Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Biochemie

Mgr. Zuzana Matoušková

Charakterizace proteinů interagujících s proteolytickým systémem klíšťat

Characterization of proteins interacting with proteolytic system of ticks

Disertační práce

Školitel: RNDr. Michael Mareš, CSc.

Praha, 2024

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, dne

……..

Podpis

Poděkování

Ráda bych na tomto místě poděkovala všem, kteří mi pomáhali a podporovali mě během celého doktorského studia. V první řadě děkuji svému školiteli Michaeli Marešovi za jeho cenné rady, vstřícnost, ochotu a čas, který mi věnoval. Rovněž děkuji všem současným i bývalým kolegům a kolegyním za všestrannou pomoc a milé pracovní prostředí. Největší poděkování patří mé rodině za trpělivost a nesmírnou podporu.

Obsah

Seznam zkratek

Å	Ångström, 10-10 m
AMC	7-amino-4-methylkumarin
Abz	kyselina aminobenzoová
BPTI	„bovine pancreatic trypsin inhibitor“, inhibitor trypsinu z hovězí slinivky
BmAP	aspartátová proteasa z klíštěte *Rhipicephalus (Boophilus) microplus*
BmCL1	cysteinová proteasa podobná katepsinu L z klíštěte *Rhipicephalus (Boophilus) microplus*
CTAP-III	„connective tissue-activating peptide III“, peptid aktivující pojivové tkáně III
DTT	dithiotreitol
E-64	N-[N-(L-3-trans-karboxyoxirin-2-karbonyl)-L-leucin]-agmatin
EDTA	kyselinu ethylendiamintetraoctová
FPLC	„fast protein liquid chromatography“, rychlá proteinová kapalinová chromatografie
FRET	„fluroescence resonance energy transfer“, fluorescenční rezonanční přenos energie
GAG	glykosaminoglykan
IC_{50}	koncentrace inhibitoru potřebná k dosažení 50% inhibice aktivity enzymu
IGF	„insulin-like growth factor“, insulinu podobný růstový faktor
IGF-BP	„insulin-like growth factor-binding protein, protein vázající insulinu podobný růstový faktor
IrAE	legumain (asparaginyl endopeptidasa) z klíštěte *Ixodes ricinus*
IrCB	katepsin B z klíštěte *Ixodes ricinus*
IrCC	katepsin C z klíštěte *Ixodes ricinus*
IrCD	katepsin D z klíštěte *Ixodes ricinus*
IrCL	katepsin L z klíštěte *Ixodes ricinus*
Ir-CPI	Kunitzův inhibitor z klíštěte *Ixodes ricinus*
IRS-2	serpin 2 z klíštšte *Ixodes ricinus*
IrSPI	Kunitzův inhibitor z klíštěte *Ixodes ricinus*
IrThy	tyropin z klíštěte *Ixodes ricinus*
K_m	Michaelisova konstanta
K_i	inhibiční konstanta
LAP	leucinová aminopeptidasa z klíštěte *Ixodes ricinus*
MEROPS	peptidasová databáze
MHC	„major histocompatibility complex“, hlavní histokompatibilní komplex
mMCP-4	„mouse mast cell protease-4“, proteasa žírných buněk z myši 4
NAP-2	„neutrophil-activating peptide-2“, peptid aktivující neutrofily 2
NMR	nukleární magnetická rezonance
OmC2	cystatin 2 z klíštěte *Ornithodorus moubata*

P41	fragment p41invariantního řetězce asociovaného s MHC třídou II
PAR1	„protease activated receptor-1“, receptor aktivovaný proteasou 1
PAR4	„protease activated receptor-4“, receptor aktivovaný proteasou 4
PDB	„protein data bank”, databáze proteinových struktur
PVDF	polyvinylidenfluorid
R	„relaxed“, relaxovaná konformace
RCL	„reactive centre loop“, smyčka reaktivního centra
S	„stressed“, stresová konformace
SCP	serinová karboxypeptidasa z klíštěte *Ixodes ricinus*
SDS-PAGE	elektroforéza na polyakrylamidovém gelu v přítomnosti SDS
Tg1	tyreoglobulinová doména 1
Tris	tris(hydroxymethyl)aminomethan
VTDCE	„vitellin-degrading cysteine endopeptidas“, cysteinová endopeptidasa degradující vitelin

Abstrakt (ČJ)

Klíšťata jsou ektoparazité s globálním rozšířením, kteří se živí krví hostitele a přenášejí řadu významných patogenů na člověka a domácí zvířata. Úspěšné sání klíšťat je umožněno bioaktivními molekulami ve slinách klíšťat, které jsou injikovány do tkáně hostitele. Mezi intenzivně studované patří proteasové inhibitory proteinového charakteru, které jsou atraktivní pro využití v biomedicíně.

Tato práce je zaměřena na nové typy proteasových inhibitorů ze slin klíštěte *Ixodes ricinus*, vektoru Lymské boreliózy a klíšťové encefalitidy, a *Ornithodorus moubata*, vektoru návratné horečky a afrického moru prasat. Výzkum se soustředil na biochemickou a strukturní charakterizaci tří zástupců proteasových inhibitorů z rodin cystatinů, serpinů a tyropinů jako základ pro vysvětlení jejich biologické funkce při interakci klíštěte s hostitelem.

Cystatin OmC2 z *O. moubata* byl identifikován jako širokospektrální inhibitor cysteinových katepsinů hostitele s endopeptidasovou i exopeptidasovou aktivitou. Vyřešení krystalové struktury umožnilo popsat vztah mezi strukturou OmC2 a jeho inhibiční specifitou. V biologických testech byla prokázána schopnost OmC2 modulovat imunitní odpověď hostitele a ve vakcinačních experimentech supresní efekt na *O. moubata*. Serpin IRS-2 z *I. ricinus* je účinným inhibitorem dvou serinových proteas imunitních buněk hostitele, chymasy a katepsinu G. Díky tomu potlačuje IRS-2 procesy agregace krevních destiček a akutního zánětu, jichž se tyto proteasy účastní. Funkční specifita reaktivního centra IRS-2 byla analyzována pomocí krystalové struktury. Tyropin IrThy z *I. ricinus* má úzkou inhibiční specifitu omezenou pouze na tři cysteinové katepsiny hostitele s endopeptidasovou aktivitou, které se účastní imunitních odpovědí. Unikátní inhibiční specifita byla vysvětlena pomocí vyřešené NMR struktury a bylo zjištěno, že specifita je dále modulována glykosaminoglykany z tkání hostitele.

Disertační práce přináší nové významné informace pro pochopení mechanismů molekulárních interakcí mezi klíštětem a hostitelem a dále bioaktivní proteiny, které mají potenciální využití jako antigeny při vývoji protiklíštěcích vakcín a léčiva s unikátními farmakologickými účinky.

Abstract (EN)

Ticks are ectoparasites found worldwide that feed on the blood of their hosts and transmit several important pathogens to humans and domestic animals. Tick saliva contains bioactive molecules that are injected into the host tissue to aid in successful blood feeding. Among these molecules, proteinaceous protease inhibitors are being extensively studied for their potential biomedical applications.

This work focuses on novel protease inhibitors from the saliva of the ticks *Ixodes ricinus*, vector of Lyme disease and tick-borne encephalitis, and *Ornithodorus moubata*, vector of relapsing fever and African swine fever. Research has focused on the biochemical and structural characterization of three members of the protease inhibitor families of cystatins, serpins, and tyropins, and has attempted to elucidate their biological function in tick-host interactions.

Cystatin OmC2 from *O. moubata* was identified as a broad-spectrum inhibitor of host cysteine cathepsins with both endopeptidase and exopeptidase activity. Crystal structure determination allowed description of the relationship between the structure of OmC2 and its inhibitory specificity. The ability of OmC2 to modulate the host immune response was demonstrated in bioassays and a suppressive effect on *O. moubata* in vaccination experiments. The serpin IRS-2 from *I. ricinus* is a potent inhibitor of two host immune cell serine proteases, chymase and cathepsin G. As a result, IRS-2 inhibits the processes of platelet aggregation and acute inflammation in which these proteases are involved. The functional specificity of the IRS-2 reactive center was analyzed by crystal structure analysis. The tyropin IrThy from *I. ricinus* has a narrow inhibitory specificity limited to only three host cysteine cathepsins with endopeptidase activity that are involved in immune responses. The unique inhibitory specificity was explained by the solved NMR structure, and the specificity was found to be further modulated by glycosaminoglycans from host tissues.

The thesis provides new insights into the mechanism of molecular interactions between ticks and hosts, as well as bioactive proteins that could serve as antigens for the development of anti-tick vaccines and drugs with unique pharmacological effects.

1. Úvod

Klíšťata jsou celosvětově rozšířenými krev sajícími parazity. Přenášejí celou řadu nebezpečných a život ohrožujících onemocnění na člověka a domácí zvířata[1-4]. Klíště může být rezervoárem několika infekčních agens a přenášet tak více onemocnění jediným kontaktem s hostitelem[3]. V Evropě, Asii a Severní Americe patří mezi vážná onemocnění přenášená klíšťaty, s nejvyšším počtem hlášených případů, Lymská borelióza způsobená spirochétou druhu *Borrelia burgdorferi*[5,6]. Dalším často přenášeným onemocněním v Evropě a východní a střední Asii je klíšťová encefalitida. Způsobují ji různé subtypy viru klíšťové encefalitidy z rodu flavivirů[7,8]. Obě onemocnění jsou přenášena klíšťaty z čeledi *Ixodidae*, a pokud nejsou včas zachycena, mohou zanechat doživotní vážné následky a způsobit i smrt infikovaného jedince[7,9]. Dále jsou klíšťata i problémem ekonomickým. Sání velkého počtu klíšťat u hospodářského dobytka snižuje produkci masa a mléka, což vede k významným finančním ztrátám[10,11].

Přestože existuje vakcinace proti některým onemocněním přenášeným klíšťaty, v současné době stále chybí tzv. protiklíštěcí vakcína, která by zabraňovala sání klíšťat na hostiteli, a tím také snižovala riziko přenosu patogenů. Tato strategie by také vedla ke snížení populace vektoru. Potenciálními molekulárními cíli pro vývoj takové vakcíny jsou bioaktivní látky obsažené ve slinách klíšťat[12,13]. U klíšťat stejně jako i dalších krev sajících parazitů mají sliny injikované do hostitele klíčovou úlohu pro potlačení fyziologické obranné reakce hostitele a zajištění nerušeného sání.

Hostitelské proteasy se účastní vrozené i získané imunitní odpovědi na parazity a řídí hemostázu, proto je pro parazity výhodné regulovat jejich aktivitu. K tomu používají proteasové inhibitory akumulované ve slinách[14]. Biologická aktivita těchto inhibitorů již byla úspěšně využita v biomedicíně. Příkladem je proteasový inhibitor hirudin z pijavice lékařské (*Hirudo medicinalis*) jako antitrombotické léčivo[15]. Poslední dekáda přinesla zásadní informace o bioaktivních proteasových inhibitorech ze slin klíšťat. Tato práce se zaměřuje na detailní biochemický popis tří zástupců těchto proteinů a přináší jejich první strukturně-funkční popis.

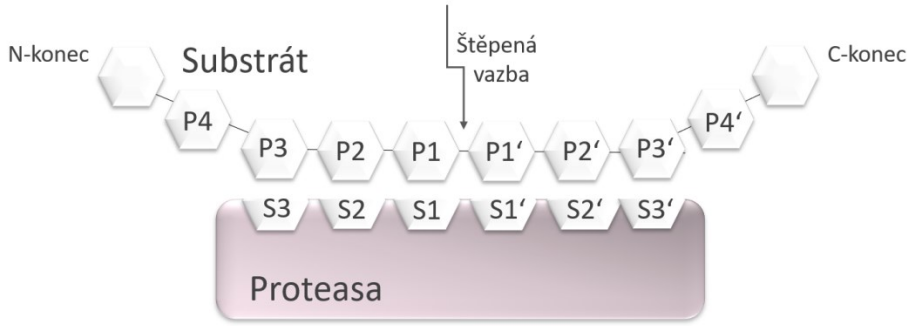
2. Literární přehled

2.1. Proteinové inhibitory proteolytických enzymů

2.1.1. Proteolytické enzymy

Proteasy, známé také jako proteinasy nebo peptidasy, jsou enzymy, které katalyzují hydrolýzu peptidové vazby mezi aminokyselinovými zbytky polypeptidového řetězce. Mohou být specifické a štěpit cíleně pouze jedno nebo více míst v proteinovém substrátu nebo nespecifické a degradovat proteiny až na úroveň jednotlivých aminokyselin. Proteasy štěpící od N- nebo C-konce jsou označovány jako exopeptidasy (aminopeptidasy, resp. karboxypeptidasy) a ty, které štěpí uprostřed polypeptidového řetězce, jako endopeptidasy. Proteasy lze na základě mechanismu katalýzy a aminokyselinových zbytků přítomných v katalytickém centru rozdělit do šesti tříd na cysteinové, serinové, treoninové, aspartátové, glutamátové proteasy a metaloproteasy[16-18]. Dále se tyto třídy klasifikují na základě homologie aminokyselinových sekvencí a strukturní podobnosti do rodin a klanů[17].

Vazba peptidového substrátu do aktivního místa proteasy je popsána konvenční terminologií[19]. Pozice aminokyselinových zbytků substrátu se označují písmenem P, ve směru k N-konci P1 až Pn a ve směru k C-konci P1' až Pn'. Podmísta proteasy, do kterých se vážou jednotlivé aminokyselinové zbytky substrátu, se označují analogicky písmenem S. Ke štěpení peptidové vazby dochází mezi zbytky P1-P1', které se vážou do podmíst S1-S1'. Schématické znázornění na Obr. 2.1 je možné použít nejen pro vazbu substrátu do aktivního místa proteasy, ale i peptidového inhibitoru.

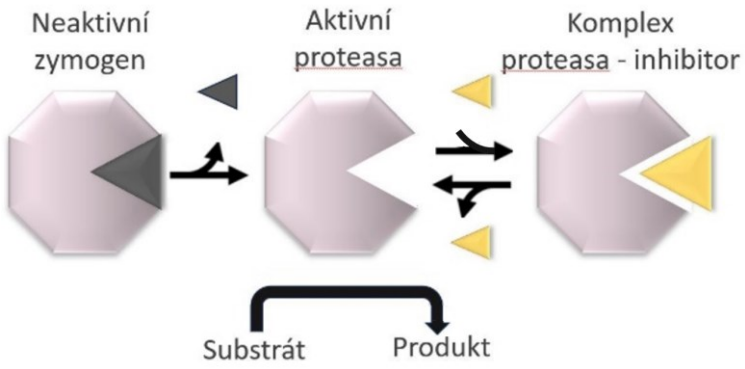


***Obr. 2.1: Schématické znázornění vazby peptidového substrátu do aktivního místa proteasy**. Aminokyselinové zbytky peptidového substrátu jsou označeny jako pozice P1–P4 ve směru k N-konci a P1'–P4' ve směru k C-konci. Aktivní místo proteasy je znázorněno jako řada podmíst S3–S3', kdy každé váže příslušný aminokyselinový zbytek substrátu. Štěpená vazba se nachází mezi pozicemi P1–P1'. Upraveno podle [19].*

Proteasy jsou přítomné ve všech organismech a jsou zapojené do všech životně důležitých funkcí, jako příklad lze uvést programovanou buněčnou smrt, buněčnou proliferaci, hemostázu, remodelaci extracelulární matrix nebo, prezentaci antigenu[20-25].

2.1.2 Regulace proteolytických enzymů

Proteolytický systém je klíčový pro fyziologické procesy tkání. Potenciální nebezpečí spočívá v neregulované aktivitě proteas a nežádoucím štěpení vitálních proteinů. Regulačních mechanismů existuje celá řada, například na úrovni genové exprese, translace nebo lokalizace proteas[26]. Dále jsou proteasy produkovány ve formě neaktivních zymogenů a k jejich aktivaci dochází proteolytickým odštěpením aktivačního peptidu (propeptidu) až v cílovém místě (Obr. 2.2). Regulace aktivity proteas je také zajištěna pomocí pH, různých kofaktorů (např. iontů, proteinů nebo sacharidů), alosterických modifikátorů nebo endogenních inhibitorů peptidového nebo proteinového charakteru[27-29]. Následující kapitoly se budou zabývat proteasovými inhibitory proteinového charakteru.



Obr. 2.2: Schématické znázornění hlavních mechanismů přirozené regulace aktivity proteasy. *Proteasa (růžově) je často syntetizována ve formě neaktivního zymogenu. Při aktivaci proteasy dochází ke zpřístupnění aktivního místa enzymu proteolytickým odštěpením propeptidu (tmavě šedě) ze struktury zymogenu. Takto aktivovaná proteasa je schopná hydrolyzovat peptidovou vazbu substrátu. Aktivní místo proteasy může být blokováno specifickým inhibitorem (žlutě) za vzniku neaktivního komplexu, který je pro většinu proteinových inhibitorů reverzibilní.*

2.1.3. Proteasové inhibitory proteinového charakteru

Proteasové inhibitory proteinového charakteru, dále jen proteasové inhibitory, se liší svou velikostí od několika kDa (např. inhibitory trypsinu CMTI z tykve, ~3 kDa) až po několik desítek kDa (např. serpin SPN93 z potemníka, 93-kDa)[30,31]. Proteasové inhibitory slouží k regulaci proteolytických enzymů na endogenní i exogenní úrovni. Pokud je narušena rovnováha mezi proteasami a jejich inhibitory, dochází k patologickým stavům[26,32-34]. Jako příklad lze uvést sníženou inhibiční aktivitu cystatinu C, která má za následek zvýšenou aktivitu cysteinových katepsinů K, L, S při ateroskleróze[35].

Proteasové inhibitory proteinového charakteru jsou díky možným aplikacím v medicíně, zemědělství nebo biotechnologiích hojně studovány. Mohou být pro tyto účely použity jako takové nebo může být využit mechanismus jejich interakce s cílovou proteasou pro návrh odvozených syntetických inhibitorů[36]. Příkladem je hirudin z pijavice lékařské (*Hirudo medicinalis)* jako silný inhibitor lidského trombinu, který byl dříve využíván jakožto antikoagulační léčivo a v současné době jsou z něho odvozeny optimalizované deriváty připravované rekombinantní technologií[15]. Další atraktivní aplikací jsou vakcinační strategie využívající proteasové inhibitory patogenů a parazitů, které bývají mezi prvními molekulami interagujícími s hostitelem, a proto představují vhodné kandidátní antigeny[37-39].

2.1.4. Klasifikace proteasových inhibitorů a mechamismus inhibice

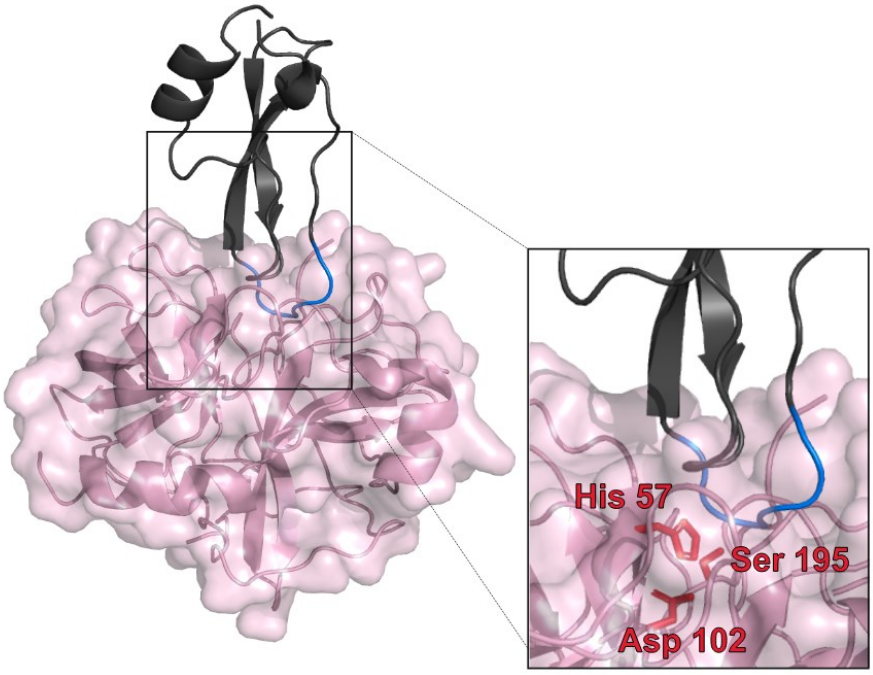
Proteasové inhibitory tvoří velkou a velmi rozmanitou skupinu proteinů, které se liší primární strukturou i 3D architekturou, typem inhibice i povahou komplexů enzym-inhibitor. Lze je klasifikovat do 108 proteinových rodin na základě sekvenční podobnosti[40]. Dále lze proteasové inhibitory dělit podle tříd proteas, které inhibují. Některé proteasové inhibitory jsou specifické pro jednu třídu (např. hirudiny), jiné jsou schopné díky své sekvenční a strukturní variabilitě inhibovat více tříd (makroglobuliny)[41,42]. Dle mechanismu inhibice lze proteasové inhibitory dělit na reverzibilní (např. cystatiny) a ireverzibilní (např. serpiny)[43]. Většina inhibitorů spadá do skupiny reverzibilních inhibitorů[44].

Mezi nejvíce početné rodiny inhibitorů patří cystatiny, které cílí cysteinové proteasy, serpiny, které inhibují serinové a cysteinové proteasy, a inhibitory Kunitzovy

rodiny z rostlin, které blokují aktivitu tří tříd proteas (serinové, cysteinové i aspartátové). Naopak mezi málo početné a prozkoumané rodiny patří chagasiny, které cílí cysteinové proteasy, tyropiny, které inhibují cysteinové a aspartátové proteasy, nebo clitocypiny inhibující cysteinové proteasy[45-47].

Interakce proteinových proteasových inhibitorů s cílovou proteasou vede k tvorbě komplexů, z nichž nejrozšířenější jsou tří následující mechanistické typy: nekovalentní Michaelisovský komplex mimikující interakci enzym-substrát, kovalentní acylenzymový intermediát a nekovalentní komplex enzymu se stéricky bránicím inhibitorem[48].

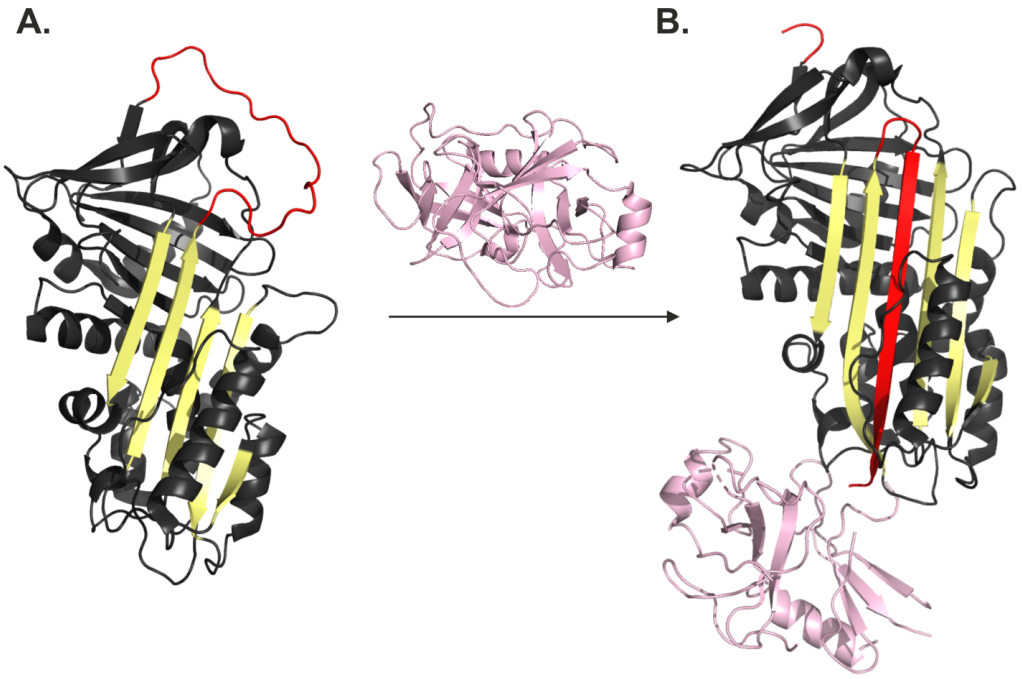
1) Inhibice tvorbou pevného Michaelisovského komplexu proteasy s inhibitorem je zprostředkována vysokoafinitní interakcí inhibitoru s aktivním místem proteasy, kdy se inhibitor váže do aktivního místa enzymu podobným způsobem jako substrát za tvorby komplexu podobného Michaelisovskému komplexu enzym-substrát. Tento typ vazby je charakteristický hlavně pro inhibitory serinových proteas s tzv. kanonickým reaktivním centrem. Ač je tato skupina inhibitorů strukturně velmi rozmanitá, tak má smyčka vázající se na proteasu podobnou kanonickou konformaci, která je komplementární s aktivním místem enzymu[49]. Jako příklad lze uvést inhibici serinových proteas hovězím pankreatickým inhibitorem trypsinu (BPTI) (Obr. 2.3)[50].



***Obr. 2.3.: Mechanismus inhibice proteasy tvorbou pevného* Michaelisovského komplexu.** *Struktura komplexu inhibitoru trypsinu z hovězí slinivky BPTI ve stužkovém modelu (černě) a hovězího chymotrypsinu ve stužkovém modelu se znázorněným povrchem (růžově). Znázorněno je zakotvení reaktivní smyčky inhibitoru (modře) do aktivního místa proteasy způsobem*

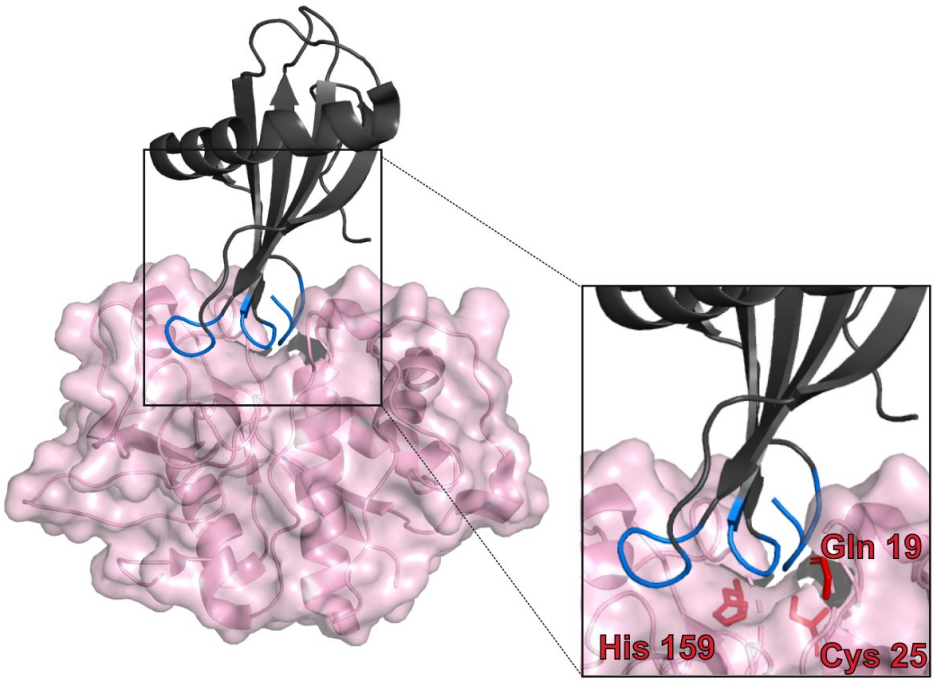
podobným interakci substrátu (P zbytky substrátu / inhibitoru se vážou do S podmíst proteasy). Výřez ukazuje kanonické reaktivní centrum inhibitoru s interakční smyčkou a katalytickou triádou enzymu: His57, Asp102 a Ser195 zvýrazněnou červeně. Upraveno podle (PDB kód: 1CBW).

2) Inhibice tvorbou acylenzymového intermediátu je dynamický proces, který vede ke tvorbě kovalentního komplexu s proteasou. Tímto mechanismem mohou být inhibovány pouze serinové, cysteinové a threoninové proteasy, které během hydrolýzy peptidové vazby tvoří kovalentní intermediát se substrátem[51]. Inhibice je zahájena štěpením dobře přístupné smyčky reaktivního centra (RCL) inhibitoru proteasou, které má za následek konformační změnu inhibitoru a vznik kovalentního komplexu inhibitor-proteasa[52]. Tento tzv. sebevražedný mechanismus využívá rodina serpinů. Jako příklad lze uvést inhibici serinové proteasy lidským α-1-antitrypsinem (Obr. 2.4)[52].



Obr. 2.4: Mechanismus inhibice proteasy tvorbou acylenzymového intermediátu. *A. Stužkový model volného inhibitoru lidského α1-antitrypsinu (černě) s intaktní smyčkou reaktivního centra (RCL, červeně) (PDB kód: 1QLP) a hlavním β-skládaným listem (žlutě). Do interakce vstupuje volná proteasa hovězí trypsin (PDB kód: 1OPH), znázorněná ve stužkovém modelu (růžově). B. Při interakci s proteasou je štěpena RCL za vzniku kovalentního acylesterového intermediátu mezi inhibitorem a proteasou. RCL (červeně) se během konformační změny molekuly vkládá do β-skládaného listu a proteasa je zachycena na druhé straně molekuly v neaktivní formě (PDB kód1EZX). Upraveno podle [52].*

3) Při mechanismu inhibice stérickým bráněním je blokováno aktivní místo proteasy, aniž by inhibitor byl v kontaktu s katalytickým centrem. Typickými zástupci tohoto mechanismu jsou cystatiny, které se vážou do aktivního místa proteasy strukturou dvou smyček a N-konce molekuly[53]. Reaktivní centrum cystatinů tak vytváří klín komplementární s aktivním místem proteasy, čímž blokuje vstup substrátu do aktivního místa. Podobnou strategií fungují také inhibitory z rodin např. tyropinů nebo chagasinů[54,55]. Jako příklad je uvedena inhibice cysteinové proteasy lidským stefinem A (Obr. 2.5)[56].



Obr. 2.5: Mechanismus inhibice proteasy stérickým bráněním. *Struktura komplexu lidského stefinu A ve stužkovém modelu (černě) a hovězího katepsinu H ve stužkovém modelu se znázorněným povrchem (růžově) zobrazuje zakotvení N-koncové části a dvou smyček inhibitoru (modře) do aktivního místa enzymu, které je tak blokováno (PDB kód: 1NB5)* [56]*. Detailní výřez ukazuje stérické bránění vstupu do aktivního místa inhibitorem a katalytickou diádu enzymu Cys 25 a His 195 podporovanou zbytkem Gln 19 (červeně).*

2.1.5. Funkce a patofyziologie proteasových inhibitorů

Proteinové inhibitory jsou významným regulačními molekulami endogenních proteas a při jejich dysregulaci dochází k celé řadě patologických jevů. Abnormalitám proteasových inhibitorů jsou mimo jiné přičítána onemocnění jako myoklonální epilepsie, dědičný angioneurotický edém a Nethertonův syndrom[57-59]. Dále mohou

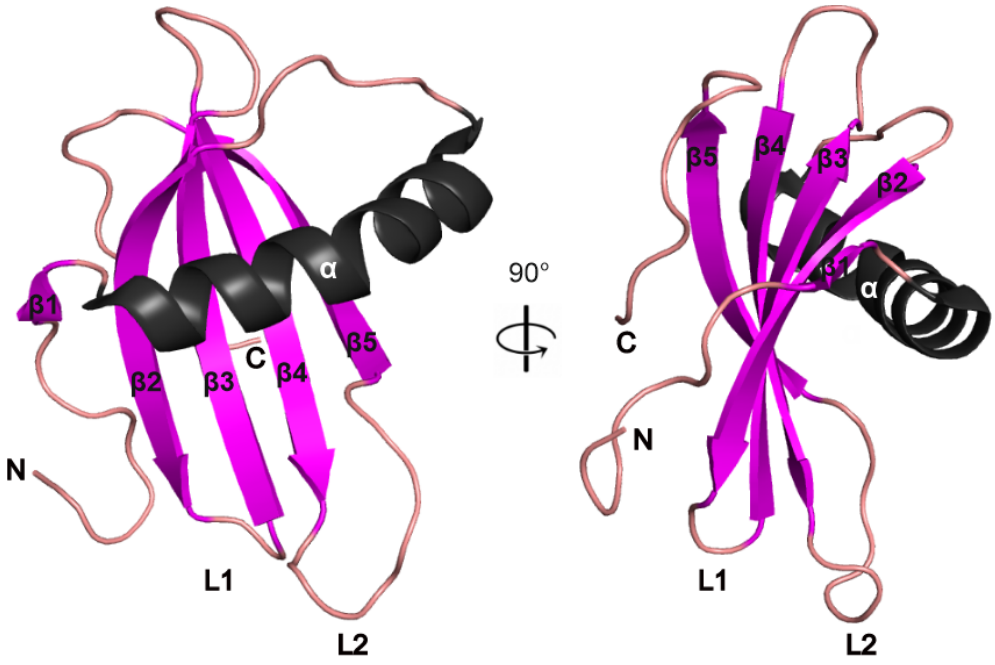
sekretované proteasové inhibitory působit jako obranné molekuly při infekci patogenem nebo parazitem. Chrání tak hostitelský systém před jejich invazivními proteasami[57-59]. Na druhou stranu inhibice proteas exogenními inhibitory mohou využít i parazité a patogeny k inhibici hostitelských proteas a ovlivnění celé řady fyziologických dějů v těle hostitele. Například jsou tak schopni ovlivnit imunitní systém hostitele ve svůj prospěch. O inhibitoru cysteinových proteas, cystatinu parazitárních nematod, bylo prokázáno, že se podílí na snížení proliferace T-buněk a pozměnění funkce makrofágů[60]. Dále inhibitor serinových proteas, serpin parazitárních nematod, zhoršuje funkci granulocytů [61].

2.1.6. Cystatiny

Cystatiny byly prvními objevenými endogenními inhibitory cysteinových proteas a patří mezi nejlépe prozkoumané. Vyskytují se u všech organismů a většina zástupců inhibuje cysteinové proteasy z rodiny C1 papainu a někteří také z rodiny C13 leguminu. Dle MEROPS databáze tvoří cystatiny rodinu I25 a dále se dělí dle počtu domén, přítomnosti disulfidových můstků, glykosylace a buněčné lokalizace na: 1) jednodoménové stefiny (cystatiny prvního typu), které nemají disulfidy, ani glykosylaci a působí intracelulárně; 2) sekretované jednodoménové pravé cystatiny (cystatiny druhého typu), které mají disulfidy, ale nejsou glykosylované a působí extracelulárně a 3) třídoménové kininogeny (cystatiny třetího typu), které jsou glykosylované a vyskytují se intra- i extracelulárně[53,62]. Kininogeny se u savců dále dělí na kininogeny typu H (vysokomolekulární), L (nízkomolekulární) a T (spojené s akutní fází)[53].

2.1.6.1. Struktura a mechanismus inhibice cystatinů

Struktura cystatinové domény je typicky tvořena pěti segmenty antiparalelního β-skládaného listu, které jsou rozmístěny kolem centrální α-šroubovice[53] (Obr. 2.6). Mechanismus inhibice byl poprvé popsán pomocí krystalové struktury komplexu stefinu B s papainem a klasifikován jako stérické bránění (kapitola 2.1.4.) vstupu substrátu do aktivního místa proteasy[63]. Reaktivní centrum inhibitoru tvoří dvě smyčky vystupující z β-skládaného listu a N-konec molekuly, celé centrum tvarem připomíná klín. Cystatiny jsou reverzibilní inhibitory s obvykle širší selektivitou (rozlišují např. katepsiny s endo- versus exopeptidasovou aktivitou) a s inhibičními konstantami v μM až pM koncentracích[64].



Obr. 2.6: Prostorová struktura reprezentativního cystatinu. *Struktura lidskeho stefinu A ve stužkovém modelu je tvořená pěti segmenty antiparalelního β-skládaného listu (fialově), které jsou umístěny kolem centrální α-šroubovice (černě). K interakci s proteasou slouží N-koncová část molekuly a dvě smyčky L1 a L2 vystupující z β-skládaného listu. Pohled na molekulu je ve dvou orientacích (PDB kód 1NB5).*

2.1.6.2. Funkce a patofyziologické změny spojené s cystatiny

Cystatiny regulují fyziologické procesy v organismu a zamezují nežádoucímu a potenciálně destruktivnímu působení cysteinových proteas. Jejich důležitou nespecifickou funkcí je ochrana před náhodně uniklými lysosomálními proteasami. Cystatiny však mají i specializované funkce jako je regulace procesů zánětu, prezentace antigenů, fagocytózy, imunitní odpovědi závislé na T-buňkách, apoptózy, zpracování hormonů a kostní resorpce[65,66]. Dále se účastní komplexní interakce mezi hostitelem a patogenem nebo parazitem, která spočívá v balancování dvou procesů, ve kterých jsou cystatiny zapojeny. Při napadení využívají patogeny nebo parazité své exogenní cystatiny k inhibici hostitelských cysteinových proteas, zejména katepsinů, které jsou spojeny s imunitní odpovědí. A na druhé straně se hostitelé brání svými endogenními cystatiny působení invazivních bakteriálních, virových a parazitárních cysteinových proteas[46].

Patofyziologické změny vznikají z narušení křehké proteolytické rovnováhy mezi cysteinovými proteasami a cystatiny. Lze uvést nedostatečnou regulaci cysteinových

katepsinů cystatiny, která přispívá k invazi a metastázám nádorů např. degradací extracelulární matrix nebo aktivací faktorů podporujících invazivnost[67,68]. Dalšími příklady důsledků dysregulace je bronchiální astma, revmatoidní artritida, osteoartritida, cukrovka a neurodegenerativní poruchy[69,70]. Nežádoucí změny však mohou být spojeny i s defektem samotného cystatinu, např. ukládání plaků oligomerního cystatinu C vede k amyloidóze[71].

2.1.6.3. Významní zástupci cystatinů

Cystatin C je nejhojněji se vyskytujícím širokospektrálním extracelulárním inhibitorem cysteinových proteas u savců. Jedná se o cystatin chránící před únikem cysteinových proteas a s nimi spojených patologií. Lidský cystatin C se používá jako biomarker ledvinových funkcí na základě rychlosti jeho glomerulární filtrace[72]. Vysoké hladiny cystatinu C byly spojeny s rizikem úmrtí a závažných kardiovaskulárních příhod u pacientů s akutním koronárním syndromem[73]. Dále vyšší koncentrace cystatinu C v séru korelovaly s vyšší závažností onemocnění COVID-19 a úmrtností[74]. Nadměrné ukládání cystatinu C v mozkových arteriích ve formě amyloidů je hlavní příčinou hereditární amyloidózy spojené s mutací v genu pro cystatin C, která vede k paralýze, demenci a opakovaným mrtvicím[71].

Cystatin B je dalším ochranným širokospektrálním cystatinem u savců, který chrání před únikem lysosomálních proteas do cytosolu. Vyskytuje se ve formě monomerů, dimerů a tetramerů a vyšších oligomerů. Avšak nadměrná exprese cystatinu B vede k oligomerizaci do nerozpustných struktur a tvorbě amyloidních fibril, které jsou spojovány s neurodegenerativními onemocněními. V jádře zabraňuje štěpení N-koncové části histonu 3 inhibicí katepsinů L a B, čímž reguluje transkripci a buněčný cyklus. Dále se účastní ochrany proti oxidačnímu stresu a jeho deficit vede ke zvýšenému zánětu[75]. Jedním ze vzácných onemocnění způsobených mutací v genu pro cystatin B je progresivní typ myoklonální epilepsie - Unverricht-Lundborgova nemoc[76].

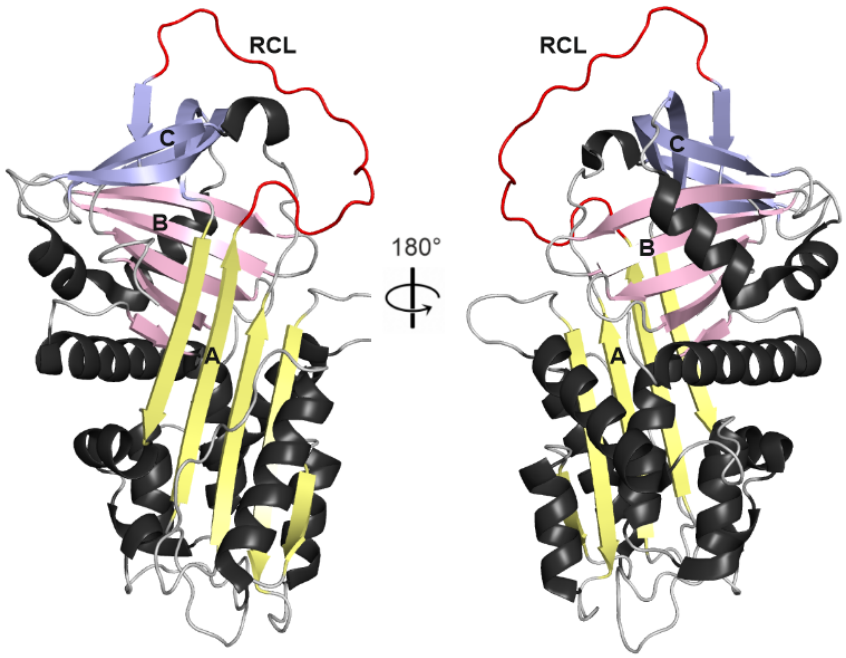
2.1.7. Serpiny

Serpinová rodina je jednou z nejrozsáhlejších skupin proteasových inhibitorů a je rozšířena u všech organismů. Dle MEROPS databáze tvoří rodinu I04. Její zástupci inhibují převážně serinové proteasy, ale mohou inhibovat i proteasy cysteinové. Dále se

mezi serpiny řadí proteiny, které nemají inhibiční aktivitu a slouží jako transportéry hormonů nebo regulátory krevního tlaku[77].

2.1.7.1. Struktura a mechanismus inhibice serpinů

3D architektura molekuly serpinů je velmi konzervativní na rozdíl od relativně variabilní aminokyselinové sekvence (Obr. 2.7). Struktura je tvořena třemi β-skládanými listy (značenými A, B a C) a 7 až 9 α-šroubovicemi. Důležitým strukturním prvkem je variabilní smyčka reaktivního centra (RCL), jejíž délka se liší insercemi nebo delecemi u jednotlivých zástupců pro optimální interakci s cílovou proteasou. Molekulová hmotnost serpinů se pohybuje od 40 do 100 kDa[78]. Častým jevem je u serpinů možnost jemné kontroly inhibičních interakcí prostřednictvím specifických modulátorů, zejména glykosaminoglykanů[79].



Obr. 2.7: Prostorová struktura reprezentativního serpinu. *Struktura α1-antitrypsinu ve stužkovém modelu. Znázorněny jsou α-šroubovice (černě), β-skládané listy A, B a C (žlutě, růžově a fialově) a intaktní smyčka reaktivního centra (RCL, červeně). Molekula je zobrazena ve dvou orientacích. (PDB kód: 1QLP).*

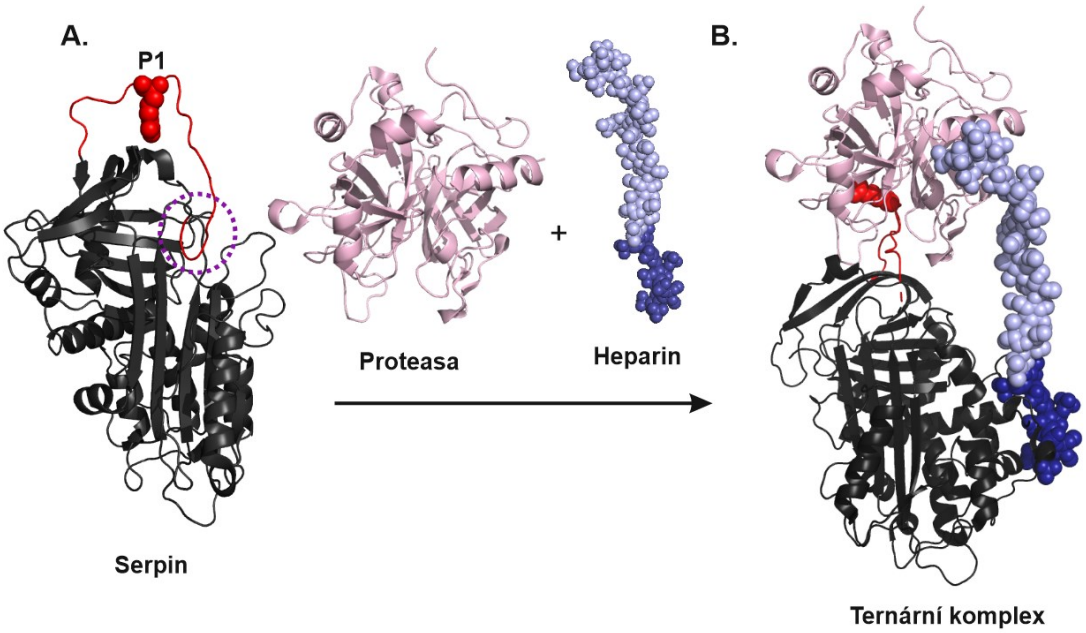
Mechanismus inhibice serpinů označovaný jako „sebevražedný“ je mezi proteasovými inhibitory unikátní. Jedná se o inhibici pomocí tvorby kovalentního acylenzymového intermediátu znázorněnou v kapitole 2.1.4.. Interakce mezi proteasou a serpinem je zahájena štěpením smyčky RCL, která spojuje β-skládaný list A a C. RCL je

exponována do prostoru a je proteasou rozpoznávána jako substrát. Po štěpení peptidové vazby v RCL vzniká kovalentní acylesterový meziprodukt mezi inhibitorem a proteasou a dochází ke konformačnímu přeskupení v molekule serpinu. N-koncová část RCL se vkládá do β-skládaného listu A[52]. Tato konformační změna serpinu se označuje jako přechod ze stresové („stressed“ - S) do energeticky výhodnější relaxované („relaxed“ - R) formy. Proteasa kovalentně vázaná v acylesterovém meziproduktu podléhá významným deformačním změnám, které ji činí katalyticky nefunkční. V některých případech může dojít k rozpadu kovalentního komplexu za vzniku funkční proteasy a inaktivovaného serpinu, ve kterém zůstává štěpená RCL vložena do β-skládaného listu A[47].

Nezreagovaný serpin je v nativní formě v S-konformaci, která je označovaná jako metastabilní stav, který je až na vzácné výjimky za fyziologických podmínek dlouhodobě stabilní. Metastabilita serpinů je klíčová pro úspěšnost inhibice, jelikož energie přechodu do R-konformace je využita k zachycení a deformaci proteasy[80]. V některých případech vede tato konformační nestabilita k jinému způsobu inaktivace serpinu, aniž by došlo k jeho ireverzibilní proteolýze. Intaktní smyčka RCL (bez interakce s proteasou) může být vložena do β-skládaného listu (způsobem podobným jako u R-konformace) za vzniku latentního stavu serpinu. Tento latentní stav se uplatňuje i při tvorbě neaktivních serpinových polymerů, kdy RCL smyčka vstupuje do β-skládaného listu sousední molekuly serpinu[51,78].

2.1.7.2. Regulace inhibiční funkce serpinů pomocí modulátorů

Inhibiční aktivita serpinů může být regulována pomocí dalších makromolekulárních modulátorů. Příkladem takového modulátoru jsou sulfatované polysacharidy, glykosaminoglykany, které zlepšují inhibiční aktivitu u serpinů jako je např. α1-antitrombin, heparinový kofaktor II a protein C inhibitor[81-83]. Samotný α1-antitrombin není účinným inhibitorem serinových proteas trombinu a faktoru Xa. K přechodu do aktivní konformace α1-antitrombinu dochází po navázání pentasacharidového segmentu přítomného v molekule heparinu[52]. Změna konformace vede ke zpřístupnění RCL a odhalení aminokyselinového zbytku v P1 pozici. Heparin dále stabilizuje vzniklý komplex serpinu a proteasy tvorbou ternárního komplexu s oběma partnery (Obr. 2.8). V dalším kroku dochází ke štěpení RCL a inhibičnímu mechanismu popsanému v obrázku 2.4 (str.15)[81].



Obr. 2.8: Modulace inhibiční interakce serpinu s proteasou pomocí glykosaminoglykanu. *A. Prostorová struktura lidského α1-antitrombinu je ve stužkovém modelu (černě, PDB kód 2ANT) s vyznačenou polohou P1 zbytku (červeně postranní řetězec v kuličkovém modelu) na smyčce RCL (červeně). Tečkovaným fialovým kruhem je vyznačená část RCL zanořená do β-skládaného listu. B. Interakce α1-antitrombinu (černě) s trombinem (růžově) je stabilizována glykosaminoglykanem heparinem (modře), který se váže na oba partnery za vzniku ternárního komplexu (PDB kód 1TB6). Pentasacharidový segment (tmavě modře) na heparinu interaguje se specifickým vazebným místem na molekule α1-antitrombinu, což vede ke konformační změně na RCL a zpřístupnění P1 místa*[81].

2.1.7.3. Funkce a patofyziologické změny spojené se serpiny

Serpiny představují klíčové regulátory mnoha fyziologických procesů, včetně krevní koagulace, fibrinolýzy, remodelace chromatinu, imunitní regulace a apoptózy[84,85]. Řada procesů regulovaných serpiny je zajímavá pro patogeny a parazity z hlediska úspěšné invaze. Ti pak svými serpiny posouvají rovnováhu v hostitelském systému ve svůj prospěch a inhibují proteasy souvisejících zejména s hemostázou a imunitní odpovědí hostitele[86,87].

Závislost inhibičního mechanismu serpinů na konformační mobilitě a metastabilním nativním stavu je činí náchylné k mutacím, které vedou k jejich inaktivaci nebo polymeraci[88]. Změny související s polymerací se výrazně projevují u neuronově

specifického serpinu, neuroserpinu a způsobují onemocnění nazvané familiální encefalopatie[89]. Dochází k hromadění neuroserpinových agregátů v neuronech, což má za následek neuronové dysfukce až buněčnou smrt. Toto onemocnění se projevuje narušenými kognitivními i motorickými funkcemi pacientů. Dalšími patofyziologickými změnami způsobenými nedostatkem serpinů jsou např. chronické obstrukční onemocnění plic vyvolané nedostatkem α1-antitrypsinu nebo vznik trombóz způsobený nedostatkem antitrombinu[77,78].

2.1.7.4. Významní zástupci serpinů

Antitrombin je klíčový regulátor krevní koagulace obratlovců. Vrozený nebo získaný deficit tohoto proteinu vede ke zvýšenému riziku trombotických onemocnění. Antitrombin přímo inhibuje serinové proteasy koagulační kaskády - hlavními cílovými proteasami jsou trombin, faktor Xa a faktor IXa. Antitrombin má nejen antikoagulační, ale i protizánětlivé účinky. Prostřednictvím interakce s vaskulárními heparansulfátovými proteoglykany indukuje syntézu prostaglandinu I2 endoteliálními buňkami vedoucí k inhibici exprese prozánětlivých cytokinů. Prostaglandin I2 je navíc silným vasodilatátorem a inhibitorem agregace destiček, čímž nepřímo podporuje antikoagulační funkci antitrombinu[90].

α1-antitrypsin je produkován v játrech a sekretován do krevního řečiště. Hlavní funkcí je inhibice neutrofilní elastasy, serinové proteasy schopné štěpit elastin, který je významnou složkou pojivových tkání. Dále působí protizánětlivě díky modulování chemotaxe a degranulace neutrofilů[91]. Deficit tohoto proteinu vede k řadě plicních onemocnění v důsledku nadměrné aktivity neutrofilní elastasy a prozánětlivých faktorů[92]. Nízké hladiny zvyšují riziko infekce a těžšího průběhu onemocněním COVID-19. α1-antitrypsin také inhibuje hostitelskou proteasu TMPRSS2 nezbytnou pro vstup SARS-CoV-2 do buněk[93]. Polymerace α1-antitrypsinu s mutací Z vede k hromadění nerozpustných polymerních agregátů v hepatocytech a hepatocelulárnímu poškození[94].

Inhibitor aktivátoru plasminogenu 1 řídí fibrinolýzu tím, že inhibuje tkáňový aktivátor plasminogenu a urokinásový aktivátor plasminogenu. Mezi serpiny je význačný svou schopností spontánně přejít do latentního stavu, což je další mechanismus udržení rovnováhy mezi proteasami a serpiny. *In vivo* je tento přechod zpomalen vazbou na plazmatický protein vitronektin a extracelulární matrix[95].

2.1.8. Tyropiny

Proteasové inhibitory rodiny tyropinů jsou součástí velké skupiny proteinů s tyreoglobulinovou doménou typu 1 (Tg1) a nebyly dosud detailně prostudovány. Společným znakem této rodiny je strukturní doména tvořená 65–80 zbytky, která je bohatá na cysteiny. Doména typu 1A nese šest cysteinových zbytků a doména typu 1B má pouze čtyři cysteinové zbytky. Všechny cysteiny tvoří disulfidy s typickým uspořádáním. Typickým rysem v sekvenci je konzervovaný zbytek Gly49 a motiv Cys-Trp-Cys-Val přítomný u většiny zástupců[96]. Tg1 doména byla poprvé popsána v tyreoglobulinu, který je biosyntetickým prekurzorem hormonů štítné žlázy. Tg1 doména se může vyskytovat samostatně nebo v několika opakováních a být součástí multidoménových proteinů[97].

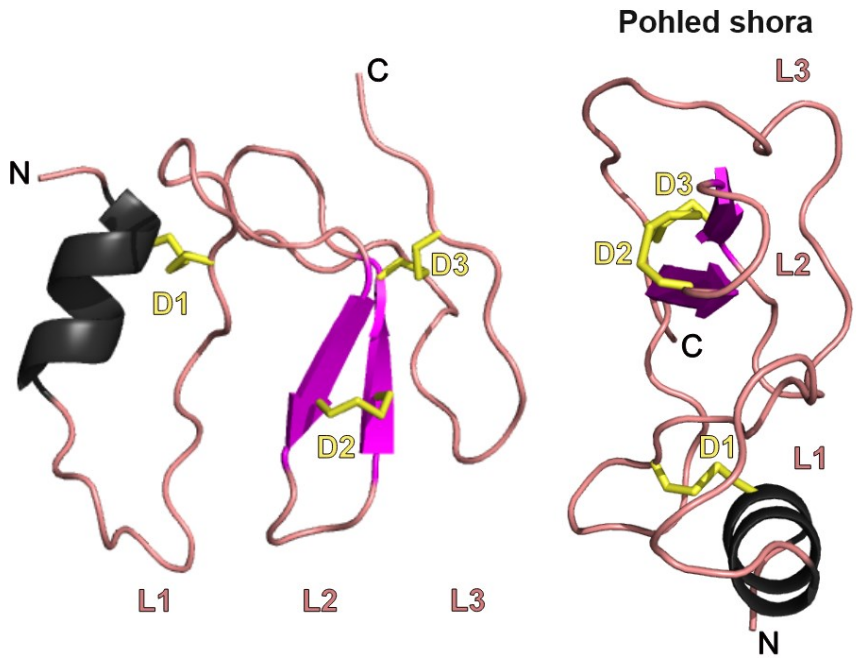
Se zástupci proteinů nesoucích Tg1 doménu se můžeme setkat u bakterií, prvoků, hub a živočichů. U savců se Tg1 doména vyskytuje u proteinů účastnících se vazebných interakcí, např. epiteliálního glykoproteinového antigenu, gastrointestinálního antigenu nebo nidogenu. Přímou účast Tg1 domény z receptoru IGF-BP při vazbě na insulinu podobný růstový faktor (IGF) se podařilo prokázat z NMR struktury komplexu[98]. Další významnou funkcí Tg1 domény je schopnost inhibovat proteasy, která byla studována u fragmentu p41 invariantního řetězce asociovaného s MHC třídou II, testikanu, ekvistatinu, saxifilinu atd.[55,96]. Proteiny obsahující Tg1 doménu s prokázanou inhibiční aktivitou jsou označované jako tyropiny a věnují se jim následující kapitoly.

Tyropiny, řazené dle MEROPS databáze do rodiny I31, jsou specifické inhibitory cysteinových proteas a někteří zástupci inhibují i proteasy aspartátové. Příkladem je třídoménový ekvistatin z mořské sasanky *Actinia equina*, jehož první doména inhibuje cysteinové proteasy a druhá doména aspartátovou proteasu katepsin D[99]. Nejlépe prostudovaným zástupcem této rodiny je fragment p41 díky analýze fyziologické úlohy a známé prostorové struktuře[55,100].

2.1.8.1. Struktura a mechanismus inhibice tyropinů

Krystalová struktura fragmentu p41 v komplexu s cysteinovou proteasou katepsinem L poskytla informaci o prostorovém uspořádání molekuly, která je tvořena N-koncovou α-šroubovicí a antiparalelním β-skládaným listem, a o vazebném módu tyropinů (Obr. 2.9)[55]. K inhibici dochází mechanismem stérického bránění jako u cystatinů (kapitola 2.1.4.). Tyropin se váže do aktivního místa proteasy prostřednictvím

tří interakčních smyček L1 až L3, které jsou stabilizovány třemi disulfidovými můstky, a tím brání vstupu substrátu[55].



Obr. 2.9: Prostorová struktura reprezentativního tyropinu. *Struktura fragmentu p41 (invariantního řetězce asociovaného s MHC třídou II) ve stužkovém modelu je tvořená N-koncovou α-šroubovicí (černě), centrálním β-skládaným listem (fialově) a třemi vazebnými smyčkami (L1, L2 a L3) pro interakci s proteasou. Tři disulfidy jsou znázorněny žlutě (D1-D3), N- a C- konec molekuly jsou vyznačeny. Molekula je zobrazena ve dvou orientacích (PDB kód 1ICF).*

2.1.8.2. Funkce a patofyziologické změny spojené s tyropiny

Biologické funkce tyropinů jsou málo prostudované. Jelikož Tg1 doména bývá součástí multidoménových proteinů, je konkrétní inhibiční role této domény známa jen u několika zástupců. Popsáno je zapojení lidského fragmentu p41 do prezentace antigenů na imunitních buňkách a adhezní molekuly epiteliálních buněk EpCAM do regulace onkogenních signálních drah[101,102]. Tyto fyziologické funkce jsou spojovány s inhibicí cysteinových katepsinů[101,102].

