

Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Biochemie



Mgr. Michal Buša

Parazitární cystatiny jako inhibitory cysteinových proteas:
strukturní aspekty funkční specifity a evoluce

Parasite cystatins as inhibitors of cysteine proteases:
structural aspects of functional specificity and their evolution

Disertační práce

Školitel: RNDr. Michael Mareš, CSc.

Praha 2023

Prohlášení

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

V Praze dne

.....

Podpis

Poděkování

Na tomto místě bych rád poděkoval všem, kteří mi pomáhali a podporovali mě během celého doktorského studia.

V první řadě děkuji školiteli Michaelu Marešovi za vedení disertační práce a věnovaný čas, ochotu a trpělivost, a Pavlíně Řezáčové za uvedení do světa strukturní biologie. Velký dík poté patří všem současným i bývalým kolegům a kolegyním za jejich vstřícnost, rady, podporu a vytvoření toho nejlepšího pracovního prostředí, které kdy měl tu čest zažít. Mnohé z nich mohu nyní nazývat svými přáteli, především jsem vděčen za vás Adriane, Petře a Kubo.

To největší poděkování ale patří Anežce, Medě, Heleně, Tereze a Heleně za všechnu vaši lásku a podporu.

Děkuji

Obsah

Prohlášení.....	2
Poděkování.....	3
Seznam zkratk.....	6
Abstrakt (ČJ).....	7
Abstract (EN).....	8
1. Teoretický úvod.....	9
1.1. Cysteinové proteasy.....	9
1.1.1. Struktura cysteinových katepsinů.....	9
1.1.2. Role cysteinových katepsinů v biologii člověka.....	13
1.1.3. Parazitární cysteinové katepsiny.....	15
1.1.4. Role cysteinových katepsinů v biologii parazitů.....	16
1.2. Regulace cysteinových katepsinů.....	17
1.3. Přirozené inhibitory cysteinových katepsinů.....	19
1.3.1. Proteinové inhibitory cysteinových katepsinů.....	20
1.3.2. Mechanismy inhibice cysteinových katepsinů proteinovými inhibitory.....	20
1.3.3. Přehled hlavních rodin inhibitorů cysteinových katepsinů.....	21
1.3.4. Rodina cystatinů.....	23
1.3.4.1. Stefiny.....	23
1.3.4.2. Právě cystatiny.....	24
1.3.4.3. Multidoménné cystatiny.....	24
1.3.4.4. Struktura a inhibiční mechanismus cystatinů.....	24
1.4. Klíště obecné (<i>Ixodes ricinus</i>).....	27
1.4.1. Cysteinové katepsiny <i>I. ricinus</i>	29
1.4.2. Proteasové inhibitory klíšťat.....	30
1.4.3. Rodina cystatinů u klíšťat.....	31
1.5. Motolice jaterní (<i>Fasciola hepatica</i>).....	32
1.5.1. Cysteinové katepsiny u <i>F. hepatica</i>	34
1.5.2. Proteasové inhibitory u <i>F. hepatica</i>	35
1.5.3. Rodina cystatinů u <i>F. hepatica</i>	36
2. Cíle práce.....	38
3. Laboratorní vybavení a metody.....	39
3.1. Materiál a laboratorní vybavení.....	39
3.2. Metodika.....	39
3.2.1. Purifikace a příprava komplexu ricistatinu a katepsinu V.....	39
3.2.2. Kinetické inhibiční testy.....	39
3.2.3. Krystalizace komplexu ricistatinu a katepsinu V.....	40
3.2.4. Řešení struktury komplexu ricistatinu a katepsinu V.....	40
4. Výsledky.....	42
4.1. Publikace č. 1: An evolutionary molecular adaptation of an unusual stefin from the liver fluke <i>Fasciola hepatica</i> redefines the cystatin superfamily.....	43

4.2.	Publikace č. 2: Mialostatin, a novel midgut cystatin from <i>Ixodes ricinus</i> ticks: crystal structure and regulation of host blood digestion.	88
4.3.	Publikace č. 3: The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin.	117
4.4.	Nepublikované výsledky.....	138
4.4.1.	Inhibiční specifita ricistatinu a její srovnání s ostatními cystatiny	138
4.4.2.	Krystalová struktura komplexu ricistatinu s katepsinem V a analýza vazebného módu	139
5.	Diskuse.....	148
5.1.	Identifikace a charakterizace nového typu cystatinů u motolice <i>F. hepatica</i>	148
5.2.	Strukturní vysvětlení inhibiční specifity cystatinů z klíštěte <i>I. ricinus</i>	150
6.	Závěry	154
6.1.	Prohlášení spoluautorů.....	156
7.	Seznam použité literatury	157

Seznam zkratek

Å	Ångström, 10 ⁻¹⁰ m
AdAPR-1	aspartátová proteasa 1 z měchovce lidského (<i>Ancylostoma duodenale</i>)
AdCP2	cysteinová proteasa 2 z měchovce lidského (<i>Ancylostoma duodenale</i>)
AMC	aminometylkumarin
BPTI	„bovine pancreatic trypsin inhibitor“, inhibitor hovězího trypsinu ze slinivky
CEW	„Chicken Egg White“,
CPA	cysteinová proteasa A z ničivky útrobní (<i>Leishmania donovani</i>)
CPB	cysteinová proteasa B z ničivky útrobní (<i>Leishmania donovani</i>)
CEW	„chicken egg white“, cystatin z kura domácího (<i>Gallus gallus</i>)
CTLA	„cytotoxic T-lymphocyte associated protein“ protein asociovaný s cytotoxickými T-lymfocyty
CyLS	„cystatin-like stefin“, cystatinu podobný stefin
DPAP1	dipeptidyl aminopeptidasa 1 ze zimničky tropické (<i>Plasmodium falciparum</i>)
DTT	dithiotreitol
EDTA	kyselinu ethylendiamintetraoctová
ESP	exkrečně sekreční produkty
FgStf1	stefin 1 z <i>Fasciola gigantica</i>
FhAE	legumain (asparagyl- endopeptidasa) z motolice jaterní (<i>Fasciola hepatica</i>)
FhCB1	katepsin B1 z motolice jaterní (<i>Fasciola hepatica</i>)
FhCL1-5	katepsiny L1-5 z motolice jaterní (<i>Fasciola hepatica</i>)
FhKT1.3	inhibitor z rodiny Kunitz-BPTI z motolice jaterní (<i>Fasciola hepatica</i>)
FhStf1-3	stefiny 1-3 z motolice jaterní (<i>Fasciola hepatica</i>)
FhSrp1-4	serpiny 1-4 z motolice jaterní (<i>Fasciola hepatica</i>)
FPLC	„fast protein liquid chromatography“, rychlá proteinová kapalinová chromatografie
FRET	Försterův rezonanční přenos energie
Hyp	hydroxyprolin
IC ₅₀	koncentrace inhibitoru potřebná k dosažení 50% inhibice aktivity enzymu
IrAE	legumain (asparagyl- endopeptidasa) z klíštěte obecného (<i>Ixodes ricinus</i>)
IrCB	katepsin B z klíštěte obecného (<i>Ixodes ricinus</i>)
IrCC	katepsin C z klíštěte obecného (<i>Ixodes ricinus</i>)
IrCL1	katepsin L1 z klíštěte obecného (<i>Ixodes ricinus</i>)
IrCL3	katepsin L3 z klíštěte obecného (<i>Ixodes ricinus</i>)
IrCD	katepsin D z klíštěte obecného (<i>Ixodes ricinus</i>)
kDa	kilodalton
NaAPR-1	aspartátová proteasa 1 z měchovce amerického (<i>Necator americanus</i>)
NaAPR-2	aspartátová proteasa 2 z měchovce amerického (<i>Necator americanus</i>)
NaCP3	cysteinová proteasa 3 z měchovce amerického (<i>Necator americanus</i>)
NEJ	nově excystování juvenilové
NK	„natural killer“, přirozené zabíječské buňky
PEG	polyetylen glykol
PDB	„protein data bank“, databáze proteinových struktur
RMSD	„root mean square deviation“, střední kvadratická odchylka
SmCB1	katepsin B z krevničky střešní (<i>Schistoma mansoni</i>)
SmCC	katepsin C z krevničky střešní (<i>Schistoma mansoni</i>)
SmCD	katepsin D z krevničky střešní (<i>Schistoma mansoni</i>)
SmCL1	katepsin L1 z krevničky střešní (<i>Schistoma mansoni</i>)
TbCB	katepsin B z trypanosomy spavičné (<i>Trypanosoma brucei</i>)
TLR	„Toll-like receptors“, receptory podobné genu Toll

Abstrakt (ČJ)

Proteiny z rodiny cystatinů jsou důležitými inhibitory cysteinových proteas typu katepsinů a účastní se řady patologií. Parazitární cystatiny jsou atraktivní cílové molekuly pro regulaci parazitů, ale je o nich dosud málo informací. Tato práce se zaměřuje na cystatiny dvou významných krevsajících parazitů: klíštěte obecného (*Ixodes ricinus*) jako hlavního vektoru lymfské boreliózy a klíšťové encefalitidy, a motolice jaterní (*Fasciola hepatica*), původce fasciolózy. Byla provedena funkční a rentgenostrukturní charakterizace čtyř nových cystatinů, která umožnila určit strukturní podstatu inhibiční specifity a vysvětlit je v kontextu evoluce cystatinů a jejich fyziologické role.

Cystatin FhCyLS-2 z *F. hepatica* má širokou inhibiční specifitu a hraje pravděpodobně duální roli při regulaci proteolytických systémů v tkáni hostitele a ve střevu parazita. FhCyLS-2 kombinuje unikátním způsobem charakteristiky dvou podrodin cystatinů a představuje modelového zástupce nové evoluční skupiny cystatinů identifikované u několika řádů parazitických motolic.

Ricistatin a iristatin jsou cystatiny ze slin *I. ricinus*, jejichž imunomodulační působení na hostitele je způsobeno výjimečně úzkou inhibiční specifitou. Ta byla vysvětlena strukturními modifikacemi ve vazebných segmentech inhibitorů. Modifikace mají jiný charakter u ricistatinu a iristatinu, jak bylo prokázáno s využitím krystalové struktury komplexu s proteasou, a odlišují dvě fylogenetické skupiny slinných cystatinů klíšťat. Mialostatin ze střeva *I. ricinus* má nemodifikované vazebné segmenty zajišťující širokou inhibiční specifitu, která umožňuje efektivně regulovat multienzymový trávicí proteolytický systém klíštěte.

Disertační práce přináší nové informace o strukturně-funkčních vztazích u parazitárních cystatinů a jejich molekulární evoluci ke specializovaným fyziologickým funkcím. Studované cystatiny představují potenciální antigeny pro vývoj antiparazitárních vakcín a mají farmakologicky zajímavé účinky s možným využitím v biomedicíně.

Abstract (EN)

Members of the cystatin family are important inhibitors of cathepsin-type cysteine proteases and are involved in a number of pathologies. Parasite cystatins are attractive target molecules for parasite control, but our knowledge about them is still limited. This work is focused on cystatins of two blood-feeding parasites: the common tick (*Ixodes ricinus*) as the main vector of Lyme disease and tick-borne encephalitis, and the liver fluke (*Fasciola hepatica*), the causative agent of fasciolosis. Four novel cystatins were functionally and structurally characterized to define the structural determinants of their inhibitory specificity and to describe them in the context of evolution and physiological role of cystatins.

The cystatin FhCyLS-2 from *F. hepatica* has broad inhibitory specificity and is suggested to play a dual role in the regulation of proteolytic systems in host tissue and the parasite gut. FhCyLS-2 combines the characteristics of two cystatin subfamilies in a unique way and is a model representative of a novel evolutionary group of cystatins identified in several orders of parasitic flukes.

Ricistatin and iristatin are salivary cystatins of *I. ricinus* with immunomodulatory effects on the host caused by an exceptionally narrow inhibitory specificity. This specificity was explained by structural modifications of the binding segments of the inhibitors. The modifications have a different pattern for ricistatin and iristatin, as was demonstrated by the crystal structure of the cystatin-protease complex and are the principal feature distinguishing two phylogenetic groups of tick salivary cystatins. Mialostatin from *I. ricinus* gut contains unmodified binding segments, providing a broad inhibitory specificity that enables effective regulation of the multienzyme proteolytic digestive system of the ticks.

The thesis provides new information on structure-function relationships in parasite cystatins and their molecular evolution towards various physiological roles. The studied cystatins represent promising antigens for the development of antiparasitic vaccines and have pharmacologically interesting properties for potential biomedical applications.

1. Teoretický úvod

1.1. Cysteinové proteasy

Proteasy, označované i jako proteiny nebo peptidasy, jsou enzymy hydrolyzující peptidovou vazbu proteinů a peptidů. Jsou přítomny u všech živých organismů a mají nezastupitelnou úlohu v celé řadě fyziologických dějů. Lidský genom obsahuje více než 600 genů kódujících různé proteasy, což představuje cca 3 % všech genů ¹.

Proteasy lze klasifikovat několika způsoby. Dle pozice štěpené peptidové vazby se dělí na endopeptidasy, štěpící uvnitř polypeptidového řetězce, a exopeptidasy štěpící peptidové vazby z C-konce (karboxypeptidasy) nebo N-konce (aminopeptidasy) substrátu. MEROPS databáze zavádí členění na základě jejich 3D architektury do 53 různých klanů a dle sekvenční příbuznosti do 266 rodin ². Hlavní rozdělení proteas lze provést dle stavby katalytického centra a mechanismu hydrolyzy štěpené peptidové vazby. Tímto způsobem se proteasy dělí do sedmi různých tříd – na serinové, cysteinové, aspartátové, glutamátové, threoninové, asparaginové proteasy a metaloproteasy ².

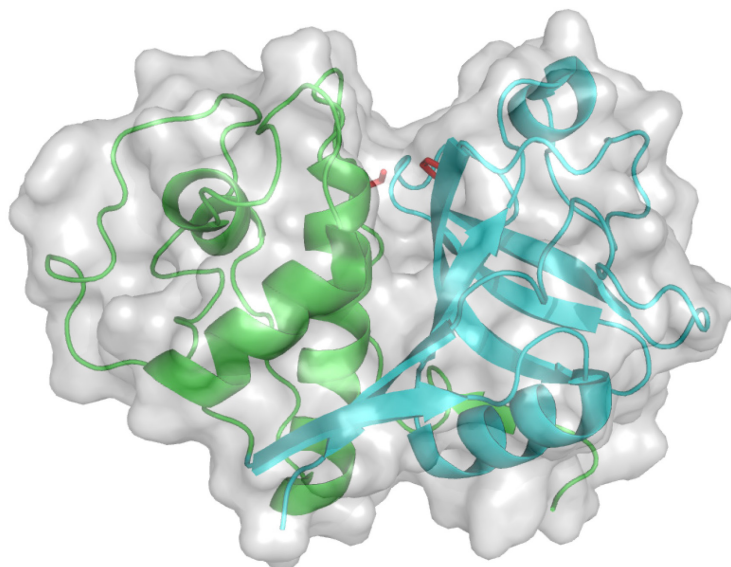
Cysteinové proteasy jsou třetí nejrozšířenější rodinou proteas po serinových proteasách a metaloproteasách. Jsou tvořeny devíti klany, které jsou složeny z celkem 67 rodin. Z nich, nejrozsáhlejší a nejvíce studovanou je rodina C1 nesoucí název podle první identifikované proteasy – papainu ². Dalšími významnými rodinami cysteinových proteas jsou kalpainy (C2), legumainy (C13) a kaspasy (C14).

Do rodiny papainu patří i skupina proteas, původně popsanych jako lysosomální, známých jako cysteinové katepsiny. U člověka je nyní termínem katepsin označováno celkem 15 proteas. Tato heterogenní skupina je tvořena 11 cysteinovými proteasami (katepsiny B, C, L, K, S, F, H, O, V, X, W z rodiny C1), 2 serinovými proteasami (katepsin A z rodiny S10 a katepsin G z rodiny S1) a 2 aspartátovými proteasami (katepsiny D a E z rodiny A1) ³. Jednotlivé lidské katepsiny slouží díky svým odlišným vlastnostem jako modelové proteasy pro klasifikaci homologních proteas identifikovaných u ostatních organismů ².

1.1.1. Struktura cysteinových katepsinů

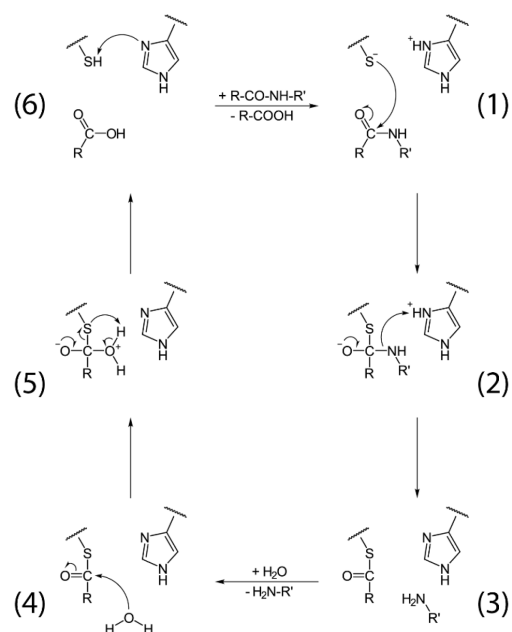
Funkční molekulou cysteinových katepsinů jsou monomery o velikosti 20–35 kDa. Výjimkou je katepsin C, jehož funkční jednotkou je homotetramer o celkové molekulové hmotnosti cca 200 kDa ⁴. Jako první byla v roce 1984 vyřešena prostorová struktura rostlinného papainu z papáji obecné (*Carica papaya*), která nyní slouží jako strukturní archetyp cysteinových katepsinů (Obr. 1). Papain je monomerní, globulární dvoudoménný protein

s molekulovou hmotností 24 kDa⁵. Domény jsou označovány jako levá a pravá podle polohy při konvenčním zobrazení⁶. Zatímco levá doména je tvořena především α -šroubovicemi, pravá obsahuje především β -struktury. Aktivní místo se nachází v prohlubni mezi doménami a na jeho dně se nalézá katalytická dyáda tvořená zbytky Cys25 a His159.



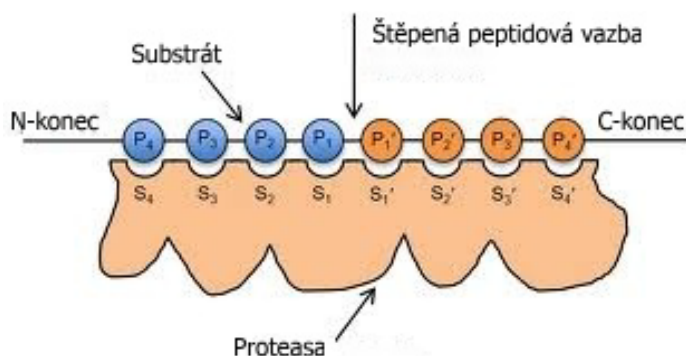
Obr. 1: Molekula papainu v konvenční orientaci. Katalytická doména papainu (znázorněna jako šedý povrch) se skládá ze dvou domén odlišných barevně ve stužkovém modelu. Levá (L) doména je označená zeleně a obsahuje pouze α -šroubovice, pravá (P) doména označená azurově je tvořena především β -strukturami. Každá doména nese jeden zbytek katalytické dyády zbytků Cys a His (červeně). Převzato z⁶ a upraveno.

Mechanismus štěpení peptidového substrátu (Obr. 2) je shodný pro všechny cysteinové katepsiny a jsou pro něj kritické především dva aminokyselinové zbytky tvořící konzervovanou katalytickou dyádu – Cys25 a His159 (číslováno dle zralého papainu)⁷. Po navázání substrátu do aktivního místa proteasy thiolová skupina cysteinu, která v deprotonované formě slouží jako nukleofil, atakuje uhlík karbonylové skupiny štěpené peptidové vazby. Dojde k vytvoření tetrahedrálního intermediátu, který je stabilizovaný postranním řetězcem Gln19. V dalším kroku je rozštěpena peptidová vazba, uvolní se první produkt z C-konce substrátu a vzniká kovalentní komplex acyl-proteasa. Dále deprotonovaná molekula vody nukleofilně atakuje karbonylový uhlík komplexu acyl-proteasa za vzniku dalšího tetrahedrálního intermediátu. Posledním krokem je deacylace proteasy, uvolnění druhého produktu a regenerace proteasy⁷.



Obr. 2: Katalytický mechanismus cysteinových katepsinů. Štěpení substrátu se odehrává v šesti krocích: 1) nukleofilní atak deprotonované thiolové skupiny Cys25 na karbonylový uhlík peptidové vazby, 2) vznik thioesterového intermediátu a následná deprotonace imidazolového kruhu His159, 3) vznik komplexu acyl-proteasy a uvolnění prvního produktu (amin), 4) nukleofilní atak molekuly vody na thioesterově vázaný uhlík, 5) vznik druhého thioesterového intermediátu a 6) uvolnění druhého produktu (karboxylové kyseliny) a regenerace aktivního místa enzymu. Převzato z ⁸ a upraveno.

Štěpený peptidový substrát se váže do aktivního místa a vyplňuje bočními řetězci aminokyselinových zbytků jednotlivá vazebná podmísta proteasy (Obr. 3). U papainu bylo definováno 5 podmíst konvenčně označovaných jako S3, S2, S1, S1' a S2'. Příslušným aminokyselinovým zbytkům substrátu potom odpovídající jednotlivé pozice označované jako P3, P2, P1, P1' a P2'.

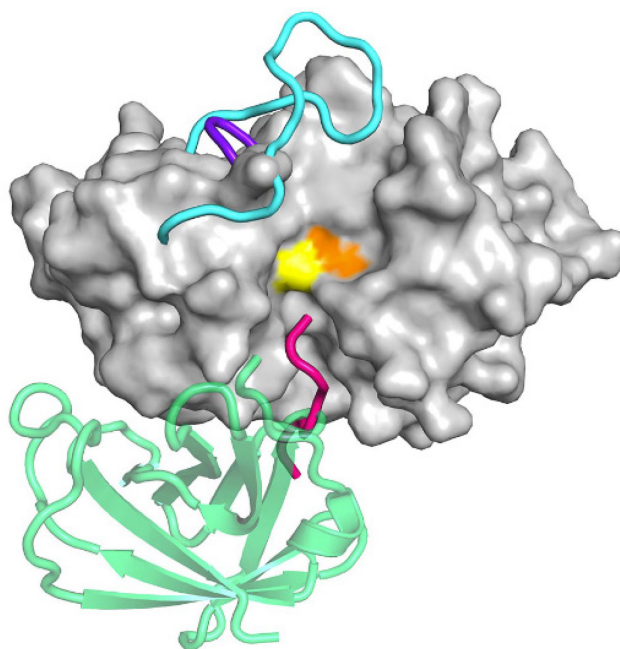


Obr. 3: Schematické znázornění vazby peptidového substrátu do aktivního místa proteasy. Zobrazeny jsou pozice jednotlivých aminokyselinových zbytků peptidového substrátu (P4–P4'), které se vážou do odpovídajících podmíst enzymu (S4–S4'). Štěpená peptidová vazba mezi zbytky P1 a P1' je znázorněna červeně. Převzato z ⁹ a upraveno.

Substrátová specifita papainu je primárně určena podmísty S1 a S2, kdy podmísto S1 preferuje velké bazické aminokyseliny (Arg, Lys) a podmísto S2 váže všechny hydrofobní aminokyseliny (především Val). Lidské cysteinové katepsiny vykazují specifitu podobnou papainu, ovšem některé z nich mají odlišné preference pro podmísto S2. Katepsiny L a V preferují aromatické aminokyseliny před hydrofobními, katepsiny F, K a S preferují alifatické aminokyseliny a katepsin B nemá vyhraněnou specifitu pro podmísto S2 ¹⁰.

Substrátová specifita je stěžejní pro biologickou funkci dané proteasy. Např. katepsin K je využíván osteoklasty pro degradaci kostní matrix, která je složená především z kolagenu typu I s typickou sekvencí Gly-Xxx-Xxx, kde Xxx může být jakákoliv aminokyselina, ale nejčastěji se jedná o Pro. Katepsin K preferuje v podmístě S2 Pro a dokáže tak konzervovanou sekvenci kolagenu vázat a efektivně degradovat ¹¹.

Cysteinové katepsiny jsou převážně endopeptidasami, pouze několik z nich, katepsiny B, C, H a X, vykazují exopeptidasovou aktivitu, která je způsobena implementací specifických strukturních segmentů: smyček, zbytků aktivačních peptidů či celých domén (Obr. 4). U exopeptidas je přístup substrátu do S nebo S' podmíst aktivního místa těmito segmenty omezen. To umožní vazbu pouze N- nebo C-konce substrátu a ke štěpení tedy dochází pouze v aminopeptidasovém resp. karboxypeptidasovém módu ¹². Konkrétně katepsin H působí jako aminopeptidasa ¹³, katepsin C jako aminodipeptidasa (dipeptidylpeptidasa) ⁴, katepsin X jako karboxypeptidasa ¹⁴ a katepsin B jako karboxydipeptidasa (peptidyldipeptidasa) ¹⁵.



Obr. 4: Struktura cysteinových katepsinů s exopeptidasovou aktivitou. Přídavné strukturální segmenty zodpovědné za exopeptidasovou specifitu modifikují základní endopeptidasovou architekturu. Klíčové strukturální segmenty exopeptidas jsou superponovány přes povrch endopeptidasy katepsinu L (šedě) s vyznačenou katalytickou dyádou (cystein – žlutě, histidin – oranžově). Segmenty jsou zobrazeny stužkovým modelem pro aminopeptidasu katepsin H („mini-řetězec“ – červeně), aminodipeptidasu katepsin C („exclusion domain“ – zeleně), karboxypeptidasu katepsin X („mini-smyčka“ – fialově) a karboxydipeptidasu katepsin B („occluding loop“ – azurově). Převzato z ¹⁶.

1.1.2. Role cysteinových katepsinů v biologii člověka

Jednou z hlavních úloh katepsinů v tkáních je štěpení buněčných proteinů v lysosomech. Do lysosomů je nespecificky směřován veškerý materiál pohlcený buňkou, membránové proteiny nebo fagocytované patogeny ^{17,18}. Zároveň jsou do lysosomů specificky transportovány cytosolární proteiny pomocí receptorů a transportérů ¹⁹. Koncentrace jednotlivých katepsinů může být v lysosomech až 1 mM ²⁰, zároveň je zde udržováno kyselé pH (rozmezí 4,5-5,0) odpovídající pH optimu těchto proteas.

Funkce katepsinů nicméně nejsou limitovány pouze na nespecifickou degradaci proteinů, ale hrají důležitou roli i při vysoce specifických procesech. Příkladem je aktivace zymogenů proteas (granzymů, neutrofilní elastasy, chymasy či tryptasy) v případě katepsinu C, H a L ^{21,22}, štěpení invariantního řetězce MHC II receptorů v endosomech antigen-prezentujících buněk ²³,

role signálních molekul díky aktivaci Mas-příbuzných GPCR receptorů katepsinem S²⁴ nebo procesování Toll-like receptorů (TLR3, TLR7 a TLR9) katepsinem B a H, které je podmínkou jejich správné funkce²⁵. Většina lysosomálních cysteinových katepsinů je exprimována všemi druhy buněk, ale několik katepsinů vykazuje tkáňově specifickou expresi. Jedná se zejména o katepsin K (osteoklasty), katepsin S (antigen-prezentující buňky), katepsin V (brzlík, rohovka a varlata) a katepsin W (NK buňky, T-lymfocyty)¹². Tkáňově specifický výskyt jednotlivých katepsinů je společně s nejběžnějšími patologiemi shrnut v Tab. 1.

Katepsiny byly původně popsány jako lysosomální proteiny, ale v posledních letech se ukázalo, že katepsiny mají přirozené funkce i mimo lysosom, např. v cytosolu při regulaci apoptosy a zánětu²⁶ nebo v buněčném jádru při regulaci buněčného cyklu²⁷. Katepsiny B a L jsou sekretovány do mimobuněčného prostoru, kde degradují proteiny extracelulární matrix, odstraňují membránové receptory a adhezní molekuly²⁸ či specificky aktivují cytokiny a růstové faktory^{29,30}. Dalším katepsinem cíleně sekretovaným do mezibuněčného prostoru je katepsin K, který je využíván osteoklasty při odbourávání kostní tkáně³¹.

Tab. 1: Přehled lidských cysteinových katepsinů, jejich funkcí a patologií. Převzato z³².

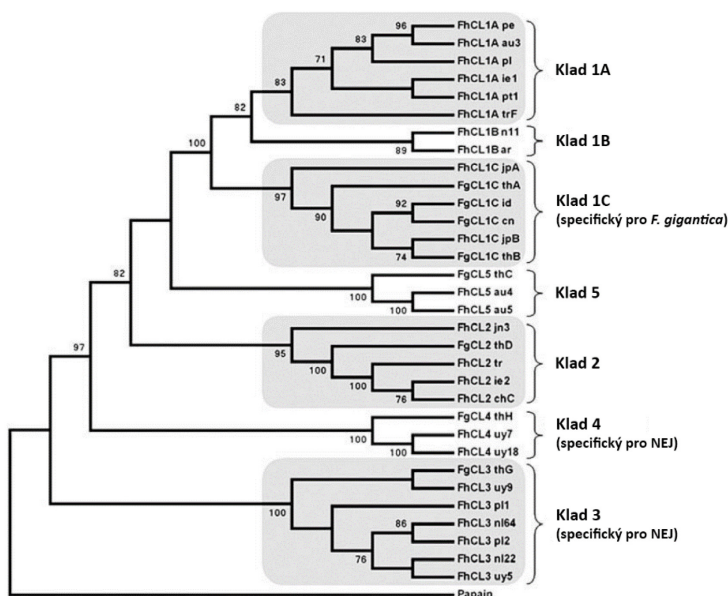
Katepsin	Proteolytická aktivita	Výskyt	Patologie
Katepsin B	endopeptidasa/ karboxydipeptidasa	nespecifický	rakovina, revmatoidní artritida, pankreatitida
Katepsin C	aminodipeptidasa	nespecifický	Papillon-Leferův syndrom, Haim-Munkův syndrom
Katepsin F	endopeptidasa	nespecifický	rakovina, neuronální lipofucinóza
Katepsin H	endopeptidasa/ aminodipeptidasa	nespecifický	rakovina, ateroskleróza
Katepsin K	endopeptidasa	osteoklast	osteoporóza, revmatoidní artritida, osteoartróza, pyknodysostóza
Katepsin L	endopeptidasa	nespecifický	rakovina, ateroskleróza, epidermální hyperplazie, kardiomyopatie, zánětlivá onemocnění, obezita
Katepsin O	endopeptidasa	nespecifický	rakovina
Katepsin S	endopeptidasa	antigen prezentující buňky	bronchiální astma, artritida, ateroskleróza, lupénka, rakovina, neurodegenerativní onemocnění
Katepsin V	endopeptidasa	brzlík, varlata, rohovka	autoimunitní onemocnění, myathenia gravis
Katepsin W	endopeptidasa	T-lymfocyty, NK buňky	není známo
Katepsin X	karboxypeptidasa	nespecifický	rakovina, zánětlivá onemocnění

Vychýlení přesně vybalancované aktivity katepsinů z rovnovážného stavu je často příčinou nebo průvodním jevem různých patologických stavů (Tab. 1). Například ztráta aktivity katepsinu C dochází ke vzniku Papillon-Lefèvreova a Haim-Munkova syndromu³³.

Zvýšená aktivita katepsinu K způsobuje osteoporózu ³⁴, a naopak ztráta jeho aktivity vede k pyknodysostóze ³⁵. Zvýšená aktivita katepsinu S v séru je spojována s diabetem 2. typu ³⁶. Dále cysteinové katepsiny hrají roli v zánětlivých onemocněních plic ³⁷, Alzheimerově chorobě ³⁸, ateroskleróze ³⁹, tvorbě aneurysmat ⁴⁰, svalové dystrofii ⁴¹, revmatoidní artritidě ⁴² a osteoartróze ⁴³. Často bývá aktivita cysteinových katepsinů zvýšena při nádorovém onemocnění, ta je potom spojována s horší prognózou, jelikož usnadňuje angiogenezi a metastazování rakovinných buněk ⁴⁴⁻⁴⁶. Několik katepsinů bylo navrženo jako prognostické markery u nádorového onemocnění krve, prsu, střev, plic, jater, prostaty nebo slinivky ^{3,47-49}.

1.1.3. Parazitární cysteinové katepsiny

Parazitární ortology katepsinů jsou označovány na základě jejich homologie k archetypálním lidským katepsinům. U parazitů prodělaly katepsiny na rozdíl od člověka bouřlivý vývoj a díky duplikaci archetypálních genů nyní parazité disponují celými genovými rodinami katepsinů. Nejvíce ortologů bylo popsáno pro katepsiny L, B, C a F. Jako příklad lze uvést katepsiny L z motolice jaterní (*F. hepatica*), která nyní disponuje celkem 26 isoenzymy katepsinu L (Obr. 5) ⁵⁰.



Obr. 5: Fylogenetická analýza genové rodiny katepsinu L u *F. hepatica* a *F. gigantea*. Z původního jednoho genu pro katepsin L u *F. hepatica* duplikacemi a následnou divergencí vznikla genová rodina nyní zahrnující 7 různých kladů s celkem 32 isoenzymy. Pro zakořenění stromu byl použit papain z *Carica papaya*. Převzato z ⁵¹ a upraveno.

Lidské katepsiny a jejich parazitární ortology jsou strukturně velmi blízké, a to navzdory poměrně nízké sekvenční identitě⁵². Velký počet aminokyselinových záměn tedy neovlivnil charakteristickou strukturu známou u archetypálních lidských katepsinů, ale byly získány jiné funkční charakteristiky, jako jsou např. širší pH optimum, vyšší stabilita či odlišná substrátová specifita⁵³. Například lidský katepsin L vykazuje pH optimum mezi 4,5 a 5,5 a při pH 7,0 je inaktivován v řádu minut. Jeho ortolog IrCL1 z klíštěte obecného (*Ixodes ricinus*) má pH optimum posunuté do velmi kyselé oblasti 3,0–4,0⁵⁴ a může tak efektivně degradovat hemoglobin, který v pH pod 3,5 denaturuje⁵⁵. Oproti tomu katepsiny L z trypanosomy spavičné (*Trypanosoma brucei*) a *F. hepatica* jsou stabilní a aktivní při velkém rozpětí pH od 3,0 do 9,0. Díky tomu mohou kombinovat funkci trávicích proteas a současně modulovat hostitelské prostředí^{56,57}.

1.1.4. Role cysteinových katepsinů v biologii parazitů

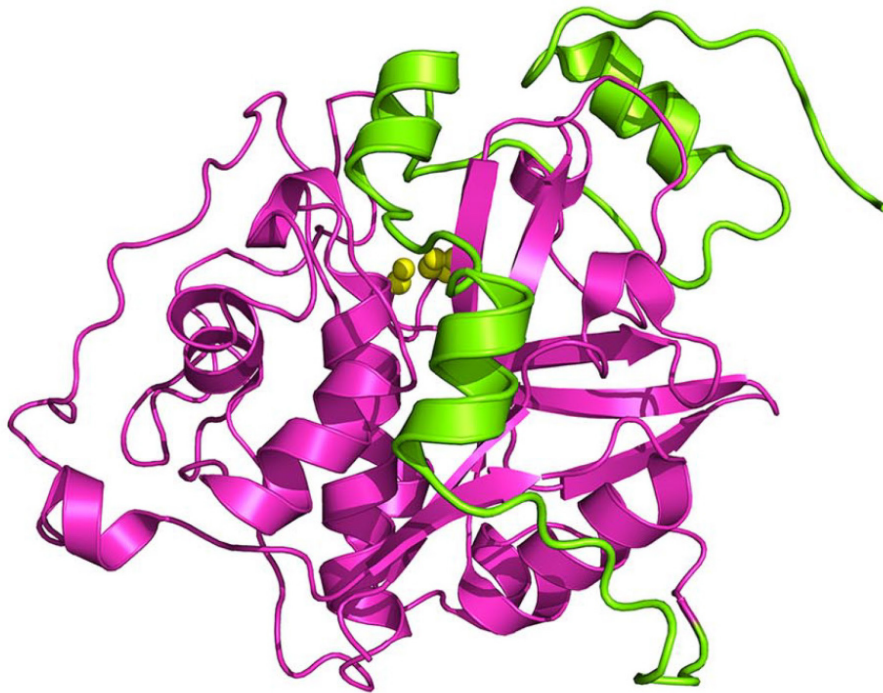
U parazitů mají cysteinové katepsiny specifické role, které souvisejí s jejich životní strategií. U všech parazitů, kromě hmyzu a korýšů, plní cysteinové katepsiny nenahraditelnou roli v trávení krevních proteinů hostitele. K degradaci hemoglobinu parazité využívají kaskádu proteolytických enzymů založenou na cysteinových katepsinech (rodina C1), legumainu (rodina C13) a aspartátové protease katepsinu D (rodina A1)^{53,58,59}. Tato proteolytická kaskáda je překvapivě konzervovaná napříč různými parazity (Tab. 2) a je společná mnoha helmintům⁶⁰⁻⁶², prvokům⁶³⁻⁶⁵ i členovcům⁶⁶ a bude ji detailně věnována pozornost v kapitolách 1.7.1 a 1.8.1. Některé parazitární cysteinové katepsiny jsou sekretované do hostitelského prostředí, kde jsou díky vyšší stabilitě nadále aktivní a mohou ovlivňovat různé molekulární procesy v tkáních hostitele, včetně imunitní odpovědi⁵⁹. Endoparazité jsou parazité žijící dlouhodobě v těle svého hostitele. Ke vstupu do hostitele a k migraci do cílového orgánu využívají specifické cysteinové katepsiny. Ty jsou schopné štěpit např. proteolyticky velmi odolnou strukturu kolagenu a umožňují parazitům pronikat přes extracelulární matrix a tkáňové bariéry^{67,68}.

Tab. 2: Přehled významných parazitů s proteolytickým trávicím systémem založeným na cysteinových a aspartátových katepsinech. Ekonomicky významná onemocnění jsou označena (a). Převzato z ¹⁶.

Parazit	Klasifikace	Trávicí proteasy		Onemocnění (milióny nakažených osob)
		cysteinové proteasy rodiny papainu	aspartátové proteasy rodiny pepsinu	
Trypanosoma americká a spavičná (<i>Trypanosoma cruzi</i> , <i>T. brucei gambiense/rhodesiense</i>)	protisté (prvoci)	cruzain, rhodesain, TbCB		Chagasova choroba, spavá nemoc ^a (10)
Ničivka útrobní (<i>Leishmania donovani</i>)	protisté (prvoci)	CPA, CPB		leishmanióza (1,5)
Zimnička tropická (<i>Plasmodium falciparum</i>)	protisté (prvoci)	falcipain 2, 3, 2', DPAP1	plasmepsin I, II, IV	malárie (350)
Krevnička střevní (<i>Schistosoma mansoni</i>)	helminți (ploštěnci)	SmCB1, SmCC, SmCL	SmCD	schistosomóza (bilharzióza) (240)
Motolice jaterní (<i>Fasciola hepatica</i>)	helminți (ploštěnci)	FhCB, FhCL		fasciolóza ^a (5)
Měchovec americký a lidský (<i>Necator americanus</i> , <i>Ancylostoma duodenale</i>)	helminți (hlístice)	necpain, NaCP3, AdCP2	NaAPR-1, 2, AdAPR-1	ankylostomóza (700)
Klíště obecné (<i>Ixodes ricinus</i>)	členovci (pavoukovci)	IrCB1, IrCL1, IrCC	IrCD1	vektor lymeské borreliózy a klíšťové encefalidity

1.2. Regulace cysteinových katepsinů

Hydrolyza peptidové vazby katalyzovaná proteasami je děj obecně nevratný, a proto musí být činnost i cysteinových katepsinů přísně regulována, a to pomocí několika různých mechanismů. Patří sem regulace genové exprese, alternativního sestřihu mRNA a kompartmentace buňky ⁶⁹. S kompartmentací souvisí i další mechanismus regulace pomocí pH. Lidské katepsiny mají pH optimum v mírně kyselé oblasti a v neutrálním pH jsou, s výjimkou katepsinu S, velmi nestabilní. V cytosolu tedy podléhají rychlé denaturaci, čímž je potlačena jejich aktivita ^{3,70,71}.



Obr. 6: Struktura zymogenu katepsinu B1 z krevničky střešní (*Schistosoma mansoni*). Stužkový model propeptidu (zeleně) blokuje aktivní místo katepsinu mezi L a R doménami (fialově) a interaguje s katalytickými aminokyselinovými zbytky (žlutě). Převzato z ⁷².

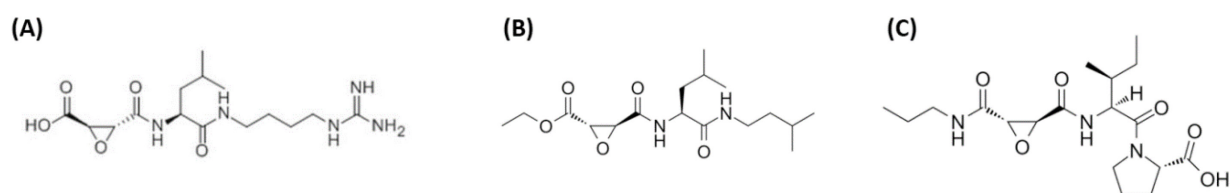
Dalším mechanismem typickým pro cysteinové katepsiny je syntéza ve formě neaktivního zymogenu. U zymogenu je aktivní místo katepsinu vyplněno propeptidem, který sekvenčně předchází katalytické doméně katepsinu (Obr. 6). Řetězec propeptidu je stabilizován proti štěpení, protože prochází aktivním místem v opačném směru než substrát ⁷³. Aktivace zymogenu může probíhat buď autokatalyticky, kdy v kyselém prostředí dochází k disociaci propeptidu z aktivního místa zymogenu a jeho fragmentaci ostatními molekulami daného katepsinu, nebo trans-aktivací pomocí jiné proteasy, typicky jiným cysteinovým nebo aspartátovým katepsinem nebo legumainem ^{74,75}. Autoaktivační procesy některých katepsinů mohou být modulovány záporně nabitými polysacharidy, např. glykosaminoglykany ^{72,76}. Posledním mechanismem využívaným pro regulaci cysteinových katepsinů jsou proteinové inhibitory proteas, kterými se tato disertační práce zabývá a bude jim věnována následující kapitola.

1.3. Přírodní inhibitory cysteinových katepsinů

Přírozně se vyskytující proteasové inhibitory mohou být látky buď proteinové, nebo nízkomolekulární. Proteinové inhibitory jsou všeobecně přítomné u všech organismů. Jejich rolí je především regulace vlastních proteas či ochrana tkání před proteolytickým poškozením a jsou detailněji popsány v další kapitole. Nízkomolekulární inhibitory jsou syntetizovány především bakteriemi, houbami a rostlinami jako ochranné molekuly proti patogenům.

Mezi první popsané přírodní nízkomolekulární inhibitory cysteinových katepsinů patří leupeptin ⁷⁷, chymostatin ⁷⁸ a antipain ⁷⁹ identifikované v aktinobakteriích. Tyto peptidyl-aldehydové inhibitory se kovalentně vážou na katalytický zbytek katepsinu. Jejich peptidová část nezaručuje dostatečnou specifitu a tak jsou silnými inhibitory jak cysteinových, tak serinových proteas ⁸⁰.

Prvním identifikovaným specifickým inhibitorem cysteinových katepsinů je epoxid E-64 (systematickým názvem N-[N-(L-3-trans-karboxyoxiran-2-karbonyl)leucyl]-agmatin) z houby *Aspergillus japonicus* ⁸¹. Struktura E-64 je uvedena na Obr. 7, agmatinová skupina vyplňuje podmísto S3, zbytek leucinu hydrofobní podmísto S2 a reaktivní epoxidová skupina se nachází v blízkosti katalytického cysteinu ⁶. Reakce probíhá nukleofilním atakem katalytického cysteinu na epoxidový C2 uhlík E-64. Díky kombinaci vysoké reaktivity a dostatečné specifity je E-64 používán k důkazu aktivity cysteinových proteas. Struktura E-64 byla dále modifikována a bylo vyvinuto mnoho derivátů se specializovanými vlastnostmi (Obr. 7), např. E-64d schopný procházet přes biologické membrány ⁸² nebo CA-074, který je specifickým inhibitorem katepsinu B ⁸³.



Obr. 7: Přírodní inhibitor cysteinových proteas E-64 (A) a jeho syntetické deriváty E-64d (Aloxistatinu, B) a CA-074 (C). Převzato z ⁸⁴ a upraveno.

V mořských houbách, cyanobakteriích a rostlinách bylo popsáno mnoho dalších inhibitorů katepsinů, z nichž nejvýznamnější jsou miraziridin A ⁸⁵, tokaramid A ⁸⁶, gallinamid A ⁸⁷,

grassypeptolidy A, B a C ⁸⁸, panduratin A a nikolaioidesin C ⁸⁹. Mezi zajímavé skupiny bioaktivních malých molekul patří triterpeny, flavonoidy a alkaloidy, které stericky snadno vyplňují aktivní místo cysteinových proteas, a jejich někteří zástupci fungují jako slabé, kompetitivní inhibitory cysteinových katepsinů ⁸⁰.

Přírodní inhibitory a substráty cysteinových katepsinů přinášejí nové strukturální vzory využitelné k vývoji silnějších či specifitějších inhibitorů a léčiv a je jim v současné době věnována velká pozornost. Inhibitory odvozené z peptidové stavby se nazývají peptidomimetika a obsahují typicky část zodpovědnou za vazbu do aktivního místa proteasy a reaktivní skupinu (tzv. warhead), která umožňuje kovalentní vazbu s katalytickým cysteinem ⁹⁰. Tyto inhibitory se dají rozdělit dle povahy kovalentní vazby vytvořené s proteasou na ireverzibilní inhibitory (halomethylketony, diazomethany, epoxysukcinylové deriváty, acyloxymethylketony a vinyl sulfony) a reverzibilní inhibitory (peptidyl aldehydy, nitrily, methylketony) ^{91,92}. Vyvinuté prototypové látky mohou být následně modifikovány tak, aby splnily farmakologická kritéria a mohly se stát registrovanými léčivy, jako např. nirmatrelvir účinný proti hlavní protease viru SARS-CoV-2 ⁹³. Potentní a specifické inhibitory lze dále využít i jako součást tzv. aktivních sond („activity based probes“), sloužících k detekci či vizualizaci cysteinových proteas ^{94,95}.

1.3.1. Proteinové inhibitory cysteinových katepsinů

Databáze Merops eviduje celkem 84 proteinových rodin inhibitorů proteolytických enzymů, z nichž 8 rodin je specializováno na inhibici cysteinových katepsinů. Nejvýznamnější a nejpočetnější je rodina cystatinů (I25), dále to jsou např. rodina CTLA (I29), thyropiny (I31), chagasiny (I42), cliticypiny (I48) nebo makrocypiny (I85) ². Některé rodiny inhibitorů serinových proteas obsahují proteiny schopné inhibovat i cysteinové proteasy. Jako příklad mohou sloužit zástupci rodiny živočišných (I2) i rostlinných (I3) Kunitzových inhibitorů nebo serpiny (I4) ⁹⁶⁻⁹⁸.

1.3.2. Mechanismy inhibice cysteinových katepsinů proteinovými inhibitory

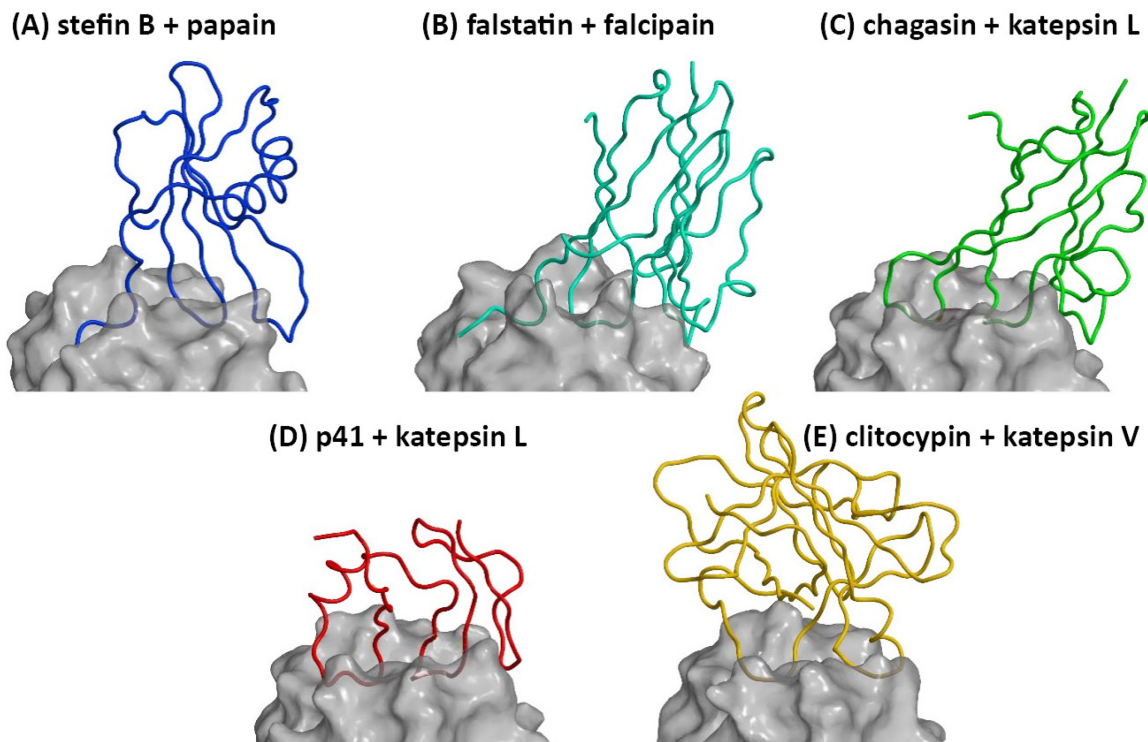
Existují dva základní principy, kterými proteinové inhibitory interagují s cysteinovými katepsiny a blokují tak jejich aktivitu. Prvním je poškození architektury aktivního místa při kovalentní interakci inhibitoru s katalytickým zbytkem. Tento princip je poměrně vzácný a je využíván pouze serpiny (I4) a inhibitory kaspas z rodiny p35 (I50) ^{99,100}. Druhým principem je sterické vyplnění aktivního místa proteasy zprostředkované pomocí

nekovalentních interakcí. Díky vysoké afinitě reaktivních segmentů reverzibilních inhibitorů k aktivnímu místu proteas jsou tyto proteiny často nazývány také jako „tight-binding“ inhibitory ¹⁰¹. Značný interakční povrch umožňuje vytvořit celou síť nekovalentních interakcí, především vodíkových můstků a hydrofobních interakcí, a díky tomu mohou dosahovat nekovalentní inhibitory jak vysoké afinity, tak selektivity díky rozlišování i malých strukturních změn u příbuzných katepsinů. Inhibitory mohou využívat i další interakční regiony na povrchu proteasy mimo její aktivní místo, tzv. exomísta, a to pro zvýšení specifity jako např. serpiny ¹⁰² nebo pro zvýšení afinity jako např. hirudiny ¹⁰³.

1.3.3. Přehled hlavních rodin inhibitorů cysteinových katepsinů

Rodina CTLA (kód I29 podle Merops databáze) je tvořena propeptidy cysteinových katepsinů a inhibitory, které vznikly genovou duplikací segmentu propeptidů z původních genů pro zymogeny cysteinových katepsinů. Jako součást zymogenů cysteinových katepsinů jsou široce rozšířené, ovšem jako samostatné inhibitory jsou vzácné a jejich funkce nebyla popsána ^{104,105}

Thyropiny (rodina I31) byly nejprve identifikovány jako součást multidoménných proteinů thyroglobulinů. Jsou stabilizované sadou disulfidických můstků a do aktivního místa katepsinů se váží pomocí trojice interakčních smyček, podobně jako cystatiny (Obr. 8). Thyropiny jsou značně selektivní v rámci cílových cysteinových katepsinů, např. protein p41 (fragment invariantního řetězce MHC II) je pikomolárním inhibitorem katepsinu L a H, ale katepsiny S, B a C neinhibuje vůbec ¹⁰⁶. Bylo prokázáno, že má stabilizační vliv na katepsin L v lysosomech i v extracelulární matrix ^{107,108}, a účastní se regulace antigen prezentujících buněk ¹⁰⁹. Thyropinová doména je často přítomna ve velkých multidoménných proteinech, jako jsou např. testikany a thyroglobuliny. Equistatin, protein ze sasanky mořské (*Actinia equina*), obsahuje tři thyropinové domény – doména 1 inhibuje cysteinové proteasy, doména 2 aspartátové proteasy a doména 3 nejspíše inhibiční funkci ztratila ¹¹⁰.



Obr. 8: Struktury komplexů různých rodin proteinových inhibitorů s cysteinovými proteasami z rodiny papainu. Všechny komplexy zobrazují stejný boční pohled na komplex proteasy (šedý povrch) a inhibitoru (barevný stužkový model bez sekundárních struktur). (A) Komplex stefinu B z rodiny I25 a papainu (modře), (B) komplex falstatinu z rodiny I71 a falcipainu (azurově), (C) komplex chagasinu z rodiny I42 a katepsinu L (světle zeleně), (D) komplex fragmentu p41 z rodiny I31 a katepsinu L (červeně) a (E) komplex clitocypinu z rodiny I48 a katepsinu V (žlutě). Převzato z ¹⁰¹ a upraveno.

Regulace cysteinových katepsinů pomocí proteinových inhibitorů je i u parazitů nesmírně důležitá. V parazitických prvocích se vyvinuly tři vzájemně příbuzné rodiny s nízkou sekvenční homologií (<25%), které sdílí strukturální vzor imunoglobulinů ¹¹¹. Jedná se o chagasiny (rodina I42), falstatiny (I71) a toxostatiny (I81), jejichž způsob interakce s aktivním místem proteasy pomocí několika segmentů je podobný jako u cystatinů (Obr. 8) ^{112,113}. Modeloví zástupci jednotlivých rodin, chagasin z trypanosomy americké (*Trypanosoma cruzi*), falstatin ze zimničky tropické (*Plasmodium falciparum*) a toxostatin z kokcidie kočičí (*Toxoplasma gondii*) jsou pikomolární inhibitory hlavních sekretovaných katepsinů daných parazitů a hrají roli při průniku parazita do hostitelských buněk, přeměně vývojových stádií parazita a regulaci imunitního systému hostitele ¹¹¹.

