University of South Bohemia in České Budějovice

Faculty of Science

# The production and purification of tick salivary serpins in bacterial expression system

Bachelor Thesis in Biological Chemistry

Simone Röhrnbacher

České Budějovice, 2018

Supervisor: RNDr. Jindřich Chmelař Ph.D.

Co-Supervisor: MSc. Jan Kotál

Faculty of Medical Biology

**Bibliographical reference:**

Röhrnbacher, S., 2018: The production and purification, of tick salivary serpins in bacterial expression system. BSc, thesis, in English. – 52 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

**Annotation:**

The aim of this thesis was to produce several tick salivary serine protease inhibitors and a chimeric mouse-tick serpin in a bacterial expression system, to refold and purify the proteins and to finally test them in protease inhibitor assays for drawing early conclusions about their function in tick-host interactions.

**Declaration:**

I hereby declare that I have worked on my bachelor's thesis independently and used only the sources listed in the bibliography. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defence in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

České Budějovice, Date Simone Röhrnbacher

**Acknowledgements:**

First, I want to thank my supervisor, RNDr. Jindřich Chmelař Ph.D., for giving me the chance to work on this interesting project and for always providing help whenever needed.

I am also very thankful to my co-supervisor, MSc. Jan Kotál, who assisted me throughout my whole project, especially with issues and problems regarding the laboratory work.

Big thanks go to everyone else, who made me feel welcome and contributed to making this project worthwhile.

# Table of contents:

# List of abbreviations:

| | |
|---|---|
| **ADP** | Adenosine diphosphate |
| **BCA** | Bicinchoninic acid |
| **BSA** | Bovine serum albumin |
| **cDNA** | Complementary DNA |
| **dNTP** | Deoxy nucleoside triphosphate |
| **DTT** | Dithiothreitol |
| **EtOH** | Ethanol |
| **FVII** | Factor VII |
| **FVIII** | Factor VIII |
| **FIX** | Factor IX |
| **FX** | Factor X |
| **FXII** | Factor XII |
| **Glu** | Glucose |
| **IB** | Inclusion bodies |
| **IEC** | Ion exchange chromatography |
| **IPTG** | Isopropyl-β-D-1-thiogalactopyranoside |
| **IRS** | *Ixodes ricinus* serpin |
| **LB** | Lysogeny broth |
| **OD** | Optical density |
| **pI** | Isoelectric point |
| **PCR** | Polymerase chain reaction |
| **RCL** | Reactive centre loop |
| **Rpm** | Rotations per minute |
| **SDS-PAGE** | Sodium dodecyl sulphate – polyacrylamide gel electrophoresis |
| **SEC** | Size exclusion chromatography |
| **tPA** | Tissue plasminogen activator |
| **uPA** | Urokinase-type plasminogen activator |

# 1 Introduction

Ticks are obligate hematophagous ectoparasites of vertebrates such as mammals, birds, reptiles and amphibians (Francischetti et al., 2009). They cause concerns for human and animal health as they are disease vectors of illnesses such as Lyme disease (Pearson, 2014) or tick-borne encephalitis (Amicizia et al., 2013). They are segregated into three families, namely Argasidae (soft ticks), Ixodidae (hard ticks) and Nuttalliellidae (Nava et al., 2009), which employ different feeding strategies. While soft ticks feed for less than an hour, hard ticks remain attached to the host from a few days up to two weeks. As hosts apply various defence mechanisms during the long feeding period, ticks evolved to counter these measures with numerous means, leading to complex tick-host interactions (Ribeiro, Francischetti, 2003).

## 1.1 Tick-host interaction

The often redundant host defence mechanisms at the cutaneous interface depend, among others, on protease cascades or bioactive peptides, like chemokines, which are activated proteolytically (Chmelar et al., 2011). Proteases cleave proteins into smaller units and play important roles in various processes, such as immunity or haemostasis (Amara et al., 2008). Other mechanisms like cytokine activation, phagocytosis or antigen presenting also rely on proteolysis (Müller et al., 2012). The proteases are tightly regulated and there is a delicate balance between activation and endogenous inhibition. Due to this sensitive equilibrium, many parasites, including ticks, have evolved to target and inhibit proteases involved in host defence (Chmelař et al., 2017).

### 1.1.1 Tick-host interaction – Host defence mechanisms

Host defence mechanisms can be divided into three parts: 1. haemostasis, which prevents blood loss, 2. innate immunity, consisting of the complement cascade and inflammation, which causes itching or pain and might make the host aware of the presence of the parasite and 3. mainly antigen-specific acquired immunity during reoccurring tick exposure (Chmelar et al., 2012, Francischetti et al., 2009).

#### 1.1.1.1 Haemostasis

Haemostasis is one of the first lines of defence and is initiated immediately after tick mouth parts have penetrated the host tissue. The first phase of haemostasis is blood coagulation, which leads to the formation of fibrin clots. It can be triggered by intrinsic or extrinsic factors, which in turn leads to activation of a complex cascade of different factors and enzymes. This cascade generally involves several compounds, such as extravascular tissue factors and serine

proteases FVII, FVIII, FIX, FX and FXII, that finally activate thrombin (Broze, 1995, Chmelar et al., 2012, Mann et al., 2003).

Thrombin in turn cleaves fibrinogen to fibrin, which is needed for clot formation. It then supports platelet aggregation that leads to thrombus formation and wound healing. It also plays a role in immunity, since platelets release inflammatory mediators such as chemokines. Triggers for platelet aggregation include collagen, thrombin, Cathepsin G and ADP, which are released by damaged cells. Those activators change the condition of fibrinogen receptors and thus a fibrin-platelet thrombus is formed (Chmelar et al., 2012, Francischetti et al., 2009).

The third process, vasoconstriction, reduces the blood flow to the site of the injury and is activated by compounds released by platelets such as serotonin and thromboxane. Other substances involved in vasoconstriction comprise compounds released from different origins, such as the endothelium or mast cells as well as compounds present in the blood (Chmelar et al., 2012, Francischetti et al., 2009).

#### 1.1.1.2 Innate Immunity

Besides haemostasis early host response also includes measures from the innate immune system. This includes the complement cascade and inflammation, which causes itching and pain, posing another challenge for successful tick feeding (Chmelar et al., 2011, Ribeiro, Francischetti, 2003). The innate immune system consists of various types of cells that are present in tissues in the vicinity of the tick bite and that secrete different types of mediators to induced inflammation and the complement cascade.

Keratinocytes, which release pro-inflammatory mediators (Heath, Carbone, 2013), as well as leucocytes are amidst the first immune cells to encounter foreign particles. Mast cells, a type of leukocyte, are found in connective tissue especially beneath the epithelium. They release mediators including serine proteases (Metcalfe et al., 1997) such as tryptase and chymase, which show inflammatory functions (Gilfillan, Metcalfe, 2011), Moreover, inflammatory cells are recruited to the site of the tick bite by other compounds released by mast cells (Metcalfe et al., 1997). Another type of leukocytes is the antigen presenting dendritic cell. They first engulf parasitic cells via phagocytosis and subsequently present antigens to stimulate T-cells (Banchereau, Steinman, 1998). Eosinophils found in tissues such as skin are a source of mast cell activators (Rothenberg, Hogan, 2006) and molecules involved in tissue repair & inflammation (Elovic et al., 1998). Finally, macrophages, another type of phagocyte, react to foreign molecules by releasing pro-inflammatory chemokines, cytokines and growth factors. They then attract inflammatory cells, including neutrophils that engulf

microorganisms and (Francischetti et al., 2009) also release serine proteases such as neutrophil elastase, proteinase 3 and cathepsin G. These proteases are involved in pathogen destruction and regulation of inflammation (Segal, 2005). Serine proteases uPA and tPA, which are released in response to certain stimuli proteolytically transform plasminogen to plasmin. Plasmin then plays a role in tissue modelling, complement activation and cell migration (Collen, 1999, Saksela, Rifkin, 1988, Schaiff, Eisenberg, 1997).

#### 1.1.1.3 Adaptive Immunity

After the host has been exposed to a tick once, the adaptive immune system, consisting of cellular and humoral immune response, is activated upon new tick infestation. Memory T-cells react by releasing antibodies in response to antigen presenting dendritic cells (Banchereau, Steinman, 1998) and B-cells release cytokines. These molecules then detect components of tick saliva and subsequently activate the complement cascade or sensitise leukocytes (Bowman, Nuttall, 2008, Brossard, Wikel, 2004, Wikel, 2013). Moreover, Cytotoxic lymphocytes release granzymes, which are also serine proteases, that induce apoptosis (Chávez-Galán et al., 2009).

### 1.1.2 Tick-Host Interaction – How ticks influence host defence mechanisms

The sophisticated host defence mechanisms explained in 1.1.1, relying on proteases, must be supressed and evaded by comparably elaborate strategies employed by ticks. This effect is mainly accomplished by the bioactive molecules of tick saliva that are injected into the host (Kotál et al., 2015). Tick saliva contains hundreds of pharmacoactive molecules that suppress haemostasis (Chmelar et al., 2012), inflammation (Chmelar et al., 2011), complement activation (Couvreur et al., 2008) or act as immunomodulators (Ramamoorthi et al., 2005, Sá-Nunes et al., 2009). Besides tick proteins actively suppressing host defence mechanisms, the properties of those proteins also help to evade the host immune system.

#### 1.1.2.1 Evasion of host immune system

Salivary proteins can be grouped into large multigenic families containing many individual proteins that only differ by a few amino acids (Schwarz et al., 2013). They show pluripotency and redundancy to avoid defence mechanisms of the host. Pluripotency means that one salivary protein has more than one target in the host, while redundancy indicates that one host defence mechanisms is targeted by more than one salivary protein. (Chmelař et al., 2016).

In tick saliva each protein is only present in low concentrations, which are too low to induce any observable outcomes in ex vivo testing, but it is theorised, that the existence of many similar proteins causes an additive effect in tick saliva. Additionally, sequential

expression of similar proteins leads to an evasion of host immune reactions, while still blocking their target pathway. Low molecular weight proteins also show lower antigen immunogenicity and post-translational modifications might also lower immunogenicity of salivary proteins. These four reasons are speculated to be responsible for the ability of ticks to evade the host immune system (Chmelař et al., 2016).

#### 1.1.2.2 Suppression of host defence – Protease inhibitors

Since proteases play a crucial role in host defence mechanisms, ticks have evolved to inhibit many of these processes with protease inhibitors (Amara et al., 2008, Chmelar et al., 2012, Couvreur et al., 2008). Protease inhibitors are the biggest group of salivary proteins. Four main types of tick protease inhibitors are currently described in tick saliva. Kunitz domain inhibitors are haemostatic modulators, which show anti-coagulative effects as well as inhibiting platelet aggregation (Corral-Rodríguez et al., 2009). One Kunitz domain is a small peptide which usually weighs 7 kDa, but many members consist of several domains (Chmelar et al., 2012). The second group is formed by cystatins (cysteine protease inhibitors) that mainly modulate the host immune system (Ibelli et al., 2013). The third group, Trypsin inhibitor like cysteine rich domain (TIL) containing proteins have conserved 5-disulhpide bridges (Bania et al., 1999). They play important roles in immune response and show anticoagulative (Chen et al., 2013) and antimicrobial properties (Fogaça et al., 2006). Serpins, the last group, also have effects on haemostasis (Stark, James, 1998) and the immune system (Ooi et al., 2015). Additionally, there are other small groups of proteases as well as types of proteins with unknown functions (Chmelař et al., 2016).

## 1.2 Serpins

Serpins (**Ser**ine **p**rotease **in**hibitors) are present in all groups of organisms, including viruses, prokaryotes, and eukaryotes and they fulfil a broad range of functions (Law et al., 2006). In humans 29 inhibitory and 7 non-inhibitory serpins have been found, while mice DNA contains 60 functional serpin genes (Heit et al., 2013). In humans serpins play a role in the immune system and in haemostasis, among others. For example α-III-antithrombin, is a human endogenous inhibitor of thrombin, FX and FIX (Chmelar et al., 2012). Human non-inhibitory serpins are angiotensin which regulates vasoconstriction and blood pressure (Lu et al., 2016), heat shock proteins (Clarke et al., 1991) or hormone transporting serpins, like cortisol and thyroxine binding proteins (Hammond et al., 1987).

### 1.2.1 Serpin structure and mode of action

Serpins are large 350-400 amino acids long proteins and possess a molecular weight of 40 to 50 kDa. They mostly inhibit serine proteases (Patston, 2000). One serpin molecule generally consists of a conserved secondary structure of 3 β-sheets (A, B, C) and 8-9 α-helixes (hA-hI) (Elliott et al., 1996). A feature of salivary tick serpins like IRS-2 is the presence of a signal peptide, which is cleaved before secretion (Chmelar et al., 2011).

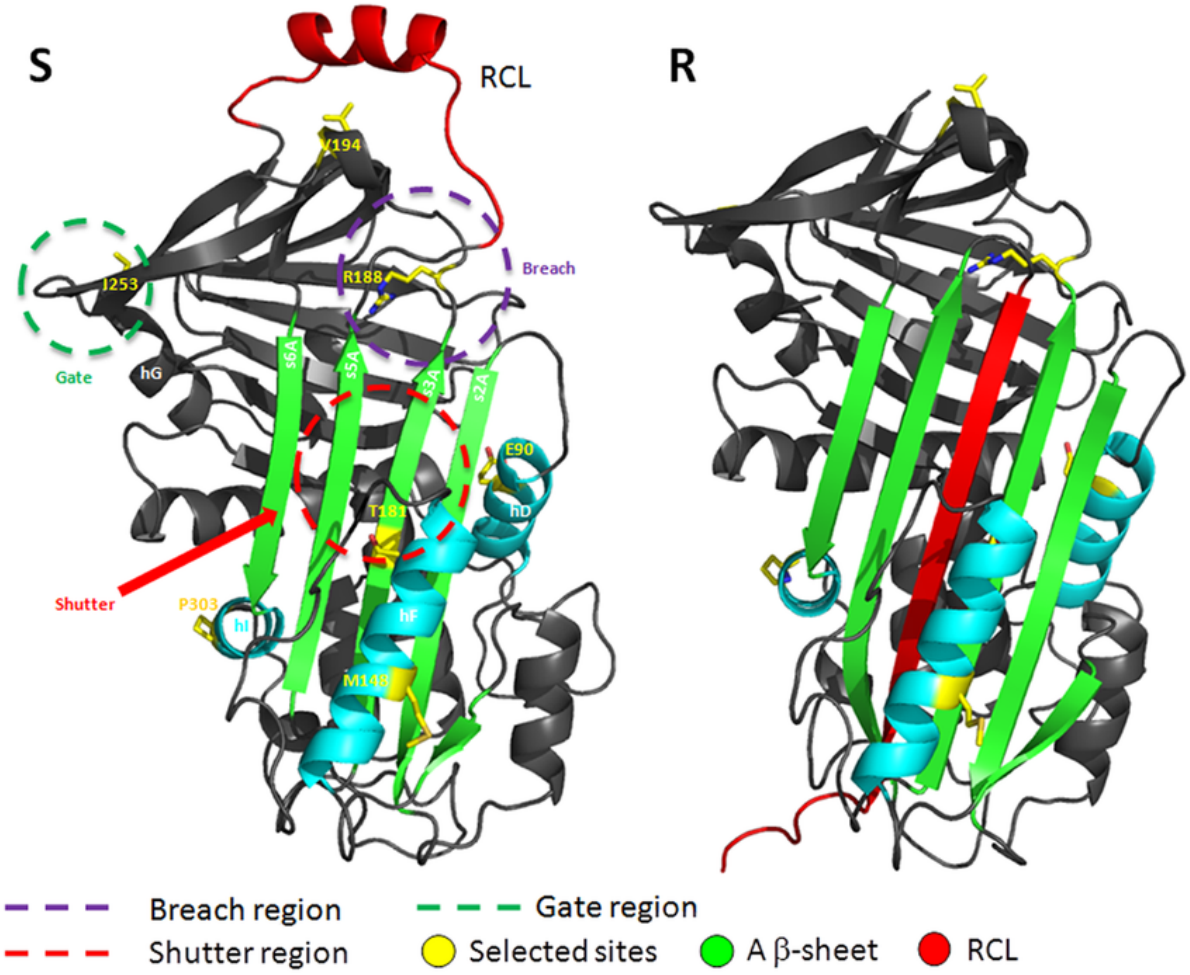


Figure 1: S-state of the active serpin with RCL extended above the protein; R-state after substrate binding which causes the RCL to move into the protein due to conformational changes. Figure adopted from (Seixas et al., 2012).

A serpin molecule contains a region, called reactive centre loop (RCL), an approximately 17 amino acids long region between beta sheets A and C, which interacts with the target proteases (Silverman et al., 2001). Especially the so called P1/P1' cleavage site within the RCL is important for the serpins' function. This site interacts with the corresponding site of the substrate, the S1 site (Schechter, Berger, 1968). Even though the P1 site alone does not fully dictate a serpin's function, predictions can be made. Serpins with Arginine (R) or Lysine (K) at the P1 position likely inhibit trypsin-like proteases, while Phenylalanine (F),

Tyrosine (Y), Leucine (L) or Isoleucine (I) indicates anti-chymotrypsin activity. Alanine (A) or Valine (V) residues predict elastase as the target protease (Waterhouse et al., 2007, Zhao et al., 2012). After inhibiting the target protease proteolytically, the serpin loses its function as it is also proteolysed at the P1/P1' site, which induces a conformational change (Huntington et al., 2000). Its conformation changes from the so-called S ("stressed") state with the (red) RCL being extended above the serpin to bait a target molecule to the R ("relaxed") state, which can be seen in Figure 1. In the R state the cleaved amino-terminus of the RCL is inserted into β-sheet A to form a fourth β-strand (red) and the inhibitor is covalently linked to the protease. A serpin-protease complex is formed (Huntington et al., 2000). The R conformation shows a higher stability than the S state, the native form of a serpin (Cabrita, Bottomley, 2004). This permanent conformational change leads to inactivity and the loss of the ability to further inhibit other target molecules. This is why serpins are called single-use or suicide inhibitors (Huntington et al., 2000).