2.1.8.3. Významní zástupci tyropinů

Fragment p41je součástí invariantního řetězce asociovaného s MHC třídy II. Inhibicí proteolytické aktivity katepsinu L v kompartmentech, kde jsou zpracovávány antigeny, chrání imunogenní epitopy proteinů před nadměrnou degradací a umožňuje jejich rozpoznání T-lymfocyty [102]. Fragment p41 je schopen inhibovat celou škálu dalších cysteinových katepsinů (F, K, V a S), které jsou přítomné v buňkách prezentujících antigen, a může tak zřejmě ovlivňovat proces prezentace antigenu komplexněji[100]. Zároveň slouží i jako chaperon zralého katepsinu L, který chrání před degradací jak v antigen prezentujících buňkách, tak po uvolnění do extracelulární matrix[103,104].

Adhezní molekula epiteliálních buněk EpCAM je multidoménový protein, který obsahuje Tg1 doménu zodpovědnou za inhibici katepsinu L. Tento katepsin se účastní procesu invaze nádorů a metastázy. U rakovinných buněk dochází k mutacím EpCAM, které vedou ke ztrátě inhibiční aktivity, a některé mutace také zabraňují správné distribuci EpCAM na povrchu buněk[101].

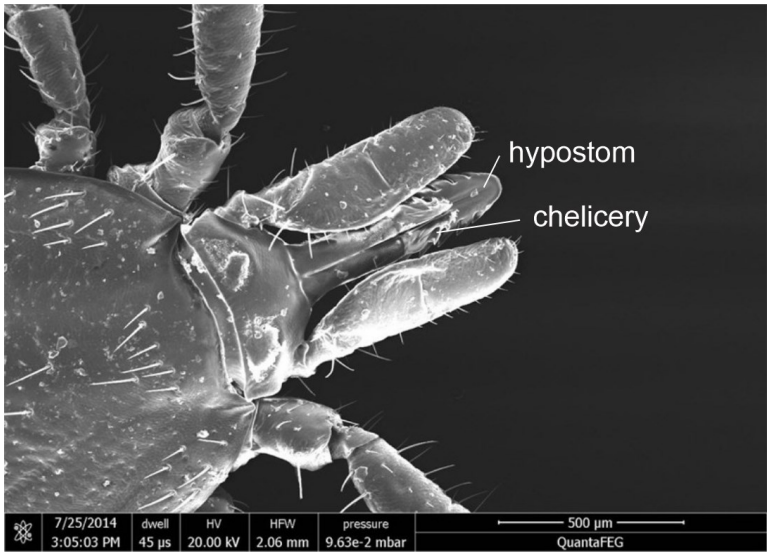
2.2. Klíšťata a jimi přenášené patogeny

2.2.1 Taxonomické zařazení

Říše	živočichové (*Animalia)*
Kmen	členovci (*Arthropoda)*
Podkmen	klepítkatci *(Chelicerata)*
Třída	pavoukovci *(Arachnida)*
Řád	roztoči *(Acari)*
Podřád	klíšťata *(Ixodida)*
Čeleď	klíšťatovití (*Ixodidae*), klíšťákovití (*Argasidae*) a *Nuttalliellidae*

Klíšťata se dělí do tří čeledí *Ixodidae* (klíšťatovití neboli tvrdá klíšťata), *Argasidae* (klíšťákovití neboli měkká klíšťata) a *Nuttalliellidae*, které se liší způsobem života a životním cyklem[105]. Tato práce se věnuje charakterizací proteasových inhibitorů ze slin klíštěte obecného (*Ixodes ricinus)* z čeledi klíšťat a *Ornithodorus moubata* z čeledi klíšťáků.

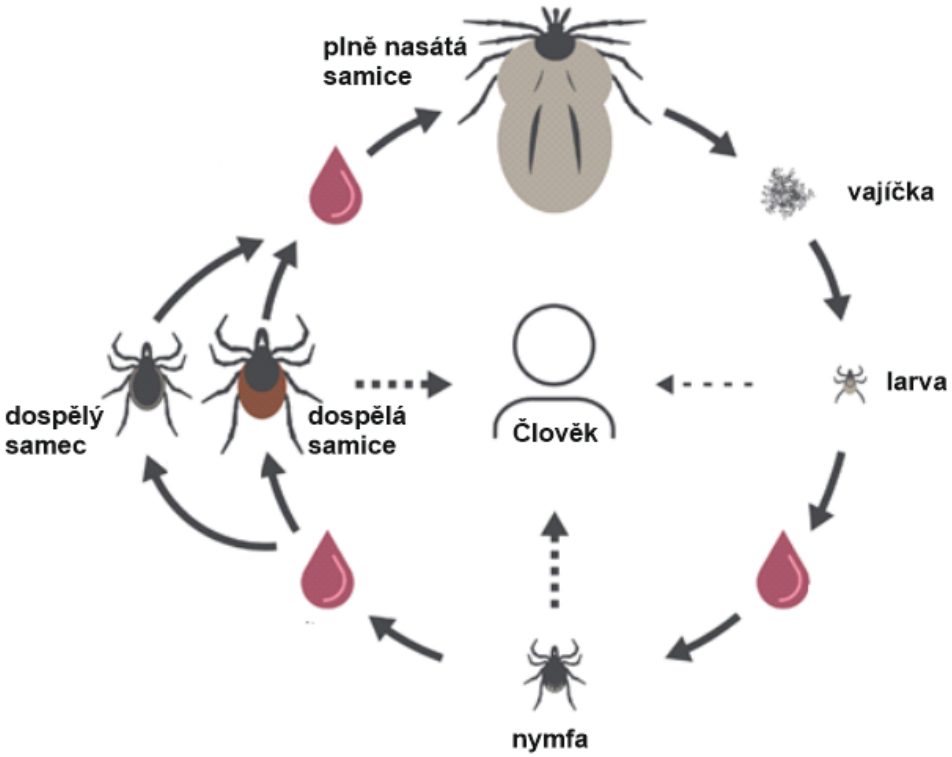
Klíště má na předních končetinách umístěný specializovaný smyslový Hallerův orgán, který má funkci mechanického, chemického a tepelného čidlo pro rozpoznání hostitele. K průniku pokožkou slouží sací ústrojí tvořené chelicery, kterými klíště prořízne kůži, a hypostomem, který vstupuje do tkáně a zafixuje se v ní. Žlábkem na hypostomu klíště vstřikuje sliny do rány (Obr. 2.10).



Obr. 2.10: Sací ústrojí klíštěte obecného (Ixodes ricinus). *Mezi dvěma chelicery je umístěn centrální hypostom pro vstřikování slin do tkáně hostitele. Fotografie z elektronového mikroskopu. Převzato z*[106].

2.2.2. Životní cyklus

Klíšťata i klíšťáci procházejí několika vývojovými stádii - larva, nymfa, dospělý jedinec- a každý přechod je spojen se sáním krve (Obr. 2.11). U klíšťáků je několik nymfálních stádií (4 až 8) oproti jedinému u klíšťat. Dalším rozdílem je, že zatímco samička klíštěte po oplodnění a nakladení vajíček hyne, u klíšťáků není kladení vajíček spojeno se smrtí samičky, která tak může sát i být oplodněna opakovaně[107].



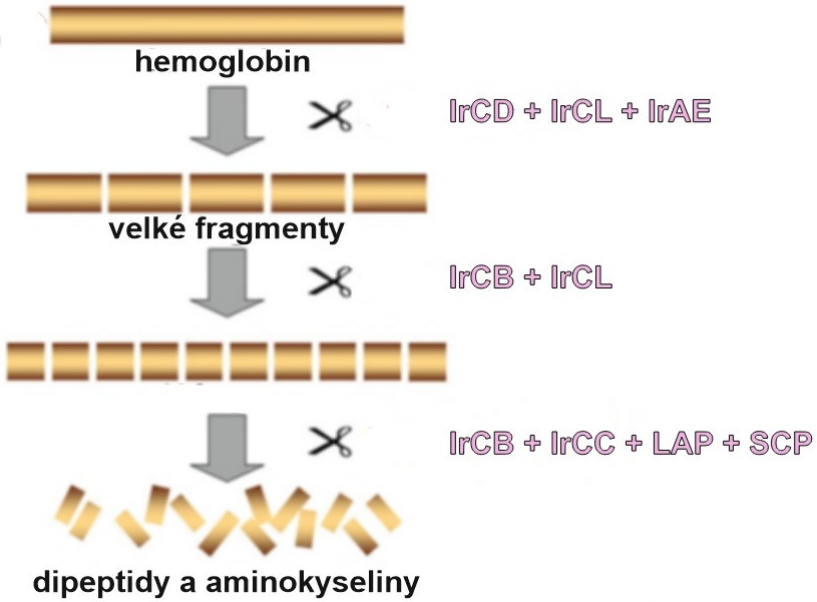
***Obr. 2.11: Schéma životního cyklu klíštěte obecného (Ixodes ricinus)**. Přechod mezi jednotlivými vývojovými stádii klíštěte je spojen s příjmem krevní potravy (červená kapka): larva se přeměňuje na nymfu, která přechází na finální stádium dospělce, kdy na hostiteli saje pouze samice. Plně nasátá samice po oplození klade 3000-5000 vajíček. Všechna sající vývojová stádia jsou schopná parazitovat na člověku. Převzato a upraveno podle*[106].

2.2.3. Proteolytický systém klíšťat

2.2.3.1. Proteasy klíšťat

Klíšťata si osvojila unikátní strategii při získávání krevní potravy ve srovnání s jinými hematofágními členovci. Například komáři proces sání dokončí během několika sekund nebo minut a krevní potravu tráví v lumen střeva s mírně alkalickým pH. Tomu odpovídá i převaha proteas serinové rodiny při trávení[108]. Naproti tomu klíšťata dokážou sát krev hostitele až celé dny, čemuž se přizpůsobují morfologickými změnami, při kterých se několikanásobně zvětší jejich objem. Trávení probíhá intracelulárně v lysozomech střevních buněk, kde je kyselé prostředí a využívají repertoár aspartátových a cysteinových proteas[109]. Primárním zdrojem potravy je hemoglobin uvolněný z erytrocytů a další krevní proteiny. Degradace hemoglobinu byla podrobně popsána u klíštěte obecného (*I. ricinus*). V první fázi je štěpen na velké fragmenty aspartátovou proteasou - katepsinem D, a cysteinovými proteasami - katepsinem L a legumainem

(asparaginyl endopeptidasou). Dále působením cysteinových katepsinů L a B vznikají malé fragmenty, které jsou v dalším kroku degradovány na vstřebatelné dipeptidy a samotné aminokyseliny cysteinovými katepsiny B a C s exopeptidasovou aktivitou, leucinovou aminopeptidasou a serinovou karboxypeptidasou (Obr. 2.12)[110]. Aktivita proteas se dynamicky mění v průběhu sání a dosahuje maxima u plně nasátých jedinců. Nejvíce dominantními proteasami jsou katepsin B a C[111].



Obr. 2.12: Schématické znázornění trávení hemoglobinu u klíštěte obecného (I. ricinus). *Hemoglobin je nejprve degradován na velké fragmenty aspartátovou proteasou katepsinem D (IrCD), a cysteinovými proteasami katepsinem L (IrCL) a legumainem (IrAE). Tyto fragmenty jsou dále štěpeny cysteinovými proteasami katepsinem B (IrCB) a L (IrCL). Finálního štěpení na dipeptidy a aminokyseliny se účastní především cysteinové proteasy katepsin B a C, metaloproteasa leucinová aminopeptidasa (LAP) a serinová karboxypeptidasa (SCP). Převzato a upraveno podle [110].*

Proteasy plní u klíšťat řadu fyziologických funkcí; kromě trávicích proteas ve střevě byly dále studovány proteasy ve slinných žlázách a ovariích[112-114]. Účastní se také embryonálního vývoje. Vitelin, zásobní protein vaječného žloutku, je štěpen u klíštěte *Rhipicephalus microplus* pomocí tzv. vitelin degradující cysteinové endopeptidasy (VTDCE)[115] a u klíštěte *Haemaphysalis longicornis* katepsinem B a D[116]. Proteasy jsou součástí vrozené imunitní odpovědi na infekční patogeny, která je spojená s generováním antimikrobiálních peptidů, aktivací fagocytózou, enkapsulací, tvorbou melaninu a koagulací hemolymfy[117]. Ochranné antimikrobiální peptidy označované hemocidiny jsou

generovány z molekuly hemoglobinu kombinovaným působením aspartátové proteasy BmAP a cysteinové proteasy BmCL1 z klíštěte *R. microplus*[118].

2.2.3.2 Klíštěcí proteasové inhibitory proteinového charakteru

Endogenní proteasové inhibitory mají širokou distribuci v tkáních klíštěte a jsou zapojeny do různých fyziologických drah včetně trávení, vrozené imunity, srážení hemolymfy a tvorby vajíček[119,120]. Často plní i více funkcí, např. vícedoménový Kunitzův inhibitor KPI z klíštěte *Dermacentor variabilis* reguluje trávení ve střevě a zároveň působí bakteriocidně[121]. Intenzivně je zkoumáno exogenní působení klíštěcích proteasových inhibitorů při regulaci hostitelských proteas. Proteasové inhibitory jsou mezi prvními molekulami, které se ve slinách klíštěte dostávají do těla hostitele a usnadňují proces sání. Při kontaktu parazita s hostitelem dochází k lokálnímu poškození epidermis, dermis a krevních cév hostitele. Takové poškození by vedlo ke srážení krve, agregaci krevních destiček, zánětlivé a adaptivní imunitní reakci hostitele a procesu hojení, což by potlačilo úspěšnou invazi parazita a vedlo k zamezení sání. Obranná reakce hostitele se opírá o proteolytické dráhy zahrnující prokoagulanty (trombin, koagulační faktory), prozánětlivé enzymy (neutrofilní elastázu, proteinázu 3, chymázu, tryptázu, kalikrein a katepsiny L, B, S, C a G, atd.) a enzymy komplementu[122-124]. Tyto skupiny hostitelských proteas dokáže klíště blokovat díky široké škále proteasových inhibitorů[125-127]. Mapování těchto inhibitorů u 26 druhů klíšťat odhalilo zástupce z 18 různých rodin proteasových inhibitorů. Většina z nich reguluje aktivitu serinových proteas, menší počet představují inhibitory cysteinových proteas a metaloproteas[128].

Kunitzovy inhibitory

Proteiny z Kunitzovy rodiny (klasifikované jako rodina I2 dle MEROPS databáze) patří k nejhojněji zastoupeným proteasovým inhibitorům ve slinách klíšťat[129]. Většinou inhibují serinové proteasy trypsinového typu. Studována byla inhibice proteas koagulační kaskády a proteas účastnících se agregace krevních destiček, zejména trombinu, faktorů Xa a XIIa a elastasy[130]. Kunitzův inhibitor Ir-CPI exprimovaný ve slinách *I. ricinus* specificky interaguje s faktory XIIa a XIa a kalikreinem a brání srážení krve[131]. Kunitzovy inhibitory využívají pro interakci s proteasami kanonické reaktivní centrum s konzervativní strukturou a lokalizací na struktuře inhibitoru (Obr. 2.3). U několika klíštěcích Kunitzových inhibitorů byla také identifikována existence

nekanonických reaktivních center jako u Ornithodorinu z *O. moubata*, který inhibuje trombin, a u Ixolarisu z *I. scapularis*, který je inhibitorem faktoru Xa[132,133]. Tyto inhibitory mohou využívat pro interakci i exomísta na povrchu cílových proteas. Dlouho byly Kunitzovy inhibitory klíšťat považovány výhradně za antihemostatické proteiny, nově se objevují studie o imunomodulačních, zejména protizánětlivých účincích. IrSPI z *I. ricinus* inhibuje elastázu a moduluje tak zánětlivou odpověď potlačením sekrece prozánětlivých cytokinů a adaptivní imunitní odpověď potlačením proliferace CD4+ T lymfocytů[134].

Serpiny

Serpiny (rodina I04 dle MEROPS databáze) jsou dalšími hojně zastoupenými proteasovými inhibitory klíšťat studovanými zejména ve slinných žlázách, střevě a ovariích. V genomu *I. scapularis* bylo nalezeno 45 serpinů[135]. Klíštěcí serpiny jsou většinou širokospektrální inhibitory inhibující více proteas účastnících se určitého biologického procesu jako je imunitní odpověď a hemostáze. Serpin IRS-2 z *I. ricinus*, který byl studován v této disertační práci, je prvním parazitárním serpinem s vyřešenou 3D strukturou. Jeho unikátní inhibiční specifita kombinující inhibici serinových proteas chymotrypsinového a trypsinového typu umožňuje protizánětlivé působení[136]. Řada serpinů je cílena na hemostatický systém hostitele. Dobře prostudovaný serpin19 z *Amblyomma americanum* inhibuje většinu faktorů krevního srážení, a tím účinně potlačuje srážení krve[137].

Cystatiny

Cystatiny (rodina I25 dle MEROPS databáze) jsou nejrozšířenějšími a nejvíce studovanými inhibitory cysteinových proteas u klíšťat. Patří mezi ně cystatiny typu 1 (stefiny) a hlavně typu 2 (pravé cystatiny). Pravé cystatiny plní buď endogenní funkce, nebo exogenní funkce při interakci s hostitelem. Nejlépe prostudovaným endogenním cystatinem je mialostatin z *I. ricinus* lokalizovaný ve střevě klíšťete, kde pravděpodobně kontroluje trávicí proteolytický systém[138]. Cystatiny s exogenní funkcí jsou sekretované do slin a mají převážně imunomodulační funkce, interferují s prezentací antigenu, zánětlivou odpovědí a proliferací imunitních buněk v tkáni hostitele[119,139]. Sialostatin L z *I. scapularis* a cystatin OmC2 z *O. moubata*, který je popisovaný v této disertační práci, jsou první klíštěcí cystatiny, které při vakcinačních experimentech výrazně omezily schopnost klíšťat sát krev nebo jejich vývoj[140,141]. Mají silné imunosupresivní vlastnosti

spojené s inhibicí cysteinových katepsinů. Byly i prvními strukturně charakterizovanými parazitárními cystatiny[140,141]. Slinné cystatiny z klíšťat rodu *Ixodes* mají unikátní inhibiční specifitu cílenou na cysteinové katepsiny s endopeptidasovou aktivitou a tato vysoká selektivita jim zřejmě umožňuje efektivní manipulaci imunitní odpovědí hostitele při dlouhodobém sání[138,140,142].

2.2.4. Onemocnění přenášená klíšťaty

Klíšťata představují vektory široké škály patogenů patřících mezi viry, bakterie a prvoky, které do hostitele přecházejí spolu s klíštěcími slinami. Imunomodulační vlastnosti bioaktivních látek ve slinách klíšťat posilují infekčnost přenášených patogenů. Z epidemiologického hlediska jsou klíšťata a komáři nejdůležitějšími parazitárními přenašeči infekčních nemocí[143]. Tabulka 1 přináší přehled hlavních patogenů, klíštěcích vektorů a geografické distribuce způsobovaných onemocnění[143]. Tradiční metody kontroly klíšťat jsou akaricidy a repelenty, ty mají však svá omezení, včetně obav o životní prostředí a vznik rezistence klíšťat[105]. Dále existují pro některá onemocnění vakcíny na bázi antigenů z přenášených patogenů, např. v ČR jsou dostupné dvě vakcíny proti klíšťové encefalitidě FSME-IMMUN a Encepur[144]. Existují i závažná onemocnění jako Lymská borelióza, proti kterým prevence dosud neexistuje. Proto je v současné době věnována značná pozornost vývoji vakcín směřovaných přímo proti klíštěcímu vektoru. Cílem těchto tzv. protiklíštěcích vakcín je potlačení sání klíštěte na hostiteli, a tím i snížení rizika přenosu patogenů. Tato disertační práce přispívá k identifikaci a charakterizaci potenciálních kandidátních antigenů pro tento typ vakcín, které představují racionální volbu z hlediska nákladové efektivity i ekologické bezpečnosti[13].

Tab. 1. Přehled onemocnění přenášených klíšťaty na člověka. Upraveno podle[143].

Onemocnění	Patogeny	Vektory	Rozšíření
Africká klíšťová horečka	*Rickettsia africae*	*Amblyomma hebraeum*, *A. variegatum*	Afrika, Indie
Horečka Skalistých hor	*Rickettsia rickettsii*	*Amblyomma americanum*, *A. aureolatum*, *A. cajennense*, *Dermacentor andersoni*, *D. variabilis*, *Rhipicephalus sanguineus*	Severní a Jižní Amerika
Lidská granulocytická anaplazmóza	*Anaplasma phagocytophilum*	*Haemaphysalis concinna*, *H. punctata*, *Ixodes ricinus*, *I. pacificus*, *I. scapularis*, *Rhipicephalus bursa*	Evropa, Severní Amerika
Lidská monocytická ehrlichióza	*Ehrlichia chaffeensis*	*Amblyomma americanum*	Severní Amerika
Lymská borelióza	*Borrelia burgdorferi*	*Ixodes hexagonus*, *I. pacificus*, *I. persulcatus*, *I. ricinus*, *I. scapularis*	Asie, Evropa, Severní Amerika
Návratná horečka	*Borrelia* spp.	*Ornithodoros* spp.	Afrika, Asie, Austrálie, Evropa, Severní Amerika
Q horečka	*Coxiella burnetii*	Různé druhy	Afrika, Asie, Austrálie, Evropa, Severní Amerika
Skvrnitý tyfus	*Rickettsia conorii*	*Rhipicephalus sanguineus*, *R. turanicus*	Afrika, Asie, Evropa
Tularemie	*Francisella tularensis*	Různé druhy	Asie, Evropa, Severní Amerika
Babesióza	*Babesia divergens*, *B. microti*	*Ixodes ricinus*, *I. scapularis*	Evropa, Severní Amerika
Klíšťová coloradská horečka	Coltivirus	*Dermacentor andersoni*	Severní Amerika
Klíšťová encefalitida	Flavivirus	*Ixodes persulcatus*, *I. ricinus*, *Haemaphysalis concinna*, *H. punctata*	Asie, Evropa
Krymsko-konžská hemoragická horečka	Naiovirus	*Amblyomma variegatum*, *Hyalomma anatolicum*, *H. punctata*, *H. marginatum*, *H. truncatum*, *Rhipicephalus bursa*	Afrika, Asie, Evropa
Nemoc Kyasanurského lesa	Flavivirus	*Haemaphysalis spinigera*, *H. turturis*	Indie
„Louping ill“	Flavivirus	*Ixodes ricinus*	Evropa
Omská hemoragická horečka	Flavivirus	*Dermacentor marginatus*, *D. reticulatus*, *Ixodes persulcatus*	Asie
Powassanská encefalitida	Flavivirus	*Dermacentor andersoni*, *Haemaphysalis longicornis*, *Ixodes cookei*, *I. scapularis*	Asie, Severní Amerika

3. Cíle disertační práce

Proteasové inhibitory proteinového charakteru jsou významnými přirozenými regulátory proteolytických procesů. Pro parazity mají kritický význam při jejich interakci s hostitelem na několika úrovních, což z nich činí potenciální cílové molekuly pro intervenci. Disertační práce se zaměřuje na nové proteasové inhibitory ze slin klíšťat, které jsou injikovány do tkáně hostitele. Jako zdroj byly zvoleny dva významné druhy klíšťat: *Ixodes ricinus* jako vektor klíšťové encefality a Lymské boreliózy a *Ornithodoros moubata* jako vektor návratné horečky a afrického prasečího moru. Obecným cílem práce je strukturně-funkční biochemická charakterizace vybraných zástupců tří rodin proteasových inhibitorů a využití výsledků k objasnění jejich biologické role.

Dílčí cíle této práce jsou následující:

1) Identifikovat tři vybrané zástupce proteasových inhibitorů ze tříd cystatinů, serpinů a tyropinů ve slinách klíšťat a připravit je jako rekombinantní proteiny.

2) Určit inhibiční specifitu studovaných inhibitorů k fyziologicky relevantním proteasám, které jsou regulovány ve tkáni hostitele.

3) Podílet se na určení prostorové struktury studovaných inhibitorů, zejména přípravou difraktujících krystalů pro rentgenostrukturní analýzu. Interpretovat 3D strukturu a popsat, jak stavba reaktivního centra inhibitorů určuje jejich inhibiční aktivitu.

4) S využitím získaných informací o strukturních a funkčních vlastnostech studovaných inhibitorů vyvodit závěry o jejich fyziologickém působení při interakci klíštěte s hostitelem.

4. Materiál a metodika

Tato kapitola stručně shrnuje základní metodiku, vybavení a materiály použité v této disertační práci. Detailní popis použitých materiálů a metodiky je vždy součástí přiložených publikací.

4.1. Materiál a laboratorní vybavení

Většina dat byla získána s využitím vybavení a přístrojů v laboratořích Ústavu organické chemie a biochemie AV ČR (ÚOCHB). Projekt byl vypracován ve spolupráci s Parazitologickým ústavem Biologického centra AV ČR v Českých Budějovicích (ParÚ). Řešení krystalových a NMR struktur bylo provedeno ve spolupráci s laboratoří Dr. Řezáčové ÚOCHB. NMR spektra byla měřena na přístroji 850 MHz Bruker Avance spectrometer, ÚOCHB. Difrakční data pro řešení krystalografických struktur byla získána na synchrotronu na pracovišti EMBL-Hamburg, DESY (Deutsches Elektronen-Synchrotron), Hamburg, Německo.

Fluorogenní peptidové substráty a nízkomolekulární inhibitory byly komerčního původu. Proteolytické enzymy použité při testování inhibitorů z klíšťat byly komerční kromě katepsinů C, H, K a V, které byly připraveny na ÚOCHB. Proteinové inhibitory OmC2 a IRS-2 byly připraveny na ParÚ.

4.2. Metodika

Základní metody použité v přiložených publikacích jsou následující:

4.2.1. Metody molekulární biologie:

Celkově bylo připraveno 6 konstruktů tyropinu IrThy a jeho domén IrThy-Cd a IrThy-Nd pro expresi jednak v hmyzích buňkách (Schneider S2 buňky) pro biochemické charakterizace a jednak v *E.coli* (SHuffle T7 Competent buňky) pro strukturní charakterizaci. Pro NMR analýzu byla rekombinantní exprese proteinů optimalizována v systému *E.coli* za použití izotopově ($^{13}C/^{15}N$) značeného média.

4.2.2. Biochemické metody:

Hlavní použité metody byly následující: (1) elektroforetická separace proteinů pomocí SDS-PAGE, (2) přenos proteinů na PVDF membránu metodou Western blot, (3) chromatografická purifikace proteinů pomocí systému FPLC, (4) určení koncentrace proteinů a peptidů pomocí aminokyselinové analýzy, (5) analýza proteinů hmotnostní spektrometrií a N-koncovým sekvenováním, (6) příprava extraktů klíštěcích tkání.

4.2.3. Enzymologické metody:

Hlavní použité metody byly následující: měření aktivity enzymů pomocí fluorogenních a FRET (Fluroescence Resonance Energy Transfer) substrátů pomocí fluorescenční čtečky (Tecan), stanovení kinetických parametrů proteas pro substráty (K_m) a inhibičních konstant pro klíštěcí inhibitory (IC_{50}, K_i) za použití nelineární regrese v programu Grafit.

4.2.4. Metody strukturní biologie:

Krystalizace klíštěcích inhibitorů byla provedena ve dvou krocích (1) vysokokapacitní testování sad komerčních krystalizačních roztoků v 96-jamkových destičkách metodou sedící kapky, (2) optimalizování úspěšné krystalizační podmínky ve visící kapce. Řešení 3D struktur rentgenostrukturní analýzou bylo provedeno metodou molekulárního nahrazení. 3D struktury byly analyzovány a graficky upraveny pomocí programů CCP4, PyMol, CorelDraw.

5. Výsledky

Výsledky této disertační práce jsou shrnuty v celkem čtyřech publikovaných pracích (kapitola 5.1.-5.4.). Publikace uvedené v této disertační práci mají za cíl biochemicky a strukturně charakterizovat proteinové proteasové inhibitory ze slin klíštěte *Ixodes ricinus* a klíštěte *Ornithodorus moubata*. V této části uvádím publikace s krátkým souhrnem rekapitulujícím výsledky.

Seznam publikací

Publikace č.1: Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick Ornithodoros moubata.
Salát J, Paesen GC, Řezáčová P, Kotsyfakis M, **Kovářová Z**, Šanda M, Majtán J, Grunclová L, Horká H, Andersen JF, Brynda J, Horn M, Nunn MA, Kopáček P, Kopecký J, Mareš M.
Biochem J. 2010 Jul 1;429(1):103-12.

Publikace č.2: A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation
Chmelař J, Oliveira CJ, Řezáčová P, Francischetti IM, **Kovářová Z**, Pejler G, Kopáček P, Ribeiro JM, Mareš M, Kopecký J, Kotsyfakis M.
Blood. 2011 Jan 13;117(2):736-44.

Publikace č.3: Crystallization and diffraction analysis of the serpin IRS-2 from the hard tick *Ixodes ricinus*
Kovářová Z, Chmelař J, Šanda M, Brynda J, Mareš M, Řezáčová P.
Acta Crystallogr Sect F Struct Biol Cryst Commun. 2010 Nov 1;66(Pt 11):1453-7.

Publikace č.4: An Unusual Two-Domain Tyropin from Tick Saliva: NMR Solution Structure and Highly Selective Inhibition of Cysteine Cathepsins Modulated by Glycosaminoglycans
Matoušková Z, Orsághová K, Srb P, Pytelková J, Kukačka Z, Buša M, Hajdušek O, Šíma R, Fábry M, Novák P, Horn M, Kopáček P, Mareš M.
Int J Mol Sci. 2024 Feb 13;25(4):2240.

5.1. Publikace č. 1: Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick *Ornithodoros moubata*.

Klíšťák *Ornithodoros moubata* je druh klíštěte, které se vyskytuje převážně v Africe a je známé jako vektor řady onemocnění, jako je například návratná horečka způsobená spirochétou *Borrelia duttoni*. Dále přenáší virus afrického moru prasat, závažné onemocnění hospodářských zvířat způsobující smrtelné hemoragie. K přenosu patogenů a nerušenému sání krve hostitele pomáhají bioaktivní molekuly obsažené ve slinách klíšťat. Obecně jsou sliny krev sajících parazitů bohaté na inhibitory proteas, ale dosud nebyly detailně studované u tohoto druhu klíštěte.

Studie je zaměřena na inhibitor proteas z *O. moubata*, který patří do rodiny cystatinů a byl označen OmC2. Prvním krokem byla úspěšná identifikace OmC2 ve slinách klíštěte pomocí proteomické analýzy hmotnostní spektrometrií. Dále byl připraven OmC2 ve formě rekombinantního proteinu, u kterého byla testována inhibiční specifita. Ta byla cílena výlučně na cysteinové proteasy klanu CA a nebyla pozorována inhibiční aktivita vůči jiným významným proteasám, jako jsou cysteinové proteasy klanu CD nebo aspartátové a serinové proteasy. U detailněji studovaného klanu CA byla zjištěna inhibiční interakce jak s endopeptidasami (savčími katepsiny L a S a archetypálním papainem), tak exopeptidasami (savčími katepsiny B, C a H). Inhibiční účinnost OmC2 se v případě papainu, katepsinu L, S a C pohybuje v oblasti subnanomolárních koncentrací, v případě katepsinu B a H v nanomolární oblasti. OmC2 je tedy vysoce účinným širokospektrálním inhibitorem. Schopnost OmC2 inhibovat exopeptidasy jej odlišuje od homologického slinného cystatinu z klíštěte *Ixodes scapularis*, sialostatinu, který není schopen katepsin B a H inhibovat vůbec a je 300krát horším inhibitorem katepsinu C. Endopeptidasy inhibuje sialostatin s podobnou účinností jako OmC2.

Dalším krokem byla strukturní charakterizace OmC2 rentgenostrukturní analýzou. Podařilo se najít podmínky pro přípravu krystalu difraktujícího do 2,45 Å a prostorovou strukturu vyřešit metodou molekulárního nahrazení. Molekula OmC2 se skládá z pěti segmentů β-skládaného listu v antiparalelní orientaci, které jsou rozmístěné kolem centrální α-šroubovice. Celá struktura je stabilizována dvěma konzervovanými disulfidy. K interakci cystatinů s aktivním místem proteasy slouží obecně dvě smyčky vystupující z β-skládaného listu (označované jako L1 a L2) a N-koncová část proteinu.

Právě detailní stavbou N-konce a smyčky L2 se OmC2 liší od sialostatinu a tento rozdíl je zřejmě zodpovědný za odlišnou inhibiční specifitu. Širokou inhibiční specifitou podobnou OmC2 se vyznačují i některé cystatiny obratlovců, jako hojně rozšířený cystatin C, a při srovnání sekvence N-konce a smyčky L2 je patrná jejich značná strukturní podobnost. Klíště by tak pomocí OmC2 mohlo mimikovat cystatiny hostitele, a tím měnit proteolytickou rovnováhu při regulaci hostitelských katepsinů, o kterých je známo, že hrají významnou roli při imunitní odpovědi.

Dále práce studovala biologickou aktivitu OmC2 pomocí testů na hostitelských imunitních buňkách a pomocí vakcinačních experimentů na myších. V prvním typu testů se sledoval vliv OmC2 na produkci cytokinů dendritickými buňkami, které hrají důležitou roli v imunitní odpovědi na antigen pomocí cytokinů a prezentováním antigenu T-lymfocytům. V přítomnosti OmC2 došlo ke snížení produkce prozánětlivých cytokinů TNFα a IL-12. Dále se zkoumal vliv OmC2 na proliferaci T-lymfocytů indukovanou dendritickými buňkami prezentujícími specifický antigen. V tomto testu rozvoje adaptivní imunitní odpovědi došlo k statisticky významnému snížení proliferace naivních CD4+ T-lymfocytů v přítomnosti OmC2. V druhém typu testů byl sledován rozdíl v sání nymf *O. moubata* na myších imunizovaných OmC2 a kontrolních myších. U nymf parazitujících na imunizovaných myších bylo pozorováno omezení sání a schopnosti transformace na další stádium životního cyklu.

Získané výsledky ukazují, že OmC2 dokáže modulovat hostitelský imunitní systém, a vyhnout se tak obranné reakci hostitele během interakce s parazitem. Tato schopnost je zásadní pro úspěšné sání klíštěte na hostiteli. Poznatky z této studie mohou být využity při přípravě tzv. protiklíštěcí vakcíny proti *O. moubata* potlačující sání klíšťat na hostiteli a omezující šíření přenášených onemocnění. Analýza struktury a biochemické aktivity OmC2 významně přispěla k pochopení mechanismu účinku tohoto bioaktivního proteasového inhibitoru ve slinách klíštěte.

Můj podíl na práci spočíval v: 1) návrhu a provádění biochemických funkčních testů s OmC2, 2) nalezení a optimalizaci krystalizačních podmínek OmC2 a příspěvku k následné rentgenostrukturní analýze, 3) interpretaci získané 3D struktury OmC2 s využitím molekulární grafiky a korelace vztahů mezi strukturou a aktivitou.

BJ www.biochemj.org

BJ Structure

Biochem. J. (2010) **429**, 103–112 (Printed in Great Britain) doi:10.1042/BJ20100280

online data

103

Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick *Ornithodoros moubata*

Jiří SALÁT*†[1], Guido C. PAESEN‡, Pavlína ŘEZÁČOVÁ§||, Michalis KOTSYFAKIS*¶, Zuzana KOVÁŘOVÁ§**, Miloslav ŠANDA§, Juraj MAJTÁN‡, Lenka GRUNCLOVÁ*†, Helena HORKÁ*†, John F. ANDERSEN¶, Jiří BRYNDA§||, Martin HORN§, Miles A. NUNN‡, Petr KOPÁČEK*†, Jan KOPECKÝ*† and Michael MAREŠ§[1]

*Institute of Parasitology, Biology Centre of the Academy of Sciences of the Czech Republic, Branišovská 31, 37005 České Budějovice, Czech Republic, †Faculty of Sciences, University of South Bohemia, Branišovská 31, 37005 České Budějovice, Czech Republic, ‡Centre for Ecology and Hydrology, Mansfield Road, Oxford OX1 3SR, U.K., §Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610 Praha 6, Czech Republic, ||Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610 Praha 6, Czech Republic, ¶Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, U.S.A., and **Department of Biochemistry, Faculty of Science, Charles University in Prague, Hlavova 8, 12843 Praha 2, Czech Republic

The saliva of blood-feeding parasites is a rich source of peptidase inhibitors that help to overcome the host's defence during host–parasite interactions. Using proteomic analysis, the cystatin OmC2 was demonstrated in the saliva of the soft tick *Ornithodoros moubata*, an important disease vector transmitting African swine fever virus and the spirochaete *Borrelia duttoni*. A structural, biochemical and biological characterization of this peptidase inhibitor was undertaken in the present study. Recombinant OmC2 was screened against a panel of physiologically relevant peptidases and was found to be an effective broad-specificity inhibitor of cysteine cathepsins, including endopeptidases (cathepsins L and S) and exopeptidases (cathepsins B, C and H). The crystal structure of OmC2 was determined at a resolution of 2.45 Å (1 Å = 0.1 nm) and was used to describe the structure–inhibitory activity relationship. The biological impact of OmC2 was demonstrated both *in vitro* and *in vivo*. OmC2 affected the function of antigen-presenting mouse dendritic cells by reducing the production of the pro-inflammatory cytokines tumour necrosis factor α and interleukin-12, and proliferation of antigen-specific $CD4^+$ T-cells. This suggests that OmC2 may suppress the host's adaptive immune response. Immunization of mice with OmC2 significantly suppressed the survival of *O. moubata* in infestation experiments. We conclude that OmC2 is a promising target for the development of a novel anti-tick vaccine to control *O. moubata* populations and combat the spread of associated diseases.

Key words: cathepsin, cystatin, *Ornithodoros moubata*, parasite, peptidase inhibitor.

INTRODUCTION

Ticks (Ixodida) are blood-feeding parasites that transmit many pathogens of medical and veterinary importance. The order Ixodida has two main families, the Ixodidae (hard ticks) and Argasidae (soft ticks), which differ in many aspects of their biology. In the hard ticks, the adult female engorges only once (for several days), lays a batch of eggs and then dies. In the soft ticks, the female lays a batch of eggs after each feeding episode, which takes minutes, and typically repeats the cycle several times. The soft tick *Ornithodoros moubata* is widely distributed throughout drier parts of south and east Africa. It is an important disease vector, since it transmits the spirochaete *Borrelia duttoni*, causing African tick-borne relapsing fever in humans, and African swine fever virus, which causes a highly lethal haemorrhagic disease of domestic swine [1,2].

Successful feeding of ticks depends on a cocktail of salivary proteins with antihaemostatic, anti-inflammatory and immunomodulatory properties that are injected into the host [3]. The tick feeding site, which is highly modified by these pharmacologically active molecules, is the place where transmission of tick-borne pathogens occurs (for a recent review, see [4]). Much of the current vector–host research focuses on the discovery of novel components of tick saliva and on the description of their physiological functions, as this may aid the development of vaccines to control tick populations and the diseases they transmit. Moreover, compounds of tick saliva have clear therapeutic potential, which have, in some cases, been evaluated in animal models [5].

The saliva of ticks and other blood-feeding parasites contains a wide variety of peptidase inhibitors belonging to several structural families. Inhibitors from the cystatin family are small proteins that interact with and block the active site of cysteine peptidases. Cystatins are found in all living organisms and regulate various biological processes, including the immune response (for a review, see [6]). They have been shown to play an important role in parasite–host interactions. Immunomodulatory effects were ascribed to cystatins of parasitic nematodes, e.g. *Onchocerca volvulus* [7], which inhibit antigen processing associated with MHC II antigen presentation. Such interference reduces the proliferation of human peripheral-blood mononuclear cells stimulated with a specific antigen [8].

Cystatins were found in the transcriptome of the salivary glands (sialome) of several hard tick species [9,10]. Silencing of a cystatin gene in the hard tick *Amblyomma americanum* by means of RNA interference significantly reduced the tick's ability to feed

Abbreviations used: CID, collision-induced dissociation; DC, dendritic cell; IL-12, interleukin-12; LC, liquid chromatography; LPS, lipopolysaccharide; NCS, non-crystallographic symmetry; OVA, ovalbumin; RMSD, root mean square deviation; TNFα, tumour necrosis factor α, T-PBS, 0.05% Tween 20 in PBS.

[1] Correspondence may be addressed to either of these authors (email george@paru.cas.cz or mares@uochb.cas.cz).

The structural co-ordinates reported for salivary cystatin from *Ornithodoros moubata* will appear in the Protein Data Bank under accession code 3L0R.

successfully [11]. Two salivary cystatins (sialostatins L and L2) have been characterized in the hard tick *Ixodes scapularis* [12]. They displayed high affinity for cathepsins L and S, which are cysteine peptidases that play an important role in the processing of antigens by DCs (dendritic cells). Subsequent studies showed that maturation of DCs is significantly impaired in the presence of sialostatin L [13], and that vaccination against sialostatin L2 can lead to decreased feeding success of *I. scapularis* [14].

In the present study, we focused on OmC2, a cystatin expressed in the salivary glands of the soft tick *O. moubata* [15]. First, we showed that OmC2 is secreted by the glands to form part of the saliva, using a proteomic approach. Secondly, we determined the crystal structure of OmC2, providing a structural basis for understanding its inhibitory specificity. Thirdly, we demonstrated the immunomodulatory effect of OmC2 on antigen-presenting cells, and evaluated the protein's vaccine potential using an animal model.

EXPERIMENTAL

Animals

C57BL/6 and C3H/HeN female mice were from Charles River Laboratories, and transgenic female mice OT-II were from the Jackson Laboratory. Mice aged 7–8 weeks were used. Animals were housed at 22 °C and a relative humidity of 65 %. Experimental procedures complied with the rules of the European Union and institutional guidelines on the use of experimental animals. *O. moubata* ticks were maintained in an established laboratory colony at the Institute of Parasitology, České Budějovice, Czech Republic, using natural feeding on mice.

Tick saliva preparation

Saliva was obtained from fasting *O. moubata* females after injecting 1 μl of a 1 % solution of pilocarpine hydrochloride in PBS into their genital pores. After stimulation, saliva was collected from the mouthpart using capillary tubes and stored at − 80 °C.

MS proteomic analysis

LC (liquid chromatography)-MS/MS analysis was performed on a hybrid mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific) coupled to a two-dimensional capillary LC system (Rheos 2000; Flux Instruments). The first-dimension column was a monolithic PS-DVB (200 μm in diameter × 10 mm long; Dionex) and the second-dimension column was a C_{18} PepMap 100 (75 μm in diameter × 150 mm long × 3 μm pore size; Dionex) with gradient elution in a 0.1 % formic acid/acetonitrile system. The identification of OmC2 in tick saliva was accomplished in two steps. In the first step, the LC-MS/MS profile was obtained from a tryptic digest of recombinant OmC2. The DDA (data-dependent analysis) experiment consisted of one full MS scan (at a resolution of 60 000), and three MS/MS CID (collision-induced dissociation) experiments of the most intensive ions from the full-scan spectra. In the second step, a tryptic digest of the protein fraction of tick saliva was searched for the presence of OmC2-derived peptides. This analysis was limited to the major indicative peptides detected in the first step. The applied method consisted of one full MS scan (at a resolution of 60 000), and five MS/MS CID experiments with high-resolution MS analysis of fragments in the Orbitrap (at a resolution of 7500). The following parameters were used for the MS/MS experiments: collision energy, 35 eV; isolation width, 1; activation Q, 250; and activation time, 250 ms. Identification of the selected peptides (Supplementary Table S1 at http://www.BiochemJ.org/bj/429/bj4290103add.htm) was confirmed by the analysis of retention time, mass accuracy and CID fragmentation spectra. The mass data were processed by Bioworks (Thermo Fisher Scientific) and Peaks (Bioinformatics Solutions) software.

Expression and purification of OmC2

OmC2 was expressed in Sf9 insect cells (Gibco) using the *flash*BAC™ one-step baculovirus protein expression system (Oxford Expression Technologies). Briefly, OmC2 cDNA (GenBank® accession no. AY547735) prepared as described previously [15] was cloned into the transfer vector pBacPAK9. Primers (forward 5′-GGAGGATCCATGTCAAGTTTTAAGG-TGG-3′ and reverse 5′-GGAGCGGCCGCCTAATGGTGAT-GGTGATGGTGTCCCTCACATCTGTATGACGTG-3′) were designed to include the tick signal sequence for export from the cell and to add a C-terminal oligohistidine tag (-Gly-His_6) to aid purification. The NotI and BamHI restriction sites were used for cloning into the transfer vector. The construct was amplified in XL1-Blue cells (Stratagene). Baculovirus containing the *OmC2* gene was prepared by homologous recombination between the *flash*BAC and transfer vector DNA following the manufacturer's instructions. For expression, Sf9 cells were infected with high-titre baculovirus-construct virus stock, in a shaker incubator at 28 °C. The recombinant protein was purified from the expression medium by heparin (HiTrap™ Heparin HP; GE Healthcare) and Co^{2+} (TALON® Superflow™ Metal Affinity Resin; Clontech) affinity chromatography using the manufacturer's protocol. This was followed by size-exclusion chromatography, using a Superdex™ 75 10/300 GL column (GE Healthcare) and running buffer (15 mM Hepes, pH 7.2, and 200 mM NaCl); OmC2 eluted close to the cytochrome *c* standard (12.4 kDa), indicating that it was monomeric. The product was concentrated with a 5 kDa cut-off ultrafiltration unit (Sartorius). Purified protein was tested for endotoxin contamination by means of Limulus Amobecyte Lysate QCL-1000® (Lonza) following the manufacturer's instructions. Endotoxin contamination did not exceed 0.1 EU (endotoxin units)/ml in any of the samples used for immunological assays. The purified protein was stored at − 80 °C.

Enzymatic assays

Inhibitory activity of recombinant cystatins OmC2 and sialostatin L (prepared as described in [16]) was determined by continuous peptidase-activity assays using specific fluorigenic substrates. The cysteine peptidases tested were as follows: human cathepsins H, S and L and human legumain (Calbiochem); and papain and bovine cathepsins B and C (Sigma–Aldrich). The peptidase substrates applied were the following: *N*-carbobenzyloxy-Leu-Arg-7-amino-4-methylcoumarin (R&D Systems) for cathepsins B and L and papain; *N*-carbobenzyloxy-Val-Val-Arg-7-amino-4-methylcoumarin, H-Gly-Arg-7-amino-4-methylcoumarin and *N*-carbobenzyloxy-Ala-Ala-Asn-7-amino-4-methylcoumarin (Bachem Bioscience) for cathepsins S and C and legumain respectively; and H-Arg-7-amino-4-methylcoumarin (Calbiochem) for cathepsin H. The assay system contained 0.25 mM substrate; cysteine peptidases were applied at the following concentrations: 0.04 nM papain, 1.0 nM legumain, 0.025 nM cathepsin L, 0.05 nM cathepsin S, 0.04 nM cathepsin C, 1.5 nM cathepsin H and 1.3 nM cathepsin B. The assay buffer used was 0.1 M sodium acetate buffer, pH 5.5 (or 5.0 for legumain), 0.1 M NaCl, 1 mM EDTA, 1 mg/ml cysteine

and 0.005 % Triton X-100. For aspartic and serine peptidases, these assay conditions were modified as described in [16]. Apparent inhibition constants were determined essentially as described in [16] by measuring the loss of enzymatic activity at increasing concentrations of inhibitor in the presence of a substrate in large excess. Briefly, each enzyme was pre-incubated for 10 min with the inhibitor, and subsequently the assay was initiated and developed by the addition of the corresponding substrate at 30 °C. The linear hydrolysis rate of the substrate was followed for 20 min in a Spectramax Gemini XPS 96-well plate fluorescence reader (Molecular Devices) using $\lambda_{ex} = 365$ nm and $\lambda_{em} = 450$ nm with a cut-off at 435 nm. All experiments were performed in triplicate. Inhibition constants were calculated by non-linear regression.

Crystallization and data collection

Screening for crystallization conditions was performed using the Crystallization Basic and Extension Kits (Sigma–Aldrich) by the hanging-drop vapour-diffusion technique. Preliminary crystals were obtained in 0.1 M Hepes buffer, pH 7.5, and 1.5 M lithium sulfate. Optimal crystals of OmC2 were prepared at 20 °C using the hanging-drop vapour-diffusion technique in 24-well Nextal plates (Qiagen). The crystallization drop consisted of 1.25 μl of the OmC2 protein solution (3.5 mg/ml in 20 mM Hepes buffer, pH 7.2) and 0.75 μl of the reservoir solution (0.1 M Hepes buffer, pH 7.5, and 1.5 M lithium sulfate). Crystals shaped as bi-pyramids reached their final size of 0.35 mm×0.15 mm×0.15 mm within 4 days.

For data collection, crystals were soaked gradually in reservoir solution supplemented with 10, 15 and 20 % (v/v) glycerol and flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K using the X12 EMBL beamline at DESY (Deutsches Elektronen-Synchrotron), Hamburg, Germany, and processed using the HKL-2000 suite of programs [17]. Crystals exhibited the symmetry of space group $P3_121$ and contained two molecules in the asymmetric unit. Crystal parameters and data-collection statistics are given in Supplementary Table S2 at http://www.BiochemJ.org/bj/429/bj4290103add.htm.

Structure determination

The phase problem was solved by molecular replacement using the MolRep program [18]. The search model was derived from the structure of sialostatin L2 from *I. scapularis* (PDB code 3LH4; to be released) sharing 37 % sequence identity with OmC2. Model refinement was carried out using the program REFMAC 5.3 [19] from the CCP4 package [20]. Manual building was performed using Coot software [21]. Tight NCS (non-crystallographic symmetry) restraints were applied during initial refinement; at later stages, NCS restraints were loosened as guided by the behaviour of R_{free}. The final refinement steps included TLS (translation/libration/screw) refinement [22]. The quality of the final models was validated with Molprobity software [23]. Final refinement statistics are given in Supplementary Table S2. Figures showing structural representations were prepared with the program PyMOL (DeLano Scientific; http://www.pymol.org). The DALI [24] and PISA [25] servers were used to analyse the crystal structure.

Isolation of immune cells

DCs

The spleen of a C57BL/6 mouse was dissected, minced with scissors, digested in RPMI 1640 medium containing 1 mg/ml collagenase D (Roche) at 37 °C for 30 min and passed through a 70 μm nylon cell strainer (BD Falcon). DCs were isolated using magnetic beads conjugated with anti-CD11c (N418) antibody and MACS® Column (Miltenyi Biotec) separation following the manufacturer's instructions. The purity of isolated DCs (~90 % $CD11c^+$ cells) was evaluated by flow cytometry.

$CD4^+$ T-cells

Whole splenocytes from OT-II mice were obtained by mechanical disruption of the spleen and washed three times in RPMI 1640 medium. Thereafter, $CD4^+$ T-cells were isolated by immunomagnetic separation using the Dynal Mouse CD4 Negative Isolation Kit (Invitrogen) following the instructions of the manufacturer. The purity of isolated $CD4^+$ T-cells (~90 % $CD4^+$ cells) was evaluated by flow cytometry (see below).

Quantification of cytokine production

Purified spleen DCs were cultured in 96-well plates, at a density of 10^5 cells per well, in 200 μl of culture medium [RPMI 1640 supplemented with 10 % (v/v) heat-inactivated fetal bovine serum, 50 μM 2-mercaptoethanol, 100 μg/ml penicillin and 100 units/ml streptomycin]. OmC2 was added to wells in 1–30 μg/ml final concentrations. After a 2 h incubation (37 °C, 5 % CO_2), DCs were activated with LPS (lipopolysaccharide) (50 ng/ml). Cell-free supernatant samples were collected 3 and 48 h later and analysed for the presence of TNFα (tumour necrosis factor α) and IL-12 (interleukin-12) respectively using Ready-SET-Go!™ ELISA kits (eBiosciences) following the manufacturer's instructions. Samples were tested in triplicate.

Antigen-specific $CD4^+$ T-cell proliferation

DCs (3×10^4/well) isolated from the spleen of C57BL/6 mouse were pre-incubated (4 h) with culture medium (see above) in the presence or absence of OmC2 (1–30 μg/ml). After the pre-incubation period, $CD4^+$ T-cells (2×10^5/well) from OT-II mouse were added to the culture; thereafter, medium with or without EndoGrade™ ovalbumin (Profos AG) (10 μg/ml) was added. The culture was incubated for 72 h at 37 °C and 5 % CO_2. Proliferation was assessed by adding Cell Counting Kit-8 reagent (Fluka) at 5 % of the incubation volume during the last 4 h of the cultivation and evaluated by reading the A_{490} of the supernatant. Samples were tested in triplicate.

Flow cytometry

Samples (0.5×10^6 cells) were incubated with CD4-specific monoclonal antibody (conjugated with FITC) and CD11c-specific monoclonal antibody (conjugated with Alexa Fluor® 488) for CD4 T-cells and DCs respectively. Isotype control antibodies were used for appropriate controls. All reagents were obtained from eBiosciences. Labelled cell samples were analysed on an Epics XL flow cytometer (Coulter) equipped with a 15 mW argon-ion laser with excitation capabilities at 488 nm using System II software (Coulter).

Vaccination with OmC2 and tick feeding success

A group of six C3H/HeN mice was vaccinated by two intraperitoneal injections of OmC2 supplemented with Freund's adjuvant (Sigma–Aldrich). Briefly, 20 μg of OmC2 was administered with complete Freund's adjuvant. The same dose of OmC2 was delivered with incomplete Freund's adjuvant 2 weeks later. Samples of blood were taken from vaccinated mice 10 days

later and the presence of specific anti-OmC2 antibodies was assessed by ELISA. A control group of six C3H/HeN mice was vaccinated with ovalbumin (Sigma–Aldrich) using the same vaccination protocol. Thereafter, the ability of ticks to feed on these animals was tested. The mice were placed into small cages to reduce the possibility of movement, and a plastic tube containing six or seven *O. moubata* nymphs (at the first nymphal stage) was fixed to the tail of each mouse using Parafilm® M (Sigma–Aldrich). Whether engorgement occurred or not was assessed after 1 h and monitored as the percentage of nymphs that were able to feed. Nymphs that fed were weighed. Survival of engorged nymphs to the next nymphal stage was assessed after 14 days.

Determination of specific anti-OmC2 antibody titre

ELISA-based systems were used to detect specific anti-OmC2 antibodies in the sera of immunized and control mice. A flat-bottomed plate (Nunc) was coated with OmC2. A 100 μl volume of coating buffer containing OmC2 (10 μg/ml) was added to each well and incubated overnight at 4 °C. After washing two times with T-PBS (0.05 % Tween 20 in PBS), unoccupied sites were blocked with sample buffer [10 % (v/v) newborn bovine serum in PBS] for 30 min at 37 °C. The plate was washed three times with T-PBS and wells were incubated with serial dilutions of the sera in sample buffer (100 μl final volumes) for 1 h at 37 °C. Subsequently the plate was washed three times with T-PBS and incubated with peroxidase-labelled goat anti-mouse antibodies (Sigma–Aldrich) diluted 1:1000 in sample buffer (100 μl/well) for 45 min at 37 °C. After washing three times with T-PBS, an enzymatic colour reaction was generated using *o*-phenylenediamine (Sigma–Aldrich) following the manufacturer's instructions. Enzymatic reactions were stopped after 10 min by adding 50 μl of 2 M H_2SO_4. The A_{490} was measured with an ELISA Multiskan MCC340 spectrophotometer (Labsystems). Specific anti-OmC2 antibody titres were determined as the reciprocal value of the serum dilution that gave an A_{490} reading twice that of the negative control.

Statistical analysis

Differences in cytokine production and proliferation of T-cells were analysed using Student's *t* test. The homogeneity of variances assumption was tested by the Levene test. Results of OmC2 affected groups were compared with corresponding controls and *P* values of 0.05 or less were considered significant. Results of tick feeding success and survival were compared by Fisher's exact test. The link between the ability of nymphs to transform to the next stage and the anti-OmC2 antibody titre of infested mice was tested by the Spearman rank correlation (percentage data were arcsine-transformed). In addition to conventional statistical significance, the effect size was reported. The effect size measured the magnitude of a treatment effect independent of sample size. Cohen's *d*, which represents the difference between means divided by S.D., was used. The Cohen's *d* was manually computed from the *t* value of the *t* test and the *r* value of the Spearman rank correlation. Small, medium and large effects for *d* were defined as 0.2, 0.5 and 0.8 [26]. All statistical analyses were conducted using Statistica 8.0 software (StatSoft).

RESULTS

Proteomic identification of OmC2 in tick saliva

Saliva collected from *O. moubata* adult females was subjected to proteomic analysis to directly determine the presence or absence of OmC2. The applied LC-MS/MS strategy is based on enzymatic digestion of a complex protein mixture and MS/MS peptide sequencing (see the Experimental section). This analysis provided ~53 % peptide coverage of the OmC2 sequence and a mass accuracy of <5 p.p.m. (Supplementary Table S1), allowing us to conclude that OmC2 is secreted in the saliva of *O. moubata*.

Preparation of recombinant OmC2

The complete cDNA of OmC2 (GenBank® accession no. AY547735) contains one open reading frame coding for 129 amino acids (including a 19-residue signal sequence), followed by a 1071 bp 3′-untranslated region. OmC2 was prepared in insect cells as a recombinant protein with an oligohistidine-tag added to its C-terminus using the baculovirus expression system. Predicted pI and molecular mass values of the protein including the tag were 6.2 and 13 115 Da respectively. The protein was purified to homogeneity from the expression medium using a three-step chromatographic procedure (see the Experimental section). Purified OmC2 migrated as a single band of 13 kDa on reducing SDS/PAGE. The identity of the purified protein was confirmed by LC-MS/MS analysis (peptide coverage 100 %).

