this inhibitor a herculean task. To purify the FXa inhibitor for further analysis, several thousands of ticks would be required. We are trying to rear more ticks to isolate the FXa inhibitor sufficient for sequencing the protein. We are also adapting alternative approaches to identify the inhibitor.

CHAPTER 7

Conclusions and future perspectives

7.1. Conclusions

Haemostasis is the physiological process that prevents blood loss following a vascular injury. Blood coagulation, which is an important part of haemostasis, involves a cascade of events which finally culminate to form a stable blood clot (Davie et al., 1991). The players of the blood coagulation cascade are activated one after the other in a stepwise manner and among all players; thrombin is the most important serine protease. Thrombin is an effector enzyme playing paradoxical roles both as a procoagulant and an anticoagulant, hence maintaining the delicate balance between uncontrolled bleeding and formation of obstructive thrombi, and allowing sufficient thrombus formation when desired.

Abnormalities in haemostasis result in unwanted clots which in turn lead to the formation of cardiovascular disease. These diseases are the single largest killer worldwide and is a heavy contributor to the burden of non-communicable diseases. Anticoagulants like direct thrombin inhibitors and direct FXa inhibitors are the most commonly used therapeutic options for the control of these unwanted clots.

Traditional development of anticoagulants from natural complex mixtures has rather been a cumbersome procedure involving lead separation, identification, characterization and development as separate procedures demanding intensive manual labour and copious amounts of starting material. Out of the two issues, obtaining large amount of starting material from natural sources has been a severe concern and has seriously hampered drug development.

To address this concern, we have developed a high throughput platform for the identification of functionally active molecules targeting the blood coagulation cascade and established the proof-of-principle concept with the two major enzymes targeted for the control of unwanted clot formation- thrombin and FXa. We have developed an on-line post-separation bioassay in which inhibitors for these two enzymes from complex mixtures can be separated, identified for functional activity and their exact masses can be determined simultaneously. In this novel approach, we have coupled a nano-HPLC to a microfluidic chip based bioassay system and a mass spectrometer. The instrumentation is designed in such a way that eluate from the nano-HPLC is split into two equal parts wherein one part is fed into a mass spectrometer (which identifies exact mass) and another part is fed into a microfluidic chip where the enzyme assay takes place (which determines the functional activity of the molecule). This nano-HPLC coupled to a microfluidic bioassay system reduces the amount of starting material and reagent consumption by about 100 times than that used by the conventional approaches.

To identify novel molecules with potent pharmacological activities, we have looked into the saliva of ticks. Ticks are obligate parasites that obtain a blood meal from their host. The host however, restrains the foreign invader from accessing its blood by various defense mechanisms such as vasoconstriction, blood coagulation and immune reaction. To counter these host defense mechanisms, the blood sucking parasites infuse their saliva which is a complex cocktail of vasodilatory, anti-clotting, anti-platelet and immune suppressors, into the host blood at the blood feeding site. The evolution of hematophagy has

occurred on at least 20 independent times in the arthropod genera and at each of these occasions, novel mechanisms and distinct protein scaffolds to target the host hemostatic and immune system have been adopted. Hence tick saliva presents a unique pharmacopoeia which can be explored for novel proteins with enormous significance.

Ixodid ticks - a specialized group among arthropods, are long term feeders that remain attached to the host and feed on the host blood for periods as long as 9-12 days. However, striking differences in the feeding behaviours between male and female ticks of the same species are observed. For example, the female ticks feed and increase in size by about 100 times of the unfed body weight, while the male ticks grow barely about 2 times their body weight after feeding.

Because it is through the saliva that ticks obtain the host blood, tick feeding triggers the expression of new protein in the saliva, and these newly synthesized proteins may be the regulators which disarm the host defense mechanisms enabling prolonged tick feeding. Therefore, we have carried out transriptomic (Illumina) and quantitative proteomic (iTRAQ) profiling of the tick salivary gland extract of an ixodid tick, *Dermacentor reticulatus* at different stages of feeding to identify sex specific differences and the special proteins implicated in tick feeding. With this deep sequencing approach, we have identified more than 30,000 transcripts in the transcriptome, and over 400 proteins in the proteome of the male and female ticks. We have provided evidence that feeding stage-specific expression occurs in tick saliva, and the alteration in the levels of these proteins is what mediates prolonged tick feeding. We are currently

carrying out recombinant expression of some of these proteins to study their structure-activity relationships. Because these proteins are mainly anticoagulants, anti-platelets, vasodilators and immune suppressors, this library of tick saliva not only presents a unique pharmacopoeia of molecules which can be developed into drugs, but also pinpoints towards specific sequences which are unique and most interesting members of these specific families.

DTIs and direct FXa inhibitors have been the most preferred anticoagulants for the prevention and control of cardiovascular disorders. Despite being the most sought after options, these anticoagulants are fraught with limitations and do not present a good safety-efficacy balance. In our quest for better and safer anticoagulants, we have identified and characterized a novel thrombin inhibitor-avathrin from the salivary glands *Ambylomma variegatum*. We have solved the 3D-crystal structure of avathrin in complex with thrombin to study detailed structure-activity relationships. We have successfully demonstrated the activities of avathrin using *in vitro* assays and *in vivo* animal models. We have also identified similar sequences from other tick species and demonstrated the presence of a family of thrombin inhibitors in ixodid ticks. These peptides are short simple sequences with a unique mechanism of inhibiting thrombin active site and exosite with affinities in the picomolar to femtomolar range. We have also conceived and synthesized structure based variants of these peptides with better potencies and different modes of inhibition than the native peptides. We are currently evaluating the safety-efficacy balance and pharmacokinetics-pharmacodynamics of some of these peptides in animal models to provide

evidence that these peptides hit the sweet spot and fulfil an unmet need in the current day anticoagulant market.

7.2. Future Perspectives

7.2.1. Development of at-line nano-fractionation assay for identifying thrombin and factor Xa inhibitors

We plan to develop an at-line nano-fractionation assay in collaboration with the Biomolecular Analysis group at VU University, Amsterdam. This approach would overcome the limitations posed by the online assay (described in Chapter 2). Several additives are used in the online assay to prevent the adsorbance of assay reagents to the walls of the microfluidic chip. Because the microfluidic chip has a narrow internal diameter, the mixing of assay reagents and additives with organic eluents from the nano-LC causes precipitation of the enzymes and substrates, when complex mixtures such as tick salivary glands are used in the online assay.

With the at-line nano-fractionation assay, we will be able to identify novel thrombin and FXa inhibitors from tick salivary gland extracts. Briefly, in this approach, we plan to split the eluate from the nano-LC in a 1:9 ratio using a post-column flow split (similar to the T-splitter used in the online assay). The smaller fraction will be directed towards the MS and the larger fraction will be directed towards a fraction collection device. The fractions eluting from the nano-LC will be collected in 384 well-microtiter plates and will be subjected to freeze drying using a vacuum centrifuge. Followed by the freeze drying, the entire plate will be subjected to a thrombin/ FXa bioassay using the fluorescence substrate- R22124. To achieve an excellent separation, and prevent mixing of the compounds eluting

from the nano-LC, the fraction collection device will be operated at the highest possible resolution (3-12 s/well).