Makrocypiny (rodina I85) a clitocypiny (I48) jsou dvě malé rodiny s pouze několika popsány inhibitory z vyšších hub sdílející obecnou strukturní architekturu s rodinou rostlinných Kunitz inhibitorů (I3). Na rozdíl od Kunitzovy rodiny neobsahují žádné disulfidické můstky. Některé inhibitory z těchto rodin jsou schopny regulovat kromě cysteinových katepsinů (C1) i legumain (C13) či serinové proteasy z rodiny chymotrypsinu S1^{114,115}. Jejich obecnou funkcí je nejspíše ochrana proti patogenům a škůdcům¹¹⁴.

1.3.4. Rodina cystatinů

Cystatiny jsou největší rodinou inhibitorů (kód I25 podle Merops databáze) cysteinových proteas a jsou přítomné u všech živých organismů s výjimkou hub a archeí. Zaměřené jsou především na cysteinové katepsiny rodiny papainu (rodina C1), některé jsou schopny inhibovat i proteasy z rodiny legumainu (C13)^{116,117}. Dle primární sekvence je lze rozdělit do tří podrodin – stefiny, pravé cystatiny a multidoménné cystatiny². Stefiny (také označované jako cystatiny 1. typu) jsou jednodoménné proteiny bez signální sekvence a disulfidických můstků s molekulovou hmotností cca 11,5 kDa. Pravé cystatiny (cystatiny 2. typu) jsou také jednodoménné proteiny s molekulovou hmotností cca 13 kDa, ale disponují signální sekvencí a dvěma páry konzervovaných disulfidických můstků. Multidoménné cystatiny (cystatiny 3. typu) jsou tvořené několika doménami pravých cystatinů či jejich kombinací s jinými proteinovými doménami¹¹⁸.

1.3.4.1. Stefiny

Stefiny jsou díky absenci signální sekvence intracelulárními proteiny, které se vyskytují v cytosolu, mitochondriích a buněčném jádru. V lidském organismu jsou dva homology – stefin A a B. Stefín A má velmi omezenou tkáňovou distribuci a je exprimován pouze buňkami epitelu, jater a lymfatickou tkání. Tato selektivní exprese odpovídá tkáním zapojeným v první linii boje proti patogenům, nicméně jeho funkce není stále přesně známá¹¹⁹. Oproti tomu stefin B je exprimovaný prakticky všemi buněčnými typy^{118,120}. Primární funkcí obou stefinů je ochrana před nežádoucí proteolýzou způsobenou cysteinovými katepsiny mimo jejich základní lysosomální kompartment¹²¹. Pro stefin B byla popsána i role v regulaci buněčného cyklu, díky regulaci aktivity katepsinu L v jádru buňky¹²⁰.

1.3.4.2. Právě cystatiny

Právě cystatiny jsou především sekretované proteiny, ale někteří zástupci mají i intracelulární funkce. V lidském organismu bylo popsáno 7 různých homologů cystatinů, z nichž nejvýznamnějším je cystatin C, který reguluje aktivitu katepsinů v hlavních tělních tekutinách¹²². Je produkován všemi jadernými buňkami a používá se jako časný biomarker akutního poškození ledvin¹²³ či kardiovaskulárních onemocnění¹²⁴. Jednobodová mutace cystatinu C způsobuje jeho oligomerizaci a následnou amyloidovou angiopatii islandského typu¹²⁵. Role cystatinu C v nádorových onemocněních není jednoznačně definována, v některých typech rakovinného bujení má cystatin C roli supresoru, v dalších naopak dělení podporuje a zdá se, že interaguje i s jinými proteiny než cysteinové katepsiny a legumain¹²⁶. Homologní cystatiny jsou součástí dalších dějů v lidském organismu, např. cystatiny D, S, SA a SN jsou exprimovány především slinnými žlázami a chrání nejspíše před exogenními katepsiny v ústní dutině¹²⁷, ale mají i jiné funkce jako například u cystatinu D, který inhibuje růst nádorů¹²⁸ nebo prezentaci antigenů APC¹²⁹. Cystatin E/M je exprimován podkožními buňkami a hraje roli v homeostázi kůže a v mnoha typech nádorových onemocnění¹³⁰. Cystatin F je produkován především buňkami imunitního systému a ovlivňuje cytotoxicitu NK buněk a T-lymfocytů skrze inhibici katepsinů C, H a L¹³¹.

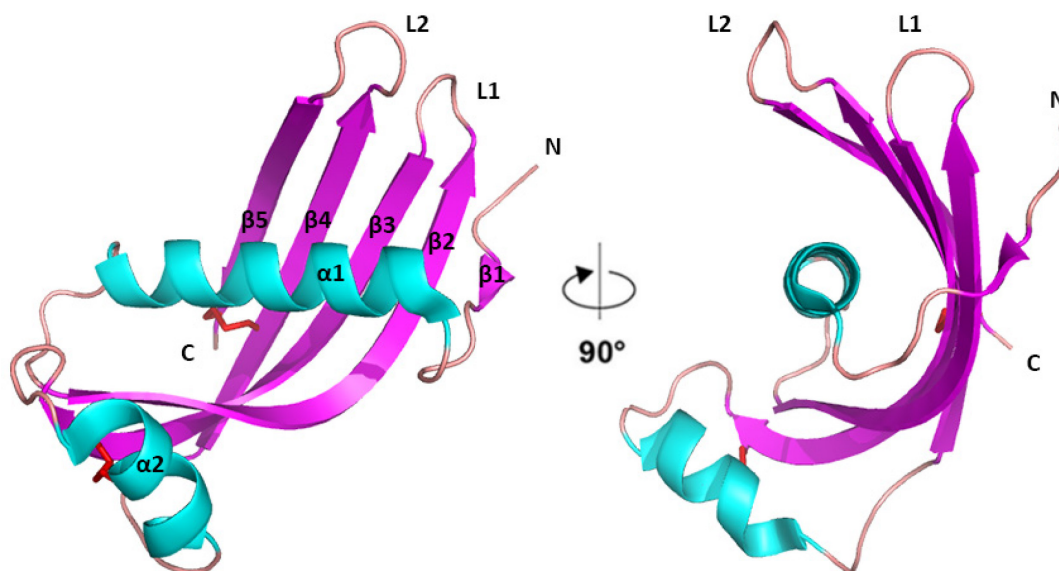
1.3.4.3. Multidoménové cystatiny

Skupina multidoménových cystatinů není monofyletická, ale sdružuje vzájemně nepříbuzné proteiny vývojově pokročilejších eukaryot tvořené dvěma (fetuiny, „histidin-rich“ proteiny), třemi (kininogeny) a až 12 (multidoménové cystatiny parazitů) doménami pravých cystatinů^{118,132,133}. Cystatinové domény těchto inhibitorů často během evoluce ztrácejí svou schopnost inhibovat cysteinové katepsiny a získávají funkce jiné, tak jako u fetuinů či „histidin-rich“ proteinů¹³⁴. Inhibiční funkce je naopak zachovaná u kininogenů, které ale kromě toho slouží i jako prekurzory hormonů, a i u multidoménových cystatinů parazitů s pravděpodobnou funkcí při gametogenezi^{133,135}

1.3.4.4. Struktura a inhibiční mechanismus cystatinů

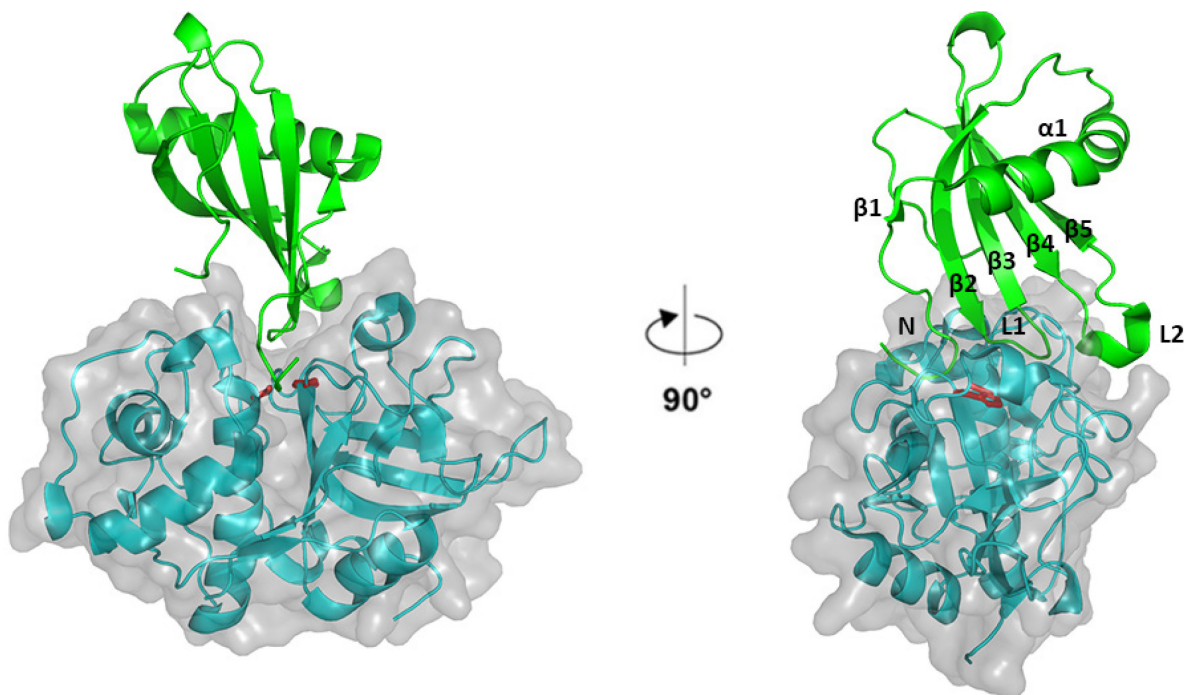
Krystalová struktura CEW („Chicken Egg White“) cystatinu byla vyřešena již v roce 1988 a je tvořena centrální α -šroubovicí, kolem které se obtáčí 5 skládaných antiparalelních β -struktur (Obr. 9). Tyto β -struktury jsou číslovány postupně od N- k C-konci a nejkratší z nich β 1 je tvořena buď pouze několika aminokyselinami, nebo není vytvořená vůbec. Obecná

architektura steffinů a cystatinů je analogická, nicméně cystatiny obsahují inserci cca 20 aminokyselin mezi $\beta 3$ a $\beta 4$, která je na opačné straně molekuly než reaktivní centrum a pár konzervovaných disulfidických můstků^{136,137}.



Obr. 9: Krystalová struktura CEW cystatinu. CEW cystatin je zobrazen ve stužkovém modelu s vyznačenou α -šroubovicí (azurová) a β -strukturami (fialová, číslované dle pořadí od N-konce), segmenty bez sekundární struktury jsou zobrazeny růžově. N- a C-konec molekuly je označen. Tři segmenty v horní části molekuly (N a smyčky L1 a L2) tvoří reaktivní centrum inhibitoru pro interakci s cysteinovými katepsiny. Konzervované disulfidové můstky jsou zobrazeny tyčkovým modelem (červeně). Převzato z¹³⁶ a upraveno.

Prostorová struktura komplexu steffinu B s papainem (Obr. 10) ukázala, že cystatiny interagují s cysteinovými katepsiny třemi interakčními segmenty, N-koncem proteinu a povrchovými smyčkami L1 a L2, které jsou lokalizovány mezi sekundárními strukturami $\beta 2$ a $\beta 3$, respektive $\beta 4$ a $\beta 5$ ¹³⁷. Tyto segmenty vstupují do aktivního místa proteasy, kde vyplňují vazebná podmiesta S3-S2', a obsahují konzervované sekvence důležité pro afinitu k cysteinovým katepsinům.



Obr 10: Strukturní analýza interakce cystatinů s proteasami rodiny papainu. V inhibičním komplexu je archetypální proteasa papain znázorněna stužkovým modelem (azurově) s šedým transparentním povrchem a lidský stefin B zeleným stužkovým modelem. Katalytické zbytky papainu jsou zobrazeny pomocí tyčkového modelu (červeně), uvedeny jsou interakční segmenty (N-konec a smyčky L1 a L2) a sekundární struktury stefinu B. Upraveno ¹³⁷.

V inhibičním segmentu N-konce jsou konzervované dva glycinové zbytky, které zajišťují flexibilitu tohoto segmentu nutnou k otočení hlavního řetězce cystatinu mezi podmísty S2 a S1. Zároveň umožní obsazení podmísta S3 a S2 bočními řetězci aminokyselinových zbytků předcházející Gly-Gly páru ¹³⁷.

Smyčka L1 obsahuje konzervovanou sekvenci Gln-Xxx-Val-Yyy-Gly, kde Xxx zbytek je nejčastěji Ile, Leu či Val a Yyy pak Ala nebo Asn. Gln z této sekvence vyplňuje podmísto S1' a Gly opět zajišťuje flexibilitu této smyčky. Konzervovaný Val je držen v nestandardní kombinaci torzních úhlů tak, aby ideálně vyplnil kavitu aktivního místa katepsinu mezi podmísty S1' a S2' ^{137,138}.

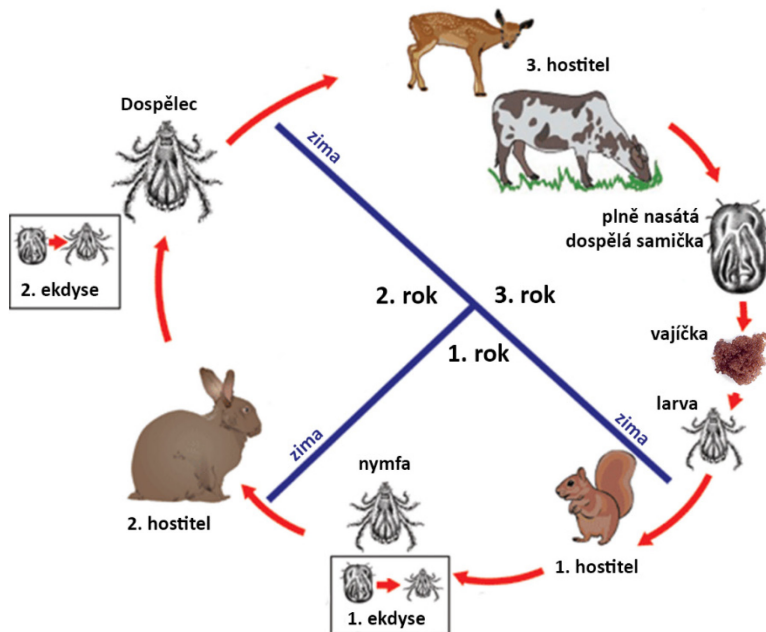
Interakční smyčka L2 má konzervovanou sekvenci Pro-Trp u cystatinů a Leu-Pro u stefinů ^{137,138}. Zbytek Pro zajišťuje smyčce L2 rigiditu, zatímco hydrofobní řetězce Leu nebo Trp vyplňují širokou hydrofobní kapsu podmíst S2' a S3'.

Příspěvek smyčky L2 k celkové vazebné energii komplexu je ve srovnání s ostatními segmenty malý, neboť přispívá pouze hydrofobními interakcemi. Většina (až 90 %) vazebné energie pochází z interakcí N-konce a smyčky L1; poměr jejich příspěvků se liší podle typu navázané proteasy a cystatinu. Vazba cystatinů do aktivního místa proteasy je pravděpodobně zahájena interakcí bočních řetězců N-konce do podmíst S2 a S3. Poté smyčka L1 vyplní aktivní místo v okolí podmísta S1' a jako poslední probíhá hydrofobní asociace smyčky L2 s podmísty S2' a S3' ¹³⁹.

Cystatiny často tvoří nekovalentně vázané dimery, které neinhibují cysteinové proteasy, jelikož k dimerizaci dochází skrze inhibiční smyčku L1. Dimery vznikají vzájemnou interakcí částečně rozvolněných cystatinů a jednotlivé monomery si vyměňují segmenty $\alpha 1$, $\beta 1$ a $\beta 2$ ¹⁴⁰. Vazebné místo pro inhibici legumainu na začátku $\alpha 1$ není porušeno a dimerní cystatiny inhibující legumain si tuto schopnost zachovávají ¹⁴¹. Cystatin F jako jediný vytváří dimer jiného typu, spojený kovalentně pomocí disulfidu. Tento inhibičně neaktivní dimer je v dendritických buňkách převeden na efektivní inhibitor pomocí aminopeptidasové aktivity katepsinu H, který odštěpí 15 zbytků z N-konce cystatinu F ¹⁴². Dynamická rovnováha vzniku dimerů je ovlivněna vrozenými mutacemi ^{125,143} a dimery mohou dále oligomerizovat a tvořit např. patologické amyloidní plaky ^{125,144}.

1.4. Klíště obecné (*Ixodes ricinus*)

Klíšťata jsou ektoparazitičtí roztoči živící se krví svých hostitelů. Lze je rozdělit na tvrdá a měkká klíšťata, která se odlišují anatomicky i životní strategií. Nejrozšířenější klíště obecné (*Ixodes ricinus*) patří mezi tvrdá klíšťata a vyskytuje se po celé Evropě. Jeho vývoj, který trvá 2 až 6 let, probíhá s proměnou nedokonalou (vajíčko, larva, nymfa, dospělec) a typicky zahrnuje tři různé savčí hostitele (Obr. 11).



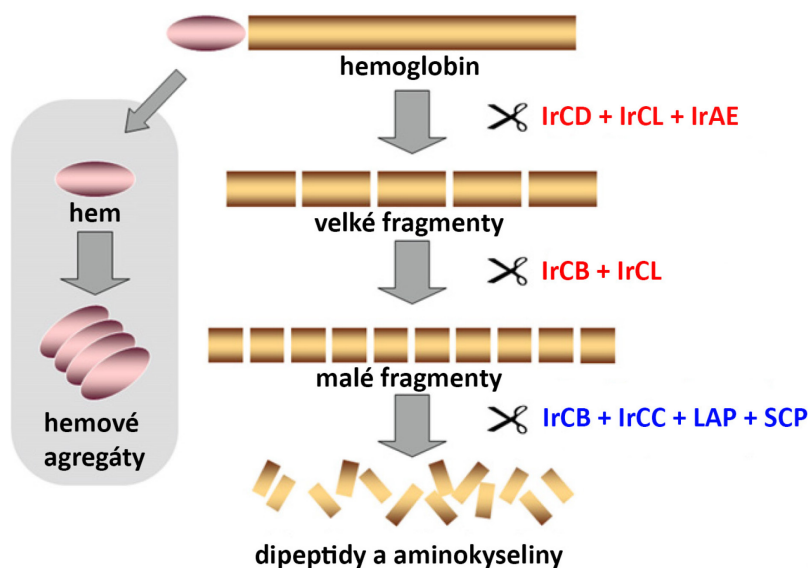
Obr. 11: Životní cyklus klíštěte *I. ricinus*. Z nakladených vajíček se vylíhne šestinohá larva, která přečká zimu v úkrytu a následně vyhledá hostitele (malé obratlovce), nasaje se a přemění se v osminohou nymfu. Ta parazituje na větších obratlovcích a po nasátí se mění v dospělce. Dospělci se páří na třetím hostiteli, nasátá samička pak odpadá a po naklazení snůšky umírá. Převzato z ¹⁴⁵ a upraveno.

Z hlediska medicíny není klíště významné jako parazit, ale především jako vektor mnoha závažných onemocnění, klíšťové encefalitidy a lymbské borreliózy ¹⁴⁶. Dále může přenášet i mnoho dalších onemocnění např. babeziózu, anaplazmózu, tularémii a rikettsiózu ¹⁴⁷.

V každé fázi životního cyklu klíštěte (instar) dochází k sání krve pouze jednou, a to nepřetržitě po několik dní (u dospělé samičky až 9 dní). Po odpadnutí klíštěte z hostitele dochází k přeměně do dalšího instaru či produkci vajíček a smrti ¹⁴⁸. Při samotném přísátí klíštěte dochází nejprve k proříznutí kůže, zanoření chobotku (hypostom) do hostitele a fixace v ráně pomocí cementu. Následuje fáze pomalého sání, během které klíště během 6 dní postupně zvyšuje příjem krve, aktivuje trávicí kaskádu a až 10x znásobí svůj objem. Takto získanou energii využije pro tvorbu dostatečně velké kutikuly pro závěrečnou periodu rychlého sání (6-8 dní). Během ní klíště opět 10x zvětší svůj objem a celkově nasaje až 1 ml krve. Poté klíště odpadne z hostitele a v případě dospělé samičky jsou veškeré požitelné krevní proteiny využity k produkci vajíček ¹⁴⁹.

1.4.1. Cysteinové katepsiny *I. ricinus*

Studium fyziologické funkce katepsinů v klíšťatech se soustředilo především na proteolytické trávení krevních proteinů. Po nasátí krve jsou hlavní proteiny hemoglobin a sérový albumin z lumen střeva transportovány do lysosomů střevních buněk klíštěte různým způsobem – hemoglobin aktivně, sérový albumin pasivně¹⁵⁰. Zde jsou oba proteiny degradovány pomocí podobného proteolytického systému^{66,151}. V kyselém pH lysosomů hemoglobin spontánně denaturuje a uvolňuje toxickou molekulu hemu, která je uložena ve specializovaných organelách – hemosomech¹⁵². Štěpení hemoglobinu bylo poprvé detailně prozkoumáno u druhu *I. ricinus*: začíná pomocí endopeptidas – cysteinového katepsinu L1 (IrCL1), aspartátového katepsinu D (IrCD1) a legumainu (asparaginylní endopeptidasa, IrAE) za vzniku velkých peptidových fragmentů o velikosti 5-7 kDa. Ty jsou následně degradovány především endopeptidasovou aktivitou katepsinu B (IrCB1) na krátké peptidy a dále štěpeny pomocí exopeptidas katepsinu C (IrCC) a IrCB1 na vstřebatelné dipeptidy a jednotlivé aminokyseliny⁶⁶. Základní proteolytická kaskáda je schematicky znázorněna na Obr. 12. V posledních letech se objevují práce naznačující, že proces trávení se po odpadnutí z hostitele přesouvá do lumenu střeva^{153,154}. Proteolytická aktivita je postupně indukována během procesu sání krve. Zejména od 5. dne dochází k rychlému nárůstu aktivity až do maxima, kterého je dosaženo 8. den sání⁶⁶. U příbuzného druhu klíštěte *Rhipicephalus microplus* byla prokázána i role klíštěcího katepsinu L při inhibici srážení krve působením thrombinu, což je důležité pro skladování krve v lumen střeva klíštěte¹⁵⁵.



Obr. 12: Schéma proteolytického trávicího systému klíšťať založeného na cysteinových a aspartátových proteasách. V lysozomech střevní tkáně je hemoglobin postupně degradován proteasami s endopeptidasovou aktivitou (červeně) a následně proteasami s exopeptidasovou aktivitou (modře) na vstřebatelné aminokyseliny a dipeptidy. Uvolněný toxický hem je odstraněn tvorbou agregátů. IrCB, IrCC, IrCD, IrCL označují jednotlivé katepsiny z *I. ricinus* a IrAE (asparaginylní endopeptidasa) legumain. LAP a SCP označují leucinovou aminopeptidasu a serinovou karboxypeptidasu. Převzato z ⁶⁶ a upraveno.

1.4.2. Proteasové inhibitory klíšťať

Proteasové inhibitory klíšťať jsou zkoumány především kvůli jejich interakci s proteasami hostitele a největší pozornost je tedy věnována inhibitorům produkovaným slinnými žlázami klíšťať a injikovaným do tkáně hostitele. U klíšťať bylo identifikováno široké spektrum proteasových inhibitorů z celkem 18 různých rodin. Nejvíce identifikovaných zástupců (asi 75 %) jsou inhibitory serinových proteas, zejména z rodin Kunitz-BPTI (rodina I2) a serpinů (I4). Druhou nejpočetnější skupinou (cca 10 %) jsou zástupci rodiny cystatinů (I25) jako inhibitory cysteinových proteas. Dále jsou ve slinách klíšťete přítomny inhibitory metaloproteas a α 2-makroglobuliny ¹⁵⁶. Nejlépe jsou popsány interakce proteinových inhibitorů se serinovými proteasami koagulační kaskády (především zástupci Kunitz-BPTI a serpinů) a imunitního systému hostitele (zástupci Kunitz-BPTI, serpinů a cystatinů) ^{157,158}. Endogenní funkce proteasových inhibitorů klíšťať nebyly dosud detailně studovány, ale byla navržena jejich role při embryogenezi, v imunitním systému klíšťete a při regulaci trávení ¹⁵⁷.

1.4.3. Rodina cystatinů u klíšťat

V klíšťatech byla prokázána produkce intracelulárních steffinů a sekretovaných pravých cystatinů. Pro steffiny byla identifikována řada sekvencí, ale jejich funkce nebyly dosud studovány. Detailně bylo popsáno pět zástupců steffinů (Rmcystatin-1b¹⁵⁹, Hlcyst-1¹⁶⁰, DvM602¹⁶¹, Bmcystatin-1¹⁶² a RHcyst-1¹⁶³), u kterých byla diskutována možná účast na regulaci trávení nebo funkce v embryogenezi^{162,164,165}. Biochemicky bylo charakterizováno zhruba 15 zástupců pravých cystatinů identifikovaných ve slinných žlázách, ale také ve střevu a dalších tkáních, přehledně jsou uvedeny v tabulce 3. K dnešnímu dni nebyly u *I. ricinus* popsány žádné inhibitory z cystatinové rodiny.

Tab. 3: Přehled zástupců rodiny cystatinů z klíšťat a jejich předpokládané funkce. Údaje o inhibiční specifitě jsou omezené vzhledem k nejednotnému spektru testovaných proteas. Převzato z^{157,166} a upraveno.

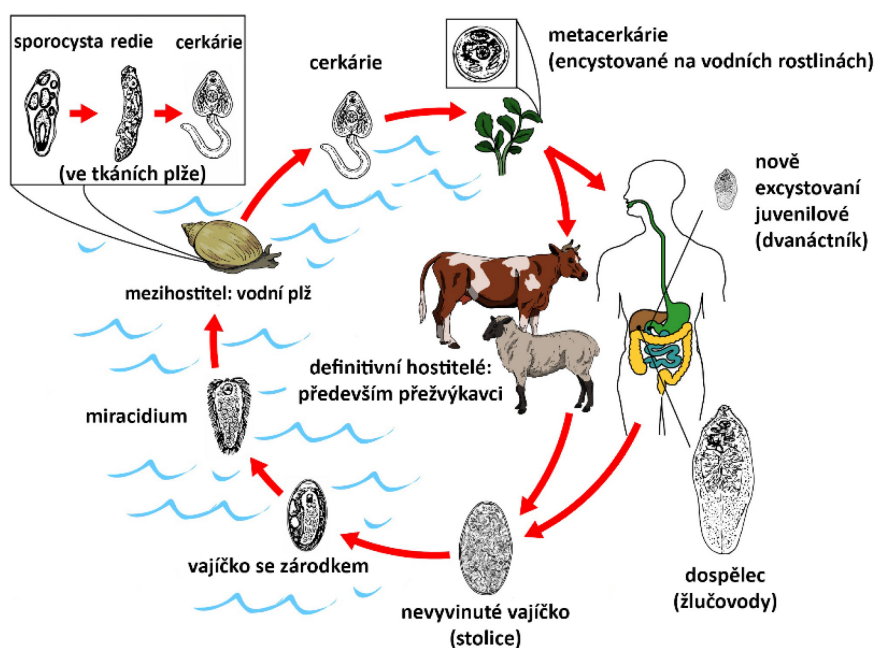
Podrodina	Jméno	Druh klíštěte	Inhibované katepsiny	Odhadovaná fyziologická funkce
steffiny	Hlcyst-1	<i>Haemaphysalis longicornis</i>	B, H, L, papain	trávení
pravé cystatiny	HISC-1	<i>Haemaphysalis longicornis</i>	L, papain	trávení, modulace imunitního systému klíštěte
pravé cystatiny	Hlcyst-2	<i>Haemaphysalis longicornis</i>	L	modulace host. imunitního systému, interakce s patogeny
pravé cystatiny	Sialostatin L	<i>Ixodes scapularis</i>	L, V, papain	modulace host. imunitního systému
pravé cystatiny	Sialostatin L2	<i>Ixodes scapularis</i>	L, V, S	modulace host. imunitního systému, interakce s patogeny
pravé cystatiny	BrBmcys2b	<i>Rhipicephalus microplus</i>	B, C, L	trávení, modulace imunitního systému klíštěte
pravé cystatiny	BrBmcys2c	<i>Rhipicephalus microplus</i>	C, L	trávení, modulace imunitního systému klíštěte
pravé cystatiny	Rmcystatin-3	<i>Rhipicephalus microplus</i>	B, L, trávicí RmCL1	modulace imunitního systému klíštěte
steffiny	Bmcystatin	<i>Rhipicephalus microplus</i>	B, L	embryogeneze
steffiny	Rmcystatin-1b	<i>Rhipicephalus microplus</i>	B, L, papain, trávicí RmCL1	trávení
pravé cystatiny	Ra-cyst-1	<i>Rhipicephalus appendiculatus</i>	nejsou známy	trávení
steffiny	RHcyst-1	<i>Rhipicephalus haemaphysaloides</i>	B, C, L, S, papain	embryogeneze
pravé cystatiny	RHcyst-2	<i>Rhipicephalus haemaphysaloides</i>	B, C, H, L, S, papain	embryogeneze
pravé cystatiny	OmC2	<i>Ornithodoros moubata</i>	B, C, H, L, S, papain	trávení, modulace host. imunitního systému
pravé cystatiny	OmC1	<i>Ornithodoros moubata</i>	B, C, H, papain	trávení

Klíšťové pravé cystatiny nejsou příliš specifické, typicky jsou jejich inhibiční konstanty vůči všem testovaným cysteinovým katepsinům velmi blízké (v rozsahu dvou řádů) a tlumí tak pravděpodobně veškerou proteolytickou aktivitu cysteinových katepsinů v jejich cílovém prostředí¹⁶⁷. Odhad jejich fyziologické funkce se tedy zakládá především na jejich lokalizaci. Střevním cystatinům je přisuzována regulační funkce v procesu trávení¹⁶⁰, cystatinům slinných žláz modulace imunitní odpovědi hostitele¹⁶⁸ a cystatiny přítomné ve vaječnicích a vejcích mají vliv na embryogenezi¹⁶⁵. Zástupce nalezený v tukovém tělisku a hematocytu nejspíše hraje roli v imunitním systému klíštěte¹⁶⁹. Nicméně je časté, že jsou jednotlivé cystatiny přítomny ve více typech tkání najednou, a tak nejspíše zastávají více zmíněných tkáňově specifických funkcí¹⁶⁴.

Sialomu klíšťat (transkriptom ze slinných žláz) je věnována velká pozornost, s cílem porozumět procesům spojeným se sáním a trávením krve, přenosem patogenů a modulací hostitelského prostředí. Proto patří k prvním nejlépe prozkoumaným cystatinům klíšťat zástupci ze slin, jako jsou homologické sialostatiny L a L2 z klíšťete *Ixodes scapularis*. Sialostatiny disponují unikátním inhibičním profilem, který je zaměřen zejména na cysteinové katepsiny s endopeptidasovou aktivitou ^{168,170}. Jejich prostorová struktura odhalila modifikovanou smyčku L2 se zbytky Asn/Ser-Leu, která připomíná konzervovaný motiv stefinů Pro-Leu, místo motivu Pro-Trp typického pro pravé cystatiny ¹¹⁸. U stefinů nicméně tato smyčka nevede ke stejnému inhibičnímu profilu. Sialostatiny jsou přítomné ve slinách *I. scapularis* a mají silné a komplexní imunomodulační účinky. Tlumí dozrávání cytotoxických T-lymfocytů skrze inhibici katepsinů S a L, zodpovědných za generaci peptidů pro MHC II receptory antigen-prezentujících buněk ^{168,171}. Inhibice katepsinů L a H může ovlivňovat produkci cytokinů s komplexním vlivem na imunitní systém hostitele ¹⁷². Významná je i inhibice katepsinu C, který je zodpovědný za aktivaci granulárních serinových proteas (granzym A a B, tryptasa, proteinasa 3, chymasa, katepsin G a elastasa) v různých druzích imunitních buněk ^{170,173}. Důležitost sialostatínů pro klíšťata dokazují i experimenty s cíleným tlumením jejich exprese pomocí RNA interference, při nichž 40 % klíšťat nebylo schopno dokončit krmení ¹⁷⁴. To dělá ze slinných cystatinů klíšťat zajímavé antigeny pro vývoj protiklíšťecích vakcín, které by zabraňovaly dlouhodobému sání klíšťat na hostiteli a přenosu patogenů.

1.5. Motolice jaterní (*Fasciola hepatica*)

Motolice jaterní (*Fasciola hepatica*) je významný, celosvětově rozšířený parazitický helmint. Má složitý dvouhostitelský životní cyklus (Obr. 13) s vodním plžem jako mezihostitelem a velkými býložravými savci, ale i člověkem jako definitivním hostitelem. Způsobuje onemocnění fasciolózu, kterou jsou nakaženy zhruba 3 miliony lidí a další desítky milionů jsou v ohrožení ¹⁷⁵. Dále má fasciolóza obrovský dopad na produkci hospodářských zvířat s odhadovanými ztrátami v řádu miliard dolarů ročně. Současná terapie spoléhá pouze na přípravek triklabendazol ¹⁷⁶, proti kterému jsou hlášeny rezistentní kmeny již od 90. let ¹⁷⁷.



Obr. 13: Životní cyklus motolice jaterní (*F. hepatica*). Schéma životního cyklu s jednotlivými vývojovými stádii, s meziphostitelem a definitivními hostiteli; lokalizace daného stádia je uvedena v závorce. Převzato z ¹⁷⁸ a upraveno.

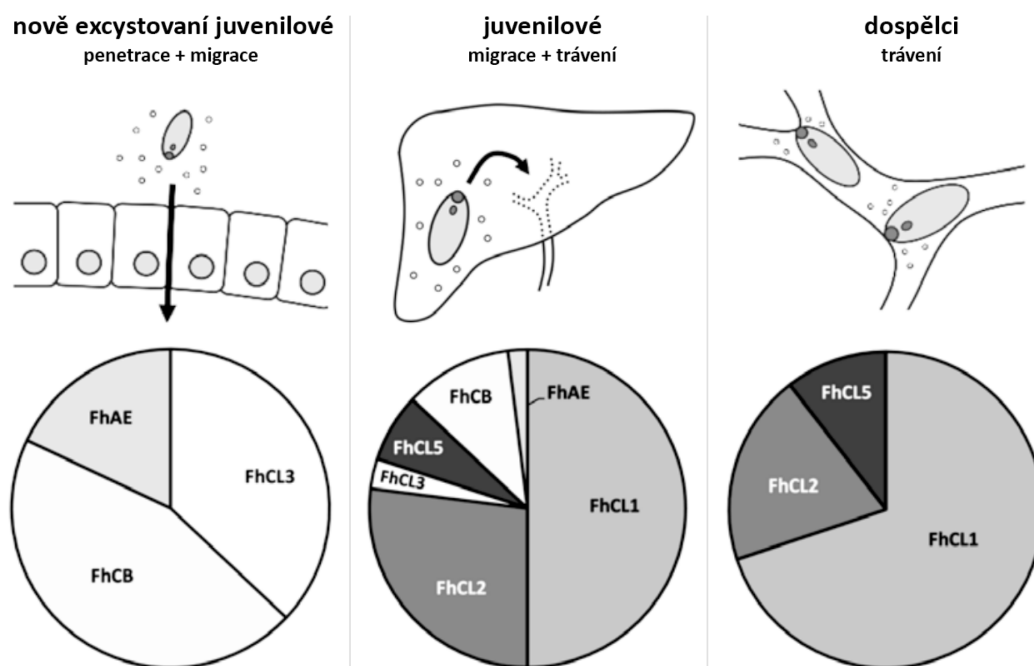
Po pozření rostliny s cystami metacerkárií definitivním hostitelem dojde ve dvanáctníku díky změně pH k excystaci juvenilních motolic označovaných NEJ (nově excystovaní juvenilové). Ty pronikají střevní stěnou do břišní dutiny a následně na zhruba 10 týdnů do jaterního parenchymu. Poté juvenilní motolice migrují do žlučovodů, kde pohlavně dospívají a začínají klást vajíčka odcházející se stolicí ¹⁷⁹. Dospělci *F. hepatica* se dožívají až 13 let a živí se krví a tkání stěn žlučovodů ¹⁸⁰. Fasciolóza má akutní a chronickou fázi. Akutní fáze začíná proniknutím larev střevní stěnou a končí dosažením pohlavní dospělosti ve žlučovodech. Je charakteristická silnou zánětlivou odpovědí organismu a může docházet k lokálním nebo systémovým alergickým reakcím a vnitřnímu krvácení ¹⁸¹. Následná chronická fáze je již bezpříznaková s mírně zvýšenou hladinou eozinofilních granulocytů ¹⁸².

Trávení proteinů u motolic probíhá zpočátku v lumen střeva, které je slepé, a proto je jeho obsah vyvrhován zhruba každé 3 hodiny, jako prevence akumulace toxického hemu ^{51,179}. Takto se do okolí parazita dostává velké množství parazitárních proteas, inhibitorů a dalších molekul. Tato směs látek vylučovaných parazitem do prostředí hostitele se nazývá sekretom či exkrečně-sekreční produkty (ESP).

1.5.1. Cysteinové katepsiny u *F. hepatica*

Cysteinové katepsiny helmintů prodělaly velkou funkční diverzifikaci spojenou často se vznikem genových rodin určitých typů katepsinů⁵¹. *F. hepatica* ztratila geny kódující katepsin C a F, ale její geny pro katepsiny L a B prodělaly silnou expanzi, a tak nyní disponuje 26 různými isoenzymy katepsinu L (FhCL) (Obr. 5, kapitola 1.4), 11 isoenzymy katepsinu B (FhCB) či 8 isoenzymy legumainu (FhAE)⁵¹.

F. hepatica je na rozdíl od *I. ricinus* endoparazitem a je tak v delším kontaktu s hostitelským prostředím. Tomu odpovídá rozšířený repertoár funkcí cysteinových katepsinů, který kromě trávení zahrnuje i degradaci tkání hostitele nebo ovlivnění hostitelského imunitního systému. Spektrum exprimovaných proteas *F. hepatica* se v průběhu života parazita zásadně mění (Obr. 14).



Obr. 14: Složení sekretovaných katepsinů v různých fázích životního cyklu *F. hepatica* u definitivního hostitele. U NEJ jsou exprimovány především katepsiny FhCL3 a FhCB zajišťující penetraci střevní stěny a migraci dutinou břišní. FhAE (legumain) slouží k trans-aktivaci FhCB. V jaterním parenchymu je exprese FhCL3 a FhCB, postupně nahrazena trávicími katepsiny FhCL1 a FhCL2. Tyto dvě proteasy spolu s FhCL5 pak u dospělců tvoří více než 80 % veškerých sekretovaných proteinů. Převzato z⁵¹ a upraveno.

Trávení hemoglobinu u *F. hepatica* probíhá v lumen střeva v mírně kyselém pH, při kterém tetramery hemoglobinu disociují⁵⁷. Předpokládá se, že nejdříve katepsiny L (FhCL1, FhCL2, FhCL5) štěpí molekulu hemoglobinu na krátké fragmenty o délce 4-14 aminokyselin, které jsou transportovány do buněk střevního epitelu¹⁸³. Další degradace probíhá v lysosomech těchto buněk pomocí katepsinů B (FhCB) a metaloproteas (leucinové aminopeptidasy) na jednotlivé aminokyseliny a dipeptidy¹⁸⁴. Hlavní trávicí proteasa motolice je FhCL1, který je specializován na štěpení hemoglobinu, s podmístem S2 specifickým pro Leu, Val, Ala a Phe, tedy zbytky, které tvoří více než 40 % molekuly hemoglobinu^{52,57}).

Některé cysteinové katepsiny motolice jaterní (např. FhCB a FhCL3) mají kolagenasovou aktivitu a umožňují tak stádiu NEJ překonat bariéry hostitelského organismu^{185,186}. Tyto bariéry, stejně jako extracelulární matrix jsou tvořeny především kolagenem, který obsahuje sekvenční motiv Gly-Xxx-Xxx s vysokým obsahem Pro a Hyp (hydroxyprolin), který běžné proteasy nejsou schopny štěpit⁶⁸. Parazitární katepsiny produkované migrujícími stádii motolic tak mají, oproti lidským homologům upravenou substrátovou specifitu, podobně jako lidský katepsin K¹⁰. Např. FhCL3 má substrátovou specifitu evolučně přizpůsobenou své kolagenolytické funkci a preferované aminokyseliny v jeho podmístech (Gly v S1 a S3, Pro v S2) jsou u cysteinových katepsinů naprosto nezvyklé, avšak odpovídají sekvenci kolagenu¹⁸⁵. Srovnání substrátové specifity FhCL3 s FhCL1 či lidským katepsinem L pak svědčí o výrazné evoluční diferenciaci genové rodiny katepsinu L u *F. hepatica*^{52,185}.

Bylo prokázáno, že sekretované katepsiny, především FhCL1 a FhCL2, modulují imunitní odpověď hostitele, a to pomocí dvou různých mechanismů. Zaprvé štěpí protilátky typu IgG a IgE tak, že se jejich Fab fragmenty (zodpovědné za rozpoznávání antigenů) vážou na povrchové proteiny *F. hepatica* a maskují parazita před rozpoznáním imunitním systémem¹⁸⁷. Zadruhé jsou tyto proteasy internalizovány do lysosomů makrofágů, kde štěpí TLR3 („toll-like receptor 3“), což působí protizánětlivě a podporuje hojivé procesy hostitele, včetně obnovy stěn žlučovodů^{188,189}.

1.5.2. Proteasové inhibitory u *F. hepatica*

Proteasové inhibitory motolic hrají stejně jako u klíšťat klíčovou roli nejen při regulaci vlastních fyziologických procesů, ale také modulují proteolýzu v hostitelském prostředí. Proto jsou v současné době zkoumány jako potenciální vakcinační antigeny. V transkriptomu motolice jaterní bylo nalezeno celkem 12 rodin proteasových inhibitorů cílených proti cysteinovým a serinovým proteasám a metaloproteasám. Podobně jako u katepsinů došlo

i u klíčových rodin inhibitorů Kunitz-BPTI (I2), serpinů (I4) a cystatinů (I25) ke vzniku genových rodin. *F. hepatica* tak např. disponuje 11 variantami monomerních a 4 variantami multidoménových Kunitz-BPTI inhibitorů. Podobně jako u proteas jsou expresní profily inhibitorů velmi rozdílné v závislosti na životním stádiu parazita⁵⁰.

U *F. hepatica* byly doposud alespoň částečně charakterizovány tyto proteasové inhibitory na proteinové úrovni: multidoménový cystatin¹³³, tři stefiny¹⁹⁰, čtyři serpiny^{191,192} a dva inhibitory z rodiny Kunitz-BPTI⁹⁸. Cystatiny a stefiny jako inhibitory cysteinových proteas jsou diskutovány v následující kapitole. Serpiny a proteiny z rodiny Kunitz-BPTI jsou typicky inhibitory serinových proteas. V případě serpinů exprimovaných NEJ stádii se předpokládá jejich role v modulaci imunitního systému hostitele a krevního srážení nebo v oogenezi^{191,192}. Pro inhibitory z rodiny Kunitz-BPTI identifikované u *F. hepatica* bylo překvapivě zjištěno, že nejsou inhibitory serinových proteas, ale jsou cíleny na cysteinové katepsiny^{50,193}. Pouze jeden isoinhibitor FhKT1.3 je schopen inhibovat oba typy proteas. Exprese zástupců Kunitz-BPTI rodiny je závislá na životním stádiu parazita. Některé jsou exprimovány v metacerkáriích a NEJ a poskytují jim pravděpodobně ochranu při migraci z tenkého střeva do jater. Další jsou produkováni dospělci a podílejí se na regulaci trávení v lumen střeva a na embryogenezi; vzhledem k sekreci nelze vyloučit ani jejich vliv na modulaci imunitního systému hostitele⁹⁸.

1.5.3. Rodina cystatinů u *F. hepatica*

U *F. hepatica* byl identifikován jeden multidoménový cystatin a tři stefiny^{133,190}. Multidoménový cystatin byl nalezen u NEJ stádia a obsahuje 6 domén pravých cystatinů, ze kterých je pouze čtvrtá doména schopna inhibovat papain a FhCL1, u ostatních domén došlo k deformaci inhibičních segmentů¹³³. U příbuzného druhu motolice *F. gigantica* byl popsán multidoménový cystatin, který obsahuje 12 domén pravých cystatinů, což je nejvyšší počet domén ze všech známých cystatinů. Čtyři domény si zachovaly schopnost inhibovat cysteinové katepsiny. Tento inhibitor díky své lokalizaci ve varlatech a ovariích hraje nejspíše roli v spermatogenezi¹³².

V roce 2017 byla publikována studie, která provedla částečnou charakterizaci tří stefinů *F. hepatica* označených FhStf-1 až 3¹⁹⁰. Jde o nanomolární inhibitory katepsinů FhCL1 a lidského katepsinu L, které se liší v inhibiční specifitě k lidským katepsinům B a C, s nimiž interaguje pouze FhStf-2. Geny FhStf-2 a FhStf-3 obsahují signální sekvenci na rozdíl od FhStf-1¹⁹⁰; vysvětlení fenoménu neobvyklé přítomnosti signálního peptidu v sekvenci

stefinů přináší tato disertační práce. Další studie prokázala, že FhStf-2 (nyní označovaný jako FhCyLS-2) je schopen modulovat lidský imunitní systém a vykazuje protizánětlivé účinky ¹⁹⁴. U *F. gigantica* byl popsán stefin FgStf-1, ortolog FhStf-1 z *F. hepatica*, tvořící významnou část ESP, jehož sekreční mechanismus zatím nebyl identifikován ¹⁹⁵.

2. Cíle práce

Proteiny z rodiny cystatinů jsou důležitými inhibitory cysteinových proteas typu katepsinů a účastní se řady patologií. Parazitární cystatiny jsou atraktivní cílové molekuly pro regulaci parazitů, ale je o nich dosud málo informací. Tato práce se zaměřuje na cystatiny dvou významných krevsajících parazitů: motolice jaterní (*Fasciola hepatica*), původce fasciolózy a klíštěte obecného (*Ixodes ricinus*) hlavního vektoru lymfské boreliózy a klíšťové encefalidity. Cílem práce je popsat funkční a strukturní vlastnosti vybraných zástupců parazitárních cystatinů a interpretovat získané výsledky z pohledu molekulární evoluce a fyziologické role.

Dílčí cíle disertační práce jsou následující:

- 1) Vybrat modelové zástupce cystatinů z obou parazitů na základě fylogenetické analýzy dostupných sekvenčních dat (v rámci spolupráce s Parazitologickým ústavem AV ČR a Parazitologickým ústavem Univerzity v Curychu).
- 2) Použít rekombinantní formy studovaných cystatinů k biochemické charakterizaci interakce s fyziologicky relevantními a modelovými proteasami s cílem určit jejich inhibiční specifitu.
- 3) Nalézt krystalizační podmínky rekombinantních cystatinů a vyřešit jejich 3D strukturu pomocí rentgenostrukturní analýzy. Pokusit se krystalizovat jeden z cystatinů v komplexu s vhodnou cílovou proteasou za účelem strukturního popisu reaktivního centra inhibitoru, který zatím není u parazitárních cystatinů k dispozici.
- 4) Analyzovat strukturně-funkční vztahy studovaných cystatinů a vyvodit závěry týkající se molekulární evoluce cystatinů u parazitů a jejich potenciálních fyziologických funkcí.

3. Laboratorní vybavení a metody

Tato kapitola popisuje materiály, vybavení a metody využité pro získání nepublikovaných výsledků, v případě publikovaných výsledků je použita metodika vždy součástí přiložených publikací.

3.1. Materiál a laboratorní vybavení

Většina výsledků byla získána s využitím laboratorního vybavení na Ústavu organické chemie a biochemie AV ČR (ÚOCHB). Syntetické peptidové substráty a inhibitory byly zakoupeny od firem Bachem (Švýcarsko). Lidský katepsin L byl zakoupen od společnosti Merck (USA), katepsiny B a S od společnosti Enzo Life Science (USA), legumain od společnosti R&D Systems (USA) a lidský katepsin V a hovězí katepsiny C a H byly připraveny rekombinantní expresí v domácí laboratoři dle ¹⁹⁶⁻¹⁹⁸. Sady komerčních krystalizačních roztoků byly od firem Molecular Dimensions (Velká Británie) Jena Bioscience (Německo) a Hampton Research (USA). Testování difrakce krystalů a sběr difrakčních dat probíhalo buď na rentgenové difrakční stanici na ÚOCHB nebo na experimentálních stanicích pro makromolekulární krystalografii na synchrotronu BESSY II provozovaném Helmholtz-Zentrum Berlin (Německo).

3.2. Metodika

3.2.1. Purifikace a příprava komplexu ricistatinu a katepsinu V

Jednotlivé rekombinantní proteiny a následně i komplex ricistatinu a mutantu Cys25Ser katepsinu V (připravený v molárním poměru 1,2:1) byly purifikovány pomocí gelové chromatografie na FPLC přístroji Äkta Basic (GE Healthcare, USA) a koloně Superdex 75 10/300 GL Tricorn (GE Healthcare, USA). Jako pufr byl použit 50 mM citrát sodný pH 5,5, 150 mM NaCl. Čistota proteinů i komplexu byla analyzována pomocí SDS-PAGE (Laemmli 1970). Frakce obsahující proteiny resp. komplex byly spojeny a koncentrovány. Krystalizační roztok komplexu byl převeden do 10 mM citrátu sodného pH 5,5, 10 mM NaCl a zakoncentrován na 3 mg/ml.

3.2.2. Kinetické inhibiční testy

Měření byla prováděna v triplikátech v bílých 96-jamkových destičkách Corning (Thermo Fisher Scientific, USA) s objemem testu 100 μ l při 37 °C na fluorescenční čtečce M1000 (Tecan, Männedorf, Švýcarsko) s FRET substráty a excitační a emisní vlnovou délkou 365 nm,

resp, 450 nm. Ricistatin byl inkubován v reakční směsi s testovanou proteasou po 10 min při 37 °C, následně byla reakce zahájena přidávkem substrátu.

Složení substrátů a proteas v reakční směsi bylo následující: 53 μM Z-Phe-Arg-AMC s 0,2 nM katepsinem B; 2,5 μM s 0,2 nM katepsinem L; 17 μM ; 3,6 μM s 0,1 nM katepsinem V; 130 μM H-Gly-Arg-AMC s 0,4 nM katepsinem C; 100 μM H-Arg-AMC s 1,9 nM katepsinem H; Z-Val-Val-Arg-AMC s 0,1 nM katepsinem S a 70 μM Z-Ala-Ala-Asn-AMC s 1,2 nM legumainem. Složení reakčních pufrů bylo následující: 100 mM Na-acetát, pH 5,0 pro legumain, pH 5,5 50 mM MES pro katepsiny B, C, L a V, nebo pH 6,5 pro katepsin H a S; reakční pufr dále obsahovaly 2,5 mM DTT, 0,1% PEG 1500 a 50 mM NaCl (pro katepsin C). Hodnoty IC_{50} byly stanoveny z residuálních rychlostí nelineární regrese vypočtených programem GraFit (Erithacus, East Grinstead, Velká Británie).

3.2.3. Krystalizace komplexu ricistatinu a katepsinu V

Pro nalezení krystalizačních podmínek komplexu ricistatinu s katepsinem V byla využita metoda difuze par v sedící či visící kapce při 18°C za využití krystalizačních robotů Gryphon (Art Robbins Instruments, USA) a OryX8 (Douglas Instruments, Velká Británie), 96-jamkových destiček Swissci (Molecular Dimensions, Velká Británie) a následujících krystalizačních sad: JCSG++ (Jena Bioscience, Německo), Morpheus (Molecular Dimensions, Velká Británie), Stura FootPrint (Molecular Dimensions, Velká Británie), SG1 (Molecular Dimensions, Velká Británie) a sady na přípravu mikrokrytalů Seed Bead (Hampton Research, USA).

Optimalizace probíhala stejnou metodou s využitím 15-jamkových destiček (Molecular Dimensions, Velká Británie). Finální krystalizační podmínka obsahovala 1 μl roztoku komplexu, 1 μl roztoku finálního krystalizačního roztoku (70 mM octan amonný, 3,17% PEG 10000) a 0,2 μl 100x ředěného roztoku rozdrcených krystalů komplexu. Připravené krystaly byly krátce vystaveny kryoprotektivnímu roztoku (70 mM octan amonný, 3,17% PEG 10000, 25% glycerol) a poté zchlazeny v kapalném dusíku.

3.2.4. Řešení struktury komplexu ricistatinu a katepsinu V

Difrakční data byla sbírána při teplotě 100 K na rentgenové difrakční stanici v laboratoři Strukturální biologie na ÚOCHB. Difrakční data byla zpracována pomocí sady programů XDS¹⁹⁹ a analyzována programem Xtriage ze sady PHENIX²⁰⁰. Krystaly vykazovaly symetrii C121 a obsahovaly jednu kopii proteinového komplexu v asymetrické jednotce. Krystalové parametry a statistiky nasbíraných difrakčních dat jsou uvedeny v kapitole 4.4.2. v Tab. 4.

Struktura komplexu ricistatinu a Katepsinu V byla vyřešena pomocí metody molekulárního nahrazení programem MolRep (Vagin and Teplyakov 2010) ze sady CCP4 (Winn et al. 2011). Jako model byl nejprve použit Katepsin V (PDB: 1TH0) a následně sialostatin L2 (PDB: 3LH4). Model byl upřesněn v programu REFMAC 5.8 (Murshudov et al. 2011) ze sady CCP4, přičemž 5 % reflexí bylo vyhrazeno pro generaci testovací sady. Molekuly vody a manuální úpravy byly provedeny v programu Coot (Emsley et al. 2010). Kvalita výsledného modelu byla ověřena pomocí programu MolProbity (Williams et al. 2018). Finální statistiky upraveného modelu byly generovány pomocí programu PHENIX (Liebschner et al., 2019) a jsou uvedeny v kapitole 4.4.2. v Tab. 4. Atomové souřadnice a strukturní faktory byly deponovány do proteinové databanky (PDB) s kódem 7PK4 a jsou již veřejně přístupné. Ilustrace byly připraveny pomocí programu PyMOL (Schrodinger 2015), interakce mezi ricistatinem a Katepsinem V byly analyzovány pomocí webových nástrojů PISA (Krissinel and Henrick 2007) a PLIP (Adasme et al. 2021).