Inhibitory specificity of OmC2

The purified recombinant OmC2 was screened *in vitro* for its inhibitory activity against a panel of cysteine peptidases. The inhibitory profile of OmC2 is summarized in Table 1, where it is compared with that of sialostatin L, a salivary cystatin from the hard tick *I. scapularis*.

In a first step, we demonstrated that both cystatins interact with peptidases of the CA but not the CD clan of the cysteine peptidase class. This was tested with papain and legumain, which are representative members of the CA and CD clan respectively and archetypes in cystatin research (Table 1) [27]. The activity of aspartic (cathepsin D) or serine (cathepsin G and trypsin) peptidases was not affected by either tick cystatin (results not shown).

In a second step, a set of papain-type peptidases (family C1, clan CA) of mammalian origin was screened, including: cathepsins L and S (endopeptidases), cathepsin B (a peptidyl dipeptidase and endopeptidase), cathepsin C (dipeptidyl peptidase I) and cathepsin H (an aminopeptidase). These enzymes were selected to cover a wide range of endo- and exo-peptidase cleavage specificities. OmC2 inhibited all of these peptidases, with IC_{50} values ranging from approx. 0.15 to 8.9 nM. Inhibition of endopeptidase cathepsins L and S by OmC2 was similar to that of sialostatin L (subnanomolar IC_{50} values), whereas OmC2 was a far more potent inhibitor than sialostatin L of the exopeptidases cathepsins B, C and H (Table 1). Three hundred times more sialostatin L than OmC2 was needed to inhibit cathepsin C based on IC_{50} values of ~314 and ~1.07 nM respectively. Moreover, even in high molar excess, sialostatin L did not block cathepsins B and H, whereas nanomolar IC_{50} values characterized the interaction of OmC2 with these peptidases.

Three-dimensional structure of OmC2

The crystal structure of OmC2 was determined by molecular replacement using the structure of sialostatin L2 as a search model and refined using data to 2.45 Å (1 Å = 0.1 nm) resolution (Supplementary Table S2). The hexagonal crystal form contains two molecules in the asymmetric unit with a relatively high solvent content of 65.4 %. All protein residues could be modelled into a well-defined electron density map (Supplementary Figure S1 at http://www.BiochemJ.org/bj/429/bj4290103add.htm) with the exception of the first two N-terminal residues (Thr^1 and Ser^2). Owing to its inherent disorder, the C-terminal oligohistidine

Table 1 Inhibitory effect of OmC2 on the activity of cysteine peptidases

The inhibitory potency of OmC2 from the soft tick *O. moubata* is compared with that of sialostatin L (SialoL) from the hard tick *I. scapularis*. The IC_{50} values (means ± S.E.M.) were determined by the peptidase activity assays using specific fluorigenic substrates. The MEROPS database classification (http://merops.sanger.ac.uk) of the cysteine peptidases tested (clan) and their specificity (activity mode) are given. NI, no significant inhibition at 10 μM inhibitor concentration.

Enzyme	Enzyme specificity, *classification*	IC_{50} of OmC2 (nM)	IC_{50} of SialoL (nM)
Papain	Endopeptidase, *CA*	0.164 ± 0.003	0.548 ± 0.037
Cathepsin L	Endopeptidase, *CA*	0.146 ± 0.016	0.178 ± 0.005
Cathepsin S	Endopeptidase, *CA*	0.149 ± 0.015	0.559 ± 0.026
Cathepsin C	Dipeptidyl peptidase, *CA*	1.07 ± 0.12	314 ± 5
Cathepsin H	Aminopeptidase, *CA*	1.21 ± 0.01	NI
Cathepsin B	Peptidyl dipeptidase/ endopeptidase, *CA*	8.81 ± 0.52	NI
Legumain	Endopeptidase, *CD*	NI	NI

tag (except for the proximal residue) is also lacking in the crystallographic model. The final OmC2 model consists of two molecules of OmC2, each containing 108 protein residues. The RMSD (root mean square deviation) for superposition of the main-chain atoms of the two molecules is 0.62 Å, a value within the range observed for different crystal structures of identical proteins [28]. Minor structural changes are localized in the loop regions, which are exposed to solvent and/or involved in crystal contacts (residues 68–72, 48–52 and 95–99).

Figure 1(a) shows the overall structure of OmC2. The molecule adopts a typical cystatin fold similar to that of vertebrate homologues, characterized by a five-stranded twisted antiparallel β-sheet, which wraps around an α-helix. OmC2 contains two conserved disulfide bridges connecting Cys^{65} with Cys^{77}, and Cys^{88} with Cys^{108} (Figure 2a). The sequence alignment together with a structural comparison with known cystatin structures clearly demonstrated that OmC2 belongs to family 2 of the cystatin superfamily (Figure 2). The closest structural homologue of OmC2 is sialostatin L2 from the hard tick *I. scapularis*, the structure of which was recently determined (PDB code 3LH4; to be released): the RMSD is 1.25 Å (for 99 aligned Cα atoms), and the sequence identity is 37%. A lower structural similarity was found to family 2 cystatins from vertebrates, namely human cystatin F (PDB code 2CH9; RMSD ~2.0 Å), chicken egg-white cystatin (PDB code 1YVB; RMSD ~2.6 Å), and human cystatin D (PDB code 1RN7; RMSD ~2.8 Å). The sequence identity of OmC2 with representative members of this group is in the range 13–24% (Figure 2a), with the highest values found for salivary S-type cystatins (represented by human cystatin SN in Figure 2a).

Interaction of the family 2 cystatins with papain-type peptidases is mediated by three regions, the N-terminal segment and two hairpin loops L1 and L2 (Figure 2a), which form a tripartite wedge-shaped edge that binds to the enzyme active-site cleft [29] (Figure 1b). The first part of the binding site is formed by the N-terminal segment extending to a totally conserved Gly^5 residue (Figure 2a). The orientation of this region in OmC2 suggests conformational flexibility, as seen in vertebrate cystatins (Figure 2b). Thus Gly^5 can function as a hinge that allows the flexible N-terminal segment to adopt a conformation that is optimal for target enzyme binding. The loop L1 of OmC2 (between β2 and β3) is quite similar in conformation to other cystatin structures and contains the conserved sequence motif Gln-Xaa-Val-Xaa-Gly (Figure 2a). The loop L2 (between β4 and β5) is characterized in OmC2 and vertebrate cystatins by the presence of conservative Pro-Trp residues; this motif is, however,

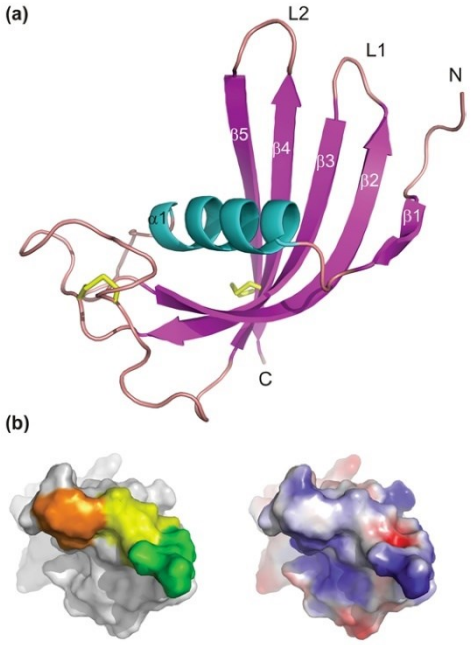


Figure 1 Crystal structure of OmC2

(**a**) A cartoon representation of the protein main chain coloured by secondary-structure elements (α1, cyan; β1–5, magenta). The N- and C-termini and two disulfide bridges, Cys^{65}–Cys^{77} and Cys^{88}–Cys^{108} (yellow sticks), are indicated. The hairpin loops L1 and L2 and the N-terminus of cystatins are involved in the binding of papain-type peptidases. (**b**) Top view of the solvent-accessible surface of the OmC2 molecule showing the binding site for papain-type peptidases. In the left-hand image, the N-terminus and L1 and L2 loops are coloured green, yellow and orange, respectively. In the right-hand image, the surface is coloured by its electrostatic potential displayed at a scale from −2 kT (red) to +2 kT (blue).

lacking in sialostatins, where the L2 loop is shortened by a one-residue deletion (Figure 2a). Furthermore, in contrast with the other cystatins depicted in Figure 2(b), the N-terminal segment of Sialostatin L2 is tightly packed against the β-sheet and stabilized there by a net of interactions.

Several members of the cystatin family 2 are able to inhibit legumain, and the legumain-binding site was localized at critical residue Asn^{31} [27]. This functional motif is absent from the OmC2 structure (Figure 2a), which is in line with the fact that OmC2 does not suppress legumain activity (Table 1).

Effects of OmC2 on host immune cells

OmC2 decreases levels of TNFα and IL-12 produced by LPS- activated DCs

The influence of OmC2 on cytokine production of LPS-activated DCs was tested. DCs are an important subset of immunocompetent cells that typically initiate the development of the innate and adaptive immune responses by producing cytokines and presenting foreign antigen to T-cells. LPS (a component of the outer membrane of Gram-negative bacteria) activates DCs by binding to Toll-like receptor 4 of these cells. TNFα and IL-12 are pro-inflammatory cytokines produced by DCs upon

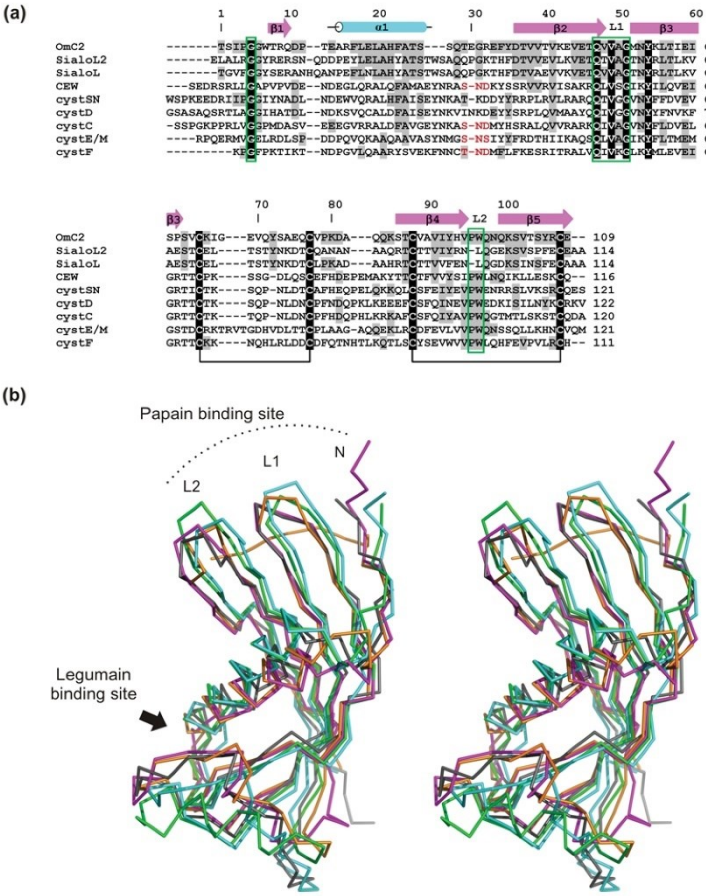


Figure 2 Comparison of the structure of OmC2 with other family 2 cystatins

(**a**) Sequence alignment of OmC2 (from the soft tick *O. moubata*) with sialostatins (SialoL and SialoL2 from the hard tick *I. scapularis*), chicken egg-white cystatin (CEW) and representative human members of family 2 cystatins (cystatins SN, D, C, E/M and F). An initial alignment was made using ClustalW and then adjusted using structural alignments where available. Residues identical with those of OmC2 are shaded grey or black (fully conserved). The secondary-structure elements of OmC2 are depicted as in Figure 1(a) (magenta for strands, cyan for helices); the L1 and L2 loops are labelled. The conserved disulfide bridges are indicated by the connecting horizontal black lines. Three conserved regions involved in the interaction with papain-type peptidases are boxed in green [32] and the putative legumain-binding site in four cystatins is highlighted in red [30]. The top line gives residue numbers for mature OmC2. Residue numbering on the right-hand side gives mature protein sequences used for the alignment (the activation processing of cystatin F is included [30]). (**b**) Stereo image showing a superposition of Cα traces of OmC2 with four other cystatin structures. The tick salivary cystatins OmC2 and SialoL2 (PDB code 3LH4; to be released) are coloured magenta and orange respectively. Chicken egg-white cystatin (PDB code 1CEW) and human cystatins F (2CH9) and D (1RN7) are shown in green, grey and cyan respectively. Positions of the binding sites for papain-type peptidases and legumains are indicated.

activation by LPS [31]. The levels of these cytokines in culture supernatants were measured by capture ELISA. The presence of OmC2 (10 μg/ml) suppressed (by 20 %) the production of TNFα in the culture ($P = 0.043$ by t test; $d = 2.9$) after a 3 h incubation (Figure 3a). Lower concentrations of OmC2 did not cause a significant reduction of TNFα. A similar suppressive effect of OmC2 was observed on the production of IL-12 (Figure 3b). At the highest concentration tested, OmC2 (30 μg/ml) was able to partially inhibit (by 25 %) the production of this cytokine by LPS-activated DCs following a 48 h incubation period ($P = 0.051$ by t test; $d = 2.8$). OmC2 had no suppressive effect on IL-6 production by LPS-activated DCs (results not shown).

OmC2 reduces DC-mediated proliferation of naive CD4$^+$ T-cells

The effect of OmC2 on the development of adaptive immunity was assessed in this experiment. The antigen-dependent T-cell proliferation induced by DCs presenting a specific antigen was

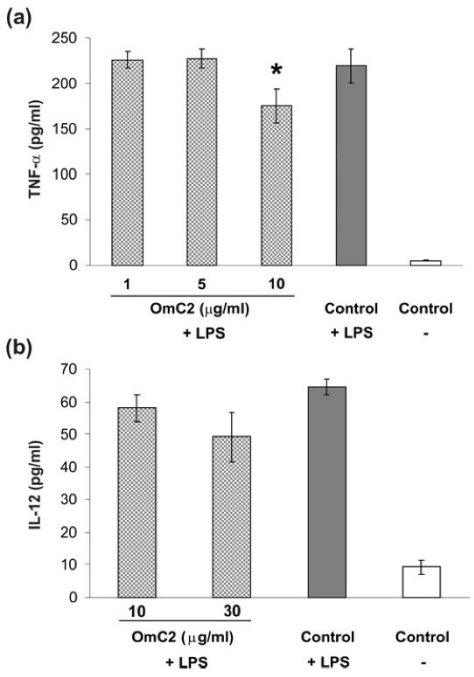


Figure 3 Effect of OmC2 on (a) TNFα and (b) IL-12 production by LPS-stimulated DCs

Enriched CD11$^+$ cells (10^5/well) were pre-incubated with OmC2 for 2 h and subsequently stimulated with LPS (50 ng/ml). Supernatants were collected 3 h and 48 h later for TNFα and IL-12 quantification respectively. Results are shown as means ± S.D.; * $P < 0.05$ compared with the LPS-activated control group.

tested using CD4$^+$ T-lymphocytes from OT-II mice. The T-cell receptor of these transgenic mice specifically recognizes the OVA (ovalbumin peptide 323–339) presented by their DCs. DC presentation of OVA in complex with the MHC II protein leads to proliferation of the transgenic CD4$^+$ T-cells, which was measured using a chromogenic substrate. Partial, statistically significant, inhibition of CD4$^+$ T-cell proliferation (by 38 %) was observed in culture wells containing OmC2 (30 μg/ml) at 72 h of incubation ($P = 0.03$ by t test; $d = 2.18$). Lower concentrations of OmC2 did not cause a significant reduction in CD4$^+$ T-cell proliferation (Figure 4).

OmC2 vaccination decreases tick feeding success

The effect of vaccinating mice with recombinant OmC2 on the feeding success of *O. moubata* ticks was assessed. The first nympal feeding stage of *O. moubata* and an inbred mouse strain (C3H/HeN) were used. Feeding success on OmC2 immunized and control mice was evaluated using three criteria: (i) the ability of nymphs to engorge; (ii) the weight of engorged nymphs; and (iii) the survival of engorged nymphs to the next developmental stage. Six mice were immunized with OmC2 and six mice (the control group) with ovalbumin. At 10 days after the last injection, sera

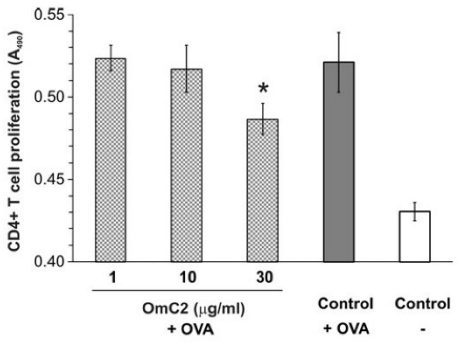


Figure 4 Effect of OmC2 on antigen specific CD4$^+$ T-cell proliferation

Enriched CD11c$^+$ cells (3×10^4/well) were pre-incubated with OmC2 (concentrations as indicated) for 4 h, before adding CD4$^+$ T-cells (2×10^5/well) from transgenic OT-II mice together with OVA (10 μg/ml). The culture was incubated for 72 h and the cell proliferation was assessed using a chromogenic assay (A_{490}). Results are shown as means ± S.D.; * $P < 0.05$ compared with the OVA-activated controls.

were collected and tested by ELISA for anti-OmC2 antibodies. Anti-OmC2 reciprocal titres ranged from 1×10^3 to 512×10^3 with an average of 203×10^3. No anti-OmC2 antibodies could be detected in the sera of the control mice. Subsequently, six or seven *O. moubata* nymphs were allowed to feed on each of these mice. The analysis of their ability to engorge showed that 62.5 % of ticks (25 out of 40 ticks) were able to feed on OmC2-immunized mice compared with 78 % of ticks (32 out of 41 ticks) fed on control mice ($P = 0.149$ by Fisher's exact test). Similarly, there was no significant difference in the weight of engorged nymphs (the second criterion for feeding success) fed on OmC2-immunized mice (1.7 ± 0.5 mg) or control mice (1.8 ± 0.3 mg) ($P = 0.55$ by t test; $d = 0.24$). The last criterion for assessment of feeding success was the survival of engorged nymphs to the next developmental stage. All ticks fed on control animals were able to moult to the second nymphal stage. In contrast, only 76 % of ticks fed on OmC2-immunized animals survived and successfully developed to the next stage. Thus the mortality rate in the group of ticks fed on OmC2-immunized animals (24 %) was significantly higher than in the control group (no mortality) ($P = 0.005$ by Fisher's exact test). The results of the vaccination experiment are summarized in Figure 5(a). Only 47.5 % of the total ticks initially allowed to feed on OmC2-immunized animals were able to successfully reach the second nymphal stage, compared with 78 % of the ticks feeding on control animals, illustrating a significant decrease in the feeding success of OmC2-immunized animals ($P = 0.006$). Detailed statistical analysis of tick survival showed that the ability of nymphs to survive and develop into the next stage was negatively correlated with the anti-OmC2 titre in the blood of the host mice ($n = 12$, $r_s = -0.66$, $P = 0.019$ by Spearman rank correlation; $d = 1.8$). Ticks fed on the mouse with the highest level of anti-OmC2 antibodies (titre 512×10^3) had the highest mortality (83.3 %); only one of six engorged ticks was able to survive and transform into the next stage (16.7 % success rate) (Figure 5b).

DISCUSSION

The present study characterizes OmC2 present in the saliva of the soft tick *O. moubata* as a cystatin with an

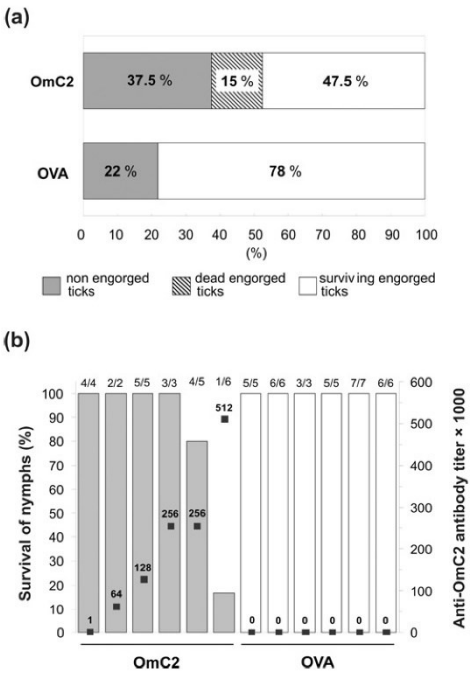


Figure 5 Effect of vaccination with OmC2 on tick feeding success

(**a**) Summary of the results of the vaccination experiment with OmC2. Mice and the first nymphal stage of *O. moubata* ticks were used as a host–parasite infestation model. The ability of ticks to engorge and survive to the next developmental stage is shown as percentages. OmC2, group of ticks fed on OmC2-immunized mice; OVA, group of ticks fed on control (OVA-immunized) mice. (**b**) Relationship between the survival of *O. moubata* ticks and anti-OmC2 antibody titres in infested mice. Column height indicates the percentage of engorged nymphs that survived and were able to develop from the first to the second nymphal stage (left-hand *y*-axis). The ratio of surviving to total number of nymphs that engorged is given above the columns (grey bars, OmC2; open bars, OVA). The anti-OmC2 reciprocal antibody titre (× 1000, right-hand *y*-axis) in each mouse is indicated by the numbers above the black squares.

unusually broad specificity. The high affinity for cysteine cathepsins with endopeptidase and exopeptidase activities clearly distinguishes the soft-tick-derived OmC2 from the hard-tick-derived sialostatins, which display weak or no inhibition of these exopeptidases [12]. The highest overall similarity of OmC2, on the level of tertiary (as well as primary) structure, is to sialostatin L2. However, local structural differences occur at the N-terminus and the L2 loop, segments that determine the cystatin's affinity and selectivity [32,33], and these are very probably the reason OmC2 targets both endo- and exo-peptidases sialostatins preferentially target endopeptidases. Interestingly, the structures of the N-terminus and the L2 loop in OmC2 resemble the corresponding segments in vertebrate cystatins. In accordance with this, the broad specificity of OmC2 is similar to that of some vertebrate cystatins, e.g. human cystatin C (for a review, see [34]). We hypothesize that the soft-tick-derived OmC2 mimics specific host-derived cystatin(s) to interfere with its/their *in vivo* functions in controlling cathepsin-mediated proteolysis.

Mammalian papain-type cathepsins play an essential role in a variety of immunological mechanisms. The biochemical properties of OmC2 that we have investigated help us to suggest aspects of the host's immune system that the inhibitor may affect.

Cathepsins S and L are critically important for the processing of antigens by antigen-presenting cells and for their presentation in complex with MHC II proteins, mechanisms that are required for the development of an adaptive immune response. In addition, these peptidases interact with the invariant chain of MHC II and are responsible for its degradation [35]. We studied the *ex vivo* effect of OmC2 on DCs, professional antigen-presenting cells, as they are an important resident cell type in the skin, where tick feeding takes place. The observed inhibition of antigen presentation and suppression of TNFα and IL-12 production agrees with the recent finding that sialostatin L suppresses DCs through the inhibition of cathepsin S [13]. TNFα and IL-12 are important messengers between immunocompetent cells, and reducing their production may impair a normal host immune response. Thus, by suppressing DC activation and antigen presentation, OmC2 can suppress the recognition of salivary antigens secreted into the host by *O. moubata* ticks. Manipulating the adaptive immunity may greatly facilitate repeated feeding by ticks and their offspring on the same host.

OmC2 was also a potent inhibitor of cathepsins B and H, which are lysosomal exopeptidases involved in intracellular protein degradation. They contribute to the variable peptide length of MHC-II-bound peptides by trimming their N- and C-termini [36]. Thus, by blocking these cathepsins, OmC2 may further suppress the recognition of tick saliva antigens. OmC2 also strongly inhibits cathepsin C (dipeptidyl peptidase I) which is a key component of innate immunity, activating serine proteases expressed by mature neutrophils [37]. Neutrophils are a major population of immunocompetent cells in blood that are part of the first line of defence in cellular immunity. By inhibiting cathepsin C, OmC2 may combat an immediate neutrophil-driven reaction against feeding ticks. It should be mentioned that cathepsin C is also involved in the activation of granzymes of cytotoxic T-lymphocytes and natural killer cells [38]. These immune mechanisms take some time to develop and soft ticks feed for a short period only. However, they may be relevant when ticks feed on a host whose immune system has been primed by the feeding of other ticks.

The vaccination experiments indicated a role of OmC2 in blood feeding, as the presence of anti-OmC2 antibodies in the blood of immunized animals reduced the tick feeding success. Interestingly, the weight of *O. moubata* nymphs was not lower when they were fed on immunized animals rather than on control animals. Immunization of guinea-pigs with sialostatin L2, on the other hand, not only promoted rejection of *I. scapularis* nymphs, but also led to a reduction in their weight [14]. This difference could be explained by a different uptake of, and exposure to, host antibodies, reflecting the specific feeding patterns of soft and hard ticks.

The principal result of our vaccination study was the finding that *O. moubata* nymphs feeding on the OmC2-immunized host showed significantly increased post-engorgement mortality that was linked to the amount of specific anti-OmC2 antibodies in the host serum. We previously reported that OmC2 is expressed not only in the salivary glands, but also in the gut of the ticks, where it is secreted into the lumen [15]. Ingested blood is stored in the tick gut lumen, whereas digestion takes place in the gut epithelium [39]. We speculated that OmC2 can protect the stored blood meal against undesired proteolysis by endogenous peptidases that are released upon lysis of digestive cells [15]. However, exogenous peptidases may also be present in the ingested blood

(including cathepsins derived from immunocompetent cells) and could be targeted by OmC2. Anti-OmC2 antibodies in the blood of sensitized hosts may block the physiological function of OmC2 in the tick gut, thereby contributing to the observed mortality of engorged *O. moubata* nymphs.

Since OmC2 occurs both in the saliva and the gut of the tick, a vaccine based on this protein may have dual activity, and may therefore be more efficient. Specific anti-OmC2 antibodies could neutralize OmC2 that is secreted into the host (interfering with its immunomodulatory function) as well as OmC2 that is expressed in the gut (interfering with a putative role in the regulation of digestion). A 'dual vaccine', which recognizes both exposed and concealed antigens, has been developed on the basis of a tick cement protein. This vaccine reportedly protects the host against the transmission of tick-borne encephalitis virus from the hard tick *Ixodes ricinus* [40].

In summary, the present study provides a comprehensive structure–function analysis of the salivary cystatin OmC2 from the soft tick *O. moubata*. This cystatin differs in its inhibitory profile and in its three-dimensional structure from the other salivary cystatins described in ticks. We demonstrate that OmC2 acts upon immunocompetent cells, and that it has potential as an anti-tick vaccine that could be used to control *O. moubata* populations and combat the spread of associated diseases.

AUTHOR CONTRIBUTION

Jiří Salát was responsible for protein expression and biological experiments; Guido Paesen, Miles Nunn, Juraj Majtán, Lenka Grunclová and Petr Kopáček participated in protein production; Helena Horká and Jan Kopecký participated in the biological studies; Michalis Kotsyfakis, Zuzana Kovářová and Martin Horn performed the biochemical characterization; Miloslav Šanda performed the proteomic analysis; Zuzana Kovářová, Pavlína Řezáčová, Jiří Brynda and John F. Andersen were involved in the crystallization experiements and structure determination; and Jiří Salát and Michael Mareš designed the project and wrote the manuscript.

ACKNOWLEDGEMENTS

We thank Markéta Ondračková for help with statistical analysis. Diffraction data were collected at beamline X12, EMBL-Hamburg outstation (at DESY, Hamburg, Germany).

FUNDING

This work was supported by the Grant Agency of the Academy of Sciences of the Czech Republic [grant numbers KJB500960702, IAA600960811], the Grant Agency of the Czech Republic [grant number P207/10/2183], Research Center LC06009 [research projects Z60220518, Z40550506 and MSM-123100003] and by the UK Natural Environment Research Council.

REFERENCES

1 Varma, M. G. (1956) Infections of *Ornithodororos* tick with relapsing fever spirochaetes, and the mechanisms of their transmission. Ann. Trop. Med. Parasitol. **50**, 18–31
2 Dixon, L. K., Abrams, C. C., Bowick, G., Goatley, L. C., Kay-Jackosn, P. C., Chapman, D., Liverani, E., Nix, R., Silk, R. and Zhang, F. Q. (2004) African swine fever virus proteins involved in evading host defence systems. Immunopathology **100**, 117–134
3 Bowman, A. S., Coons, L. B., Needham, G. R. and Sauer, J. R. (1997) Tick saliva: recent advances and implications for vector competence. Med. Vet. Entomol. **11**, 277–285
4 Francischetti, I. M. B, Sa-Nunes, A., Mans, B. J., Santos, I. M. and Ribeiro, J. M. C. (2009) The role of saliva in tick feeding. Front. Biosci. **14**, 2051–2088
5 Coullin, I., Maillet, I., Vargaftig, B. B., Jacobs, M., Paesen, G. C., Nuttall, P. A., Lefort, J., Moser, R., Weston-Davies, W. and Ryffel, B. (2004) Arthropod-derived histamine binding protein prevents murine allergic asthma. J. Immunol. **173**, 3281–3286
6 Turk, V., Stoka, V. and Turk, D. (2008) Cystatins: biochemical and structural properties, and medical relevance. Front. Biosci. **13**, 5406–5420
7 Schonemeyer, A., Lucius, R., Sonnenburg, B., Brattig, N., Sabat, R., Schilling, K., Bradley, J. and Hartmann, S. (2001) Modulation of human T cell responses and macrophage functions by onchocystatin, a secreted protein of the filarial nematode *Onchocerca volvulus*. J. Immunol. **167**, 3207–3215
8 Hartmann, S. and Lucius, R. (2003) Modulation of host immune response by nematode cystatins. Int. J. Parasitol. **33**, 1291–1302
9 Valenzuela, J. G., Francishetti, I. M. B., Pham, V. M., Garfield, M. K., Mather, T. N. and Ribeiro, J. M. C. (2002) Exploring the sialome of the tick *Ixodes scapularis*. J. Exp. Biol. **205**, 2843–2864
10 Ribeiro, J. M., Alarcon-Chaidez, F. B., Francishetti, I. M., Mans, B. J., Mather, T. N., Valenzuela, J. G. and Wikel, S. K. (2006) An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. Insect Biochem. Mol. Biol. **36**, 111–129
11 Karim, S., Miller, N. J., Valenzuela, J., Sauer, J. R. and Mather, T. N. (2005) RNAi-mediated gene silencing to assess the role of synaptobrevin and cystatin in tick blood feeding. Biochem. Biophys. Res. Commun. **334**, 1336–1342
12 Kotsyfakis, M., Karim, S., Andersen, J. F., Mather, T. N. and Ribeiro, J. M. C. (2007) Selective cysteine protease inhibition contributes to blood-feeding success of the tick *Ixodes scapularis*. J. Biol. Chem. **282**, 29256–29263
13 Sa-Nunes, A., Bafica, A., Antonelli, L. R., Choi, E. Y., Francischetti, I. M. B., Andersen, J. F., Shi, G., Chavakis, T., Ribeiro, J. M. C. and Kotsyfakis, M. (2009) The immunomodulatory action of Sialostatin L on dendritic cells reveals its potential to interfere with autoimmunity. J. Immunol. **182**, 7422–7429
14 Kotsyfakis, M., Anderson, J. M., Andersen, J. F., Calvo, E., Francischetti, I. M. B., Mather, T. N., Valenzuela, J. G. and Ribeiro, J. M. C. (2008) Cutting edge: immunity against a 'silent' salivary antigen of the Lyme vector *Ixodes scapularis* impairs its ability to feed. J. Immunol. **181**, 5209–5212
15 Grunclova, L., Horn, M., Vancova, M., Sojka, D., Franta, Z., Mares, M. and Kopacek, P. (2006) Two secreted cystatins of the soft tick *Ornithodoros moubata*: differential expression pattern and inhibitory specificity. Biol. Chem. **387**, 1635–1644
16 Kotsyfakis, M., Sa-Nunes, A., Francischetti, I. M., Mather, T. N., Andersen, J. F. and Ribeiro, J. M. (2006) Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*. J. Biol. Chem. **281**, 26298–26307
17 Minor, W., Cymborowski, M., Otwinowski, Z. and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution – from diffraction images to an initial model in minutes. Acta Crystallogr. Sect. D Biol. Crystallogr. **62**, 859–866
18 Vagin, A. and Teplyakov, A. (2000) An approach to multi-copy search in molecular replacement. Acta Crystallogr. Sect. D Biol. Crystallogr. **56**, 1622–1624
19 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect. D Biol. Crystallogr. **53**, 240–255
20 Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. **50**, 760–763
21 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. **60**, 2126–2132
22 Winn, M. D., Isupov, M. N. and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. Sect. D Biol. Crystallogr. **57**, 122–133
23 Lovell, S. C., Davis, I. W., Arendall, III, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. and Richardson, D. C. (2003) Structure validation by Cα geometry: θ, ψ and Cβ deviation. Proteins **50**, 437–450
24 Holm, L. and Sander, C. (1994) Searching protein structure databases has come of age. Proteins **19**, 165–173
25 Krissinel, E. and Henrick, K. (2005) Detection of protein assemblies in crystals. In CompLife (Berthold, M. R., Glen, R., Diederichs, K., Kohlbacher, O. and Fischer, I., eds), pp. 163–174, Springer-Verlag, Berlin
26 Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences, 2nd edn, pp. 8–108, Lawrence Earlbaum Associates, Hillsdale
27 Alvarez-Fernandez, M., Barrett, A. J., Gerhartz, B., Dando, P. M., Ni, J. and Abrahamson, M. (1999) Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. J. Biol. Chem. **274**, 19195–19203
28 Betts, M. J. and Sternberg, M. J. (1999) An analysis of conformational changes on protein protein association: implications for predictive docking. Protein Eng. **12**, 271–283
29 Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B. and Turk, V. (1990) The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. EMBO J. **9**, 1939–1947
30 Hamilton, G., Colbert, J. D., Schuettelkopf, A. W. and and Watts, C. (2008) Cystatin F is a cathepsin C-directed protease inhibitor regulated by proteolysis. EMBO J. **27**, 499–508
31 Kaisho, T. and Akira, S. (2001) Dendritic-cell function in Toll-like receptor- and MyD88-knockout mice. Trends Immunol. **22**, 78–83

32 Hall, A., Håkansson, K., Mason, R. W., Grubb, A. and Abrahamson, M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. J. Biol. Chem. **270**, 5115–5121
33 Hall, A., Ekiel, I., Mason, R. W., Kasprzykowski, F., Grubb, A. and Abrahamson, M. (1998) Structural basis for different inhibitory specificities of human cystatins C and D. Biochemistry **37**, 4071–4079
34 Dickinson, D. P. (2002) Salivary (SD-type) cystatins: over one billion years in the making – but to what purpose? Crit. Rev. Oral Biol. Med. **13**, 485–508
35 Honey, K. and Rudensky, A. Y. (2003) Lysosomal cysteine proteases regulate antigen presentation. Nat. Rev. Immunol. **3**, 472–482
36 Chapman, H. A. (2006) Endosomal proteases in antigen presentation. Curr. Opin. Immunol. **18**, 78–84
37 Adkison, A. M., Raptis, S. Z., Kelley, D. G. and Pham, C. T. (2002) Dipeptidyl peptidase activates neutrophil-derived serine proteases and regulates the development of acute experimental arthritis. J. Clin. Invest. **109**, 363–371
38 Pham, C. T. and Ley, T. J. (1999) Dipeptidyl peptidase I is required for the processing and activation of granzymes A and B *in vivo*. Proc. Natl. Acad. Sci. U.S.A. **96**, 8627–8632
39 Lara, F. A., Lins, U., Paiva-Silva, G., Almeida, I. C., Braga, C. M., Miguens, F. C., Oliveira, P. L. and Dansa-Petretski, M. (2003) A new intracellular pathway of haem detoxification in the midgut of the cattle tick *Boophilus microplus*: aggregation inside a specialized organelle, the hemosome. J. Exp. Biol. **206**, 1707–1715
40 Labuda, M., Trimnell, A. R., Lickova, M., Kazimirova, M., Davies, G. M., Lissina, O., Hails, R. S. and Nuttall, P. A. (2006) An antivector vaccine protects against a lethal vector-borne pathogen. PLoS Pathog. **2**, 251–259

Received 22 February 2010/21 April 2010; accepted 28 April 2010
Published on the Internet 14 June 2010, doi:10.1042/BJ20100280

Biochem. J. (2010) **429**, 103–112 (Printed in Great Britain) doi:10.1042/BJ20100280

online data

SUPPLEMENTARY ONLINE DATA

Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick *Ornithodoros moubata*

Jiří SALÁT*†[1], Guido C. PAESEN‡, Pavlína ŘEZÁČOVÁ§‖, Michalis KOTSYFAKIS*¶, Zuzana KOVÁŘOVÁ§**, Miloslav ŠANDA§, Juraj MAJTÁN‡, Lenka GRUNCLOVÁ*†, Helena HORKÁ*†, John F. ANDERSEN¶, Jiří BRYNDA§‖, Martin HORN§, Miles A. NUNN‡, Petr KOPÁČEK*†, Jan KOPECKÝ*† and Michael MAREŠ§[1]

*Institute of Parasitology, Biology Centre of the Academy of Sciences of the Czech Republic, Branišovská 31, 37005 České Budějovice, Czech Republic, †Faculty of Sciences, University of South Bohemia, Branišovská 31, 37005 České Budějovice, Czech Republic, ‡Centre for Ecology and Hydrology, Mansfield Road, Oxford, OX1 3SR, U.K., §Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610 Praha 6, Czech Republic, ‖Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610 Praha 6, Czech Republic, ¶Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, U.S.A., and **Department of Biochemistry, Faculty of Science, Charles University in Prague, Hlavova 8, 12843 Praha 2, Czech Republic

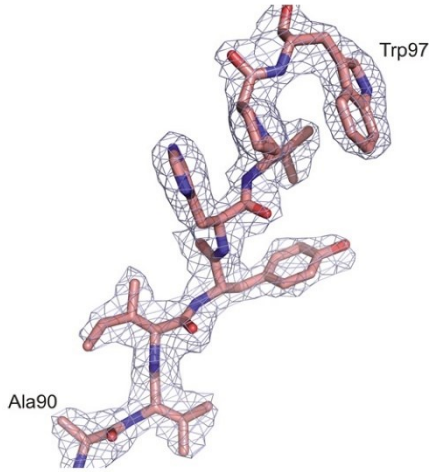


Figure S1 Representative $2F_o-F_c$ electron density map contoured at 1.5σ for residues 90–97 belonging to the L2 loop

Table S1 Peptides derived from the native OmC2 identified by LC-MS/MS analysis

Δ difference between experimental and theoretical mass; *n*, ion charge.

$[M+nH]^{n+}_{theor.}$ (Da)	*n*	Δ (p.p.m.)	Peptide (amino acid sequence)
775.87960	2	−2.2	IGEVQYSAEQCVPK
726.68804	3	−3.6	IGEVQYSAEQCVPKDAQQK
624.98569	3	−3.3	STCVAVIYHVPWQNQK
623.83921	2	−0.1	LTIEISPSVCK
598.63208	3	−0.4	FLELAHFATSSQTEGR

Table S2 Crystallographic (a) dataset and (b) refinement statistics

The data in parentheses refer to the highest-resolution shell. $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i\ I_i(hkl)$, where the $I_i(hkl)$ is an individual intensity of the *i*th observation of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of reflection hkl with summation over all data. R value $= \||F_o| - |F_c|\|/|F_o|$, where F_o and F_c are the observed and calculated structure factors respectively. R_{free} is equivalent to the R value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process [1].

Parameter	Value
(a) Data collection statistics	
Space group	$P3_121$
Cell parameters (Å)	69.8, 69.8, 133.9 90.0, 90.0, 120.0
Number of molecules in AU	2
Wavelength (Å)	0.953
Resolution (Å)	50–2.45 (2.54–2.45)
Number of unique reflections	14369 (1285)
Redundancy	5.6 (3.8)
Completeness (%)	98.8 (90.0)
R_{merge}	5.7 (33.4)
Average I/σ(I)	27.3 (2.43)
Wilson B (Å^2)	53.4
(b) Refinement statistics	
Resolution range (Å)	33.7–2.45 (2.51–2.45)
No. of reflections in working set	13588 (899)
No. of reflections in test set	744 (39)
R value (%)	18.7 (26.2)
R_{free} value (%)	22.3 (42.2)
RMSD bond length (Å)	0.012
RMSD angle (°)	1.256
Number of atoms in AU	1865
Number of protein atoms in AU	1735
Number of water molecules in AU	112
Mean B value protein/solvent (Å^2)	37.8/66.8
Ramachandran plot statistics[a]	
Residues in favoured regions (%)	90.6
Residues in allowed regions (%)	9.4

[a]As determined by PROCHECK [2].

[1] Correspondence may be addressed to either of these authors (email george@paru.cas.cz or mares@uochb.cas.cz).
The structural co-ordinates reported for salivary cystatin from *Ornithodoros moubata* will appear in the Protein Data Bank under accession code 3L0R.

REFERENCES

1 Brunger, A. T. (1992) Free R-value – a novel statistical quantity for assessing the accuracy of crystal-structures. Nature **355**, 472–475
2 Laskowski, R. A., Macarthur, M. W., Moss, D. S. and Thornton, J. M. (1993) Procheck – a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291

Received 22 February 2010/21 April 2010; accepted 28 April 2010
Published on the Internet 14 June 2010, doi:10.1042/BJ20100280

5.2. Publikace č. 2: A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation

Klíště obecné (*Ixodes ricinus*) je jedním z nejrozšířenějších druhů klíštěte ve střední Evropě. Je známé jako hlavní vektor dvou významných onemocnění přenášených na člověka - Lymské boreliózy a klíšťové encefalitidy. Sliny krev sajících parazitů přicházejí jako první do kontaktu s hostitelem a bioaktivní látky v nich obsažené interakci parazit-hostitel umožňují. V současné době je proto věnována značná pozornost jejich mapování a komplexní analýze u medicinálně významného druhu *I. ricinus.*

Tato práce je zaměřena na novou proteinovou molekulu ze slin *I. ricinus*, která patří do serpinové rodiny proteasových inhibitorů a byla nazvána *I. ricinus* serpin 2 (IRS-2). Sekvence IRS-2 byla určena z cDNA knihovny připravené ze slinných žláz izolovaných z polonasátých samic *I. ricinus*. Dynamika exprese IRS-2 byla analyzována pomocí metody qPCR a prokázala zvýšení hladin exprese mRNA v průběhu sání na 36násobek u sajících klíšťat oproti nenasátým klíšťatům, což nasvědčuje úloze tohoto proteinu v průběhu interakce s hostitelem. Tato funkce byla dále studována pomocí biologických testů.

Na myším modelu byl analyzován proces akutního zánětu, který byl zvolen na základě znalosti o protizánětlivých účincích slin klíšťat. Test je založen na vstříknutí polysacharidu karagenanu, což indukuje akutní zánět v tkáni, při kterém dochází k tvorbě otoku a influxu neutrofilů. V experimentu provedeném v přítomnosti IRS-2 docházelo k výraznému snížení otoku až na úroveň kontroly. Podobně byl také silně redukován příliv neutrofilů monitorovaný měřením jejich myeloperoxidasové aktivity.

V dalším kroku se detailně studovala inhibiční specifita IRS-2 vůči 14 zástupcům fyziologicky relevantních serinových proteas a dvěma prototypovým serinovým proteasám - chymotrypsinu a trypsinu. IRS-2 nejlépe inhiboval trypsin a trombin z proteas trypsinového typu a také chymotrypsin, katepsin G a chymasu žírných buněk z proteas chymotrypsinového typu. Pro senzitivní fyziologicky relevantní proteasy byly navrženy podrobné testy. V případě chymasy se pracovalo s myším homologem mMCP-4, který je hlavní proteasou myších žírných buněk pojivové tkáně. IRS-2 inhiboval mMCP-4 v testech *in vitro* i *ex vivo*. V testu se sledoval vliv IRS-2 na agregaci krevních destiček. Agregace jakožto fyziologická odpověď na poranění, je pro klíště nežádoucí, protože brání sání krve a může způsobit uvěznění klíštěte v hojící se ráně. V přítomnosti IRS-2 došlo ke kompletní inhibici agregace indukované katepsinem G přes receptor

PAR4 a při vyšší koncentraci IRS-2 byl inhibován proces agregace indukované trombinem přes receptor PAR1.

Součástí projektu byla krystalografická analýza IRS-2. Byly určeny krystalizační podmínky a vyřešena prostorová struktura s rozlišením 1,8 Å pomocí metody molekulárního nahrazení. IRS-2 má typickou architekturu serpinů tvořenou třemi β-skládanými listy a devíti α-šroubovicemi. IRS-2 byl vykrystalizován v tzv. relaxovaném R-stavu, kdy je smyčka reaktivního centra proteolyticky štěpena (viz navazující publikace č. 3 prezentovaná v této disertační práci) a vložena do β-skládaného listu. Tato změna konformace byla dříve popsána jako klíčová součást „sebevražedného" inhibičního mechanismu serpinů při ataku proteasou. Štěpená vazba je za zbytkem Tyr341, který odpovídá pozici P1 podmísta substrátu pro proteasy. Přítomnost této aromatické aminokyseliny v pozici P1 reaktivního centra je specifickým rysem IRS-2. Většina serpinů v genomu klíštěte *Ixodes scapularis* (nejbližšího příbuzného *I. ricinus*) má v této pozici arginin, což odpovídá typickému štěpenému místu pro proteasy trypsinového typu. Naopak přítomnost experimentálně ověřeného tyrosinu v poloze P1 činí IRS-2 vhodný k inhibici chymotrypsinových proteas jako je chymasa a katepsin G. Dokonce lze z dostupných publikovaných dat usuzovat, že IRS-2 je jediný slinný serpin z *I. ricinus* schopný inhibovat jak chymasu, tak katepsin G. Obě proteasy jsou zapojené do vzniku akutního zánětu - katepsin G sekretují po aktivaci neutrofily a chymasu žírné buňky. Mají v tomto procesu řadu úloh, mezi které patří na molekulární úrovni také např. proteolytická aktivace chemokinů a zpracování dalších substrátů, jako jsou velké endoteliny. Ty jsou prekurzory několika vazokonstriktorů a současně i chemoatraktanty pro neutrofily a monocyty. Reaktivní centrum IRS-2 se významně podobá chymasovému štěpenému místu v endotelinech: pozice P3 až P3' v IRS-2 obsahuje sekvenci VPY*SLG (* označuje štěpenou vazbu) homologickou se sekvencí VPY*GLG v endotelinu. IRS-2 tak představuje první parazitární protein, který mimikuje hostitelské endogenní substráty za účelem inhibice cílových hostitelských proteas. Zároveň se svou strukturou IRS-2 také podobá α-1-antichymotrypsinu, který je přirozeným serpinovým regulátorem katepsinu G a chymasy u savců. Druhou strukturně homologickou molekulou je savčí serpin antitrombin, který je důležitým inhibitorem koagulační proteolytické kaskády. Na strukturní a funkční úrovni je tedy IRS-2 blízký hostitelským proteasovým inhibitorům, což umožňuje cílený zásah do proteolytické rovnováhy mezi proteasami hostitele a jejich endogenními inhibitory s cílem modulovat procesy koagulace a zánětu ve prospěch parazita.

IRS-2 jako funkčně významný slinný protein představuje kandidátní molekulu pro vývoj protiklíštěcí vakcíny proti *I. ricinus*, která by snižovala riziko přenosu zejména Lymské boreliózy. Dále mohou být poznatky o biochemických vlastnostech IRS-2 využity při navrhování nových protizánětlivých a antitrombotických léčiv. K tomuto účelu má značný význam vyřešená 3D struktura IRS-2 umožňující racionální odvození derivátů s optimalizovanými funkčnímu vlastnostmi.

Můj podíl na práci spočíval v: 1) nalezení a optimalizaci krystalizačních podmínek IRS2 a příspěvku k následné rentgenostrukturní analýze, 2) interpretaci získané 3D struktury IRS2 s využitím molekulární grafiky a korelace vztahů mezi strukturou a aktivitou.

blood

2011 117: 736-744
Prepublished online Oct 12, 2010;
doi:10.1182/blood-2010-06-293241

A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation

Jindrich Chmelar, Carlo J. Oliveira, Pavlina Rezacova, Ivo M. B. Francischetti, Zuzana Kovarova, Gunnar Pejler, Peter Kopacek, José M. C. Ribeiro, Michael Mares, Jan Kopecky and Michail Kotsyfakis

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation

Jindrich Chmelar,[1,2] Carlo J. Oliveira,[3] Pavlina Rezacova,[4] Ivo M. B. Francischetti,[5] Zuzana Kovarova,[4] Gunnar Pejler,[6] Peter Kopacek,[1] José M. C. Ribeiro,[5] Michael Mares,[4] Jan Kopecky,[1] and Michail Kotsyfakis[1]

[1]Institute of Parasitology, Biology Centre of the Academy of Sciences of Czech Republic, Ceske Budejovice, Czech Republic; [2]Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic; [3]Department of Biochemistry and Immunology, School of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto, Brazil; [4]Institute of Organic Chemistry and Biochemistry of the Academy of Sciences of Czech Republic, Prague, Czech Republic; [5]National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD; and [6]Department of Anatomy, Physiology and Biochemistry, Biomedical Centre, Swedish University of Agricultural Sciences, Uppsala, Sweden

Platelet aggregation and acute inflammation are key processes in vertebrate defense to a skin injury. Recent studies uncovered the mediation of 2 serine proteases, cathepsin G and chymase, in both mechanisms. Working with a mouse model of acute inflammation, we revealed that an exogenous salivary protein of *Ixodes ricinus*, the vector of Lyme disease pathogens in Europe, extensively inhibits edema formation and influx of neutrophils in the inflamed tissue. We named this tick salivary gland secreted effector as *I ricinus* serpin-2 (IRS-2), and we show that it primarily inhibits cathepsin G and chymase, while in higher molar excess, it affects thrombin activity as well. The inhibitory specificity was explained using the crystal structure, determined at a resolution of 1.8 Å. Moreover, we disclosed the ability of IRS-2 to inhibit cathepsin G-induced and thrombin-induced platelet aggregation. For the first time, an ectoparasite protein is shown to exhibit such pharmacological effects and target specificity. The stringent specificity and biological activities of IRS-2 combined with the knowledge of its structure can be the basis for the development of future pharmaceutical applications. (*Blood.* 2011; 117(2):736-744)

Introduction

The first response of a vertebrate organism to a skin injury involves blood coagulation, vasoconstriction, platelet aggregation, and innate immune mechanisms, such as inflammation and complement activation, that have an important antibacterial role. Upon a tick bite, these mechanisms are activated as a first line of defense and are dependent, among others, on serine protease cascades (in blood coagulation and complement activation) or on the proteolytic activation of bioactive peptides, such as chemokines, vasoconstrictors, or protease-activated receptors (PARs). Previous studies demonstrated that these serine protease–dependent processes are regulated by endogenous inhibitors. To prevent pathophysiological conditions, proteases and inhibitors must be in balance; in the case of an exogenous inhibitor (eg, of parasitic origin) infiltrating the system, this balance will be impaired.

This phenomenon is also observed in the saliva of hematophagous arthropod ectoparasites, including ticks.[1] The saliva of ixodid ticks plays a crucial role at the tick-host interface not only because it contains numerous proteins that affect many host nonspecific defensive mechanisms, such as inflammation, blood coagulation, and platelet aggregation, but also because it can target acquired immunity.[2] Activation of these host physiological mechanisms immediately after a tick bite would be detrimental for tick feeding success, and therefore it is vital for the tick to overcome them. Investigation of the underlying mechanism, although time consuming, revealed salivary constituents with pharmacologic activities targeting vertebrate host defense, including molecules with anti-inflammatory potential.[3,4] This process was accelerated by recent transcript-sequencing projects on tick salivary glands (ie, sialotranscriptomics), where numerous salivary protein candidates were identified that may account for the pharmacologic properties of tick saliva.[5-7]

Of interest, a number of genes encoding for potential serine protease inhibitors were identified in *Ixodes spp* salivary glands, including serpins, the largest, most diverse family of protease inhibitors. The mode of action of serpins is unique and depends on their folding and the primary structure of specific conserved domains. The literature dealing with the structural and functional properties of serpins and the evolution of this protein family is reviewed elsewhere.[8,9] More than 60 serpins were identified at the sequence level in ixodid ticks, largely due to the completed *I scapularis* genome[10-14]; however, only 1 tick serpin originating from *I ricinus*, the vector of Lyme disease pathogens in Europe, has been functionally characterized.[14] The serpin was named *I ricinus* immunosuppressant (IRIS), and it targets the serine protease, elastase. Despite not bearing a classical secretion signal, it can affect vertebrate hemostasis apart from vertebrate immunity.[15] Recently, Prevot and colleagues showed the involvement of an exosite domain in the anti-inflammatory activity of IRIS. Surprisingly, unlike its antihemostatic activity, the anti-inflammatory properties of IRIS are independent of its inhibitory nature.[16] Thus, the inhibitory function of IRIS was not directly shown to be responsible for its observed immunomodulatory properties, so the answer—whether salivary serpins can mimic the function of vertebrate regulators and thus immunomodulate the host—

Submitted June 25, 2010; accepted September 29, 2010. Prepublished online as *Blood* First Edition paper, October 12, 2010; DOI 10.1182/blood-2010-06-293241.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

remained undisclosed. The same is true for the structural basis of IRIS function. Apart from IRIS, no other serpin from *I ricinus* has been described to date.

Here, we present the functional characterization of a novel inhibitory serpin from the saliva of the tick, *I ricinus*, that we named *I ricinus* serpin-2 (IRS-2). It bears a clear secretion signal and inhibits edema formation and neutrophil influx in the inflamed tissues in a mouse model of acute inflammation. The protein targets primarily 2 proinflammatory serine proteases, cathepsin G and mast cell chymase, and, in higher molar excess, thrombin. Because of its inhibitory activity, IRS-2 blocks cathepsin G– and thrombin-induced platelet aggregation, thus playing a dual role during tick feeding, as it can interfere with both inflammation and wound healing. We also determined a high-resolution crystal structure of IRS-2 that provides structural insight into the observed inhibitory specificity. This is the first report of the crystal structure of a serpin isolated from a parasitic organism. Moreover, such a mechanism in which the parasite uses inhibition of cathepsin G and chymase to overcome the host defense system, has not been shown for any blood-feeding arthropod salivary constituent to date.

Methods

Unless otherwise indicated, standard procedures were followed,[17] and experiments were performed at room temperature (25 ± 1°C). All water used was of 18-MΩ quality produced by a MilliQ apparatus (Millipore). If not otherwise stated, all reagents were purchased from Sigma-Aldrich, and all cells were cultured at 37°C under an atmosphere of 5% CO_2. The procedures of gene cloning, sequence analysis, real-time polymerase chain reaction (PCR), protein expression, and biochemical methods are detailed in supplemental Methods (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

Paw edema assay

Female C57BL/6 mice (6-8 weeks old) were used. The mice were purchased from The Jackson Laboratory and maintained in the National Institute of Allergy and Infectious Diseases (NIAID) Animal Care Facility under pathogen-free conditions in temperature-controlled rooms, receiving water and food ad libitum. All treatments were performed in accordance with the National Institutes of Health [NIH]'s *Guide for the Care and Use of Laboratory Animals* , and the animal study protocols were approved by the Division of Intramural Research/NIAID Animal Care and Use Committee. The carrageenan-induced hind-paw inflammation model was used to investigate the potential anti-inflammatory role of IRS-2. Before each injection, the basal footpad thickness of each mouse was recorded using a caliper (Mitutoyo America Corp). Subsequently, 40 μL of carrageenan (2% in saline) was administered by intraplantar injection in each footpad in the absence or presence of different concentrations of lipopolysaccharide-free IRS-2. As a control, each group of mice received the same volume of saline (vehicle) in the presence of IRS-2 only. The same experimental design and concentrations were used for chymostatin, a generic protease inhibitor, and indomethacin, a nonsteroid anti-inflammatory compound. Each effector was coadministered with carrageenan in the mouse footpad. As an index of edema formation, paw thickness (in millimeters) was measured at 4 and 24 hours after injection, and myeloperoxidase (MPO) activity was measured to estimate the potential effect of IRS-2 on neutrophil migration. For details, please see supplemental Methods.

Isolation and cultivation of PCMCs

Animal experiments were performed in accordance with protocols approved by the local ethical committee. Peritoneal cells were obtained from 6-8-week-old C57BL/6 mice by lavage of the peritoneal cavity with ice-cold phosphate-buffered saline (PBS), washed in fresh PBS, and resuspended in growth medium. The growth medium was Dulbecco modified Eagle medium with GlutaMAX (Gibco), 10% of conditioned media from Chinese hamster ovary cells transfected with mouse stem cell factor (a gift from Dr Marc Daeron), 10% fetal calf serum, 1% penicillin-streptomycin (Invitrogen), 1% minimum essential medium with nonessential amino acids (Gibco), and 50μM mercaptoethanol. Nonadherent cells (mainly peritoneal cell–derived mast cells [PCMCs]) were collected and resuspended in fresh medium to a concentration 10^5 cells/mL every third day. After 6 weeks of cultivation, the enriched PCMCs were used for experiments.

Preparation of washed human platelets and platelet aggregation assays

Platelet-rich plasma was obtained by plateletpheresis from medication-free, adult, healthy platelet donors at the Department of Transfusion Medicine/NIH blood bank under the direction of Dr S. Leitmann as described elsewhere.[18] Briefly, after the addition of 0.2 U/mL apyrase, platelet-rich plasma was centrifuged at 1100 × *g* for 15 minutes and washed twice by centrifugation in Tyrode buffer (137mM NaCl, 27mM KCl, 12mM $NaHCO_3$, 0.42mM NaH_2PO_4, 1mM $MgCl_2$, 5.55mM glucose, 0.25% bovine serum albumin; pH 7.4). Platelets were resuspended in apyrase-free Tyrode buffer and adjusted to a concentration of 200 000-400 000 platelets/μL. Washed human platelets (300 μL) were placed in a Chrono-Log Lumi-aggregometer (Chrono-Log Corp) and stirred at 1200 rpm at 37°C for 1 minute, followed by the addition of reagents, as indicated in the figure legends. In all experiments, 1.5μM IRS-2 was preincubated for 15 minutes with the various effectors tested, and the mixture was added to platelets.