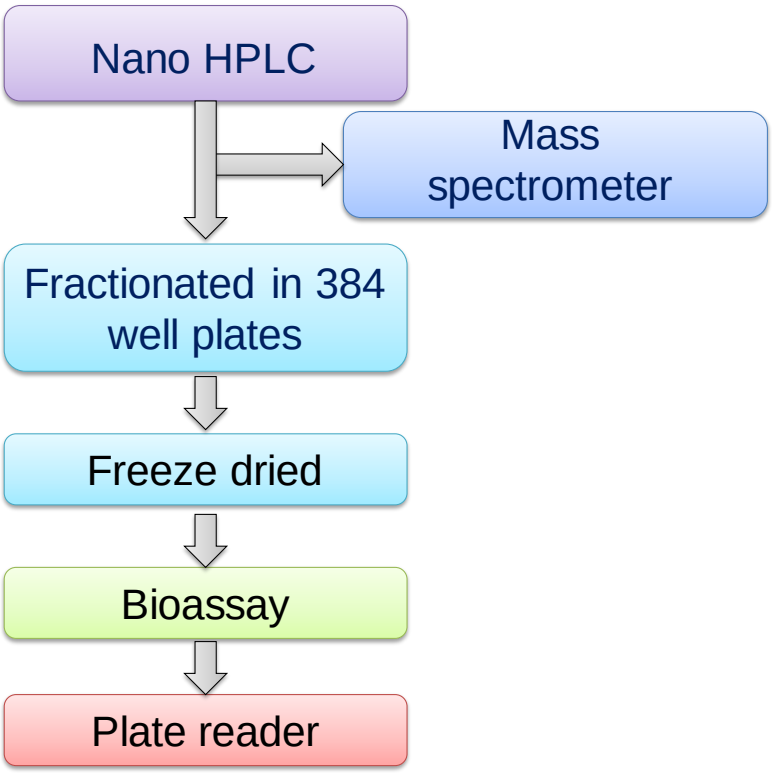


Figure 7.1. At-line nano fractionation set-up. Complex mixtures will be separated using a reverse phase column connected to the nano-LC. The eluate from the nano-LC will be split in a 1:9 ratio using a T-splitter. The smaller fraction will be directed towards the MS where the masses of the eluting compounds will be identified and the larger fraction will be directed towards a fractionantion device, where the eluate will be fractionated into 384-well microtiter plates. The microtiter plates will be freeze dried and a fluorescence based thrombin/FXa assay will be carried to identify wells in which the inhibitors have been collected. By correlating the times at which the inhibitors are identified in the MS and in the bioassay, their accurate masses will be identified.

With information from the fluorescence assay, the wells containing inhibitors will be identified, and the parallel MS data will reveal the masses of the eluting compounds. By correlating the times at which the inhibitor is identified in the bioassay and in the MS, the exact masses of the inhibitors will be identified.

Subsequently, inhibitors in those wells will be subjected to in-solution tryptic digestion and using the transcriptome library as the database, peptides from the tryptic digest which match with proteins in the library will be identified.

7.2.2. Recombinant expression and functional studies of monolaris proteins

The transcriptome and proteome library of male and female salivary gland extracts of D. reticulatus represents a pharmacopeia of novel proteins which could serve as potential drug leads for the control of several cardiovascular disorders. In our study we have identified a novel subclass of monolaris proteins which contain two extra cysteines at the N-terminus of the molecules and these proteins are specifically overexpressed in the females (e.g. SigP-383232, DrIxod-9948 and DrIxod-595427 are 1200, 1300 and 2800 time over-expressed in the females compared to the males). With one extra disulphide bond, these proteins may have a fold that is different than the usual Kunitz proteins, and this novel fold may confer an interesting functional property on the protein. We will carry out recombinant expression of three such monolaris proteins in insect cell lines that would allow appropriate disulphide bond formation of the expressed proteins. Briefly, the synthetic genes for these transcripts will be ordered and the genes will be first be cloned and amplified in plasmids in bacterial cells. The plasmids will be sequenced to check for the sequence and the orientation of the insert. These plasmids will then be transfected into High Five^{TM} insect cell lines and the protein expression will be carried out in these cells. Once the proteins are expressed and purified from these cell lines, the masses and disulphide bonding patterns of these proteins will be identified using ESI-MS and disulphide mapping experiments.

Further, a serine protease screning assay will be carried out using the pro- and anticoagulant enzymes of the blood coagulation cascade to identify the targets of these Kunitz proteins. Having a target serine protease that is inhibited by these proteins identified, all the kinetics and structure activity relationships of these proteins with the target enzyme will be studied. We will also try to understand detailed structure activity relationships by solving the co-crystal structure of the Kunitz inhibitor with the serine protease. Further site directed mutagenesis to identify key residues important for the binding of the inhibitor to the serine protease will also be carried out. Site directed mutagenesis will also be used to improve the selectivity and affinity of the inhibitors to target specific enzymes while leaving the other serine proteases unaffected.

Based on the crystal structure and the information about the key residues we will try to design shorter pepdide sequences that would retain the inhibitory effects of the full length Kunitz proteins. Being short peptides, these would be easy to produce in larger amounts, and we will be able to study the efficacy and safety profiles of these inhibitors in smaller animals models and develop them further as therapeutics.

7.2.3. Preclinical studies of avathrin and other peptides of the variegin family

For the development of avathrin and ultravareigin as therapeutics, their *in vivo* efficacy and safety in rodent models will be studied. For the efficacy models, we will use the well established ferric chloride induced thrombosis of the carotid arteries. Briefly, in this set up, the carotid artery of rats will be exposed using blunt end dissection, and increasing amounts of the peptide will be infused intravenously into the circulation by cannulating the femoral vein. A probe will be placed on the carotid artery and the blood flow will be monitored. After the stabilization of the blood flow, a filter paper saturated with ferric chloride will be placed on the carotid artery to induce thrombosis. The time taken for an occlusion will be monitored by observing the blood flow. An ED_{50} curve will be plotted and the most efficacious dose will be determined.

For determining a safe dose of the anticoagulant, a tail bleeding rat model will be used. In this model, the peptide will be infused intravenously into the femoral vein. Subsequently an insicion in the tail vein will be made using a Surgicutt device and the blood will blotted using a filter paper and the time taken for the bleeding to stop will be noted. Plotting the bleeding times, and the ED_{50} values on the same graph, a therapeutic window for avathrin and ultravariegin will be identified. This therapeutic window will be compared to the therapeutic window of other commercially available and most extensively used drugs such as unfractionated heparin and hirulog-1. This would give a head-to-head comparison of the already available drugs with our peptides. Preliminary experiments for this part are already being carried out using our first peptide- variegin and that data

already shows that doses at which our peptides possess an efficient antithrombotic effect do not cause significant bleeding in comparison to the other existing drugs.

Pharmacokinetics/ pharmacodynamics experiments will be carried out to determine the clearance of the peptide from the animals and duration of the efficacy of the antithrombotic effect. All of this data will together indicate the doses that would serve as therapeutic doses in animals. These doses will then be translated to other higher animals like pigs and similar and more detailed experiments will be carried out in larger animal models. Other models like the cardiopulmonary bypass model will also be studied with our peptides in pigs. These results will be give us insights about the doses of avathrin and ultravariegin that could be initially tried in humans after obtaining all the necessary approvals.

References

Alarcon-Chaidez, F.J., Sun, J., and Wikel, S.K. (2007). Transcriptome analysis of the salivary glands of Dermacentor andersoni Stiles (Acari: Ixodidae). Insect Biochem. Mol. Biol. *37*, 48–71.

Apte, S.S. (2004). A disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motifs: The ADAMTS family. Int. J. Biochem. Cell Biol. *36*, 981–985.

Avila, F.W., Sirot, L.K., LaFlamme, B. a, Rubinstein, C.D., and Wolfner, M.F. (2011). Insect seminal fluid proteins: identification and function. Annu. Rev. Entomol. *56*, 21–40.

Badimon, L., and Vilahur, G. (2007). Platelets, arterial thrombosis and cerebral ischemia. Cerebrovasc. Dis. *24*, 30–39.

Barker, S.C., and Murrell, a (2004). Systematics and evolution of ticks with a list of valid genus and species names. Parasitology *129 Suppl*, S15–S36.

Bates, S.M., and Weitz, J.I. (2005). Coagulation assays. Circulation *112*, e53–e60.

Bauer, K. a (2013). Pros and cons of new oral anticoagulants. Hematology Am. Soc. Hematol. Educ. Program *2013*, 464–470.

Bergman, D.K., Palmer, M.J., Caimano, M.J., Radolf, J.D., and Wikel, S.K. (2000). Isolation and molecular cloning of a secreted immunosuppressant protein from Dermacentor andersoni salivary gland. J. Parasitol. *86*, 516–525.

Bern, A., Freiburg, R., Dundee, B., Norwich, R., and Martinsried, W. (1991). Mechanism and functions of ATP-dependent proteases in bacterial and animals cells. Eur. J. Biochem. *203*, 9–23.

