UNIVERSITY OF SOUTH BOHEMIA IN ČESKÉ BUDĚJOVICE
FACULTY OF SCIENCE

Structural insight into the salivary serpins of *Ixodes ricinus*

Ph.D. Thesis

Mgr. Barbora Kaščáková

Supervisor: prof. Mgr. Ivana Kutá Smatanová, Ph.D.
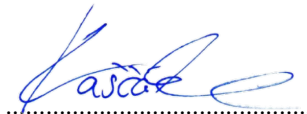
České Budějovice 2022

Annotation

The knowledge of the detailed structure of proteins and their complexes with other proteins, such as serpins, helps to understand the mechanism of action. Serpins, a large protein group of protease inhibitors that possess almost identical secondary-structural folds, represent the perfect example of expanding the knowledge of their inhibition process through detailed structural analyses. The universal process of serpin inhibition is known, but also is known that serpins are structurally similar whereas their functional diversity is significant. Therefore, each serpin will have some properties specific to its own and knowledge of the serpin structures can explain high functional diversity. X-ray crystallography was one of the most common tools used for serpin structural analysis.

This thesis describes the structural information of serpins found in *Ixodes ricinus* ticks. Serpins of this species are mainly responsible for the modulation of the host immune response via inhibiting involved proteases. Serpins with proteases form covalent complexes. This process leads to a suicide mechanism that inactivates the protease as well as serpin. Here are presented results of the X-ray structural analysis of four *I. ricinus* serpins named Iripin-3, Iripin-5, Iripin-4, and Iripin-1. All of them help the tick in different ways to stay attached to the host for sufficient time for feeding by inhibiting the proteases involved in host immune defense responses to a tick bite.

Declaration

I hereby declare that I am the author of this dissertation and that I have used only those sources and literature detailed in the list of references.

České Budějovice, 15.04.2022 ....

Barbora Kaščáková

This thesis originated from the Faculty of Science, University of South Bohemia supporting doctoral studies in the Biochemistry study programme.

Dedicated to my family

List of papers and author's contribution

The thesis is based on the following papers (listed chronologically):

I.   Chlastakova, A., Kotal, J., Berankova, Z., **Kascakova, B.,** Martins, L. A., Langhansova, H., Prudnikova, T., Ederova, M., Kuta Smatanova, I., Kotsyfakis, M., & Chmelar, J. (2021). Iripin-3, a New Salivary Protein Isolated From *Ixodes ricinus* Ticks, Displays Immunomodulatory and Anti-Hemostatic Properties In Vitro. Frontiers in immunology, 12, 626200. (IF= 6.429) https://doi.org/10.3389/fimmu.2021.626200.
Barbora Kaščáková participated in crystallization, X-ray data measurements, data solving and manuscript preparation. Her contribution was 20 %.

II.   **Kascakova, B.,** Kotal, J., Martins, L. A., Berankova, Z., Langhansova, H., Calvo, E., Crossley, J. A., Havlickova, P., Dycka, F., Prudnikova, T., Kuty, M., Kotsyfakis, M., Chmelar, J., & Kuta Smatanova, I. (2021). Structural and biochemical characterization of the novel serpin Iripin-5 from *Ixodes ricinus*. Acta crystallographica. Section D, Structural biology, 77(Pt 9), 1183–1196. (IF= 7.652) https://doi.org/10.1107/S2059798321007920.
Barbora Kaščáková participated in crystallization, X-ray data measurements, data solving, and she was responsible for manuscript writing. Her contribution was 50 %.

III.   **Kascakova, B.,** Kotal, J., Havlickova, P., Prudnikova, T., Grinkiewich, P., Kuty, M., Chmelar, J. & Kuta Smatanova, I. Conformational transition of Iripin-4, *Ixodes ricinus* salivary serpin. *Manuscript.*
Barbora Kaščáková participated in crystallization, X-ray data measurements, data solving, and she was responsible for manuscript writing. Her contribution was 50 %.

IV.   **Kascakova, B.,** Havlickova, P., Prudnikova, T., Kuta Smatanova, I., & Chmelar, J. Crystal structure of *Ixodes ricinus* serpin Iripin-1. *Manuscript.*
Barbora Kaščáková was responsible for crystallization, X-ray data measurements, data solving and manuscript preparation. Her contribution was 50%.

Co-author agreement

Ivana Kutá Smatanová, the supervisor of this Ph.D. thesis and co-author of all stated papers, fully acknowledges the stated contribution of Barbora Kaščáková to these manuscripts.

……………………………………………

Ivana Kutá Smatanová

Contents

List of abbreviations

3D –tree-dimensional

ADP –atomic displacement parameter

Ala, A –Alanine

APC –activated protein C

aPTT –activated partial thromboplastin time

Arg, R –Arginine

Asn, N –Asparagine

Asp, D –Aspartic acid

BMDMs–bone marrow-derived macrophages

Cys, C –Cysteine

DLS –dynamic light scattering

FBS –fetal bovine serum

FVIIa –factor VIIa

fXa –factor Xa

fXIa –factor XIa

Gln, Q –Glutamine

Glu, E –Glutamic acid

Gly, G –Glycine

His, H –Histidine

IFN-γ –Interferon-γ

IL-6 –Interleukin-6

Ile, I –Isoleucine

Leu, L –Leucine

Lys, K –Lysine

MAD –multi-wavelength anomalous dispersion

MALS –static/multiangle light scattering

Met, M –Methionine

MR –Molecular replacement

NO –nitric oxide

OVA –ovalbumin (

PDB –Protein data bank

PEG –polyethylene glycol

Phe, F –Phenylalanine

Pro, P –Proline

PT –prothrombin time

RCL –reactive centre loop

s3A –3rd strand of β-sheet A

s3C –3rd strand of β-sheet C

s4A –4th strand of β-sheet A

s4C –4th strand of β-sheet C

s5A –5th strand of β-sheet A

SAD –single-wavelength anomalous dispersion

Ser, S –Serine

SFX – serial femtosecond crystallography

Thr, T –Threonine

Trp, W –Tryptophan

TT –thrombin time

Tyr, Y –Tyrosine

Val, V –Valine

XFEL – X-ray free-electron lasers

1 Introduction

1.1 Prologue and aims of the research

This thesis briefly introduces tick physiology and the related characterisation of serine protease inhibitors – serpins. Serpins are interesting proteins with even more unique inhibition mechanism that is discussed in this chapter. Serpins are one of the largest superfamilies of structurally conserved protease inhibitors. They are ubiquitously distributed in nature and have many regulatory functions making them one of the most studied protein families. During the evolution, many serpins lost their inhibitory function and work as chaperons or storage proteins (Fig. 1).
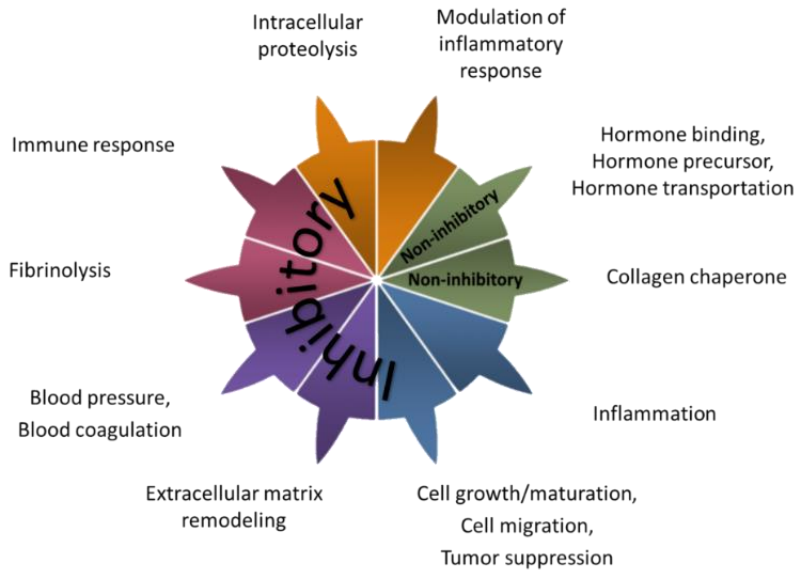
Figure. 1   Multiple regulatory functions of serpins (Modified from: [1,2])

Moreover, serpin's possible structural conformations involved in inhibition mechanism or results of mutations or serpinopathies are presented. The importance of serpins is underlined by the existence of these serpinopathies, the diseases that are caused by serpin dysfunction or deficiency. For example, Emphysema, Cirrhosis, Angioedema, Hypertension, and even familial Dementia

are caused at least in part by serpin dysfunction. Next, a brief description of known *Ixodes ricinus* tick serpins is stated.

Many recent studies focus on the revelation of serpin's structure and their unique and highly intriguing mechanism of inhibition. Macromolecular X-ray crystallography is the most common experimental method used in serpin structure determination. X-ray crystallography can establish the three-dimensional atomic structure of macromolecules. The most used methods for crystallization of macromolecules are described as well. Further, the basics for understanding the principles of X-ray crystallography from data measurements to structure deposition are specified. X-ray crystallography is frequently used for the investigation of the structural characteristics of biomolecules, and their complexes with other molecules. Moreover, it is used for elucidation of basic biochemical mechanisms and disease pathways important for biology, medicine and related sciences. In this work, the structural characterization of representatives of serpins from *I. ricinus* ticks is discussed in detail.

1.2 *Ixodes ricinus* tick physiology

Hard ticks (*Ixodidae*) are hematophagous parasites that attack animals as well as humans and feed on them for several days [3]. *I. ricinus* ticks are waiting for a host and after attaching to them, they are distributed by and fed on a broad range of warm- or cold-blooded vertebrate hosts. *I. ricinus* has a three-host cycle that lasts two to three years. They need to feed on blood for proper progress to the next stage, from larvae to nymphs (found on small mammals and birds) and from nymphs to adult stage (found on larger hosts) [4]. The adult female is after completing the feeding. Ticks are opportunistic parasites that feed on humans, at any life stage, when the opportunity arises, which makes them the effective vectors of tick-borne diseases [5]

The occurrence of *I. ricinus* is very wide throughout Europe [6], which increases the importance of ticks as vectors for many pathogenic organisms of medical and veterinary importance [7,8]. Some of the most important pathogens with *Ixodes ricinus* as a vector are *Borrelia burgdorferi s.l.* causing Lyme borreliosis and tick-borne encephalitis virus, but many others are also reported, such as *Rickettsia helvetica* and *Rickettsia monacensis* responsible for spotted fever rickettsiosis, *Babesia divergens* and *Babesia microti* causing babesiosis, *Anaplasma phagocytophilum* responsible for human granulocytic anaplasmosis, *Francisella tularensis* causing tularaemia, and also Louping ill virus and Tribec virus [3].

The host defense mechanism at the bite site is suppressed by many pharmacologically active molecules that are secreted by tick salivary glands [3]. These salivary molecules are essential for successful feeding because inhibit host proteases involved in physiological processes such as hemostasis and immune responses [9]. An extensive understanding of tick protease inhibitors and their physiological roles in facilitating blood-feeding are crucial for clarifying how ticks overcome the host defenses and could reveal the potentiality of these molecules for tick control use [10].

Tick saliva contains many inhibitory molecules that belong to the Kunitz-type family, serpin family, cystatin families and several small peptide inhibitors

[11]. The next paragraphs discuss a large family of serine protease inhibitors - serpins with a unique inhibition mechanism.

1.3   Serpins

Serine protease inhibitors-serpins are a group of ancient proteins widely distributed in nature [12]. Serpins function as serine protease inhibitors but during the evolution, some serpins lost their inhibitory function and serve as molecular chaperones (Heat shock serpin 47), tumor suppressors (Maspin), and hormone transporters (Cortisol-binding globulin) or storage proteins (Ovalbumin) [13]. Inhibitory serpins vary in functions according to their specificity. Serpins play crucial roles during some vital processes such as fibrinolysis, tumor suppression angiogenesis, inflammation, blood coagulation, and complement activation [13]. The importance of serpins is stressed also by serpinopathies, diseases caused by serpin dysfunction or deficiency. Many of today's well-known diseases, such as emphysema, cirrhosis, angioedema, hypertension and familial dementia, are associated at least partially with serpin dysfunction [14]. This makes serpins interesting candidates for drug development and knowledge of detailed serpin structure is necessary for that.

1.3.1   Mechanism of inhibition

Serpin's mechanism of action is also known as suicide-mechanism because at the end of successful inhibition both target protease and serpin are inactivated [12]. Serpins are capable to inhibit multiple enzymes, but these target proteases are usually part of the selected biological mechanism and contribute to the modulation of the selected proteolytic cascade [2]. Serpins operate by a two-step cascade catalytic mechanism that is "sequentially activated" [15]. Firstly, serpins imitate the structure of the usual protease substrate, therefore protease active site recognizes the bait sequence of the serpin in the reactive centre loop (RCL) and binds to it by forming the Michaelis-Menten complex (see later Structure and structural conformations of serpins), exactly forming an acyl-enzyme intermediate. Then enzyme recognition site (P1-P1' scissile bond) [15,16] of the serpin located on its RCL (Fig. 2) is cleaved by a protease and

conformational change is occurred, preventing protease to finish catalysis. The whole protease is then translocated to the opposite side of the serpin and RCL is inserted as a new additional β-strand of β-sheet A [15].
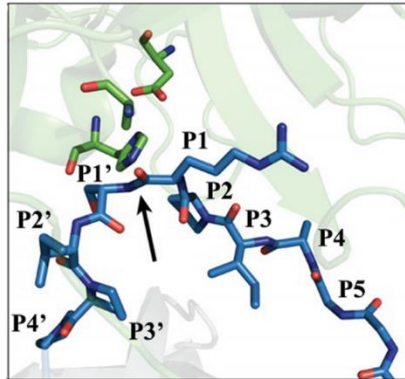
Figure. 2   Interaction between the serpin RCL (blue) and target protease active site (green) in the Michaelis-Menten complex. Residues of RCL are marked according to nomenclature developed by Schechter and Berger, 1967 [16](Modified from [17])

The protease is inhibited by a formation of a covalent bond between the main-chain carbonyl carbon of the P1 residue of the RCL and the protease active site, linking two molecules together and disrupting the protease active site [15]. According to the reaction rate of the RCL insertion into β-sheet A, the serpin can form a covalent complex, when a reaction is quick enough, or otherwise is consumed as protease substrate and released in its inactive, cleaved form (Fig. 3) [18]. This one-to-one inhibitory mechanism is very effective even though requires a large number of proteins and is an irreversible and complex process [19].
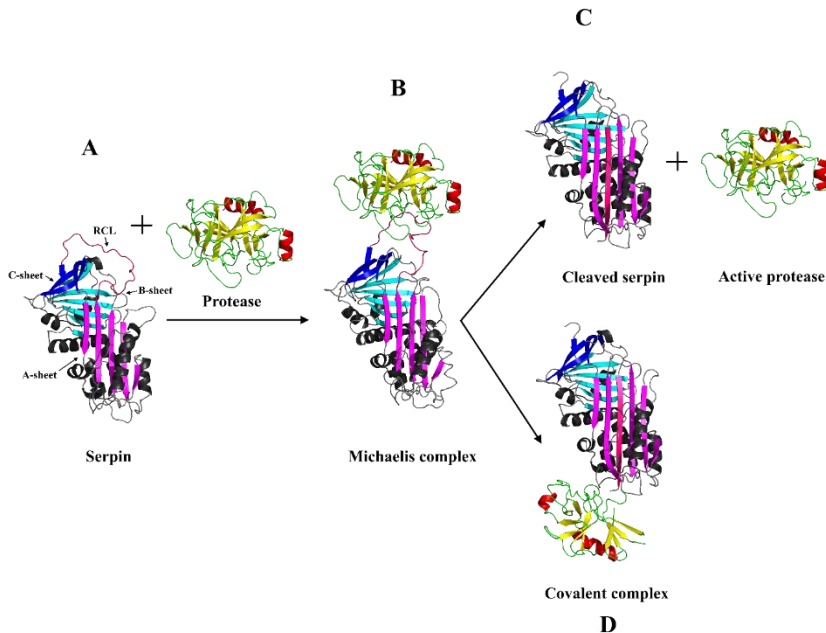
Figure. 3   Serpin possible conformations are shown during the inhibitory pathway of trypsin. From left to right: **A** Native serpin conformation – α1AT (PDB entry 1QLP) in reaction with protease; **B** Michaelis-Menten complex between Serpin 1 and trypsin (PDB entry 1I99); **C** Cleaved serpin conformation – α1AT (PDB entry 7API); **D** Covalent complex between α1AT and trypsin (PDB entry 1EZX). Secondary structure elements are coloured as α-helices are grey; β-sheet A is magenta, β-sheet B is cyan and β-sheet C is blue; the RCL is highlighted as hotpink. (author's work)

For correct inhibitory function, two structural features of the RCL namely mobility and length, as well as sequence-based motifs, protease recognition sequence and serpin exosites are important.

- The loop mobility and insertion are affected by a hinge region (P15-P9 N-terminal sequence before the cleavage site) that is alanine-rich and well-conserved among inhibitory serpins. In this region, the uncharged residues are preferable for enzyme inhibition; charged residues do not affect the recognition of proteases only their inhibition [20].

- The length of the RCL can affect inhibition function, especially the N-terminal part that is inserted inside the β-sheet A during inhibition. The stability of the formed complex can vary by modifying the RCL length. The extension of RCL, by adding one or two residues, dramatically decreases the complex stability, unlike its shortening, by deletion of one

or two residues, which doubled the stability but decreases serpin inhibitory efficiency [21].

- The sequence motif of the RCL is variable between serpins and corresponds to its specific inhibitory functions. The P1-P4 sequence is responsible for the recognition and efficient cleavage. The inhibition of defined proteases can be change to inhibition of different proteases by changing the residues in this site [22].

- Exosites are away from the RCL cleavage site, and their role is as secondary binding sites that improve serpin specificity. This is done by assisting during the docking of the target protease for its binding to P1 residue. Further, the exosites located at extended N- and C-termini help to increase inhibition by binding to the target protease. Moreover, exosites allow interaction with cofactors that boosts the inhibition rate and improve its specificity [2].

1.3.2 Structure and structural conformations of serpins

All serpins include ~380 residues (40-100 kDa) forming one conserved core domain. This domain is made up of three β-sheets (A, B and C) and 7-9 α-helices (hA – hI). Another typical characteristic is the presence of an exposed, extended loop called RCL that is important for serpin inhibitory function. The RCL has formed of ~17 residues that are located between β-sheet A and β-sheet C. The serpin's fold uniformity is confirmed by X-ray structures of many serpins (Fig. 3) [23]. Different structural conformations of serpins were found accordingly to the phase of the inhibition reaction and some improper conformations were caused by mutations.

Native conformation

Native structural conformation is required for inhibitory activity and represents intermediate to a metastable or stressed state of the protein. Thus, the native conformation is an exception to Anfinsen´s conjecture that predicts the folding of proteins to a single structure with the lowest free-energy state. In

native conformation, the RCL is exposed to target protease and serpins are more vulnerable to mutations or misfolding (Fig. 3A, Fig. 4) [13,23].

Some regions of serpin are involved in protease transfer during the inhibition mechanism. The hinge region of RCL is an N-terminal sequence before the cleavage site (P15-P9), which promotes flexibility and its insertion. The conserved sequence of inhibitory serpins is significant for inhibition compared with non-inhibitory serpins that have hinge regions less conserved. The β-sheet A carry two regions, the upper breach and bottom shutter regions, both essential for controlling the conformational change of RCL insertion. The gate region is the β-turn of the β -sheet C which is important because RCL needs to pass around the gate region for full insertion into β-sheet A [23].
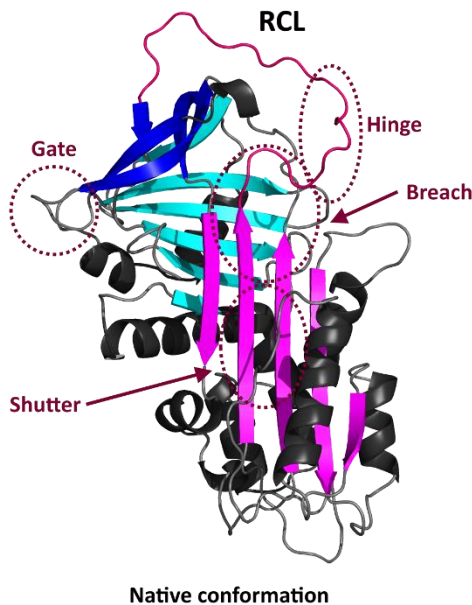
Figure. 4   Native, active conformation of serpin α1AT (PDB entry 1QLP) with marked regions important for inhibition mechanism. Secondary structure elements are coloured as α-helices are grey; β-sheet A is magenta, β-sheet B is cyan and β-sheet C is blue; the RCL is highlighted as hotpink. Regions of serpin assisting during the inhibition are in dashed ellipses. (author's work)

Michaelis-Menten complex

Michaelis-Menten complex is a noncovalent complex between serpin and protease formed through the residues of a scissile bond (P1-P1´) of the RCL and the target protease (Fig. 3B). The target protease binds to the active site of the RCL scissile bond then cleaves the peptide bond and forms a covalent ester linkage with main-chain carbonyl carbon of the P1 residue of serpin (Fig. 5). Due to this cleavage, the serpin is released from its metastable conformation by inserting RCL into β-sheet A. During this process, the protease is translocated over 70 Å and the active site of protease is distorted and thus deactivated. This transition is used by serpins to target protease inhibition and the cleavage of the scissile bond of RCL and subsequent translocation of RCL are crucial steps for protease inhibition. From this point, according to the success of the inhibitory mechanism, there is a possibility of two different conformations/serpin products: cleaved conformation or covalent complex (see next) [2,12].
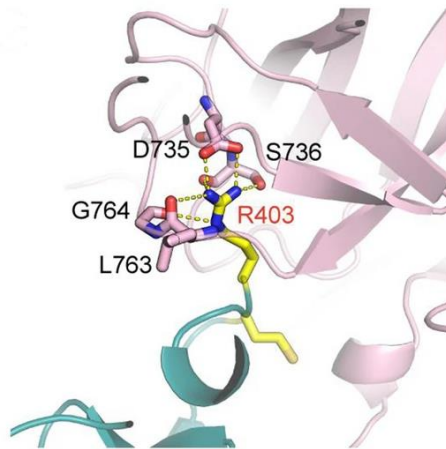
Figure. 5   Detailed view of Michaelis-Menten complex of μPlm and a2AP from homology modelling. The active site of protease (light pink), catalytic triad - S741, H603 and D646, form covalent linkage with R403 and M404 (yellow) of inhibitor RCL (green) (Modified from [24]).

Covalent complex

The covalent complex is the final complex of covalently linked protease to serpin, where the protease is trapped at the acyl-intermediate rather than the tetrahedral intermediate stage of the catalytic cycle. This is done by the release of the newly formed N terminus (from P1 residue) and insertion of presently cleaved RCL into β-sheet A (Fig. 3D). This prevents the release of the protease from the complex. Nevertheless, the protease may be released after hours or days [23].

Cleaved conformation

The cleavage of serpins is a result of proteolysis and the RCL is inserted into the center of β-sheet A that dramatically increases in thermal stability. This conformational transition to the most stable state is called the "stressed to relaxed transition" (Fig. 3C) [13,25].

Serpins can be inactivated through polymerization and transition to the latent state. The shutter region, the region underneath β-sheet A in the center of the serpin, together with the breach region, above the shutter region (Fig. 4), are two regions that after their destabilization (through mutations) prefer the transition to latent state or polymerization of serpins [26].

This state is characterised by the spontaneous insertion of the RCL into β-sheet A, like a cleavage state, which is blocking the inhibition function – auto-inactivation. There is no cleavage of the scissile bond, and this spontaneous transition can be an important control mechanism of serpin polymerization (docking of RCL of one serpin into β-sheet A of another) (Fig. 8) or mutation that results in serpin deficiency or serpinopathy. Serpins can be converted from latent conformation to the active state by denaturation and refolding (Fig. 6) [12].
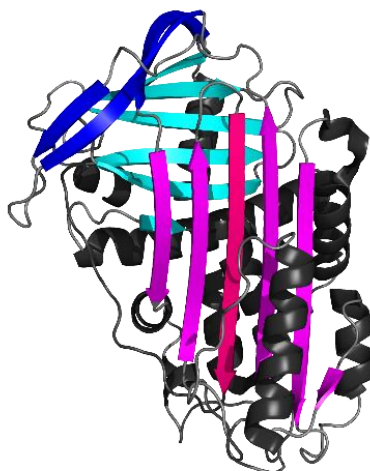
Latent conformation

Figure. 6   Latent conformation of PAI-1 (PDB entry 1DVN). Secondary structure elements are coloured, α-helices are grey; β-sheet A is magenta, β-sheet B is cyan and β-sheet C is blue; the RCL is highlighted as hotpink. (author's work)

δ-conformation and polymerization

This inappropriate conformational change is a result of mutation. The serpin structure has the cleaved RCL partially inserted into the top part of β-sheet A and the bottom part is filled with the last turn of α-helix F that partially unwinding to form a central part of β-sheet A. Some hypothesis was claimed for the origin of this conformation and thus that it's the intermediate state in forming

cleaved form from native conformation or as one of the steps in polymeric formation (Fig. 7) [26,27].

δ-form

Figure. 7 δ-conformation of antichymotripsin (PDB entry 1QMN). Secondary structure elements are coloured, as α-helices are grey; β-sheet A is magenta, β-sheet B is cyan and β-sheet C is blue; the RCL is highlighted as hotpink/lightpink. (author's work)

From the structural data known to date, we can state that serpins need conformational change for a controllable inhibition mechanism. However, many uncontrolled conformational changes such as dimerization (domain swapping) (Fig. 8) or polymerization led to the development of serpinopathies [26].
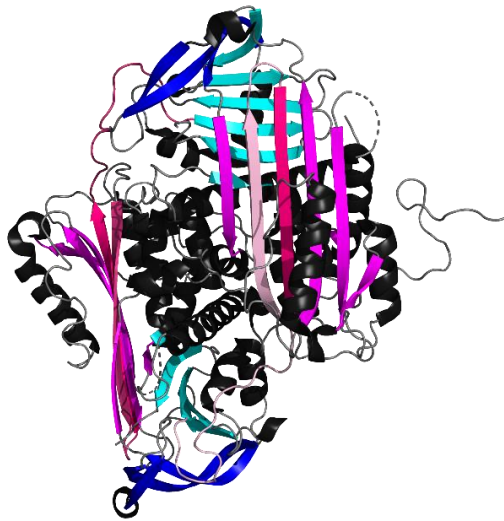
Serpin dimer

Figure. 8 Dimerization of Antithrombin-III (PDB entry 2ZNH). The insertion of the RCL and 5th strand of β-sheet A of one serpin into β-sheet A of another serpin is marked by a lightpink colored loop and 5th β-strand following on hotpink colored RCL. The insertion from another molecule is colored by warmpink loop and 5th β-strand. Secondary structure elements are coloured, as α-helices are grey; β-sheet A is magenta, β-sheet B is cyan and β-sheet C is blue. (author's work)

1.3.3 *Ixodes ricinus* serpins

I. ricinus, which form a family of hard ticks (*Ixodidae*), have a wide geographical distribution mainly in Europe that points toward its resistance to various environmental conditions. Tick is a long-time bloodsucking parasite in all developmental stages (larval, nymphal and adult form), which modulate host defense mechanism due to the feeding process. The process takes several days until repletion. This is the reason, why tick saliva contains antihemostatic and immunomodulatory molecules (such as serpins) that irreversibly inhibit their targets. Serpins modulate host responses during biting, act on hemostasis, and modulate immune responses and angiogenesis. Serpins can indirectly influence the spreading of pathogens causing public and veterinary health problems such as Tick-borne diseases such as Lyme disease, Helvetica spotted fever, Tick-borne meningoencephalitis, and Babesiosis or Tick paralysis [28].

The high conservation of serpins across tick species was observed together with the fact that tick serpins are species-specific inhibitors [29]. *I. ricinus* ticks exhibit at least 36 serpins, with different functional specificities, from which only two (IRIS and IRS-2) were fully described functionally and only IRS-2 structurally [9].

IRIS, *Ixodes ricinus* immunosuppressor, has many important functions in various processes during tick feeding. The peak expression of the protein was observed on the fourth day of feeding when female ticks ingest the largest amount of blood. The most notable function is thrombin inhibition when almost 30% is inhibited. The additional function is hindering fibrinolysis, preventing platelet adhesion but IRIS is not a powerful anticoagulant. The immunosuppressive properties are supported by strong inhibition of elastase-like proteases and by inhibition of T lymphocyte proliferation [30]. Another function is the suppression of the tumor necrosis factor TNF secretion by binding to monocytes or macrophages and its anti-hemostatic properties. Protein proteolytic activity is independent of its anti-inflammatory properties that are mediated through the exosite domain. IRIS can also modulate innate and adaptive immunity [31].

IRS-2, *Ixodes ricinus* serpin-2 (Fig. 9), expression was observed from the second day of feeding with a peak on the sixth day that indicates its important role in the early stages of feeding. The main function is the inhibition of tissue swelling and neutrophil migration into inflamed tissue. IRS-2 inhibits both chymase and cathepsin G, of neutrophils and mast cells respectively, during inflammatory responses. Moreover, IRS-2 inhibits platelet aggregation induced only by cathepsin G and neutrophil migration. IRS-2 is exhibiting anti-chymotrypsin activity and affects thrombin-induced platelet aggregation [32].

Figure. 9   Crystal structure of cleaved *I. ricinus* serpin IRS-2 (PDB entry 3NDA). Secondary structure elements are coloured as α-helices are cyan; β-sheet A is red, β-sheet B is salmon and β-sheet C is chocolate; loops are magenta. Inserted RCL into β-sheet A is highlighted by firebrick colour. (author's work)

Serpins with more polar or basic charged amino acids at the P1 site of RCL seem to be more conserved [29]. IRIS belongs to the group of serpins with methionine and cysteine in their RCL (Fig. 10) [33]. On the other site, the IRS-2 has tryptophan in its P1 site (Fig. 10) [32]. Some serpins have extended termini that can be related to their additional functions [26].


```
Iris    --MEASLSNHILNFSVDLYKRLKPSGKDTAGNVFCSPFSIAAALSMALAGARGNTAKQIA 58
IRS-2   MQEEAKLTKANNRFGLRLLRA-LP--SGPEKNVFFSPYSVSTAMGMAFAGARGQTQQELS 57
           **.*::   .*.: *  :  .   *** **:*:::*:.**:*****:* ::::

Iris    AILHSNDDK-----IHDHFSSFLCKLPSYAPDVALHIANRMYSEQTFHPKAEYTTLLQKS 113
IRS-2   QGLGFSDVDLTDAGVLDAYTHHTERLKSTPSNSTLDVANAAAIQRTLALLNSYESALQSS 117
           *  .* .  : * ::. :* *  : :*.:** ::*:   .* : **.*

Iris    YDSTIKAVDFAGNADRVRLEVNAWVEEVTRSKIRDLLAPGTVDSSTSLILVNAIYFKGLW 173
IRS-2   FGAELHKVDFAGEPQAAVDFVNNWVKRKTHDKIEKLFN-EPLDPDTLLVLLNAIYFKGEW 176
           :.: :: *****: :  .  ** **:. *:.**.,*:   :* .* *:*:*******  *

Iris    DSQFKPSATKPGDFHLTPQTSKKVDMMHQKGDFKMGHCSDLKVTALEIPYKGNKTSMVIL 233
IRS-2   NTAFVKEHTEKRQFFNGGVTPVEVDTMRLEARIKYRFFDDLQVEVVELPYRGLDYTMAIL 236
           :: *  .*:  :*.    *  :** *:: .:*  .  ,**:* .:*:**:* . :*.**

Iris    LPEDVEGLSVLEEHLTAPKLSALLGGMYVTSDVNLRLPKFKLEQSIGLKDVLMAMGVKDF 293
IRS-2   LPKENTGVEGLKQNLTIDRFQNYLSDLR-ERKITVLLPKFKLETKYSLKAPLQSLGIKQI 295
           **:: *:. *:,*** :: .  .:  :  ****** .:**  *  :*:*:*:.

Iris    FTSLADLSGISAAGNLCASDIIHKAFVEVNEEGTEAAAATAIPIMLMCARFPQVVNFFVD 353
IRS-2   FESGADLSGIND-GSLRVSAVEHKAVVEVNEEGTVAAATTGVVIVP-YSLGPEPVVFRVD 353
           * * ******. *.*  .*.: ***.*******  ***.*:  *:  *: * * **
                                         ┌──────────────────────────┐
                                         └──────────RCL─────────────┘

Iris    RPFMFLIHSHDPDVVLFMGSIREL 377
IRS-2   HPFLFFIRNTRTDDIFFVGQVNKL 377
           :**:*:*:.  * ::*:*.:.:*
```

Figure. 10   Sequence alignment of IRIS and IRS-2 serpins. This alignment was obtained using the ClustalW algorithm. Dots indicate similar amino acids, and stars indicate identical amino acids. Indents indicate gaps. The RCL of both proteins is in the square. (author's work)

2 Material and Methods

2.1 X-ray crystallography

X-ray crystallography is a tool to study high-resolution three-dimensional structures of individual molecules by integrating the elements of physics, chemistry, biology, medicine, and mathematics and computing. Crystallography can be used to identify the structure of either small molecules, proteins or large macromolecular complexes, studied in this thesis. Contrary to former years, when an extensive effort was invested to identify mainly secondary structures [34], in modern biology an atomic-resolution 3D structure is obligatory for a better understanding of protein functions.

2.1.1 Protein crystallization

Before initiating any crystallization experiment, various molecular biology techniques are required. Firstly, it is necessary to clone the gene of interest into an expression vector and subsequently express and purify the protein. These steps are critical for every crystallization experiment since the high level of protein homogeneity and purity is the key step to obtaining crystals diffracting to a high resolution. The protein is dissolved in an appropriate stabilizing solution for obtaining a stable protein suitable for crystallization [35].

Advances in the molecular-biology field improve and accelerate the process of protein sample preparation not only for protein crystallization. Most nowadays obtained protein crystals are expressed as recombinant proteins in bacterial, yeast, insect, or mammalian cells [36,37]. Another breakthrough is the use of autotrophic strains that are able to incorporate seleno-methionine into recombinant protein. This is useful in the case when there is missing information about a phase problem that is solved by the multiwavelength anomalous dispersion (MAD) method. The use of different tags or fusions with highly soluble proteins is used to enhance protein isolation and purification by various chromatographic techniques [36]. Furthermore, the combination of size-exclusion chromatography with static/multiangle light scattering (MALS) and dynamic light scattering (DLS) brings information about protein sample

homogeneity and its oligomeric state [37]. This boosts the chances of obtaining the diffraction-quality protein crystals because the crystalized sample is as pure as possible and in a single oligomeric state.

An appropriate condition forming at least microcrystals must be detected in any crystallization experiment and then optimized to improve the crystal quality [38]. The process of crystal formation involves spontaneous clustering of dissolved molecules and ions into the crystal nuclei, which is well-known process for small-molecule growth. The chemical bonds formation and molecule interactions provide an assembly of molecules into lattices, which has a stabilization role in crystal formation during the transportation of solution away from the equilibrium [39]. Moreover, a solution to protein crystallization must be stabilizing properties when solvent molecules fill gaps in the protein structure because intermolecular interactions supporting the crystalline lattice are weak [40]. The solvent can fill 25-90% of the crystal, surrounding the molecules [41].

After the addition of precipitants (e.g., neutral salts or polymers) to the solution with the protein, the solution is brought to the supersaturated state. The ordered state of nuclei formation is slightly visualized by aggregates at the beginning of the crystallization process [38]. The supersaturated state must be reduced to achieve a suitable number of crystals; otherwise, a large number of nuclei will be formed, which results in the creation of many small crystals [42]. Protein supersaturation can be adjusted by changing the protein concentration, precipitant concentration, pH or by adding specific additives and also temperature can be changed [43].

Since there is not yet available to predict the specific crystallization conditions for obtaining crystals of certain proteins, the screening of a large number of conditions is necessary. The development of the crystallization robots that can test thousands of conditions in a short time as well as saves the protein sample by dispensing low-volume drops; and the use of "crystal hotels" customized to grow crystals at defined temperatures while detecting newly grown crystals by automated crystal recognition software, facilitated the crystallization experiment [37].

2.1.2   Crystallization methods

Several different methods are used in protein crystallization. The main principle of these methods is to get a protein solution into the supersaturation state, which is essential for the formation of stable crystal nuclei.

Common crystallization methods are briefly described below:

- Vapor diffusion – the method is based on solvent (water) evaporation from the drop with protein to reservoir solution thus from lower concentration to higher concentration of precipitant. This process ensures that the drop with the protein solution goes to equilibrium. Three crystallization techniques, namely sitting drop, hanging drop and sandwich drop vapor diffusion (Fig. 11) are based on this method. The most common technique is sitting drop vapour diffusion, which is frequently used for crystallization with robots [38].
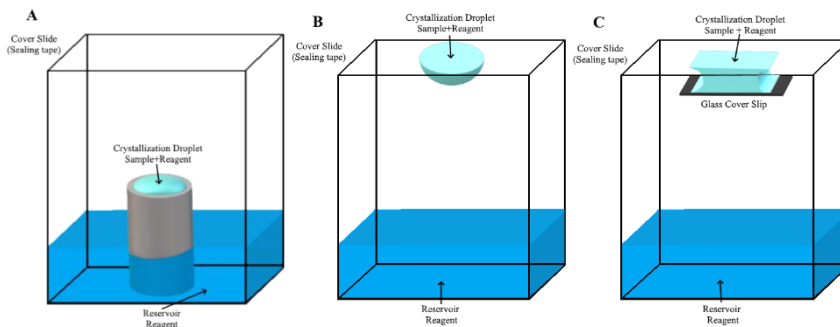
Figure. 11   Sitting drop vapour diffusion (A) and hanging drop vapour diffusion (B-hanging drop, C-sandwich drop) methods visualization. (author's work)

- Microbatch crystallization – this method is based on the same principle as vapor diffusion methods with one exception that the drop with protein is placed under mineral (paraffin) oil, silicon oil or their mixture and water can evaporate from the drop (Fig. 12). This evaporation ensures a small achievement of supersaturation, which has a benefit for crystal formation [38].
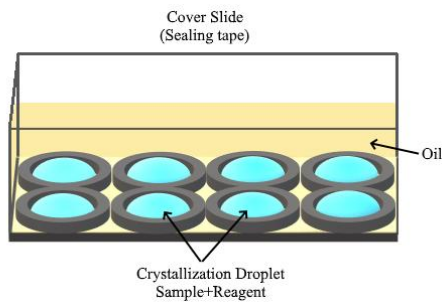
Figure. 12   Microbatch under oil crystallization (author's work)

- Microdialysis – this method is based on sample separation from the precipitant solution with the use of a semipermeable membrane located on the dialysis button immersed under the precipitant solution (Fig. 13). The membrane is permeable to water and small molecules. The semipermeable membrane can have different sizes of pores ensuring that the protein of interest stays inside the microdialysis chamber of the dialysis button [38].
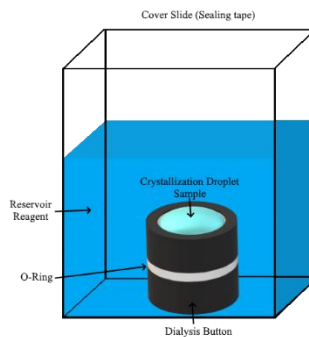
Figure. 13   Microdialysis crystallization method with described parts (author's work)

- Free-interface diffusion – the glass capillaries are commonly used to create a gradient of concentrations between protein and precipitant solution that are pipetted from each side. These two solutions diffuse from each other (Fig. 14). This method requires only a small volume of protein solution (up to 1μl), which is helpful when there is low amount of protein [38].

Precipitant solution   Protein solution
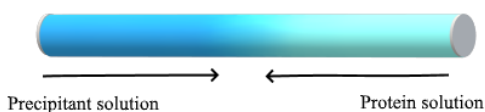
Figure. 14   Capillary free-interface diffusion visualization (author's work)

The main goal of each crystallization experiment is to obtain a few large-homogenous-high-quality crystals, but in most cases, the crystals are size-limited, soft and crushed easily. Moreover, macromolecule crystals are often surrounded by water thus their biochemical properties are the same as for solvated molecules. Crystal polymorphism is also very frequent and means the formation of multiple crystal forms of the same protein [38].

A formed protein crystal is a homogeneous solid unit with an arranged structure of atoms. Following that many molecules in the crystal are identical creating an array of identical points forming the imaginary structure lattice. A parallelogram of four chosen lattice points creates a unit cell with a certain dimension of three lengths (a, b and c) and three angles (α, β, and γ). The unit cells in an ideal crystal are in the same orientation. The asymmetric unit is the smallest unit that forms a unit cell. The asymmetric unit fills a unit cell with different symmetry orientations, which means that the asymmetric macromolecules rotate, translate or screw axes to fill one unit cell. The possible symmetries (Fig. 15) of the protein unit cell are summarized in 65 space groups [44].
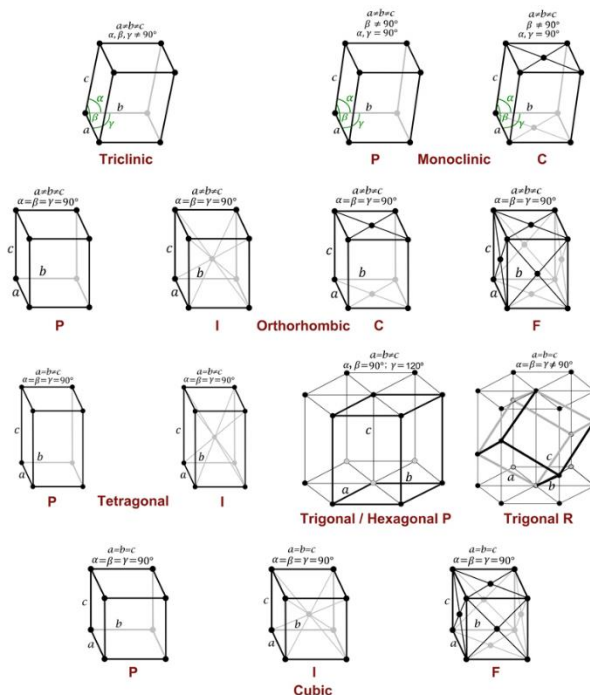
Figure. 15   The symmetry of crystals. The black and grey points of lattices represent places in crystal space that are indistinguishable from each other. The letters represent different lattice types: P-primitive, C-centred in the faces perpendicular to the cell c axis, I-centered in the body of the cell, F- centered in all facets of the cell, R- primitive, identical cell axes and cell angles, or hexagonal two times body-centred (From https://www.xtal.iqfr.csic.es/Cristalografia/parte_03_4-en.html)

2.1.3   X-ray diffraction and data processing

Nowadays, advances in the X-ray crystallography field made the process of structure solving much easier. Some of the recent updates are brighter synchrotron X-ray sources, faster detectors, and automatization of refinement pipelines [36].

The theory behind X-ray diffraction

The X-ray electromagnetic wavelength interacts with an atom's electrons, and this phenomenon is needed for atom visualization and for capturing their distances from each other. Once the beam of X-rays is physically interrupted by a crystalline (periodically ordered molecular object) sample, X-rays scatter [34,40] and the diffraction is initiated when the wave incidents on

electrons. The electrons subsequently oscillate resulting in the emission of radiation with the same frequency as the incident radiation. Each crystal generates a unique diffraction pattern for an X-ray sensitive detector. The crystals can be damaged after prolonged exposure to radiation consequently the measurement must be done carefully [40]. Currently, most of the solved crystal structures are measured at cryotemperatures (~100 K) because it decreases the radiation damage and prevents dynamic disorder in the protein as protein is "fixed" in a crystal lattice. Moreover, the overall alpha-amino acid fold is temperature independent and thus cryotemperature measurements are not affecting the results, on the contrary, more defined side-chain alternative conformations and more ordered water molecules were found in structures solved at cryotemperatures [37,45].

X-ray sources and detectors

The storage ring of X-ray beams at synchrotrons are the main sources where diffraction data measurements take place. Synchrotrons are capable to distribute the X-ray beams with powerful intensity, a high degree of collimation and the possibility to adjust the wavelength. On the contrary, the "home sources" have rotating anodes as the bases that only deliver wavelength typical for molybdenum (0.71 Å), copper (1.54 Å), or chromium (2.29 Å) [36]. The breakthrough is an invention of X-ray free-electron lasers (XFELs) that are extremely intense X-ray flashes, which immediately destroy a sample of microcrystals in a stream of mother liquor measured in a vacuum but result in a large number of diffraction images [46]. This method is called serial femtosecond crystallography (SFX) and represents a pioneering method in a new era of X-ray crystallography. Every X-ray data measurement needs a fast and accurate detector for a recording of refection intensities. The majority of synchrotrons are equipped with single-photon counting pixel detectors that are able to rapidly transfer measured data in milliseconds and are more position sensitive devices [37,47,48].

New improvements of synchrotron beamlines are sample-mounting robots that are used for safer crystal transfer to the goniometer with constant crystal cooling. Moreover, the possibility of remote data collection is in high

demand because of the opportunity to comfortable measure data from "home" [37].

Data processing

The diffraction pattern is a summation of diffracted waves, especially their intensities and direction, from electrons and each electron [43]. These reflections are characterized as a complex number, structure factor (F_{hkl}) [49]. The X-ray diffraction pattern is recorded as an array of spots (thus raw data) and after that follows indispensable data processing [34]. Depending on the quality of structure uniformity and periodic arrangement of molecules in the crystal, the diffraction pattern exhibits a vast area of resolution and a higher level of atomic position details [38]. The data processing programs for diffraction data processing such as *HKL*2000 [50] *Mosfilm* [51] and *XDS* [52] are efficient to bring high quality and accurate data for further processing [36].

The main product of the X-ray diffraction experiment is an electron density map showing the atoms electrons distribution of the molecule/molecules in the asymmetric unit of measured crystal. The position of electrons is averaged position calculated from molecules present in the crystal. Some parts of the molecules such as loops are flexible, and this can lead to a poor to no electron density map in this region thus these parts are not modelled into a final model. The chemical composition of the crystal must be known before structure solving because the resulting map generally does not provide the precise location of defined atoms (C, N, O and even S) as they have a comparable number of electrons, not to mention H atom that is barely visible due to its single electron [53].

2.1.4 Structure solving strategies

To imagine simply what crystal structure means, we can say that it is averaged time and space of single macromolecule calculated from millions of molecules forming the crystal. The crystallization conditions are responsible for the limited dynamical behaviour of molecules within the crystal [37] and thus crystal structure captures the most "accurate" structural information.

Phasing

The structure factor can be expressed by Fourier transformation to obtain an electron density map:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl=-\infty}^{+\infty} |F_{hkl}| \, e^{i\varphi} e^{-2\pi i(hx+ky+lz)}$$

where $\rho(xyz)$ represents the electron density map at position xyz; V is the volume of the unit cell; hkl are Miller reflection indices and φ is a phase. The phase information is not obtained from the experiment. The so-called phase problem must be solved by using one or several methods such as molecular replacement, isomorphous replacement, and single- and multi-wavelength anomalous diffraction (SAD and MAD) [49].