4. Výsledky

Výsledky disertační práce jsou prezentovány ve třech článcích publikovaných v mezinárodních recenzovaných časopisech, kterým se věnují kapitoly 4.1. až 4.3. s komentáři hlavních výsledků a uvedením podílu autora na publikacích. Kapitola 4.4. obsahuje dosud nepublikované výsledky, které se v současné době připravují ke zveřejnění.

Tyto publikace a nepublikovaná data se zabývají strukturní a funkční charakterizací čtyř nových proteinů z rodiny cystatinů, které jsou produkovány parazity jako inhibitory cysteinových proteas. V první publikaci je charakterizován inhibitor FhCyLS-2 z motolice jaterní (*F. hepatica*) a na jeho základě je definována nová evoluční skupina cystatinů. V druhé a třetí publikaci jsou popsány dva cystatiny z klíštěte obecného (*I. ricinus*) – mialostatin ze střeva klíštěte a iristatin ze slin klíštěte, které se účastní regulace odlišných proteolytických systémů klíštěte, resp. hostitele. Kapitola 4.4. se zabývá inhibitorem ricistatinem z *I. ricinus* a přináší první strukturní charakterizaci vazebného módu slinných cystatinů.

Seznam publikací:

Publikace č. 1: **An evolutionary molecular adaptation of an unusual stefin from the liver fluke *Fasciola hepatica* redefines the cystatin superfamily.**, Buša M., Matoušková Z., Bartošová-Sojková P., Páchl P., Řezáčová P., Eichenberger R. M., Deplazes P., Horn M., Štefanić S., Mareš, M., *Journal of Biological Chemistry*, 2023, (IF = 5,5)

Publikace č. 2: **Mialostatin, a novel midgut cystatin from *Ixodes ricinus* Ticks: crystal structure and regulation of host blood digestion.**, Kotál J*, Buša M*, Urbanová V, Řezáčová P, Chmelař J, Langhansová H, Sojka D, Mareš M, Kotsyfakis M., *International Journal of Molecular Sciences*, 22(10):5371, 2021, (IF = 6,2) * Oba autoři se podíleli stejným dílem

Publikace č. 3: **The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin.**, Kotál J, Stergiou N, Buša M, Chlastáková A, Beránková Z, Řezáčová P, Langhansová H, Schwarz A, Calvo E, Kopecký J, Mareš M, Schmitt E, Chmelař J, Kotsyfakis M. *Cellular and Molecular Life Sciences*, 76(10):2003-2013, 2019, (IF = 9,2)

4.1. Publikace č. 1: An evolutionary molecular adaptation of an unusual stefin from the liver fluke *Fasciola hepatica* redefines the cystatin superfamily.

Sekretované inhibitory z podrodiny pravých cystatinů hrají u parazitů důležitou roli při interakci s hostitelem. *F. hepatica* je jedním z celosvětově nejvýznamnějších parazitů člověka a hospodářských zvířat, nicméně žádné pravé cystatiny u ní nebyly identifikovány. Vysvětlení umožňuje tato studie, která popisuje evoluci nové skupiny parazitárních cystatinů a detailně charakterizuje jejich modelového zástupce FhCyLS-2.

FhCyLS-2 byl exprimován v dospělých parazitech získaných z těla hostitele. Pomocí imunofluorescenční mikroskopie a imunoblotování bylo prokázáno, že FhCyLS-2 je lokalizován ve střevu parazita a je též uvolňován v exkrementně-sekrecních produktech. Bylo také zjištěno, že sekretovaný FhCyLS-2 je rozpoznán imunitním systémem hostitele a dochází ke vzniku specifických protilátek v séru experimentálně infikovaných zvířat.

FhCyLS-2 byl připraven jako rekombinantní protein a byly analyzovány jeho funkční biochemické vlastnosti. V kinetických inhibičních testech bylo zjištěno, že FhCyLS-2 je širokospektrým inhibitorem cysteinových proteas z rodiny papainu. Je schopen efektivně inhibovat (v nanomolárních koncentracích) jednak hlavní trávicí katepsiny *F. hepatica* (katepsiny FhCL1 až FhCL3) a jednak sadu hostitelských cysteinových katepsinů s endopeptidasovou i exopeptidasovou aktivitou. Ze získaných výsledků lze předpokládat, že FhCyLS-2 má duální funkci při ochraně trávicí tkáně parazita proti proteolytickému poškození a při regulaci hostitelských katepsinových proteas na rozhraní parazita s hostitelem.

Analýza sekvence ukázala, že FhCyLS-2 patří do podrodiny stefinů a nikoli pravých cystatinů. Překvapivě ale obsahuje dva hlavní strukturní znaky typické pro pravé cystatiny – signální sekvenci a dva disulfidy. Přítomnost signální sekvence u FhCyLS-2 vysvětluje, proč byl nalezen v exkrementně-sekrecních produktech *F. hepatica*, ačkoli zástupci podrodiny stefinů jsou typické intracelulární proteiny. Přesné párování disulfidů bylo určeno pomocí hmotnostní spektrometrie. Dále bylo zjištěno, že FhCyLS-2 se od ostatních stefinů odlišuje absencí C-terminálního motivu Asp/Glu-Xxx-Leu-Xxx-Tyr/His-Phe. V dalším kroku byla určena prostorová struktura FhCyLS-2 rentgenostrukturní analýzou s rozlišením 1,6 Å s využitím metody isomorfního nahrazení. FhCyLS-2 má typickou architekturu stefinů s konzervovaným vazebným centrem, které je tvořeno N-koncem molekuly a dvěma povrchovými smyčkami L1 a L2. Molekula je stabilizována dvěma disulfidickými můstky, které mají lokalizaci blízkou, nikoli ale identickou jako konzervované disulfidy pravých cystatinů.

Fylogenetická analýza parazitárních cystatinů a stefinů odhalila, že existuje série sekvencí homologických s FhCyLS-2 ze střevních a jaterních motolic řádu Plagiorchiida

a Opisthorchiida. Tyto proteiny vytvářejí v podrodině stefinů sesterskou větev, která nyní umožnila definovat novou skupinu s názvem CyL-stefiny („Cystatin-Like stefins“). Všechny druhy motolic řádu Plagiorchiida disponující CyL-stefiny postrádají geny kódující pravé cystatiny a lze tedy předpokládat, že CyL-stefiny jsou evoluční úpravou stefinů, která kompenzuje absenci pravých cystatinů.

Můj podíl na práci zahrnoval: (1) rekombinantní expresi a purifikaci FhCyLS-2, (2) rekombinantní expresi a purifikaci katepsinů FhCL1 a FhCL2 z motolice jaterní, (3) určení inhibičních konstant pro FhCyLS-2 pomocí kinetických testů s fluorescenčními substráty a různými proteasami, (4) testování inhibičního efektu FhCyLS-2 na nativní parazitární proteasy v testech s hostitelskými proteinovými substráty, (5) krystalizaci FhCyLS-2, včetně přípravy derivovaných krystalů a sběru nativních a derivovaných difrakčních dat na synchrotronu, (6) řešení a analýzu krystalové struktury FhCyLS-2, (7) přípravu sekvencí pro fylogenetickou analýzu, (8) přípravu manuskriptu.

Journal Pre-proof

An evolutionary molecular adaptation of an unusual stefin from the liver fluke *Fasciola hepatica* redefines the cystatin superfamily

Michal Buša, Zuzana Matoušková, Pavla Bartošová-Sojková, Petr Páchl, Pavlína Řezáčová, Ramon Marc Eichenberger, Peter Deplazes, Martin Horn, Saša Štefanić, Michael Mareš

PII: S0021-9258(23)00102-3

DOI: <https://doi.org/10.1016/j.jbc.2023.102970>

Reference: JBC 102970

To appear in: *Journal of Biological Chemistry*

Received Date: 14 November 2022

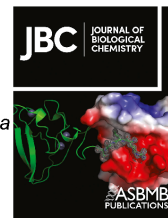
Revised Date: 26 January 2023

Accepted Date: 29 January 2023

Please cite this article as: Buša M, Matoušková Z, Bartošová-Sojková P, Páchl P, Řezáčová P, Eichenberger RM, Deplazes P, Horn M, Štefanić S, Mareš M, An evolutionary molecular adaptation of an unusual stefin from the liver fluke *Fasciola hepatica* redefines the cystatin superfamily, *Journal of Biological Chemistry* (2023), doi: <https://doi.org/10.1016/j.jbc.2023.102970>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023 THE AUTHORS. Published by Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology.



An evolutionary molecular adaptation of an unusual stefin from the liver fluke *Fasciola hepatica* redefines the cystatin superfamily

Michal Buša^{1,2}, Zuzana Matoušková^{1,2}, Pavla Bartošová-Sojková³, Petr Páchl¹, Pavlína Řezáčová¹, Ramon Marc Eichenberger⁴, Peter Deplazes⁴, Martin Horn¹, Saša Štefanić^{4,*} and Michael Mareš^{1,*}

From the ¹Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague, Czechia; ²Department of Biochemistry, Faculty of Science, Charles University, Prague, Czechia; ³Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Budweis, Czechia; ⁴Institute of Parasitology, University of Zurich, Zurich, Switzerland.

*Corresponding authors: sasa.stefanic@uzh.ch, mares@uochb.cas.cz

Running title:

Cystatin-like stefins redefine the cystatin superfamily

Keywords:

cystatin, stefin, protease inhibitor, cysteine cathepsin, protein structure, protein evolution, helminth parasite

Abstract

Fasciolosis is a worldwide parasitic disease of ruminants and an emerging human disease caused by the liver fluke *Fasciola hepatica*. The cystatin superfamily of cysteine protease inhibitors is composed of distinct families of intracellular stefins and secreted true cystatins. FhCyLS-2 from *F. hepatica* is an unusual member of the superfamily, where our sequence and 3D structure analyses in this study revealed that it combines characteristics of both families. The protein architecture demonstrates its relationship to stefins, but FhCyLS-2 also contains the secretion signal peptide and disulfide bridges typical of true cystatins. The secretion status was confirmed by detecting the presence of FhCyLS-2 in excretory/secretory products, supported by immunolocalization. Our high-resolution crystal structure of FhCyLS-2 showed a distinct disulfide bridging pattern and functional reactive center. We determined that FhCyLS-2 is a broad specificity inhibitor of cysteine cathepsins from both the host and *F. hepatica*, suggesting a dual role in the regulation of exogenous and endogenous proteolysis. Based on phylogenetic analysis that identified several FhCyLS-2 homologs in liver/intestinal foodborne flukes, we propose a new group within the cystatin superfamily called cystatin-like stefins.

Introduction

The common liver fluke, *Fasciola hepatica*, is a globally spread parasitic flatworm of the class Trematoda that infects mainly sheep and cattle, causing the disease fasciolosis. It negatively affects livestock production with estimated worldwide losses of €2.5 billion annually (1). Fasciolosis is an important zoonosis with 2.4 million people infected, and 150 million in over seventy countries at risk, and is classified by the WHO as a neglected tropical disease. The control of fasciolosis relies on a single drug, triclabendazole, and reports of resistant strains of *F. hepatica* are on the rise (2).

The superfamily of cystatins (Merops database ID: I25) is ubiquitous in a wide range of organisms, including animals, plants, fungi, and protista. The superfamily members are reversible, tight-binding inhibitors of cysteine proteinases; they possess a reactive center against papain-family proteases, and some members contain an additional reactive center against legumains (3-5). Based on their sequence and spatial structure, cystatins can be divided into three families: the type 1 family of stefins (Merops ID: I25A) with intracellular single-domain proteins devoid of the signal sequence and disulfide bridges, the type 2 family of true cystatins (Merops ID: I25B) with secreted single-domain proteins containing two disulphide bridges, and the type 3 family of kininogens (Merops ID: I25C) with multidomain proteins composed of I25B repeats (5,6). The most characterized are mammalian cystatins, which were shown to regulate a broad range of physiological and pathological processes, including protein catabolism, hormone processing, bone resorption, antigen processing, inflammation, and tumor metastasis (for review, see e.g. (5,7-9)).

In parasites, cystatins are not only essential in the regulation of physiological processes during parasite development, but they also represent important pathogenicity factors. Various reports indicated that parasite cystatins have evolved exceptional immunomodulatory properties. Most of the examples studied originated from ticks and helminths. In ticks, it was shown that gut-associated cystatins regulate endogenous digestive proteases, while salivary cystatins that are injected into the host can inhibit the secretion of pro-inflammatory cytokines, reduce T-cell proliferation or disrupt dendritic cell maturation and differentiation (10,11). Helminth cystatins were demonstrated to block antigen processing and presentation, interfere with the processing of pattern recognition receptors in innate immunity, modulate production of cytokines and nitric oxide, and suppress T-cell proliferation (12-16).

F. hepatica cysteine proteases from the papain family (cysteine cathepsins) represent about 80% of secreted proteolytic activity of the adult parasite, and belong to cathepsins L and B. They are involved in invasion, migration through the host body, feeding, immune evasion and are regulated by several mechanisms, including activation processing, compartmentalization, and interaction with protease inhibitors (17-20).

Previous studies identified several inhibitors of cysteine cathepsins in *F. hepatica*, including stefins, a multidomain cystatin, and a Kunitz protein (21-23), which were suggested as candidate targets for a fasciolosis vaccine (24,25). A similar distribution of inhibitors has been reported for a closely related species, *F. gigantica* (26,27). No *Fasciola*-derived member of the cystatin superfamily has yet been structurally characterized and in general, our knowledge of helminth cystatins at the structure-function level is very limited (28,29).

In this work, we investigated the cysteine cathepsin inhibitor FhCyLS-2 from *F. hepatica*, belonging to the cystatin superfamily and analyzed its crystal structure, inhibitory properties, localization, and phylogeny. Based on distinct molecular features of FhCyLS-2 and its homologs, we defined a new evolutionary group called cystatin-like stefins (CyL-stefins, CyLS) within the cystatin superfamily, with FhCyLS-2 as a prototype member characterized at the structural and functional levels.

Results

FhCyLS-2 sequence combines characteristics of stefins and true cystatins

The full-length cDNA available in GenBank under accession no. AY647146.1 contains an open reading frame coding for a 116-amino acid protein (GenBank accession no. AAV68752), which we designated here as *Fasciola hepatica* cystatin-like stefin 2 (FhCyLS-2). In Figures 1 and S1 we compared the sequence of FhCyLS-2 with that of other members of the cystatin superfamily, including representative vertebrate members of the type 1 family (stefins) and type 2 family (true cystatins) (Fig. 1) and selected invertebrate members of both families from various blood feeding parasites (Fig. S1). The sequence alignments revealed several important features in the FhCyLS-2 sequence, which allowed us to assess the relationship between FhCyLS-2 and both families.

Our results showed that FhCyLS-2 and stefins did not contain a large insert in the central part of the molecule, which is seen in true cystatins (between strands $\beta 2$ and $\beta 3$, Figs. 1 and S1). On the other hand, FhCyLS-2 lacked a C-terminal sequence motif Asp/Glu-Xxx-Leu-Xxx-Tyr/His-Phe (30) typical for stefins (Figs. 1 and S1). The interaction of all cystatin superfamily members with cysteine cathepsins is mediated by three regions, including the N-terminal segment and two hairpin loops L1 and L2 (3,31). In FhCyLS-2, the mature N-terminus carried the Gly-Gly motif and the L1 loop bore the Gln-Xxx-Val-Xxx-Gly motif; both motifs are conserved among stefins and true cystatins (Figs. 1 and S1). The L2 loop was characterized in stefins and FhCyLS-2 by the presence of a conserved central Leu residue, in contrast to a Trp residue in true cystatins (Figs. 1 and S1). To summarize, FhCyLS-2 possessed the sequence insertion/deletion pattern and signature of a functionally competent reactive center for inhibition that were related to those present in stefins.

Analysis of the FhCyLS-2 sequence identified a signal peptide composed of 20 N-terminal residues with a predicted cleavage site between residues Gly and Glu1 (the first residue in the mature protein) (Fig. S1). The signal peptide is a hallmark of true cystatins, which are secreted proteins, and distinguish them from stefins localized intracellularly due to the lack of the signal peptide (6,32).

The FhCyLS-2 sequence also contained four cysteine residues, and we investigated their redox status (i.e., free or disulfide-bonded). Disulfide bridges are a typical feature of true cystatins, however they are absent in stefins, which are usually devoid of cysteine residues or contain unpaired cysteines with a free thiol (6,32). We prepared recombinant FhCyLS-2 and performed

LC-MS/MS peptide mapping. Analysis of disulfide-linked peptide clusters revealed that the cysteine residues formed two disulfide bridges with connectivity Cys39–Cys66 and Cys69–Cys95 (Table S1). They are labelled as D1 and D2, respectively, in the FhCyLS-2 sequence in Figure 1. Thus, FhCyLS-2 resembles true cystatins containing two disulfides; a comparison of their disulfide patterns is discussed in the chapter on the 3D structure of FhCyLS-2.

In conclusion, we demonstrated that FhCyLS-2 was a close homolog of the stefin family due to sequence similarity. However, at the same time, FhCyLS-2 also displayed two hallmarks of the family of true cystatins, namely two disulfides and the N-terminal signal sequence for secretion. Furthermore, in the sequence databases we found a set of stefin sequences of trematode origin, which were highly homologous to FhCyLS-2 and also contained the signal peptide as well as putative disulfides (Figs. S1). Based on this finding, we propose a new group of the cystatin superfamily named cystatin-like stefins (CyL-stefins) with FhCyLS-2 as a first member. In the following chapters, we will analyze FhCyLS-2 in detail; regarding the signal sequence function, disulfide-bridged 3D structure, and phylogeny of CyL-stefins.

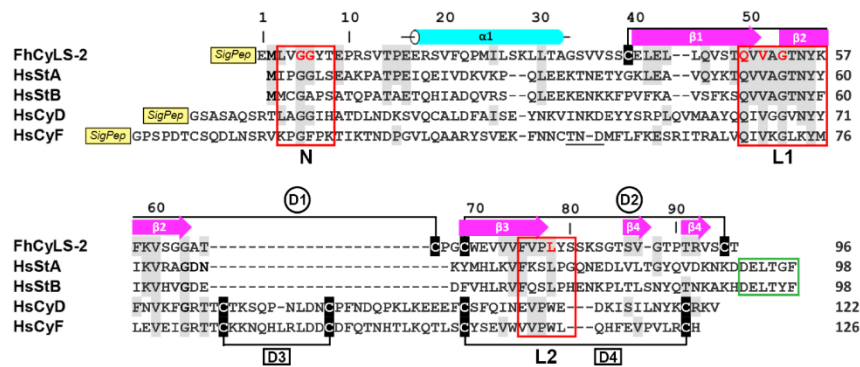


Figure 1. Structure-based sequence alignment of FhCyLS-2 with representative members of cystatin superfamily. FhCyLS-2 is compared with representative human members of the type 1 family, stefins A and B (HsStA and HsStB, respectively), and the type 2 family, cystatins D and F (HsCyD and HsCyF, respectively). Residues identical with those of FhCyLS-2 are shaded grey, cysteine residues are shaded black. The disulfide bridges are indicated by the connecting black lines and labeled (circled D1 and D2 for FhCyLS-2, boxed D3 and D4 for cystatins D/F). The secondary structure elements of FhCyLS-2 are depicted in cyan for α -helix and magenta for β -strands and labeled as in Fig. 5A. Three segments forming the reactive center for inhibition of cysteine cathepsins are boxed in red and labeled as in Fig. 5A (the region size was selected based on the predominant binding residues in the available complex structures); the critical

consensus positions are highlighted in red for FhCyLS-2. The conserved C-terminal segment of classical stefins is boxed in green. Mature protein sequences without signal peptides (indicated by a SigPep box) were used in the alignment; residue numbering on the top line is according to FhCyLS-2. GenBank accessions: FhCyLS-2: AAV68752, HsStA: AAH10379, HsStB: AAH10532, HsCyD: AAH62678, HsCyF: CAG46658.

***F. hepatica* developmental stages express and secrete FhCyLS-2**

The mRNA transcript levels for FhCyLS-2 were evaluated in eggs, miracidia, metacercariae, newly excysted juveniles (NEJ) and adults by qRT-PCR (Fig. 2); intra-molluscan stages and cercariae were not investigated. The highest expression of FhCyLS-2 was recorded in adults followed by NEJ and metacercariae. Expression in miracidia was below 1% of the FhCyLS-2 level in adults, and no expression was detected in eggs. The transcript level increased gradually with development from miracidia to adult.

The excretory/secretory (E/S) products collected from adult *F. hepatica* were subjected to proteomic analysis to directly demonstrate the presence of FhCyLS-2. The LC-MS/MS strategy was based on enzymatic digestion of a complex protein mixture and MS/MS peptide sequencing. This analysis provided 22% peptide coverage of the FhCyLS-2 sequence, allowing us to conclude that FhCyLS-2 was secreted into the E/S products of *F. hepatica*. Furthermore, using this approach we determined the N-terminus of native FhCyLS-2 (**Table S1**) and demonstrated that mature FhCyLS-2 was produced by cleavage of the signal peptide at the Gly(-1)-Glu1 bond (Fig. S1).

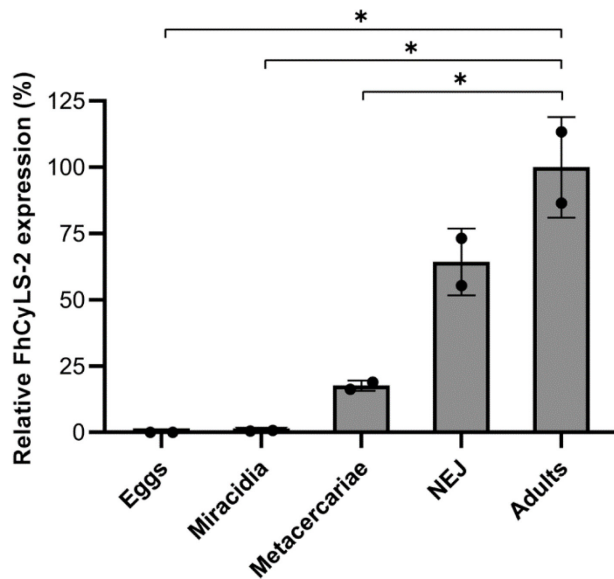


Figure 2. Transcriptional profiling of FhCyLS-2 in the developmental stages of *F. hepatica*. The expression of FhCyLS-2 was evaluated by qRT-PCR. mRNA transcriptional levels are presented as percentage of expression relative to that in adults. The mean values \pm SD for pooled biological samples are given; statistical significance compares adults with other stages (* $p < 0.05$).

FhCyLS-2 is localized in the worm gut, reproductive system, and tegument surface

Indirect immunofluorescence microscopy on semi-thin sections using polyclonal antibodies raised against recombinant FhCyLS-2 demonstrated that FhCyLS-2 was expressed in distinct tissues of adult worms (Fig. 3). FhCyLS-2 was observed in the intestinal epithelium and the reproductive tract of adult parasites; a much weaker signal was present on tegument surface. FhCyLS-2 was absent in parenchyma or vitelline cells. Pre-immune serum was used as a negative control and only faint background fluorescence was detected (Fig. S2).

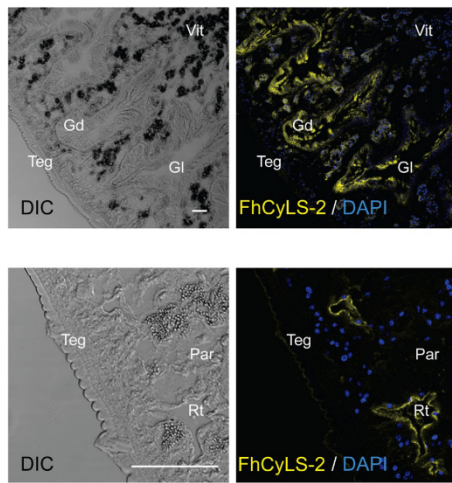


Figure 3. FhCyLS-2 immunolocalization in sections of an adult *F. hepatica*. Semi-thin tissue sections of adult worms were probed with an anti-FhCyLS-2 antiserum followed by reaction with an anti-mouse IgG Alexa 633-labeled secondary antibody (yellow). DAPI was used to label nuclear DNA (blue). The left column shows differential interference contrast (DIC), the right column fluorescent channels. Upper panels: FhCyLS-2 localized to the gastroderm of the adult worm (Gd: gastrodermal cells, Gl: gut lumen, Teg: tegument, Vit: vitelline cells). Lower panels: Higher magnification revealed a faint signal on the tegument surface and an intense signal lining the ducts of presumably reproductive organs (Rt: reproductive tract, Par: parenchymal tissue, Teg: tegument). The scale bars represent 100 μm ; negative control staining is shown in Figure S2.

Preparation of recombinant FhCyLS-2

Recombinant FhCyLS-2 was prepared as a mature protein with an oligohistidine tag appended to its C-terminus (predicted mass of 11698 Da) and expressed using a *Pichia pastoris* system. The protein purified to homogeneity (see Experimental procedures) migrated on SDS-PAGE as a single band of ~ 12 kDa (Fig. 4). Its identity was confirmed by LC-MS/MS analysis with complete peptide coverage. Mouse polyclonal antibodies, raised against recombinant FhCyLS-2 were used for the detection of native FhCyLS-2. In the E/S products and tissue extract of adult *F. hepatica*, the antibody recognized a band of ~ 12 kDa corresponding to FhCyLS-2 (Fig. 4), which indicated that FhCyLS-2 was released into the host environment.

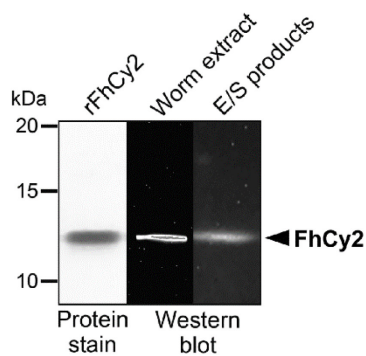


Figure 4. Preparation of recombinant FhCyLS-2 and identification of native FhCyLS-2. The recombinant FhCyLS-2 (rFhCyLS-2) expressed in *P. pastoris* was resolved by SDS-PAGE and protein-stained. Protein extract of adult *F. hepatica* and the excretory/secretory (E/S) products were visualized by western blotting using mouse polyclonal antibodies raised against rFhCyLS-2. The position of FhCyLS-2 is indicated.

FhCyLS-2 is a potent inhibitor of mammalian and *F. hepatica* cysteine cathepsins

The purified recombinant FhCyLS-2 was screened *in vitro* for its inhibitory activity against a panel of cysteine proteases; its inhibitory profile is shown in Table 1. We demonstrated that FhCyLS-2 interacts with proteases of the CA but not the CD clan from the cysteine protease class. This was tested with papain and legumain (Table 1), which are representative members of the CA and CD clans, respectively, and archetypes in cystatin research (3,33). A set of papain-family proteases (family C1, clan CA) of mammalian origin was then screened, including cathepsins F, K, L, S, and V (endopeptidases), cathepsins B and X (a peptidyl dipeptidase/endopeptidase and carboxypeptidase, respectively), and cathepsins C and H (a dipeptidyl peptidase and aminopeptidase, respectively). These enzymes were selected to cover a wide range of endo- and exopeptidase cleavage specificities. FhCyLS-2 strongly inhibited all these cysteine cathepsins, with IC_{50} values ranging from approx. 0.9 to 8.6 nM, except for cathepsins F and C (IC_{50} of 49 and 66 nM, respectively) and cathepsin X (IC_{50} of 8.4 μ M), which were inhibited with lower potency (Table 1). Furthermore, FhCyLS-2 was tested against three major *F. hepatica* cathepsin L proteases, which are important *Fasciola* virulence factors: cathepsins L1 (FhCL1) and L2 (FhCL2) expressed in immature and mature flukes, and cathepsin L3 (FhCL3)

from infectious larvae (18). These recombinant cysteine cathepsins from *F. hepatica* were highly sensitive to inhibition, with IC₅₀ values in the low nanomolar range (2.5 to 18.1 nM) (Table 1).

We also investigated the effect of FhCyLS-2 on the proteolytic activity of native cysteine cathepsins in the E/S products of adult *F. hepatica*, which were attributed mainly to FhCL1 and FhCL2 (18). Activity measured using a specific peptide substrate was almost completely inhibited by FhCyLS-2 in agreement with the data obtained for recombinant *F. hepatica* cathepsins L (Table S2). This result was further supported by evaluation of FhCyLS-2 in an assay using general protein substrates, including physiologically relevant collagen and elastin. It showed that FhCyLS-2 was capable of effectively blocking most proteolytic activity of the E/S products (Table S2).

In summary, FhCyLS-2 was identified as a broad-specificity inhibitor targeting various cysteine cathepsins, including exogenous mammalian enzymes and endogenous digestive enzymes secreted by *F. hepatica*.

Table 1. Inhibitory effect of FhCyLS-2 on the activity of cysteine proteases. The inhibitory potency of FhCyLS-2 was determined against papain-family proteases from the CA clan and legumain from the CD clan. The IC₅₀ values (means ± SEM) were determined by a kinetic activity assay using specific fluorogenic peptide substrates. The Merops database classification of the cysteine proteases tested (clan/family) and their mode of action are given. n.i. indicates no significant inhibition at 10 μM inhibitor concentration.

Enzyme	IC ₅₀ (nM)	Enzyme specificity, protease clan/family
mammalian cathepsins		
Cathepsin L	0.92 ± 0.01	Endopeptidase, CA/C1
Cathepsin S	1.1 ± 0.1	Endopeptidase, CA/C1
Cathepsin B	2.1 ± 0.2	Peptidyl dipeptidase/endopeptidase, CA/C1
Cathepsin H	2.5 ± 0.2	Aminopeptidase, CA/C1
Cathepsin V	5.9 ± 0.3	Endopeptidase, CA/C1
Cathepsin K	8.6 ± 0.7	Endopeptidase, CA/C1
Cathepsin F	49 ± 5	Endopeptidase, CA/C1
Cathepsin C	66 ± 7	Dipeptidyl peptidase, CA/C1
Cathepsin X	8400 ± 800	Carboxypeptidase, CA/C1
<i>Fasciola hepatica</i> cathepsins		
Cathepsin L1	2.5 ± 0.5	Endopeptidase, CA/C1
Cathepsin L2	10 ± 1	Endopeptidase, CA/C1
Cathepsin L3	18 ± 1	Endopeptidase, CA/C1
others		
Papain	1.1 ± 0.1	Endopeptidase, CA/C1
Legumain	n.i.	Endopeptidase, CD/C13

Three-dimensional structure of FhCyLS-2: a unique disulfide-bridged stefin-like architecture

The crystal structure of recombinant FhCyLS-2 was determined by single isomorphous replacement and refined using native data to 1.60 Å resolution. FhCyLS-2 crystallized in an orthorhombic space group $C222_1$ with one molecule in the asymmetric unit and a solvent content of 48.3% (Table S3). All protein residues could be modeled into a well-defined electron density map except for the first three N-terminal residues (Glu1–Leu3) and the C-terminal oligohistidine tag (Ala97–His105). The final model of FhCyLS-2 contained 93 residues (Val4–Thr96).

Figure 5A shows the overall structure of FhCyLS-2 formed by a four-stranded twisted antiparallel β -sheet, which were wrapped around a central α -helix. The molecule adopted a typical fold of classical stefins like the human homologs stefins A and B (Fig. 5B). The closest structural homolog of FhCyLS-2 was stefin B with the lowest RMSD for C α (1.98 Å), followed by stefin A (2.14 Å RMSD) (Fig. 5B); lower structural similarities were found with true cystatins represented by human cystatins D and F (3.90 Å and 3.34 Å RMSD, respectively) (Fig. 5C).

In contrast to stefins A and B, the intact C-terminal strand β 4 was split into two smaller β -strands in FhCyLS-2 (Figs. 1 and 5A). Importantly, FhCyLS-2 incorporated two disulfide bridges into the disulfide-free scaffold of stefins. They were located in loops at the edge of the β -sheet and indicated as D1 (Cys39–Cys66) and D2 (Cys69–Cys95) (Fig. 5A). However, the two-disulfide pattern of FhCyLS-2 differed from that conserved in true cystatins (their disulfides are indicated as D3 and D4 in Figs. 1 and 5C). While the positions of the C-terminal disulfides (D2 and D4) were similar, the upstream disulfides formed an interloop bridge (D1) in FhCyLS-2 but an intraloop bridge (D3) in the cystatins D and F (Figs. 1 and 5C). All disulfides were located on the side of the molecules opposite to the reactive center (Fig. 5C).

In both stefins and true cystatins, the reactive center creates a tripartite wedge-shaped edge formed by the N-terminal trunk and two hairpin loops L1 and L2, which binds into the active site cleft of cysteine cathepsins (3,31). In FhCyLS-2 (Figs. 1 and 5A), the N-terminal trunk was located around the conserved pair of glycines (Gly5, Gly6) providing conformational flexibility to this reactive segment in order to adopt an optimal conformation for target binding. The L1 loop (between β 1 and β 2) of FhCyLS-2 exposed the segment Gln49-Val50-Val51-Ala52-Gly53 corresponding to the critical binding motif Gln-Xaa-Val-Xaa-Gly. The L2 loop (between β 3 and β 4)

was characterized in FhCyLS-2 and classical stefins by the presence of a conserved Leu78 residue.

In conclusion, the crystallographic analysis demonstrated that FhCyLS-2 is a structurally distinct member of the type 1 family of stefins equipped with a unique pattern of disulfide bridges and a functionally competent reactive site.

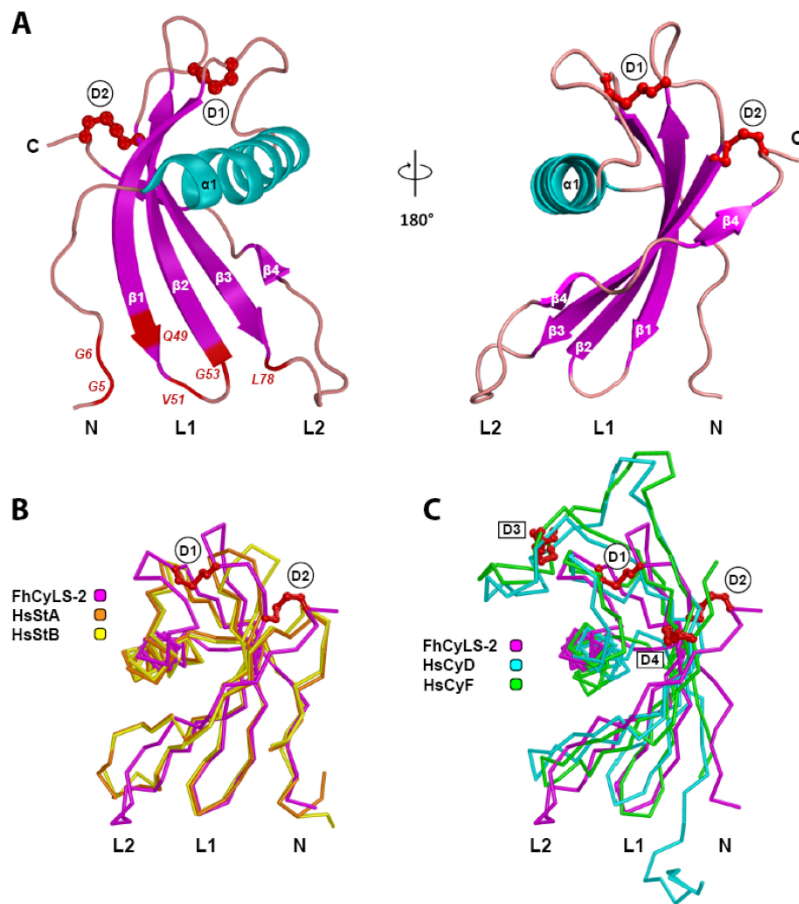


Figure 5. The crystal structure of FhCyLS-2 and its comparison with cystatins from type 1 and type 2 families. *A*, The 3D structure of FhCyLS-2 (PDB code 6l1M) is shown in a cartoon representation colored by secondary structure elements ($\alpha 1$ helix in cyan, $\beta 1$ – $\beta 4$ strands in magenta). The N- and C-termini (labelled N and C, respectively) and two disulfide bridges Cys39–Cys66 (D1) and Cys69–Cys95 (D2) (represented by red sticks and balls) are indicated. The hairpin loops L1 and L2 and the N-terminus are involved in the binding of inhibitors from the cystatin superfamily to cysteine cathepsins; the predicted critical residues are highlighted and labeled in red. *B and C*, Superposition of Ca traces of FhCyLS-2 (magenta) with human stefins A and B (orange and yellow, respectively) and with human cystatins D and F (cyan and green, respectively). Disulfide bridges are shown as red sticks and balls and labeled as in Fig. 1.

The segments N, L1, and L2 forming the reactive center against cysteine cathepsins are defined in Fig. 1. The orientation of FhCyLS-2 is as in A (the right-hand view). PDB accessions: FhCyLS-2: 6I1M, stefin A (protease bound): 3KSE, stefin B (protease bound): 1STF, cystatin D: 1RN7, cystatin F: 2CH9.

Phylogenetic analysis of FhCyLS-2 and cystatin-like stefins

Phylogenetic analysis of FhCyLS-2 and other monomeric members of the cystatin superfamily is presented in Figure 6. It shows two main monophyletic clades of type 1 and type 2 families, corresponding to stefins and true cystatins, respectively. Within the major clades, all sequences were grouped into several taxonomy-defined subclades. Importantly, stefins were divided into a monophyletic clade of classical stefins and a stem group of cystatin-like stefins (CyL-stefins) including FhCyLS-2. Sequences of CyL-stefins were distinct from classical stefins by containing the signal sequence (as indicated in Fig. 6); this feature was also present in true cystatins. All identified CyL-stefins originated from certain taxa of parasitic Trematoda and Cnidaria, namely the foodborne flukes (Trematoda) infecting the gastrointestinal, biliary and respiratory tract of their definitive vertebrate hosts (fasciolids, echinostomatids, psilostomids, paramphistomids, microphalids, troglotrematids, opisthorchiids, heterophyids) and the myxozoans (Cnidaria). Interestingly, no CyL-stefins were identified in the blood flukes (schistosomatids) (Fig. 6).

F. hepatica FhCyLS-2 (Genbank: AY64714) firmly grouped with two paralogs from the same species (Genbank: THD26684, THD28140) and with the orthologs from other liver and intestinal foodborne flukes (Plagiorchiida: fasciolids, echinostomatids, psilostomids), forming the largest clade of CyL-stefins (Fig. 6). The only other sequence from the cystatin superfamily found in *F. hepatica* (Genbank: THD19897) belonged to the clade of classical stefins (Figs. S1 and S3). Thus, no true cystatin homolog was identified for *F. hepatica* and for other Plagiorchiida. In contrast, members of Opisthorchiida and Schistosomatida possessed homologs that clustered in the clade of true cystatins.

In conclusion, our phylogenetic analysis revealed that FhCyLS-2 is formally a member of the type 1 family, stefins. However, FhCyLS-2 and its homologs, defined as CyL-stefins from Trematoda and Myxozoa species, were evolutionarily different from classical stefins and formed a distinct stem group in the phylogenetic branch of the type 1 family. *F. hepatica* and all other species from the order Plagiorchiida encoded both CyL-stefins and classical stefins but lacked genes for true cystatins.

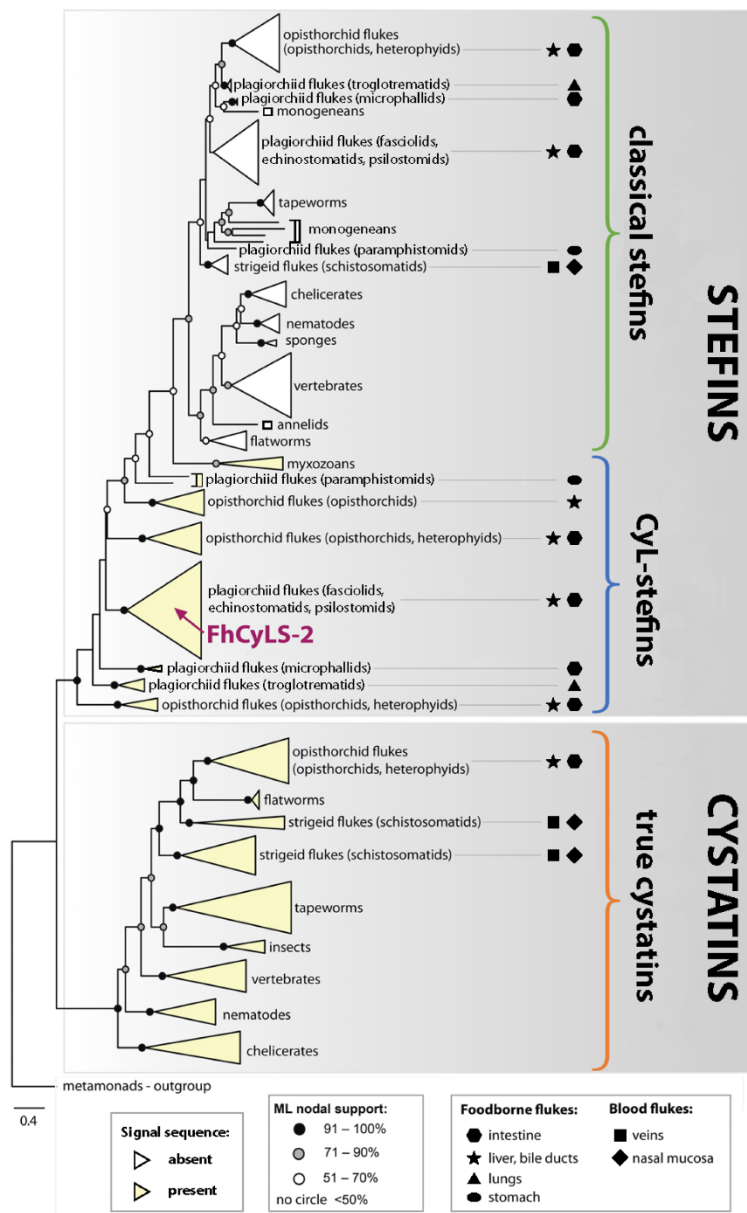


Figure 6. Maximum likelihood phylogenetic tree of the cystatin superfamily. The rooted tree of vertebrate and invertebrate members shows two clades of the type 1 family of stefins and the type 2 family of true cystatins; the former is divided into classical stefins and cystatin-like stefins (CyL-stefins). The position of *F. hepatica* FhCyLS-2 is marked. Note that sequences of CyL-stefins and true cystatins contain the signal peptide (yellow-colored taxa/clades). The stem group of CyL-stefins covers foodborne flukes (Trematoda) and myxozoans (Cnidaria). The organ location of mature fluke stages in the definitive host is indicated by symbols for each group. The metamonad *Giardia intestinalis* cystatin, related to an ancestral gene of the superfamily (6), was used as the outgroup. Nodal supports calculated from 1000 bootstrap replicates are depicted as dots on each node. The detailed phylogenetic tree with taxa names and input sequences are provided in Figure S3 and File S1.

FhCyLS-2 is recognized by serum from naturally and experimentally infected ruminants

Sera from experimentally infected sheep and from naturally infected cows were examined for reactivity with recombinant FhCyLS-2 in an ELISA assay. Experimentally infected sheep recognized FhCyLS-2 three weeks post infection, and the ELISA signals steadily increased until week 5 post infection, when the experiment had to be terminated (Fig. 7A). In addition, sera from naturally infected cattle, in which *F. hepatica* infections were demonstrated by the presence of *F. hepatica* eggs in the bile of slaughtered animals, reacted significantly stronger with FhCyLS-2 than sera from non-infected cattle (Fig. 7B).

These results further support the conclusion that FhCyLS-2 is an actively secreted protein from *F. hepatica*. Moreover, secreted FhCyLS-2 is recognized by the host during *F. hepatica* infection, which suggests that FhCyLS-2 may have potential value as a diagnostic marker for fasciolosis.

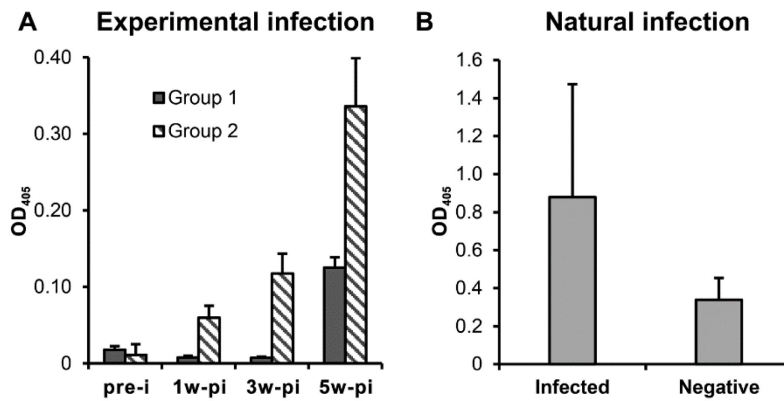


Figure 7. Recognition of FhCyLS-2 by serum antibodies from *F. hepatica* infected ruminants. Sera response to recombinant FhCyLS-2 was analyzed by ELISA and expressed as means \pm SD (* $p < 0.05$, ** $p < 0.01$). **A**, Sera from two groups of experimentally infected sheep were drawn prior to infection (pre-i) with *F. hepatica* metacercaria and 1, 3, and 5 weeks post inoculation (1 to 5w-pi). Statistical significance compares pre-i and 5w-pi data. **B**, Sera from cattle naturally infected with *F. hepatica* collected from a slaughterhouse are compared with sera from non-infected cattle).

Discussion

In this work, we provide a comprehensive structural, functional, and phylogenetic analysis of FhCyLS-2, an unusual member of the cystatin superfamily from *F. hepatica*. The sequence and spatial structure revealed that FhCyLS-2 combines, in an unprecedented way, hallmarks of two major families in the cystatin superfamily, stefins and true cystatins. Homology in the protein structure indicated a close relationship of FhCyLS-2 to stefins; however, FhCyLS-2 also contains the secretion signal peptide and disulfides, two structural features typical of true cystatins. We demonstrated that FhCyLS-2 was present in the E/S products of adult parasites and was released most likely from their gut. The secreted FhCyLS-2 was recognized by the immune system of the host during *F. hepatica* infection, leading to production of specific antibodies detected in the serum. The crystallographic analysis showed that FhCyLS-2 incorporated two disulfide bridges into the disulfide-free scaffold of stefins, and the spatial pattern of disulfides was related to but not identical with that of true cystatins. The acquisition of disulfides might reflect the secretion status of FhCyLS-2 that differs from that of typical intracellular stefins. In general,

disulfides stabilize proteins in the oxidizing extracellular environment, and therefore they are present in the majority of secreted proteins while they are rather rare in cytosolic proteins (34).

The FhCyLS-2 molecule bears the reactive center with an overall architecture common to its superfamily and is functionally competent for protease binding. FhCyLS-2 acts as an effective multitarget inhibitor of cysteine cathepsins of mammalian as well as *F. hepatica* origins. Figure 8 compares inhibitory specificity of FhCyLS-2 with that of other members of the superfamily. It shows that FhCyLS-2 effectively inhibits cysteine cathepsins with a wide range of endo- and exopeptidase cleavage specificities and thus resembles several broad-specificity inhibitors, including previously well-characterized human and tick stefins, gut-associated tick cystatins, but also some secreted human cystatins. Mammalian cysteine cathepsins sensitive to FhCyLS-2 play an essential role in a variety of immunological mechanisms such as antigen processing and activation of neutrophils, cytotoxic T-lymphocytes, and natural killer cells (for review, see (35,36)).

Inhibitory potency
IC₅₀ or K_i (nM)

Protease	FhCyLS-2	Stefin A	Stefin B	Rhcyst-1	Cystatin C	Cystatin D	Cystatin F	Mialostatin	OmC2	Sialostatin L2	Sialostatin L	Iristatin
	hCatL	0.92	1.3	0.23	34	0.005	25	0.31	1	0.1	1.1	1
hCatS	1.1	0.05	0.07	0.3	0.008	0.24	33	2.2	0.1	200	0.2	n.i.
hCatC	66	33	0.23	280	3.5	n.i.	>1000	2.1	1.1	680	670	500
hCatH	2.5	0.31	0.58	24	0.28	8.5	36	29	1.2	n.i.	n.i.	n.i.
hCatB	2.1	8.2	73	590	0.25	>1000	>1000	10	8.8	n.i.	n.i.	n.i.
Legumain	n.i.	n.i.	n.i.	n.i.	0.2	n.i.	10	n.i.	n.i.	n.i.	n.i.	n.i.
	trematoda	human	tick		human					tick		
	CyL-stefins	stefins			cystatins							

Figure 8. A comparison of the inhibitory potency of FhCyLS-2 and other members of the cystatin superfamily. Well documented classical stefins from the type 1 family and true cystatins from the type 2 family are presented, including members of human origin (stefins A and B, cystatins C, D, and F) and those from parasitic ticks (salivary sialostatins L and L2 (sialo L/L2), and iristatin; salivary gland-associated Rhcyst-1; gut-associated mialostatin and OmC2). Inhibition data against various cysteine proteases, including human papain-family cathepsins L to B (hCatL to hCatB) and mammalian legumains, are presented as IC₅₀ (tick inhibitors (10,37-40), FhCyLS-2

data are from Table 1) or K_i values (human inhibitors (4,41-45) and colored as a heat map (green scale); n.i. means no inhibition.

Therefore, we hypothesize that secreted FhCyLS-2 mimics specific host-derived cystatins and interferes with their functions *in vivo*, controlling cathepsin-mediated proteolysis as previously found for cystatins with various immunomodulatory activities produced by helminths or ticks (for review, see ((11,16)). Furthermore, FhCyLS-2 inhibits secreted digestive cathepsins L from *F. hepatica* and may function as a physiological regulator in order to protect parasite tissues from autoproteolytic damage. Thus FhCyLS-2 might have a dual role in the control of both exogenous and endogenous proteolytic systems during host-parasite interactions. Of note is that a similar combined function was proposed for the tick cystatin OmC2, which is secreted into the host for immunomodulatory action as well as regulates proteolysis in the tick gut (39).

The sequence and structural features and phylogenetic analysis revealed that FhCyLS-2 is a member of the type 1 family of stefins. We identified a set of FhCyLS-2 homologs in certain trematode and myxozoan parasite groups. They are evolutionarily different from classical stefins and form a distinct stem group in the phylogenetic branch of the type 1 family. Based on this, we defined a new group called cystatin-like stefins (CyL-stefins) within the cystatin superfamily, with FhCyLS-2 as a prototype member. They possess an obligatory signal peptide on the stefin-like sequence lacking the C-terminal motif Asp/Glu-Xxx-Leu-Xxx-Tyr/His-Phe. The structure of CyL-stefins typically contains two disulfides; however, their number can be reduced to one or, rarely, altogether absent.

A vast majority of animals express both classical stefins and true cystatins as important antiproteolytic molecules (5,6). We analyzed the repertoire of cystatin superfamily genes, including CyL-stefins, among individual fluke groups and found significant differences when comparing Plagiorchiida, Opisthorchiida, and Schistosomatida. *F. hepatica* and all other liver and intestinal foodborne flukes (Plagiorchiida) encode both CyL-stefins and classical stefins but lack genes for true cystatins. Conversely, no CyL-stefins were identified in the blood flukes (Schistosomatida) that possess classical stefins and true cystatins; however, opisthorchiid flukes combine all three inhibitor groups. In addition, Plagiorchiida and Opisthorchiida also express multicystatins. We hypothesize that CyL-stefins represent an evolutionary upgrade of classical stefins that occurred as a molecular adaptation to compensate for the absence of, or insufficient functional range of secreted true cystatins. CyL-stefins can form a multigene group (e.g. three homologs

present in *F. hepatica*) and this diversification burst, also known in true cystatins, is unusual compared to classical stefins, which remain rather evolutionarily stable without dynamic gene duplications (6). What drives the differences in the repertoire of superfamily genes among fluke taxa? It is tempting to speculate that they reflect the complexity of life cycles of flukes and strategies associated with parasite development and survival in host organs, and avoidance of host immunity. In general, a broader panel of superfamily members is associated with a higher complexity of multi-host life cycles, the pronounced involvement of vertebrate hosts possessing more advanced immunity than invertebrate hosts, and diverse types of invaded tissues with different spectra of host proteases (46,47). The overall complexity of these factors decreases from Opisthorchiida to Plagiorchiida and Schistosomatida, in line with their inhibitor repertoire (File S1).

In summary, we present a comprehensive structural and functional characterization of FhCyLS-2, a secreted inhibitor of cysteine cathepsins from *F. hepatica*, and propose its potential role in host-parasite interactions. The unique sequence and structural pattern of FhCyLS-2, together with phylogenetic analysis of its homologs, allowed us to define cystatin-like stefins as a new distinct group belonging to the cystatin superfamily.

Experimental procedures

Parasites and parasite-derived materials

Adult worms of *F. hepatica* were collected from the livers of naturally infected cattle acquired from the local abattoir. The worms were washed three times in wash medium (RPMI 1640 medium (Gibco) supplemented with Penicillin-Streptomycin-Amphotericin B (Antibiotic Antimycotic 100x, Gibco) and 10 mM HEPES, pH 7.3) and then transferred to complete culture medium (RPMI 1640 medium supplemented with 2 mM L-glutamine supplemented with 5% calf serum, 55 mM glucose, 30 mM HEPES, penicillin (100 U/ml), streptomycin (100 µg/ml), and gentamicin (25 µg/ml), pH 7.3). After 4 h incubation at 37 °C, the medium with freshly laid eggs were collected. Metacercariae were obtained from Ridgeway Research Ltd. NEJ material was prepared by excystation of metacercariae (19) and harvested after 24 h cultivation. For protein extraction, the worm tissue in PBS containing protease inhibitor cocktail (539131, Calbiochem) was mechanically disrupted by metal beads (Qiagen TissueLyser II), following by ultrasonication. The lysate was centrifuged at 16000g at 4 °C for 10 min, and the supernatant was stored at -80°C. Medium containing the E/S products of adult flukes was collected after cultivation for 4 h at 37 °C in serum-free medium, filtered through an Ultrafree-MC 0.22 µm filter (Millipore), concentrated, and stored at -80 °C.