Crystallization and data collection

Details of the crystallization procedure and diffraction data collection are reported elsewhere.[19] Crystals of IRS-2 were prepared at 20°C using the hanging-drop, vapor-diffusion technique. The crystallization drop consisted of 2 μL of the IRS-2 solution (3.5 mg/mL in 20mM HEPES [N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid]; pH 7.2) and 1 μL of the reservoir solution (75mM MES [2-(N-morpholino)ethanesulfonic acid], pH 6.5, 9% [wt/vol] polyethylene glycol 20 000). For data collection, crystals were soaked in reservoir solution supplemented with 20% (vol/vol) polyethylene glycol 400 and flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K using the X12 EMBL beamline (DESY) and processed using the HKL-2000 suite of programs.[20] Crystals exhibited the symmetry of space group $P4_3$ and contained 2 molecules in the asymmetric unit. Crystal parameters and data collection statistics are given in supplemental Table 1.

Structure determination

The structure of IRS-2 was solved by molecular replacement using the MoLrep 9.2 program.[21] The search model was derived from the structure of the equine leukocyte elastase inhibitor in R-state conformation (Protein Data Bank [PDB] code 1HLE).[22] Model refinement was carried out using the program REFMAC 5.3[23] from the CCP4 package.[24] Manual building was done using Coot.[25] Tight noncrystallographic symmetry restraints were applied during initial refinement; in later stages, the restraints were loosened as guided by the behavior of R_{free}. The final steps included translation, libration, and screw refinement.[26] The quality of the final models was validated with MoLProbity server.[27] Final refinement statistics are given in supplemental Table 1. Figures showing structural representations were prepared with the program PyMOL 0.99 (DeLano Scientific).[28] The DALI server was used to search for structural homologs.[29] Ramachandran plot statistics were determined by PROCHECK.[30]

Results

Cloning of serpins from *I ricinus*

The sequences of numerous serine protease inhibitors from the serpin superfamily have been identified from various ixodid tick

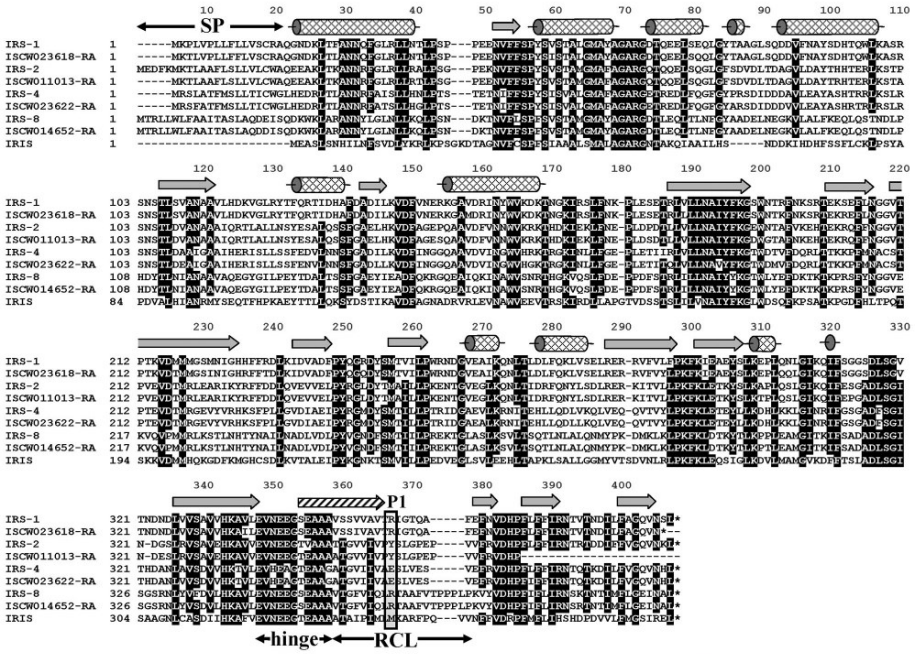


Figure 1. Alignment of the 4 full-length *I ricinus* serpin clones with their orthologs from *I scapularis* obtained from the completed *I scapularis* genome. All serpins are compared with IRIS, the only functionally characterized tick serpin so far. Unlike IRIS, IRS-2 has a clear signal peptide (residues 1-21). The secondary structure elements are indicated according to the IRS-2 crystal structure: α-helixes (cylinders), β-strands (arrows); the hatched arrow represents the region that undergoes a conformational change from a loop to a β-strand after proteolytic cleavage of RCL. SP, signal peptide; hinge, hinge region, an important determinant of serpin inhibitory potential; RCL, reactive center loop; P1 (rectangle), predicted residue behind which the target protease cleaves the RCL.

species, and it is also known that the family is represented in *I ricinus* salivary glands.[14] However, the knowledge of *I ricinus* salivary serpins, and the function of tick serpins in general, is limited to IRIS, a serpin that does not follow a classical secretion pathway from the salivary glands. Therefore, with the aim to identify classically secreted *I ricinus* serpins (ie, those containing a clear signal peptide), we designed degenerate primers based on the 2 most conserved domains (NAIYFKG and PFLFFI) found in the alignment of 10 arthropod serpins, for which sequence data were available at that time. The obtained PCR-amplified fragment was radioactively labeled and used for the screening of a cDNA library prepared from salivary glands isolated from adult female ticks after 5 days of attachment to the host to find as many serpin clones as possible. Four different full-length serpin cDNA clones that contained signal peptide, namely, IRS-1, -2, -4, and -8 (National Center for Biotechnology Information accession numbers DQ915842, DQ915843, DQ915844, and DQ915845), were identified. Orthologs for all 4 genes can be found in the genome of *I scapularis* (http://iscapularis.vectorbase.org/index.php), the vector of Lyme disease pathogens in the Eastern and Central parts of the United States. Amino-acid identity between the corresponding serpin orthologs varies from 94.9% for IRS-4 to 98.8% for IRS-8. The alignment in Figure 1 compares the 4 discovered serpins with the homologous proteins from *I scapularis* and with IRIS.[14] All proteins, including IRIS, display putative inhibitory features in the serpin hinge region, and all proteins except IRIS bear a clear signal peptide, suggesting a classical secretory mechanism from the salivary glands. Based on their amino acid (aa) sequence and on the preliminary prediction of the P1 position of each serpin, we proceeded with further analysis of IRS-2. More specifically, the predicted P2-P2′ sequence of the IRS-2–reactive center loop is Pro-Tyr-Ser-Leu with the tyrosine at the P1 position, suggesting antichymotrypsin, rather than antitrypsin, specificity. Mature IRS-2 is a 376-aa protein, weakly acidic with pI 5.53; the predicted molecular weight was 41.9 kDa. The unprocessed protein contains a 21-aa predicted signal peptide, according to the SignalP 3.0 server program.

IRS-2 expression is up-regulated after tick attachment

Next, we determined the expression profile of IRS-2 in different tick tissues throughout the feeding period of adult ticks using quantitative real-time PCR (Figure 2). IRS-2 displayed 36 times higher mRNA expression in tick salivary glands at day 6 after tick attachment to the host, compared with salivary glands derived from unfed ticks. Already, at day 2 after attachment, the expression was 9-fold higher than in unfed ticks, suggesting a role in the tick salivary glands, even in the early stage of tick feeding. The increase of IRS-2 gene expression was also notable in tick ovaries, with a 4-fold increase at day 6 after attachment. Finally, IRS-2 transcript abundance fluctuated in the tick midgut as feeding progressed.

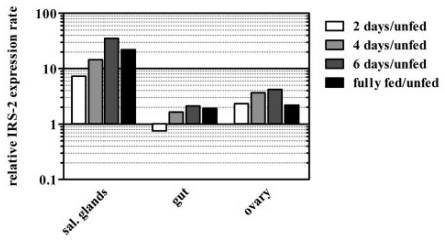


Figure 2. Expression of IRS-2 is up-regulated after tick attachment to the host. IRS-2 transcripts accumulate rapidly in salivary glands after the attachment of an adult *I ricinus* female, less rapidly in its ovaries, and they remain almost unchanged in its midgut. Each bar in the graph represents the ratio between IRS-2 transcript abundance in a certain tick tissue and a given tick feeding stage, compared with the corresponding tissue from unfed ticks (ie, the black bar in the salivary glands part of the graph shows that the transcript abundance of IRS-2 mRNA is 20× higher in the salivary glands of fully fed adult female ticks than in the salivary glands of unfed females). The scale in the y-axis is logarithmic.

IRS-2 inhibits acute inflammation

Induction of IRS-2 transcription in the salivary glands upon blood feeding and the known anti-inflammatory properties of tick saliva led us to investigate whether IRS-2 could contribute to the immunomodulatory activity of tick saliva. For this purpose, we produced recombinant IRS-2 in a bacterial expression system and proceeded with lipopolysaccharide decontamination. Subsequently, we assessed its effect in a mouse model of acute inflammation. It is well described that when carrageenan is inoculated in the mouse footpad, it induces an acute inflammatory response characterized by edema formation in the paw accompanied by neutrophil influx.[31] To investigate whether IRS-2 could modulate carrageenan-induced inflammation, carrageenan was administered in the mouse footpads in the presence or absence of IRS-2. When carrageenan was injected in the presence of IRS-2, a dose-dependent inhibition of edema formation was observed (Figure 3 A). More specifically, in the presence of 1 mg/kg of IRS-2, the thickness of the mouse paw was similar to that of negative controls; that is, PBS-injected mice (data not shown) or mice injected with IRS-2 only (Figure 3A). The decrease in edema formation reached 67.2% ($P < .001$) at 4 hours after injection, compared with positive control (carrageenan-injected footpads). This was not 100% because there was a background increase in paw thickness even when injecting vehicle alone, since the inoculum (40 μL of liquid) could not be completely absorbed by the footpad within 4 hours. The inhibition observed after 4 hours of carrageenan injection with IRS-2 at the concentration 0.3 mg/kg was 23.4%, compared with positive control, and it was not statistically significant (Figure 3A). At 24 hours after injection, there were no differences in edema formation in any of the experimental groups of mice. Next, we analyzed carrageenan-induced recruitment of neutrophils in the footpads by measuring tissue MPO activity. The experimental design was the same as in the paw edema experiment. MPO activity in the tissue was evaluated at 4 hours after injection, the time point at which edema peaks. Statistically significant inhibition ($P < .05$) of MPO activity (34.3%) was observed even when 0.3 mg/kg of IRS-2 was coadministered with carrageenan. Inhibition reached 71.6% when 1 mg/kg of IRS-2 was coadministered with carrageenan ($P < .001$; Figure 3B). Finally, we normalized the detected MPO activity for the number of neutrophils per milligram of inflamed tissue, and we further evaluated the effect of chymostatin (a generic inhibitor of primarily chymotrypsin-like serine proteases) and indomethacin (a nonsteroidal anti-inflammatory compound) on MPO activity/neutrophil migration. Both effectors displayed similar inhibitory effect with IRS-2 on neutrophil migration (Figure 3C).

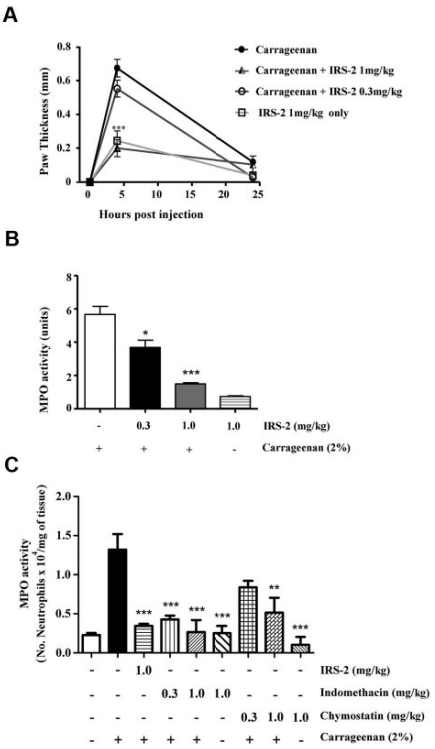


Figure 3. IRS-2 inhibits carrageenan-induced acute inflammation. Mice received carrageenan injections intraplantary in either the absence or presence of 0.3 or 1 mg/kg IRS-2. (A) Edema formation was evaluated at 4 and 24 hours after injection (abscissa) as the increase in paw thickness (in millimeters). (B) Neutrophil recruitment in inflamed footpads was evaluated by measuring tissue myeloperoxidase activity, expressed as units of activity/g of tissue (ordinate). Bar 1 (numbering left to right), activity detected when only carrageenan was administered to mice; bars 2 and 3, effect of coadministration of carrageenan with 0.3 and 1 mg of IRS-2 per 1 kg of body weight, respectively. Bar 4, IRS-2 injected without carrageenan. (C) Mice received injections of saline (-), carrageenan, carrageenan plus IRS-2 (1 mg/kg), indomethacin (0.3 and 1 mg/kg) or chymostatin (0.3 and 1 mg/kg). At 4 hours after injection, the hind paws were collected for MPO analysis and the amount of neutrophils per milligram of tissue was estimated by comparison with purified neutrophils. Asterisks represent statistically significant differences in MPO activity (*$P < .05$; **$P < .01$; and ***$P < .001$), compared with groups injected with carrageenan only (1-way analysis of variance followed by Tukey post-hoc test; n:4 in each group).

IRS-2 specifically targets cathepsin G and chymase

Given the observed anti-inflammatory effect of IRS-2, we proceeded to a more detailed analysis of the IRS-2 mechanism of action. Considering that serpins are potential inhibitors of serine proteases that play a role in inflammation, and that the bioinformatic analysis of IRS-2 reactive center loop (RCL) suggests an

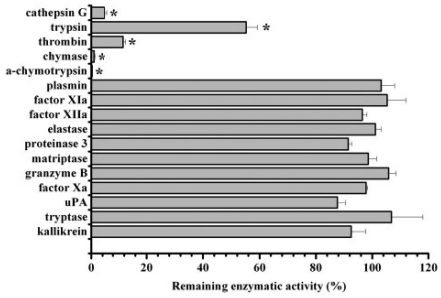


Figure 4. Inhibitory specificity of IRS-2. IRS-2 (400nM) was tested against 16 different serine proteases in triplicates. The enzyme concentration is stated in supplemental Table 2. Bars represent the mean remaining enzymatic activity in the presence of IRS-2, while error bars represent the SEM. Enzymes with an asterisk were inhibited with a statistical significance (*t* test; $P < .05$).

inhibitory specificity against chymotrypsin-like, rather than trypsin-like, serine proteases, we tested the protein for inhibitory activity against a panel of pure recombinant human serine proteases. Recombinant IRS-2 was tested against an array of 14 different physiologically relevant serine proteases, in addition to α-chymotrypsin and trypsin, the 2 archetypes for chymotrypsin- and trypsin-like serine proteases. As shown in Figure 4, IRS-2 inhibited 2 trypsin-like (trypsin and thrombin) and 3 chymotrypsin-like (α-chymotrypsin, cathepsin G, and mast cell chymase) proteases significantly ($P < .05$), with higher activity against the latter group. IRS-2 did not inhibit a series of serine proteases related with the coagulation or inflammation, such as plasmin, factor Xa, factor XIa, or elastase and proteinase 3, suggesting a stringent specificity (Figure 4). Table 1 and Figure 5A-B summarize the concentration-dependence of IRS-2 inhibition for all 5 targeted enzymes. Table 1 further describes the binding characteristics for thrombin, cathepsin G, and chymase, 3 physiologically relevant proteases. For further details on the related biochemical analysis, please see supplemental Results and supplemental Figures 1 and 2.

IRS-2 inhibits mMCP-4

Having dissected the target specificity of IRS-2, and considering that chymase (1 of the 2 targets of IRS-2) is released from mast cells upon acute inflammation, we next investigated whether IRS-2 would bind to mouse (the animal model used for the anti-inflammatory experiments) mast cell protease-4 (mMCP-4). Furthermore, mMCP-4 is the main chymotrypsin-like serine protease produced by connective tissue–type mouse mast cells and is a functional homolog to the human chymase.[32] Therefore, we evaluated the effect of IRS-2 on mMCP-4. Indeed, IRS-2 was found to inhibit mMCP-4 both in vitro and ex vivo. Analogous to the inhibition observed for human chymase, purified mMCP-4 was inhibited by IRS-2 in equimolar concentrations. The estimated inhibitory concentration at half-maximum of IRS-2 against purified mMCP-4 was 5.21nM, when using a 5nM concentration of the enzyme (Figure 5C). Moreover, IRS-2 inhibited the chymotryptic activity present in the suspension of ionomycin-activated PCMCs in a dose-dependent manner (Figure 5C). We also detected the formation of covalent mMCP-4/IRS-2 complexes by Western blot. Notably, mMCP-4/IRS-2 complexes were only detected in association with the cell layer, while no complexes were detected in the conditioned medium. The number of produced complexes was dependent on the concentration of IRS-2 (Figure 5D). More specifically, at 20nM concentration of IRS-2, most of the mMCP-4 remained in the free state, while at 50 and 200nM IRS-2, higher amounts of mMCP-4 were recovered in complex with IRS-2.

IRS-2 inhibits cathepsin G– and thrombin-induced platelet aggregation

Cathepsin G, the second target of IRS-2, plays a role in platelet aggregation, which is crucial in wound-healing processes that are essential for host ability to reject a blood-feeding tick. Therefore, we evaluated the effects of IRS-2 on cathepsin G–induced platelet aggregation, which is mediated by protease-activated receptor 4 (PAR4).[33] Figure 6A shows that cathepsin G–induced shape change and platelet aggregation was completely inhibited by 15-minute incubation with IRS-2 (Figure 6A). In higher excess to the enzyme, IRS-2 inhibited also thrombin-induced platelet aggregation (Figure 6B), which is mediated by PAR1 and PAR4. Collagen-induced (Figure 6C), convulxin-induced[34] (Figure 6D), U46619-induced (Figure 6E), and arachidonic acid-induced (Figure 6F) platelet aggregation remained unaffected, demonstrating the specificity in IRS-2–driven inhibition of platelet aggregation. These results suggest that IRS-2 interferes with platelet aggregation by blocking their activation through PARs, thus implicating IRS-2 as an inhibitor of wound-healing processes.

Structural analysis of IRS-2

Finally, we resolved the crystal structure of IRS-2, which was determined by molecular replacement and refined using data to 1.8 Å resolution. The structure of equine leukocyte elastase inhibitor (PDB code 1HLE) was used as a search model; this homolog had the highest sequence similarity with IRS-2 (35% identity) among the serpin structures available in the PDB. The tetragonal crystal form of IRS-2 contains 2 molecules in the asymmetric unit with solvent content of 45%. The root mean square deviation (RMSD) for superposition of the main-chain atoms of these

Table 1. A tabular representation of IRS-2 inhibition characteristics for the targeted serine proteases

Enzyme	Amount of enzyme used, nM	IRS2 IC_{50}, nM	IC_{50}/enzyme ratio	Inhibitor characteristics
Cathepsin G	12.2	11.5 ± 0.7	0.94	Fast binding tight
Chymase	3.6	4 ± 0.2	1.11	Slow binding tight
Thrombin	0.02	170.3 ± 11	8515	Slow binding classical
α-chymotrypsin	0.038	0.38 ± 0.02	10	N/A
Trypsin	0.25	562.7 ± 27.3	2251	N/A

The amount of enzyme used in the assays as well as the inhibitory concentration at half-maximum (IC_{50}) of IRS-2 for these enzymes is stated in the first and second columns. The third column represents the ratio between IRS-2 IC_{50} and enzyme concentration, and the last column shows the binding characteristics of the inhibitor to the targeted enzymes with the relevance in host physiology (ie, tight or classical inhibitor, fast binding, or slow binding to the enzyme). Because trypsin and α-chymotrypsin were not physiologically relevant to our studies, we did not further study their inhibition characteristics, as depicted in by N/A.

NA indicates not applicable.

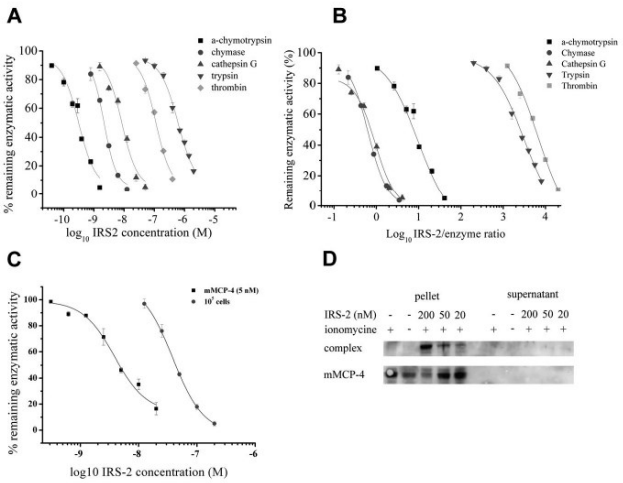


Figure 5. IRS-2 is a serpin with activity against chymotrypsin-like, rather than trypsin-like, serine proteases. (A) Proteases targeted by IRS-2. α-chymotrypsin, chymase, cathepsin G, trypsin, and thrombin are inhibited by IRS-2. The amount of enzyme used is stated in Table 1. The mean remaining enzymatic activity in the presence of various concentrations of IRS-2 is represented, while the error bars represent the SEM in triplicate assays. (B) Inhibition data are normalized by plotting the remaining enzymatic activity (y-axis) against the IRS-2/enzyme ratio (x-axis). (C) Inhibition of purified mMCP-4 (5nM) or present in the suspension of activated mouse PCMCs by IRS-2. The mean remaining enzymatic activity in the presence of various concentrations of IRS-2 is presented (± SEM). (D) Western blot analysis showing covalent complex formation between IRS-2 and mMCP-4 produced by PCMCs. Notably, enzymatic activity and complexes between mMCP-4 and IRS-2 were predominantly cell associated, rather than being present in cell-free supernatants. Cells activated by ionomycin were used as positive control (+), and untreated cells were used as negative control (−). All samples with IRS-2 were activated by ionomycin.

2 molecules is 0.69 Å, a value within the range observed for different crystal structures of identical proteins.[35] Minor structural changes are localized in solvent-exposed loop regions. The N-terminal residue, Met1, represents a cloning artifact that is not part of the native protein sequence. Atomic coordinates of IRS-2 and experimental structure factors have been deposited in the PDB, with the accession code 3NDA.

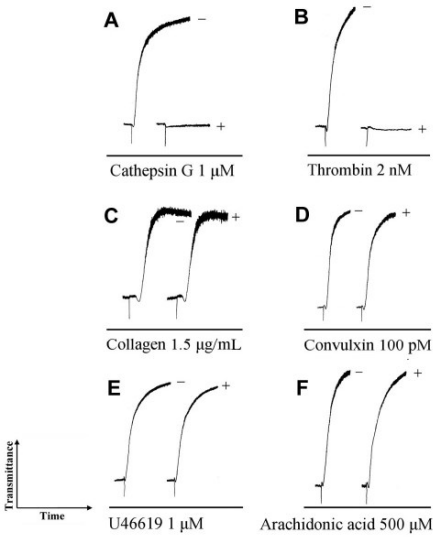


Figure 6. IRS-2 inhibits platelet aggregation induced by cathepsin G. In all experiments, 1.5µM IRS-2 was incubated with the indicated amount of platelet aggregation activator for 15 minutes and the mixture was added to platelets. (−) activator only, (+) activator plus IRS-2.

Figure 7 shows the overall structure of IRS-2. It adopts a typical serpin fold composed of 3 large β-sheets and 9 α-helices. Both molecules in the asymmetric unit adopt a conformation known as the relaxed or R state of the serpins, in which the RCL is cleaved and inserted into the central β-sheet A as a strand, S4 (Figure 7A). Cleavage of RCL occurs during the crystallization process and is catalyzed by traces of contaminating proteases, as demonstrated in the crystallization study.[19] The cleavage site is positioned at residue Tyr341, which represents the P1 substrate residue. In our structure, this residue has a well-defined electron density for its C-terminal carboxyl group. The $P1'$–$P5'$ residues of the cleaved RCL are disordered to various extents in the crystal structure; more specifically, residues 342-344 and 342-346 in molecules A and B, respectively, could not be modeled and are missing in the final structure.

A structural comparison of IRS-2 with structures from the serpin superfamily deposited in the PDB identified bovine antithrombin III (PDB code 1ATT) as the closest structural homolog (RMSD approximately 1.4 Å for 369 aligned residues). A lower structural similarity was found to typical representatives of the mammalian serpins, human α-1-antichymotrypsin (PDB code 2ACH, RMSD approximately 1.6 Å for 337 residues), human α-1-antitrypsin (PDB code 9API; RMSD approximately 1.8 Å for 335 residues), and to human corticosteroid-binding globulin and human plasminogen activator inhibitor-1 (PDB codes 2VDX and 3CVM; both RMSD approximately 1.7 Å for 363 residues). All the compared structures are in the R-state conformation; a superposition of selected structures with IRS-2 is presented in Figure 7B.

Discussion

Many *I scapularis* serpins possess Arg at the P1 position,[10] which corresponds to typical cleavage sites for trypsin-like enzymes. This suggests their involvement in modulation or regulation of trypsin-like proteases involved in host-blood coagulation or tick hemolymph, probably with some redundancy in their function. On the

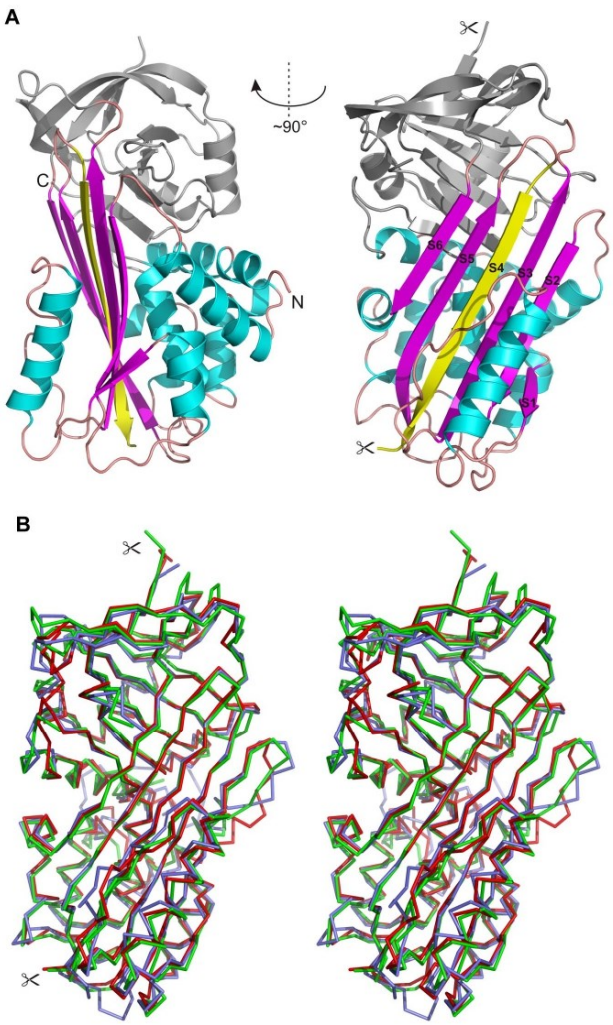


Figure 7. Crystal structure of IRS-2. (A) Overall 3-dimensional structure of IRS-2 in a cartoon representation. The central β-sheet A (magenta) and surrounding helices (cyan) are highlighted. The reactive center loop is cleaved in the relaxed (R-state) conformation of IRS-2 and forms the S4 β-strand (yellow) inserted into the β-sheet A; the termini generated by this proteolytic cleavage are marked by scissors. N- and C-termini of IRS-2 are labeled (N, C). (B) Stereo image showing a superposition of Cα traces of IRS-2 (red) with 2 homologous mammalian serpins in the R-state conformation. Antithrombin III (blue; 1ATT) and α-1-antichymotrypsin (green; 2ACH) display a high level of similarity to IRS-2 with regard to structural homology and inhibitory specificity, respectively.

other hand, only 2 putative serpins that possess the aromatic aa (ie, Trp, Tyr, or Phe) at the P1 position can be found in the *I scapularis* genome, thus having the potential to inhibit chymase and other chymotrypsin-like proteases.[36] According to the nomenclature used by the VectorBase server (http://iscapularis.vectorbase.org), these serpins are encoded by the transcripts, ISCW004156-RA and ISCW011013-RA. The latter transcript is the homolog of IRS-2. Considering the preferential aa residues in the cleavage site for cathepsin G,[37] we believe that IRS-2 is the only known salivary serpin from *I ricinus* that can target both cathepsin G and chymase.

IRS-2 interacts with proteases via a classical "suicide inhibition" mechanism of serpins. It involves cleavage of RCL and formation of S4 β-strand that is inserted in the middle of β-sheet A (Figure 7A). We determined a high-resolution crystal structure of the protease-cleaved IRS-2 that provided an experimental evidence of the Tyr residue at the P1 position. These structural data will be valuable for designing smaller peptides that may mimic IRS-2 inhibitory activities and its pharmacological actions. IRS-2 targets specifically the chymotrypsin-like proteases, cathepsin G and chymase. Both of these proteases are secreted after neutrophil

(cathepsin G) and mast cell (chymase) activation, and they are involved in a whole range of physiological processes associated with the development of an acute inflammatory response and, in particular, in the cross-talk between neutrophils and platelets in the hemostatic process.[38] Cathepsin G is produced mainly by neutrophils, where it is stored in azurophile granules, but it may also be produced in low amounts by mast cells. Cathepsin G is known primarily for its ability to kill engulfed pathogens and for its role in tissue remodeling during inflammation. Furthermore, it is involved in the proteolytic activation of various chemoattractants and hormones as well as in cell signaling by cleaving PAR4, the latter having been identified as an important signaling receptor in inflammation and platelet activation.[39] Cleavage of PAR4 by cathepsin G is responsible for the platelet activation that leads to their aggregation and clot formation. Here, we showed, for the first time, that a tick salivary protein inhibits cathepsin G–induced platelet aggregation, which apparently helps the tick in obtaining its blood meal. In addition, we revealed that in higher molar excess, IRS-2 affects thrombin-induced platelet aggregation as well, further disclosing its multipotential role in inflammation and hemostasis through the modulation of PAR activation.

Besides preventing blood loss after an injury, activated platelets produce several chemokines and their precursors, which are further proteolyticaly processed. In this respect, the role of cathepsin G overlaps with that of chymase. Chymase is produced almost exclusively by mast cells that are resident in mucosal and connective tissues and it processes many substrates, such as angiotensin I, extracellular matrix components, and also several proinflammatory substances, such as interleukin-1β and interleukin-18 precursors.[32] Both cathepsin G and chymase activate connective tissue–activating peptide-III (CTAP-III), which is secreted by activated platelets, into an active neutrophil-activating peptide-2 (NAP-2), the chemokine responsible for further activation of neutrophils and their attraction to the site of injury. It was also shown that chymase is mainly responsible for CTAP-III/NAP-2 conversion, and that activated mast cells displayed a 1000-fold higher conversion rate than activated neutrophils.[40] Other natural substrates for both enzymes are the big endothelins (ETs), the precursors for several vasoconstrictors with different potency. Their cleavage products are 31-aa fragments, denoted Ets (1-31), that, apart from being among the most potent vasoconstrictors, also act as chemoattractants for neutrophils and monocytes.[41] Interestingly, the RCL of IRS-2 strongly resembles the chymase cleavage site in big ETs. The sequence in the P3–P3′ region consists of VPY–SLG for IRS-2 and VPY–GLG for big ETs. We can therefore hypothesize that this is the result of convergent evolution that created almost identical recognition sites that target both cathepsin G and chymase on otherwise unrelated proteins. To our knowledge, this is the first time that a parasitic protein is shown to inhibit host proteases by mimicking host endogenous substrates. Moreover, structural comparison revealed significant structural homology of IRS-2 with α-1-antichymotrypsin, which is supposed to be a natural vertebrate host regulator of cathepsin G and mast cell chymase.

Neutrophil-derived cathepsin G and, in particular, mast cell chymase are present at the early stages of an acute inflammatory response, probably with some level of redundancy and/or cooperation in their functions. By inhibiting both enzymes, IRS-2 can significantly alter a defensive response to tissue destruction caused by a tick bite. Indeed, we demonstrated that IRS-2 inhibits carrageenan-induced acute inflammation and cathepsin G and thrombin-induced platelet aggregation. According to recent findings relating to the biologic function of cathepsin G and mast cell chymase, we can hypothesize that both described effects may be mediated by proteolyticaly activated transducers and receptors during the early steps of inflammatory response that occurs after a tick bite. To explain the observed effects of IRS-2, we suggest that IRS-2 targets neutrophils, platelets, and mast cells at the early steps of their activation by preventing cathepsin G– or thrombin-driven platelet activation, which may lead to reduced CTAP-III release. Subsequently, IRS-2, similar to chymase inhibitors,[40] inhibits the cathepsin G– and chymase-catalyzed conversion of CTAP-III into NAP-2, resulting in impaired extravasation of neutrophils. The direct proof of the suggested, literature-based hypothesis exceeds the purpose of this work.

Here, we describe a mechanism of vertebrate host modulation that is novel not only for a tick salivary component, but to our knowledge for any parasite. Moreover, the dual specificity of IRS-2 for both cathepsin G and chymase is rare even among natural serine protease inhibitors in general. The saliva of *I ricinus* is a complex mixture of many proteins, including protease inhibitors with unknown specificity and function. This work describes a novel salivary protein that is unique among other salivary serpins of *Ixodes spp*—and proteases inhibitors in general—regarding both specificity and the mode of action in the vertebrate host. Therefore, the herein described findings contribute to our understanding of vertebrate physiology and parasite-host interaction, while the herein disclosed structural basis of IRS-2 activity can lead to the development of pharmaceutical applications in the near future.

Acknowledgments

We thank Ida Waern and Elin Rönnberg from the Swedish University of Agricultural Sciences for help with mast cell experiments and the Swedish Research Council for financial support. We also thank Jan Erhart (Institute of Parasitology, Biology Center of the Academy of Sciences of Czech Republic) for providing ticks; Anderson Sá-Nunes (Institute of Biomedical Sciences, University of São Paulo) for discussions and help with the experimental design; and NIAID intramural editor Brenda Rae Marshall for assistance. We also thank the anonymous reviewers for critical and constructive comments and the editorial team of *Blood* for efficiently processing the manuscript.

This work was supported by grant no. IAA600960811 from the Grant Agency of the Academy of Sciences of the Czech Republic, Research Center no. LC06009 from the Ministry of Education, Youth and Sports of the Czech Republic, and by grant no. Z60220518 of the Academy of Sciences of the Czech Republic. The structural element was supported by grant no. P207/10/2183 from the Grant Agency of the Czech Republic and research project Z40550506. Diffraction data were collected at beam line X12, EMBL-Hamburg outstation at DESY (Hamburg, Germany). I.M.B.F., J.M.C.R., and M.K. were supported by the Intramural Research Program of the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, NIH, and M.K. by a Jan Evengelista Purkyne fellowship of the Academy of Sciences of the Czech Republic.

Because I.M.B.F. and J.M.C.R. are government employees and this is a government work, the work is in the public domain in the United States. Notwithstanding any other agreements, the NIH reserves the right to provide the work to PubMedCentral for display and use by the public, and PubMedCentral may tag or modify the work consistent with its customary practices. One can establish

Authorship

Contribution: J.C., C.J.O., I.M.B.F., M.M., and M.K. designed and performed experiments, analyzed data, and wrote the manuscript; P.R. and Z.K. designed and performed experiments; and G.P., P.K., J.M.C.R., and J.K. designed experiments and analyzed data.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Michail Kotsyfakis, Institute of Parasitology, Biology Center of the Academy of Sciences of Czech Republic, Branisovska 31, 37005 Ceske Budejovice, Czech Republic; e-mail: kotsyfakis@paru.cas.cz.

References

1. Ribeiro JM, Makoul GT, Levine J, Robinson DR, Spielman A. Antihemostatic, anti-inflammatory, and immunosuppressive properties of the saliva of a tick, Ixodes dammini. *J Exp Med.* 1985;161(2):332-344.
2. Francischetti IM, Sa-Nunes A, Mans BJ, Santos IM, Ribeiro JM. The role of saliva in tick feeding. *Front Biosci.* 2009;14:2051-2088.
3. Deruaz M, Frauenschuh A, Alessandri AL, et al. Ticks produce highly selective chemokine binding proteins with anti-inflammatory activity. *J Exp Med.* 2008;205(9):2019-2031.
4. Kotsyfakis M, Sa-Nunes A, Francischetti IM, Mather TN, Andersen JF, Ribeiro JM. Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick, Ixodes scapularis. *J Biol Chem.* 2006;281(36):26298-26307.
5. Ribeiro JM, Alarcon-Chaidez F, Francischetti IM, et al. An annotated catalog of salivary gland transcripts from Ixodes scapularis ticks. *Insect Biochem Mol Biol.* 2006;36(2):111-129.
6. Chmelar J, Anderson JM, Mu J, Jochim RC, Valenzuela JG, Kopecky J. Insight into the sialome of the castor bean tick, Ixodes ricinus. *BMC Genomics.* 2008;9:233.
7. Francischetti IM, My Pham V, Mans BJ, et al. The transcriptome of the salivary glands of the female western black-legged tick, Ixodes pacificus (Acari: Ixodidae). *Insect Biochem Mol Biol.* 2005;35(10):1142-1161.
8. Potempa J, Korzus E, Travis J. The serpin superfamily of proteinase inhibitors: structure, function, and regulation. *J Biol Chem.* 1994;269(23):15957-15960.
9. Irving JA, Pike RN, Lesk AM, Whisstock JC. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res.* 2000;10(12):1845-1864.
10. Mulenga A, Khumthong R, Chalaire KC. Ixodes scapularis tick serine proteinase inhibitor (serpin) gene family: annotation and transcriptional analysis. *BMC Genomics.* 2009;10:217.
11. Mulenga A, Khumthong R, Blandon MA. Molecular and expression analysis of a family of the Amblyomma americanum tick, Lospins. *J Exp Biol.* 2007;219(18):3188-3198.
12. Sugino M, Imamura S, Mulenga A, 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(21-22):2844-2851.
13. 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(2):267-276.
14. Leboulle G, Crippa M, Decrem Y, et al. Characterization of a novel salivary immunosuppressive protein from Ixodes ricinus ticks. *J Biol Chem.* 2002;277(12):10083-10089.
15. Prevot PP, Adam B, Boudjeltia KZ, et al. Antihemostatic effects of a serpin from the saliva of the tick, Ixodes ricinus. *J Biol Chem.* 2006;281(36):26361-26369.
16. Prevot PP, Beschin A, Lins 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(12):3235-3246.
17. Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning: A Laboratory Manual.* 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Press; 1989.
18. Andersen JF, Francischetti IM, Valenzuela JG, Schuck P, Ribeiro JM. Inhibition of hemostasis by a high affinity biogenic amine-binding protein from the saliva of a blood-feeding insect. *J Biol Chem.* 2003;278(7):4611-4617.
19. Kováľová Z, Chmelař J, Šanda M, Brynda J, Mareš M, Řezáčová P. Crystallization and diffraction analysis of the serpin, IRS-2, from the hard tick, *Ixodes ricinus*. *Acta Crystallogr F Struct Biol Cryst Commun.* In press.
20. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr.* 2006;62(8):859-866.
21. Vagin A, Teplyakov A. An approach to multi-copy search in molecular replacement. *Acta Crystallogr D Biol Crystallogr.* 2000;56(12):1622-1624.
22. Baumann U, Bode W, Huber R, Travis J, Potempa J. Crystal structure of cleaved equine leucocyte elastase inhibitor determined at 1.95 A resolution. *J Mol Biol.* 1992;226(4):1207-1218.
23. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 1997;53(3):240-255.
24. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994;50(5):760-763.
25. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004;60(12,1):2126-2132.
26. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr.* 2001;57(1):122-133.
27. Davis IW, Leaver-Fay A, Chen VB, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 2007;35(Web Server issue):W375-W383.
28. De Lano WL. *The PyMOL Molecular Graphics System.* San Carlos, CA: DeLano Scientific LLC; 2002.
29. Holm L, Sander C. Searching protein structure databases has come of age. *Proteins.* 1994;19(3):165-173.
30. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck—a program to check the stereochemical quality of protein structures. *J Appl Cryst.* 1993;26:283-291.
31. Levy L. Carrageenan paw edema in the mouse. *Life Sci.* 1969;8(11):601-606.
32. Pejler G, Abrink M, Ringvall M, Wernersson S. Mast cell proteases. *Adv Immunol.* 2007;95:167-255.
33. Sambrano GR, Huang W, Faruqi T, Mahrus S, Craik C, Coughlin SR. Cathepsin G activates protease-activated receptor-4 in human platelets. *J Biol Chem.* 2000;275(10):6819-6823.
34. Francischetti IM, Saliou B, Leduc M, et al. Convulxin, a potent platelet-aggregating protein from Crotalus durissus terrificus venom, specifically binds to platelets. *Toxicon.* 1997;35(8):1217-1228.
35. Betts MJ, Sternberg MJ. An analysis of conformational changes on protein-protein association: implications for predictive docking. *Protein Eng.* 1999;12(4):271-283.
36. Andersson MK, Enoksson M, Gallwitz M, Hellman L. The extended substrate specificity of the human mast cell chymase reveals a serine protease with well-defined substrate recognition profile. *Int Immunol.* 2009;21(1):95-104.
37. Rehault S, Brillard-Bourdet M, Juliano MA, Juliano L, Gauthier F, Moreau T. New, sensitive fluorogenic substrates for human cathepsin G based on the sequence of serpin-reactive site loops. *J Biol Chem.* 1999;274(20):13810-13817.
38. Zarbock A, Polanowska-Grabowska RK, Ley K. Platelet-neutrophil-interactions: linking hemostasis and inflammation. *Blood Rev.* 2007;21(2):99-111.
39. McDougall JJ, Zhang C, Cellars L, Joubert E, Dixon CM, Vergnolle N. Triggering of proteinase-activated receptor 4 leads to joint pain and inflammation in mice. *Arthritis Rheum.* 2009;60(3):728-737.
40. Schiemann F, Grimm TA, Hoch J, et al. Mast cells and neutrophils proteolytically activate chemokine precursor CTAP-III and are subject to counter-regulation by PF-4 through inhibition of chymase and cathepsin G. *Blood.* 2006;107(6):2234-2242.
41. Cui P, Tani K, Kitamura H, et al. A novel bioactive 31-amino-acid, endothelin-1, is a potent chemotactic peptide for human neutrophils and monocytes. *J Leukoc Biol.* 2001;70(2):306-312.

Isolation of IRS-2 cDNA. Ten different arthropod serpins whose sequences were publicly disclosed at that time were used in multiple alignments to reveal two conserved domains in arthropod serpins. A radioactively labeled DNA probe was prepared from the PCR product, obtained with degenerated primers designed based on the sequences of these two domains (Fwd: 5′-AACGC(C/T)(A/G)TCTACTT(C/T)AA(A/G)GG-3′, Rev: 5′-GGATGAAGAA (G/C)A(G/T)GAA(A/T/C/G)GG-3′) using a DecaLabel DNA kit (Fermentas Life Sciences). The probe was used for the screening of a λ phage cDNA library constructed from the salivary glands of partially fed (d 5 after attachment) *I. ricinus* female adults. The cDNA library was prepared using the SMART™ cDNA library construction kit (ClonTech) according to the manufacturer's protocol. Phages were plated on 150-mm Petri dishes, and the print was taken onto the nylon membrane Hybond™ (Amersham Biosciences). The DNA was linked to the membranes using Stratalinker™ 1800 (Stratagene), the membranes were preincubated (30 min at 63.5°C with 2% salmon sperm in 0.1% SDS, 300 mM NaCl, 30 mM natrium citrate, pH 7), and then 25 µl of the denatured (97°C; 10 min) probe was added to each hybridization tube. Hybridization was done overnight at 63.5°C. After hybridization, the membranes were washed 3 × 30 min with decreasing concentrations of NaCl (450, 150, and 75 mM) in the presence of 0.1% SDS. After washing, the membranes were dried and exposed on Kodak X-Omat™ LS film using a Kodak X-Omat™ cassette at –80°C. Positive single clones were isolated after a second screen, excised, and subcloned into pTriplEx vector according to the manufacturer's protocol and sequenced in an ABI PRISM 3130xl sequencer (Applied Biosystems).

Bioinformatics tools. BLASTx algorithm on the server http://blast.ncbi.nlm.nih.gov/Blast.cgi [1] was used for similarity searches of IRS-2 with other members of the serpin superfamily. The presence of signal sequences on the proteins was determined using SignalP 3.0 server http://www.cbs.dtu.dk/services/SignalP/, [2] and their mol wt and pI were determined using ProtParam program on ExPASy proteomic server http://www.expasy.org/tools/protparam.html.[3]

cDNA preparation and quantitative PCR. Ticks used for tissue isolation were fed on guinea pigs complying with Act No. 207/2004 Coll. and approval AVCR 51/2005 given by the respective committee of Czech Academy of Sciences. Tissues used for RNA isolation were dissected from five feeding stages of adult females: for unfed ticks, 20 females were used; for next stages, 10–15 females were dissected. RNA was isolated with a NucleoSpin RNA XS kit (Macherey-Nagel). The amount of total RNA used for cDNA preparation was set to 0.3 μg for all samples. cDNA was prepared using an Avian Enhanced First-strand Synthesis Kit and diluted 20 times in deionized water to achieve optimal amplification conditions (linear amplification between 20th and 30th cycle). The program for quantitative PCR was: 95°C for 10 min and then 40 cycles of 95°C for 15 s and 60°C for 60 s. All samples were prepared in triplicates. Dual-labeled probe from Universal probe library (Roche) and IRS-2-specific primers were used for quantitative RT-PCR. Data were obtained with Rotor-Gene RG3000 (Corbett Research) and analyzed using the Pfaffl equation. [4] Based on our previous experiences with various housekeeping genes (data not shown), we decided on the tick ferritin (NCBI GenBank Acc. No. AF068224) [5] for use as a reference gene.

Preparation of recombinant IRS-2. Primers for amplifying IRS-2 cDNA for overexpression in bacteria were designed to amplify the gene without its signal peptide sequence and to introduce restriction sites for *Nde*I and *Xho*I on its 5′ and 3′ ends, respectively, for subcloning purposes. To improve gene expression levels in bacteria, codons for rare tRNAs in *Escherichia coli* were replaced by introducing synonymous mutations into the forward

primer. (Fwd: 5′-GCCATATGCAAGAAGAAGCCAAGCTCACCAAGGCCAACAACCG T-3′, Rev: 5′-GCCTCGAGTTATCACAGCTTGTTAACCTGTCCCACGAAAAAAATGT CATC-3′). The gene was PCR amplified using cDNA isolated from adult female tick salivary glands 5 d after attachment to guinea pigs as a template (see below). The amplified product was cloned into Zero Blunt™ TOPO PCR vector (Invitrogen) and sequenced to verify the integrity of the frame. Then the IRS-2 gene was subcloned into the expression vector pET 17b (Novagen) using *Nde*I and *Xho*I restriction enzymes (Invitrogen). Gene overexpression was done in BL21(DE3)pLysS cells (Invitrogen) grown in ampicillin (100 μg/ml) and chloramphenicol (34 μg/ml) selection and isopropyl β-D-1-thiogalactopyranoside (IPTG) was used for gene expression induction (0.5 mM). IPTG was added when the OD_{600} of the culture reached approximately 0.7 and bacterial growth in the presence of IPTG continued for 3.5 h. Subsequently, the cells were harvested by centrifugation (10,000 × g, 5 min), and the resulting pellet was resuspended in 20 mM Tris-HCl, pH 8. The cells were disrupted by sonication (medium amplitude, 3 × 1 min), and centrifuged (10,000 × g; 10 min), resulting in a pellet that, among other insoluble material, contained the rIRS-2 in inclusion bodies. The pellet was resuspended in inclusion bodies isolation buffer (20 mM Tris-HCl, 1% Triton X-100; pH 8) and sonicated. Next, the suspension was centrifuged (10,000 × g; 10 min), and the resulting pellet was washed four times with 20 mM Tris-HCl, pH 8, resulting in a final pellet that contained mainly inclusion bodies of nonsoluble, overexpressed recombinant IRS-2. Inclusion bodies were dissolved by stirring in 6 M guanidine hydrochloride, pH 8, with 10 mM dithiotreitol for 1 h at room temperature. The solution was centrifuged at 10,000 × g for 15 min, the resulting pellet was discarded, and the resulting supernatant containing the denatured protein content of the inclusion bodies was dissolved into 150-fold volume of refolding buffer (20 mM Tris-HCl, 0.25 M L-arginine; pH 8). The buffer remained overnight at 4°C, and the precipitated protein was removed by filtration through filtration paper and Steritop-GP (Millipore). The protein solution was then

concentrated using a stirred chamber concentrator (Millipore) down to the volume of 10 ml. The concentrated protein was dialyzed against 20 mM Tris-HCl, pH 8. Refolded and concentrated IRS-2 was purified on a Mono-Q column with 0–1 M gradient of NaCl using ÄKTA FPLC system (Pharmacia) with purity > 99% according to SDS-PAGE. During Mono Q purification step, a major symmetric peak, corresponding to properly folded IRS-2 (as verified on SDS-PAGE and by comparing inhibitory activity of isolated fractions) was separated from other contaminants and potentially misfolded protein. Endotoxin was removed by phase separation using Triton X-114 (Sigma-Aldrich), and traces of detergent were removed using SM2 BioBeads (BioRad) as described elsewhere [6]. Endotoxin level was estimated using Limulus amobecyte lysate QCL-1000® (Lonza), following the manufacturer's instructions. Endotoxin contamination did not exceed 0.1 EU/ml in any of the protein samples used for immunological assays, and the protein samples did not induce production of nitric oxide by Dendritic Cells (data not shown), further verifying the LPS decontamination in the IRS-2 preparation.

Serine protease inhibition assays. All assays were performed at 30°C in triplicates. For the initial screen, 400 nM of protein was pre-incubated with each enzyme for 10 min before the addition of the corresponding substrate; *t*-test was used for statistical analysis of the observed inhibition in the presence of 400 nM IRS-2, and statistical significance was considered when $p < 0.05$ when comparing the enzymatic activity in the presence or absence of the inhibitor. For IC_{50} estimation of IRS-2, increasing concentrations of the inhibitor were pre-incubated for 10 min with the enzymes targeted by IRS-2, and the reaction was started by adding the corresponding substrates. All experiments were performed in triplicate (for each enzyme and each concentration of IRS-2). The mean percentage of remaining enzymatic activity in the presence of various IRS-2 concentrations compared with the control enzymatic activity (in the absence of IRS-2) was plotted against the concentration of IRS-2 used in the assays and in logarithmic scale. Finally, sigmoidal fit of the data gave the estimate of the IC_{50} of IRS-2 for

the various enzymes. To determine whether IRS-2 is a fast- or slow-binding inhibitor for the targeted enzymes, either the enzyme was pre-incubated with the inhibitor for 10 min and the addition of substrate followed, or the substrate was pre-incubated with the inhibitor for 10 min and the addition of the enzyme followed. In both cases, the hydrolysis rate of the fluorescent substrate was estimated from the slope that results from the linear fit (arbitrary fluorescence units per sec; $r^2 > 0.95$) of the data (each experiment was performed in triplicate, and the mean of the three experiments was plotted).

All enzymes used were of human origin, purified or recombinant. Thrombin, α-chymotrypsin, plasmin, and chymase were purchased from Sigma; β-tryptase was purchased from Promega; factor Xa was purchased from EMD Biosciences; factor XIIa was purchased from Haematologic Technologies Inc.; kallikrein was purchased from Fitzgerald Industries International; elastase was purchased from Elastin Products; cathepsin G, Factor XIa, uPA, and tPA were from Molecular Innovations; matriptase was from R&D Systems; proteinase 3 was from Merck; and sequencing-grade trypsin was purchased from Roche. The amount of targeted enzyme used in each assay is shown in Table I. Assay buffers were: for elastase, proteinase 3 and chymase, 50 mM Hepes buffer, pH 7.4, 100 mM NaCl, 0.01% Triton X-100; for trypsin and α-chymotrypsin, factor XIa, factor XIIa, and thrombin, 50 mM Tris-HCl, pH 8, 150 mM NaCl, 20 mM $CaCl_2$, 0.01% Triton X-100; for β-tryptase, 50 mM Tris-HCl, pH 8, 50 mM NaCl, 0.05% Triton X-100; for kallikrein, matriptase, and plasmin, 20 mM Tris–HCl, pH 8.5, 150 mM NaCl, 0.02% Triton X-100; for factor Xa, 20 mM Tris-HCl, pH 8, 200 mM NaCl, 5 mM $CaCl_2$, 0.1% BSA; for uPA and tPA, 20 mM Tris-HCl, pH 8.5, 0.05% Triton X-100; and for cathepsin G, 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.01% Triton X-100. The substrates used were Suc-Ala-Ala-Pro-Val-AMC for elastase and proteinase 3, Boc-Asp-Pro-Arg-AMC for thrombin and plasmin, Boc-Gln-Ala-Arg-AMC for trypsin, factor XIa and uPA (Sigma), Boc-Phe-Ser-Arg-AMC for β-tryptase, Suc-Leu-Leu-Val-Tyr-AMC for chymase

(Bachem Bioscience, Inc.), Suc-Ala-Ala-Pro-Val-AMC for α-chymotrypsin and chymase (EMD Biosciences), and methylsulfonyl-D-cyclohexylalanyl-Gly-Arg-AMC acetate for factor Xa, factor XIIa, t-PA, matriptase, and kallikrein (American Diagnostica Inc.). All substrates were used in 250 μM final concentration in all assays. Substrate hydrolysis rate was followed in a Spectramax Gemini XPS 96-well plate fluorescence reader (Molecular Devices) using 365 nm excitation and 450 nm emission wavelength with a cutoff at 435 nm.

MPO activity. MPO was performed as an index of granulocyte recruitment into the hind paw inflamed tissue according to a modified protocol of Bradley et al. [7]. Briefly, the injected paws were cut at 4 h post-injection (the time point at which edema peaks) and weighed prior to their homogenization in 2 ml of 0.5% hexadecyl-trimethylammonium bromide phosphate-buffered solution (pH 6.0). The homogenates were then centrifuged at 13,000×g for 3 min, and the supernatants were collected and frozen until use. Three aliquots of each supernatant were then transferred into 96-well plates before the addition of a solution containing o-dianisidine (0.2 mg/ml) and H_2O_2 (0.0006%). In parallel, dilutions of pure myeloperoxidase (Sigma) were used for the construction of a standard curve (OD as a function of units of enzyme activity). OD readings at 450 nm were taken at 1 min (time point corresponding to the linear portion of the enzymatic reaction) using a spectrofluorometer linked to SOFTmax Pro 3.0 software (Molecular Devices). MPO activity detected in the paws was expressed as units of enzyme/g of tissue. A unit of MPO activity was defined as that converting 1 μm of hydrogen peroxide to water in 1 min at 22°C. Alternatively The MPO activity of samples was compared to a standard curve of neutrophils [8]. Briefly, 10 ml of sample was mixed with 200 ml of 50mM phosphate buffer pH 6.0, containing 0.167mgml^{-1} O-dianisidine dihydrochloride and 0.0005% hydrogen peroxide. The results are presented as the MPO activity (number of neutrophils per mg of tissue).

Analysis of PCMC derived proteases activity. On d 2 after the change of growth medium, PCMCs were washed and resuspended in Tyrode's buffer (137 mM NaCl, 2.7 mM KCl, 0.36

mM Na_2HPO_4, 5.55 mM glucose; pH 7.4) to a concentration of 10^6 cells/ml. 10^5 of PCMCs were incubated at 37°C and 5% CO_2 for 40 min with different concentrations of IRS-2 in a total volume of 200 µl. Then the cells were activated with 2 µM of ionomycin in the presence or absence of IRS-2 and incubated for additional 30 min. Untreated cells with ionomycin were used as a negative control, and cells treated with ionomycin only (in the absence of IRS-2) were used as a positive control. For enzymatic assays 20 µl (final concentration 100 µM) of the chromogenic substrate MeO-Suc-Arg-Pro-Tyr-pNA (Chromogenix) was added after 30 min of preincubation with IRS-2, and the substrate hydrolysis rate was measured at 405 nM using Sunrise™ reader (TECAN) with the methodology described above. For western blot analysis, cells were preincubated for 30 min with different concentrations of IRS-2, centrifuged at 1100 rpm for 7 min, and the supernatant transferred into a clean tube. The pellet was resuspended in 20 µl of 1× Laemmli sample buffer (2% SDS, 5% 2-mercaptoehtanol, 10% glycerol, 0.002% bromophenol blue, 60 mM Tris-HCl). Twenty µl of both supernatant and resuspended cells was loaded onto 12% SDS polyacrylamide gel and transferred to nitrocellulose filters. Filters were blocked with 5% BSA in TBS-T (10 mM Tris-HCl, pH 7.7, 0.15 M NaCl, 2 % Tween-20) for 1 h at room temperature. Custom polyclonal rabbit antisera raised against mMCP-4 were diluted 1:500 in TBS-T and incubated overnight at 4°C with the filters. After three washing steps with TBS-T for 3 × 5 min, donkey anti-rabbit Ig conjugated to horseradish peroxidase (Amersham Biosciences) was diluted 1:2000 in TBS-T and incubated for 1 h with the filters. After 4 washing steps with TBS-T for 4 × 5 min, an ECL detection kit (Amersham Biosciences) was used for results visualization.

Results

Characterization of chymase, cathepsin G and thrombin inhibition by IRS-2

IRS-2 inhibited cathepsin G and chymase in equimolar ratio. Furthermore, when lowering the concentrations of cathepsin G or chymase, lower concentrations of IRS-2 were necessary for

50% inhibition of these two enzymes. In fact, there was a linear correlation between the amount of enzyme used in the assays and the observed IC_{50}. Therefore, IRS-2 can be considered as a tight binding inhibitor of these two enzymes (Fig. S1 A, B).