### 1.2.2 Serpins in ticks

In arthropods, including ticks, serpins regulate the immune system and haemostatic processes in addition to being modulators of parasite-host interactions in the case of parasitic arthropods such as ticks or mosquitos (Meekins et al., 2017). Approximately 20 serpins from various tick species have been characterised. Many show anti-inflammatory effects or anti-haemostatic effects by preventing platelet aggregation and prolonging clotting time. They target different serine proteases involved in the haemostatic process, such as thrombin, cathepsin G, elastase, Plasmin or FX (Chalaire et al., 2011, Kim et al., 2015, Porter et al., 2015, Tirloni et al., 2014).

Two of the characterised tick serpins were described in *I. ricinus*, namely Iris (*I. ricinus* immunosuppressor) and IRS-2 (*I. ricinus* serpin 2). Iris inhibits leukocyte elastase and influences contact phase coagulation, fibrinolysis as well as platelet aggregation. Additionally, it also modulates the innate and adaptive immune system by altering T-lymphocyte and macrophage responsiveness (Leboulle et al., 2002).

The structure of IRS-2 has already been elucidated. It has a molecular weight of 41.9 kDa and is comprised of 3 large β-sheets and 9 α-helices. It has a predicted signal peptide of 21 amino acids and tyrosine was predicted to be at the P1 position in the RCL, indicating it to be an inhibitor of α-chymotrypsin like proteases (Chmelar et al., 2011). The speicificity of inhibition was shown by Chemlař et al. and indicates that IRS-2 is an inhibitor of chymase and cathepsin G. Thus, IRS-2 suppresses inflammatory response, swelling and the migration of neutrophils to the site of the tick bite (Chmelar et al., 2011).

### 1.2.3 Tick serpins as novel drugs

Human serine proteases targeted by tick protease inhibitors play an important role in many diseases. Serine proteases are linked to the metastasis of cancer to the lungs (El Rayes et al., 2015), some lung conditions such as cystic fibrosis (Wagner et al., 2016) or chronic obstructive pulmonary disease (Owen, 2008). They also play a role in metabolic syndrome and obesity affected by neutrophil proteases (Mansuy-Aubert et al., 2013). In the future, medicine might take advantage of the large and diverse amount of protease inhibitors present in tick saliva to improve treatment options for some diseases (Chmelař et al., 2017). These molecules are especially interesting since they target only specific molecular pathways in contrast to present drugs, that induce many unwanted effects (Chmelar et al., 2012). Creating chimeric proteins from these molecules might increase their potential even more as immunogenicity could be reduced.

#### 1.2.3.1 Chimeric or fusion proteins

Domains are the units of a protein and each protein can consist of one or multiple domains. If one domain is inserted into another domain, a chimeric protein or fusion protein is created, either through a natural process or through protein engineering (Russell, 1994). This can improve the properties of the protein or add a new function (Yu et al., 2015). In research fusion proteins have already been used for different applications such as protein tags for purification (Bell et al., 2013) or for monitoring protein expression levels with GFP (Tsien, 1998). They can also be used as new therapeutic agents. There are already a few approved drugs, such as immunosuppressive and anti-inflammatory medications or cancer drugs which are fusion proteins. Generally, one domain interacts with the target substrate, while the other domain adds or improves a feature such as stability (Yu et al., 2015) or lower immunogenicity (Marcet-Palacios et al., 2015).

An example of experiments performed with a chimeric serine protease inhibitor is Serpina3n, a mouse extracellular serpin, fused with human anti-chymotrypsin. Serpina3n is an inhibitor of Granzyme B, which plays a role in inflammation and autoimmunity, thus making this chimeric protein a potential anti-inflammatory agent. By replacing the RCL of human anti-chymotrpysin with the one of serpina3n, the serine inhibitor properties of serpina3n were retained. The main part of the fusion protein, human anti-chymotrypsin, should prevent immune reaction (Marcet-Palacios M. et al, 2015).

As tick saliva contains serine proteins inhibitors, chimeric tick serpins could play a role in the treatment of diseases involving serine proteases. IRS-2, which has been shown to inhibit

cathepsin G and chymase, prevents host inflammation and platelet aggregation (Chmelar et al., 2011). Thus it has the potential to be used as a drug by creating a chimeric protein in a similar manner Marcet-Palacios M. et al have done.

# 2 Aims

1. To clone various *I. ricinus* salivary serpins into pET-17b expression vector
2. To overexpress the serpins in an *E. coli* expression system
3. To isolate the serpins from bacterial inclusion bodies and to purify the recombinant proteins
4. To test the basic function of the serpins in serine protease inhibitor assays

# 3 Materials and methods

## 3.1 Materials

Table 1: Purchased substances and materials with manufacturer and catalogue number as well as composition of prepared chemicals.

| General chemicals | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| Acetic Acid 99.8 % G.R. | lach:ner | 10047-A9B |
| Calcium chloride anhydrous powder | lach:ner | 30096-AP1 |
| DL-Dithiothreitol | Sigma-Aldrich | D9779 |
| Ethanol for molecular biology | Millipore | 108543 |
| Ethylenediaminetetraacetic acid | Sigma-Aldrich | EDS |
| Guanidine hydrochloride | Sigma-Aldrich | G4505 |
| Hydrochloric acid | lach:ner | 10033-A35 |
| Magnesium chloride hexahydrate | lach:ner | 30080-AP0 |
| Potassium chloride | lach:ner | 30076-FP0 |
| Sodium chloride | lach:ner | 30093-AP0 |
| Sodium hydroxide | lach:ner | 10006-AP2 |
| Triton™ X-100 | Sigma-Aldrich | X100 |
| Trizma® base | Sigma-Aldrich | T1503 |

| Agarose gel electrophoresis | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| Agarose SERVA for DNA Electrophoresis | SERVA | 11404 |
| Ethidium bromide aqueous solution 1 % w/v | SERVA | 21251 |
| DNA Gel Loading Dye (6X) | Thermo Fisher Scientific | R0611 |
| GeneRuler 100 bp Plus DNA Ladder | Thermo Fisher Scientific | SM0323 |
| TrackIt™ 100 bp DNA Ladder | Invitrogen | 10488058 |
| | | |
| **Chemical name** | **Composition** | |
| 1x TAE Buffer | 40mM Tris, 20mM Acetate, 1mM EDTA | |

| DNA agarose gel extraction | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| GeneJET Gel Extraction Kit | Thermo Fisher Scientific | K0691 |

| Transformation | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| One Shot™ TOP10 Chemically Competent E. coli | Invitrogen | C404003 |
| One Shot™ BL21(DE3)pLysS Chemically Competent E. coli | Invitrogen | C606010 |
| S.O.C. Medium | Invitrogen | 15544034 |
| pET-17b DNA | Novagen | 69663 |

| Plasmid isolation | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| GeneJET Plasmid Miniprep Kit | Thermo Fisher Scientific | K0502 |

| Bacterial cultivation | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| Agar, Bacteriological | Amresco | J637 |
| LB Broth, Miller (Luria-Bertani) | Amresco | J106 |
| Ampicillin sodium salt | Sigma-Aldrich | A9518 |
| Chloramphenicol | Sigma-Aldrich | C0378 |
| Glycerol anhydrous | lach:ner | 40058-AT0 |
| | | |
| **Chemical name** | **Composition** | |
| Ampicillin 1000x stock | 50 mg $mL^{-1}$ in $dH_2O$ | |
| Chloramphenicol 1000x stock | 34 mg $mL^{-1}$ in Ethanol | |
| LB medium | 2.5 % LB broth in $dH_2O$, autoclaved | |
| LB-Agar petri dishes | 2.5 % LB, 1.5 % Agar, autoclaved | |
| Glycerol solution | 80 % in $dH_2O$, autoclaved | |

| SDS-PAGE | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| NuPAGE™ 4-12 % Bis-Tris Protein Gels, 1.5 mm, 10-well | Invitrogen | NP0335BOX |
| NuPAGE™ MES SDS Running Buffer (20X) | Invitrogen | NP000202 |
| NuPAGE™ LDS Sample Buffer (4X) | Invitrogen | NP0007 |
| NuPAGE™ Sample Reducing Agent (10X) | Invitrogen | NP0004 |
| SeeBlue™ Plus2 Pre-stained Protein Standard | Invitrogen | LC5925 |
| Methanol | lach:ner | 20038-AT0 |
| Coomassie® Brilliant blue R 250 | Merck | 112553 |

| Chemical name | Composition |
|---|---|
| Destaining solution | 25 % MeOH, 10 % Acetic acid in $dH_2O$ |
| CBB solution | 0.1 % CBB powder in destaining solution |

| Polymerase chain reaction | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| PCR Master Mix (2X) | Thermo Fisher Scientific | K0171 |
| Water, nuclease-free | Thermo Fisher Scientific | R0581 |
| Phusion DNA Polymerase, 2 U/µL | Thermo Fisher Scientific | Included in F534S |
| 5X Phusion Green HF Buffer | Thermo Fisher Scientific | Included in F534S |
| dNTP Mix (10 mM each) | Thermo Fisher Scientific | R0192 |
| T7 Sequencing Primer | Invitrogen | included in K20001 |
| T7 Reverse Primer | Invitrogen | Included in K20001 |
| cDNA | Prepared from 1 µg RNA from fed nymphs | |

| Restriction | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| FastDigest NdeI | Thermo Fisher Scientific | FD0583 |
| FastDigest XhoI | Thermo Fisher Scientific | FD0694 |
| FastDigest Green Buffer (10X) | Thermo Fisher Scientific | B72 |
| Water, nuclease-free | Thermo Fisher Scientific | R0581 |
| IRS-2 chimeric serpin DNA | Generi Biotech | |

| Ligation | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| T4 DNA Ligase (5U $\mu L^{-1}$) | Invitrogen | 15224041 |
| 5X reaction buffer | Invitrogen | Included in 15224041 |

| Overexpression | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| OmniPur® IPTG | Calbiochem | 5800-OP |
| D-(+)-Glucose | Sigma-Aldrich | G7021 |
| Tunair™ shake flask | Sigma-Aldrich | Z710822 |
| Cap for Tunair™ shake flask | Sigma-Aldrich | Z710849 |
| Dri-gauze™ filter liner for Tunair™ shake flask | Sigma-Aldrich | Z710881 |

| Chemical name | Composition |
|---|---|
| D-Glucose solution | 20 % in $dH_2O$, autoclaved |
| IPTG 1000x stock | 1 M IPTG in $dH_2O$ |

| Inclusion body isolation & protein refolding | |
|---|---|
| **Chemical name** | **Composition** |
| Tris HCl | 20 mM Tris-HCl, pH = 8 |
| Guanidine solution | 6 M Guanidine, 20 mM Tris-HCl, pH = 7.8 |
| DTT 100x stock | 1 M DTT |
| Refolding Buffer 1 | 300 mM NaCl, 20 mM Na-Acetate, pH = 5.5 |
| Refolding Buffer 2 | 20 mM Tris-HCl, 300 mM NaCl, pH = 8.5 |
| Refolding Buffer 3 | 20 mM Tris-HCl, 300 mM NaCl, pH = 6.8 |
| Refolding Buffer 4 | 20 mM Tris-HCl, 300 mM NaCl, pH = 8.0 |
| Refolding Buffer 5 | 20 mM Tris-HCl, 10 mM NaCl, 0.5 mM KCl, 1 mM EDTA, pH = 8.0 |
| Refolding Buffer 6 | 20 mM Tris-HCl, 10 mM NaCl, 0.5 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, pH = 8.0 |
| Refolding Buffer 7 | 20 mM Tris-HCl, 240 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, pH = 8.0 |
| Refolding Buffer 8 | 20 mM Tris-HCl, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, pH = 8.0 |

| BCA assay | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| Pierce™ BCA Protein Assay Kit | Thermo Fisher Scientific | 23227 |
| Bovine Serum Albumin | Sigma-Aldrich | A7906 |
| Nunc™ MicroWell™ 96-Well Microplates | Thermo Fisher Scientific | 269620 |

| Protein concentrating | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-10 membrane | Millipore | UFC901024 |
| Vacuum Filtration "rapid"-Filtermax | TPP | 99500 |
| Amicon® Stirred Cell 400mL | Millipore | UFSC40001 |
| Ultrafiltration Discs, PLGC, Ultracel regenerated cellulose, 10 kDa NMWL, 76 mm | Millipore | PLGC07610 |

| Protein purification | | |
|---|---|---|
| **Product name** | **Manufacturer** | **Catalogue number** |
| HiPrep 16/60 Sephacryl S-100 HR | GE Healthcare | 17116501 |
| Superdex 75 10/300 GL | GE Healthcare | 17517401 |
| **Chemical name** | **Composition** | |
| Mobile phase for SEC | 20 mM Tris-HCl, pH = 8.5 | |
| Mobile phase for IEC | A: 20 mM Tris-HCl, pH = 8.5<br>B: 20 mM Tris-HCl, 1 M NaCl, pH = 8.5, | |

| Serine protease inhibitor assays | | |
|---|---|---|
| **Enzyme** | | |
| **Product name** | **Manufacturer** | **Catalogue number** |
| Trypsin | Sigma | T8802 |
| α-Chymotrypsin | Sigma | C3142 |
| Chymase | Sigma | C8118 |
| Cathepsin G | Molecular Innovations | IHCG |
| fX | EMD Millipore | 69036 |
| fXII | Haematologic Technologies Inc | HCXII-0155 |
| Plasmin | Sigma | P1867 |
| | | |
| **Substrate** | | |
| **Product name** | **Manufacturer** | **Catalogue number** |
| Boc-Val-Pro-Arg-AMC hydrochloride salt | Bachem | I-1120 |
| Chymotrypsin Substrate II, Fluorogenic | Calbiochem | 230914 |
| Suc-Leu-Leu-Val-Tyr-AMC | Bachem | I-1395 |
| SPECTROFLUOR™ FXa | Sekisui Diagnostics | 222F |
| Boc-β-benzyl-Asp-Pro-Arg-7-amido-4-methylcoumarin hydrochloride | Sigma | B4028 |
| | | |
| **Chemical name** | **Composition** | |
| Trypsin assay buffer | 50 mM Tris-HCl, pH = 8, 150 mM NaCl, 20 mM $CaCl_2$, 0.01 % Triton X-100 | |
| Kallikrein assay buffer | 20 mM Tris–HCl, pH = 8.5, 150 mM NaCl, 0.02 % Triton X-100 | |
| Elastase assay buffer | 50 mM Hepes buffer, pH = 7.4, 100 mM NaCl, 0.01 % Triton X-100 | |

### 3.1.1 Agarose gel and SDS-PAGE size / weight markers

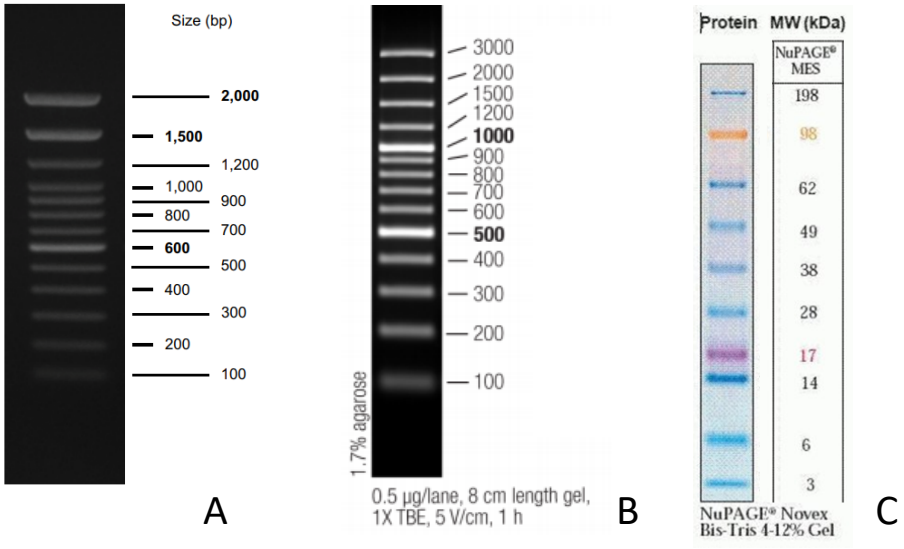


Figure 2: Fragment lengths of A - TrackIt™ 100 bp DNA Ladder (Invitrogen, 2017); B - GeneRuler 100 bp Plus DNA ladder (Thermo Scientific, 2018); protein size of C - SeeBlue™ Plus2 Pre-stained Protein Standard (Novex, 2014).

## 3.2 Methods

### 3.2.1 General methods

#### 3.2.1.1 Agarose gel electrophoresis

1 % agarose gels were prepared by dissolving and heating 0.75 g agarose in 75 mL 1x TAE buffer. After cooling the mixture down to 50 °C, 1.5 µL ethidium bromide was added to the solution before pouring it into the casting frame. The set gel was covered with 1x TAE buffer. Sample preparation involved combining 20 µL of the sample and 4 µL DNA loading dye, which was then pipetted into the wells of the gel. 10 µL of GeneRuler 100 bp Plus DNA ladder or TrackIt™ 100 bp DNA Ladder was used as a marker. Electrophoresis was carried out at a constant voltage of 100 V for 45 minutes after which the bands were visualised under UV-light.