Ethic statement

Animal experiments were carried out at the experimental units of the Vetsuisse Faculty at the University of Zurich after approval by the Cantonal Veterinary Office of Zurich (permission numbers 162/2009 for mice and 234/2012 for sheep) according to Swiss animal rights and regulation standards.

Isolation of mRNA, cDNA synthesis, and qRT-PCR

Total RNA was extracted from *F. hepatica* developmental stages using Trizol reagent (Invitrogen) and purified using the RNA Isolation Kit RNeasy (Qiagen) as reported previously (48). Single-stranded cDNA was synthesized from 1 µg of total RNA using SuperScript III reverse transcriptase (Invitrogen) and an oligo d(T)18 reverse primer according to the manufacturer's protocol, and the final cDNA product was purified (QIAquick PCR Purification Kit, Qiagen). The gene expression profile of the FhCyLS-2 was assessed using qRT-PCR. The primers (2 sets of forward and reverse primers, Table S4) were designed to amplify 160–180 bp fragments using

the Primer 3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and their efficiency was evaluated as described previously (49,50). Triplicate reactions, containing SYBR-green MasterMix (Roche), were carried out in a final volume of 25 µl in 96-well plates in a MX 3005P Real-Time PCR cycler (Biorad). The amplification profile consisted of an initial hot start (95 °C for 10 min) followed by 40 cycles comprising 95 °C for 30 s, 55 °C for 60 s and 72 °C for 30 s. Carboxyrhodamine and *F. hepatica* cytochrome C oxidase 9, subunit I (GenBank: KF111595.1) were used as a reference dye and gene, respectively. Upon completion of the amplification, the dissociation curve was examined for potential primer dimerization. The cycle threshold (CT) values were averaged, and the standard deviation was determined. The relative expression levels were calculated using the delta-delta CT method (51). The PCR reactions were performed in duplicate for each cDNA sample obtained from a pool of multiple parasites of individual developmental stages. Statistical significance was determined with the unpaired Student's t-test.

Cloning and expression of recombinant FhCyLS-2

FhCyLS-2 (GenBank: AY647146) was expressed in the X-33 strain of the methylotrophic yeast *Pichia pastoris* (Thermo Fisher). The full-length FhCyLS-2 gene without the N-terminal signal sequence predicted by SignalP v5.0 (52) was amplified from the adult stage cDNA using the forward primer, 5'-CGGAATTCGAGGTGAAATGCTCGTGGGTG-3' (EcoRI restriction site underlined) and reverse primer, 5'-CGTCTAGAGCTCAATGATGATGATGATGATGAGCTGCAGCAGTGCAGGATACCCGAGTC (XbaI restriction site underlined, 6xHis-tag in bold) with Phusion DNA polymerase (Fermentas). The insert sequence was cloned into pPICZαB vector (Thermo Fisher) using EcoRI and XbaI restriction sites. Recombinant FhCyLS-2 contained an N-terminal extension of residues AGIRG and a C-terminal extension of residues AAAHHHHH compared to the native mature sequence. The vector sequence was verified by DNA sequencing. Transformation of *P. pastoris* cells and protein expression were carried out as described previously (53). Media supernatants from induced *P. pastoris* were lyophilized and stored at -20 °C until use.

Purification of recombinant FhCyLS-2

The recombinant His-tagged FhCyLS-2 was purified from the concentrated and desalted expression medium by Ni-affinity chromatography on a HiTrap IMAC HP column (GE Healthcare) using the manufacturer's protocol for elution with imidazole at pH 8. It was followed

by size-exclusion chromatography, using a HiLoad 16/600 Superdex 75 pg column (GE Healthcare) equilibrated with 50 mM Tris-HCl, 200 mM NaCl, pH 8. The purified protein was concentrated and buffer-exchanged into 10 mM Tris-HCl, 10 mM NaCl, pH 8, using an Amicon Ultra-3K centrifugal unit (Millipore). Purification was monitored by Laemmli SDS-PAGE on 15% polyacrylamide gels stained by Coomassie Brilliant Blue R-250. Edman sequencing of purified recombinant FhCyLS-2 demonstrated the N-terminal sequence starting with IRG, which indicated a proteolytic trimming and removal of the N-terminal dipeptide AG during protein expression.

Protein crystallization, data collection, and structure determination

Crystals were grown using vapor-diffusion setup in hanging drops at 18 °C. The ratio of protein to reservoir solution in the drops was 1:1. The drops were equilibrated over 300 µl of reservoir solution consisting of 2 M ammonium sulfate, 0.1 M HEPES, 10% PEG 400, pH 7.5. The protein concentration of the FhCyLS-2 solution was 17 mg/ml. Crystals shaped as thin needles were soaked overnight in 2 M ammonium sulfate, 0.1 M Na-iodide, 0.1 M HEPES, 10% PEG 400, pH 7.5 and flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K on the MX14.1 beamline at the BESSY electron-storage ring, Berlin, Germany (54). The anomalous data set was collected at a wavelength of 1.500 Å; the native data set was collected at a wavelength of 0.9184 Å. Diffraction data were integrated, reduced, and scaled using XDS integration package (55). Phasing was performed according to the SIRAS method (56), using anomalous and native datasets with help of hkl2map (57) and SHELX (58). The initial model was improved and rebuilt by Buccaneer (59), refined by REFMAC5 (60), manually rebuild in Coot (61) and validated with MolProbity (62), all programs originating from the CCP4 package (63). The structure was refined at 1.60 Å resolution, with $R_{\text{work}} = 0.22$ and $R_{\text{free}} = 0.26$. The effective resolution d_{eff} of the structure calculated using $d_{\text{eff}} = d_{\text{min}}C^{-1/3}$ (64) was 1.61 Å, and the resolution at the mean $I/\sigma(I) = 2$ was 1.74 Å. Figures showing structural representations were generated by PyMOL (65). Table S3 was generated by PHENIX (66). The atomic coordinates and experimental structure factors were deposited in the Protein Data Bank with accession code 6I1M.

Protease inhibition assays

Inhibition measurements were performed in triplicate in 96-well microplates (100 µL assay volume) at 37 °C. Recombinant FhCyLS-2 was 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) 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 substrate). IC₅₀ values were determined from residual velocities using dose-response plots; nonlinear regression was fitted using GraFit software (Erithacus, East Grinstead, UK). All substrates were purchased from Bachem with the exception of Abz-Phe-Arg-Phe(NO₂)-OH from MP Biomedicals. Human cathepsins L, K, and V were purchased from Merck, cathepsins F, B, and S from Enzo Life Science, and cathepsin X and legumain from R&D Systems. *Carica papaya* papain was from Merck, bovine cathepsins C and H were prepared as described in (67,68), and *F. hepatica* cathepsins L1, L2, and L3 as described in (69,70). Substrates and proteases were applied in assays as follows: Z-Phe-Arg-AMC in 2.5 μM concentration 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, 20 μM with 2 nM FhCL1, and 20 μM with 5 nM FhCL2; 17 μM Z-Val-Val-Arg-AMC with 0.1 nM cathepsin S; 100 μM H-Arg-AMC with 1.9 nM cathepsin H; 87 μM Abz-Phe-Arg-Phe(NO₂)-OH with 1.2 nM cathepsin X; 130 μM H-Gly-Arg-AMC with 0.4 nM cathepsin C; 70 μM Z-Ala-Ala-Asn-AMC with 1.2 nM legumain; and 20 μM Z-Gly-Pro-Arg-AMC with 6 nM FhL3. The assay buffers were as follows: 100 mM Na-acetate, pH 5.0 (for cathepsin X, legumain, and E/S products) or pH 5.5 (for cathepsins L, K, V, B, F, C, and papain); 50 mM MES, pH 5.5 (FhCL1 and FhCL3) or pH 6.0 (FhCL2) or pH 6.5 (for cathepsins S and H); the buffers contained 2.5 mM dithiothreitol, 0.1% polyethylene glycol 1500, and 50 mM NaCl (for cathepsin C).

Immune assays

Production of anti-FhCyLS-2 serum: Polyclonal serum was raised in three 12 weeks-old NMRI mice initially immunized with 60 μg recombinant FhCyLS-2 in Freund's complete adjuvant (100 μL s.c.) followed by a single booster injection with the same amount of antigen in Freund's incomplete adjuvant after three weeks. The collected serum was tested in ELISA for immunoreactivity against recombinant FhCyLS-2, as described for sheep serum below, with the only difference being the use of mouse serum (1:200) and anti-mouse-IgG for detection (1:10000, A3562, Sigma) and in western blot for specificity against the *F. hepatica* E/S products and total worm extract containing native FhCyLS-2.

Immunofluorescence microscopy: Adult *F. hepatica* worms were prepared for immunofluorescence microscopy as described previously (71,72). Antibody dilutions were

1:500 for mouse polyclonal anti-FhCyLS-2 serum and 1:200 for anti-mouse IgG Alexa 633-conjugated secondary antibody (A21052, Molecular Probes). Antibodies were incubated for 45 min at 25 °C on sections with 3 washes between the primary and secondary antibody and 4 times after secondary antibody; the third wash contained 1 µg/ml DAPI. The sections embedded in mowiol were visualized using a Leica SP2 AOBS confocal laser scanning microscope (Leica Microsystems) using a 20x oil immersion objective, and images were processed using the Huygens deconvolution software package version 2.7 (Scientific Volume Imaging).

Immunoreactivity of sera from infected animals: Sheep were assigned to groups 1 and 2, consisting of 4 and 5 animals, respectively, that were experimentally infected with 150 metacercariae in a gelatin capsule administered orally. Prior to infection, group 2 was immunized with ovalbumin (300 µg) as an additional and irrelevant immunogen three times at two-weekly intervals. The infection was allowed to proceed for 5 weeks, and sera were collected. Sera from cattle (10 animals) were collected at abattoir, and bile was taken from gall bladders and examined for the presence of *F. hepatica* eggs (73).

ELISA assay: Microtiter plates were coated with recombinant FhCyLS-2 (0.5 µg per well), blocked with PBS, containing 2% dry milk and 0.05% Tween-20. After 3 washes with PBS containing 0.05% Tween-20 (PBS-T), the antigen was incubated with sera (1:200 sheep, 1:100 cattle) for 1 h at 22 °C. Unbound antibody was removed by washing 4 times with PBS-T, and the plate was incubated with alkaline phosphatase-conjugated secondary antibodies (donkey anti-sheep IgG 1:10000, A24564, Invitrogen or goat anti-bovine IgG 1:2000, 5220-0383, KPL-SeraCare) for 1 h at 22 °C. The plate was washed 4 times with PBS-T and incubated with the substrate (pNPP disodium hexahydrate, Sigma) for 15 min and read at 405 nm (with reference at 630 nm). All measurements were performed in duplicate. Statistical significance was determined with the two-sided Mann-Whitney U test.

Western blot: For testing mice sera, protein samples (10 µg of *F. hepatica* E/S products and adult worm lysate, or 1 µg of FhCyLS-2) were separated by reducing Laemmli SDS-PAGE and electroblotted onto a nitrocellulose membrane. The membrane was blocked with PBS-T, containing 5% dry milk, and incubated with polyclonal serum (1:200) in blocking solution for 4 h at 22 °C. After 6 washes with PBS-T, bound antibodies were detected with horseradish peroxidase-conjugated secondary antibodies (goat anti-mouse IgG, 1:20000, 1706516, Bio-Rad)

for 2 h at 22 °C. Signal was visualized by Western Lightning Chemiluminescence Reagent (PerkinElmer Life Sciences).

MS proteomic analysis

For identification of native FhCyLS-2, proteins of the E/S products from adult *F. hepatica* were reduced, alkylated, and digested with trypsin and chymotrypsin. Recombinant FhCyLS-2 was analyzed analogously using trypsin digest. The LC-MS/MS analysis of the digests was performed on an UltiMate 3000 RSLCnano system (Dionex) coupled to a TripleTOF 5600 mass spectrometer with a NanoSpray III source (AB Sciex). The peptides were separated on an Acclaim PepMap100 analytical column (3 µm, 15 cm × 75 µm; Thermo Scientific) with gradient elution in a 0.1% formic acid/acetonitrile system. Full MS scans were recorded from 350 to 1250 m/z, up to 25 candidate ions per cycle were subjected to fragmentation; in MS/MS mode the fragmentation spectra were acquired from 100 to 1600 m/z. The mass data were processed by the ProteinPilot 4.5 software (AB Sciex).

For disulfide pairing analysis, the recombinant FhCyLS-2 was digested with trypsin and subsequently with GluC. The LC-MS/MS analysis of the digests was performed using the same setup as described above. A peptide of molecular weight 3769.656 in the form of multiply charged (6+) ion (629.286) was detected and its MS/MS spectrum was recorded (Table S2). This peptide matches the sequence of a cluster of three peptides (LLTAGSVVSSC^{39E}, VSGGATC⁶⁹PGC⁶⁹WE, VSC⁹⁵TAAAHHHHHH) connected by two disulfide bonds (Table S1). The MS/MS fragments with m/z 1016.932 (doubly charged) and 678.291 (triply charged) demonstrated the breakage of the parental cluster forming a cluster of two peptides (PGC⁶⁹WE, VSC⁹⁵TAAAHHHHHH) connected by the disulfide Cys69-Cys95. This indicated that the disulfide connectivity in the original peptide cluster was provided by disulfide bonds of Cys69-Cys95 and Cys39-Cys66.

Phylogenetic and bioinformatic analyses

The dataset used for the phylogenetic analysis consisted of 160 amino acid sequences of vertebrate and invertebrate single-domain cystatins. It includes the recently reported dataset of the superfamily members (30) and newly mined helminth homologues. Sequences were aligned in MAFFT version 7.450 (74) implemented in Geneious Prime version 2019.0.4 (75) using the G-INS-i method, with default settings for gap opening penalty and offset value. The sequences

were retrieved as GenBank/WormBase Parasite annotated entries or were mined from the publicly available genome/ transcriptome assemblies using the tBLASTn algorithm and E-value cut off 10^{-5} (File S1). Multidomain cystatins (listed for helminths in File S1) were not included in the analysis due to difficulties with their alignment to single-domain molecules and classification of their secretion status. Non-homologous terminal regions of the alignment were removed manually in Geneious Prime so the final alignment comprised 164 positions, principally corresponding to the conserved cystatin domain Pfam PF00031. The metamonad *Giardia intestinalis* cystatin (GenBank: XP_001705664) resembling the most ancestral eukaryotic cystatin (6) was used as the outgroup. The phylogenetic tree was reconstructed by the maximum likelihood (ML) method in IQ-TREE version 1.6.12 (76) using the WAG+F+G4 protein model selected by ModelFinder (77). Nodal supports were based on 1000 bootstrap replicates. The visualization and graphics of the resulting tree was performed with the FigTree version 1.4.2. The presence/absence of signal peptide in the cystatin sequences was predicted in SignalP version 5.0 (52).

Data availability

3D structure of FhCyL2-2 obtained by X-ray crystallography was deposited in the PDB database the accession code 6I1M.

Supporting information — This article contains supporting information.

Acknowledgements — The authors thank John P. Dalton for providing Fasciola proteases, Martin Hubálek for technical assistance with mass spectrometry analysis and John D. Brooker for language editing. Diffraction data were collected on BL14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin.

Funding and additional information — The study was supported by the European Regional Development Fund (ChemBioDrug project CZ.02.1.01/0.0/0.0/16_019/0000729), grant LTAUSA19109 from the Ministry of Education, Youth and Sports of the Czech Republic, institutional project RVO 61388963, Czech Science Foundation grants 19-17269S (to M.H.) and 21-16565S (to P.B.S.), and by Forschungskredit der Universität Zürich grant FKUZH 55080505 (to S.S.).

References

1. Mehmood, K., Zhang, H., Sabir, A. J., Abbas, R. Z., Ijaz, M., Durrani, A. Z., Saleem, M. H., Ur Rehman, M., Iqbal, M. K., Wang, Y., Ahmad, H. I., Abbas, T., Hussain, R., Ghori, M. T., Ali, S., Khan, A. U., and Li, J. (2017) A review on epidemiology, global prevalence and economical losses of fasciolosis in ruminants. *Microb Pathog* **109**, 253-262
2. Castro-Hermida, J. A., Gonzalez-Warleta, M., Martinez-Sernandez, V., Ubeira, F. M., and Mezo, M. (2021) Current Challenges for Fasciolicide Treatment in Ruminant Livestock. *Trends Parasitol* **37**, 430-444
3. 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
4. Alvarez-Fernandez, M., Barrett, A. J., Gerhartz, B., Dando, P. M., Ni, J. A., and Abrahamson, M. (1999) Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. *Journal of Biological Chemistry* **274**, 19195-19203
5. Turk, V., Stoka, V., and Turk, D. (2008) Cystatins: biochemical and structural properties, and medical relevance. *Front Biosci* **13**, 5406-5420
6. Kordis, D., and Turk, V. (2009) Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. *Bmc Evol Biol* **9**
7. Stoka, V., Turk, V., and Turk, B. (2016) Lysosomal cathepsins and their regulation in aging and neurodegeneration. *Ageing Res Rev* **32**, 22-37
8. Magister, S., and Kos, J. (2013) Cystatins in immune system. *J Cancer* **4**, 45-56
9. Zavasnik-Bergant, T. (2008) Cystatin protease inhibitors and immune functions. *Front Biosci* **13**, 4625-4637
10. Kotal, J., Busa, M., Urbanova, V., Rezacova, P., Chmelar, J., Langhansova, H., Sojka, D., Mares, M., and Kotsyfakis, M. (2021) Mialostatin, a Novel Midgut Cystatin from Ixodes ricinus Ticks: Crystal Structure and Regulation of Host Blood Digestion. *Int J Mol Sci* **22**, 5371
11. Chmelar, J., Kotal, J., Langhansova, H., and Kotsyfakis, M. (2017) Protease Inhibitors in Tick Saliva: The Role of Serpins and Cystatins in Tick-host-Pathogen Interaction. *Front Cell Infect Microbiol* **7**, 216
12. Manoury, B., Gregory, W. F., Maizels, R. M., and Watts, C. (2001) Bm-CPI-2, a cystatin homolog secreted by the filarial parasite *Brugia malayi*, inhibits class II MHC-restricted antigen processing. *Curr Biol* **11**, 447-451
13. 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
14. Dainichi, T., Maekawa, Y., Ishii, K., Zhang, T., Nashed, B. F., Sakai, T., Takashima, M., and Himeno, K. (2001) Nippocystatin, a cysteine protease inhibitor from *Nippostrongylus brasiliensis*, inhibits antigen processing and modulates antigen-specific immune response. *Infect Immun* **69**, 7380-7386
15. Schnoeller, C., Rausch, S., Pillai, S., Avagyan, A., Wittig, B. M., Loddenkemper, C., Hamann, A., Hamelmann, E., Lucius, R., and Hartmann, S. (2008) A helminth immunomodulator reduces allergic and inflammatory responses by induction of IL-10-producing macrophages. *J Immunol* **180**, 4265-4272
16. Maizels, R. M., Smits, H. H., and McSorley, H. J. (2018) Modulation of Host Immunity by Helminths: The Expanding Repertoire of Parasite Effector Molecules. *Immunity* **49**, 801-818

17. Cwiklinski, K., Donnelly, S., Drysdale, O., Jewhurst, H., Smith, D., De Marco Verissimo, C., Pritsch, I. C., O'Neill, S., Dalton, J. P., and Robinson, M. W. (2019) The cathepsin-like cysteine peptidases of trematodes of the genus *Fasciola*. *Adv Parasitol* **104**, 113-164
18. McVeigh, P., Maule, A. G., Dalton, J. P., and Robinson, M. W. (2012) *Fasciola hepatica* virulence-associated cysteine peptidases: a systems biology perspective. *Microbes Infect* **14**, 301-310
19. Cancela, M., Acosta, D., Rinaldi, G., Silva, E., Duran, R., Roche, L., Zaha, A., Carmona, C., and Tort, J. F. (2008) A distinctive repertoire of cathepsins is expressed by juvenile invasive *Fasciola hepatica*. *Biochimie* **90**, 1461-1475
20. Law, R. H., Smooker, P. M., Irving, J. A., Piedrafita, D., Ponting, R., Kennedy, N. J., Whisstock, J. C., Pike, R. N., and Spithill, T. W. (2003) Cloning and expression of the major secreted cathepsin B-like protein from juvenile *Fasciola hepatica* and analysis of immunogenicity following liver fluke infection. *Infect Immun* **71**, 6921-6932
21. Cancela, M., Corvo, I., da Silva, E., Teichmann, A., Roche, L., Diaz, A., Tort, J. F., Ferreira, H. B., and Zaha, A. (2017) Functional characterization of single-domain cystatin-like cysteine proteinase inhibitors expressed by the trematode *Fasciola hepatica*. *Parasitology* **144**, 1695-1707
22. Khaznadji, E., Collins, P., Dalton, J. P., Bigot, Y., and Moire, N. (2005) A new multi-domain member of the cystatin superfamily expressed by *Fasciola hepatica*. *Int J Parasitol* **35**, 1115-1125
23. Smith, D., Cwiklinski, K., Jewhurst, H., Tikhonova, I. G., and Dalton, J. P. (2020) An atypical and functionally diverse family of Kunitz-type cysteine/serine proteinase inhibitors secreted by the helminth parasite *Fasciola hepatica*. *Sci Rep-Uk* **10**
24. Cwiklinski, K., Drysdale, O., Corrales, J. L., Corripio-Miyar, Y., Verissimo, C. D., Jewhurst, H., Smith, D., Lalor, R., McNeilly, T. N., and Dalton, J. P. (2022) Targeting Secreted Protease/Anti-Protease Balance as a Vaccine Strategy against the Helminth *Fasciola hepatica*. *Vaccines-Basel* **10**
25. Zhang, K., Liu, Y., Zhang, G., Wang, X., Li, Z., Shang, Y., Ning, C., Ji, C., Cai, X., Xia, X., Qiao, J., and Meng, Q. (2022) Molecular Characteristics and Potent Immunomodulatory Activity of *Fasciola hepatica* Cystatin. *Korean J Parasitol* **60**, 117-126
26. Siricoon, S., Grams, S. V., and Grams, R. (2012) Efficient inhibition of cathepsin B by a secreted type 1 cystatin of *Fasciola gigantica*. *Mol Biochem Parasit* **186**, 126-133
27. Geadkaew, A., Kosa, N., Siricoon, S., Grams, S. V., and Grams, R. (2014) A 170kDa multi-domain cystatin of *Fasciola gigantica* is active in the male reproductive system. *Mol Biochem Parasitol* **196**, 100-107
28. Park, S. Y., Jeong, M. S., Park, S. A., Ha, S. C., Na, B. K., and Jang, S. B. (2018) Structural basis of the cystein protease inhibitor *Clonorchis sinensis* Stefin-1. *Biochem Biophys Res Commun* **498**, 9-17
29. Mei, G. Q., Dong, J. M., Li, Z. T., Liu, S. L., Liu, Y. F., Sun, M. Z., Liu, G. Y., Su, Z., and Liu, J. S. (2014) Structural Basis for the Immunomodulatory Function of Cysteine Protease Inhibitor from Human Roundworm *Ascaris lumbricoides*. *Plos One* **9**
30. Bartosova-Sojkova, P., Kyslik, J., Alama-Bermejo, G., Hartigan, A., Atkinson, S. D., Bartholomew, J. L., Picard-Sanchez, A., Palenzuela, O., Faber, M. N., Holland, J. W., and Holzer, A. S. (2021) Evolutionary Analysis of Cystatins of Early-Emerging Metazoans Reveals a Novel Subtype in Parasitic Cnidarians. *Biology (Basel)* **10**
31. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J* **7**, 2593-2599
32. Rawlings, N. D., and Barrett, A. J. (1990) Evolution of proteins of the cystatin superfamily. *J Mol Evol* **30**, 60-71
33. Dall, E., Fegg, J. C., Briza, P., and Brandstetter, H. (2015) Structure and Mechanism of an Aspartimide-Dependent Peptide Ligase in Human Legumain. *Angew Chem Int Edit* **54**, 2917-2921

34. Robinson, P. J., and Bulleid, N. J. (2020) Mechanisms of Disulfide Bond Formation in Nascent Polypeptides Entering the Secretory Pathway. *Cells* **9**
35. Nanut, M. P., Fonovic, U. P., Jakos, T., and Kos, J. (2021) The Role of Cysteine Peptidases in Hematopoietic Stem Cell Differentiation and Modulation of Immune System Function. *Frontiers in Immunology* **12**
36. Olson, O. C., and Joyce, J. A. (2015) Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nat Rev Cancer* **15**, 712-729
37. Kotal, J., Stergiou, N., Busa, M., Chlastakova, A., Berankova, Z., Rezacova, P., Langhansova, H., Schwarz, A., Calvo, E., Kopecky, J., Mares, M., Schmitt, E., Chmelar, J., and Kotsyfakis, M. (2019) The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin. *Cell Mol Life Sci* **76**, 2003-2013
38. Kotsyfakis, M., Horka, H., Salat, J., and Andersen, J. F. (2010) 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
39. Salat, J., Paesen, G. C., Rezacova, P., Kotsyfakis, M., Kovarova, Z., Sanda, M., Majtan, J., Grunclova, L., Horka, H., Andersen, J. F., Brynda, J., Horn, M., Nunn, M. A., Kopacek, P., Kopecky, J., and Mares, M. (2010) Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick *Ornithodoros moubata*. *Biochem J* **429**, 103-112
40. Wang, Y., Zhou, Y., Gong, H., Cao, J., Zhang, H., Li, X., and Zhou, J. (2015) Functional characterization of a cystatin from the tick *Rhipicephalus haemaphysaloides*. *Parasit Vectors* **8**, 140
41. Abrahamson, M., Barrett, A. J., Salvesen, G., and Grubb, A. (1986) Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. *J Biol Chem* **261**, 11282-11289
42. Barrett, A. J., Rawlings, N. D., Davies, M. E., Machleidt, W., Salvesen, G., and Turk, V. (1986) *Cysteine proteinase inhibitors of the cystatin superfamily*, Elsevier, Amsterdam
43. Abrahamson, M. (1994) Cystatins. *Methods Enzymol* **244**, 685-700
44. Ni, J., Fernandez, M. A., Danielsson, L., Chillakuru, R. A., Zhang, J. L., Grubb, A., Su, J., Gentz, R., and Abrahamson, M. (1998) Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor. *Journal of Biological Chemistry* **273**, 24797-24804
45. Langerholc, T., Zavasnik-Bergant, V., Turk, B., Turk, V., Abrahamson, M., and Kos, J. (2005) Inhibitory properties of cystatin F and its localization in U937 promonocyte cells. *FEBS J* **272**, 1535-1545
46. Cribb, T. H., Bray, R. A., Olson, P. D., and Littlewood, D. T. (2003) Life cycle evolution in the digenea: a new perspective from phylogeny. *Adv Parasitol* **54**, 197-254
47. Netea, M. G., Schlitzer, A., Placek, K., Joosten, L. A. B., and Schultze, J. L. (2019) Innate and Adaptive Immune Memory: an Evolutionary Continuum in the Host's Response to Pathogens. *Cell Host Microbe* **25**, 13-26
48. Stefanic, S., Dvorak, J., Horn, M., Braschi, S., Sojka, D., Ruelas, D. S., Suzuki, B., Lim, K. C., Hopkins, S. D., McKerrow, J. H., and Caffrey, C. R. (2010) RNA Interference in *Schistosoma mansoni* Schistosomula: Selectivity, Sensitivity and Operation for Larger-Scale Screening. *Plos Neglect Trop D* **4**
49. Stefanic, S., Dvorak, J., Horn, M., Braschi, S., Sojka, D., Ruelas, D. S., Suzuki, B., Lim, K.-C., Hopkins, S. D., McKerrow, J. H., and Caffrey, C. R. (2010) RNA Interference in *Schistosoma mansoni* Schistosomula: Selectivity, Sensitivity and Operation for Larger-Scale Screening. *PLoS Negl. Trop. Dis.* **4**

50. Horn, M., Fajtova, P., Arreola, L. R., Ulrychova, L., Bartosova-Sojkova, P., Franta, Z., Protasio, A. V., Opavsky, D., Vondrasek, J., McKerrow, J. H., Mares, M., Caffrey, C. R., and Dvorak, J. (2014) Trypsin- and Chymotrypsin-Like Serine Proteases in *Schistosoma mansoni* - 'The Undiscovered Country'. *PLoS Negl. Trop. Dis.* **8**
51. Livak, K. J., and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408
52. Almagro Armenteros, J. J., Tsirigos, K. D., Sonderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420-423
53. Jilkova, A., Horn, M., and Mares, M. (2020) Structural and Functional Characterization of *Schistosoma mansoni* Cathepsin B1. *Methods Mol Biol* **2151**, 145-158
54. Mueller, U., Darowski, N., Fuchs, M. R., Forster, R., Hellmig, M., Paithankar, K. S., Puhlinger, S., Steffien, M., Zocher, G., and Weiss, M. S. (2012) Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J Synchrotron Radiat* **19**, 442-449
55. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* **66**, 133-144
56. Dauter, Z., Dauter, M., and Rajashankar, K. R. (2000) Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallogr D Biol Crystallogr* **56**, 232-237
57. Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J Appl Crystallogr* **37**, 843-844
58. Sheldrick, G. M. (2008) A short history of SHELX. *Acta Crystallogr A* **64**, 112-122
59. Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* **62**, 1002-1011
60. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367
61. Moriarty, N. W., Liebschner, D., Klei, H. E., Echols, N., Afonine, P. V., Headd, J. J., Poon, B. K., and Adams, P. D. (2018) Interactive comparison and remediation of collections of macromolecular structures. *Protein Sci* **27**, 182-194
62. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., and Richardson, D. C. (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315
63. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242
64. Weiss, M. (2001) Global indicators of X-ray data quality. *J Appl Crystallogr* **34**, 130-135
65. Janson, G., Zhang, C., Prado, M. G., and Paiardini, A. (2017) PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL. *Bioinformatics* **33**, 444-446
66. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J., and Adams, P. D. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D* **75**, 861-877

67. Horn, M., Baudys, M., Voburka, Z., Kluch, I., Vondrasek, J., and Mares, M. (2002) Free-thiol Cys331 exposed during activation process is critical for native tetramer structure of cathepsin C (dipeptidyl peptidase I). *Protein Sci.* **11**, 933-943
68. Horn, M., Doleckova-Maresova, L., Rulisek, L., Masa, M., Vasiljeva, O., Turk, B., Gan-Erdene, T., Baudys, M., and Mares, M. (2005) Activation processing of cathepsin H impairs recognition by its propeptide. *Biol. Chem.* **386**, 941-947
69. Stack, C. M., Caffrey, C. R., Donnelly, S. M., Seshadri, A., Lowther, J., Tort, J. F., Collins, P. R., Robinson, M. W., Xu, W., McKerrow, J. H., Craik, C. S., Geiger, S. R., Marion, R., Brinen, L. S., and Dalton, J. P. (2008) Structural and functional relationships in the virulence-associated cathepsin L proteases of the parasitic liver fluke, *Fasciola hepatica*. *J Biol Chem* **283**, 9896-9908
70. Corvo, I., Cancela, M., Cappetta, M., Pi-Denis, N., Tort, J. F., and Roche, L. (2009) The major cathepsin L secreted by the invasive juvenile *Fasciola hepatica* prefers proline in the S2 subsite and can cleave collagen. *Mol Biochem Parasitol* **167**, 41-47
71. Martini, F., Eckmair, B., Stefanic, S., Jin, C., Garg, M., Yan, S., Jimenez-Castells, C., Hykollari, A., Neupert, C., Venco, L., Varon Silva, D., Wilson, I. B. H., and Paschinger, K. (2019) Highly modified and immunoactive N-glycans of the canine heartworm. *Nat Commun* **10**, 75
72. Fajtova, P., Stefanic, S., Hradilek, M., Dvorak, J., Vondrasek, J., Jilkova, A., Ulrychova, L., McKerrow, J. H., Caffrey, C. R., Mares, M., and Horn, M. (2015) Prolyl Oligopeptidase from the Blood Fluke *Schistosoma mansoni*: From Functional Analysis to Anti-schistosomal Inhibitors. *PLoS Negl Trop Dis* **9**, e0003827
73. Eichenberger, R. M., Lewis, F., Gabriel, S., Dorny, P., Torgerson, P. R., and Deplazes, P. (2013) Multi-test analysis and model-based estimation of the prevalence of *Taenia saginata* cysticercus infection in naturally infected dairy cows in the absence of a 'gold standard' reference test. *Int J Parasitol* **43**, 853-859
74. Katoh, K., Kuma, K., Miyata, T., and Toh, H. (2005) Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome Inform* **16**, 22-33
75. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649
76. Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., and Minh, B. Q. (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* **44**, W232-235
77. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587-589

Abbreviations:

NEJ - newly excysted juveniles
 RMSD - root mean square deviation
 E/S products - excretory/secretory products
 OCT - optimal cutting temperature medium.

Supporting Information

An evolutionary molecular adaptation of an unusual steffin from the liver fluke *Fasciola hepatica* redefines the cystatin superfamily

Michal Buša, Zuzana Matoušková, Pavla Bartošová-Sojková, Petr Páchl, Pavlína Řezáčová, Ramon Marc Eichenberger, Peter Deplazes, Martin Horn, Saša Štefanić, and Michael Mareš

Figure S1. Multiple sequence alignment of FhCyLS-2 with selected members of the cystatin superfamily from blood feeding parasites. The groups of sequences correspond to classical stefins from the type 1 family, true cystatins from the type 2 family, and cystatin-like stefins. The sequence of FhCyLS-2 is in magenta. The predicted signal peptides are highlighted in yellow, and cysteine residues in mature protein sequences are shaded black. The conserved C-terminal segment of classical stefins is green. Three segments N, L1, and L2 forming the reactive center for inhibition of cysteine cathepsins are indicated by red bars and labeled; the critical consensus positions are shown in red. The sequence length is indicated on the right (additional C-terminal extensions are marked with a cross). The sequences are indicated with an abbreviated species name (Av: *Acanthocheilonema viteae*, Bm: *Brugia malayi*, Cs: *Clonorchis sinensis*, Em: *Echinococcus multilocularis*, Eg: *Echinococcus granulosus*, Fg: *Fasciola gigantica*, Fh: *Fasciola hepatica*, Hl: *Haemaphysalis longicornis*, Ir: *Ixodes ricinus*, Is: *Ixodes scapularis*, Ov: *Onchocerca volvulus*, Of: *Opisthorchis felineus*, Om: *Ornithodoros moubata*, Rm: *Rhipicephalus microplus*, Sj: *Schistosoma japonicum*, Sm: *Schistoma mansoni*, Sp: *Sphaeridiotrema pseudoglobulus*, Ta: *Taenia asiatica*, Ts: *Taenia saginata*, To: *Taenia solium*, Tt: *Trichuris trichiura*, Tz: *Trichinella zimbabwensis*), phylogenetic affiliation in parentheses (a: arthropoda, n: nematoda, c: cestoda, t: trematoda), and GenBank accession number.

Table S1. Mass spectrometry analysis of posttranslational modifications of FhCyLS-2. Listed are selected peptides derived from the native and recombinant FhCyLS-2 that cover the mature N-terminal region (tryptic fragment 1–11 and chymotryptic fragment 8–21) and the regions containing disulfide-bridged cysteine residues. The peptides were analyzed by LC-MS/MS as described in Experimental procedures. Cysteine residues (red) forming a disulfide bond in the peptide cluster are aligned; the disulfide bonds connect residues Cys39 to Cys66 and Cys69 to Cys95. Delta indicates the difference between experimental and theoretical mass.

FhCyLS-2	Theoretical peptide mass [M] (Da)	Measured peptide mass [M] (Da)	Delta (Da)	Peptide sequence (residues No.)
Native	1250.597	1250.594	0.002	EMLVGGYTEPR (1-11)
Native	1632.811	1632.815	0.005	TEPRSVTPPEERSVF (8-21)
Recombinant	3769.656	3769.671	0.015	LLTAGSVVSSCE (29-40) VSGGATCPGCWE (59-71) VSC TAAAHHHHHH (93-105)

Table S2. Inhibitory effect of FhCyLS-2 on the proteolytic activity of the E/S products. The proteolytic activity of the E/S products of adult *F. hepatica* was assayed using the fluorogenic peptide substrate Z-Phe-Arg-AMC and three different chromogenic protein substrates. The inhibition by 0.5 μ M FhCyLS-2 is expressed as remaining activity relative to the uninhibited control (100%) and is compared with the effect of 0.5 μ M E-64, a general peptide inhibitor of papain-family proteases. The mean values \pm SE are given.

Substrate	Relative inhibition (%)	
	E-64	FhCyLS-2
Z-Phe-Arg-AMC	99.7 \pm 0.8	99.1 \pm 0.4
Azocasein	99.2 \pm 0.7	96.3 \pm 1.9
Azocollagen	99.4 \pm 4.1	93.9 \pm 0.8
Elastin congo red	98.4 \pm 1.2	98.2 \pm 3.0

Table S3. Crystal parameters, data collection statistics, and refinement statistics. All statistics were generated by PHENIX

Dataset	Native	I ⁺ soaked
Wavelength (Å)	0.918	1.5
Resolution range (Å)	46.67 - 1.60 (1.66 - 1.60)	33.6 - 1.90 (1.97 - 1.90)
Space group	C222 ₁	C222 ₁
a, b, c (Å)	64.9, 67.2, 44.5	64.7, 67.2, 44.6
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Total reflections	88303 (8518)	89387 (5006)
Unique reflections	13131 (1297)	7925 (727)
Multiplicity	6.7 (6.6)	11.3 (6.5)
Completeness (%)	98.43 (99.69)	97.97 (91.50)
Mean I/sigma(I)	7.83 (0.80)	17.81 (2.77)
Wilson B-factor (Å ²)	21.33	22.25
R _{merge}	0.16 (2.14)	0.11 (0.63)
R _{meas}	0.17 (2.32)	0.12 (0.69)
R _{pim}	0.07 (0.90)	0.03 (0.27)
CC _{1/2}	1.00 (0.43)	1.00 (0.91)
CC ⁺	1.00 (0.77)	1.00 (0.98)
Reflections used in refinement	12986 (1297)	
Reflections used for R _{free}	650 (65)	
R _{work}	0.22 (0.44)	
R _{free}	0.26 (0.36)	
CC _{work}	0.95 (0.41)	
CC _{free}	0.94 (0.24)	
Number of non-hydrogen atoms	807	
macromolecules	707	
ligands	2	
solvent	98	
protein residues	93	
RMS bond length (Å)	0.014	
RMS angle (°)	1.82	
Ramachandran favored (%)	97.8	
Ramachandran allowed (%)	1.1	
Ramachandran outliers (%)	1.1	
Rotamer outliers (%)	1.18	
Clashscore	3.5	
Average B-factor (Å ²)	28.47	
macromolecules (Å ²)	27.44	
ligands (Å ²)	25.70	
solvent (Å ²)	35.93	

Table S4. Primer sets used for qRT-PCR analysis.

Primer name	sequence 5'-3'
QFheCyLS2-A-s	CCCAGTGAGGTTGGATGACT
QFheCyLS-2-A-as	CAGTCTGCAATGACAGCACA
QFheCyLS-2-B-s	TGTTGCCGCACTACTTGGAA
QFheCyLS-2-B-as	AGCAACTCCAACACACGA
FheCoxl-A-s	TGGCCTGTTGTGACTGGTTA
FheCoxl-A-as	ACTACCTGCCGAGACAAGAG
FheCoxl-B-s	GGTAGTGTGTTGGGCTCA
FheCoxl-B-as	CCACACAACAGGATCCATA

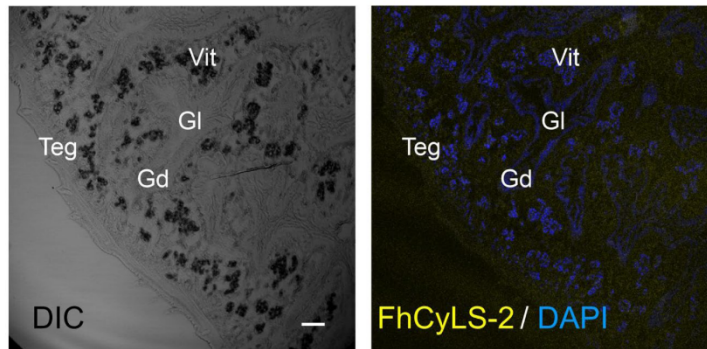


Figure S2. Negative control staining of section of adult *F. hepatica*. Semi-thin tissue section of the adult worm was probed with pre-immune mice serum followed by reaction with an anti-mouse IgG Alexa 633-labeled secondary antibody (yellow). DAPI was used to label nuclear DNA (blue). The left panel shows differential interference contrast (DIC), the right panel fluorescent channels. Positive staining with antiserum is presented in Figure 3. Gd: gastodermal cells, Gl: gut lumen, Teg: tegument, Vit: vitelline cells; scale bar represents 100 μ m.

Figure S3. Maximum likelihood phylogenetic tree of the cystatin superfamily. The rooted, non-collapsed tree of vertebrate and invertebrate members shows two clades of the type 1 family of stefins and the type 2 family of true cystatins; the former is divided into classical stefins and cystatin-like stefins. The highest taxonomic rank is indicated. FhCyLS-2 is in magenta. The *Giardia intestinalis* cystatin was used as the outgroup. Nodal supports calculated from 1000 bootstrap replicates are indicated.

Figure S3. Maximum likelihood phylogenetic tree of the cystatin superfamily. The rooted, non-collapsed tree of vertebrate and invertebrate members shows two clades of the type 1 family of stefins and the type 2 family of true cystatins; the former is divided into classical stefins and cystatin-like stefins. The highest taxonomic rank is indicated. FhCyLS-2 is in magenta. The *Giardia intestinalis* cystatin was used as the outgroup. Nodal supports calculated from 1000 bootstrap replicates are indicated.

4.2. Publikace č. 2: Mialostatin, a novel midgut cystatin from *Ixodes ricinus* ticks: crystal structure and regulation of host blood digestion.

Klíště obecné (*Ixodes ricinus*) je ektoparazit a hlavní vektor lymfské boreliózy a klíšťové encefalitidy v Evropě. Jako zdroj živin klíště využívá krevní proteiny hostitele, které jsou štěpené proteolytickou kaskádou, využívající cysteinové proteasy. Nicméně naše znalosti o regulaci těchto klíšťových trávicích enzymů jsou nedostatečné. Tato práce popisuje první střevní inhibitor cysteinových proteas z *I. ricinus* nazvaný mialostatin, přináší jeho první strukturní a funkční charakterizaci a navrhuje fyziologickou roli.

Mialostatin z rodiny cystatinů byl identifikován ve střevní tkáni *I. ricinus* a byl připraven jako rekombinantní protein. Kinetické inhibiční testy ukázaly, že je širokospektrým inhibitorem modelových i klíšťových cysteinových proteas a to jak s endopeptidasovou, tak exopeptidasovou aktivitou. Nejvýraznější inhibice (v subnanomolární oblasti koncentrací) byla určena pro dva trávicí katepsiny L z *I. ricinus* – IrCL1 a IrCL3. In vitro experimenty prokázaly stabilitu mialostatinu v prostředí trávicích proteas *I. ricinus* a jeho schopnost efektivně inhibovat degradaci hostitelských krevních proteinů.

V dalším kroku byly připraveny krystaly mialostatinu a vyřešena jeho 3D struktura rentgenovou difrakcí s finálním rozlišením 1,6 Å. Mialostatin má architekturu podrodiny pravých cystatinů (včetně páru konzervovaných disulfidů), která nese vazebné centrum tvořené třemi interakčními regiony: N-konec molekuly a povrchové smyčky L1 a L2. Analýza struktury ukázala plně kompetentní vazebné centrum odpovídající jeho inhibičnímu profilu. Mialostatin je prvním širokospektrým cystatinem z *I. ricinus*, který byl strukturně charakterizován. Jeho funkční vlastnosti ho výrazně odlišují od slinných cystatinů z *I. ricinus*, u kterých byla schopnost inhibovat exopeptidasu silně potlačena. To je v souladu i s fylogenetickou analýzou, která prokázala evoluční odlišnost mialostatinu od cystatinů ze slin klíšťat.

Expres mialostatinu i trávicích proteas je ve střevu klíštěte indukována během trávení hostitelské krve. Pomocí imunohistochemie byla prokázána lokalizace mialostatinu ve střevním epitelu a pomocí hmotnostní spektrometrie i v obsahu lumen střeva. Mialostatin se kolokalizoval především s trávicí proteasou IrCL3, která představuje zřejmě hlavní cílovou proteasu působící v pozdní fázi trávení. Lze tedy předpokládat roli mialostatinu při regulaci časových kroků ve fyziologii trávicího procesu.

Tato studie umožňuje nový pohled na mechanismy, které kontrolují trávicí proteolýzu klíšťat a mohou být využity při vývoji molekulárních vakcín proti klíšťatům.

Můj podíl na práci zahrnoval: (1) určení inhibičních konstant pro mialostatin pomocí kinetických testů s fluorescenčními substráty a modelovými a klíšťovými proteasami, (2) testování inhibičního efektu mialostatinu na nativní klíšťové proteasy v experimentech s hostitelskými proteinovými substráty, (3) krystalizaci mialostatinu a sběr difrakčních dat na synchrotronu, (4) řešení a analýzu krystalové struktury mialostatinu, (5) příprava manuskriptu.



Article

Mialostatin, a Novel Midgut Cystatin from *Ixodes ricinus* Ticks: Crystal Structure and Regulation of Host Blood Digestion

Jan Kotál ^{1,2,†} , Michal Buša ^{3,4,†} , Veronika Urbanová ¹, Pavlína Řezáčová ³, Jindřich Chmelář ², Helena Langhansová ² , Daniel Sojka ¹, Michael Mareš ^{3,*} and Michail Kotsyfakis ^{1,*}

- ¹ Institute of Parasitology, Biology Centre, Academy of Sciences of the Czech Republic, Branišovská 1160/31, 37005 České Budějovice, Czech Republic; jankotal@gmail.com (J.K.); veronika@paru.cas.cz (V.U.); sojka@paru.cas.cz (D.S.)
- ² Department of Medical Biology, Faculty of Science, University of South Bohemia in České Budějovice, Branišovská 1760c, 37005 České Budějovice, Czech Republic; chmelar@prf.jcu.cz (J.C.); hlanghansova@prf.jcu.cz (H.L.)
- ³ Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, 16610 Praha, Czech Republic; michal.busa@uochb.cas.cz (M.B.); rezacova@img.cas.cz (P.Ř.)
- ⁴ Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, 12800 Prague, Czech Republic
- * Correspondence: mares@uochb.cas.cz (M.M.); mich_kotsyfakis@yahoo.com (M.K.)
- † These authors contributed equally to this work.



Citation: Kotál, J.; Buša, M.; Urbanová, V.; Řezáčová, P.; Chmelář, 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. <https://doi.org/10.3390/ijms22105371>

Academic Editor: Andrea Battistoni

Received: 18 April 2021

Accepted: 17 May 2021

Published: 20 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The hard tick *Ixodes ricinus* is a vector of Lyme disease and tick-borne encephalitis. Host blood protein digestion, essential for tick development and reproduction, occurs in tick midgut digestive cells driven by cathepsin proteases. Little is known about the regulation of the digestive proteolytic machinery of *I. ricinus*. Here we characterize a novel cystatin-type protease inhibitor, mialostatin, from the *I. ricinus* midgut. Blood feeding rapidly induced mialostatin expression in the gut, which continued after tick detachment. Recombinant mialostatin inhibited a number of *I. ricinus* digestive cysteine cathepsins, with the greatest potency observed against cathepsin L isoforms, with which it co-localized in midgut digestive cells. The crystal structure of mialostatin was determined at 1.55 Å to explain its unique inhibitory specificity. Finally, mialostatin effectively blocked in vitro proteolysis of blood proteins by midgut cysteine cathepsins. Mialostatin is likely to be involved in the regulation of gut-associated proteolytic pathways, making midgut cystatins promising targets for tick control strategies.

Keywords: cathepsin; crystal structure; cysteine protease; digestion; *Ixodes ricinus*; midgut; parasite

1. Introduction

Ticks are globally distributed ectoparasitic arthropods that strictly feed on host blood. While soft ticks (family Argasidae) feed only for a few hours, hard ticks (family Ixodidae) usually attach to their hosts for several days to fully engorge and proceed to their next developmental stage. The hard tick *Ixodes ricinus* is found mainly in Europe but also in neighboring parts of Africa and the Middle East, where it is a major vector of pathogens such as Lyme disease spirochetes (*Borrelia burgdorferi sensu lato*), tick-borne encephalitis virus [1] or *Babesia* spp. [2]. Adult *I. ricinus* females feed for 6–9 days on a vertebrate host to enlarge over 100 times in weight [3].

Since blood is a highly specific and sole source of nutrients for these ticks, they have adapted to efficiently process large amounts of host blood. Blood degrades in the acidic endolysosomes of digestive cells of the tick midgut. Gut lumen uptake of the two main blood constituents, albumin and hemoglobin, is facilitated by two different mechanisms [4]: albumin is taken up non-specifically by fluid-phase endocytosis, while hemoglobin is recognized by specific receptor-mediated endocytosis. Subsequently, albumin is directed to small acidic vesicles and hemoglobin to a population of large digestive vesicles [4]. Despite

these differences, both albumin and hemoglobin are cleaved and processed to single amino acids and short peptides by the same proteolytic system [5,6]. The degradation pathway for hemoglobin is described in detail elsewhere [4]. Briefly, the initial phase is catalyzed by three *I. ricinus* digestive endopeptidases at low pH (3.5 to 4.5) including a cysteine protease legumain (IrAE) [7] and aspartic protease cathepsin D (IrCD1) [8]. Two cysteine protease cathepsin L isoforms, IrCL1 [9] (GenBank: EF428205) and IrCL3 (GenBank: QBK51063), complement the initial phase: IrCL1 expression in tick gut cells peaks at the end of tick feeding [9], while its ortholog, IrCL3, is present in the tick midgut predominantly after feeding, where it complements the activity of IrCL1 (D. Sojka, personal communication, December 2020). Cysteine proteases with exopeptidase activity, cathepsins B and C (IrCB and IrCC), continue hemoglobin degradation to dipeptides at an optimal pH of 5.5–6.0 in digestive cells [6,10,11]. Digestion to single amino acids is facilitated by carboxypeptidase and leucine aminopeptidase [6]. Blood processing by ticks and the roles of individual proteases are reviewed in detail elsewhere [12,13].

Under physiological conditions, cysteine protease activity is regulated by proteinaceous inhibitors, including those in the cystatin family [14,15]. Cystatins are tight binding, reversible inhibitors of legumain and papain-like cysteine proteases [16]. According to MEROPS nomenclature, cystatins are subdivided into three subfamilies: I25A (type 1, stefins), I25B (type 2 and type 3, kininogens), and I25C (type 4, fetuins) [17]. Only type 1 and 2 cystatins have so far been identified in ticks [18]. Cystatins are mostly associated with the regulation of proteases involved in blood digestion and heme detoxification in the tick midgut [18] and with the modulation of the host immune system as components of tick saliva [19,20], although they have also been detected in other tick tissues [21,22].

In soft ticks, only two midgut cystatins have been functionally characterized: Om-cystatins 1 and 2 from *Ornithodoros moubata* [23]. While Om-cystatin 1 is exclusively expressed in the midgut, Om-cystatin 2 can be found in all tissues and has immunomodulatory properties when secreted into the host [24,25]. Both inhibit cathepsins B, C, and H and are involved in blood processing [23]. Gut-associated cystatins from only two *Ixodes* species have been reported to date: a gut-secreted cystatin Jplocys2a from *Ixodes ovatus* was shown to inhibit cathepsins B, C, and L [26], while the expression of three *Ixodes persulcatus* cystatins, JpIpcys2a, b, and c, was demonstrated in almost all tissues and instars [27].

Despite the relatively good characterization of the digestive proteases present in the *I. ricinus* midgut [13], there has been little functional characterization of their inhibitors and regulatory mechanisms. Here we report a novel cystatin from the *I. ricinus* midgut, mialostatin, and present its crystal structure, inhibitory specificity, tissue localization, and role in the regulation of blood digestion.

2. Results

2.1. Mialostatin Transcript Predominantly Accumulates in the Tick Midgut

In order to clone mialostatin, we used primers based on available cystatin sequences identified in *Ixodes scapularis* tick genome. To obtain the longest possible reads, we also focused on the 5' UTRs and 3' UTRs regions. In the course of our study, an *I. ricinus* transcriptome was published with a transcript of an identical sequence to mialostatin (Genbank accession number GFVZ01041806.1) [28]. However, since this particular transcript was obtained from whole body tick sequencing, we used BLAST to search for highly similar sequences in other transcriptomic studies to specifically localize its expression (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed on 20 May 2021). As a result we found a highly similar transcript SigP-158801 upregulated mainly in the tick midgut [29]; similarly to another transcript GCJO1026918.1 identified in a study focusing on the tick gut [30]. To verify the localization of mialostatin, we examined its expression in different tick tissues and feeding stages and confirmed its predominantly midgut expression.

Increased transcription of mialostatin over the feeding course implies an important role in tick metabolism [18]. Figure 1 shows the expression of mialostatin in the tick midgut, ovaries, and salivary glands before, during, and after feeding. Mialostatin transcript was

predominantly present in the tick midgut, where expression oscillated throughout feeding and after detachment but at consistently higher expression than the other examined tissues. The presence and upregulation of mialostatin transcript in tick salivary glands and ovaries were low, peaking in fully fed ticks and at the early phase of detachment at maximum levels of only 10–20% of midgut expression (Figure 1).

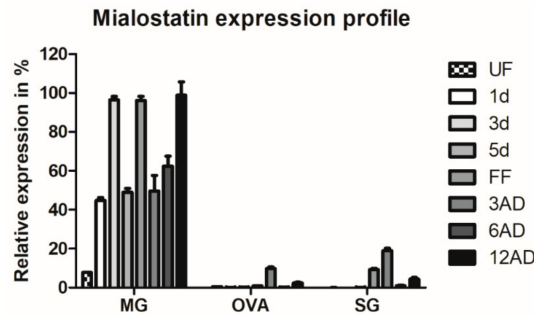


Figure 1. Mialostatin is predominantly produced in the tick midgut and its expression is upregulated by tick feeding. Expression maxima are prior to rapid engorgement, in fully fed females, and at two weeks post tick detachment from the host. Mialostatin expression was determined by quantitative PCR using cDNA templates prepared from a pool of three tissues from female ticks (MG—midgut, OVA—ovaries, SG—salivary glands). The qPCR output was normalized to the *I. ricinus* elongation factor 1 gene and compared across all values with the highest expression set to 100%. Data show an average of three biological replicates \pm SEM. Categories: UF—unfed ticks; 1d, 3d, 5d—ticks after 1, 3, or 5 days of feeding; FF—fully fed ticks after 7–8 days of feeding; 3AD, 6AD, 12AD—ticks 3, 6, or 12 days after detachment.