Special-atom phasing

This phasing technique introduces a way how to solve the crystal structure of novel proteins. The initial reflection phases are predicted by the addition of a few heavy atoms into proteins. From them, the anomalous signal is measured either from single-wavelength anomalous diffraction (SAD) or from multi-wavelength anomalous diffraction (MAD), which is more challenging but also more efficient. The complexity of MAD lies in the necessity to collect many data sets at multiple wavelengths which leads to severe radiation damage to the sample and subsequently the small anomalous signal is deteriorating [54]. Thanks to the aforementioned selenomethionine incorporation into recombinant proteins the majority of novel structures are solved by SAD. Also, some structures were solved by SAD thanks to naturally occurring sulfur or phosphorus atoms in native proteins [55]. The *SHELX* suite [56] is one of the phasing software adapted to process very weak anomalous signals that provide better results [37].

Molecular replacement phasing

The molecular replacement (MR) method works in the case when the crystal asymmetric unit consists of more identical molecules when the novel molecule has some supposed search models that have already been solved. The majority of deposited structures are solved by the MR approach since there is a

high chance to find a search model among other deposited structures [36]. The most used MR programs are *MOLREP* [57], *Phaser* [58] and *AMoRe* [59]. Other programs employing different approaches are *ARCIMBOLDO* [60], which searches for small fragments of molecule in form of polyalanine helices, or *MrBUMP* [61], *BALBES* [62] and *phenix.mr_rosetta* [63] that use available search models from databases and then optionally can optimize the model automatically [36].

Model building

Model building was automatized and programs such as *ARP/wARP* (Automated Refinement Procedure) [64], *Buccaneer* [65], *PHENIX AutoBuild* [66], *SHELXE* [67] are based on the positioning of water molecules or structural fragments on the electron density maxima and then these water molecules or fragments, that are part of protein structure, are interpreted as connected amino acids residues. The remaining solvent molecules can be also interpreted around the protein in this step [68]. Mentioned programs are powerful enough to be able to build up the models even at low resolution. However, there is no model building suit capable to generate perfect structure and thus human eye review is so far irreplaceable [36].

Refinement

The refinement process is evaluated by R and R_{free}, which quantity the internal agreement of experimental and calculated data [36]. *REFMAC* [69] and *Phenix.refine* [70] are frequently used automated programs for structure model refinement. These use many methods to improve and to be able to obtain the best possible model [36]. Some of the mentioned methods are the maximum-likelihood minimization method, use of the prior information about experimental phases, "jelly body" refinement [69], and TLS motion refinement [71] or Rosetta-sampling [72]. One of the most recognized manual building programs is *Coot* [73].

Parameters of the refined model

The refined model obtained from the calculated electron density map carries out some important information about the structure. Every atom of the

modes is described by the atom type, the Cartesian coordinates, the occupancy, and the atomic displacement parameter - ADP (atomic vibration parameter / thermal parameter / B-factor) [53,74].

- The position of atoms is given by the Cartesian coordinates that are defined by three mutually perpendicular axes (x, y, z) provided with the Å as a unit [74].

- The occupancy introduces a perceptual representation (0.0-1.0) of an atom at its mean position. The occupancy can be detected as a dynamic disorder when the atom is occupying a certain position for defined time or as a static disorder when the occupancy of the atom at the same position of the crystal unit cells is observed [53,75].

- The ADP represents information about atomic motions and generates the displacement ellipsoids (spherical regions of atom displacement) [76]. The ADPs can be isotropic explained by a single B-factor number in Å, or anisotropic which expresses the weakening of scattering caused by atomic thermal motion and static disorder [75,77]. The crystallographic data are better in the case of anisotropic atoms when the ellipses are smaller and more spherical, thus the shape and size of ADPs can provide information on data quality [78].

Final model validation

The validation of the structural model is necessary during the refinement process and especially after the finalisation of the investigated model before its deposition. Ramachandran plot, which captures the main-chain torsion angles that must be qualitatively validated by programs such as *PROCHECK* [79] and *WHATCHECK* [79]. Main-chain torsion angles are normally not restrained during refinement and for that reason should be verified. Another thing that must be inspected is structure geometry, more precisely bond lengths and angles. Even though they are restrained, the clashes between fully occupied atoms could be detected that way. For these purposes, the *MolProbity* webserver [80] offers an all-atom calculation of clashscore, rotamer, and Ramachandran evaluations. Finally, yet importantly, things to validate are ligands or metals modelled into the final model that is sometimes troublesome to model in the right place. Several programs and webservers were created to check and correct the

placement of ligands and metals such as *TWILIGHT* [81] and *CheckMyMetal* [36,82].

Structure deposition

The Protein Data Bank (PDB) is an open access archive for the deposition of solved structures, especially by X-ray crystallography. The deposition process inspects the structure and diffraction data for ensuring the best possible result, even though the PDB can simply recommend the corrections as it is only the repository of the resulting models [36,83].

3   Results and Discussion

3.1   Iripin-3, a New Salivary Protein Isolated From *Ixodes ricinus* Ticks, Displays Immunomodulatory and Anti-Hemostatic Properties *In Vitro*.

ABSTRACT

Tick saliva is a rich source of pharmacologically and immunologically active molecules. These salivary components are indispensable for successful blood feeding on vertebrate hosts and are believed to facilitate the transmission of tick-borne pathogens. Here we present the functional and structural characterization of Iripin-3, a protein expressed in the salivary glands of the tick Ixodes ricinus, a European vector of tick-borne encephalitis and Lyme disease. Belonging to the serpin superfamily of protease inhibitors, Iripin-3 strongly inhibited the proteolytic activity of serine proteases kallikrein and matriptase. In an in vitro setup, Iripin-3 was capable of modulating the adaptive immune response as evidenced by reduced survival of mouse splenocytes, impaired proliferation of CD4+ T lymphocytes, suppression of the T helper type 1 immune response, and induction of regulatory T cell differentiation. Apart from altering acquired immunity, Iripin-3 also inhibited the extrinsic blood coagulation pathway and reduced the production of pro-inflammatory cytokine interleukin-6 by lipopolysaccharide-stimulated bone marrow-derived macrophages. In addition to its functional characterization, we present the crystal structure of cleaved Iripin-3 at 1.95 Å resolution. Iripin-3 proved to be a pluripotent salivary serpin with immunomodulatory and anti-hemostatic properties that could facilitate tick feeding via the suppression of host anti-tick

defenses. Physiological relevance of Iripin-3 activities observed in vitro needs to be supported by appropriate in vivo experiments.

INTRODUCTION

The European tick *Ixodes ricinus* (Acari: Ixodidae) is an obligate blood-sucking ectoparasite that transmits several medically important pathogens such as Lyme disease spirochetes from the *Borrelia burgdorferi* sensu lato complex and tick-borne encephalitis virus (1). The insertion of the tick hypostome and two chelicerae into host skin disrupts the surrounding tissue and capillaries, to which the host responds by activating a series of physiological defense processes including hemostasis and innate and adaptive immune responses (2–5). Cutaneous tissue injury and tick antigens are sensed by cells in the vicinity of the tick attachment site, such as keratinocytes, fibroblasts endothelial cells, mast cells, macrophages and dendritic cells (3). These cells release pro- inflammatory and chemotactic molecules that stimulate the recruitment of neutrophils and other immune cells to the area of tick feeding (3, 4, 6). Moreover, Langerhans cells and macrophages trap tick antigens and present them to T cells, which triggers T cell proliferation and ultimately results in the development of the acquired immune response (7). If unopposed, the host defense reaction rejects the tick via detrimental effects on tick viability and reproduction (8). Therefore, ticks surpass the host response by secreting hundreds of bioactive molecules via their saliva into the wound (9–11). Since these salivary molecules can target hemostasis and almost every branch of the immune response, they might be useful in the development of novel pharmaceuticals for the treatment of immune-mediated inflammatory diseases, hypercoagulable states, diseases associated with excessive complement activation, or even cancer (11–14). Moreover, tick salivary proteins represent potential targets for the development of anti-tick and/or transmission blocking vaccines (15).

Protease inhibitors form the largest functional group of tick salivary proteins (16). Based on their specificity, tick protease inhibitors can be divided

into inhibitors of cysteine proteases (e.g., cystatins) and inhibitors of serine proteases (e.g., Kunitz domain-containing proteins and serpins) (17). Serpins (serine protease inhibitors) are mid-sized proteins consisting of about 330–500 amino acids (18, 19) with a conserved serpin domain and an exposed region near the carboxyl-terminal end referred to as the reactive center loop (RCL) (20). Cleavage of the scissile P1- P1′ bond in the RCL by a target serine protease result in the formation of a covalent serpin-protease complex and permanent inactivation of both the serpin and the protease (18, 20).

Serpins have been identified in many species of hard-bodied ticks of medical and veterinary importance such as *Amblyomma americanum* (21), *Haemaphysalis longicornis* (22), *I. ricinus* (23), *I. scapularis* (24), *Rhipicephalus appendiculatus* (25), and *Rhipicephalus microplus* (26, 27). Some of the functionally characterized tick serpins have been shown to suppress the enzymatic activity of blood clotting factors (mainly thrombin and factor Xa) and consequently inhibit the intrinsic and common coagulation pathways (28–31). Tick serpins that inhibit thrombin and cathepsin G can block platelet aggregation triggered by these two serine proteases (30–33). In addition to anti-hemostatic activities, many of the functionally characterized tick serpins interfere with the host innate immunity, since they inhibit the enzymatic activity of mast cell and neutrophil serine proteases, reduce vascular permeability and paw edema formation, suppress neutrophil migration *in vivo* and attenuate the production of pro-inflammatory cytokines by activated innate immune cells, such as macrophages and dendritic cells (32, 34–37). Last but not least, tick serpins can modify the host adaptive immune response via suppression of T lymphocyte proliferation and inhibition of Th1 and Th17 cell differentiation (35, 37–40). A number of RNA interference and vaccination experiments have demonstrated the important role of tick serpins in successful completion of a blood meal by prolonging the feeding period, reducing engorgement weight, or resulting in higher mortality rates or impaired oviposition (41–45).

To date, only two serpins from the tick *I. ricinus* have been assigned functions: Iris (*I. ricinus* immunosuppressor) (38) and IRS-2 (*I. ricinus* serpin-2) (32). Due to possible confusion arising from the previously used abbreviation

IRS for *I. ricinus* serpins (32) (with insulin receptor substrates), we decided to name *I. ricinus* serpins Iripins (*Ixodes ricinus* serpins). Here we present the structural and functional characterization of Iripin- 3 (*I. ricinus* serpin-3). Iripin-3 primarily inhibited two trypsin- like serine proteases, kallikrein and matriptase. When tested in various *in vitro* assays, Iripin-3 displayed several distinct functions: it inhibited the extrinsic blood coagulation pathway, attenuated interleukin-6 (IL-6) production by LPS-activated bone marrow-derived macrophages (BMDMs), impaired the survival and proliferation of CD4$^+$ T cells, and suppressed the Th1 immune response. The presence of Iripin-3 protein in tick saliva suggests that this serpin could play a role at the tick-host interface by suppressing various aspects of the host defense to *I. ricinus* feeding. Further *in vivo* studies, however, are necessary to confirm herein presented results. Finally, we determined the crystal structure of cleaved Iripin-3 at 1.95 Å resolution.

MATERIALS AND METHODS
Animals

C57BL/6N mice were purchased from Velaz, Ltd (Praha- Lysolaje, Czechia). C3H/HeN mice and OT-II transgenic mice were obtained from Charles River Laboratories (Wilmington, MA). Mice were maintained under standard, pathogen-free conditions in the animal house facility of the Department of Medical Biology, Faculty of Science, University of South Bohemia in ČeskéBudějovice, Czech Republic. Guinea pigs utilized for *I. ricinus* feeding and a rabbit used for the production of anti-Iripin-3 antibodies were bred and maintained at the Institute of Parasitology, Biology Centre of the Czech Academy of Sciences (IP BC CAS), Czech Republic. All animal experiments were performed in accordance with the Animal Protection Law of the Czech Republic No. 246/1992 Sb. (ethics approval No. 34/2018) and protocols approved by the Ministry of Education, Youth and Sports of the Czech Republic (protocol No. 19085/2015-3) and the responsible committee of the IP BC CAS. Pathogen-free *I. ricinus* ticks were obtained from the tick colony maintained at the IP BC CAS.

Bioinformatics Analyses

The molecular weight and isoelectric point of Iripin-3 were computed by ProtParam (46). The presence of a signal peptide was predicted using the SignalP 4.1 server (47). The ScanProsite tool (48) was utilized to identify the serpin signature motif PS00284 as well as two other consensus amino acid motifs N- [AT]-[VIM]-[YLH]-F-[KRT]-[GS] and [DERQ]-[VL]-[NDS]-E-[EVDKQ]-G (26, 49). The reactive central loop together with the amino acid residue at the P1 site were determined based on the eight-residue pattern p17[E]-p16[E/K/R]-p15[G]-p14[T/S]-p13[X]-p12-9[AGS]-p8-1[X]-p1′-4′ [X] (26, 49).

NetNGlyc 1.0 (Gupta et al., unpublished) and NetOGlyc 4.0 (50) servers were used to predict potential N-glycosylation and O-glycosylation sites, respectively. To compare Iripin-3 with other known serpins, the Iripin-3 protein sequence was tested against the GenBank database of non-redundant protein sequences using BLASTP (51). Alignment of IRS-2 and Iripin- 3 amino acid sequences was conducted with ClustalW (52). Visualization of the alignment and addition of secondary structure elements were performed using ESPript 3.0 (53).

Crystal Structure Determination

The production of recombinant Iripin-3 in an *Escherichia coli* expression system is detailed in the Supplementary Materials. Crystallization experiments were conducted using the sitting- drop vapor diffusion technique, and the obtained crystals were used to collect X-ray diffraction data on the beamline BL14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin (54). The structure of Iripin-3 was solved by the molecular replacement method, in which the known structure of IRS-2 (Protein Data Bank (PDB) code 3NDA) (32) was used as a search model. The whole procedure of Iripin-3 structure determination, starting with crystallization and ending with structure refinement and validation, is described in detail in the Supplementary Materials. Complete data processing and refinement statistics are summarized in Supplementary Table 1. Atomic coordinates were deposited in the PDB under accession code 7AHP.

Phylogenetic Analysis

For the purpose of phylogenetic analysis, the amino acid sequences of 27 tick serpins and one human serpin were retrieved from GenBank. Accession numbers of these sequences are provided in Supplementary Table 2. Retrieved sequences were aligned and edited manually using BioEdit 7.2.5 (55). Evolutionary history was deduced from the protein sequences without a signal peptide by using the maximum likelihood method and Jones-Taylor-Thornton (JTT) matrix- based model (56). Initial trees for the heuristic search were obtained automatically by applying the neighbor-joining (57) and BIONJ (58) algorithms to a matrix of pairwise distances estimated using the JTT model, and then the topology with a superior log likelihood value was selected. The reliability of individual branches was determined by bootstrapping. Bootstrap values were calculated for 1000 replicates. Evolutionary analyses were conducted in MEGA X (59).

Iripin-3 Expression in Ticks

I. ricinus nymphs were fed on C3H/HeN mice for 1 day, 2 days, and until full engorgement (3–4 days). *I. ricinus* adult females were fed on guinea pigs for 1, 2, 3, 4, 6, and 8 days. Tick removal from host animals at given time points was followed by the dissection of nymphs and adult female salivary glands, midguts, and ovaries under RNase-free conditions. RNA was isolated from tick tissues using TRI Reagent (Molecular Research Center, Inc., Cincinnati, OH), and 1 mg of total RNA was reverse transcribed into cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche Applied Science, Penzberg, Germany) according to the manufacturer's instructions. Five-fold diluted cDNA mixed with FastStart Universal SYBR Green Master (Roche Applied Science) and gene-specific primers were used for the analysis of *iripin-3* expression in the Rotor-Gene 6000 thermal cycler (Corbett Research, Saffron Walden, UK). Cycling conditions were 95°C for 10 min followed by 45 cycles of 95°C for 15 s, 60°C for 10 s and 72°C for 30 s. The relative quantification of *iripin-3* transcripts in tick tissues was performed using the DDCt method (60). The *I. ricinus* gene encoding ribosomal protein S4 (*rps4*, GenBank accession

number MN728897.1) was utilized as a reference gene for the calculation of relative expression ratios (61, 62). Nucleotide sequences of forward and reverse primers as well as amplicon lengths are provided in Supplementary Table 3.

Presence of Iripin-3 in Tick Saliva

Polyclonal antibodies against Iripin-3 were produced in a rabbit injected subcutaneously with 100 mg of purified Iripin-3 in 500 ml of complete Freund's adjuvant. The first immunization was followed by another two injections of Iripin-3 in 500 ml of incomplete Freund's adjuvant at 14-day intervals. On day 14 after the last injection, the rabbit was sacrificed, and its blood was collected. Prepared rabbit antiserum to Iripin-3 was subsequently utilized for the detection of Iripin-3 in tick saliva by indirect ELISA and western blotting. The saliva was collected from *I. ricinus* ticks feeding for 6–7 days on guinea pigs as described previously (63). ELISA and western blot analyses are detailed in the Supplementary Materials.

Inhibition of Serine Proteases

Preliminary screening of Iripin-3 inhibitory activity against a set of 17 serine proteases was performed as described previously (32), with the exception of factor VIIa (FVIIa). Human FVIIa (Haematologic Technologies, Inc., Essex Junction, VT) at 20 nM concentration was pre-incubated for 10 min at 30°C with 400 nM Iripin-3 before the addition of 250 μM fluorogenic substrate Boc- QAR-AMC. The assay buffer used consisted of 20 mM Tris, 150 mM NaCl, 0.01% Triton X-100, 5 mM CaCl$_2$, and 0.1% polyethylene glycol 6000, pH 8.0. After the determination of the substrate hydrolysis rate, the six most strongly inhibited proteases were chosen for more detailed analysis. The assessment of covalent complex formation between Iripin-3 and selected serine proteases and the determination of second-order rate constants of protease inhibition are detailed in the Supplementary Materials.

Blood Coagulation

The effect of Iripin-3 on blood coagulation was tested by prothrombin time (PT), activated partial thromboplastin time (aPTT), and thrombin time (TT) assays. All chemicals were purchased from Technoclone (Vienna, Austria). Citrated human plasma (Coagulation Control N) was mixed either with 6 μM Iripin-3 or with five different Iripin-3 concentrations and then incubated for 10 min at room temperature. To perform the PT test, 100 ml of plasma with added Iripin-3 was incubated for 1 min at 37°C before the addition of 200 ml of Technoplastin HIS pre-warmed to 37°C. Plasma clotting time was measured on the Ceveron four coagulometer (Technoclone). In the aPTT test, the incubation of 100 ml of plasma mixed with Iripin-3 at 37°C for 1 min was followed by the addition of 100 ml of Dapttin TC. After incubating the mixture of plasma and Dapttin at 37°C for 2 min, 100 ml of 25 mM $CaCl_2$ was added to initiate the coagulation cascade. Plasma clotting time was determined as described above. To perform the TT test, 200 ml of plasma mixed with Iripin-3 was incubated at 37°C for 1 min. At the end of incubation, 200 ml of thrombin reagent was added, and plasma clotting time was measured as in the PT and aPTT assays.

Pro-Inflammatory Cytokine Production by BMDMs

Bone marrow cells were isolated from femurs and tibias of C57BL/6N mice. Both ends of the bones were cut with scissors, and bone marrow was flushed with complete medium. The complete medium was prepared by supplementation of RPMI 1640 medium containing glutamine (Biosera) with 10% heat-inactivated fetal bovine serum (FBS, Biosera), 50 μM 2-mercaptoethanol (Sigma Aldrich, St Louis, MO), 100 U/ml penicillin G (Biosera, Kansas City, MO) and 100 mg/ ml streptomycin (Biosera). After erythrocyte lysis in RBC lysis buffer (eBioscience, San Diego, CA), bone marrow cells resuspended in complete medium were seeded into 10 cm Petri dishes and incubated in the presence of 10 ng/ml granulocyte-macrophage colony- stimulating factor (GM-CSF, Sigma Aldrich) at 37°C and 5% CO_2 for 10 days. On days 4 and 7, non-adherent cells were removed, and the medium was replaced with fresh complete medium containing 10 ng/ml GM-CSF. On

day 10, adherent cells (macrophages) were collected, resuspended in RPMI 1640 medium supplemented only with 0.5% bovine serum albumin (BSA, Biosera), and seeded into 24- well culture plates (2×10^5 cells in 500 µl of culture medium per well). After 5 h incubation at 37°C and 5% CO_2, the medium was replaced with fresh RPMI 1640 medium containing 0.5% BSA, and BMDMs were pre-incubated for 40 min with 3 µ*M* or 6 µ*M* Iripin-3. Finally, 100 ng/ml of LPS (Sigma Aldrich; *E. coli* serotype O111:B4) was added, and macrophages were incubated in the presence of Iripin-3 and LPS for another 24 h. At the end of incubation, cells and cell-free supernatants were collected for RNA isolation and protein quantification, respectively. Relative expression of *Tnf*, *Il6*, and *Il1b* in macrophages was determined by RT-qPCR and concentrations of tumor necrosis factor (TNF), IL-6, and interleukin-1b (IL-1b) cytokines in collected supernatants were measured by DuoSet ELISA Development Kits (R&D Systems, Minneapolis, MN) according to the manufacturer's instructions with only minor modifications. The RT-qPCR analysis is described in detail in the Supplementary Materials.

Splenocyte Isolation and Culture in the Presence of Iripin-3

Spleens harvested from OT-II mice were forced through a Corning 70 µm cell strainer to obtain a single cell suspension. Red blood cells (RBCs) were removed from the suspension by the addition of 1× RBC lysis buffer (eBioscience), and the erythrocyte-free spleen cells were resuspended in RPMI 1640 medium with stable glutamine (Biosera) supplemented with 10% heat-inactivated FBS (Biosera), 50 µM 2-mercaptoethanol (Sigma Aldrich), 100 U/ml penicillin G (Biosera), and 100 µg/ml streptomycin (Biosera). Splenocytes were then seeded into 24-well or 96-well culture plates and pre-incubated with 3 µ*M* or 6 µ*M* Iripin-3 for 2 h. Pre-incubation with Iripin-3 was followed by the addition of ovalbumin (OVA) peptide 323–339 (Sigma Aldrich) at a concentration of 100 ng/ml. Splenocytes were incubated in the presence of Iripin-3 and OVA peptide at 37°C and 5% CO_2 for either 20 h (assessment of cell survival) or 72 h (analysis of cell proliferation and transcription factor expression).

Survival of B and T Cells

Mouse splenocytes were seeded into 96-well culture plates (5×10^5 cells in 200 ml of complete medium per well), pre- incubated with Iripin-3, and stimulated with OVA peptide. After 20 h incubation at 37°C and 5% CO_2, cells were harvested for flow cytometry analysis. First, splenocytes were stained with fixable viability dye eFluor 780 (eBioscience). Subsequently, Fc receptors were blocked with anti-CD16/CD32 antibody (eBioscience, clone 93), and surface antigen staining was performed with following monoclonal antibodies purchased from eBioscience: anti-CD45-PerCP-Cyanine5.5 (clone 30- F11), anti-CD19-PE (clone eBio1D3(1D3)), and anti-CD3e- APC (clone 145-2C11). Finally, the active form of caspase 3 in splenocytes was labeled using the FITC Active Caspase-3 Apoptosis Kit (BD Biosciences). The percentage of live $CD19^+$ and $CD3e^+$ splenocytes as well as the level of active caspase 3 were analyzed on the BD FACSCanto II flow cytometer using BD FACSDiva software version 6.1.3 (BD Biosciences).

Proliferation of $CD4^+$ T Cells

Erythrocyte-free splenocytes were stained with red fluorescent dye eFluor 670 (eBioscience), which allows monitoring of individual cell divisions. The stained splenocytes were seeded into 96-well culture plates (5×10^5 cells in 200 ml of complete medium per well), pre-incubated with Iripin-3, and stimulated with OVA peptide. Cells were allowed to proliferate for 72 h and then were harvested for flow cytometry analysis. Collected cells were stained with FITC-labelled anti-CD4 monoclonal antibody (clone GK1.5, eBioscience) and propidium iodide (eBioscience), and the percentage of proliferating live $CD4^+$ splenocytes was measured on the BD FACSCanto II flow cytometer using BD FACSDiva software version 6.1.3 (BD Biosciences).

Transcription Factor Expression in $CD4^+$ T Cells (RT-qPCR)

Splenocytes were seeded into 24-well culture plates (4.5 x 106 cells in 500 ml of complete medium per well), pre-incubated with Iripin- 3, and stimulated with OVA peptide. At the end of 72 h incubation, non-adherent cells

were collected, stained with FITC-labeled anti- CD4 monoclonal antibody (clone GK1.5, eBioscience), and CD4+ splenocytes were separated from the rest of the cell population using the S3e Cell Sorter (Bio-Rad Laboratories, Hercules, CA). RNA was extracted from CD4+ cells with the help of NucleoSpin RNA isolation kit (Macherey-Nagel, Düren, Germany), and 1 mg of total RNA was reverse transcribed into cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche Applied Science). RT-qPCR was performed in the CFX384 Touch thermal cycler (Bio-Rad) by utilizing five-fold diluted cDNA, SsoAdvanced Universal SYBR Green Supermix (Bio-Rad), and gene-specific primers. The PCR cycling conditions were 95°C for 3 min followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. The relative quantification of Tbx21 (Tbet), Gata3, Rorc, and Foxp3 transcripts in CD4+ splenocytes was performed using Pfaffl's mathematical model (64). Based on the results of geNorm analysis (65), Actb and Gapdh were utilized as reference genes for the calculation of relative expression ratios. Nucleotide sequences of forward and reverse primers as well as amplicon lengths are given in Supplementary Table 3.

Transcription Factor Expression in CD4$^+$ T Cells (Flow Cytometry)

Splenocytes were seeded into 24-well culture plates (2 x 10^6 cells in 500 ml of complete medium per well), pre-incubated with Iripin-3, and stimulated with OVA peptide. After 68 h incubation at 37°C and 5% CO_2, 20 ng/ml of phorbol 12-myristate 13-acetate (PMA; Sigma Aldrich) together with 1 μM ionomycin (Sigma Aldrich) were added to re-stimulate the cells. Brefeldin A (eBioscience) at a concentration of 3 mg/ml was added 1 h later, and splenocytes were incubated in the presence of PMA, ionomycin, and brefeldin A for another 4 h. At the end of incubation, non-adherent cells were collected and stained with fixable viability dyes eFluor 520 and eFluor 780 (eBioscience). Subsequently, Fc receptors were blocked with anti-CD16/CD32 antibody (eBioscience, clone 93), and surface antigen staining was performed with anti-CD4- Alexa Fluor 700 (BD Biosciences, clone RM4-5) and anti-CD25- PerCP-Cyanine5.5 (eBioscience, clone PC61.5) monoclonal antibodies. Surface antigen staining was followed by intracellular staining of transcription factors and cytokine IFN-

γ, for which the Foxp3/Transcription Factor Staining Buffer Set (eBioscience) was used in conjunction with following monoclonal antibodies: anti-T- bet-APC (clone eBio4B10 (4B10)), anti-GATA-3-PE (clone TWAJ), anti-RORgt-PE-CF594 (clone Q31-378), anti- Foxp3-PE-Cyanine7 (clone FJK-16s), and anti-IFN-γ-PE (clone XMG1.2). All antibodies were purchased from eBioscience except for the anti-RORgt antibody, which was obtained from BD Biosciences. Analysis was performed on the BD FACSCanto II flow cytometer using BD FACSDiva software version 6.1.3 (BD Biosciences).

Statistical Analyses

Data are presented in all graphs as mean ± the standard error of the mean (SEM). Differences between the mean values of two groups were analyzed by the unpaired two-tailed t-test. Differences between the mean values of three or more groups were analyzed by one-way ANOVA or randomized block ANOVA, which involved two variables: a fixed effect factor (treatment) and a random effect factor/block (an experimental run) (66). In the case of a statistically significant result ($p < 0.05$), Dunnett's post hoc test was performed to compare the mean of a control group with the means of experimental groups. All statistical tests were conducted using the software package STATISTICA 12 (StatSoft, Inc.). Statistically significant differences between groups are marked with asterisks (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

RESULTS

Iripin-3 Belongs to the Serpin Superfamily

A full-length nucleotide sequence of Iripin-3 was obtained during a salivary gland transcriptome project (16) and was submitted to GenBank under accession number GADI01004776.1. This sequence, consisting of 1182 base pairs, encodes a 377-amino acid (AA) protein with predicted molecular weight of approximately 42 kDa and with theoretical isoelectric point (pI) 5.23. The SignalP 4.1 server found a 16-AA signal peptide at the N terminus of the protein sequence (Figure 1A), which indicates that Iripin-3 is a potentially secreted protein. Using ScanProsite, the serpin signature motif PS00284 was identified

at AA positions 366-376 (Figure 1A). Moreover, two other serpin consensus AA motifs N-[AT]-[VIM]- [YLH]-F-[KRT]-[GS] and [DERQ]-[VL]-[NDS]-E-[EVDKQ]-G were recognized: NAMYFKG at AA positions 183-189 and EVNEEG at AA positions 338-343 (Figure 1A), suggesting that Iripin-3 belongs to the serpin superfamily. The hinge region of the Iripin-3 RCL has glycine at the P15 position, threonine at the P14 position, and residues with short side chains (alanine and valine) at positions P12-P9 (Figure 1A), which correspond to the RCLs of inhibitory serpins (68). The P1 site is occupied with the basic amino acid residue arginine (Figure 1A), suggesting Iripin-3 might target trypsin-like rather than chymotrypsin-like or elastase-like serine proteases (69). Using NetNGlyc 1.0 and NetOGlyc 4.0 servers, the Iripin-3 AA sequence was predicted to contain two potential N- glycosylation sites (N-X-[S/T]) and one putative O-glycosylation site (Figure 1A).

Iripin-3 Adopts a Typical Serpin Fold

Employing X-ray crystallography, we determined the 3D structure of Iripin-3 at 1.95 Å resolution. The crystal used exhibited symmetry of the $P6_222$ space group and contained one molecule in the asymmetric unit with a solvent content of 42.68%. The tertiary structure of Iripin-3 matched the 3D structures of other serpins, including the tick serpin IRS-2 (Figure 1B), with which it had the highest sequence similarity of all the serpin structures currently deposited in the PDB. More specifically, the Iripin-3 tertiary structure was composed of ten α-helices and three β-sheets, which were sequentially arranged in the order a1-b1-a2-a3-a4-a5-b2-a6-b3-a7-b4-b5-b6-b7-b8- a8-a9-b9-b10-a10-b11-b12-b13-b14-b15 (Figures 1A, 2). The sheet A consisted of six β-strands (b2, b3, b4, b10, b11, b12), sheet B of five β-strands (b1, b7, b8, b14, b15), and sheet C of four β-strands (b5, b6, b9, b13) (Figure 2). Iripin-3 in the crystal adopted a conformation known as the relaxed (R) state, since its RCL was probably cleaved by some contaminating proteases before or during the crystallization experiment. A protein sample can contain traces of contaminating cysteine and serine proteases, as demonstrated previously (70). The cleavage of the RCL led to the insertion of the RCL hinge region into the β-sheet A as an additional β-

strand S4 (Figure 2). The 3D structure of Iripin-3 contained 367 amino acid residues. The first 19 residues, which basically corresponded to the signal peptide of the protein, were missing. Moreover, the region $_{356}$LRSGSFD$_{362}$, in which the cleavage occurred, could not be modelled in the Iripin-3 structure due to its absence in the electron-density map. To compare the tertiary structure of Iripin-3 with that of IRS-2, the molecular structure of Iripin-3 was superposed with Ca atoms of IRS-2 with root-mean-square deviation of 0.8085 Å. The secondary structure elements were well conserved in both serpins, but there was a certain degree of divergence in disordered loop regions (Figure 1B).
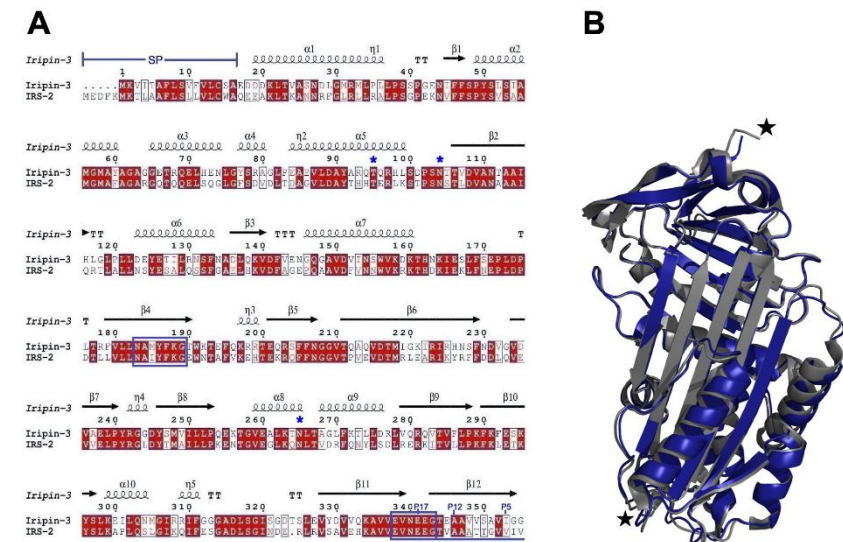
Figure. 1 A comparison of the primary, secondary and tertiary structures of Iripin-3 and IRS-2. (A) Structure-based sequence alignment of Iripin-3 and IRS-2. Secondary structure elements, which are shown above the aligned sequences, are depicted as spirals (α-helices, 3$_{10}$-helices) and arrows (β-sheets). Both Iripin-3 and IRS-2 possess a signal peptide (SP) at the N terminus of their sequences. Conserved AA motifs PS00284, N-[AT]-[VIM]-[YLH]-F-[KRT]-[GS], and [DERQ]-[VL]-[NDS]-E-[EVDKQ]-G are boxed in blue. The RCLs of both serpins are double underlined. Numbering of amino acid residues in the RCL is based on the standard nomenclature developed by Schechter and Berger (67). Putative N-glycosylation and O-glycosylation sites are marked with blue asterisks. (B) Superposition of the cleaved Iripin-3 structure (blue) on the structure of cleaved IRS-2 (gray). Cleavage sites are marked with black stars.
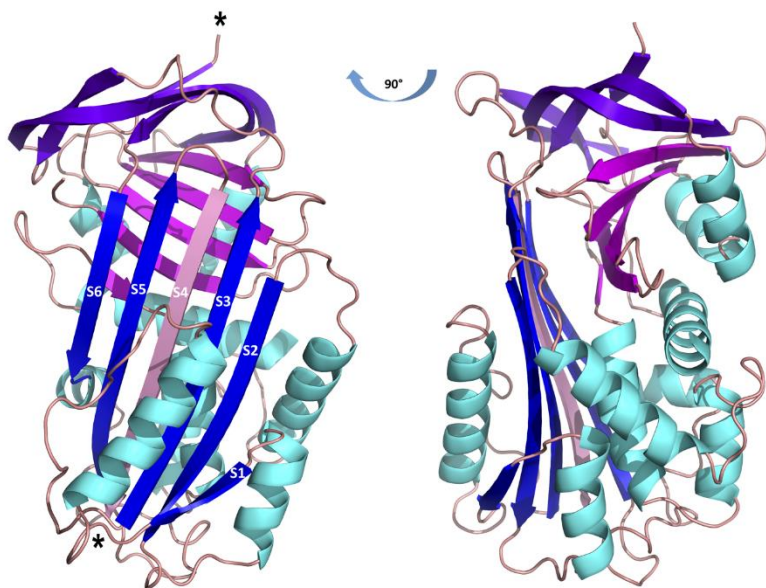
Figure. 2   Cartoon representation of the structure of cleaved Iripin-3. α-helices are colored cyan, β-sheet A is blue, β-sheet B is magenta, β-sheet C is purple, and loops are colored wheat. The insertion of the RCL hinge region between β-strands S3 and S5 (depicted in blue) resulted in the formation of an additional β-strand S4 (depicted in pink). Cleavage sites are marked with asterisks.

Iripin-3 Is Most Closely Related to Serpins From *I. scapularis*

The BLASTP search of the GenBank non-redundant protein sequences identified three *I. scapularis* serpins (accession numbers XP_029826754.1, EEC19555.1, and AAV80788.1) whose sequences were highly similar to the Iripin-3 sequence (percentage identities 95.4%, 94.9%, and 93.6%, respectively). These homologs have not been functionally characterized. The phylogenetic relationship of Iripin-3 with 26 tick serpins, whose function was deciphered either by using recombinant protein or at least by gene knockdown via RNA interference in ticks, was determined by using the maximum likelihood method and JTT matrix-based model. The resulting phylogenetic tree, with human alpha-1-antitrypsin as an outgroup, showed two distinct groups of tick serpins (Figure 3A). The first group at the bottom of the tree included eight serpins without a signal peptide with presumably intracellular function (Figure 3A). Notably, these serpins usually contained one or more cysteines and

methionines in their RCL (Figure 3B). The second, larger group at the top of the tree comprised 19 serpins with a signal peptide, including Iripin-3 (Figure 3A). Iripin-3 formed a small branch with one serpin from *I. scapularis* (IxscS-1E1) and one serpin from *I. ricinus* (IRS-2) (Figure 3A). In addition to the construction of the phylogenetic tree, we aligned the RCLs of the serpins used in the phylogenetic analysis (Figure 3B). Serpins that clustered together usually had similar RCLs, and the RCL of Iripin-3 resembled that of IxscS-1E1 (Figure 3B).
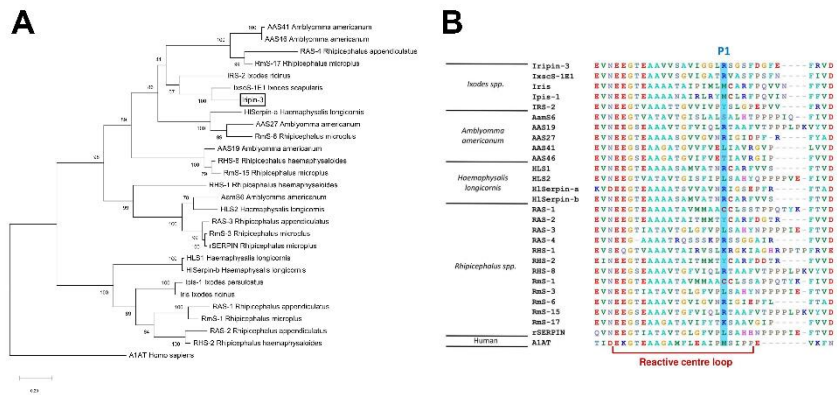
Figure. 3 Phylogenetic analysis of selected tick serpins. Protein sequences of previously characterized tick serpins were aligned and analyzed to determine phylogenetic relationships. (A) A phylogenetic tree was built using the maximum likelihood method and JTT matrix-based model. Alpha-1-antitrypsin (A1AT) was utilized as an outgroup to root the tree. The branch length represents the number of substitutions per site. The reliability of individual branches, assessed by bootstrapping, is expressed as a percentage of trees in which a given topology was present out of 1,000 replications. Iripin-3 is boxed. (B) Alignment of reactive center loop (RCL) regions of 27 tick serpins and one human serpin was performed using BioEdit. RCLs were determined based on the eight-residue pattern p17[E]-p16[E/K/R]-p15[G]-p14[T/S]-p13[X]-p12-9[AGS]- p8-1[X]-p1′-4′ [X] typical for inhibitory serpins (68). Amino acid residues at the predicted P1 site are highlighted in blue.

Iripin-3 Is Expressed in Feeding Ticks and Is Secreted Into Tick Saliva

In order to see how *iripin-3* expression changes during blood feeding, nymphal and adult ticks were allowed to feed on blood from host animals for various periods of time, and the amount of *iripin-3* transcript in tick tissues was subsequently determined by RT-qPCR. Overall, *iripin-3* expression was significantly induced in response to blood feeding in nymphs as well as in the salivary glands and ovaries of adult females (Figure 4A). In adults, the highest

- 43 -

levels of *iripin-3* mRNA were detected in the salivary glands (Figure 4A). To prove the presence of Iripin-3 protein in tick saliva, we collected saliva from ticks that were feeding for 6 to 7 days on guinea pigs. By ELISAs, markedly higher optical density values were obtained after exposure of tick saliva to anti-Iripin-3 serum than to pre-immune serum (Figure 4B), suggesting that Iripin-3 is a salivary protein. This result was further confirmed by western blotting. Rabbit pre-immune serum did not recognize recombinant Iripin-3, and there was no band of appropriate size (around 42 kDa) in tick saliva (Figure 4C). Conversely, the use of anti-Iripin-3 serum led to the recognition of recombinant Iripin-3 and appearance of an approximately 45 kDa band in tick saliva, which might represent native Iripin-3 (Figure 4D). The difference in the sizes of native and recombinant Iripin-3 was probably caused by the fact that native Iripin-3 is glycosylated, whereas recombinant Iripin-3 was prepared in the *E. coli* expression system and therefore lacks glycosylation. The other bands with sizes greater or less than 45 kDa that appeared in the lanes with tick saliva after exposure of membranes to either pre-immune serum or anti-Iripin-3 serum are most likely a result of non-specific binding of antibodies to some components
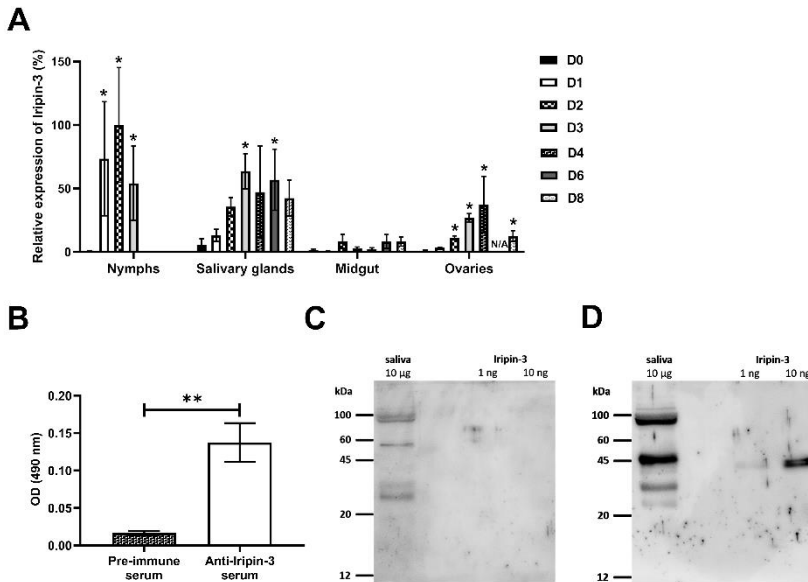
of tick saliva (Figures 4C, D).

Figure. 4 *Iripin-3* transcription in *I. ricinus* ticks is increased in response to blood feeding, and Iripin-3 protein is present in the saliva of feeding ticks. (A)

Iripin-3 mRNA expression in nymphs and in the salivary glands, midguts and ovaries of adult females feeding for 1 (D1), 2 (D2), 3 (D3), 4 (D4), 6 (D6), and 8 (D8) days or not feeding at all (D0). In nymphs, the last column represents fully engorged ticks that completed their blood meal in 3 or 4 days. N/A – data not available. Relative expression values were calculated using the DDCt (Livak) method (60), with *rps4* serving as a reference gene. A group with the highest *iripin-3* expression (nymphs feeding for 2 days) was utilized as a calibrator during calculations, and its expression value was set to 100%. Data are presented as mean of three biological replicates ± SEM. Statistically significant induction ($p < 0.05$) of *iripin-3* expression as compared to unfed ticks is marked with an asterisk. (B) ELISA results expressed as optical density (OD) values measured after exposure of tick saliva to either rabbit pre-immune serum or rabbit antiserum to Iripin-3. Data are presented as mean ± SEM of three values (**$p < 0.01$). (C, D) Tick saliva (10 mg) and Iripin-3 (1 ng or 10 ng) were resolved by SDS-PAGE and transferred to PVDF membranes. The membranes were incubated with rabbit pre-immune serum (C) or rabbit antiserum to Iripin-3 (D).