#### 3.2.1.2 Agarose gel extraction

Desired DNA fragments were cut out of the gel with a surgical blade and GeneJET Gel Extraction Kit was used, according to the included manual to purify the retrieved fragments: All centrifugations were carried out at 13000 g. The gel slices were weighed and 1:1 volume of Binding buffer was added. They were incubated at 50-60 °C and occasionally inverted until fully dissolved. The solution was then transferred to the purification column, centrifuged for 1 minute and the flow through was discarded. After adding 700 µL Wash Buffer another centrifugation for 1 minute was performed. The empty column tube was again centrifuged for

1 minute to remove residual Wash Buffer. Finally, the column was transferred to a microtube and the DNA was eluted by centrifugation with 50 µL of Elution Buffer for 1 minute.

#### 3.2.1.3 Plasmid transformation

50 µL of chemically competent cells, either TOP10 for cloning or BL21-pLysS for protein expression, were thawed on ice and 1-5 µL of purchased plasmid or ligation product was pipetted to the cells. The cells were kept on ice for 20 minutes, heat shocked at 42 °C for 1 minute, ice cooled for 2 minutes and finally, 125 µL S.O.C medium was added. The mixture was incubated and shaken at 37 °C, 250 rpm for 1 hour. Finally, the resulting solution could then be spread on LB agar plates or grown as a liquid culture explained in the next section 3.2.1.4.

#### 3.2.1.4 Bacteria cultivation – Overnight culture and petri dish culture

All bacteria cultivations were performed in LB medium containing the appropriate antibiotics, which was ampicillin at a final concentration of 50 µg $mL^{-1}$ for cloning, while all overexpression experiments were conducted in the presence of ampicillin and chloramphenicol at a final concentration of 50 µg $mL^{-1}$ and 34 µg $mL^{-1}$ respectively. Bacteria in S.O.C medium grown according to 3.2.1.3 could either be spread on LB agar petri dishes or directly added into liquid LB medium (overnight culture). The bacteria were then incubated at 37° C for 16 to 18 hours and additionally liquid cultures were constantly shaken at 250 rpm. Glycerol stocks could then be prepared by mixing 750 µL overnight culture and 250 µL sterile glycerol solution. They were stored at -80 °C.

#### 3.2.1.5 Plasmid isolation

Plasmids were isolated from overnight cultures of bacteria using GeneJET Plasmid Miniprep Kit according to its manual: 5 mL of overnight culture was centrifuged at 6800 g for 2 minutes in a microcentrifuge tube and the supernatant was removed. All subsequent centrifugation steps were performed at 13000 g. The cell pellet was resuspended in 250 µL Resuspension Solution. 250 µL of Lysis Solution was added and the tube was inverted several times until a clear, viscous solution was observed. After added 350 µL of Neutralisation solution and inverting multiple times, the solution was centrifuged to pellet the white precipitate. The supernatant was transferred to a spin column and centrifuged for 1 minutes. After adding 500 µL Wash Solution, the column was centrifuged for 1 minutes. This step was repeated one more time. An additional centrifugation of the empty tube was performed for 1 minute to remove any remaining Wash Solution. Finally, 50 µL Elution Buffer was added and the solution was

incubated for 2 minutes. The DNA within the Elution Buffer was then collected in a fresh microcentrifuge tube by centrifuging for 2 minutes.

#### 3.2.1.6 SDS-PAGE

For the sample preparation, 4 µL reducing agent, 10 µL sample buffer and 26 µL protein sample were combined. If the sample was a pellet, it was first suspended by pipetting after adding 26 µL $dH_2O$. Subsequently, all samples were heated to 70 °C for 10 minutes, pellet samples had to be centrifuged at 10000 g for 10 minutes and the supernatant was used for SDS-PAGE analysis. Then, 20 µL of each sample or 10 µL of the marker were loaded into the wells. The gels were run at 160 V for 45 minutes in 1x MES running buffer. They were then cooled, removed from the casting frames and rinsed with deionised water a few times to fully remove the running buffer. They were covered in 0.1 % CBB solution for 40 minutes. After removing the staining solution and rinsing them well with deionised water, they were left in destaining solution until the background became transparent.

### 3.2.2 Cloning

#### 3.2.2.1 Primers

IRS-7 fwd.: 5'-AAT CAT ATG GGA GAT GAG GAT AAA GTG ACG CTG-3'

IRS-7 rev.: 5'-ATT CTC GAG TTA GAG CTT GGT GAC CTG TC-3'

IRS-9 fwd.: 5'-ATA CAT ATG ATC CAC TTT CGT CGG AAA ACA GC-3'

IRS-9 rev.: 5'-AAT CTC GAG TCA CAC ATC CAT CGA AAC AAT ATG C-3'

IRS-10-11-12 fwd.: 5'-ACT CAT ATG CAA GAA GAA CTG AAG CTG ACC-3'

IRS-10-11-12 rev.: 5'-AAT CTC GAG CTA CAG TTT GTT GAC TTG ACC-3'

IRS-18 fwd.: 5'-AAT CAT ATG GAC CGT GAT GCC GTC AGC C-3'

IRS-18 rev.: 5'-AGT CTC GAG TCA CAC CTC TTG TAG CCT TC-3'

IRS-19 fwd.: 5'-AAT CAT ATG CAA ATG GAC GAC CGA CTG ACG TTG G-3'

IRS-19 rev.: 5'-AAT CTC GAG CTA GAG TGC GTT GAT ATG TCC CAC G-3'

IRS-20 fwd.: 5'-AAT CAT ATG GAG GAA GAG GAC AAA CTC ACC ACC-3'

IRS-20 rev.: 5'-AAT CTC GAG TCA GAG CTT GTT CAC TTG CC-3'

IRS-25 fwd.: 5'-AAT CAT ATG GAG CCT AAC GAG CTT GAG GAC C-3'

IRS-25 rev.: 5'-AAT CTC GAG TCA CAA GTC CAG CAC ACG G-3'

IRS-27 fwd.: 5'-AGT CAT ATG GGA GAA AAC AAG CTA ACT ACA GCG-3'

IRS-27 rev.: 5'-AAT CTC GAG TCA GAG TTT GTT CAC CTG CC-3'

#### 3.2.2.2 Gradient PCR

The first step after designing primers was a PCR with different annealing temperatures to investigate the optimal annealing temperatures to amplify genes of interest from tick cDNA. Sample preparation included combining 10 µL PCR Master Mix, 1 µL of each of the two 0.01 mM primers, 1 µL cDNA and 7 µL nuclease free water on ice to form a total volume of 20 µL. The thermocycler settings for the gradient PCR are shown in Table 2. Four different temperatures were used in step 3 to find the optimal annealing temperatures.

Table 2: Thermocycler settings for gradient PCR

| Step | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Temperature | 95 | 95 | 60 / 54.2 / 50.7 / 45 | 72 | 72 | 12 |
| Time | 3 min | 30 s | 30 s | 90 s | 5 min | - |
| Number of Cycles | 1 | 34 | | | 1 | - |

#### 3.2.2.3 High fidelity PCR

After confirming the optimal annealing temperatures for each pair of primers on an agarose gel, high fidelity polymerase was used in the next step to obtain an amplicon without base pairing errors which is suitable to be inserted into a vector. For this 4 µL Phusion Green HF Buffer, 1 µL of each of the two 0.01 mM Primers, 1 µL cDNA, 0.2 µL Phusion DNA Polymerase, 0.4 µL dNTP Mix and 12.4 µL nuclease free water were combined on ice. The used settings are shown in Table 3. The annealing temperature in step 3 was chosen individually for each pair of primers based on the gradient PCR and is one of the four shown options.

Table 3: Thermocycler settings for high fidelity PCR

| Step | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Temperature | 98 | 98 | 60 / 54.2 / 50.7 / 45 | 72 | 72 | 12 |
| Time | 30s | 10 s | 30 s | 60 s | 10 min | - |
| Number of Cycles | 1 | 34 | | | 1 | - |

#### 3.2.2.4 Digestion of PCR fragments and plasmids

Following gel extraction the amplicons as well as the IRS-2 chimera construct and the pET-17b vector were all digested with NdeI and XhoI. The procedure required the combination of all reaction components on ice according to Table 4 and an incubation at 37 °C. PCR fragments were incubated for 90 minutes, while plasmids were incubated for 15 minutes. Heat shock was performed at 65 °C for 5 minutes on PCR fragments and plasmid digests to

inactivate the restriction enzymes. Visualisation on an agarose gel and purification from the gel were the final steps.

Table 4: Reaction components of double digest

| Component | PCR fragment digest | Plasmid digest |
|---|---|---|
| FastDigest Green Buffer | 3 µL | 2 µL |
| NdeI | 1 µL | 1 µL |
| XhoI | 1 µL | 1 µL |
| DNA | 0.2 µg | 1 µg |
| Nuclease-free water | Up to 30 µL | Up to 30 µL |

#### 3.2.2.5 Ligation

All cut and purified inserts were then ligated into the pET-17b vector according to the following procedure. 4 µL 5x reaction buffer, 5 µL pET-17b vector (45 ng) and 10.8 µL (50 ng – 120 ng) insert and 0.2 µL T4 Ligase (5U $\mu L^{-1}$) were combined on ice and incubated at room temperature for 5 minutes. Ligation mixtures were then transformed into TOP10 cells following the procedure in 3.2.1.3 and grown according to 3.2.1.4.

#### 3.2.2.6 Colony PCR

To confirm the presence of plasmids containing the desired amplicon, single colonies from the agar plates were chosen. DNA from these bacterial colonies was picked up with a pipette tip and added into a mixture of 10 µL PCR Master Mix, 1 µL of each of the two 0.01mM T7 primers and 8 µL nuclease free water on ice. The thermocycler settings for the gradient PCR are displayed in Table 5. After confirming the presence of the plasmid with the insert on an agarose gel, bacteria from the specific colonies were grown as an overnight culture, the plasmid was isolated and sent for sequencing.

Table 5: Thermocycler settings for colony PCR.

| Step | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Temperature | 95 | 95 | 50 | 72 | 72 | 12 |
| Time | 5 min | 30 s | 30 s | 90 s | 5 min | - |
| Number of Cycles | 1 | 34 | | | 1 | - |

### 3.2.3 Protein production

After sequencing results confirmed the presence of the amplicon within the plasmid, overnight cultures from the corresponding colonies on the agar plate were grown. The plasmid was isolated and transformed into different *E. coli* cell strains, BL21-pLysS, for protein expression, according to 3.2.1.3. For the overexpression in BL21-pLysS cells containing the plasmid with

the desired genes were grown in all LB medium containing Ampicillin (50 µg $mL^{-1}$) and Chloramphenicol (34 µg $mL^{-1}$).

### 3.2.3.1 Small scale pilot expression

Small scale pilot expressions were performed to asses, if the bacteria express the protein, after what time the expression maximum is reached and in what quantities the protein is expressed. For this 5 mL overnight cultures were grown and then transferred to 50 mL media containing the appropriate antibiotics. They were incubated at 37 °C, 200 rpm until an OD of 0.6 – 0.8 was reached. Then, IPTG was added to a final concentration of 1 mM. Directly after the induction 1 mL of the solution was removed and centrifuged at 10000 g for 2 minutes. The supernatant was removed, and the pellet was frozen. This was repeated at different time points for up to 24 hours.

If little or no overexpression was observed, the conditions were modified by first altering IPTG concentrations and the OD at which IPTG induction was started. Further optimisations include adding glucose or ethanol in addition to all other components of the growth medium according to (Chhetri G. et al, 2015 and Novagen, n.d.). The details about specific medium composition are shown in Table 6.

Table 6: List of substances in addition to the basic LB medium for optimised overexpression.

| | Overnight culture | Growth before induction |
|---|---|---|
| Condition 1 (EtOH) | | 3 % EtOH (v/v%) |
| Condition 2 (EtOH + Glu) | 1 % glucose | 1 % glucose + 3 % EtOH (v/v%) |
| Condition 3 (Glu) | 1 % glucose | 1 % glucose |

For the conditions in Table 6 the procedure was also modified slightly, which involved cells being grown to an OD of 0.8 – 1.1. Additionally, after reaching the right OD, cells in glucose containing medium were spun down at 500g for 5 minutes and the supernatant was discarded. The cell pellet was then resuspended in the same solution, without glucose. The addition of IPTG and sample taking at time points followed the same procedure as explained above.

### 3.2.3.2 Small scale isolation of inclusion bodies

After the pilot expression, the presence of protein had to be verified on an SDS-PAGE. Before this, inclusion bodies (IB), which most likely formed and contained the protein had to be isolated. For IB isolation the cell pellets were dissolved by vortexing and heating to 55 °C in 500 µL 20 mM Tris-HCl, pH = 8. After each pellet was fully dissolved, the test tubes were immersed into liquid nitrogen until fully frozen and then reheated to 55 °C. This step was carried out 4 times. The solutions were then centrifuged at 10000 g for 10 minutes and the

supernatant and the pellet (containing the inclusion bodies with the recombinant protein) were separated. The pellets were visualised on an SDS-PAGE according to the procedure in point 3.2.1.6.

#### 3.2.3.3 Big scale protein expression

Once optimal conditions for protein expression had been assessed, the method was scaled up to 8 L of LB medium with appropriate antibiotics. The medium was distributed to 8 baffled flasks with filter caps and 50 mL overnight culture per flask was added. The flasks were shaken at 100 rpm, 37 °C until an optical density of 0.6 to 0.8 was reached. Then, IPTG was added to a final concentration of 1 mM and the cultures were incubated for the optimised time period. Then they were centrifuged at 2000 g for 10 minutes at 4 °C. The supernatant was discarded, and the pellet could be frozen, or inclusion bodies could be isolated immediately.

#### 3.2.3.4 Big scale isolation of inclusion bodies

The cell pellets were resuspended in 500 mL Tris-HCl and the solution was stirred for one hour. Then the suspension was sonicated for 2 minutes, followed by centrifugation at 10000 g, 4 °C for 10 minutes. The supernatant, the cytosolic fraction, was removed and the pellet was again resuspended in 500 mL Tris-HCl and 1 % Triton and stirred for 1 hour. After another sonication and centrifugation at the before mentioned settings, the membrane fraction (supernatant) and a pellet with the desired inclusion bodies were obtained. Each of the fractions was visualised on an SDS-PAGE to verify the presence of the protein in the inclusion bodies once again.

### 3.2.4 Refolding

#### 3.2.4.1 Small scale refolding

Before further processing of the IB, they were washed four times with 20 mL of Tris-HCl to remove any traces of Triton. Up to 0.16 g of wet inclusion bodies were dissolved in 2000 µL 6 M Guanidine solution and 20 µL 1 M DTT and rotated for 1 hour until the pellet had dissolved. After centrifugation at 10000 g for 10 minutes, 200 µL of the solution was added quickly to 40 mL of each of the eight refolding buffers and immediately inverted multiple times. All refolding solutions were put on a shaker for 3 hours with open lids and afterwards moved to the fridge overnight. The samples were then centrifuged at 10000 g, 4 °C for 10 minutes, to deposit all incorrectly folded and aggregated protein at the bottom of the test tubes. The supernatant with the refolded proteins was concentrated to about 700 µL – 1000 µL using 10 kDa Amicon filter units. Comparison of refolding efficiency of the eight refolding buffers

was assessed by running all samples on an SDS-PAGE and performing an estimation of the protein concentration explained in section 3.2.4.2 below.

#### 3.2.4.2 Protein concentration estimation – Bicinchoninic acid assay (BCA assay)

Protein concentration was estimated using the Pierce™ BCA Protein Assay Kit according to its microplate protocol. BSA solutions of different concentrations (0, 25, 125, 250, 500, 750, 1000, 1500, 2000 ug $mL^{-1}$) were used as standards. 25 µL of all standards and triplicates of the samples were pipetted into the wells of a transparent 96-well microplate, 200 µL working reagent was added, the plate was shaken for 30 seconds and incubated at 37 °C for 30 minutes. After cooling to room temperature, the absorbance at 562 nm was measured with a Tecan Infinite M200 96-well plate reader. Regression analysis was then performed on the measured values from the external calibration, which was used to calculated protein concentrations of the samples.

#### 3.2.4.3 Big scale refolding

Using SDS-PAGE and BCA assay results, the optimal refolding conditions were determined and with these conditions, big scale refolding could be performed. After washing the IB with Tris-HCl, pH = 8 between 500 mg and 700 mg of wet IB were dissolved in 25 mL 6 M Guanidine solution and 250 µL DTT solution. The tubes were rotated for 2 hours until fully dissolved. After centrifugation at 10000 g for 10 minutes the supernatant was poured into 4 L of the selected refolding buffer under constant, vigorous stirring. The solution was kept stirring for at least 3 hours and sometimes precipitates in the whirl were removed. After this the solution was stored in the fridge overnight.

#### 3.2.4.4 Concentrating of protein solutions

Solutions of refolded proteins were vacuum filtered with Vacuum Filtration "rapid"-Filtermax. Then the solutions were concentrated to approximately 50 mL at 4 °C using Amicon 8400 stirred cell and 10 kDa ultrafiltration disks. For further concentration of volumes of 50 mL or less, 10 kDa amicon filter units were used to obtain final volumes of 0.5 – 2 mL. Protein concentration was then again estimated with the BCA assay explained in 3.2.4.2.

### 3.2.5 Protein purification and testing

#### 3.2.5.1 Size exclusion chromatography (SEC)

Size exclusion chromatography was performed with a HiPrep 16/60 Sephacryl S-100 HR column after centrifugation for 10 minutes at 10000 g. The fractions were eluted using 20 mM Tris-HCl pH 8.5 with a flow rate of 0.5 mL $min^{-1}$ in 1 mL steps per fraction. Each fraction

showing UV absorbance at 280 nm was then run on an SDS-PAGE to confirm the presence of the desired protein.