2.2. Mialostatin Is a Broad-Spectrum Inhibitor of Cysteine Cathepsins and Is Highly Effective against Cathepsin L

Purified recombinant mialostatin was screened *in vitro* for its inhibitory potential against major endogenous digestive proteases present in the *I. ricinus* gut [6]. These proteases were tested in the form of recombinant enzymes or proteolytic activities in the gut tissue extract (Table 1, left and middle panels). The strongest inhibition was found for recombinant *I. ricinus* cathepsins L1 and L3 (IC_{50} s of 0.071 and 0.39 nM, respectively), which are papain-type cysteine proteases and consistent with sub-nanomolar inhibition of cathepsin L-like activity by the extract (IC_{50} of 0.18 nM). The cathepsin B-like and cathepsin C-like activities of the extract were inhibited with lower potency, with IC_{50} values in the double-digit nanomolar range (12.1 and 91.7 nM, respectively). Mialostatin did not inhibit *I. ricinus* digestive proteases out of the papain family, including the aspartic protease cathepsin D1 and the clan CD cysteine protease legumain (asparaginyl endopeptidase).

We next expanded a spectrum of papain-type cysteine proteases and screened mialostatin against a representative panel of human cathepsins selected to cover a wide range of endo- and exopeptidase activities (Table 1, right panel) including endopeptidases cathepsins L, K, and S and exopeptidases cathepsin B (a peptidyl dipeptidase and endopeptidase), cathepsin C (a dipeptidyl peptidase), and cathepsin H (an aminopeptidase). Human cathepsin L was inhibited at subnanomolar concentrations (IC_{50} 0.38 nM), similar to its *I. ricinus* homologs, and all other human cysteine cathepsins were inhibited with IC_{50} values in a narrow range from 2.2 to 24 nM.

In conclusion, mialostatin displays an unusually broad inhibitory specificity against cysteine cathepsins, with a particularly strong interaction with cathepsin L isoforms. The high affinity for cysteine cathepsins with endopeptidase and exopeptidase activities clearly distinguishes mialostatin from other described *Ixodes* cystatins, which display weak or no inhibition of these exopeptidases (see below).

Table 1. Inhibitory effect of mialostatin on the activity of tick and human proteases. The inhibitory potency of mialostatin was determined against: (i) native cysteine cathepsins present in *I. ricinus* gut tissue extract using protease-specific assays (left panel); (ii) selected digestive proteases of *I. ricinus* prepared as recombinant proteins (middle panel); and (iii) a representative set of human cysteine cathepsins (right panel). The IC₅₀ values (mean values ± SE) were measured by kinetic activity assays using specific fluorogenic peptide substrates (for details, see Methods).

Inhibition of <i>I. ricinus</i> Midgut Homogenate		Inhibition of Recombinant Digestive <i>I. ricinus</i> Proteases		Inhibition of Human Cysteine Cathepsins	
Targeted Activity	IC ₅₀ (nM)	Proteases	IC ₅₀ (nM)	Protease	IC ₅₀ (nM)
Cathepsin L	0.18 ± 0.02	<i>Ir</i> -cathepsin L1 (IrCL1)	0.071 ± 0.01	<i>Hs</i> -cathepsin L	0.38 ± 0.03
Cathepsins L and B	3.1 ± 0.4	<i>Ir</i> -cathepsin L3 (IrCL3)	0.39 ± 0.18	<i>Hs</i> -cathepsin C	2.1 ± 0.8
Cathepsin B	12.1 ± 1.5	<i>Ir</i> -legumain (IrAE)	n.i.	<i>Hs</i> -cathepsin S	2.2 ± 0.4
Cathepsin C	91.7 ± 5.5	<i>Ir</i> -cathepsin D1 (IrCD1)	n.i.	<i>Hs</i> -cathepsin B	9.0 ± 0.3
				<i>Hs</i> -cathepsin K	9.7 ± 1.3
				<i>Hs</i> -cathepsin H	24.0 ± 3.5

Abbreviation: n.i.—no significant inhibition at 10 μM mialostatin concentration.

2.3. Mialostatin Is Present in the Tick Gut Wall and Lumen

We further investigated mialostatin's distribution within the tick midgut. Gut epithelia and lumina were collected from fully fed *I. ricinus* adult females and subjected to proteomic analysis to directly determine the presence or absence of mialostatin. The LC-MS/MS strategy was based on the enzymatic digestion of a complex protein mixture and MS/MS peptide sequencing. This analysis provided 11–71% peptide coverage of the mialostatin sequence and a mass accuracy of <5 p.p.m. (Table S1), allowing us to conclude that mialostatin is present in both the gut tissue and luminal contents of *I. ricinus* ticks.

Immunolabeling was used to evaluate the potential biological selectivity of mialostatin towards different papain-like enzymes present in tick gut tissue. Localization of mialostatin with IrCB, IrCL1, and IrCL3 was examined using multicolor immunohistochemistry (Figure 2), with the sample collection and section preparation time points selected based on qPCR-determined dynamic expression profiles of individual proteases (sixth day of feeding for IrCLB and IrCL1; eleventh day post feeding for IrCL3) to establish the availability of these proteases for co-localization with mialostatin at these timepoints. IrCL3 was the most probable target protease for mialostatin, as co-localization signals at the surface of large vesicles in tick gut cells (specific ring patterns) was nearly complete. However, there was also some co-localization of mialostatin with IrCL1 but not the cysteine protease cathepsin B (IrCB).

Immunoblot analyses of tick gut tissue were performed to (i) confirm mialostatin selectivity for IrCL3 and further evaluate potential interactions with IrCL1; and (ii) evaluate potential secretion of mialostatin and the cathepsin-L-like tick proteases into the gut lumen. The latter could not be observed by immunohistochemical labeling due to the rapid dilution of gut epithelial cell secretions with the large amount of imbibed host blood in the lumen. However, mialostatin was detected in the gut wall (Figure 3A) at all collected time-points during feeding. Gut tissue originating from ticks membrane-fed on pure bovine blood serum (without erythrocytes) [31] was used to avoid interference between mialostatin-specific signals and host hemoglobin proteins of identical molecular weight. IrCL1 and IrCL3 signals were also detected in both the gut epithelium (cell wall) and the gut lumen (Figure 3B). The multiple IrCL1 bands corresponded to the proenzyme and mature enzyme forms [9].

2.4. Mialostatin Inhibits Blood-Protein Digestion Catalyzed by Tick Gut Cysteine Cathepsins

In tick gut tissue, cysteine cathepsins play a critical role in the acidic degradation of the two most abundant host blood proteins, hemoglobin and serum albumin. In particular, cathepsin L is involved in the initial phase of the degradation pathway, which is continued by the action of cathepsins B and C [5,6]. We evaluated the effect of mialostatin on the in vitro degradation of hemoglobin and serum albumin by the proteolytically active extract

of *I. ricinus* gut tissue (limited to cysteine proteases by treatment with inhibitors of other protease classes). Both blood proteins were digested at optimal acidic pH, and SDS-PAGE analysis demonstrated highly efficient degradation of these substrates (Figure 4A,B). These processes were effectively blocked by mialostatin in a dose-dependent manner, with complete inhibition at a low nanomolar concentration of mialostatin. A similar effect was achieved by adding E-64, a small molecule general inhibitor of cysteine cathepsins. Further, we tested the stability of mialostatin exposed to the complex proteolytic environment of the gut tissue extract (Figure 4C), which revealed that mialostatin was generally stable and showed only partial degradation over long-term treatment. In summary, the blood protein digestion catalyzed by cysteine cathepsins of *I. ricinus* can be effectively controlled by mialostatin under native-like conditions.

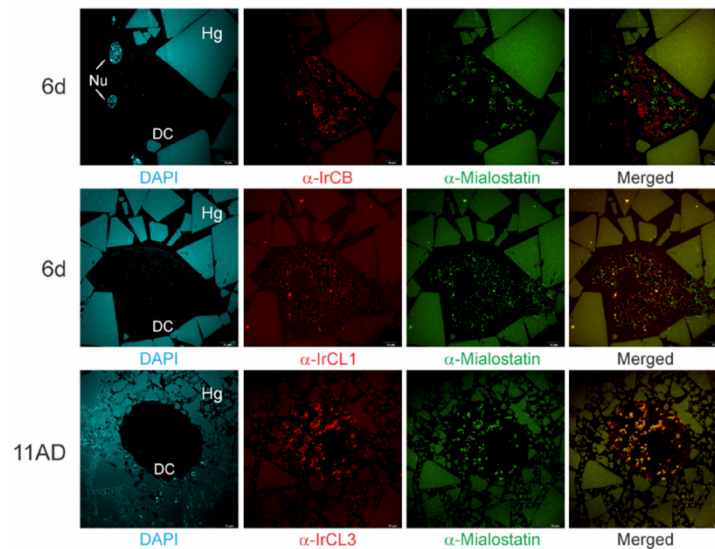


Figure 2. Mialostatin co-localizes with IrCL3 inside gut digestive cells of female *I. ricinus* ticks. Multicolored confocal immunofluorescence indicates variable colocalization of mialostatin (green signal) with the cathepsin-type proteases IrCB, IrCL1, and IrCL3 (red signal) in female tick gut sections at the sixth day of feeding (6d) and eleventh day post detachment from the host (11AD). Mialostatin and IrCL3 show the greatest co-localization (yellow signal in merged images), thus IrCL3 represents the most probable target protease. DAPI counterstaining is shown in cyan. DC—digestive cells; Nu—nucleus; Hg—hemoglobin crystals in gut lumen, scale bar—10 μ m.

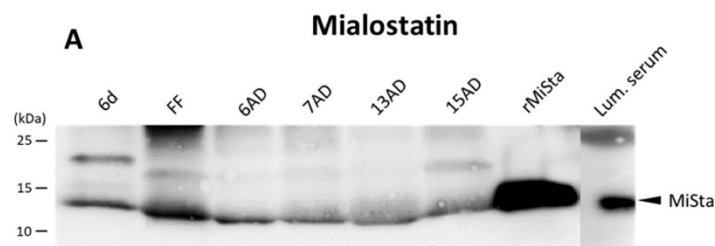


Figure 3. Cont.

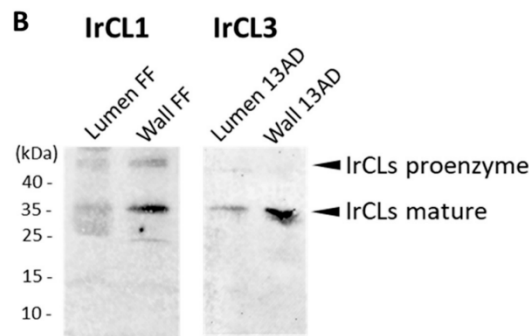


Figure 3. Western blot analysis of mialostatin, IrCL1, and IrCL3 in the tick midgut. (A) Tick midgut wall tissue extracts from various stages of tick feeding and midgut lumina from fully fed ticks were analyzed by SDS-PAGE and Western blotting. Mialostatin was labeled with a mouse monoclonal antibody and its signal detected using a fluorescently labeled secondary antibody. (B) Tick midgut wall and midgut lumen homogenates from fully fed ticks were analyzed by SDS-PAGE and immunoblotting with α -IrCL1 and α -IrCL3 rabbit polyclonal antibodies. Goat α -rabbit IgG Alexa 488 fluorescent secondary antibody was used to visualize protein bands using ChemiDoc MP imager. 6d—sixth day of feeding; FF—fully fed; 6, 7, 13, 15 AD—days post detachment from the host. Lum. serum—luminal fluid from ticks fed on erythrocyte-free serum; rMiSta—recombinant mialostatin. Full view of presented Western blots can be found in the Supplementary Materials, Figures S2–S4.

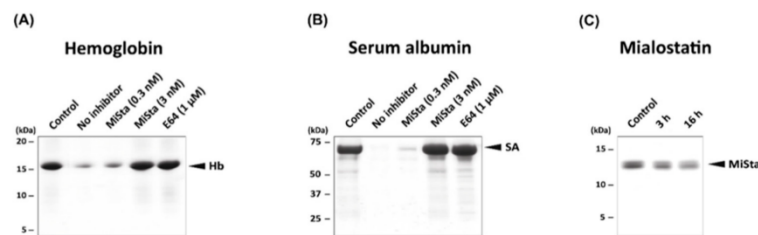


Figure 4. Blood protein digestion with tick gut proteases is inhibited by mialostatin. Hemoglobin (A) and serum albumin (B) were digested in vitro with *I. ricinus* gut tissue extracts in the presence and absence of mialostatin. Blood protein substrate (5 μ g of hemoglobin or 10 μ g of serum albumin) was incubated with 0.4 μ g gut tissue extract of cysteine proteases at pH 3.6 for 16 h. The extract was pre-incubated with mialostatin (MiSta) or the general cysteine protease inhibitor E-64 at the indicated concentrations prior to initiation of digestion. The digests were subjected to Tricine-SDS-PAGE (A) or Laemmli-SDS-PAGE (B) and visualized by protein staining. The hemoglobin (Hb) and serum albumin (SA) substrates are marked; the non-digested control is indicated. (C) Proteolytic stability of mialostatin in the gut tissue extract. Mialostatin (5 μ g) was incubated with 0.4 μ g gut extract protein under the same conditions as in (A,B), subjected to Tricine-SDS-PAGE, and visualized by protein staining. Mialostatin (MiSta) is marked; the non-digested control is indicated.

2.5. Phylogenetic Analysis and Three-Dimensional Structure of Mialostatin and Its Reactive Site

Phylogenetic analysis clearly demonstrated that mialostatin belongs to the cystatin superfamily. According to the maximum likelihood method, the tick cystatin phylogenetic tree contained three separate prostriate clades (Figure 5A and Figure S1). As shown in the simplified tree in Figure 5A, mialostatin fell into a strongly supported group with four other cystatins from the genus *Ixodes*. This clade was distant from other clades, including the recently described iristatin [32] and previously characterized sialostatins I.

Cys69–Cys82 and Cys93–Cys113 (yellow sticks), are indicated. The hairpin loops L1 and L2 and the N-terminus of cystatins are involved in the binding of papain-type cysteine proteases. (C) Superposition of C α traces of the mialostatin structure with five other cystatin structures including OmC2 from the soft tick *O. moubata* (PDB code 3L0R), iristatin from the hard tick *I. ricinus* (5O46), sialostatin L2 from the hard tick *I. scapularis* (3LH4), and representative vertebrate members of family 2 cystatins: chicken egg white (CEW) cystatin (1CEW) and human cystatin D (1RN7). The orientation of mialostatin is as in (B). Color coding of the structures and positions of the binding sites for papain-type cysteine proteases and legumains are indicated. (D) Structure-based sequence alignment of mialostatin with OmC2, iristatin, sialostatins L1 and L2, CEW cystatin, and human cystatin D. Residues identical to those of mialostatin are shaded grey. The secondary structural elements of mialostatin are depicted in magenta for β -strands and cyan for α -helices. The conserved disulfide bridges are indicated by the connecting black lines. Three regions involved in the interaction between cystatins and papain-type cysteine proteases are boxed in green and labeled (the region size was selected based on the predominant binding residues in the available complex structures); the consensus core residues are highlighted in red. The legumain binding site in CEW cystatin is highlighted in blue. Mature protein sequences were used in the alignment; residue numbering is according to mialostatin. (E) A comparison of the inhibitory potency of mialostatin with the other family 2 cystatins (shown in D) against various cysteine proteases including human papain-type cathepsins L to B (hCatL to hCatB) and mammalian legumains. IC₅₀ values are presented [23,32,33,35–37] and displayed as a heat map (green scale); n.i.—not inhibited; x—no literature data are available.

The crystal structure of mialostatin was determined by molecular replacement using the structure of the tick cystatin OmC2 as a search model and refined using data to 1.55 Å resolution (Table S2). The hexagonal prism crystal form contained two molecules in the asymmetric unit with a solvent content of about 57%. All protein residues could be modeled into a well-defined electron density map with the exception of the first nine residues, which formed a flexible N-terminus (Ser1 to Gly9), and the last two C-terminal residues (Asn118, Val119) of chain A. The final model consisted of two mialostatin molecules, chains A and B, containing 108 and 110 residues, respectively. The root-mean-square deviation (RMSD) for superposition of the C α atoms of the two chains was 0.14 Å, a low value within the range observed for different crystal structures of identical proteins.

Figure 5B shows the overall structure of mialostatin. The molecule adopts a typical cystatin fold (so called ‘hot dog’ fold [38]) characterized by a five-stranded twisted antiparallel β -sheet wrapped around a central α -helix. Mialostatin contains two conserved disulfide bridges connecting Cys69 with Cys82 and Cys93 with Cys113. Structural comparison and sequence alignment with other known cystatin structures clearly demonstrated that mialostatin belongs to family 2 of the cystatin superfamily (Figure 5C,D). The closest structural homolog of mialostatin was the salivary/gut cystatin OmC2 from the soft tick *O. moubata* [24] with the highest sequence identity (53%) and lowest RMSD for C α (0.86 Å), followed by salivary homologs iristatin (41% identity, 1.56 Å RMSD) from the hard tick *I. ricinus* and sialostatins L1 (42%, 2.09 Å) and L2 (40%, 2.50 Å) from the hard tick *I. scapularis* [32,33]. Lower structural similarity was found with vertebrate members, namely human cystatin D (35%, 4.80 Å) and chicken egg white (CEW) cystatin (23%, 3.80 Å) (Figure 5D) [39,40].

The interaction between family 2 cystatins and papain-type cysteine proteases is mediated by three regions, the N-terminal segment and two hairpin loops L1 and L2, which form a tripartite wedge-shaped edge that binds to the enzyme active site cleft (Figure 5B,C) [40–42]. In mialostatin, the first part of the binding site is formed by the N-terminal segment around Gly10, which is the first visible residue in the electron density map. The conserved pair of glycines (Gly9, Gly10) provide conformational flexibility to the N-terminal segment to adopt an optimal conformation for target binding. The L1 loop (between β 1 and β 2) of mialostatin exposes the segment Gln51-Ile52-Val53-Ala54-Gly55 corresponding to the critical binding motif Gln-Xaa-Val-Xaa-Gly conserved in cystatins (Figure 5B). The L2 loop (between β 3 and β 4) is characterized in mialostatin and other cystatins, except sialostatins, by the presence of a conservative Pro101-Trp102 segment. To

conclude, the structural analysis of mialostatin demonstrated a functionally competent reactive site against papain-type cysteine proteases. The binding motif for legumain-type cysteine proteases, which has been characterized in several cystatins (e.g., CEW cystatin), was absent in mialostatin (Figure 5C,D), consistent with the fact that mialostatin and other tick cystatins do not suppress legumain activity (Table 1, Figure 5E).

The inhibitory selectivity of the structurally analyzed cystatins is illustrated in Figure 5E. Mialostatin and OmC2 represent broad-spectrum inhibitors of papain-type cathepsins and are the most versatile in terms of their interactions of the analyzed tick cystatins. However, the other tick homologs displayed a distinct selectivity profile limited to effective inhibition of only some cathepsins. This may reflect structural changes in the conserved motifs on the L1 and L2 loops of iristatin and sialostatins, respectively, and in their N-terminal sequence potentially clashing with the partially occluded active sites of exopeptidases such as cathepsins B or H. Conversely, binding events to, for example, cathepsins B and C, can be supported by the electrostatic interactions formed by a positively charged basic patch (residues 12, 20, 106, 107) located at the reactive site of mialostatin and OmC2.

3. Discussion

Ixodes ricinus has previously been used as a model tick species to investigate and describe the complex intestinal digestive proteolytic mechanisms occurring in hematophagous arthropods. Blood proteins have been shown to be processed intracellularly by a multi-enzyme network of cysteine and aspartic proteases, with major involvement of cysteine cathepsin-type proteases from the CA clan [6]. However, previous studies have not investigated the regulation of digestive proteolysis, including the control mechanisms that protect the gut epithelium from excessive proteolysis and potential cell damage. Cystatins, naturally occurring cysteine protease inhibitors, are among the primary molecules of interest in the *I. ricinus* anti-proteolytic system, as they have been previously proposed to interact with digestive proteases in several other tick species [26,43,44].

In this study, we identified mialostatin as the first gut-associated cystatin to be identified in *I. ricinus* and present its comprehensive functional and structural characterization. Mialostatin was a potent inhibitor of *I. ricinus* digestive cysteine proteases of clan CA, covering both exopeptidases cathepsins B and C and endopeptidases cathepsins L1 and L3 (named IrCB, IrCC, IrCL1, and IrCL3, respectively). Its broad inhibitory selectivity is in clear contrast with *Ixodes* salivary cystatins such as sialostatins L1, L2, and iristatin, which have much narrower selectivity and mainly target endopeptidases [32–34]. On the other hand, similar broad anti-protease activity has been reported for OmC2 and partially also for OmC1 [23], cystatins present in the midgut of *O. moubata* soft ticks, or rBrBmcys2b from *Rhipicephalus microplus* [26] hard ticks. The 3D structural analysis identified mialostatin as a close homolog of OmC2 and provided a structural explanation for its binding selectivity through comparison of the architecture of the reactive site of mialostatin with other publicly available tick cystatin structures. Specifically, we highlight a combination of structural changes in three segments forming a tripartite wedge on mialostatin and OmC2 that slots into the cathepsin active site cleft. Based on structure-activity relationships and phylogenetic data, we propose that well-characterized mialostatin and OmC2 represent a new evolutionary subgroup of tick gut-associated cystatins that differ from salivary cystatins modulating host immune responses. Functional diversification of the cystatin superfamily is described in vertebrates [45]. It is likely that similar process occurs in ticks due to fast evolution of secreted proteins, therefore the phylogenetic tree reflects both localization and function of the cystatins. It is interesting to note that OmC2 also exhibits immunomodulatory properties, which correlate with its dual expression pattern in both the salivary glands and gut of *O. moubata* ticks, while OmC1 and mialostatin are expressed predominantly in tick midguts [23].

The biological role(s) of mialostatin in the tick gut can be inferred in the context of tick feeding behavior and associated physiological processes. Adult *I. ricinus* females engorge an enormous amount of host blood that exceeds the weight of the unfed tick

more than a hundred-fold. The current model of the multienzyme digestive protease network responsible for blood protein processing is based mainly on investigations of the well-developed digestive midgut cells occurring in partially engorged *I. ricinus* females at the end of the slow feeding period at day 6–7 [13]. This period is followed by a rapid engorgement phase lasting 12–24 h, which accounts for about two-thirds of the total blood volume ingested before detachment from the host. Most blood proteins are used for vitellogenesis and massive egg production during several weeks off-host [5,46]. The molecular mechanisms underlying the associated protein turnover and long-term blood meal storage in the tick gut lumen remain unexplored, mainly due to technical limitations in studying fully fed females. Nevertheless, advances in the field and initial results led to the hypothesis that off-host digestion may include extracellular proteolysis of blood proteins in the gut lumen, which supports or replaces intracellular digestive proteolysis in the gut epithelium [47]. Despite the broad biochemical selectivity of mialostatin, its biological selectivity is limited due to compartmentalization in tick midgut cells. Our immunohistochemistry results demonstrated that mialostatin is localized to the same population of intracellular vesicles as IrCL3 on the 11th day post detachment, suggesting that mialostatin predominantly targets IrCL3. Mialostatin is stored in these vesicles in some cells even during tick feeding. Forming an inhibitory complex between mialostatin and IrCL3 might be relevant for intracellular trafficking of enzymatically inactive IrCL3 in tick gut cells. The localization of the mialostatin-IrCL3 complex to the surface of the large dense granules two weeks post detachment is probably associated with an excretion/secretion mechanism allowing translocation of the complex to the gut lumen. We hypothesize that IrCL3 might partially restore its proteolytic activity in the diluted contents of the gut lumen, where mialostatin can competitively interact with other secreted cysteine cathepsins including IrCL1 as the strongest mialostatin binder. This would enable cathepsin-mediated luminal proteolysis of blood proteins or the generation of antimicrobial peptides under general mialostatin control [48]. Luminal IrCL3 might also act as an anti-coagulation factor, as recently reported for a related *R. microplus* cathepsin L [49].

In conclusion, mialostatin is the first gut-associated cystatin characterized from *I. ricinus* at the functional and structural levels. Mialostatin localized to both digestive cells and the gut lumen, where it targets cathepsin L isoforms and regulates their activity during trafficking and processing of host blood proteins. As components of gut-associated proteolytic pathways, mialostatin and homologous cystatins in other tick species represent potential vaccination antigens for novel anti-tick interventions targeting tick reproduction. The vaccination efficacy of proteins derived from the tick gut (“concealed” antigens) in controlling tick infestations has already been successfully demonstrated [50], and new candidate antigens are increasingly in demand to combat tick infestations and to limit the global spread of tick-borne diseases.

4. Materials and Methods

4.1. Ticks and Laboratory Animals

All animal experiments were carried out in accordance with the Animal Protection Law of the Czech Republic No. 246/1992 Sb., ethics approval No. 34/2018, and protocols approved by the responsible committee of the Institute of Parasitology, Biology Centre of the Czech Academy of Sciences. Male and female adult *I. ricinus* ticks were collected by flagging in a forest near České Budějovice in the Czech Republic and then kept in 95% humidity chambers under a 12 h light/dark cycle at room temperature. Female BALB/c mice were purchased from Velaz (Prague, Czech Republic). Mice were housed in individually ventilated cages maintained under a 12 h light/dark cycle. Mice were used at 8–12 weeks of age. Laboratory rabbits were purchased from RABBIT CZ a. s. (Trhový Štěpánov, Czech Republic) and housed individually in cages in the animal facility of the Institute of Parasitology. Guinea pigs were bred and housed in cages in the animal facility of the Institute of Parasitology. All mammals were fed a standard pellet diet and provided with water ad libitum.

4.2. Quantitative Real-Time PCR

Female *I. ricinus* ticks were fed on rabbits and allowed to mate with male ticks. Salivary glands, midguts, and ovaries from five ticks per time point were dissected on a petri dish under a drop of ice-cold DEPC-treated PBS. Total RNA was isolated from dissected tissue using the NucleoSpin RNA kit (Macherey-Nagel, Düren, Germany) and its quality checked by agarose gel electrophoresis before storing the RNA at -80°C . cDNA was prepared from 500 ng of total RNA from independent biological triplicates using the Transcriptor High-Fidelity cDNA Synthesis Kit (Roche Applied Science, Penzberg, Germany). The cDNAs served as templates for subsequent quantitative expression analyses of mialostatin transcription by qRT-PCR. Samples were analysed with a LightCycler 480 (Roche Applied Science, Penzberg, Germany) using FastStart Universal SYBR Green Master Mix (Roche Applied Science, Penzberg, Germany). Reaction conditions over 50 cycles were as follows: denaturation, $95^{\circ}\text{C}/10\text{ s}$; annealing, $60^{\circ}\text{C}/10\text{ s}$; extension, $72^{\circ}\text{C}/10\text{ s}$. Relative expression values were standardized to a reference gene, *I. ricinus* elongation factor 1 (*efl*; GenBank: GU074828) [51–53], and normalized to the sample with the highest level of expression. The primers sequences for mialostatin and *efl* RT-PCR are shown in Table S3.

4.3. Mialostatin Cloning, Expression, Refolding, and Purification and Antibody Production

The full cDNA sequence of the gene encoding mialostatin was amplified using primers designed based on the GFVZ01041806.1 [28] transcript from NCBI GenBank. The primer sequences used for the final cloning of mialostatin are presented in Table S3. A pool of *I. ricinus* cDNA prepared from the salivary glands of female ticks fed for three and six days on rabbits was used as a template. The 372 base pair DNA fragment encoding mialostatin without a signal peptide and with an inserted ATG codon was cloned into a pET-17b vector (Novagen, Darmstadt, Germany) and transformed into *Escherichia coli* strain BL21(DE3)pLysS (Novagen) for expression. Bacterial cultures were grown in LB medium with 100 $\mu\text{g}/\text{mL}$ ampicillin and 34 $\mu\text{g}/\text{mL}$ chloramphenicol to an OD₆₀₀ of 0.8, when protein expression was induced by the addition of isopropyl 1-thio- β -D-galactopyranoside to a final concentration of 1 mM. Cultures were harvested after 2 h of incubation at 37°C at 200 rpm shaking speed. Isolated inclusion bodies were dissolved in 6 M guanidine hydrochloride, 20 mM Tris, and 10 mM DTT, pH 8 for 1 h followed by centrifugation (10 min, $10,000\times g$) to remove undissolved impurities. Refolding was performed by rapid dilution in $160\times$ excess of 20 mM Tris and 300 mM NaCl, pH 8.5. The resulting refolded protein was purified by HiLoad Superdex 200 26/60 gel filtration chromatography and HiPrep Q FF 16/10 ion exchange chromatography. Endotoxin was removed using a detergent-based method. Purified recombinant mialostatin was used to raise antibodies in a mouse and rabbit as described previously [54,55]. The immunoglobulin (Ig) fraction of rabbit serum was obtained by caprylic acid precipitation of serum proteins as described previously [56]. Hybridoma cells were raised by fusing splenocytes from immunized mice and mouse myeloma SP 2/0-Ag14 cells. Monoclonal antibodies were produced in cell culture following the previously described protocol [55].

4.4. Preparation of Tick Gut Samples

I. ricinus midguts were dissected from female *I. ricinus* fed on laboratory guinea pigs (samples for proteolysis analysis and Western blotting) or from females' membrane fed on erythrocyte-depleted blood serum (samples for mass spectrometry analysis) [5]. The gut contents were carefully removed without disrupting the epithelium, and the gut tissue was washed in phosphate buffered saline (PBS). For mass spectrometry analysis, the gut contents were processed as described previously [6]. Gut tissue extract (150 mg protein/mL) was prepared by homogenization of the pooled gut tissue in 0.1 M Na acetate pH 4.5, 1% CHAPS on ice. The extract was cleared by centrifugation ($16,000\times g$, 10 min, 4°C), filtered through Ultrafree MC 0.22 μm (Millipore, Bedford, MA, USA), and stored at -80°C .

4.5. Protease Inhibition Assays

Inhibition measurements were performed in triplicate in 96-well microplates (100 μ L assay volume) at 37 °C. Recombinant mialostatin was preincubated with protease for 10 min followed by the addition of specific fluorogenic substrate (see Sections 4.5.1–4.5.3). The kinetics of product release were continuously monitored using an Infinite M1000 (Tecan, Männedorf, Switzerland) microplate fluorescence reader at 360 nm excitation and 465 nm emission wavelengths (for AMC-containing substrates) or at 320 nm excitation and 420 nm emission wavelengths (for Abz-containing substrate). IC₅₀ values were determined from residual velocities using dose-response plots; nonlinear regression was fitted using GraFit software (Erithacus, East Grinstead, UK).

4.5.1. Inhibition of Proteases in Tick Gut Homogenates

To prevent interference of non-target proteases, homogenates (80 ng) were treated with specific low molecular weight inhibitors (final assay concentrations are indicated) including 1 μ M pepstatin and 1 mM EDTA (against aspartic proteases and metalloproteases; all assays), 1 μ M E-64 (against cathepsins L/B; cathepsin C assay), 1 μ M CA-074 (against cathepsin B; cathepsin L assay), and 1 μ M Z-Phe-Phe-DMK (against cathepsin L; cathepsin B assay) [6]. The assay substrates and buffers were as follows: 20 μ M Z-Phe-Arg-AMC substrate and 0.1 M Na acetate pH 4.5 or 5.0 in the cathepsin L and L/B assays, respectively; 20 μ M Z-Arg-Arg-AMC substrate and 0.1 M MES pH 6.5 in the cathepsin B assay; 20 μ M Gly-Arg-AMC substrate in 0.1 M Na acetate pH 5.5, 25 mM NaCl in the cathepsin C assay; all assay buffers contained 2.5 mM DTT and 0.1% PEG 1500.

4.5.2. Inhibition Assays of Recombinant Tick Proteases

The assay conditions for individual proteases were as follows: 1.2 nM IrCD1 and 20 μ M Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu substrate in 0.1 M Na acetate pH 4.0; 1.25 nM IrAE and 20 μ M Z-Ala-Ala-Asn-AMC substrate in 0.1 M MES pH 5.0, 2.5 mM DTT, 1 μ M E-64; 0.1 nM IrCL1 or 20 pM IrCL3 and 20 μ M Z-Phe-Arg-AMC substrate in 0.1 M Na acetate pH 4.5, 2.5 mM DTT; all assay buffers contained 0.1% PEG 1500. The tick proteases were prepared as described elsewhere [7,9,10,57,58].

4.5.3. Inhibition Assays of Human Proteases

Inhibition assays were performed following the same protocol used in our previous publications [24,32]. The assay conditions for individual proteases were as follows: 35 pM cathepsin B or 33 pM cathepsin L or 5 nM cathepsin K and 250 μ M Z-Leu-Arg-AMC substrate in 0.1 M Na acetate pH 5.5, 0.1 M NaCl; 350 pM cathepsin S and 250 μ M Z-Val-Val-Arg-AMC substrate in the same buffer; 0.5 nM cathepsin C and 250 μ M Gly-Arg-AMC substrate in the same buffer; 20 nM cathepsin H and 40 μ M Z-Leu-Arg-AMC substrate in 0.1 M Na/K phosphate pH 6.8; all assay buffers contained 1 mM EDTA, 2.5 mM DTT, and 0.01% Triton X-100. The human proteases were purchased from Merck (Kenilworth, NJ, USA) and Biomol (Hamburg, Germany).

4.6. Protein Digestion Assay

Digestion of 10 μ g human serum albumin (Sigma Aldrich, St Louis, MO, USA), 5 μ g bovine hemoglobin (Sigma Aldrich, St Louis, MO, USA), and 5 μ g of mialostatin was performed with the tick gut tissue homogenate (0.4 μ g protein) in 50 mM Na citrate pH 3.6, 2.5 mM DTT, in a total volume of 100 μ L for 16 h at 26 °C. In the albumin and hemoglobin digestion assays, the homogenate was preincubated (15 min) in the same buffer with non-cysteine protease inhibitors: 1 μ M pepstatin, 100 μ M Pefablock, and 1 mM EDTA. The albumin digest was resolved with Laemmli SDS-PAGE gels (15%) and the hemoglobin and mialostatin digests by Tricine-SDS-PAGE gels (16% T/6% C) containing 6 M urea [59]. Electrophoresis was performed under reducing conditions, and protein was stained with Coomassie Blue G250.

4.7. Reducing SDS-PAGE and Western Blotting

Tick tissue homogenates were separated by reducing SDS-PAGE using 4–20% Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-Rad Laboratories, Hercules, CA, USA). Separated protein loads were visualized using TGX stain-free chemistry in the ChemiDoc MP imager (Bio-Rad, Hercules, CA, USA). After protein load documentation, separated proteins were electro-transferred from the gel onto an Immun-Blot® LF PVDF membrane using the Trans-Blot Turbo system (Bio-Rad, Hercules, CA, USA). Prior to Western blot analyses, membranes were blocked with 3% non-fat milk in PBS with 0.05% Tween 20 (PBS-Tween) for 1 h at room temperature. Blocked membranes were incubated with the rabbit Ig fraction of α -IrCL1 or α -IrCL3 polyclonal sera diluted 1:1000 in PBS-Tween containing 1% milk. Goat anti-rabbit IgG Alexa 488-labeled antibody (1:1000, Thermo Fisher Scientific, Waltham, MA, USA) was used as a secondary antibody. For mialostatin detection, α -mialostatin monoclonal antibody (1:30) diluted in in PBS-Tween containing 1% milk and the goat anti-mouse Alexa 546-labeled antibody (1:1000, Thermo Fisher Scientific, Waltham, MA, USA) were used. In between individual steps of the whole procedure, membranes were washed 3 \times 5 min in PBS-Tween on a rotating shaker platform at room temperature. Labeling with primary antibodies was performed on a rotating shaker platform at 4 °C overnight. Labeling with secondary antibodies was performed on a rotating shaker platform at room temperature for 1 h. Fluorescent signals were again visualized using the ChemiDoc MP imager and analyzed using Image Lab Software (Bio-Rad, Hercules, CA, USA).

4.8. Immunohistochemistry

Samples of *I. ricinus* gut tissues were prepared as described previously [11]. Briefly, the gut was dissected from adult females at specific days of feeding on the host and days post-attachment and fixed in 4% formaldehyde and 0.1% glutaraldehyde solution, washed with PBS, dehydrated using ascending ethanol dilutions, then infiltrated with LR White resin (London Resin Company, Stansted, UK) and polymerized. Semi-thin sections (0.5 μ m) were blocked with 1% BSA and 1% milk in PBS-Tween (0.3% (v/v) Tween 20) for 45 min. Immunohistochemical double-staining was performed gradually, with the initial antibody labeling of the respective intestinal protease (*I. ricinus* cathepsin L1 IrCL1 [9]; cathepsin L3 IrCL3; cathepsin B IrCB [11]) subsequently followed with immunolabeling of mialostatin. First, semi-thin tick gut tissue sections were blocked with blocking solution (1% BSA, 1% milk solution in PBS-Tween) for 45 min at room temperature. For protease immunostaining, sections were first labeled (4 °C overnight) with primary antibodies: (i) rabbit α -IrCL1 affinity-purified polyclonal serum diluted 1:5 in PBS-Tween; (ii) rabbit α -IrCB affinity-purified polyclonal serum diluted 1:5 in PBS-Tween; (iii) isolated Ig fraction of α -IrCL3 polyclonal serum diluted 1:5 (IrCL1) in PBS-Tween. After washing 3 \times 5 min in PBS-Tween, sections were subsequently labeled with Alexa Fluor® 647 goat α -rabbit secondary antibody (diluted 1:500 in PBS-Tween; Thermo Fisher Scientific, Waltham, MA, USA). Sections were subsequently used for mialostatin immunolabeling: sections were once again washed 3 \times 5 min in PBS-Tween and incubated with mouse α -mialostatin monoclonal antibody diluted 1:50 in PBS-Tween. Incubation was performed in a humid chamber at room temperature for 90 min. Sections were once again washed (3 \times 5 min in PBS-Tween) and incubated with secondary goat α -mouse Alexa Fluor® 488 (Thermo Fisher Scientific, Waltham, MA, USA) diluted 1:500 in PBS-Tween for 1 h at room temperature. Finally, all sections were washed in PBS-Tween and counterstained with DAPI (4',6'-diamidino-2-phenylindole; 2.5 μ g/mL; Sigma Aldrich, St Louis, MO, USA) for 7 min, washed again with PBS-Tween, mounted in Fluoromount medium (Sigma Aldrich, St Louis, MO, USA), and examined with the IX83 confocal microscope (Olympus, Tokyo, Japan). Images were processed with FluoView FV3000 software (Olympus, Tokyo, Japan).

4.9. Evolutionary Analysis by the Maximum Likelihood Method

The evolutionary history was inferred using the maximum likelihood method and JTT matrix-based model [60]. The bootstrap consensus tree inferred from 1000 replicates

was taken to represent the evolutionary history of the taxa analyzed [61]. Branches corresponding to partitions reproduced in less than 20% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [61]. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model and then selecting the topology with the superior log-likelihood value. This analysis involved 71 amino acid sequences. There were 108 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [62].

4.10. Crystallization and Data Collection

Screening for crystallization conditions was performed using the JCSG-plus kit (Molecular Dimensions, Sheffield, UK) by the sitting drop vapor diffusion technique. Preliminary crystals of mialostatin were obtained in 0.1 M citric acid pH 3.5, 0.8 M ammonium sulfate. Optimal crystals were prepared at 18 °C using the hanging drop vapor diffusion technique in 15-well NeXtal plates (Qiagen, Hilden, Germany). The crystallization drop consisted of 2 µL of the mialostatin protein solution (12.5 mg/mL in 10 mM Tris buffer, pH 8.0) and 1 µL of the precipitant solution equilibrated over a reservoir containing 300 µL precipitant solution (0.1 M citric acid pH 4.0, 0.8 M ammonium sulfate). Crystals shaped as hexagonal prisms reached their final size of 0.6 × 0.3 × 0.3 mm within 1 month. For data collection, crystals were soaked in reservoir solution supplemented with 20% glycerol and flash cooled in liquid nitrogen. Diffraction data at 100 K were collected using a BL14.1 beamline operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron storage ring (Berlin-Adlershof, Germany) [63] and processed using the XDS suite of programs [64]. Crystals exhibited the symmetry of space group P6222 and contained two molecules in the asymmetric unit. Crystal parameters and data collection statistics are shown in Table S2.

4.11. Structure Determination

The phase problem was solved by molecular replacement using Molrep [65] from the CCP4 package [66]. The search model was derived from the structure of cystatin OmC2 (PDB code 3LOR) [24] sharing 53% sequence identity with mialostatin. Model refinement was carried out using REFMAC 5.8 [66] from the CCP4 package with 5% of the reflections reserved for cross-validation. Manual building and addition of water molecules was performed using Coot [67]. The quality of the final model was validated with Molprobit [68]. Final refinement statistics are given in Supporting Information Table S2. Figures showing structural representations were prepared with the PyMOL Molecular Graphics System (Schrödinger, New York, NY, USA). Atomic coordinates and structure factors were deposited in the PDB under accession code 6ZTK.

4.12. Statistical Analysis

All experiments were performed in biological triplicate. Data are presented as mean ± standard error of mean (SEM) in all graphs. Student's t-test or one-way ANOVA were used to calculate statistical differences between two or more groups, respectively. Statistically significant results are marked: * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.0001$.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms22105371/s1>, Figure S1: The phylogenetic tree of 71 cystatins from both Ixodidae and Argasidae tick species, Figure S2: Detection of mialostatin in midgut of blood fed ticks, Figure S3: Detection of mialostatin in midgut of serum fed ticks, Figure S4: Detection of *I. ricinus* cathepsins L in midgut of blood fed ticks, Table S1: Identification of mialostatin by mass spectrometry, Table S2: X-ray data collection and refinement statistics, Table S3: Primer sequences.

Author Contributions: J.K., M.B., J.C., D.S., M.M. and M.K. designed and performed the experiments, performed the analyses, and wrote the manuscript; V.U., P.Ř. and H.L. designed and performed the experiments and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Grant Agency of the Czech Republic (No. 19-382 07247S to M.K.), the project CePaViP OPVVV (No. CZ.02.1.01/0.0/0.0/16_019/0000759 to M.K.) from the European Regional Development Fund and institutional project RVO 60077344 to M.K., M.B., P.Ř., and M.M. were supported by project ChemBioDrug CZ.02.1.01/0.0/0.0/16_019/0000729 from the European Regional Development Fund (OP RDE) and the institutional project RVO 61388963. M.M., D.S. and V.U. were supported by the Grant Agency of the Czech Republic (No. 21-08826S). J.C. was supported by the Grant Agency of the Czech Republic (No. 19-14704Y).

Institutional Review Board Statement: All animal experiments were carried out in accordance with the Animal Protection Law of the Czech Republic No. 246/1992 Sb., ethics approval No. 34/2018, and protocols approved by the responsible committee of the Institute of Parasitology, Biology Centre of the Czech Academy of Sciences.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are either contained within the manuscript and supporting information or available from the corresponding author on reasonable request.

Acknowledgments: The authors thank Jiří Brynda and Martin Hubálek for technical assistance with crystallography and mass spectrometry analysis, respectively. Diffraction data were collected on BL14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin. The funders had no role in the design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

References

- Lindgren, E.; Talleklint, L.; Polfeldt, T. Impact of climatic change on the northern latitude limit and population density of the disease-transmitting European tick *Ixodes ricinus*. *Environ. Health Perspect.* **2000**, *108*, 119–123. [\[CrossRef\]](#)
- Yang, Y.; Christie, J.; Köster, L.; Du, A.; Yao, C. Emerging Human Babesiosis with “Ground Zero” in North America. *Microorganisms* **2021**, *9*, 440. [\[CrossRef\]](#)
- Sonenshine, D.E.; Roe, R.M. *Biology of Ticks*, 2nd ed.; Oxford University Press: New York, NY, USA, 2014.
- Lara, F.A.; Lins, U.; Bechara, G.H.; Oliveira, P.L. Tracing heme in a living cell: Hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. *J. Exp. Biol.* **2005**, *208*, 3093–3101. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sojka, D.; Pytelková, J.; Perner, J.; Horn, M.; Konvičková, J.; Schrenková, J.; Mareš, M.; Kopáček, P. Multienzyme degradation of host serum albumin in ticks. *Ticks Tick-Borne Dis.* **2016**, *7*, 604–613. [\[CrossRef\]](#) [\[PubMed\]](#)
- Horn, M.; Nussbaumerová, M.; Šanda, M.; Kovářová, Z.; Srba, J.; Franta, Z.; Sojka, D.; Bogyo, M.; Caffrey, C.R.; Kopáček, P.; et al. Hemoglobin Digestion in Blood-Feeding Ticks: Mapping a Multi-peptidase Pathway by Functional Proteomics. *Chem. Biol.* **2009**, *16*, 1053–1063. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sojka, D.; Hajdušek, O.; Dvořák, J.; Sajid, M.; Franta, Z.; Schneider, E.L.; Craik, C.S.; Vancová, M.; Burešová, V.; Bogyo, M.; et al. IrAE—An asparaginyl endopeptidase (legumain) in the gut of the hard tick *Ixodes ricinus*. *Int. J. Parasitol.* **2007**, *37*, 713–724. [\[CrossRef\]](#)
- Sojka, D.; Franta, Z.; Frantová, H.; Bartoňová, P.; Horn, M.; Václavová, 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. [\[CrossRef\]](#) [\[PubMed\]](#)
- Franta, Z.; Sojka, D.; Frantová, H.; Dvořák, J.; Horn, M.; Srba, J.; Talacko, P.; Mareš, M.; Schneider, E.; Craik, C.S.; et al. IrCL1—The haemoglobinolytic cathepsin L of the hard tick, *Ixodes ricinus*. *Int. J. Parasitol.* **2011**, *41*, 1253–1262. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sojka, D.; Franta, Z.; Horn, M.; Hajdušek, O.; Caffrey, C.R.; Mareš, M.; Kopáček, P. Profiling of proteolytic enzymes in the gut of the tick *Ixodes ricinus* reveals an evolutionarily conserved network of aspartic and cysteine peptidases. *Parasites Vectors* **2008**, *1*, 7. [\[CrossRef\]](#)
- 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*. *Parasites Vectors* **2010**, *3*, 119. [\[CrossRef\]](#)
- Sojka, D.; Francischetti, I.M.B.; Calvo, E.; Kotsyfakis, M. Cysteine Proteases from Bloodfeeding Arthropod Ectoparasites. *Adv. Exp. Med. Biol.* **2011**, *712*, 177–191. [\[CrossRef\]](#)

13. Sojka, D.; Franta, Z.; Horn, M.; Caffrey, C.R.; Mareš, M.; Kopáček, P. New insights into the machinery of blood digestion by ticks. *Trends Parasitol.* **2013**, *29*, 276–285. [[CrossRef](#)] [[PubMed](#)]
14. Caffrey, C.R.; Goupil, L.; Rebello, K.M.; Dalton, J.P.; Smith, D. Cysteine proteases as digestive enzymes in parasitic helminths. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0005840. [[CrossRef](#)] [[PubMed](#)]
15. Novinec, M.; Lenarčič, B.; Turk, B. Cysteine Cathepsin Activity Regulation by Glycosaminoglycans. *BioMed Res. Int.* **2014**, *2014*, 1–9. [[CrossRef](#)]
16. Turk, V.; Bode, W. The cystatins: Protein inhibitors of cysteine proteinases. *FEBS Lett.* **1991**, *285*, 213–219. [[CrossRef](#)]
17. Rawlings, N.D.; Waller, M.; Barrett, A.J.; Bateman, A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **2014**, *42*, D503–D509. [[CrossRef](#)]
18. Schwarz, A.; Valdés, J.J.; Kotsyfakis, M. The role of cystatins in tick physiology and blood feeding. *Ticks Tick-Borne Dis.* **2012**, *3*, 117–127. [[CrossRef](#)] [[PubMed](#)]
19. Chmelař, J.; Kotál, J.; Langhansová, H.; Kotsyfakis, M. Protease Inhibitors in Tick Saliva: The Role of Serpins and Cystatins in Tick-host-Pathogen Interaction. *Front. Cell. Infect. Microbiol.* **2017**, *7*, 216. [[CrossRef](#)] [[PubMed](#)]
20. Martins, L.A.; Kotál, J.; Bensaoud, C.; Chmelař, J.; Kotsyfakis, M. Small protease inhibitors in tick saliva and salivary glands and their role in tick-host-pathogen interactions. *Biochim. Biophys. Acta BBA Proteins Proteom.* **2020**, *1868*, 140336. [[CrossRef](#)] [[PubMed](#)]
21. Lima, C.A.; Sasaki, S.D.; Tanaka, A.S. Bmcystatin, a cysteine proteinase inhibitor characterized from the tick *Boophilus microplus*. *Biochem. Biophys. Res. Commun.* **2006**, *347*, 44–50. [[CrossRef](#)] [[PubMed](#)]
22. Parizi, L.F.; Githaka, N.W.; Acevedo, C.; Benavides, U.; Seixas, A.; Logullo, C.; Konnai, S.; Ohashi, K.; Masuda, A.; Vaz, I.D.S. Sequence characterization and immunogenicity of cystatins from the cattle tick *Rhipicephalus (Boophilus) microplus*. *Ticks Tick-Borne Dis.* **2013**, *4*, 492–499. [[CrossRef](#)] [[PubMed](#)]
23. Grunclová, L.; Horn, M.; Vancová, M.; Sojka, D.; Franta, Z.; Mares, M.; Kopáček, P. Two secreted cystatins of the soft tick *Ornithodoros moubata*: Differential expression pattern and inhibitory specificity. *Biol. Chem.* **2006**, *387*, 1635–1644. [[CrossRef](#)]
24. Salát, J.; Paesen, G.C.; Řezáčová, P.; Kotsyfakis, M.; Kovářová, Z.; Šanda, 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.* **2010**, *429*, 103–112. [[CrossRef](#)] [[PubMed](#)]
25. 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.* **2017**, *7*, 288. [[CrossRef](#)] [[PubMed](#)]
26. Parizi, L.F.; Sabadin, G.A.; Alzugaray, M.F.; Seixas, A.; Logullo, C.; Konnai, S.; Ohashi, K.; Masuda, A.; da Sliva Vaz, I., Jr. *Rhipicephalus microplus* and *Ixodes ovatus* cystatins in tick blood digestion and evasion of host immune response. *Parasites Vectors* **2015**, *8*, 122. [[CrossRef](#)]
27. Rangel, C.K.; Parizi, L.F.; Sabadin, G.A.; Costa, E.P.; Romeiro, N.C.; Isezaki, M.; Githaka, N.W.; Seixas, A.; Logullo, C.; Konnai, S.; et al. Molecular and structural characterization of novel cystatins from the taiga tick *Ixodes persulcatus*. *Ticks Tick-Borne Dis.* **2017**, *8*, 432–441. [[CrossRef](#)] [[PubMed](#)]
28. Charrier, N.P.; Couton, M.; Voordouw, M.J.; Rais, O.; Durand-Hermouet, A.; Hervet, C.; Plantard, O.; Rispe, C. Whole body transcriptomes and new insights into the biology of the tick *Ixodes ricinus*. *Parasites Vectors* **2018**, *11*, 1–15. [[CrossRef](#)]
29. Kotsyfakis, M.; Schwarz, A.; Erhart, J.; Ribeiro, J.M.C. Tissue- and time-dependent transcription in *Ixodes ricinus* salivary glands and midguts when blood feeding on the vertebrate host. *Sci. Rep.* **2015**, *5*, srep09103. [[CrossRef](#)]
30. Cramaro, W.J.; Revets, D.; Hunewald, O.E.; Sinner, R.; Reye, A.L.; Muller, C.P. Integration of *Ixodes ricinus* genome sequencing with transcriptome and proteome annotation of the naïve midgut. *BMC Genom.* **2015**, *16*, 871. [[CrossRef](#)]
31. Perner, J.; Provazník, J.; Schrenková, J.; Urbanová, V.; Ribeiro, J.M.C.; Kopáček, P. RNA-seq analyses of the midgut from blood- and serum-fed *Ixodes ricinus* ticks. *Sci. Rep.* **2016**, *6*, 36695. [[CrossRef](#)]
32. 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](#)] [[PubMed](#)]
33. Kotsyfakis, M.; Horka, H.; Salát, 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.* **2010**, *77*, 456–470. [[CrossRef](#)] [[PubMed](#)]
34. Kotsyfakis, M.; Sá-Nunes, A.; Francischetti, I.M.B.; Mather, T.N.; Andersen, J.F.; Ribeiro, J.M.C. Antiinflammatory and Immunosuppressive Activity of Sialostatin L, a Salivary Cystatin from the Tick *Ixodes scapularis*. *J. Biol. Chem.* **2006**, *281*, 26298–26307. [[CrossRef](#)]
35. Balbin, M.; Hall, A.; Grubb, A.; Mason, R.W.; Lopez-Otin, C.; Abrahamson, M. Structural and functional characterization of two allelic variants of human cystatin D sharing a characteristic inhibition spectrum against mammalian cysteine proteinases. *J. Biol. Chem.* **1994**, *269*, 23156–23162. [[CrossRef](#)]
36. Anastasi, A.; Brown, M.A.; Kembhavi, A.A.; Nicklin, M.J.H.; Sayers, C.A.; Sunter, D.C.; Barrett, A.J. Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *Biochem. J.* **1983**, *211*, 129–138. [[CrossRef](#)] [[PubMed](#)]