On the contrary, IRS-2 behaves like a classical Michaelis-Menten inhibitor (large excess of inhibitor compared with the amount of the enzyme used in the assays) for thrombin (Table I; Fig. S1 C). Furthermore, inhibition of chymase was higher when the enzyme was preincubated with IRS-2, and a competition between the inhibitor and the substrate was observed when they were added simultaneously to the reaction, suggesting slow binding (Table I; Fig. S2 A). Preincubation of IRS-2 with cathepsin G did not change the observed inhibition rate, indicating fast binding (Table I; Fig. S2 B).

Preincubation of IRS-2 with thrombin improved the observed enzymatic inhibition, indicating that IRS-2 is a slow-binding inhibitor of thrombin with low ability to compete with the artificial substrate used for detection of enzymatic activities (Table I; Fig. S2 C).

REFERENCES

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* Oct 5 1990;215(3):403-410.
2. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol.* Jul 16 2004;340(4):783-795.
3. Gasteiger E, Hoogland C, Gattiker A, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, ed. *The Proteomics Protocols Handbook*: Humana Press; 2005:571-607.
4. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-

PCR. *Nucleic Acids Res.* May 1 2001;29(9):e45.

5. Kopacek P, Zdychova J, Yoshiga T, Weise C, Rudenko N, Law JH. Molecular cloning, expression and isolation of ferritins from two tick species--Ornithodoros moubata and Ixodes ricinus. *Insect Biochem Mol Biol.* Jan 2003;33(1):103-113.

6. Aida Y, Pabst MJ. Removal of endotoxin from protein solutions by phase separation using Triton X-114. *J Immunol Methods.* Sep 14 1990;132(2):191-195.

7. Bradley PP, Priebat DA, Christensen RD, Rothstein G. Measurement of cutaneous inflammation: estimation of neutrophil content with an enzyme marker. *J Invest Dermatol.* Mar 1982;78(3):206-209.

8. Guerrero AT, Verri WA, Jr., Cunha TM, et al. Hypernociception elicited by tibio-tarsal joint flexion in mice: a novel experimental arthritis model for pharmacological screening. *Pharmacol Biochem Behav.* Jun 2006;84(2):244-251.

9. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr.* Aug 2006;62(Pt 8):859-866.

Table S1. Crystallographic dataset and refinement statistics

Data collection statistics	
Space group	$P4_3$
Cell parameters (Å)	84.6, 84.6, 124.4
Number of molecules in AU	2
Wavelength (Å)	0.953
Resolution (Å)	30.0 – 1.80 (1.86 – 1.80)
Number of unique reflections	80,249 (7,790)
Redundancy	3.8 (3.3)
Completeness (%)	99.6 (97.3)
R_{merge} [a]	4.6 (35.8)
Average I/σ(I)	32.8 (2.3)
Wilson B ($Å^2$)	23.2
Refinement statistics	
Resolution range (Å)	30.0 – 1.80 (1.86 – 1.80)
No. of reflections in working set	76,177 (5,415)
No. of reflections in test set	4,026 (285)
R value (%) [b]	17.1 (24.7)
R_{free} value (%) [c]	21.2 (31.5)
RMSD bond length (Å)	0.011
RMSD angle (°)	1.22
Number of atoms in AU	7,003
Number of protein atoms in AU	6,111
Number of water molecules in AU	875
Mean B value protein / solvent ($Å^2$)	23.6/38.6
Ramachandran plot statistics	
Residues in favored regions (%)	91.3
Residues in allowed regions (%)	5.7

Data in parentheses refer to the highest-resolution shell.

$$R_{merge} = \Sigma_{hkl}\Sigma_i |I_i - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl),$$

where the $I_i(hkl)$ is an individual intensity of the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of reflection hkl with summation over all data.

$$R \text{ value} = ||F_o| - |F_c|| / |F_o|,$$

where F_o and F_c are the observed and calculated structure factors, respectively. R_{free} is equivalent to R value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process.

Table S2. Concentration of used enzymes and remaining enzymatic activity. The values represent the mean remaining enzymatic activity in the presence of IRS-2 with the standard error of the mean. Enzymes in bold were inhibited with a statistical significance (*t*-test; $P < 0.05$).

Enzyme tested	Final concentration of enzyme used [nM]	% remaining enzymatic activity in the presence of 0.4µM IRS-2
kallikrein	0.8	92.5 ± 5.2
tryptase	0.27	106.8 ± 11.2
uPA	0.7	87.6 ± 2.9
factor Xa	0.8	97.8 ± 0.6
granzyme B	20	105.8 ± 2.7
matriptase	1.2	98,6 ± 3
elastase	0.01	101.1 ± 2.1
proteinase 3	10	91.45 ± 1.3
a-chymotrypsin	**0.05**	**0.3 ± 0.2**
chymase	**1.8**	**1 ± 0.5**
factor XIIa	1.2	96,5 ± 1.6
factor XIa	0.06	105.3 ± 6.7
plasmin	0.25	103.2 ± 4.8
thrombin	**0.01**	**11.6 ± 0.9**
trypsin	**0.25**	**55.2 ± 4**
cathepsin G	**10**	**4.9 ± 0.9**

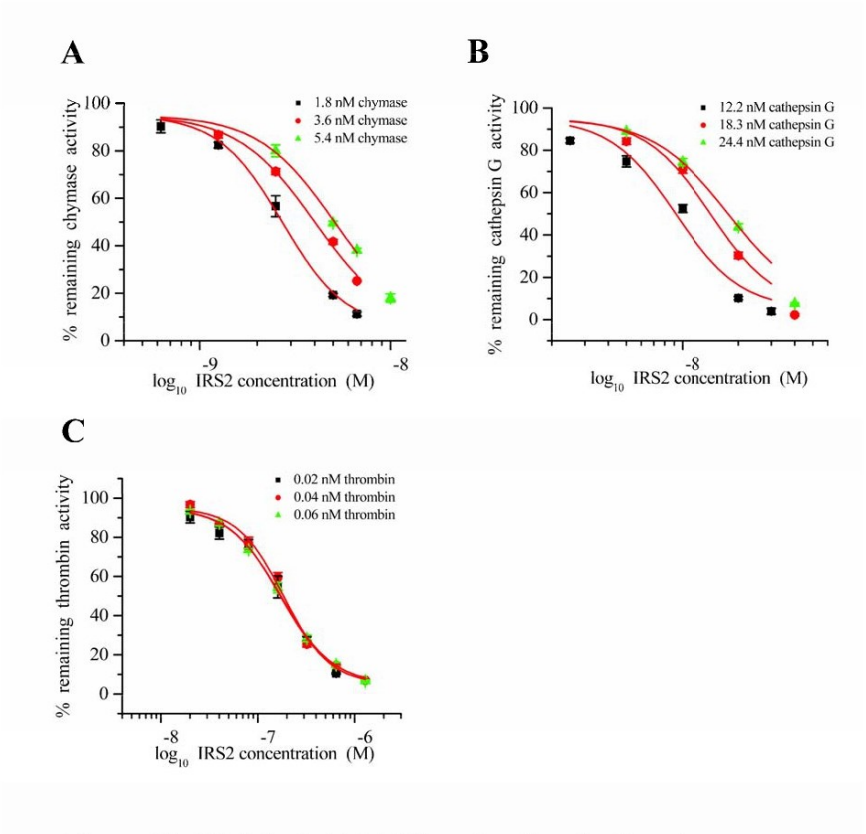


Figure S1. IRS-2 is tight inhibitor of cathepsin G and chymase and classical inhibitor of thrombin. The inhibition curve changed when using different concentrations of chymase (A) and cathepsin G (B).The increase of used enzyme had no effect on the inhibitory profile of IRS-2 in the case of thrombin (C).

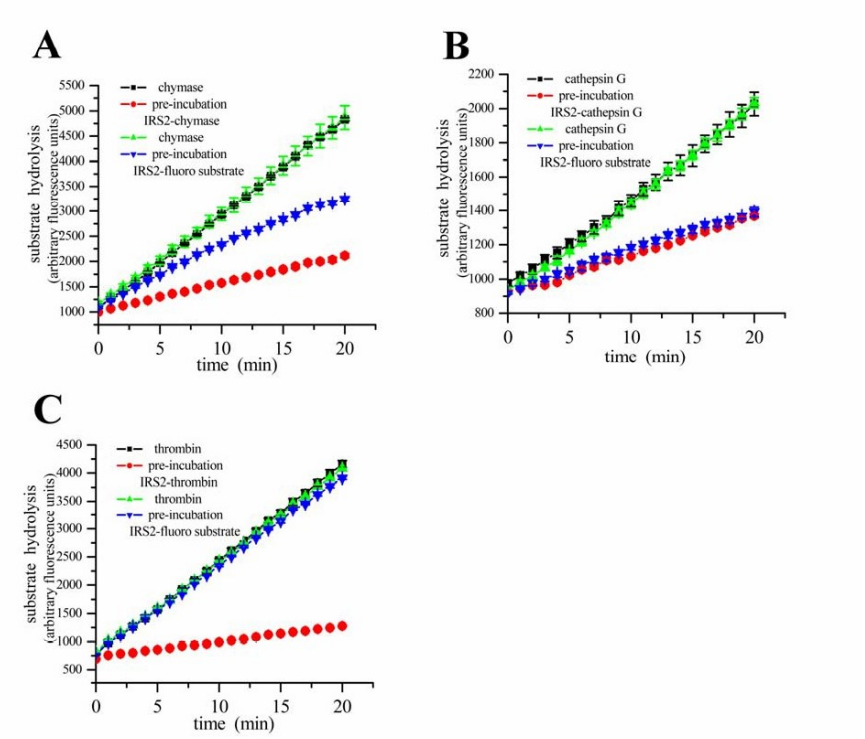


Figure S2. IRS-2 is a fast inhibitor of cathepsin G, slow inhibitor of chymase, which can compete with the substrate for the enzyme and slow inhibitor of thrombin, which cannot compete with the substrate for the enzyme.

5.3. Publikace č. 3: Crystallization and diffraction analysis of the serpin IRS-2 from the hard tick *Ixodes ricinus*

V této publikaci je popsán detailní postup krystalizace serpinu IRS-2 z klíštěte obecného (*I. ricinus*) a regulace tohoto procesu působením proteolytických enzymů na IRS-2. Jedná se o doprovodnou studii k publikaci č. 2. IRS-2 patří do serpinové rodiny proteasových inhibitorů s unikátním mechanismem účinku. Tyto proteiny fungují jako „sebevražedné" inhibitory vystavující na povrchu reaktivního centra smyčku, která mimikuje substrát cílové proteasy. Po štěpení této smyčky je inhibovaná proteasa kovalentně vázána inhibitorem, nicméně může dojít i k uvolnění proteasy z komplexu, přičemž serpin zůstává v inaktivované konformaci se štěpenou smyčkou. Posledně uvedená situace byla pozorována v případě IRS-2.

IRS-2 byl připraven expresí v bakterii *E. coli* ve formě inkluzních tělísek. Aktivní inhibitor byl získán renaturací a purifikován jako intaktní jednořetězcový protein. Tento materiál byl použit pro vyhledání krystalizačních podmínek. Úvodní vysokokapacitní testování identifikovalo primární krystalizační podmínku, která byla dále optimalizována na výslednou směs, ve které narostl krystal difraktující na 1,8 Å. Nicméně při analýze krystalů pomocí SDS-PAGE se zjistilo, že obsahují molekulární formu IRS-2, která byla specificky modifikována limitovanou proteolýzou. Proteomická analýza hmotnostní spektrometrií identifikovala štěpenou vazbu mezi Tyr341 a Ser342. Toto místo se nachází ve smyčce reaktivního centra a zřejmě odpovídá místu přirozeného štěpení cílovými proteasami při jejich ataku. Proteolytické štěpení vedlo ke konformačnímu přechodu molekuly IRS-2 z méně stabilního S-(stresového) stavu na více stabilní R-(relaxovaný) stav, který lépe krystalizoval.

Studie se dále zaměřila na určení proteolytické aktivity, která štěpení IRS-2 katalyzovala a která pravděpodobně pocházela z expresního systému jako stopová kontaminace. K měření této aktivity byly využity fluorescenční peptidové substráty specifické pro relevantní proteasy a jejich hydrolýza byla testována za pH odpovídajícího krystalizačním podmínkám. Aktivitní test se dále prováděl v přítomnosti nízkomolekulárních selektivních proteasových inhibitorů k detailní identifikaci typu proteas. Tímto způsobem byly prokázány stopové aktivity několika cysteinových a serinových proteas v preparátu IRS-2. V dalším kroku bylo ověřeno působení uvedených proteas na IRS-2 pomocí simulovaného fragmentačního testu. V přítomnosti selektivních

proteasových inhibitorů pro detekované kontaminační proteasy byla potlačena konverze IRS-2 z formy S na R.

Poznatky z této studie mohou být využity při krystalizaci serpinů (a obecně řady proteinů citlivých na proteolýzu) a poukazují na důležitou roli proteas přítomných ve stopovém množství, které mohou regulovat molekulární formu serpinu. Proces konverze na relaxovanou formu R může být přínosem a umožnit efektivní získání krystalů. A naopak v případech, kdy je přítomnost kontaminujících proteas nežádoucí, lze jejich aktivitu analyzovat a účinně potlačit použitím nízkomolekulárních selektivních proteasových inhibitorů v krystalizačních experimentech.

Můj podíl na práci spočíval v: 1) návrhu a provádění biochemických a enzymologických testů pro identifikaci proteas štěpících serpin IRS-2, 2) realizaci krystalizačních experimentů a optimalizaci krystalů IRS-2 pro strukturní analýzu a 3) přípravě manuskriptu.

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and diffraction analysis of the serpin IRS-2 from the hard tick *Ixodes ricinus*

Zuzana Kovářová,[a] Jindřich Chmelař,[b] Miloslav Šanda,[a] Jiří Brynda,[a,c] Michael Mareš[a] and Pavlína Řezáčová[a,c]*

[a]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., Flemingovo nám. 2, 16610 Praha 6, Czech Republic, [b]Institute of Parasitology, Biology Centre of the Academy of Sciences of the Czech Republic, Czech Republic, and [c]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, v.v.i., Vídeňská 1083, 142 20 Praha 4, Czech Republic

Correspondence e-mail: rezacova@uochb.cas.cz

Received 7 June 2010
Accepted 11 August 2010

IRS-2 from the hard tick *Ixodes ricinus* belongs to the serpin family of protease inhibitors. It is produced in the salivary glands of the tick and its anti-inflammatory activity suggests that it plays a role in parasite–host interaction. Recombinant IRS-2 prepared by heterologous expression in a bacterial system was crystallized using the hanging-drop vapour-diffusion method. The crystals belonged to the primitive tetragonal space group $P4_3$ and diffracted to 1.8 Å resolution. Mass-spectrometric and electrophoretic analyses revealed that IRS-2 was cleaved by contaminating proteases during crystallization. This processing of IRS-2 mimicked the specific cleavage of the serpin by its target protease and resulted in a more stable form (the so-called relaxed conformation), which produced well diffracting crystals. Activity profiling with specific substrates and inhibitors demonstrated traces of serine and cysteine proteases in the protein stock solution.

1. Introduction

Serpins (serine protease inhibitors) are a broadly distributed family of protease inhibitors (Irving *et al.*, 2000). The majority of serpins specifically inhibit serine proteases, but some serpins that inhibit papain-like cysteine proteases (Schick *et al.*, 1998) and caspases (Ray *et al.*, 1992) have also been identified. In rare cases, the serpins have lost their inhibitory function and act, for example, as hormone transporters (Pemberton *et al.*, 1988), chaperones (Nagata, 1996) or tumour suppressors (Zou *et al.*, 1994).

Serpins are relatively large molecules (comprising 330–500 amino-acid residues) which act as 'single-use' or 'suicide' inhibitors that undergo an extensive conformational change to inhibit proteases (Huntington *et al.*, 2000). The serpin family is structurally well characterized; over 70 serpin structures have been determined (reviewed, for example, in Law *et al.*, 2006). The remarkable structural change is induced by proteolysis of the reactive-centre loop (RCL), which in the native state (also called the S state) has an extended conformation that protrudes from the serpin domain. Upon cleavage, the amino-terminal part of the RCL inserts into the central β-sheet to form an additional β-strand. This structural rearrangement is crucial for protease inhibition and results in the so-called R state of serpin, which is more stable compared with the S state. In the final serpin–protease complex the protease remains covalently linked to the serpin. However, the protease can escape this trap and dissociate over time, leaving active protease and inactive serpin with its RCL cleaved in a substrate-like manner (Huntington *et al.*, 2000).

Ticks are blood-feeding parasites of a variety of vertebrates, including domestic animals and humans. They are important (second only to mosquitoes) pathogen vectors worldwide (Sauer *et al.*, 1995). The hard tick *Ixodes ricinus* is the most important European vector of *Borrelia burgdorferi* s.l., the agent of Lyme disease and tick-borne encephalitis virus.

In hard ticks, serpins have been identified in all tissues and have been proposed to be involved in both tick physiology and tick–host interaction. The salivary glands in particular produce a number of serpins, which may play roles in the modulation of immune response,

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

No. of crystals	1
Beamline	X12, DESY, Germany
Wavelength (Å)	0.953
Detector	MAR Mosaic 225
Crystal-to-detector distance (mm)	185
Rotation range per image (°)	0.5
Total rotation range (°)	54
Exposure time per image (s)	10
Resolution range (Å)	30.0–1.80 (1.86–1.80)
Space group	$P4_2$
Unit-cell parameters (Å)	$a = b = 84.6$, $c = 124.4$
Mosaicity (°)	0.33
Total no. of measured intensities†	301874
No. of unique reflections	80249 (7790)
Multiplicity	3.8 (3.3)
Average $I/\sigma(I)$	32.8 (2.3)
Completeness (%)	99.6 (97.3)
$R_{\rm merge}$‡ (%)	4.6 (35.8)
$R_{\rm r.i.m.}$§ (%)	7.2 (53.7)
$R_{\rm p.i.m.}$¶ (%)	3.7 (29.2)
Overall B factor from Wilson plot (Å^2)	23.2

† The criterion used for observed reflections was $I/\sigma(I) > 0$. ‡ $R_{\rm merge} = 100\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is an individual intensity of the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of reflection hkl with summation over all data. § $R_{\rm r.i.m.} = 100\sum_{hkl}[N/(N-1)]^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where N is the number of times the given reflection hkl was observed (Weiss, 2001). ¶ $R_{\rm p.i.m.} = 100\sum_{hkl}\sum_i I_i(hkl) / \sum_{hkl}\sum_i I_i(hkl)$, where N is the number of times the given reflection hkl was observed (Weiss, 2001).

coagulation, complement regulation and inflammation (Prevot *et al.*, 2009). Tick serpins are thus considered to be promising antigens for the induction of host protective immunity against ticks (Andreotti *et al.*, 2002; Imamura *et al.*, 2005; Sugino *et al.*, 2003).

The serpin IRS-2 has recently been identified in the salivary glands of the tick *I. ricinus*. Functional characterization has shown that its protease-inhibitor activity is directed against chymotrypsin-like proteases, resulting in an anti-inflammatory effect (Chmelar *et al.*, manuscript in preparation).

No crystal structure of a serpin isolated from a parasitic organism is available to date. The closest homologues of IRS-2 among the serpin structures deposited in the Protein Data Bank are equine leukocyte elastase inhibitor (PDB code 1hle; Baumann *et al.*, 1992) and human squamous cell carcinoma antigen 1 (PDB entry 2zv6; Zheng *et al.*, 2009), with 35 and 32% sequence identity, respectively.

In order to gain structural information on this newly discovered serpin, we initiated structural studies on this protein; here, we present the crystallization of recombinant IRS-2, analysis of its diffraction and an initial molecular-replacement solution.

2. Materials and methods

2.1. Protein expression and purification

The coding sequence of IRS-2 (UniProt entry Q06B74) was amplified from cDNA isolated from the salivary glands of an *I. ricinus* adult female and was cloned into pET-17b vector (Novagen).

Details of cloning and purification have been described elsewhere (Chmelar *et al.*, manuscript in preparation). The 377 amino-acid residue IRS-2 protein was overexpressed in *Escherichia coli* BL21 (DE3) pLysS cells (Invitrogen) at 310 K upon induction by 0.5 m*M* IPTG. The expressed protein accumulated in inclusion bodies, which were separated. The inclusion bodies were dissolved in 6 *M* guanidine hydrochloride and the supernatant was diluted into a 150-fold volume of refolding buffer (20 m*M* Tris–HCl pH 8, 0.25 *M* L-arginine). The protein solution was then concentrated using a stirred-chamber concentrator (Millipore) and dialyzed against 20 m*M* Tris–HCl pH 8. Refolded and concentrated IRS-2 was purified on a Mono Q column with a 0–1 *M* gradient of NaCl using an ÄKTA FPLC system (Pharmacia) to a purity of >99% according to SDS–PAGE analysis. For crystallization, IRS-2 was dialyzed into 20 m*M* HEPES pH 7.2.

2.2. Protein crystallization

Screening for crystallization conditions was performed by sparse-matrix screening (Jancarik & Kim, 1991) using the commercially available Crystallization Basic and Extension Kits (Sigma–Aldrich) and the hanging-drop vapour-diffusion technique in 24-well Linbro plates at 293 K. Subsequent optimization of the initial crystallization conditions was performed in NeXtal plates (Qiagen). The reservoir contained 0.5 ml reservoir solution and the crystallization drop consisted of 2 µl IRS-2 protein solution (3.5 mg ml^{-1} in 20 m*M* HEPES pH 7.2) and 1 µl reservoir solution. Preliminary needle-shaped crystals grew in condition No. 22 of the Extension Kit: 0.1 *M* MES pH 6.5, 12%(*w*/*v*) PEG 20 000. Optimal crystals were obtained from the original kit solution diluted with water to a final composition of 75 m*M* MES pH 6.5, 9%(*w*/*v*) PEG 20 000.

2.3. Diffraction data collection

For data collection, crystals were soaked for 30 s in reservoir solution supplemented with 20%(*v*/*v*) PEG 400 and flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K on the X12 EMBL beamline at DESY, Hamburg, Germany. A set of 107 images was recorded with a 0.5° oscillation angle, an exposure time of 10 s per image and a crystal-to-detector distance of 185 mm. Diffraction data were processed using the *HKL*-2000 suite of programs (Minor *et al.*, 2006). The redundancy-independent merging *R* factor $R_{\rm r.i.m.}$ as well as the precision-indicating merging *R* factor $R_{\rm p.i.m.}$ were calculated using the program *RMERGE* (Weiss, 2001). Crystal parameters and data-collection statistics are summarized in Table 1.

2.4. Mass spectrometry

Mass-spectrometric characterization of IRS-2 was performed by LC-MS/MS analysis and *de novo* sequencing of tryptic and chymotryptic digests. LC-MS/MS analysis was performed on a LTQ Orbitrap XL hybrid mass spectrometer (Thermo Scientific) coupled to a Rheos 2000 two-dimensional capillary LC system (Flux Instruments). The first-dimension column was a monolithic PS-DVB (200 µm × 10 mm; Dionex) and the second-dimension column was a C18 PepMap100 (75 µm × 150 mm × 3 µm; Dionex) with gradient elution in a 0.1% formic acid/acetonitrile system. The LC-MS/MS data were processed with *SEQUEST* and *BioWorks* (Thermo Scientific) and *PEAKS* (Bioinformatics Solutions) software.

2.5. Profiling of proteolytic activities

The residual proteolytic activities that were present in the IRS-2 protein were detected by hydrolysis of the following fluorogenic peptidic substrates containing a 7-amino-4-methylcoumarin group (AMC; Bachem): Z-Phe-Arg-AMC, MeoSuc-Ala-Ala-Pro-Val-AMC, Z-Gly-Gly-Leu-AMC, Suc-Ala-Ala-Pro-Phe-AMC and Suc-Leu-Tyr-AMC (Horn *et al.*, 2009). The reaction mixture contained 15 µg ml^{-1} protein stock solution (final concentration), 40 µ*M* substrate and 0.1 *M* MES buffer pH 6.5. A blank was prepared without addition of the IRS-2 protein sample. The reaction mixtures were incubated for 24 h at 310 K followed by reading the fluorescence of the liberated AMC using a GENios Plus fluorescence reader at 360 nm excitation and 465 nm emission wavelengths. For activity assays in the presence

Table 2
Profiling of proteolytic activities in IRS-2 protein preparation.

Substrate	Protease type (catalytic class)	Proteolytic activity	Inhibitor sensitivity
Z-Phe-Arg-AMC	Papain-like (cysteine)	Yes	E64
	Trypsin-like (serine)		Pefabloc
MeoSuc-Ala-Ala-Pro-Val-AMC	Elastase-like (serine)	No	
Z-Gly-Gly-Leu-AMC	Chymotrypsin-like (serine)	Yes	Pefabloc
Suc-Ala-Ala-Pro-Phe-AMC	Chymotrypsin-like (serine)	No	
Suc-Leu-Tyr-AMC	Calpain-like (cysteine)	Yes	E64
	Chymotrypsin-like (serine)		Pefabloc

of protease inhibitors, the following inhibitors were added to the reaction mixture: 10 µ*M* E-64 for cysteine proteases and 1 m*M* Pefabloc for serine proteases.

2.6. Serpin-cleavage assay

A 10 µl reaction mixture containing 0.4 µg protein from the IRS-2 stock solution in 0.1 *M* MES buffer pH 6.5 was incubated at 310 K for 6 d in the presence or absence of protease inhibitors (10 µ*M* E-64 and 1 m*M* Pefabloc). The reaction was stopped by heating (at 343 K for 5 min) in reducing Laemmli sample buffer. Nonincubated IRS-2 was used as a control. The reaction mixture was separated by Laemmli SDS–PAGE (15% gels) and visualized by silver staining.

3. Results

IRS-2, a serpin from the hard tick *I. ricinus*, was refolded from inclusion bodies, producing an active inhibitor, and purified using ion-exchange chromatography. The recombinant IRS-2 contained 377 amino-acid residues with a molecular weight of 42.2 kDa. The purity of the sample was confirmed by the presence of a single band on silver-stained SDS–PAGE.

IRS-2 protein solution at 3.5 mg ml^{-1} in 20 m*M* HEPES pH 7.2 was used for crystallization at 293 K. Preliminary needle-shaped crystals were obtained in 5 d using reservoir solution consisting of 0.1 *M* MES pH 6.5, 12%(*w*/*v*) PEG 20 000 (Fig. 1*a*). During optimization, larger crystals of final dimensions of about 0.3 × 0.1 × 0.1 mm appeared within 21 d using reservoir solution consisting of 75 m*M* MES pH 6.5, 9%(*w*/*v*) PEG 20 000 (Fig. 1*b*).

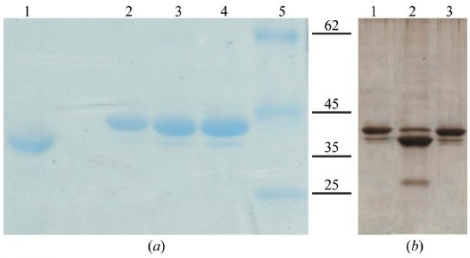


Figure 2
Analysis of IRS-2 proteolysis by SDS–PAGE. (*a*) 15% SDS–PAGE of an IRS-2 crystal (lane 1) and the protein sample used for crystallization experiments (lanes 2–4). Lane 1, dissolved protein crystal; lane 2, 0.75 µg IRS-2; lane 3, 1.1 µg IRS-2; lane 4, 1.5 µg IRS-2; lane 5, protein molecular-weight markers (Fermentas; labelled in kDa). (*b*) 15% SDS–PAGE of IRS-2 protein (0.4 µg) stored at 277 K (lane 1) and incubated for 6 d at 310 K without protease inhibitors (lane 2) and in the presence of 10 µ*M* E-64 and 1 m*M* Pefabloc (lane 3).

Analysis of dissolved crystals using SDS–PAGE indicated that the IRS-2 protein had been proteolyzed during the crystallization at a temperature of 293 K (Fig. 2*a*, lane 1). In order to identify the cleavage site, IRS-2 was subjected to enzymatic digestion followed by LC-MS/MS analysis, which provided ∼85% peptide coverage. This revealed that the cleavage site was located between Tyr341 and Ser342, as the unexpected peptide 342–361 was identified in the tryptic digest. Cleavage at the identified site produced an N-terminal fragment of the serpin molecule with a molecular weight of 38 kDa corresponding to the band identified in the protein crystal (Fig. 2*a*, lane 1) and a 4.2 kDa C-terminal fragment. The identified cleavage site is located within a region homologous to the reactive-centre loop (RCL) of other serpin inhibitors and is most likely to correspond to a natural cleavage site for target proteases.

As proteolysis was also obvious in protein sample stored at 277 K (Fig. 2*a*, lanes 2–4), we concluded that the proteolysis was caused by contaminating proteases that were present in the protein sample and not in the solutions used for crystallization. In order to identify and characterize the contaminating proteases, we further investigated the proteolytic activity in the purified IRS-2 protein sample by testing a

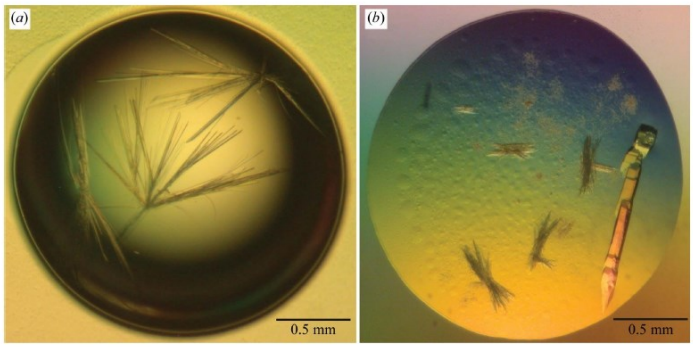


Figure 1
Crystals of IRS-2. (*a*) Initial crystals obtained from screening for crystallization conditions. (*b*) Crystals grown from optimized conditions. For data collection, part of the large crystal cluster on the right was used.

panel of specific peptidic substrates for serine and cysteine proteases (Table 2). The hydrolysis of these substrates was monitored using a continuous fluorimetric assay at the pH of the crystallization mixture. In addition, an activity assay was performed in the presence of selective protease inhibitors, namely Pefabloc and E64, which inhibit serine and cysteine proteases, respectively. Based on the activity profiling (Table 2), the detected activities were attributed to several protease types including the papain-like and calpain-like activities of the cysteine protease class and the trypsin-like and chymotrypsin-like activities of the serine protease class.

To verify that the cleavage of IRS-2 protein was indeed caused by the residual protease activities detected in the protein sample, we incubated the protein sample for an extended period of time (6 d) at 310 K in the presence or absence of serine and cysteine protease inhibitors. SDS–PAGE analysis (Fig. 2*b*) revealed specific cleavage of IRS-2 which is inhibited in the presence of the Pefabloc/E64 mixture. The cleavage produced an IRS-2 major fragment of ∼38 kDa corresponding to the fragment detected in the protein crystal. Further proteolysis to smaller fragments (∼30 kDa) was also observed.

Analysis of the crystals and protein sample led us to the conclusions that IRS-2 was cleaved by contaminating serine and cysteine proteases in the canonical serpin cleavage site and that the protein species which produced crystals is the cleaved and more stable R state.

IRS-2 crystals were used to collect diffraction data after cryoprotection with 20% PEG 400. A complete data set was collected from a single crystal on the X13 beamline at BESSY (Hamburg) to 1.8 Å resolution (Fig. 3). The crystal exhibited the symmetry of space group $P4$, with systematic absences indicating the presence of a 4_1 or 4_3 screw axis. Crystal parameters and data-collection statistics are summarized in Table 1. Evaluation of the crystal-packing parameters indicated the presence of two molecules (containing 377 amino-acid residues each) in the asymmetric unit, with a solvent content of 53% and a Matthews coefficient of 2.62 Å^3 Da^{-1} (Matthews, 1968).

Determination of the structure by molecular replacement with *MOLREP* (Vagin & Teplyakov, 2000) was attempted using the structures of the two serpins that share the highest sequence identity with IRS-2: human squamous cell carcinoma antigen 1 (PDB entry 2zv6; Zheng *et al.*, 2009) and equine leukocyte elastase inhibitor (PDB entry 1hle; Baumann *et al.*, 1992). The solution in space group $P4_3$ using a polyalanine model derived from the structure of the cleaved (R-state) equine leukocyte elastase inhibitor gave the best results (*R* factor 0.55, score 0.252). The two molecules pack sensibly in the unit cell with no clashes and the electron densities calculated from the molecular-replacement solution appeared to be suitable for model building. Examination of the electron-density maps calculated from the initial phases permitted us to identify that the structure indeed represents an R form of serpin with the RCL inserted into the central β-sheet to form as an additional β-strand (Fig. 4). Model building and refinement of the model is currently in progress and the structure analysis will be described in a subsequent manuscript (Chmelar *et al.*, manuscript in preparation).

4. Discussion

Partial proteolysis has been shown to be an important event that is required for the successful crystallization of many protein samples. Successful crystallization of proteins after proteolysis by a contaminating protease originating from the protein sample or crystallization solution has been reported (Mandel *et al.*, 2006) and the inclusion of a protease in the crystallization experiment (*in situ* proteolysis) is becoming a widely used technique (Dong *et al.*, 2007). For most proteins, the positive effect of limited proteolysis on crystallization

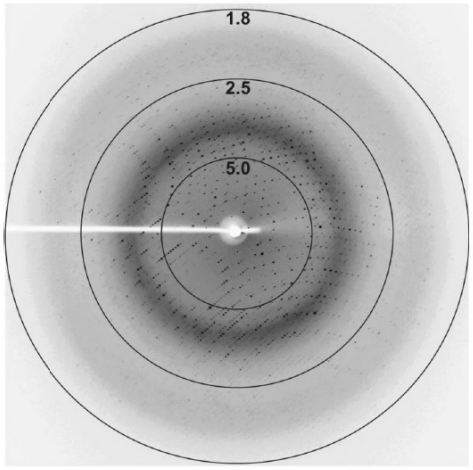


Figure 3
A diffraction image from an IRS-2 crystal. The numbers represent the resolution in Å.

Figure 4
Electron-density map ($2F_{obs} - F_{calc}$, contoured at 1σ) in the region of the central β-sheet of IRS-2 calculated from the initial phases after molecular replacement. The protein model (a polyalanine model derived from the structure of the cleaved equine leukocyte elastase inhibitor) is shown as a ribbon, with the region of the reactive-centre loop (RCL; residues 362–377) coloured red. The β-strand conformation of the RCL is well supported by the electron-density map. This figure was generated using the program *PyMOL* (DeLano, 2002).

stems from the removal of flexible parts or highly hydrophobic segments from the protein surface.

In our case, we observed a specific cleavage of the serpin IRS-2 from the tick *I. ricinus* by contaminating proteases originating from the bacterial expression system. Interestingly, the cleavage was located in the RCL, which is the region of the serpin that is responsible for interaction with the target protease. This proteolytic treatment of IRS-2 mimicked specific cleavage by a target protease and caused a conformational transition of the serpin molecule from the S form into a more stable form (the R form) which produced crystals that were suitable for diffraction analysis.

The cleavage of serpins during crystallization has been reported previously. The first crystal structure of a cleaved serpin, that of human α1-proteinase inhibitor, was the result of an unsuccessful attempt to crystallize its complex with the protease zymogen (Loebermann *et al.*, 1984). During long-term crystallization, the serpin reacted with traces of active proteinase to produce the covalent complex, which then dissociated to give the cleaved serpin, which crystallized. We show here that serpins can be efficiently cleaved by proteases originating from the expression host and that these contaminating activities can be identified using specific substrates and inhibitors as diagnostic tools.

This work was supported by the Grant Agency of the Czech Republic (grant No. P207/10/2183) and in part by research centre No. LC06009 from the Ministry of Education and projects Z40550506 and Z50520514 awarded by the Academy of Sciences of the Czech Republic. Diffraction data were collected on beamline X12 at the EMBL Hamburg Outstation at DESY in Hamburg, Germany; beamline MX14.2 at BESSY, Berlin, Germany was used for preliminary diffraction analysis.

References

Andreotti, R., Gomes, A., Malavazi-Piza, K. C., Sasaki, S. D., Sampaio, C. A. & Tanaka, A. S. (2002). *Int. Immunopharmacol.* **2**, 557–563.
Baumann, U., Bode, W., Huber, R., Travis, J. & Potempa, J. (1992). *J. Mol. Biol.* **226**, 1207–1218.
DeLano, W. L. (2002). *PyMOL*. DeLano Scientific, Palo Alto, California, USA.
Dong, A. *et al.* (2007). *Nature Methods*, **4**, 1019–1021.
Horn, M., Nussbaumerova, M., Sanda, M., Kovarova, Z., Srba, J., Franta, Z., Sojka, D., Bogyo, M., Caffrey, C. R., Kopacek, P. & Mares, M. (2009). *Chem. Biol.* **16**, 1053–1063.
Huntington, J. A., Read, R. J. & Carrell, R. W. (2000). *Nature (London)*, **407**, 923–926.
Imamura, S., da Silva Vaz Junior, I., Sugino, M., Ohashi, K. & Onuma, M. (2005). *Vaccine*, **23**, 1301–1311.
Irving, J. A., Pike, R. N., Lesk, A. M. & Whisstock, J. C. (2000). *Genome Res.* **10**, 1845–1864.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Law, R. H., 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.
Loebermann, H., Tokuoka, R., Deisenhofer, J. & Huber, R. (1984). *J. Mol. Biol.* **177**, 531–557.
Mandel, C. R., Gebauer, D., Zhang, H. & Tong, L. (2006). *Acta Cryst.* F**62**, 1041–1045.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* D**62**, 859–866.
Nagata, K. (1996). *Trends Biochem. Sci.* **21**, 22–26.
Pemberton, P. A., Stein, P. E., Pepys, M. B., Potter, J. M. & Carrell, R. W. (1988). *Nature (London)*, **336**, 257–258.
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.
Ray, C. A., Black, R. A., Kronheim, S. R., Greenstreet, T. A., Sleath, P. R., Salvesen, G. S. & Pickup, D. J. (1992). *Cell*, **69**, 597–604.
Sauer, J. R., McSwain, J. L., Bowman, A. S. & Essenberg, R. C. (1995). *Annu. Rev. Entomol.* **40**, 245–267.
Schick, C., Pemberton, P. A., Shi, G. P., Kamachi, Y., Cataltepe, S., Bartuski, A. J., Gornstein, E. R., Bromme, D., Chapman, H. A. & Silverman, G. A. (1998). *Biochemistry*, **37**, 5258–5266.
Sugino, M., Imamura, S., Mulenga, A., Nakajima, M., Tsuda, A., Ohashi, K. & Onuma, M. (2003). *Vaccine*, **21**, 2844–2851.
Vagin, A. & Teplyakov, A. (2000). *Acta Cryst.* D**56**, 1622–1624.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
Zheng, B., Matoba, Y., Kumagai, T., Katagiri, C., Hibino, T. & Sugiyama, M. (2009). *Biochem. Biophys. Res. Commun.* **380**, 143–147.
Zou, Z., Anisowicz, A., Hendrix, M. J., Thor, A., Neveu, M., Sheng, S., Rafidi, K., Seftor, E. & Sager, R. (1994). *Science*, **263**, 526–529.

5.4. Publikace č. 4: An Unusual Two-Domain Tyropin from Tick Saliva: NMR Solution Structure and Highly Selective Inhibition of Cysteine Cathepsins Modulated by Glycosaminoglycans

Publikace se zaměřuje na nový proteasový inhibitor ze slin klíštěte obecného (*I. ricinus*) nazvaný Ir-tyropin (IrThy). Tento protein patří do tyropinové rodiny inhibitorů, jejíž zástupci dosud nebyli detailně studováni u parazitů a patogenů. Předložená práce přináší první komplexní strukturně-funkční charakterizaci parazitárního tyropinu. *I. ricinus* je důležitým vektorem Lymské boreliózy a klíšťové encefalitidy. Bioaktivní látky z jeho slin, jako jsou proteasové inhibitory, jsou přímými účastníky interakce s hostitelem a IrThy představuje významný příspěvek k mapování těchto molekul a poznání mechanismů jejich působení v interakci parazit–hostitel.

Sekvence IrThy byla určena z cDNA knihovny připravené ze slinných žláz polonasátých samic *I. ricinus*. Přítomnost ve slinách a extraktu ze slinných žláz byla následně potvrzena pomocí proteomické analýzy hmotnostní spektrometrií. Expresní profil IrThy byl určen metodou qPCR ve dvou tkáních účastnících se primárně interakce s hostitelem, slinných žlázách a střevě. IrThy byl výrazně více exprimován ve slinných žlázách a jeho exprese v průběhu sání klesá, což nasvědčuje jeho funkci ve velmi časné fázi interakce s hostitelem.

Pro účely biochemické charakterizace byl IrThy připraven jako rekombinantní protein v expresním systému hmyzích buněk. Analýza sekvence IrThy prokázala, že obsahuje dvě tyropinové domény (také označované jako tyreoglobulinové domény typu 1). Kromě kompletní rekombinantní molekuly IrThy byly proto připraveny i jednotlivé samostatné domény (označované jako N- a C-koncová doména). Pro navržení těchto konstruktů bylo analyzováno párování disulfidů u dvoudoménového IrThy pomocí hmotnostní spektometrie a správné párování poté ověřeno i u produkovaných rekombinantních domén.

Inhibiční specifita IrThy byla určena v aktivitních testech s širokým panelem zástupců proteolytických enzymů z hlavních tříd se zaměřením na relevantní hostitelské proteasy. Bylo zjištěno, že IrThy je velmi selektivní nanomolární inhibitor tří lidských cysteinových proteas s endopeptidasovou aktivitou – katepsinu V, K a L. Při testování samostatných domén si C-doména zachovala obdobnou inhibiční specifitu jako celý protein, zatímco volná N-doména byla schopná inhibovat i dvě další cysteinové proteasy, lidský katepsin F a archetypální proteasu papain. Je známo, že katepsin V a K mají na

svém povrchu vazebná místa pro glykosaminoglykany (GAG) a interakce s těmito negativně nabitými polysacharidy ovlivňuje jejich funkční vlastnosti. V dalším kroku byl proto testován vliv GAG na inhibiční interakci mezi cílovými katepsiny a IrThy. Zatímco inhibice katepsinu L byla ovlivněna minimálně, inhibice katepsinu V byla v přítomnosti GAG výrazně posílena a inhibice katepsinu K naopak snížena. Tento efekt je závislý na typu GAG a jeho koncentraci. Lze předpokládat, že GAG, které jsou široce rozšířené v tkáních hostitele, představují modulační faktor zvyšující selektivitu působení IrThy.

Dále se publikace zabývala strukturní charakterizací IrThy. Z důvodu značné flexibility molekuly nebyly úspěšné krystalizační experimenty, a proto byl k řešení prostorové struktury použit přístup NMR spektroskopie. IrThy a jeho dvě samostatné domény byly připraveny v expresním systému *E. coli* umožňujícím izotopové značení a pro rekombinantní produkty bylo ověřeno správné sbalení i párování disulfidů. Pomocí NMR byla úspěšně určena 3D struktura C-domény (IrThy-Cd) vykazující příznivou vnitřní dynamiku pro získání NMR dat. IrThy-Cd má typický klínovitý tvar charakteristický pro tyropiny tvořený N-koncovou α-šroubovicí a centrálním antiparalelním β-skládaným listem. Strukturním motivem reaktivního centra jsou tři smyčky (L1–L3) zapojené do interakce s proteasou. Oproti jediným dvěma dalším strukturně charakterizovaným tyropinům, fragmentu p41 a saxifilinu, obsahuje IrThy-Cd navíc v C-koncové oblasti krátkou α-šroubovici a dále disulfidové propojení smyček L2 a L3, které zvyšuje rigiditu molekuly IrThy-Cd. Vyřešená 3D struktura umožňuje vysvětlit pozorovanou úzkou selektivitu IrThy-Cd, která je omezena na několik málo cysteinových katepsinů s endopeptidasovou aktivitou. Uvedené specifické strukturní změny v oblasti smyček L2/L3 a substituce v sekvenci smyčky L1 jsou příčinou, proč IrThy neinteraguje s cysteinovými katepsiny s exopeptidasovou aktivitou.

Publikace otevírá cestu k navazujícím studiím o biologické aktivitě IrThy a obecně tyropinů z klíšťat. Podle inhibiční specifity lze v analogii s klíštěcími cystatiny předpokládat účast na imunomodulaci reakce hostitele. Dalším vodítkem je homologie IrThy a ctenitoxinu z jedu pavouků s antinociceptivními účinky, což by mohlo naznačovat potenciální funkci tyropinů jako lokálních anestetik v místě sání klíštěte.

Můj podíl na práci spočíval v: 1) klonování a přípravě rekombinantního proteinu IrThy a jeho jednotlivých domén v expresním systému hmyzích buněk pro funkční studie a v *E.coli* pro strukturní analýzu, 2) přípravě vybraných rekombinantních katepsinových

proteas, 3) návrhu a provádění biochemických a enzymologických funkčních testů se studovanými inhibitory zaměřenými na interakci s proteasami a GAG, 4) interpretaci získané 3D struktury IrThy-Cd s využitím molekulární grafiky a korelace vztahů mezi strukturou a aktivitou a 5) přípravě manuskriptu.

International Journal of
Molecular Sciences

MDPI

Article

An Unusual Two-Domain Thyropin from Tick Saliva: NMR Solution Structure and Highly Selective Inhibition of Cysteine Cathepsins Modulated by Glycosaminoglycans

Zuzana Matoušková [1,2], Katarína Orsághová [1,3], Pavel Srb [1], Jana Pytelková [1], Zdeněk Kukačka [4], Michal Buša [1], Ondřej Hajdušek [5], Radek Šíma [5,6], Milan Fábry [1], Petr Novák [4], Martin Horn [1], Petr Kopáček [5] and Michael Mareš [1,*]

[1] Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Flemingovo n. 2, 16610 Praha, Czech Republic; zuzana.matouskova@uochb.cas.cz (Z.M.); katarina.orsaghova@uochb.cas.cz (K.O.); michal.busa@uochb.cas.cz (M.B.); milan.fabry@uochb.cas.cz (M.F.); martin.horn@uochb.cas.cz (M.H.)
[2] Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, 12800 Praha, Czech Republic
[3] First Faculty of Medicine, Charles University, Katerinska 32, 12108 Praha, Czech Republic
[4] Institute of Microbiology, Czech Academy of Sciences, Prumyslova 595, 25250 Vestec, Czech Republic
[5] Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branisovska 31, 37005 Ceske Budejovice, Czech Republic
[6] Biopticka Laborator, Mikulasske Namesti 4, 32600 Plzen, Czech Republic
* Correspondence: michael.mares@uochb.cas.cz

Abstract: The structure and biochemical properties of protease inhibitors from the thyropin family are poorly understood in parasites and pathogens. Here, we introduce a novel family member, Ir-thyropin (IrThy), which is secreted in the saliva of *Ixodes ricinus* ticks, vectors of Lyme borreliosis and tick-borne encephalitis. The IrThy molecule consists of two consecutive thyroglobulin type-1 (Tg1) domains with an unusual disulfide pattern. Recombinant IrThy was found to inhibit human host-derived cathepsin proteases with a high specificity for cathepsins V, K, and L among a wide range of screened cathepsins exhibiting diverse endo- and exopeptidase activities. Both Tg1 domains displayed inhibitory activities, but with distinct specificity profiles. We determined the spatial structure of one of the Tg1 domains by solution NMR spectroscopy and described its reactive center to elucidate the unique inhibitory specificity. Furthermore, we found that the inhibitory potency of IrThy was modulated in a complex manner by various glycosaminoglycans from host tissues. IrThy was additionally regulated by pH and proteolytic degradation. This study provides a comprehensive structure–function characterization of IrThy—the first investigated thyropin of parasite origin—and suggests its potential role in host–parasite interactions at the tick bite site.

Keywords: cathepsin; cysteine protease; tick; parasite; saliva; thyropin; protease inhibitor; protein structure

Citation: Matoušková, Z.; Orsághová, K.; Srb, P.; Pytelková, J.; Kukačka, Z.; Buša, M.; Hajdušek, O.; Šíma, R.; Fábry, M.; Novák, P.; et al. An Unusual Two-Domain Thyropin from Tick Saliva: NMR Solution Structure and Highly Selective Inhibition of Cysteine Cathepsins Modulated by Glycosaminoglycans. *Int. J. Mol. Sci.* **2024**, *25*, 2240. https://doi.org/10.3390/ijms25042240

Academic Editor: Antonio Lucacchini

Received: 12 December 2023
Revised: 2 February 2024
Accepted: 10 February 2024
Published: 13 February 2024

1. Introduction

The thyroglobulin type-1 (Tg1) domain is typically a 7 kDa protein module found in a variety of unrelated multidomain proteins in the animal kingdom. It was first identified in thyroglobulin, which carries eleven Tg1 repeats [1]. The Tg1 domains, which feature a conserved pattern of cysteine residues, are divided into subgroups 1A and 1B based on the number of cysteines [1]. The variable loops in the molecular architecture of the Tg1 domain make it highly adaptive and account for the multiple functions it has acquired through evolution, including protease inhibition. Tg1 domain-containing proteins with antiproteolytic activity are known as thyropins [2], classified as protease inhibitor family I31 by the Merops database [3]. While Tg1 domains have been shown to directly inhibit cysteine and aspartic proteases, mainly those of the cathepsin type, individual thyropins differ greatly in their inhibitory specificities [4].

Most of our knowledge about thyropins and their inhibitory activity derives from the following members: the p41 invariant chain fragment, involved in major histocompatibility complex (MHC) class II maturation and antigen processing [5,6]; saxiphilin, a neurotoxin-binding protein from bullfrogs [7,8]; equistatin, derived from sea anemones, which contains domains that selectively inhibit cysteine or aspartic cathepsins [9,10]; human testican-1, a multi-domain proteoglycan from the brain [11,12]; human epithelial cell adhesion molecule (EpCAM), a carcinoma cell marker [13]; and the cysteine protease inhibitor ECI from chum salmon eggs [14]. Additionally, Tg1 domains in insulin-like growth factor binding proteins (IGFBPs) have been proposed as proteolysis regulators [15,16].

Of the thyropins whose inhibitory properties have been demonstrated, only two have been characterized at the structural level. These include crystal structures of saxiphilin, which has two Tg1 domains embedded in a large transferrin architecture [8], and the p41 fragment, a single Tg1 domain thyropin, in complex with cathepsin L [6]. This complex identified the reactive center on the Tg1 domain, which is formed by three loops that bind to the active site of cathepsin L. In addition to their inhibitory effects, Tg1 domains may establish more complex relationships with the target cathepsin. For instance, as demonstrated for cathepsin L complexed with the p41 fragment, the Tg1 domain may stabilize the protease at a neutral pH [17]. Another example is the proteolytic degradation of testican-1, an inhibitor of cathepsin L, by the same protease at high concentrations [12].

There is very little information on the structure–function relationships of thyropins in invertebrates, particularly those found in parasites. To address this gap, we performed a molecular characterization of thyropin from the hard tick *Ixodes ricinus*, a well-known vector of Lyme borreliosis and tick-borne encephalitis. Although thyropins have been recently identified by transcriptomic and proteomic analyses in several tick species [18–21], their biological roles remain unknown. Our interest in thyropins builds upon successful research into another family of protease inhibitors targeting cysteine cathepsins, the cystatins. These proteins found in tick saliva act as powerful immunomodulators of the host response, while those in the tick gut regulate blood digestion [22–25].

In this study, we provide a comprehensive functional and structural characterization of Ir-thyropin (IrThy), an unusual thyropin secreted in the saliva of *I. ricinus* ticks. Its small molecule consists of only two Tg1 domains, which are atypically disulfide-bonded. The narrow inhibitory specificity of IrThy was explained by NMR structural analysis and was further shown to be modulated by glycosaminoglycans (GAGs). Finally, we propose a role for IrThy in the initial stages of host–parasite interactions at the tick bite site.

2. Results

2.1. IrThy—A Salivary Protein Secreted by the I. ricinus Tick

Based on the analysis of the salivary gland transcriptome (sialome) of the hard tick *I. ricinus* [26], we identified a set of homologous sequences belonging to the I31 family of thyropin proteins. Using a consensus sequence, we designed primers for PCR amplification from the cDNA of salivary glands of half-fed *I. ricinus* females. The resulting amplicon was cloned and sequenced, yielding a sequence with an open reading frame of 501 bp (Figure S1). This sequence was deposited in GenBank under accession number PP107940. The encoded sequence of the protein, named IrThy, exhibited ~97% and ~96% identity to GenBank accessions JAA67758.1 and JAP74612.1, respectively, from *I. ricinus* [26,27].

To demonstrate the specific localization of IrThy, we applied an LC–MS/MS strategy based on enzymatic digestion of a complex proteome in conjunction with MS/MS peptide sequencing. Based on our analysis of salivary gland extract and saliva from *I. ricinus* females (Table S1), we identified IrThy in both materials, indicating that IrThy was a secreted salivary protein.

Next, we compared the mRNA transcript levels of IrThy in two tissues that play a key role in the tick feeding process (Figure 1A). Specifically, we observed substantially higher expression in the salivary glands of half-fed females than in the midgut, which is responsible for blood digestion. Figure 1B shows the dynamics of IrThy expression

during tick feeding on the host. Unfed females displayed the highest level of expression, which then gradually decreased to reach an expression level approximately one order of magnitude lower than in fed females.

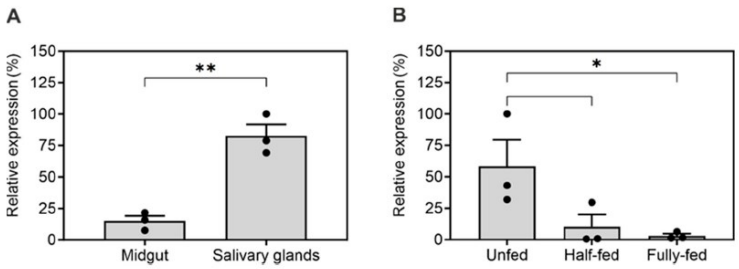


Figure 1. Transcriptional profiling of IrThy. Expression of IrThy was evaluated by qRT-PCR in adult female *Ixodes ricinus* ticks in (**A**) salivary glands and midgut from half-fed ticks and (**B**) whole-body homogenates at different stages of blood feeding. The mRNA transcript levels were normalized to the housekeeping gene *ferritin1*. Results represent the mean ± SD of biological triplicates (pooled samples), expressed relative to the highest measured value (100%); * $p < 0.05$, ** $p < 0.005$.

In summary, we cloned IrThy from *I. ricinus* as a new member of the thyropin family. It is predominantly expressed in the salivary glands of unfed female ticks and is secreted into the tick saliva, suggesting its involvement in the early stages of tick–host interaction.

2.2. IrThy Is a Two-Domain Thyropin: Sequence Analysis, Evolutionary Distribution, and Recombinant Production

IrThy cDNA contains an open reading frame encoding a protein of 166 amino acid residues. It contains a 19-residue signal peptide followed by a mature IrThy sequence of 147 residues consisting of two consecutive Tg1 domains from the I31 thyropin family (Figure 2A). The N- and C-terminal domains have 79 and 68 residues, respectively, and share 26% sequence identity. Both domains belong to Tg1 subtype 1A [1], which features three disulfide bonds. By evaluating the connectivity of the disulfides by LC–MS/MS peptide mapping (Table S2), we found that all of the cysteines formed disulfide bridges. The connectivity is as follows: Cys17–Cys37, Cys48–Cys59, and Cys61–Cys81 in the N-domain, and Cys79–Cys103, Cys115–Cys139, and Cys122–Cys147 in the C-domain (Figure 2A). Two of them are associated with the central sequence Cys79(N-domain)–Lys80–Cys81(C-domain), where Cys79 is cross-linked to the C-domain and Cys81 to the N-domain; therefore, they can be classified as interdomain disulfide bridges (Figure 2A). Sequence comparison shows that the N-domain of IrThy shares the disulfide pattern with the prototype thyropin, the human p41 fragment, while the pattern in the C-domain is modified (Figure S2).

Given that IrThy consists of two Tg1 domains, we decided to investigate the phylogenetic distribution of IrThy homologs with analogous domain compositions. The Pfam database (hosted by InterPro) [28] lists over 20,000 proteins containing the Tg1 domain (Pfam entry: PF00086). These proteins are typically multi-domain, with up to more than a dozen Tg1 repeats, often combined with domains of other types. Figure 2C shows the distribution of Tg1 domain-containing proteins for three different search parameters. Selection I presents all sequences that have two Tg1 domains along with other Pfam domains. Most of them, 77%, belong to vertebrates, and 23% are mammalian. Selection II differs from Selection I in that it excludes proteins containing other Pfam domains. This resulted in a predominance of invertebrate sequences (70%), with over half coming from arthropods (57%). In Selection III, we restricted the sequence length to 180 residues spanning two consecutive Tg1 domains; this is consistent with the domain pattern of IrThy. These short two-domain thyropins were mostly found in invertebrates (63%), including ticks; none

were found in mammals. This indicated that IrThy-like thyropins from ticks have no analogs in mammalian hosts.