#### 3.2.5.2 Ion exchange chromatography (IEC)

The fractions containing the desired protein as determined by SDS-PAGE were further purified using ion exchange chromatography with a Superdex 75 10/300 GL column. The fractions were eluted with buffer A: 20 mM Tris-HCl, pH = 8.5, and buffer B: 20 mM Tris-HCl, 1 M NaCl, pH = 8.5, by increasing the amount of buffer B gradually and thus increasing the ionic strength. Fractions were collected in 1 mL steps. Each fraction showing UV absorbance at 280 nm was then run on an SDS-PAGE to confirm the presence of the desired protein.

#### 3.2.5.3 Protease inhibitor assays

All assays were performed in triplicates with a no-inhibitor control, which was assumed to show 100 % protease activity. The total reaction volume was 50 µL. It included 1 µL enzyme with and 4 – 6 µL of inhibitor at final concentrations shown in Table 7. The wells were filled up with assay buffer to a final volume of 47.5 µL. After an incubation period of 10 minutes, 2.5 µL of substrate was added. A Texan Infinite M200 96-well plate fluorescence reader with 365 nm excitation and 450 nm emission wavelength and a cut off at 535 nm was used to follow the inhibition of the enzyme at 30 °C for 20 minutes. Statistical analysis involved a two-sided, heteroscedastic t-test.

Table 7: Detailed reaction conditions for each serine protease.

| | c enzyme in solution | Buffer | c inhibitor in solution | substrate |
|---|---|---|---|---|
| Trypsin | 20 pM | Trypsin AB | 2 µM | Boc-Val-Pro-Arg-AMC hydrochloride salt |
| α-Chymotrypsin | 30 pM | Trypsin AB | 2 µM | Chymotrypsin Substrate II |
| Chymase | 2 nM | Elastase AB | 2 µM | Suc-Leu-Leu-Val-Tyr-AMC |
| Cathepsin G | 15 nM | Elastase AB | 2 µM | Chymotrypsin Substrate II |
| fXa | 60 pM | Trypsin AB | 2 µM | Spectrofluor™ FXa |
| fXIIa | 1.2 nM | Trypsin AB | 2 µM | Spectrofluor™ FXa |
| Plasmin | 222 pM | Kallikrein AB | 2 µM | Boc-β-benzyl-Asp-Pro-Arg-7-amido-4-methylcoumarin hydrochloride |

# 4 Results

## 4.1 Generation of serpin gene amplicons

After designing primers, the first step was a gradient PCR according to 3.2.2.2 with four different annealing temperatures to find the optimal conditions for each pair of primers to amplify genes of interest from tick cDNA. The resulting products were then visualised on an agarose gel. All pairs of primers, except the ones for IRS-19, generated an amplicon at the expected size of approximately 1200 bp as shown in Figure 3.

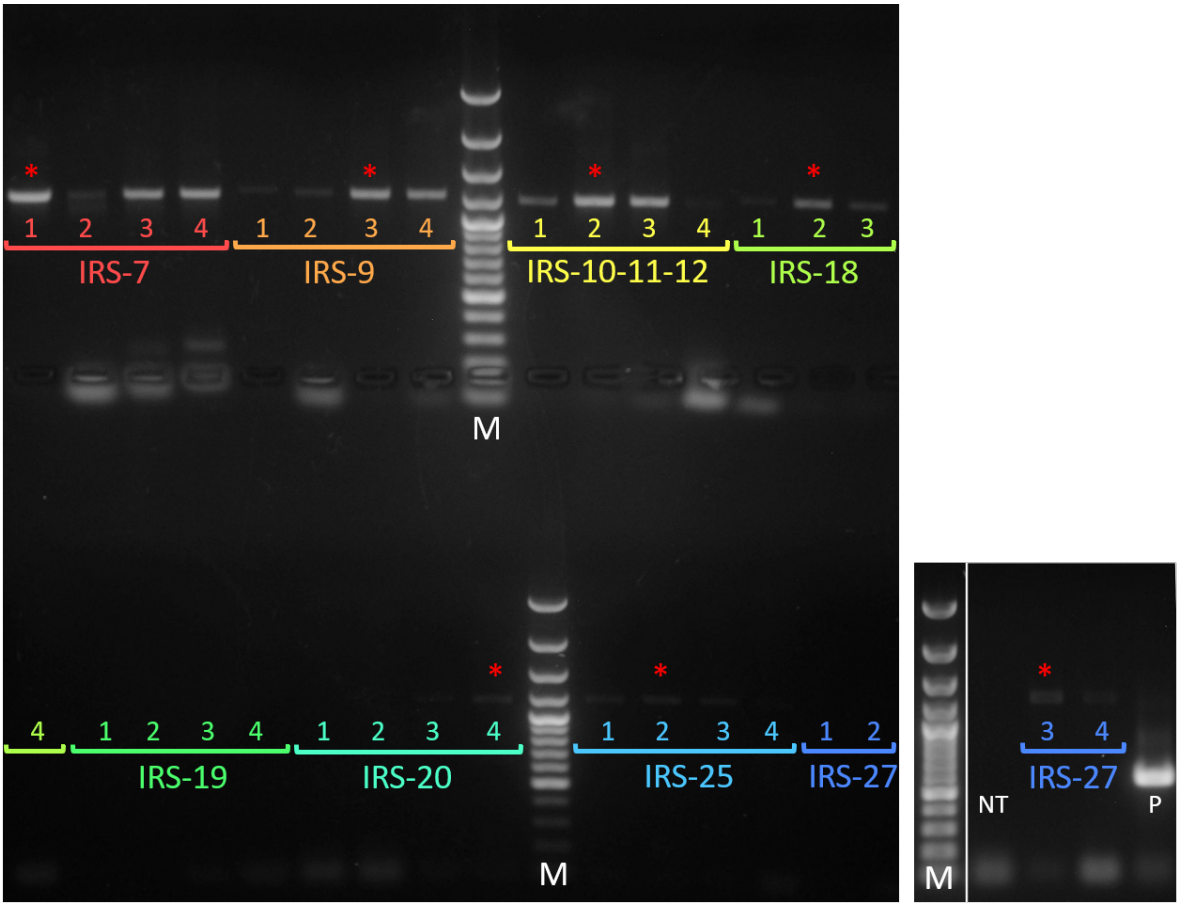


Figure 3: Gradient PCR results; the best annealing temperature for each pair of primers among different annealing temperature 60°C (1), 54.2°C (2), 50.7°C (3) and 45°C (4) is marked with an asterisk. M - molecular weight marker, NT - no template control, P - positive control (tick actin).

Subsequently, the appropriate annealing temperature for each pair of primers as indicated in Figure 3 was chosen and high-fidelity polymerase was then used to replicate the amplicon for further cloning. The gel showing the outcome of this PCR is presented in Figure 4. The primers for IRS-7, -9, -10-11-12 and -18 generated a well visible band, while IRS-20 and -25 only show very weak bands and no detectable amount of the amplicon was generated in the case of IRS-27. All visible bands at approximately 1200 bp were cut from the gel and purified.

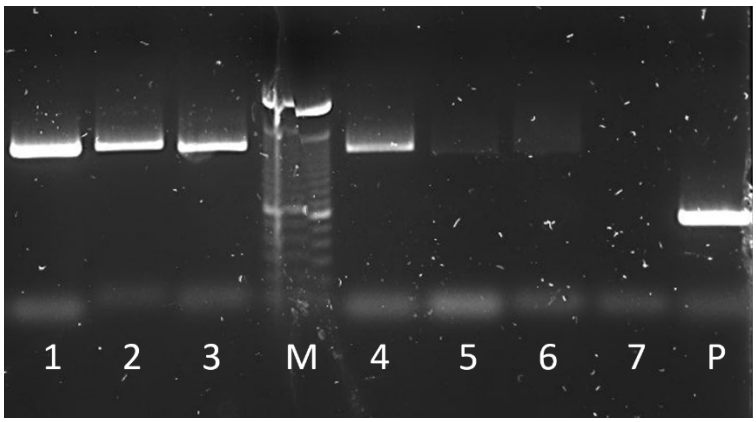


Figure 4: High fidelity PCR results; 1 – IRS-7, 2 – IRS-9, 3 – IRS-10-11-12, 4 – IRS-18, 5 – IRS-20, 6 – IRS-25, 7 – IRS-27, M – molecular weight marker P – positive control (tick actin)

Next, the pET-17b plasmid, into which the amplicons were later ligated, was transformed into TOP10 cells and then purified from the overnight culture to increase the amount of starting material. The same was also done with the plasmid containing the chimeric IRS-2 sequence. The two plasmids and the amplicons from the high-fidelity PCR were cut with NdeI and XhoI, and visualised on an agarose gel, which is shown in Figure 5. The amplicons of the primers for IRS-7, -9 and -10-11-12 showed prominent bands at the expected length of 1200 bp, while the IRS-20 amplicon only showed a very weak band. Moreover, the gel displays that IRS-18 and -25 amplicons were digested at more than one site. The chimeric serpin and pET-17b also showed bright brands at 1200 bp and 3300 bp, respectively. All fragments marked with an asterisk were cut from the gel and purified.

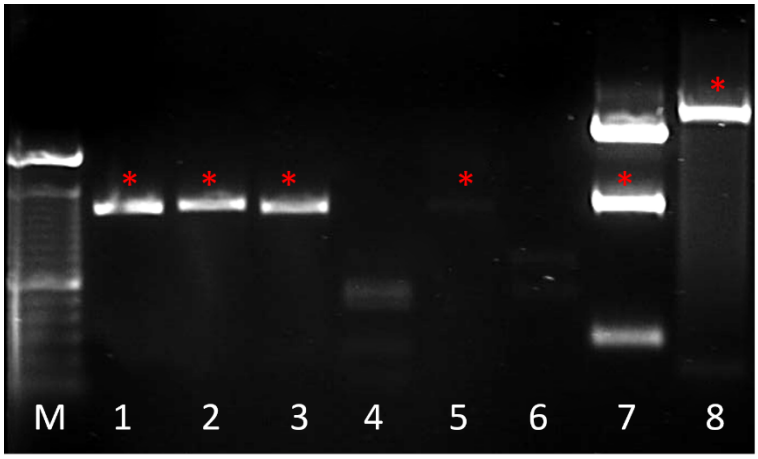


Figure 5: Enzyme restriction; bands with asterisks selected for further ligation; 1 – IRS-7, 2 – IRS-9, 3 – IRS-10-11-12, 4 – IRS-18, 5 – IRS-20, 6 – IRS-25, 7 – IRS-2 chimera, 8 – pET-17b, M – molecular weight marker

## 4.2 Ligation and cloning

After ligating each insert into the pET-17b vector and subsequent transformation into TOP10 cell, cultures were grown on petri dishes and finally colony PCR was performed to determine the success of ligation and transformation. Pictures of the agarose gel in Figure 6 indicate that a positive clone, containing the plasmid with the insert between 1200 and 1500 bp was obtained for IRS-7, -9, -10-11-12 and the chimeric serpin, while no positive colony was found for IRS-20 (not shown). Bands at 300 bp indicate a negative colony without insert. Positive colonies were grown in an overnight culture, purified and sent for sequencing to confirm the right DNA sequence.

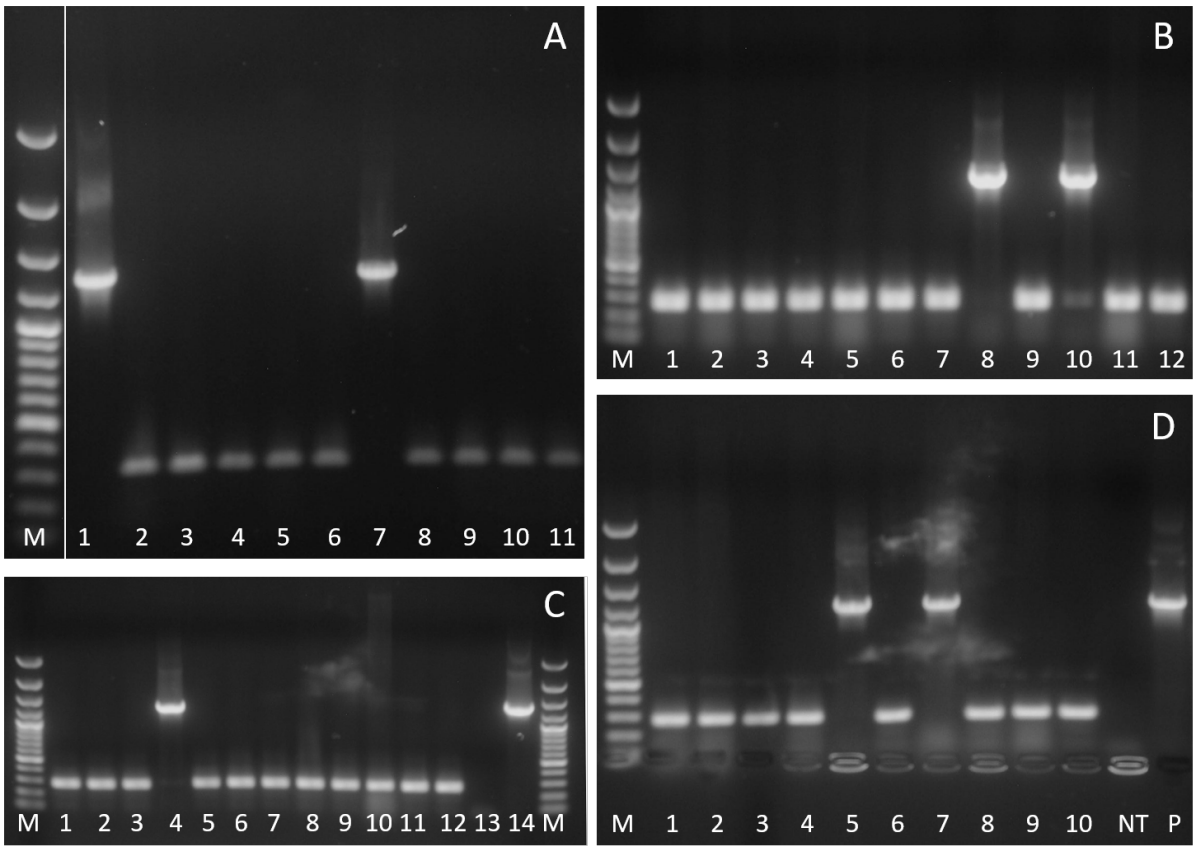


Figure 6: Colony PCR of IRS-9 (A), IRS-2 chimera(B), IRS-10-11-12 (C) and IRS-7 (D); M – molecular weight marker, NT – no template control, P – positive control (already sequenced pET-17b plasmid containing IRS-4)

### 4.2.1 Sequencing results

The nucleotide sequences were translated, aligned with ClustalW and compared to the original sequences the primers were designed for. The resulting alignments can be viewed in Figure 7.

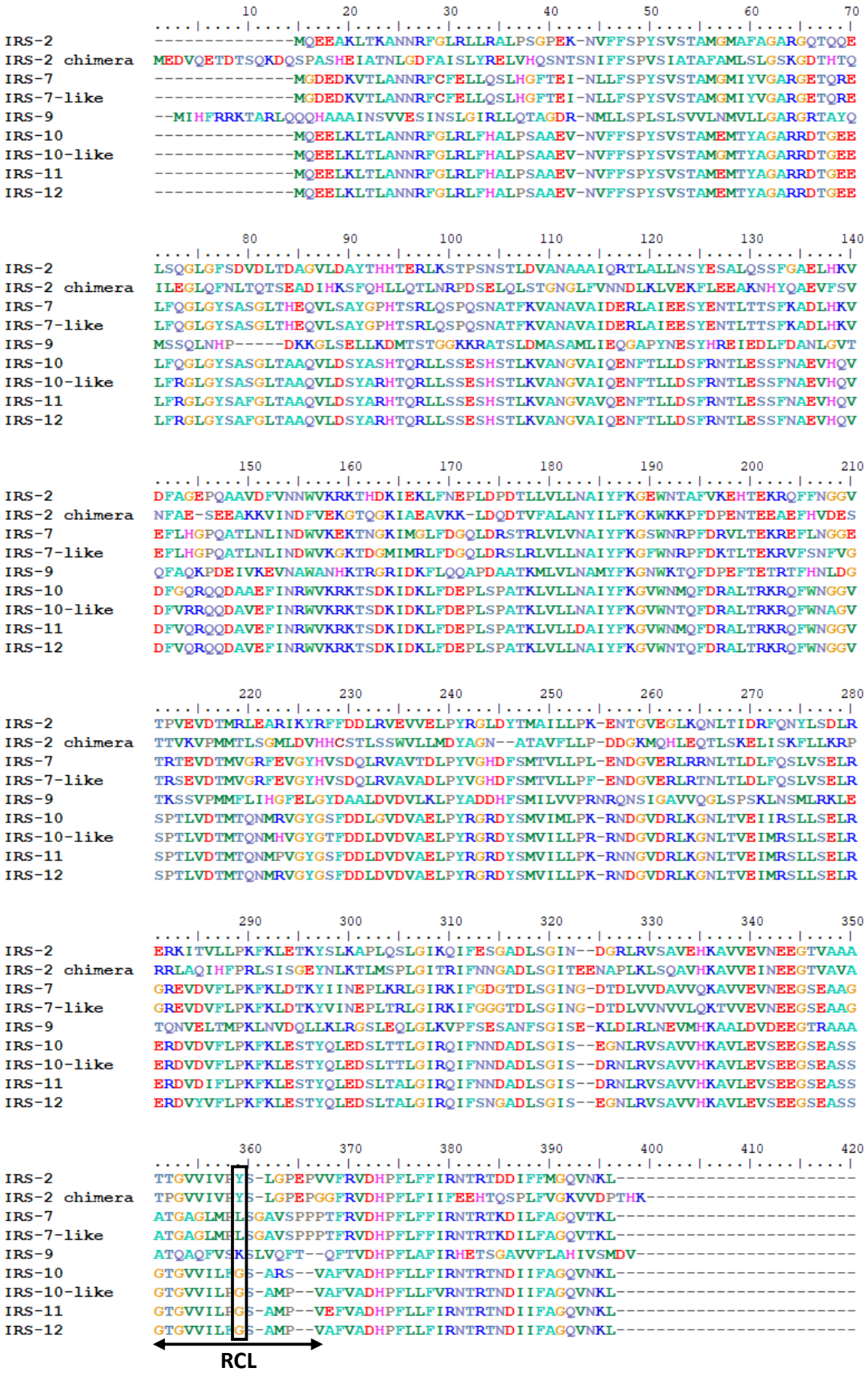


Figure 7: Aligned sequencing results; RCL and P1 position are indicated; IRS-2 chimera, IRS-7-like, IRS-9 and IRS-10-like obtained from cloning compared to other serpins.