Blankenship, T., Brankamp, R., Manley, G., and Cardin, D. (1990). Amino acid sequence of Ghilanten: Anticoagulant-antimetastatic principle of the South American Leech, Haementeria ghilianii. Biochem. Biophys. Res. Commun. *166*, 1384–1389.

De Boer, A.R., Bruyneel, B., Krabbe, J.G., Lingeman, H., Niessen, W.M. a, and Irth, H. (2005). A microfluidic-based enzymatic assay for bioactivity screening combined with capillary liquid chromatography and mass spectrometry. Lab Chip *5*, 1286–1292.

Borissoff, J.I., Spronk, H.M.H., Heeneman, S., and Ten Cate, H. (2009). Is thrombin a key player in the “coagulation-atherogenesis” maze? Cardiovasc. Res. *82*, 392–403.

Bourdon, P. (1991). Hirulog peptides with sciccile bond replacements resistant to thrombin cleavage. Biochem. Biophys. Res. Commun. *177*, 1049–1055.

Bowman, a S., and Sauer, J.R. (2004). Tick salivary glands: function, physiology and future. Parasitology *129 Suppl*, S67–S81.

Bowman, a. S., Nuttall, P. a., and Chappell, L.H. (2008). Ticks Biology, Disease, and Control.

Bowman, A., Coons, L.B., Needham, G.R., and Sauer, J.R. (1997). Tick saliva: recent advances and implications for vector competence. Med. Vet. Entomol. *11*, 277–285.

Brass, L.F. (2003). Thrombin and Platelet Activation * Thrombin and Platelet Activation *. Chest *124*, 18S – 25S.

Calvo, E., Mizurini, D.M., Sá-Nunes, A., Ribeiro, J.M.C., Andersen, J.F., Mans, B.J., Monteiro, R.Q., Kotsyfakis, M., and Francischetti, I.M.B. (2011). Alboserpin, a factor Xa inhibitor from the mosquito vector of yellow fever, binds heparin and membrane phospholipids and exhibits antithrombotic activity. J. Biol. Chem. *286*, 27998–28010.

Campos, I.T.N., Amino, R., Sampaio, C. a M., Auerswald, E. a., Friedrich, T., Lemaire, H.G., Schenkman, S., and Tanaka, a. S. (2002). Infestin, a thrombin inhibitor presents in Triatoma infestans midgut, a Chagas' disease vector: Gene cloning, expression and characterization of the inhibitor. Insect Biochem. Mol. Biol. *32*, 991–997.

Cappello, M., Vlasuk, G.P., Bergum, P.W., Huang, S., and Hotez, P.J. (1995). Ancylostoma caninum anticoagulant peptide: a hookworm-derived inhibitor of human coagulation factor Xa. Proc. Natl. Acad. Sci. U. S. A. *92*, 6152–6156.

Cavassani, K. a., Aliberti, J.C., Dias, A.R. V, Silva, J.S., and Ferreira, B.R. (2005). Tick saliva inhibits differentiation, maturation and function of murine bone-marrow-derived dendritic cells. Immunology *114*, 235–245.

Di Cera, E. (2008). Thrombin. Mol. Aspects Med. *29*, 203–254.

Champagne, D.E. (2006). Antihemostatic molecules from saliva of blood-feeding arthropods. Pathophysiol. Haemost. Thromb. *34*, 221–227.

Chang, J.Y. (1983). The functional domain of hirudin, a thrombin-specific inhibitor. FEBS Lett. *164*, 307–313.

Chaudhari, K., Hamad, B., and Syed, B. a (2014). Antithrombotic drugs market. Nat. Rev. Drug Discov. *13*, 571–572.

Chmelar, J., Oliveira, C.J., Rezacova, P., Francischetti, I.M.B., Kovarova, Z., Pejler, G., Kopacek, P., Ribeiro, J.M.C., Mares, M., Kopecky, J., et al. (2011). A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. Blood *117*, 736–744.

Clemetson, K.J. (1999). Platelet collagen receptors: a new target for inhibition? Haemostasis *29*, 16–26.

Collin, N., Assumpção, T.C.F., Mizurini, D.M., Gilmore, D.C., Dutra-Oliveira, A., Kotsyfakis, M., Sá-Nunes, A., Teixeira, C., Ribeiro, J.M.C., Monteiro, R.Q., et al. (2012). Lufaxin, a novel factor Xa inhibitor from the salivary gland of the sand fly lutzomyia longipalpis blocks protease-activated receptor 2 activation and inhibits inflammation and thrombosis in vivo. Arterioscler. Thromb. Vasc. Biol. *32*, 2185–2196.

Copeland, R. a (2000). ENZYMES A Practical Introduction.

Coppens, M., Eikelboom, J.W., Gustafsson, D., Weitz, J.I., and Hirsh, J. (2012). Translational success stories: Development of direct thrombin inhibitors. Circ. Res. *111*, 920–929.

Corral-Rodríguez, M.Á., Macedo-Ribeiro, S., Barbosa Pereira, P.J., and Fuentes-Prior, P. (2009). Tick-derived Kunitz-type inhibitors as antihemostatic factors. Insect Biochem. Mol. Biol. *39*, 579–595.

Cupp, M.S., Ribeiro, J.M., Champagne, D.E., and Cupp, E.W. (1998). Analyses of cDNA and recombinant protein for a potent vasoactive protein in saliva of a blood-feeding black fly, Simulium vittatum. J. Exp. Biol. *201*, 1553–1561.

Curiale, G., Lara, L., Singla, A., and Fisher, M. (2011). Warfarin versus Warfarin and Aspirin inAtrial fibrillation. Neurol. Bull. *1*, 1–10.

Dahlbäck, B. (2000). Blood coagulation. Lancet *355*, 1627–1632.

Davie, E., Fujikawa, K., and Kisiel, W. (1991). The Coagulation Cascade: Initiation, Maintenance, and Regulation. Biochemistry *30*, 10364–10370.

Decrem, Y., Rath, G., Blasioli, V., Cauchie, P., Robert, S., Beaufays, J., Frère, J.-M., Feron, O., Dogné, J.-M., Dessy, C., et al. (2009). Ir-CPI, a coagulation contact phase inhibitor from the tick Ixodes ricinus, inhibits thrombus formation without impairing hemostasis. J. Exp. Med. *206*, 2381–2395.

Dodt, J., Otte, M., Strube, K.H., and Friedrich, T. (1996). Thrombin inhibitors of bloodsucking animals. Semin. Thromb. Hemost. *22*, 203–208.

Donohue, K. V, Khalil, S.M., Mitchell, R.D., Sonenshine, D.E., and Roe, R.M. (2008). Molecular characterization of the major hemelipoglycoprotein in ixodid ticks. Insect Mol.Biol. *17*, 197–208.

Draxler, D.F., and Medcalf, R.L. (2014). The Fibrinolytic System—More Than Fibrinolysis? Transfus. Med. Rev. *29*, 102–109.

Duckert, P., Brunak, S., and Blom, N. (2004). Prediction of proprotein convertase cleavage sites. Protein Eng. Des. Sel. *17*, 107–112.

Dunwiddie, C.T., Neeper, M.P., Nutt, E.M., Waxman, L., Smith, D.E., Hofmann, K.J., Lumma, P.K., Garsky, V.M., and Vlasuk, G.P. (1992). Site-directed analysis of the functional domains in the factor Xa inhibitor tick anticoagulant peptide: identification of two distinct regions that constitute the enzyme recognition sites. Biochemistry *31*, 12126–12131.

Eckly, a, Hechler, B., Freund, M., Zerr, M., Cazenave, J.-P., Lanza, F., Mangin, P.H., and Gachet, C. (2011). Mechanisms underlying FeCl3-induced arterial thrombosis. J. Thromb. Haemost. *9*, 779–789.

Emsley, P., and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. *60*, 2126–2132.

Entzeroth, M. (2003). Emerging trends in high-throughput screening. Curr. Opin. Pharmacol. *3*, 522–529.

