STRUCTURAL
BIOLOGY

Volume 77 (2021)

Supporting information for article:

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

Barbora Kascakova, Jan Kotal, Larissa Almeida Martins, Zuzana Berankova, Helena Langhansova, Eric Calvo, Joel A. Crossley, Petra Havlickova, Filip Dycka, Tatyana Prudnikova, Michal Kuty, Michail Kotsyfakis, Jindrich Chmelar and Ivana Kuta Smatanova

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 %.

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-chymotrypsin**	**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**

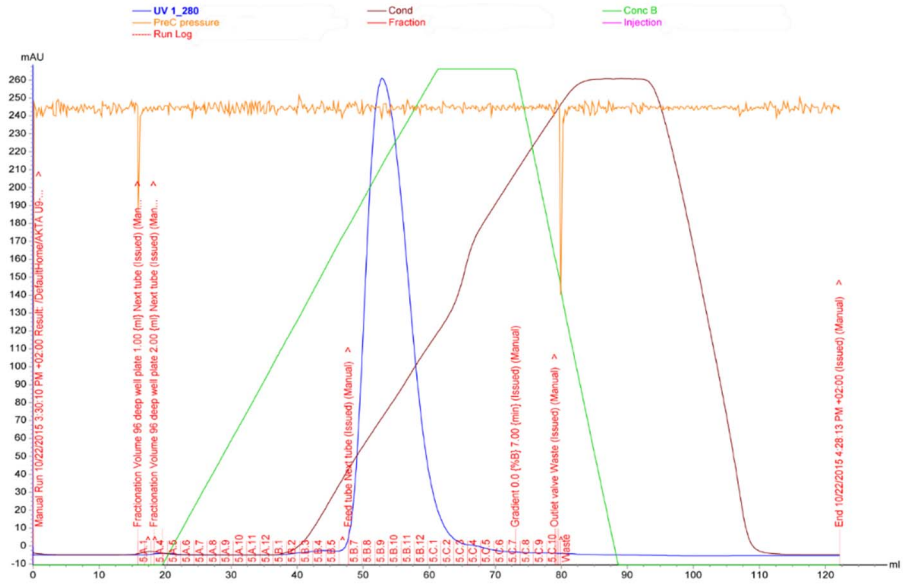

Figure S1 Anion exchange chromatography (HiTrap® Q High Performance - GE17-1154-01 column) with use of 0.5 M NaCl as an eluent. The equilibration buffer composition was 20 mM Tris, pH 8.5 and elution buffer was 20mM Tris, 500mM NaCl, pH 8.5. Iripin-5 started to elute at ~68% of 0.5M NaCl gradient and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

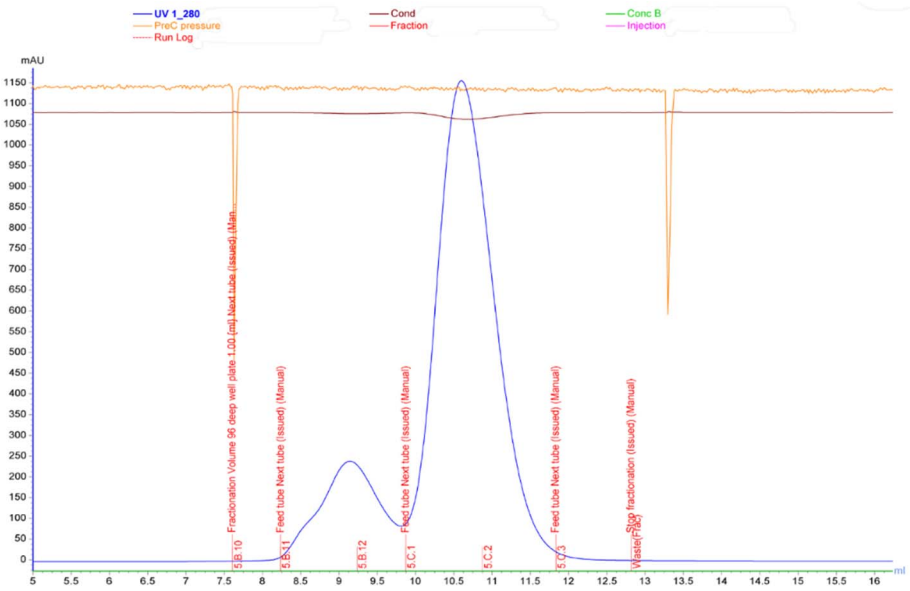

Figure S2 Size exclusion chromatography (Superdex 75 10/300). The buffer composition was 20mM Tris, 150mM NaCl, pH 8.5. The first peak probably represents the multimeric state of protein and second, larger peak monomeric Iripin-5 that was used for further analyses and for crystallization.

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.