Iripin-3 Primarily Inhibits Kallikrein and Matriptase

An initial screen for Iripin-3 inhibitory activity was carried out against 17 different serine proteases. Statistically significant reductions in enzymatic activity were observed for ten proteases, but only six of these, namely kallikrein, matriptase, trypsin, plasmin, thrombin, and FVIIa, had their proteolytic activity reduced by >20% (Figure 5A). Iripin-3 formed covalent complexes, typical for the serpin "suicide" mechanism of inhibition (71), with kallikrein, matriptase, thrombin, and trypsin, as shown by SDS-PAGE (Figure 5B). There was no visible complex between Iripin-3 and plasmin on the gel (Figure 5B). It is possible that the complex was hidden within an approximately 70 kDa protein band, which was also present in the lane with plasmin only (Figure 5B). Moreover, no SDS- and heat-stable complex was formed between Iripin-3 and FVIIa in the absence or presence of tissue factor under given conditions (Supplementary Figure 1), suggesting Iripin-3 probably does not reduce the proteolytic activity of FVIIa through the classic serpin inhibitory mechanism. Finally, the second-order rate constants k_2 for the interactions between Iripin-3 and kallikrein, matriptase, thrombin, and trypsin were measured by a discontinuous method under pseudo first-order conditions. Iripin- 3 most potently inhibited kallikrein with $k_2 = 8.46 \pm 0.51 \times 10^4 \, M^{-1} \, s^{-1}$ (Figure 5C). The k_2 for the interactions between Iripin-3 and matriptase and trypsin were determined as $5.93 \pm 0.39 \times 10^4 \, M^{-1} \, s^{-1}$ and $4.65 \pm 0.32 \times 10^4 \, M^{-1} \, s^{-1}$, respectively (Figures 5D, F). Thrombin was inhibited by Iripin-3 with the lowest potency (k_2

$= 1.37 \pm 0.21 \times 10^3$ M^{-1} s^{-1}) (Figure 5E). Interface analysis between the active sites of matriptase, thrombin, kallikrein and trypsin and the P4-P4′ part of Iripin-3 RCL revealed possible polar interactions that could indicate the binding selectivity of Iripin-3 for target proteases (Supplementary Figure 2). The strongest interaction with the catalytic triad was calculated for matriptase, followed by trypsin, kallikrein and thrombin (data not shown). According to this analysis, thrombin and kallikrein should be inhibited by Iripin-3 with similar potency. This, however, was not supported by enzyme-substrate kinetic analyses (Figures 5C–F), in which kallikrein displayed 60 times higher k_2 value than thrombin. Therefore, the specificity of Iripin-3 is probably dependent on more factors. As shown in Supplementary Figure 3, matriptase and trypsin have open and shallow active sites, easily accessible to various substrates, including Iripin-3 RCL. Thrombin and kallikrein, on the other hand, possess narrower and deeper cavities with the catalytic triad (Supplementary Figure 3). It is possible that some subtle differences in spatial arrangement hinder the access of Iripin-3 RCL to the thrombin's active site, while facilitating its access to the kallikrein's active site cleft.
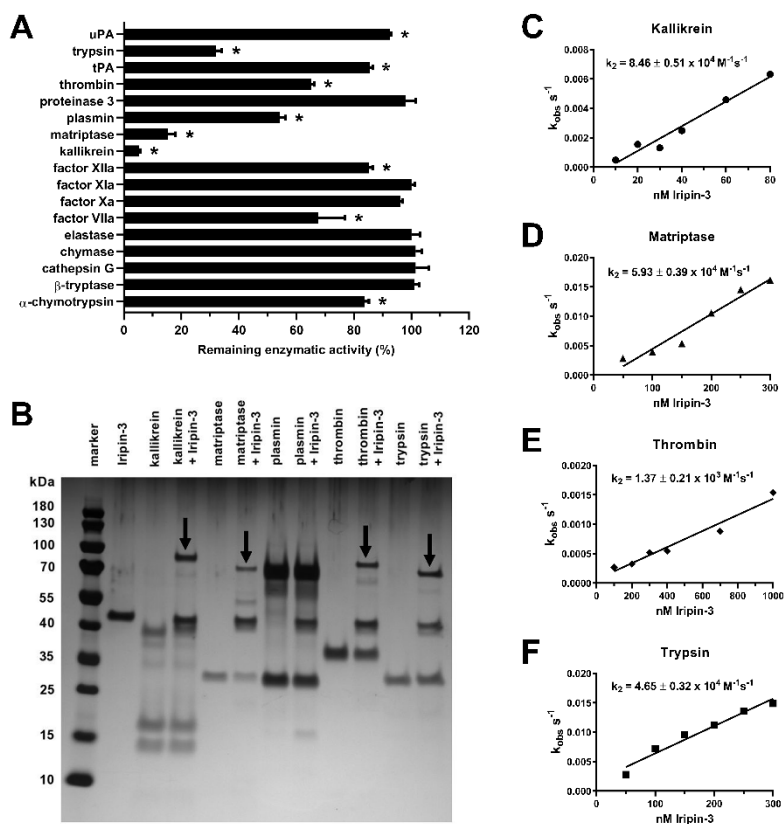
Figure. 5 Iripin-3 suppresses the enzymatic activities of kallikrein, matriptase, thrombin, and trypsin through the classic serpin inhibitory mechanism. (A) The residual enzymatic activities of 17 selected serine proteases in the presence of 400 n*M* Iripin-3. The experiment was performed in triplicate, and data are expressed as mean ± SEM. The enzymatic activities of individual proteases in the absence of Iripin-3 (control groups) were considered as 100%, and differences between control groups and Iripin-3- treated groups were analyzed by the unpaired two-tailed *t*-test. Enzymes labelled with an asterisk were inhibited with statistical significance ($p < 0.05$). (B) Formation of SDS- and heat-stable complexes between Iripin-3 and kallikrein, matriptase, plasmin, thrombin, and trypsin. Proteins were resolved on 4 to 12% NuPAGE Bis-Tris gels and visualized by silver staining. Covalent complexes between Iripin-3 and target proteases are marked with black arrows. (C–F) The apparent first-order rate constant k_{obs} was plotted against Iripin-3 concentration, and linear regression was performed to obtain the line of best fit. The slope of the line represents the second-order rate constant k_2 for the inhibition of kallikrein (C), matriptase (D), thrombin (E), and trypsin (F) by Iripin-3. For each determination, the standard error of the slope is given.

Iripin-3 Prolongs Plasma Clotting Time in the Prothrombin Time Assay

Since tick serpins commonly inhibit the host coagulation system (72), we tested the effect of Iripin-3 on the extrinsic coagulation pathway, intrinsic coagulation pathway, and common coagulation pathway by using prothrombin time (PT), activated partial thromboplastin time (aPTT), and thrombin time (TT) tests, respectively (73). Iripin-3 at 6 μM final concentration did not significantly prolong plasma clotting time in the aPTT and TT assays (data not shown). However, there was a statistically significant delay in blood clot formation in the PT test when plasma was treated with 1.5, 3, and 6 μM Iripin-3 (Figure 6). The highest Iripin-3 concentration prolonged the prothrombin time by 8.8 s when compared to control plasma (Figure 6). These results therefore indicate that Iripin-3 slightly inhibits the extrinsic pathway while not affecting the intrinsic and common pathways of blood coagulation.
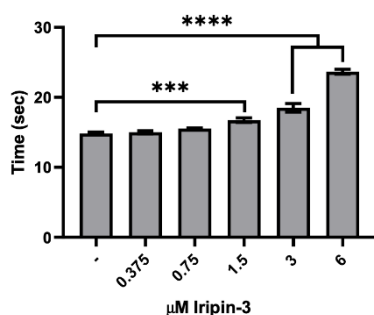
Figure. 6   Iripin-3 inhibits the extrinsic pathway of blood coagulation. Human plasma was treated with no Iripin-3 or with 0.375, 0.75, 1.5, 3, and 6 μM Iripin-3 and the time required for blood clot formation in the prothrombin time assay was subsequently determined on a coagulometer. Data are presented as mean ± SEM of three independent experiments (***p < 0.001, ****p < 0.0001).

Iripin-3 Decreases Production of IL-6 by BMDMs

Serpins secreted in tick saliva can facilitate blood meal uptake not only by inhibiting coagulation but also by suppressing host inflammatory responses (37, 72, 74). Therefore, we next investigated whether Iripin-3 attenuates pro-inflammatory cytokine production by LPS-stimulated BMDMs. The production of TNF, IL-6, and IL-1b was assessed at the mRNA level by RT-qPCR as well as at the protein level by ELISA. Iripin-3 caused a dose-dependent and

statistically significant reduction in the transcription of all three genes (Figures 7A–C). However, decreases in the transcription of *Tnf* and *Il1b* did not result in corresponding changes in the concentrations of these two pro- inflammatory cytokines at the protein level (Figures 7D, F). Conversely, Iripin-3 was an efficient inhibitor of both IL-6 synthesis and secretion (Figure 7E).
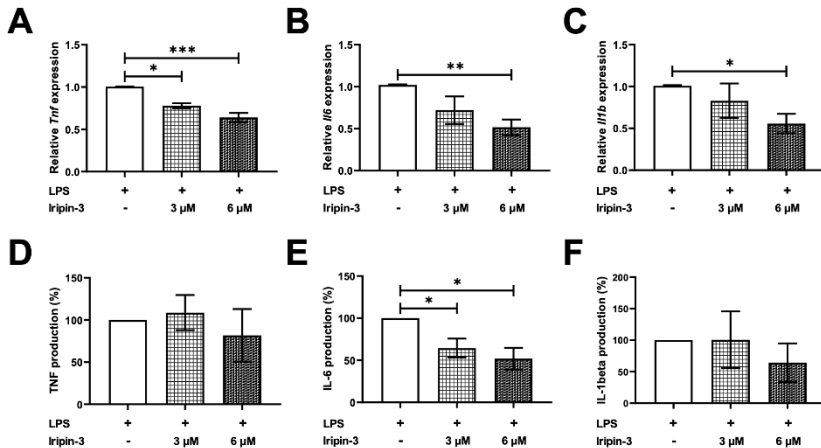
Figure. 7   Iripin-3 inhibits the expression of pro-inflammatory cytokines in LPS-stimulated BMDMs. Macrophages derived from bone marrow cells isolated from C57BL/6N mice were pre-incubated with 3 μM or 6 μM Iripin-3 for 40 min and were then stimulated with LPS (100 ng/ml) for 24 h. (A–C) At the end of 24 h incubation, cells were harvested for RNA extraction and the expression of *Tnf* (A), *Il6* (B), and *Il1b* (C) was determined by RT-qPCR. Relative expression values were calculated using the delta-delta Ct (Livak) method (60), with *Gapdh* serving as a reference gene. Cells incubated only in the presence of LPS were utilized as a calibrator during calculations. Data are presented as mean ± SEM of four independent experiments (*p < 0.05, **p < 0.01, ***p < 0.001). (D–F) Supernatants were collected, and TNF, IL-6, and IL-1b concentrations in these supernatants were measured by sandwich ELISA. TNF (D), IL-6 (E), and IL-1b (F) production by Iripin-3- treated BMDMs is expressed as the percentage of the cytokine production by control macrophages, since there were large differences in the concentrations of the same cytokine between three independent repeats of the experiment. Data are expressed as mean ± SEM, and statistically significant differences (p < 0.05) are marked with an asterisk.

Iripin-3 Impairs B and T Cell Viability *In Vitro*

In addition to inhibiting innate immune mechanisms, tick serpins can alter the host adaptive immune response (35, 37, 72). First, we tested whether Iripin-3 had an effect on B and T lymphocyte viability. Incubation of splenocytes derived from OT-II mice for 20 h in the presence of two different concentrations of Iripin-3 (3 μM and 6 μM) resulted in a pronounced dose-

dependent reduction in the viability of both B cells (CD45$^+$ CD19$^+$ splenocytes) and T cells (CD45$^+$ CD3e$^+$ splenocytes), with B cell survival more negatively affected by the serpin presence than T cell survival (Figures 8A–D). B and T cell viability was impaired irrespective of whether the splenocytes were left unstimulated or were stimulated with OVA peptide (Figures 8C, D). Conversely, Iripin-3 did not reduce the viability of BMDMs or dendritic cells (Supplementary Figures 4A, B), and the viability of LPS-activated neutrophils was impaired only in the presence of the highest (6 mM) concentration of Iripin-3 (Supplementary Figure 4C). Therefore, Iripin-3 might selectively induce B and T cell death. To investigate the possibility that Iripin-3 triggers lymphocyte apoptosis, we measured active caspase-3 levels in both unstimulated and OVA peptide-stimulated splenocytes. Treatment of splenocytes with Iripin-3 did not lead to a statistically significant increase in the level of active caspase-3 (Figures 8E, F). Therefore, Iripin-3 probably does not induce B and T cell death through
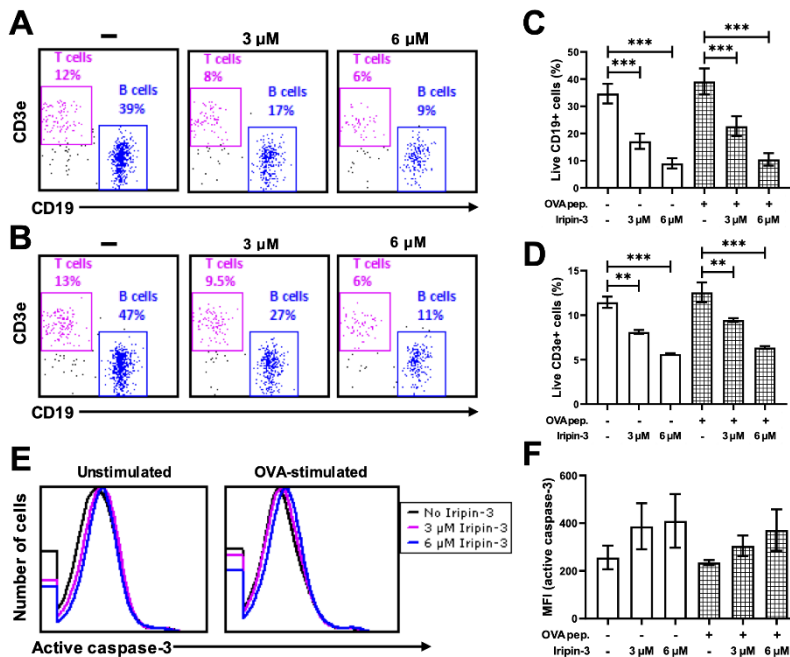
activation of a caspase-3-dependent pathway.

Figure. 8   Iripin-3 reduces B and T cell viability and does not significantly alter active caspase-3 levels. (A, B) Dot plots depicting the percentage of live CD45$^+$CD19$^+$ cells (B cells) and live CD45$^+$CD3e$^+$ cells (T cells) in unstimulated splenocytes (A) or OVA peptide-stimulated splenocytes (B). Splenocytes were not treated with Iripin-3 (left) or were treated with 3 μM (middle) or 6 μM (right) Iripin-

3. (C, D, F) The percentage of live B cells (C), live T cells (D), and median fluorescence intensity (MFI) corresponding to the level of active caspase-3 (F) after incubating the splenocytes for 20 h in the absence of Iripin-3 or in the presence of 3 μM and 6 μM Iripin-3. The cells were left either unstimulated or were stimulated with 100 ng/ml of OVA peptide. Data are presented as mean ± SEM of three independent experiments (**p < 0.01, ***p < 0.001). (E) Histograms showing the level of active caspase-3 in either unstimulated splenocytes (left) or splenocytes stimulated with OVA peptide (right). Splenocytes were incubated for 20 h without Iripin-3 or were treated with 3 μM or6 μM Iripin-3.

Iripin-3 Inhibits *In Vitro* CD4$^+$ T Cell Proliferation

Since Iripin-3 reduced T cell viability, we tested whether it also affected the survival and proliferation of CD4$^+$ helper T cells. OT-II splenocytes were pre-incubated with 3 μM or6 μM Iripin- 3 for 2 h before being stimulated with OVA peptide for 72 h. Propidium iodide staining in combination with the application of anti-CD4 antibody revealed a lower percentage of live CD4$^+$ cells in Iripin-3-treated groups than in the control group (Figure 9A), suggesting Iripin-3 has a negative effect on CD4$^+$ T cell viability. After the exclusion of dead cells, we assessed the proliferation of CD4$^+$ T cells. Unstimulated CD4$^+$ cells did not proliferate at all (Figure 9C), whereas addition of OVA peptide triggered proliferation in approximately 95% of cells (Figures 9B, D). Treatment with Iripin-3 caused a dose-dependent decrease in CD4$^+$ splenocyte proliferation (Figure 9B). While about 84% of cells proliferated in the presence of 3 μM Iripin-3 (Figures 9B, E), only 35% of cells were capable of proliferation after addition of 6 μM Iripin-3 (Figures 9B, F). Therefore, Iripin-3 impairs both the viability and proliferation of CD4$^+$ T cells.
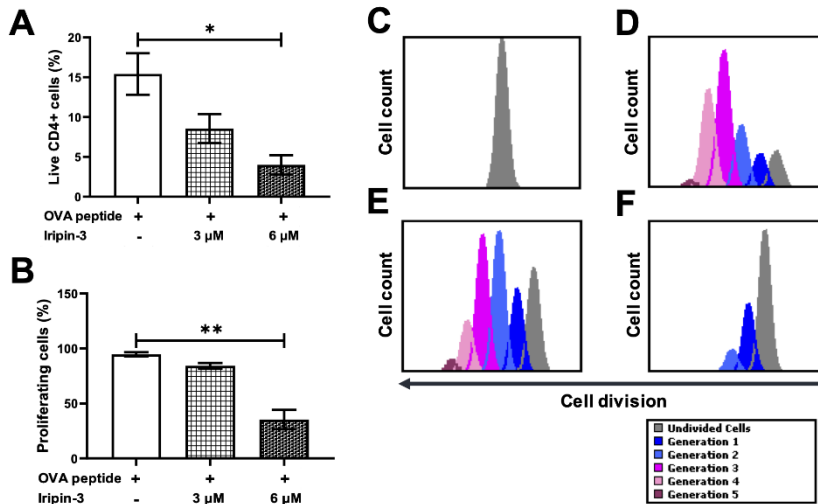
Figure. 9 Iripin-3 impairs the survival and proliferation of CD4$^+$ splenocytes. (A, B) The percentage of live CD4$^+$ cells (A) and the percentage of proliferating live CD4$^+$ cells (B) after exposure to 3 μM or 6μM Iripin-3. Cells not treated with Iripin-3 were used as control. After 2 h pre-incubation with Iripin-3, cells were cultured in the presence of OVA peptide (100 ng/ml) for 72 h. Data are presented as mean ± SEM of three independent experiments (*p < 0.05, **p < 0.01). (C–F) Histograms showing the number of live CD4$^+$ cells that managed to divide once (blue), twice (light blue), 3 times (pink), 4 times (rose), 5 times (plum), or did not divide at all (gray) within the 72 h culture period. Cells were incubated in the absence of Iripin-3 and OVA peptide (C), in the presence of OVA peptide only (D), or were treated with the combination of 3 μM Iripin-3 and OVA peptide (E) or 6 μM Iripin-3 and OVA peptide (F).

Iripin-3 Inhibits a Th1 Immune Response and Promotes Differentiation of Regulatory T Cells (Tregs) *In Vitro*

To examine whether Iripin-3 alters the differentiation of naïve CD4$^+$ T cells into Th1, Th2, Th17, or Treg subpopulations, we evaluated the expression of transcription factors T-bet, GATA-3, RORgt, and Foxp3 in OVA peptide-stimulated CD4$^+$ splenocytes by RT-qPCR and flow cytometry. T-bet, GATA-3, RORgt, and Foxp3 are considered lineage-specifying transcription factors that govern Th1, Th2, Th17, and Treg differentiation, respectively (75–79). Iripin-3 markedly and dose-dependently inhibited the expression of T-bet in CD4$^+$ T cells at both the mRNA and protein levels (Figures 10A–C). Since T-bet controls *Ifng* transcription (76), we also tested the ability of Iripin-3 to inhibit the production of this hallmark Th1 cytokine. As with T- bet, Iripin-3 induced a

pronounced and dose-dependent reduction in the percentage of CD4$^+$ T cells producing IFN-γ (Figures 10D, E). Despite the inhibition of the Th1 immune response, we did not observe significant changes in the differentiation of T cells into Th2 or Th17 subpopulations (Figures 10F–K). GATA-3 expression was slightly increased only in CD4$^+$ T cells treated with 3 μM Iripin-3 (Figures 10G, H). Similarly, both Iripin-3 concentrations induced only a small and non-significant increase in the percentage of CD4$^+$ T cells expressing RORgt (Figures 10J, K). Finally, Iripin-3 moderately stimulated the expression of Foxp3 at both the mRNA and protein levels (Figures 10L–N). Therefore, Iripin-3 might induce the differentiation of Tregs in addition to inhibiting Th1 cell development.
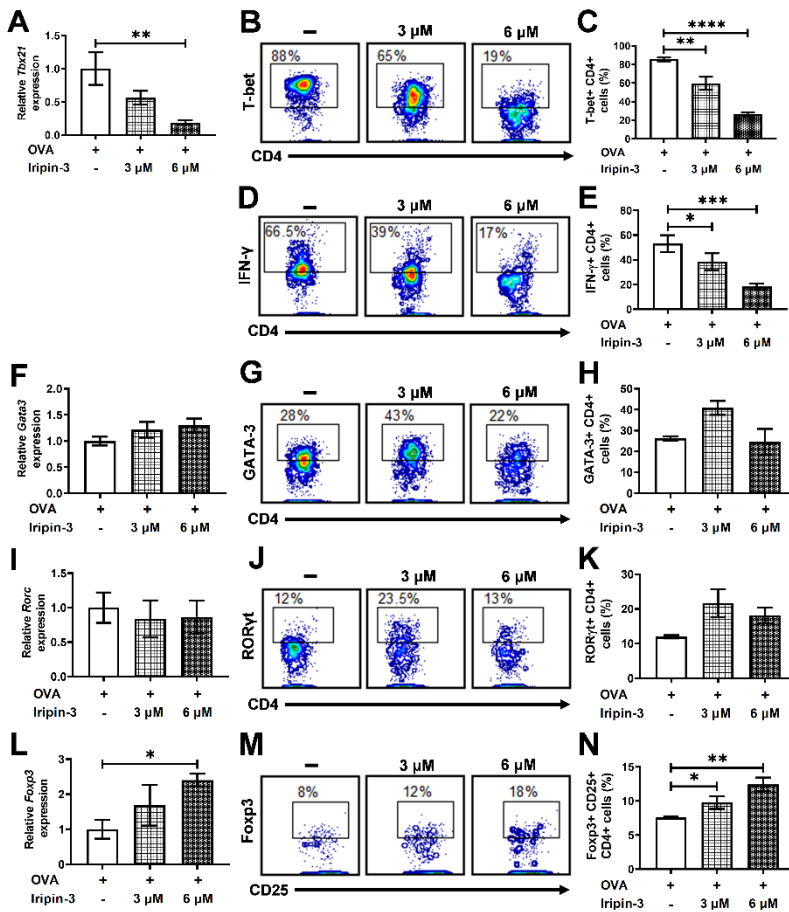
Figure. 10 Iripin-3 alters the expression of CD4[+] T cell transcription factors at both the mRNA and protein levels. (A, F, I, L) Expression of *Tbx21* (A), *Gata3* (F), *Rorc* (I), and *Foxp3* (L) in CD4[+] cells stimulated with OVA peptide for 72 h. Cells were untreated with Iripin-3 or were treated with 3 μM or 6 μM Iripin-3. Cells incubated only in the presence of OVA peptide were utilized as a calibrator during calculations of relative expression values. Data are presented as mean ± SEM of four independent experiments (* p < 0.05, ** p < 0.01). (B, D, G, J, M) Representative contour plots showing the proportion of OVA peptide-stimulated CD4[+] splenocytes expressing T-bet (B), IFN-γ (D), GATA-3 (G), RORgt (J) and the combination of CD25 and Foxp3 (M). The cells were incubated in the absence of Iripin- 3 (left) or in the presence of two different Iripin-3 concentrations: 3 μM (middle) and 6 μM (right). (C, E, H, K, N) The percentage of CD4[+] T cells producing the cytokine IFN-γ (E) and expressing transcription factors T-bet (C), GATA-3 (H), RORgt (K), and Foxp3 together with CD25 (N). Cells were cultured in the presence of Iripin-3 (3 μM or 6 μM) and OVA peptide for 72 h. Cells incubated without Iripin-3 were used as control. Data are presented as mean ± SEM of three or four independent experiments (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001).

Iripin-3 Is Not Essential for Feeding Success of *I. ricinus* Nymphs

Since *iripin-3* expression is induced in nymphs in response to blood feeding, we decided to assess the role of this serpin in the blood-feeding process by silencing *iripin-3* expression in nymphs via RNA interference. *Iripin-3* expression in *iripin-3* dsRNA- treated ticks was 34% when compared to *gfp* dsRNA-treated ticks (data not shown), suggesting that the knockdown of the target gene was successful. Despite diminished *iripin-3* expression, the time course of blood feeding and overall feeding success (*i.e*., the number of nymphs that reached full engorgement) did not significantly differ between control ticks and *iripin-3* dsRNA- treated ticks (Supplementary Table 4). The weight of fully engorged nymphs was not significantly affected by *iripin-3* silencing as well (Supplementary Table 4). Therefore, we can conclude that the deficiency of Iripin-3 alone is not sufficient to impair the blood meal acquisition and processing by nymphal *I. ricinus* ticks.

DISCUSSION

Tick saliva contains hundreds to thousands of proteins from diverse protein families (80). These salivary proteins are differentially expressed over the course of blood feeding and enable ticks to feed to repletion by maintaining blood fluidity and suppressing host defense responses (80). Serpins form one of four serine protease inhibitor families that have been discovered in ticks (72). Serpins are particularly intriguing to study, not only due to their unique trapping inhibitory mechanism but also because they regulate a variety of physiological processes in many organisms. The functional diversity of the serpin superfamily is exemplified by the widely studied human serpins, which have been shown to regulate blood pressure, transport hormones, and control blood coagulation, fibrinolysis, angiogenesis, programmed cell death, inflammation, or complement activation (81–84). We presume that ticks employ some of their serpins to modulate host defenses, as evidenced by several tick serpins with anti-platelet, anti-coagulant, anti-inflammatory, and/or immunomodulatory properties that have been shown to be secreted via saliva into the host (34–37, 72).

Here we determined the structure and partially deciphered the function of *Ixodes ricinus* serpin Iripin-3 by using several *in vitro* models. The size (377 amino acids), molecular weight (42 kDa), and 3D structure of Iripin-3, consisting of three β-sheets, ten α-helices, and a cleaved RCL, correspond to the structural parameters of typical serpins (18, 20, 71). *Iripin-3* expression was induced by blood feeding in both nymphs and adult females, suggesting Iripin-3 contributes to feeding success in both developmental stages. Of the three organs of adult ticks, the highest levels of *iripin-3* transcript were detected in the salivary glands. The presence of Iripin-3 protein in the saliva of partially engorged adults was confirmed by immunodetection. Thus, we can assume that Iripin-3 is secreted via saliva into the tick attachment site where it interferes with host anti-tick defenses. Statistically significant increase of *iripin-3* expression in response to blood feeding occurred not only in the salivary glands but also in the ovaries of adult ticks, which indicates that Iripin-3 might be somehow involved in the reproductive process.

The role of serpins in tick reproduction has been evidenced recently by *Rhipicephalus haemaphysaloides* serpin RHS-8, the knockdown of which impaired oocyte maturation due to the inability of oocytes to uptake adequate amount of vitellogenin (45). The presence of the basic amino acid residue arginine at the P1 site of the Iripin-3 RCL indicates that Iripin-3 might inhibit trypsin-like rather than chymotrypsin-like or elastase-like serine proteases (69, 85). Indeed, out of 17 selected serine proteases,

Iripin-3 most potently inhibited trypsin-like serine proteases kallikrein and matriptase and exhibited weaker inhibitory activity against trypsin, thrombin, plasmin, and factor VIIa. Kallikrein participates in the activation of the intrinsic blood coagulation pathway, promotes fibrinolysis, and is also responsible for the release of the potent inflammatory mediator bradykinin, which further induces vasodilation, increases vascular permeability, and evokes pain and itch (86, 87). Matriptase is a type II transmembrane serine protease that is primarily expressed in epithelial cells and is essential for the maintenance of skin barrier function (88). Moreover, matriptase seems to be involved in cutaneous wound healing (89, 90) and might contribute to the amplification and

perpetuation of the inflammatory response through the activation of protease-activated receptor-2 (PAR-2) (91). Therefore, we speculate that Iripin-3-mediated inhibition of kallikrein and matriptase contributes to tick feeding success by suppressing the inflammatory response and consequent itch and pain and by impairing wound healing.

A phylogenetic analysis of 27 functionally characterized tick serpins revealed a close phylogenetic relationship between Iripin-3 and *I. scapularis* serpin IxscS-1E1. Both serpins possess arginine at the P1 site and inhibit trypsin and thrombin (30). However, while IxscS-1E1 prolonged plasma clotting time in aPTT and TT assays and had no effect on blood clot formation in the PT assay (30), Iripin-3 inhibited only the extrinsic coagulation pathway. This indicates that the Iripin-3-mediated inhibition of kallikrein and thrombin was not sufficient to significantly impair the intrinsic and common coagulation pathways. Other blood clotting factors (XIIa, XIa, Xa) involved in the intrinsic and common pathways were not markedly inhibited by Iripin-3. Several tick serpins are capable of inhibiting the common (and perhaps intrinsic) pathway of blood coagulation (28–31, 41, 92); however, none have shown any effect on the extrinsic coagulation pathway. The extrinsic coagulation pathway is initiated by damage to a blood vessel and subsequent formation of a FVIIa/tissue factor (TF) complex, which further activates factor X (93). In view of the fact that Iripin-3 exhibited weak inhibitory activity only in the PT test and not in the aPTT test or TT test, we hypothesized that it might target either FVIIa or TF, since these two proteins are the only unique components of the extrinsic pathway. FVIIa seemed to be a more likely target for Iripin-3 given that it is a serine protease (94), and some human serpins, such as antithrombin III or protein C inhibitor, have been shown to inhibit the proteolytic activity of FVIIa (95–97). In our hands, Iripin-3 did not form a covalent complex with FVIIa either in the absence or in the presence of TF. However, the proteolytic activity of FVIIa was reduced by approximately 30% in the presence of 400 nM Iripin-3 in the kinetic enzyme-substrate assay. Therefore, the prolongation of blood clot formation in the PT assay might be caused by the non-canonical inhibition of FVIIa by Iripin-3. Alternatively, a possible interaction between Iripin-3 and TF could also prevent

FVIIa/TF complex formation, leading to a lower rate of FXa generation and inhibition of blood coagulation. In addition to the inhibition of blood coagulation, Iripin-3 displayed anti-inflammatory activity *in vitro*, since it significantly and dose-dependently attenuated the production of pro-inflammatory cytokine IL-6 by LPS-stimulated bone marrow- derived macrophages. The decreased IL-6 production was probably caused by the inhibition of *Il6* transcription and not by reduced viability of macrophages, since the metabolic activity of macrophages remained unchanged in the presence of Iripin-3. Several tick serpins have been shown to inhibit IL-6 transcription and secretion (37–39, 74, 98), which can occur as a result of serpin-mediated inhibition of proteases such as cathepsin G and cathepsin B (37). However, the inhibition of pro-inflammatory cytokine production does not have to be dependent on serpin anti-protease activity because some serpins, like Iris and α-1- antitrypsin, can alter pro-inflammatory cytokine production by binding to immune cells via exosites (98, 99). An inflammatory environment with reduced IL-6 might favor differentiation of Tregs (100–102). Splenocytes, incubated in the presence of Iripin-3 for 72 h, increased the expression of Treg-specific transcription factor Foxp3 (77, 78), suggesting that Iripin-3 indeed induces the differentiation of naïve CD4[+] T cells into anti-inflammatory Tregs. Tregs would facilitate the suppression of the host immune response (103), which would be beneficial for feeding ticks. There is scarce evidence that tick saliva induces Treg differentiation (104, 105). The results of our *in vitro* assay indicate that salivary serpins could contribute to this particular activity of tick saliva.

Besides the reduction in IL-6 production and increase in Foxp3 expression, Iripin-3 caused a pronounced, dose- dependent decrease in B and T cell viability *in vitro*. This effect appears to be B and T cell-specific since macrophage and dendritic cell survival was not affected by Iripin-3 and the viability of LPS-stimulated neutrophils was slightly impaired only at the highest (6 mM) concentration of Iripin-3. Serpins usually protect cells from dying by reducing the proteolytic activity of enzymes (such as granzymes and caspases) involved in programmed cell death (106). However, certain serpins, e.g., kallikrein-binding protein, pigment epithelium-derived factor, or maspin, induce

apoptosis of endothelial cells and some cancer cells through distinct mechanisms such as the activation of the Fas/FasL/caspase-8 signaling pathway or the permeabilization of the outer mitochondrial membrane followed by a loss of transmembrane potential (107–111). Active caspase-3 levels were only slightly and non-significantly increased in Iripin-3- treated splenocytes. Therefore, the induction of caspase- dependent apoptosis was not the main cause of impaired splenocyte viability. Various forms of caspase-independent cell death have been described such as autophagy, paraptosis, necroptosis, or necrosis (112, 113). Elucidation of the exact mechanism behind the extensive splenocyte death in the presence of Iripin-3 is, however, beyond the scope of this paper.

I. ricinus saliva and salivary gland extracts inhibit T cell proliferation and suppress Th1 cell differentiation while simultaneously augmenting the Th2 immune response (114– 117). Iripin-3 might contribute to this immunomodulatory effect of saliva, since in our *in vitro* assays it inhibited CD4$^+$ T lymphocyte proliferation and impaired the differentiation of naïve CD4$^+$ T cells into Th1 cells. Impaired Th1 cell generation was evidenced by decreased expression of the Th1 lineage- specifying transcription factor T-bet and a reduced percentage of CD4$^+$ T cells producing the hallmark Th1 cytokine IFN-γ. Several studies have reported inhibition of splenocyte and peripheral blood mononuclear cell proliferation in the presence of tick serpins (35, 37, 38, 40). Interestingly, the inhibition of mitosis observed in these studies was usually accompanied by decreased IFN-γ production (35, 38, 40), which might indicate, among other things, the suppression of Th1 cell differentiation. The causative mechanism of reduced cell proliferation and impaired Th1 cell differentiation in the presence of tick serpins remains unknown, but it could be associated with decreased production of certain cytokines such as IL-2, IL-12, and IFN-γ. In the case of Iripin-3, there might be a connection between the inhibition of cell proliferation and impaired viability of splenocytes, *i.e.*, the mechanism behind B and T cell death could be also responsible for the suppression of CD4$^+$ T cell division. Iripin-3-mediated differentiation of naïve CD4$^+$ T cells into Tregs might also contribute to the reduction in CD4$^+$ T cell proliferation, since Tregs

can inhibit cell multiplication by various mechanisms including the production of immunosuppressive cytokines TGF-b and IL-35, consumption of IL-2, and conversion of ATP to adenosine (103, 118).

It is worth mentioning that the Iripin-3 concentrations used in *in vitro* experiments (3 μM and 6 μM) are probably higher than the amount of Iripin-3 at the tick feeding site. This fact, however, does not make the anticoagulant, ant-inflammatory and immunomodulatory activities of Iripin-3 observed *in vitro* physiologically irrelevant. Tick saliva is a complex mixture of proteins from the same or different protein families, and some of these salivary proteins can share the same function (119). Therefore, even a low concentration of one tick protein may be sufficient to achieve a desired effect at the tick attachment site if this protein acts in concert with other tick proteins (119). For instance, the ability of *I. ricinus* saliva to inhibit CD4$^+$ T cell proliferation is probably a result of combined action of more proteins with anti- proliferative properties, such as the serpins Iripin-3 and Iris, the cystatin Iristatin and the Kunitz domain-containing protein IrSPI (38, 120, 121). That *I. ricinus* saliva may contain other proteins possessing Iripin-3-like activities was demonstrated by the RNA interference experiment. *Iripin-3* knockdown did not significantly affect the overall feeding success, time course of blood feeding and weight of fully engorged nymphs, which indicates that other similarly acting salivary proteins might compensate for the loss of *iripin-3* expression. It is also important to note that native Iripin-3 is most likely glycosylated. However, recombinant Iripin-3 was prepared in an *E. coli* expression system, and therefore it lacks glycosylation. Glycosylation has been shown to reduce the propensity of serpins for polymerization (122) and increase the stability and half-life of circulating serpins by conferring resistance to proteolytic degradation (123, 124). The impact of glycosylation on the biological function of serpins is less clear. Recombinant Iripin- 3 inhibited the proteolytic activity of some serine proteases, suggesting that its functions dependent on anti-protease activity (like anticoagulant properties) may not be affected by missing glycosylation. However, the absence of glycosylation might have an impact on anti-inflammatory, and immunomodulatory activities of Iripin-3 mediated by its binding to cell surfaces and soluble immune mediators.

For example, only glycosylated, but not non-glycosylated, α-1-antitrypsin was capable of binding IL-8, thus inhibiting IL-8-CXCR1 interaction (125).

CONCLUSION

To conclude, Iripin-3 is a pluripotent salivary protein secreted by *I. ricinus* ticks via saliva into the feeding site, where it might suppress various aspects of host anti-tick defenses. The attenuation of IL-6 production, suppression of CD4$^+$ T cell proliferation, and inhibition of Th1 immune responses have also been observed with other tick serpins and are consistent with the previously reported immunomodulatory effects of *I. ricinus* saliva and salivary gland extracts (114–117). On the other hand, our study is the first to describe the inhibition of the extrinsic pathway of blood coagulation, impaired B and T cell survival, and the induction of Treg differentiation by a tick serpin. The pluripotency and redundancy in Iripin-3 functions are consistent with the theory about the importance of these protein features for successful tick feeding (119). Although several distinct *in vitro* activities of Iripin-3 were observed in this study, their physiological relevance, mechanisms behind them and potential of Iripin-3 to be a candidate for drug or vaccine development remain to be determined. Therefore, further *in vivo* experiments and mechanistic studies are needed to validate and elucidate the Iripin-3 functions described in this work.

DATA AVAILABILITY STATEMENT

The data sets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ Supplementary Material.

ETHICS STATEMENT

All animal experiments were performed in accordance with the Animal Protection Law of the Czech Republic No. 246/1992 Sb. (ethics approval no. 34/2018) and protocols approved by the Ministry of Education, Youth and Sports of the Czech Republic (protocol no. 19085/2015-3) and the responsible

committee of the IP BC CAS. Pathogen-free *I. ricinus* ticks were obtained from the tick colony maintained at the IP BC CAS.

AUTHOR CONTRIBUTIONS

AC designed and performed experiments, analyzed data, and wrote the manuscript. JK, ZB, BK, LAM, HL, TP, ME, and IKS designed and performed experiments and analyzed data. MK edited the manuscript. JC directed the study, designed experiments, analyzed data, and edited the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.626200/full#supplementary-material

REFERENCES

1. Rizzoli A, Silaghi C, Obiegala A, Rudolf I, Hubálek Z, Földvári G, et al. *Ixodes ricinus* and its transmitted pathogens in urban and peri-urban areas in Europe: new hazards and relevance for public health. *Front Public Health* (2014) 2:251. doi: 10.3389/fpubh.2014.00251

2. Castelli E, Caputo V, Morello V, Tomasino RM. Local reactions to tick bites. *Am J Dermatopathol* (2008) 30:241–8. doi: 10.1097/DAD.0b013e3181676b60

3. Heinze DM, Carmical JR, Aronson JF, Thangamani S. Early immunologic events at the tick-host interface. *PLoS One* (2012) 7:e47301. doi: 10.1371/journal.pone.0047301

4. Boppana DKV, Wikel SK, Raj DG, Manohar MB, Lalitha J. Cellular infiltration at skin lesions and draining lymph nodes of sheep infested with adult *Hyalomma anatolicum anatolicum* ticks. *Parasitology* (2005) 131:657–67. doi: 10.1017/S0031182005008243

5. Mbow ML, Rutti B, Brossard M. Infiltration of CD4+ CD8+ T cells, and expression of ICAM-1, Ia antigens, IL-1 alpha and TNF-alpha in the skin lesion of BALB/c mice undergoing repeated infestations with nymphal *Ixodes ricinus* ticks. *Immunology* (1994) 82:596–602.

6. Glatz M, Means T, Haas J, Steere AC, Müllegger RR. Characterization of the early local immune response to *Ixodes ricinus* tick bites in human skin. *Exp Dermatol* (2017) 26:263–9. doi: 10.1111/exd.13207

7. Nithiuthai S, Allen JR. Langerhans cells present tick antigens to lymph node cells from tick-sensitized guinea-pigs. *Immunology* (1985) 55:157–63.

8. KazimírováM, ŠtibrániováI. Tick salivary compounds: their role in modulation of host defences and pathogen transmission. *Front Cell Infect Microbiol* (2013) 3:43. doi: 10.3389/fcimb.2013.00043

9. Šimo L, Kazimirova M, Richardson J, Bonnet SI. The essential role of tick salivary glands and saliva in tick feeding and pathogen transmission. *Front Cell Infect Microbiol* (2017) 7:281. doi: 10.3389/fcimb.2017.00281

10. Mans BJ. Chemical equilibrium at the tick–host feeding interface: a critical examination of biological relevance in hematophagous behavior. *Front Physiol* (2019) 10:530. doi: 10.3389/fphys.2019.00530

11. ŠtibrániováI, Bartı́kováP, Holı́kováV, KazimírováM. Deciphering biological processes at the tick-host interface opens new strategies for treatment of human diseases. *Front Physiol* (2019) 10:830. doi: 10.3389/ fphys.2019.00830

12. Hovius JWR, Levi M, Fikrig E. Salivating for knowledge: potential pharmacological agents in tick saliva. *PLoS Med* (2008) 5:e43. doi: 10.1371/journal.pmed.0050043

13. Chmelar J, Calvo E, Pedra JHF, Francischetti IMB, Kotsyfakis M. Tick salivary secretion as a source of antihemostatics. *J Proteomics* (2012) 75:3842–54. doi: 10.1016/j.jprot.2012.04.026

14. ChmelařJ, Kotál J, KovaříkováA, Kotsyfakis M. The use of tick salivary proteins as novel therapeutics. *Front Physiol* (2019) 10:812. doi: 10.3389/ fphys.2019.00812

15. Rego ROM, Trentelman JJA, Anguita J, Nijhof AM, Sprong H, Klempa B, et al. Counterattacking the tick bite: towards a rational design of anti-tick vaccines targeting pathogen transmission. *Parasit Vectors* (2019) 12:229. doi: 10.1186/s13071-019-3468-x

16. Schwarz A, von Reumont BM, Erhart J, Chagas AC, Ribeiro JMC, Kotsyfakis M. *De novo Ixodes ricinus* salivary gland transcriptome analysis using two next-generation sequencing methodologies. *FASEB J* (2013) 27:4745–56. doi: 10.1096/fj.13-232140

17. Martins LA, Kotál J, Bensaoud C, ChmelařJ, Kotsyfakis M. Small protease inhibitors in tick saliva and salivary glands and their role in tick-host- pathogen interactions. *Biochim Biophys Acta Proteins Proteom* (2020) 1868:140336. doi: 10.1016/j.bbapap.2019.140336

18. Law RH, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, et al. An overview of the serpin superfamily. *Genome Biol* (2006) 7:216. doi: 10.1186/gb-2006-7-5-216

19. Heit C, Jackson BC, McAndrews M, Wright MW, Thompson DC, Silverman GA, et al. Update of the human and mouse SERPIN gene superfamily. *Hum Genomics* (2013) 7:22. doi: 10.1186/1479-7364-7-22

20. Khan MS, Singh P, Azhar A, Naseem A, Rashid Q, Kabir MA, et al. Serpin inhibition mechanism: a delicate balance between native metastable state and polymerization. *J Amino Acids* (2011) 2011:606797. doi: 10.4061/2011/ 606797

21. Porter L, Radulovic Z, Kim T, Braz GRC, Da Silva Vaz I, Mulenga A. Bioinformatic analyses of male and female *Amblyomma americanum* tick expressed serine protease inhibitors (serpins). *Ticks Tick Borne Dis* (2015) 6:16–30. doi: 10.1016/j.ttbdis.2014.08.002

22. Tirloni L, Islam MS, Kim TK, Diedrich JK, Yates JR, Pinto AFM, et al. Saliva from nymph and adult females of *Haemaphysalis longicornis*: a proteomic study. *Parasit Vectors* (2015) 8:338. doi: 10.1186/s13071-015-0918-y

23. Kotsyfakis M, Schwarz A, Erhart J, Ribeiro JMC. Tissue- and time- dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Sci Rep* (2015) 5:9103. doi: 10.1038/srep09103

24. Mulenga A, Khumthong R, Chalaire KC. *Ixodes scapularis* tick serine proteinase inhibitor (serpin) gene family; annotation and transcriptional analysis. *BMC Genomics* (2009) 10:217. doi: 10.1186/1471-2164-10-217

25. de Castro MH, de Klerk D, Pienaar R, Latif AA, Rees DJG, Mans BJ. *De novo* assembly and annotation of the salivary gland transcriptome of *Rhipicephalus appendiculatus* male and female ticks during blood feeding. *Ticks Tick Borne Dis* (2016) 7:536–48. doi: 10.1016/j.ttbdis.2016.01.014

26. Tirloni L, Seixas A, Mulenga A, da Silva Vaz I, Termignoni C. A family of serine protease inhibitors (serpins) in the cattle tick *Rhipicephalus (Boophilus) microplus*. *Exp Parasitol* (2014) 137:25–34. doi: 10.1016/ j.exppara.2013.12.001

27. Rodriguez-Valle M, Xu T, Kurscheid S, Lew-Tabor AE. *Rhipicephalus microplus* serine protease inhibitor family: annotation, expression and functional characterisation assessment. *Parasit Vectors* (2015) 8:7. doi: 10.1186/s13071-014-0605-4

28. Prevot P-P, Adam B, Boudjeltia KZ, Brossard M, Lins L, Cauchie P, et al. Anti-hemostatic effects of a serpin from the saliva of the tick *Ixodes ricinus*. *J Biol Chem* (2006) 281:26361–9. doi: 10.1074/jbc.M604197200

29. Yu Y, Cao J, Zhou Y, Zhang H, Zhou J. Isolation and characterization of two novel serpins from the tick *Rhipicephalus haemaphysaloides*. *Ticks Tick Borne Dis* (2013) 4:297–303. doi: 10.1016/j.ttbdis.2013.02.001

30. Ibelli AMG, Kim TK, Hill CC, Lewis LA, Bakshi M, Miller S, et al. A blood meal-induced *Ixodes scapularis* tick saliva serpin inhibits trypsin and thrombin, and interferes with platelet aggregation and blood clotting. *Int J Parasitol* (2014) 44:369–79. doi: 10.1016/j.ijpara.2014.01.010

31. Kim TK, Tirloni L, Radulovic Z, Lewis L, Bakshi M, Hill C, et al. Conserved *Amblyomma americanum* tick Serpin19, an inhibitor of blood clotting factors Xa and XIa, trypsin and plasmin, has anti-haemostatic functions. *Int J Parasitol* (2015) 45:613–27. doi: 10.1016/j.ijpara.2015.03.009

32. Chmelar J, Oliveira CJ, Rezacova P, Francischetti IMB, Kovarova Z, Pejler G, et al. A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood* (2011) 117:736–44. doi: 10.1182/blood-2010-06-293241

33. Tirloni L, Kim TK, Coutinho ML, Ali A, Seixas A, Termignoni C, et al. The putative role of *Rhipicephalus microplus* salivary serpins in the tick-host relationship. *Insect Biochem Mol Biol* (2016) 71:12–28. doi: 10.1016/ j.ibmb.2016.01.004

34. Tirloni L, Kim TK, Berger M, Termignoni C, da Silva Vaz I, Mulenga A. *Amblyomma americanum* serpin 27 (AAS27) is a tick salivary anti-inflammatory protein secreted into the host during feeding. *PloS Negl Trop Dis* (2019) 13:e0007660. doi: 10.1371/journal.pntd.0007660

35. Coutinho ML, Bizzarro B, Tirloni L, Berger M, Freire Oliveira CJ, Sá-Nunes A, et al. *Rhipicephalus microplus* serpins interfere with host immune responses by specifically modulating mast cells and lymphocytes. *Ticks Tick Borne Dis* (2020) 11:101425. doi: 10.1016/j.ttbdis.2020.101425

36. Kim TK, Tirloni L, Berger M, Diedrich JK, Yates JR, Termignoni C, et al. *Amblyomma americanum* serpin 41 (AAS41) inhibits inflammation by targeting chymase and chymotrypsin. *Int J Biol Macromol* (2020) 156:1007–21. doi: 10.1016/j.ijbiomac.2020.04.088

37. Wang F, Song Z, Chen J, Wu Q, Zhou X, Ni X, et al. The immunosuppressive functions of two novel tick serpins, HlSerpin-a and HlSerpin-b, from *Haemaphysalis longicornis*. *Immunology* (2020) 159:109–20. doi: 10.1111/ imm.13130

38. Leboulle G, Crippa M, Decrem Y, Mejri N, Brossard M, Bollen A, et al. Characterization of a novel salivary immunosuppressive protein from *Ixodes ricinus* ticks. *J Biol Chem* (2002) 277:10083–9. doi: 10.1074/jbc.M111391200

39. Páleniḱová J, Lieskovská J, Langhansová H, Kotsyfakis M, Chmelař J, Kopecký J. *Ixodes ricinus* salivary serpin IRS-2 affects Th17 differentiation via inhibition of the interleukin-6/STAT-3 signaling pathway. *Infect Immun* (2015) 83:1949–56. doi: 10.1128/IAI.03065-14

40. Toyomane K, Konnai S, Niwa A, Githaka N, Isezaki M, Yamada S, et al. Identification and the preliminary *in vitro* characterization of IRIS homologue from salivary glands of *Ixodes persulcatus* Schulze. *Ticks Tick Borne Dis* (2016) 7:119–25. doi: 10.1016/j.ttbdis.2015.09.006

41. Imamura S, Da Silva Vaz I, Sugino M, Ohashi K, Onuma M. A serine protease inhibitor (serpin) from *Haemaphysalis longicornis* as an anti-tick vaccine. *Vaccine* (2005) 23:1301–11. doi: 10.1016/j.vaccine.2004.08.041

42. Prevot P-P, Couvreur B, Denis V, Brossard M, Vanhamme L, Godfroid E. Protective immunity against *Ixodes ricinus* induced by a salivary serpin. *Vaccine* (2007) 25:3284–92. doi: 10.1016/j.vaccine.2007.01.008

43. Kim TK, Radulovic Z, Mulenga A. Target validation of highly conserved *Amblyomma americanum* tick saliva serine protease inhibitor 19. *Ticks Tick Borne Dis* (2016) 7:405–14. doi: 10.1016/j.ttbdis.2015.12.017

44. Xu Z, Yan Y, Cao J, Zhou Y, Zhang H, Xu Q, et al. A family of serine protease inhibitors (serpins) and its expression profiles in the ovaries of *Rhipicephalus haemaphysaloides*. *Infect Genet Evol* (2020) 84:104346. doi: 10.1016/ j.meegid.2020.104346

45. Xu Z, Yan Y, Zhang H, Cao J, Zhou Y, Xu Q, et al. A serpin from the tick *Rhipicephalus haemaphysaloides*: involvement in vitellogenesis. *Vet Parasitol* (2020) 279:109064. doi: 10.1016/j.vetpar.2020.109064

46. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: JM Walker, editor. *The Proteomics Protocols Handbook Springer Protocols Handbooks*. Totowa, NJ: Humana Press (2005). p. 571–607. doi: 10.1385/ 1-59259-890-0:571

47. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* (2011) 8:785–6. doi: 10.1038/nmeth.1701

48. de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* (2006) 34:W362–5. doi: 10.1093/nar/gkl124

49. Mulenga A, Khumthong R, Blandon MA. Molecular and expression analysis of a family of the *Amblyomma americanum* tick Lospins. *J Exp Biol* (2007) 210:3188–98. doi: 10.1242/jeb.006494

50. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT-BG, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J* (2013) 32:1478–88. doi: 10.1038/emboj.2013.79

51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* (1990) 215:403–10. doi: 10.1016/S0022-2836(05) 80360-2

52. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford Engl)* (2007) 23:2947–8. doi: 10.1093/bioinformatics/btm404

53. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* (2014) 42:W320–4. doi: 10.1093/ nar/gku316

54. Mueller U, Förster R, Hellmig M, Huschmann FU, Kastner A, Malecki P, et al. The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: current status and perspectives. *Eur Phys J Plus* (2015) 130:141. doi: 10.1140/epjp/i2015-15141-2

55. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* (1999) 41:95–8. doi: 10.14601/Phytopathol_Mediterr-14998u1.29

56. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* (1992) 8:275–82. doi: 10.1093/bioinformatics/8.3.275

57. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* (1987) 4:406–25. doi: 10.1093/oxfordjournals.molbev.a040454

58. Gascuel O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol* (1997) 14:685–95. doi: 10.1093/oxfordjournals.molbev.a025808

59. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* (2018) 35:1547–9. doi: 10.1093/molbev/msy096

60. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* (2001) 25:402–8. doi: 10.1006/meth.2001.1262

61. Vechtova P, Fussy Z, Cegan R, Sterba J, Erhart J, Benes V, et al. Catalogue of stage-specific transcripts in *Ixodes ricinus* and their potential functions during the tick life-cycle. *Parasit Vectors* (2020) 13:311. doi: 10.1186/ s13071-020-04173-4

62. Koči J, Šimo L, Park Y. Validation of internal reference genes for real-time quantitative polymerase chain reaction studies in the tick, *Ixodes scapularis* (Acari: Ixodidae). *J Med Entomol* (2013) 50:79–84. doi: 10. 1603/me12034

63. KýckováK, KopeckýJ. Effect of tick saliva on mechanisms of innate immune response against *Borrelia afzelii*. *J Med Entomol* (2006) 43:1208–14. doi: 10.1093/jmedent/43.6.1208

64. Pfaffl MW. A new mathematical model for relative quantification in real- time RT–PCR. *Nucleic Acids Res* (2001) 29:e45. doi: 10.1093/nar/29.9.e45

65. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* (2002) 3:research0034.1. doi: 10.1186/gb-2002-3-7-research0034

66. Lew M. Good statistical practice in pharmacology Problem 2. *Br J Pharmacol* (2007) 152:299–303. doi: 10.1038/sj.bjp.0707372

67. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* (1967) 27:157–62. doi: 10.1016/s0006-291x (67)80055-x

68. Hopkins PC, Carrell RW, Stone SR. Effects of mutations in the hinge region of serpins. *Biochemistry* (1993) 32:7650–7. doi: 10.1021/bi00081a008

69. Olsen JV, Ong S-E, Mann M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* (2004) 3:608–14. doi: 10.1074/mcp.T400003-MCP200

70. Kovář#ováZ, ChmelařJ, Šanda M, Brynda J, MarešM, ŘezáčováP.Crystallization and diffraction analysis of the serpin IRS-2 from the hard tick *Ixodes ricinus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* (2010) 66:1453–7. doi: 10.1107/S1744309110032343

71. Gettins PGW. Serpin structure, mechanism, and function. *Chem Rev* (2002) 102:4751–804. doi: 10.1021/cr010170+

72. ChmelařJ, Kotál J, LanghansdocáH, Kotsyfakis M. Protease inhibitors in tick saliva: the role of serpins and cystatins in tick-host-pathogen interaction. *Front Cell Infect Microbiol* (2017) 7:216. doi: 10.3389/fcimb.2017.00216

73. Raber MN. Coagulation Tests, in: *Clinical Methods: The History, Physical, and Laboratory Examinations*. Boston: Butterworths. Available at: http://www.ncbi.nlm.nih.gov/books/NBK265/ (Accessed October 12, 2020).