The sequencing data for the chimeric serpin and IRS-9 completely matches the template sequences. However, the other two serpins don't match it completely, which is why they will be referred to as IRS-7-like and IRS-10-like from now on.

Generally, different *I. ricinus* serpins such as IRS-1, IRS-2, IRS-4 mentioned in (Chmelar et al., 2011) show 40 – 60 % similarity. Compared to that, IRS-10, IRS-11, IRS-12 and IRS-10-like are very similar and have between 95 – 97 % identical amino acids. The same holds for IRS-7 and IRS-7-like showing 93 % sequence similarity. In contrast, IRS-9 only shows 24 – 28 % sequence similarity to most other *I. ricinus* serpins. However, a close homologue with 97 % sequence similarity was found in another tick species, *I. scapularis* with the NCBI accession number XP_002433376.1. IRS-7-like and IRS-9 have Leucine and Lysine at their P1 position, respectively.

## 4.3 Pilot expression and optimisation

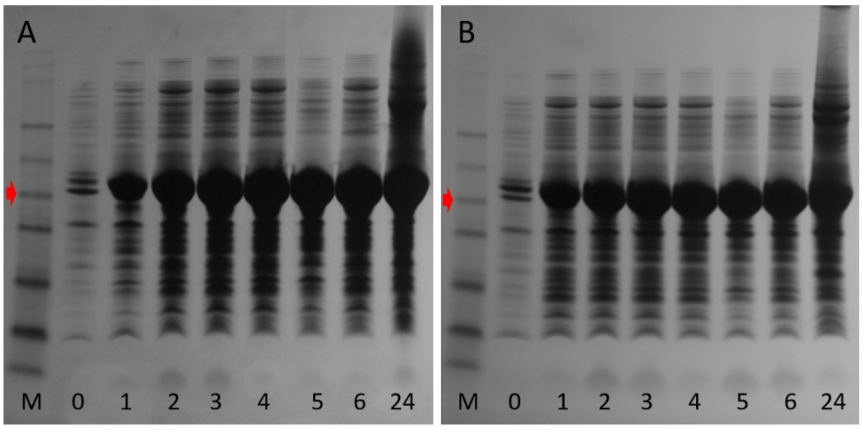


Figure 8: Pilot expression of IRS-7-like (A), IRS-9 (B); M – molecular weight marker; the numbering corresponds to time points after IPTG induction

After sequencing, the plasmids were transformed into BL21-pLysS cells. Small scale overexpression was performed according to 3.2.3.1. IRS-7-like and IRS-9 were induced with IPTG at an OD of 0.6, while IRS-2 chimera and IRS-10-like had an OD of 0.8 at induction. Samples were taken at time t = 0, 1, 2, 3, 4, 5, 6, 24 hours after induction. Inclusion bodies obtained according to 3.2.3.2 were visualised on an SDS-PAGE presented in Figure 8. SDS-PAGE for IRS-10-like and chimeric IRS-2 are not shown since no significant amount of protein was produced, posing the need for further expression optimisation. In Figure 8 A and B a significant amount of protein produced from IRS-7-like and -9 can be seen around the 38 kDa weight marker indicated with a red arrow. The obtained results correlate with the expected molecular weight of the proteins of about 42 kDa. The experiment also showed that for both, IRS-7-like and -9 the amount of protein produced beyond 4 hours after induction does not

increase significantly. Thus, a 4 hour time span was chosen as the optimum for large scale overexpression of IRS-7-like and -9 presented in section 4.4.

As the level of induction for IRS-2 chimera and IRS-10-like were found to be insufficient further optimisations were performed. The first optimisation was a change in IPTG concentration and OD for induction, which did not improve the expression of the two serpins. Next, another optimisation by adding glucose and / or ethanol to the medium was performed according to 3.2.3.1. Overexpression in medium containing glucose was induced at OD 1.05, for ethanol at 0.85 and for glucose with ethanol at 0.94. Samples were taken at time points 0, 2, 4 and the results can be seen in Figure 9. Information from Figure 9 indicates that ethanol alone does not lead to a significant overexpression. By adding glucose to the medium a higher expression could be achieved for both serpins. Adding both glucose and ethanol gave the best results with an increased protein expression in comparison to before and a lower background compared to glucose alone.

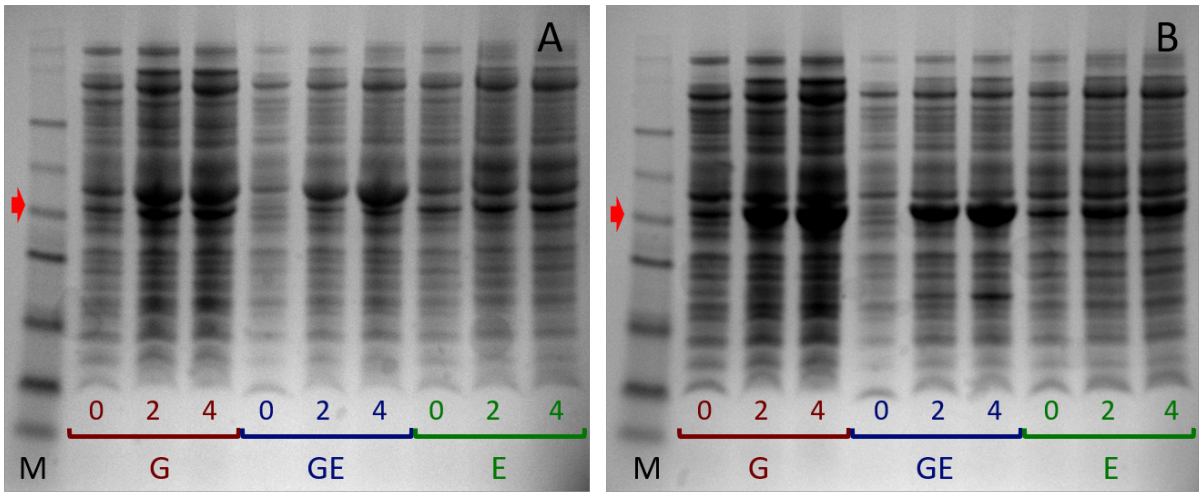


Figure 9: Overexpression in glucose- (G), glucose-and-ethanol- (GE) and ethanol-containing (E) medium of IRS-2 chimera (A) and IRS-10-like (B); M – molecular weight marker; numbers correlate to elapsed time after IPTG induction.

## 4.4 Large scale overexpression

All further steps in 3.2.3.3 were only performed with IRS-7-like and IRS-9, since the optimisation of expression for the other serpins is still incomplete. Bacterial cultures of the scale up grew slowly, taking one or two days to reach an OD of 0.6 at which overexpression was induced. Inclusion body isolation was then performed according to 3.2.3.4. All three fractions, cytosolic fraction, membrane fraction and Inclusion bodies, were run on an SDS-PAGE, shown in Figure 10, which confirms the presence of a large amount of desired protein in the IB, while nearly no protein was found in the other fractions.

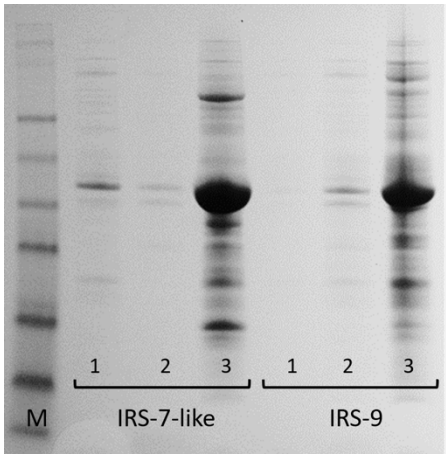


Figure 10: SDS-PAGE of protein in the 3 fractions obtained from sonication of IRS-7-like and IRS-9; M – molecular weight marker, 1 – cytosolic fraction, 2 – membrane fraction, 3 – IB

## 4.5 Refolding

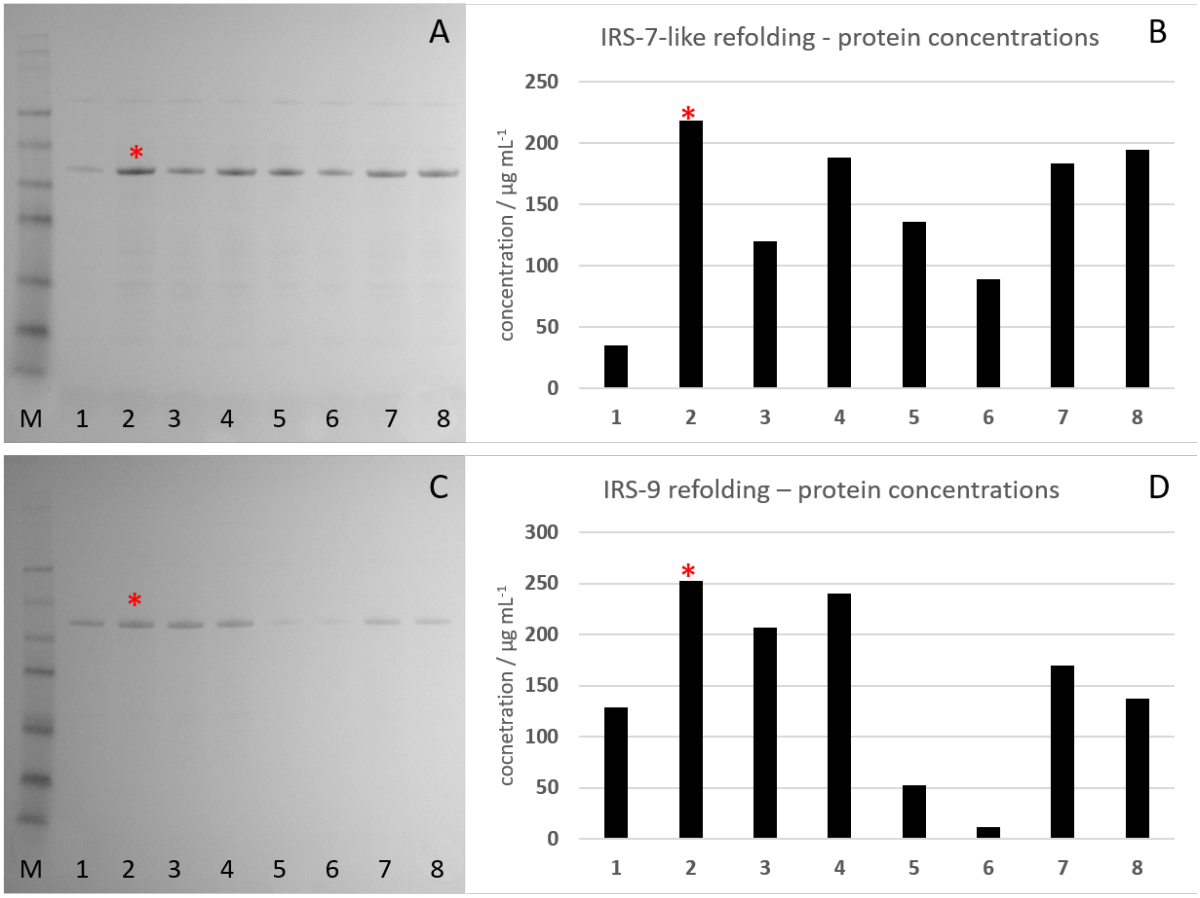


Figure 11: Refolding evaluation of IRS-7-like (A, B) and IRS-9 (C, D) with SDS-PAGE on the left and concentrations measured with BCA assay on the right; numbers indicate the refolding buffer; M – marker; asterisk point out the optimal refolding buffer

After washing the IB with Tris-HCl, pH = 8, 0.1495 g of wet IRS-7-like IB and 0.1611g of wet IRS-9 IB were used according to 3.2.4.1 to test refolding efficiency in 8 different refolding buffers. Incorrectly folded proteins forming aggregates were removed by centrifugation, the supernatant, containing mostly correctly folded and therefore soluble

protein, was concentrated to 700 – 1000 µL, and then filled up to 1 mL with Tris-HCl, pH = 8. For determining the best refolding conditions, SDS-PAGE and results from a BCA assay were evaluated.

The results for IRS-7-like shown in Figure 11 A and B indicated that buffer 2 provides the best conditions for refolding since SDS-PAGE as well as the concentration measurements clearly indicate the highest protein concentration of 218 µg $mL^{-1}$ in this buffer. In the case of IRS-9 the results highlight that the highest protein concentration is also achieved in buffer 2 with a final concentration of 252 µg $mL^{-1}$.

For big scale refolding, approximately 700 mg and 500 mg of wet IB of IRS-7-like and IRS-9, respectively, were refolded in refolding buffer 2 according to the procedure in 3.2.4.3. Between 1 and 2 mL of protein solution of IRS-7-like and IRS-9 was obtained after concentrating the refolding solutions.

## 4.6 Purification

To purify serpins IRS-7-like and IRS-9, size exclusion chromatography was performed according to 3.2.5.1. All fractions that showed detectable UV absorbance at 280 nm were run on an SDS-PAGE gel to confirm the presence of the proteins. Fractions containing the protein were further purified with ion exchange chromatography to 3.2.5.2. and subsequent SDS-PAGE was again used to reveal the presence of the protein.

### 4.6.1 IRS-7-like

A chromatogram of the SEC of IRS-7-like is presented in Figure 12. Two overlapping peaks between fraction 6A9 and 6A12 can be seen. Fractions 6A9 – 6A10 and 6A11 – 6A12 were pooled. Another small peak at 6B8 was collected, which did not contain any protein. Subsequent analysis of the SDS-PAGE shown in Figure 14 revealed that the majority of the protein was contained in fraction 2 and only a small amount in fraction 1. Regardless, both, fraction 1 and 2, were further purified with IEC. Each IEC run produced three peaks, which were named fraction A, B, C for SEC fraction 1 and D, E, F for SEC fraction 2. Only the chromatogram for SEC fraction 2 is shown in Figure 13, because SDS-PAGE of fractions A, B and C only showed very weak or no bands. Two big peaks between 6D7 – 6E2 and 6F4 – 6G1 as well as a smaller peak between 6E6 - 6E12 can be observed. The SDS-PAGE in Figure 14 confirmed that most of the protein was contained in fraction D, while all other fractions only contained negligible amounts of the IRS-7-like.

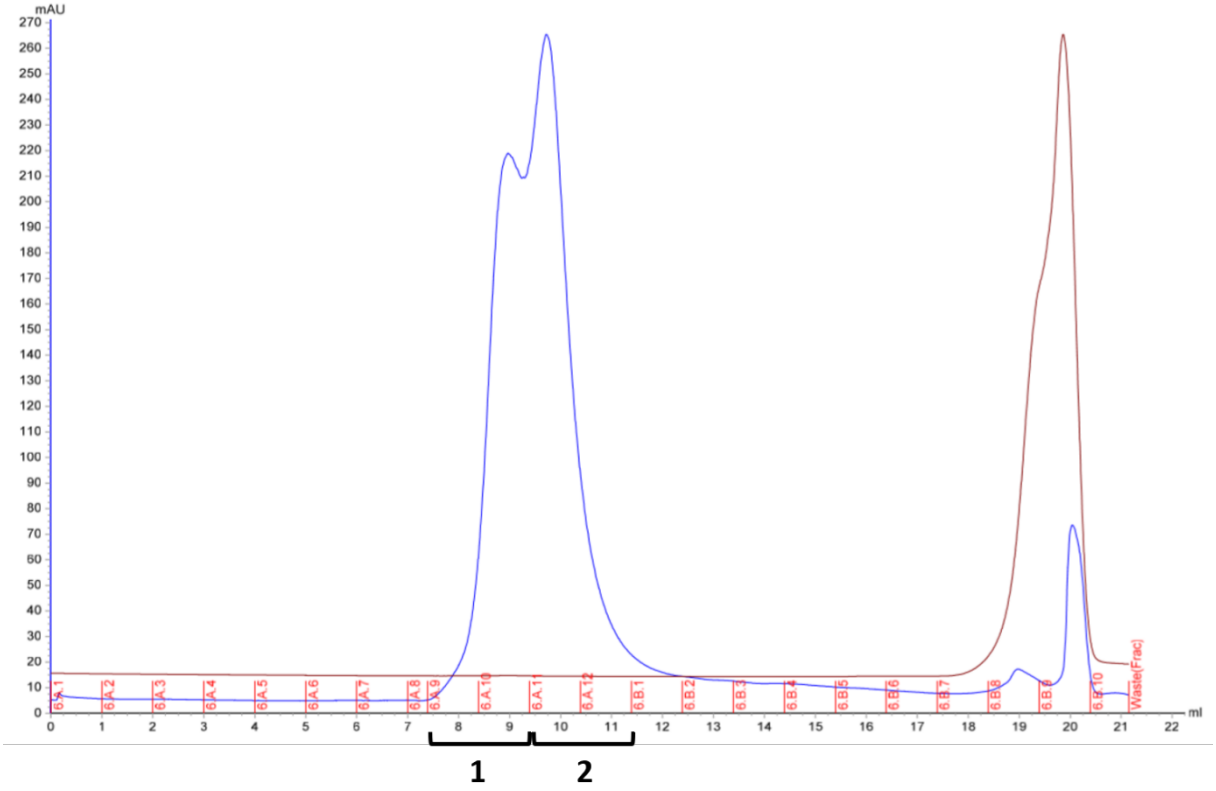


Figure 12: IRS-7-like SEC Chromatogram; blue line – UV absorbance in mAU, brown line – conductivity; 1 and 2 pooled fractions for further analysis.