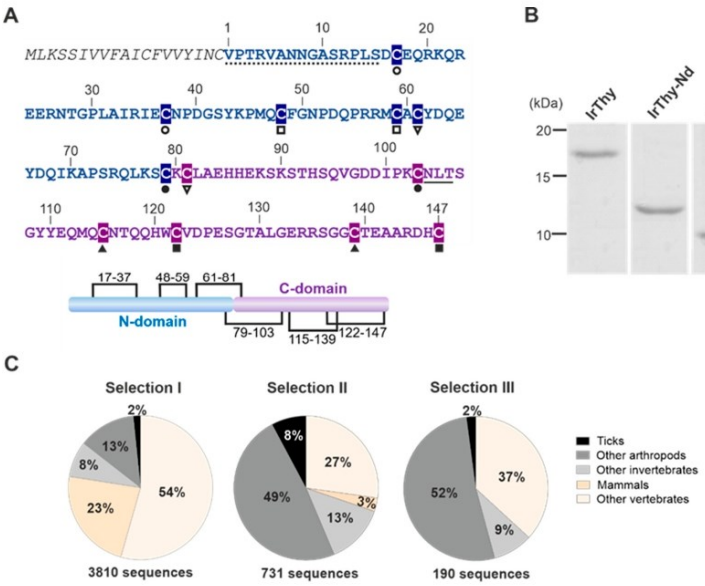


Figure 2. Sequence, evolution, and recombinant production of IrThy. (**A**) Amino acid sequence of IrThy featuring a signal peptide (italics), an unstructured N-terminal region (dotted), an N-terminal thyroglobulin-type 1 (Tg1) domain (blue), and a C-terminal Tg1 domain (purple). Residue numbering is according to the mature protein. The connectivity of cysteine residues (highlighted) is indicated by pairs of black symbols below the sequence. The predicted N-glycosylation site is underlined. The schematic diagram shows the organization of the domains (N- and C-domains) with six disulfide bridges (the black lines indicate cysteine residue connectivity). The signal peptide, the unstructured region, and the two Tg1 domains were predicted using SignalP 6.0, PrDOS, and InterPro, respectively. (**B**) Purified recombinant full-length IrThy and its individual N- and C-terminal domains (IrThy-Nd and IrThy-Cd), produced in the insect cell system, were resolved by SDS–PAGE and visualized by protein staining. (**C**) Phylogenetic distribution of proteins containing two Tg1 domains in invertebrates (including ticks and other arthropods) and vertebrates (including mammals). The sequences from three searches in the InterPro-hosted Pfam database are specified as follows: Selections I and II—the sequence length is not restricted, and other domain types may (Selection I) or may not (Selection II) be present in the molecules; Selection III—the sequence length is restricted to 180 residues, corresponding to proteins with only two consecutive Tg1 domains, such as IrThy.

Recombinant full-length IrThy was prepared as a mature protein with a removable oligohistidine tag (ultimately cleaved by a TEV protease) and expressed using the S2 insect cell system. The same procedure was used to produce individual single domains of IrThy, namely the N-domain protein IrThy-Nd and the C-domain protein IrThy-Cd. For simplicity, interdomain disulfides linked to the central sequence Cys79–Lys80–Cys81 were replaced by intradomain disulfides spanning a sequence two amino acids shorter, namely Cys61–Cys79 in IrThy-Nd and Cys81–Cys103 in IrThy-Cd. These disulfides formed correctly in the expressed single-domain proteins, as confirmed by LC–MS/MS peptide mapping (Table S2). All three recombinants were purified to homogeneity from expression media (see Section 4). They migrated on SDS–PAGE as single bands corresponding to the

predicted molecular masses of 16.9, 9.6, and 7.9 kDa for IrThy, IrThy-Nd, and IrThy-Cd, respectively (Figure 2B). To increase the homogeneity of the recombinants for structural analysis, the N-glycosylation motif on the C-domain was mutated (Figure 2A); this motif is located on the opposite side of the domain from the predicted reactive center for protease binding (see Section 2.6.).

2.3. The Unique, Narrow Inhibitory Specificity of IrThy

Purified recombinant IrThy and its individual domains, IrThy-Nd and IrThy-Cd, were screened in vitro for inhibitory activity against two panels of proteases, including archetypal representatives of major protease classes and relevant cysteine cathepsins of the mammalian host. The latter panel, selected to cover a wide range of endo- and exopeptidase cleavage specificities of cysteine cathepsins, consisted of cathepsins F, K, L, S, and V (endopeptidases), cathepsins B and X (a carboxydipeptidase/endopeptidase and carboxypeptidase, respectively), and cathepsins C and H (an aminodipeptidase and aminopeptidase, respectively).

Inhibition constants (K_i) were determined using a fluorescence activity assay; the specificity profiles are compared in Table 1. No inhibition was observed for members of aspartic and serine proteases or for cysteine proteases of the CD clan. Inhibition was observed only for certain cysteine cathepsins with endopeptidase activities from the CA clan of cysteine proteases. In general, IrThy and its two individual domains selectively inhibited cathepsins V, K, and L as follows: K_i values were within the range of 27–41 nM and 44–57 nM for cathepsin V and cathepsin K, respectively; cathepsin L was more sensitive to inhibition by IrThy and IrThy-Nd (K_i of 201 and 179 nM, respectively) than by IrThy-Cd (K_i of 795 nM). However, IrThy-Nd exhibited a broader inhibitory specificity and, in addition to the three cathepsins mentioned above, it effectively inhibited two other proteases with endopeptidase activities from the CA clan, namely cathepsin F and the archetypal plant protease papain, with K_i values of 207 and 153 nM, respectively.

Table 1. Inhibitory specificity of IrThy and its individual domains. The inhibitory potency of full-length IrThy and its N- and C-terminal domains (IrThy-Nd and IrThy-Cd) was screened against mammalian host-derived cysteine cathepsins of the CA clan and archetypal representatives of protease clans. K_i values (mean ± SD) were determined by a kinetic activity assay using specific fluorogenic peptide substrates (see Section 4). The Merops database classification of the tested proteases (class/clan) and their mode of action are given. n.i.—no significant inhibition at 2 μM inhibitor concentration.

Enzymes	Enzyme Specificity, Protease Class/Clan	K_i (nM) IrThy	IrThy-Nd	IrThy-Cd
Host cathepsin proteases				
Cathepsin V	Endopeptidase, Cys/CA	34.7 ± 2.9	27.2 ± 3.1	40.7 ± 3.8
Cathepsin K	Endopeptidase, Cys/CA	56.9 ± 4.3	43.7 ± 5.4	53.1 ± 7.7
Cathepsin L	Endopeptidase, Cys/CA	201.4 ± 15.8	178.6 ± 6.7	795.1 ± 19.8
Cathepsin F	Endopeptidase, Cys/CA	n.i.	207.2 ± 14.2	n.i.
Cathepsin S	Endopeptidase, Cys/CA	n.i.	n.i.	n.i.
Cathepsin B	Endo- and carboxydipeptidase, Cys/CA	n.i.	n.i.	n.i.
Cathepsin X	Carboxypeptidase, Cys/CA	n.i.	n.i.	n.i.
Cathepsin C	Aminodipeptidase, Cys/CA	n.i.	n.i.	n.i.
Cathepsin H	Aminopeptidase, Cys/CA	n.i.	n.i.	n.i.
Model proteases				
Papain	Endopeptidase, Cys/CA	n.i.	153.4 ± 11.1	n.i.
Legumain	Endopeptidase, Cys/CD	n.i.	n.i.	n.i.
Cathepsin D	Endopeptidase, Asp/AA	n.i.	n.i.	n.i.
Chymotrypsin	Endopeptidase, Ser/PA	n.i.	n.i.	n.i.
Trypsin	Endopeptidase, Ser/PA	n.i.	n.i.	n.i.

We demonstrated that IrThy is capable of inhibiting the activity of model target proteases across the entire pH activity range and, moreover, that the degree of inhibition is

pH-dependent (Figure 3A). The most effective inhibition was observed at a mildly acidic pH of ~6, with reduced inhibitory potency at more acidic and neutral/alkaline pH values. The regulatory effect of pH was more pronounced for cathepsin V than for cathepsin K. We also analyzed the inhibition mode of IrThy, classifying it as competitive, which suggests that IrThy binds in the protease active site as previously described for the p41 fragment [6]. The inhibition mode was determined using a Lineweaver–Burk plot with a series of straight lines intersecting on the 1/v axis (Figure 3B, C).

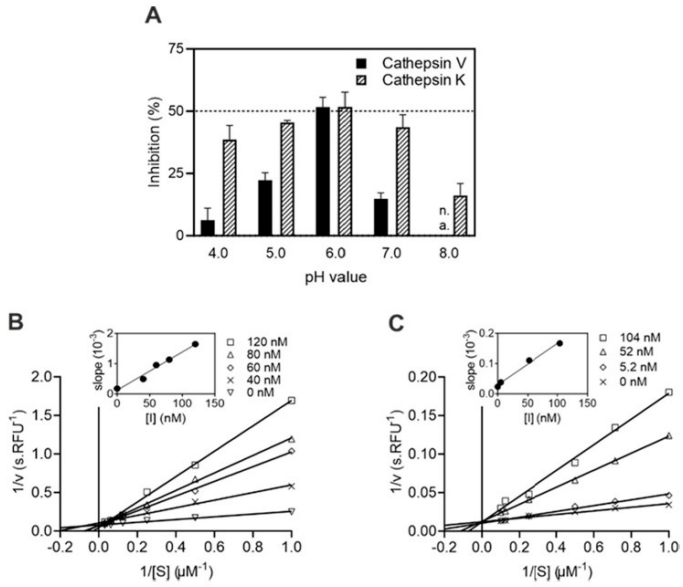


Figure 3. Characteristics of cysteine cathepsin inhibition by IrThy. (**A**) Effect of pH on the inhibitory potency of IrThy. The kinetic activity assay for cathepsins K and V with fluorogenic peptide substrates was performed at different pH values in the presence and absence of IrThy. The inhibitor was applied at a concentration providing ~50% inhibition at pH 6.0, and the % inhibition was calculated relative to an uninhibited control (0%) at the pH values indicated; means ± SD are given. Note that the pH value of 8.0 is outside the functional range of cathepsin V (n.a.). (**B**,**C**) Competitive mode of inhibition by IrThy. Lineweaver–Burk plots are presented together with secondary plots of the same data (inset) for (**B**) cathepsin V and (**C**) cathepsin K. The kinetic activity assay with fluorogenic peptide substrates was performed at pH 5.5. Means ± SD are given for triplicates.

In summary, we demonstrated that IrThy is a thyropin with a unique and narrow inhibitory specificity toward mammalian cysteine cathepsins. It targeted only the endopeptidase cathepsins V and K within nanomolar range, followed by cathepsin L; the other cathepsins with the spectrum of endopeptidase and exopeptidase activities proved insensitive to IrThy.

2.4. IrThy Is Prone to Proteolytic Degradation by Target Proteases at High Concentrations

The Tg1 domains proved susceptible to cleavage by cysteine cathepsins, as reported in [29]. Human testican, a cathepsin L inhibitor from the thyropin family, undergoes degradation by its target protease, cathepsin L, at high concentrations [12]. In this context, we chose to examine the sensitivity of IrThy to the proteolytic activity of cathepsins V, K, and L, which are selectively inhibited by IrThy (see Section 2.3). For comparison, we also

included an experiment with cathepsin B as a model endopeptidase/exopeptidase, which is not targeted by IrThy.

First, we applied the concentration conditions used in the kinetic assay for measuring the inhibitory potency of IrThy (Figure 4). No apparent degradation of IrThy was observed using these catalytic concentrations of cathepsins, with an enzyme:inhibitor ratio of 1:20,000 (*w*/*w*) (corresponding to 0.2 nM:6 µM concentrations). On the contrary, incubation of IrThy with cathepsin concentrations increased by three orders of magnitude (1:5 ratio, *w*/*w*) (corresponding to 0.8 µM:6 µM concentrations) led to the substantial degradation of IrThy by all tested cathepsins. The time course of the degradation reaction is shown in Figure S3, which illustrates the generation of IrThy fragments. Furthermore, we used mass spectrometry for mapping the cleavage sites, which were found to be widely distributed on the IrThy molecule (Figure S4).

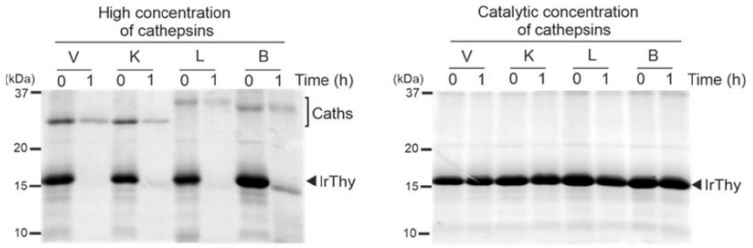


Figure 4. Proteolytic degradation of IrThy by target cathepsins V, K, and L, and a model non-target cathepsin B. IrThy was treated with high concentrations of cathepsins (an enzyme:inhibitor ratio of 1:5, *w*/*w*) or with catalytic concentrations of these cathepsins, corresponding to conditions in a kinetic inhibition assay (an enzyme:inhibitor ratio of 1:20,000, *w*/*w*). The reaction mixture was incubated at pH 5.5; the aliquots at time points 0 and 1 h were resolved by SDS–PAGE and visualized by protein staining. The positions of IrThy and cathepsins (Caths) are indicated.

To conclude, we provide evidence that IrThy can be proteolytically degraded by both its target and non-target cysteine cathepsins. We speculate that this may represent a concentration-dependent regulatory mechanism that controls the physiological inhibitory effect of IrThy.

2.5. Complex Glycosaminoglycan Modulation of IrThy Inhibitory Activity

As previously reported, GAGs, which are widely distributed in tissues, modulate interactions between several protease inhibitors of proteinaceous character and their target proteases, including cathepsins [30,31]. However, this relationship has not been investigated in thyropins. To address this shortcoming, we tested the effect of GAGs on IrThy action against cathepsins V, K, and L, which proved sensitive to IrThy inhibition (Table 1). For the screening, we used a panel of GAGs that included (i) heparin (17–19 kDa) and its disaccharide fragment (0.7 kDa), (ii) chondroitin-4-sulfate, chondroitin-6-sulfate, dermatan sulfate (all in the 20–60 kDa range), and (iii) dextran sulfate (9–20 kDa), which served as a model sulfated polysaccharide GAG analog. Individual GAGs at a concentration of 10 µg/mL were incubated with IrThy and a target cathepsin. The resulting change in cathepsin activity was measured using a fluorescence activity assay and compared to the control experiment without GAG.

The results in Figure 5A show the dramatic effects of GAGs on the inhibition of cathepsins K and V. The inhibitory potency of IrThy against cathepsin K was substantially reduced from ~50% inhibition to ~7, 16, and 23% in the presence of heparin, chondroitin-4-sulfate, and chondroitin-6-sulfate, respectively, and was completely abolished by dermatan and dextran sulfates. The opposite effect was observed for cathepsin V, whose inhibitory potency was greatly enhanced from ~25% to within an inhibitory range of 55–70% in the

presence of all tested GAGs. The only exception was the heparin disaccharide fragment, which had no significant effect on IrThy inhibition of cathepsin V or the other cathepsins. In contrast to cathepsins K and V, the inhibition of cathepsin L by IrThy was not substantially affected by GAGs in general; changes in inhibitory potency were within 15% (Figure 5A).

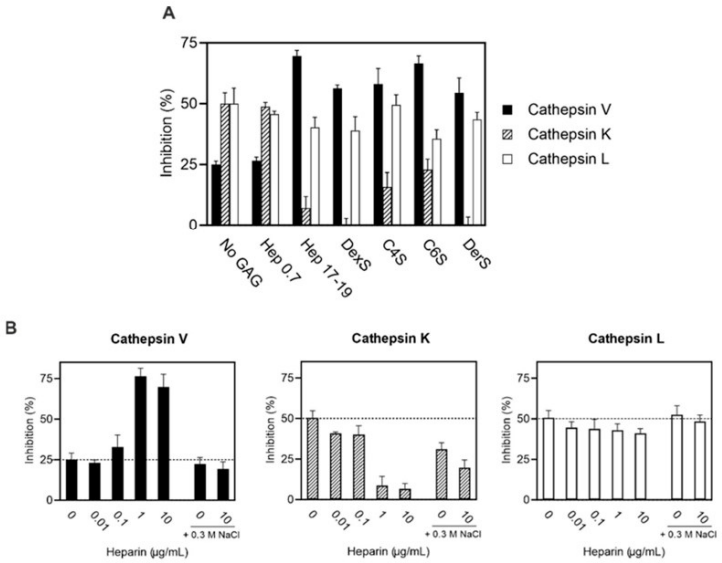


Figure 5. Effect of glycosaminoglycans (GAGs) on the inhibitory potency of IrThy. IrThy was applied at a concentration providing ~50% or ~25% inhibition in the absence of heparin. The kinetic activity assay with fluorogenic peptide substrates was performed at pH 5.5. The % inhibition was calculated relative to uninhibited controls (0%) with the same GAG concentrations (Figure S5); means ± SD are given. Comparative experiments without GAG are indicated as No GAG and 0 heparin. (**A**) The inhibitory potency of IrThy against human cathepsins V, K, and L in the presence of different GAGs, including heparin (Hep 17-19) and its disaccharide fragment (Hep 0.7), chondroitin-4-sulfate (C4S), chondroitin-6-sulfate (C6S), dermatan sulfate (DerS), and the GAG analog dextran sulfate (DexS). GAGs were applied at 10 µg/mL. (**B**) The inhibitory potency of IrThy against human cathepsins V, K, and L in the presence of various concentrations of heparin (0–10 µg/mL) and 0.3 M NaCl (where indicated).

Figure 5B shows the dependence of the modulatory effect on the GAG concentration as demonstrated by heparin. For the inhibition of cathepsins K and V by IrThy, heparin was highly effective at concentrations of 1 µg/mL and above, where the inhibitory potency of IrThy reached its minimum value against cathepsin K and its maximum value against cathepsin V. Heparin concentrations at and below 0.1 µg/mL had a much less pronounced effect (<17% compared to the control without heparin). Inhibition of cathepsin L by IrThy, which exhibited low sensitivity to heparin, was associated with a decrease in inhibition of up to ~10%.

The electrostatic character of the heparin interactions involved in the modulatory effect was examined by increasing the ionic strength in the inhibition assay (Figure 5B). The presence of 0.3 M NaCl either substantially or completely suppressed the heparin-induced changes observed in the inhibitory potency of IrThy. In addition, the control experiment with NaCl and without heparin suggests that electrostatic interactions also contributed

to the interactions between IrThy and cathepsin K, as indicated by the ~20% reduction in inhibition.

In conclusion, GAGs modulate the inhibitory potency of IrThy against cysteine cathepsins in a complex manner. This modulatory effect is highly variable and depends on the type of target cathepsin, GAG, and GAG concentration employed. Most notably, IrThy inhibition of cathepsins K and V, both of which contain GAG-binding sites [32–35], was down- and upregulated by GAGs in opposite ways.

2.6. Spatial Structure of the IrThy C-Domain Determined by NMR Spectroscopy

To investigate the spatial structure of IrThy by NMR spectroscopy, we prepared full-length IrThy and its individual domains, IrThy-Nd and IrThy-Cd, in an *Escherichia coli* expression system for ^{13}C and ^{15}N isotope labeling. Correct protein folding was confirmed by comparing their 1D NMR spectra and inhibitory activity with recombinants produced in insect cells used in functional studies.

For IrThy and IrThy-Nd, the poor quality of all standard 3D spectra indicated that the internal dynamics of both proteins contributed to a faster relaxation of the NH resonances, preventing the acquisition of reliable data. In contrast, the 3D spectra of IrThy-Cd were of sufficient quality to enable the assignment of ~85% of the backbone atoms and ~64% of all the atoms. Of the 68 IrThy-Cd residues, resonances for 17 residues were not detected. These included a stretch of Lys88−Val96, which is probably flexible and lacks contacts with more rigid structural elements. The other residues without backbone assignment were found in the proximity of cysteines Cys115 (Thr117), Cys139 (Ser136, Gly137, and Thr140), and Cys147 (Arg144, Asp145, His146, and Cys147 itself); this can be attributed to the dynamics of disulfide conformation on a millisecond timescale, resulting in a decrease in signal intensities.

Interproton distance restraints were obtained from NOESY spectra (see Section 4). A dense network of inter-residue restraints was observed within the β-sheet-containing region Tyr109−Ala130; 18, 15, and 10 restraints were generated for residues Val123, Asp124, and Pro125, respectively. This points to a significant stabilization of this region, including the L2 loop (Figure 6B). As additional restraints, we also incorporated experimental mass spectrometry data on the pairing of cysteines in IrThy-Cd, namely Cys81–Cys103, Cys115–Cys139, and Cys122–Cys147, into the structure calculation protocol (Table S2). The final refinement yielded 30 solution structures of IrThy-Cd selected on the basis of low overall energy and minimal constraint violations (Figure 6C). Atomic coordinates and experimental constraints were deposited in the Protein Data Bank (accession code: 8R6T) and the Biological Magnetic Resonance Data Bank (BMRB) (accession code: 34883).

A representative spatial structure of IrThy-Cd is presented in Figure 6A. The molecule adopts a wedge-shaped fold typical of thyropins and similar to those of the human p41 fragment (PDB code: 1ICF [6]) and bullfrog saxiphilin (PDB code: 6O0F [8]), the only thyropins whose structures are known. Figure 6B,D shows the structure-based sequence alignment and structural comparison of IrThy-Cd with p41 and the saxiphilin domains 1 and 2; the Cα RMSD values are within the range of 4.9−5.6 Å, and the sequence identities are within the range of 26−32%. While the architecture of IrThy-Cd is clearly similar to that of the other thyropins (Figure 6D), the high RMSD values can be attributed to the considerable flexibility of the loops and orientation of the structural regions. The fold is characterized by a conserved N-terminal α-helix (α1), an antiparallel β-sheet (β1–β2), and three loops (L1–L3), which are involved in the protease interaction (Figure 6A). In addition, IrThy-Cd incorporates another short C-terminal α-helix (α2). In terms of spatial arrangement, these segments constitute two subdomains: the first subdomain contains α1 and L1, and the second the rest of the thyropin molecule [6].

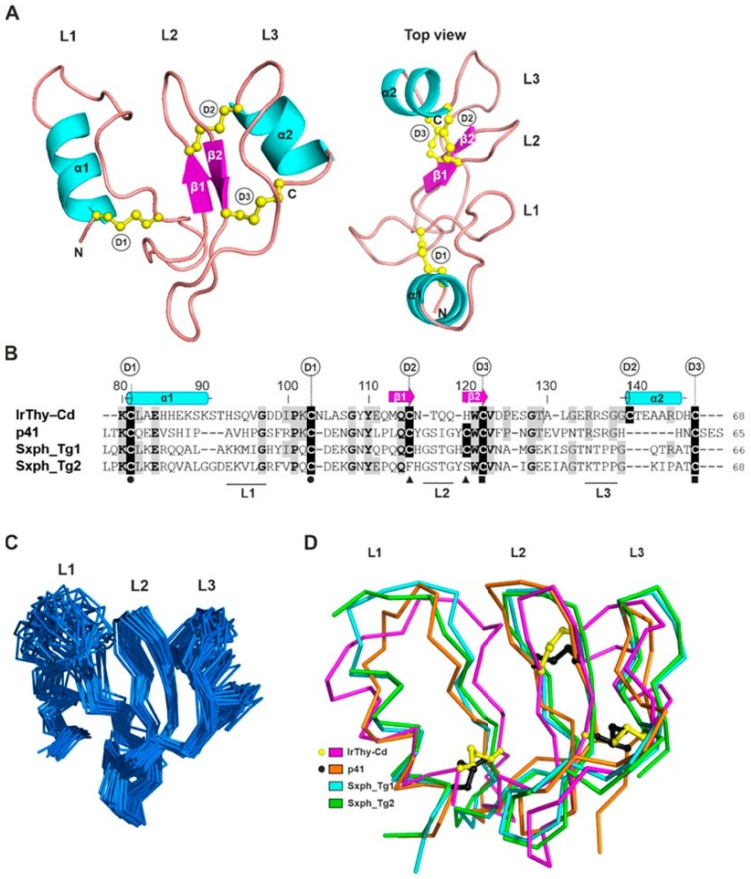


Figure 6. NMR solution structure of the IrThy C-domain (IrThy-Cd) and its comparison with other structurally characterized thyropins. (**A**) Representative three-dimensional structure of IrThy-Cd (PDB code: 8R6T) is depicted in a cartoon representation colored by secondary structural elements (α1–2 helixes, cyan; β1–2 strands, magenta). The N- and C-termini (N, C) and the three disulfide bridges (yellow sticks and balls) Cys81–Cys103 (D1), Cys115–Cys139 (D2), and Cys122–Cys147 (D3) are indicated. The L1, L2, and L3 loops are involved in the binding of the thyropin inhibitors to cysteine cathepsins [6]. (**B**) Structure-based sequence alignment of IrThy-Cd with structurally characterized thyropins, human p41 fragment (p41), and bullfrog saxiphilin domains (Sxph_Tg1 and Sxph_Tg2). Residues identical to those of IrThy-Cd are shaded in grey; fully conserved residues are in bold. Residue numbering is according to IrThy-Cd. Cysteine residues forming disulfide bridges are shaded in black and labeled D1–3 for IrThy-Cd; disulfide connectivity for the other thyropins is indicated by black circles, triangles, and squares below the alignment; note changes in the pairings compared to D2 of IrThy-Cd. The secondary structural elements of IrThy-Cd are depicted in cyan or magenta above the sequence and labeled the same as in (**A**). Three regions topped with the L1, L2,

and L3 loops, which are responsible for the protease interaction, are indicated by horizontal black lines below the alignment based on the structure of the p41 complex with cathepsin L [6]. The alignment was generated using ClustalQ [36] and edited based on the structural superposition. (**C**) An ensemble of the 30 lowest-energy solution structures of IrThy-Cd in the ribbon representation is superimposed. The orientation of the molecule is the same as in (A) (left panel); the binding loops L1 to L3 are marked. (**D**) Overlay of IrThy-Cd (magenta) with structurally characterized thyropins in a ribbon representation, including the p41 fragment (orange, PDB code: 1ICF) and the saxiphilin domains 1 and 2 (cyan and green, respectively, PDB code: 6O0F). The positions of the disulfides are compared for IrThy-Cd (yellow sticks and balls) and p41 (black sticks and balls). The orientation of the molecules is the same as in (**A**) (left panel); the binding loops L1 to L3 are marked.

The disulfide bridges are critical to the overall fold of IrThy-Cd, considering the absence of a hydrophobic core, low content of secondary structures, and substantial conformational flexibility. Both IrThy-Cd subdomains are stabilized by the disulfide bridges Cys81–Cys103, Cys115–Cys139, and Cys122–Cys147, designated D1 through D3, respectively, in Figure 6A. Thus, IrThy-Cd conforms to the thyroglobulin subtype 1A domain, which has three disulfides [1]. In the first subdomain, the conserved D1 disulfide bridge near the N-terminus supports the formation of the L1 loop. The D2 and D3 bridges of IrThy-Cd, located in the second subdomain, connect the L2-containing segment with the L3-containing segment. Interestingly, the interloop disulfide bond D2 in IrThy-Cd is replaced by an intraloop disulfide in p41 and saxiphilin domain 1, where it forms a clamp within the L2 loop; this bridge is absent in saxiphilin domain 2. The higher number of interloop disulfides may provide greater rigidity for the second subdomain of IrThy-Cd.

Loops L1, L2, and L3 form the reactive centers of the thyropins (Figure 6A). The tripartite wedge-shaped edge fills the active site cleft of the target proteases, as previously reported for the crystal structure of the p41 fragment complex with cathepsin L [6]. As the lowest-energy NMR structures of IrThy-Cd demonstrate, the L1- and L3-containing segments are the most flexible parts of the unbound thyropin molecule in solution (Figure 6C). The large L1 loop of IrThy-Cd features the conserved residues Gly101 and Pro105 (Figure 6B), which may be important for protease recognition specificity, as previously proposed for the L1 loops of thyropins [8]. The L2 loop, rather rigid in IrThy-Cd, is stabilized by the β1–β2 sheet and two disulfides. The β2 strand contains a part of the Trp–Cys–Val sequence motif conserved in the Tg1 domains (Figure 6B), which form the core of the second subdomain. The L3 loop is flanked by the D2 interloop disulfide followed by the C-terminal α2 helix; both of these features are unique to IrThy-Cd among the known thyropin structures.

3. Discussion

In this study, we provide a comprehensive structural and functional analysis of IrThy, an unusual member of the thyropin family of protease inhibitors that is expressed and secreted by the salivary glands of the hard tick *I. ricinus*. In general, the scientific understanding of thyropins in invertebrates is limited. To our knowledge, this is the first study to characterize parasite-derived thyropin at the molecular level. The IrThy molecule consists of two consecutive Tg1 domains, the N-domain and the C-domain. We identified an atypical disulfide pattern within the C-domain and also between both domains, forming an interdomain linkage. Intriguingly, the IrThy molecule lacks the complexity of structural modules common to many other multidomain Tg1-containing proteins. Our phylogenetic distribution analysis revealed that these notably short thyropin molecules (up to 180 residues), containing only two tandem Tg1 domains, are widespread in arthropods (including ticks) and entirely absent in mammals.

IrThy is a potent inhibitor of only three human cysteine cathepsins, namely V, K, and L, which belong to the cathepsin L subfamily of endopeptidases. IrThy does not inhibit other related cathepsins that possess diverse endo- and exopeptidase activities. Therefore, IrThy possesses a unique, narrow inhibitory specificity compared to other thyropins. In

general, thyropins can be effective against endopeptidases, but also against exopeptidases. A wide range of cysteine cathepsins with endopeptidase activities, including cathepsins V, K, L, F, and S, are inhibited by the p41 fragment [5]. Inhibition of exopeptidases, such as cathepsins B and H, has been reported for the p41 fragment, equistatin, saxiphilin, and salmon egg thyropin [7,14,37,38].

The recombinant C-domain of IrThy (IrThy-Cd) exhibits the same inhibitory profile as full-length IrThy. We determined the spatial structure of IrThy-Cd by NMR spectroscopy, which allowed us to analyze the structural basis of the inhibitory specificity by comparing the IrThy-Cd structure with that of the p41 fragment complexed with cathepsin L [6]. The binding mode of the p41 fragment, which carries three inhibitory loops (L1, L2, and L3), enables important interactions with the S2–S1 subsite area as well as the surrounding loops located on the protease R-domain [6]. These interactions are understood to be negatively affected by two structural changes in IrThy-Cd, including the increased size of the L1 loop in the first subdomain and the shape of the second subdomain induced by the unique disulfide D2 (Figure 6B,D). In particular, these changes prevent binding to exopeptidases such as cathepsin H (an aminopeptidase) and cathepsin B (a carboxydipeptidase), whose partially obstructed active sites restrict access [39,40].

We observed that the recombinant N-domain (IrThy-Nd) displayed a broader selectivity pattern compared to the narrow inhibitory specificity of IrThy-Cd and the parental IrThy. In addition to cathepsins V, K and L, IrThy-Nd inhibited two other homologous proteases with endopeptidase activities, cathepsin F and papain. This suggests that the binding interactions of the N-domain are somewhat restricted by the closely positioned C-domain in the full-length IrThy scaffold. Nevertheless, our AlphaFold-derived model of IrThy (Figure S6) revealed that the two reactive centers of the N- and C-domains are oriented in opposite directions and can function simultaneously. A similar arrangement has been previously documented in the tandem Tg1 domains of saxiphilin, which bind two protease molecules with different inhibitory specificities [7].

Furthermore, we performed a preliminary cross-linking mass spectrometry (XL-MS) experiment with IrThy and cathepsin V to identify proximal structural regions in the inhibitor−protease complex (Table S3). Our results indicated that IrThy preferentially binds to the active site of cathepsin V (added in an equimolar amount) through the reactive center on the N-domain, which exhibits a lower K_i value than the C-domain. However, further comparative XL-MS data collected under different concentration conditions and with other target cathepsins is required in order to better understand the IrThy binding process and the discrimination between domains.

Our results show that several GAGs in their macromolecular forms modulated the inhibitory potency of IrThy in a complex manner. Widely distributed in host tissues, these sulfated polysaccharides regulate cysteine cathepsins, particularly their autocatalytic processing, pH stability, and interactions with specific macromolecular substrates [32,35,41–45]. In the presence of GAGs, the inhibitory potency of IrThy generally decreased against cathepsin K, increased against cathepsin V, and remained relatively unaffected against cathepsin L. As previously reported, GAGs participate in a variety of functional interactions with cathepsins K and V. For example, while the collagenolytic activity of cathepsin K requires interaction with chondroitin-4-sulfate, GAGs suppress the elastinolytic activities of both cathepsins K and V [32,35,46]. The molecular surfaces of cathepsins K and V contain positively charged patches for GAG binding [32–34,45]; however, these regions are less distributed in cathepsin L, leading to reduced interactions with GAGs [33]. We hypothesize that GAGs regulate IrThy inhibition by interacting with distinct GAG-binding sites on the target cathepsins, and that GAGs may also interact with positively charged regions found on the predicted surface model of IrThy, particularly on its N-domain (Figure S6). The arrangement of these intermolecular interactions can either facilitate or hinder IrThy binding.

Protease inhibitors from tick saliva injected into the host have been shown to modulate the host immune response and suppress blood clotting at the site of tick attachment and blood feeding [47]. To date, the only studied tick salivary inhibitors of cysteine cathepsins

are members of the cystatin family, which exert immunosuppressive and anti-inflammatory effects on the host [23–25]. A recent comparative analysis of inhibitory specificity revealed that tick salivary cystatins differ from other family members in their high affinity for endopeptidases and limited effect on exopeptidases [24]. This biochemical feature makes tick salivary cystatins functionally similar to IrThy, exhibiting narrow inhibitory specificity against cysteine cathepsins with endopeptidase activities. It is therefore reasonable to assume that IrThy may play a role in immunomodulation and, based on the expression dynamics, that it acts in the initial phases of the tick–host interaction. In this context, cathepsins V, K, and L targeted by IrThy are known to be involved in the immune response and inflammation and are also present in the host skin [48–53]. Another clue to the potential physiological function of IrThy comes from a recent report on U24-ctenitoxin-Pn1a, a neurotoxin found in spider venom, that targets voltage-gated sodium channels (VGSCs) [54]. Interestingly, this protein contains two tandem Tg1 domains and is a close homolog of IrThy. The spider VGSC toxins are generally antinociceptive in an animal model [55], and it is tempting to speculate about the potential analgesic effect of IrThy at the tick bite site. Our future research will focus on identifying the exact physiological role of IrThy in tick–host interactions.

4. Materials and Methods

4.1. Materials

All protease substrates were purchased from Bachem (Bubendorf, Switzerland) with the exception of Abz-Phe-Arg-Phe(NO_2)-OH from MP Biomedicals (Irvine, CA, USA). Human cathepsins L, K, and V were produced in the *Pichia pastoris* expression system as described previously [12,56,57]. Bovine cathepsins C and H were prepared as described in [58,59]; human cathepsin D was prepared as described in [60]. Bovine trypsin, chymotrypsin, and *Carica papaya* papain were purchased from Merck (Kenilworth, NJ, USA); human cathepsins F, B, and S were obtained from Enzo Life Sciences (New York, NY, USA); and human cathepsin X and legumain were obtained from R&D Systems (Minneapolis, MN, USA). GAGs were purchased from Sigma-Aldrich (St Louis, MO, USA), including heparin (H3393), chondroitin-4-sulfate (C9819), chondroitin-6-sulfate (C4384), dermatan sulfate (C3788), and dextran sulfate (D6924); heparin disaccharide (H1002) was purchased from Dextra (Reading, UK).

Ticks and Tick-Derived Materials

Adult *I. ricinus* ticks were collected in a forest near České Budějovice, Czech Republic. Ticks were kept at 24 °C and 95% humidity under a 15/9 h day/night regime. Adult *I. ricinus* females were fed on guinea pigs and either forcibly removed on the sixth day of feeding (half-fed) or allowed to fully feed (days 7 to 8). Saliva was collected from half-fed adult *I. ricinus* females as previously described [61]. Salivation was induced by applying pilocarpine solution to the tick scutum. Saliva was collected in 10-µL capillaries for 2 h in a wet chamber kept at 30 °C, pooled, and stored at −80 °C. The salivary glands were carefully removed without disrupting the epithelium and washed in phosphate-buffered saline (PBS). Tissue extracts (150 mg protein/mL) were prepared by homogenizing pooled tissues in 0.1 M sodium acetate pH 4.5, 1% CHAPS on ice. The extract was cleared by centrifugation (16,000× *g*, 10 min, 4 °C), filtered through Ultrafree MC 0.22 µm (Millipore, Bedford, MA, USA), and stored at −80 °C.

4.2. IrThy Cloning and Sequencing

Total RNA was isolated from the salivary glands of five adult *I. ricinus* females fed for six days, as previously described [62]. The RNA was reverse-transcribed into cDNA (0.5 µg RNA per 20 µL reaction) using the Transcriptor High-Fidelity cDNA Synthesis Kit (Roche Applied Science, Penzberg, Germany) with random hexamers and diluted 20 times in water. A 501 bp long fragment encoding IrThy was amplified from cDNA using specific forward 5′-atgctgaagtcaagtatagtagtg-3′ and reverse 5′-ctttcttgattagcagtggtcgc-3′

primers, which were designed based on a consensus sequence of thyropin homologs from *I. ricinus* sialome [26,63]. The sequence was deposited in GenBank under accession number PP107940. The resulting amplicon was isolated from the gel and cloned into the pCR4-TOPO vector (TOPO TA Cloning Kit, Thermo Fisher Scientific, Waltham, MA, USA). The isolated plasmid DNA was sequenced using the sequencing primers provided in the kit; the IrThy coding sequence resulted from the sequencing of seven individual clones.

4.3. Production of Recombinant IrThy and its Domains in Insect Cells

S2 insect cells were used to produce full-length IrThy (residues 1–147) and its individual domains, IrThy-Nd (1–80) and IrThy-Cd (80–147). They were produced as fusion proteins containing an oligohistidine tag (6xHis) and a TEV protease cleavage site with an extension of Ser-Asn-Ala-Ala-Ser residues at the N-terminus. IrThy cDNA was PCR-amplified from the pCR4-TOPO vector containing the complete IrThy insert using the forward 5′-gctagcgtgccaactcgagtggcc-3′ and reverse 5′-gcggccgctaacagtggtcgcgc-3′ primers, which introduced the 5′ NheI and 3′ NotI restriction cloning sites (underlined). The PCR product was ligated into the pUC19 vector (Thermo Fisher Scientific) using SmaI restriction endonuclease. The N-glycosylation motif Asn-Leu-Thr106 was disrupted by Thr106 to Ala106 mutagenesis as described in [64]. pUC19 vectors containing individual single-domain inserts were prepared from full-length IrThy by inverse PCR [65] using the following primers: the forward 5′-tagcggccgcagggtaccgagctcg-3′ and reverse 5′-cttgcaggacttcagctgcctcgatgg-3′ primers for the C-domain deletion; the forward 5′ aagtgtctggcagagcatcacgagaag-3′ and reverse 5′-gctagcgggggggatcctctagagtcg-3′ primers for the N-domain deletion; all primers were 5′-phosphorylated. The resulting vectors were digested with NheI and NotI restriction endonucleases and ligated into the pMT/BiP/V5-His A plasmid (Thermo Fisher Scientific), which was modified to contain at the N-terminus a NheI restriction site and a sequence encoding the BiP signal sequence, His6 tag, and TEV cleavage site.

Stably transfected insect cells were prepared as follows: Schneider 2 (S2) cells (Thermo Fisher Scientific) were cultured at 28 °C in Sf-900 II SFM medium (Thermo Fisher Scientific) containing 10% fetal bovine serum (Gibco, Thermo Fisher Scientific). Cells were co-transfected with the pMT/BiP/V5 vector containing IrThy-derived inserts (19 ng) and the pCoBlast selection vector containing the blasticidin resistance gene (1 ng) (Thermo Fisher Scientific) using the Calcium Phosphate Transfection Kit (Sigma Aldrich) according to the Drosophila Expression System protocol (Thermo Fisher Scientific). Stably transfected cells were selected with blasticidin (25 μg/mL) two weeks after transfection. S2 transfectants were expanded to a large volume (1.5 L) and cultured in Sf-900 II SFM medium to a density of 10^7 cells/mL in a 6 L spinner flask, stirred at 120 rpm at 28 °C. Expression was induced with 1 mM $CuSO_4$, and after 10 days, cells were removed by centrifugation (1000× *g*, 5 min, 10 °C). The supernatant was concentrated by lyophilization to a final volume of 150 mL and desalted on a Sephadex G25 column equilibrated with 10 mM Tris-HCl pH 8.0.

Recombinant His-tagged proteins were purified by Ni^{2+} affinity chromatography on a HiTrap IMAC HP column (Cytiva, Marlborough, MA, USA) following the manufacturer's protocol for elution with imidazole at pH 8.0. The His-tags were cleaved by TEV protease (enzyme:protein ratio 1:10, *w*/*w*) in 50 mM Tris-HCl pH 8.0, 0.5 mM EDTA, and 1 mM DTT for 12 h at 26 °C and removed by Ni^{2+} affinity chromatography on a HiTrap IMAC HP column (see above). This was followed by size-exclusion chromatography on a HiLoad 16/600 Superdex 75 column (Cytiva) equilibrated with 50 mM Tris-HCl, 300 mM NaCl, pH 8.0 (for IrThy and IrThy-Cd) or pH 7.0 (for IrThy-Nd). Protein purity was monitored by Laemmli SDS-PAGE on 15% polyacrylamide gels stained with Coomassie Brilliant Blue R-250. The purified proteins were buffer-exchanged into 20 mM Tris-HCl, 20 mM NaCl, pH 8.0 (for IrThy and IrThy-Cd) or pH 7.0 (for IrThy-Nd), concentrated using an Amicon Ultra-3K centrifugal unit (Millipore), and stored at −80 °C. These recombinants were used in all experiments except for the NMR structure determination.

4.4. Production of Recombinant IrThy and its Domains in E. coli

The pUC19 vectors containing IrThy, IrThy-Nd, or IrThy-Cd inserts (see above) were digested with NcoI and NotI restriction endonucleases and ligated into the pETM-60 plasmid (EMBL, Heidelberg, Germany) in a frame with the NusA chaperone protein, His6-tag, and TEV protease cleavage site. Recombinant proteins were produced in SHuffle T7 Competent *E. coli* cells (New England BioLabs, Ipswich, MA, USA). Bacterial cultures were grown in ^{15}N or $^{13}C/^{15}N$ isotope-containing optimized M9 minimal media (0.5 L) with 50 μg/mL kanamycin to an OD_{600} of 0.6. Expression was induced by the addition of isopropyl 1-thio-β-D-galactopyranoside to a final concentration of 0.4 mM. Cells were harvested by centrifugation (4000× *g*, 10 min, 4 °C) after 16 h of incubation at 16 °C with shaking at 200 rpm. Harvested *E. coli* cells were lysed using an EmulsiFlex-C3 homogenizer (Avestin, Ottawa, ON, Canada), and the lysate was clarified by centrifugation (12,000× *g*, 10 min, 4 °C). The recombinant His-tagged proteins were purified by Ni^{2+} affinity chromatography on a HiTrap IMAC HP column (Cytiva) using the manufacturer's protocol for elution with imidazole at pH 8.0. The fusion tags were cleaved by TEV protease (see above) and removed by Ni^{2+} affinity chromatography on a HiTrap IMAC HP column (see above). The purified proteins were buffer-exchanged into 50 mM Tris-HCl pH 8.0, 150 mM NaCl, concentrated using an Amicon Ultra-3K centrifugal unit (Millipore), and stored at −80 °C. Protein purity was monitored by Laemmli SDS-PAGE (see above).

4.5. Expression Analysis by Quantitative Real-Time PCR

Total RNA was isolated from adult *I. ricinus* females during different stages of blood feeding and from different tissues of females fed for six days, as previously described [62]. The IrThy gene-specific qRT-PCR primers were designed using Primer3. The expression level of IrThy mRNA was measured by qRT-PCR using LightCycler 480 (Roche Applied Science) and SYBR green chemistry [62], with the forward 5′-atgagcaaatgcagtgcaac-3′ and reverse 5′-gcggttccactttctggat-3′ primers. Relative expression was normalized to *I. ricinus ferritin1* (AF068224) [66] according to the mathematical model of Pfaffl [67]. Data were obtained from three independent biological replicates (five females per replicate). Statistical significance was analyzed using the unpaired Student's *t*-test and the one-way ANOVA test.

4.6. Mass Spectrometry Proteomic Analysis

For the identification of native IrThy, proteins from the salivary gland extract and saliva of adult *I. ricinus* female were reduced, alkylated, and digested with trypsin. LC–MS/MS analysis of the digests was performed on an Ulti-Mate 3000 RSLCnano system (Dionex, Thermo Fisher Scientific) coupled to a Triple-TOF 5600 mass spectrometer equipped with a NanoSpray III source (ABSciex, Framingham, MA, USA). The peptides were separated on an Acclaim Pep-Map100 analytical column (3 μm, 150 × 0.75 mm; Thermo Fisher Scientific) by gradient elution in a 0.1% formic acid–acetonitrile mobile phase. Full mass spectrometry scans were recorded from 350 to 1250 *m/z*; up to 25 candidate ions per cycle were subjected to fragmentation; in MS/MS mode, fragmentation spectra were acquired from 100 to 1600 *m/z*. Mass data were processed using ProteinPilot 4.5 software (ABSciex).

For disulfide pairing analysis, recombinant IrThy, IrThy-Nd, or IrThy-Cd was digested with trypsin or Asp-N in the presence of cystamine; LC–MS/MS analysis of the digests was performed according to the protocol [68] described in Table S2.

4.7. Phylogenetic Distribution Analysis

Phylogenetic distribution was analyzed for proteins with Tg1 domains (Pfam entry: PF00086) in the InterPro-hosted Pfam database [28]. Proteins containing two Tg1 domains and meeting the additional criteria specified in Section 2.2. were filtered using Python scripts.

4.8. Protease Inhibition Assays

Inhibition measurements were performed in triplicates in 96-well microplates (100 µL assay volume) at 37 °C. Recombinant inhibitors were preincubated with protease for 15 min, followed by the addition of a specific fluorogenic substrate (see below). The kinetics of product release were continuously monitored using an Infinite M1000 (Tecan, Männedorf, Switzerland) microplate reader at 365 nm excitation and 450 nm emission wavelengths (for AMC-containing substrates) or at 330 nm excitation and 410 nm emission wavelengths (for Abz-containing substrates). The progress curves of kinetic measurements were linear, indicating that the analysis was not affected by inhibitor degradation. IC_{50} values were determined from residual velocities using dose–response plots; nonlinear regression was fitted using GraFit software Version 7 (Erithacus, East Grinstead, UK), and the inhibition constants K_i were calculated using the Cheng–Prusoff equation [41,60].

Substrates and proteases were applied in assays as follows: Z-Phe-Arg-AMC in a concentration of 2.5 µM with 0.2 nM cathepsin L; 4.3 µM with 0.2 nM cathepsin K; 3.6 µM with 0.1 nM cathepsin V; 53 µM with 0.2 nM cathepsin B; 27 µM with 1.2 nM cathepsin F; 100 µM with 0.5 nM papain; 17 µM Z-Val-Val-Arg-AMC with 0.1 nM cathepsin S; 100 µM Arg-AMC with 1.9 nM cathepsin H; 87 µM Abz-Phe-Arg-Phe(NO_2)-OH with 1.2 nM cathepsin X; 130 µM Gly-Arg-AMC with 0.4 nM cathepsin C; 70 µM Z-Ala-Ala-Asn-AMC with 1.2 nM legumain; 33 µM Abz-Lys-Pro-Ala-Glu-Phe-Phe(NO_2)-Ala-Leu-NH_2 with 3.5 nM cathepsin D; 67.8 µM Z-Phe-Arg-AMC with 0.6 nM trypsin; and 41.5 µM Suc-Ala-Ala-Phe-Arg-AMC with 0.02 nM chymotrypsin. The assay buffers were as follows: 50 mM sodium acetate pH 4.0 (for cathepsin D), pH 5.0 (for cathepsin X and legumain), or pH 5.5 (for cathepsins L, K, V, B, F, C, and papain); 50 mM MES pH 6.5 (for cathepsins S and H); and 50 mM Tris-HCl pH 8.0 (for trypsin and chymotrypsin). The buffers contained 0.1% polyethylene glycol 1500 and 2.5 mM dithiothreitol (for cysteine cathepsins and legumain), 50 mM NaCl (for cathepsin C), and 10 mM $CaCl_2$ (for trypsin and chymotrypsin).

The inhibition mode was determined using an analogous activity assay with substrate concentrations of 1–10 µM for cathepsin K and 1–32 µM for cathepsin V; initial velocities of the product release were interpreted using a Lineweaver–Burk plot. The effect of pH on the inhibitory potency of IrThy was measured with cathepsins K and V (as described above) in 100 mM Britton–Robinson buffers within a pH range of 4.0–8.0. The effect of GAGs on the inhibitory potency of IrThy was measured with cathepsins K, L, and V (as described above) in the presence of various GAGs added to the preincubation mixture at a final concentration of 10 µg/mL. Heparin was also tested in a concentration range of 0.01–10 µg/mL. Where indicated, incubation was performed in the presence of 0.3 M NaCl.

4.9. Proteolytic Degradation of IrThy

Recombinant IrThy was incubated with cathepsins B, K, L, or V in 50 mM sodium acetate pH 5.5, 2.5 mM DTT, 0.1% polyethylene glycol 1500 at 37 °C. The reaction was stopped by the addition of 10 µM E-64 at different time points. A 100 µL aliquot contained 10 µg of IrThy (final concentration of 6 µM) and 2 µg or 0.5 ng of cathepsin (final concentrations of 0.8 µM or 0.2 nM, respectively). The reaction mixture was acetone-precipitated, separated by Laemmli SDS-PAGE (15% polyacrylamide gel), and stained with Coomassie Brilliant Blue R-250.

4.10. NMR Spectroscopy

NMR spectra were acquired at 25 °C using an 850 MHz Bruker Avance spectrometer equipped with a triple resonance ($^{15}N/^{13}C/^{1}H$) cryoprobe. The protein sample (0.35 mL in a Shigemi tube) was measured in 50 mM deuterated Tris-HCl pH 8.0, 150 mM NaCl, and 5% D_2O/95% H_2O. A series of double- (Figure S7) and triple-resonance spectra [69,70] were recorded to obtain sequence-specific resonance assignment in NMRFAM-Sparky [71]. $^{1}H–^{1}H$ distance restraints were derived from spectra obtained by 3D $^{15}N/^{1}H$ NOESY-HSQC and $^{13}C/^{1}H$ NOESY-HMQC spectra using a NOE mixing time of 100 ms. The structure calculation was performed in CYANA [72] using NOESY data in combination

with backbone torsion angle restraints generated from assigned chemical shifts using the TALOS+ program [73].

The combined automated NOE assignment and structure determination protocol (CANDID) was used for automatic NOE cross-peak assignment. Five cycles of simulated annealing combined with redundant dihedral angle restraints were then used to calculate a set of converged structures with no significant restraint violations (distance and van der Waals violations < 0.5 Å and dihedral angle constraint violations < 5°). The 30 structures with the least restraint violations were further analyzed using the Protein Structure Validation Software suite [www.nesg.org] (accessed on 11 December 2023).

Statistics for the resulting structure are summarized in Table S4. The structures, NMR restraints, and resonance assignments were deposited in the PDB (accession code: 8R6T) and BMRB (accession code: 34883) databases. Figures with structural representations were generated using PyMOL v2.3.2 (Schrödinger, New York, NY, USA) [74]. The structural comparison of IrThy-Cd with other thyropins provided the following Cα RMSD values (number of aligned Cα atoms is indicated): 5.36 Å (58 atoms) for the p41 fragment, 4.92 Å (50 atoms) for saxiphilin domain 1, and 5.62 Å (49 atoms) for saxiphilin domain 2.

5. Conclusions

Ticks are parasite vectors for a variety of viral and bacterial diseases in humans and domestic animals. Our study focuses on *Ixodes ricinus* ticks, vectors of Lyme borreliosis and tick-borne encephalitis. Tick saliva injected into the host modulates the physiological response at the tick bite site and facilitates the transmission of pathogens. Protease inhibitors from tick saliva are important effector proteins in this process and thus represent promising vaccination targets and pharmacological agents. Here, we characterize at the molecular level a new protease inhibitor, IrThy, found in *I. ricinus* saliva. It belongs to the thyropin family, which is poorly understood in parasites and pathogens. We provide a comprehensive analysis of IrThy, including its unique inhibitory specificity against only three human cysteine proteases, cathepsins V, K, and L, involved in immunity and inflammation, and the NMR structure that explains the functional properties. Based on the results obtained, we propose potential physiological roles for IrThy in host–parasite interaction at the tick bite site.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25042240/s1 References [6,24,68,75–79] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, M.M. and Z.M.; investigation, Z.M., K.O., P.S., J.P., Z.K., M.B., O.H., R.Š. and M.F.; writing—original draft preparation, Z.M., M.M., K.O., P.S. and M.B.; writing—review and editing, M.M., Z.M., M.H. and K.O.; visualization, Z.M., K.O. and M.M.; supervision, M.M., M.H., P.K., and P.N.; funding acquisition, M.M., P.K.; O.H., R.Š. and M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grant 21-08826S and 22-27695S from the Czech Science Foundation (CSF) (to P.K. and M.M.) and Institutional project RVO 61388963. O.H. and P.K. were supported by the Centre for Research of Pathogenicity and Virulence of Parasites (CZ.02.1.01/0.0/0.0/16_019/0000759) from the European Regional Development Fund (ERDF) and the Ministry of Education, Youth and Sports of the Czech Republic (MEYS); R.S. was supported by grant 22-30920S from the CSF; Z.K. and P.N. were supported by grant 22-27695S from the CSF, and M.H. was supported by grant LUC23037 from the MEYS.

Institutional Review Board Statement: All laboratory animals were treated in accordance with the Animal Protection Law of the Czech Republic No. 246/1992 Sb., ethics approval No. 25/2018. The study was approved by the Institute of Parasitology, Biology Centre of the Czech Academy of Sciences and Central Committee for Animal Welfare, Czech Republic (protocol no. 1/2015).

Data Availability Statement: Atomic coordinates and experimental constraints were deposited in the Protein Data Bank (accession code: 8R6T) and the Biological Magnetic Resonance Data Bank (BMRB) (accession code: 34883).