74. Bakshi M, Kim TK, Porter L, Mwangi W, Mulenga A. *Amblyomma americanum* ticks utilizes countervailing pro and anti-inflammatory proteins to evade host defense. *PLoS Pathog* (2019) 15:e1008128. doi: 10.1371/journal.ppat.1008128

75. Zheng W, Flavell RA. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4+ T cells. *Cell* (1997) 89:587–96. doi: 10.1016/S0092-8674(00)80240-8

76. Szabo SJ, Kim ST, Costa GL, Zhang X, Fathman CG, Glimcher LH. A novel transcription factor, T-bet, directs Th1 lineage commitment. *Cell* (2000) 100:655–69. doi: 10.1016/S0092-8674(00)80702-3

77. Fontenot JD, Gavin MA, Rudensky AY. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat Immunol* (2003) 4:330– 6. doi: 10.1038/ni904

78. Hori S, Nomura T, Sakaguchi S. Control of regulatory T cell development by the transcription factor Foxp3. *Science* (2003) 299:1057–61. doi: 10.1126/science.1079490

79. Ivanov II, McKenzie BS, Zhou L, Tadokoro CE, Lepelley A, Lafaille JJ, et al. The orphan nuclear receptor RORgt directs the differentiation program of proinflammatory IL-17+ T helper cells. *Cell* (2006) 126:1121–33. doi: 10.1016/j.cell.2006.07.035

80. Francischetti IMB, Sá-Nunes A, Mans BJ, Santos IM, Ribeiro JMC. The role of saliva in tick feeding. *Front Biosci* (2009) 14:2051–88. doi: 10.2741β363

81. O'Reilly MS. Antiangiogenic antithrombin. *Semin Thromb Hemost* (2007) 33:660–6. doi: 10.1055/s-2007-991533

82. Rau JC, Beaulieu LM, Huntington JA, Church FC. Serpins in thrombosis, hemostasis and fibrinolysis. *J Thromb Haemost* (2007) 5:102–15. doi: 10.1111/j.1538-7836.2007.02516.x

83. Gál P, DobóJ, Beinrohr L, Pál G, Závodszky P. Inhibition of the serine proteases of the complement system. *Adv Exp Med Biol* (2013) 735:23–40. doi: 10.1007/978-1-4614-4118-2_2

84. Mkaouar H, Akermi N, Kriaa A, Abraham A-L, Jablaoui A, Soussou S, et al. Serine protease inhibitors and human wellbeing interplay: new insights for old friends. *PeerJ* (2019) 7:e7224. doi: 10.7717/peerj.7224

85. Harris JL, Backes BJ, Leonetti F, Mahrus S, Ellman JA, Craik CS. Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc Natl Acad Sci U S A* (2000) 97:7754–9. doi: 10.1073/ pnas.140132697

86. Björkqvist J, Jämsä A, RennéT. Plasma kallikrein: the bradykinin-producing enzyme. *Thromb Haemost* (2013) 110:399–407. doi: 10.1160/TH13-03-0258

87. Paterson KJ, Zambreanu L, Bennett DLH, McMahon SB. Characterisation and mechanisms of bradykinin-evoked pain in man using iontophoresis. *Pain* (2013) 154:782–92. doi: 10.1016/j.pain.2013.01.003

88. List K, Bugge TH, Szabo R. Matriptase: potent proteolysis on the cell surface. *Mol Med* (2006) 12:1–7. doi: 10.2119/2006-00022.List

89. Chen Y-W, Wang J-K, Chou F-P, Wu B-Y, Hsiao H-C, Chiu H, et al. Matriptase regulates proliferation and early, but not terminal, differentiation of human keratinocytes. *J Invest Dermatol* (2014) 134:405–14. doi: 10.1038/ jid.2013.320

90. Bardou O, Menou A, François C, Duitman JW, von der Thüsen JH, Borie R, et al. Membrane-anchored serine protease matriptase is a trigger of pulmonary fibrogenesis. *Am J Respir Crit Care Med* (2016) 193:847–60. doi: 10.1164/rccm.201502-0299OC

91. Seitz I, Hess S, Schulz H, Eckl R, Busch G, Montens HP, et al. Membrane- type serine protease-1/matriptase induces interleukin-6 and -8 in endothelial cells by activation of protease-activated receptor-2: potential implications in atherosclerosis. *Arterioscler Thromb Vasc Biol* (2007) 27:769–75. doi: 10.1161/01.ATV.0000258862.61067.14

92. Sugino M, Imamura S, Mulenga A, Nakajima M, Tsuda A, Ohashi K, et al. A serine proteinase inhibitor (serpin) from ixodid tick *Haemaphysalis longicornis*; cloning and preliminary assessment of its suitability as a candidate for a tick vaccine. *Vaccine* (2003) 21:2844–51. doi: 10.1016/ S0264-410X(03)00167-1

93. Palta S, Saroa R, Palta A. Overview of the coagulation system. *Indian J Anaesth* (2014) 58:515–23. doi: 10.4103/0019-5049.144643

94. Larsen KS, Østergaard H, Bjelke JR, Olsen OH, Rasmussen HB, Christensen L, et al. Engineering the substrate and inhibitor specificities of human coagulation Factor VIIa. *Biochem J* (2007) 405:429–38. doi: 10.1042/ BJ20061901

95. Lawson JH, Butenas S, Ribarik N, Mann KG. Complex-dependent inhibition of factor VIIa by antithrombin III and heparin. *J Biol Chem* (1993) 268:767– 70. doi: 10.1016/S0021-9258(18)53998-3

96. Rao LV, Rapaport SI, Hoang AD. Binding of factor VIIa to tissue factor permits rapid antithrombin III/heparin inhibition of factor VIIa. *Blood* (1993) 81:2600–7. doi: 10.1182/blood.V81.10.2600.2600

97. Fortenberry YM, Hlavacek AC, Church FC. Protein C inhibitor inhibits factor VIIa when bound to tissue factor. *J Thromb Haemost* (2011) 9:861–3. doi: 10.1111/j.1538-7836.2011.04196.x

98. Prevot P-P, Beschin A, Lins L, Beaufays J, Grosjean A, Bruys L, et al. Exosites mediate the anti-inflammatory effects of a multifunctional serpin from the saliva of the tick *Ixodes ricinus*. *FEBS J* (2009) 276:3235–46. doi: 10.1111/ j.1742-4658.2009.07038.x

99. Aldonyte R, Jansson L, Janciauskiene S. Concentration-dependent effects of native and polymerised a1-antitrypsin on primary human monocytes, *in vitro*. *BMC Cell Biol* (2004) 5:11. doi: 10.1186/1471-2121-5-11

100. Bettelli E, Carrier Y, Gao W, Korn T, Strom TB, Oukka M, et al. Reciprocal developmental pathways for the generation of pathogenic effector Th17 and regulatory T cells. *Nature* (2006) 441:235–8. doi: 10.1038/nature04753

101. Korn T, Mitsdoerffer M, Croxford AL, Awasthi A, Dardalhon VA, Galileos G, et al. IL-6 controls Th17 immunity *in vivo* by inhibiting the conversion of conventional T cells into Foxp3+ regulatory T cells. *Proc Natl Acad Sci U S A* (2008) 105:18460–5. doi: 10.1073/pnas.0809850105

102. Chen X, Das R, Komorowski R, Beres A, Hessner MJ, Mihara M, et al. Blockade of interleukin-6 signaling augments regulatory T-cell reconstitution and attenuates the severity of graft-versus-host disease. *Blood* (2009) 114:891–900. doi: 10.1182/blood-2009-01-197178

103. Plitas G, Rudensky AY. Regulatory T cells: differentiation and function. *Cancer Immunol Res* (2016) 4:721–5. doi: 10.1158/2326-6066.CIR-16-0193

104. Ferreira BR, Silva JS. Successive tick infestations selectively promote a T-helper 2 cytokine profile in mice. *Immunology* (1999) 96:434–9. doi: 10.1046/ j.1365-2567.1999.00683.x

105. Heinze DM, Wikel SK, Thangamani S, Alarcon-Chaidez FJ. Transcriptional profiling of the murine cutaneous response during initial and subsequent infestations with *Ixodes scapularis* nymphs. *Parasit Vectors* (2012) 5:26. doi: 10.1186/1756-3305-5-26

106. Ashton-Rickardt PG. An emerging role for Serine Protease Inhibitors in T lymphocyte immunity and beyond. *Immunol Lett* (2013) 152:65–76. doi: 10.1016/j.imlet.2013.04.004

107. Gao G, Shao C, Zhang SX, Dudley A, Fant J, Ma J-X. Kallikrein-binding protein inhibits retinal neovascularization and decreases vascular leakage. *Diabetologia* (2003) 46:689–98. doi: 10.1007/s00125-003-1085-9

108. Latha K, Zhang W, Cella N, Shi HY, Zhang M. Maspin mediates increased tumor cell apoptosis upon induction of the mitochondrial permeability transition. *Mol Cell Biol* (2005) 25:1737–48. doi: 10.1128/MCB.25.5.1737- 1748.2005

109. Chen L, Zhang SS-M, Barnstable CJ, Tombran-Tink J. PEDF induces apoptosis in human endothelial cells by activating p38 MAP kinase dependent cleavage of multiple caspases. *Biochem Biophys Res Commun* (2006) 348:1288–95. doi: 10.1016/j.bbrc.2006.07.188

110. Becerra SP, Notario V. The effects of PEDF on cancer biology: mechanisms of action and therapeutic potential. *Nat Rev Cancer* (2013) 13:258–71. doi: 10.1038/nrc3484

111. Yao Y, Li L, Huang X, Gu X, Xu Z, Zhang Y, et al. SERPINA3K induces apoptosis in human colorectal cancer cells via activating the Fas/FasL/caspase-8 signaling pathway. *FEBS J* (2013) 280:3244–55. doi: 10.1111/febs.12303

112. Bröker LE, Kruyt FAE, Giaccone G. Cell death independent of caspases: a review. *Clin Cancer Res* (2005) 11:3155–62. doi: 10.1158/1078-0432.CCR-04- 2223

113. Dhuriya YK, Sharma D. Necroptosis: a regulated inflammatory mode of cell death. *J Neuroinflamm* (2018) 15:199. doi: 10.1186/s12974-018-1235-0

114. Kovár L, KopeckýJ, RıhováB. Salivary gland extract from *Ixodes ricinus* tick polarizes the cytokine profile toward Th2 and suppresses proliferation of T lymphocytes in human PBMC culture. *J Parasitol* (2001) 87:1342–8. doi: 10.1645/0022-3395(2001)087[1342:SGEFIR]2.0.CO;2

115. Kovár L, KopeckýJ, RıhováB. Salivary gland extract from *Ixodes ricinus* tick modulates the host immune response towards the Th2 cytokine profile. *Parasitol Res* (2002) 88:1066–72. doi: 10.1007/s00436-002-0714-4

116. Mejri N, Rutti B, Brossard M. Immunosuppressive effects of *Ixodes ricinus* tick saliva or salivary gland extracts on innate and acquired immune response of BALB/c mice. *Parasitol Res* (2002) 88:192–7. doi: 10.1007/ s00436-001-0515-1

117. SkallováA, Iezzi G, Ampenberger F, Kopf M, Kopecky J. Tick saliva inhibits dendritic cell migration, maturation, and function while promoting development of Th2 responses. *J Immunol* (2008) 180:6186–92. doi: 10.4049/jimmunol.180.9.6186

118. Arce-Sillas A, Álvarez-LuquínDD, Tamaya-Domınguez B, Gomez-FuentesS, Trejo-GarcıáA, Melo-Salas M, et al. Regulatory T cells: molecular actions on effector cells in immune regulation. *J Immunol Res* (2016) 2016:1720827. doi: 10.1155/2016/1720827

119. ChmelařJ, Kotál J, Kopecky J, Pedra JH, Kotsyfakis M. All for one and one for all on the tick-host battlefield. *Trends Parasitol* (2016) 32:368–77. doi: 10.1016/j.pt.2016.01.004

120. Kotál J, Stergiou N, BusăM, ChlastákováA, BeránkováZ, ŘezáčováP, et al. The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin. *Cell Mol Life Sci* (2019) 76:2003–13. doi: 10.1007/s00018- 019-03034-3

121. Blisnick AA, Šimo L, Grillon C, Fasani F, BrûléS, Le Bonniec B, et al. The immunomodulatory effect of IrSPI, a tick salivary gland serine protease inhibitor involved in *Ixodes ricinus* tick feeding. *Vaccines (Basel)* (2019) 7:148. doi: 10.3390/vaccines7040148

122. Visentin C, Broggini L, Sala BM, Russo R, Barbiroli A, Santambrogio C, et al. Glycosylation tunes neuroserpin physiological and pathological properties. *Int J Mol Sci* (2020) 21:3235. doi: 10.3390/ijms21093235

123. Sarkar A, Wintrode PL. Effects of glycosylation on the stability and flexibility of a metastable protein: the human serpin a(1)-antitrypsin. *Int J Mass Spectrom* (2011) 302:69–75. doi: 10.1016/j.ijms.2010.08.003

124. Kwon KS, Yu MH. Effect of glycosylation on the stability of alpha1- antitrypsin toward urea denaturation and thermal deactivation. *Biochim Biophys Acta* (1997) 1335:265–72. doi: 10.1016/s0304-4165 (96)00143-2

125. Bergin DA, Reeves EP, Meleady P, Henry M, McElvaney OJ, Carroll TP, et al. a-1 Antitrypsin regulates human neutrophil chemotaxis induced by soluble immune complexes and IL-8. *J Clin Invest* (2010) 120:4236–50. doi: 10.1172/ JCI41196

3.1.1 Supplementary Material

MATERIALS AND METHODS

Production of recombinant Iripin-3

A full-length Iripin-3 sequence was obtained during a salivary gland transcriptome project (1) and was submitted to GenBank under accession number GADI01004776.1. The Iripin-3 nucleotide sequence without a signal peptide and with an inserted ATG codon to its 5'-terminus was cloned into pET-17b vector (Novagen), and the resulting plasmid was transformed into BL21(DE3) pLysS chemically competent *E. coli cells* (Thermo Fisher Scientific). The cells were grown in LB medium containing ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml), and when the OD600 of the culture reached approximately 0.7, isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM) was added to induce gene expression. The cells were harvested after 3.5-h incubation in the presence of IPTG, and the cell pellet was resuspended in 20 mM Tris-HCl (pH 8). Cell disruption by sonication in inclusion bodies isolation buffer (20 mM Tris-HCl, 1% Triton X-100, pH 8) and repeated washing with 20 mM Tris-HCl (pH 8) resulted in a pellet that contained mainly inclusion bodies of insoluble recombinant Iripin-3. The inclusion bodies were dissolved by stirring in 6 M guanidine hydrochloride (pH 8) and 10 mM dithiotreitol for 1 h at room temperature. Following centrifugation, the supernatant, which contained denatured Iripin-3 released from the inclusion bodies, was diluted 150-fold in refolding buffer (20 mM Tris-HCl, 150 mM NaCl, pH 8), and the mixture was incubated overnight at 4 °C. The precipitated protein was removed by filtration through filter paper and Steritop-GP (Millipore) and concentrated with a stirred chamber concentrator (Millipore). Properly refolded Iripin-3 was then dialyzed against 20 mM Tris-HCl (pH 8) and purified on a HiLoad 26/60 Superdex 200 pg gel filtration column (Cytiva) and a Mono Q column (Cytiva) with the 0-1 M gradient of NaCl. Endotoxin was removed by the company ARVYS Proteins, Inc. via a detergent-based method.

Crystallization

Crystallization experiments were performed in Swissci 96-well 2-drop MRC crystallization plates (Molecular Dimensions Ltd.) using the sitting-drop vapor diffusion technique and OryxNano crystallization robot (Douglas Instruments Ltd.). Suitable Iripin-3 concentration (1.88 mg/ml) was determined by the PCT Pre-Crystallization Test (Hampton Research). Iripin-3 crystals were grown with the precipitant composed of 0.2 M potassium thiocyanate, 0.1 M sodium cacodylate and 8% w/v γ-polyglutamic acid, pH 6.5 at 21 °C. The protein-to-precipitant solution ratios 2 µl:1 µl or 1 µl:1 µl were equilibrated against 50 µl of reservoir solution.

X-ray data collection and structure determination

Freshly grown crystals were flash frozen in a liquid nitrogen stream without additional cryoprotection, and X-ray diffraction data were collected at

the BESSY II electron storage ring on the beamline BL14.1 operated by the Helmholtz-Zentrum Berlin (2). Data were processed using the *XDS* Program Package (3) with the *XDSAPP* graphical user interface (4). The best diffracting crystal exhibited symmetry of the $P6_222$ space group and contained one molecule in the asymmetric unit. The structure of Iripin-3 was solved by the molecular replacement method using *MOLREP* (5). Of all the structures deposited in the PDB, IRS-2 (PDB code 3NDA) (6) displayed the highest sequence identity (56%) to Iripin-3 and was therefore used as a search model. The Iripin-3 structure was refined with the program *REFMAC5* (7) from the *CCP4* suite (8) and manually rebuilt in *Coot* (9). *MolProbity* (10) and wwPDB (11) were used for final qualitative validation of the model. Figures with the Iripin-3 structure were made in the *PyMOL* Molecular Graphic System (Schrödinger, LLC). Atomic coordinates were deposited in the PDB under accession code 7AHP. Data collection, processing and refinement statistics are summarized in **Supplementary Table 1**.

Presence of Iripin-3 in tick saliva (ELISA)

Each well of a Corning 96-well microplate was coated overnight at 4 °C with 50 µl of tick saliva diluted with coating buffer (15 mM Na_2CO_3, 35 mM $NaHCO_3$, pH 9.6) to a final concentration of 10 µg/ml. The unoccupied binding sites of the plate were blocked by the addition of 300 µl/well of blocking buffer (5% precolostral calf serum in PBS) for 1 h at 37 °C. After washing the plate three times with wash buffer (0.05% Tween 20 in PBS), 50 µl of pre-immune serum or antiserum against Iripin-3, pre-diluted 1:1000 with 2% precolostral calf serum in PBS, was added and incubated for 1 h at 37 °C. Another three washes with wash buffer were followed by the addition of 100 µl/well of peroxidase-conjugated goat antibody recognizing rabbit immunoglobulin G (Sigma-Aldrich). The antibody was pre-diluted 1:1000 with 2% precolostral calf serum in PBS and incubated for 1 h at 37 °C. At the end of incubation, the plate was again washed three times with wash buffer, and then 100 µl of a substrate solution (51.4 mM $Na_2HPO_4 \cdot 12 H_2O$, 24.3 mM $C_6H_8O_7 \cdot H_2O$, 3.7 mM *o*-phenylenediamine, and 0.012% H_2O_2, pH 5) was applied to each well. The enzymatic reaction was stopped by the addition of 2 M H_2SO_4, and the optical density was measured at 490 nm on Synergy H1 microplate reader (BioTek Instruments, Inc.).

Presence of Iripin-3 in tick saliva (Western blot)

Tick saliva (10 µg) and Iripin-3 (1 ng or 10 ng) were subjected to SDS-PAGE using 10% gel, and the separated proteins were transferred to polyvinylidene difluoride (PVDF) membranes (Thermo Fisher Scientific). Following blocking in Tris-buffered saline containing 5% fat-free milk and 0.1% Tween 20 for 1 h at room temperature, the blots were incubated overnight at 4 °C with pre-immune serum or antiserum against Iripin-3. Both sera were pre-diluted 1:100 in Tris-buffered saline containing 5% milk and

0.1% Tween 20. After washing, the membranes were incubated with goat anti-rabbit antibody conjugated with horseradish peroxidase (Cell Signaling Technology) for 1 h at room temperature. The secondary antibody was pre-diluted 1:2000 in the same solution as primary antibodies. The proteins were visualized using the enhanced chemiluminescent substrate WesternBright Quantum (Advansta), and the signal was detected using a charge-coupled device (CCD) imaging system (Uvitec Ltd.).

Inhibition of serine proteases

Firstly, the formation of SDS- and heat-stable complexes between Iripin-3 and selected serine proteases (kallikrein, matriptase, plasmin, thrombin, trypsin, factor VIIa) was tested. All enzymes used were of human origin. Kallikrein and thrombin were purchased from Sigma-Aldrich, matriptase and trypsin were purchased from R&D Systems and plasmin and factor VIIa were obtained from Haematologic Technologies, Inc. Iripin-3 and proteases were diluted in assay buffer corresponding to each protease (described below), and then each of the six serine proteases was incubated with Iripin-3 at equimolar concentrations (1 μM) for 1 h at room temperature. Factor VIIa was incubated with Iripin-3 in the absence or presence of human tissue factor (1 μM, BioLegend). The 1-h incubation was followed by the addition of NuPAGE LDS Sample Buffer (Thermo Fisher Scientific) together with dithiotreitol and boiling of samples for 10 min. Finally, samples were analyzed by SDS-PAGE, and protein bands were visualized by silver staining.

Secondly, second-order rate constants of protease inhibition were measured by a discontinuous method under pseudo first-order conditions, using at least a 50-fold molar excess of Iripin-3 over serine proteases. Reactions were incubated at room temperature and were stopped for each time point by the addition of the fluorogenic substrate appropriate for the protease used. The slope of the linear part of fluorescence increases over time gave the residual protease activity at each time point. The apparent (observed) first-order rate constant k_{obs} was calculated from the slope of the plot of the natural log of residual protease activity against time. K_{obs} was measured for six different Iripin-3 concentrations and plotted against the serpin concentration. The slope of the line of best fit gave an estimate of the second-order rate constant k_2. The assay buffer was 20 mM Tris, 150 mM NaCl, 0.02% Triton X-100, pH 8.5 for kallikrein and plasmin; 50 mM Tris, 50 mM NaCl, 0.01% Tween 20, pH 9.0 for matriptase; 20 mM Tris, 150 mM NaCl, 0.01% Triton X-100, 5 mM $CaCl_2$, 0.1% polyethylene glycol 6000, pH 8.0 for thrombin and factor VIIa; and 50 mM Tris, 150 mM NaCl, 10 mM $CaCl_2$, 0.05% Brij 35, pH 7.5 for trypsin. The fluorogenic substrates used were Z-FR-AMC for kallikrein, Boc-QAR-AMC for matriptase and Boc-VPR-AMC for trypsin and thrombin. All substrates were used at 250 μM final concentration. Kallikrein, matriptase, thrombin and trypsin were used in 200 pM, 500 pM, 20 pM and 2 pM final concentrations, respectively.

Pro-inflammatory cytokine expression in bone marrow-derived macrophages (RT-qPCR)

Total RNA was isolated from macrophages using TRI Reagent (Molecular Research Center, Inc.) according to the manufacturer's instructions. Extracted RNA (500 ng) was treated with DNase I (Thermo Fisher Scientific) and reverse transcribed into cDNA using M-MLV Reverse Transcriptase (Thermo Fisher Scientific), as detailed in the protocol of the manufacturer. The resulting cDNA mixed with Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher Scientific) and gene-specific primers was used for the analysis of *Tnf*, *Il6* and *Il1b* expression in the thermal cycler QuantStudio 6 (Thermo Fisher Scientific). Cycling conditions were 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 60 s. Relative gene expression was calculated using the delta-delta Ct method since the amplification efficiencies of target genes and a reference gene (*Gapdh*) were approximately equal (12). Nucleotide sequences of forward and reverse primers as well as amplicon lengths are provided in **Supplementary Table 3**.

Viability of macrophages, dendritic cells and neutrophils

Dendritic cells and macrophages were obtained as described before (13). Briefly, dendritic cells were derived from bone marrow cells of a C57BL/6N mouse by incubating the cells for 8 days in the presence of 20 ng/ml of GM-CSF (Sigma-Aldrich). Macrophages were obtained from bone marrow cells by 7-day incubation in the presence of L929 cell-conditioned medium. Neutrophils were isolated from bone marrow cells by magnetic separation using Neutrophil Isolation Kit (Miltenyi Biotec). Dendritic cells, macrophages and neutrophils were resuspended in RPMI 1640 medium with stable glutamine (Biosera) supplemented with 10% heat-inactivated FBS (Biosera), 50 μM 2-mercaptoethanol (Sigma-Aldrich), 100 U/ml penicillin G (Biosera) and 100 μg/ml streptomycin (Biosera) and then were treated with four different concentrations of Iripin-3 for 1 h at 37 °C and 5% CO_2. Subsequently, macrophages and neutrophils were stimulated by the addition of LPS (100 ng/ml, Sigma-Aldrich, *E. coli* serotype O111:B4). Dendritic cells were left unstimulated. After incubating the cells for 20 h at 37 °C and 5% CO_2, alamarBlue HS Cell Viability Reagent (Thermo Fisher Scientific) was added. The fluorescence intensity was measured on Synergy H1 microplate reader (BioTek Instruments, Inc.; excitation 550 nm; emission 590 nm) 4 h (a 24-h incubation period) and 28 h (a 48-h incubation period) following alamarBlue addition.

Supplementary Table 1. X-ray data collection, processing and refinement statistics.

Data collection	
X-ray source	BL14.1, BESSY II, Germany
Wavelength (Å)	0.9184
Detector	PILATUS 6M
Crystal-detector distance (mm)	222.687
Rotation range per image (°)	0.1
Total rotation range (°)	240
Exposure time per image (s)	0.25
Resolution range (Å)	48.32-1.95 (2.07-1.95)
Space group	$P6_222$
Unit-cell dimensions: a, b, c (Å)	132.94, 132.94, 88.89
Unit-cell dimensions: α, β, γ (°)	90.0, 90.0, 120.0
Mosaicity (°)	0.135
Total number of reflections	889221 (130341)
Number of unique reflections	34278 (5397)
Multiplicity	25.94 (24.15)
Average I/σ(I)	13.45 (2.45)
Completeness (%)	99.9 (99.4)
CC ½	99.8 (82.6)
R_{meas} (%)[a]	25.9 (157.3)
Overall B factor from Wilson plot (Å^2)	29.14
Refinement	
Resolution range (Å)	48.32-1.95 (2.07-1.95)
Number of reflections in working set	32559 (2340)
Final R value (%)[b] / Final R_{free} value (%)[c]	19.16 / 22.28
Mean B value (Å)	24.26
Number of atoms in the asymmetric unit	
Protein	2881
Ligand-Tris ion	3
Water	243
Total	3148
Root-mean-square deviations	
Bonds (Å)	0.015
Angles (°)	1.709
Average B factors (Å^2) Overall	24.264
Ramachandran plot	
Most favored (%)	98.90
Allowed (%)	100.00
PDB code	7AHP

The data in parentheses refer to the highest-resolution shell.

[a] $R_{meas} = (|I_{hkl} - \langle I \rangle|)/I_{hkl}$, where the average intensity $\langle I \rangle$ is taken over all symmetry equivalent measurements, and I_{hkl} is the measured intensity for any given reflection.

[b] R value = $||F_o| - |F_c||/|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

[c] R_{free} is equivalent to R value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process.

Supplementary Table 2. GenBank accession numbers of serpins used in the phylogenetic analysis.

| Serpin	Species	GenBank accession number	Reference
A1AT	*Homo sapiens*	AAB59495.1	(14)
AamS6		ABS87358.1	(15)
AAS19	*Amblyomma americanum*	JAI08902.1	(16)
AAS27		JAI08961.1	
AAS41[a]		JAI08957.1	(16,17)
AAS46[a]		JAI08784.1	
HLS1[b]	*Haemaphysalis longicornis*	Not found	(18)
HLS2		BAD11156.1	(19)
HlSerpin-a		QFQ50847.1	(20)
HlSerpin-b		QFQ50848.1	
Ipis-1	*Ixodes persulcatus*	BAP59746.1	(21)
Iripin-3	*Ixodes ricinus*	JAA69032.1	
Iris		CAB55818.2	(22)
IRS-2		ABI94056.2	(6)
IxscS-1E1	*Ixodes scapularis*	AID54718.1	(23)
RAS-1	*Rhipicephalus appendiculatus*	AAK61375.1	(24)
RAS-2		AAK61376.1	
RAS-3		AAK61377.1	
RAS-4		AAK61378.1	
RHS-1	*Rhipicephalus haemaphysaloides*	AFX65224.1	(25)
RHS-2		AFX65225.1	
RHS8		QHU78941.1	(26)
RmS-1	*Rhipicephalus microplus*	AHC98652.1	(27)
RmS-3		AHC98654.1	
RmS-6		AHC98657.1	
RmS-15		AHC98666.1	
RmS-17		AHC98668.1	
rSERPIN[b]		Not found	(28)

[a] Full-length protein sequences of AAS41 and AAS46 were obtained from the cited article (17) since GenBank contains only partial amino acid sequences of these two *A. americanum* serpins.

[b] In the case of two tick serpins (HLS1 and rSERPIN), no accession number was found, and therefore the amino acid sequences needed for phylogenetic analysis were derived directly from the cited articles.

Supplementary Table 3. Sequences of primers used in the study.

Gene	Species	Sequence	Amplicon length (bp)
Iripin-3 expression in ticks			
EF-1	*Ixodes ricinus*	Forward: 5´-CTGGGTGTGAAGCAGATGAT-3´ Reverse: 5´-GTAGGCAGACACTTCCTTCTG-3´	105
Iripin-3		Forward: 5´-CACAGCGGCAATTCATTTAGG-3´ Reverse: 5´-CGGTACGTCTCTTCTGAAACTC-3´	269
Pro-inflammatory cytokine expression in macrophages			
Gapdh	*Mus musculus*	Forward: 5'-TGTGTCCGTCGTGGATCTGA-3' Reverse: 5'-TTGCTGTTGAAGTCGCAGGAG-3'	150
Il1b		Forward: 5'-TGACCTGGGCTGTCCTGATG-3' Reverse: 5'-GGTGCTCATGTCCTCATCCTG-3'	160
Il6		Forward: 5'-CTGCAAGAGACTTCCATCCAG-3' Reverse: 5'-AGTGGTATAGACAGGTCTGTTGG-3'	131
Tnf		Forward: 5'-CCCCAAAGGGATGAGAAGTTC-3' Reverse: 5'-GGCTTGTCACTCGAATTTTGAGA-3'	101
Transcription factor expression in CD4+ T cells			
Actb	*Mus musculus*	Forward: 5'-CTCTGGCTCCTAGCACCATGAAGA-3' Reverse: 5'-GTAAAACGCAGCTCAGTAACAGTCCG-3'	200
Foxp3		Forward: 5'-CAGCTCTGCTGGCGAAAGTG-3' Reverse: 5'-TCGTCTGAAGGCAGAGTCAGGA-3'	190
Gapdh		Forward: 5'-TGTGTCCGTCGTGGATCTGA-3' Reverse: 5'-TTGCTGTTGAAGTCGCAGGAG-3'	150
Gata3		Forward: 5'-CTCGGCCATTCGTACATGGAA-3' Reverse: 5'-GGATACCTCTGCACCGTAGC-3'	134
Rorc		Forward: 5'-ACGGCCCTGGTTCTCATCA-3' Reverse: 5'-CCAAATTGTATTGCAGATGTTCCAC-3'	79
Tbx21		Forward: 5'-TCAACCAGCACCAGACAGAGA-3' Reverse: 5'-TCCACCAAGACCACATCCAC-3'	130

RESULTS
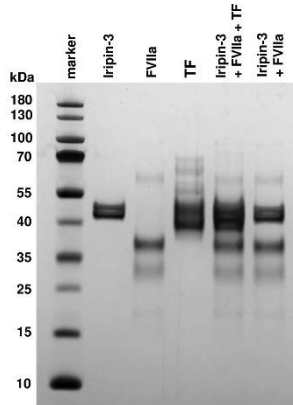
Supplementary Figure 1. Iripin-3 does not form a covalent complex with activated factor VII (FVIIa). A high molecular weight complex formation between FVIIa and Iripin-3 in the absence or presence of tissue factor (TF) was analyzed by using SDS-PAGE. Proteins were resolved on 4 to 12% NuPAGE Bis-Tris gel and were visualized by silver staining.
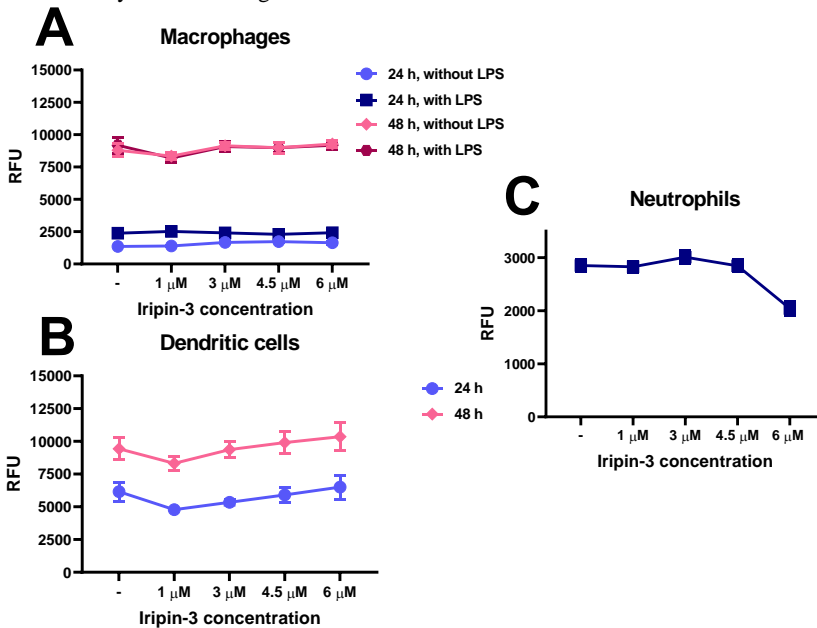
Supplementary Figure 2. Iripin-3 does not negatively affect the viability of macrophages, dendritic cells and neutrophils. The viability of the cells was evaluated by their ability to reduce virtually non-fluorescent resazurin, the active ingredient of alamarBlue, to highly fluorescent resorufin. **(A, B)** The viability of unstimulated or LPS-stimulated macrophages **(A)** and unstimulated dendritic cells **(B)** after exposure to four different concentrations of Iripin-3 for either 24 h or 48 h. Macrophages and dendritic cells unexposed to Iripin-3 were used as control. **(C)** The viability of LPS-stimulated neutrophils untreated with Iripin-3 or treated with four different concentrations of Iripin-3 for 24 h. All data in **(A, B, C)** are presented as mean ± SEM. The experiment was performed only once. RFU, relative fluorescence unit.

REFERENCES

1. Schwarz A, von Reumont BM, Erhart J, Chagas AC, Ribeiro JMC, Kotsyfakis M. *De novo Ixodes ricinus* salivary gland transcriptome analysis using two next-generation sequencing methodologies. *FASEB J* (2013) **27**:4745–4756. doi:10.1096/fj.13-232140

2. Gerlach M, Mueller U, Weiss MS. The MX beamlines BL14.1-3 at BESSY II. *JLSRF* (2016) **2**:47. doi:10.17815/jlsrf-2-64

3. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr* (2010) **66**:125–132. doi:10.1107/S0907444909047337

4. Sparta KM, Krug M, Heinemann U, Mueller U, Weiss MS. XDSAPP2.0. *J Appl Cryst* (2016) **49**:1085–1092. doi:10.1107/S1600576716004416

5. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J Appl Cryst* (1997) **30**:1022–1025. doi:10.1107/S0021889897006766

6. Chmelar J, Oliveira CJ, Rezacova P, Francischetti IMB, Kovarova Z, Pejler G, Kopacek P, Ribeiro JMC, Mares M, Kopecky J, et al. A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood* (2011) **117**:736–744. doi:10.1182/blood-2010-06-293241

7. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* (2011) **67**:355–367. doi:10.1107/S0907444911001314

8. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* (2011) **67**:235–242. doi:10.1107/S0907444910045749

9. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* (2010) **66**:486–501. doi:10.1107/S0907444910007493

10. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* (2010) **66**:12–21. doi:10.1107/S0907444909042073

11. Gore S, Velankar S, Kleywegt GJ. Implementing an X-ray validation pipeline for the Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* (2012) **68**:478–483. doi:10.1107/S0907444911050359

12. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* (2001) **25**:402–408. doi:10.1006/meth.2001.1262

13. Lieskovská J, Páleníková J, Širmarová J, Elsterová J, Kotsyfakis M, Campos Chagas A, Calvo E, Růžek D, Kopecký J. Tick salivary cystatin sialostatin L2 suppresses IFN responses in mouse dendritic cells. *Parasite Immunol* (2015) **37**:70–78. doi:10.1111/pim.12162

14. Long GL, Chandra T, Woo SL, Davie EW, Kurachi K. Complete sequence of the cDNA for human alpha 1-antitrypsin and the gene for the S variant. *Biochemistry* (1984) **23**:4828–4837. doi:10.1021/bi00316a003

15. Mulenga A, Khumthong R, Blandon MA. Molecular and expression analysis of a family of the *Amblyomma americanum* tick Lospins. *J Exp Biol* (2007) **210**:3188–3198. doi:10.1242/jeb.006494

16. Porter L, Radulovic Z, Kim T, Braz GRC, Da Silva Vaz I, Mulenga A. Bioinformatic analyses of male and female *Amblyomma americanum* tick expressed serine protease inhibitors (serpins). *Ticks Tick Borne Dis* (2015) **6**:16–30. doi:10.1016/j.ttbdis.2014.08.002

17. Kim TK, Tirloni L, Berger M, Diedrich JK, Yates JR, Termignoni C, da Silva Vaz I, Mulenga A. *Amblyomma americanum* serpin 41 (AAS41) inhibits inflammation by targeting chymase and chymotrypsin. *Int J Biol Macromol* (2020) **156**:1007–1021. doi:10.1016/j.ijbiomac.2020.04.088

18. Sugino M, Imamura S, Mulenga A, Nakajima M, Tsuda A, Ohashi K, Onuma M. A serine proteinase inhibitor (serpin) from ixodid tick *Haemaphysalis longicornis*; cloning and preliminary assessment of its suitability as a candidate for a tick vaccine. *Vaccine* (2003) **21**:2844–2851. doi:10.1016/S0264-410X(03)00167-1

19. Imamura S, Da Silva Vaz I, Sugino M, Ohashi K, Onuma M. A serine protease inhibitor (serpin) from *Haemaphysalis longicornis* as an anti-tick vaccine. *Vaccine* (2005) **23**:1301–1311. doi:10.1016/j.vaccine.2004.08.041

20. Wang F, Song Z, Chen J, Wu Q, Zhou X, Ni X, Dai J. The immunosuppressive functions of two novel tick serpins, HlSerpin-a and HlSerpin-b, from *Haemaphysalis longicornis*. *Immunology* (2020) **159**:109–120. doi:10.1111/imm.13130

21. Toyomane K, Konnai S, Niwa A, Githaka N, Isezaki M, Yamada S, Ito T, Takano A, Ando S, Kawabata H, et al. Identification and the preliminary *in vitro* characterization of IRIS homologue from salivary glands of *Ixodes persulcatus* Schulze. *Ticks Tick Borne Dis* (2016) **7**:119–125. doi:10.1016/j.ttbdis.2015.09.006

22. 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. *J Biol Chem* (2002) **277**:10083–10089. doi:10.1074/jbc.M111391200

23. Ibelli AMG, Kim TK, Hill CC, Lewis LA, Bakshi M, Miller S, Porter L, Mulenga A. A blood meal-induced *Ixodes scapularis* tick saliva serpin inhibits trypsin and thrombin, and interferes with platelet aggregation and blood clotting. *Int J Parasitol* (2014) **44**:369–379. doi:10.1016/j.ijpara.2014.01.010

24. Mulenga A, Tsuda A, Onuma M, Sugimoto C. Four serine proteinase inhibitors (serpin) from the brown ear tick, *Rhiphicephalus appendiculatus*; cDNA cloning and preliminary characterization. *Insect Biochem Mol Biol* (2003) **33**:267–276. doi:10.1016/s0965-1748(02)00240-0

25. Yu Y, Cao J, Zhou Y, Zhang H, Zhou J. Isolation and characterization of two novel serpins from the tick *Rhipicephalus haemaphysaloides*. *Ticks Tick Borne Dis* (2013) **4**:297–303. doi:10.1016/j.ttbdis.2013.02.001

26. Xu Z, Yan Y, Zhang H, Cao J, Zhou Y, Xu Q, Zhou J. A serpin from the tick *Rhipicephalus haemaphysaloides*: involvement in vitellogenesis. *Vet Parasitol* (2020) **279**:109064. doi:10.1016/j.vetpar.2020.109064

27. Tirloni L, Seixas A, Mulenga A, da Silva Vaz I, Termignoni C. A family of serine protease inhibitors (serpins) in the cattle tick *Rhipicephalus (Boophilus) microplus*. *Exp Parasitol* (2014) **137**:25–34. doi:10.1016/j.exppara.2013.12.001

28. Kaewhom P, Sirinarumitr T, Chantakru S, Jittapalapong S. Cloning and characterization of cDNA encoding a serine protease inhibitor from salivary glands of Thai cattle tick (*Boophilus microplus*). *Kasetsart J (Nat Sci)* (2007) **41**:74–80.