Falck, D., de Vlieger, J.S.B., Niessen, W.M. a, Kool, J., Honing, M., Giera, M., and Irth, H. (2010). Development of an online p38α mitogen-activated protein kinase binding assay and integration of LC-HR-MS. Anal. Bioanal. Chem. *398*, 1771–1780.

Feldman-Muhsam, B., Borut, S., and Saliternik-Givant, S. (1970). Salivary secretion of the male tick during copulation. J. Insect Physiol. *16*, 1945–1949.

Figueiredo, A.C., de Sanctis, D., Gutiérrez-Gallego, R., Cereija, T.B., Macedo-Ribeiro, S., Fuentes-Prior, P., and Pereira, P.J.B. (2012). Unique thrombin inhibition mechanism by anophelin, an anticoagulant from the malaria vector. Proc. Natl. Acad. Sci. U. S. A. *109*, E3649–E3658.

Figueiredo, A.C., de Sanctis, D., and Pereira, P.J.B. (2013). The Tick-Derived Anticoagulant Madanin Is Processed by Thrombin and Factor Xa. PLoS One *8*, 1–11.

Flower, D.R., North, a. C.T., and Sansom, C.E. (2000). The lipocalin protein family: Structural and sequence overview. Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. *1482*, 9–24.

Fogaça, A.C., Almeida, I.C., Eberlin, M.N., Tanaka, A.S., Bulet, P., and Daffre, S. (2006). Ixodidin, a novel antimicrobial peptide from the hemocytes of the cattle tick Boophilus microplus with inhibitory activity against serine proteinases. Peptides *27*, 667–674.

Francischetti, I.M.B. (2010). Platelet aggregation inhibitors from Hematophagous animals. Toxicon *56*, 1130–1144.

Francischetti, I.M.B., Valenzuela, J.G., Andersen, J.F., Mather, T.N., and Ribeiro, M.C. (2002). Ixolaris , a novel recombinant tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick , Ixodes scapularis : identification of factor X and factor Xa as scaffolds for the inhibition of factor VIIa / tissue factor complex. Blood *99*, 3602–3612.

Francischetti, I.M.B., Mather, T.N., and Ribeiro, J.M.C. (2004). Penthalaris, a novel recombinant five-Kunitz tissue factor pathway inhibitor (TFPI) from the salivary gland of the tick vector of Lyme disease, Ixodes scapularis. Thromb. Haemost. *91*, 886–898.

Francischetti, I.M.B., Sa-Nunes, A., Mans, B.J., Santos, I.M., and Ribeiro, J.M.C. (2009). The role of saliva in tick feeding. Front. Biosci. *14*, 2051–2088.

Friesen, K.J., and Kaufman, W.R. (2009). Salivary gland degeneration and vitellogenesis in the ixodid tick Amblyomma hebraeum: Surpassing a critical weight is the prerequisite and detachment from the host is the trigger. J. Insect Physiol. *55*, 936–942.

Fuentes-Prior, P., Iwanaga, Y., Huber, R., Pagila, R., Rumennik, G., Seto, M., Morser, J., Light, D.R., and Bode, W. (2000). Structural basis for the anticoagulant activity of the thrombin- thrombomodulin complex. Nature *404*, 518–525.

Fuller, H.R., and Morris, G.E. (2012). Quantitative Proteomics Using iTRAQ Labeling and Mass Spectrometry. Integr. Proteomics 442.

Furie, B., and Furie, B.C. (2007). In vivo thrombus formation. J. Thromb. Haemost. *5*, 12–17.

Furie, B., and Furie, B.C. (2008). Mechanisms of thrombus formation. N. Engl. J. Med. *359*, 938–949.

Gailani, D., and Renné, T. (2007). Intrinsic pathway of coagulation and arterial thrombosis. Arterioscler. Thromb. Vasc. Biol. *27*, 2507–2513.

Garber, M., Grabherr, M.G., Guttman, M., and Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. Nat. Methods *8*, 469–477.

Giera, M., Heus, F., Janssen, L., Kool, J., Lingeman, H., and Irth, H. (2009). Microfractionation revisited: A 1536 well high resolution screening assay. Anal. Chem. *81*, 5460–5466.

Gonias, S.L., and Pizzo, S. V (1986). The biochemistry of haemostasis. Clin. Lab. Haematol. *8*, 281–306.

Gould, W.R., McClanahan, T.B., Welch, K.M., Baxi, S.M., Saiya-Cork, K., Chi, L., Johnson, T.R., and Leadley, R.J. (2006). Inhibitors of blood coagulation factors Xa and IIa synergize to reduce thrombus weight and thrombin generation in vivo and in vitro. J. Thromb. Haemost. *4*, 834–841.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D. a, Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

Graßmann, J., Scheerle, R.K., and Letzel, T. (2012). Functional proteomics: Application of mass spectrometry to the study of enzymology in complex mixtures. Anal. Bioanal. Chem. *402*, 625–645.

Greinacher, A., Volpel, H., Janssens, U., and Hach-Wunderle, V. (1999). Recombinant Hirudin (Lepirudin) Provides Safe and Effective Anticoagulation in Patients With Heparin-Induced Thrombocytopenia. Circulation 73–81.

Grütter, M.G., Priestle, J.P., Rahuel, J., Grossenbacher, H., Bode, W., Hofsteenge, J., and Stone, S.R. (1990). Crystal structure of the thrombin-hirudin complex: a novel mode of serine protease inhibition. EMBO J. *9*, 2361–2365.

Hajnická, V., Kocáková, P., Sláviková, M., Slovák, M., Gašperík, J., Fuchsberger, N., and Nuttall, P. a. (2001). Anti-interleukin-8 activity of tick salivary gland extracts. Parasite Immunol. *23*, 483–489.

Halos, L., Lebert, I., Abrial, D., Danlois, F., Garzik, K., Rodes, D., Schillmeier, M., Ducrot, C., and Guillot, J. (2014). Questionnaire-based survey on the distribution and incidence of canine babesiosis in countries of Western Europe. Parasite *21*, 13.

Heindl, B., and Kupatt, C. (2000). Endothelial function and hemostasis. *167*, 160–167.

Heus, F., Giera, M., De Kloe, G.E., Van Iperen, D., Buijs, J., Nahar, T.T., Smit, A.B., Lingeman, H., De Esch, I.J.P., Niessen, W.M. a, et al. (2010). Development of a microfluidic confocal fluorescence detection system for the hyphenation of nano-LC to on-line biochemical assays. Anal. Bioanal. Chem. *398*, 3023–3032.

Heus, F., Otvos, R. a., Aspers, R.L.E.G., van Elk, R., Halff, J.I., Ehlers, A.W., Dutertre, S., Lewis, R.J., Wijmenga, S., Smit, A.B., et al. (2014). Miniaturized bioaffinity assessment coupled to mass spectrometry for guided purification of bioactives from toad and cone snail. Biology (Basel). *3*, 139–156.

Hong, K. (2014). Dual Antiplatelet Therapy after Noncardioembolic Ischemic Stroke or Transient Ischemic Attack : Pros and Cons Dual Antiplatelet Therapy in CHD Short-Term Dual Antiplatelet. *10*, 189–196.

Hovius, J.W.R., Levi, M., and Fikrig, E. (2008). Salivating for knowledge: Potential pharmacological agents in tick saliva. PLoS Med. *5*, 0202–0208.

Hultin-Rosenberg, L., Forshed, J., Branca, R.M.M., Lehtiö, J., and Johansson, H.J. (2013). Defining, comparing, and improving iTRAQ quantification in mass spectrometry proteomics data. Mol. Cell. Proteomics *12*, 2021–2031.

Huntington, J. a. (2014). Natural inhibitors of thrombin. Thromb. Haemost. *111*, 583–589.

Irth, H. (2007). Continuous-flow Systems for Ligand Binding and Enzyme Inhibition Assays Based on Mass Spectrometry.

Isawa, H., Orito, Y., Jingushi, N., Iwanaga, S., Morita, A., Chinzei, Y., and Yuda, M. (2007). Identification and characterization of plasma kallikrein-kinin system inhibitors from salivary glands of the blood-sucking insect Triatoma infestans. FEBS J. *274*, 4271–4286.