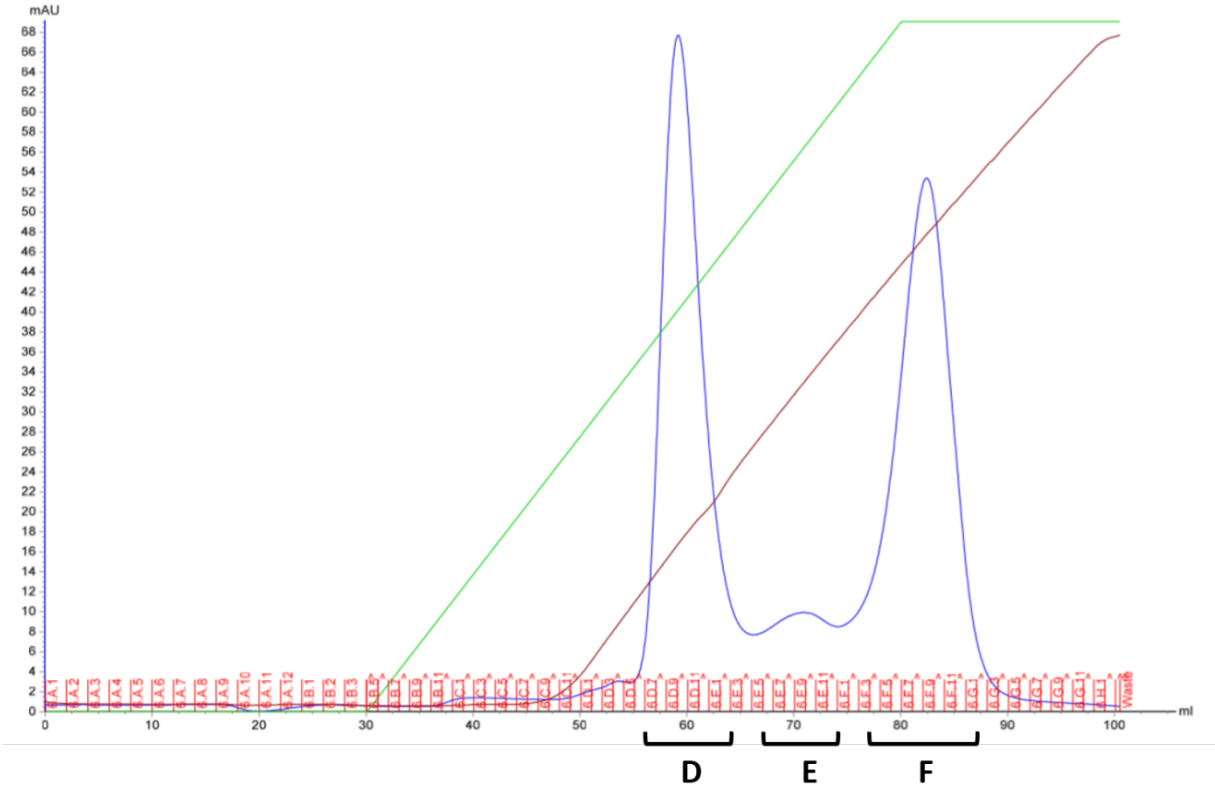


Figure 13: IRS-7-like IEC Chromatogram; blue line – UV absorbance in mAU, brown line – conductivity; green line – NaCl gradient, D, E, F – pooled fractions.

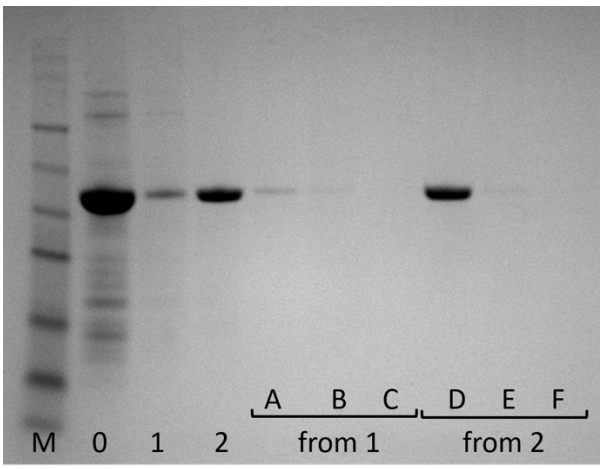


Figure 14: SDS-PAGE of all purification steps of IRS-7-like; M – molecular weight marker; 0 – pre SEC, 1 – SEC fraction 1, 2 – SEC fraction 2 (see Figure 12); A, B, C – three fractions of SEC fraction 1 after IEC; D, E, F – three fractions of SEC fraction 2 after IEC (see Figure 13)

In the SDS-PAGE in lane 0, which is the protein solution before purification, other proteins with different weights were also present in the solution. After purification with SEC, most of the protein was found in fraction 2, and a lower amount and some contamination was found in solution 1. Further purification of fraction 2 with IEC resulted in a protein solution (D) with a clear band at around 42 kDa without any other visible proteins in the SDS-PAGE. This solution was then concentrated to a final concentration of 1.2 mg $mL^{-1}$.

### 4.6.2 IRS-9

For IRS-9 the same steps were performed. The Chromatogram of SEC is shown in Figure 15. It indicates one big peak between 5A10 – 5B1 with a very small peak at fraction 5A9 right in front of it and two small peaks from 5B8 to 5B9 which did not contain the serpin. The SDS-PAGE in Figure 17 reveals that the target serpin, as well as other proteins are present in fraction 2, but in much lower concentration when compared to the SEC of IRS-7-like. Since detectable amounts of protein were only found in fraction 2, only this fraction was further purified with IEC. Figure 16 presents the chromatogram of IEC. It shows a few overlapping peaks between fraction 4F9 and 4H4, which have been divided into four fractions. SDS-PAGE reveals the presence of the target serpin as well as some residual contamination in fractions A, B and C, indicating incomplete purification.

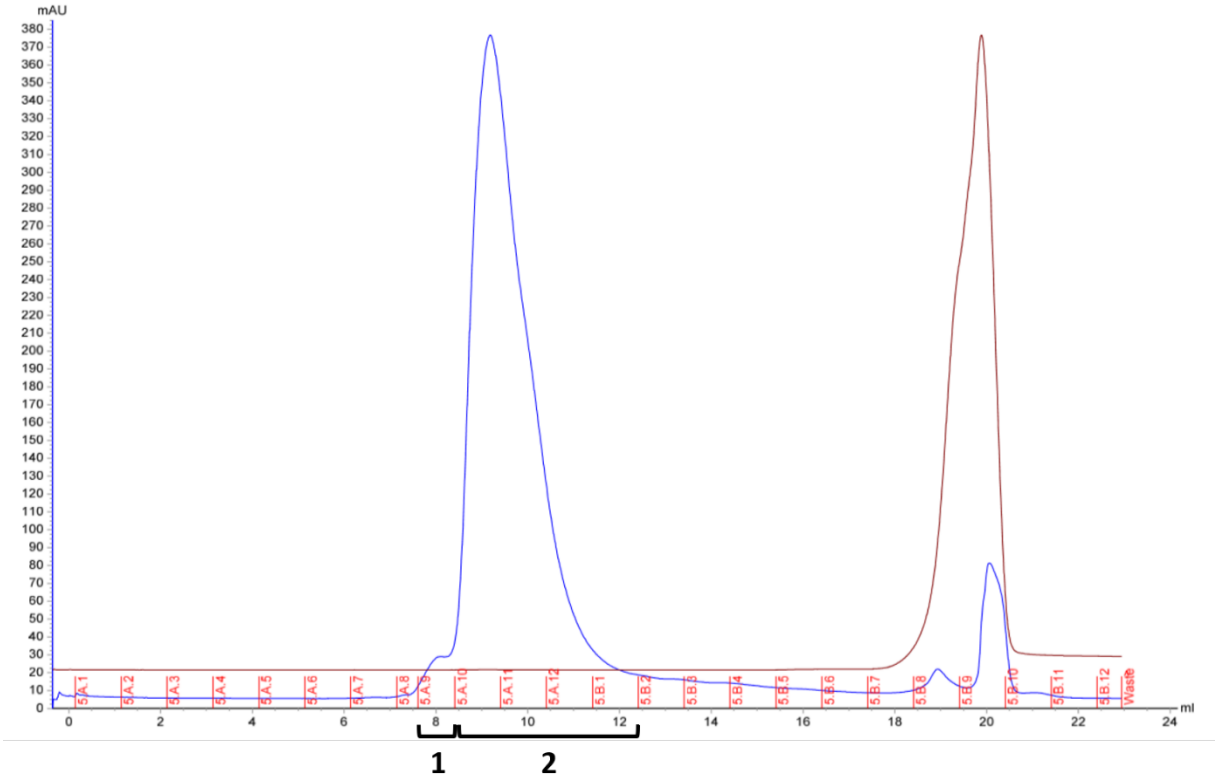


Figure 15: IRS-9 SEC Chromatogram; blue line – UV absorbance in mAU, brown line – conductivity; 1 and 2 pooled fractions for further analysis

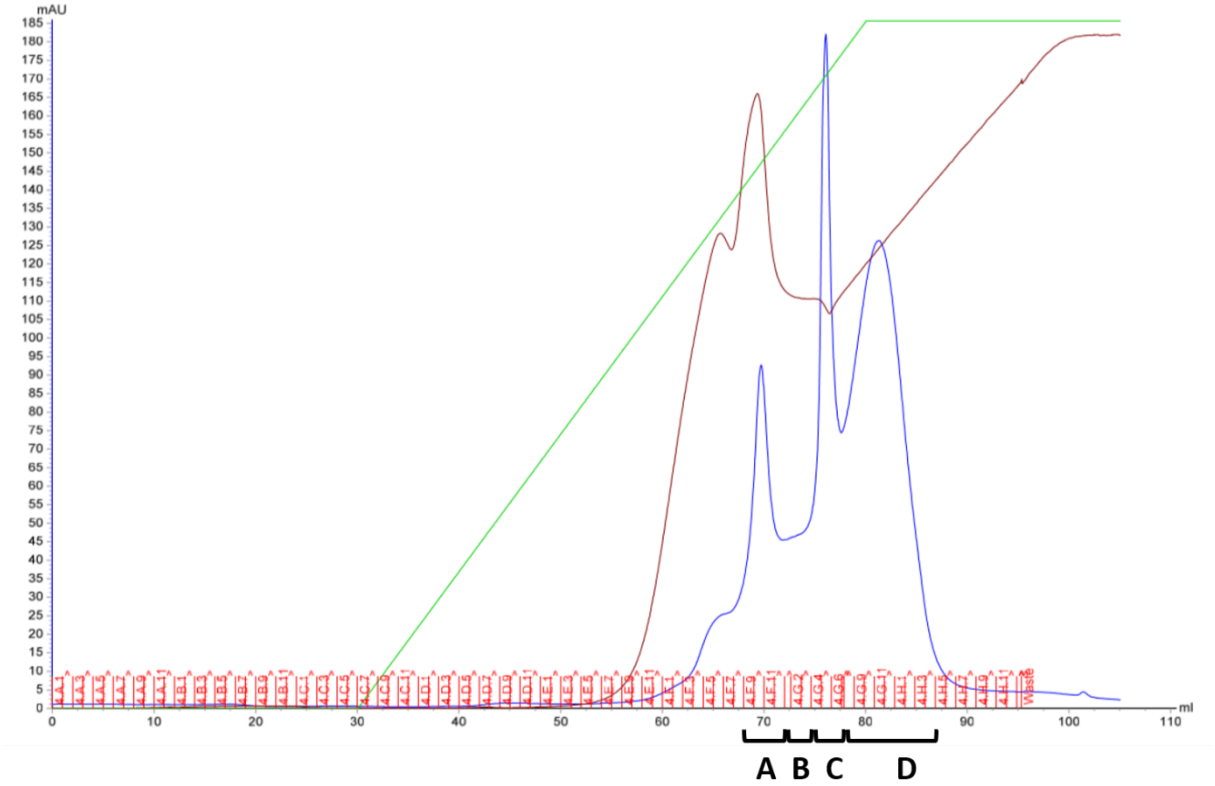


Figure 16: IRS-9 IEC Chromatogram; blue line – UV absorbance in mAU, brown line – conductivity; green line – NaCl gradient, A, B, C, D – pooled fractions.

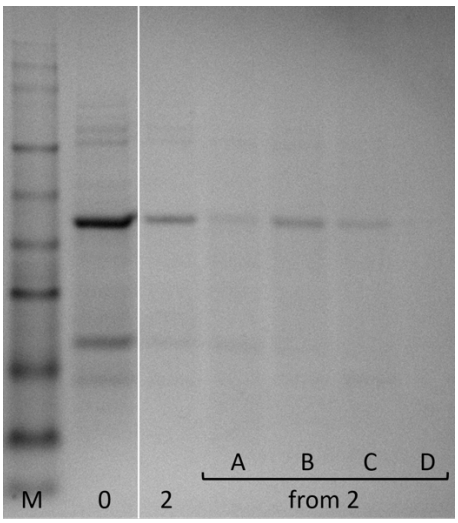


Figure 17: SDS-PAGE of all purification steps of IRS-9; M – molecular weight marker; 0 – pre-SEC, 2 – SEC fraction 2 (see Figure 15); A, B, C, D – four fractions of SEC fraction 2 after IEC

As with IRS-7-like the SDS-PAGE of IRS-9 clearly shows a band at around 42 kDa and proteins with other molecular weights before any purification (0). However, in contrast to the purification of IRS-7-like, there possibly still is some contamination from other proteins in the fractions from IEC. Moreover, quality of separation is insufficient, and the amount of protein seems to be lower. Fractions B and C were combined and concentrated to 0.689 mg $mL^{-1}$.

## 4.7 Protease inhibitor assay

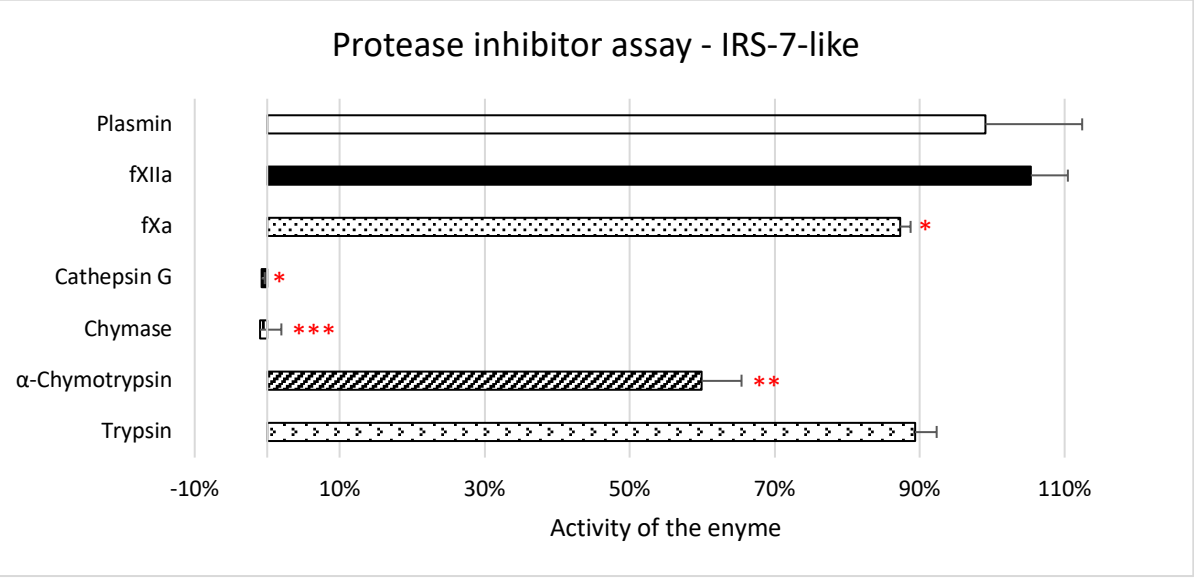


Figure 18: IRS-7-like – Protease inhibitor assay; percentage of enzyme activity in presence of inhibitor compared to uninhibited enzyme; two-sided, heteroscedastic t-test; statistical significance for $p < 0.05$ (* $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$)

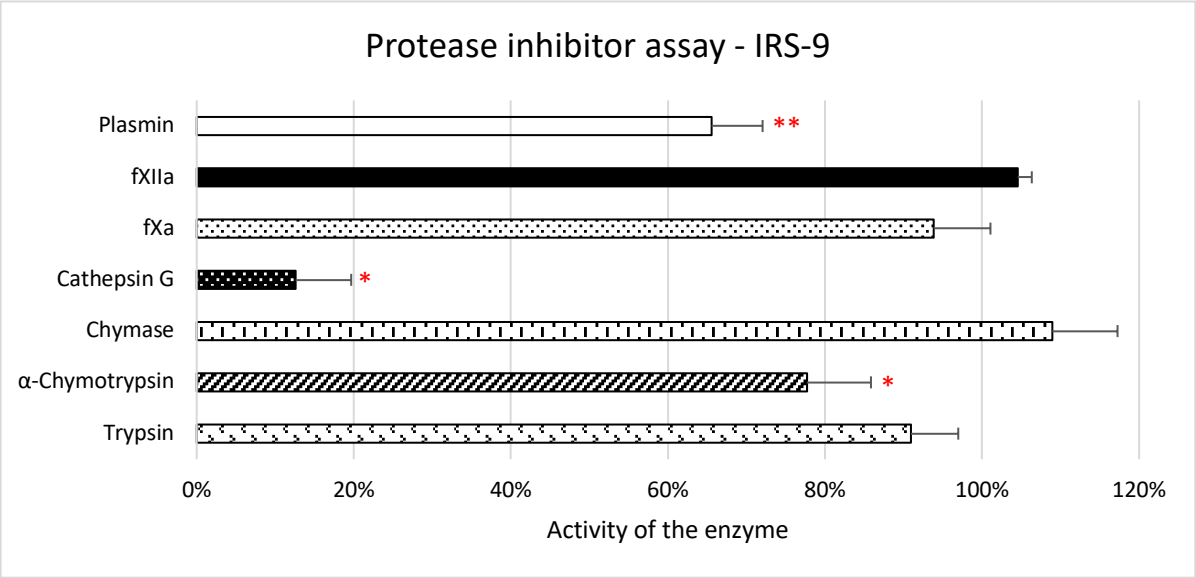


Figure 19: IRS-9 – Protease inhibitor assay; percentage of enzyme activity in presence of inhibitor compared to uninhibited enzyme; two-sided, heteroscedastic t-test; statistical significance for $p < 0.05$ (* $p < 0.05$; ** $p < 0.01$)

The ability of the purified and concentrated serpins to inhibit serine proteases was tested in protease inhibitor assays. Seven enzymes were tested according to the procedure explained in 3.2.5.3. The results for IRS-7-like are presented in Figure 18, while those for IRS-9 in Figure 19. The reaction rate was compared to a control without a serpin, which was considered as 100 % activity.