3.2 Structural and biochemical characterization of the novel serpin Iripin-5 from *Ixodes ricinus*

ABSTRACT

Iripin-5 is the main *Ixodes ricinus* salivary serpin, which acts as a modulator of host defence mechanisms by impairing neutrophil migration, suppressing nitric oxide production by macrophages and altering complement functions. Iripin-5 influences host immunity and shows high expression in the salivary glands. Here, the crystal structure of Iripin-5 in the most thermodynamically stable state of serpins is described. In the reactive-centre loop, the main substrate-recognition site of Iripin-5 is likely to be represented by Arg342, which implies the targeting of trypsin-like proteases. Furthermore, a computational structural analysis of selected Iripin-5–protease complexes together with interface analysis revealed the most probable residues of Iripin-5 involved in complex formation.

Keywords: serpins; serine protease inhibitors; Iripin-5; X-ray structure; *Ixodes ricinus*; tick saliva.

INTRODUCTION

The castor bean tick (*Ixodes ricinus*) has a wide geographical distribution throughout the Northern Hemisphere of Europe, Asia and Africa that points toward its resistance to various environmental conditions. This has helped this ticks to become one of the major factor in the spread of zoonotic diseases, as it serves as a vector for multiple vector-borne pathogens (Tirloni *et al.*, 2014; Francischetti *et al.*, 2009). These include tick-borne diseases such as Lyme disease, Helvetica spotted fever, tick-borne meningoencephalitis, babesiosis and tick paralysis (Sprong *et al.*, 2018). *I. ricinus* represents a model organism used

in the development of new sustainable tick control approaches such as acaricides and repellents. The saliva of ticks helps them to stay attached to the host until finishing the long-lasting blood feeding process is finished. This is facilitated by many immunomodulatory, anti-inflammatory and anti-hemostatic proteins, peptides and non-peptide molecules in the saliva (Francischetti *et al.*, 2009; Kotál *et al.*, 2015).

Serpins (<u>ser</u>ine <u>p</u>rotease <u>in</u>hibitors) are the largest superfamily of protease inhibitors and are broadly distributed in nature (Silverman *et al.*, 2001; Spence *et al.*, 2021). The vast majority of serpins act as serine protease inhibitors, but during the evolution, some serpins switched to non-inhibitory functions such as molecular chaperones (for example heat shock serpin 47; Nagata, 1996), tumour suppressors (for example Maspin; Zou *et al.*, 1994), storage proteins (for example Ovalbumin ;Mellet *et al.*, 1996; Law *et al.*, 2006) and hormone binding globulins (for example thyroxine-binding globulin and cortisol-binding globulin; Pemberton *et al.*, 1988). The typical process of serpin inhibition is irreversible and leads to substrate suicide: inactivation of both the serpin and the target protease. At the beginning of the inhibitory pathway, serpins form a Michaelis complex with the protease (Huntington, 2011). Subsequently, translocation of the reactive-centre loop (RCL) with the bound protease takes place and leads to the formation of a covalent complex with the trapped protease and the addition of a new strand in β-sheet A (Silverman *et al.*, 2001). Inhibitory serpins vary in function according to their specificities, and their importance can be illustrated by serpinopathies, diseases caused by serpin dysfunction or deficiency (Belorgey *et al.*, 2007). Many well-known diseases, for example emphysema, cirrhosis, angioedema, hypertension and familial dementia, are caused at least partially by serpin dysfunction (Law *et al.*, 2006; Huntington, 2011). This makes serpins interesting candidates for drug design and development, for which a high-resolution structure is necessary. All serpins possess a structurally similar core domain consisting of ~ 380 residues. This domain is made up of three β-sheets (A, B and C) and eight or more α-helices (hA – hI; Gettins, 2002). Another typical characteristic feature of serpins is the presence of an exposed, extended RCL that acts as a bait for the target protease

during the inhibition. The RCL consisting of ~ 17 residues and is located between β-sheet A and β-sheet C (Dunstone & Whisstock, 2011). It was found that serpins show different structural conformations such as native (S, stressed state), cleaved (R, relaxed state), latent (a result of auto inactivation due to a mutation or self-stabilization) and the δ-conformation (inappropriate partial insertion of the RCL due to a mutation) as well as the possible formation of complexes as a result of the inhibitory mechanism (Dunstone & Whisstock, 2011). The inhibitory mechanism can result in successful inhibition by covalent complex formation with the target protease or cleaved conformation. During conformational change of both states, thus incorporation of the RCL into β-sheet A, energy release occurs and a rise in serpin stability is reported as a consequence of this transition. In the case where this process is not sufficiently fast, it results in unsuccessful inhibition of the protease and its release from the acyl-intermediate, followed by the formation of a cleaved conformation of the serpin (Gettins, 2002; Gettins & Olson, 2016; Yamasaki *et al.*, 2002).

Tick salivary serpins play important roles in tick physiology. They are necessary to modulate the immune-system responses of the host and to inhibit various defence mechanisms such as hemostasis, which can result in the facilitated transmission of the aforementioned tick-borne pathogens (Kotál *et al.*, 2015). High structural conservation of serpins across tick species has been observed (Porter *et al.*, 2015). The tick *I. ricinus* ticks expresses over 30 serpins with different specificities, of which only Iris, IRS-2 and Iripin-3 have been characterized in details functionally, while IRS-2 and Iripin-3 have also been structurally characterized (Prevot *et al.*, 2006; Chmelař *et al.*, 2017; Páleníková *et al.*, 2015; Chlastáková *et al.*, 2021).

Here, we present the structural, biochemical, and functional characterization of the serpin from *I. ricinus* named Iripin-5 (*I. ricinus* serpin-5) that is highly expressed in the salivary glands of the tick; its expression is induced by feeding on blood and it displays anti-inflammatory and anti-complement features. Structural analysis revealed that Iripin-5 crystallized in a cleaved conformation and its structure was solved at 1.50 Å resolution. The structure was used for interface and computational analyses of its complexes with chosen proteases.

MATERIALS AND METHODS

Protein cloning, expression and purification

The full-length Iripin-5 sequence was cloned into a pET-17b vector and transformed into *Escherichia. coli* strain BL21-pLysS (Novagen, USA). 6 l LB medium (100 µg*ml^{-1} ampicillin, 34 µg*ml^{-1} chloramphenicol) was inoculated with an overnight culture of BL21-pLysS cells containing the Iripin-5 gene. Protein overexpression was induced by 1 mM isopropyl β-$_D$-1-thiogalactopyranoside (IPTG) on reaching OD_{600} of 0.6 and the cells were harvested 3 h after induction. Inclusion bodies were isolated by sonication in 20 mM Tris-buffered saline (TBS), 150 mM NaCl, pH 8.0 with 1 % (v/v) Triton X-100 buffer and washed three times with TBS to remove traces of Triton X-100. Inclusion bodies were dissolved in 20 mM TBS, 6 M guanidine-HCl, pH 8 and undissolved impurities were removed by centrifugation (12 000g). Refolding was achieved by rapid dilution in a 160-fold excess of 50 mM Tris, 300 mM NaCl, 0.8 mM KCl, 250 mM $_L$-arginine, pH 8.5 with the 0.25 g of wet inclusion bodies per litre of refolding buffer. After filtration, Iripin-5 was purified by ion exchange and size exclusion chromatography (Supplementary Figs. S1 and S2). Pure protein was sent and decontaminated from Lipopolysaccharide (LPS) by Arvys Proteins company (Trumbull, USA) using a detergent-based method. The LPS was removed from the sample becouse of its proven activation effect on cells, especially the stimulation of cells responsible for immune responses. This would interfere in subsequent experiments. The final concentration of protein was 1.14 mg ml^{-1} in 20 mM Tris, 150 mM NaCl, pH 8.0 buffer and the protein was stored at -80°C.

Nitric oxide production by IC-21 macrophages

Macrophages of the IC-21 cell line were pre-incubated with various concentrations of Iripin-5 for 4 h. After stimulation with 100 ng*ml^{-1} LPS and 5 ng*ml^{-1} interferon γ (IFNγ), the cells were incubated for 24 h or 48 h. The nitric oxide (NO) concentration was assessed after incubation with a modified Griess reagent (Sigma-Aldrich, Germany).

Antiprotease selectivity

Assays were performed according to a previous publication (Chmelar *et al.*, 2011). The enzyme concentrations do not reflect their ratio in the plasma or skin of the tick host. The used concentrations were chosen based on the biochemical properties of particular proteases in order to detect substrate hydrolysis and do not reach saturation of reaction at the same time. Generally, the assay conditions were chosen as half of the V_{max} of each particular protease. Briefly, assays were performed at 30°C and tested in triplicates. The used protein concentration in the reaction was from 400 nM and the serpin was pre-incubated with the target enzyme (listed in Table 1) for 10 min before adding a substrate (250 μM final concentration). For each target enzyme, appropriate buffers at different final concentrations were used. The substrate-hydrolysis rate was determined using an Infinite 200 PRO 96-well plate fluorescence reader (Tecan, Switzerland; excitation at 365 nm, emission at 450 nm).

Table 1 Antiprotease selectivity of Iripin-5.

Enzyme	Amount of enzyme used (nM)	Remaining enzymatic activity (%)
Thrombin	0.01	95.2 ± 3.2
Factor Xa	0.33	97.6 ± 4
Kallikrein	0.04	100.9 ± 2.5
Chymase	0.45	81.1 ± 3.3
Trypsin	0.1	55.9 ± 1.5
α-Chymotrypsin	0.05	68.6 ± 1.4
β-Tryptase	0.01	104.2 ± 1.4
Human neutrophil elastase	0.06	**13 ± 2.2**
Cathepsin G	8.8	80 ± 1.8
U-PA	0.5	101 ± 1.5
Plasmin	1.2	94.1 ± 2.2
Matriptase	0.03	100 ± 1.9
Factor xia	0.06	98.8 ± 3
Factor xiia	0.1	98.8 ± 1.2
T-PA	0.02	100.8 ± 3.7
Proteinase 3	1.7	**4.6 ± 0.8**

Complement assay

Fresh rabbit erythrocytes were collected in Alsever's solution from the rabbit marginal ear artery, washed three times in an excess of PBS buffer (1.8 mM KH2PO4, 137 mM NaCl, 10 mM Na2HPO4-7H2O) and finally diluted to a final 2% suspension (*v/v*). Fresh human serum was obtained from three healthy individuals. The assay was performed in a 96-well round-bottomed microtiter plate (Nunc, Denmark). In each well a concentration of 50% human serum in PBS premixed with different concentrations of Iripin-5 (156 n*M* to 5 μ*M*) was added to a volume of 100 μl. After 10 min incubation at room temperature, 100 μl of erythrocyte suspension was added. Since the human serum lysed rabbit erythrocytes immediately after their addition to the reaction, we used only 50% concentration (*i.e.*, a 25% final serum concentration after addition of the erythrocyte suspension) the final dilution had been empirically established as optimal. Reaction wells were observed individually under an Olympus SZX7 stereomicroscope with oblique illumination (Olympus KL 1500) using an aluminium pad. The time needed for erythrocytes lysis was measured using a chronometer. When full lysis was achieved, the reaction mixture turned from opaque to transparent. Negative controls did not contain either serpin or human serum. Additional controls were performed with heat-inactivated serum (56ºC, 30 min) and the serpin, Iripin-3 (156 n*M* to 10 μ*M*). The assay was evaluated in technical and biological triplicates.

Neutrophil-migration assay

Neutrophils were obtained from the bone marrow of C57BL/6J mice by magnetic separation using a Neutrophil Isolation Kit (Miltenyi Biotec, Germany). Isolated neutrophils were pre-incubated in growth RPMI1640 growth medium containing 0.5 % (m/v) bovine serum albumin (BSA) in the presence or absence of Iripin-5 (3 μ*M*) for 1 h at 37°C and 5% CO_2. The cells were then seeded on the upper inserts of 3.0 μm pore Corning® Transwell® chambers (24-well format; Sigma-Aldrich, Germany). Chemoattractant solution (1 μM N-formyl-L-methionyl-L-leucyl-phenylalanine -fMLP in RPMI1640 with 0.5 % BSA) was placed in the lower compartments. After incubation for 1 h at

37°C and 5% CO_2, migration was determined by counting the cells in the lower chamber using a hemocytometer (Meopta, Czech Republic).

Iripin-5 expression profiles

I. ricinus nymphs were fed on C3H/HeN mice for one day, two days and until full engorgement (3-4 days); *I. ricinus* females were fed on guinea pigs for one, two, three, four, six and eight days. Adult salivary glands, midguts and ovaries, as well as nymph whole bodies, were dissected under RNAse-free conditions and total RNA was isolated using TriReagent (MRC). cDNA preparations were made from 1 μg total RNA from independent biological triplicates using a Transcriptor First Strand cDNA Synthesis kit (Roche, Czech Republic) according to the manufacturer's instructions. The cDNA was subsequently used for the analysis of Iripin-5 transcription by qPCR in a RotorGene 6000 cycler (Corbett Research Ltd, UK) using Fast Start Universal SYBR® Green Master Mix (Roche, Czech Republic), forward primer 5'-CGA GAA CGC AAC CAC TAA GA-3' and reverse primer 5'-GCT CAA CGT GAC CAA TGT AAT C -3'. Iripin-5 expression profiles were calculated using Livak's mathematical model (Livak & Schmittgen, 2001) and normalized to *I. ricinus* elongation factor 1α (ef1α; GU074829.1; forward primer 5'-CTG GGT GTG AAG CAG ATG AT-3' and reverse primer 5'-GTA GGC AGA CAC TTC CTT CTG-3'). The amplicon lengths were ef-1α, 105 bp; Iripin-5, 251 bp.

Protein crystallization, X-ray data collection and processing

Crystallization screening using commercial kits (JCSG++ from Jena Bioscience, SG1™ and PGA Screen™ from Molecular Dimensions, and PEGRx™ and Peg/Ion from Hampton Research, USA) was carried out at room temperature (20°C) and at 4°C by the sitting-drop vapour diffusion method using an OryxNano crystallization robot (Douglas Instruments). A suitable protein concentration for crystallization screening was determined using the Pre-Crystallization Test (Hampton Research, California, USA) as 1.14 mg*ml⁻¹. Drops of protein solution composed of 20 mM Tris, 150 mM NaCl, pH 8.0 buffer (1 μl) mixed with reservoir solution (1 or 0.5 μl) were equilibrated against

50 µl of reservoir solution and sealed in 96-well Swissci MRC 2-drop crystallization plates (Molecular Dimensions).

For data collection, crystals of Iripin-5 that grew for about one month were flash-cooled in liquid nitrogen with 20% (*v/v*) glycerol as an additional cryoprotectant. Measurements were carried out on beamline BL14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin (Mueller *et al.*, 2012). Collection of diffraction data was performed at 100 K with a 295.165 mm crystal-to-detector (PILATUS 6M) distance. Diffraction intensity data were processed using *XDS* program (Kabsch, 2010) with the *XDSAPP* graphical user interface (Sparta *et al.*, 2016). Data-collection statistics are summarized in Table 2.

Structure determination and refinement

Crystallographic and structural analyses were performed using the *CCP4* package (Winn et al., 2011). The structure of Iripin-5 was solved by the molecular-replacement method using *MOLREP* (Vagin & Teplyakov, 2010) with the structure of the serpin IRS-2 (PDB entry 3NDA; Chmelar *et al.*, 2011) as the search model. The structure was refined with the program *REFMAC5* (Murshudov *et al.*, 2011) and further manually in *Coot* (Emsley *et al.*, 2010) from evaluation of the electron-density peaks. The improvement during refinement was monitored by structure validation throughout the refinement process. Water molecules were added to the model using the *REFMAC5* interface. Accepted solvent molecules had tolerable hydrogen-bonding geometry contacts of 2.5-3.5 Å with protein atoms or with existing solvent. At this point, the residues with two possible conformations were included, and their alternative conformations were added for further refinement. In the last steps of refinement, the glycerol was built into the appropriate $(2F_o – F_c)$ and $(F_o – F_c)$ electron-density maps using coordinates from the ligand data bank in *Coot* (Emsley *et al.*, 2010). The *MolProbity* server (Williams *et al.*, 2018) and wwPDB validation server (Berman *et al.*, 2003) were used for final qualitative validation of the model. All figures were prepared using the *PyMOL* (DeLano, 2002). A summary of the data-collection and refinement statistics is given in Table 2.

Table 2 X-ray data-collection and refinement statistics. Values in the parentheses are for the highest-resolution shell.

Data collection	
X-ray diffraction source	BL14.1, BESSY II, Germany
Wavelength (Å)	0.9184
Detector	PILATUS 6M
Crystal-detector distance (mm)	295.165
Rotation range per image (°)	0.1
Total rotation range (°)	360
Exposure time per image (s)	0.1
Resolution range (Å)	48.09-1.50 (1.59-1.50)
Space group	$P12_11$ (4)
Asymmetric Unit	2 molecules
Unit-cell dimensions: a, b, c (Å)	76.24, 63.78, 81.99
Unit-cell dimensions: α, β, γ (°)	90.0, 116.78, 90.0
Mosaicity (°)	0.199
Total No. of reflections	752984 (117495)
No. of unique reflections	112133 (17637)
Multiplicity	6.72
Average I/σ(I)	11.17 (1.41)
Completeness (%)	98.7 (96.5)
CC ½	99.8 (61.3)
R_{meas} (%)[a]	11.0 (124.1)
Overall B factor from Wilson plot (Å^2)	24.46
Refinement	
Resolution range (Å)	48.09-1.50
No. of reflections in working set	110024 (7612)
Final R value[b](%) / Final R_{free}[c] value(%)	0.153 / 0.185
Mean B value (Å)	17.725
No. of nonhydrogen atoms in the asymmetric unit	
Protein	6010
Water	1097
Magnesium	4
Chlorine	6
Total	7117
R.m.s. deviations	
Bonds (Å)	0.012
Angles (°)	1.672
Average B factors (Å^2) Overall	17.517
Ramachandran plot	
Most favoured (%)	98.64
Allowed (%)	100.00
PDB ID:	7b2t

[a] $R_{meas} = \dfrac{\sum_h \left(\frac{n_h}{n_h-1}\right)^{1/2} \sum |\langle I_h \rangle - I_{h,i}|}{\sum_h \sum_i I_{h,I}}$, where the average intensity $\langle I\ (hkl)\rangle$ is taken over all symmetry-equivalent measurements, and $I_i(hkl)$ is the measured intensity for i^{th} observation of reflection hkl.

[b] $R = \dfrac{\sum ||F_{obs}| - |F_{calc}||}{2a \sum |F_{obs}|}$ where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[c] R_{free} is equivalent to R value but is calculated for 1.87% of the reflections that were chosen at random and omitted from the refinement process.

Structural analysis and molecular dynamics of the modelled Michaelis complexes

The structures of the predominantly inhibited proteases – proteinase 3 and neutrophil elastase were fetched form the PDB as PDB entries 1fuj at 2.20 Å resolution (Fujinaga *et al.*, 1996) and 3q76 at 1.86 Å resolution (Hansen *et al.*, 2011), respectively. The ceaved Iripin-5 crystal structure was modelled to match the native conformation of serpins. The inserted RCL from β-sheet A was modelled above the Iripin-5 structure and the missing residues (Leu-343, Ile-344, Glu-345, Val-346 and Pro-347) were modelled into the structure to complete the native structure. The crystal structures of the chosen proteases were modified by removing of alternative conformations of the amino-acid side chains, ligands and ions from the structures as required for further docking calculations with prepared native Iripin-5 model. The *HADDOCK2.2* web server (van Zundert *et al.*, 2016) was used for Michaelis complex docking and the best results were used to run molecular-dynamics (MD) simulations. MD simulations were performed by using the *GROMACS* simulation suite (Berendsen *et al.*, 1995) with the CHARMM27 all-atom force field and the SPC/E (extended simple point charge) model for water (Feller & MacKerell, 2000; Klauda *et al.*, 2005). Both Michaelis complex models were prepared for simulation by removing the solvent and were solvated by using the SPC/E water model (Berendsen *et al.*, 1987) in a rhombic dodecahedral box. The protein was centered in the box and the size of the box was such that the protein was at least 1 nm from all edges. Na^+ ions were added to the system at a concentration of 150 m*M* as well as an appropriate amount of Cl^- ions to neutralize the system. The entire system was minimized using a steepest-descent minimization procedure. The energy minimized structure was then further equilibrated in two phases for 100 ps each: first under an NVT ensemble (constant number of particles, volume, and temperature) followed by an NPT ensemble (constant number of particles, pressure and temperature) to ensure that the system remained stable. Simulations were then performed for 100 ns each, during which time equilibrium of the system was achieved. MD simulations were performed fully in triplicate (*i.e.*, from the minimization to production run) to ensure

reproducibility. The results of the molecular simulations were analyzed using *VMD* (Humphrey *et al.*, 1996) with the use of the r.m.s.d. trajectory tool. An interface analysis of the resulting structures (the final frames of each simulation) was then performed using the *PDBePISA* web server (Krissinel & Henrick, 2007).

Structural analysis and protein docking of modelled covalent complex conformations

Structures were analyzed and compared with those of the other *I. ricinus* serpins using *PyMOL* version 2.0 (DeLano, 2002; Schrödinger,). The *HADDOCK2.2* web server (van Zundert *et al.*, 2016) was used for protein docking to generate covalent complexes. The possible target proteases selected as the best candidates from anti-protease selectivity assays, namely proteinase 3, human neutrophil elastase, trypsin, α-chymotrypsin, cathepsin G and chymase, were used for analysis. The crystal structures of the human proteases were taken from the Protein Data Bank: PDB entries 1fuj at 2.20 Å resolution (Fujinaga *et al.*, 1996), 3q76 at 1.86 Å resolution (Hansen *et al.*, 2011), 1h4w at 1.70 Å resolution (Katona *et al.*, 2002), 4cha at 1.68 Å resolution (Tsukada & Blow, 1985), 1au8 at 1.90 Å resolution (to be published) and 3n7o at 1.80 Å resolution (Kervinen *et al.*, 2010). Alternative conformations of the amino-acid side chains, ligands and ions were removed from the structures as required. Interface analysis was performed using the *PDBePISA* (Krissinel & Henrick, 2007) and *COCOMAPS* (Vangone *et al.*, 2011) web servers.

PDB deposition

The atomic coordinates of Iripin-5 have been deposited in the Protein Data Bank with accession code 7b2t.

Statistical analyses

All immunological experiments were performed as at least in three biological replicates. Data are presented as mean ± standard error of mean (SEM) in all graphs. Student's t-test or one-way ANOVA were used to calculate statistical differences between two or more groups, respectively. Statistically

significant results are marked as follows in the figures: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$; n.s., not significant.

RESULTS

Iripin-5 expression profiles

Expression of the Iripin-5 gene was upregulated during tick feeding in all tested tissues. The highest expression was found in semi-engorged nymphs (D2), and expression was also high in fully engorged nymphs (D3) as well as in female salivary glands during finishing of their blood meal (D6 and D8, Fig. 1). Furthermore, the Iripin-5 transcripts in engorged nymphs and female salivary glands were the most abundant among all tested tick serpins (data not shown). Thus, Iripin-5 is likely to be the most abundant serpin that is secreted from the salivary glands to the host.
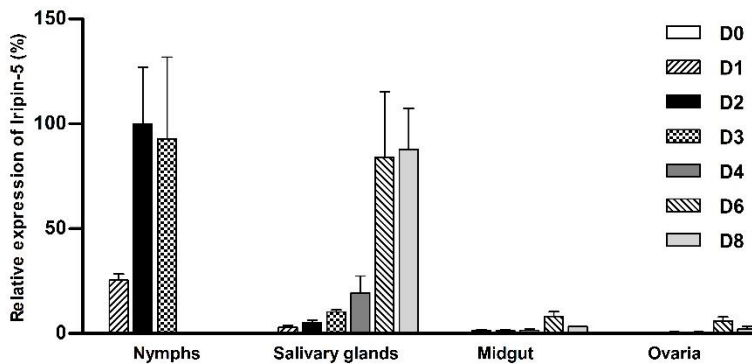

Figure. 1 The Iripin-5 expression is upregulated in *I. ricinus* nymphs and adults during feeding. The analysis was performed on tissues of flat, semi-engorged and fully engorged nymphs and female salivary glands, midguts and ovaries. RT-qPCR expression data are normalized against elongation factor 1α (ef1α) and the highest expression was set up as 100 %. Shown data represent mean + SEM from three biological replicates. D0 – D8: days of feeding.

Antiprotease selectivity and neutrophil migration

In the protease selectivity assay, Iripin-5 needed to be in a high excess compared with the target enzymes in order to obtain even a low level of inhibition. The remaining protease inhibition after 10 min incubation with 200

nM Iripin-5 is given in Table 1. Iripin-5 showed the highest inhibitory specificity against two neutrophil proteases: human neutrophil elastase (87 % inhibition) and proteinase 3 (95 % inhibition). Statistically significant results are noted in bold. Based on physiologically relevant proteases for tick-host interaction, it was found that only chymase and cathepsin G were inhibited significantly, and only very weakly. Another two inhibited proteases, trypsin and α-chymotrypsin, show importance during digestion.

Since Iripin-5 primarily inhibited neutrophil proteases, the effect on neutrophils functions was also studied. Static migration was tested using a Transwel chamber and purified mouse neutrophils isolated from bone marrow. Pre-incubation with 3 μM Iripin-5 led to greater than 70 % decrease in neutrophils migration, thus showing a significant antineutrophil effect of Iripin-5 (Fig. 2).
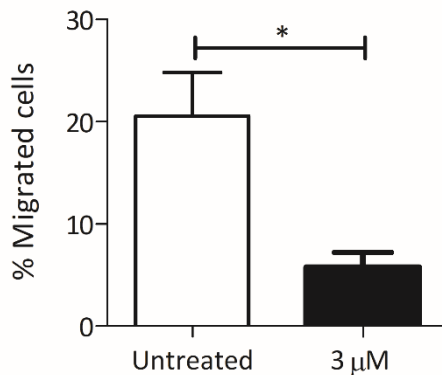

Figure. 2   Iripin-5 inhibits neutrophil migration. Mouse bone-marrow neutrophils were pre-incubated with 3 μM Iripin-5 and subjected to migration towards fMLP in a Transwell ™ chamber. The average of three independent experiments (±SEM) is shown. * $p \leq 0.05$

NO production by IC-21 macrophages

The incubation of macrophages in the presence of Iripin-5 led to a decrease in NO production in a dose-dependent manner. At a concentration of 1 μM, Iripin-5 inhibited NO production slightly, but not significantly, at 24 h, but

not at 48 h. At a higher concentration of 5 μ*M*, Iripin-5 significantly decreased the amount of NO at both time points: by 35% and 36%, respectively (Fig. 3).
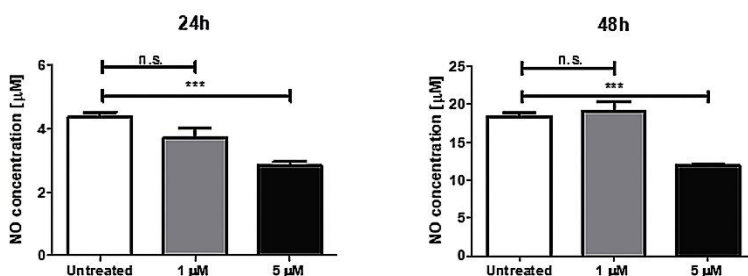

Figure. 3   Iripin-5 decreased NO production by activated IC-21 macrophages. Iripin-5 inhibited NO production by IC-21 macrophages when used at high concentration. Macrophages were preincubated with 1 and 5 μ*M* Iripin-5, stimulated with LPS and IFN-γ, and the NO concentration was assessed after 24 or 48 h. The mean of three independent experiments (± SEM) is shown. ***$p \leq 0.001$; n.s. not significant.

Complement assay

Since Iripin-5 affected two major immune-cell types involved in innate immune response, the interference of the tested serpin with another innate immune mechanism involved in anti-tick immunity of the complement was tested. Iripin-5 inhibited the lysis of erythrocytes by human complement. Human plasma was pre-incubated with different concentrations of Iripin-5, from 156 n*M* to 5 μ*M*. After the addition of rabbit erythrocytes, their lysis time by complement was measured. A statistically significant reduction in complement-driven lysis activity against erythrocytes when incubating human plasma with Iripin-5 at concentrations of 625 n*M* and higher was observed. No lysis of any erythrocytes was detected when using 5 μ*M* Iripin-5. The results were compared with those for another serpin, Iripin-3 (Chlastáková *et al.*, 2021), which had no effect on complement activity, demonstrating the specificity of our assay. The lysis of rabbit erythrocytes in the presence of 25% human serum was achieved within 7 min 57 s ± 0.12 s on average in the control group, corresponding to the zero value in the graph (Fig. 4).
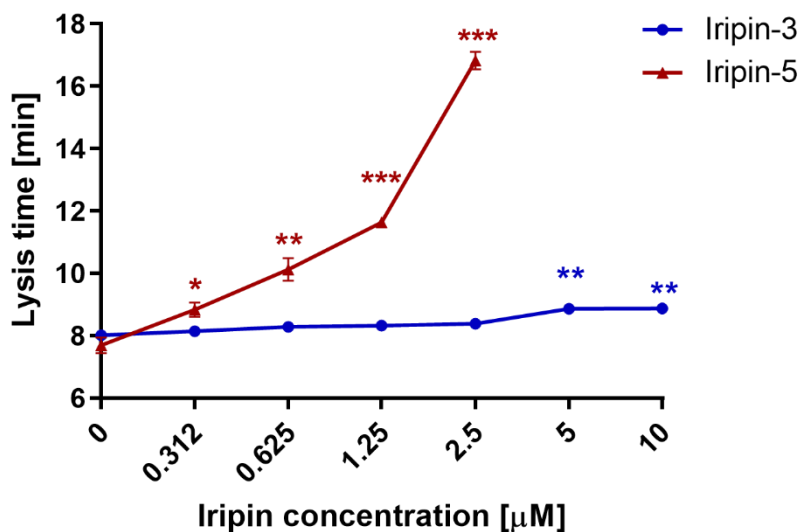
Figure. 4   Inhibition of complement by Iripin-5 compared with another *I. ricinus* salivary serpin, Iripin-3. Human plasma was pre-incubated with an increasing concentration of Iripin-5 (156 n*M* to 5 μ*M*) and Iripin-3 (312 n*M* to 10 μ*M*). After the addition of rabbit erythrocytes, their lysis time by complement was measured. For each point in the graph, the mean of three independent experiments (± SEM) is shown. * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

Crystal structure of Iripin-5

In order to obtain a deeper view into the mechanisms of Iripin-5 activity, a detailed structural analysis was performed. To generate a protein structure of Iripin-5, crystallization experiments were performed and the structure of the serpin was solved form the best-diffracting crystals to a resolution of 1.5 Å. Iripin-5 crystals with a monoclinic shape (Fig. 5) grew after one month at 4°C in a precipitant composed of 0.2 *M* magnesium chloride hexahydrate, pH 8.5, 0.1 *M* Tris, 30 % (*w/v*) PEG 4000 (condition No. 1-1 of SG1from Molecular Dimensions). The crystal space group and unit-cell parameters are reported in Table 2.
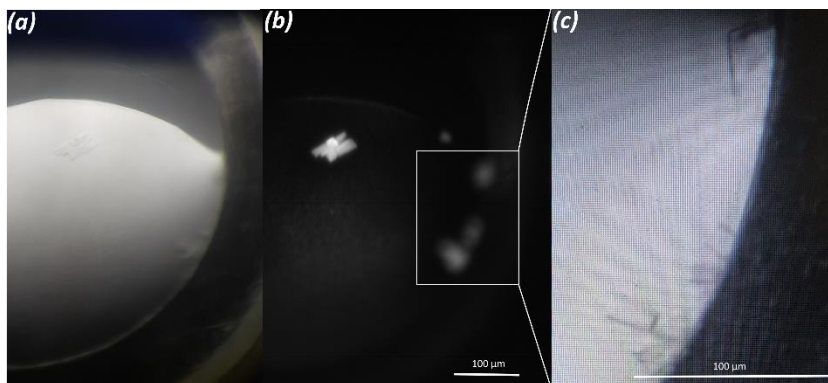
Figure. 5   Crystals of Iripin-5 from *I. ricinus*. (*a*) Crystals of protein grown in the 0.2 *M* magnesium chloride hexahydrate, pH 8.5, 0.1 *M* Tris, 30 % (*w/v*) PEG 4000. (*b*) The same crystallization droplet is shown under UV light. (*c*) Focus on the best-shaped crystals for diffraction measurements. (*a*) was taken using an Olympus SZX9 microscope and (*b*) and (*c*) were taken using a Minstrel Desktop Crystal Imaging System (Rigaku, Japan). The scale bar represents 100 μm.

The structure of Iripin-5 was solved by molecular replacement using the previously published structure of the serpin IRS-2 (PDB entry 3nda) as the model structure, which has a sequence identity of 55.70% (Chmelar *et al.*, 2011). The crystal structure contains two molecules per asymmetric unit, with a solvent content of 39.97 % and a Matthews coefficient of 2.05 $Å^3Da^{-1}$. The Iripin-5 structure has a typical cleaved serpin secondary-structure fold in both molecules. The structure consists of a mixed αβ secondary structure with an N-terminal helical region and a C-terminal β-sheet fold (Huntington, 2011; Fig. 6). The structure is composed of eight α-helices and three β-sheets sequentially arranged in the order α1-β1-α2-α3-β2-α4-β3-α5-β4-β5-β6-β7-β8-α6-α7-β9-β10-α8-β11-β12-β13-β14-β15. Sheet A consists of six β-strands (β2, β3, β4, β10, β11 and β12), sheet B of five β-strands (β1, β7, β8, β14 and β15) and sheet C of four β-strands (β5, β6, β9 and β13) (Fig. 6).

The final model of Iripin-5 contains 373 residues in chain A and chain B out of a total 378, with the five missing residues (Leu-343 – Thr-347) in both chains. The missing residues in the crystal structure were detected as an absence of electron density due to the high flexibility of the cleaved regions (Fig. 6), and thus these regions were not modelled in the final structure. The cleavage is probably a consequence of the presence of protease, most probably during

storage. The cleavage of the sample used for crystallization was confirmed by MALDI mass-spectrometry protein analysis (Supplementary Fig. S3) and led to the structural change and thus the insertion of the cleavage site inside the β-sheet to form the extra β-strand (S4). The cleavage site is homologous to the RCL of other serpin inhibitors and the cleaved state is the most stable, so-called hyperstable or R form of inhibitory serpins (Huntington, 2011; Fig. 6). Moreover, the analysis of the protein interfaces by *PDBePisa* (Krissinel & Henrick, 2007) did not reveal any specific interactions resulting in the formation of stable quaternary structures. Most probably the structures do not form any complexes in solution (Schlee *et al.*, 2019).
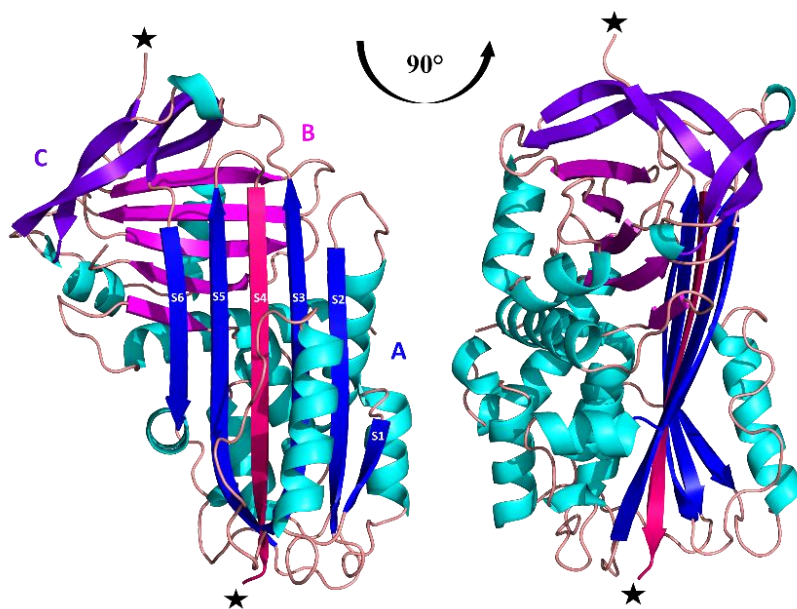

Figure. 6   Cleaved protein form with colour-distinguished β-sheets: sheet A (blue), sheet B (magenta) and sheet C (violet). The insertion of the RCL between β-strands S3 and S5 (blue) is marked as S4 β-strand (dark pink). The location of protease cleavage is marked with black stars.

Structural analysis and molecular dynamics of the theoretical Michaelis complex

To identify the specific interactions that are potentially responsible for the mechanism of inhibition between the target proteases and Iripin-5, protein-docking and subsequent MD simulations of the Michaelis complexes were

performed. Three simulations for modelled Michaelis complexes with both neutrophil elastase and proteinase 3, each 100 ns long, were performed.

The stability of the complex was monitored by the r.m.s.d. evaluation (Fig. 7). The results showed that triplicates of both simulated complexes reached equilibrium within the simulation time and the average r.m.s.d.s from the initial starting structure for the Michaelis complexes were 5.3 Å (Fig. 7*b*, orange), 6.7 Å (Fig. 7*b*, grey) and 8.0 Å (Fig. 7*b*, yellow) for Iripin-5 – proteinase 3 complex and 8.4 Å (Fig. 7*a*, blue), 9.6 Å (Fig. 7*a*, yellow) and 5.8 Å (Fig. 7*a*, green) for the Iripin-5 – neutrophil elastase complex. The difference between the average r.m.s.d. of the Iripin-5 – proteinase 3 complex triplicates was 3.8 Å and that for Iripin-5 – neutrophil elastase complex was 2.7 Å.
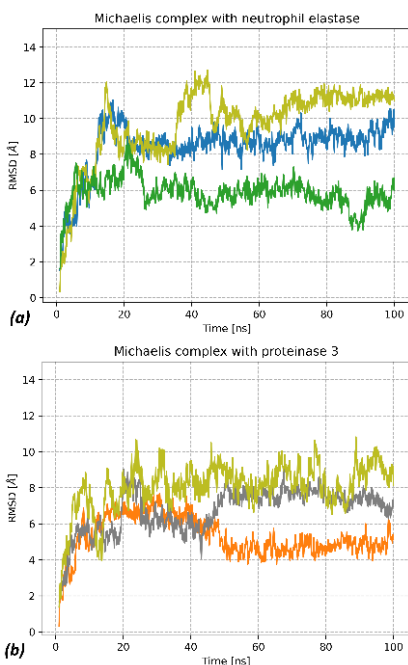

Figure. 7   All-atom r.m.s.d. of MD simulations of Michaelis complex models. *(a)* R.m.s.d. for the Iripin-5 – neutrophil elastase Michaelis complex and *(b)* r.m.s.d. for the Iripin-5 – proteinase 3 Michaelis complex, each for 100 ns simulation. Triplicates are distinguished by different colours (corresponding to the visualization of Michaelis complex models on Fig. 8).

Representations of the Michaelis complexes between Iripin-5 and neutrophil elastase and proteinase 3 are shown in Fig. 8. Structural alignment performed by *PyMOL* (DeLano, 2002) showed that the average r.m.s.d. between

the Iripin-5 – neutrophil elastase triplicates was 1.709 Å and that between the Iripin-5 – proteinase 3 triplicates was 1.958 Å. These results show the flexibility of the Michaelis complex conformation and, more precisely, the flexibility of the Iripin-5 RCL (Fig. 8).

The interface analysis of the Michaelis complex triplicates is summarized in Table 3. The data in bold indicate the importance of the interface for complex formation (*PDBePISA*; Krissinel & Henrick, 2007)). A more detailed summary of the interacting residues is presented in Supplementary Table 3.
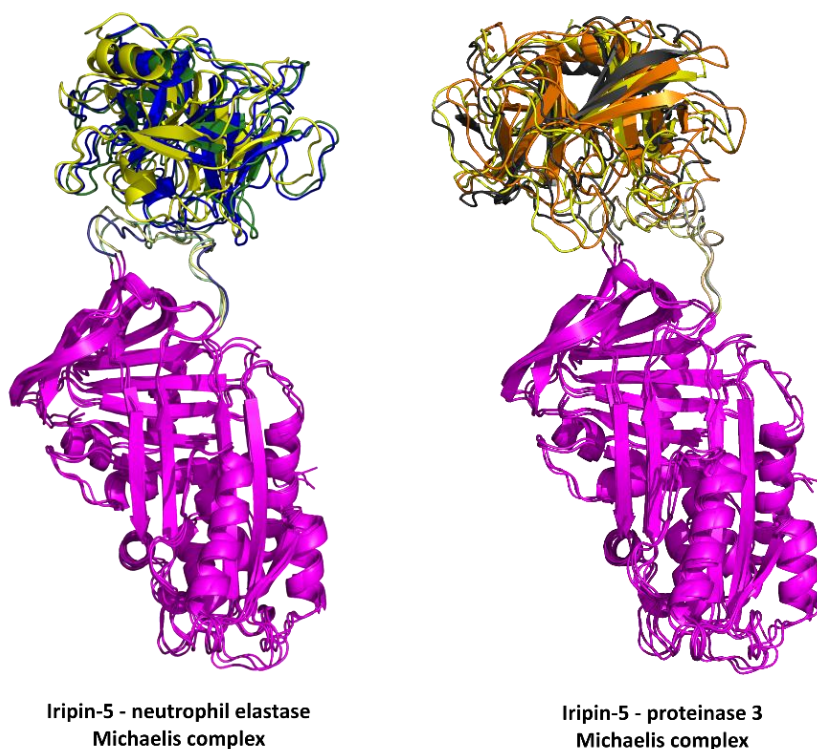

Iripin-5 - neutrophil elastase
Michaelis complex

Iripin-5 - proteinase 3
Michaelis complex

Figure. 8   Results of MD simulation of Michaelis complex. The structures are shown at 100 ns point of simulation for each triplicate of chosen target protease. The Iripin-5 (magenta) structures are aligned to show RCL dynamics. Triplicates are distinguished by different colours for the target protease: neutrophil elastase (blue, green and yellow); and proteinase 3 (grey, orange and yellow). The Iripin-5 RCL is also distinguished in a corresponding colour to the interacting protease. A detailed view on Michaelis complex interfaces is presented in Supplementary Fig. S6.

Table 3 Area of the accessible surface interface between Iripin-5 and tested proteases in the Michaelis complex conformation, the number of hydrogen bonds and the number of salt bridges formed after 100 ns of MD simulation (from *PDBePISA*; Krissinel & Henrick, 2007)).

Protease	Surface interface (Å^2)	No. of hydrogen bonds	No. of salt bridges
Human	**684.9**	**5**	**4**
neutrophil	662.9	4	5
elastase	**655.7**	**2**	**2**
Proteinase 3	864.2	4	2
	804.2	3	2
	827.0	**2**	**2**

Structural analysis of theoretical protein–protease covalent complex conformation

To test the hypothesis of the presence of polar contacts between Iripin-5 and six chosen proteases (proteinase 3, human neutrophil elastase, trypsin, α-chymotrypsin, cathepsin G and chymase), docking calculations of protein-protein interaction were performed using *HADDOCK* (van Zundert *et al.*, 2016). The results of the docking studies of interacting residues at the Iripin-5 – protease complex interfaces, listed in the Table 4 and shown in Fig. 9, show different character for the interactions in complexes.
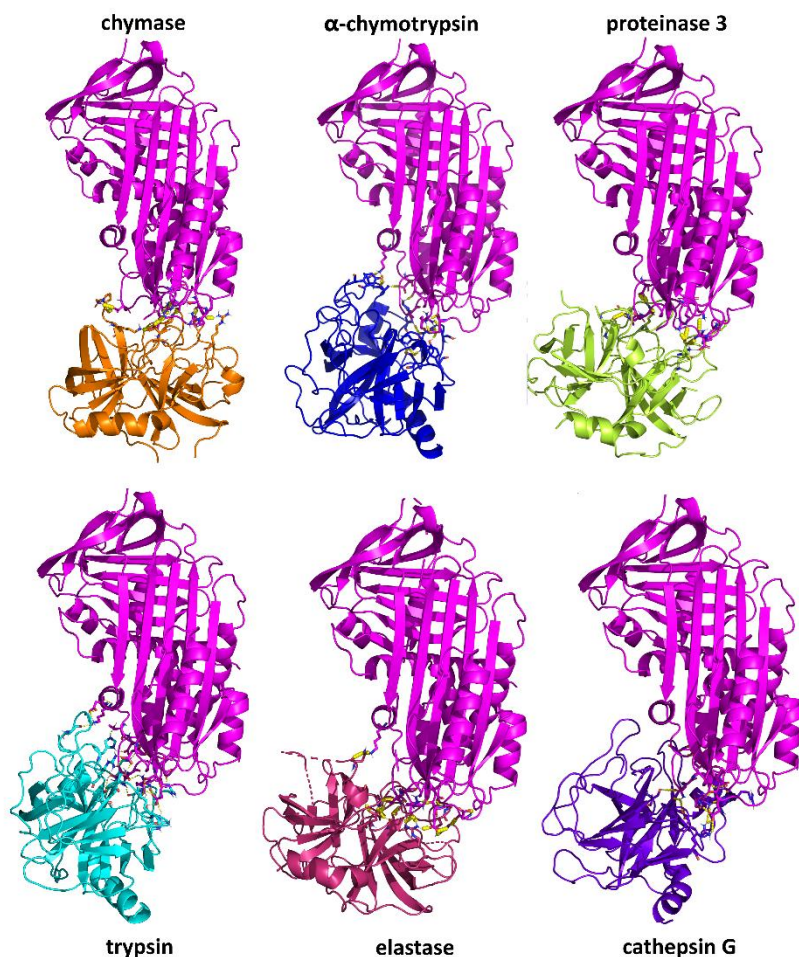
Figure. 9 Cartoon representation of the docking results of Iripin-5 (magenta) with chosen proteases: cathepsin G (violet), trypsin (cyan), elastase (hot pink), α-chymotrypsin (blue), chymase (orange) and proteinase 3 (lemon). The residues interacting with protease catalytic triad are shown in detail in Fig. S5.

Only the α-chymotrypsin catalytic triad interacted with Arg-342; thus, the potential Iripin-5 P1 site was a candidate for binding the protease. However, the proteinase 3 and chymase side-chain residues of the catalytic triad were not in contact with Arg-342 of Iripin-5. The remaining proteases (cathepsin G, elastase, and trypsin) interacted with Arg-342 of Iripin-5 via side-chain residues other than the catalytic triad (Supplementary Fig. S5). Detailed information

about atomic interface analysis is shown in Supplementary Table S2. These results were calculated using the *PDBePISA* (Krissinel & Henrick, 2007).