Jagadeeswaran, P. (2005). Zebrafish: a tool to study hemostasis and thrombosis. Curr. Opin. Hematol. *12*, 149–152.

Jennings, L.K. (2009). Mechanisms of platelet activation: Need for new strategies to protect against platelet-mediated atherothrombosis. Thromb. Haemost. *102*, 248–257.

Johns, R., Sonenshine, D.E., and Hynes, W.L. (2001). Identification of a defensin from the hemolymph of the American dog tick, Dermacentor variabilis. Insect Biochem. Mol. Biol. *31*, 857–865.

Karbowiak, G. (2014). The occurrence of the Dermacentor reticulatus tick – its expansion to new areas and possible causes. An. Parasitol. *60*, 37–47.

Karczewski, J., Endris, R., and Connolly, T.M. (1994). Disagregin is a fibrinogen receptor antagonist lacking the Arg-Gly-Asp sequence from the tick, Ornithodoros moubata. J. Biol. Chem. *269*, 6702–6708.

Karim, S., Singh, P., and Ribeiro, J.M.C. (2011). A deep insight into the sialotranscriptome of the gulf coast tick, Amblyomma maculatum. PLoS One *6*.

Karp, N. a., and Lilley, K.S. (2007). Design and analysis issues in quantitative proteomics studies. Proteomics *7 Suppl 1*, 42–50.

Kazimírová, M., and Štibrániová, I. (2013). Tick salivary compounds: their role in modulation of host defences and pathogen transmission. Front. Cell. Infect. Microbiol. *3*, 43.

Khan, M.S., Singh, P., Azhar, A., Naseem, A., Rashid, Q., Kabir, M.A., and Jairajpuri, M.A. (2011). Serpin Inhibition Mechanism: A Delicate Balance between Native Metastable State and Polymerization. J. Amino Acids *2011*, 606797.

Koh, C.Y., and Kini, R.M. (2008). Anticoagulants from hematophagous animals. Expert Rev. Hematol. *1*, 135–139.

Koh, C., Kazimirova, M., Trimnell, A., Takac, P., Labuda, M., Nuttall, P. a., and Kini, R.M. (2007). Variegin, a novel fast and tight binding thrombin inhibitor from the tropical bont tick. J. Biol. Chem. *282*, 29101–29113.

Koh, C.Y., Kazimirova, M., Nuttall, P. a., and Kini, R.M. (2009). Noncompetitive inhibitor of thrombin. ChemBioChem *10*, 2155–2158.

Koh, C.Y., Kumar, S., Kazimirova, M., Nuttall, P. a., Radhakrishnan, U.P., Kim, S., Jagadeeswaran, P., Imamura, T., Mizuguchi, J., Iwanaga, S., et al. (2011). Crystal structure of thrombin in complex with s-variegin: Insights of a novel mechanism of inhibition and design of tunable thrombin inhibitors. PLoS One *6*.

Kool, J., De Kloe, G.E., Bruyneel, B., De Vlieger, J.S., Retra, K., Wijtmans, M., Van Elk, R., Smit, A.B., Leurs, R., Lingeman, H., et al. (2010). Online fluorescence enhancement assay for the acetylcholine binding protein with parallel mass spectrometric identification. J. Med. Chem. *53*, 4720–4730.

Kool, J., Giera, M., Irth, H., and Niessen, W.M. a (2011). Advances in mass spectrometry-based post-column bioaffinity profiling of mixtures. Anal. Bioanal. Chem. *399*, 2655–2668.

Kotsyfakis, M., Karim, S., Andersen, J.F., Mather, T.N., and Ribeiro, J.M.C. (2007). Selective cysteine protease inhibition contributes to blood-feeding success of the tick Ixodes scapularis. J. Biol. Chem. *282*, 29256–29263.

Kottke-marchant, K. (2010). Antithrombotic Drug Therapy in Cardiovascular Disease. Contemp. Cardiol. Antithrombotic Drug Ther. Cardiovasc. Dis. 19–38.

Kroll, M.H., Harris, T.S., Moake, J.L., Handin, R.I., and Schafer, a I. (1991). von Willebrand factor binding to platelet GpIb initiates signals for platelet activation. J. Clin. Invest. *88*, 1568–1573.

Krowarsch, D., Cierpicki, T., Jelen, F., and Otlewski, J. (2003). Canonical protein inhibitors of serine proteases. Cell. Mol. Life Sci. *60*, 2427–2444.

Kubes, M., Fuchsberger, N., Labuda, M., Zuffová, E., and Nuttall, P. a (1994). Salivary gland extracts of partially fed Dermacentor reticulatus ticks decrease natural killer cell activity in vitro. Immunology *82*, 113–116.

LaFlamme, B. a., Ravi Ram, K., and Wolfner, M.F. (2012). The Drosophila melanogaster seminal fluid protease "Seminase" regulates proteolytic and post-mating reproductive processes. PLoS Genet. *8*, 30–32.

Lane, D. a., Philippou, H., and Huntington, J. a. (2005). Directing thrombin. Blood *106*, 2605–2612.

Laskowski, M., and Kato, I. (1980). Protein Inhibitors of proteinases. Annu. Rev. Biochem. *49*, 593–626.

Lee, A.. Y., and Vlasuk, G.P. (2003). Recombinant nematode anticoagulant protein c2 and other inhibitors targeting blood coagulation factor VIIa/tissue factor. J. Intern. Med. *254*, 313–321.

Lee, C.J., and Ansell, J.E. (2011). Direct thrombin inhibitors. Br. J. Clin. Pharmacol. *72*, 581–592.

Lentz, B.R. (2003). Exposure of platelet membrane phosphatidylserine regulates blood coagulation. Prog. Lipid Res. *42*, 423–438.

Lerner, E. a., Ribeiro, J.M.C., Nelson, R.J., and Lerner, M.R. (1991). Isolation of maxadilan, a potent vasodilatory peptide from the salivary glands of the sand fly Lutzomyia longipalpis. J. Biol. Chem. *266*, 11234–11236.

Leslie, A.G.W., and Powell, H.R. (2007). Processing diffraction data with MOSFLM.

Lijfering, W.M., Ph, D., Flinterman, L.E., Sc, M., Vandenbroucke, J.P., Ph, D., Rosendaal, F.R., Ph, D., Cannegieter, S.C., and Ph, D. (2011). Relationship between Venous and Arterial Thrombosis: A Review of the Literature from a Casual Perspective. *1*, 885–896.

Lim-Wilby, M.S., Hallenga, K., de Maeyer, M., Lasters, I., Vlasuk, G.P., and Brunck, T.K. (1995). NMR structure determination of tick anticoagulant peptide (TAP). Protein Sci. *4*, 178–186.

Linhardt, R.J., and Liu, J. (2012). Synthetic heparin. Curr. Opin. Pharmacol. *12*, 217–219.

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

Van de Locht, a, Stubbs, M.T., Bode, W., Friedrich, T., Bollschweiler, C., Höffken, W., and Huber, R. (1996). The ornithodorin-thrombin crystal structure, a key to the TAP enigma? EMBO J. *15*, 6011–6017.

Macedo-Ribeiro, S., Almeida, C., Calisto, B.M., Friedrich, T., Mentele, R., Stürzebecher, J., Fuentes-Prior, P., and Pereira, P.J.B. (2008). Isolation, cloning and structural characterization of boophilin, a multifunctional Kunitz-type proteinase inhibitor from the cattle tick. PLoS One *3*, 1–17.

Malovichko, M. V., Sabo, T.M., and Maurer, M.C. (2013). Ligand binding to anion-binding exosites regulates conformational properties of thrombin. J. Biol. Chem. *288*, 8667–8678.

Mans, B.J. (2011). Evolution of vertebrate hemostatic and inflammatory control mechanisms in blood-feeding arthropods. J. Innate Immun. *3*, 41–51.

Mans, B.J., and Neitz, a. W.H. (2004). Adaptation of ticks to a blood-feeding environment: Evolution from a functional perspective. Insect Biochem. Mol. Biol. *34*, 1–17.