Shown in Figure 18, 2 µM IRS-7-like fully inhibits the activity of Chymase and Cathepsin G, while a significant reduction in the activity of the enzymes was also observed for α-Chymotrypsin showing a remaining activity of 60% and 90% for fXa. In comparison, IRS-9 shows a significant inhibition of Cathpesin G with only 10% of active protease remaining and a lower inhibition of Plasmin and α-Chymotrypsin.

# 5 Discussion

## 5.1 Cloning

The first objective of cloning several serpin genes, namely the genes of IRS-7, IRS-9, IRS-10, IRS-11, IRS-12, IRS-18, IRS-19, IRS-20, IRS-25, IRS-27 and a chimeric IRS-2, into a pET-17b expression vector was achieved for a few, but not all genes.

In the first step, gradient PCR, IRS-19 had to be excluded, due to unsuccessful generation of an amplicon, while no amplicon was observed for IRS-27 in the high-fidelity PCR also leading to an exclusion. Designing new primers with different parameters, choosing a different polymerase or different PCR conditions, such as amount of cycles and changing temperatures, might all be possible solutions to also obtain amplicons for these primers.

The remaining amplicons were cut with restriction enzymes which led to the realisation that two genes, the ones for IRS-18 and IRS-25, contain restriction sites for one of the used enzymes, XhoI. This led to the generation of multiple fragments. A different restriction enzyme without a cutting site in the serpin gene must be chosen. New primers must then be designed with the new enzyme cutting site to solve this issue.

The performed colony PCR did not give positive results for the IRS-20 gene, also leading to an exclusion of this serpin. All four remaining clones from the primers for IRS-7, IRS-9, IRS-10-11-12 and the chimeric IRS-2 were sent for sequencing. The structure of the chimeric serpin and of IRS-9 was confirmed to be identical to the expected sequence. The fragments from primers designed for IRS-7 and IRS-10-11-12 showed between 4 and 7 % amino acid difference from the template sequences, which means that IRS-7-like and IRS-9-like are possibly homologues of the original templates. This theory seems very plausible, since Chmelař et al explain that tick serpins have homologues differing in only a several amino acids to reduce immunogenicity of each single serpin (Chmelař et al., 2016).

## 5.2 Protein expression

Pilot expressions showed that IRS-7-like and IRS-9 were sufficiently expressed in standard LB medium, while the chimeric serpin and IRS-10-like were not expressed in high enough amounts, which led to the need for further optimisations. By adding glucose and ethanol to the growth medium, good results were achieved.

The low expression level in usual LB medium might have been caused by toxicity of the serpins to *E. coli* cells. Even without IPTG induction, proteins are produced at a low, basal

level. If the recombinant protein is toxic to *E. coli* cells, it might interfere with normal cell functions and cause a lack of establishment of the plasmids in the bacterial cells. This low basal expression can be suppressed by supplying the cells with glucose until induction, which led to improved expression levels (Novagen). Additionally, ethanol changes the environment causing changes in protein expression, which showed as a lower production of endogenous *E. coli* proteins (Chhetri et al., 2015).

Even though small-scale expression showed good results for IRS-7-like and IRS-9, the bacteria cells grew very slowly when protein expression was scaled up. Overexpression of bacteria is a complex topic with many optimisable parameters such as growth medium, pH, temperature, culture volume to flask volume ratio, agitation rate, IPTG concentration or induction period (Collins et al., 2013). Which means that many possible reasons could have led to slow bacteria growth and low protein yield of only around 1 mg for several litres of bacteria cultures.

Collins et al. found that a culture volume to flask volume ratio of 1:10 provided the best results, as sufficient aeration is vital for *E. coli* growth (Collins et al., 2013). Since, for this overexpression 1 L of medium was used in 2 L flasks, aeration was likely not sufficient. Another factor leading to poor protein production might be that β-lactamase accumulates in the medium during cell growth and degrades antibiotics. This causes a possible loss of plasmids due to a lack of selective pressure and in turn leads to a decreased protein yield (Collins et al., 2013). Combining an improper aeration leading to a significantly slowed growth, which causes a high accumulation of β-lactamase, and the resulting reduced selective pressure might both have contributed to poor yield.

Overall, the second objective to overexpress the serpins in *E. coli*, was, in the end, achieved for all four serpins, IRS-2 chimera, IRS-7-like, IRS-9 and IRS-10-like. However, only pilot expression could be performed for IRS-2 chimera and IRS-10-like due to time constraints. Moreover, although the obtained protein amount was sufficient for preliminary inhibitor assays, for studying the serpins in greater detail, more protein is necessary. Thus, the many variables for optimised protein overexpression must be explored.

## 5.3 Protein refolding and purification

Generally, only very low protein yields of 1 mg were obtained, for which many factors can possibly be the reason. First, contamination with proteases, that degraded the serpins during the process of concentrating, is possible, since the whole procedure took a few days. The

formation of inclusion bodies, due to the lack of mechanisms to modify eukaryotic proteins post-translationally in *E. coli*, might be one of the reasons and poor for yield (Bird et al., 2004). On the one hand, IB can greatly simplify purification of the recombinant protein, since most endogenous *E. coli* proteins remain in soluble fractions, while recombinant proteins accumulate in inclusion bodies (Baneyx, 1999). On the other hand, the incorrectly folded, aggregated proteins must be refolded and resolubilised, which might not be possible at all sometimes or lead to decreased protein yield.

The refolding of serpins is an empirical process and many conditions must be tested, as every serpin refolds best in different environments. Additives such as glycerol, sucrose, arginine, or methionine can be included in the refolding buffer to increase native protein yield. In his work Kotál found that two *I. ricinus* serpins gave the highest yield in refolding buffers with a neutral pH and with high amounts of Arginine (Kotál, 2013). Proteins with an acidic pI refold better in basic refolding buffers and vice versa (Coutard et al., 2012), which is valid for IRS-7-like with a pI of 5.69 showing best results in a refolding buffer with a pH of 8.5. The second refolded protein has a pI of 7.23, which is uncommon and might make it trickier to refold (Coutard et al., 2012). The refolding of IRS-9 might therefore benefit a lot from testing different pH conditions. Kotál also optimised reducing agent concentrations, IB concentrations, Arginine concentrations and ionic strength of the refolding buffer, which I propose should all also be tested for the refolding of IRS-7-like and IRS-9 in further experiments.

One limitation of the presented results is that refolding efficiency was determined by measuring the protein concentration after removing precipitated, and therefore inactive, proteins. This means that higher concentrations indicate less precipitated, inactive proteins and more correctly refolded molecules. However, misfolded proteins also form small aggregates which are not removed by centrifugation. Therefore, also measuring remaining aggregated protein in the solution, like Kotál did, would give more precise results about refolding efficiency in further work (Kotál, 2013).

The HPLC purification of IRS-7 gave good results, while the same process did not give satisfying results for IRS-9. After SEC and IEC, the gel still showed proteins at higher and lower molecular weight than the target protein. A possible explanation is obviously contamination with *E. coli* proteins, for lower weight proteins the target protein could have been degraded by proteases.

Generally, even though the third goal of refolding and purifying the serpins, showed some promising results, many further optimisation possibilities have to be tested to improve these initial refolding experiments.

## 5.4 Inhibitor assay

The last objective, to test the serpins in protease inhibitor assays, for some preliminary functional characterisation, already provided some interesting results.

### 5.4.1 IRS-7-like

IRS-7-like shows Leucine at its P1 position, which, like Tyrosine at the P1 position of IRS-2 (Chmelar et al., 2011), inhibits chymotrypsin like serine proteases (Zhao et al., 2012). This specificity is also shown in the performed protease inhibitor assay with a significant inhibition of α-chymotrypsin. Generally, the targeted proteases of IRS-7-like are very similar to the target proteases of IRS-2. Cathepsin G is completely inhibited by IRS-7-like. It has trypsin-like and chymotrypsin-like properties by preferring basic as well as aromatic residues at the P1 position (Polanowska et al., 1998). Even though IRS-7-like shows Leucine at its P1 position, it inhibits Cathepsin G. Cathepsin G released by neutrophils plays a role in inflammation, tissue remodelling and in the crosstalk between platelets and neutrophils. Thus, it helps the tick obtain its meal by suppressing the innate immune system and haemostasis (Chmelar et al., 2011). Another by IRS-7-like completely inhibited protease, Chymase, plays a role in the recruitment of more neutrophils (Chmelar et al., 2011), meaning that it also inhibits the same host processes. Last, factor fXa is also suppressed by IRS-7-like, thus indicating that the serpin has an effect on the coagulation cascade (Mann et al., 2003).

### 5.4.2 IRS-9-like

IRS-9 shows a basic Lysine residue at its P1 position, which indicates that it inhibits trypsin-like proteases. However, surprisingly it does not significantly inhibit Trypsin, the model trypsin-like protease, but it slightly inhibits α-Chymotrypsin. It strongly inhibits Cathepsin G, with 10 % remaining enzymes activity. This can be explained by the basic residue at the P1 position, for which Cathepsin G was shown to have an affinity for (Polanowska et al., 1998). Thus, IRS-9, like IRS-7-like, also inhibits host inflammation and platelet aggregation (Chmelar et al., 2011). IRS-9 also inhibits Plasmin, which was shown to prefer basic amino acids at its P1 position, such as Lysine (Hervio et al., 2000). The suppression of plasmin indicates effects on the coagulation cascade, tissue remodelling and inflammation (Li et al., 2005).

# 6 Conclusion

The first aim, cloning of various serpins into an expression vector was achieved for IRS-2 chimera, IRS-7-like, IRS-9 and IRS-10-like. Subsequent overexpression of the serpins showed sufficient protein production of IRS-7-like and IRS-9. The expression conditions for the other two serpins were optimised, but due to time constraints, no further work on the chimeric serpin and IRS-10-like could be done. IRS-7-like and IRS-9 were, despite low amounts, refolded and purified by HPLC. The final protease inhibitor assays served as preliminary tests for the functional characterisation of the serpins. The assays showed that IRS-7-like inhibits Cathepsin G, Chymase and fX while IRS-9 inhibits Cathpesin G and Plasmin. Thus, both serpins most likely play a role in the suppression of host haemostasis and inflammation.

# 7 References


Amara, U.; Rittirsch, D.; Flierl, M.; Bruckner, U.; Klos, A.; Gebhard, F.; Lambris, J. D.; Huber-Lang, M. Interaction between the coagulation and complement system. *Advances in experimental medicine and biology* **2008,** *632*, 71–79.

Amicizia, D.; Domnich, A.; Panatto, D.; Lai, P. L.; Cristina, M. L.; Avio, U.; Gasparini, R. Epidemiology of tick-borne encephalitis (TBE) in Europe and its prevention by available vaccines. *Human vaccines & immunotherapeutics* **2013,** *9* (5), 1163–1171. DOI: 10.4161/hv.23802.

Banchereau, J.; Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **1998,** *392* (6673), 245–252. DOI: 10.1038/32588.

Baneyx, F. Recombinant protein expression in Escherichia coli. *Current opinion in biotechnology* **1999,** *10* (5), 411–421.

Bania, J.; Stachowiak, D.; Polanowski, A. Primary structure and properties of the cathepsin G/chymotrypsin inhibitor from the larval hemolymph of Apis mellifera. *European journal of biochemistry* **1999,** *262* (3), 680–687.

Bell, M. R.; Engleka, M. J.; Malik, A.; Strickler, J. E. To fuse or not to fuse: what is your purpose? *Protein science : a publication of the Protein Society* **2013,** *22* (11), 1466–1477. DOI: 10.1002/pro.2356.

Bird, P. I.; Pak, S. C.; Worrall, D. M.; Bottomley, S. P. Production of recombinant serpins in Escherichia coli. *Methods (San Diego, Calif.)* **2004,** *32* (2), 169–176.

*Ticks*. *Biology, disease and control;* Bowman, A. S., Nuttall, P. A., Eds.; Cambridge University Press: Cambridge, 2008.

Brossard, M.; Wikel, S. K. Tick immunobiology. *Parasitology* **2004,** *129 Suppl*, S161-76.

Broze, G. J. Tissue factor pathway inhibitor and the revised theory of coagulation. *Annual review of medicine* **1995,** *46*, 103–112. DOI: 10.1146/annurev.med.46.1.103.

Cabrita, L. D.; Bottomley, S. P. How do proteins avoid becoming too stable? Biophysical studies into metastable proteins. *European biophysics journal : EBJ* **2004,** *33* (2), 83–88. DOI: 10.1007/s00249-003-0356-1.

Chalaire, K. C.; Kim, T. K.; Garcia-Rodriguez, H.; Mulenga, A. Amblyomma americanum (L.) (Acari: Ixodidae) tick salivary gland serine protease inhibitor (serpin) 6 is secreted into tick saliva during tick feeding. *The Journal of experimental biology* **2011,** *214* (Pt 4), 665–673. DOI: 10.1242/jeb.052076.

Chávez-Galán, L.; Arenas-Del Angel, M. C.; Zenteno, E.; Chávez, R.; Lascurain, R. Cell death mechanisms induced by cytotoxic lymphocytes. *Cellular & molecular immunology* **2009,** *6* (1), 15–25. DOI: 10.1038/cmi.2009.3.

Chen, Z.; Wang, B.; Hu, J.; Yang, W.; Cao, Z.; Zhuo, R.; Li, W.; Wu, Y. SjAPI, the first functionally characterized Ascaris-type protease inhibitor from animal venoms. *PloS one* **2013,** *8* (3), e57529. DOI: 10.1371/journal.pone.0057529.

Chhetri, G.; Kalita, P.; Tripathi, T. An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli. *MethodsX* **2015,** *2*, 385–391. DOI: 10.1016/j.mex.2015.09.005.

Chmelar, J.; Calvo, E.; Pedra, J. H. F.; Francischetti, I. M. B.; Kotsyfakis, M. Tick salivary secretion as a source of antihemostatics. *Journal of proteomics* **2012,** *75* (13), 3842–3854. DOI: 10.1016/j.jprot.2012.04.026.

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.; Kotsyfakis, M. A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood* **2011,** *117* (2), 736–744. DOI: 10.1182/blood-2010-06-293241.

Chmelař, J.; Kotál, J.; Kopecký, J.; Pedra, J. H. F.; Kotsyfakis, M. All For One and One For All on the Tick-Host Battlefield. *Trends in parasitology* **2016,** *32* (5), 368–377. DOI: 10.1016/j.pt.2016.01.004.

Chmelař, J.; Kotál, J.; Langhansová, H.; Kotsyfakis, M. Protease Inhibitors in Tick Saliva: The Role of Serpins and Cystatins in Tick-host-Pathogen Interaction. *Frontiers in cellular and infection microbiology* **2017,** *7*, 216. DOI: 10.3389/fcimb.2017.00216.

Clarke, E. P.; Cates, G. A.; Ball, E. H.; Sanwal, B. D. A collagen-binding protein in the endoplasmic reticulum of myoblasts exhibits relationship with serine protease inhibitors. *The Journal of biological chemistry* **1991,** *266* (26), 17230–17235.

Collen, D. The plasminogen (fibrinolytic) system. *Thrombosis and haemostasis* **1999,** *82* (2), 259–270.

Collins, T.; Azevedo-Silva, J.; da Costa, A.; Branca, F.; Machado, R.; Casal, M. Batch production of a silk-elastin-like protein in E. coli BL21(DE3): key parameters for optimisation. *Microbial cell factories* **2013,** *12*, 21. DOI: 10.1186/1475-2859-12-21.