Table 4 Area of the accessible surface interface between Iripin-5 and tested proteases, the number of hydrogen bonds and the number of salt bridges formed (from PDBePISA (Krissinel & Henrick, 2007))

Protease	Surface interface ($Å^2$)	No. of hydrogen bonds	No. of salt bridges
Proteinase 3	892.8	13	10
Human neutrophil elastase	733.6	8	4
Trypsin	919.5	12	10
α-Chymotrypsin	787.5	7	1
Cathepsin G	947.1	8	4
Chymase	849.8	5	6

DISCUSSION

The *I. ricinus* sialome (transcriptome from the salivary glands) contains four major types of protease inhibitors such as cystatins, TIL-domain inhibitors, Kunitz inhibitors and serpins, which are proven or presumed to be modulators of host-defense mechanisms (Chmelař *et al.*, 2017). Among them, the serpins stand out thanks to their omnipresence across all living organisms and their indispensability for many crucial biochemical pathways, such as coagulation or complement and other fundamental functions (Huntington, 2011; Law *et al.*, 2006). Considering the fact that tick serpins usually do not form multigenic families, as are typical for other salivary protease inhibitors such as Kunitz-domain and TIL-domain inhibitors, they seem to be suitable candidates for targeting in tick-control attempts. Moreover, the structural conservation and use of serpins by vertebrates makes them promising candidates for novel drug development combined with the use of protein engineering (Chmelař *et al.*, 2017). Tick serpins can be utilized as specific regulators of dysregulated processes, such as inflammation, immune-system regulation or hemostasis. Several tick serpins have been shown to interfere with vertebrate immunity (Chmelař *et al.*, 2017). To date, three of them have been functionally characterized in *I. ricinus*. It has been shown that the salivary serpin Iris

modulates host innate and acquired immunity (Leboulle *et al.*, 2002). Likewise, IRS-2 and Iripin-3 modulated adaptive immune responses (Chmelar *et al.*, 2011; Chlastáková *et al.*, 2021). Moreover, crystal structures were determined for the last two, which are the only two tick serpins with resolved 3D structures to date.

Iripin-5 belongs to the salivary serpins, the role of which is considered to be as modulators of host defense mechanisms. Iripin-5 seems to be one of the main salivary serpins since its mRNA expression is by far the highest compared with other *I. ricinus* serpins. This serpin is massively induced by the blood meal. Here, several effects supporting the immunomodulatory and anti-inflammatory roles of Iripin-5 are reported. The observed inhibition of neutrophils migration suggests the anti-inflammatory activity at the very beginning of the immune reaction. Macrophages play an important role in the interaction between ticks, the immune system of the host and transmitted pathogens. Activated macrophages secrete signalling molecules such as cytokines or NO to recruit immune cells to sites of inflammation or towards pathogen (Laroux *et al.*, 2001). The saliva of different tick species has been shown to suppress the ability of macrophage to produce NO (Kýcková & Kopecký, 2006). Since Iripin-5 inhibits this very feature of macrophages, Iripin-5 is likely to be at least partially responsible for this activity, observed in *I. ricinus* saliva.

The inhibition of complement described here is interesting, but not surprising, as vertebrate serpins are natural regulators of the complement cascade (Bos *et al.*, 2002). There are other tick salivary protein families in which the members were have been described as complement inhibitors (Daix *et al.*, 2007; Tyson *et al.*, 2008), but our case is the first observation of complement inhibition by a tick serpin. This finding confirms the hypotheses about the functional redundancy of tick salivary proteins (Chmelař *et al.*, 2016).

Structural analysis of Iripin-5 shows the typical serpin fold in the relaxed state that was observed in other known crystal structures of *I. ricinus* serpins (IRS-2 and Iripin-3; (Chmelař *et al.*, 2017; Chlastáková *et al.*, 2021). The relaxed cleaved state of Iripin-5 model was caused by the presence of contaminating proteases, probably during protein storage, and this cleavage has been observed previously (Kovářová *et al.*, 2010). The crystal structure of Iripin-

5 was compared with those of IRS-2 (PDB entry 3nda; Chmelar *et al.*, 2011) and Iripin-3 (PDB entry 7ahp; (Chlastáková *et al.*, 2021) both by sequence alignment (Fig. 10*a*) and structural superimposition (Fig. 10*b*). The comparison of *I. ricinus* serpins with known structures reveals an almost identical fold (Fig. 10*b*) with some divergence in the loop regions. The r.m.s.d. between molecules was calculated by *PyMOL* (DeLano, 2002). On alignment of Iripin-5 and Iripin-3 the r.m.s.d. was 0.616Å, while the r.m.s.d. between Iripin-5 and IRS-2 was 0.804 Å across all atoms. In contrast, the sequence alignments of Iripin-3 and IRS-2 with Iripin-5 showed only 53.89% and 55.70% sequence identity, respectively (Fig. 10*a*). Electrostatic surface potentials support complex formation and stability and consequently the inhibition of proteases. This can be achieved by charge – charge repulsion or attraction in accordance with their function as protease substrate or inhibitor (Marijanovic *et al.*, 2019). A comparison of surface electrostatics among *I. ricinus* serpins reveals that Iripin-5 has a more negatively charged surface than the other two aforementioned serpins; Iripin-3 has only a slightly more negatively charged surface than Iripin-5 but shows much greater inhibition (Fig. 10*c*).
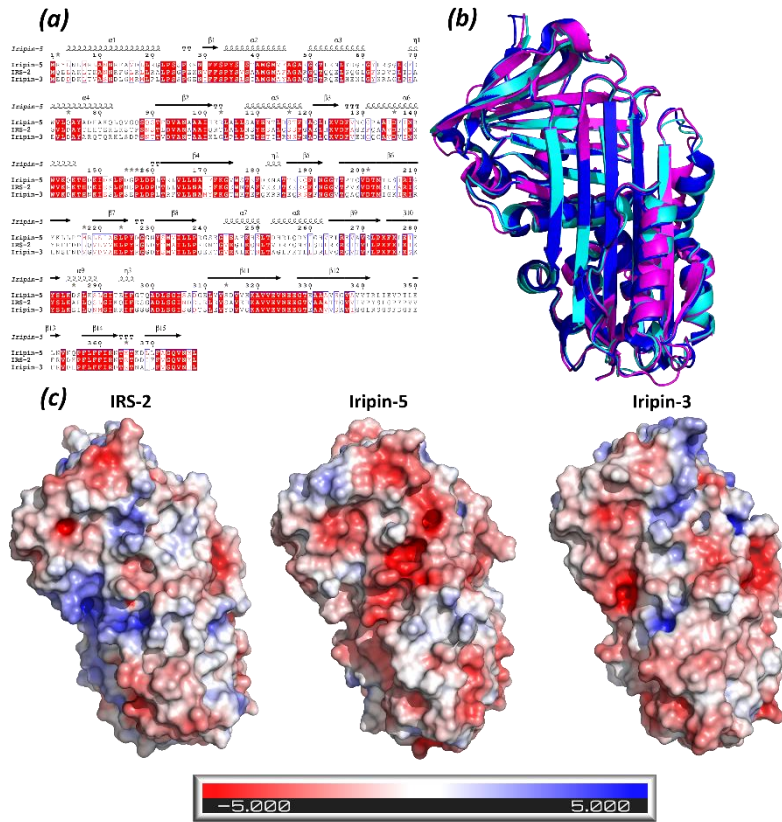
Figure. 10 (*a*) Amino acid sequence alignment between IRS-2, Iripin-3 and Iripin-5 serpins. Well-conserved amino acid motifs are indicated by red and the P1 site of the RCL is marked as a bold rectangle. This sequence alignment was obtained using the Clustal Omega (Madeira *et al.*, 2019) and ESPript (http://espript.ibcp.fr, (Robert & Gouet, 2014)). (*b*) Superposition of *I. ricinus* crystal structures, namely Iripin-5 (magenta), IRS-2 (blue) and Iripin-3 (cyan). (*c*) Comparison of electrostatic potentials of IRS-2 (PDB ID: 3NDA), Iripin-5 (PDB ID: 7B2T) and Iripin-3 (PDB ID: 7AHP). As shown in the picture, blue indicates positive, and red indicates negative potential.

The amino acids of the RCL, specifically the P1 residue, determine the protease specificity (Marijanovic *et al.*, 2019). This was confirmed by structural analysis of the *I. ricinus* salivary serpins IRS-2, Iripin-3 and Iripin-5. Iris, with Met-340 at the P1 site, is an inhibitor of leukocyte elastase and elastase-like serine proteases (Prevot *et al.*, 2007), although its inhibition is managed by several exosites in α-helices A and D (Prevot *et al.*, 2009). However, IRS-2 has Tyr-341 at its P1 site, which signifies the inhibition of chymotrypsin-like

proteases (Chmelař *et al.*, 2017) and Iripin-3 has Arg-342 at the P1 site, indicating its trypsin-like protease inhibition (Chlastáková *et al.*, 2021). Nevertheless, diverse RCL residues can represent potential cleavage sites, but only a few residues (16-17 residues from the C-terminal β-sheet) manage to successfully inhibited the target protease (Gettins, 2002). For the Iripin-5, the last visible residue of the inserted RCL is Arg-342 (Supplementary Table S1 and Fig. S4), which is the potential P1 site, suggesting the targeting of trypsin-like proteases preferring Arg or Lys side chains at the P1 site rather than elastase-like (Ala, Gly and Val) or chymotrypsin-like (Tyr, Phe and Trp) proteases (Barrett *et al.*, 2004). However, Iripin-5 mainly inhibited neutrophil elastase and proteinase 3, which is previously described behaviour of some serpins that inhibited serine proteases, despite that fact that these serpins have an inappropriate P1 recognition site and should have inhibited different proteases (Gettins, 2002). The presence of Arg at the P1 site is common for salivary serpins from prostrite ticks (Mulenga *et al.*, 2009) and led to the proposal of an interaction with blood-coagulation proteases.

Michaelis and covalent complex studies were performed to reveal the possible residues responsible for the inhibition of target proteases. The Michaelis complex is the initial step of protease inhibition; more specifically, it enable the cleavage of the scissile bond and the subsequent acylation step, and therefor represents the most informative structural conformation of serpins (Gettins, 2002). Apart from the primary recognition site of the serpin, some serpins also employ specific surface regions called exosites that can specify the protease inhibition (Gettins & Olson, 2016). For this reason, MD simulations of Michaelis complexes were performed. No exosites were found to be directly involved in formation of the Michaelis complex formation (Fig. 8). In Iripin-5 – neutrophil elastase the Michaelis complex was observed to involve engagement of Glu-330 in the Iripin-5 RCL to form salt bridges with Arg-36 of neutrophil elastase. Similarly, in Iripin-5 – proteinase 3 the Michaelis complex was observed to involve the formation of the salt bridges between Glu-345 in the Iripin-5 RCL and Lys-103 of proteinase 3 and between Val 340 of the RCL and Glu-101 after the MD simulation in all triplicates. The two resulting structures

of the Iripin-5 – neutrophil elastase Michaelis complex and the single structure of Iripin-5 – proteinase 3 Michaelis complex were confirmed to have involved interface that played important roles in complex formation (*PDBePISA*; Krissinel & Henrick, 2007)). Previously, it was observed that not only the position of the specific residues in RCL but also the dynamics of the RCL play an important role in the protease inhibition by serpins (Marijanovic *et al.*, 2019). It is probable that these two aspects are responsible for protease inhibition of the *I. ricinus* serpin Iripin-5.

Docking studies of covalent complexes revealed probable interactions between the chosen proteases and Iripin-5. The docking covalent complexes exhibit quite a large interface area, as observed previously for serpin-trypsin covalent complexes, with around 12 interacting interface residues. In the Iripin-5 - trypsin complex more residues were involved in the formation of hydrogen bonds compared with other Iripin-5 - protease complexes. These results are similar to the results of interface interaction comparison of antithrombin – trypsin and antithrombin – elastase complexes, in which the complex with trypsin made more hydrogen bonds (Rashid *et al.*, 2015). This could probably explain the important role of Glu-310 in Iripin-5, which forms salt bridges in the complex with protease. Moreover, some residues of Iripin-5 were involved in hydrogen-bond formation more frequently, namely Gln-299, Asp-301, Glu-51, Lys-288, Glu-294 and the abovementioned Glu-310 and Arg-342. We propose that these residues should play an important role in the formation of a covalent complex between Iripin-5 and protease.

CONCLUSIONS

The continuing structural studies of arthropod (ectoparasite) serpins provide an understanding of their specific functions and protease targets. Structural information on complexes with targets and cofactors would help to understand the exact mechanism of action of these functionally diverse serpins. Iripin-5 is the third described crystal structure of a tick serpin, and despite its cleaved form; it provides an important experimental proof of the specificity of Iripin-5 and its possible interactions with proteases. Iripin-5 appears to be an

immunomodulatory and anti-inflammatory protein used by *I. ricinus* ticks to overcome host defensive mechanisms. The presence of Arg at the P1 site led to the proposal of an interaction with blood-coagulation proteases. MD simulations of the Michaelis complex revealed flexibility of the RCL to be one of the factors responsible for inhibition. A more detailed study of the dynamic behavior of Iripin-5 during the inhibition mechanism may be beneficial for better understanding of inhibition. The residues with the most important roles in the formation of a covalent complex between Iripin-5 and proteases were proposed based on docking and MD simulations and it was found that Glu-310 should play a crucial role in the interaction between Iripin-5 and proteases, with the exception of α-chymotrypsin.

RELATED LITERATURE

The following references are cited in the supporting information for this article: Cox & Mann (2008), Cox *et al.* (2011), Rappsilber *et al.* (2007) and Shevchenko *et al.* (2006).

FUNDING INFORMATION

REFERENCES

Barrett, A. J., Rawlings, N. D. & Woessner, J. F. (2004). Handbook of Proteolytic Enzymes, 2nd ed, Vol. 1. London: Academic Press.

Belorgey, D., Ha¨ gglo¨ f, P., Karlsson-Li, S. & Lomas, D. A. (2007). Prion, 1, 15–20.

Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. (1987). J. Phys. Chem. 91, 6269–6271.

Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. (1995). Comput. Phys. Commun. 91, 43–56.

Berman, H., Henrick, K. & Nakamura, H. (2003). Nat. Struct. Mol. Biol. 10, 980.

Bos, I. G. A., Hack, C. E. & Abrahams, J. P. (2002). Immunobiology, 205, 518–533.

Chlastáková, A., Kota´l, J., Bera´nkova´, Z., Kasˇcˇaʹkovaʹ, B., Martins, L. A., Langhansovaʹ, H., Prudnikova, T., Ederovaʹ, M., Kutaʹ Smatanovaʹ, I., Kotsyfakis, M. & Chmelarˇ, J. (2021). Front. Immunol. 12, 626200.

Chmelarˇ, J., Kotaʹl, J., Kopeckyʹ, J., Pedra, J. H. F. & Kotsyfakis, M. (2016). Trends Parasitol. 32, 368–377.

Chmelarˇ, J., Kotaʹl, J., Langhansovaʹ, H. & Kotsyfakis, M. (2017). Front. Cell. Infect. Microbiol. 7, 216.

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. (2011). Blood, 117, 736–744.

Cox, J. & Mann, M. (2008). Nat. Biotechnol. 26, 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V. & Mann, M. (2011). J. Proteome Res. 10, 1794–1805.

Daix, V., Schroeder, H., Praet, N., Georgin, J.-P., Chiappino, I., Gillet, L., de Fays, K., Decrem, Y., Leboulle, G., Godfroid, E., Bollen, A., Pastoret, P.-P., Gern, L., Sharp, P. M. & Vanderplasschen, A. (2007). Insect Mol. Biol. 16, 155–166.

DeLano, W. L. (2002). PyMOL. http://www.pymol.org.

Dunstone, M. A. & Whisstock, J. C. (2011). Methods Enzymol. 501, 63–87.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Acta Cryst. D66, 486–501.

Feller, S. E. & MacKerell, A. D. (2000). J. Phys. Chem. B, 104, 7510– 7515.

Francischetti, I. M. B., Sa-Nunes, A., Mans, B. J., Santos, I. M. & Ribeiro, J. M. C. (2009). Front. Biosci. 14 2051–2088.

Fujinaga, M., Chernaia, M. M., Halenbeck, R., Koths, K. & James, M. N. G. (1996). J. Mol. Biol. 261, 267–278.

Gettins, P. G. W. (2002). Chem. Rev. 102, 4751–4804.

Gettins, P. G. W. & Olson, S. T. (2016). Biochem. J. 473, 2273–2293. Hansen, G., Gielen-Haertwig, H., Reinemer, P., Schomburg, D., Harrenga, A. & Niefind, K. (2011). J. Mol. Biol. 409, 681–691. Humphrey, W., Dalke, A. & Schulten, K. (1996). J. Mol. Graph. 14, 33–38.

Huntington, J. A. (2011). J. Thromb. Haemost. 9, Suppl. 1, 26–34. Kabsch, W. (2010). Acta Cryst. D66, 133–144.

Katona, G., Berglund, G. I., Hajdu, J., Gra´f, L. & Szila´gyi, L. (2002). J. Mol. Biol. 315, 1209–1218.

Kervinen, J., Crysler, C., Bayoumy, S., Abad, M. C., Spurlino, J., Deckman, I., Greco, M. N., Maryanoff, B. E. & de Garavilla, L. (2010). Biochem. Pharmacol. 80, 1033–1041.

Klauda, J. B., Brooks, B. R., MacKerell, A. D., Venable, R. M. & Pastor, R. W. (2005). J. Phys. Chem. B, 109, 5300–5311.

Kota´l, J., Langhansova´, H., Lieskovska´, J., Andersen, J. F., Francischetti, I. M. B., Chavakis, T., Kopecky´, J., Pedra, J. H. F., Kotsyfakis, M. & Chmelarˇ, J. (2015). J. Proteomics, 128, 58–68.

Kovářˇová, Z., Chmelarˇ, J., Sˇanda, M., Brynda, J., Maresˇ, M. & Rˇ eza´cˇova´, P. (2010). Acta Cryst. F66, 1453–1457.

Krissinel, E. & Henrick, K. (2007). J. Mol. Biol. 372, 774–797.

Ky´cˇkova´, K. & Kopecky´, J. (2006). J. Med. Entomol. 43, 1208–1214. Laroux, F. S., Pavlick, K. P., Hines, I. N., Kawachi, S., Harada, H.,

Bharwani, S., Hoffman, J. M. & Grisham, M. B. (2001). Acta Physiol. Scand. 173, 113–118.

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. (2006). Genome Biol. 7, 216.

Leboulle, G., Crippa, M., Decrem, Y., Mejri, N., Brossard, M., Bollen, A. & Godfroid, E. (2002). J. Biol. Chem. 277, 10083–10089.

Livak, K. J. & Schmittgen, T. D. (2001). Methods, 25, 402–408. Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D. & Lopez, R. (2019). Nucleic Acids Res. 47, W636–W641.

Marijanovic, E. M., Fodor, J., Riley, B. T., Porebski, B. T., Costa, M. G. S., Kass, I., Hoke, D. E., McGowan, S. & Buckle, A. M. (2019). Sci. Rep. 9, 3870.

Mellet, P., Michels, B. & Bieth, J. G. (1996). J. Biol. Chem. 271, 30311– 30314.

Mueller, U., Darowski, N., Fuchs, M. R., Fo¨ rster, R., Hellmig, M., Paithankar, K. S., Pu¨ hringer, S., Steffien, M., Zocher, G. & Weiss, M. S. (2012). J. Synchrotron Rad. 19, 442–449.

Mulenga, A., Khumthong, R. & Chalaire, K. C. (2009). BMC Genomics, 10, 217.

Murshudov, G. N., Skuba´k, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). Acta Cryst. D67, 355–367.

Nagata, K. (1996). Trends Biochem. Sci. 21, 23–26.

Pa´lenıkova´, J., Lieskovska´, J., Langhansova´, H., Kotsyfakis, M., Chmelarˇ, J. & Kopecky´, J. (2015). Infect. Immun. 83, 1949–1956. Pemberton, P. A., Stein, P. E., Pepys, M. B., Potter, J. M. & Carrell, R. W. (1988). Nature, 336, 257–258.

Porter, L., Radulovic´, Zˇ ., Kim, T., Braz, G. R. C., Da Silva Vaz, I. & Mulenga, A. (2015). Ticks Tick Borne Dis. 6, 16–30.

Prevot, P.-P., Adam, B., Boudjeltia, K. Z., Brossard, M., Lins, L., Cauchie, P., Brasseur, R., Vanhaeverbeek, M., Vanhamme, L. & Godfroid, E. (2006). J. Biol. Chem. 281, 26361–26369.

Prevot, P.-P., Beschin, A., Lins, L., Beaufays, J., Grosjean, A., Bruys, L., Adam, B., Brossard, M., Brasseur, R., Zouaoui Boudjeltia, K., Vanhamme, L. & Godfroid, E. (2009). FEBS J. 276, 3235–3246.

Prevot, P.-P., Couvreur, B., Denis, V., Brossard, M., Vanhamme, L. & Godfroid, E. (2007). Vaccine, 25, 3284–3292.

Rappsilber, J., Mann, M. & Ishihama, Y. (2007). Nat. Protoc. 2, 1896– 1906.

Rashid, Q., Kapil, C., Singh, P., Kumari, V. & Jairajpuri, M. A. (2015). J. Biomol. Struct. Dyn. 33, 1352–1362.

Robert, X. & Gouet, P. (2014). Nucleic Acids Res. 42, W320–W324. Schlee, S., Straub, K., Schwab, T., Kinateder, T., Merkl, R. & Sterner, R. (2019). Proteins, 87, 815–825.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. (2006). Nat. Protoc. 1, 2856–2860.

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. (2001). J. Biol. Chem. 276, 33293–33296. Sparta, K. M., Krug, M., Heinemann, U., Mueller, U. & Weiss, M. S. (2016). J. Appl. Cryst. 49, 1085–1092.

Spence, M. A., Mortimer, M. D., Buckle, A. M., Minh, B. Q. & Jackson, C. J. (2021). Mol. Biol. Evol. 38, 2915–2929.

Sprong, H., Azagi, T., Hoornstra, D., Nijhof, A. M., Knorr, S., Baarsma, M. E. & Hovius, J. W. (2018). Parasites Vectors, 11, 145. Tirloni, L., Seixas, A., Mulenga, A., da Silva Vaz, I. Jr & Termignoni, C. (2014). Exp. Parasitol. 137, 25–34.

Tsukada, H. & Blow, D. M. (1985). J. Mol. Biol. 184, 703–711. Tyson, K. R., Elkins, C. & de Silva, A. M. (2008). J. Immunol. 180, 3964–3968.

Vagin, A. & Teplyakov, A. (2010). Acta Cryst. D66, 22–25. Vangone, A., Spinelli, R., Scarano, V., Cavallo, L. & Oliva, R. (2011). Bioinformatics, 27, 2915–2916.

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (2018). Protein Sci. 27, 293–315.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). Acta Cryst. D67, 235–242.

Yamasaki, M., Arii, Y., Mikami, B. & Hirose, M. (2002). J. Mol. Biol. 315, 113–120.

Zou, Z., Anisowicz, A., Hendrix, M. J., Thor, A., Neveu, M., Sheng, S., Rafidi, K., Seftor, E. & Sager, R. (1994). Science, 263, 526–529.

Zundert, G. C. P. van, Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., Melquiond, A. S. J., van Dijk, M., de Vries, S. J. & Bonvin, A. M. J. J. (2016). J. Mol. Biol. 428, 720– 725.

3.2.1 Supporting information for an article

S1. Materials and Methods

S1.1. MALDI-MS protein analysis

Protein in solution was purified using ZipTip C18 tips (Merck Millipore, MA, USA). The sample was placed onto an MTP 384 target plate ground steel BC (Bruker Daltonics, Bremen, Germany) previously modified using saturated sinapinic acid (Bruker Daltonics) in ethanol and mixed with saturated sinapinic acid in 30% acetonitril/0.1% formic acid. The sample was also applied onto an MTP AnchorChipTM 384 target plate (Bruker Daltonics) and mixed with α-cyano-4-hydroxycinnamic acid (Bruker Daltonik) in 88.8% acetonitril/0.1% formic acid. Mass spectrometric measurements were performed on an Autoflex Speed MALDI-TOF/TOF (Bruker Daltonics). Mass spectra were acquired in positive linear ion mode (acceleration voltage: 19.5 kV; extraction voltage: 16.65 kV; lens voltage: 9 kV; delayed extraction time: 1200 ns) and positive reflectron mode (acceleration voltage: 19 kV; extraction voltage: 16.6 kV; lens voltage: 8.6 kV; delayed extraction time: 140 ns), respectively. Mass spectra were accumulated from up to 500 laser shots; a mixture of protein or peptide standards (Bruker Daltonics) were used for external calibration.

S1.2. NanoLC-ESI-MS/MS protein analysis

In-gel digestion of protein was performed according to a standard protocol (Shevchenko *et al.*, 2006). The extracted peptides were purified using C18 Empore disks (3M, St. Paul, USA) (Rappsilber *et al.*, 2007). The nanoLC-ESI-MS/MS analysis was carried out on an UltiMate 3000 RLSCnano system (Thermo Fisher Scientific, MA, USA) coupled online to mass spectrometer timsTOF Pro (Bruker Daltonics, Bremen, Germany). Peptides were dissolved in 30 µl of 3 % acetonitrile/0.1 % formic acid and 2 µl of 20-fold diluted peptide solution was injected onto an Acclaim PepMap 100 C18 trapping column (300 µm i.d., 5 mm length, particle size 5 µm, pore size 100 Å; Thermo Fisher Scientific) at a 2.5 µl/min flow rate. Bound peptides were eluted from the trapping column onto an Acclaim PepMap 100 C18 analytical column (75 µm i.d., 150 mm length, particle size 2 µm, pore size 100 Å; Thermo Fisher Scientific) and separated by a 20 min long linear gradient of 5-35 % acetonitrile/0.1 % formic acid at a constant rate of 0.3 µl/min. The column oven temperature was set to 35 °C. The MS analysis was operated in PASEF scan mode with positive polarity. Electrospray ionization was performed using a CaptiveSpray (Bruker Daltonics) with capillary voltage at 1500 V, dry gas at 3 l/min and dry temperature at 180 °C. Ions were accumulated for 100 ms and 10 PASEF MS/MS scans were acquired per topN acquisition cycle. An ion mobility range (1/K0) was set at 0.6-1.6 Vs/cm2. Mass spectra were collected over a m/z range of 100 to 1700. Polygon filtering was applied to exclude the low m/z of singly charged ions. Target intensity was set to 20 000 to repeatedly select

precursor for PASEF MS/MS repetitions. The precursors that reached the target intensity were then excluded for 0.4 min. Collision energies were changed from 20 to 59 eV in 5 steps of equal width between 0.6 and 1.6 Vs/cm2 of 1/K0 values.

Acquired MS and MS/MS data were submitted for database searching using MaxQuant software (version 1.6.14) (Cox & Mann, 2008) with an integrated Andromeda search engine (Cox *et al.*, 2011). A protein sequence database supplemented with a contaminant database included in the MaxQuant software was used to identify proteins. Default parameters for TIMS-DDA search type and Bruker TIMS instrument were applied. Trypsin/P was set as an enzyme allowing up to two missed cleavages in specific digestion mode; carbamidomethylation of cysteine was used as fixed modification; methionine oxidation and protein N-term acetylation were set as variable modifications; the minimum required peptide length was set to five amino acids. Precursor ion tolerance was set at 20 and 10 ppm in the first and main peptide search, respectively; the mass tolerance for MS/MS fragment ions was set at 40 ppm; peptide spectrum match (PSM) and protein identifications were filtered using a target-decoy approach at a false discovery rate (FDR) of 1 %.

S Data Tables

Table S1 Intact protein analysis using MALDI-TOF MS. Detected peptides resulting from protease digestion are listed from the smallest theoretical mass. The results from intact protein analysis using MALDI-TOF MS are shown in Fig. 4 (supplementary material).

peptide	theoretical mass
VPTLELNVNQPFLFFIRNTHTKDLLFAGQVNHL	3 835,06 Da
EVPTLELNVNQPFLFFIRNTHTKDLLFAGQVNHL	3 964,10 Da
IEVPTLELNVNQPFLFFIRNTHTKDLLFAGQVNHL	4 077,18 Da
LIEVPTLELNVNQPFLFFIRNTHTKDLLFAGQVNHL	4 190,27 Da
RLIEVPTLELNVNQPFLFFIRNTHTKDLLFAGQVNHL	4 346,37 Da

Table S2 Summary of the interactions between Iripin-5 and chosen proteases residues from docking studies of theoretical covalent complex conformations.

Iripin-5	cathepsin G	trypsin	neutrophil elastase	α-chymo trypsin	chymase	proteinase 3
Lys-49						Trp-218
Glu-51	Ser-218	Lys-175		Ser-218		
Lys-54	Ser-218					
Glu-102					Arg-174	
Arg-103					Arg-174	Glu-97
Asp-159		Ser-39				
Asp-159		Arg-193				
Arg-163					Thr-96	Glu-97
Lys-288		Ser-146	Asn-61	Thr-219, Ser-218, Tyr-146		
Asp-289		Gly-148, Ser-214				
Thr-293		His-217				
Glu-294		Lys-175, His-217		Lys-175		Arg-143
Thr-298			Asp-102		Gly-37, Pro-38	Asp-61
Gln-299	Ser-214, Asp-102, Ser-195	Thr-98	His-57, Ser-214	Thr-98		His-40, Gln-299
Ala-300	Gln-96					
Asp-301	Gln-96	Arg-96	Gly-218, Gly-219		Lys-192	
Ser-303					Lys-192	
Ser-306					Ser-218	
Asp-308			Arg-217			Lys-99
Glu-310	Arg-90	Tyr-94, Arg-90	Arg-177, Arg-217	Tyr-94	Lys-40	
Thr-341		Lys-60				
Arg-342	Gln-96	Lys-60	Arg-177	Ser-96		

Table S3 Summary of the interactions between Iripin-5 and neutrophil elastase or proteinase 3 residues from MD studies of theoretical Michaelis complexes

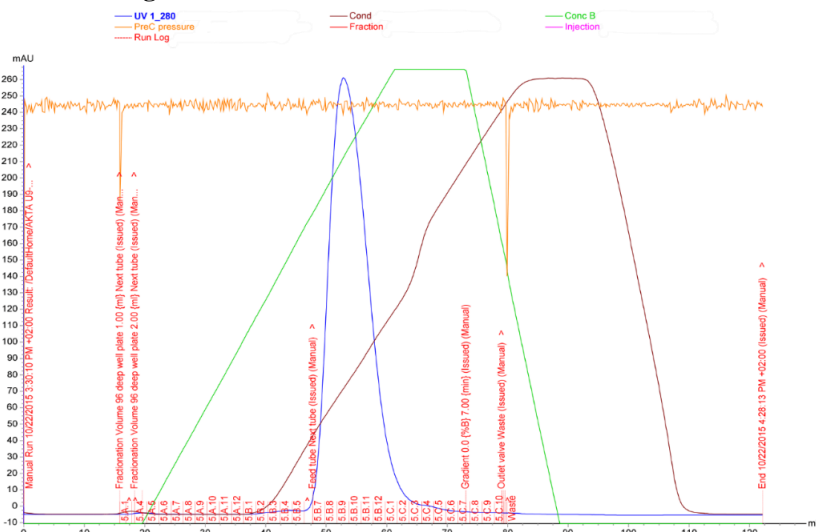| Iripin-5	neutrophil elastase	Iripin-5	proteinase 3
Glu-330	**Arg-36**	**Ser-335**	Trp-210
Ala-333	**Arg-36, Gly-37**	**Glu-345**	Lys-103
Glu-330	Arg-36	**Val-340**	Glu-101
Ala-333	Arg-36, Gly-37	**Arg-342**	His-59
Glu-345	His-56	**Ser-335**	Trp-210
Glu-330	**Arg-36**	**Glu-345**	Lys-103
Ala-332	**Asn-62**	**Val-340**	Glu-101
		GLU-345	**LYS-103**
		VAL-340	**GLU-101**

S Data Figures

Figure. S1 Anion exchange chromatography (HiTrap® Q High Performance - GE17-1154-01 column) with use of 0.5 M NaCl as an eluent. The equilibration buffer composition was 20 mM Tris, pH 8.5 and elution buffer was 20mM Tris, 500mM NaCl, pH 8.5. Iripin-5 started to elute at ~68% of 0.5M NaCl gradient and fractions B7-C1 were used for further purification step.
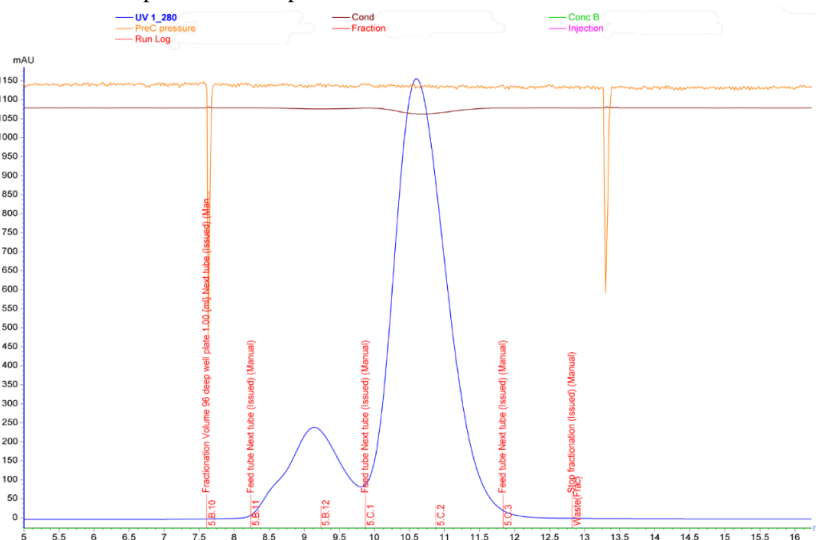

Figure. S2 Size exclusion chromatography (Superdex 75 10/300). The buffer composition was 20mM Tris, 150mM NaCl, pH 8.5. The first peak probably represents the multimeric state of protein and second, larger peak monomeric Iripin-5 that was used for further analyses and for crystallization.

MRYENEMRLANNRFAVDLLRGLPSSPEKNIFFSPYSISTAMGMVFAGAKGETLKNLYDGF
GYLRSGLKEDWVLQAYADHAKQLQVGQSQSTFDVANAAAIHERLALLSAYENTLDSTFHA
QLLKVDFVNGGPAAIDEINRWVKQKTHDKIDKLFDGPLDPLTRLVLLNAIFFKGVWSTKF
DENATTKKQFLNGGTTPTQVDTMTKSIRIGYKLLPTMRLEIAELPYDGGNYSMVILLPRG
SEGIEAFKHSLTDHRLQDYIGHVELREVAVSLPKFKLETEYSLKDSLKSLGITEIFGTQA
DLSGISSDGELVVSDVVHKAVVEVNEEGTEAAAVSGVAVVTRLIEVPTLELNVNQPFLFF
IRNTHTKDLLFAGQVNHL

Figure. S3 Results of protein Iripin-5 identification from gel pieces using timsTOF Pro. The analysis of protein sequence by MS/MS has shown the presence of Iripin-5 (marked red) but missing the C-terminal part (LEU-343 – LEU-378). The presence of theoretical P1 site is highlighted by square around representing residue.
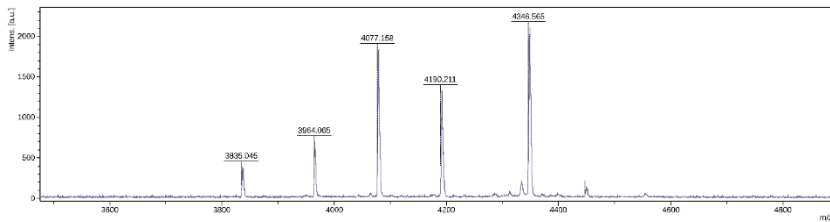

Figure. S4 Intact protein analysis using MALDI-TOF MS. Each peak represents the mass of the observed peptides from analysis. These peptides can be results from protease digestion probably during the sample storage. The sequence and theoretical mass for each peptide are listed in Table 1 (supplementary material).
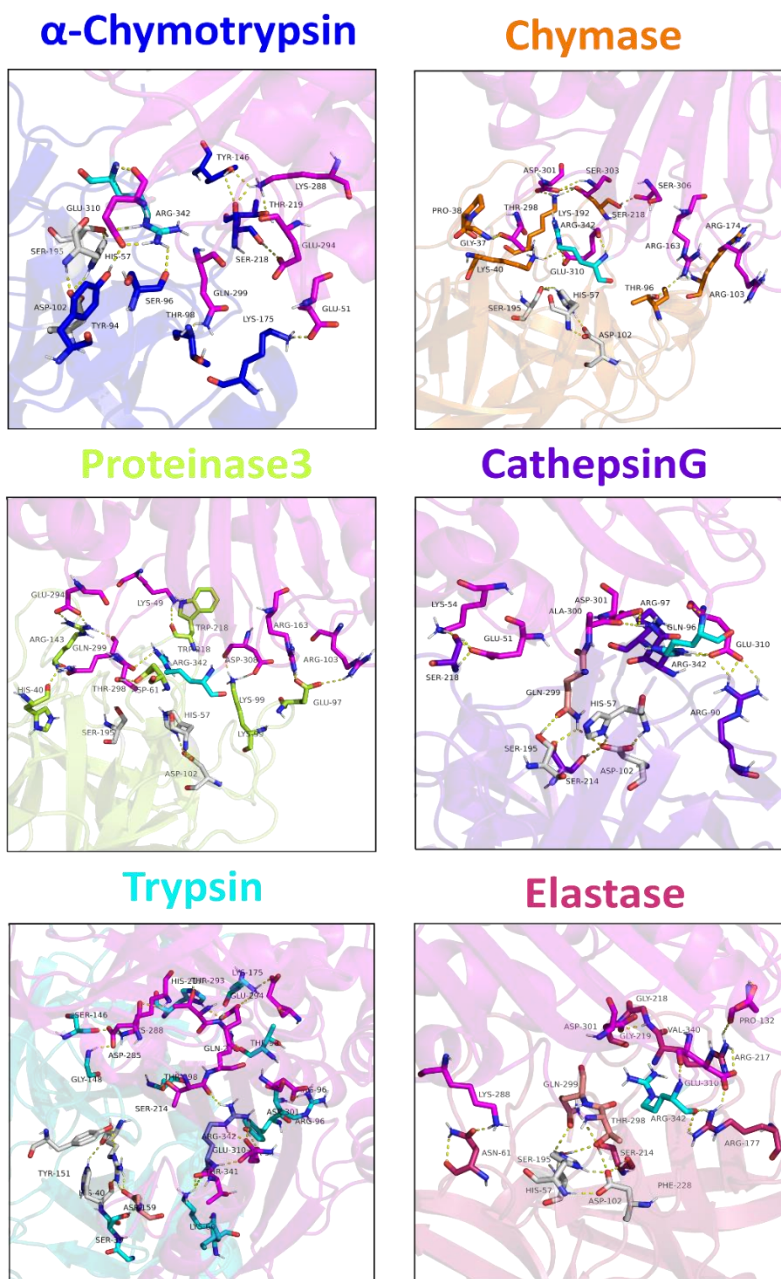
Figure. S5 The focus and visualization of Iripin-5 and chosen proteases [α-Chymotrypsin (blue), Chymase (orange), Proteinase 3 (lemon), Cathepsin G (violet), Trypsin (cyan), Elastase (hot pink)] are rotated for a better view. The catalytic residues side chains of proteases are presented like grey sticks; theoretical P1 site of Iripin-5 (ARG-342) is cyan; the residues interacting with protease catalytic triad are wheat; H-bond interactions are shown as dashed lines.

Michaelis complex

Iripin-5 - proteinase 3

Iripin-5 - neutrophil elastase

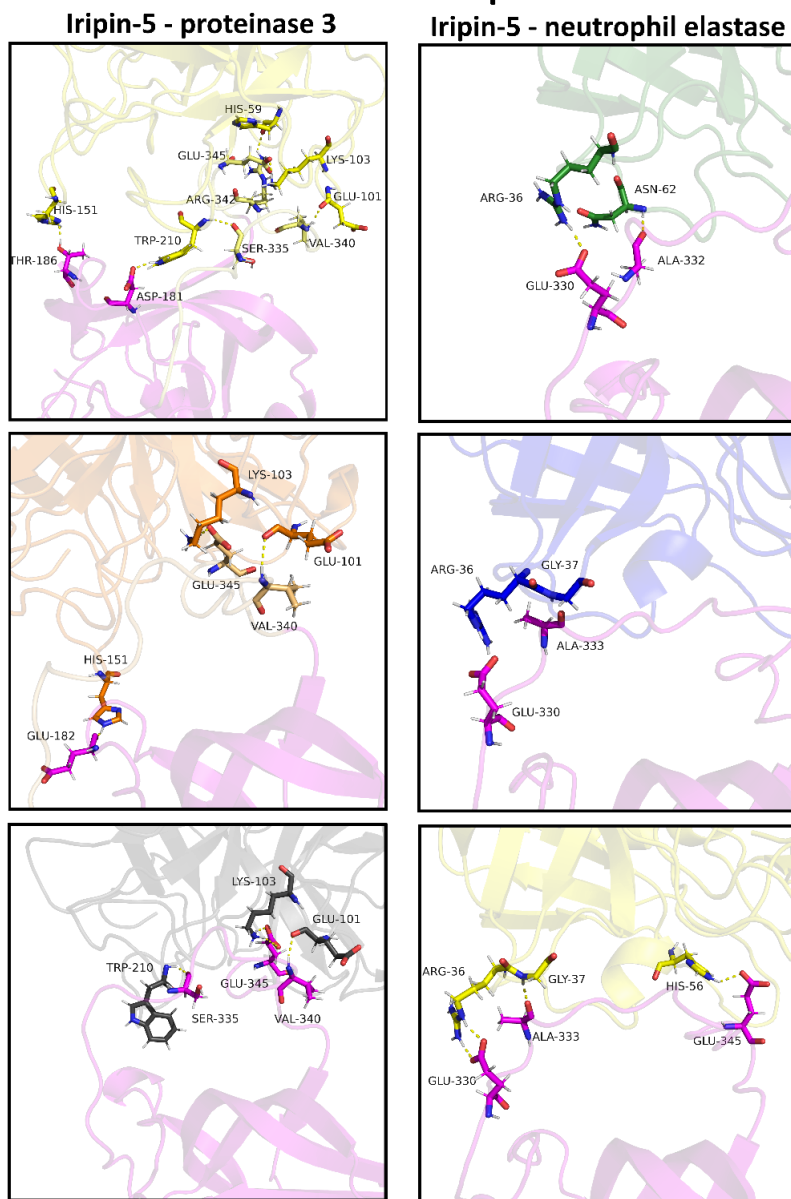**Figure. S6** The focus and visualization of Michaelis complex interface of Iripin-5 and chosen proteases, the zoom is rotated for a better view. The side chain residues of proteases are presented by different colour (proteinase 3 – yellow, orange, and grey; neutrophil elastase – green, blue, and yellow) and residues of Iripin-5 RCL are magenta; H-bond interactions are shown as dashed lines.

REFERENCES

Cox, J. & Mann, M. (2008). *Nature biotechnology*. **26**, 1367–1372, doi:10.1038/nbt.1511.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V. & Mann, M. (2011). *Journal of proteome research*. **10**, 1794–1805, doi:10.1021/pr101065j.

Rappsilber, J., Mann, M. & Ishihama, Y. (2007). *Nature protocols*. **2**, 1896–1906, doi:10.1038/nprot.2007.261.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. (2006). *Nature protocols*. **1**, 2856–2860, doi:10.1038/nprot.2006.468.

3.3 Conformational transition of Iripin-4, the *Ixodes ricinus* salivary serpin

This chapter is based on Paper III.:

ABSTRACT

Iripin-4, one of the many salivary serpins of *Ixodes ricinus* ticks, with hitherto unexplained function, crystallized in two different structural conformations. This paper describes the structural information of the native partially relaxed state and cleaved serpin conformations. The native structure was solved at 2.3 Å resolution and the structure of cleaved conformation at 2.0 Å resolution. Furthermore, structural changes during reactive-centre loop transition from native to cleaved conformation were observed. In addition to this finding, we confirm that the main substrate-recognition site for the inhibitory mechanism is represented by Glu341. The presence of glutamate instead of typical arginine at the P1 recognition site for all structurally described *I. ricinus* serpins (7B2T, 7PMU and 7AHP) except tyrosine in IRS-2 P1 site (3NDA) would explain no protease inhibition of tested proteases that cleave their substrate after arginine. Further research on Iripin-4 should focus on functional analysis of this interesting serpin.

Keywords: serpin; X-ray structure; native conformation; cleaved conformation; *Ixodes ricinus*.

INTRODUCTION

Ixodes ricinus (Acari: Ixodidae) is one of the most important European vectors that transmits many pathogens and is of both veterinary and medical significance. This hard tick species attacks a large variety of cold-blooded and warm-blooded vertebrate hosts including humans. The two most frequently reported tick-borne infections of humans are Lyme borreliosis caused by

Borrelia burgdorferi s.l. and the Western European subtype of tick-borne encephalitis virus of which *I. ricinus* is a main vector in Europe [1,2].

Because of the large area of *I. ricinus* spread across Europe, there were attempts to decrease tick abundance and thus reduce the risk of infection. The use of acaricides is one option, however in the long-term run, ticks can develop resistance to these products. Another strategy is the use of repellents that, however, often contain toxic substances that can cause environmental contamination. Moreover, the repellents offer only short-term protection [3]. Biological approach to tick control, such as host vaccination, seems to be valid choice as it can induce protective immunity leading to tick rejection by the host. Tick saliva contains many important substances that could be key for transmission-blocking vaccine development Among [4]. these substances, serine protease inhibitors (serpins) have been tested in several studies and recombinant serpins from several tick species were described as a candidate for vaccine design [5]. For example*, I. ricinus* serpin named IRIS utilised a protective immunity against nymphs and adults fed on vaccinated rabbits but not nymphs fed on mice [6,7].

Serpins form the largest family of protease inhibitors with more than 6000 members widely distributed across all organisms and function in many physiological processes[8,9] Serpins differ from other groups of protease inhibitors by a unique mode of action. In order to inhibit target protease, serpins undergo large conformational change that leads to both protease and serpin inactivation. Thus, serpins are referred to as "one use only" or "suicide" inhibitors [10,11]. There is rich evidence that some serpins, such as viral CrmA or plant serpin1 can inhibit also cysteine proteases in addition to serine ones [12,13]. Serpins contain also non-inhibitory members with roles in hormone transport, protein folding, regulation of blood pressure, chromatin condensation or tumour progression [14].

The main feature of all serpins is a conserved tertiary structure composed of ~350 amino acids core domain. The secondary structure fold of serpins consists of 3 β-sheets (A-, B- and C- β-sheet), at least 8 α-helices (A-I) and a reactive centre loop (RCL) [15]. The amino acids of RCL, especially the

residues at P1 and P1' positions, determine the protease selectivity of serpins by mimicking the target protease substrate [16]. Thanks to excessive crystallographic studies, there is structural data on possible conformations of serpins that are independent from the process of inhibition. The conformations can be divided into a range of monomeric serpin structures (native fully stressed state, native partially relaxed state, latent conformation, abnormal δ-conformation, and cleaved conformation) and complexes with inhibited protease (Michaelis-Menten complex and covalent complex) [11].