Mans, B.J., Gaspar, a R., Louw, a I., and Neitz, a W. (1998). Apyrase activity and platelet aggregation inhibitors in the tick Ornithodoros savignyi (Acari: Argasidae). Exp. Appl. Acarol. *22*, 353–366.

Mans, B.J., Louw, A.I., and Neitz, A.W.H. (2002). Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus Ornithodoros. Mol. Biol. Evol. *19*, 1695–1705.

Mans, B.J., Ribeiro, J.M.C., and Andersen, J.F. (2008). Structure, function, and evolution of biogenic amine-binding proteins in soft ticks. J. Biol. Chem. *283*, 18721–18733.

Mao, S.-S. (1993). Factor Xa inhibitors. Perspect. Drug Discov. Des. *1*, 423–430.

Marder, V.J., Rosove, M.H., and Minning, D.M. (2004). Foundation and sites of action of antithrombotic agents. Best Pract. Res. Clin. Haematol. *17*, 3–22.

Mavrakanas, T. a., and Chatzizisis, Y.S. (2015). Bivalirudin in stable angina and acute coronary syndromes. Pharmacol. Ther. *152*, 1–10.

Maynard, J.R., Heckman, C.A., Pitlick, F.A., and Nemerson, Y. (1975). Association of Tissue Factor Activity with the Surface of Cultured Cells. *55*, 814–824.

McCoy, A.J. (2006). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. Sect. D Biol. Crystallogr. *63*, 32–41.

Mende, K., Petoukhova, O., Koulitchkova, V., Schaub, G. a., Lange, U., Kaufmann, R., and Nowak, G. (1999). Dipetalogastin, a potent thrombin inhibitor from the blood-sucking insect Dipetalogaster maximus. cDNA cloning, expression and characterization. Eur. J. Biochem. *266*, 583–590.

Mieszczanek, J., Harrison, L.M., and Cappello, M. (2004). Ancylostoma ceylanicum anticoagulant peptide-1: Role of the predicted reactive site amino acid in mediating inhibition of coagulation factors Xa and VIIa. Mol. Biochem. Parasitol. *137*, 151–159.

Monroe, D.M., and Hoffman, M. (2006). What does it take to make the perfect clot? Arterioscler. Thromb. Vasc. Biol. *26*, 41–48.

Monroe, D.M., Hoffman, M., and Roberts, H.R. (2002). Platelets and thrombin generation. Arterioscler. Thromb. Vasc. Biol. *22*, 1381–1389.

Mousa, S.A. (2008). Pharmacological Assays in Thrombosis and Haemostasis. J. Obstet. Gynaecol. (Lahore). *24*, 398–453.

Myles, T., and Leung, L.L.K. (2008). Thrombin hydrolysis of human osteopontin is dependent on thrombin anion-binding exosites. J. Biol. Chem. *283*, 17789–17796.

Myles, T., Le Bonniec, B.F., Betz, a., and Stone, S.R. (2001). Electrostatic steering and ionic tethering in the formation of thrombin-hirudin complexes: The role of the thrombin anion-binding exosite-I. Biochemistry *40*, 4972–4979.

Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J., and Bishop, R. (2002). AvGI, an index of genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum. Int. J. Parasitol. *32*, 1447–1456.

Nielsen, H., Brunak, S., and von Heijne, G. (1999). Machine learning approaches for the prediction of signal peptides and other protein sorting signals. Protein Eng. *12*, 3–9.

Noeske-Jungblut, C., Haendler, B., Donner, P., Alagon, a., Possani, L., and Schleuning, W.D. (1995). Triabin, a highly potent exosite inhibitor of thrombin. J. Biol. Chem. *270*, 28629–28634.

Oaks, J.F., McSwain, J.L., Bantle, J. a, Essenberg, R.C., and Sauer, J.R. (1991). Putative new expression of genes in ixodid tick salivary gland development during feeding. J. Parasitol. *77*, 378–383.

Ogata, S., Hayashi, Y., Misumi, Y., and Ikehara, Y. (1990). Membrane-anchoring domain of rat liver 5'-nucleotidase: Identification of the COOH-terminal serine-523 covalently attached with a glycolipid. Biochemistry *29*, 7923–7927.

Okafor, O.N., and Gorog, D. a. (2015). Endogenous Fibrinolysis. J. Am. Coll. Cardiol. *65*, 1683–1699.

Otvos, R. a., Heus, F., Vonk, F.J., Halff, J., Bruyneel, B., Paliukhovich, I., Smit, A.B., Niessen, W.M. a, and Kool, J. (2013). Analytical workflow for rapid screening and purification of bioactives from venom proteomes. Toxicon *76*, 270–281.

Page, M.J., Macgillivray, R.T. a, and Di Cera, E. (2005). Determinants of specificity in coagulation proteases. J. Thromb. Haemost. *3*, 2401–2408.

Pechik, I., Madrazo, J., Mosesson, M.W., Hernandez, I., Gilliland, G.L., and Medved, L. (2004). Crystal structure of the complex between thrombin and the central "E" region of fibrin. Proc. Natl. Acad. Sci. U. S. A. *101*, 2718–2723.

Perona, J.J., Hedstrom, L., Rutter, W.J., and J, R.J.F. (1995). Structural Origins of Substrate Discrimination in Trypsin and Chymotrypsin? 1489–1499.

Phillips, D.R., Conley, P.B., Sinha, U., and Andre, P. (2005). Therapeutic approaches in arterial thrombosis. J. Thromb. Haemost. *3*, 1577–1589.

Piccolo, F., Popowicz, N., Wong, D., Chor, Y., and Lee, G. (2015). Intrapleural tissue plasminogen activator and deoxyribonuclease therapy for pleural infection. J. Thorac. Dis. *7*, 999–1008.

Polgár, L. (2005). The catalytic triad of serine peptidases. Cell. Mol. Life Sci. *62*, 2161–2172.

Preston, S., Majtan, J., Rysnik, O., and Burger, L. (2013). Novel immunomodulators from hard ticks selectively reprogramme human dendritic cell responses. PLoS Pathog. *9*, e1003450.

Pütz, S., Reinders, J., Reinders, Y., and Sickmann, A. (2005). Mass spectrometry-based peptide quantification: applications and limitations. Expert Rev. Proteomics *2*, 381–392.

Raskob, G. (2014). Thrombosis : A major contributor to global disease burden. ISTH Steer. Committe World Thromb. Day Thromb. Glob. Burd. *2010*, 843–852.

Reed, G.L. (2004). Platelet secretory mechanisms. Semin. Thromb. Hemost. *30*, 441–450.

Renné, T., and Gailani, D. (2007). Role of Factor XII in hemostasis and thrombosis: clinical implications. Expert Rev. Cardiovasc. Ther. *5*, 733–741.

Reno, H.E., and Novak, R.J. (2005). Characterization of apyrase-like activity in Ochlerotatus triseriatus, Ochlerotatus hendersoni, and Aedes aegypti. Am. J. Trop. Med. Hyg. *73*, 541–545.

Ribeiro, J.M.C., and Francischetti, I.M.B. (2003). Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Annu. Rev. Entomol. *48*, 73–88.

Ribeiro, J.M., Marinotti, O., and Gonzales, R. (1990a). A salivary vasodilator in the blood-sucking bug, Rhodnius prolixus. Br. J. Pharmacol. *101*, 932–936.

Ribeiro, J.M., Endris, T.M., and Endris, R. (1991). Saliva of the soft tick, Ornithodoros moubata, contains anti-platelet and apyrase activities. Comp. Biochem. Physiol. A. Comp. Physiol. *100*, 109–112.

Ribeiro, J.M., Schneider, M., and Guimarães, J. a (1995). Purification and characterization of prolixin S (nitrophorin 2), the salivary anticoagulant of the blood-sucking bug Rhodnius prolixus. Biochem. J. *308 (Pt 1*, 243–249.

Ribeiro, J.M., Anderson, J.M., Manoukis, N.C., Meng, Z., and Francischetti, I.M. (2011). A further insight into the sialome of the tropical bont tick, Amblyomma variegatum. BMC Genomics *12*, 136.