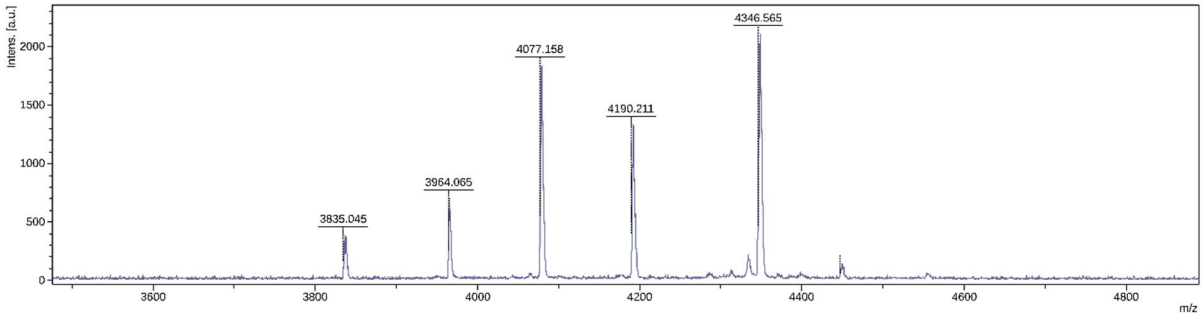

Figure S4 Intact protein analysis using MALDI-TOF MS. Each peak represents the mass of the observed peptides from analysis. These peptides can be results from protease digestion probably during the sample storage. The sequence and theoretical mass for each peptide are listed in Table 1 (supplementary material).

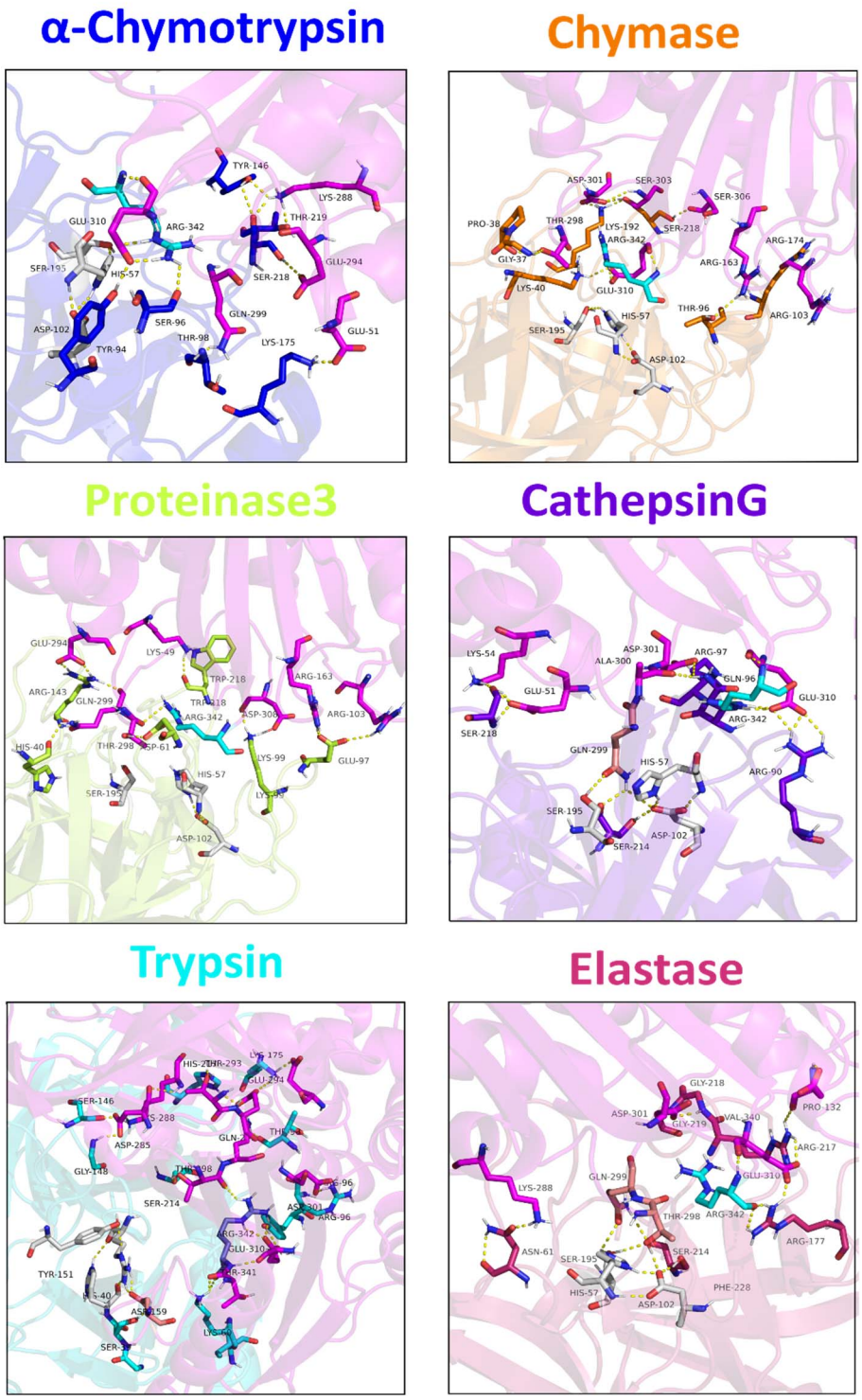

Figure S5 The focus and visualization of Iripin-5 and chosen proteases [α-Chymotrypsin (blue), Chymase (orange), Proteinase 3 (lemon), Cathepsin G (violet), Trypsin (cyan), Elastase (hot pink)] are rotated for a better view. The catalytic residues side chains of proteases are presented like grey sticks; 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.