The mechanism of protease inhibition is well explained because all inhibitory serpins undergo a similar molecular process. The protease recognizes the exposed RCL of native, metastable serpin as the "substrate" and forms the Michaelis-Menten complex with it. Subsequently, the RCL is cleaved between P1 and P1′ sites by protease, which leads to the binding of RCL, which is transported to the other side of the serpin. This transition to the more stable serpin conformation results in RCL insertion into the β-sheet A as an additional β-strand and simultaneous inactivation of serpin and protease by forming a covalent complex. When the reaction is not fast enough, the result of the process is inactivated serpin in its most thermally stable cleaved confirmation and released active protease [11,17,18]. The process of inhibition represents a very flexible structural change that can make serpins vulnerable to mutations. Such mutations can lead to pathologies called serpinopathies that manifest as thrombosis, cirrhosis, emphysema, immune hypersensitivity, and other diseases that are results of serpin dysfunction [19].

To fully understand the role of tick serpins there is a necessity to uncover the structural information of serpins and their complexes with proteases. This information is essential starting point for the design and development of novel pharmaceuticals. Therefore, X-ray crystallography was used for solving *I. ricinus* serpin Iripin-4 structures for initial native conformation and resulting unsuccessful inhibition represented by cleaved conformation.

RESULTS

Recombinant Iripin-4 production and purification

The amplified transcript of the Iripin-4 gene most probably encoded the homologous protein with three amino acid substitutions: H78Q, G155E, G307D compared to the original Iripin-4 sequence (GADI01002650) that was used for primer design. This happened most probably due to the high inter-tick variability of this multigenic family as reported in Kotal et al., 2021 [20]. The acquired protein has no substitutions in RCL or hinge region, therefore it should retain the same inhibitory function and specificity as the original Iripin-4. The final yield of IRS-4 was 17mg (~2mg per litre of medium).
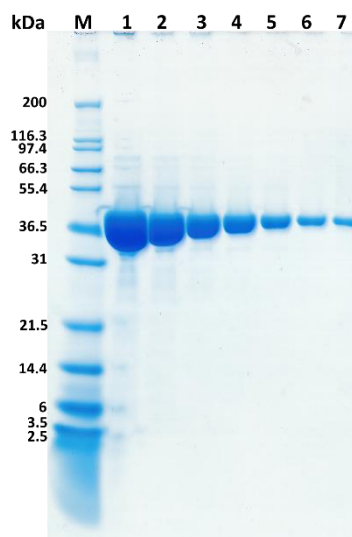
Figure. 1   Iripin-4was analyzed by a reducing SDS-PAGE gel. M: Molecular weight marker, 1-7: Iripin-4 with load of 50, 25, 12.5, 6.2, 3.1, 1.55, 0.8 µg per well.

Analysis of complex formation of Iripin-4 with proteases

Serpin Iripin-4 is not an inhibitor of chosen serine proteases. We did not observe any complex formation between Iripin-4 and any of tested serine proteases. Furthermore, Iripin-4 acted as a substrate for pork elastase, as we can see a double band of a native and nicked serpin.

Figure. 2   Analysis of SDS and heat-stable complex formation between Iripin-4 and tested proteases. No complex formation was observed. There is visible cleavage of Iripin-4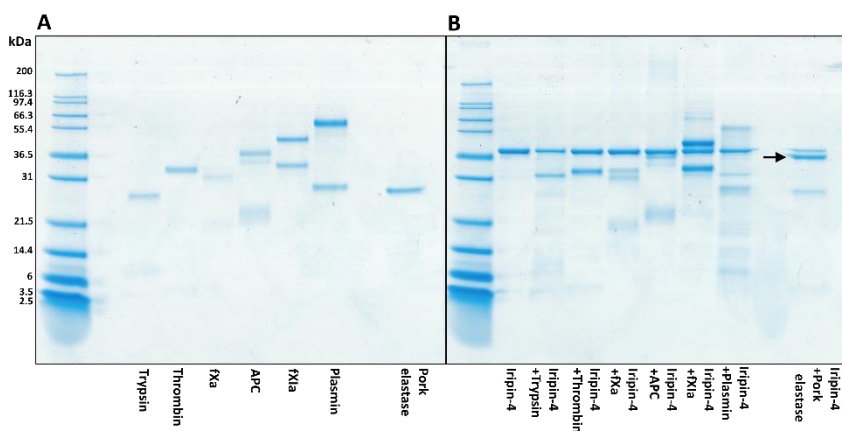 by pork elastase showed by black arrow. Proteins were resolved on 12% NuPAGE Bis-Tris gels and visualized by Coomassie staining.

Crystallization of Iripin-4 conformations

From the initial screening, several successful crystallization conditions were found for both protein concentrations. Preliminary data collection showed that crystal optimization is needed and thus the 1.17 mg/ml protein concentration was used for further optimization with a larger drop volume (2μl). The crystals of native Iripin-4 grew in condition consisting of 25%(w/v) PEG 3350, 0.1 M Bis-Tris pH 5.5 and 0.2 M ammonium acetate (Fig. 3). The difference in crystallization condition of cleaved conformation was the presence of 0.2 M sodium chloride instead of 0.2 M ammonium acetate used in native serpin crystallization (Fig. 4). One of the most probable reasons for obtaining the crystals of both conformations was the storage time of the protein sample. In the case of native Iripin-4 crystals, the protein sample was freshly obtained, whereas cleaved Iripin-4 crystals were formed from the protein sample stored at -20°C for 6 months before starting optimization.
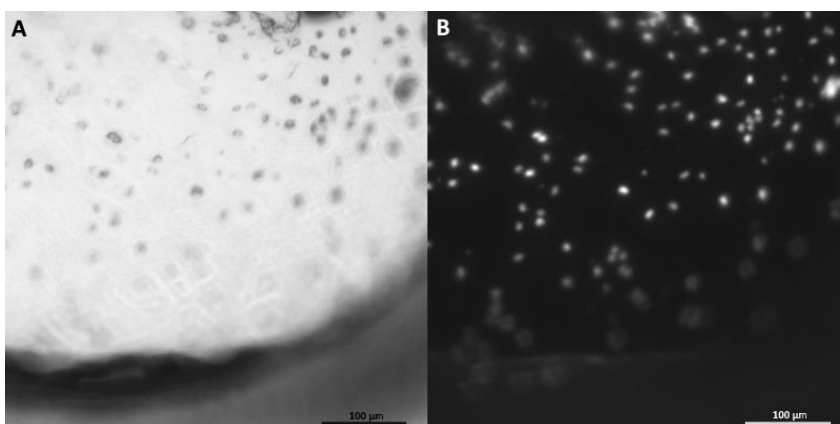
Figure. 3   Crystals of native Iripin-4. (A) Crystals of protein grown in 25%(w/v) PEG 3350, 0.1 M Bis-Tris pH 5.5, 0.2 M ammonium acetate. (B) The same crystallization droplet is shown under UV light. (A) and (B) were taken using a JANSi UVEXm (SWISSCI, UK). The scale bar represents100 μm.
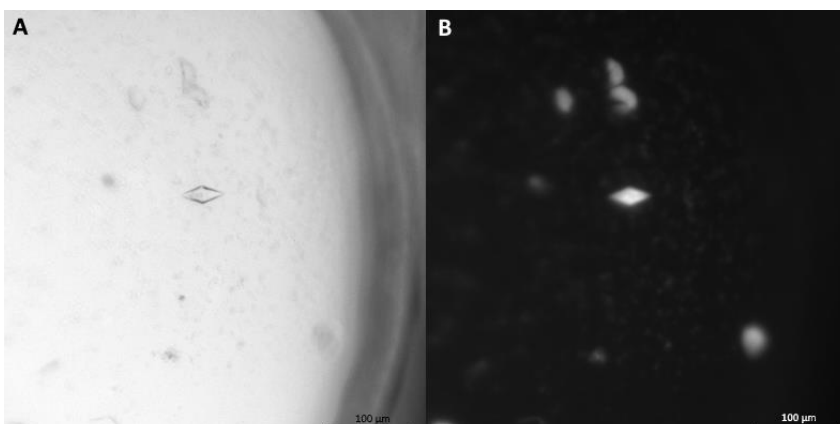
Figure. 4   Crystals of cleaved Iripin-4. (A) Crystals of protein grown in 25%(w/v) PEG 3350, 0.1 M Bis-Tris pH 5.5, 0.2 M sodium chloride. (B) The same crystallization droplet is shown under UV light. (A) and (B) were taken using a JANSi UVEXm (SWISSCI, UK). The scale bar represents100 μm.

The partially relaxed state of native conformation

The crystal of native Iripin-4 was solved at 2.3 Å resolution and crystal belongs to P3$_1$21 (152) space group, with unit-cell parameters a=77.78, b=77.78, c=109.49 (Å), α= 90.0, β=90.0, γ=120.0 (°). Crystal contains one molecule in the asymmetric unit with 51.33% solvent content and a Matthews coefficient of 2.53 Å^3 Da^{-1}. Detailed refinement statistic is listed in Tab. 1.

Overall Iripin-4 structure consists of all 376 residues with a typical native serpin secondary structure fold with RCL exposed on top of β-sheet C prepared as protease "bait (substrate)" (Fig. 7A) [17]. The structure is composed of 9 α-helices and 3 β-sheets (A, B and C) sequentially arranged in the order αA-β1-αB-αC-αD-β2-αE-β3-αF-β4-β5-β6-β7-β8-αG-αH-β9-β10-αI-β11-β12-β13-β14 (Fig. 7B). β-sheet A consists of 5 β-strands, β-sheet B of 5 β-strands and β-sheet C of 4 β-strands. The crystal structure contains the Nickel cation (Ni^{2+}) located between two Iripin-4 molecules in crystallographic symmetry and interacting with His253 and Asp257 of both molecules (Fig. 5). The probable presence of Ni^{2+} ion is caused by the release of ion during affinity purification.
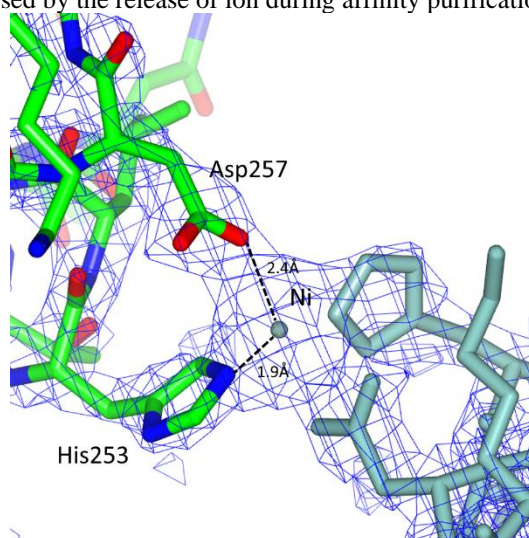
Figure. 5   Visualization of the Nickel cation between two molecules of Iripin-4 in crystal. The Asp257 and His253 interacting with Ni^{2+} are marked. The distances between interacting atoms are 1.9 Å for His253 $N^{\epsilon 2}$ and Ni^{2+}, and 2.4 Å for Asp257 $O^{\delta 2}$ and Ni^{2+}

The Iripin-4 crystal structure shows the typical partially relaxed state of native serpins, when the RCL, situated between β-sheet A and β-sheet C, is partially inserted into the breach region (Fig. 8) (top part of β-sheet A) [20] whereas the remaining part of RCL is exposed (Fig. 7A). The result of the crystallization process trapped serpin in one of two possible states considering high RCL flexibility and its dynamic equilibrium between partially inserted and fully exposed native state. The six amino acids namely: Ala326,

Gly327, Thr328, Glu329, Ala330 and Ala331 are inserted between the 3rd β-strand (s3A) and 5th β-strand (s5A) of β-sheet A (Fig. 6). Three amino acids of the exposed part of RCL, more specifically Thr334, Gly335 that are part of the hinge region and Leu343 were modelled into the structure even though there was lower observed electron density that indicates high flexibility of these two regions of RCL (Fig. 6).
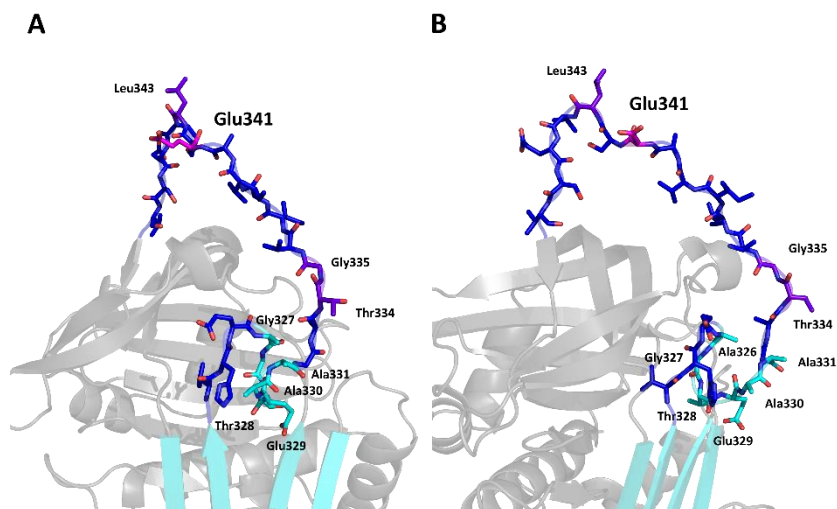
Figure. 6   The RCL of native Iripin-4. (A) Front view of RCL (blue) with marked inserted residues into β-sheet A (cyan) and highly flexible residues of the hinge region of RCL (purple). (B) Rotated view by 50°.

Structural transition to cleaved conformation

The crystal of cleaved Iripin-4 diffracted to the 2.00 Å resolution and belongs to $P12_11$ (4) space group, with unit-cell parameters a=65.70, b=138.41, c= 80.22 (Å), α= 90.0, β=107.7, γ= 90.0 (°). Crystal contains four molecules in the asymmetric unit with a solvent content of 40.65% and a Matthews coefficient of 2.07 $Å^3$ Da^{-1}. Detailed refinement statistics are listed in Tab. 1.

The structure of cleaved conformation consists of a serpin´s typical mixed α-β secondary structure with an N-terminal helical region and a C-terminal β-sheet fold. The overall structure is composed of 9 α-helices (A-I) and 3 β-sheets sequentially arranged in the order αA-β1-αB-αC-αD-β2-αE-β3-αF-β4-β5-β6-β7-β8-αG-αH-β9-β10-αI-β11-β12-β13-β14-β15   (Fig 7C). The additional β12-strand is a consequence of RCL insertion between the two β-

strands of β-sheet A and the formation of an additional strand called s4A, more precisely positioned between β11-strand (s5A) and β4-strand (s3A) (Fig. 7D). A chloride ion (Cl⁻) is present in the model at the same position for all chains, interacting with gate region residue Leu215. The gate region is responsible for RCL stabilization and prevents its premature insertion into β-sheet A (Fig. 8) [21].

Four molecules in the final model differ from each other in the number of modelled amino acids near the P1 site. Chain A and chain D contain 373 residues with missing Ser342, Leu343 and Val344 compared to chain B with missing Ser342 and Leu343 composed of 374 residues, and chain C with missing only Ser342 consisting of 375 residues (of 376-residues monomeric protein). The missing residues were observed as an absence of the electron density map after P1 residue Glu341 located at the end of the inserted RCL rather strand s4A (Fig. 5). Alignment of the chains using *PyMOL* [22] showed minor differences confirmed by r.m.s.d. calculation between chain C (3014 atoms) and other three chains: chain A r.m.s.d. = 0.181 (2351 atoms), chain B r.m.s.d. = 0.270 (2424 atoms) and chain D r.m.s.d. = 0.271 (2367 atoms).
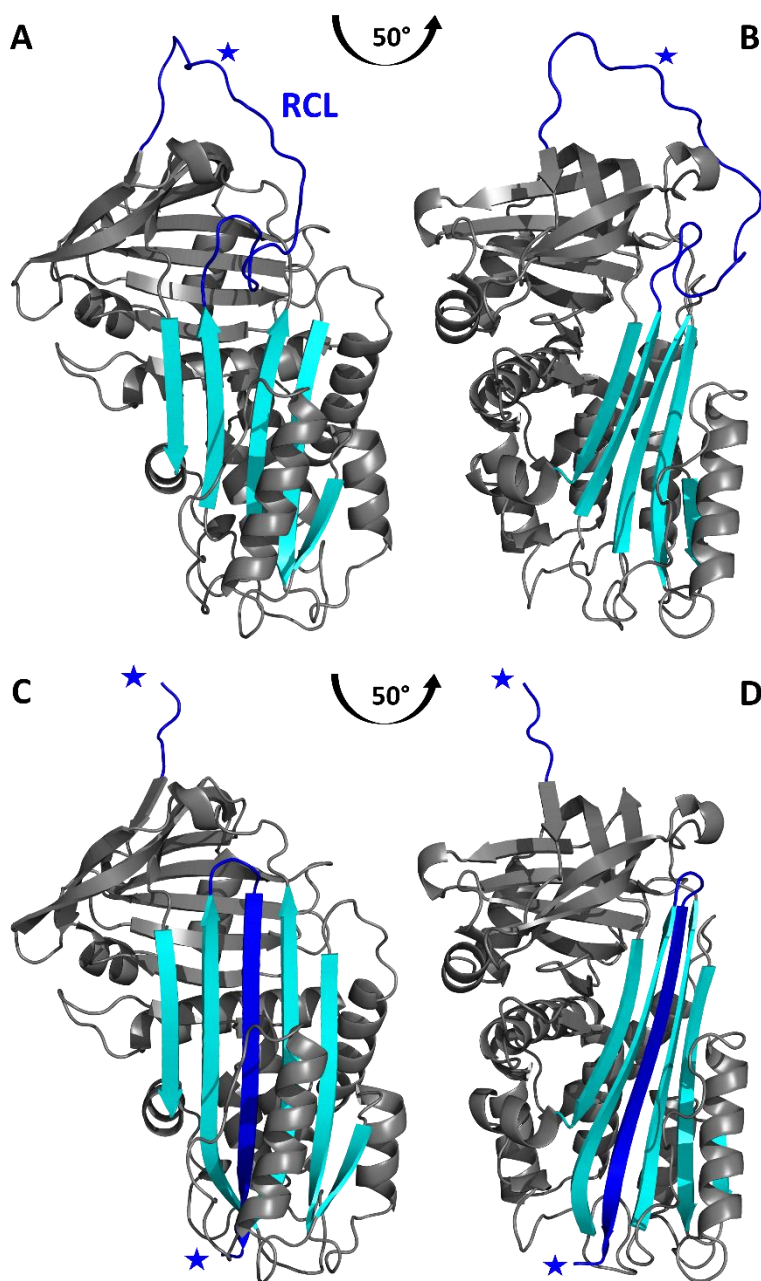
Figure. 7   Crystal structures of Iripin-4. A Native Iripin-4 with a highlighted position of RCL partially inserted into breach region of β-sheet A. B Side view of native conformation. C Cleaved Iripin-4 with displayed additional β-strand in β-sheet A as a result of RCL insertion. D Side view of cleaved conformation. Both crystal structures are displayed as a cartoon, β-sheet A is cyan and RCL is blue. The position of the P1 cleavage site is marked by a blue asterisk.

Table 1 X-ray data-collection and refinement statistics

Data collection	Native Iripin-4	Cleaved Iripin-4
X-ray diffraction source	BL14.2, BESSY II, Germany	BL14.2, BESSY II, Germany
Wavelength (Å)	0.9184	0.9184
Detector	PILATUS 6M	PILATUS 6M
Crystal-detector distance (mm)	175.66	373.61
Rotation range per image (°)	0.10	0.10
Total rotation range (°)	270.00	360.00
Exposure time per image (s)	0.10	0.10
Resolution range (Å)	44.61-2.30 (2.44-2.30)	46.42-2.00 (2.12-2.00)
Space group	$P3_121$ (152)	$P12_11$ (4)
Molecules in Asymmetric unit	1	4
Unit-cell dimensions: a, b, c (Å)	78.93 78.93 117.78	65.70 138.41 80.22
Unit-cell dimensions: α, β, γ (°)	90.00 90.00 120.00	90.00 107.70 90.00
Mosaicity (°)	0.105	0.128
Total No. of reflections	289932 (43011)	634544 (98469)
No. of unique reflections	19373 (3085)	91690 (14692)
Multiplicity	14.97	6.92
Average I/σ(I)	22.47 (4.30)	9.10 (2.32)
Completeness (%)	99.70 (99.70)	99.40 (98.60)
CC ½	99.90 (79.70)	99.70 (77.30)
R_{meas} (%)[a]	9.60 (62.40)	16.70 (85.30)
Overall B factor from Wilson plot (Å^2)	39.20	23.00

Refinement	Native Iripin-4	Cleaved Iripin-4
Resolution range (Å)	44.62-2.30	46.42-2.00
Space group	$P3_121$ (152)	$P12_11$ (4)
No. of reflections in working set	18404	89590
Final R_{value}[b] (%) / Final R_{free} value[c] (%)	20.30 / 25.40	19.30 / 25.30
Mean B value (Å)	39.80	28.00
No. of atoms in the asymmetric unit		
Protein	2962	11869
Ligand-Chloride ion		4
Ligand-Nickel ion	1	
Water	196	1276
Total	3168	13196
R.m.s. deviations		
Bonds (Å)	0.69	0.68
Angles (°)	0.81	0.81
Average B factors (Å^2) Overall	42.0	28.0
Ramachandran plot		
Most favoured (%)	95.99	98.17
Allowed (%)	4.01	1.83
Molprobity score	1.46 (99th percentile)*	1.27 (99th percentile)*
PDB entry	7ZBF	7ZAS

The data in parentheses refer to the highest-resolution shell.

[a]$R_{meas} = \dfrac{\sum_h \left(\frac{n_h}{n_h-1}\right)^{1/2} \sum |\langle I_h \rangle - I_{h,i}|}{\sum_h \sum_i I_{h,I}}$, where the average intensity $\langle I \rangle$ is taken over all symmetry equivalent measurements, and I_{hkl} is the measured intensity for any given reflection.

Structural differences in native and cleaved Iripin-4

Structural comparison of the native and cleaved conformations of Iripin-4 reveals that β-strands of β-sheet A together with α-helix D, α-helix E, α-helix F, and α-helix I shift to make space for the inserted RCL (Fig. 8). Considering the fact that cleaved conformation of serpin is structurally mostly related to the structure of a covalent complex between serpin and protease, it can be stated that these structural changes will be the same or at least similar to changes necessary for covalent complex formation. These observations also show how a large conformational change is needed for the transition from native to cleaved state and for protease inhibition as well. Except for structural changes in shutter and hinge regions, there is a noticeable shift of the gate region, the β-sheet C loop part between s3C and s4C strand (Fig. 8).
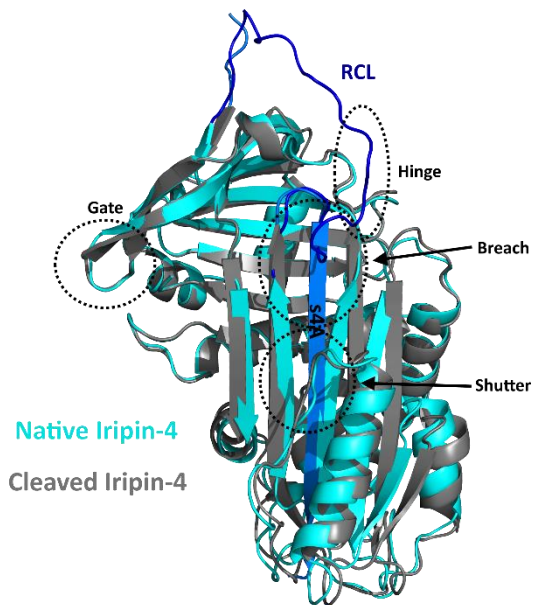

Figure. 8 Structural superposition of Iripin-4 conformations. The native Iripin-4 (cyan) with exposed RCL (blue) is aligned with cleaved Iripin-4 (grey) containing new additional β-strand (marine) marked as s4A. The three regions responsible for insertion of RCL into β-sheet A are marked in circles [99] The hinge region is marked with an oval.

DISCUSSION

Iripin-4 shows no inhibition by a typical mechanism of enzyme inhibition through forming a covalent complex with tested target protease. The most prevalent residue at P1 site differs from Arg residue occurring in other *I. ricinus* serpins with known structures, namely Iripin-5, Iripin-8 and Iripin-3 (7B2T-Arg342, 7PMU-Arg364 and 7AHP-Arg357). Other known *I. ricinus* serpins has Glu341 (Iripin-4), Tyr341 (IRS-2, 3NDA) or Met340 (IRIS) at the P1 position (Fig. 9) [20,23–26]. The first serpin characterized in *I. ricinus* saliva is the immunosuppressant IRIS that inhibits elastase-like proteases. IRIS prevents blood clotting by inhibiting factor Xa and thrombin, delays fibrinolysis by inhibiting elastase and tPA and inhibits the secretion of pro-inflammatory cytokines such as IFNγ and IL-6. Some of the mentioned activities are the result of several IRIS exosites in the αA and αD area [26–28]. Another well described serpin from the same tick species is IRS-2 that targets chymotrypsin-like proteases. IRS-2 inhibits pro-inflammatory proteinases cathepsin G and mast cell chymase. By inhibiting thrombin, IRS-2 blocked platelet aggregation and was shown to act as an immunomodulator by inhibiting the IL-6 production via STAT3 signaling alteration [23,29]. Another *I. ricinus* serpin Iripin-3 inhibits trypsin-like proteases with highest inhibition rates against kallikrein and matriptase and was shown to suppress the inflammation and wound healing. Iripin-3 also suppresses the proliferation and differentiation of CD4[+] T cells into Th1 pro-inflammatory subpopulation hence regulates adaptive and acquired immune responses [24]. Iripin-5, the most abundant salivary serpin of *I. ricinus*, displayed an anti-inflammatory role, as it inhibited neutrophil migration, complement activity and suppressed the ability of macrophages to produce NO [25]. The Iripin-8 is an inhibitor of multiple coagulation proteases and it strongly inhibited intrinsic pathway of coagulation. Moreover, it inhibited complement related lysis of erythrocytes. Thus, Iripin-8 functions as an anticoagulant and an inhibitor of complement system [20] The presence of partially inserted RCL in Iripin-4 could lead to observed low protease inhibition. This was reported for antithrombin that had in this state the P1 site not accessible for target protease. However, it is likely that Iripin-4 targets different proteases than we tested.

Glutamate at P1 site could suggest granzyme B as target. Since we have no functional data to date, we cannot rule out exosites mediated activity, similarly to IRIS. The function of Iripin-4 should be addressed in further studies.
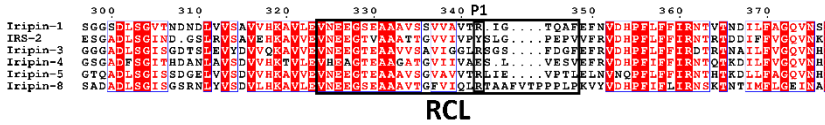

Figure. 9   Comparison of RCL region of *I. ricinus* serpins. The RCL is in the black box with prompted P1 recognition site in the smaller black box.

Figure. 10   During the years of crystallographic study of serpins, many structural conformations and states of serpins were identified and the importance of some structural regions such as breach, shutter, or hinge regions was described [11].

The only serpin conformation with inhibitory function is native conformation that is metastable and stressed [30]. This type of conformation was solved for Iripin-4 at 2.3 Å resolution and it represents the active serpin ready to "trap" the target protease. The breach region of native Iripin-4 differs from breach regions of other solved native *I. ricinus* serpin structures (Fig. 10). This is mainly because Iripin-1 (unpublished data) and Iripin-8 (PDB ID: 7PMU) [20] native structures were solved at a fully stressed state in contrast with Iripin-4 solved in a partially relaxed state (Fig. 8). The breach region, which is the upper part of s3A and s5A β-strands of β-sheet A, is opened to the sides especially by s3A movement to the side to make space for RCL partial insertion. On the other hand, Iripin-1 and Iripin-8 do not have this gap in the breach region (Fig. 10). The shutter region (middle part of β-sheet A) of all three serpins is almost identical. Another region that is part of the RCL (P15-P9), so-called hinge region, differs structurally among these three serpins because this region is responsible for RCL mobility and its insertion into β-sheet A [16,31]. Hinge regions differ also sequentially; while Iripin-4 has threonine at P14 and glycine at P10 site, Iripin-1 and Iripin-8 have serine at P14 and alanine at P10 site. Nevertheless, this does not suggest any functional variation since it was shown that uncharged residues are suitable for good loop insertion but do not affect the protease selectivity [32]. It was reported that the use of cofactors during crystallization can lead to a fully exposed RCL state by stabilizing the hinge region [33,34] but it was not observed for the mentioned proteins. The most

remarkable part of native *I. ricinus* serpins is RCL length, specifically the length of Iripin-8 RCL that is longer by 5 residues. Another difference is that Iripin-4 has glutamate at its P1 recognition site, while there is arginine in Iripin-1 and Iripin-8. The Glu residues located at P4' and P5' were reported as exosites increasing the serpin inhibition rate [35]. The r.m.s.d. between molecules was calculated by *PyMOL* [22] and results showed that the r.m.s.d. of Iripin-4 alignment with the Iripin-1 model was 0.894 (2008 atoms) and with Iripin-8 was 1.143 (1943 atoms). Iripin-8 inhibits serine proteases involved in coagulation and complement function. These proteases include thrombin, factor Xa, factor XIa, factor VIIa, factor IXa, kallikrein and plasmin [20] In contrast, Iripin-4 shows no noticeable inhibition of tested proteases.

Figure. 11 *I. ricinus* native serpin conformations superposition and alignment. A Superposition of the Iripin-4 (PDB ID: 7ZBF), Iripin-1 (unpublished data) and Iripin-8 (PDB ID: 7PMU) crystal structures are displayed as a cartoon. The structure models are colour differentiated as follows, Iripin-4 is blue, Iripin-1 is firebrick and

Iripin-8 is green. The structure of Iripin-1 was modified because of the absence of residues in its RCL that are coloured salmon for emphasis. B Close look on RCLs residues displayed as sticks and colour difference is same as previously described. C Structure-based sequence alignment of *I. ricinus* native serpins. The secondary structure elements are depicted as arrows for β-strands and spirals for α-helices. The RCL is stressed in a black box and P1 and P1' residues are marked according to Schechter and Berger nomenclature [36].

The breach, shutter, hinge and gate regions mentioned before are involved in the incorporation of RCL specifically hinge region of RCL into β-sheet A by coordinated movements of this complex process [30,37]. This process is called S to R transition, which means a transition from a native stressed state to a more relaxed cleaved state with more thermal stability [38]. The comparison of Iripin-4 native and cleaved structures (Fig. 8) shows significant movement in breach and shutter regions. The breach region that is the starting point of the inhibitory mechanism was slightly opened compared to other native serpins (Fig. 10) and thus is more flexible than the shutter region responsible for controlling or preventing undesired conformational change. The shutter region has over 85% of conserved residues and must open for incorporation of RCL that helps to stabilize this transition process [39]. The polar contacts between s3A and s5A, more precisely 316Glu-167Asn, 318Lys-167Asn and 319Thr-169Ile, were disrupted by RCL insertion when compared to a cleaved structure with the additional β-s4A strand. For the correct inhibition of RCL with protease the transition must pass off the gate region [37] and resulted in serpin cleavage instead of the formation of a covalent complex with protease.

The comparison of the cleaved *I. ricinus* structures namely IRS-2 (PDB ID: 3NDA) [40], Iripin-3 (PDB ID: 7AHP) [24] and Iripin-5 (PDB ID: 7B2T) [26] with Iripin-4 cleaved structure showed almost identical structural fold (Fig. 9). The results of alignment calculations showed that the r.m.s.d of Iripin-4 alignment with IRS-2 was 0.750 (2212 atoms), with Iripin-3 was 0.525 (2090 atoms) and with Iripin-5 was 0.690 (2132 atoms). The differences were mainly in loop regions and most notably in the αD helix.
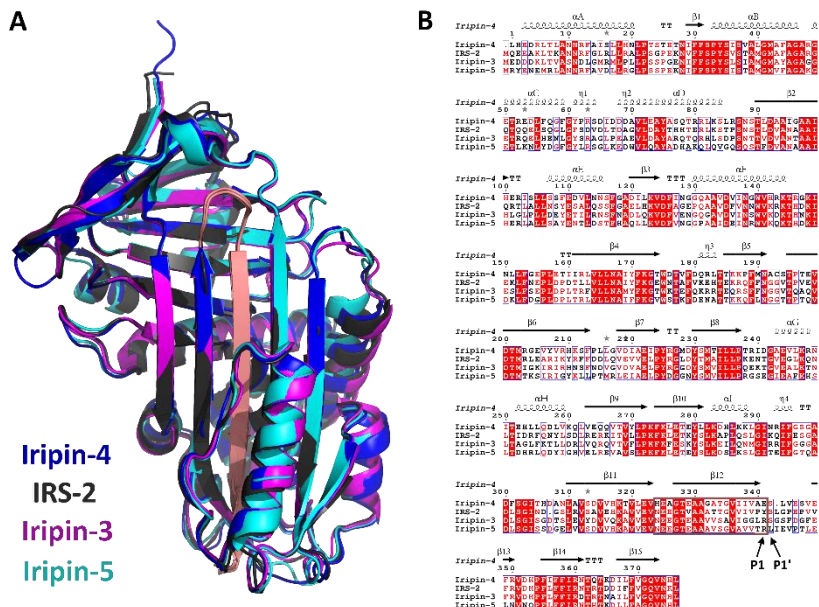
Figure. 12   *I. ricinus* cleaved serpin conformations superposition and alignment. A Superposition of the Iripin-4 (PDB ID: 7ZAS), IRS-2 (PDB ID: 3NDA), Iripin-3 (PDB ID: 7AHP) and Iripin-5 (PDB ID: 7B2T) crystal structures displayed as a cartoon. The structure models are colour differentiated as follows, Iripin-4 is blue, IRS-2 is grey, Iripin-3 is magenta, and Iripin-5 is cyan. The RCL of all serpins is colour marked as salmon. B Structure-based sequence alignment of *I. ricinus* cleaved serpins. The secondary structure elements are marked as arrows for β-strands and spirals for α-helices. The inserted RCL is stressed in the black box and P1 and P1' residues are marked according to Schechter and Berger nomenclature [36].

MATERIALS AND METHODS

Protein Cloning, Expression and Purification

The full-length Iripin-4 sequence was cloned into a Champion™ pET-SUMO expression vector (Life Technologies, Carlsbad, CA, USA) and transformed into Rosetta 2 (DE3) pLysS competent cells (Novagen, Merck Life Science, Darmstadt, Germany). The autoinduction TB medium, (100 µg*ml⁻¹ kanamycin, chloramphenicol 34 µg*ml⁻¹) was inoculated with an overnight culture (10ml/l), incubated for 24 h at 25°C. Bacteria cells were then harvested and disrupted by cell disruptor. The soluble SUMO-tagged Iripin-4 was purified by the HisTrap FF column (GE Healthcare, Chicago, IL, USA) and eluted with 200 mM imidazole. After the first purification, His and SUMO tags were cleaved using a SUMO protease (1:100 w/w) overnight at laboratory

temperature. Samples were then loaded onto the HisTrap column again to separate tags from the native serpin. This step was followed by Ion Exchange Chromatography using the HiTrap Q HP column (GE Healthcare) that provided sufficient purity of the protein. The final concentration of protein was 1.17 mg/ml in 20 mM Tris pH 7.4, 20 mM NaCl and protein was stored at -80°C.

Analysis of the complex formation of Iripin-4 with proteases

Selected serine proteases (Trypsin, Thrombin, fXa, APC, fXIa, Plasmin, Elastase) were incubated with an equimolar amount of Iripin-4 for 1h at RT (final concentration 1 μM). The formation of serpin-protease complexes was analysed in reducing SDS-PAGE with the use of 12% gels, followed by Coomassie staining.

Protein Crystallization, X-Ray Data Collection and Processing

The suitable protein concentration for crystallization screening was determined by Pre-Crystallization Test (PCT^{TM}, Hampton Research, Aliso Viejo, California, USA). Crystallization experiments were performed using the sitting-drop vapor diffusion technique in 96-well crystallization plates Swissci MRC 2-drop (Molecular Dimension, Newmarket, Suffolk, UK) or Swissci MRC 3-drop (Molecular Dimension, Newmarket, Suffolk, UK). Initial screening for crystallization conditions and further crystallization experiments were carried out using the OryxNano crystallization robot (Douglas Instruments, Hungerford, UK). Crystallization condition screening was done using the commercially available crystallization kits (Crystal Screen, PegRX, PEGIon, Index; JBScreen JCSG++, Jena Bioscience). The two protein-to-well solution ratios 1:0.5 and 1:1 (0.4μl: 0.2μl and 0.4μl: 0.4μl) were applied for protein concentrations 1.17 and 2.3 mg-ml^{-1}, and drop solution was equilibrated against 50 μl reservoir solution at 4°C. Applicable crystallization conditions were identified in the conditions: JBScreen JCSG++, no. H8 (25%(w/v) PEG 3350, 0.1 M Bis-Tris pH 5.5, 0.2 M sodium chloride) and JBScreen JCSG++, no. H10 (25%(w/v) PEG 3350, 0.1 M Bis-Tris pH 5.5, 0.2 M ammonium acetate). For these conditions, a Simple 2D

gradient was used for crystal quality optimization in a larger volume (0.5μl:0.5μl).

For data collection, crystals of both native and cleaved Iripin-4 were mounted directly from the crystallization drop and flash-cooled in liquid nitrogen without any additional cryoprotection before measurement. Collection of X-ray diffraction data was performed at the BESSY II electron storage ring on the beamline BL14.2 operated by the Helmholtz-Zentrum Berlin [41]. Data for all crystals were processed using the *XDS* program [42] with the *XDSAPP* graphical user interface [43]. The crystal solvent content was analysed by *MATTHEWS_COEF* program from the *CCP4* suite [44] Data-collection statistics for both data sets are summarized in Table 1.

Structure Determination and Refinement

The structures of serpin Iripin-4 were solved by using the *CCP4* suite [44] namely: *MOLREP* [45] molecular-replacement method using the *MrBUMP:* a molecular-replacement pipeline [46] for cleaved conformation and structure of IRS-2 serpin (PDB entry 3NDA; [23]) with 56.5% sequence identity generated from the automated model search by *BALBES:* a molecular-replacement pipeline [47] for the native conformation. The resolution cut-off for native Iripin-4 was performed to 2.3 Å resolution because of the presence of two ice rings at 2.25 Å and 1.95 Å resolution. The structures were refined with the program *REFMAC5* [48] from the *CCP*4 package [44] and manually rebuilt in *Coot* [49] from evaluation of the electron-density peaks. The improvement during refinement was monitored by structure validation throughout the refinement process. Water molecules were added to the model using the *REFMAC5* interface. Accepted solvent molecules had tolerable hydrogen-bonding geometry contacts of 2.5–3.5 Å with protein atoms or with existing solvent. At this point, residues, and water molecules with two possible conformations were included and their alternative conformations were added for further refinement. In the last steps of refinement, Nickel^{2+} ion (native structure) and four Chloride ions (cleaved structure) were built into the appropriate ($2F_o -$ F_c) and ($F_o - F_c$) electron-density maps using coordinates from the ligand data

bank in *Coot* [49]. *MolProbity* server [50] and *wwPDB OneDep* validation system [51] were used for final model qualitative validation. All figures were created using *PyMOL* [22] or *CCP4* Molecular Graphics [52]. For the visualisation of the cleaved conformation chain C was used. Refinement statistics are summarized in Table 1.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA ACCESSIBILITY

The atomic coordinates of Iripin-4 native and cleaved conformations are available in the Protein Data Bank (PDB) database under accession numbers: 7ZBF and 7ZAS respectively.

AUTHOR CONTRIBUTIONS

IKS and JC designed the project. JK performed the cloning, expression and purification and determined the antiprotease selectivity. BK, PH, PG, MK and TP carried out crystallization experiments and performed X-ray diffraction analysis. BK analysed the crystallographic data, solved the structures, and drafted the manuscript.

REFERENCES

1. Rizzoli A, Hauffe HC, Carpi G, Vourc'h GI, Neteler M & Rosà R (2011) Lyme borreliosis in Europe. *Eurosurveillance* **16**.

2. Madison-Antenucci S, Kramer LD, Gebhardt LL & Kauffman E (2020) Emerging Tick-Borne Diseases. *Clin Microbiol Rev* **33**, e00083-18.

3. Rosario-Cruz R, Almazan C, Miller RJ, Dominguez-Garcia DI, Hernandez-Ortiz R & de La Fuente J (2009) Genetic basis and impact of tick acaricide resistance. *Frontiers in Bioscience* **14**.

4. Boulanger N, Boyer P, Talagrand-Reboul E & Hansmann Y (2019) Ticks and tick-borne diseases. *Médecine et Maladies Infectieuses* **49**, 87–97.

5. de la Fuente J, Kopáček P, Lew-Tabor A & Maritz-Olivier C (2016) Strategies for new and improved vaccines against ticks and tick-borne diseases. *Parasite Immunology* **38**.

6. Prevot PP, Couvreur B, Denis V, Brossard M, Vanhamme L & Godfroid E (2007) Protective immunity against Ixodes ricinus induced by a salivary serpin. *Vaccine* **25**, 3284–3292.

7. Meekins DA, Kanost MR & Michel K (2017) Serpins in arthropod biology. *Semin Cell Dev Biol* **62**, 105–119.

8. Silverman GA, Whisstock JC, Bottomley SP, Huntington JA, Kaiserman D, Luke CJ, Pak SC, Reichhart JM & Bird PI (2010) Serpins flex their muscle: I. Putting the clamps on proteolysis in diverse biological systems. *Journal of Biological Chemistry* **285**, 24299–24305.

9. Spence MA, Mortimer MD, Buckle AM, Minh BQ & Jackson CJ (2021) A Comprehensive Phylogenetic Analysis of the Serpin Superfamily. *Molecular Biology and Evolution* **38**, 2915–2929.

10. Whisstock J, Skinner R & Lesk AM (1998) An atlas of serpin conformations. *Trends in Biochemical Sciences* **23**, 63–67.

11. Dunstone MA & Whisstock JC (2011) Crystallography of serpins and serpin complexes. In *Methods in Enzymology* pp. 63–87. Academic Press Inc.

12. Komiyama T, Ray CA, Pickup DJ, Howard AD, Thornberry NA, Peterson EP & Salvesen G (1994) Inhibition of interleukin-1 beta converting enzyme by the cowpox virus serpin CrmA. An example of cross-class inhibition. *Journal of Biological Chemistry* **269**, 19331–19337.

13. Vercammen D, Belenghi B, van de Cotte B, Beunens T, Gavigan J-A, de Rycke R, Brackenier A, Inzé D, Harris JL & van Breusegem F (2006) Serpin1 of Arabidopsis thaliana is a Suicide Inhibitor for Metacaspase 9. *Journal of Molecular Biology* **364**, 625–636.

14. Kaiserman D, Whisstock JC & Bird PI (2006) Mechanisms of serpin dysfunction in disease. *Expert Reviews in Molecular Medicine* **8**.

15. Gettins PGW (2002) Serpin structure, mechanism, and function. *Chemical Reviews* **102**, 4751–4803.

16. Nieman MT, Lawrence DA, Renne T, De PA, Mourão S, de Maat S, Sanrattana W & Maas C (2019) SERPINs-From Trap to Treatment. *Frontiers in Medicine | www.frontiersin.org* **6**, 25.

17. Huntington JA (2011) Serpin structure, function and dysfunction. *Journal of Thrombosis and Haemostasis* **9**, 26–34.

18. Huntington JA & Carrell RW (2001) The Serpins: Nature's Molecular Mousetraps. *Science Progress* **84**, 125–136.

19. Whisstock JC, Silverman GA, Bird PI, Bottomley SP, Kaiserman D, Luke CJ, Pak SC, Reichhart JM & Huntington JA (2010) Serpins flex their muscle: II. Structural insights into target peptidase recognition, polymerization, and transport functions. *Journal of Biological Chemistry* **285**, 24307–24312.

20. Kotál J, Polderdijk SGI, Langhansová H, Ederová M, Martins LA, Beránková Z, Chlastáková A, Hajdušek O, Kotsyfakis M, Huntington JA & Chmelař J (2021)

Ixodes ricinus salivary serpin iripin-8 inhibits the intrinsic pathway of coagulation and complement. *International Journal of Molecular Sciences* **22**.

21. Irving JA, Miranda E, Haq I, Perez J, Kotov VR, Faull S v, Motamedi-Shad N & Lomas DA (2015) An antibody raised against a pathogenic serpin variant induces mutant-like behaviour in the wild-type protein. *Biochem J* **468**, 99–108.

22. Khan MS, Singh P, Azhar A, Naseem A, Rashid Q, Kabir MA & Jairajpuri MA (2011) Access to. *Research Journal of Amino Acids* **2011**.

23. Chmelar J, Oliveira CJ, Rezacova P, Francischetti IMB, Kovarova Z, Pejler G, Kopacek P, Ribeiro JMC, Mares M, Kopecky J & Kotsyfakis M (2011) A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood* **117**, 736–744.

24. Chlastáková A, Kotál J, Beránková Z, Kaščáková B, Martins LA, Langhansová H, Prudnikova T, Ederová M, Kutá Smatanová I, Kotsyfakis M & Chmelař J (2021) Iripin-3, a New Salivary Protein Isolated From Ixodes ricinus Ticks, Displays Immunomodulatory and Anti-Hemostatic Properties In Vitro. *Frontiers in Immunology* **12**.

25. Kascakova B, Kotal J, Martins LA, Berankova Z, Langhansova H, Calvo E, Crossley JA, Havlickova P, Dycka F, Prudnikova T, Kuty M, Kotsyfakis M, Chmelar J & Smatanova IK (2021) Structural and biochemical characterization of the novel serpin Iripin-5 from Ixodes ricinus. *Acta Crystallographica Section D: Structural Biology* **77**.

26. Prevot PP, Adam B, Boudjeltia KZ, Brossard M, Lins L, Cauchie P, Brasseur R, Vanhaeverbeek M, Vanhamme L & Godfroid E (2006) Anti-hemostatic effects of a serpin from the saliva of the tick Ixodes ricinus. *Journal of Biological Chemistry* **281**, 26361–26369.

27. Leboulle G, Rochez C, Louahed J, Rutti B, Brossard M, Bollen A & Godfroid E (2002) Isolation of Ixodes ricinus salivary gland mRNA encoding factors induced during blood feeding. *American Journal of Tropical Medicine and Hygiene* **66**.

28. Prevot P-P, Beschin A, Lins L, rô me Beaufays J, lie Grosjean A, Bruys L, Adam B, Brossard M, Brasseur R, Zouaoui Boudjeltia K, Vanhamme L & Godfroid E Exosites mediate the anti-inflammatory effects of a multifunctional serpin from the saliva of the tick Ixodes ricinus. .