Ribeiro, J.M.C., Alarcon-Chaidez, F., Ivo, I.M., Mans, B.J., Mather, T.N., Valenzuela, J.G., and Wikel, S.K. (2006). An annotated catalog of salivary gland transcripts from Ixodes scapularis ticks. Insect Biochem. Mol. Biol. *36*, 111–129.

Ribeiro, M.C., Wms, J.J., and Iii, S. a M.R.T. (1990b). Saliva of the Tick lxodes dammini Inhibits Neutrophil Function cells and actively moving out of the vascu-

growth and reproduction . During their lar space to the area of injury . Enzymes into this host ’ skin from several days to lysozyme , glucuronidase. Exp. Parasitol. *70*, 382–388.

Richardson, J.L., Kröger, B., Hoeffken, W., Sadler, J.E., Pereira, P., Huber, R., Bode, W., and Fuentes-Prior, P. (2000). Crystal structure of the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. EMBO J. *19*, 5650–5660.

Richardson, J.L., Fuentes-Prior, P., Sadler, J.E., Huber, R., and Bode, W. (2002). Characterization of the residues involved in the human thrombin-haemadin complex: An exosite II-binding inhibitor. Biochemistry *41*, 2535–2542.

Richter, D., Matuschka, F., Spielman, A., and Mahadevan, L. (2013). How ticks get under your skin : insertion mechanics of the feeding apparatus of Ixodes ricinus ticks How ticks get under your skin : insertion mechanics of the feeding apparatus of Ixodes ricinus ticks Authors for correspondence : Proc. R. Soc. Biol. Sci.

Roche, W.R., Montefort, S., Baker, J., and Holgate, S.T. (1993). Cell adhesion molecules and the bronchial epithelium. Am. Rev. Respir. Dis. *148*, S79–S82.

Rodriguez-Valle, M., Vance, M., Moolhuijzen, P.M., Tao, X., and Lew-Tabor, A.E. (2012). Differential recognition by tick-resistant cattle of the recombinantly expressed Rhipicephalus microplus serine protease inhibitor-3 (RMS-3). Ticks Tick. Borne. Dis. *3*, 159–169.

Rosendaal, F.R. (1999). Venous thrombosis: A multicausal disease. Lancet *353*, 1167–1173.

Sadler, E., Lentz, S., Sheehan, J., Tsiang, M., and Wu, Q. (1993). Sturcture-Function Relationships of the Thrombin-Thrombomodulin Interaction. Haemostasis *23*, 183–193.

Saito, T., Kawamura, Y., Sato, N., Kano, K., Takahashi, K., Asanome, A., Sawada, J., Katayama, T., and Hasebe, N. (2015). Non-Vitamin K Antagonist Oral Anticoagulants Do Not Increase Cerebral Microbleeds. J. Stroke Cerebrovasc. Dis. *24*, 1373–1377.

Sangamnatdej, S., Paesen, G.C., Slovak, M., and Nuttall, P. a. (2002). A high affinity serotonin- and histamine-binding lipocalin from tick saliva. Insect Mol. Biol. *11*, 79–86.

Santarpia, G., De Rosa, S., Polimeni, A., Giampà, S., Micieli, M., Curcio, A., and Indolfi, C. (2015). Efficacy and Safety of Non-Vitamin K Antagonist Oral Anticoagulants versus Vitamin K Antagonist Oral Anticoagulants in Patients

Undergoing Radiofrequency Catheter Ablation of Atrial Fibrillation: A Meta-Analysis. PLoS One *10*, e0126512.

Sauer, J.R., McSwain, J.L., Bowman, a S., and Essenberg, R.C. (1995). Tick salivary gland physiology. Annu. Rev. Entomol. *40*, 245–267.

Schwarz, A., Cabezas-Cruz, A., Kopecký, J., and Valdés, J.J. (2014). Understanding the evolutionary structural variability and target specificity of tick salivary Kunitz peptides using next generation transcriptome data. BMC Evol. Biol. *14*, 4.

She, Z., Liu, X., Kotamraju, V.R., and Ruoslahti, E. (2014). Clot-Targeted Micellar Formulation Improves Anticoagulation E ffi cacy of Bivalirudin. 10139–10149.

Shi, S.Y., Zhang, Y.P., Jiang, X.Y., Chen, X.Q., Huang, K.L., Zhou, H.H., and Jiang, X.Y. (2009). Coupling HPLC to on-line, post-column (bio)chemical assays for high-resolution screening of bioactive compounds from complex mixtures. TrAC - Trends Anal. Chem. *28*, 865–877.

Shin, Y.G., and van Breemen, R.B. (2001). Analysis and screening of combinatorial libraries using mass spectrometry. Biopharm. Drug Dispos. *22*, 353–372.

Šimo, L., Kocáková, P., and Sláviková, M. (2004). Dermacentor reticulatus (Acari, Ixodidae) female feeding in laboratory. Biol. … 655–660.

Simone, G.D.E., Lombardi, A., Galdero, S., Nastri, F., Morte, R. Della, Staian, N., Pedone, C., Bolognesi, M., and Pavone, V. (1998). The crystal structure of human a-thrombin-hirunom V complex. October.

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M., and Birol, I. (2009). ABySS: A parallel assembler for short read sequence data. Genome Res. *19*, 1117–1123.

Skrzypczak-Jankun, E., Carperos, V.E., Ravichandran, K.G., Tulinsky, a, Westbrook, M., and Maraganore, J.M. (1991). Structure of the hirugen and hirulog 1 complexes of alpha-thrombin. J. Mol. Biol. *221*, 1379–1393.

Slovak, M., Labuda, M., and Marley, S.E. (2002). Mass laboratory rearing of Dermacentor reticulatus ticks (Acarina, Ixodidae). Biol. Bratislava *57*, 261–266.

Smith, J., and Stanfield, G.M. (2012). A seminal fluid protease activates sperm motility in C. elegans males. Worm *1*, 0–3.

Soares, T.S., Watanabe, R.M.O., Tanaka-Azevedo, A.M., Torquato, R.J.S., Lu, S., Figueiredo, A.C., Pereira, P.J.B., and Tanaka, A.S. (2012). Expression and functional characterization of boophilin, a thrombin inhibitor from Rhipicephalus (Boophilus) microplus midgut. Vet. Parasitol. *187*, 521–528.

Stassens, P., Bergum, P.W., Gansemans, Y., Jespers, L., Laroche, Y., Huang, S., Maki, S., Messens, J., Lauwereys, M., Cappello, M., et al. (1996). Anticoagulant repertoire of the hookworm Ancylostoma caninum. Proc. Natl. Acad. Sci. U. S. A. *93*, 2149–2154.

Stone, S.R., and Hofsteenge, J. (1986). Kinetics of the inhibition of thrombin by hirudin. Biochemistry *25*, 4622–4628.

Strube, K., Krogers, B., Bialojan, S., Otteq, M., and Dodtq, J. (1993). Isolation , Sequence Analysis , and Cloning of Haemadin. *268*, 8590–8595.

Stubbs, M.T., and Bode, W. (1993). A player of many parts: The spotlight falls on thrombin's structure. Thromb. Res. *69*, 1–58.

Szabó, C., Ischiropoulos, H., and Radi, R. (2007). Peroxynitrite: biochemistry, pathophysiology and development of therapeutics. Nat. Rev. Drug Discov. *6*, 662–680.

Szlam, F. (2007). Thrombin Generation Assay and Viscoelastic Coagulation Monitors Demonstrate Differences in the Mode of Thrombin Inhibition Between Unfractionated Heparin and Bivalirudin. *105*, 933–939.

Tan, A.W.L., Francischetti, I.M.B., Slovak, M., Kini, R.M., and Ribeiro, J.M.C. (2015). Sexual differences in the sialomes of the zebra tick, Rhipicephalus pulchellus. J. Proteomics *117*, 120–144.

Tanaka, a S., Andreotti, R., Gomes, a, Torquato, R.J., Sampaio, M.U., and Sampaio, C. a (1999). A double headed serine proteinase inhibitor--human plasma kallikrein and elastase inhibitor--from Boophilus microplus larvae. Immunopharmacology *45*, 171–177.