Acknowledgments: The authors thank Lenka Grunclová, Daniel Sojka, and Jakub Benýšek for providing clone and enzyme products used in the study; Kateřina Faltejsková for phylogenetic analysis; Martin Hubálek for mass spectrometry analysis; Milana Štajfová and Anna Fuller for technical assistance; and Michael FitzGerald for language editing.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Molina, F.; Bouanani, M.; Pau, B.; Granier, C. Characterization of the type-1 repeat from thyroglobulin, a cysteine-rich module found in proteins from different families. *Eur. J. Biochem.* **1996**, *240*, 125–133. [CrossRef] [PubMed]
2. Lenarcic, B.; Bevec, T. Thyropins—New structurally related proteinase inhibitors. *Biol. Chem.* **1998**, *379*, 105–111.
3. Rawlings, N.D.; Barrett, A.J.; Thomas, P.D.; Huang, X.; Bateman, A.; Finn, R.D. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **2017**, *46*, D624–D632. [CrossRef]
4. Mihelic, M.; Turk, D. Two decades of thyroglobulin type-1 domain research. *Biol. Chem.* **2007**, *388*, 1123–1130. [CrossRef] [PubMed]
5. Mihelic, M.; Dobersek, A.; Guncar, G.; Turk, D. Inhibitory fragment from the p41 form of invariant chain can regulate activity of cysteine cathepsins in antigen presentation. *J. Biol. Chem.* **2008**, *283*, 14453–14460. [CrossRef]
6. Guncar, G.; Pungercic, G.; Klemencic, I.; Turk, V.; Turk, D. Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. *EMBO J.* **1999**, *18*, 793–803. [CrossRef] [PubMed]
7. Lenarcic, B.; Krishnan, G.; Borukhovich, R.; Ruck, B.; Turk, V.; Moczydlowski, E. Saxiphilin, a saxitoxin-binding protein with two thyroglobulin type 1 domains, is an inhibitor of papain-like cysteine proteinases. *J. Biol. Chem.* **2000**, *275*, 15572–15577. [CrossRef]
8. Yen, T.J.; Lolicato, M.; Thomas-Tran, R.; Du Bois, J.; Minor, D.L. Structure of the saxiphilin: Saxitoxin (STX) complex reveals a convergent molecular recognition strategy for paralytic toxins. *Sci. Adv.* **2019**, *5*, eaax2650. [CrossRef]
9. Lenarcic, B.; Turk, V. Thyroglobulin type-1 domains in equistatin inhibit both papain-like cysteine proteinases and cathepsin D. *J. Biol. Chem.* **1999**, *274*, 563–566. [CrossRef]
10. Galesa, K.; Pain, R.; Jongsma, M.A.; Turk, V.; Lenarcic, B. Structural characterization of thyroglobulin type-1 domains of equistatin. *FEBS Lett.* **2003**, *539*, 120–124. [CrossRef]
11. Bocock, J.P.; Edgell, C.J.; Marr, H.S.; Erickson, A.H. Human proteoglycan testican-1 inhibits the lysosomal cysteine protease cathepsin L. *Eur. J. Biochem.* **2003**, *270*, 4008–4015. [CrossRef] [PubMed]
12. Meh, P.; Pavsic, M.; Turk, V.; Baici, A.; Lenarcic, B. Dual concentration-dependent activity of thyroglobulin type-1 domain of testican: Specific inhibitor and substrate of cathepsin L. *Biol. Chem.* **2005**, *386*, 75–83. [CrossRef] [PubMed]
13. Sankpal, N.V.; Brown, T.C.; Fleming, T.P.; Herndon, J.M.; Amaravati, A.A.; Loynd, A.N.; Gillanders, W.E. Cancer-associated mutations reveal a novel role for EpCAM as an inhibitor of cathepsin-L and tumor cell invasion. *BMC Cancer* **2021**, *21*, 541. [CrossRef] [PubMed]
14. Yamashita, M.; Konagaya, S. A novel cysteine protease inhibitor of the egg of chum salmon, containing a cysteine-rich thyroglobulin-like motif. *J. Biol. Chem.* **1996**, *271*, 1282–1284. [CrossRef] [PubMed]
15. Fowlkes, J.L.; Thrailkill, K.M.; Serra, D.M.; Nagase, H. Insulin-like growth factor binding protein (IGFBP) substrate zymography. A new tool to identify and characterize IGFBP-degrading proteinases. *Endocrine* **1997**, *7*, 33–36. [CrossRef] [PubMed]
16. Moreno, M.J.; Ball, M.; Rukhlova, M.; Slinn, J.; L'Abbe, D.; Iqbal, U.; Monette, R.; Hagedorn, M.; O'Connor-McCourt, M.D.; Durocher, Y.; et al. IGFBP-4 anti-angiogenic and anti-tumorigenic effects are associated with anti-cathepsin B activity. *Neoplasia* **2013**, *15*, 554–567. [CrossRef] [PubMed]
17. Fiebiger, E.; Maehr, R.; Villadangos, J.; Weber, E.; Erickson, A.; Bikoff, E.; Ploegh, H.L.; Lennon-Duménil, A.M. Invariant chain controls the activity of extracellular cathepsin L. *J. Exp. Med.* **2002**, *196*, 1263–1269. [CrossRef]
18. Porter, L.M.; Radulović, Ž.M.; Mulenga, A. A repertoire of protease inhibitor families in *Amblyomma americanum* and other tick species: Inter-species comparative analyses. *Parasit. Vectors* **2017**, *10*, 152. [CrossRef]
19. Oliveira, C.J.; Anatriello, E.; de Miranda-Santos, I.K.; Francischetti, I.M.; Sá-Nunes, A.; Ferreira, B.R.; Ribeiro, J.M. Proteome of *Rhipicephalus sanguineus* tick saliva induced by the secretagogues pilocarpine and dopamine. *Ticks Tick Borne Dis.* **2013**, *4*, 469–477. [CrossRef]
20. Bensaoud, C.; Tenzer, S.; Poplawski, A.; Medina, J.M.; Jmel, M.A.; Voet, H.; Mekki, I.; Aparicio-Puerta, E.; Cuveele, B.; Distler, U.; et al. Quantitative proteomics analysis reveals core and variable tick salivary proteins at the tick-vertebrate host interface. *Mol. Ecol.* **2022**, *31*, 4162–4175. [CrossRef]
21. Kozelková, T.; Dyčka, F.; Lu, S.; Urbanová, V.; Frantová, H.; Sojka, D.; Šíma, R.; Horn, M.; Perner, J.; Kopáček, P. Insight Into the Dynamics of the *Ixodes ricinus* Nymphal Midgut Proteome. *Mol. Cell Proteom.* **2023**, *22*, 100663. [CrossRef]
22. Kotál, J.; Buša, M.; Urbanová, V.; Řezáčová, P.; Chmelař, J.; Langhansová, H.; Sojka, D.; Mareš, M.; Kotsyfakis, M. Mialostatin, a Novel Midgut Cystatin from *Ixodes ricinus* Ticks: Crystal Structure and Regulation of Host Blood Digestion. *Int. J. Mol. Sci.* **2021**, *22*, 5371. [CrossRef]

23. Kotál, J.; Stergiou, N.; Buša, M.; Chlastáková, A.; Beránková, Z.; Řezáčová, P.; Langhansová, H.; Schwarz, A.; Calvo, E.; Kopecký, J.; et al. The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin. *Cell Mol. Life Sci.* **2019**, *76*, 2003–2013. [CrossRef]
24. Martins, L.A.; Buša, M.; Chlastáková, A.; Kotál, J.; Beránková, Z.; Stergiou, N.; Jmel, M.A.; Schmitt, E.; Chmelař, J.; Mareš, M.; et al. Protease-bound structure of Ricistatin provides insights into the mechanism of action of tick salivary cystatins in the vertebrate host. *Cell Mol. Life Sci.* **2023**, *80*, 339. [CrossRef]
25. Kotsyfakis, M.; Sá-Nunes, A.; Francischetti, I.M.; Mather, T.N.; Andersen, J.F.; Ribeiro, J.M. Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*. *J. Biol. Chem.* **2006**, *281*, 26298–26307. [CrossRef]
26. Schwarz, A.; von Reumont, B.M.; Erhart, J.; Chagas, A.C.; Ribeiro, J.M.; Kotsyfakis, M. De novo *Ixodes ricinus* salivary gland transcriptome analysis using two next-generation sequencing methodologies. *FASEB J.* **2013**, *27*, 4745–4756. [CrossRef]
27. Perner, J.; Provazník, J.; Schrenková, J.; Urbanová, V.; Ribeiro, J.M.; Kopáček, P. RNA-seq analyses of the midgut from blood- and serum-fed *Ixodes ricinus* ticks. *Sci. Rep.* **2016**, *6*, 36695. [CrossRef] [PubMed]
28. Paysan-Lafosse, T.; Blum, M.; Chuguransky, S.; Grego, T.; Pinto, B.L.; Salazar, G.A.; Bileschi, M.L.; Bork, P.; Bridge, A.; Colwell, L.; et al. InterPro in 2022. *Nucleic Acids Res.* **2022**, *51*, D418–D427. [CrossRef] [PubMed]
29. Pungercic, G.; Dolenc, I.; Dolinar, M.; Bevec, T.; Jenko, S.; Kolaric, S.; Turk, V. Individual recombinant thyroglobulin type-1 domains are substrates for lysosomal cysteine proteinases. *Biol. Chem.* **2002**, *383*, 1809–1812. [CrossRef] [PubMed]
30. Denamur, S.; Chazeirat, T.; Maszota-Zieleniak, M.; Vivès, R.R.; Saidi, A.; Zhang, F.; Linhardt, R.J.; Labarthe, F.; Samsonov, S.A.; Lalmanach, G.; et al. Binding of heparan sulfate to human cystatin C modulates inhibition of cathepsin L: Putative consequences in mucopolysaccharidosis. *Carbohydr. Polym.* **2022**, *293*, 119734. [CrossRef]
31. Rein, C.M.; Desai, U.R.; Church, F.C. Serpin-glycosaminoglycan interactions. *Methods Enzymol.* **2011**, *501*, 105–137. [CrossRef] [PubMed]
32. Li, Z.; Kienetz, M.; Cherney, M.M.; James, M.N.G.; Brömme, D. The Crystal and Molecular Structures of a Cathepsin K: Chondroitin Sulfate Complex. *J. Mol. Biol.* **2008**, *383*, 78–91. [CrossRef] [PubMed]
33. David, A.; Chazeirat, T.; Saidi, A.; Lalmanach, G.; Lecaille, F. The Interplay of Glycosaminoglycans and Cysteine Cathepsins in Mucopolysaccharidosis. *Biomedicines* **2023**, *11*, 810. [CrossRef] [PubMed]
34. Chazeirat, T.; Denamur, S.; Bojarski, K.K.; Andrault, P.M.; Sizaret, D.; Zhang, F.; Saidi, A.; Tardieu, M.; Linhardt, R.J.; Labarthe, F.; et al. The abnormal accumulation of heparan sulfate in patients with mucopolysaccharidosis prevents the elastolytic activity of cathepsin V. *Carbohydr. Polym.* **2021**, *253*, 117261. [CrossRef] [PubMed]
35. Aguda, A.H.; Panwar, P.; Du, X.; Nguyen, N.T.; Brayer, G.D.; Brömme, D. Structural basis of collagen fiber degradation by cathepsin K. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 17474–17479. [CrossRef] [PubMed]
36. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [CrossRef] [PubMed]
37. Bevec, T.; Stoka, V.; Pungercic, G.; Dolenc, I.; Turk, V. Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L. *J. Exp. Med.* **1996**, *183*, 1331–1338. [CrossRef]
38. Lenarcic, B.; Ritonja, A.; Strukelj, B.; Turk, B.; Turk, V. Equistatin, a new inhibitor of cysteine proteinases from *Actinia equina*, is structurally related to thyroglobulin type-1 domain. *J. Biol. Chem.* **1997**, *272*, 13899–13903. [CrossRef]
39. Musil, D.; Zucic, D.; Turk, D.; Engh, R.A.; Mayr, I.; Huber, R.; Popovic, T.; Turk, V.; Towatari, T.; Katunuma, N.; et al. The refined 2.15 A X-ray crystal structure of human liver cathepsin B: The structural basis for its specificity. *EMBO J.* **1991**, *10*, 2321–2330. [CrossRef]
40. Guncar, G.; Podobnik, M.; Pungercar, J.; Strukelj, B.; Turk, V.; Turk, D. Crystal structure of porcine cathepsin H determined at 2.1 A resolution: Location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. *Structure* **1998**, *6*, 51–61. [CrossRef]
41. Horn, M.; Jílková, A.; Vondrášek, J.; Marešová, L.; Caffrey, C.R.; Mareš, M. Mapping the pro-peptide of the *Schistosoma mansoni* cathepsin B1 drug target: Modulation of inhibition by heparin and design of mimetic inhibitors. *ACS Chem. Biol.* **2011**, *6*, 609–617. [CrossRef]
42. Jílková, A.; Horn, M.; Řezáčová, P.; Marešová, L.; Fajtová, P.; Brynda, J.; Vondrášek, J.; McKerrow, J.H.; Caffrey, C.R.; Mareš, M. Activation route of the *Schistosoma mansoni* cathepsin B1 drug target: Structural map with a glycosaminoglycan switch. *Structure* **2014**, *22*, 1786–1798. [CrossRef]
43. Caglic, D.; Pungercar, J.R.; Pejler, G.; Turk, V.; Turk, B. Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions. *J. Biol. Chem.* **2007**, *282*, 33076–33085. [CrossRef]
44. Almeida, P.C.; Nantes, I.L.; Chagas, J.R.; Rizzi, C.C.; Faljoni-Alario, A.; Carmona, E.; Juliano, L.; Nader, H.B.; Tersariol, I.L. Cathepsin B activity regulation. Heparin-like glycosaminogylcans protect human cathepsin B from alkaline pH-induced inactivation. *J. Biol. Chem.* **2001**, *276*, 944–951. [CrossRef]
45. Novinec, M.; Lenarčič, B.; Turk, B. Cysteine cathepsin activity regulation by glycosaminoglycans. *Biomed. Res. Int.* **2014**, *2014*, 309718. [CrossRef] [PubMed]
46. Yasuda, Y.; Li, Z.; Greenbaum, D.; Bogyo, M.; Weber, E.; Brömme, D. Cathepsin V, a Novel and Potent Elastolytic Activity Expressed in Activated Macrophages. *J. Biol. Chem.* **2004**, *279*, 36761–36770. [CrossRef] [PubMed]

47. Jmel, M.A.; Aounallah, H.; Bensaoud, C.; Mekki, I.; Chmelař, J.; Faria, F.; M'Ghirbi, Y.; Kotsyfakis, M. Insights into the Role of Tick Salivary Protease Inhibitors during Ectoparasite-Host Crosstalk. *Int. J. Mol. Sci.* **2021**, *22*, 892. [CrossRef] [PubMed]
48. Panwar, P.; Hedtke, T.; Heinz, A.; Andrault, P.M.; Hoehenwarter, W.; Granville, D.J.; Schmelzer, C.E.H.; Brömme, D. Expression of elastolytic cathepsins in human skin and their involvement in age-dependent elastin degradation. *Biochim. Biophys. Acta Gen. Subj.* **2020**, *1864*, 129544. [CrossRef] [PubMed]
49. Dennemärker, J.; Lohmüller, T.; Mayerle, J.; Tacke, M.; Lerch, M.M.; Coussens, L.M.; Peters, C.; Reinheckel, T. Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis. *Oncogene* **2010**, *29*, 1611–1621. [CrossRef] [PubMed]
50. Lecaille, F.; Chazeirat, T.; Saidi, A.; Lalmanach, G. Cathepsin V: Molecular characteristics and significance in health and disease. *Mol. Aspects Med.* **2022**, *88*, 101086. [CrossRef] [PubMed]
51. Brömme, D.; Wilson, S. Role of Cysteine Cathepsins in Extracellular Proteolysis. In *Extracellular Matrix Degradation*; Parks, W.C., Mecham, R.P., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 23–51.
52. Vizovišek, M.; Fonović, M.; Turk, B. Cysteine cathepsins in extracellular matrix remodeling: Extracellular matrix degradation and beyond. *Matrix Biol.* **2019**, *75–76*, 141–159. [CrossRef]
53. Senjor, E.; Kos, J.; Nanut, M.P. Cysteine Cathepsins as Therapeutic Targets in Immune Regulation and Immune Disorders. *Biomedicines* **2023**, *11*, 476. [CrossRef] [PubMed]
54. Khamtorn, P.; Peigneur, S.; Amorim, F.G.; Quinton, L.; Tytgat, J.; Daduang, S. De Novo Transcriptome Analysis of the Venom of *Latrodectus geometricus* with the Discovery of an Insect-Selective Na Channel Modulator. *Molecules* **2021**, *27*, 47. [CrossRef]
55. Emerich, B.L.; Ferreira, R.C.M.; Cordeiro, M.N.; Borges, M.H.; Pimenta, A.M.C.; Figueiredo, S.G.; Duarte, I.D.G.; De Lima, M.E. δ-Ctenitoxin-Pn1a, a Peptide from *Phoneutria nigriventer* Spider Venom, Shows Antinociceptive Effect Involving Opioid and Cannabinoid Systems, in Rats. *Toxins* **2016**, *8*, 106. [CrossRef] [PubMed]
56. Benýšek, J.; Buša, M.; Rubešová, P.; Fanfrlík, J.; Lepšík, M.; Brynda, J.; Matoušková, Z.; Bartz, U.; Horn, M.; Gütschow, M.; et al. Highly potent inhibitors of cathepsin K with a differently positioned cyanohydrazide warhead: Structural analysis of binding mode to mature and zymogen-like enzymes. *J. Enzyme Inhib. Med. Chem.* **2022**, *37*, 515–526. [CrossRef] [PubMed]
57. Brömme, D.; Li, Z.; Barnes, M.; Mehler, E. Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization. *Biochemistry* **1999**, *38*, 2377–2385. [CrossRef] [PubMed]
58. Horn, M.; Baudyš, M.; Voburka, Z.; Kluh, I.; Vondrášek, J.; Mareš, M. Free-thiol Cys331 exposed during activation process is critical for native tetramer structure of cathepsin C (dipeptidyl peptidase I). *Protein Sci.* **2002**, *11*, 933–943. [CrossRef] [PubMed]
59. Horn, M.; Dolečková-Marešová, L.; Rulíšek, L.; Máša, M.; Vasiljeva, O.; Turk, B.; Gan-Erdene, T.; Baudys, M.; Mareš, M. Activation processing of cathepsin H impairs recognition by its propeptide. *Biol. Chem.* **2005**, *386*, 941–947. [CrossRef] [PubMed]
60. Máša, M.; Marešová, L.; Vondrášek, J.; Horn, M.; Ježek, J.; Mareš, M. Cathepsin D propeptide: Mechanism and regulation of its interaction with the catalytic core. *Biochemistry* **2006**, *45*, 15474–15482. [CrossRef]
61. Perner, J.; Helm, D.; Haberkant, P.; Hatalová, T.; Kropáčková, S.; Ribeiro, J.M.; Kopáček, P. The Central Role of Salivary Metalloproteases in Host Acquired Resistance to Tick Feeding. *Front. Cell Infect. Microbiol.* **2020**, *10*, 563349. [CrossRef]
62. Urbanová, V.; Šíma, R.; Šauman, I.; Hajdušek, O.; Kopáček, P. Thioester-containing proteins of the tick *Ixodes ricinus*: Gene expression, response to microbial challenge and their role in phagocytosis of the yeast *Candida albicans*. *Dev. Comp. Immunol.* **2015**, *48*, 55–64. [CrossRef]
63. Perner, J.; Kropáčková, S.; Kopáček, P.; Ribeiro, J.M.C. Sialome diversity of ticks revealed by RNAseq of single tick salivary glands. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0006410. [CrossRef]
64. Jílková, A.; Řezáčová, P.; Lepšík, M.; Horn, M.; Váchová, J.; Fanfrlík, J.; Brynda, J.; McKerrow, J.H.; Caffrey, C.R.; Mareš, M. Structural basis for inhibition of cathepsin B drug target from the human blood fluke, *Schistosoma mansoni*. *J. Biol. Chem.* **2011**, *286*, 35770–35781. [CrossRef]
65. Qi, D.; Scholthof, K.B. A one-step PCR-based method for rapid and efficient site-directed fragment deletion, insertion, and substitution mutagenesis. *J. Virol. Methods* **2008**, *149*, 85–90. [CrossRef]
66. Sojka, D.; Franta, Z.; Frantová, H.; Bartošová, P.; Horn, M.; Váchová, J.; O'Donoghue, A.J.; Eroy-Reveles, A.A.; Craik, C.S.; Knudsen, G.M.; et al. Characterization of gut-associated cathepsin D hemoglobinase from tick *Ixodes ricinus* (IrCD1). *J. Biol. Chem.* **2012**, *287*, 21152–21163. [CrossRef]
67. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **2001**, *29*, e45. [CrossRef]
68. Fojtík, L.; Fiala, J.; Pompach, P.; Chmelík, J.; Matoušek, V.; Beier, P.; Kukačka, Z.; Novák, P. Fast Fluoroalkylation of Proteins Uncovers the Structure and Dynamics of Biological Macromolecules. *J. Am. Chem. Soc.* **2021**, *143*, 20670–20679. [CrossRef] [PubMed]
69. Renshaw, P.S.; Lightbody, K.L.; Veverka, V.; Muskett, F.W.; Kelly, G.; Frenkiel, T.A.; Gordon, S.V.; Hewinson, R.G.; Burke, B.; Norman, J.; et al. Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J.* **2005**, *24*, 2491–2498. [CrossRef] [PubMed]
70. Veverka, V.; Lennie, G.; Crabbe, T.; Bird, I.; Taylor, R.J.; Carr, M.D. NMR assignment of the mTOR domain responsible for rapamycin binding. *J. Biomol. NMR* **2006**, *36* (Suppl. 1), 3. [CrossRef] [PubMed]
71. Lee, W.; Tonelli, M.; Markley, J.L. NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015**, *31*, 1325–1327. [CrossRef] [PubMed]

72. Herrmann, T.; Güntert, P.; Wüthrich, K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **2002**, *319*, 209–227. [CrossRef] [PubMed]
73. Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **2009**, *44*, 213–223. [CrossRef] [PubMed]
74. Janson, G.; Zhang, C.; Prado, M.G.; Paiardini, A. PyMod 2.0: Improvements in protein sequence-structure analysis and homology modeling within PyMOL. *Bioinformatics* **2017**, *33*, 444–446. [CrossRef] [PubMed]
75. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef] [PubMed]
76. Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L.E.; Brookes, D.H.; Wilson, L.; Chen, J.; Liles, K.; et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **2018**, *27*, 112–128. [CrossRef]
77. Meng, E.C.; Goddard, T.D.; Pettersen, E.F.; Couch, G.S.; Pearson, Z.J.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **2023**, *32*, e4792. [CrossRef]
78. Kukačka, Z.; Rosůlek, M.; Jelínek, J.; Slavata, L.; Kavan, D.; Novák, P. LinX: A Software Tool for Uncommon Cross-Linking Chemistry. *J. Proteome Res.* **2021**, *20*, 2021–2027. [CrossRef]
79. Götze, M.; Pettelkau, J.; Fritzsche, R.; Ihling, C.H.; Schäfer, M.; Sinz, A. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 83–97. [CrossRef]

Supplementary Materials

An Unusual Two-Domain Thyropin from Tick Saliva: NMR Solution Structure and Highly Selective Inhibition of Cysteine Cathepsins Modulated by Glycosaminoglycans

Zuzana Matoušková, Katarína Orsághová, Pavel Srb, Jana Pytelková, Zdeněk Kukačka, Michal Buša, Ondřej Hajdušek, Radek Šíma, Milan Fábry, Petr Novák, Martin Horn, Petr Kopáček, Michael Mareš

Supplementary Figures

```
atg ctg aag tca agt ata gta gtg ttt gct att tgt ttc gtg gtt tac atc aac tgc gtg   60
 M   L   K   S   S   I   V   V   F   A   I   C   F   V   V   Y   I   N   C   V   20
cca act cga gtg gcc aac aat gga gct tcc cga ccc ctg agt gac tgc gag caa cgt aaa   120
 P   T   R   V   A   N   N   G   A   S   R   P   L   S   D   C   E   Q   R   K   40
cag agg gag gaa cgt aac acg ggt ccg ctg gcc atc cgc atc gag tgc aac ccc gat gga   180
 Q   R   E   E   R   N   T   G   P   L   A   I   R   I   E   C   N   P   D   G   60
agc tac aag ccg atg caa tgc ttt gga aat ccc gat caa ccc cgc agg atg tgt gcc tgc   240
 S   Y   K   P   M   Q   C   F   G   N   P   D   Q   P   R   R   M   C   A   C   80
tac gat caa gag tac gac cag atc aag gcc cca tcg agg cag ctg aag tcc tgc aag tgt   300
 Y   D   Q   E   Y   D   Q   I   K   A   P   S   R   Q   L   K   S   C   K   C   100
ctg gca gag cat cac gag aag tca aag tcc aca cat agt caa gtt gga gac gac att ccc   360
 L   A   E   H   H   E   K   S   K   S   T   H   S   Q   V   G   D   D   I   P   120
aag tgt aat ctg acg agc gga tat tat gag caa atg cag tgc aac act cag cag cat tgg   420
 K   C   N   L   T   S   G   Y   Y   E   Q   M   Q   C   N   T   Q   Q   H   W   140
tgc gtg gat cca gaa agt gga acc gca ctt gga gaa agg cgt tcc gga ggc tgc acc gaa   480
 C   V   D   P   E   S   G   T   A   L   G   E   R   R   S   G   G   C   T   E   160
gct gcg cgc gac cac tgc taa                                                     501
 A   A   R   D   H   C   *                                                      166
```

Figure S1. Nucleotide and deduced protein sequence of IrThy. This sequence was deposited in GenBank under accession number PP107940.

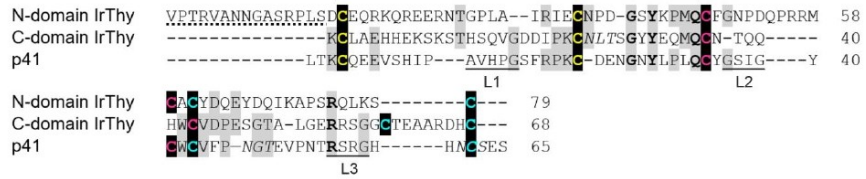


Figure S2. Sequence alignment of the N- and C-domains of IrThy. The human p41 fragment (p41) is included as a prototype member of the thyropin family. The unstructured extension of the N-domain is marked by a dotted line; the N-glycosylation motif is in italics. Cysteine residues forming disulfide bridges are shaded in black; disulfide connectivity is indicated by color coding. Fully conserved residues are in bold and shaded in grey; other identical residues are shaded in grey. The alignment was generated using ClustalQ and edited based on the structural superposition of the IrThy C-domain and p41 (Figure 6B). Residue numbering of IrThy is according to the mature protein. Three regions topped with the binding loops L1–L3 are underlined in the p41 sequence based on the structure of the p41–protease complex (Figure 6D). Sequence comparison shows that the N-domain of IrThy shares the disulfide pattern with p41, while that in the C-domain is modified, and that both IrThy domains differ in two insertions/deletions in the vicinity of the L2 and L3 loops.

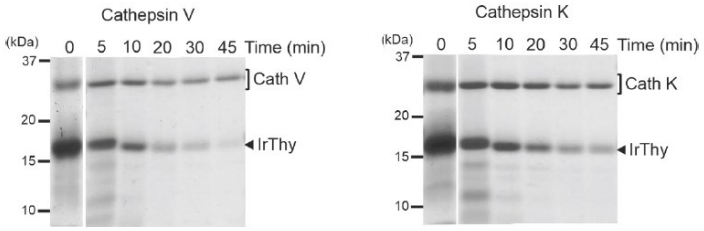


Figure S3. Time course of the proteolytic degradation of IrThy by target cathepsins. IrThy was treated with high concentrations of cathepsins V and K (an enzyme:inhibitor ratio of 1:5, w/w). The reaction mixture was incubated for 45 min at pH 5.5; the aliquots at indicated time points were resolved by SDS–PAGE and visualized by protein staining. The positions of IrThy and cathepsins (Cath V and Cath K) are indicated.

Figure S4. Proteolytic cleavage map of the IrThy molecule. Recombinant IrThy was digested by cathepsin V, K, or L under the same conditions as in Figure 4 at a high enzyme concentration. The fragments were identified by mass spectrometry, and the corresponding cleavage sites are indicated in the IrThy sequence (the N- and C-domains of IrThy are in blue and purple, respectively) as follows: cleavage by cathepsin V (▲), cathepsin K (▲), and cathepsin L (▲). Mass spectrometry analysis of the reaction mixtures after cathepsin digestion was performed as described in the Methods section 4.6. (without additional trypsin digestion).

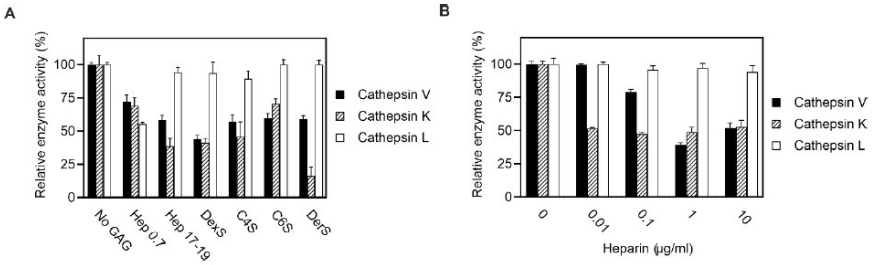


Figure S5. Effect of glycosaminoglycans (GAGs) on human cathepsins V, K, and L (control experiments for Figure 5). The kinetic activity assay with fluorogenic peptide substrates was performed at pH 5.5, and the activity was expressed relative to a control without GAG; means ± SD are given. (A) The following GAGs were used in the assay at 10 µg/mL: heparin (Hep 17-19) and its disaccharide fragment (Hep 0.7), chondroitin-4-sulfate (C4S), chondroitin-6-sulfate (C6S), dermatan sulfate (DerS), and the GAG analog dextran sulfate (DexS). (B) The assay contained various concentrations of heparin (0–10 µg/mL).

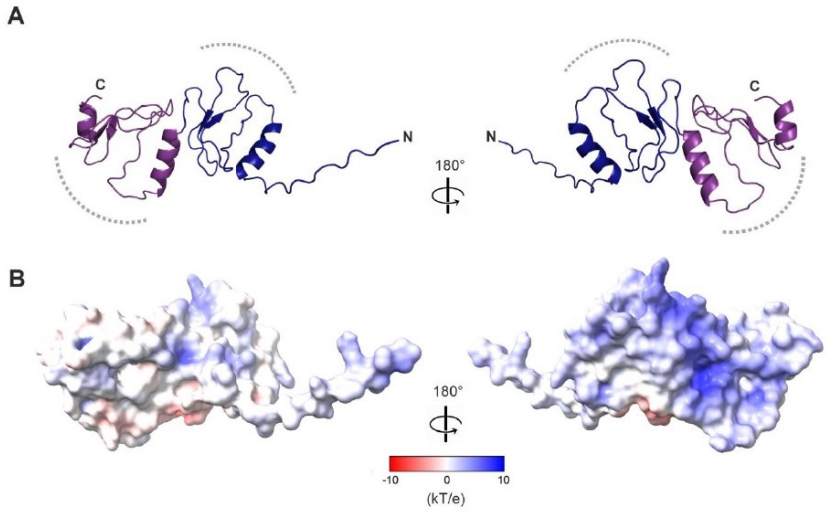


Figure S6. Predicted spatial structure of full-length IrThy. The structure was predicted by AlphaFold2 using the ColabFold web server [1]. (**A**) The model is depicted in a cartoon representation colored by domains: the N-terminal domain is in blue, the C-terminal domain in purple. The N- and C-termini are labeled (N, C); the N-terminus is located on a short unstructured region. On each domain, the dotted line marks three binding loops that form the reactive center involved in the interaction of thyropin inhibitors with cysteine cathepsins [2]. (**B**) Molecular surface of the IrThy model is colored by its electrostatic potential, which is displayed on a scale from −10 kT/e (red) to +10 kT/e (blue). The orientation of IrThy is the same as in (**A**). The electrostatic potential was calculated for pH 5.5 using the APBS-PDB2PQR software suite [3], and the figure was generated using ChimeraX [4].

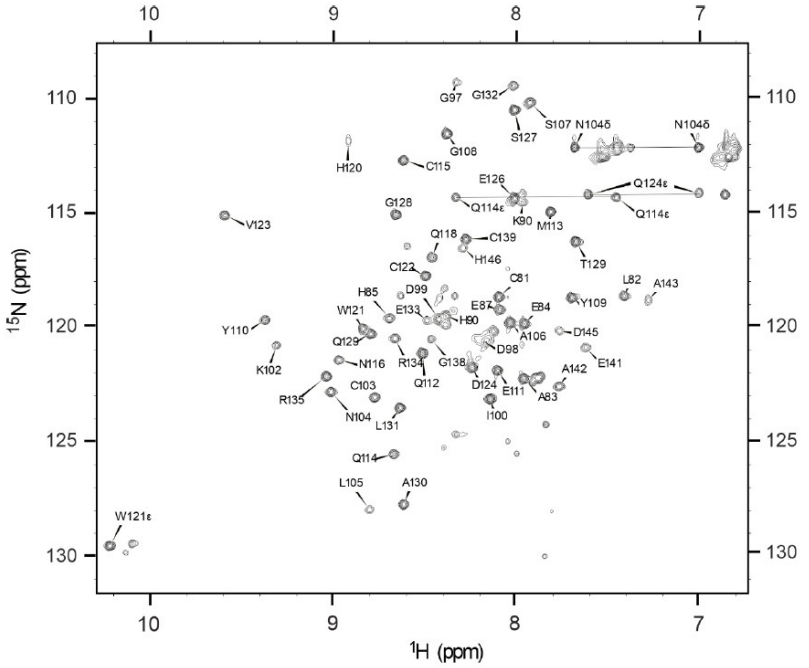


Figure S7. 2D ^{1}H-^{15}N NMR spectrum of IrThy-Cd with the assigned residues indicated. Dotted lines connect signals from the same side-chain amino groups. Residue numbering is according to full-length IrThy.

Supplementary Tables

Table S1. Mass spectrometry identification of native IrThy. Listed are peptides identified in the tryptic digest of proteins from saliva and those extracted from salivary glands of half-fed females of *Ixodes ricinus*. Peptides were analyzed by LC–MS/MS as described in Materials and Methods. Mass deviation (ppm) between experimental and theoretical values and residue numbers of identified peptides are indicated.

Source	Theoretical Mass [M] (Da)	Experimental Mass [M] (Da)	Deviation (ppm)	Peptide Sequence
Salivary gland	2133.901	2133.892	4.22	**MCACYDQEYDQIKAPSR** (58–74)
Salivary gland	2290.004	2290.010	2.62	**RMCACYDQEYDQIKAPSR** (57–74)
Saliva	1722.669	1722.669	0.00	**MCACYDQEYDQIK** (58–70)
Saliva	2133.908	2133.892	7.50	**MCACYDQEYDQIKAPSR** (58–74)
Saliva	1878.772	1878.770	1.07	**RMCACYDQEYDQIK** (57–70)
Saliva	2290.004	2289.993	4.80	**RMCACYDQEYDQIKAPSR** (57–74)

Table S2. Mass spectrometry analysis of disulfide connectivity. Listed are disulfide-bonded peptides identified in proteolytic digests of recombinant IrThy, IrThy-Nd, and IrThy-Cd by LC-MS/MS, as described in a note below the table. Mass deviation (ppm) between experimental and theoretical values, residue numbers of identified peptides and disulfide-bonded cysteine residues are indicated. In the peptide sequences, the disulfide bridges are presented schematically in red. Note that IrThy contains two interdomain disulfide bridges, Cys61-Cys81 and Cys79-Cys103. Residue numbering is according to full-length IrThy.

Experimental Mass [a] [M] (Da)	Theoretical Mass [M] (Da)	Deviation (ppm)	Protein	Connected Peptides	Peptide Sequences	Disulfide Bond
1222.4845	1222.4838	0.57	IrThy–Nd	(16–20)-(35–39)	DCEQR IECNP	C17–C37
2364.8898	2364.8906	0.34	IrThy–Nd	(40–52)-(58–62)-(78–80)	DGSYKPMQCFGNP MCACY SCK	C48–C59; C61–C79
1194.5041	1194.5049	0.67	IrThy–Cd	(81–83)-(103–110)	CLA CNLASGYY	C81–C103
2854.0878	2854.0876	0.07	IrThy–Cd	(111–123)-(136–144)-(145–147)	EQMQCNTQQHWCV SGGCTEAAR DHC	C115–C139; C122–C147
1222.4834	1222.4838	0.33	IrThy	(16–20)-(35–39)	DCEQR IECNP	C17–C37
2333.8845	2333.8848	0.13	IrThy	(40–52)-(58–62)-(81–83)	DGSYKPMQCFGNP MCACY CLA	C48–C59; C61–C81
1224.5027	1224.5035	0.65	IrThy	(78–80)-(103–110)	SCK CNLASGYY	C79–C103
2854.0882	2854.0876	0.14	IrThy	(111–123)-(136–144)-(145–147)	EQMQCNTQQHWCV SGGCTEAAR DHC	C115–C139; C122–C147

[a]Disulfide pairing analysis: The protein sample was digested with Asp-N protease or trypsin (enzyme:protein ratio 1:15, w/w) for 10 h in 50 mM ethylmorpholine pH 8.5, 10% acetonitrile, and 100 µM cysteamine. LC–MS/MS analysis was performed according to a previous protocol [5]. Briefly, the peptide mixture was injected onto a C18 desalting column (Luna Omega Polar C18 5 µm, 0.3 × 30 mm) and then separated on a C18 column (Luna Omega Polar C18, 3 µm, 0.3 × 150 mm) by a linear acetonitrile gradient of 5–35% (v/v) at 50 °C. An Agilent 1290 UPLC system was coupled to the electrospray ionization source of a 15T solariX XR FT-ICR mass spectrometer. The eluted peptides were analyzed in a positive mode with one million transient data points over the 250–2500 *m/z* range. The final spectrum was created by averaging the four following spectra with accumulation ions in the collision cell for 0.2 s. The mass spectrometer was operated in data independent mode. Fragmentation data were acquired with 12 eV collision voltage, 0.2 s MS/MS ion accumulation; two scans were accumulated per spectrum. The raw data were processed using Data Analysis 4.4 software. The exported txt files were uploaded into LinX 2.0 [6], and the search was performed with the following setup: protease – AspN or trypsin with two missed cleavages; variable modification – oxidation of methionine, modification of cysteine by cysteamine; disulfide bridge of cysteines; mass error – 1 ppm. All identified peptides with disulfide bridges were checked manually.

Table S3. Mass spectrometry analysis of intermolecular cross-links in the equimolar complex of IrThy with cathepsin V. The chemical cross-linking reaction was performed with disuccinimidyl dibutyric urea (DSBU). Listed are cross-linked peptides identified in the proteolytic digest of the complex by LC–MS/MS, as described in a note below the table. Mass deviation (ppm) between experimental and theoretical values, residue numbers of identified peptides, and cross-linked residues are indicated. In the peptide sequences, the cross-linked residues are in red. The data obtained indicate that the cross-linked residues of IrThy belong to its N-domain. The homology model of the complex of IrThy (bound via its N-domain) with cathepsin V was constructed using an AlphaFold model of IrThy (Figure SX) and the crystal structure of cathepsin V (PDB code: 1FH0); these structures were superimposed on the structure of the complex of the p41 fragment with cathepsin L (PDB code: 1ICF). In the homology model, the identified cross-links are formed between (i) the segments adjacent to the binding loops of IrThy and (ii) the region surrounding the active site of cathepsin V. This result is consistent with the predicted homology model of the IrThy–cathepsin V complex and suggests that IrThy binds preferentially via its N-domain to cathepsin V under equimolar conditions.

Experimental Mass [a] [M] (Da)	Theoretical Mass [M] (Da)	Deviation (ppm)	Connected Peptides	Peptide Sequences	Cross-Linked Residues
5722.7024	5722.700	0.43	CatV (125–159) × IrThy (13–26)	AVATVGPISVAMDAGHSSFQFYKSGIYFEPDCS**S**K × PLSDCEQR**K**QREER	S158 × K21
6609.0710	6609.072	0.12	CatV (125–159) × IrThy (57–77)	AVATVGPISVAMDAGHSSFQFY**K**SGIYFEPDCSSK × RMCACYDQEYDQI**K**APSRQLK	K147 × K70
6437.9502	6437.945	0.80	CatV (21–58) × IrThy (58–74)	QCG**S**SWAFSATGALEGQMFRKTGKLVSLSEQNLVDCSR × MCACYDQEYDQI**K**APSR	S24 × K70
7560.5138	7560.512	0.21	CatV (125–159) × IrThy (58–80)	AVATVGPISVAMDAGHSSFQFY**K**SGIYFEPDCSSK × MCACYDQEYDQIKAPSRQL**K**SCK	K147 × K77
7720.5163	7720.517	0.13	CatV (148–181) × IrThy (24–56)	SGIYFEPDCSS**K**NLDHGVLVVGYGFEGANSDNSK × EERNTGPLAIRIECNPDGSY**K**PMQCFGNPDQPR	K159 × K44

[a] Cross-linking mass spectrometry (XL-MS) method: An equimolar mixture of the IrThy–cathepsin V complex (12 μM, in sodium acetate pH 6.0) was cross-linked using a 100-fold molar excess of DSBU; an active site mutant of cathepsin V was used for this experiment [7]. The cross-linked samples were diluted twice with 100 mM ethylmorpholine pH 8.5 and then reduced and alkylated with 5 mM tris(2-carboxyethyl)phosphine (TCEP) and chloroacetamide (CAA) for 10 min at 70 °C. Subsequently, the samples were digested twice with a combination of trypsin/Lys-C (enzyme:protein ratio 1:20, w/w) for 4 h, followed by the addition of 0.1% TFA. Peptides were analyzed by LC–MS/MS using a Vanquish chromatography system coupled to a timsTOF SCP mass spectrometer equipped with a captive spray source. The peptide mixture was injected onto a C18 trap column (Pepmap Neo C18, 5 μm, 0.3 × 5 mm) and then separated on a C18 column (Pepsep C18, 1.5 μm, 150 × 0.15 mm) by a linear acetonitrile gradient of 5–35% (v/v) at 50 °C. Parameters from the standard proteomics PASEF method were used to set a timsTOF SCP. The target intensity per individual PASEF precursor was set to 20,000 and the intensity threshold was set to 1500. The scan range was set between 0.6 and 1.6 V s/cm^2 with a ramp time of 100 ms. The number of PASEF MS/MS scans was 10. Precursor ions in the *m*/*z* range between 100 and 1,700 with charge states ≥ 2$^+$ and ≤ 6$^+$ were selected for fragmentation. The active exclusion was enabled for 24 s. The raw data were processed using Data Analysis 5.2 software and exported to Mascot generic files. The search was performed using Merox 2.0 software [8]with the following setup: protease – trypsin/LysC with three missed cleavages; variable modification – oxidation of methionine, stable modification – carbamylation of cysteine; mass shift for DSBU and reporter ions – 196.085, 85.053 and 111.032; mass error – 10 ppm for precursor and 15 ppm for fragments. All cross-links identified in triplicate were manually checked in the raw spectra.

Table S4. NMR constraints and statistics for the final set of structures of IrThy-Cd.

Non-Redundant Distance and Angle Constraints	
Total number of NOE restraints	364
Intraresidue (i = j)	124
Sequential (| i – j | = 1)	108
Medium-range NOEs (1 < | i – j | < 5)	30
Long-range NOEs (| i – j | ≥ 5)	102
Torsion angles	82
Hydrogen bond restraints	32
Total number of restricting restraints	478
Total restricting restraints per restrained residue	8.1
Residual Constraint Violations	
Distance violations per structure	
0.1–0.2 Å	0
0.2–0.5 Å	0
> 0.5 Å	0
r.m.s. of distance violation per constraint	0.02 Å
Maximum distance violation	0 Å
Dihedral angle violation per structure	
1–10 °	7.33
> 10 °	0
r.m.s. of dihedral violations per constraint	1.45 °
Ramachandran Plot Summary	
Most favoured regions	87%
Additionally allowed regions	13%
Generously allowed regions	0%
Disallowed regions	0%
r.m.s.d. to the mean structure	All / ordered
All backbone atoms	1.4 / 0.9 Å
All heavy atoms	2.0 / 1.3 Å
PDB Entry	8R6T
BMRB Accession Code	34883

References

1. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583-589, doi:10.1038/s41586-021-03819-2.
2. Guncar, G.; Pungercic, G.; Klemencic, I.; Turk, V.; Turk, D. Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. *EMBO J* **1999**, *18*, 793-803, doi:10.1093/emboj/18.4.793.
3. Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L.E.; Brookes, D.H.; Wilson, L.; Chen, J.; Liles, K.; et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **2018**, *27*, 112-128, doi:10.1002/pro.3280.
4. Meng, E.C.; Goddard, T.D.; Pettersen, E.F.; Couch, G.S.; Pearson, Z.J.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **2023**, *32*, e4792, doi:10.1002/pro.4792.
5. Fojtík, L.; Fiala, J.; Pompach, P.; Chmelík, J.; Matoušek, V.; Beier, P.; Kukačka, Z.; Novák, P. Fast Fluoroalkylation of Proteins Uncovers the Structure and Dynamics of Biological Macromolecules. *J Am Chem Soc* **2021**, *143*, 20670-20679, doi:10.1021/jacs.1c07771.
6. Kukačka, Z.; Rosůlek, M.; Jelínek, J.; Slavata, L.; Kavan, D.; Novák, P. LinX: A Software Tool for Uncommon Cross-Linking Chemistry. *J Proteome Res* **2021**, *20*, 2021-2027, doi:10.1021/acs.jproteome.0c00858.
7. Martins, L.A.; Buša, M.; Chlastáková, A.; Kotál, J.; Beránková, Z.; Stergiou, N.; Jmel, M.A.; Schmitt, E.; Chmelař, J.; Mareš, M.; Kotsyfakis, M. Protease-bound structure of Ricistatin provides insights into the mechanism of action of tick salivary cystatins in the vertebrate host. *Cell Mol Life Sci* **2023**, *80*, 339, doi:10.1007/s00018-023-04993-4.
8. Götze, M.; Pettelkau, J.; Fritzsche, R.; Ihling, C.H.; Schäfer, M.; Sinz, A. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J Am Soc Mass Spectrom* **2015**, *26*, 83-97, doi:10.1007/s13361-014-1001-1.

6. Diskuse

Interakce mezi krev sajícím parazitem a hostitelem je na molekulární úrovni vysoce komplexní proces a kritickým způsobem se ho účastní proteolytický systém parazita i hostitele. Jedním z hlavních přirozených mechanismů regulace proteolýzy jsou proteasové inhibitory proteinového charakteru, které působí jak na endogenní úrovni, kdy např. pomáhají při načasování a kompartmentaci působení proteas, tak na exogenní úrovni, kdy parazitům umožňují invazi hostitele pomocí suprese hostitelských proteas účastnících se procesů jako je imunitní odpověď a hemostáza. Tato disertační práce se zaměřuje na nové zástupce proteinových proteasových inhibitorů z rodiny cystatinů, serpinů a tyropinů, kteří pocházejí ze slin dvou druhů klíšťat. Klíště *Ixodes ricinus* je významným vektorem klíšťové encefalitidy a Lymské boreliózy a klíšťák *Ornithodoros moubata* je vektorem návratné horečky a afrického prasečího moru. Výzkum se soustředil zejména na strukturně-funkční analýzu těchto bioaktivních proteinů injikovaných do těla hostitele a přináší významné nové informace pro pochopení mechanismů interakce mezi klíštětem a hostitelem, které mají využití při vývoji protiklíštěcích vakcín a biofarmak. Diskuse je rozdělena do tří kapitol věnovaných jednotlivým proteasovým inhibitorům.

6.1. Proteasový inhibitor OmC2 z rodiny cystatinů

Cystatin OmC2 ze slin klíšťáka Ornithodoros moubata byl biochemicky charakterizován jako účinný inhibitor cysteinových katepsinů s neobvykle širokou inhibiční specifitou cílenou proti katepsinům jak s endopeptidasovou, tak exopeptidasovou aktivitou. To jej odlišuje od široké skupiny sialostatinů, slinných cystatinů z klíštěte *I. scapularis* a *I. ricinus*, které exopeptidasové katepsiny inhibují velmi slabě nebo vůbec[138,140,142]. OmC2 se tak spíše podobá nově identifikovanému mialostatinu, cystatinu ze střeva *I. ricinus*[145]. Vyřešení 3D struktury OmC2 umožnilo srovnání těchto dvou typů cystatinů, které obecně vykazují značnou podobnost jak na úrovni primární, tak terciární struktury. U OmC2 byly nicméně zjištěny lokální sekvenční a strukturní odlišnosti na N-konci molekuly a smyčce L2. Oba tyto segmenty jsou zodpovědné za afinitu a selektivitu cystatinů k cílovým proteasám a jsou proto pravděpodobně hlavní příčinou široké inhibiční specifity OmC2[146]. Podobný design obou segmentů je možné pozorovat u cystatinů z obratlovců, jako je např. široce rozšířený cystatin C. Proto se lze domnívat, že OmC2 mimikuje hostitelské cystatiny a tím

interferuje s jemnou rovnováhou hostitelského proteolytického systému mezi katepsiny a jejich endogenními inhibitory.

O mechanismu imunomodulačního účinku OmC2 lze spekulovat na základě jeho inhibiční specifity a dostupných informací o roli inhibovaných proteas v imunologických procesech. Katepsin S a L jsou zapojeny do procesování antigenů pomocí antigen-prezentujících buněk v komplexu s MHC II proteiny, což umožňuje rozvoj adaptivní imunitní odpovědi[147]. Experimenty studující vliv OmC2 na antigen-prezentující buňky, konkrétně dendritické buňky (hojně přítomné v kůži jako místě sání klíštěte), prokázaly inhibici prezentace antigenu a snížení produkce mediátorů zánětu, cytokinů TNFα a IL-12. Suprese dendritických buněk spojená s inhibicí katepsinu S byla také prokázána u příbuzného klíštěcího cystatinu, sialostatinu L[148]. Manipulování adaptivní imunity může usnadnit opakované sání klíšťat a jejich potomků na hostiteli, jelikož jsou slinné antigeny lépe maskovány před imunitním systémem hostitele. K potlačení rozpoznání klíštěcích antigenů může OmC2 přispět i inhibicí katepsinů B a H s exopeptidasovou aktivitou, které se podílejí na úpravě peptidů vázaných na MHC II zkracováním jejich koncových úseků[149]. Další proteasou silně inhibovanou OmC2 je katepsin C, který je klíčový pro vrozenou imunitu, jelikož aktivuje serinové proteasy produkované zralými neutrofily[150]. Po přisátí klíštěte, jsou to právě neutrofily, které jsou v první linii obrany buněčné imunity a které by mohly negativně ovlivnit proces sání.

Vakcinace pomocí OmC2 omezila schopnost nymf efektivně sát na imunizovaných zvířatech a zvýšila i mortalitu nymf po nasátí, která byla úměrná titru specifických anti-OmC2 protilátek v séru hostitele. OmC2 je exprimován nejen ve slinách, ale i ve střevě klíštěte, kde probíhá proces trávení hostitelské krve[151]. Je pravděpodobné, že OmC2 zde může hrát protektivní roli při zamezení předčasné proteolýzy přijaté potravy, která je určena pro skladování, a to jednak inhibicí endogenních proteas z poškozených buněk střeva a jednak hostitelských proteas přijatých s krví hostitele. Pokud je tato funkce ve střevě blokována anti-OmC2 protilátkami, může to vysvětlit zvýšenou mortalitu nasátých nymf. Protiklíštěcí vakcína založená na OmC2 by tedy mohla působit na dvou úrovních a v rozdílných kompartmentech během interakce klíštěte s hostitelem. Specifické protilátky proti OmC2 budou interferovat s imunomodulační funkcí OmC2 injikovaného do hostitele ve slinách klíštěte a zároveň budou tyto protilátky blokovat předpokládanou fyziologickou funkcí OmC2 ve střevě klíštěte. To by mohlo zvyšovat účinnost této duální vakcíny.

6.2. Proteasový inhibitor IRS-2 z rodiny serpinů

Další část disertační práce se zabývala inhibitorem serinových proteas, IRS-2, ze slin klíštěte *I. ricinus*, který patří do rodiny serpinů. Serpiny působí jako „sebevražedné" inhibitory, u nichž cílová proteasa štěpí smyčku reaktivního centra inhibitoru za vzniku primárního kovalentního komplexu proteasy s inhibitorem. Analýza krystalové struktury IRS-2 prokázala, že u tohoto serpinu dochází při ataku proteasou ke štěpení reaktivní smyčky v unikátní poloze s tyrosinovým zbytek v poloze P1. Většina serpinů z genomu příbuzného klíštěte *I. scapularis* má v poloze P1 bazický (především argininový) zbytek vhodný pro štěpení enzymy trypsinového typu, např. proteasami koagulační kaskády. Naopak aromatické aminokyseliny se v této poloze nacházejí pouze u dvou predikovaných serpinů *I. scapularis*, z nichž jeden je homologem IRS-2. Tyrosinový zbytek je zde unikátní a lze tak předpokládat, že IRS-2 je jediný slinný serpin *I. ricinus* schopný inhibovat dvě cílové proteasy chymotrypsinového typu, katepsin G a chymasu.

Katepsin G je sekretován po aktivaci neutrofilů a chymasa je sekretována po aktivaci žírných buněk. Tyto buňky se přímo účastní akutní zánětlivé odpovědi. O katepsinu G je známo, že se podílí na odstranění pohlcených patogenů a remodelaci tkání během zánětu, proteolytické aktivaci různých chemoaktraktantů a hormonů a buněčné signalizaci štěpením proteasou aktivovaného receptoru 4 (PAR4)[152,153]. Štěpení PAR4 katepsinem G aktivuje krevní destičky a vede k agregaci a tvorbě krevní zátky[154]. Tato studie o IRS-2 jako první prokázala, že bioaktivní protein ze slin klíšťat inhibuje agregaci krevních destiček indukovanou katepsinem G a také trombinem, a tím klíštěti usnadňuje příjem krve.

Aktivované krevních destičky produkují řadu chemokinů, jejich prekurzorů a bioaktivních peptidů, které jsou proteolyticky zpracovávány katepsinem G a chymasou. Příkladem je chemokin CTAP-III sekretovaný aktivovanými krevními destičkami, který je konvertován na chemokin NAP-2, jenž aktivuje neutrofily a přivolává je do místa poranění[155]. Katepsin G a chymasa dále zpracovávají velké endoteliny, prekurzory několika vazokonstrikčních látek jako jsou peptidové fragmenty označované ET(1-31), které současně působí i jako chemoaktraktanty pro neutrofily a monocyty[156]. Zajímavým zjištěním je, že štěpená oblast smyčky reaktivního centra na IRS-2 (VPY*SLG) se výrazně podobá místu štěpenému chymasou ve velkých endotelinech (VPY*GLG). IRS-2 je tak první popsaný parazitární protein, který inhibuje hostitelské proteasy tím, že mimikuje jejich endogenní substráty. Dále se IRS-2 strukturně značně podobá

přirozenému inhibičnímu regulátoru katepsinu G a chymasy, kterým je hostitelský serpin α-1-antichymotrypsin.

Inhibováním katepsinu G neutrofilů a chymasy žírných buněk přítomných v časné fázi zánětu může klíště pozměnit obrannou reakci na poškození kůže způsobené jeho přisátím. V této práci se podařilo experimentálně prokázat, že IRS-2 inhibuje akutní zánět vyvolaný polysacharidem karagenanem a agregaci krevních destiček indukovanou katepsinem G nebo trombinem. Domníváme se, že IRS-2 cílí neutrofily, krevní destičky a žírné buňky v časných fázích jejich aktivace tím, že zamezuje aktivaci krevních destiček vyvolané katepsinem G nebo trombinem. Tím se uvolňuje méně CTAP-III a ten také není účinně konvertován na NAP-2, čímž se do místa poranění zhorší přístup neutrofilů.

Doprovodná studie se detailně zabývala procesem krystalizace IRS-2 v přítomnosti stopových množství proteas. Jejich katalytické působení mimikovalo specifické štěpení smyčky reaktivního centra cílovými hostitelskými proteasami. Identifikované štěpení indukovalo celkovou konformační přestavbu molekuly IRS-2 ze stresového stavu S do relaxovaného stavu R. Tato konverze typická pro reakční mechanismus serpinů[52] umožnila nalezení účinných krystalizačních podmínek vedoucích k finálním krystalům IRS-2 s kvalitním rozlišením. Proteasy účastnící se konverze IRS-2 byly detekovány a klasifikovány za použití diagnostických selektivních substrátů a inhibitorů. Vliv stopových proteas na krystalizaci serpinů nebyl dříve studován. Popsaný efekt lze srovnat s používanou technikou limitované proteolýzy krystalizačních směsí proteinů[157], která vede k odstranění flexibilních částí nebo nežádoucích segmentů z proteinových povrchů, což napomáhá procesu krystalizace.

Strukturně-funkční analýza IRS-2 je první publikovaná pro serpin pocházející z klíšťat (a obecně z parazitů). Následující studie slinných serpinů z *I. ricinus* podporují komplexnost funkce serpinů při interakci s hostitelem jako bylo zjištěno pro Iripin-1 ovlivňující zánětlivé odpovědi[158], Iripin-3 působící imunomodulačně a antihemostaticky[159] a Iripin-5 a Iripin-8 působí proti komplementu [160,161]. Inhibiční specifita uvedené skupiny Iripinů je cílena proti proteasám trypsinového typu a IRS-2 je jediným dosud podrobně studovaným slinným serpinem klíšťat s unikátní inhibiční specifitou vůči medicinálně významným proteasám chymotrypsinového typu.

6.3. Proteasový inhibitor IrThy z rodiny tyropinů

IrThy sekretovaný slinnými žlázami klíštěte *I. ricinus* byl charakterizován na úrovni struktury a funkce jako první parazitární zástupce rodiny tyropinů. Molekula tohoto proteinu se skládá ze dvou tyroglobulinových (Tg1) domén a neobsahuje další strukturní moduly, které jsou běžné u velké řady multidoménových proteinů obsahujících Tg1 doménu. Jak ukázala fylogenetická analýza dvoudoménové uspořádání tyropinů je rozšířené u členovců, včetně klíšťat, ale nevyskytuje se u savců. Molekula IrThy dále obsahuje atypické uspořádání disulfidů v C-koncové doméně a na rozhraní obou domén.

Ve srovnání s jinými tyropiny má IrThy unikátní úzkou inhibiční specifitu, kdy inhibuje pouze katepsiny V, K a L vykazující endopeptidasovou aktivitu. Řada v literatuře popsaných tyropinů, jako je fragment p41, equistatin, saxifilin nebo tyropin z vajíček lososa, je schopná inhibovat katepsiny s jak endopeptidasovou, tak exopeptidasovou aktivitou (např. katepsin B nebo H)[99,100,162,163]. Rekombinantní C-koncová doména IrThy (IrThy-Cd) vykazuje stejnou inhibiční specifitu jako celá dvoudoménová molekula. Vyřešení její 3D struktury pomocí NMR spektroskopie umožnilo analyzovat strukturní příčinu inhibiční specifity porovnáním struktury IrThy-Cd s již známou strukturou p41 v komplexu s katepsinem L[55]. Reaktivní centrum p41 nese tři inhibiční smyčky (L1, L2 a L3), které umožňují důležité interakce s oblastí vazebných podmíst S2-S1 a okolními smyčkami na R-doméně proteasy. Takové interakce jsou ale v případě IrThy-Cd negativně ovlivněny dvěma specifickými strukturními změnami v její struktuře: zvětšením velikosti smyčky L1 a změnou tvaru molekuly, která je indukovaná atypickým disulfidem D2 mezi smyčkami L2 a L3. Tyto změny zabraňují vazbě IrThy-Cd na exopeptidasy, jako je katepsin H (aminopeptidasa) a katepsin B (karboxydipeptidasa), jejichž prostorově omezená aktivní místa se stávají nepřístupná pro vazbu studovaného inhibitoru[164,165].

Rekombinantní N-koncová doména (IrThy-Nd) měla širší inhibiční specifitu o dvě další endopeptidasy (katepsin F a papain), než kompletní dvoudoménový IrThy. To naznačuje, že vazebné interakce N-domény jsou poněkud omezeny těsně umístěnou C-doménou ve dvoudoménové molekule. Model celé struktury IrThy byl navržen programem AlphaFold, ze kterého je patrné, že obě reaktivní centra N- a C-domény jsou orientována v opačných směrech a mohou fungovat současně. Obdobné uspořádáním dvou Tg1 domén je popsáno u saxifilinu, kde domény vážou dvě molekuly proteas s různou inhibiční specifitou[163]. K lepšímu pochopení orientace molekuly IrThy v

komplexu s proteasou byly také provedeny předběžné experimenty chemického zesítění IrThy s katepsinem V analyzované hmotnostní spektrometrií. Výsledky ukázaly, že IrThy se přednostně váže do aktivního místa katepsinu V (přidaného v ekvimolárním množství) prostřednictvím reaktivního centra na N-doméně, která vykazuje nižší hodnotu K_i než C-doména. Pro podrobný popis procesu vazby IrThy a využití jednotlivých domén budou zapotřebí systematické experimenty za ruzných koncentračních podmínek a s různými proteasami.

Inhibiční účinnost IrThy je modulována několika glykosaminoglykany (GAG) hojně zastoupenými v tkáních hostitelů. Je známo, že tyto sulfatované polysacharidy regulují cysteinové katepsiny, zejména jejich autokatalytickou aktivaci, pH stabilitu a interakci se specifickými makromolekulárními substráty[166-169]. V přítomnosti GAG se inhibiční účinnost IrThy obecně snižovala vůči katepsinu K, zvyšovala vůči katepsinu V a zůstávala relativně neovlivněná vůči katepsinu L. Molekulární povrchy katepsinů K a V obsahují kladně nabité oblasti pro vazbu GAG[166,169,170], tyto oblasti jsou však méně rozšířené u katepsinu L, což vede k poklesu jeho interakcí s GAG. Lze předpokládat, že GAG regulují inhibici IrThy pomocí kombinované interakce jednak s různými vazebnými místy pro GAG na cílových katepsinech a jednak s pozitivně nabitými oblastmi, které se nacházejí na navrženém modelu povrchu IrThy, zejména na jeho N-doméně. Konkrétní uspořádání těchto mezimolekulárních interakcí může buď posilovat nebo potlačovat vazbu IrThy.