29. Páleníková J, Lieskovská J, Langhansová H, Kotsyfakis M, Chmelař J & Kopeckỳ J (2015) Ixodes ricinus salivary serpin IRS-2 affects Th17 differentiation via inhibition of the interleukin-6/STAT-3 signaling pathway. *Infection and Immunity* **83**.

30. Pearce MC, Pike RN, Lesk AM & Bottomley SP (2007) Serpin Conformations. *Molecular and Cellular Aspects of the Serpinopathies and Disorders in Serpin Activity*, 35–66.

31. Hopkins PCR, Carrell RW & Stone SR (1993) Effects of Mutations in the Hinge Region of Serpins. *Biochemistry* **32**.

32. Lawrence DA, Olson ST, Muhammad S, Day DE, Kvassman JO, Ginsburg D & Shore JD (2000) Partitioning of serpin-proteinase reactions between stable inhibition and substrate cleavage is regulated by the rate of serpin reactive center loop insertion into β-sheet A. *Journal of Biological Chemistry* **275**.

33. Whisstock JC & Bottomley SP (2006) Molecular gymnastics: serpin structure, folding and misfolding. *Current Opinion in Structural Biology* **16**, 761–768.

34. Li W, Johnson DJD, Esmon CT & Huntington JA (2004) Structure of the antithrombin–thrombin–heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nature Structural & Molecular Biology* **11**, 857–862.

35. Ibarra CA, Blouse GE, Christian TD & Shore JD (2004) The Contribution of the Exosite Residues of Plasminogen Activator Inhibitor-1 to Proteinase Inhibition. *Journal of Biological Chemistry* **279**.

36. Schechter I & Berger A (1967) On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications* **27**.

37. Khan MS, Singh P, Azhar A, Naseem A, Rashid Q, Kabir MA & Jairajpuri MA (2011) Serpin Inhibition Mechanism: A Delicate Balance between Native Metastable State and Polymerization. *Journal of Amino Acids* **2011**, 1–10.

38. Whisstock JC, Skinner R, Carrell RW & Lesk AM (2000) Conformational changes in serpins: I. The native and cleaved conformations of α1-antitrypsin. *Journal of Molecular Biology* **296**.

39. Irving JA, Pike RN, Lesk AM & Whisstock JC (2000) Phylogeny of the serpin superfamily: Implications of patterns of amino acid conservation for structure and function. *Genome Research* **10**, 1845–1864.

40. Kováóvá Z, Chmelař J, Šanda M, Brynda J, Mareš M & Řezáčová P (2010) Crystallization and diffraction analysis of the serpin IRS-2 from the hard tick Ixodes ricinus. In *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* pp. 1453–1457.

41. Mueller U, Förster R, Hellmig M, Huschmann FU, Kastner A, Malecki P, Pühringer S, Röwer M, Sparta K, Steffien M, Ühlein M, Wilk P & Weiss MS (2015) The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *The European Physical Journal Plus* **130**.

42. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.

43. Sparta KM, Krug M, Heinemann U, Mueller U & Weiss MS (2016) XDSAPP2.0. *Journal of Applied Crystallography* **49**, 1085–1092.

44. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A & Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**.

45. Vagin A & Teplyakov A (1997) *MOLREP*: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography* **30**, 1022–1025.

46. Krissinel E, Uski V, Lebedev A, Winn M & Ballard C (2018) Distributed computing for macromolecular crystallography. *Acta Crystallogr D Struct Biol* **74**, 143–151.

47. Long F, Vagin AA, Young P & Murshudov GN (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125–132.

48. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.

49. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.

50. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2012) *MolProbity*: all-atom structure validation for macromolecular crystallography. *International Tables for Crystallography*, 694–701.

51. Young JY, Westbrook JD, Feng Z, Sala R, Peisach E, Oldfield TJ, Sen S, Gutmanas A, Armstrong DR, Berrisford JM, Chen L, Chen M, di Costanzo L, Dimitropoulos D, Gao G, Ghosh S, Gore S, Guranovic V, Hendrickx PMS, Hudson BP, Igarashi R, Ikegawa Y, Kobayashi N, Lawson CL, Liang Y, Mading S, Mak L, Mir MS, Mukhopadhyay A, Patwardhan A, Persikova I, Rinaldi L, Sanz-Garcia E, Sekharan MR, Shao C, Swaminathan GJ, Tan L, Ulrich EL, van Ginkel G, Yamashita R, Yang H, Zhuravleva MA, Quesada M, Kleywegt GJ, Berman HM, Markley JL, Nakamura H, Velankar S & Burley SK (2017) OneDep:

Unified wwPDB System for Deposition, Biocuration, and Validation of Macromolecular Structures in the PDB Archive. *Structure* **25**, 536–545.

52. McNicholas S, Potterton E, Wilson KS & Noble MEM (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr D Biol Crystallogr* **67**, 386–394.

3.4 Crystal structure of *Ixodes ricinus* serpin Iripin-1

This chapter is based on Paper IV.:

Kaščáková, B., Havlíčková, P., Prudnikova, T., Chmelař, J & Kutá Smatanová, I. (2022). Crystal structure of *Ixodes ricinus* serpin Iripin-1. *Manuscript in preparation.*

INTRODUCTION

Ixodes ricinus ticks developed many strategies to be able to feed on their host. One protein group owing to facilitates tick feeding and thus reproduction is serpins which are found all over kingdoms of life [1]. Serpins are interesting proteins due to their drastic fold change necessary for their functional regulation as serine protease inhibitors [2]. Where other protein retains the initial native structural fold as their most thermodynamically stable conformation, serpins shift to their most stable state after their "cleavage" [3,4].

The native, metastable, active conformation of serpins has a conserved secondary structure fold composed of three β-sheets (A, B and C), ∼ 9 α-helices and a flexible, exposed reactive center loop (RCL) [5,6]. The RCL is a loop near the protein C-terminus positioned on the top of the serpin backbone functioning as the "substrate" for target protease. The β-sheets A consist of 5 β-strands important during the inhibition mechanism and carrying two crucial structural regions: breach (top) and shutter (bottom) responsible for RCL insertion into β-sheets A during protease transition to the serpin another site [7].

The serpin RCL cleavage is crucial for their inhibitory mechanism which is also called the suicide mechanism since serpins are at the end of the mechanism inactivated together with the target enzyme [8,9]. The inhibition starts with the enzyme recognizing serpin RCL as substrate and then enzyme catalytic triad intrude the enzyme cleavage site (P1-P1´, [10]) located on RCL, resulting in the formation of tetrahedral intermediate [8,11]. The structural conformation of this state is called the Michaelis-Menten complex and from this moment there is a possibility of two results and thus: (I.) the formation of the covalently linked complex where both protease and serpin are inactivated, or (II.) protease hydrolyzes the tetrahedral intermediate and is released before inactivation that results with only serpin inactivation. The result is determined

by the speed of the reaction thus the faster reaction the faster the enzyme transition and at the end the distortion of the active site is done well [11–13].

MATERIAL AND METHODS

Crystallization

Pre-Crystallization Test (PCT^{TM}, Hampton Research, Aliso Viejo, California, USA) was applied to determine the suitable protein concentration for crystallization screening. Crystallization experiments were realized using the sitting-drop vapor diffusion technique in 96-well crystallization plates Swissci MRC 2-drop (Molecular Dimension, Newmarket, Suffolk, UK). Crystallization experiments were performed using the OryxNano crystallization robot (Douglas Instruments, Hungerford, UK). Crystallization conditions screening was performed using the commercially available crystallization kits (PEGIon, Morpheus II, SG1). The protein concentration was determined to 0.93 mg·ml^{-1}, two protein-to-well solution ratios 1:0.5 and 1:1 (1μl: 0.5μl and 1 μl: 1 μl) were applied and drop solution was equilibrated against 50 μl reservoir solution at 4°C.

X-ray data collection, structure determination and refinement

For data collection, crystals were flash-frozen in liquid nitrogen without any additional cryoprotection before measurement. The collection of X-ray diffraction data was performed at the BESSY II electron storage ring on the beamline BL14.2 operated by the Helmholtz-Zentrum Berlin [14]. Data were processed using the *XDS* program [15] with the *XDSAPP* graphical user interface [16].

The structure of serpin Iripin-1 was solved by using the *CCP4* suite [17] namely: *POINTLESS* and *SCALA* [18,19] for scaling and *MOLREP* [20] molecular replacement method using the automated model search by *BALBES: a molecular-replacement pipeline* [21]. The structure was then refined with the program *REFMAC*5 from the *CCP*4 package [22] and manually rebuilt in *Coot* [23]. *MolProbity* server [24] and *wwPDB OneDep* validation system [25] were used for the final model qualitative validation. All figures were made using *PyMOL* [26] and atomic coordinates have been deposited in the Protein Data Bank.

RESULTS AND DISCUSSIONS

Crystal structure of Iripin-1

Appropriate crystallization condition was identified in the MemGold™ screen MD1-41 (Molecular Dimension, Newmarket, Suffolk, UK), namely condition No. 1-35 (C11) (0.02 M Nickel (II) sulfate hexahydrate pH 7.0, 0.01 M HEPES, 33 % v/v Jeffamine® M-600). The crystals of good diffraction quality were grown in conditions with Iripin-1 (0.93 mg-ml^{-1}) in a 1:1 ratio (Fig. 1).
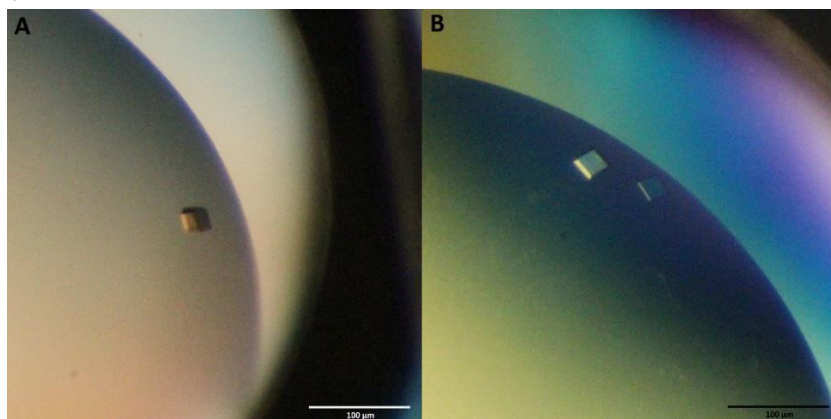

Figure. 1 The crystals of Iripin-1. (A) The crystal was grown in 0.02 M Nickel (II) sulfate hexahydrate pH 7.0, 0.01 M HEPES, 33 % v/v Jeffamine® M-600 in a 1:1 ratio with protein. (B) The crystal growth in the same conditions but in a 0.5:1 ratio with protein.

The structure of native Iripin-1 was solved at the resolution of 2.10 Å. The crystal had a monoclinic P12$_1$1 space group and the asymmetric unit of the crystal contained two molecules of Iripin-1 with a calculated solvent content of 50.82% (Matthew's coefficient). The *BALBES* – automated Molecular Replacement pipeline was used to solve the structure and generated the model by using the crystal structure of human squamous cell carcinoma antigen 1 (2ZV6) with 35.54% sequence identity (46.15% similarity) as starting model.

The structure consists of a mixed α-β secondary structure with an N-terminal helical region and a C-terminal β-sheet fold that is a typical serpin secondary structure fold. Also, the structure of Iripin-1 possesses a typical structure of native serpin where RCL is fully exposed on top of the protein "body" and β-sheet A consisting of 5 β-strands [27] (Fig. 2). The overall

structure is composed of 10 α-helices and 3 β-sheets sequentially arranged in the order α1-β1-α2-α3-α4-α5-β2-α6-β3-α7-β4-β5-β6-β7-β8-β9-α8-α9-β10-β11-α10-β12-β13-β14-β15. The final model contains 359 residues in chain A and 351 residues in chain B (of 377-residues monomeric protein). The missing residues 180Phe – 184Arg (both chains) and 324Val – 336Ser (Chain A) / 326Glu – 346Gln (Chain B) were observed as an absence of the electron density map in these two loop regions most probably due to the high flexibility of these loops (Fig. 2). The RCL part that was not modelled is called the hinge region which is the N-terminal part of RCL before the cleavage site (P15-P9) is responsible for RCL mobility and its insertion into β-sheet A [28]. This explains why this highly flexible RCL part was not modelled.

The atomic coordinates have been deposited in the Protein Data Bank under accession code 7QTZ. Data collection and refinement statistics are summarized in Table 1.

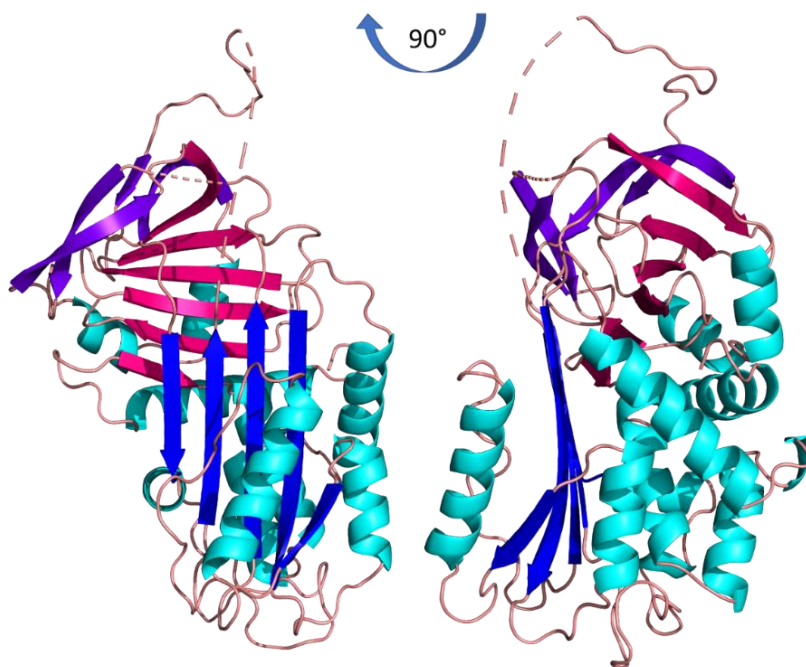**Figure. 2**   The overall structure of the crystal structure of native Iripin-1. Cartoon representation of protein secondary structure; N-terminus and C-terminus of Iripin-1 are labelled (N, C); α-helices are coloured by cyan, β-sheets: sheet A (blue), sheet B (magenta) and sheet C (purple) and loops are salmon. The loop regions with missing residues due to poor to no electron density map are depicted as dashes.

Table 1 Data collection and refinement statistics

Data collection

X-ray diffraction source	BL14.2, BESSY II, Germany
Wavelength (Å)	0.9184
Detector	PILATUS 6M
Crystal-detector distance (mm)	346.829
Rotation range per image (°)	0.10
Total rotation range (°)	360.0
Exposure time per image (s)	0.10
Resolution range (Å)	50.00-2.10 (2.33-2.10)
Space group	$P12_11$ (4)
Unit-cell dimensions: a, b, c (Å)	48.82 91.00 95.83
Unit-cell dimensions: α, β, γ (°)	90.00 97.53 90.00
Mosaicity (°)	0.193
Total No. of reflections	331018 (52480)
No. of unique reflections	48350 (7680)
Average I/σ(I)	11.12 (1.62)
Completeness (%)	99.1 (97.7)
CC ½	99.8 (68.9)
R_{meas} (%)[a]	116.8 (14.3)
Overall B factor from Wilson plot (Å²)	31.0

Refinement

Resolution range (Å)	48.44-2.10
No. of reflections in working set	45891
Final R value (%)[b] / Final R_{free} value (%)[c]	0.190/0.251
Mean B value (Å)	38.591
No. of atoms in the asymmetric unit	
Protein	5669
Ligand-Magnesium ion	1
Water	502
Total	6191
R.m.s. deviations	
Bonds (Å)	0.007
Angles (°)	1.417
Average B factors (Å²) Overall	41.0
Ramachandran plot	
Most favoured (%)	97.42
Allowed (%)	0.14
Molprobity score	1.37 (99th percentile) *
PDB code	7QTZ

The data in parentheses refer to the highest-resolution shell.

[a] $R_{meas} = (|I_{hkl} -)I*|)/I_{hkl}$, where the average intensity $)I*$ is taken over all symmetry equivalent measurements, and I_{hkl} is the measured intensity for any given reflection.
[b] R value $= ||F_o| - |F_c||/|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.
[c] R_{free} is equivalent to R value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process.
* (N=11758, 2.10Å ± 0.25Å); 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst.

Structural comparison of *Ixodes ricinus* serpins solved in native conformation

When Iripin-1 is compared to other *I. ricinus* serpins in native conformation two important differences can be spotted (Fig. 3). (I) The length of RCL significantly differs compared to the 5-residue longer RCL of Iripin-8. The length of RCL relates to the kinetic stability of the formed complex with protease because RCL needs to become an additional β-strand of β-sheet A. The extension of RCL leads to decreased complex stability while shortening of RCL increases the stability but reduces the ability to inhibit proteases [13,29]. Iripin-1 has a frequent length of RCL that imply normal RCL insertion during the inhibition mechanism. (II) The breach region located on the top of the β-sheet A is responsible for hinge region insertion when enzymes form acyl intermediate with serpin RCL [30]. The breach region of Iripin-1 can be compared to that of Iripin-8 because RCL is fully exposed/stressed over serpin scaffold but when compared to Iripin-4 it differs because Iripin-4 has a partially inserted/relaxed hinge region of RCL [31].
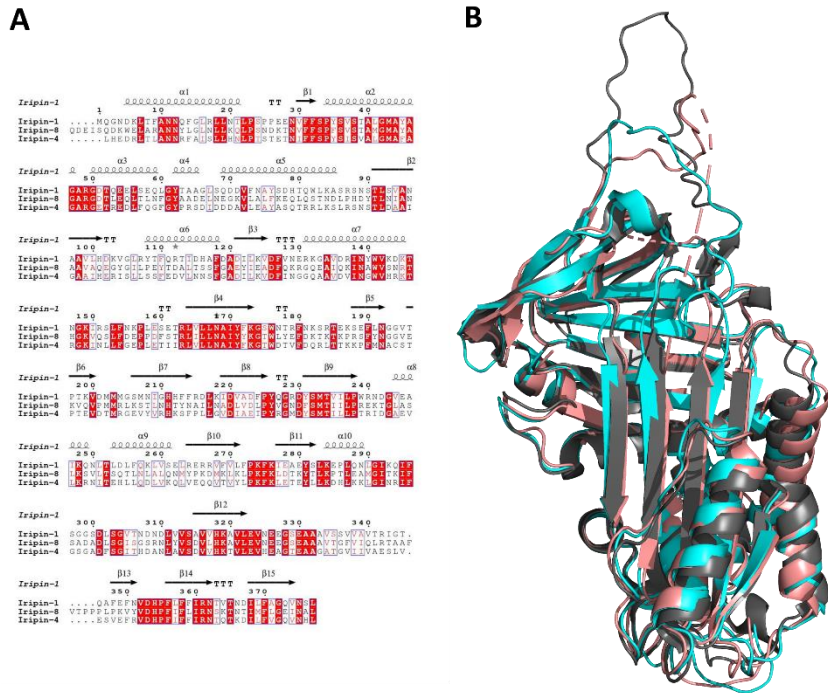
Figure. 3 **Structural alignment of known native structures of *I. ricinus* serpins.** (A) Structure-based sequence alignment of Iripin-1, Iripin-4 and Iripin-8. Secondary structure elements, which are shown above the aligned sequences, are depicted as spirals (α-helices) and arrows (β-sheets). (B) Sequence-based alignment and superposition of the native Iripin-1 structure (salmon) on the structure of native Iripin-4 (cyan) and native Iripin-8 (grey).

It was assumed that electrostatics of protein surface support complex formation [6] and thus the analysis of surface electrostatic potentials was performed for three native serpin structures from *I. ricinus.* From the generated figures from *PyMOL* [26], there is a visible difference in some regions. The large surface area of Iripin-1 breach and shutter regions are electropositive compared to both Iripin4 and Iripin-8. Next, the surface potentials of Iripin-1 β-sheet C are much more electronegative than Iripin-4 which is more neutral in charge and Iripin-8 has overly electropositive surface potential in this area. These regions contribute to protease translocation during inhibition that can helps Iripin-1 with successful inhibition of the enzyme. Further, the RCL itself is more electropositively charged than neutral Iripin-8 RCL and slightly positively charged Iripin-4 RCL. Marijanovic et al., 2019 considered the electropositivity of the loop region as a good indicator for considering the serpin as s substrate of

proteases with the electronegatively charged binding pocket such as trypsin. The stability of the complex should be achieved by charge-charge repulsion between two electropositive counterparts. This hypothesis needs to be confirmed by protease selectivity analysis.
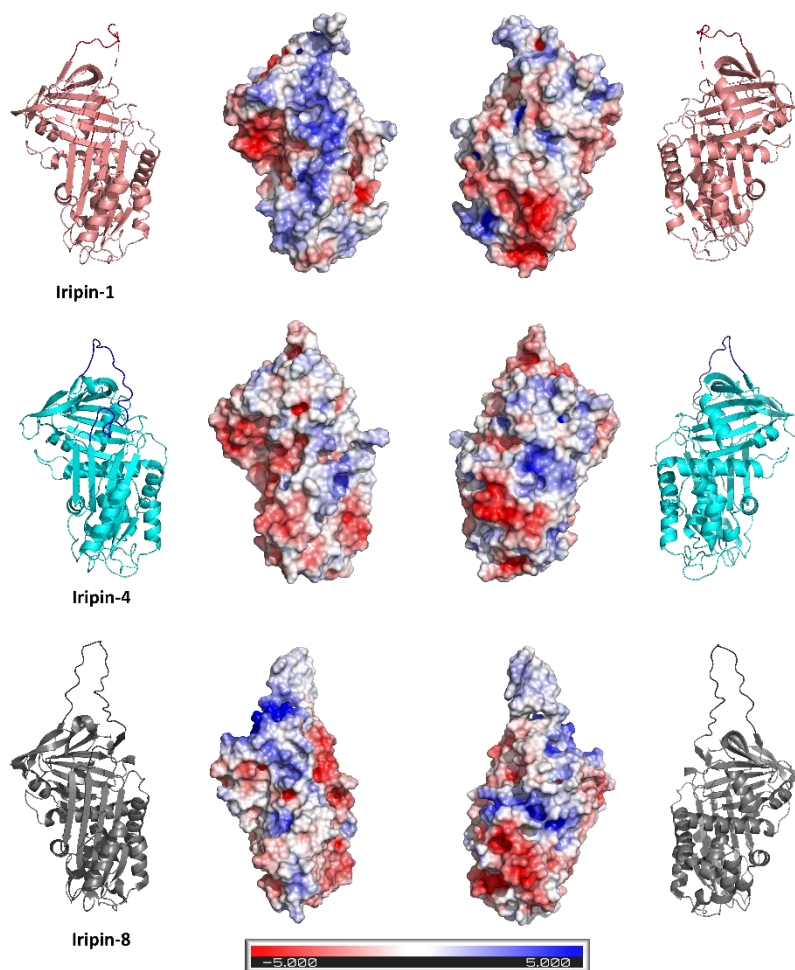

Figure. 4 Cartoon and electrostatic surface comparison of native conformations of *I. ricinus* serpins: Iripin-1, Iripin-4 and Iripin-8. The reactive center loop (RCL) is highlighted in red (Iripin-1), blue (Iripin-4) or black (Iripin-8). The loop regions with missing residues due to poor to no electron density map in the Iripin-1 structure were modelled into the structure. The electrostatic potentials were calculated using the APBS Electrostatics plugin of *PyMOL* [26]. Different colours on the electrostatic surface models represent regions with a potential value (-5.0 kT – +5.0 kT) as shown in the picture legend.

REFERENCES:

1. Silverman GA, Bird PI, Carrell RW, Church FC, Coughlin PB, Gettins PGW, Irving JA, Lomas DA, Luke CJ, Moyer RW, Pemberton PA, Remold-O'Donnell E, Salvesen GS, Travis J & Whisstock JC (2001) The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *Journal of Biological Chemistry* **276**, 33293–33296.
2. Hemanth Giri Rao V v. & Gosavi S (2018) On the folding of a structurally complex protein to its metastable active state. *Proc Natl Acad Sci U S A* **115**.
3. Bryan PN & Orban J (2010) Proteins that switch folds. *Current Opinion in Structural Biology* **20**.
4. Best RB (2018) Race to the native state. *Proc Natl Acad Sci U S A* **115**.
5. Law RHP, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, Rosado CJ, Langendorf CG, Pike RN, Bird PI & Whisstock JC (2006) An overview of the serpin superfamily. *Genome Biology* **7**, 1–11.
6. Marijanovic EM, Fodor J, Riley BT, Porebski BT, Costa MGS, Kass I, Hoke DE, McGowan S & Buckle AM (2019) Reactive centre loop dynamics and serpin specificity. *Scientific Reports* **9**.
7. Whisstock JC, Skinner R, Carrell RW & Lesk AM (2000) Conformational changes in serpins: I. The native and cleaved conformations of α1-antitrypsin. *Journal of Molecular Biology* **296**.
8. Maas C & de Maat S (2021) Therapeutic SERPINs: Improving on Nature. *Frontiers in Cardiovascular Medicine* **8**.
9. Huntington JA & Carrell RW (2001) The Serpins: Nature's Molecular Mousetraps. *Science Progress* **84**, 125–136.
10. Schechter I & Berger A (1967) On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications* **27**.
11. Lawrence DA, Ginsburg D, Day DE, Berkenpas MB, Verhamme IM, Kvassman JO & Shore JD (1995) Serpin-protease complexes are trapped as stable acyl-enzyme intermediates. *Journal of Biological Chemistry* **270**.
12. Maddur AA, Swanson R, Izaguirre G, Gettins PGW & Olson ST (2013) Kinetic intermediates en route to the final serpin-protease complex: Studies of complexes of α1-protease inhibitor with Trypsin. *Journal of Biological Chemistry* **288**.
13. Nieman MT, Lawrence DA, Renne T, De PA, Mourão S, de Maat S, Sanrattana W & Maas C (2019) SERPINs-From Trap to Treatment. *Frontiers in Medicine | www.frontiersin.org* **6**, 25.
14. Mueller U, Förster R, Hellmig M, Huschmann FU, Kastner A, Malecki P, Pühringer S, Röwer M, Sparta K, Steffien M, Ühlein M, Wilk P & Weiss MS (2015) The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *The European Physical Journal Plus* **130**.
15. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
16. Sparta KM, Krug M, Heinemann U, Mueller U & Weiss MS (2016) Xdsapp2.0. *Journal of Applied Crystallography* **49**, 1085–1092.
17. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A & Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**.
18. Evans P (2005) Scaling and assessment of data quality. *Acta Crystallographica Section D Biological Crystallography* **62**, 72–82.
19. Evans PR (2011) An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D: Biological Crystallography* **67**.

20. Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallographica Section D: Biological Crystallography* **66**.

21. Long F, Vagin AA, Young P & Murshudov GN (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125–132.

22. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography* **67**, 355–367.

23. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.

24. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2012) *MolProbity*: all-atom structure validation for macromolecular crystallography. *International Tables for Crystallography*, 694–701.

25. Young JY, Westbrook JD, Feng Z, Sala R, Peisach E, Oldfield TJ, Sen S, Gutmanas A, Armstrong DR, Berrisford JM, Chen L, Chen M, di Costanzo L, Dimitropoulos D, Gao G, Ghosh S, Gore S, Guranovic V, Hendrickx PMS, Hudson BP, Igarashi R, Ikegawa Y, Kobayashi N, Lawson CL, Liang Y, Mading S, Mak L, Mir MS, Mukhopadhyay A, Patwardhan A, Persikova I, Rinaldi L, Sanz-Garcia E, Sekharan MR, Shao C, Swaminathan GJ, Tan L, Ulrich EL, van Ginkel G, Yamashita R, Yang H, Zhuravleva MA, Quesada M, Kleywegt GJ, Berman HM, Markley JL, Nakamura H, Velankar S & Burley SK (2017) OneDep: Unified wwPDB System for Deposition, Biocuration, and Validation of Macromolecular Structures in the PDB Archive. *Structure* **25**, 536–545.

26. DeLano WL (2002) Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography* **40**.

27. Huntington JA (2011) Serpin structure, function and dysfunction. *Journal of Thrombosis and Haemostasis* **9**, 26–34.

28. Hopkins PCR, Carrell RW & Stone SR (1993) Effects of Mutations in the Hinge Region of Serpins. *Biochemistry* **32**.

29. Zhou A, Carrell RW & Huntington JA (2001) The Serpin Inhibitory Mechanism Is Critically Dependent on the Length of the Reactive Center Loop. *Journal of Biological Chemistry* **276**.

30. Khan MS, Singh P, Azhar A, Naseem A, Rashid Q, Kabir MA & Jairajpuri MA (2011) Serpin Inhibition Mechanism: A Delicate Balance between Native Metastable State and Polymerization. *Journal of Amino Acids* **2011**, 1–10.

31. Dunstone MA & Whisstock JC (2011) Crystallography of serpins and serpin complexes. *Methods in Enzymology* **501**, 63–87.

4   Conclusions

Serpins are widely distributed inhibitors that are necessary for many crucial biological processes. In hard tick (*Ixodidae*) species, serpins facilitate prolonged blood feeding necessary for reproduction and distribution of ticks. It is believed that serpins of tick *I. ricinus* modulate host defense mechanism cascade by inhibition of proteases involved in host defense. These serpins are mainly expressed in salivary glands and thus are present at the site of first contact with host. Serpins are good candidates for drug development in combination with protein engineering.

This thesis was aimed to gain structural insight into *I. ricinus* serpins. Based on the structural as well as functional analyses of serpins of interest, it was found that structural differences are important factors to consider when trying to explain functional divergence of these serpins. The results partially explain the connections between primary and secondary protein structure and define serpin function.

Iripin-3 serpin structure was solved at 1.95 Å resolution. The cleaved Iripin-3 has Arg at its P1 recognition site that indicates inhibition of trypsin-like proteases, what was confirmed by detailed antiprotease selectivity analysis against 17 different proteases. Iripin-3 mainly inhibits kallikrein and matriptase but shows this effect also to other substrates. The inhibition of kallikrein and matriptase indicates that Iripin-3 is responsible for inhibition of blood coagulation, fibrinolysis, inflammatory responses, suppressing itch and pain together with wound healing. These results show how much Iripin-3 contributes with blood feeding of ticks. Other immunomodulatory effects such as suppression of CD4$^+$ proliferation, inhibition of Th1 immune responses and depletion of IL-6 production were shown as well.

Iripin-5 was solved in cleaved conformation at 1.5 Å resolution. Iripin-5 is main *I. ricinus* salivary serpin that is highly produced after blood meal. This together with reported inhibition of neutrophil migration indicates its involvement in anti-inflammatory activity at a start of feeding. Moreover, Iripin-5 is the first *I. ricinus* serpin that regulates the complement cascade. The P1 residue is Arg (the same as for Iripin-3) and Iripin-5 inhibited primary neutrophil

elastase and proteinase 3 connected to regulation of inflammation and immune responses. This behaviour of targeting of proteases that are not "attacking" the P1 residue was previously observed and indicates involvement not only RCL, but also other parts so called exosites of serpin needed for successful inhibition.

Iripin-4 was solved in both native and cleaved conformations. The cleaved conformation solved at 2.0 Å resolution did not differ significantly from Iripin-3 and Iripin-5 cleaved structures. Nevertheless, the native structure was interesting because the structure was in partially relaxed state solved at 2.3 Å resolution. The presence of Glu instead of typical arginine at the P1 recognition site for the most of characterized *I. ricinus* serpins can explain low protease inhibition of tested proteases. The use of proteases that target Glu would be beneficial for better explanation of Iripin-4 function. Further research on Iripin-4 should focus on functional analysis of this interesting serpin.

X-ray crystallographic analysis of Iripin-1 revealed its native conformation in fully exposed state. The only does not resolved part of protein was hinge region of RCL that is responsible for RCL flexibility, and this was the reason why this part was not modelled into the structure. The P1 recognition site is represented by Arg residue that suggests inhibition of trypsin-like proteases responsible for regulation of immune responses and inflammation, but further functional analyses are required to confirm this assumption.

By analysing the sequences of studied serpins there is only ~50% sequence identity compared to almost identical structural fold when compared same structural conformations. This confirmed well-known fact that serpins are structurally conserved but functionally diverse proteins.

This research illustrates that not only the RCL can be responsible for the inhibition function, but it also raises the question how much surface electrostatics, exosites and other parameters are involved during inhibition mechanism.

Further research will be necessary to determine relationship between RCL sequence and function. The crystallographic study of formed covalent complexes or Michaelis-Menten complexes could bring better view on how serpin structure contributes to inhibition process. The better understanding of

serpins can be also achieved by mutation of RCL residues and further structural and functional studies could bring new insight into its function.

5 References

1. Koiou E, Tziomalos K, Dinas K, Katsikis I, Kandaraki EA, Tsourdi E, Mavridis S & Panidis D (2012) Plasma plasminogen activator inhibitor-1 levels in the different phenotypes of the polycystic ovary syndrome. *Endocrine Journal* **59**.

2. Nieman MT, Lawrence DA, Renne T, De PA, Mourão S, de Maat S, Sanrattana W & Maas C (2019) SERPINs-From Trap to Treatment. *Frontiers in Medicine | www.frontiersin.org* **6**, 25.

3. Francischetti IMB, Sa-Nunes A, Mans BJ, Santos IM & Ribeiro JMC (2009) The role of saliva in tick feeding. *Frontiers in Bioscience* **14**.

4. Sauer JR, Essenberg RC & Bowman AS (2000) Salivary glands in ixodid ticks: Control and mechanism of secretion. *Journal of Insect Physiology* **46**.

5. Medlock JM, Hansford KM, Bormane A, Derdakova M, Estrada-Peña A, George J-C, Golovljova I, Jaenson TGT, Jensen J-K, Jensen PM, Kazimirova M, Oteo JA, Papa A, Pfister K, Plantard O, Randolph SE, Rizzoli A, Santos-Silva MM, Sprong H, Vial L, Hendrickx G, Zeller H & van Bortel W (2013) Driving forces for changes in geographical distribution of Ixodes ricinus ticks in Europe. *Parasit Vectors* **6**, 1.

6. Estrada-Peñ A, Farkas R, T Jaenson TG, Koenen F, Madder M, Pascucci I, Salman M, Tarrés-Call J, Jongejan F, Farkas R, T Jaenson TG, Koenen F, Madder M, Madder Á Jongejan MF & Pascucci I (2013) Association of environmental traits with the geographic ranges of ticks (Acari: Ixodidae) of medical and veterinary importance in the western Palearctic. A digital data set. *Exp Appl Acarol* **59**, 351–366.

7. de la Fuente J, Kopáček P, Lew-Tabor A & Maritz-Olivier C (2016) Strategies for new and improved vaccines against ticks and tick-borne diseases. *Parasite Immunology* **38**.

8. Hajdušek O, Šíma R, Ayllón N, Jalovecká M, Perner J, de la Fuente J & Kopáček P (2013) Interaction of the tick immune system with transmitted pathogens. *Frontiers in Cellular and Infection Microbiology* **4**.

9. Chmelař J, Kotál J, Langhansová H & Kotsyfakis M (2017) Protease inhibitors in tick saliva: The role of serpins and cystatins in tick-host-pathogen interaction. *Frontiers in Cellular and Infection Microbiology* **7**, 1–16.

10. Chudzinski-Tavassi AM, Morais KLP, Pacheco MTF, Pasqualoto KFM & de Souza JG (2016) Tick salivary gland as potential natural source for the discovery of promising antitumor drug candidates. *Biomedicine and Pharmacotherapy* **77**.

11. Parizi LF, Ali A, Tirloni L, Oldiges DP, Sabadin GA, Coutinho ML, Seixas A, Logullo C, Termignoni C & da Silva Vaz I (2018) Peptidase inhibitors in tick physiology. *Medical and Veterinary Entomology* **32**, 129–144.

12. Silverman GA, Bird PI, Carrell RW, Church FC, Coughlin PB, Gettins PGW, Irving JA, Lomas DA, Luke CJ, Moyer RW, Pemberton PA, Remold-O'Donnell E, Salvesen GS, Travis J & Whisstock JC (2001) The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *Journal of Biological Chemistry* **276**, 33293–33296.

13. Law RHP, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, Rosado CJ, Langendorf CG, Pike RN, Bird PI & Whisstock JC (2006) An overview of the serpin superfamily. *Genome Biology* **7**, 1–11.

14. Gettins PGW (2002) Serpin structure, mechanism, and function. *Chemical Reviews* **102**, 4751–4803.

15. HUNTINGTON JA (2011) Serpin structure, function and dysfunction. *Journal of Thrombosis and Haemostasis* **9**, 26–34.

16. Schechter I & Berger A (1967) On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications* **27**.

17. Song J, Matthews AY, Reboul CF, Kaiserman D, Pike RN, Bird PI & Whisstock JC (2011) Predicting serpin/protease interactions. In *Methods in Enzymology*.

18. Maddur AA, Swanson R, Izaguirre G, Gettins PGW & Olson ST (2013) Kinetic intermediates en route to the final serpin-protease complex: Studies of complexes of α1-protease inhibitor with Trypsin. *Journal of Biological Chemistry* **288**.

19. Lucas A, Yaron JR, Zhang L & Ambadapadi S (2018) Overview of serpins and their roles in biological systems. In *Methods in Molecular Biology*.

20. Lawrence DA, Olson ST, Muhammad S, Day DE, Kvassman JO, Ginsburg D & Shore JD (2000) Partitioning of serpin-proteinase reactions between stable inhibition and substrate cleavage is regulated by the rate of serpin reactive center loop insertion into β-sheet A. *Journal of Biological Chemistry* **275**.

21. Zhou A, Carrell RW & Huntington JA (2001) The Serpin Inhibitory Mechanism Is Critically Dependent on the Length of the Reactive Center Loop. *Journal of Biological Chemistry* **276**.

22. Rashid Q, Kapil C, Singh P, Kumari V & Jairajpuri MA (2015) Understanding the specificity of serpin-protease complexes through interface analysis. *Journal of Biomolecular Structure and Dynamics* **33**, 1352–1362.

23. Dunstone MA & Whisstock JC (2011) Crystallography of serpins and serpin complexes. *Methods in Enzymology* **501**, 63–87.

24. Yang B, Zhang M & Luo T (2020) Identification of Potential Core Genes Associated With the Progression of Stomach Adenocarcinoma Using Bioinformatic Analysis. *Frontiers in Genetics* **11**.

25. Dunstone MA & Whisstock JC (2011) Crystallography of serpins and serpin complexes. In *Methods in Enzymology* pp. 63–87. Academic Press Inc.

26. Gettins PGW & Olson ST (2016) Inhibitory serpins. New insights into their folding, polymerization, regulation and clearance. *Biochemical Journal* **473**, 2273–2293.

27. Yamasaki M, Li W, Johnson DJD & Huntington JA (2008) Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* **455**, 1255–1258.

28. Tirloni L, Seixas A, Mulenga A, da Silva Vaz I & Termignoni C (2014) A family of serine protease inhibitors (serpins) in the cattle tick Rhipicephalus (Boophilus) microplus. *Experimental Parasitology* **137**, 25–34.

29. Porter L, Radulović Z, Kim T, Braz GRC, da Silva Vaz I & Mulenga A (2015) Bioinformatic analyses of male and female amblyomma americanum tick expressed serine protease inhibitors (serpins). *Ticks and Tick-borne Diseases* **6**, 16–30.

30. Prevot PP, Adam B, Boudjeltia KZ, Brossard M, Lins L, Cauchie P, Brasseur R, Vanhaeverbeek M, Vanhamme L & Godfroid E (2006) Anti-hemostatic effects of a serpin from the saliva of the tick Ixodes ricinus. *Journal of Biological Chemistry* **281**, 26361–26369.

31. Prevot P-P, Beschin A, Lins L, rô me Beaufays J, lie Grosjean A, Bruys L, Adam B, Brossard M, Brasseur R, Zouaoui Boudjeltia K, Vanhamme L & Godfroid E Exosites mediate the anti-inflammatory effects of a multifunctional serpin from the saliva of the tick Ixodes ricinus.

32. Chmelar J, Oliveira CJ, Rezacova P, Francischetti IMB, Kovarova Z, Pejler G, Kopacek P, Ribeiro JMC, Mares M, Kopecky J & Kotsyfakis M (2011) A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood* **117**, 736–744.

33. Prevot PP, Couvreur B, Denis V, Brossard M, Vanhamme L & Godfroid E (2007) Protective immunity against Ixodes ricinus induced by a salivary serpin. *Vaccine* **25**, 3284–3292.

34. Helliwell JR (2017) X-Ray Crystallography . Second Edition. By William Clegg. Oxford University Press, 2015. Pp. 128. Price GBP 14.99 (paperback). ISBN 9780198700975. . *Acta Crystallographica Section A Foundations and Advances* **73**.

35. Dessau MA & Modis Y (2010) Protein crystallization for X-ray crystallography. *Journal of Visualized Experiments*.
36. Dauter Z & Wlodawer A *Progress in protein crystallography*.
37. Garman EF (2014) Developments in X-ray crystallographic structure determination of biological macromolecules. *Science (1979)* **343**, 1102–1108.
38. Alexander McPherson & Jose A. Gavira (2014) Introduction to protein crystallization. *Acta Crystallographica Section F:Structural Biology Communications* .
39. Mcpherson A *INTRODUCTION TO MACROMOLECULAR CRYSTALLOGRAPHY Second Edition*.
40. Dawe LN (2018) Principles of X-ray crystallography. *Crystallography Reviews* **24**.
41. Miller RJD (2014) Femtosecond crystallography with ultrabright electrons and x-rays: Capturing chemistry in action. *Science (1979)* **343**.
42. McPherson A, Malkin AJ & Kuznetsov YG (2000) Atomic force microscopy in the study of macromolecular crystal growth. *Annual Review of Biophysics and Biomolecular Structure* **29**.
43. Drenth J & Mesters J (2007) *Principles of protein X-ray crystallography: Third edition*.
44. Patrick MH (2007) Crystallography made crystal clear: A guide for users of macromolecular models (3rd Ed.). *Biochemistry and Molecular Biology Education* **35**.
45. Nave C & Garman EF (2005) Synchrotron Radiation Towards an understanding of radiation damage in cryocooled macromolecular crystals. *J Synchrotron Rad* **12**, 257–260.
46. Schlichting I (2015) Serial femtosecond crystallography: The first five years. *IUCrJ* **2**.
47. Gruner SM, Eikenberry EF & Tate MW (2012) *Chapter 7.1. Comparison of X-ray detectors*.
48. Jakubek J (2009) Energy-sensitive X-ray radiography and charge sharing effect in pixelated detector. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **607**.
49. Rupp B (2009) *Biomolecular crystallography: principles, practice, and application to structural biology.* Garland Science.
50. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**.
51. Leslie AGW (2006) The integration of macromolecular diffraction data. In *Acta Crystallographica Section D: Biological Crystallography*.
52. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
53. Maveyraud L & Mourey L (2020) Protein X-ray crystallography and drug discovery. *Molecules* **25**.
54. Blow DM (2003) How Bijvoet Made the Difference: The Growing Power of Anomalous Scattering. *Methods in Enzymology* **374**.
55. Liu Q, Liu Q & Hendrickson WA (2013) Robust structural analysis of native biological macromolecules from multi-crystal anomalous diffraction data. *Acta Crystallographica Section D: Biological Crystallography* **69**.
56. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallographica Section A: Foundations of Crystallography* **64**.
57. Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallographica Section D: Biological Crystallography* **66**.
58. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *Journal of Applied Crystallography* **40**.
59. Navaza J (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallographica Section A* **50**.

60. Rodríguez D, Sammito M, Meindl K, de Ilarduya ÍM, Potratz M, Sheldrick GM & Usón I (2012) Practical structure solution with ARCIMBOLDO. *Acta Crystallographica Section D: Biological Crystallography* **68**.

61. Keegan RM & Winn MD (2007) MrBUMP: An automated pipeline for molecular replacement. In *Acta Crystallographica Section D: Biological Crystallography*.

62. Long F, Vagin AA, Young P & Murshudov GN (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125–132.

63. Terwilliger TC, di Maio F, Read RJ, Baker D, Bunkóczi G, Adams PD, Grosse-Kunstleve RW, Afonine P v. & Echols N (2012) Phenix.mr-rosetta: Molecular replacement and model rebuilding with Phenix and Rosetta. *Journal of Structural and Functional Genomics* **13**.

64. Langer G, Cohen SX, Lamzin VS & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* **3**.

65. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D: Biological Crystallography* **62**.

66. Terwilliger TC, Grosse-Kunstleve RW, Afonine P v., Moriarty NW, Zwart PH, Hung LW, Read RJ & Adams PD (2007) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. In *Acta Crystallographica Section D: Biological Crystallography*.

67. Usón I & Sheldrick GM (2018) An introduction to experimental phasing of macromolecules illustrated by SHELX; New autotracing features. *Acta Crystallographica Section D: Structural Biology* **74**.

68. Langer GG, Hazledine S, Wiegels T, Carolan C & Lamzin VS (2013) Visual automated macromolecular model building. *Acta Crystallographica Section D: Biological Crystallography* **69**.

69. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.

70. Afonine P v., Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D: Biological Crystallography* **68**.

71. Painter J & Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *Journal of Applied Crystallography* **39**.

72. DiMaio F, Tyka MD, Baker ML, Chiu W & Baker D (2009) Refinement of Protein Structures into Low-Resolution Density Maps Using Rosetta. *Journal of Molecular Biology* **392**.

73. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.

74. McREE DE (1999) COMPUTATIONAL TECHNIQUES. *Practical Protein Crystallography*, 91-cp1.

75. Carugo O (1999) *Correlation between occupancy and B factor of water molecules in protein crystal structures*.

76. Harlow RL *Troublesome Crystal Structures: Prevention, Detection, and Resolution*.

77. Sun Z, Liu Q, Qu G, Feng Y & Reetz MT (2019) Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chemical Reviews*.

78. Powell HR (2021) A Beginner's Guide to X-Ray Data Processing. *Biochemist* **43**.

79. Laskowski RA, MacArthur MW, Moss DS & Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**.

80. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2012) *MolProbity*: all-atom structure

validation for macromolecular crystallography. *International Tables for Crystallography*, 694–701.

81. Weichenberger CX, Pozharski E & Rupp B (2013) Visualizing ligand molecules in twilight electron density. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **69**.

82. Zheng H, Cooper DR, Porebski PJ, Shabalin IG, Handing KB & Minor W (2017) CheckMyMetal: A macromolecular metal-binding validation tool. *Acta Crystallographica Section D: Structural Biology* **73**.

83. Abola EE, Manning NO, Prilusky J, Stampf DR & Sussman JL (1996) The Protein Data Bank: Current status and future challenges. *Journal of Research of the National Institute of Standards and Technology* **101**.

Structural insight into the salivary serpins of *Ixodes ricinus*
Ph.D. Thesis Series, 2022, No. 8