Tanaka, K. a., Szlam, F., Sun, H.Y., Taketomi, T., and Levy, J.H. (2007). Thrombin generation assay and viscoelastic coagulation monitors demonstrate differences in the mode of thrombin inhibition between unfractionated heparin and bivalirudin. Anesth. Analg. *105*, 933–939.

Tuszynski, G., Gasic, T., and Gasic, G. (1987). Isolation and characterization of Antistatin. J. Biol. Chem. *168*, 493–501.

Valenzuela, J.G. (2004). Exploring tick saliva: from biochemistry to "sialomes" and functional genomics. Parasitology *129 Suppl*, S83–S94.

Vančová, I., Hajnická, V., Slovák, M., Kocáková, P., Paesen, G.C., and Nuttall, P. a. (2010). Evasin-3-like anti-chemokine activity in salivary gland extracts of ixodid ticks during blood-feeding: A new target for tick control: Brief Definitive Report. Parasite Immunol. *32*, 460–463.

Versteeg, H.H., Heemskerk, J.W.M., Levi, M., and Reitsma, P.H. (2013). New fundamentals in hemostasis. Physiol. Rev. *93*, 327–358.

Veveris-Lowe, T.L., Kruger, S.J., Walsh, T., Gardiner, R. a., and Clements, J. a. (2007). Seminal fluid characterization for male fertility and prostate cancer: Kallikrein-related serine proteases and whole proteome approaches. Semin. Thromb. Hemost. *33*, 87–99.

Vyasa, B.M., Dave, R.D., Daniel, P.S., Anand, I.S., and Patel, C.N. (2011). A view on combination antiplatelet agents in ischemic stroke. Int. J. Pharm. Res. *3*, 15–19.

Wagner, G.P., Kin, K., and Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. *131*, 281–285.

Wakefield, T.W., Myers, D.D., and Henke, P.K. (2008). Mechanisms of venous thrombosis and resolution. Arterioscler. Thromb. Vasc. Biol. *28*, 387–391.

Wallace, E.L., and Smyth, S.S. (2013). Targeting platelet thrombin receptor signaling to prevent thrombosis. Pharmaceuticals *6*, 915–928.

Wang, H., and Nuttall, P. a (1995). Immunoglobulin G binding proteins in male Rhipicephalus appendiculatus ticks. Parasite Immunol. *17*, 517–524.

Warkentin, T.E. (2004). Bivalent direct thrombin inhibitors: Hirudin and bivalirudin. Best Pract. Res. Clin. Haematol. *17*, 105–125.

Waxman, L., Smith, D.E., Arcuri, K.E., and Vlasuk, G.P. (1990). Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. Science *248*, 593–596.

Wesselschmidt, R., Likert, K., Huang, Z., MacPhail, L., and Broze, G.J. (1993). Structural requirements for TFPI interactions with fXa and heparin.pdf. Blood Coagul. Fibrinolysis *4*, 661–669.

Wiese, S., Reidegeld, K. a., Meyer, H.E., and Warscheid, B. (2007). Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. Proteomics *7*, 340–350.

Wigglesworth, M.J., Murray, D.C., Blackett, C.J., Kossenjans, M., and Nissink, J.W.M. (2015). Increasing the delivery of next generation therapeutics from high throughput screening libraries. Curr. Opin. Chem. Biol. *26*, 104–110.

Wójcik-Fatla, A., Cisak, E., Zajac, V., Zwoliński, J., and Dutkiewicz, J. (2011). Prevalence of tick-borne encephalitis virus in Ixodes ricinus and Dermacentor reticulatus ticks collected from the Lublin region (eastern Poland). Ticks Tick. Borne. Dis. *2*, 16–19.

Wu, K.K., and Matijevic-Aleksic, N. (2005). Molecular aspects of thrombosis and antithrombotic drugs. Crit. Rev. Clin. Lab. Sci. *42*, 249–277.

Xu, X., Yang, H., Ma, D., Wu, J., Wang, Y., Song, Y., Wang, X., Lu, Y., Yang, J., and Lai, R. (2008). Toward an understanding of the molecular mechanism for successful blood feeding by coupling proteomics analysis with pharmacological testing of horsefly salivary glands. Mol. Cell. Proteomics *7*, 582–590.

Yadav, V.K., Chhikara, N., Gill, K., Dey, S., Singh, S., and Yadav, S. (2013). Three low molecular weight cysteine proteinase inhibitors of human seminal fluid: Purification and enzyme kinetic properties. Biochimie *95*, 1552–1559.

Yeh, C.H., Hogg, K., and Weitz, J.I. (2015). Overview of the New Oral Anticoagulants: Opportunities and Challenges. Arterioscler. Thromb. Vasc. Biol. 1056–1066.

Zhang, J., Chung, T., and Oldenburg, K. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J. Biomol. Screen. *4*, 67–73.

Zhao, Q.-Y., Wang, Y., Kong, Y.-M., Luo, D., Li, X., and Hao, P. (2011). Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. BMC Bioinformatics *12*, S2.

LIST OF PUBLICATIONS

Journal articles

- **Iyer, J.K**., Otvos, R. A., Kool, J., Kini R.M. Microfluidic chip-based online screening coupled to mass spectrometry: Identification of inhibitors of thrombin and factor Xa. Accepted: J. Biomol. Screen.

- Otvos, R.A., **Iyer, J.K**., van Elk, R., Ulens, C., Niessen, W.M., Somsen, G.W., Kini, R.M., Smit, A.B., Kool, J., Development of a plate reader and on-line microfluidic fluorescence enhancement assay for screening of pure compounds and mixtures to identify ligands of the 5-hydroxytryptamine binding protein. Toxins, 2015: 24(7), 2336-53.

- Sharma, M., Das, D., **Iyer, J.K**., Kini, R.M., Doley, R., Unveiling the complexities of Daboia russelii venom, a medically important snake of India, by tandem mass spectrometry. Toxicon, 2015 Jul 6. pii: S0041-0101(15)30010-6.

Patents

- Iyer, J.K., Koh, C. Y., Kazimirova, M., Roller, L., Imamura and Kini R.M. Novel thrombin inhibitors. *Provisional patent filed*.

Conference presentations

- **Iyer, J.K**., Kazimirova, M., Koh, C. Y., Jobichen, C., Imamura, T., Mizuguchi, J., Iwanaga, S., Swaminathan, K., Chan, M.Y., Nuttall, P., Kini R.M., Avathrin, a novel thrombin inhibitor from *Amblyomma variegatum.*

ISTH 2015 Congress, Toronto, Canada, Oral presentation, June 2015, (Chairperson of session: Natural anticoagulants)

- **Iyer, J.K**., Slovak., M., Francischetti. I., Ribeiro, J., and Kini, R. M., Transcriptomic and proteomic profiling of *Dermacentor reticulatus* salivary glands for identification of novel feeding stage specific proteins. ISTH 2015 Congress, Toronto, Canada, Poster presentation, June 2015

- **Iyer, J.K**., Otvos, R. A., Kool, J., Kini R.M., Novel screening techniques for identification of thrombin and FXa inhibitors. 18th BSGC, Kuala Lumpur, Malaysia, Oral presentation, January 2014

- **Iyer, J.K**., Otvos, R. A., Kool, J., Kini R.M., Rapid high throughput bioassay for identification of anticoagulants from complex mixtures. ISTH 2013 Congress, Amsterdam, The Netherlands, Poster presentation, June 2013

- **Iyer, J.K**., Slovak., M., and Kini R. M., Novel anticoagulants from hematophagous animals. EFATH 2013 Conference, Amsterdam, The Netherlands, Oral presentation, June 2013

- **Iyer, J.K**., Slovak., M., and Kini R. M., Antiplatelets from salivary gland extracts of *Rhipicephalus pulchellus*. 17th BSGC, Bangkok, Thailand, Poster presentation, December 2011

APPENDIX

- DVD contains PDB co-ordinate file of avathrin thrombin crystal structure
- Transcriptome sequences and classification of *Dermacentor reticulatus*
- iTRAQ proteome and sequences and classification of *Dermacentor reticulatus*