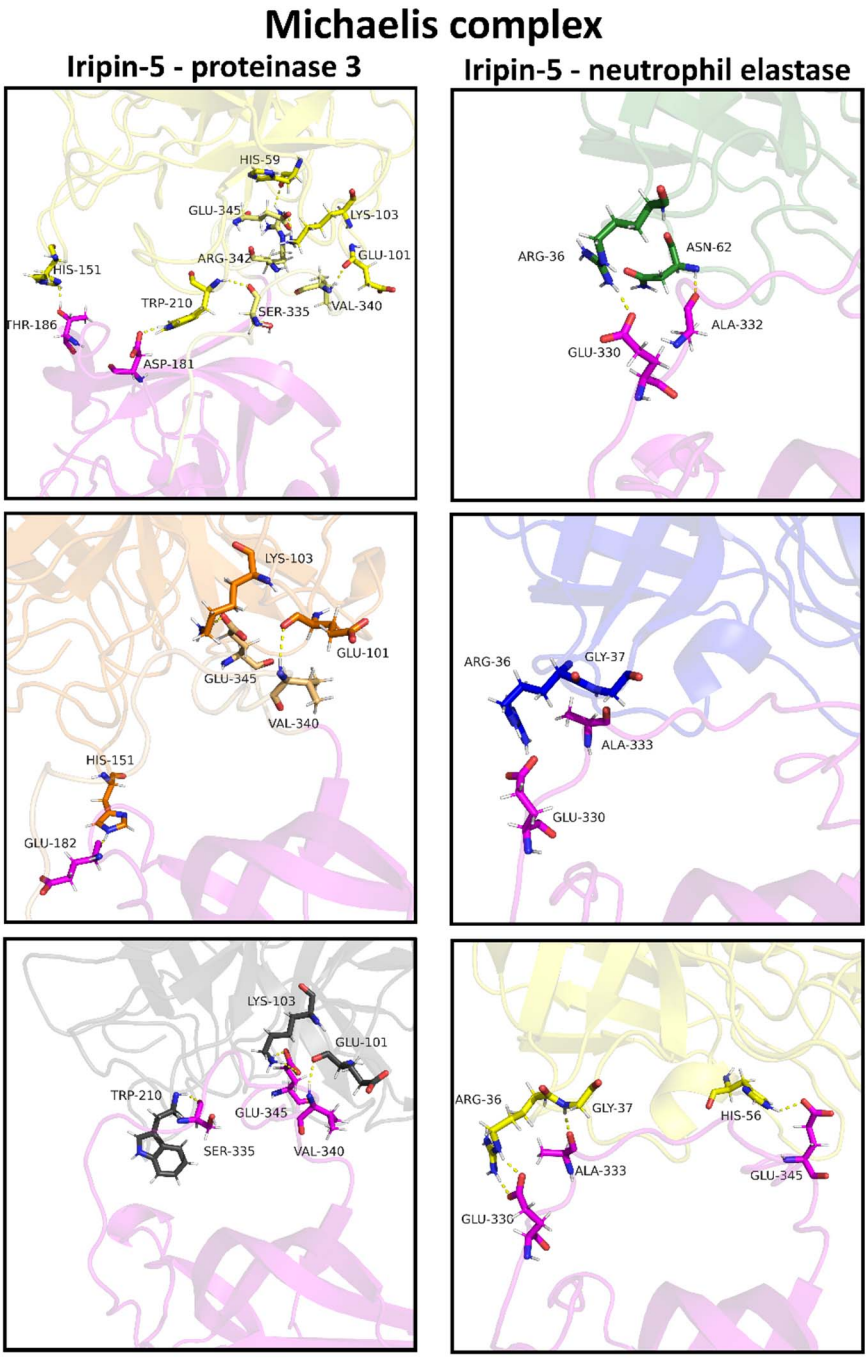

Figure S6 The focus and visualization of Michaelis complex interface of Iripin-5 and chosen proteases, the zoom is rotated for a better view. The side chain residues of proteases are presented by different colour (proteinase 3 – yellow, orange, and grey; neutrophil elastase – green, blue, and yellow)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

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

Rappsilber, J., Mann, M. & Ishihama, Y. (2007). *Nature protocols*. **2**, 1896–1906.

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