Bylo prokázáno, že inhibitory proteáz ze slin klíštěte injikované do hostitele modulují imunitní odpověď hostitele a potlačují srážení krve v místě přisátí klíštěte [171]. Dosud jedinými studovanými inhibitory cysteinových katepsinů ze slin klíšťat jsou zástupci rodiny cystatinů, kteří mají imunosupresivní a protizánětlivé účinky [138,140,142]. Nedávná srovnávací analýza inhibiční specifity prokázala, že klíštěcí slinné cystatiny se od ostatních členů rodiny cystatinů liší vysokou afinitou k endopeptidasám a omezeným účinkem na exopeptidasy [142]. Tato biochemická vlastnost činí klíštěcí slinné cystatiny funkčně podobné IrThy, který také vykazuje úzkou inhibiční specifitu vůči cysteinovým katepsinům s endopeptidasovou aktivitou. Lze proto předpokládat, že IrThy může hrát roli v imunomodulaci a že na základě dynamiky exprese působí v počátečních fázích interakce klíštěte s hostitelem. V této souvislosti je známo, že katepsiny V, K a L, na které je IrThy cílen, se podílejí na procesech imunitní odpovědi a zánětu a jsou rovněž přítomny v kůži hostitele[172-175]. Další vodítko k potenciální fyziologické funkci IrThy pochází z nedávné studie o U24-ctenitoxinu-Pn1a, neurotoxinu obsaženém v jedu pavouků, který

reguluje napěťově řízené sodíkové kanály (VGSC)[176]. Je zajímavé, že tento protein obsahuje dvě tandemové domény Tg1 a je blízkým homologem IrThy. Toxiny pavouků s VGSC aktivitou jsou obecně antinociceptivní[177] a je lákavé spekulovat o možném analgetickém účinku IrThy v místě přisátí klíštěte.

7. Závěry

Disertační práce se zabývala funkční a strukturní charakterizací tří nových proteasových inhibitorů z rodiny cystatinů, serpinů a tyropinů. Tyto inhibitory byly identifikovány ve slinách klíštěte druhu *Ixodes ricinus*, který je vektorem klíšťové encefalitidy a Lymské boreliózy, a *Ornithodorus moubata*, který je vektorem návratné horečky a afrického prasečího moru. Disertační práce obsahuje výsledky uveřejněné ve čtyřech původních publikacích.

V rámci disertační práce byly splněny zadané cíle s následujícími hlavními závěry:

Cystatin OmC2 z *O. moubata*

- Pro rekombinantní OmC2 byla určena široká inhibiční specifita, která zahrnuje inhibici cysteinových katepsinů jak s endopeptidasovou, tak exopeptidasovou aktivitou. To výrazně odlišuje OmC2 od dosud známých slinných cystatinů klíšťat s typickou úzkou inhibiční specifitou.
- Byly nalezeny krystalizační podmínky a vyřešena prostorová struktura OmC2, která umožnila objasnit jeho inhibiční specifitu. Reaktivní centrum OmC2 vykazuje značnou strukturní podobnost s některými cystatiny hostitele.
- Byl korelován vztah mezi zjištěnými imunomodulačními vlastnostmi OmC2 a inhibovanými cílovými katepsiny hostitele, které hrají významnou roli při imunitní odpovědi. Lze předpokládat, že OmC2 interferuje s aktivitou hostitelských cystatinů a tak mění proteolytickou rovnováhu v tkáni hostitele a spojené relevantní fyziologické procesy ve prospěch klíštěte.

Serpin IRS-2 z *I. ricinus*

- Pro rekombinantní IRS-2 byla určena inhibiční specifita, která cílí unikátní kombinaci serinových proteas trypsinového a chymotrypsinového typu, včetně trombinu, katepsinu G a chymasy. Naopak většina ostatních serpinů ze slin *I. ricinus* inhibuje pouze proteasy trypsinového typu.
- Byly nalezeny krystalizační podmínky a vyřešena prostorová struktura IRS-2 jako první pro serpin parazitárního původu. Tato struktura odpovídala tzv. relaxovanému R-stavu konformace serpinové molekuly po štěpení reaktivního

centra cílovou proteasou. To umožnilo podrobně popsat „sebevražedný“ inhibiční mechanismus a korelovat stavbu reaktivního centra se specifitou IRS-2.

- IRS-2 má výrazné antikoagulační a protizánětlivé vlastnosti výhodné pro sání klíštěte, které lze vysvětlit na základě jeho inhibiční specifity. Ta je zaměřena na regulaci agregace krevních destiček a dále na relevantní imunitní buňky, zejména neutrofily a žírné buňky produkující katepsin G a chymasu.

Tyropin IrThy z *I. ricinus*

- Rekombinantní expresí byl připraven IrThy a jeho dvě separátní tyropinové domény. IrThy má unikátní úzkou inhibiční specifitu a inhibuje pouze tři cysteinové katepsiny K, L a V. Obě domény jsou funkční, ale liší se specifitou.
- Inhibiční aktivita IrThy je modulována komplexním způsobem glykosaminoglykany z tkání hostitele. To zvyšuje selektivitu působení IrThy na katepsiny K a V, které mají vazebná místa pro glykosaminoglykany.
- Prostorová struktura domény IrThy byla určena NMR spektroskopií jako první struktura pro tyropin parazitárního původu. Určená struktura umožnila vysvětlit úzkou inhibiční specifitu IrThy, která je řízena modifikacemi ve stavbě tří vazebných smyček, které neumožňují účinnou interakci s exopeptidasami.
- Na základě strukturně-funkční analýzy a expresního profilu IrThy u *I. ricinus* byly navrženy potenciální fyziologické funkce IrThy v časné fázi interakce klíštěte s hostitelem.

Disertační práce přináší významné informace o nových proteasových inhibitorech ze slin klíšťat, které hrají klíčovou roli v molekulárních mechanismech interakce mezi tímto parazitem s hostitelem. Potenciální využití studovaných bioaktivních inhibitorů a získaných výsledků je v biomedicinských aplikacích při vývoji např. protizánětlivých a antikoagulačních léčiv. Dále představují studované inhibitory zajímavé kandidáty pro vývoj protiklíštěcích vakcín, které narušují sání klíštěte a snižují riziko přenosu patogenů.

Prohlášení spoluautorů

Zuzana (Kovářová) Matoušková se významně podílela na přípravě publikací uvedených v této disertační práci.

Publikace č.1: Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick Ornithodoros moubata.
Salát J, Paesen GC, Řezáčová P, Kotsyfakis M, **Kovářová Z**, Šanda M, Majtán J, Grunclová L, Horká H, Andersen JF, Brynda J, Horn M, Nunn MA, Kopáček P, Kopecký J, Mareš M.
Biochem J. 2010 Jul 1;429(1):103-12.

Publikace č.2: A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation
Chmelař J, Oliveira CJ, Řezáčová P, Francischetti IM, **Kovářová Z**, Pejler G, Kopáček P, Ribeiro JM, Mareš M, Kopecký J, Kotsyfakis M.
Blood. 2011 Jan 13;117(2):736-44.

Publikace č.3: Crystallization and diffraction analysis of the serpin IRS-2 from the hard tick *Ixodes ricinus*
Kovářová Z, Chmelař J, Šanda M, Brynda J, Mareš M, Řezáčová P.
Acta Crystallogr Sect F Struct Biol Cryst Commun. 2010 Nov 1;66(Pt 11):1453-7.

Publikace č.4: An Unusual Two-Domain Tyropin from Tick Saliva: NMR Solution Structure and Highly Selective Inhibition of Cysteine Cathepsins Modulated by Glycosaminoglycans
Matoušková Z, Orsághová K, Srb P, Pytelková J, Kukačka Z, Buša M, Hajdušek O, Šíma R, Fábry M, Novák P, Horn M, Kopáček P, Mareš M.
Int J Mol Sci. 2024 Feb 13;25(4):2240.

……
RNDr. Michael Mareš, CSc.

školitel a spoluautor všech uvedených publikací

8. Seznam použité literatury

1. Baneth, G. Tick-borne infections of animals and humans: a common ground. *Int J Parasitol* 44, 591-596, doi:10.1016/j.ijpara.2014.03.011, (2014).
2. Jongejan, F.; Uilenberg, G. The global importance of ticks. *Parasitology* 129 Suppl, S3-14, (2004).
3. Rocha, S.C.; Velásquez, C.V.; Aquib, A.; Al-Nazal, A.; Parveen, N. Transmission Cycle of Tick-Borne Infections and Co-Infections, Animal Models and Diseases. *Pathogens* 11, doi:10.3390/pathogens11111309, (2022).
4. Wu, X.B.; Na, R.H.; Wei, S.S.; Zhu, J.S.; Peng, H.J. Distribution of tick-borne diseases in China. *Parasit Vectors* 6, 119, doi:10.1186/1756-3305-6-119, (2013).
5. Bush, L.M.; Vazquez-Pertejo, M.T. Tick borne illness-Lyme disease. *Dis Mon* 64, 195-212, doi:10.1016/j.disamonth.2018.01.007, (2018).
6. Nau, R.; Christen, H.J.; Eiffert, H. Lyme disease--current state of knowledge. *Dtsch Arztebl Int* 106, 72-81; quiz 82, I, doi:10.3238/arztebl.2009.0072, (2009).
7. Yoshii, K. Epidemiology and pathological mechanisms of tick-borne encephalitis. *J Vet Med Sci* 81, 343-347, doi:10.1292/jvms.18-0373, (2019).
8. Worku, D.A. Tick-Borne Encephalitis (TBE): From Tick to Pathology. *J Clin Med* 12, doi:10.3390/jcm12216859, (2023).
9. Smit, R.; Postma, M.J. The Burden of Tick-Borne Encephalitis in Disability-Adjusted Life Years (DALYs) for Slovenia. *PLoS One* 10, e0144988, doi:10.1371/journal.pone.0144988, (2015).
10. Grisi, L.; Leite, R.C.; Martins, J.R.; Barros, A.T.; Andreotti, R.; Cancado, P.H.; Leon, A.A.; Pereira, J.B.; Villela, H.S. Reassessment of the potential economic impact of cattle parasites in Brazil. *Rev Bras Parasitol Vet* 23, 150-156, (2014).
11. Narladkar, B.W. Projected economic losses due to vector and vector-borne parasitic diseases in livestock of India and its significance in implementing the concept of integrated practices for vector management. *Vet World* 11, 151-160, doi:10.14202/vetworld.2018.151-160, (2018).
12. Merino, O.; Alberdi, P.; Perez de la Lastra, J.M.; de la Fuente, J. Tick vaccines and the control of tick-borne pathogens. *Front Cell Infect Microbiol* 3, 30, doi:10.3389/fcimb.2013.00030, (2013).
13. Abbas, M.N.; Jmel, M.A.; Mekki, I.; Dijkgraaf, I.; Kotsyfakis, M. Recent Advances in Tick Antigen Discovery and Anti-Tick Vaccine Development. *Int J Mol Sci* 24, doi:10.3390/ijms24054969, (2023).
14. Ali, A.; Zeb, I.; Alouffi, A.; Zahid, H.; Almutairi, M.M.; Ayed Alshammari, F.; Alrouji, M.; Termignoni, C.; Vaz, I.D.S., Jr.; Tanaka, T. Host Immune Responses to Salivary Components - A Critical Facet of Tick-Host Interactions. *Front Cell Infect Microbiol* 12, 809052, doi:10.3389/fcimb.2022.809052, (2022).
15. Singh, A.P. Medicinal leech therapy (hirudotherapy): a brief overview. *Complement Ther Clin Pract* 16, 213-215, doi:10.1016/j.ctcp.2009.11.005, (2010).
16. Fujinaga, M.; Cherney, M.M.; Oyama, H.; Oda, K.; James, M.N. The molecular structure and catalytic mechanism of a novel carboxyl peptidase from Scytalidium lignicolum. *Proc Natl Acad Sci U S A* 101, 3364-3369, doi:10.1073/pnas.0400246101, (2004).
17. Rawlings, N.D.; Barrett, A.J. Evolutionary families of peptidases. *Biochem J* 290 (Pt 1), 205-218, (1993).
18. Seemuller, E.; Lupas, A.; Stock, D.; Lowe, J.; Huber, R.; Baumeister, W. Proteasome from Thermoplasma acidophilum: a threonine protease. *Science* 268, 579-582, (1995).
19. Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* 27, 157-162, doi:10.1016/s0006-291x(67)80055-x, (1967).

20. Ai, Y.; Meng, Y.; Yan, B.; Zhou, Q.; Wang, X. The biochemical pathways of apoptotic, necroptotic, pyroptotic, and ferroptotic cell death. *Mol Cell* 84, 170-179, doi:10.1016/j.molcel.2023.11.040, (2024).
21. Al-Amer, O.M. The role of thrombin in haemostasis. *Blood Coagul Fibrinolysis* 33, 145-148, doi:10.1097/mbc.0000000000001130, (2022).
22. Lu, P.; Takai, K.; Weaver, V.M.; Werb, Z. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harb Perspect Biol* 3, doi:10.1101/cshperspect.a005058, (2011).
23. Turk, B. Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov* 5, 785-799, doi:10.1038/nrd2092, (2006).
24. van den Eshof, B.L.; Medfai, L.; Nolfi, E.; Wawrzyniuk, M.; Sijts, A. The Function of Immunoproteasomes-An Immunologists' Perspective. *Cells* 10, doi:10.3390/cells10123360, (2021).
25. Vidak, E.; Javoršek, U.; Vizovišek, M.; Turk, B. Cysteine Cathepsins and their Extracellular Roles: Shaping the Microenvironment. *Cells* 8, doi:10.3390/cells8030264, (2019).
26. López-Otín, C.; Bond, J.S. Proteases: multifunctional enzymes in life and disease. *J Biol Chem* 283, 30433-30437, doi:10.1074/jbc.R800035200, (2008).
27. Brix, K.; Dunkhorst, A.; Mayer, K.; Jordans, S. Cysteine cathepsins: cellular roadmap to different functions. *Biochimie* 90, 194-207, doi:10.1016/j.biochi.2007.07.024, (2008).
28. Shen, A. Allosteric regulation of protease activity by small molecules. *Mol Biosyst* 6, 1431-1443, doi:10.1039/c003913f, (2010).
29. Wiederanders, B.; Kaulmann, G.; Schilling, K. Functions of propeptide parts in cysteine proteases. *Curr Protein Pept Sci* 4, 309-326, doi:10.2174/1389203033487081, (2003).
30. Chakraborty, S.; Bhattacharya, S.; Ghosh, S.; Bera, A.K.; Haldar, U.; Pal, A.K.; Mukhopadhyay, B.P.; Banerjee, A. Structural and interactional homology of clinically potential trypsin inhibitors: molecular modelling of cucurbitaceae family peptides using the X-ray structure of MCTI-II. *Protein Eng* 13, 551-555, doi:10.1093/protein/13.8.551, (2000).
31. Jiang, R.; Zhang, B.; Kurokawa, K.; So, Y.I.; Kim, E.H.; Hwang, H.O.; Lee, J.H.; Shiratsuchi, A.; Zhang, J.; Nakanishi, Y.; et al. 93-kDa twin-domain serine protease inhibitor (Serpin) has a regulatory function on the beetle Toll proteolytic signaling cascade. *J Biol Chem* 286, 35087-35095, doi:10.1074/jbc.M111.277343, (2011).
32. Khezri, M.R.; Ghasemnejad-Berenji, M. The Role of Caspases in Alzheimer's Disease: Pathophysiology Implications and Pharmacologic Modulation. *J Alzheimers Dis* 91, 71-90, doi:10.3233/jad-220873, (2023).
33. Slack, M.A.; Gordon, S.M. Protease Activity in Vascular Disease. *Arterioscler Thromb Vasc Biol* 39, e210-e218, doi:10.1161/atvbaha.119.312413, (2019).
34. Heutinck, K.M.; ten Berge, I.J.; Hack, C.E.; Hamann, J.; Rowshani, A.T. Serine proteases of the human immune system in health and disease. *Mol Immunol* 47, 1943-1955, doi:10.1016/j.molimm.2010.04.020, (2010).
35. Lafarge, J.C.; Naour, N.; Clément, K.; Guerre-Millo, M. Cathepsins and cystatin C in atherosclerosis and obesity. *Biochimie* 92, 1580-1586, doi:10.1016/j.biochi.2010.04.011, (2010).
36. Rawlings, N.D.; Tolle, D.P.; Barrett, A.J. Evolutionary families of peptidase inhibitors. *Biochem J* 378, 705-716, doi:10.1042/bj20031825, (2004).
37. Ranasinghe, S.L.; Duke, M.; Harvie, M.; McManus, D.P. Kunitz-type protease inhibitor as a vaccine candidate against schistosomiasis mansoni. *Int J Infect Dis* 66, 26-32, doi:10.1016/j.ijid.2017.10.024, (2018).
38. Knox, D.P. Proteinase inhibitors and helminth parasite infection. *Parasite Immunol* 29, 57-71, doi:10.1111/j.1365-3024.2006.00913.x, (2007).
39. Sacks, D.; Baxter, B.; Campbell, B.C.V.; Carpenter, J.S.; Cognard, C.; Dippel, D.; Eesa, M.; Fischer, U.; Hausegger, K.; Hirsch, J.A.; et al. Multisociety Consensus Quality

Improvement Revised Consensus Statement for Endovascular Therapy of Acute Ischemic Stroke. *Int J Stroke* 13, 612-632, doi:10.1177/1747493018778713, (2018).
40. Rawlings, N.D.; Barrett, A.J.; Thomas, P.D.; Huang, X.; Bateman, A.; Finn, R.D. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res* 46, D624-d632, doi:10.1093/nar/gkx1134, (2018).
41. Borth, W. Alpha 2-macroglobulin, a multifunctional binding protein with targeting characteristics. *Faseb j* 6, 3345-3353, doi:10.1096/fasebj.6.15.1281457, (1992).
42. Fischer, K.G. The role of recombinant hirudins in the management of thrombotic disorders. *BioDrugs* 18, 235-268, doi:10.2165/00063030-200418040-00003, (2004).
43. Ye, S.; Goldsmith, E.J. Serpins and other covalent protease inhibitors. *Curr Opin Struct Biol* 11, 740-745, doi:10.1016/s0959-440x(01)00275-5, (2001).
44. Farady, C.J.; Craik, C.S. Mechanisms of macromolecular protease inhibitors. *Chembiochem* 11, 2341-2346, doi:10.1002/cbic.201000442, (2010).
45. Bendre, A.D.; Ramasamy, S.; Suresh, C.G. Analysis of Kunitz inhibitors from plants for comprehensive structural and functional insights. *Int J Biol Macromol* 113, 933-943, doi:10.1016/j.ijbiomac.2018.02.148, (2018).
46. Dubin, G. Proteinaceous cysteine protease inhibitors. *Cell Mol Life Sci* 62, 653-669, doi:10.1007/s00018-004-4445-9, (2005).
47. Law, R.H.; Zhang, Q.; McGowan, S.; Buckle, A.M.; Silverman, G.A.; Wong, W.; Rosado, C.J.; Langendorf, C.G.; Pike, R.N.; Bird, P.I.; Whisstock, J.C. An overview of the serpin superfamily. *Genome Biol* 7, 216, doi:10.1186/gb-2006-7-5-216, (2006).
48. Otlewski, J.; Jelen, F.; Zakrzewska, M.; Oleksy, A. The many faces of protease-protein inhibitor interaction. *Embo j* 24, 1303-1310, doi:10.1038/sj.emboj.7600611, (2005).
49. Krowarsch, D.; Cierpicki, T.; Jelen, F.; Otlewski, J. Canonical protein inhibitors of serine proteases. *Cellular and Molecular Life Sciences CMLS* 60, 2427-2444, doi:10.1007/s00018-003-3120-x, (2003).
50. Scheidig, A.J.; Hynes, T.R.; Pelletier, L.A.; Wells, J.A.; Kossiakoff, A.A. Crystal structures of bovine chymotrypsin and trypsin complexed to the inhibitor domain of Alzheimer's amyloid beta-protein precursor (APPI) and basic pancreatic trypsin inhibitor (BPTI): engineering of inhibitors with altered specificities. *Protein Sci* 6, 1806-1824, doi:10.1002/pro.5560060902, (1997).
51. Gettins, P.G.W.; Olson, S.T. Inhibitory serpins. New insights into their folding, polymerization, regulation and clearance. *The Biochemical journal* 473, 2273-2293, doi:10.1042/BCJ20160014, (2016).
52. Huntington, J.A.; Read, R.J.; Carrell, R.W. Structure of a serpin–protease complex shows inhibition by deformation. *Nature* 407, 923-926, doi:10.1038/35038119, (2000).
53. Turk, V.; Stoka, V.; Turk, D. Cystatins: biochemical and structural properties, and medical relevance. *Front Biosci* 13, 5406-5420, doi:10.2741/3089, (2008).
54. Wang, S.X.; Pandey, K.C.; Scharfstein, J.; Whisstock, J.; Huang, R.K.; Jacobelli, J.; Fletterick, R.J.; Rosenthal, P.J.; Abrahamson, M.; Brinen, L.S.; et al. The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. *Structure* 15, 535-543, doi:10.1016/j.str.2007.03.012, (2007).
55. Guncar, G.; Pungercic, G.; Klemencic, I.; Turk, V.; Turk, D. Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. *Embo j* 18, 793-803, doi:10.1093/emboj/18.4.793, (1999).
56. Jenko, S.; Dolenc, I.; Guncar, G.; Dobersek, A.; Podobnik, M.; Turk, D. Crystal structure of Stefin A in complex with cathepsin H: N-terminal residues of inhibitors can adapt to the active sites of endo- and exopeptidases. *J Mol Biol* 326, 875-885, doi:10.1016/s0022-2836(02)01432-8, (2003).

57. Bitoun, E.; Chavanas, S.; Irvine, A.D.; Lonie, L.; Bodemer, C.; Paradisi, M.; Hamel-Teillac, D.; Ansai, S.; Mitsuhashi, Y.; Taieb, A.; et al. Netherton syndrome: disease expression and spectrum of SPINK5 mutations in 21 families. *J Invest Dermatol* 118, 352-361, doi:10.1046/j.1523-1747.2002.01603.x, (2002).
58. Lehesjoki, A.E. Molecular background of progressive myoclonus epilepsy. *Embo j* 22, 3473-3478, doi:10.1093/emboj/cdg338, (2003).
59. Ritchie, B.C. Protease inhibitors in the treatment of hereditary angioedema. *Transfus Apher Sci* 29, 259-267, doi:10.1016/j.transci.2003.08.004, (2003).
60. Schierack, P.; Lucius, R.; Sonnenburg, B.; Schilling, K.; Hartmann, S. Parasite-specific immunomodulatory functions of filarial cystatin. *Infect Immun* 71, 2422-2429, doi:10.1128/iai.71.5.2422-2429.2003, (2003).
61. Zang, X.; Yazdanbakhsh, M.; Jiang, H.; Kanost, M.R.; Maizels, R.M. A novel serpin expressed by blood-borne microfilariae of the parasitic nematode Brugia malayi inhibits human neutrophil serine proteinases. *Blood* 94, 1418-1428, (1999).
62. Ochieng, J.; Chaudhuri, G. Cystatin superfamily. *J Health Care Poor Underserved* 21, 51-70, doi:10.1353/hpu.0.0257, (2010).
63. Stubbs, M.T.; Laber, B.; Bode, W.; Huber, R.; Jerala, R.; Lenarcic, B.; Turk, V. The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *Embo j* 9, 1939-1947, (1990).
64. Turk, B.; Turk, D.; Salvesen, G.S. Regulating cysteine protease activity: essential role of protease inhibitors as guardians and regulators. *Curr Pharm Des* 8, 1623-1637, (2002).
65. Henskens, Y.M.; Veerman, E.C.; Nieuw Amerongen, A.V. Cystatins in health and disease. *Biol Chem Hoppe Seyler* 377, 71-86, (1996).
66. Magister, S.; Kos, J. Cystatins in immune system. *J Cancer* 4, 45-56, doi:10.7150/jca.5044, (2013).
67. Olson, O.C.; Joyce, J.A. Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nat Rev Cancer* 15, 712-729, doi:10.1038/nrc4027, (2015).
68. Breznik, B.; Mitrović, A.; T, T.L.; Kos, J. Cystatins in cancer progression: More than just cathepsin inhibitors. *Biochimie* 166, 233-250, doi:10.1016/j.biochi.2019.05.002, (2019).
69. Vasiljeva, O.; Reinheckel, T.; Peters, C.; Turk, D.; Turk, V.; Turk, B. Emerging roles of cysteine cathepsins in disease and their potential as drug targets. *Curr Pharm Des* 13, 387-403, (2007).
70. Hasanbasic, S.; Jahic, A.; Karahmet, E.; Sejranic, A.; Prnjavorac, B. THE ROLE OF CYSTEINE PROTEASE IN ALZHEIMER DISEASE. *Mater Sociomed* 28, 235-238, doi:10.5455/msm.2016.28.235-238, (2016).
71. Levy, E.; Jaskolski, M.; Grubb, A. The role of cystatin C in cerebral amyloid angiopathy and stroke: cell biology and animal models. *Brain Pathol* 16, 60-70, doi:10.1111/j.1750-3639.2006.tb00562.x, (2006).
72. Benoit, S.W.; Ciccia, E.A.; Devarajan, P. Cystatin C as a biomarker of chronic kidney disease: latest developments. *Expert Rev Mol Diagn* 20, 1019-1026, doi:10.1080/14737159.2020.1768849, (2020).
73. Sun, Y.; Lu, Q.; Cheng, B.; Tao, X. Prognostic value of cystatin C in patients with acute coronary syndrome: A systematic review and meta-analysis. *Eur J Clin Invest* 51, e13440, doi:10.1111/eci.13440, (2021).
74. Zinellu, A.; Mangoni, A.A. Cystatin C, COVID-19 severity and mortality: a systematic review and meta-analysis. *J Nephrol* 35, 59-68, doi:10.1007/s40620-021-01139-2, (2022).
75. Žerovnik, E. Human stefin B: from its structure, folding, and aggregation to its function in health and disease. *Front Mol Neurosci* 15, 1009976, doi:10.3389/fnmol.2022.1009976, (2022).

76. Singh, S.; Hämäläinen, R.H. The Roles of Cystatin B in the Brain and Pathophysiological Mechanisms of Progressive Myoclonic Epilepsy Type 1. *Cells* 13, doi:10.3390/cells13020170, (2024).
77. Irving, J.A.; Pike, R.N.; Lesk, A.M.; Whisstock, J.C. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res* 10, 1845-1864, (2000).
78. Gettins, P.G. Serpin structure, mechanism, and function. *Chem Rev* 102, 4751-4804, (2002).
79. Rein, C.M.; Desai, U.R.; Church, F.C. Serpin-glycosaminoglycan interactions. *Methods Enzymol* 501, 105-137, doi:10.1016/b978-0-12-385950-1.00007-9, (2011).
80. Khan, M.S.; Singh, P.; Azhar, A.; Naseem, A.; Rashid, Q.; Kabir, M.A.; Jairajpuri, M.A. Serpin Inhibition Mechanism: A Delicate Balance between Native Metastable State and Polymerization. *J Amino Acids* 2011, 606797, doi:10.4061/2011/606797, (2011).
81. Li, W.; Johnson, D.J.; Esmon, C.T.; Huntington, J.A. Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat Struct Mol Biol* 11, 857-862, doi:10.1038/nsmb811, (2004).
82. Liaw, P.C.; Austin, R.C.; Fredenburgh, J.C.; Stafford, A.R.; Weitz, J.I. Comparison of heparin- and dermatan sulfate-mediated catalysis of thrombin inactivation by heparin cofactor II. *J Biol Chem* 274, 27597-27604, doi:10.1074/jbc.274.39.27597, (1999).
83. Li, W.; Adams, T.E.; Kjellberg, M.; Stenflo, J.; Huntington, J.A. Structure of native protein C inhibitor provides insight into its multiple functions. *J Biol Chem* 282, 13759-13768, doi:10.1074/jbc.M701074200, (2007).
84. Rau, J.C.; Beaulieu, L.M.; Huntington, J.A.; Church, F.C. Serpins in thrombosis, hemostasis and fibrinolysis. *J Thromb Haemost* 5 Suppl 1, 102-115, doi:10.1111/j.1538-7836.2007.02516.x, (2007).
85. 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.; et al. The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *J Biol Chem* 276, 33293-33296, doi:10.1074/jbc.R100016200, (2001).
86. Mebius, M.M.; van Genderen, P.J.; Urbanus, R.T.; Tielens, A.G.; de Groot, P.G.; van Hellemond, J.J. Interference with the host haemostatic system by schistosomes. *PLoS Pathog* 9, e1003781, doi:10.1371/journal.ppat.1003781, (2013).
87. Kobpornchai, P.; Reamtong, O.; Phuphisut, O.; Malaitong, P.; Adisakwattana, P. Serine protease inhibitor derived from Trichinella spiralis (TsSERP) inhibits neutrophil elastase and impairs human neutrophil functions. *Front Cell Infect Microbiol* 12, 919835, doi:10.3389/fcimb.2022.919835, (2022).
88. Gooptu, B.; Lomas, D.A. Conformational pathology of the serpins: themes, variations, and therapeutic strategies. *Annu Rev Biochem* 78, 147-176, doi:10.1146/annurev.biochem.78.082107.133320, (2009).
89. Godinez, A.; Rajput, R.; Chitranshi, N.; Gupta, V.; Basavarajappa, D.; Sharma, S.; You, Y.; Pushpitha, K.; Dhiman, K.; Mirzaei, M.; et al. Neuroserpin, a crucial regulator for axogenesis, synaptic modelling and cell-cell interactions in the pathophysiology of neurological disease. *Cell Mol Life Sci* 79, 172, doi:10.1007/s00018-022-04185-6, (2022).
90. Rezaie, A.R.; Giri, H. Anticoagulant and signaling functions of antithrombin. *Journal of Thrombosis and Haemostasis* 18, 3142-3153, doi.org/10.1111/jth.15052, (2020).
91. Bergin, D.A.; Hurley, K.; McElvaney, N.G.; Reeves, E.P. Alpha-1 antitrypsin: a potent anti-inflammatory and potential novel therapeutic agent. *Arch Immunol Ther Exp (Warsz)* 60, 81-97, doi:10.1007/s00005-012-0162-5, (2012).
92. Toumpanakis, D.; Usmani, O.S. Small airways disease in patients with alpha-1 antitrypsin deficiency. *Respir Med* 211, 107222, doi:10.1016/j.rmed.2023.107222, (2023).

93. Bai, X.; Schountz, T.; Buckle, A.M.; Talbert, J.L.; Sandhaus, R.A.; Chan, E.D. Alpha-1-antitrypsin antagonizes COVID-19: a review of the epidemiology, molecular mechanisms, and clinical evidence. *Biochem Soc Trans* 51, 1361-1375, doi:10.1042/bst20230078, (2023).
94. Teckman, J.H. Liver disease in alpha-1 antitrypsin deficiency: current understanding and future therapy. *Copd* 10 Suppl 1, 35-43, doi:10.3109/15412555.2013.765839, (2013).
95. Sillen, M.; Declerck, P.J. A Narrative Review on Plasminogen Activator Inhibitor-1 and Its (Patho)Physiological Role: To Target or Not to Target? *Int J Mol Sci* 22, doi:10.3390/ijms22052721, (2021).
96. Novinec, M.; Kordis, D.; Turk, V.; Lenarcic, B. Diversity and evolution of the thyroglobulin type-1 domain superfamily. *Mol Biol Evol* 23, 744-755, doi:10.1093/molbev/msj082, (2006).
97. Molina, F.; Bouanani, M.; Pau, B.; Granier, C. Characterization of the type-1 repeat from thyroglobulin, a cysteine-rich module found in proteins from different families. *Eur J Biochem* 240, 125-133, (1996).
98. Headey, S.J.; Keizer, D.W.; Yao, S.; Brasier, G.; Kantharidis, P.; Bach, L.A.; Norton, R.S. C-Terminal Domain of Insulin-Like Growth Factor (IGF) Binding Protein-6: Structure and Interaction with IGF-II. *Molecular Endocrinology* 18, 2740-2750, doi:10.1210/me.2004-0248, (2004).
99. Lenarcic, B.; Turk, V. Thyroglobulin type-1 domains in equistatin inhibit both papain-like cysteine proteinases and cathepsin D. *J Biol Chem* 274, 563-566, doi:10.1074/jbc.274.2.563, (1999).
100. Mihelic, M.; Dobersek, A.; Guncar, G.; Turk, D. Inhibitory fragment from the p41 form of invariant chain can regulate activity of cysteine cathepsins in antigen presentation. *J Biol Chem* 283, 14453-14460, doi:10.1074/jbc.M801283200, (2008).
101. Brown, T.C.; Sankpal, N.V.; Gillanders, W.E. Functional Implications of the Dynamic Regulation of EpCAM during Epithelial-to-Mesenchymal Transition. *Biomolecules* 11, doi:10.3390/biom11070956, (2021).
102. Fineschi, B.; Sakaguchi, K.; Appella, E.; Miller, J. The proteolytic environment involved in MHC class II-restricted antigen presentation can be modulated by the p41 form of invariant chain. *J Immunol* 157, 3211-3215, (1996).
103. Lennon-Duménil, A.-M.; Roberts, R.A.; Valentijn, K.; Driessen, C.; Overkleeft, H.S.; Erickson, A.; Peters, P.J.; Bikoff, E.; Ploegh, H.L.; Bryant, P.W. The p41 isoform of invariant chain is a chaperone for cathepsin L. *The EMBO Journal* 20, 4055-4064, doi.org/10.1093/emboj/20.15.4055, (2001).
104. Fiebiger, E.; Maehr, R.; Villadangos, J.; Weber, E.; Erickson, A.; Bikoff, E.; Ploegh, H.L.; Lennon-Duménil, A.M. Invariant chain controls the activity of extracellular cathepsin L. *J Exp Med* 196, 1263-1269, doi:10.1084/jem.20020762, (2002).
105. Nepveu-Traversy, M.E.; Fausther-Bovendo, H.; Babuadze, G.G. Human Tick-Borne Diseases and Advances in Anti-Tick Vaccine Approaches: A Comprehensive Review. *Vaccines (Basel)* 12, doi:10.3390/vaccines12020141, (2024).
106. AGES. Health for humans, animals & plants Available online: https://www.ages.at/en/human/disease/ticks-diseases-info (accessed on 21.1.2024).
107. Nogueira, B.C.F.; Campos, A.K.; Muñoz-Leal, S.; Pinter, A.; Martins, T.F. Soft and hard ticks (Parasitiformes: Ixodida) on humans: A review of Brazilian biomes and the impact of environmental change. *Acta Trop* 234, 106598, doi:10.1016/j.actatropica.2022.106598, (2022).
108. Billingsley, P.F.; Hecker, H. Blood digestion in the mosquito, Anopheles stephensi Liston (Diptera: Culicidae): activity and distribution of trypsin, aminopeptidase, and alpha-glucosidase in the midgut. *J Med Entomol* 28, 865-871, doi:10.1093/jmedent/28.6.865, (1991).

109. Mendiola, J.; Alonso, M.; Marquetti, M.C.; Finlay, C. Boophilus microplus: multiple proteolytic activities in the midgut. *Exp Parasitol* 82, 27-33, doi:10.1006/expr.1996.0004, (1996).
110. Horn, M.; Nussbaumerová, M.; Sanda, M.; Kovárová, Z.; Srba, J.; Franta, Z.; Sojka, D.; Bogyo, M.; Caffrey, C.R.; Kopácek, P.; Mares, M. Hemoglobin digestion in blood-feeding ticks: mapping a multipeptidase pathway by functional proteomics. *Chem Biol* 16, 1053-1063, doi:10.1016/j.chembiol.2009.09.009, (2009).
111. Franta, Z.; Frantová, H.; Konvičková, J.; Horn, M.; Sojka, D.; Mareš, M.; Kopáček, P. Dynamics of digestive proteolytic system during blood feeding of the hard tick Ixodes ricinus. *Parasit Vectors* 3, 119, doi:10.1186/1756-3305-3-119, (2010).
112. Xavier, M.A.; Tirloni, L.; Pinto, A.F.M.; Diedrich, J.K.; Yates, J.R., 3rd; Mulenga, A.; Logullo, C.; da Silva Vaz, I., Jr.; Seixas, A.; Termignoni, C. A proteomic insight into vitellogenesis during tick ovary maturation. *Sci Rep* 8, 4698, doi:10.1038/s41598-018-23090-2, (2018).
113. Schwarz, A.; von Reumont, B.M.; Erhart, J.; Chagas, A.C.; Ribeiro, J.M.; Kotsyfakis, M. De novo Ixodes ricinus salivary gland transcriptome analysis using two next-generation sequencing methodologies. *Faseb j* 27, 4745-4756, doi:10.1096/fj.13-232140, (2013).
114. Yu, X.; Zhou, Y.; Cao, J.; Zhang, H.; Gong, H.; Zhou, J. Caspase-1 participates in apoptosis of salivary glands in Rhipicephalus haemaphysaloides. *Parasit Vectors* 10, 225, doi:10.1186/s13071-017-2161-1, (2017).
115. Seixas, A.; Dos Santos, P.C.; Velloso, F.F.; Da Silva Vaz, I., Jr.; Masuda, A.; Horn, F.; Termignoni, C. A Boophilus microplus vitellin-degrading cysteine endopeptidase. *Parasitology* 126, 155-163, doi:10.1017/s0031182002002731, (2003).
116. Qiu, Z.X.; Li, Y.; Li, M.M.; Wang, W.Y.; Zhang, T.T.; Liu, J.Z. Investigation of three enzymes and their roles in the embryonic development of parthenogenetic Haemaphysalis longicornis. *Parasit Vectors* 13, 46, doi:10.1186/s13071-020-3916-7, (2020).
117. Santiago, P.B.; de Araújo, C.N.; Motta, F.N.; Praça, Y.R.; Charneau, S.; Bastos, I.M.; Santana, J.M. Proteases of haematophagous arthropod vectors are involved in blood-feeding, yolk formation and immunity - a review. *Parasit Vectors* 10, 79, doi:10.1186/s13071-017-2005-z, (2017).
118. Cruz, C.E.; Fogaça, A.C.; Nakayasu, E.S.; Angeli, C.B.; Belmonte, R.; Almeida, I.C.; Miranda, A.; Miranda, M.T.; Tanaka, A.S.; Braz, G.R.; et al. Characterization of proteinases from the midgut of Rhipicephalus (Boophilus) microplus involved in the generation of antimicrobial peptides. *Parasit Vectors* 3, 63, doi:10.1186/1756-3305-3-63, (2010).
119. Schwarz, A.; Valdés, J.J.; Kotsyfakis, M. The role of cystatins in tick physiology and blood feeding. *Ticks Tick Borne Dis* 3, 117-127, doi:10.1016/j.ttbdis.2012.03.004, (2012).
120. Meekins, D.A.; Kanost, M.R.; Michel, K. Serpins in arthropod biology. *Semin Cell Dev Biol* 62, 105-119, doi:10.1016/j.semcdb.2016.09.001, (2017).
121. Ceraul, S.M.; Dreher-Lesnick, S.M.; Mulenga, A.; Rahman, M.S.; Azad, A.F. Functional characterization and novel rickettsiostatic effects of a Kunitz-type serine protease inhibitor from the tick Dermacentor variabilis. *Infect Immun* 76, 5429-5435, doi:10.1128/iai.00866-08, (2008).
122. Mans, B.J. Evolution of vertebrate hemostatic and inflammatory control mechanisms in blood-feeding arthropods. *J Innate Immun* 3, 41-51, doi:10.1159/000321599, (2011).
123. Mans, B.J.; Neitz, A.W. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect Biochem Mol Biol* 34, 1-17, doi:10.1016/j.ibmb.2003.09.002, (2004).
124. Zavašnik-Bergant, T.; Vidmar, R.; Sekirnik, A.; Fonović, M.; Salát, J.; Grunclová, L.; Kopáček, P.; Turk, B. Salivary Tick Cystatin OmC2 Targets Lysosomal Cathepsins S and C in Human Dendritic Cells. *Front Cell Infect Microbiol* 7, 288, doi:10.3389/fcimb.2017.00288, (2017).

125. Andrade, B.B.; Teixeira, C.R.; Barral, A.; Barral-Netto, M. Haematophagous arthropod saliva and host defense system: a tale of tear and blood. *An Acad Bras Cienc* 77, 665-693, doi:/S0001-37652005000400008, (2005).
126. Ribeiro, J.M.; Francischetti, I.M. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annu Rev Entomol* 48, 73-88, doi:10.1146/annurev.ento.48.060402.102812, (2003).
127. Simo, L.; Kazimirova, M.; Richardson, J.; Bonnet, S.I. The Essential Role of Tick Salivary Glands and Saliva in Tick Feeding and Pathogen Transmission. *Front Cell Infect Microbiol* 7, 281, doi:10.3389/fcimb.2017.00281, (2017).
128. Porter, L.M.; Radulović Ž, M.; Mulenga, A. A repertoire of protease inhibitor families in Amblyomma americanum and other tick species: inter-species comparative analyses. *Parasit Vectors* 10, 152, doi:10.1186/s13071-017-2080-1, (2017).
129. Jmel, M.A.; Voet, H.; Araújo, R.N.; Tirloni, L.; Sá-Nunes, A.; Kotsyfakis, M. Tick Salivary Kunitz-Type Inhibitors: Targeting Host Hemostasis and Immunity to Mediate Successful Blood Feeding. *Int J Mol Sci* 24, doi:10.3390/ijms24021556, (2023).
130. Chmelar, J.; Calvo, E.; Pedra, J.H.; Francischetti, I.M.; Kotsyfakis, M. Tick salivary secretion as a source of antihemostatics. *J Proteomics* 75, 3842-3854, doi:10.1016/j.jprot.2012.04.026, (2012).
131. Decrem, Y.; Rath, G.; Blasioli, V.; Cauchie, P.; Robert, S.; Beaufays, J.; Frère, J.M.; Feron, O.; Dogné, J.M.; Dessy, C.; et al. Ir-CPI, a coagulation contact phase inhibitor from the tick Ixodes ricinus, inhibits thrombus formation without impairing hemostasis. *J Exp Med* 206, 2381-2395, doi:10.1084/jem.20091007, (2009).
132. van de Locht, A.; Stubbs, M.T.; Bode, W.; Friedrich, T.; Bollschweiler, C.; Höffken, W.; Huber, R. The ornithodorin-thrombin crystal structure, a key to the TAP enigma? *Embo j* 15, 6011-6017, (1996).
133. De Paula, V.S.; Sgourakis, N.G.; Francischetti, I.M.B.; Almeida, F.C.L.; Monteiro, R.Q.; Valente, A.P. NMR structure determination of Ixolaris and factor X(a) interaction reveals a noncanonical mechanism of Kunitz inhibition. *Blood* 134, 699-708, doi:10.1182/blood.2018889493, (2019).
134. Blisnick, A.A.; Šimo, L.; Grillon, C.; Fasani, F.; Brûlé, S.; Le Bonniec, B.; Prina, E.; Marsot, M.; Relmy, A.; Blaise-Boisseau, S.; et al. The Immunomodulatory Effect of IrSPI, a Tick Salivary Gland Serine Protease Inhibitor Involved in Ixodes ricinus Tick Feeding. *Vaccines (Basel)* 7, doi:10.3390/vaccines7040148, (2019).
135. Mulenga, A.; Khumthong, R.; Chalaire, K.C. Ixodes scapularis tick serine proteinase inhibitor (serpin) gene family; annotation and transcriptional analysis. *BMC Genomics* 10, 217, doi:10.1186/1471-2164-10-217, (2009).
136. Chmelar, J.; Oliveira, C.J.; Rezacova, P.; Francischetti, I.M.; Kovarova, Z.; Pejler, G.; Kopacek, P.; Ribeiro, J.M.; Mares, M.; Kopecky, J.; Kotsyfakis, M. A tick salivary protein targets cathepsin G and chymase and inhibits host inflammation and platelet aggregation. *Blood* 117, 736-744, doi:10.1182/blood-2010-06-293241, (2011).
137. Kim, T.K.; Tirloni, L.; Radulovic, Z.; Lewis, L.; Bakshi, M.; Hill, C.; da Silva Vaz, I., Jr.; Logullo, C.; Termignoni, C.; Mulenga, A. Conserved Amblyomma americanum tick Serpin19, an inhibitor of blood clotting factors Xa and XIa, trypsin and plasmin, has anti-haemostatic functions. *Int J Parasitol* 45, 613-627, doi:10.1016/j.ijpara.2015.03.009, (2015).
138. Kotál, J.; Stergiou, N.; Buša, M.; Chlastáková, A.; Beránková, Z.; Řezáčová, P.; Langhansová, H.; Schwarz, A.; Calvo, E.; Kopecký, J.; et al. The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin. *Cell Mol Life Sci* 76, 2003-2013, doi:10.1007/s00018-019-03034-3, (2019).
139. Gao, X.; Tian, Y.; Liu, Z.L.; Li, D.; Liu, J.J.; Yu, G.X.; Duan, D.Y.; Peng, T.; Cheng, T.Y.; Liu, L. Tick salivary protein Cystatin: structure, anti-inflammation and molecular mechanism. *Ticks Tick Borne Dis* 15, 102289, doi:10.1016/j.ttbdis.2023.102289, (2024).

140. Kotsyfakis, M.; Horka, H.; Salat, J.; Andersen, J.F. The crystal structures of two salivary cystatins from the tick Ixodes scapularis and the effect of these inhibitors on the establishment of Borrelia burgdorferi infection in a murine model. *Mol Microbiol* 77, 456-470, doi:10.1111/j.1365-2958.2010.07220.x, (2010).
141. Salát, J.; Paesen, G.C.; Rezácová, P.; Kotsyfakis, M.; Kovárová, Z.; Sanda, M.; Majtán, J.; Grunclová, L.; Horká, H.; Andersen, J.F.; et al. Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick Ornithodoros moubata. *Biochem J* 429, 103-112, doi:10.1042/bj20100280, (2010).
142. Martins, L.A.; Buša, M.; Chlastáková, A.; Kotál, J.; Beránková, Z.; Stergiou, N.; Jmel, M.A.; Schmitt, E.; Chmelař, J.; Mareš, M.; Kotsyfakis, M. Protease-bound structure of Ricistatin provides insights into the mechanism of action of tick salivary cystatins in the vertebrate host. *Cell Mol Life Sci* 80, 339, doi:10.1007/s00018-023-04993-4, (2023).
143. Dantas-Torres, F.; Chomel, B.B.; Otranto, D. Ticks and tick-borne diseases: a One Health perspective. *Trends Parasitol* 28, 437-446, doi:10.1016/j.pt.2012.07.003, (2012).
144. Angulo, F.J.; Zhang, P.; Halsby, K.; Kelly, P.; Pilz, A.; Madhava, H.; Moïsi, J.C.; Jodar, L. A systematic literature review of the effectiveness of tick-borne encephalitis vaccines in Europe. *Vaccine* 41, 6914-6921, doi:10.1016/j.vaccine.2023.10.014, (2023).
145. Kotál, J.; Buša, M.; Urbanová, V.; Řezáčová, P.; Chmelař, J.; Langhansová, H.; Sojka, D.; Mareš, M.; Kotsyfakis, M. Mialostatin, a Novel Midgut Cystatin from Ixodes ricinus Ticks: Crystal Structure and Regulation of Host Blood Digestion. *Int J Mol Sci* 22, doi:10.3390/ijms22105371, (2021).
146. Tušar, L.; Usenik, A.; Turk, B.; Turk, D. Mechanisms Applied by Protein Inhibitors to Inhibit Cysteine Proteases. *Int J Mol Sci* 22, doi:10.3390/ijms22030997, (2021).
147. Honey, K.; Rudensky, A.Y. Lysosomal cysteine proteases regulate antigen presentation. *Nat Rev Immunol* 3, 472-482, doi:10.1038/nri1110, (2003).
148. Sá-Nunes, A.; Bafica, A.; Antonelli, L.R.; Choi, E.Y.; Francischetti, I.M.; Andersen, J.F.; Shi, G.P.; Chavakis, T.; Ribeiro, J.M.; Kotsyfakis, M. The immunomodulatory action of sialostatin L on dendritic cells reveals its potential to interfere with autoimmunity. *J Immunol* 182, 7422-7429, doi:10.4049/jimmunol.0900075, (2009).
149. Chapman, H.A. Endosomal proteases in antigen presentation. *Curr Opin Immunol* 18, 78-84, doi:10.1016/j.coi.2005.11.011, (2006).
150. Adkison, A.M.; Raptis, S.Z.; Kelley, D.G.; Pham, C.T. Dipeptidyl peptidase I activates neutrophil-derived serine proteases and regulates the development of acute experimental arthritis. *J Clin Invest* 109, 363-371, doi:10.1172/jci13462, (2002).
151. Grunclová, L.; Horn, M.; Vancová, M.; Sojka, D.; Franta, Z.; Mares, M.; Kopácek, P. Two secreted cystatins of the soft tick Ornithodoros moubata: differential expression pattern and inhibitory specificity. *Biol Chem* 387, 1635-1644, doi:10.1515/bc.2006.204, (2006).
152. Zamolodchikova, T.S.; Tolpygo, S.M.; Svirshchevskaya, E.V. Cathepsin G-Not Only Inflammation: The Immune Protease Can Regulate Normal Physiological Processes. *Front Immunol* 11, 411, doi:10.3389/fimmu.2020.00411, (2020).
153. McDougall, J.J.; Zhang, C.; Cellars, L.; Joubert, E.; Dixon, C.M.; Vergnolle, N. Triggering of proteinase-activated receptor 4 leads to joint pain and inflammation in mice. *Arthritis Rheum* 60, 728-737, doi:10.1002/art.24300, (2009).
154. Zarbock, A.; Polanowska-Grabowska, R.K.; Ley, K. Platelet-neutrophil-interactions: linking hemostasis and inflammation. *Blood Rev* 21, 99-111, doi:10.1016/j.blre.2006.06.001, (2007).
155. Cohen, A.B.; Stevens, M.D.; Miller, E.J.; Atkinson, M.A.; Mullenbach, G. Generation of the neutrophil-activating peptide-2 by cathepsin G and cathepsin G-treated human platelets. *Am J Physiol* 263, L249-256, doi:10.1152/ajplung.1992.263.2.L249, (1992).
156. Cui, P.; Tani, K.; Kitamura, H.; Okumura, Y.; Yano, M.; Inui, D.; Tamaki, T.; Sone, S.; Kido, H. A novel bioactive 31-amino acid endothelin-1 is a potent chemotactic peptide for human neutrophils and monocytes. *J Leukoc Biol* 70, 306-312, (2001).

157. Dong, A.; Xu, X.; Edwards, A.M.; Chang, C.; Chruszcz, M.; Cuff, M.; Cymborowski, M.; Di Leo, R.; Egorova, O.; Evdokimova, E.; et al. In situ proteolysis for protein crystallization and structure determination. *Nat Methods* 4, 1019-1021, doi:10.1038/nmeth1118, (2007).
158. Chlastáková, A.; Kaščáková, B.; Kotál, J.; Langhansová, H.; Kotsyfakis, M.; Kutá Smatanová, I.; Tirloni, L.; Chmelař, J. Iripin-1, a new anti-inflammatory tick serpin, inhibits leukocyte recruitment in vivo while altering the levels of chemokines and adhesion molecules. *Front Immunol* 14, 1116324, doi:10.3389/fimmu.2023.1116324, (2023).
159. Chlastáková, A.; Kotál, J.; Beránková, Z.; Kaščáková, B.; Martins, L.A.; Langhansová, H.; Prudnikova, T.; Ederová, M.; Kutá Smatanová, I.; Kotsyfakis, M.; Chmelař, J. Iripin-3, a New Salivary Protein Isolated From Ixodes ricinus Ticks, Displays Immunomodulatory and Anti-Hemostatic Properties In Vitro. *Front Immunol* 12, 626200, doi:10.3389/fimmu.2021.626200, (2021).
160. Kascakova, B.; Kotal, J.; Martins, L.A.; Berankova, Z.; Langhansova, H.; Calvo, E.; Crossley, J.A.; Havlickova, P.; Dycka, F.; Prudnikova, T.; et al. Structural and biochemical characterization of the novel serpin Iripin-5 from Ixodes ricinus. *Acta Crystallogr D Struct Biol* 77, 1183-1196, doi:10.1107/s2059798321007920, (2021).
161. Kotál, J.; Polderdijk, S.G.I.; Langhansová, H.; Ederová, M.; Martins, L.A.; Beránková, Z.; Chlastáková, A.; Hajdušek, O.; Kotsyfakis, M.; Huntington, J.A.; Chmelař, J. Ixodes ricinus Salivary Serpin Iripin-8 Inhibits the Intrinsic Pathway of Coagulation and Complement. *Int J Mol Sci* 22, doi:10.3390/ijms22179480, (2021).
162. Bevec, T.; Stoka, V.; Pungercic, G.; Dolenc, I.; Turk, V. Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L. *J Exp Med* 183, 1331-1338, doi:10.1084/jem.183.4.1331, (1996).
163. Lenarcic, B.; Krishnan, G.; Borukhovich, R.; Ruck, B.; Turk, V.; Moczydlowski, E. Saxiphilin, a saxitoxin-binding protein with two thyroglobulin type 1 domains, is an inhibitor of papain-like cysteine proteinases. *J Biol Chem* 275, 15572-15577, doi:10.1074/jbc.M001406200, (2000).
164. Musil, D.; Zucic, D.; Turk, D.; Engh, R.A.; Mayr, I.; Huber, R.; Popovic, T.; Turk, V.; Towatari, T.; Katunuma, N.; et al. The refined 2.15 A X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *Embo j* 10, 2321-2330, doi:10.1002/j.1460-2075.1991.tb07771.x, (1991).
165. Guncar, G.; Podobnik, M.; Pungercar, J.; Strukelj, B.; Turk, V.; Turk, D. Crystal structure of porcine cathepsin H determined at 2.1 A resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. *Structure* 6, 51-61, doi:10.1016/s0969-2126(98)00007-0, (1998).
166. Li, Z.; Kienetz, M.; Cherney, M.M.; James, M.N.; Brömme, D. The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex. *J Mol Biol* 383, 78-91, doi:10.1016/j.jmb.2008.07.038, (2008).
167. Aguda, A.H.; Panwar, P.; Du, X.; Nguyen, N.T.; Brayer, G.D.; Brömme, D. Structural basis of collagen fiber degradation by cathepsin K. *Proc Natl Acad Sci U S A* 111, 17474-17479, doi:10.1073/pnas.1414126111, (2014).
168. Almeida, P.C.; Nantes, I.L.; Chagas, J.R.; Rizzi, C.C.; Faljoni-Alario, A.; Carmona, E.; Juliano, L.; Nader, H.B.; Tersariol, I.L. Cathepsin B activity regulation. Heparin-like glycosaminogylcans protect human cathepsin B from alkaline pH-induced inactivation. *J Biol Chem* 276, 944-951, doi:10.1074/jbc.M003820200, (2001).
169. Novinec, M.; Lenarčič, B.; Turk, B. Cysteine cathepsin activity regulation by glycosaminoglycans. *Biomed Res Int* 2014, 309718, doi:10.1155/2014/309718, (2014).
170. Chazeirat, T.; Denamur, S.; Bojarski, K.K.; Andrault, P.M.; Sizaret, D.; Zhang, F.; Saidi, A.; Tardieu, M.; Linhardt, R.J.; Labarthe, F.; et al. The abnormal accumulation of heparan

sulfate in patients with mucopolysaccharidosis prevents the elastolytic activity of cathepsin V. *Carbohydr Polym* 253, 117261, doi:10.1016/j.carbpol.2020.117261, (2021).
171. Jmel, M.A.; Aounallah, H.; Bensaoud, C.; Mekki, I.; Chmelař, J.; Faria, F.; M'Ghirbi, Y.; Kotsyfakis, M. Insights into the Role of Tick Salivary Protease Inhibitors during Ectoparasite-Host Crosstalk. *Int J Mol Sci* 22, doi:10.3390/ijms22020892, (2021).
172. Panwar, P.; Hedtke, T.; Heinz, A.; Andrault, P.M.; Hoehenwarter, W.; Granville, D.J.; Schmelzer, C.E.H.; Brömme, D. Expression of elastolytic cathepsins in human skin and their involvement in age-dependent elastin degradation. *Biochim Biophys Acta Gen Subj* 1864, 129544, doi:10.1016/j.bbagen.2020.129544, (2020).
173. Vizovišek, M.; Fonović, M.; Turk, B. Cysteine cathepsins in extracellular matrix remodeling: Extracellular matrix degradation and beyond. *Matrix Biol* 75-76, 141-159, doi:10.1016/j.matbio.2018.01.024, (2019).
174. Dennemärker, J.; Lohmüller, T.; Mayerle, J.; Tacke, M.; Lerch, M.M.; Coussens, L.M.; Peters, C.; Reinheckel, T. Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis. *Oncogene* 29, 1611-1621, doi:10.1038/onc.2009.466, (2010).
175. Senjor, E.; Kos, J.; Nanut, M.P. Cysteine Cathepsins as Therapeutic Targets in Immune Regulation and Immune Disorders. *Biomedicines* 11, doi:10.3390/biomedicines11020476, (2023).
176. Khamtorn, P.; Peigneur, S.; Amorim, F.G.; Quinton, L.; Tytgat, J.; Daduang, S. De Novo Transcriptome Analysis of the Venom of Latrodectus geometricus with the Discovery of an Insect-Selective Na Channel Modulator. *Molecules* 27, doi:10.3390/molecules27010047, (2021).
177. Emerich, B.L.; Ferreira, R.C.; Cordeiro, M.N.; Borges, M.H.; Pimenta, A.M.; Figueiredo, S.G.; Duarte, I.D.; de Lima, M.E. δ-Ctenitoxin-Pn1a, a Peptide from Phoneutria nigriventer Spider Venom, Shows Antinociceptive Effect Involving Opioid and Cannabinoid Systems, in Rats. *Toxins (Basel)* 8, 106, doi:10.3390/toxins8040106, (2016).