37. Vasiljeva, O.; Dolinar, M.; Turk, V.; Turk, B. Recombinant Human Cathepsin H Lacking the Mini Chain Is an Endopeptidase. *Biochemistry* **2003**, *42*, 13522–13528. [[CrossRef](#)] [[PubMed](#)]
38. Pidugu, L.S.; Maity, K.; Ramaswamy, K.; Surolia, N.; Suguna, K. Analysis of proteins with the ‘Hot dog’ fold: Prediction of function and identification of catalytic residues of hypothetical proteins. *BMC Struct. Biol.* **2009**, *9*, 37. [[CrossRef](#)]
39. Alvarez-Fernandez, M.; Liang, Y.-H.; Abrahamson, M.; Su, X.-D. Crystal Structure of Human Cystatin D, a Cysteine Peptidase Inhibitor with Restricted Inhibition Profile. *J. Biol. Chem.* **2005**, *280*, 18221–18228. [[CrossRef](#)]
40. Bode, W.; Engh, R.; Musil, D.; Thiele, U.; Huber, R.; Karshikov, A.; Brzin, J.; Kos, J.; Turk, V. The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **1988**, *7*, 2593–2599. [[CrossRef](#)]
41. Turk, V.; Stoka, V.; Vasiljeva, O.; Renko, M.; Sun, T.; Turk, B.; Turk, D. Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochim. Biophys. Acta BBA Proteins Proteom.* **2012**, *1824*, 68–88. [[CrossRef](#)]
42. Nandy, S.K.; Seal, A. Structural Dynamics Investigation of Human Family 1 & 2 Cystatin-Cathepsin L1 Interaction: A Comparison of Binding Modes. *PLoS ONE* **2016**, *11*, e0164970. [[CrossRef](#)]
43. Wang, Y.; Yu, X.; Cao, J.; Zhou, Y.; Gong, H.; Zhang, H.; Li, X.; Zhou, J. Characterization of a secreted cystatin from the tick *Rhipicephalus haemaphysaloides*. *Exp. Appl. Acarol.* **2015**, *67*, 289–298. [[CrossRef](#)] [[PubMed](#)]
44. Lu, S.; da Rocha, L.A.; Torquato, R.J.; Junior, I.D.S.V.; Florin-Christensen, M.; Tanaka, A.S. A novel type 1 cystatin involved in the regulation of *Rhipicephalus microplus* midgut cysteine proteases. *Ticks Tick-Borne Dis.* **2020**, *11*, 101374. [[CrossRef](#)] [[PubMed](#)]
45. Kordiš, D.; Turk, V. Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. *BMC Evol. Biol.* **2009**, *9*, 266. [[CrossRef](#)]
46. Kopáček, P.; Perner, J.; Sojka, D.; Šíma, R.; Hajdušek, O. Molecular Targets to Impair Blood Meal Processing in Ticks. In *Ectoparasites*; Wiley: New York, NY, USA, 2018; Volume 8, pp. 139–165.
47. Reyes, J.; Ayala-Chavez, C.; Sharma, A.; Pham, M.; Nuss, A.B.; Gúlia-Nuss, M. Blood Digestion by Trypsin-Like Serine Proteases in the Replete Lyme Disease Vector Tick, *Ixodes scapularis*. *Insects* **2020**, *11*, 201. [[CrossRef](#)] [[PubMed](#)]
48. Cruz, C.E.; Fogaça, A.C.; Nakayasu, E.S.; Angeli, C.B.; Belmonte, R.; Almeida, I.C.; Miranda, A.; Miranda, M.T.M.; 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. *Parasites Vectors* **2010**, *3*, 63. [[CrossRef](#)]
49. Xavier, M.A.; Tirloni, L.; Torquato, R.; Tanaka, A.; Pinto, A.F.M.; Diedrich, J.K.; Yates, J.R., 3rd; da Silva Vaz, I., Jr.; Seixas, A.; Termignoni, C. Blood ant clotting activity of a *Rhipicephalus microplus* cathepsin L-like enzyme. *Biochimie* **2019**, *163*, 12–20. [[CrossRef](#)]
50. Rodríguez-Mallon, A. Developing Anti-tick Vaccines. *Methods Mol. Biol.* **2016**, *1404*, 243–259. [[CrossRef](#)]
51. Nijhof, A.M.; Balk, J.A.; Postigo, M.; Jongejan, F. Selection of reference genes for quantitative RT-PCR studies in *Rhipicephalus (Boophilus) microplus* and *Rhipicephalus appendiculatus* ticks and determination of the expression profile of Bm86. *BMC Mol. Biol.* **2009**, *10*, 112. [[CrossRef](#)]
52. Urbanová, V.; Hartmann, D.; Grunclová, L.; Šíma, R.; Flemming, T.; Hajdušek, O.; Kopáček, P. IrFC—An *Ixodes ricinus* injury-responsive molecule related to Limulus Factor C. *Dev. Comp. Immunol.* **2014**, *46*, 439–447. [[CrossRef](#)]
53. Vechtova, P.; Fussy, Z.; Cegan, R.; Sterba, J.; Erhart, J.; Benes, V.; Grubhoffer, L. Catalogue of stage-specific transcripts in *Ixodes ricinus* and their potential functions during the tick life-cycle. *Parasites Vectors* **2020**, *13*, 1–19. [[CrossRef](#)]
54. Kopáček, P.; Zdychová, J.; Yoshiga, T.; Weise, C.; Rudenko, N.; Law, J.H. Molecular cloning, expression and isolation of ferritins from two tick species—*Ornithodoros moubata* and *Ixodes ricinus*. *Insect Biochem. Mol. Biol.* **2003**, *33*, 103–113. [[CrossRef](#)]
55. Hurrell, J.G.R. *Monoclonal Hybridoma Antibodies: Techniques and Applications*; CRC Press: Boca Raton, FL, USA, 2017; p. 239.
56. Russo, C.; Callegaro, L.; Lanza, E.; Ferrone, S. Purification of IgG monoclonal antibody by caprylic acid precipitation. *J. Immunol. Methods* **1983**, *65*, 269–271. [[CrossRef](#)]
57. Hartmann, D.; Šíma, R.; Konvičková, J.; Perner, J.; Kopáček, P.; Sojka, D. Multiple legumain isoenzymes in ticks. *Int. J. Parasitol.* **2018**, *48*, 167–178. [[CrossRef](#)]
58. Hánová, I.; Brynda, J.; Houštická, R.; Alam, N.; Sojka, D.; Kopáček, P.; Marešová, L.; Vondrášek, J.; Horn, M.; Schueler-Furman, O.; et al. Novel Structural Mechanism of Allosteric Regulation of Aspartic Peptidases via an Evolutionarily Conserved Exosite. *Cell Chem. Biol.* **2018**, *25*, 318–329.e4. [[CrossRef](#)]
59. Schagger, H. Tricine-SDS-PAGE. *Nat. Protoc.* **2006**, *1*, 16–22. [[CrossRef](#)] [[PubMed](#)]
60. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* **1992**, *8*, 275–282. [[CrossRef](#)] [[PubMed](#)]
61. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evol. Int. J. Org. Evol.* **1985**, *39*, 783–791. [[CrossRef](#)]
62. Kumar, S.; Stecher, G.; Li, M.; Nkya, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
63. Mueller, U.; Darowski, N.; Fuchs, M.R.; Förster, R.; Hellmig, M.; Paithankar, K.S.; Pühringer, S.; Steffien, M.; Zocher, G.; Weiss, M.S. Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J. Synchrotron Radiat.* **2012**, *19*, 442–449. [[CrossRef](#)] [[PubMed](#)]
64. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst. Sect. D Biol. Cryst.* **2010**, *66*, 133–144. [[CrossRef](#)] [[PubMed](#)]

65. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Int. Tables Crystallogr.* **2012**, *66*, 364–366. [[CrossRef](#)]
66. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; et al. Overview of theCCP4 suite and current developments. *Acta Cryst. Sect. D Biol. Cryst.* **2011**, *67*, 235–242. [[CrossRef](#)] [[PubMed](#)]
67. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development ofCoot. *Acta Cryst. Sect. D Biol. Cryst.* **2010**, *66*, 486–501. [[CrossRef](#)] [[PubMed](#)]
68. Chen, V.B.; Arendall, W.B., 3rd; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Cryst. D Biol. Cryst.* **2010**, *66*, 12–21. [[CrossRef](#)] [[PubMed](#)]

Supplementary Materials

Table S1. Identification of Mialostatin by mass spectrometry

Equal protein amounts (2.5 µg) of gut tissue and gut lumen contents of *I. ricinus* adult females were reduced with DTT, alkylated with iodacetamide, and digested in parallel with trypsin or chymotrypsin. Peptide digests were reconstituted in 0.1% formic acid and separated on a PepMap RSLC C18 analytical column (2 µm, 15 cm × 75 µm, Thermo Fisher Scientific) using a linear gradient of acetonitrile. The LC-MS/MS analysis was performed on the UltiMate 3000 RSLCnano system (Thermo Fisher Scientific) coupled to Orbitrap Fusion Lumos (Thermo Fisher Scientific). MS scans were recorded from 350 to 2000 m/z in orbitrap; in MS/MS mode the fragmentation spectra were acquired within the mass range of 100–2000 m/z. Proteome Discoverer 2.4 (Thermo Fisher Scientific) was used for protein identification against the UniProtKB/Swiss-Prot and TrEMBL databases. Peptide coverage of the mature Mialostatin sequence at a mass accuracy of <5 p.p.m. is indicated.

Ticks	Sample	Sequence coverage (%)
Fully fed	Gut tissue	71
	Gut lumen content	48
5 Days after detachment	Gut tissue	49
	Gut lumen content	11

Table S2. X-ray data collection and refinement statistics

Data collection	
Space group	P2 ₁ 2 ₁ 2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	81.11, 81.12, 162.33
α , β , γ (°)	90.00, 90.00, 120.00
Number of molecules in AU	2
Wavelength (Å)	0.918
Resolution range (Å)	42.9 - 1.55 (1.65 - 1.55)
Number of unique reflections	46381 (4528)
Redundancy/multiplicity	21.2 (21.5)
Completeness (%)	99.90 (99.25)
R _{merge}	0.084 (1.203)
Average I/s(I)	21.95 (2.00)
Wilson B (Å ²)	19.54
Refinement	
Resolution range (Å)	42.9 - 1.55 (1.65 - 1.55)
No. of reflections in working set	44052 (4528)
No. of reflections in test set	2328 (237)
R value (%)	17.76
R _{free} value (%)	20.83
RMSD bond length (Å)	0.017
RMSD angle (°)	1.94
Number of atoms in AU	2243
Number of protein atoms in AU	1809
Number of water molecules in AU	371
Ramachandran plot	
Residues in favored regions (%)	2.34
Residues in allowed regions (%)	97.66

Table S3. Primer sequences

Amplicon name	Forward primer 5' - 3'	Reverse primer 5' - 3'
Mialostatin cloning	CATATGATGGGGTCGGCGAGCA	CTCGAGCTAGACATCATTAGGA
Mialostatin RT-PCR	GAGGTGCAGACTCAGATTGTGG	GCATACAGCTTCAACCTTGTTTGTC
E11 RT-PCR	ACGAGGCTCTGACGGAAG	CACGACGCAACTCCTTCAC



Figure S1: The phylogenetic tree of 71 cystatins from both Ixodidae and Argasidae tick species was prepared by using the maximum likelihood method and JTT matrix-based model (1). The bootstrap consensus tree is inferred from 1000 replicates (2). Branches corresponding to partitions reproduced in less than 20% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the

bootstrap test (1000 replicates) are shown next to the branches. Evolutionary analyses were conducted in MEGA X (3).

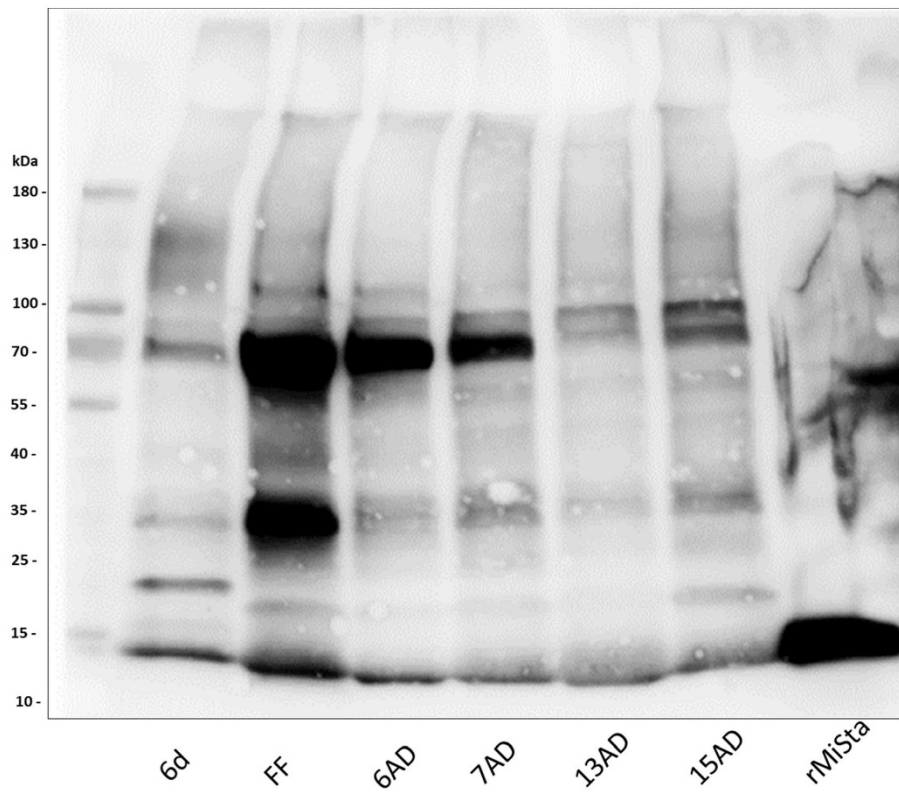


Figure S2: Detection of Mialostatin in midgut of blood fed ticks
Full view on Western blot presented in Figure 3A, left. Bands at higher molecular weight than 13kDa probably represent Mialostatin multimers or complexes.

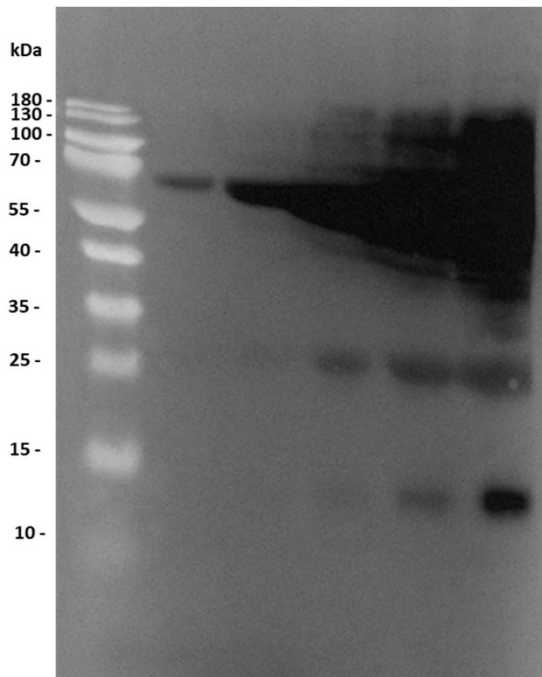


Figure S3: Detection of Mialostatin in midgut of serum fed ticks

Full view on Western blot presented in Figure 3A, right. Lanes differ in protein load and only lane in the right was used for Figure 3A. All samples represent luminal fluid collected from tick midgut after fully fed. Bands at higher molecular weight than 13kDa probably represent Mialostatin multimers or complexes.

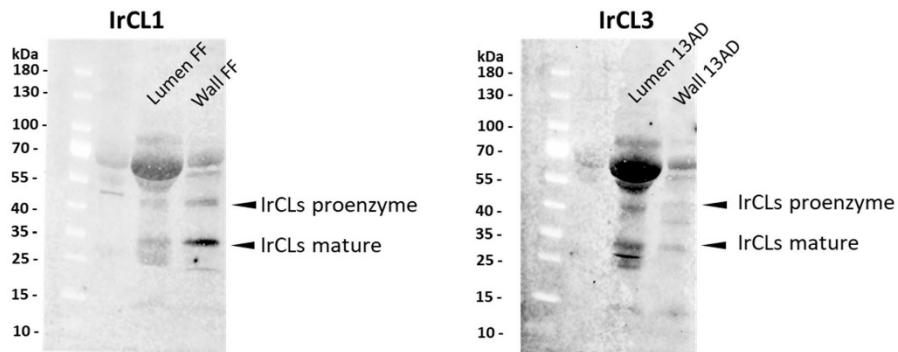


Figure S4: Detection of *I. ricinus* cathepsins L in midgut of blood fed ticks

Full view on Western blot presented in Figure 3B.

References

1. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275-282
2. Felsenstein, J. (1985) Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* **39**, 783-791
3. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547-1549

4.3. Publikace č. 3: The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin.

Sliny klíšťat obsahují velké množství bioaktivních proteinů, jejichž úkolem je modulace fyziologických procesů hostitelské tkáně, aby bylo zajištěno úspěšného sání parazita. Tyto molekuly jsou v současné době intenzivně studovány, protože mají farmakologicky významné vlastnosti. Tato studie se zabývala novým cystatinem iristatinem ze slin klíštěte obecného (*Ixodes ricinus*).

Iristatin je sekretovaný protein, který je exprimován ve slinných žlázách *I. ricinus* během sání. Iristatin byl připraven jako rekombinantní protein a jeho funkční antiproteolytické vlastnosti byly studovány v kinetických inhibičních testech. Ty prokázaly velmi úzkou inhibiční specifitu, kdy byly významně regulovány pouze dva cysteinové katepsiny hostitele v submikromolární a mikromolární oblasti koncentrací. Konkrétně jde o katepsin C s aminodipeptidasovou aktivitou a katepsin L s endopeptidasovou aktivitou. To z iristatinu činí jeden z nejselektivnějších zástupců cystatinů. Identifikovaný inhibiční profil nese příbuzné rysy s několika dříve popsányými slinnými cystatiny z *I. scapularis*, zejména omezenou interakci s cysteinovými katepsiny s exopeptidasovou aktivitou.

Součástí projektu byla krystalografická analýza struktury iristatinu. Byly určeny krystalizační podmínky a vyřešena 3D struktura iristatinu s rozlišením 1,8 Å. Architektura molekuly ukazuje obecný strukturní motiv typický pro podrodinu pravých cystatinů se dvěma páry konzervovaných disulfidů. Iristatin nese tři hlavní vazebné segmenty reaktivního centra pravých cystatinů, konkrétně N-konec molekuly a dvě povrchové smyčky L1 a L2. Nicméně v nich byly identifikovány významné strukturní změny, které jsou zodpovědné za extrémně úzkou inhibiční specifitu iristatinu: 1) N-koncový segment cystatinů obsahuje konzervovanou sekvenci Gly-Gly, která je u iristatinu nahrazena sekvencí Gly-Val, 2) povrchová smyčka L2 nese konzervovaný sekvenční motiv Gln-Xxx-Val-Xxx-Gly, který je v případě iristatinu modifikován záměnou Val za Ile. Smyčka L2 zachovává Pro-Trp motiv konzervovaný u pravých cystatinů, takže za inhibiční specifitu iristatinu jsou zjevně zodpovědné kombinace strukturních efektů na smyčce L1 a N-konci.

Fylogenetická analýza iristatinu ukázala, že není blízkým homologem jiných známých slinných cystatinů klíšťat, ale přesto mají všechny tyto inhibitory společné rysy inhibiční specifity, což naznačuje příbuznou fyziologickou funkci. Biologická aktivita iristatinu byla studována pomocí několika typů testů, které prokázaly silné protizánětlivé a imunosupresní efekty v tkáňových kulturách a v myším modelu. Iristatin snižoval hladinu prozánětlivých interleukinů, počet imunitních buněk v místě zánětu a produkci NO makrofágy.

Tato biologická aktivita dělá z iristatinu farmakologicky vysoce zajímavý protein a dále jej lze považovat, vzhledem k důležité roli v procesu sání klíšťat, za potenciální antigen pro vývoj protiklíšťových vakcín.

Můj podíl na práci zahrnoval: (1) stanovení inhibičních konstant pro iristatin pomocí kinetických testů s fluorescenčními substráty a proteasami (2) krystalizaci iristatinu a sběr difrakčních dat na synchrotronu (3) řešení struktury iristatinu a její analýzu, (4) přípravu manuskriptu.



The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin

Jan Kotál^{1,2} · Natascha Stergiou³ · Michal Buša⁴ · Adéla Chlastáková² · Zuzana Beránková² · Pavlína Řezáčová⁴ · Helena Langhansová² · Alexandra Schwarz¹ · Eric Calvo⁵ · Jan Kopecký² · Michael Mares⁴ · Edgar Schmitt³ · Jindřich Chmelař² · Michail Kotsyfakis¹

Received: 29 November 2018 / Revised: 14 January 2019 / Accepted: 28 January 2019 / Published online: 12 February 2019
© Springer Nature Switzerland AG 2019

Abstract

To successfully feed, ticks inject pharmacologically active molecules into the vertebrate host including cystatin cysteine protease inhibitors. However, the molecular and cellular events modulated by tick saliva remain largely unknown. Here, we describe and characterize a novel immunomodulatory cystatin, Iristatin, which is upregulated in the salivary glands of feeding *Ixodes ricinus* ticks. We present the crystal structure of Iristatin at 1.76 Å resolution. Purified recombinant Iristatin inhibited the proteolytic activity of cathepsins L and C and diminished IL-2, IL-4, IL-9, and IFN- γ production by different T-cell populations, IL-6 and IL-9 production by mast cells, and nitric oxide production by macrophages. Furthermore, Iristatin inhibited OVA antigen-induced CD4⁺ T-cell proliferation and leukocyte recruitment in vivo and in vitro. Our results indicate that Iristatin affects wide range of anti-tick immune responses in the vertebrate host and may be exploitable as an immunotherapeutic.

Keywords Cathepsin · Crystal structure · Immune responses · *Ixodes ricinus* · Saliva

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00018-019-03034-3>) contains supplementary material, which is available to authorized users.

✉ Michail Kotsyfakis
mich_kotsyfakis@yahoo.com

¹ Laboratory of Genomics and Proteomics of Disease Vectors, Biology Centre CAS, Institute of Parasitology, Branišovská 1160/31, 37005 České Budějovice, Czech Republic

² Department of Medical Biology, Faculty of Science, University of South Bohemia in České Budějovice, Branišovská 1760c, 37005 České Budějovice, Czech Republic

³ Institute for Immunology, University Medical Center of the Johannes Gutenberg-University Mainz, Langenbeckstrasse 1, Mainz 55131, Germany

⁴ Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610 Prague, Czech Republic

⁵ Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12735 Twinbrook Parkway, Rockville, MD 20852, USA

Introduction

Ticks are obligatory ectoparasites that feed on the blood of their vertebrate hosts. Hard ticks (family Ixodidae) feed continuously for days to weeks during each life stage, so must overcome specific host antigen-specific immune responses, non-specific innate responses, and hemostasis to successfully finish their blood meal [1]. The hard tick *Ixodes ricinus* is an important arthropod disease vector of several pathogens in Europe [1]. To counteract vertebrate host anti-tick responses, *I. ricinus* secretes saliva rich in biomolecules that facilitate tick feeding and pathogen transmission [2].

Tick saliva affects blood coagulation, complement activation, and immune reaction in terms of immune cell recruitment, cytokine production, and cell maturation [2]. It also facilitates the transmission of *Borrelia*, *Anaplasma*, *Rickettsia*, and various viruses to the vertebrate host [2]. Protease inhibitors are an important group of tick salivary effectors that are divided according to specificity into serine and cysteine protease inhibitors, and according to structure into Kunitz domain inhibitors, serpins, cystatins, and other less abundant families [1]. Kunitz inhibitors are thought to be mainly anti-hemostatic, serpins both anti-hemostatic and

immunomodulatory, and cystatins mainly anti-inflammatory and immunosuppressive [1, 3].

Cystatins are tight binding, reversible legumain, and papain-like cysteine protease inhibitors [4]. Cystatins are subdivided into three subfamilies according to the MEROPS nomenclature: I25A (type 1, stefins), I25B (type 2 and type 3, kininogens), and I25C (type 4, fetuins) [5]. Only type 1 and type 2 cystatins have, thus, far been identified in ticks [6]. Type 1 cystatins lack a signaling peptide for secretion and are, therefore, thought to regulate intracellular blood digestion in the tick midgut, while type 2 cystatins are secreted and expressed in both tick salivary glands and the midgut, and are, therefore, thought to play pleiotropic roles in both ticks and vertebrate hosts [6]. For example, sialostatin L, a cystatin identified in the hard tick *Ixodes scapularis*, inhibits cathepsins C, L, S, V, X, and papain, modulates cytokine production by lymphocytes, dendritic cells, and mast cells, and impairs T-cell proliferation [7, 8]. A similar cystatin in *I. scapularis*, sialostatin L2, inhibits cathepsins C, L, S, and V [9], diminishes IL-1 β and IL-18 secretion by macrophages, and inhibits caspase-1 maturation [10]. Furthermore, both sialostatins alter dendritic cell signaling [11], and inhibition of sialostatin by RNA interference and immunization of guinea pigs impairs tick feeding [9, 12, 13]. Therefore, salivary cystatins may be useful targets for anti-tick vaccines.

However, until the first *I. ricinus* genome was released [14], tick genomic and proteomic studies have been hampered by a lack of full genomic sequences. Here, we report the structural and functional characterization of a novel type 2 cystatin in the hard tick *I. ricinus*, which we name Iristatin. We present the crystal structure of Iristatin, which inhibits the vertebrate cathepsins C and L. Furthermore, we report the anti-inflammatory and immunomodulatory activities of Iristatin. Rather than being target specific, Iristatin appears to affect many immune mechanisms and is a broad-spectrum immunosuppressor that may be useful in the treatment of immune-mediated diseases.

Materials and methods

Quantitative real-time PCR

Female *I. ricinus* ticks were fed on rabbits for 1, 2, 4, 6, or 7 days. Tick salivary glands were dissected, and total RNA isolated and transcribed to cDNA for quantitative analysis of Iristatin by qRT-PCR. Expression profiles were normalized to ferritin mRNA, the levels of which are independent of blood feeding [15]. Detailed methods can be found in the supplement, and the primer and probe sequences are in Supplementary Table S3.

Crystallization, data collection, and structure determination

Screening for crystallization conditions was performed using the JCSG-plus kit (Molecular Dimensions Ltd., Newmarket, UK) and the sitting drop vapor diffusion technique. Preliminary crystals were obtained in 0.1 M Bis-Tris, pH 5.5, 1 M ammonium sulfate, 1% PEG 3350. Optimal Iristatin crystals were prepared at 18 °C using the hanging drop vapor diffusion technique in 15-well NeXtal plates (Qiagen, Hilden, Germany). Experiments, crystal parameters, data collection statistics, and structure determination are detailed in the Supplementary Methods and in Supplementary Table S2.

Enzyme assays

Iristatin inhibition constants against various proteases were determined by measuring the loss of enzymatic activity in the presence of increasing Iristatin concentrations, the corresponding enzyme and a fluorogenic substrate. The enzymes tested were: human liver cathepsin B (BiomolGmbH, Hamburg, Germany); human recombinant cathepsins C, L, S (Calbiochem, Merck Millipore, Burlington, MA, USA); human cathepsin G (Molecular Innovations Inc., Novi, MI, USA); and human factor Xa (Calbiochem). *I. ricinus* legumain IrAE [16] was kindly provided by Daniel Sojka, Ph.D. Experimental details are provided in the Supplementary Methods.

Measurement of cytokine production

T cells were differentiated as follows. For Th1 T cells, naïve CD4⁺ T cells from BALB/c mice were stimulated with anti-CD3/CD28 (4 μ g/ml each) under Th1-skewing conditions (IL-12, anti-IL-4). For Th2 cells, naïve CD4⁺ T cells were stimulated with anti-CD3/CD28 (4 μ g/ml each) under Th2-skewing conditions (IFN- γ , anti-IL-4). For Th9 cells, naïve CD4⁺ T cells from BALB/c mice were stimulated with anti-CD3/CD28 (4 μ g/ml each) under Th9-skewing conditions (IL-4, TGF- β , anti-IFN- γ). Fully differentiated cells were then re-stimulated solely with plate-bound CD3 mAb for 48 h in the presence or absence of 6 μ M LPS-free Iristatin. IL-2, IL-4, IL-9, and IFN- γ production were determined by Ready-SET-Go! ELISA (eBioscience, Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions.

Bone marrow-derived mast cells (BMMCs) were stimulated with ionomycin (Iono, 0.75 μ M) in the presence or absence of 6 μ M Iristatin. IL-4, IL-6, and IL-9 production was determined by ELISA after 48 h of stimulation.

For viability screening, naïve CD4⁺ T cells from BALB/c mice were stimulated with anti-CD3/CD28 (4 µg/ml each) in the presence or absence of different concentrations of LPS-free Iristatin (6, 3, 1.5, and 0.75 µM) under Th9-skewing conditions for 72 h. T cells were stained with a fixable viability dye and cell viability determined by flow cytometry.

Nitric oxide (NO) measurement

Macrophages of the PMJ2-R cell line were preincubated for 4 h with 3 and 6 µM Iristatin and then stimulated by adding LPS to final concentration 100 ng/ml and IFN-γ to final concentration 5 ng/ml. NO concentration was assessed 24 h after stimulation with modified Griess reagent (Sigma-Aldrich, St. Louis, MO, USA).

OVA antigen-induced proliferation of CD4⁺ splenocytes

OT-II mouse spleens were disintegrated through a 70 µm cell strainer to obtain a single-cell suspension, and red blood cells were removed using RBC lysis buffer (eBioscience). Splenocytes were then stained with eFluor™ 670 cell proliferation dye (eBioscience). Stained splenocytes were seeded in a 96-well plate (5 × 10⁵ cells in 200 µl complete RPMI) and preincubated for 2 h in the presence or absence of 3 µM Iristatin. Cells were then stimulated using ovalbumin (OVA) peptide 323–339 (100 ng/ml; Sigma-Aldrich), and splenocytes were incubated for 72 h at 37 °C in 5% CO₂. Cells were then stained with FITC-labeled anti-CD4 antibody and propidium iodide and analyzed by flow cytometry on a BD FACSCanto™ II using BD FACSDiva™ Software v. 6.1.3.

Thioglycollate-induced peritonitis

Female C57BL/6N mice were purchased from Velaz (Prague, Czech Republic). Mice were housed in individually ventilated cages maintained in a 12 h light/dark cycle and given a standard pellet diet and water ad libitum. All animals were used at 8–12 weeks of age. All experiments were approved by the local ethical committee and the Ministry of Education and Sports in accordance with law 246/1992 Sb (ethical approval number MSMT-19085/2015-3).

Control group mice were injected intraperitoneally (i.p.) with saline (10 ml/kg of body weight) and, 1 h later, acute peritonitis was induced by i.p. injection of 200 µl 3% sterile, fully oxidized Difco™ thioglycollate medium (BD Biosciences, Franklin Lakes, NJ, USA). Mice in the experimental group were first injected i.p. with Iristatin (2 mg/kg of body weight in saline). One hour later, mice were treated i.p. with Iristatin (2 mg/kg of body weight) together with 200 µl of 3% thioglycollate medium.

Four hours after thioglycollate medium injection, mice were killed by cervical dislocation, and peritoneal cavities were washed with 10 ml cold PBS to harvest cells. Red blood cells were lysed with RBC lysis buffer (eBioscience). Collected peritoneal cells were counted using a hemocytometer and light microscope. The percentage of live myeloid cells (CD11b⁺), neutrophils (CD11b⁺Ly-6g⁺), monocytes (CD11b⁺Ly-6c⁺), and eosinophils (CD11b⁺Siglec-F⁺) was assessed by flow cytometry (see Supplementary Methods for details). Absolute cell counts were obtained by combining flow cytometry data with cell counting under a light microscope.

Neutrophil in vitro migration assay

Neutrophils were obtained from the bone marrow of C57BL/6J mice by magnetic separation using a Neutrophil Isolation Kit (Miltenyi Biotec). Isolated neutrophils were preincubated in the RPMI medium-containing 0.5% BSA in the presence or absence of 3 µM Iristatin for 1 h at 37 °C and 5% CO₂. Cells were then seeded into the upper inserts of 3 µm pore Corning® Transwell® chambers (24-well format; Sigma-Aldrich, St. Louis, MO, USA) and allowed to migrate towards 1 µM fMLP in RPMI with 0.5% BSA in the lower chamber. After incubation for 1 h at 37 °C and 5% CO₂, cell migration was determined by counting cells in the lower chamber using a hemocytometer.

Statistical analysis

All experiments were performed in biological triplicates. Data are presented as mean ± standard error of mean (SEM) in all graphs. Student's *t* test or one-way ANOVA was used to calculate statistical differences between two or more groups, respectively. Statistically significant results were marked: **P* ≤ 0.05; ***P* ≤ 0.01; ****P* ≤ 0.001; *****P* ≤ 0.0001.

Results

Expression and functional and structural analyses of Iristatin, a novel tick cystatin

Consistent with the other proteins that play important roles in the tick lifecycle [9], Iristatin mRNA expression increased significantly over time in the salivary glands of *I. ricinus* ticks fed on rabbits: 15–20-fold over the first 1–4 days; 50-fold between days 4 and 6, and 80-fold by day 7, when compared with unfed ticks (Supplementary Figure S1). Iristatin (GenBank accession number KY348759) BLAST search revealed three genomic contigs (Genbank accession numbers JXMZ02144755.1; JXMZ02161024.1, and JXMZ02194599.1) [17] with 92.6%, 90.3%, and 96.4%

nucleotide identity, respectively, each representing a unique exon of a single gene. Furthermore, the search in available *I. ricinus* transcriptomes revealed 98.8% identity of Iristatin nucleotide sequence with its best match (Genbank accession number GFVZ01039973.1). Sequence analysis showed that Iristatin belongs to the cystatin superfamily [7–9, 13], specifically to a clade-containing only tick cystatins from the genus *Ixodes* (Fig. 1a). For structural and functional analyses, Iristatin was overexpressed in a prokaryotic system [7] to produce > 95% pure recombinant protein with 119 amino acids, molecular weight 13.8 kDa, and a *pI* of 7.67, as predicted (Supplementary Figure S2 and S3).

The crystal structure of Iristatin was determined by molecular replacement using the soft tick *Ornithodoros moubata* cystatin OmC2 structure as a search model and

refined using data to 1.76 Å resolution (Supplementary Table S2). The orthorhombic crystal form contained two molecules in the asymmetric unit with a solvent content of 50.8%. All protein residues could be modeled into a well-defined electron density map with the exception of the first two N-terminal residues (Gly1 and Met2) and the last four C-terminal residues (Lys116 to Glu119). The final model consisted of two Iristatin molecules, each containing 114 amino acid residues. The root-mean-square deviation (RMSD) for superposition of the C_{α} atoms of the two molecules was 0.52 Å, a value within the range observed for different crystal structures of identical proteins. Minor structural changes were localized to loop regions exposed to solvent and/or involved in crystal contacts (residues 1–2, 47–48, 77–80, 95–96, and 111–114).

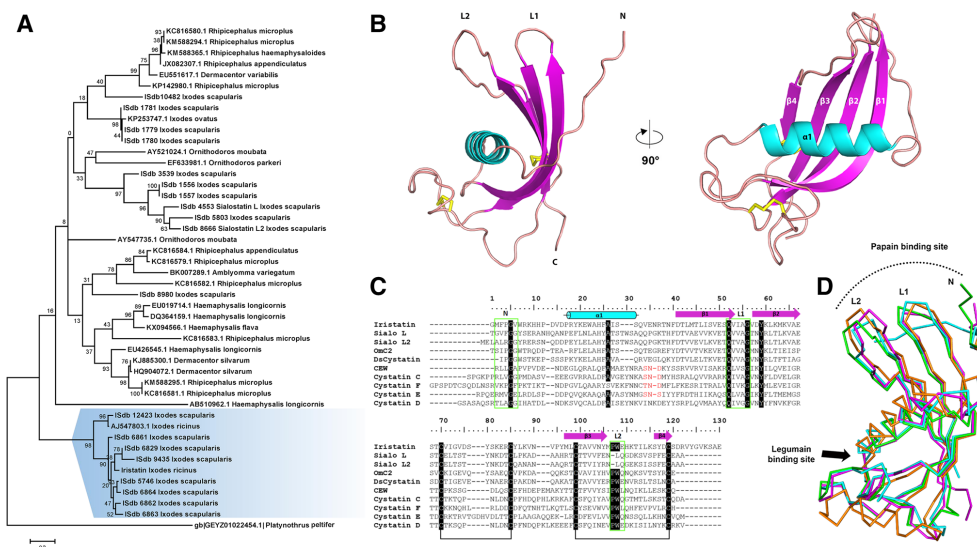


Fig. 1 Crystal structure of Iristatin and its comparison with the other family 2 cystatins. **a** Molecular phylogenetic analysis (maximum-likelihood model) of the secreted tick cystatins. Iristatin clusters with the other cystatins in the genus *Ixodes* (highlighted in blue). *Platynothis* cystatin was used as an outgroup. The tree with the highest log likelihood (–4870, 9711) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. **b** The three-dimensional structure of Iristatin is shown as a cartoon representation colored by secondary structure elements (α 1, cyan; β 1–4, magenta). The N- and C-termini and two disulfide bridges, Cys64–Cys76 and Cys87–Cys107 (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. **c** Structure-based sequence alignment of Iristatin (Iris) with OmC2 (from the soft tick *O. moubata*), sialostatin L and L2 (from the hard tick *I. scapularis*), DsCystatin (from the hard tick *D. silva-*

rum), chicken egg-white cystatin (CEW), and representative human members of family 2 cystatins (cystatins D, C, E/M, and F). Residues identical to those of Iristatin are shaded black. The secondary structure elements of Iristatin are depicted as in **b** (magenta for strands, cyan for helices). The conserved disulfide bridges are indicated by the connecting black lines. Three regions involved in the interaction of cystatins with papain-type peptidases are boxed in green and labeled; the region size was selected based on predominant binding residues in the available complex structures. The putative legumain-binding site in four cystatins is highlighted in red. Mature sequences (i.e., without signal peptide) were used in the alignment; residue numbering is according to Iristatin. **d** A superposition of C_{α} traces of Iristatin with three other cystatin structures. The tick salivary cystatins Iristatin (PDB code 5O46), OmC2 (3LOR), and sialostatin L2 (3LH4) are colored magenta, green, and cyan, respectively. Chicken egg-white cystatin (1CEW) is shown in orange. Positions of the binding sites for papain-type peptidases and legumains are indicated

Figure 1b shows the overall structure of Iristatin. The molecule adopts a typical cystatin fold similar to that of the other homologs, characterized by a twisted antiparallel β -sheet wrapped around an α -helix. However, the Iristatin β -sheet is four-stranded instead of five-stranded, lacking the N-terminal β -strand. Iristatin contains two conserved disulfide bridges connecting Cys64 with Cys76 and Cys87 with Cys107. The structure-based sequence alignment demonstrated that Iristatin displays all the characteristics of family 2 cystatins, including disulfide pattern, the Gln-Xaa-Xaa-Xaa-Gly motif, and a secretion signal removed from the mature protein (Fig. 1c) [18].

The closest structural homologs of Iristatin were salivary cystatins OmC2 from the soft tick *O. moubata* and sialostatin L2 from the hard tick *I. scapularis*; the RMSDs for C α were 1.10 Å and 1.18 Å (without flexible N-termini) and sequence identity was 42% and 36%, respectively (Fig. 1d). Lower structural similarity was found with vertebrate family 2 cystatins, namely human cystatins C, D, F and E/M and chicken egg-white cystatin (RMSDs from 1.53 to 2.21 Å); their sequence identity with Iristatin was between 16 and 24%.

Interaction of family 2 cystatins with papain-type peptidases is mediated by three regions, the N-terminal segment and two hairpin loops L1 and L2, which form a tripartite wedge-shaped edge that binds to the enzyme active site cleft (Fig. 1d). In Iristatin, the first part of the binding site is formed by the N-terminal segment around a conserved Gly5 residue, the orientation suggesting conformational flexibility, as with the other cystatins: Gly5 can function as a hinge that allows the flexible N-terminal segment to adopt an optimal conformation for target enzyme binding. The L1 loop of Iristatin (between β 1 and β 2) is similar in conformation to the other cystatins, only with an isoleucine instead of valine in the conserved sequence motif Gln-Xaa-Val-Xaa-Gly (Fig. 1c). The L2 loop (between β 3 and β 4) is characterized in Iristatin and other cystatins, except sialostatins, by the presence of conservative Pro-Trp residues.

We next analyzed which relevant representative proteases recombinant Iristatin inhibited. As predicted by the crystal structure and the absence of the legumain-binding site localized at a critical Asn residue in Iristatin (Fig. 1c, d), there was no activity against legumain. Among papain-type peptidases, cathepsins B and S were not significantly inhibited under given experimental conditions (Supplementary Table S3). Iristatin was active only against two tested enzymes, cathepsins C and L, displaying similar sub-micromolar affinity as the *I. scapularis* cystatins sialostatins L and L2 to cathepsin C but much lower micromolar affinity to cathepsin L (Supplementary Figure S4). We did not observe any effect of Iristatin against two representative serine peptidases—cathepsin G and factor Xa (Supplementary Table S3).

Iristatin affects cytokine production by T cells and mast cells

Tick salivary cystatins are known to inhibit T-cell cytokine production [3]. To elucidate Iristatin's influence on host immunity and inflammation, we activated different immune cell populations and measured the effect of Iristatin on the production of characteristic cytokines for a given sub-population. As predicted, Iristatin was a potent inhibitor of T-cell-derived cytokines (Fig. 2). In cell cultures, recombinant Iristatin inhibited the production of pro-inflammatory cytokines IFN- γ and IL-2 by polyclonally stimulated [CD3/CD28 monoclonal antibody (mAB)] Th1 cells after 48 h of incubation (Fig. 2a, b). Iristatin also suppressed the production of the anti-inflammatory cytokine IL-4 by Th2 cells (Fig. 2c) and IL-2 and IL-9 by Th9 cells (Fig. 2d, e). The inhibition of IL-9 may be an indirect effect, because IL-9 production is known to be IL-2 dependent [19]. Iristatin had no effect on IL-17 production by Th17 cells (Fig. 2f).

Mast cell numbers are positively correlated with resistance to tick feeding [2]. The only tick cystatin reported to have a direct effect on mast cells is *I. scapularis* sialostatin L, which indirectly reduced IL-9 expression [20]. We stimulated Iristatin pretreated mast cells with ionomycin in cell cultures and measured their cytokine production. While IL-4 production by mast cells was not affected (Fig. 2g), IL-6 levels decreased significantly (Fig. 2h) and IL-9 production was almost blocked by preincubation with 6 μ M Iristatin (Fig. 2i).

Iristatin inhibits antigen-specific CD4⁺ T-cell proliferation and impairs leukocyte recruitment in vitro and in vivo

Saliva, salivary gland extract, or even individual salivary molecules from many tick species inhibit CD4⁺ T-cell proliferation [1]. Cystatins from *I. scapularis* [7] and *O. moubata* [21] are known inhibitors of T lymphocyte proliferation. Accordingly, we investigated whether Iristatin has similar properties in the OVA antigen-specific CD4⁺ T-cell proliferation model using splenocytes isolated from OT-II mice. Iristatin significantly decreased dendritic cell-dependent CD4⁺ T-cell proliferation upon OVA treatment from 88% in PBS-treated controls to 72% in the Iristatin-treated group (Fig. 3a; $p \leq 0.05$) without affecting cell viability (Fig. 3b).

We next investigated whether Iristatin alters inflammatory responses in vivo in a mouse model of thioglycollate-induced peritonitis. Iristatin significantly impaired recruitment of total immune cells to the peritoneum (Fig. 4a) without affecting the proportion of living cells (Fig. 4b), excluding the possibility of Iristatin cytotoxicity. When individual cell populations were examined, Iristatin significantly inhibited the migration of myeloid cells and neutrophils,

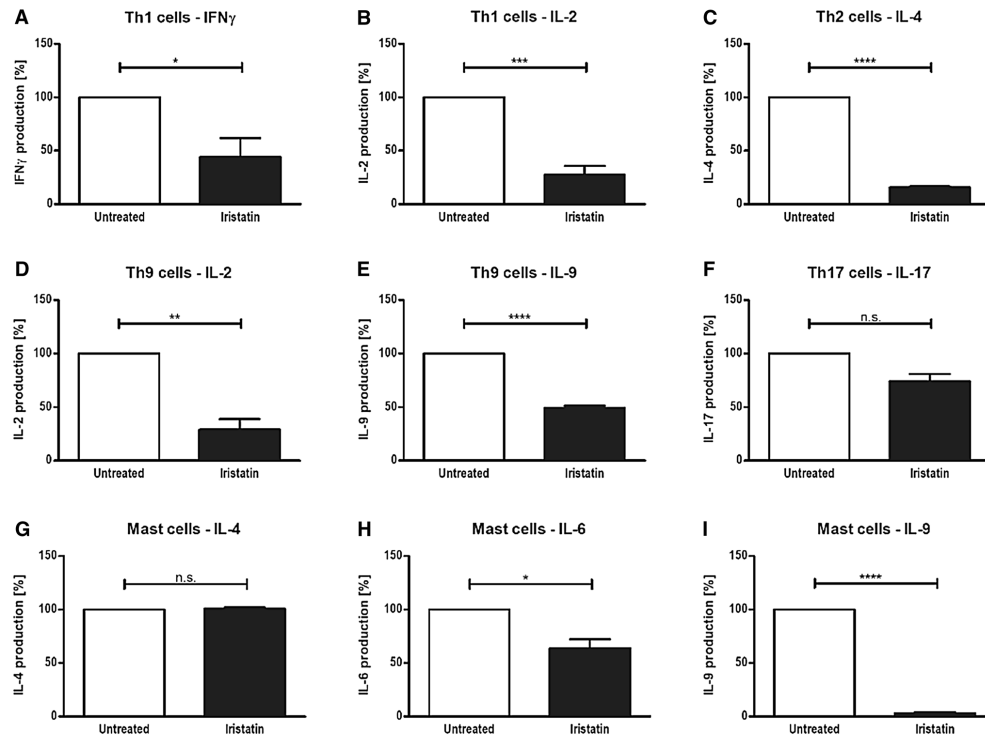


Fig. 2 The effect of Iristatin on T-cell and mast cell cytokine production. **a–f** Different subpopulations of T-helper cells were preincubated with 6 μ M Iristatin and polyclonally stimulated with a combination of CD3 and CD28 mAbs. Cytokine levels were determined 48 h after stimulation. Th cells activated in the absence of Iristatin were used as controls and set as 100% in all experiments; all other values are expressed as percentages of these controls. Iristatin inhibited IFN- γ (**a**) and IL-2 (**b**) production by Th1 cells. **c** Iristatin strongly reduced IL-4 production by Th2 cells. Iristatin inhibited IL-2 (**d**) and IL-9

(**e**) production by Th9 cells. **f** The inhibition of IL-17 production by Th-17 cells was not significant. **g–i** Mast cells were pretreated with Iristatin and stimulated with ionomycin. Cytokine levels were measured 48 h after stimulation. **g** IL-4 production was not affected by Iristatin treatment. **h** 6 μ M Iristatin inhibited IL-6 production by mast cells. **i** Iristatin strongly inhibited IL-9 production by mast cells. The mean of three independent experiments (\pm SEM) is shown in all graphs. * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$; n.s. not significant

and showed a trend to decreasing monocyte and eosinophil migration. Consistent with these *in vivo* findings, the migration of neutrophils pretreated with Iristatin was significantly less than untreated controls towards an fMLP gradient (8.4% vs. 12.1%, $p \leq 0.05$; Fig. 4g).

Finally, macrophages play an important role in the interaction between the host immune system, ticks, and transmitted pathogens. Activated macrophages are crucially involved in immune cell recruitment to sites of inflammation or towards pathogens by secreting signaling molecules such as chemokines or nitric oxide (NO) [22]. Saliva (or salivary gland extracts) from different tick species has been shown to reduce NO production by macrophages [23]. Accordingly, the incubation of monocyte/macrophage PMJ2-R cells with

Iristatin reduced *in vitro* production of NO in a dose-dependent manner to nearly 40% of controls in the presence of 6 μ M Iristatin (Fig. 4h), suggesting a considerable suppression of macrophage activation, perhaps, explaining the reduced recruitment of other immune cell types.

Discussion

Hard ticks feed for several days on their vertebrate host. To feed successfully, ticks control and evade the host immune response and maintain blood flow by secreting saliva into the feeding cavity. As *I. scapularis* cystatins are known to be strong immunomodulators [2], we focused on cloning a

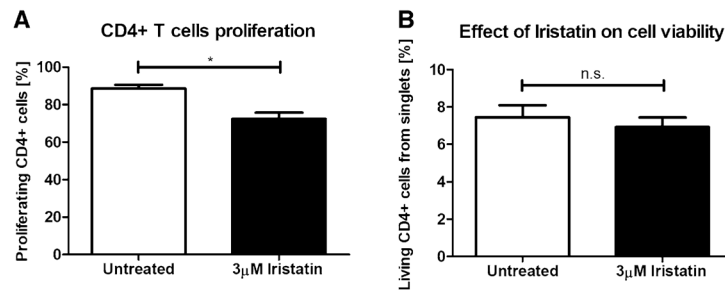


Fig. 3 Iristatin inhibits CD4⁺ lymphocyte proliferation. **a** Splenocytes from OT-II mice were preincubated with Iristatin and subsequently stimulated with OVA peptide. The percentage of proliferating cells was evaluated after 72 h. Incubation with 3 μ M Iristatin decreased

the percentage of proliferating cells from 88% in the control group to 72% in the experimental group. The mean of three independent experiments (\pm SEM) is shown. * $P \leq 0.05$ (two-tailed, unpaired t test). **b** Iristatin had no effect on cell viability

cystatin from the closely related tick *I. ricinus*, which we named Iristatin. At the time of this project initiation, *I. ricinus* genome has not yet been sequenced, so the traditional cloning procedures were needed to reveal this first immunomodulatory cystatin from this important disease vector.

The three-dimensional structural analysis of Iristatin provided useful insights into its biochemistry and function. Three conserved cystatin domains mediate their specificity to papain-like proteases [4, 24], including the N-terminal domain and two hairpin loops L1 and L2. Iristatin differed by over 50% in the N-terminal domain sequence but by only one amino acid in hairpin L1 (Ile50 instead of Val) or L2 (Glu99 instead of Gln) compared to the most structurally similar tick cystatin OmC2. This hairpin loop similarity, perhaps, explains why both cystatins inhibit cathepsins C and L. Iristatin showed a major difference in affinity to cathepsin L compared to sialostatins L and L2, in which target specificity is attributed to the lack of a conserved Pro-Trp motif in hairpin L2, at least with regard to the lower affinity to cathepsin B and no increased inhibition of cathepsin L [7, 9, 25]. While the only difference between Iristatin and the sialostatins in hairpin L1 is a Val/Ile substitution, we speculate that the significant difference in cathepsin L affinity can be explained by the different N-termini of these cystatins or in structures outside the conserved domains. Similarly, both Iristatin and DsCystatin [26] inhibited cathepsin L, but differed in their affinity to cathepsins B and C. While both of these cystatins are almost identical in their L1 and L2 hairpin sequences, the difference in their inhibitory specificity probably originates in the N-terminal region or outside the conserved domains. Similar to sialostatin L, we speculate that inhibition of cathepsins C and L by Iristatin impairs the maturation of other proteases from their proenzymes by blocking the cleavage of their N-terminal propeptides by cathepsins C and L [7]. This effect could reduce granzyme

activity in cytotoxic T lymphocytes and natural killer cells, trypsin and chymase in mast cells, cathepsin G, proteinase 3, and elastase in neutrophils, or impair the maturation of cathepsins D and B in various cell types [2].

Iristatin suppressed immune responses both in vitro and in vivo. *I. ricinus* saliva which has previously been shown to polarize immune responses towards the Th2 pathway [27], although saliva-driven inhibition of both Th1 and Th2 pathways has been observed in dendritic cells [28]. However, Iristatin appears to have a more general effect on vertebrate immunity. Inhibition of the production of pro-inflammatory cytokines TNF and IL-12 by the tick cystatin OmC2 or IL-1 β , IFN- γ , TNF, and IL6 by DsCystatin has been described in dendritic cells [21] and macrophages [26]. Moreover, impaired T-cell production of IFN- γ by sialostatin L has also been reported [8] together with a decrease in IL-17 production, which was not observed with Iristatin. Similar to sialostatin L, Iristatin suppressed IL-2 and IL-9 production by Th9 and mast cells, respectively, suggesting that Iristatin could also have a similar inhibitory effect on experimental asthma [29]. To our knowledge, no secreted tick cystatin has been reported to inhibit both Th1 and Th2 cytokines as much as Iristatin. It has been observed in ticks and other blood-feeding parasites that whole saliva can have general immunosuppressive effects, with individual proteins inhibiting specific elements. Saliva from the *Aedes aegypti* mosquito has been reported to suppress both Th1 and Th2 cytokines [30]. In contrast, a single *A. aegypti* salivary protein, SAAG-4, potentially polarized a Th2 immune response by reducing expression of the Th1 cytokine IFN- γ and upregulating the Th2 cytokine IL-4 [31]. Such immunomodulation is not limited to arthropods, as the parasitic trematode *Schistosoma japonicum* also polarized the vertebrate immune response towards Th2 upon infection, with rSjCystatin increasing IL-4 production by splenocytes [32].