Corral-Rodríguez, M. A.; Macedo-Ribeiro, S.; Barbosa Pereira, P. J.; Fuentes-Prior, P. Tick-derived Kunitz-type inhibitors as antihemostatic factors. *Insect biochemistry and molecular biology* **2009,** *39* (9), 579–595. DOI: 10.1016/j.ibmb.2009.07.003.

Coutard, B.; Danchin, E. G. J.; Oubelaid, R.; Canard, B.; Bignon, C. Single pH buffer refolding screen for protein from inclusion bodies. *Protein expression and purification* **2012,** *82* (2), 352–359. DOI: 10.1016/j.pep.2012.01.014.

Couvreur, B.; Beaufays, J.; Charon, C.; Lahaye, K.; Gensale, F.; Denis, V.; Charloteaux, B.; Decrem, Y.; Prévôt, P.-P.; Brossard, M.; Vanhamme, L.; Godfroid, E. Variability and action mechanism of a family of anticomplement proteins in Ixodes ricinus. *PloS one* **2008,** *3* (1), e1400. DOI: 10.1371/journal.pone.0001400.

El Rayes, T.; Catena, R.; Lee, S.; Stawowczyk, M.; Joshi, N.; Fischbach, C.; Powell, C. A.; Dannenberg, A. J.; Altorki, N. K.; Gao, D.; Mittal, V. Lung inflammation promotes metastasis through neutrophil protease-mediated degradation of Tsp-1. *Proceedings of the*

*National Academy of Sciences of the United States of America* **2015,** *112* (52), 16000–16005. DOI: 10.1073/pnas.1507294112.

Elliott, P. R.; Lomas, D. A.; Carrell, R. W.; Abrahams, J. P. Inhibitory conformation of the reactive loop of alpha 1-antitrypsin. *Nature structural biology* **1996,** *3* (8), 676–681.

Elovic, A. E.; Ohyama, H.; Sauty, A.; McBride, J.; Tsuji, T.; Nagai, M.; Weller, P. F.; Wong, D. T. IL-4-dependent regulation of TGF-alpha and TGF-beta1 expression in human eosinophils. *Journal of immunology (Baltimore, Md. : 1950)* **1998,** *160* (12), 6121–6127.

Fogaça, A. C.; Almeida, I. C.; Eberlin, M. N.; Tanaka, A. S.; Bulet, P.; Daffre, S. Ixodidin, a novel antimicrobial peptide from the hemocytes of the cattle tick Boophilus microplus with inhibitory activity against serine proteinases. *Peptides* **2006,** *27* (4), 667–674. DOI: 10.1016/j.peptides.2005.07.013.

Francischetti, I. M. B.; Sa-Nunes, A.; Mans, B. J.; Santos, I. M.; Ribeiro, J. M. C. The role of saliva in tick feeding. *Frontiers in bioscience (Landmark edition)* **2009,** *14*, 2051–2088.

*Mast Cell Biology. Contemporary and Emerging Topics;* Gilfillan, A. M., Metcalfe, D. D., Eds.; Advances in Experimental Medicine and Biology 716; Landes Bioscience and Springer Science+Business Media LLC: Boston, MA, 2011.

Hammond, G. L.; Smith, C. L.; Goping, I. S.; Underhill, D. A.; Harley, M. J.; Reventos, J.; Musto, N. A.; Gunsalus, G. L.; Bardin, C. W. Primary structure of human corticosteroid binding globulin, deduced from hepatic and pulmonary cDNAs, exhibits homology with serine protease inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **1987,** *84* (15), 5153–5157.

Heath, W. R.; Carbone, F. R. The skin-resident and migratory immune system in steady state and memory: innate lymphocytes, dendritic cells and T cells. *Nature immunology* **2013,** *14* (10), 978–985. DOI: 10.1038/ni.2680.

Heit, C.; Jackson, B. C.; McAndrews, M.; Wright, M. W.; Thompson, D. C.; Silverman, G. A.; Nebert, D. W.; Vasiliou, V. Update of the human and mouse SERPIN gene superfamily. *Human genomics* **2013,** *7*, 22. DOI: 10.1186/1479-7364-7-22.

Hervio, L. S.; Coombs, G. S.; Bergstrom, R. C.; Trivedi, K.; Corey, D. R.; Madison, E. L. Negative selectivity and the evolution of protease cascades: the specificity of plasmin for peptide and protein substrates. *Chemistry & Biology* **2000,** *7* (6), 443–452. DOI: 10.1016/S1074-5521(00)00125-3.

Huntington, J. A.; Read, R. J.; Carrell, R. W. Structure of a serpin-protease complex shows inhibition by deformation. *Nature* **2000,** *407* (6806), 923–926. DOI: 10.1038/35038119.

Ibelli, A. M. G.; Hermance, M. M.; Kim, T. K.; Gonzalez, C. L.; Mulenga, A. Bioinformatics and expression analyses of the Ixodes scapularis tick cystatin family. *Experimental & applied acarology* **2013,** *60* (1), 41–53. DOI: 10.1007/s10493-012-9613-2.

Invitrogen. TrackIt™ 100 bp DNA Ladder: Product Information Sheet. https:// assets.thermofisher.com/TFS-Assets/LSG/manuals/trackit_100bp_man.pdf.

Kim, T. K.; Tirloni, L.; Radulovic, Z.; Lewis, L.; Bakshi, M.; Hill, C.; da Silva Vaz, I.; Logullo, C.; Termignoni, C.; Mulenga, A. Conserved Amblyomma americanum tick Serpin19, an inhibitor of blood clotting factors Xa and XIa, trypsin and plasmin, has anti-haemostatic functions. *International journal for parasitology* **2015,** *45* (9-10), 613–627. DOI: 10.1016/j.ijpara.2015.03.009.

Kotál. Production and functional characterization of tick salivary protease inhibitors. Master Thesis; Jihočeská univerzita v Českých Budějovicích, České Budějovice, 2013.

Kotál, J.; Langhansová, H.; Lieskovská, J.; Andersen, J. F.; Francischetti, I. M. B.; Chavakis, T.; Kopecký, J.; Pedra, J. H. F.; Kotsyfakis, M.; Chmelař, J. Modulation of host immunity by tick saliva. *Journal of proteomics* **2015,** *128*, 58–68. DOI: 10.1016/j.jprot.2015.07.005.

Law, R. H. P.; Zhang, Q.; McGowan, S.; Buckle, A. M.; Silverman, G. A.; Wong, W.; Rosado, C. J.; Langendorf, C. G.; Pike, R. N.; Bird, P. I.; Whisstock, J. C. An overview of the serpin superfamily. *Genome biology* **2006,** *7* (5), 216. DOI: 10.1186/gb-2006-7-5-216.

Leboulle, G.; Crippa, M.; Decrem, Y.; Mejri, N.; Brossard, M.; Bollen, A.; Godfroid, E. Characterization of a novel salivary immunosuppressive protein from Ixodes ricinus ticks. *The Journal of biological chemistry* **2002,** *277* (12), 10083–10089. DOI: 10.1074/jbc.M111391200.

Li, J.; Ny, A.; Leonardsson, G.; Nandakumar, K. S.; Holmdahl, R.; Ny, T. The plasminogen activator/plasmin system is essential for development of the joint inflammatory phase of collagen type II-induced arthritis. *The American journal of pathology* **2005,** *166* (3), 783–792. DOI: 10.1016/S0002-9440(10)62299-7.

Lu, H.; Cassis, L. A.; Kooi, C. W. V.; Daugherty, A. Structure and functions of angiotensinogen. *Hypertension research : official journal of the Japanese Society of Hypertension* **2016,** *39* (7), 492–500. DOI: 10.1038/hr.2016.17.

Mann, K. G.; Butenas, S.; Brummel, K. The dynamics of thrombin formation. *Arteriosclerosis, thrombosis, and vascular biology* **2003,** *23* (1), 17–25.

Mansuy-Aubert, V.; Zhou, Q. L.; Xie, X.; Gong, Z.; Huang, J.-Y.; Khan, A. R.; Aubert, G.; Candelaria, K.; Thomas, S.; Shin, D.-J.; Booth, S.; Baig, S. M.; Bilal, A.; Hwang, D.; Zhang, H.; Lovell-Badge, R.; Smith, S. R.; Awan, F. R.; Jiang, Z. Y. Imbalance between neutrophil elastase and its inhibitor α1-antitrypsin in obesity alters insulin sensitivity, inflammation, and energy expenditure. *Cell metabolism* **2013,** *17* (4), 534–548. DOI: 10.1016/j.cmet.2013.03.005.

Marcet-Palacios, M.; Ewen, C.; Pittman, E.; Duggan, B.; Carmine-Simmen, K.; Fahlman, R. P.; Bleackley, R. C. Design and characterization of a novel human Granzyme B inhibitor. *Protein engineering, design & selection : PEDS* **2015,** *28* (1), 9–17. DOI: 10.1093/protein/gzu052.

Meekins, D. A.; Kanost, M. R.; Michel, K. Serpins in arthropod biology. *Seminars in cell & developmental biology* **2017,** *62*, 105–119. DOI: 10.1016/j.semcdb.2016.09.001.

Metcalfe, D. D.; Baram, D.; Mekori, Y. A. Mast cells. *Physiological reviews* **1997,** *77* (4), 1033–1079. DOI: 10.1152/physrev.1997.77.4.1033.

Müller, S.; Dennemärker, J.; Reinheckel, T. Specific functions of lysosomal proteases in endocytic and autophagic pathways. *Biochimica et biophysica acta* **2012,** *1824* (1), 34–43. DOI: 10.1016/j.bbapap.2011.07.003.

Nava, S.; Guglielmone, A. A.; Mangold, A. J. An overview of systematics and evolution of ticks. *Frontiers in bioscience (Landmark edition)* **2009,** *14*, 2857–2877.

Novagen. pET System Manual. http://lifeserv.bgu.ac.il/wp/zarivach/wp-content/uploads/2017/11/Novagen-pET-system-manual-1.pdf.

Novex. SeeBlue® Plus2 Pre-Stained Protein Standard Protocol. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/SeeBluePlus2PreStainedStnd_man.pdf.

Ooi, C.-P.; Haines, L. R.; Southern, D. M.; Lehane, M. J.; Acosta-Serrano, A. Tsetse GmmSRPN10 has anti-complement activity and is important for successful establishment of trypanosome infections in the fly midgut. *PLoS neglected tropical diseases* **2015,** *9* (1), e3448. DOI: 10.1371/journal.pntd.0003448.

Owen, C. A. Roles for proteinases in the pathogenesis of chronic obstructive pulmonary disease. *International journal of chronic obstructive pulmonary disease* **2008,** *3* (2), 253–268.

Patston, P. A. Serpins and other serine protease inhibitors. *Immunology today* **2000,** *21* (7), 354.

Pearson, S. Recognising and understanding Lyme disease. *Nursing standard (Royal College of Nursing (Great Britain) : 1987)* **2014,** *29* (1), 37–43. DOI: 10.7748/ns.29.1.37.e9073.

Polanowska, J.; Krokoszynska, I.; Czapinska, H.; Watorek, W.; Dadlez, M.; Otlewski, J. Specificity of human cathepsin G. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1998,** *1386* (1), 189–198. DOI: 10.1016/S0167-4838(98)00085-5.

Porter, L.; Radulović, Ž.; Kim, T.; Braz, G. R. C.; da Silva Vaz, I.; Mulenga, A. Bioinformatic analyses of male and female Amblyomma americanum tick expressed serine protease inhibitors (serpins). *Ticks and tick-borne diseases* **2015,** *6* (1), 16–30. DOI: 10.1016/j.ttbdis.2014.08.002.

Ramamoorthi, N.; Narasimhan, S.; Pal, U.; Bao, F.; Yang, X. F.; Fish, D.; Anguita, J.; Norgard, M. V.; Kantor, F. S.; Anderson, J. F.; Koski, R. A.; Fikrig, E. The Lyme disease agent exploits a tick protein to infect the mammalian host. *Nature* **2005,** *436* (7050), 573–577. DOI: 10.1038/nature03812.

Ribeiro, J. M. C.; Francischetti, I. M. B. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annual review of entomology* **2003,** *48*, 73–88. DOI: 10.1146/annurev.ento.48.060402.102812.

Rothenberg, M. E.; Hogan, S. P. The eosinophil. *Annual review of immunology* **2006,** *24*, 147–174. DOI: 10.1146/annurev.immunol.24.021605.090720.

Russell, R. B. Domain insertion. *Protein engineering* **1994,** *7* (12), 1407–1410.

Saksela, O.; Rifkin, D. B. Cell-associated plasminogen activation: regulation and physiological functions. *Annual review of cell biology* **1988,** *4*, 93–126. DOI: 10.1146/annurev.cb.04.110188.000521.

Sá-Nunes, A.; Bafica, A.; Antonelli, L. R.; Choi, E. Y.; Francischetti, I. M. B.; Andersen, J. F.; Shi, G.-P.; Chavakis, T.; Ribeiro, J. M.; Kotsyfakis, M. The immunomodulatory action of sialostatin L on dendritic cells reveals its potential to interfere with autoimmunity. *Journal of immunology (Baltimore, Md. : 1950)* **2009,** *182* (12), 7422–7429. DOI: 10.4049/jimmunol.0900075.

Schaiff, W. T.; Eisenberg, P. R. Direct induction of complement activation by pharmacologic activation of plasminogen. *Coronary artery disease* **1997,** *8* (1), 9–18.

Schechter, I.; Berger, A. On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochemical and biophysical research communications* **1968,** *32* (5), 898–902.

Schwarz, A.; Reumont, B. M. von; Erhart, J.; Chagas, A. C.; Ribeiro, J. M. C.; Kotsyfakis, M. De novo Ixodes ricinus salivary gland transcriptome analysis using two next-generation sequencing methodologies. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **2013,** *27* (12), 4745–4756. DOI: 10.1096/fj.13-232140.

Segal, A. W. How neutrophils kill microbes. *Annual review of immunology* **2005,** *23*, 197–223. DOI: 10.1146/annurev.immunol.23.021704.115653.

Seixas, S.; Ivanova, N.; Ferreira, Z.; Rocha, J.; Victor, B. L. Loss and gain of function in SERPINB11: an example of a gene under selection on standing variation, with implications for host-pathogen interactions. *PloS one* **2012,** *7* (2), e32518. DOI: 10.1371/journal.pone.0032518.

Silverman, G. A.; Bird, P. I.; Carrell, R. W.; Church, F. C.; Coughlin, P. B.; Gettins, P. G.; Irving, J. A.; Lomas, D. A.; Luke, C. J.; Moyer, R. W.; Pemberton, P. A.; Remold-O'Donnell, E.; Salvesen, G. S.; Travis, J.; Whisstock, J. C. The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *The Journal of biological chemistry* **2001,** *276* (36), 33293–33296. DOI: 10.1074/jbc.R100016200.

Stark, K. R.; James, A. A. Isolation and characterization of the gene encoding a novel factor Xa-directed anticoagulant from the yellow fever mosquito, Aedes aegypti. *The Journal of biological chemistry* **1998,** *273* (33), 20802–20809.

Thermo Scientific. Thermo Scientific GeneRuler 100 bp Plus DNA Ladder. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0013008_GeneRuler_100bp_Plus_DNALadder_50ug_UG.pdf.

Tirloni, L.; Seixas, A.; Mulenga, A.; Vaz, I. d. S.; Termignoni, C. A family of serine protease inhibitors (serpins) in the cattle tick Rhipicephalus (Boophilus) microplus. *Experimental parasitology* **2014,** *137*, 25–34. DOI: 10.1016/j.exppara.2013.12.001.

Tsien, R. Y. The green fluorescent protein. *Annual review of biochemistry* **1998,** *67*, 509–544. DOI: 10.1146/annurev.biochem.67.1.509.

Wagner, C. J.; Schultz, C.; Mall, M. A. Neutrophil elastase and matrix metalloproteinase 12 in cystic fibrosis lung disease. *Molecular and cellular pediatrics* **2016,** *3* (1), 25. DOI: 10.1186/s40348-016-0053-7.

Waterhouse, R. M.; Kriventseva, E. V.; Meister, S.; Xi, Z.; Alvarez, K. S.; Bartholomay, L. C.; Barillas-Mury, C.; Bian, G.; Blandin, S.; Christensen, B. M.; Dong, Y.; Jiang, H.; Kanost, M. R.; Koutsos, A. C.; Levashina, E. A.; Li, J.; Ligoxygakis, P.; Maccallum, R. M.; Mayhew, G. F.; Mendes, A.; Michel, K.; Osta, M. A.; Paskewitz, S.; Shin, S. W.; Vlachou, D.; Wang, L.; Wei, W.; Zheng, L.; Zou, Z.; Severson, D. W.; Raikhel, A. S.; Kafatos, F. C.; Dimopoulos, G.; Zdobnov, E. M.; Christophides, G. K. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science (New York, N.Y.)* **2007,** *316* (5832), 1738–1743. DOI: 10.1126/science.1139862.

Wikel, S. Ticks and tick-borne pathogens at the cutaneous interface: host defenses, tick countermeasures, and a suitable environment for pathogen establishment. *Frontiers in microbiology* **2013,** *4*, 337. DOI: 10.3389/fmicb.2013.00337.

Yu, K.; Liu, C.; Kim, B.-G.; Lee, D.-Y. Synthetic fusion protein design and applications. *Biotechnology advances* **2015,** *33* (1), 155–164. DOI: 10.1016/j.biotechadv.2014.11.005.

Zhao, P.; Dong, Z.; Duan, J.; Wang, G.; Wang, L.; Li, Y.; Xiang, Z.; Xia, Q. Genome-wide identification and immune response analysis of serine protease inhibitor genes in the silkworm, Bombyx mori. *PloS one* **2012,** *7* (2), e31168. DOI: 10.1371/journal.pone.0031168.