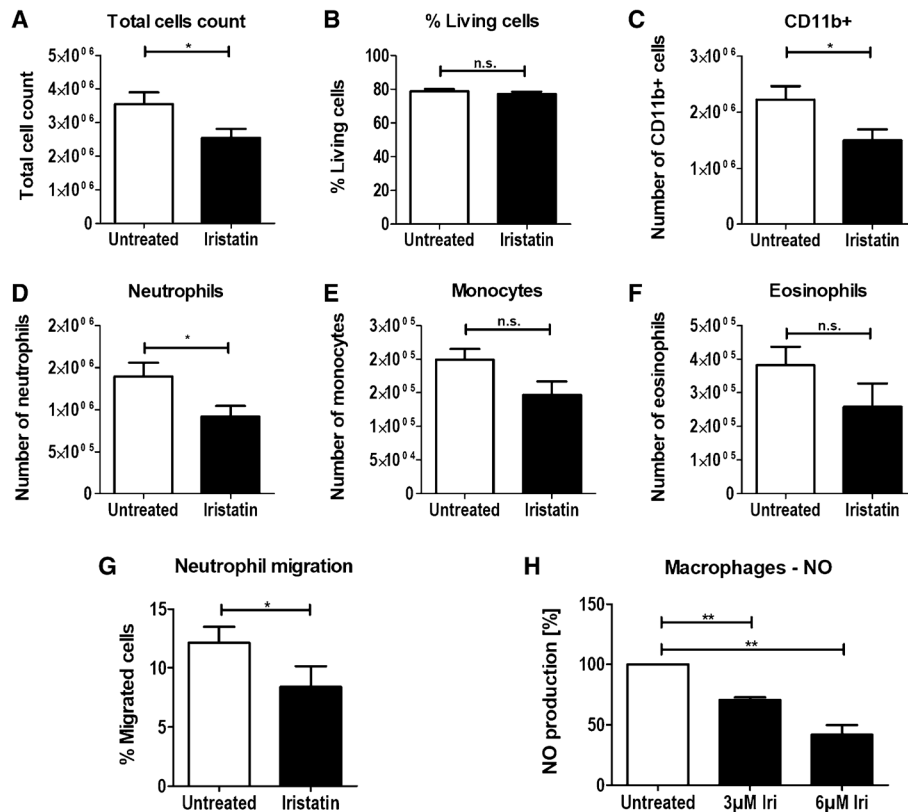


Fig. 4 Iristatin's effect on leukocyte recruitment and migration and macrophage production of NO. **a–f** Mice were injected with 3% TGM with saline or Iristatin. After 4 h, peritoneal lavage was performed and infiltrating cells analyzed by flow cytometry. In all figures, mice injected with TGM and saline are marked as Untreated, while all mice injected with TGM and Iristatin (2 mg/kg of mouse weight) are labeled as Iristatin. **a** Total number of cells in the peritoneum in tested mice. **b** Percentage of all living cells in mouse peritoneum. **c–f** Total number of all living CD11b⁺ cells, neutrophils, monocytes, or

eosinophils, respectively. **g** Mouse bone marrow neutrophils were preincubated with 3 μM Iristatin and subjected to migration towards fMLP in a Boyden chamber. **h** Iristatin inhibited NO production by PMJ2-R macrophages in a dose-dependent manner. Macrophages were preincubated with 3 and 6 μM Iristatin, stimulated with LPS and IFN-γ, and NO concentration was assessed after 24 h. The mean of three independent experiments (± SEM) is shown. * $P \leq 0.05$; ** $P \leq 0.01$; n.s. not significant

Macrophages play a crucial role in inflammatory responses through high cytokine and NO production and represent known cystatin targets [26]. Most cystatins tend to increase macrophage NO production [33]. Conversely, rSjCystatin from *S. japonicum* [34] and Iristatin decreased macrophage NO levels. Therefore, Iristatin has a rather unique effect on NO production, which is consistent with the previous studies, showing that *I. ricinus* saliva inhibits NO production [23]. Since NO is an important regulator of many processes in various immune and inflammatory cell types, we propose that its reduction by Iristatin leads to further

immunosuppression. NO can be associated with decreased T-cell proliferation [35]. Moreover, the dose-dependent differences in NO-mediated polarization of immune response have been described. Low NO levels selectively enhanced Th1 polarization, while Th1 differentiation was suppressed at higher NO levels [36]. In contrast to these studies, we observed that Iristatin decreased NO production by macrophages and also reduced T-cell proliferation; however, since both experiments were performed separately in vitro, the in vivo milieu may be different. Furthermore, we used whole splenocytes in the proliferation assays, so it is unclear

whether the observed lower CD4⁺ cell proliferation was a direct or an indirect effect via APCs.

Finally, Iristatin suppressed immune cell recruitment to the site of inflammation. Several parasitic nematode cystatins have been reported to affect inflammatory cell migration while polarizing towards Th2 immune responses and recruiting IL-10-producing macrophages or T cells [37–39]. Conversely, cystatin C inhibited T-cell and monocyte transmigration [40]. Until now, only one tick cystatin, DsCystatin, has been reported to have an impact on immune cell recruitment in a mouse arthritis model [26]. Another tick cystatin, RHCyst-1 from *Rhipicephalus haemaphysaloides*, suppressed tumor cell migration and invasion [41]. We speculate that the Iristatin effect on cell recruitment can partially be mediated by the suppression of macrophage activation [22] as shown by NO inhibition.

In conclusion, Iristatin is a potent immunomodulator of the host immune system. Iristatin is a specific cathepsin C and L inhibitor, as evidenced by both structure and function. Iristatin attenuated both Th1 and Th2 vertebrate host immune responses and inhibited T-cell proliferation and leukocyte recruitment. Our data clearly demonstrate that individual molecules contribute differentially to the overall effect of arthropod saliva on blood feeding. Our model for the action of tick salivary immunomodulators [42] stresses the importance of pluripotency and redundancy in their action; specifically, the function of each tick salivary immunomodulator may overlap with that of another one and the same tick salivary immunomodulator may simultaneously affect different vertebrate immune system functions [42]. We speculate that synergy exists between individual salivary proteins that might increase the activity of tick saliva; however, there is currently no direct evidence for such an effect. Pluripotency and redundancy seem to be essential for potent immunomodulation by tick saliva, and Iristatin displays both of these features. These structural and functional data further increase our understanding of vertebrate host immunomodulation by tick saliva. Furthermore, these properties could potentially be exploited for the development of novel immune-related disease drugs or vaccines.

Acknowledgements We thank Professor Jose MC Ribeiro and Drs. Petr Kopáček and Daniel Sojka for constructive discussions, Dr. Daniel Sojka for kindly providing published reagent, ARVYS Proteins for service provision, Dr. Andreza Campos-Chagas and Mr. Jan Erhart for technical assistance and Nextgenediting for editorial assistance. We also thank the anonymous reviewers for constructive comments about the draft.

Author contributions JKot designed and performed experiments, performed the analyses, and wrote the manuscript; NS, MB, AC, ZB, PŘ, HL, and AS designed and performed experiments; JC and MK designed experiments, performed analyses, and edited the manuscript; MM and ES designed experiments and revised the manuscript; EC and JKot revised the manuscript.

Funding This work was supported by the Grant Agency of the Czech Republic (Grant 19-07247S to MK and Grant 16-07117Y to JC), by the Grant Agency of the University of South Bohemia (Grant 038/2016/P to JKot), and by ERD Funds, project CePaViP OPVVV (No. CZ.02.1.01/0.0/0.0/16_019/0000759 to MK) and institutional project RVO 60077344 to MK. MB, MM, and PŘ were supported by project ChemBioDrug (No. CZ.02.1.01/0.0/0.0/16_019/0000729) from ERD Funds and institutional project RVO 61388963.

Data availability The data set supporting the conclusions of this article is available in GenBank, accession number KY348759; PDB code 5O46.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

- Francischetti IM, Sa-Nunes A, Mans BJ, Santos IM, Ribeiro JM (2009) The role of saliva in tick feeding. *Front Biosci (Landmark Ed)* 14:2051–2088
- Kotal J, Langhansova H, Lieskova J, Andersen JF, Francischetti IM, Chavakis T, Kopecky J, Pedra JH, Kotsyfakis M, Chmelar J (2015) Modulation of host immunity by tick saliva. *J Proteomics* 128:58–68. <https://doi.org/10.1016/j.jprot.2015.07.005>
- Chmelar J, Kotal J, Langhansova H, Kotsyfakis M (2017) Protease inhibitors in tick saliva: the role of serpins and cystatins in tick–host–pathogen interaction. *Front Cell Infect Microbiol* 7:216. <https://doi.org/10.3389/fcimb.2017.00216>
- Turk V, Bode W (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett* 285(2):213–219
- Rawlings ND, Waller M, Barrett AJ, Bateman A (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 42(Database issue):D503–D509. <https://doi.org/10.1093/nar/gkt953>
- Schwarz A, Valdes JJ, Kotsyfakis M (2012) The role of cystatins in tick physiology and blood feeding. *Ticks Tick Borne Dis* 3(3):117–127. <https://doi.org/10.1016/j.ttbdis.2012.03.004>
- Kotsyfakis M, Sa-Nunes A, Francischetti IM, Mather TN, Andersen JF, Ribeiro JM (2006) Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*. *J Biol Chem* 281(36):26298–26307. <https://doi.org/10.1074/jbc.M513010200>
- Sa-Nunes A, Bafica A, Antonelli LR, Choi EY, Francischetti IM, Andersen JF, Shi GP, Chavakis T, Ribeiro JM, Kotsyfakis M (2009) The immunomodulatory action of sialostatin L on dendritic cells reveals its potential to interfere with autoimmunity. *J Immunol* 182(12):7422–7429. <https://doi.org/10.4049/jimmunol.0900075>
- Kotsyfakis M, Karim S, Andersen JF, Mather TN, Ribeiro JM (2007) Selective cysteine protease inhibition contributes to blood-feeding success of the tick *Ixodes scapularis*. *J Biol Chem* 282(40):29256–29263. <https://doi.org/10.1074/jbc.M703143200>
- Chen G, Wang X, Severo MS, Sakhon OS, Sohail M, Brown LJ, Sircar M, Snyder GA, Sundberg EJ, Ulland TK, Olivier AK, Andersen JF, Zhou Y, Shi GP, Sutterwala FS, Kotsyfakis M, Pedra JH (2014) The tick salivary protein sialostatin L2 inhibits caspase-1-mediated inflammation during *Anaplasma phagocytophilum* infection. *Infect Immun* 82(6):2553–2564. <https://doi.org/10.1128/IAI.01679-14>

11. Lieskovska J, Palenikova J, Langhansova H, Campos Chagas A, Calvo E, Kotsyfakis M, Kopecky J (2015) Tick sialostatin L and L2 differentially influence dendritic cell responses to *Borrelia burgdorferi*. *Parasit Vectors* 8:275. <https://doi.org/10.1186/s13071-015-0887-1>
12. Kotsyfakis M, Horka H, Salat J, Andersen JF (2010) 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(2):456–470. <https://doi.org/10.1111/j.1365-2958.2010.07220.x>
13. Kotsyfakis M, Anderson JM, Andersen JF, Calvo E, Francischetti IM, Mather TN, Valenzuela JG, Ribeiro JM (2008) Cutting edge: immunity against a “silent” salivary antigen of the Lyme vector *Ixodes scapularis* impairs its ability to feed. *J Immunol* 181(8):5209–5212
14. Cramaro WJ, Revets D, Hunewald OE, Sinner R, Reye AL, Muller CP (2015) Integration of *Ixodes ricinus* genome sequencing with transcriptome and proteome annotation of the naive midgut. *BMC Genomics* 16:871. <https://doi.org/10.1186/s12864-015-1981-7>
15. Kopacek P, Zdychova J, Yoshiga T, Weise C, Rudenko N, Law JH (2003) Molecular cloning, expression and isolation of ferritins from two tick species—*Ornithodoros moubata* and *Ixodes ricinus*. *Insect Biochem Mol Biol* 33(1):103–113
16. Sojka D, Hajdusek O, Dvorak J, Sajid M, Franta Z, Schneider EL, Craik CS, Vancova M, Buresova V, Bogyo M, Sexton KB, McKerrow JH, Caffrey CR, Kopacek P (2007) IrAE: an asparaginyl endopeptidase (legumain) in the gut of the hard tick *Ixodes ricinus*. *Int J Parasitol* 37(7):713–724. <https://doi.org/10.1016/j.ijpara.2006.12.020>
17. Cramaro WJ, Hunewald OE, Bell-Sakyi L, Muller CP (2017) Genome scaffolding and annotation for the pathogen vector *Ixodes ricinus* by ultra-long single molecule sequencing. *Parasites Vectors* 10(1):71. <https://doi.org/10.1186/s13071-017-2008-9>
18. Vray B, Hartmann S, Hoebeke J (2002) Immunomodulatory properties of cystatins. *Cell Mol Life Sci* 59(9):1503–1512
19. Schmitt E, Germann T, Goedert S, Hoehn P, Huels C, Koelsch S, Kuhn R, Muller W, Palm N, Rude E (1994) IL-9 production of naive CD4⁺ T cells depends on IL-2, is synergistically enhanced by a combination of TGF-beta and IL-4, and is inhibited by IFN-gamma. *J Immunol* 153(9):3989–3996
20. Klein M, Bruhl TJ, Staudt V, Reuter S, Grebe N, Gerlitzki B, Hoffmann M, Bohn T, Ulges A, Stergiou N, de Graaf J, Lower M, Taube C, Becker M, Hain T, Dietzen S, Stassen M, Huber M, Lohoff M, Campos Chagas A, Andersen J, Kotál J, Langhansova H, Kopecky J, Schild H, Kotsyfakis M, Schmitt E, Bopp T (2015) Tick salivary sialostatin L represses the initiation of immune responses by targeting IRF4-dependent transcription in murine mast cells. *J Immunol* 195(2):621–631. <https://doi.org/10.4049/jimmunol.1401823>
21. Salat J, Paesen GC, Rezacova P, Kotsyfakis M, Kovarova Z, Sanda M, Majtan J, Grunclova L, Horka H, Andersen JF, Brynda J, Horn M, Nunn MA, Kopacek P, Kopecky J, Mares M (2010) Crystal structure and functional characterization of an immunomodulatory salivary cystatin from the soft tick *Ornithodoros moubata*. *Biochem J* 429(1):103–112. <https://doi.org/10.1042/BJ20100280>
22. Laroux FS, Pavlick KP, Hines IN, Kawachi S, Harada H, Bhawani S, Hoffman JM, Grisham MB (2001) Role of nitric oxide in inflammation. *Acta Physiol Scand* 173(1):113–118. <https://doi.org/10.1046/j.1365-201X.2001.00891.x>
23. Kyckova K, Kopecky J (2006) Effect of tick saliva on mechanisms of innate immune response against *Borrelia afzelii*. *J Med Entomol* 43(6):1208–1214
24. Bode W, Huber R (2000) Structural basis of the endoprotease-protein inhibitor interaction. *Biochim Biophys Acta* 1477(1–2):241–252
25. Bjork I, Brieditis I, Raub-Segall E, Pol E, Hakansson K, Abrahamson M (1996) The importance of the second hairpin loop of cystatin C for proteinase binding. Characterization of the interaction of Trp-106 variants of the inhibitor with cysteine proteinases. *Biochemistry* 35(33):10720–10726. <https://doi.org/10.1021/bi960420u>
26. Sun T, Wang F, Pan W, Wu Q, Wang J, Dai J (2018) An immunosuppressive tick salivary gland protein DsCystatin interferes with toll-like receptor signaling by downregulating TRAF6. *Front Immunol* 9:1245. <https://doi.org/10.3389/fimmu.2018.01245>
27. Skalova A, Iezzi G, Ampenberger F, Kopf M, Kopecky J (2008) Tick saliva inhibits dendritic cell migration, maturation, and function while promoting development of Th2 responses. *J Immunol* 180(9):6186–6192
28. Slamova M, Skalova A, Palenikova J, Kopecky J (2011) Effect of tick saliva on immune interactions between *Borrelia afzelii* and murine dendritic cells. *Parasite Immunol* 33(12):654–660. <https://doi.org/10.1111/j.1365-3024.2011.01332.x>
29. Horka H, Staudt V, Klein M, Taube C, Reuter S, Dehzad N, Andersen JF, Kopecky J, Schild H, Kotsyfakis M, Hoffmann M, Gerlitzki B, Stassen M, Bopp T, Schmitt E (2012) The tick salivary protein sialostatin L inhibits the Th9-derived production of the asthma-promoting cytokine IL-9 and is effective in the prevention of experimental asthma. *J Immunol* 188(6):2669–2676. <https://doi.org/10.4049/jimmunol.1100529>
30. Wasserman HA, Singh S, Champagne DE (2004) Saliva of the Yellow Fever mosquito, *Aedes aegypti*, modulates murine lymphocyte function. *Parasite Immunol* 26(6–7):295–306. <https://doi.org/10.1111/j.0141-9838.2004.00712.x>
31. Boppana VD, Thangamani S, Adler AJ, Wikel SK (2009) SAAG-4 is a novel mosquito salivary protein that programmes host CD4 T cells to express IL-4. *Parasite Immunol* 31(6):287–295. <https://doi.org/10.1111/j.1365-3024.2009.01096.x>
32. Chen L, He B, Hou W, He L (2017) Cysteine protease inhibitor of *Schistosoma japonicum*—a parasite-derived negative immunoregulatory factor. *Parasitol Res* 116(3):901–908. <https://doi.org/10.1007/s00436-016-5363-0>
33. Verdot L, Lalmanach G, Vercruyse V, Hartmann S, Lucius R, Hoebeke J, Gauthier F, Vray B (1996) Cystatins up-regulate nitric oxide release from interferon-gamma-activated mouse peritoneal macrophages. *J Biol Chem* 271(45):28077–28081
34. Yang X, Liu J, Yue Y, Chen W, Song M, Zhan X, Wu Z (2014) Cloning, expression and characterisation of a type II cystatin from *Schistosoma japonicum*, which could regulate macrophage activation. *Parasitol Res* 113(11):3985–3992. <https://doi.org/10.1007/s00436-014-4064-9>
35. Sato K, Ozaki K, Oh I, Meguro A, Hatanaka K, Nagai T, Muroi K, Ozawa K (2007) Nitric oxide plays a critical role in suppression of T-cell proliferation by mesenchymal stem cells. *Blood* 109(1):228–234. <https://doi.org/10.1182/blood-2006-02-002246>
36. Niedbala W, Cai B, Liew FY (2006) Role of nitric oxide in the regulation of T cell functions. *Ann Rheum Dis* 65(Suppl 3):iii37–iii40. <https://doi.org/10.1136/ard.2006.058446>
37. Schuijs MJ, Hartmann S, Selkirk ME, Roberts LB, Openshaw PJ, Schnoeller C (2016) The helminth-derived immunomodulator AvCystatin reduces virus enhanced inflammation by induction of regulatory IL-10⁺ T cells. *PLoS One* 11(8):e0161885. <https://doi.org/10.1371/journal.pone.0161885>
38. Danilowicz-Luebert E, Steinfelder S, Kuhl AA, Drozdenko G, Lucius R, Worm M, Hamelmann E, Hartmann S (2013) A nematode immunomodulator suppresses grass pollen-specific allergic responses by controlling excessive Th2 inflammation. *Int J Parasitol* 43(3–4):201–210. <https://doi.org/10.1016/j.ijpara.2012.10.014>
39. Schnoeller C, Rausch S, Pillai S, Avagyan A, Wittig BM, Lodenkemper C, Hamann A, Hamelmann E, Lucius R, Hartmann S (2008) A helminth immunomodulator reduces allergic and

- inflammatory responses by induction of IL-10-producing macrophages. *J Immunol* 180(6):4265–4272
40. Staun-Ram E, Miller A (2011) Cathepsins (S and B) and their inhibitor Cystatin C in immune cells: modulation by interferon-beta and role played in cell migration. *J Neuroimmunol* 232(1–2):200–206. <https://doi.org/10.1016/j.jneuroim.2010.10.015>
 41. Wei N, Lin Z, Xu Z, Cao J, Zhou Y, Zhang H, Gong H, Zhou J, Li G (2018) A tick cysteine protease inhibitor RHcyst-1 exhibits antitumor potential. *Cell Physiol Biochem* 46(6):2385–2400. <https://doi.org/10.1159/000489645>
 42. Chmelar J, Kotal J, Kopecky J, Pedra JHF, Kotsyfakis M (2016) All for one and one for all on the tick-host battlefield. *Trends Parasitol* 32(5):368–377. <https://doi.org/10.1016/j.pt.2016.01.004>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Material

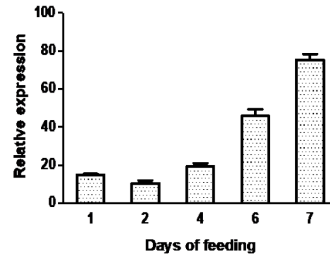


Figure S1 - *Iristatin* transcripts accumulate in *I. ricinus* salivary glands upon tick feeding. Pools of *I. ricinus* salivary glands from female ticks were dissected under RNase-free conditions. cDNA was subsequently prepared as a template for qRT-PCR. *Iristatin* expression was normalized to the reference gene ferritin and compared to unfed ticks. Values on the y-axis show the fold increase in transcript (unfed = 1). Each bar represents the ratio between the number of *Iristatin* transcripts in salivary glands at a specific tick feeding stage (shown on the x-axis) compared to unfed ticks. The data show an average of three biological replicates (\pm SEM).

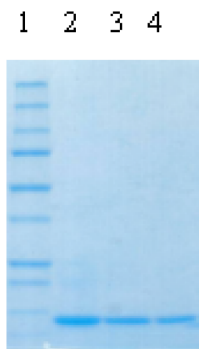


Figure S2 - *Iristatin* purification. Lane 1: MW standards (250, 150, 100, 75, 50, 37, 25, 20, 15, 10 kDa); Lane 2: *Iristatin*, load 6 μ g; Lane 3: *Iristatin*, load 3 μ g; Lane 4: *Iristatin*, load 1.5 μ g.

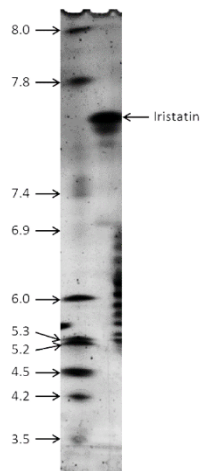


Figure S3 - Iristatin isoelectric focusing. The pI of Iristatin was between 7.6 and 7.7.

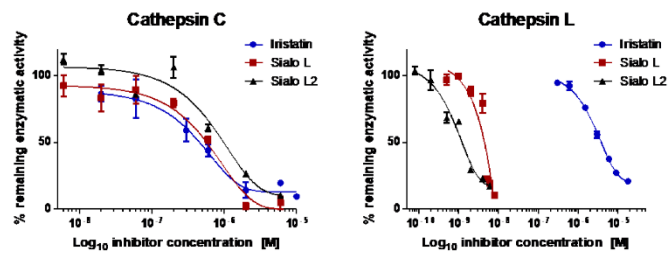


Figure S4 - Iristatin inhibits cathepsins C and L. Iristatin and sialostatins L and L2 differ in affinity for two of their common enzymatic targets cathepsins C and L. The three cystatins were allowed to interact with the same amount of enzyme under the same assay conditions. The resulting reduction in enzymatic activity was plotted against the corresponding inhibitor concentration. Each experiment was performed in triplicate.

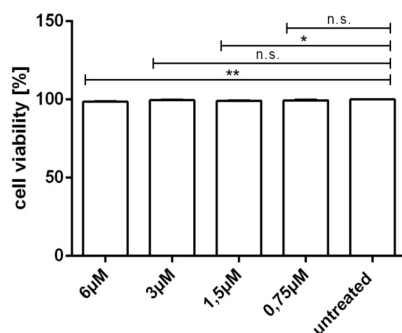


Figure S5 - Iristatin does not alter cell viability. The mean of three independent experiments (\pm SD) is shown. * \leq P 0.05; ** \leq P 0.01; *** \leq P 0.001, **** \leq P 0.0001 (two-tailed unpaired t-test). ns=not significant.

Table S1 - Primer and probe sequences.

Amplicon name	Forward primer 5' - 3'	Reverse Reverse primer 5' - 3'
Iristatin full clone	ATGAGTATCGTGAAGGCAGCGC	CGGTG TCATTCAGCTGGCTTGA
Iristatin no SigP	ATCATATGTTTCCCGGGTCTGGAGGAAGCAC	CGGTG ATCTCGAGTCATTCAGCTGACTTCA
Iristatin RT-PCR	CACACTTCGCCATCTCCTT	GCCTC CATGTAGTTGACCACGGCTGTAC
Iristatin RT-PCR probe	YAK-TCATCAGCGTCGAGTCTCAGGTAATTGC-BBQ	

Table S2 - Crystallography data collection parameters.

Data collection statistics	
Space group	$P2_12_12_1$
Cell parameters (Å)	39.03 76.93 92.12
Number of molecules in AU	2
Wavelength (Å)	0.918
Resolution range (Å)	46.06 - 1.76 (1.82 - 1.76)
Number of unique reflections	27667 (2700)
Redundancy/multiplicity	3.7 (3.7)
Completeness (%)	97.76 (97.90)

R_{merge}^a	0.0377 (0.79)
Average $I/\sigma(I)$	19.32 (1.39)
Wilson B (\AA^2)	29.42
Refinement statistics	
Resolution range (\AA)	46.06 - 1.76 (1.82 - 1.76)
No. of reflections in working set	27631 (2700)
No. of reflections in test set	1381 (135)
R value (%) ^b	19.38
R_{free} value (%) ^c	25.68
RMSD bond length (\AA)	0.016
RMSD angle ($^\circ$)	1.71
Number of atoms in AU	2196
Number of protein atoms in AU	1922
Number of water molecules in AU	239
Ramachandran plot statistics ^d	
Residues in favored regions (%)	1.71
Residues in allowed regions (%)	97.77

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

^a $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, 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.

^b R-value = $\frac{|F_o| - |F_c|}{|F_o|}$, where F_o and F_c are the observed and calculated structure factors, respectively.

^c R_{free} is equivalent to R value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process [1].

^d as determined by PROCHECK ([2])

Table S3 - Effect of Iristatin on protease activity

Protease	% of remaining enzymatic activity
Cathepsin B	98.5 ± 7.6
Cathepsin C	0 *

Cathepsin G	103.1 ± 2.8
Cathepsin L	37.7 ± 2.3 *
Cathepsin S	108.7 ± 1.1
Factor Xa	102.0 ± 3.8
Legumain	100.4 ± 1.8

Remaining enzymatic activity of tested enzymes in presence of 6 μ M Iristatin. * $p \leq 0.05$

Supplementary Methods

RT-PCR and primer design

Tick salivary glands were dissected on a petri dish under a drop of ice-cold DEPC-treated PBS. Total RNA was isolated from dissected tissue using the NucleoSpinRNA II kit (Macherey-Nagel, Düren, Germany), with quality checked by agarose gel electrophoresis before storing the RNA at -80°C . cDNA preparations were made from 0.5 μg of total RNA from independent biological triplicates using the Transcriptor High-Fidelity cDNA Synthesis Kit (Roche Diagnostics, Risch-Rotkreuz, Switzerland) according to the manufacturer's instructions. The cDNA served as templates for subsequent quantitative expression analyses of Iristatin transcription by qRT-PCR. Samples were analysed by a LightCycler 480 (Roche) using the LightCycler® 480 Probes Master kit (Roche). Relative expression was calculated using the $\Delta\Delta\text{Ct}$ method [3]. Expression profiles were normalised to *I. ricinus* ferritin mRNA, levels of which are independent of blood feeding [4]. The primer and probe sequences for Iristatin cloning and RT-PCR are shown in **Table S1**.

Iristatin cloning, expression, refolding, and purification

The full cDNA sequence of the gene encoding Iristatin was amplified using primers designed based on the orthologous *I. scapularis* cystatin genes (**Figure 2**). The exact primer sequences used for the final cloning of Iristatin are presented in **Table S1**. *I. ricinus* cDNA prepared from the salivary glands of female ticks fed for 3 and 6 days on rabbits was used as a template. The Iristatin gene without a signal peptide and with an inserted ATG codon (360 bp) was cloned into a pET-17b vector and transformed into *Escherichia coli* strain BL21(DE3)pLysS for expression. Bacterial cultures in LB medium with 100 mg/ml ampicillin and 34 mg/ml chloramphenicol were grown to an OD₆₀₀ of 0.8, and protein expression was induced by the addition of isopropyl 1-thio- β -D-galactopyranoside at a final concentration 1 mM. The cultures were harvested after 5 h of fermentation at 37°C . Following washing and recovery of inclusion bodies, DEAE Sepharose FF chromatography was performed under denaturing conditions, followed by three-step refolding by dialysis in 50 mM Tris, pH 8.5, 0.5 mM KCl, 100 mM NaCl, 50 mM betaine, 1 mM EDTA, 10% glycerol, and <0.02% Triton X114. The resulting refolded protein was purified by HiLoad Superdex 200 26/60 gel filtration chromatography and endotoxin removal was performed using a detergent-based method according to [5].

Phylogeny tree calculation

The evolutionary history of Iristatin was inferred using the maximum likelihood model based on the Dayhoff matrix-based model [6] using cDNA sequences without signal peptides. Initial tree(s) for the heuristic search were obtained by applying the neighbor-joining method to a matrix of pairwise distances estimated using a JTT model. There were 112 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [7].

Flow cytometry

Cell suspensions were incubated with rat anti-mouse CD16/CD32 monoclonal antibody (clone 93, eBioscience, Thermo Fisher Scientific, Waltham, MA) for 5 minutes on ice to block Fc gamma II/III receptors. Subsequently, cells were stained with population-specific antibodies for 30 minutes in the dark at 4°C. The following rat anti-mouse monoclonal antibodies were used: CD11b-FITC (clone M1/70, BD Biosciences, Franklin Lakes, NJ), Ly-6G-PE (clone 1A8; BD Biosciences), Ly-6C-APC (clone HK1.4; eBioscience), and Siglec-F-PerCP-CyTM5.5 (clone E50-2440, BD Pharmingen). Propidium iodide (eBioscience) was added to distinguish living and dead cells. Cell suspensions were analyzed by flow cytometry using BD FACSCantoTM II and BD FACSDivaTM Software v. 6.1.3. Neutrophils were identified as CD11b⁺Ly-6G⁺ cells, monocytes as CD11b⁺Ly-6C⁺ cells and eosinophils as CD11b⁺Siglec-F⁺.

Enzymatic assays

All assays were performed in triplicate. The mean percentage of remaining enzymatic activity in the presence of various Iristatin concentrations compared to control enzymatic activity (in the absence of Iristatin) was plotted against the concentration of Iristatin used in the assays and in logarithmic scale. Finally, sigmoidal fitting of the data gave an estimate of the IC₅₀ of Iristatin for the various enzymes.

Assay buffers were as follows: 100 mM Na-acetate, 100 mM NaCl, 1 mM EDTA, 0.01% Triton X-100, 100 µg/ml cysteine, pH 5.5 for cathepsins B, L, S; 50 mM Na-acetate, 50 mM NaCl, 5 mM DTT, pH 5.5 for cathepsin C; 20 mM Tris-HCl, pH 8, 200 mM NaCl, 5 mM CaCl₂, 0.1% BSA for factor Xa, 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.01% Triton X-100 for cathepsin G; 50 mM Na-acetate, 100 mM NaCl, 4 mM DTT, pH 5.2 for legumain.

Enzymes were used at the following final concentration: 10 pM cathepsin B, 500 pM cathepsin C, 10 nM cathepsin G, 33 pM cathepsin L, 350 pM cathepsin S, 800 pM factor Xa.

Substrates were as follows: Z-LR-AMC, 250 µM final concentration for cathepsins B, L; Z-VVR-AMC, 250 µM final concentration for cathepsin S; H-GR-AMC, 250 µM final concentration for cathepsin C; methylsulfonyl-D-cyclohexylalanyl-GR-AMC, 250 µM final concentration for factor Xa; Suc-AAPF-AMC, 250 µM final concentration for cathepsin G; Z-AAN-AMC, 250 µM final concentration for legumain.

The substrate hydrolysis rate was followed in an Infinite® 200 PRO 96-well plate fluorescence reader (Tecan, Männedorf, Switzerland) using 365 nm excitation and 450 nm emission wavelengths with a cutoff at 435 nm.

Crystallization, data collection, and structure determination

The crystallization drop consisted of 1 μ l of Iristatin protein solution (5 mg/ml in 10 mM Tris buffer, pH 8.0) and 1.5 μ l of the reservoir solution (0.1 M Bis-Tris, pH 5.5, 1 M ammonium sulfate, 1% PEG 3350). Brick-shaped crystals reached their final size of $0.4 \times 0.1 \times 0.1$ mm within 7 days. For data collection, crystals were soaked in reservoir solution supplemented with 12.5% glycerol and flash cooled in liquid nitrogen. Diffraction data were collected at 100 K using the MX-14-1 beamline at BESSY, Hamburg, Germany, and processed using the HKL-3000 suite of programs. Crystals exhibited the symmetry of space group $P2_12_12_1$ and contained two molecules in the asymmetric unit.

The phase problem was solved by molecular replacement using the program Molrep. The search model was derived from the structure of cystatin OmC2 from *O. moubata* (PDB code 3L0R), sharing 41% sequence identity with Iristatin. Model refinement was carried out using the program REFMAC 5.5 from the CCP4 package. Manual building was performed using Coot. Tight non-crystallographic symmetry (NCS) restraints were applied during initial refinement, and, in later stages, NCS restraints were loosened as guided by the behavior of Rfree. The final refinement steps included TLS refinement. The quality of the final model was validated with Molprobity. Final refinement statistics are given in **Table S2**. Figures showing structural representations were prepared with the PyMOL Molecular Graphics System (Schrödinger). Atomic coordinates and structure factors were deposited in the PDB under accession code 5O46.

Isoelectric focusing

Isoelectric focusing (IEF) was performed on precast Criterion™ IEF gels (Bio-Rad, Hercules, MA) and separated by IEF cathode buffer 2 mM lysine (free base), 2 mM arginine (free base) and IEF anode buffer 0.7 mM phosphoric acid at increased voltage modes: 100 V 60 min, 250 V 60 min, 500 V 30 min. IEF markers of pI range 3–10 were used (SERVA 39212.01).

The value of Iristatin pI predicted to be 7.67 (by <https://web.expasy.org/protparam>) was verified by isoelectric focusing. Iristatin showed a clear band between pI values 7.6 and 7.7 (**Figure S2**).

Iristatin shows no toxic effect on CD4⁺ cells

Naïve CD4⁺ T cells from Balb/c mice were stimulated with anti-CD3/CD28 (4 μ g/mL) in the presence and absence of different concentrations of LPS-free Iristatin (6, 3, 1.5, and 0.75 μ M) under Th9-skewing conditions for 72 h. T cells were stained with a fixable viability dye, and cell viability was determined by flow cytometry (**Figure S5**). The value of the untreated cells is considered 100%; the other values are expressed as percentages.

References

1. Brunger AT (1992) Free R-Value - a Novel Statistical Quantity for Assessing the Accuracy of Crystal-Structures. *Nature* 355 (6359):472-475

2. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography* 26:283-291
3. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29 (9):e45
4. Kopacek P, Zdychova J, Yoshiga T, Weise C, Rudenko N, Law JH (2003) Molecular cloning, expression and isolation of ferritins from two tick species--*Ornithodoros moubata* and *Ixodes ricinus*. *Insect Biochem Mol Biol* 33 (1):103-113
5. Sa-Nunes A, Bafica A, Antonelli LR, Choi EY, Francischetti IM, Andersen JF, Shi GP, Chavakis T, Ribeiro JM, Kotsyfakis M (2009) The immunomodulatory action of sialostatin L on dendritic cells reveals its potential to interfere with autoimmunity. *J Immunol* 182 (12):7422-7429. doi:10.4049/jimmunol.0900075
6. Schwarz R, Dayhoff MO (1979) Matrices for Detecting Distant Relationships. In: *Atlas of Protein Sequence and Structure*. vol sv. 5. National Biomedical Research Foundation.,
7. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33 (7):1870-1874. doi:10.1093/molbev/msw054

4.4. Nepublikované výsledky

Předcházející publikace č. 3 se zabývala iristatinem, imunomodulačním inhibitorem z podrodiny pravých cystatinů ze slin klíštěte obecného (*I. ricinus*). V návaznosti na tuto publikaci je připravena pro publikování další komplexní studie o homologickém inhibitoru ricistatinu ze slin *I. ricinus*, který vykazuje také imunomodulační účinky. Autor má na této publikaci sdílené první autorství.

Můj podíl na této studii zahrnoval:

(1) Analýzu inhibiční specifity ricistatinu, která umožňuje tento inhibitor zařadit do skupiny slinných klíšťových cystatinů s unikátní úzkou inhibiční specifitou cílenou preferenčně na cysteinové katepsiny s endopeptidasovou aktivitou.

(2) Řešení krystalové struktury komplexu ricistatinu s modelovým cysteinovým katepsinem. Tato strukturní analýza je první, kterou se podařilo získat pro slinný klíšťový cystatin (a parazitární cystatin obecně) a umožňuje vysvětlit strukturní příčinu inhibiční specifity těchto inhibitorů. Strukturní koordináty byly deponovány do databáze PDB s kódem 7PK4 a v současné době jsou již veřejně přístupné.

Tyto výsledky jsou uvedeny ve dvou následujících kapitolách.

4.4.1. Inhibiční specifita ricistatinu a její srovnání s ostatními cystatiny

Funkční biochemické vlastnosti rekombinantního ricistatinu byly analyzovány v kinetických inhibičních testech s hostitelskými cysteinovými proteasami z rodiny papainu a rodiny legumainu. Profil inhibiční specifity ricistatinu byl porovnán s reprezentativními zástupci cystatinové rodiny. Analýza zahrnovala dostupná systematická data publikovaná pro pravé cystatiny ze slin a střev klíšťat a modelové pravé cystatiny a stefiny obratlovců (Obr. 15). Ricistatin a ostatní slinné pravé cystatiny z klíšťat vykazují odlišný inhibiční profil, který je limitován pouze na některé cysteinové proteasy. Typicky inhibují endopeptidasy, jako je katepsin L nebo S (iristatin mezi nimi dokáže rozlišit) a exopeptidasu katepsin C. Na druhou stranu neinhibují exopeptidasové katepsiny B a H, které jsou ale velmi dobře inhibovány střevními cystatiny klíšťat – mialostatinem z *I. ricinus* a OmC2 z *Ornithodoros moubata* a dále všemi ostatními cystatiny obratlovců. Tato analýza umožňuje lépe definovat skupinu pravých cystatinů ze slin klíštěte s unikátním úzkým inhibičním profilem a klasifikovat ricistatin jako nového člena této skupiny.

		cystatiny									
		IC ₅₀						K _i			
		ricistatin	sialostatin L2	sialostatin L	iristatin	mialostatin	OmC2	CEW cystatin	cystatin C	stefin A	stefin B
proteasy	hCatL	10	1.1	1	10000	1	0.1	0.005	0.005	1.3	0.23
	hCatS	343	200	0.2	n.i.	2.2	0.1	0.11	0.008	0.05	0.07
	hCatC	748	680	670	500	2.1	1.1	0.35	3.5	33	0.23
	hCatH	n.i.	n.i.	n.i.	n.i.	29	1.2	0.064	0.28	0.31	0.58
	hCatB	n.i.	n.i.	n.i.	n.i.	10	8.8	1.7	0.25	8.2	73
	Legumain	n.i.	n.i.	n.i.	n.i.	n.i.	n.i.	5	0.2	n.i.	n.i.

Obr. 15: Inhibiční specifita ricistatinu a její porovnání s vybranými cystatiny. Inhibiční data pro ricistatin jsou porovnána s ostatními pravými cystatiny ze slin klíšťat, zahrnující sialostatin L and L2 (*I. scapularis*) a iristatin (*I. ricinus*), a s pravými cystatiny ze střev klíšťat, zahrnující mialostatin (*I. ricinus*) a OmC2 (*O. moubata*). Dále jsou uvedeny modelové cystatiny obratlovců: pravé cystatiny: CEW („Chicken Egg White“) cystatin a cystatin C a stefiny A a B. Inhibiční data proti lidským katepsinům L až B a savčímu legumainu jsou prezentována jako IC₅₀ (cystatiny z klíšťat (Kotál et al., 2021, 2019; Kotsyfakis et al., 2010; Salát et al., 2010)) nebo K_i (cystatiny obratlovců (Barrett, 1987; Maciewicz et al., 1987)) s intenzitou podbarvení odpovídající síle inhibice (zelená škála); n.i. – bez inhibice.

4.4.2. Krystalová struktura komplexu ricistatinu s katepsinem V a analýza vazebného módu

Byly připraveny krystaly komplexu ricistatinu s lidským katepsinem V (modelový cysteinový katepsin s endopeptidasovou aktivitou). K zajištění stability komplexu byl použit katalyticky neaktivní mutant Cys25Ser rekombinantního katepsinu V. Komplex krystalizoval v prostorové grupě C121 s jednou molekulou v asymetrické jednotce. Fázový problém byl vyřešen molekulárním nahrazením s využitím struktury katepsinu V (kód PDB: 1FH0)²⁰¹. Struktura komplexu byla vyřešena při rozlišení 1.9 Å (Tab. 4) a deponována do databáze PDB pod kódem 7PK4.

Molekula ricistatinu byla modelována pro oblast sekvence 4i–113i („i“ označuje zbytky inhibitoru), nedostatek elektronové hustoty neumožnil modelovat celkem 4 segmenty. Jedná se o smyčky 38i–47i a 68i–90i, které se nacházejí na opačné straně molekuly než interakční smyčky a jsou exponovány do prostředí. Dále chybí 3 zbytky na N-konci a 2 zbytky na C-konci

ricistatinu. Molekula katepsinu V byla modelována v celém rozsahu sekvence 0e–221e („e“ označuje zbytky enzymu, zbytek 0e pochází z nekompletně odštěpeného propeptidu), nedostatek elektronové hustoty neumožnil modelovat pouze povrchovou smyčku 175e–181e.

Na Obr. 16B je znázorněn model celkové struktury komplexu ricistatinu a katepsinu V. Ricistatin má typickou architekturu molekuly pravých cystatinů s pěti antiparalelními β -strukturami obtočenými kolem centrální α -šroubovice. Sekvenční a strukturální porovnání ricistatinu se známými cystatiny prokázalo, že ricistatin náleží do podrodiny pravých cystatinů (Obr. 16C a 16D). Jeho nejbližšími homology jsou slinné cystatiny z *I. scapularis*: sialostatins L2 (90 % sekvenční identity, RMSD pro C α atomy = 2,2 Å) a sialostatin L (78 % identity, monomerní struktura neexistuje). Molekula zralého katepsinu V je tvořena dvěma doménami s aktivním místem mezi nimi. Ve spodní části aktivního místa se nachází u přirozeného katepsinu V katalytický zbytek Cys25e, který byl u krystalizované inaktivní varianty mutovaný na Ser25e. Porovnáním se třemi dostupnými strukturami z databáze PDB²⁰¹⁻²⁰³ byly získány hodnoty RMSD pro C α atomy v rozsahu 0,2–0,3 Å, což ukazuje, že u katepsinu L nedošlo k výrazným celkovým konformačním změnám.

U cystatinů je interakce s cysteinovými katepsiny zprostředkována pomocí tří interakčních segmentů, N-konce a povrchových smyček L1 a L2. Dohromady tyto tři segmenty tvoří klín, který se zasouvá do aktivního místa proteasy (Obr. 16A a 16B). U ricistatinu je prvním modelovaným zbytkem N-konce Leu4i, který vyplňuje svým bočním řetězcem podmísto S3 katepsinu V, následuje Arg5i, který bočním řetězcem vyplňuje podmísto S2 a s podmístem S1 interagují zbytky Gly6i, Gly7i a Tyr9i. Konzervovaný pár glycinů Gly6i–Gly7i poskytuje N-konci důležitou konformační flexibilitu. U ricistatinu je N-konec zpevněn Arg9i, který stabilizuje jeho konformaci pomocí intramolekulárních vodíkových vazeb k Leu4i a Gly7i. Hlavní interakční zbytky N-konce (region Leu4i–Arg9i) tvoří s katepsinem V celkem 4 vodíkové můstky a 2 hydrofobní interakce. Jejich partnery jsou zbytky Asn66e, Gly68e–Met70e a Leu162e–Asp163e. Vybrané interakce jsou znázorněny na Obr. 16B.

Smyčka L1 ricistatinu spojující β 1- a β 2-struktury má sekvenci Gln53i–Thr54i–Val55i–Ala56i–Gly57i, která odpovídá motivu Gln–Xaa–Val–Xaa–Gly konzervovanému u cystatinů (Obr. 16C). Interakční oblast smyčky L1 je delší než konzervovaný motiv a končí až zbytkem Arg61i. Tato smyčka vyplňuje podmísto S1 zbytkem Gln53i, podmísto S1' zbytkem Thr54i a interaguje i s podmístem S3'. Hlavní interakční zbytky tohoto segmentu inhibitoru jsou Gln53i, Thr54i, Ala56i, Asn59i a Arg61i tvořící celkem 5 vodíkových můstků a 1 hydrofobní interakci. Ty interagují s následujícími zbytky katepsinu V: Gln19e–Gln21e, Gly23e, Cys65e–Asn66e, Asp163e a Trp190e (Obr. 16B). Na vrcholu smyčky L1 je Val55i, který sice nevytváří žádnou

interakci se zbytky katepsinu V, ale přesně vyplňuje danou část aktivního místa. Gly57i zajišťuje strukturní flexibilitu.

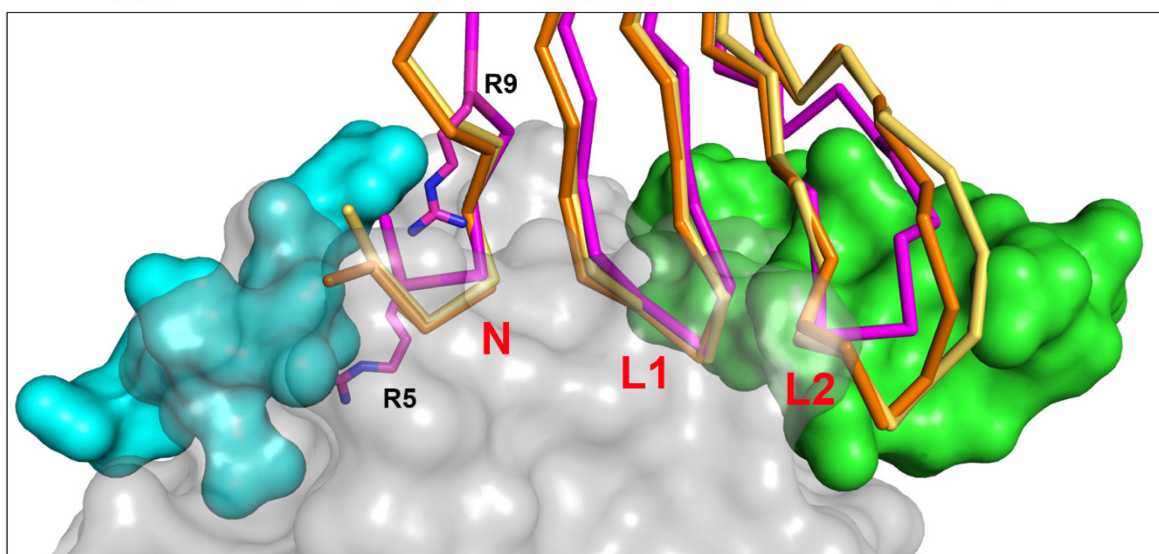
Smyčka L2 mezi β 4- a β 5-strukturami ricistatinu se váže do široké hydrofobní oblasti obsahující špatně definovaná podmísta S2'až S4'. Tento segment nevytváří žádné elektrostatické interakce, ale zbytky Tyr98i a Leu101i vytvářejí 3 hydrofobní interakce se zbytky Gln21e, Gln145e a Trp190e z katepsinu V (Obr. 16B).

Interakční plocha mezi ricistatinem a katepsinem V činí 751 \AA^2 (Obr. 16F). Hydrofobní interakce jsou rozděleny mezi obě domény rovnoměrně, zatímco všechny vodíkové můstky (Obr. 16B) mají své partnery pouze v L-doméně katepsinu V. Z těchto celkem 9 vodíkových můstků jsou 4 tvořeny hlavním řetězcem ricistatinu a tyto interakce jsou konzervované v dostupných komplexech cystatinů. Zbýlých 5 vodíkových můstků vychází z bočních řetězců a jsou u ostatních komplexů variabilní. Žádný ze zbytků ricistatinu neinteraguje přímo s katalytickými zbytky katepsinu V (Ser25e a His164e), které se nacházejí v prostoru pod segmenty N-konce a L1 smyčky.

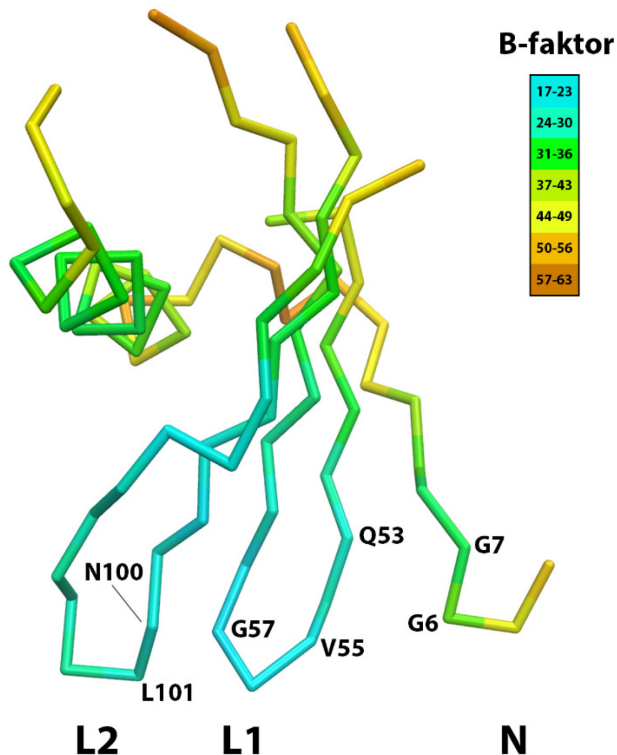
Dále byla provedena analýza konformačních změn probíhajících v molekule ricistatinu během vazby do aktivního místa katepsinu V (Obr. 16E) a analýza strukturních příčin zodpovědných za úzkou inhibiční specifitu ricistatinu (Obr. 17 a 18) – tento rozbor je uveden v kapitole 5.2. této disertační práce.

(ii) strukturně charakterizované cystatiny v komplexu s cysteinovými katepsiny (CEW cystatin, lidský cystatin C, lidské stefiny A a B a rostlinný tarocystatin z *Colocasia esculenta*). Byly použity sekvence zralých proteinů (+ označuje pokračování sekvence); číslování odpovídá ricistatinu. Sekundární struktury ricistatinu jsou vyznačeny nad sekvencí a popsány jako v panelu (A). Tři interakční segmenty zodpovědné za interakci s cysteinovými katepsiny jsou rámovány červeně a popsány jako v panelu (A); konzervované zbytky těchto segmentů jsou červeně, zbytky identické s ricistatinem jsou podbarveny šedě. Motiv pro vazbu legumainu u CEW cystatinu je modře. Cysteiny tvořící disulfidické můstky jsou podbarveny černě a jejich párování je vyznačeno zelenými symboly. **(D)** Superpozice hlavního řetězce ricistatinu (fialová) s ostatními cystatiny strukturně charakterizovanými v komplexu s cysteinovými katepsiny: CEW cystatin (zelená; komplex s falcipainem-2, PDB: 1YVB), stefin A (oranžová; komplex s katepsinem V, PDB: 3KFQ), stefin B (černá; komplex s papainem, PDB: 1STF) a tarocystatin (růžová; komplex s papainem, PDB: 3IMA). Orientace ricistatinu je jako v panelu (B). **(E)** Strukturní porovnání reaktivního centra cystatinů vázaných v komplexu – ricistatinu (fialová) a CEW cystatinu (zelená) s volným sialostatinem L2 (modrá) jako nejbližším homologem ricistatinu. Orientace molekuly je jako v panelu (B). Vybrané interakční zbytky s významně odlišnou konformací jsou znázorněny v tyčkovém modelu; atomy kyslíku jsou červeně, dusíku modře. N-konce jednotlivých cystatinů jsou označeny hvězdičkou s odpovídající barvou; aminokyselinové zbytky se stejnou pozicí v panelu (C) jsou označeny shodnými horními indexy (a-c). **(F)** Interakční povrch mezi ricistatinem a katepsinem V. Struktura komplexu je symetricky otevřena a ukazuje interakční povrch obou proteinů. Barva odpovídá skóre zanoření povrchu („buried surface area“ – b.s.a.). Peptidomimetický inhibitor K11777 (tyčkový model, azurová; PDB: 1FH0) je vložen superpozicí do aktivního místa katepsinu V a umožňuje lokalizovat jednotlivá podmísta S3-S1'. Na povrchu ricistatinu jsou označeny tři interakční segmenty: N-konec a smyčky L1 a L2.

- | | |
|--------------------------------|-----------------------------------|
| ■ katepsin V | ■ ricistatin (katepsin V komplex) |
| ■ mini-řetězec katepsinu H | ■ stefin A (katepsin H komplex) |
| ■ „occluding loop“ katepsinu B | ■ stefin A (katepsin B komplex) |



Obr. 17: Strukturní modifikace určující úzký inhibiční profil ricistatinu. Superpozice ricistatinu (fialově) z komplexu s endopeptidasou katepsinem V (šedý povrch) a stefinu A z komplexu s aminopeptidasou katepsinem H (žlutě, PDB: 1NB5) a stefinu A z komplexu s karboxypeptidasou katepsinem B (oranžově, PDB: 3K9M). Inhibiční segmenty (N-konec a smyčky L1 a L2) a zbytky Arg5 a Arg9 ricistatinu jsou popsány. Strukturní elementy limitující vazbu substrátu do aktivního místa katepsinu H (mini-řetězec, azurově) a katepsinu B („occluding loop“, zeleně) z komplexů se stefinem A jsou zobrazeny v povrchové reprezentaci přes šedý povrch endopeptidasy katepsinu V.



Obr. 18: Distribuce krystalografických B-faktorů pro ricistatin z komplexu s katepsinem V. Molekula ricistatinu je zobrazena jako hlavní řetězec ricistatinu a zbarvena dle hodnot B-faktorů pro $C\alpha$ atomy od nejnižší (azurová) po nejvyšší (hnědá). Nejnižší B-faktory indikující strukturní rigiditu vykazují smyčky L1 a L2 vázané v aktivním místě katepsinu V. U širokospektrých cystatinů je naopak L2 smyčka velmi flexibilní s relativními B-faktory více než dvojnásobnými v porovnání s L1 smyčkou (např. u vázaného CEW cystatinu a stefinu A; PDB kódy: 1YVB respektive 1NB5). Interakční segmenty (N-konec, smyčky L1 a L2), konzervované zbytky a barevná škála B-faktorů je popsána.

Tab. 4: Krystalové parametry, statistiky difrakčních dat a statistiky upřesněné struktury.

Statistiky byly generovány programem PHENIX²⁰⁰. Hodnoty v závorkách popisují slupku nejvyššího rozlišení.

Vlnová délka (Å)	1,54
Rozsah rozlišení (Å)	38,89–1,92 (1,99–1,92)
Prostorová grupa	C121
Základní buňka	
a, b, c (Å)	117,49 45,32 58,52
α, β, γ (°)	90,00 96,89 90,00
Celkový počet reflexí	70678 (5135)
Počet unikátních reflexí	23274 (2237)
Multiplicita	3,0 (2,3)
Kompletnost (%)	99,03 (95,47)
Průměrné I/sigma(I)	7,81 (0,98)
Wilsonův B-faktor (Å ²)	22,16
R _{merge}	0,11 (0,92)
R _{meas}	0,14 (1,16)
R _{pim}	0,077 (0,70)
CC _(1/2)	0,99 (0,45)
CC*	0,99 (0,79)
Počet reflexí použitý při upřesňování	23397 (2236)
Počet reflexí generaci R _{free} sady	1151 (137)
R _{work}	0,20 (0,32)
R _{free}	0,22 (0,35)
CC _{work}	0,95 (0,58)
CC _{free}	0,95 (0,51)
Počet nevodíkových atomů	2518
Makromolekuly	2294
Ligandy	37
Solvent	187

Počet aminokyselinových zbytků	298
Průměrná kvadratická odchylka vazeb (Å)	0,015
Průměrná kvadratická odchylka úhlů (°)	1,81
Ramachandranem preferované zbytky (%)	97,6
Ramachandranem povolené zbytky (%)	2,1
Ramachandranem zakázané zbytky (%)	0,4
Odlehlé rotamery (%)	2,1
Průměrný B-faktor (Å ²)	29,55
Makromolekuly	28,78
Ligandy	51,43
Solvent	34,78

5. Diskuse

Cysteinové katepsiny mají širokou distribuci a zajišťují mnoho fyziologicky významných dějů v organismu parazitů i v tělech jejich hostitelů. Parazité produkují inhibitory cysteinových katepsinů k regulaci jak endogenních procesů, tak je sekretují do tkání hostitele k ovlivnění proteolytických rovnováh. Významnou rodinou inhibitorů cysteinových katepsinů jsou cystatiny (ID databáze Merops: I25), které jsou přítomny v proteomu parazitů i savčích hostitelů včetně člověka. Rodina cystatinů obsahuje dvě hlavní podrodiny monomerních proteinových inhibitorů: intracelulární stefiny a sekretované pravé cystatiny, které se liší od stefinů přítomností signální sekvence a disulfidických můstků. Tato disertační práce se zabývá zástupci rodiny cystatinů ze dvou významných krevsajících parazitů. Prvním je motolice jaterní (*Fasciola hepatica*), parazitický helmint způsobující onemocnění fasciolózu s obrovským hospodářským dopadem. Druhým je klíště obecné (*Ixodes ricinus*), ektoparazitický členovec a hlavní evropský vektor klíšťové encefalitidy a lymfské boreliózy. Diskuse je rozdělena do dvou kapitol věnovaných cystatinům z těchto dvou parazitů.

5.1. Identifikace a charakterizace nového typu cystatinů u motolice *F. hepatica*.

Součástí disertační práce je studie, která se zabývala FhCyLS-2, neobvyklým členem cystatinové rodiny z *F. hepatica*. Sekvenční a strukturní analýza odhalila, že FhCyLS-2 bezprecedentním způsobem kombinuje charakteristické znaky dvou hlavních podrodin cystatinové rodiny, stefinů a pravých cystatinů. Sekvenční homologie prokázala příslušnost FhCyLS-2 ke stefinům; FhCyLS-2 však obsahuje také signální peptid pro sekreci a disulfidy, tedy dva typické molekulární znaky pro pravé cystatiny. Bylo zjištěno, že FhCyLS-2 je přítomen v ESP dospělých motolic a je zjevně uvolňován sekrecí ze střevních buněk. Vylučovaný FhCyLS-2 byl rozpoznán imunitním systémem hostitele během infekce *F. hepatica*, což vedlo k produkci specifických protilátek detekovaných v séru. Byla určena prostorová struktura FhCyLS-2 rentgenostrukturní analýzou a ukázala, že do obecné architektury stefinů byly inkorporovány dva disulfidické můstky. Nacházejí se v oblastech molekuly, které jsou blízké, nikoliv však totožné s umístěním disulfidů u pravých cystatinů. Akvizice disulfidů nejspíše souvisí se sekrečním charakterem FhCyLS-2, který jej významně odlišuje od typicky intracelulárních stefinů. Obecná úloha disulfidických můstků spočívá ve stabilizaci proteinů ve více oxidativním extracelulárním prostředí, a jsou proto přítomny u většiny sekretovaných proteinů, zatímco v cytosolárních proteinech jsou vzácné²⁰⁴.

Molekula FhCyLS-2 nese funkční reaktivní centrum pro interakci s proteasami, které má analogickou stavbu jako u dalších členů rodiny. FhCyLS-2 je efektivním inhibitorem řady testovaných cysteinových katepsinů, a to jak z hostitele, tak z *F. hepatica*. Porovnání inhibiční specifity FhCyLS-2 s ostatními členy rodiny ukázalo, že FhCyLS-2 inhibuje cysteinové katepsiny s endopeptidasovou i exopeptidasovou aktivitou. Podobá se tak některým již popsáním inhibitorům s širokou specifitou, jako jsou lidské a klíšťové stefiny, klíšťové střevní cystatiny i někteří zástupci lidských pravých cystatinů. Cysteinové katepsiny citlivé na FhCyLS-2 hrají u savců klíčovou roli v řadě imunologických procesů, např. zpracování antigenů a aktivace neutrofilů, cytotoxických NK buněk a T-lymfocytů^{205,206}.

Proto lze předpokládat, že sekretovaný FhCyLS-2 napodobuje určité cystatiny produkované hostitelem a zasahuje do jejich funkcí in vivo. Tím mění proteolytickou rovnováhu podobně, jako bylo dříve zjištěno u cystatinů helmintů a klíšťat s různými imunomodulačními účinky^{166,207}. Navíc FhCyLS-2 inhibuje hlavní sekretované trávicí katepsiny L z *F. hepatica* a může tak sloužit jako endogenní regulátor s cílem chránit parazitární tkáň před autoproteolytickým poškozením. FhCyLS-2 tedy může mít dvojí roli v regulaci exogenních i endogenních proteolytických systémů při interakci parazita s hostitelem. Za zmínku stojí, že podobná kombinovaná funkce byla navrhována pro klíšťový cystatin OmC2, který je vylučován do hostitele za účelem imunomodulačního působení a zároveň reguluje proteolýzu ve střevu klíšťe²⁰⁸.

Z kombinace sekvenčních a strukturních znaků a fylogenetické analýzy je zřejmé, že FhCyLS-2 patří do podrodiny stefinů. U parazitických trematod a myxozoí byla identifikována sada homologů FhCyLS-2. Ty se evolučně liší od klasických stefinů a tvoří samostatnou skupinu na fylogenetické větvi stefinů. Na základě toho je nyní možné v rámci cystatinové rodiny definovat novou skupinu nazvanou CyL-stefiny („Cystatin-Like stefins“) s FhCyLS-2 jako modelovým zástupcem. CyL-stefiny nesou signální peptid na stefinové sekvenci, která postrádá C-terminální motiv Asp/Glu-Xxx-Leu-Xxx-Tyr/His-Phe. Struktura CyL-stefinů obvykle obsahuje dva disulfidy, jejich počet však může být snížen na jeden nebo vzácně zcela chybět.

Naprostá většina organismů exprimuje jak klasické stefiny, tak pravé cystatiny jako nezbytné antiproteolytické molekuly^{118,209}. Analýza repertoáru genů rodiny cystatinů, včetně CyL-stefinů, u jednotlivých skupin motolic ukázala významné rozdíly při srovnání řádů Plagiorchiida, Opisthorchiida a Schistosomatida. *F. hepatica* a všechny ostatní jaterní a střevní motolice (Plagiorchiida) mají v genomu jak CyL-stefiny, tak klasické stefiny, ale postrádají geny pro pravé cystatiny. Naopak u krevních motolic (Schistosomatida), které disponují

klasickými stefiny a pravými cystatiny, nebyly identifikovány žádné CyL-stefiny. Motolice z rodu *Opisthorchiida* ovšem kombinují všechny tři skupiny inhibitorů. Kromě toho *Plagiorchiida* a *Opisthorchiida* navíc exprimují multicystatiny. Lze předpokládat, že CyL-stefiny představují evoluční vylepšení klasických stefinů, které nejspíše představuje molekulární adaptaci na absenci nebo nedostatečný funkční rozsah sekretovaných pravých cystatinů. Lze spekulovat, že různé zastoupení cystatinů ve výše zmíněných řádech odráží složitost životních cyklů motolic a strategie spojené s přežitím parazita v hostiteli. Obecně je širší panel zástupců rodiny cystatinů spojen s vyšší komplikovaností vícehostitelských životních cyklů, s výraznějším zapojením obratlovců jako hostitelů (disponujících pokročilejší imunitou než bezobratlí hostitelé) a více typy napadených tkání obsahujících rozdílné spektrum hostitelských proteas^{210,211}.

5.2. Strukturní vysvětlení inhibiční specifity cystatinů z klíštěte *I. ricinus*.

Druhá část disertační práce se věnovala třem novým zástupcům podrodiny pravých cystatinů, které byly identifikovány ve dvou tkáních klíčových pro proces sání klíšťat. Iristatin a ricistatin jsou produkovány ve slinných žlázách a sekretovány do slin klíštěte, zatímco mialostatin se nachází ve střevu klíštěte. Všechny tři inhibitory byly strukturně, funkčně a fylogeneticky charakterizovány a byly navrženy jejich biologické role.

Mialostatin byl imunohistochemicky lokalizován ve střevní tkáni *I. ricinus*. Tato tkáň je zodpovědná za intracelulární (a zřejmě i extracelulární) zpracování krevních proteinů hostitele prostřednictvím multienzymového systému využívajícího především cysteinové katepsiny⁶⁶. Mialostatin byl charakterizován jako efektivní inhibitor trávicích cysteinových katepsinů *I. ricinus* zahrnujících jednak katepsiny L1 a L3 (IrCL1 a IrCL3) s endopeptidasovou aktivitou a jednak katepsiny B a C (IrCB, IrCC) s exopeptidasovou aktivitou. Jeho široká inhibiční specifita je ve zjevném kontrastu se slinnými cystatiny klíšťat z rodu *Ixodes*, jako jsou dříve studované sialostatiny L a L2, které mají podstatně užší spektrum inhibovaných cysteinových katepsinů^{168,170}. Na druhou stranu podobně široká antiproteasová aktivita byla zaznamenána u OmC2 a částečně také u OmC1 přítomných ve střevu klíštěte *O. moubata*^{208,212}. Krystalová struktura identifikovala mialostatin jako blízký homolog OmC2. Dále poskytla vysvětlení vazebné selektivity mialostatinu, když ukázala analogickou architekturu jeho reaktivního centra, jako mají ostatní dostupné struktury zástupců pravých cystatinů obratlovců, které také vykazují širokou specifitu zahrnující cysteinové katepsiny s endopeptidasovou a exopeptidasovou aktivitou.

Na základě získaných funkčních dat a informací o lokalizaci lze navrhnout fyziologickou roli mialostatinu. Jde o účinný inhibitor všech trávicích cysteinových katepsinů *I. ricinus* včetně nedávno identifikovaného izoenzymu IrCL3²¹³, který působí v pozdní fázi trávení a se kterým se mialostatin výrazně kolokalizuje. Dále bylo zjištěno v buněčných testech, že mialostatin nevykazuje potenciální imunomodulační účinky na hostitele, které jsou dobře známé u cystatinů sekretovaných ve slinách klíšťat rodu *Ixodes*^{168,170}. Lze předpokládat, že mialostatin reguluje endogenní proteolytický systém ve střevu klíštěte, jehož regulace nebyla zatím studována. Pravděpodobně chrání střevní tkáň před proteolytickým poškozením a pomáhá při řízení a načasování jednotlivých kroků trávicího procesu.

Další dva studované pravé cystatiny jsou ricistatin a iristatin, které jsou produkované slinnými žlázami *I. ricinus*. U obou těchto cystatinů byly nalezeny podobné protizánětlivé a imunomodulační účinky na imunitní systém hostitele. Ricistatin a iristatin vykazují podobné, ačkoli ne identické inhibiční profily, jejichž společným rysem je selektivita pouze k některým cysteinovým katepsinům. Konkrétně inhibují endopeptidasy, jako je katepsin L nebo S (iristatin mezi nimi dokáže rozlišit), a exopeptidasu katepsin C. Na druhou stranu neinhibují exopeptidasové katepsiny B a H. Tato analýza umožnila lépe definovat skupinu pravých cystatinů ze slin klíštěte s unikátním úzkým inhibičním profilem, který je výrazně odlišuje od většiny dosud detailně charakterizovaných zástupců podrodiny pravých cystatinů z parazitů i jejich hostitelů.

Získaná krystalová struktura ricistatinu v komplexu s katepsinem V (modelovou endopeptidasou) poprvé umožňuje přímým způsobem vysvětlit strukturní podstatu úzké inhibiční specifity slinných klíšťových cystatinů. U cystatinů je interakce s cysteinovými katepsiny zprostředkována pomocí tří interakčních segmentů, N-konce a povrchových smyček L1 a L2, které se zasouvají do aktivního místa katepsinu. Díky porovnání inhibičního módu ricistatinu s ostatními dostupnými komplexy cystatinů s katepsiny byly nalezeny tři strukturní mechanismy, které omezují interakci úzkospektrých pravých cystatinů a exopeptidas. Ty totiž, na rozdíl od endopeptidas, disponují strukturními doménami limitujícími vazbu substrátu do S' resp. S podmíst.

U karboxydiptidasy katepsinu B se tato doména nazývá „occluding loop“ a blokuje přístup substrátu do podmísta S3'^{72,214}, do kterého se váže interakční smyčka L2 cystatinů. Ze strukturní analýzy dostupných komplexů cystatinů vyplývá, že smyčka L2 ricistatinu (a homologních sialostatinů) je o jeden či dva zbytky kratší a hůře vyplňuje hydrofobní podmísta S2' a S3' v porovnání se širokospektrými cystatiny (pravé cystatiny resp. stefiny). Dále, dle analýzy relativních krystalografických B-faktorů je tato smyčka navíc mnohem

rigidnější než u všech ostatních cystatinů v komplexu. Upravená smyčka L2 ricistatinu tak nejspíš není schopna odsunout „occluding loop“, jako je tomu např. u komplexu katepsinu B a stefinu A²¹⁵. Kombinace vysoké rigidity a nedokonalého vyplnění podmíst S2' a S3' kratší smyčkou L2 tak nejspíše způsobuje ztrátu afinity slinných klíšťových cystatinů ke katepsinu B.

Aminopeptidasa katepsin H využívá k omezení přístupu substrátu do podmíst S3 a S2 mini-retězec, oktapeptid, který pochází z propeptidu a je připojen k protease disulfidickým můstkem¹³. Komplex katepsinu H se stefinem A ukazuje, že N-konec stefinu je velmi flexibilní a dokáže se přizpůsobit i stericky velmi omezenému prostoru v okolí mini-řetězce. Nicméně u ricistatinu je toto konformační přizpůsobení velmi nepravděpodobné, jelikož N-konec ricistatinu obsahuje příliš objemný zbytek Arg5i, který se váže do pozice S2 již obsazení mini-řetězcem katepsinu H. Další zbytek Arg9i navíc zpevňuje konformaci celého N-konce ricistatinu pomocí intramolekulárních vodíkových můstků. Přítomnost velkých objemných zbytků lemující konzervovaný motiv Gly-Gly, je společná pro všechny pravé cystatiny ze slin klíšťat s úzkou inhibiční specifitou (Obr. 16C, kap. 4.4.2.).

Druhým studovaným slinným cystatinem byl iristatin. Jeho inhibiční specifita má hlavní rysy ricistatinu a dalších slinných cystatinů klíšťat a je omezena pouze na katepsiny L a C. Analýza sekvence a krystalové struktury iristatinu nicméně ukázala na jiný strukturní mechanismus, kterým je dosaženo limitované inhibiční specifity. N-koncový segment cystatinů obsahuje konzervovanou sekvenci Gly-Gly, zajišťující flexibilitu nezbytnou pro efektivní interakci se širokým spektrem cílových proteas, včetně exopeptidas. Iristatin disponuje sekvencí Gly-Val a lze tedy předpokládat, že záměna Val za Gly způsobí zvýšení rigidity N-konce, který tak bude kompatibilní jen s určitými interagujícími proteasami, jelikož i on nese na některých pozicích velké, objemné zbytky. Smyčka L2 má u všech cystatinů konzervovaný sekvenční motiv Gln-Xxx-Val-Xxx-Gly, který efektivně vyplňuje kavitu aktivního místa. U iristatinu je tento motiv modifikován záměnou Val za prostorově objemnější Ile. Val naprosto přesně vyplňuje prostor mezi aktivním místem a ostatními zbytky smyčky L2, jeho nahrazení za větší zbytek bude mít za následek silné omezení interakce s proteasami. Smyčka L2 zachovává Pro-Trp motiv konzervovaný u pravých širokospektrých cystatinů, takže za úzkou inhibiční specifitu iristatinu jsou zjevně zodpovědné kombinace strukturních efektů na N-konci a smyčce L1. Strukturní základ úzké inhibiční specifity iristatinu je tedy odlišný od ricistatinu, který k omezení inhibiční specifity využívá kombinace strukturních efektů na N-konci a smyčce L2.

Provedená fylogenetická analýza podrodiny pravých cystatinů z klíšťat ukázala, že u klíšťat rodu *Ixodes* existují tři různé skupiny pravých cystatinů. (1) Skupina ricistatinu

obsahuje imunomodulační slinné cystatiny s úzkou inhibiční specifitou, jejíž strukturní podstatu vysvětluje nyní získaná první krystalová struktura komplexu s proteasou. Dalšími významnými zástupci jsou dva dříve charakterizované sialostatiny z *I. scapularis* ^{168,170}. (2) Skupina iristatinu je evolučně odlišná od skupiny ricistatinu. Na základě analýzy jejího prvního modelového zástupce iristatinu jde opět o cystatiny s úzkou inhibiční specifitou, která je zjevně příčinou jejich imunomodulačních vlastností. Krystalová struktura iristatinu ukazuje, že inhibiční specifita má jiný strukturní základ, než je tomu u skupiny ricistatinu. Funkční vlastnosti byly získány pravděpodobně konvergentní evolucí, která souvisí s rolí ve slinách klíšťat, kde doprovázejí zástupce skupiny ricistatinu. (3) Skupina mialostatatinu obsahuje střevní cystatiny, které jsou z pohledu sekvence a prostorové struktury blíží ostatním zástupcům podrodiny pravých cystatinů z obratlovců než slinným klíšťovým cystatinům. Široká inhibiční specifita mialostatatinu z něj činí efektivní nástroj k regulaci multienzymového proteolytického systému ve střevu klíštěte.

6. Závěry

Disertační práce se zabývala charakterizací nových proteasových inhibitorů z proteinové rodiny cystatinů. Byla zaměřena na celkem čtyři parazitární cystatiny pocházející jednak z klíštěte obecného (*Ixodes ricinus*), hlavního vektoru lymské boreliózy a klíšťové encefalitidy, jednak z motolice jaterní (*Fasciola hepatica*), původce fasciolózy. Disertační práce obsahuje výsledky uveřejněné ve třech původních publikacích a dále nepublikovaná data, která se v současné době připravují ke zveřejnění.

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

Cystatin FhCyLS-2 z *Fasciola hepatica*

- FhCyLS-2 byl lokalizován ve střevní tkáni a sekretech *F. hepatica*. Rekombinantní FhCyLS-2 je efektivním inhibitorem jak trávicích cysteinových katepsinů motolice, tak cysteinových katepsinů hostitele a hraje pravděpodobně duální roli při regulaci endogenních i exogenních proteolytických systémů.
- Analýzou sekvence a krystalové struktury bylo zjištěno, že FhCyLS-2 patří do podrodiny stefinů, ale přitom přejímá unikátním způsobem některé typické znaky podrodiny pravých cystatinů. Jde o evoluční modifikaci, která umožnila sekreci a stabilizaci molekuly FhCyLS-2.
- Fylogenetická analýza prokázala existenci homologů FhCyLS-2 u několika taxonů motolic a umožnila definovat novou skupinu nazvanou CyL-stefiny („Cystatin-Like stefins“) v rámci cystatinové rodiny a navrhnout důvody jejího vzniku.

Cystatiny ricistatin, iristatin a mialostatin z *Ixodes ricinus*

- Byly identifikovány tři nové pravé cystatiny z *I. ricinus*, které se liší lokalizací a fylogenezí. U rekombinantních proteinů byly určeny jejich funkční vlastnosti, zejména inhibiční specifita k cílovým cysteinovým katepsinům, a byla vyřešena jejich krystalová struktura. Získané výsledky umožnily definovat následující tři evoluční skupiny pravých cystatinů u klíšťat rodu *Ixodes*.

- Skupina ricistatinu obsahuje cystatiny ze slin klíštěte a s imunomodulačními vlastnostmi. Jejich unikátní úzkou inhibiční specifitu se podařilo vysvětlit na strukturní úrovni pomocí krystalografie komplexu ricistatinu s katepsinem V, která umožnila první přímou analýzu interakce u parazitárního cystatinu. Omezená schopnost ricistatinu inhibovat katepsiny s exopeptidasovou aktivitou je dána strukturními modifikacemi dvou segmentů (N a L2) reaktivního centra ricistatinu.
- Iristatin jako modelový zástupce další skupiny slinných cystatinů má také imunomodulační vlastnosti. Ty jsou způsobeny úzkou inhibiční specifitou iristatinu, která je ještě omezenější než u ricistatinu. Krystalová struktura iristatinu ukazuje, že inhibiční specifita má jiný strukturní základ, než je tomu u skupiny ricistatinu a je založena na stavbě vazebných segmentů N a L1. Obě skupiny slinných cystatinů tedy evolučně konvergují k úzké inhibiční specifitě zajišťující fyziologickou roli při interakci klíšťat s hostitelem.
- Mialostatin je první detailně popsáný zástupce skupiny střevních cystatinů z rodu *Ixodes*. Má širokou inhibiční specifitu, která umožňuje interakci s cysteinovými katepsiny klíšťat s endo- a exopeptidasovou aktivitou. Tu vysvětluje i architektura reaktivního centra, jež nemá strukturní omezení nalezené u slinných cystatinů. Funkční vlastnosti činí z mialostatinu efektivní nástroj k regulaci trávicího multienzymového proteolytického systému ve střevu klíštěte.

Disertační práce přináší nové významné informace o evoluci a strukturně-funkčních vztazích v rodině cystatinových inhibitorů proteas. Studované parazitární cystatiny představují vzhledem ke svým významným biologickým rolím, potenciální antigeny pro vývoj antiparazitárních vakcín. Dále mají identifikované slinné cystatiny farmakologicky zajímavé aktivity s možným využitím v biomedicině.

6.1. Prohlášení spoluautorů

Michal Buša se významně podílel na přípravě publikací uvedených v této disertační práci.

Publikace č. 1: **An evolutionary molecular adaptation of an unusual stefin from the liver fluke *Fasciola hepatica* redefines the cystatin superfamily.**, Buša M., Matoušková Z., Bartošová-Sojková P., Páchl P., Řezáčová P., Eichenberger R. M., Deplazes P., Horn M., Štefanić S., Mareš, M., *Journal of Biological Chemistry*, 2023, (IF = 5,5)

Publikace č. 2: **Mialostatin, a Novel Midgut Cystatin from *Ixodes ricinus* Ticks: Crystal Structure and Regulation of Host Blood Digestion.**, Kotál J, Buša M, Urbanová V, Řezáčová P, Chmelař J, Langhansová H, Sojka D, Mareš M, Kotsyfakis M., *International Journal of Molecular Sciences*, 22(10):5371, 2021

Publikace č. 3: **The structure and function of Iristatin, a novel immunosuppressive tick salivary cystatin.**, Kotál J, Stergiou N, Buša M, Chlastáková A, Beránková Z, Řezáčová P, Langhansová H, Schwarz A, Calvo E, Kopecký J, Mareš M, Schmitt E, Chmelař J, Kotsyfakis M., *Cellular and Molecular Life Sciences*, 76(10):2003-2013, 2019

.....
RNDr. Michael Mareš, CSc.

školitel a spoluautor všech uvedených publikací

7. Seznam použité literatury

- 1 Bond, J. S. Proteases: History, discovery, and roles in health and disease. *J Biol Chem* **294**, 1643-1651, doi:10.1074/jbc.TM118.004156 (2019).
- 2 Rawlings, N. D. *et al.* 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).
- 3 Yadati, T. *et al.* The Ins and Outs of Cathepsins: Physiological Function and Role in Disease Management. *Cells* **9**, doi:ARTN 1679 10.3390/cells9071679 (2020).
- 4 Turk, D. *et al.* Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO J* **20**, 6570-6582, doi:10.1093/emboj/20.23.6570 (2001).
- 5 Kamphuis, I. G. *et al.* Structure of Papain Refined at 1.65 a Resolution. *J Mol Biol* **179**, 233-256, doi:Doi 10.1016/0022-2836(84)90467-4 (1984).
- 6 Yamamoto, D. *et al.* Refined X-Ray Structure of Papain.E-64-C Complex at 2.1-a Resolution. *J Biol Chem* **266**, 14771-14777 (1991).
- 7 Storer, A. C. *et al.* Catalytic mechanism in papain family of cysteine peptidases. *Methods Enzymol* **244**, 486-500, doi:10.1016/0076-6879(94)44035-2 (1994).
- 8 Shamsi, A. *et al.* Journey of cystatins from being mere thiol protease inhibitors to at heart of many pathological conditions. *Int J Biol Macromol* **102**, 674-693, doi:10.1016/j.ijbiomac.2017.04.071 (2017).
- 9 Song, J. *et al.* Bioinformatic approaches for predicting substrates of proteases. *J Bioinform Comput Biol* **9**, 149-178, doi:10.1142/s0219720011005288 (2011).
- 10 Choe, Y. *et al.* Substrate Profiling of Cysteine Proteases Using a Combinatorial Peptide Library Identifies Functionally Unique Specificities. *J Biol Chem* **281**, 12824-12832, doi:/10.1074/jbc.M513331200 (2006).
- 11 Bromme, D. *et al.* Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O2 in *Spodoptera frugiperda* and characterization of the enzyme. *J Biol Chem* **271**, 2126-2132, doi:10.1074/jbc.271.4.2126 (1996).
- 12 Turk, V. *et al.* Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1824**, 68-88, doi:10.1016/j.bbapap.2011.10.002 (2012).
- 13 Guncar, G. *et al.* Crystal structure of porcine cathepsin H determined at 2.1 angstrom resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. *Structure* **6**, 51-61 (1998).
- 14 Guncar, G. *et al.* Crystal structure of cathepsin X: a flip-flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease. *Structure* **8**, 305-313, doi:10.1016/s0969-2126(00)00108-8 (2000).
- 15 Turk, D. *et al.* Crystal structure of cathepsin B inhibited with CA030 at 2.0-A resolution: A basis for the design of specific epoxysuccinyl inhibitors. *Biochemistry* **34**, 4791-4797, doi:10.1021/bi00014a037 (1995).
- 16 Horn, M., Jílková, Adéla., Mareš, Michael. Katepsinové proteasy v patologii. *Chem Listy* **108**, 358–363 (2014).
- 17 Mason, R. W. *Subcellular Biochemistry: Biology of the Lysosome*. (Plenum Press, 1996).
- 18 Bursch, W. The autophagosomal–lysosomal compartment in programmed cell death. *Cell Death Differ* **8**, 569-581, doi:10.1038/sj.cdd.4400852 (2001).

- 19 Cuervo, A. M. *et al.* A receptor for the selective uptake and degradation of proteins by lysosomes. *Science* **273**, 501-503, doi:10.1126/science.273.5274.501 (1996).
- 20 Turk, B. *et al.* Lysosomal cysteine proteases: more than scavengers. *Biochim Biophys Acta* **1477**, 98-111, doi:10.1016/s0167-4838(99)00263-0 (2000).
- 21 Pham, C. T. *et al.* 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, doi:10.1073/pnas.96.15.8627 (1999).
- 22 John, D. S. *et al.* Deficiency of cathepsin C ameliorates severity of acute pancreatitis by reduction of neutrophil elastase activation and cleavage of E-cadherin. *J Biol Chem* **294**, 697-707, doi:doi.org/10.1074/jbc.RA118.004376 (2019).
- 23 Nakagawa, T. Y. *et al.* Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. *Immunity* **10**, 207-217, doi:10.1016/s1074-7613(00)80021-7 (1999).
- 24 Reddy, V. B. *et al.* Redefining the concept of protease-activated receptors: cathepsin S evokes itch via activation of Mrgprs. *Nat Commun* **6**, 7864, doi:10.1038/ncomms8864 (2015).
- 25 Ewald, S. E. *et al.* Nucleic acid recognition by Toll-like receptors is coupled to stepwise processing by cathepsins and asparagine endopeptidase. *J Exp Med* **208**, 643-651, doi:10.1084/jem.20100682 (2011).
- 26 Repnik, U. *et al.* Lysosomes and lysosomal cathepsins in cell death. *Biochim Biophys Acta* **1824**, 22-33, doi:10.1016/j.bbapap.2011.08.016 (2012).
- 27 Tamhane, T. *et al.* Nuclear cathepsin L activity is required for cell cycle progression of colorectal carcinoma cells. *Biochimie* **122**, 208-218, doi:10.1016/j.biochi.2015.09.003 (2016).
- 28 Olson, O. C. *et al.* Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nat Rev Cancer* **15**, 712-729, doi:10.1038/nrc4027 (2015).
- 29 Mohamed, M. M. *et al.* Cysteine cathepsins: multifunctional enzymes in cancer. *Nature Reviews Cancer* **6**, 764-775, doi:10.1038/nrc1949 (2006).
- 30 Sobic, B. *et al.* Proteomic Identification of Cysteine Cathepsin Substrates Shed from the Surface of Cancer Cells. *Mol Cell Proteomics* **14**, 2213-2228, doi:10.1074/mcp.M114.044628 (2015).
- 31 Bromme, D. *et al.* Role of Cysteine Cathepsins in Extracellular Proteolysis. *Extracellular Matrix Degradation*, 23-51, doi:10.1007/978-3-642-16861-1_2 (2011).
- 32 Horn, M. *et al.* Katepsinové proteazy v patologii. *Chemické Listy* **108**, 358-363 (2014).
- 33 Reiser, J. *et al.* Specialized roles for cysteine cathepsins in health and disease. *J Clin Invest* **120**, 3421-3431, doi:10.1172/Jci42918 (2010).
- 34 Thompson, S. K. *et al.* Design of potent and selective human cathepsin K inhibitors that span the active site. *Proc Natl Acad Sci U S A* **94**, 14249-14254, doi:10.1073/pnas.94.26.14249 (1997).
- 35 Gelb, B. D. *et al.* Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science* **273**, 1236-1238, doi:10.1126/science.273.5279.1236 (1996).
- 36 Jobs, E. *et al.* Serum cathepsin S is associated with decreased insulin sensitivity and the development of type 2 diabetes in a community-based cohort of elderly men. *Diabetes Care* **36**, 163-165, doi:10.2337/dc12-0494 (2013).
- 37 Brown, R. *et al.* Cathepsin S: investigating an old player in lung disease pathogenesis, comorbidities, and potential therapeutics. *Respir Res* **21**, 111, doi:10.1186/s12931-020-01381-5 (2020).
- 38 Lowry, J. R. *et al.* Emerging roles of microglial cathepsins in neurodegenerative disease. *Brain Res Bull* **139**, 144-156, doi:10.1016/j.brainresbull.2018.02.014 (2018).

- 39 Lutgens, E. *et al.* Disruption of the cathepsin K gene reduces atherosclerosis progression and induces plaque fibrosis but accelerates macrophage foam cell formation. *Circulation* **113**, 98-107, doi:10.1161/circulationaha.105.561449 (2006).
- 40 Liu, J. *et al.* Lysosomal cysteine proteases in atherosclerosis. *Arterioscler Thromb Vasc Biol* **24**, 1359-1366, doi:10.1161/01.Atv.0000134530.27208.41 (2004).
- 41 Tjondrokoesoemo, A. *et al.* Cathepsin S Contributes to the Pathogenesis of Muscular Dystrophy in Mice. *J Biol Chem* **291**, 9920-9928, doi:10.1074/jbc.M116.719054 (2016).
- 42 Schedel, J. *et al.* Targeting cathepsin L (CL) by specific ribozymes decreases CL protein synthesis and cartilage destruction in rheumatoid arthritis. *Gene Ther* **11**, 1040-1047, doi:10.1038/sj.gt.3302265 (2004).
- 43 Lindström, E. *et al.* The selective cathepsin K inhibitor MIV-711 attenuates joint pathology in experimental animal models of osteoarthritis. *J Transl Med* **16**, 56, doi:10.1186/s12967-018-1425-7 (2018).
- 44 Tu, C. *et al.* Lysosomal cathepsin B participates in the podosome-mediated extracellular matrix degradation and invasion via secreted lysosomes in v-Src fibroblasts. *Cancer Res* **68**, 9147-9156, doi:10.1158/0008-5472.Can-07-5127 (2008).
- 45 Zhang, W. *et al.* Overexpression of cysteine cathepsin L is a marker of invasion and metastasis in ovarian cancer. *Oncol Rep* **31**, 1334-1342, doi:10.3892/or.2014.2967 (2014).
- 46 Rudzińska, M. *et al.* The Role of Cysteine Cathepsins in Cancer Progression and Drug Resistance. *Int J Mol Sci* **20**, doi:10.3390/ijms20143602 (2019).
- 47 Ruan, J. *et al.* Increased Expression of Cathepsin L: A Novel Independent Prognostic Marker of Worse Outcome in Hepatocellular Carcinoma Patients. *Plos One* **9**, e112136, doi:10.1371/journal.pone.0112136 (2014).
- 48 Pandey, G. *et al.* Prognostic significance of cathepsin L expression in pediatric acute myeloid leukemia. *Leuk Lymphoma* **59**, 2175-2187, doi:10.1080/10428194.2017.1422865 (2018).
- 49 Yang, L. *et al.* Glycosylated Cathepsin V Serves as a Prognostic Marker in Lung Cancer. *Front Oncol* **12**, 876245, doi:10.3389/fonc.2022.876245 (2022).
- 50 Cwiklinski, K. *et al.* Exploiting Comparative Omics to Understand the Pathogenic and Virulence-Associated Protease: Anti-Protease Relationships in the Zoonotic Parasites *Fasciola hepatica* and *Fasciola gigantica*. *Genes (Basel)* **13**, doi:10.3390/genes13101854 (2022).
- 51 Stack, C. *et al.* The phylogeny, structure and function of trematode cysteine proteases, with particular emphasis on the *Fasciola hepatica* cathepsin L family. *Adv Exp Med Biol* **712**, 116-135, doi:10.1007/978-1-4419-8414-2_8 (2011).
- 52 Stack, C. M. *et al.* Structural and functional relationships in the virulence-associated cathepsin L proteases of the parasitic liver fluke, *Fasciola hepatica*. *J Biol Chem* **283**, 9896-9908, doi:10.1074/jbc.M708521200 (2008).
- 53 Sajid, M. *et al.* Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol* **120**, 1-21 (2002).
- 54 Franta, Z. *et al.* IrCL1-The haemoglobinolytic cathepsin L of the hard tick, *Ixodes ricinus*. *Int J Parasitol* **41**, 1253-1262, doi:10.1016/j.ijpara.2011.06.006 (2011).
- 55 Boys, B. L. *et al.* Symmetric Behavior of Hemoglobin α - and β - Subunits during Acid-Induced Denaturation Observed by Electrospray Mass Spectrometry. *Biochemistry* **46**, 10675-10684, doi:10.1021/bi701076q (2007).
- 56 Caffrey, C. R. *et al.* Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of *Trypanosoma brucei*

- rhodesiense. *Mol Biochem Parasitol* **118**, 61-73, doi:10.1016/s0166-6851(01)00368-1 (2001).
- 57 Lowther, J. *et al.* The importance of pH in regulating the function of the *Fasciola hepatica* cathepsin L1 cysteine protease. *PLoS Negl Trop Dis* **3**, e369, doi:10.1371/journal.pntd.0000369 (2009).
- 58 Knox, D. Proteases in Blood-Feeding Nematodes and Their Potential as Vaccine Candidates. *Cysteine Proteases of Pathogenic Organisms* **712**, 155-176, doi:Book_Doi 10.1007/978-1-4419-8414-2 (2011).
- 59 Caffrey, C. R. *et al.* Cysteine proteases as digestive enzymes in parasitic helminths. *PLoS Negl Trop Dis* **12**, doi:10.1371/journal.pntd.0005840 (2018).
- 60 Delcroix, M. *et al.* A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. *J Biol Chem* **281**, 39316-39329, doi:10.1074/jbc.M607128200 (2006).
- 61 Na, B. K. *et al.* CsCF-6, a novel cathepsin F-like cysteine protease for nutrient uptake of *Clonorchis sinensis*. *Int J Parasitol* **38**, 493-502, doi:10.1016/j.ijpara.2007.09.001 (2008).
- 62 Choi, J. H. *et al.* Molecular and biochemical characterization of hemoglobinase, a cysteine proteinase, in *Paragonimus westermani*. *Korean J Parasitol* **44**, 187-196, doi:10.3347/kjp.2006.44.3.187 (2006).
- 63 Rosenthal, P. J. Cysteine proteases of malaria parasites. *Int J Parasitol* **34**, 1489-1499, doi:10.1016/j.ijpara.2004.10.003 (2004).
- 64 Que, X. *et al.* Cathepsin Cs are key for the intracellular survival of the protozoan parasite, *Toxoplasma gondii*. *J Biol Chem* **282**, 4994-5003, doi:10.1074/jbc.M606764200 (2007).
- 65 Alvarez, V. E. *et al.* The peptidases of *Trypanosoma cruzi*: digestive enzymes, virulence factors, and mediators of autophagy and programmed cell death. *Biochim Biophys Acta* **1824**, 195-206, doi:10.1016/j.bbapap.2011.05.011 (2012).
- 66 Horn, M. *et al.* Hemoglobin digestion in blood-feeding ticks: mapping a multi-peptidase pathway by functional proteomics. *Chem Biol* **16**, 1053-1063, doi:10.1016/j.chembiol.2009.09.009 (2009).
- 67 Robinson, M. W. *et al.* Collagenolytic activities of the major secreted cathepsin L peptidases involved in the virulence of the helminth pathogen, *Fasciola hepatica*. *PLoS Negl Trop Dis* **5**, e1012, doi:10.1371/journal.pntd.0001012 (2011).
- 68 Grote, A. *et al.* Cysteine proteases during larval migration and development of helminths in their final host. *PLoS Negl Trop Dis* **12**, e0005919, doi:10.1371/journal.pntd.0005919 (2018).
- 69 Brix, K. *et al.* Cysteine cathepsins: Cellular roadmap to different functions. *Biochimie* **90**, 194-207, doi:10.1016/j.biochi.2007.07.024 (2008).
- 70 Vasiljeva, O. *et al.* Recombinant human procathepsin S is capable of autocatalytic processing at neutral pH in the presence of glycosaminoglycans. *FEBS Lett* **579**, 1285-1290, doi:10.1016/j.febslet.2004.12.093 (2005).
- 71 Schomburg, I. *et al.* The BRENDA enzyme information system - From a database to an expert system. *J Biotechnol* **261**, 194-206, doi:10.1016/j.jbiotec.2017.04.020 (2017).
- 72 Jilkova, A. *et al.* Activation Route of the *Schistosoma mansoni* Cathepsin B1 Drug Target: Structural Map with a Glycosaminoglycan Switch. *Structure* **22**, 1786-1798, doi:10.1016/j.str.2014.09.015 (2014).
- 73 Verma, S. *et al.* Cysteine Proteases: Modes of Activation and Future Prospects as Pharmacological Targets. *Front Pharmacol* **7**, doi:10.3389/fphar.2016.00107 (2016).

- 74 Dahl, S. W. *et al.* Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing. *Biochemistry* **40**, 1671-1678, doi:10.1021/bi001693z (2001).
- 75 Dalton, J. P. *et al.* Schistosome asparaginyl endopeptidase SM32 in hemoglobin digestion. *Parasitol Today* **12**, 125, doi:10.1016/0169-4758(96)80676-4 (1996).
- 76 Pungercar, J. R. *et al.* Autocatalytic processing of procathepsin B is triggered by proenzyme activity. *Febs j* **276**, 660-668, doi:10.1111/j.1742-4658.2008.06815.x (2009).
- 77 Aoyagi, T. *et al.* Leupeptins New Protease Inhibitors from Actinomycetes. *J Antibiot* **22**, 283-286, doi:DOI 10.7164/antibiotics.22.283 (1969).
- 78 Umezawa, H. *et al.* Chymostatin, a new chymotrypsin inhibitor produced by actinomycetes. *J Antibiot (Tokyo)* **23**, 425-427, doi:10.7164/antibiotics.23.425 (1970).
- 79 Suda, H. *et al.* Antipain, a new protease inhibitor isolated from actinomycetes. *J Antibiot (Tokyo)* **25**, 263-266, doi:10.7164/antibiotics.25.263 (1972).
- 80 Vidal-Albalat, A. *et al.* Natural Products as Cathepsin Inhibitors. *Studies in Natural Products Chemistry, Vol 50: Bioactive Natural Products* **50**, 179-213, doi:10.1016/B978-0-444-63749-9.00006-2 (2016).
- 81 Hanada, K. *et al.* Isolation and Characterization of E-64, a New Thiol Protease Inhibitor. *Agric Biol Chem* **42**, 523-528, doi:10.1080/00021369.1978.10863014 (1978).
- 82 McGowan, E. B. *et al.* Inhibition of Calpain in Intact Platelets by the Thiol Protease Inhibitor-E-64d. *Biochem Biophys Res Commun* **158**, 432-435, doi:Doi 10.1016/S0006-291x(89)80065-8 (1989).
- 83 Towatari, T. *et al.* Novel Epoxysuccinyl Peptides - a Selective Inhibitor of Cathepsin-B, *In vivo*. *FEBS Lett* **280**, 311-315, doi:Doi 10.1016/0014-5793(91)80319-X (1991).
- 84 Yoon, M. C. *et al.* Molecular Features of CA-074 pH-Dependent Inhibition of Cathepsin B. *Biochemistry* **61**, 228-238, doi:10.1021/acs.biochem.1c00684 (2022).
- 85 Nakao, Y. *et al.* Miraziridine A, a novel cysteine protease inhibitor from the marine sponge *Theonella aff. mirabilis*. *Journal of the American Chemical Society* **122**, 10462-10463 (2000).
- 86 Fusetani, N. *et al.* Tokaramide A, a new cathepsin B inhibitor from the marine sponge *Theonella aff. mirabilis*. *Bioorg Med Chem Lett* **9**, 3397-3402, doi:Doi 10.1016/S0960-894x(99)00618-6 (1999).
- 87 Boudreau, P. D. *et al.* Design of Gallinamide A Analogs as Potent Inhibitors of the Cysteine Proteases Human Cathepsin L and *Trypanosoma cruzi* Cruzain. *J Med Chem* **62**, 9026-9044, doi:10.1021/acs.jmedchem.9b00294 (2019).
- 88 Kwan, J. C. *et al.* Grassypeptolides A-C, cytotoxic bis-thiazoline containing marine cyclodepsipeptides. *J Org Chem* **75**, 8012-8023, doi:10.1021/jo1013564 (2010).
- 89 Deb Majumdar, I. *et al.* Synthetic cyclohexenyl chalcone natural products possess cytotoxic activities against prostate cancer cells and inhibit cysteine cathepsins *in vitro*. *Biochem Biophys Res Commun* **416**, 397-402, doi:10.1016/j.bbrc.2011.11.052 (2011).
- 90 Klein, P. *et al.* Naphthoquinones as Covalent Reversible Inhibitors of Cysteine Proteases-Studies on Inhibition Mechanism and Kinetics. *Molecules* **25**, doi:10.3390/molecules25092064 (2020).
- 91 Fear, G. *et al.* Protease inhibitors and their peptidomimetic derivatives as potential drugs. *Pharmacol Ther* **113**, 354-368, doi:10.1016/j.pharmthera.2006.09.001 (2007).
- 92 Powers, J. C. *et al.* Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev* **102**, 4639-4750, doi:10.1021/cr010182v (2002).

- 93 Dryden-Peterson, S. *et al.* Nirmatrelvir Plus Ritonavir for Early COVID-19 in a Large U.S. Health System : A Population-Based Cohort Study. *Ann Intern Med*, doi:10.7326/m22-2141 (2022).
- 94 Lemke, C. *et al.* An Activity-Based Probe for Cathepsin K Imaging with Excellent Potency and Selectivity. *J Med Chem* **64**, 13793-13806, doi:10.1021/acs.jmedchem.1c01178 (2021).
- 95 Kato, D. *et al.* Activity-based probes that target diverse cysteine protease families. *Nat Chem Biol* **1**, 33-38, doi:10.1038/nchembio707 (2005).
- 96 Law, R. H. P. *et al.* An overview of the serpin superfamily. *Genome Biol* **7**, 216, doi:10.1186/gb-2006-7-5-216 (2006).
- 97 Mishra, M. Evolutionary Aspects of the Structural Convergence and Functional Diversification of Kunitz-Domain Inhibitors. *J Mol Evol* **88**, 537-548, doi:10.1007/s00239-020-09959-9 (2020).
- 98 Smith, D. *et al.* An atypical and functionally diverse family of Kunitz-type cysteine/serine proteinase inhibitors secreted by the helminth parasite *Fasciola hepatica*. *Sci Rep* **10**, doi:10.1038/s41598-020-77687-7 (2020).
- 99 G. Xu, M. C., Y. Huang, R. L. Rich, *et al.* Covalent inhibition revealed by the crystal structure of the caspase-8/p35 complex. *Nature* **410**, 494-497 (2001).
- 100 Guo, P. C. *et al.* Structural insights into the unique inhibitory mechanism of the silkworm protease inhibitor serpin18. *Sci Rep* **5**, doi:10.1038/srep11863 (2015).
- 101 Tusar, L. *et al.* Mechanisms Applied by Protein Inhibitors to Inhibit Cysteine Proteases. *Int J Mol Sci* **22**, doi:10.3390/ijms22030997 (2021).
- 102 Gettins, P. G. *et al.* Exosite determinants of serpin specificity. *J Biol Chem* **284**, 20441-20445, doi:10.1074/jbc.R800064200 (2009).
- 103 JW Fenton , G. V., FA Oforu , JM Maraganore . Thrombin inhibition by hirudin: how hirudin inhibits thrombin. *Haemostasis* **21**, 27-31 (1991).
- 104 Yamamoto, Y. *et al.* Novel cysteine proteinase inhibitors homologous to the proregions of cysteine proteinases. *Curr Protein Pept Sci* **3**, 231-238, doi:10.2174/1389203024605331 (2002).
- 105 Kurata, M. *et al.* Bombyx cysteine proteinase inhibitor (BCPI) homologous to propeptide regions of cysteine proteinases is a strong, selective inhibitor of cathepsin L-like cysteine proteinases. *J Biochem* **130**, 857-863, doi:10.1093/oxfordjournals.jbchem.a003058 (2001).
- 106 Bevec, T. *et al.* 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).
- 107 Lennon-Dumenil, A. M. *et al.* The p41 isoform of invariant chain is a chaperone for cathepsin L. *EMBO J* **20**, 4055-4064, doi:DOI 10.1093/emboj/20.15.4055 (2001).
- 108 Fiebiger, E. *et al.* Invariant Chain Controls the Activity of Extracellular Cathepsin L. *J Exp Med* **196**, 1263-1270, doi:10.1084/jem.20020762 (2002).
- 109 Zavašnik-Bergant, T. *et al.* Exogenous Thyropin from p41 Invariant Chain Diminishes Cysteine Protease Activity and Affects IL-12 Secretion during Maturation of Human Dendritic Cells. *Plos One* **11**, e0150815, doi:10.1371/journal.pone.0150815 (2016).
- 110 Lenarcic, B. *et al.* 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).
- 111 Costa, T. F. R. *et al.* Natural cysteine protease inhibitors in protozoa: Fifteen years of the chagasin family. *Biochimie* **122**, 197-207, doi:10.1016/j.biochi.2015.11.002 (2016).

- 112 Ljunggren, A. *et al.* Crystal structure of the parasite protease inhibitor chagasin in
complex with a host target cysteine protease. *J Mol Biol* **371**, 137-153,
doi:10.1016/j.jmb.2007.05.005 (2007).
- 113 Wang, S. X. *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).
- 114 Sabotic, J. *et al.* Macrocypins, a family of cysteine protease inhibitors from the
basidiomycete *Macrolepiota procera*. *FEBS J* **276**, 4334-4345, doi:10.1111/j.1742-
4658.2009.07138.x (2009).
- 115 Renko, M. *et al.* Versatile loops in mycocypins inhibit three protease families. *Acta*
Crystallographica a-Foundation and Advances **67**, C788-C788, doi:DOI
10.1107/S0108767311080056 (2011).
- 116 Abrahamson, M. *et al.* Isolation of six cysteine proteinase inhibitors from human urine.
Their physicochemical and enzyme kinetic properties and concentrations in biological
fluids. *J Biol Chem* **261**, 11282-11289 (1986).
- 117 Alvarez-Fernandez, M. *et al.* Inhibition of mammalian legumain by some cystatins is
due to a novel second reactive site. *J Biol Chem* **274**, 19195-19203, doi:DOI
10.1074/jbc.274.27.19195 (1999).
- 118 Kordis, D. *et al.* Phylogenomic analysis of the cystatin superfamily in eukaryotes and
prokaryotes. *BMC Evol Biol* **9**, doi:10.1186/1471-2148-9-266 (2009).
- 119 Magister, S. *et al.* Cystatins in immune system. *J Cancer* **4**, 45-56,
doi:10.7150/jca.5044 (2013).
- 120 Čeru, S. *et al.* Stefin B interacts with histones and cathepsin L in the nucleus. *J Biol*
Chem **285**, 10078-10086, doi:10.1074/jbc.M109.034793 (2010).
- 121 Stoka, V. *et al.* Lysosomal cathepsins and their regulation in aging and
neurodegeneration. *Ageing Res Rev* **32**, 22-37, doi:10.1016/j.arr.2016.04.010 (2016).
- 122 Mussap, M. *et al.* Biochemistry and clinical role of human cystatin C. *Crit Rev Clin*
Lab Sci **41**, 467-550, doi:10.1080/10408360490504934 (2004).
- 123 Kuo, J. *et al.* Serum and urinary biomarkers to predict acute kidney injury in premature
infants: a systematic review and meta-analysis of diagnostic accuracy. *J Nephrol* **35**,
2001-2014, doi:10.1007/s40620-022-01307-y (2022).
- 124 Zethelius, B. *et al.* Use of multiple biomarkers to improve the prediction of death from
cardiovascular causes. *N Engl J Med* **358**, 2107-2116, doi:10.1056/NEJMoa0707064
(2008).
- 125 Palsdottir, A. *et al.* Hereditary cystatin C amyloid angiopathy: genetic, clinical, and
pathological aspects. *Brain Pathol* **16**, 55-59 (2006).
- 126 Leto, G. *et al.* On the role of cystatin C in cancer progression. *Life Sci* **202**, 152-160,
doi:10.1016/j.lfs.2018.04.013 (2018).
- 127 Ganeshnarayan, K. *et al.* Human salivary cystatin SA exhibits antimicrobial effect
against *Aggregatibacter actinomycetemcomitans*. *J Periodontal Res* **47**, 661-673,
doi:10.1111/j.1600-0765.2012.01481.x (2012).
- 128 Alvarez-Díaz, S. *et al.* Cystatin D is a candidate tumor suppressor gene induced by
vitamin D in human colon cancer cells. *J Clin Invest* **119**, 2343-2358,
doi:10.1172/jci37205 (2009).
- 129 Nashida, T. *et al.* Antigen-presenting cells in parotid glands contain cystatin D
originating from acinar cells. *Arch Biochem Biophys* **530**, 32-39,
doi:doi.org/10.1016/j.abb.2012.12.009 (2013).
- 130 Zeeuwen, P. L. *et al.* The biology of cystatin M/E and its cognate target proteases. *J*
Invest Dermatol **129**, 1327-1338, doi:10.1038/jid.2009.40 (2009).

- 131 Prunk, M. *et al.* Extracellular Cystatin F Is Internalised by Cytotoxic T Lymphocytes and Decreases Their Cytotoxicity. *Cancers* **12**, 3660 (2020).
- 132 Geadkaew, A. *et al.* A 170kDa multi-domain cystatin of *Fasciola gigantica* is active in the male reproductive system. *Mol Biochem Parasitol* **196**, 100-107, doi:10.1016/j.molbiopara.2014.08.004 (2014).
- 133 Khaznadji, E. *et al.* A new multi-domain member of the cystatin superfamily expressed by *Fasciola hepatica*. *Int J Parasitol* **35**, 1115-1125, doi:10.1016/j.ijpara.2005.05.001 (2005).
- 134 Lee, C. *et al.* Type 3 cystatins; fetuins, kininogen and histidine-rich glycoprotein. *Front Biosci (Landmark Ed)* **14**, 2911-2922, doi:10.2741/3422 (2009).
- 135 Lalmanach, G. *et al.* Cysteine cathepsins and cystatins: from ancillary tasks to prominent status in lung diseases. *Biol Chem* **396**, 111-130, doi:10.1515/hsz-2014-0210 (2015).
- 136 Bode, W. *et al.* The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J* **7**, 2593-2599 (1988).
- 137 Stubbs, M. T. *et al.* 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 (1990).
- 138 Wang, S. X. *et al.* Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease. *Proc Natl Acad Sci U S A* **103**, 11503-11508, doi:10.1073/pnas.0600489103 (2006).
- 139 Nandy, S. K. *et al.* Structural Dynamics Investigation of Human Family 1 & 2 Cystatin-Cathepsin L1 Interaction: A Comparison of Binding Modes. *PLoS One* **11**, doi:10.1371/journal.pone.0164970 (2016).
- 140 Janowski, R. *et al.* 3D domain-swapped human cystatin C with amyloidlike intermolecular beta-sheets. *Proteins* **61**, 570-578, doi:10.1002/prot.20633 (2005).
- 141 Dall, E. *et al.* Structural and functional analysis of cystatin E reveals enzymologically relevant dimer and amyloid fibril states. *J Biol Chem* **293**, 13151-13165, doi:10.1074/jbc.RA118.002154 (2018).
- 142 Hamilton, G. *et al.* Cystatin F is a cathepsin C-directed protease inhibitor regulated by proteolysis. *Embo j* **27**, 499-508, doi:10.1038/sj.emboj.7601979 (2008).
- 143 Abrahamson, M. *et al.* Increased body temperature accelerates aggregation of the Leu-68-->Gln mutant cystatin C, the amyloid-forming protein in hereditary cystatin C amyloid angiopathy. *Proc Natl Acad Sci U S A* **91**, 1416-1420, doi:10.1073/pnas.91.4.1416 (1994).
- 144 Zerovnik, E. *et al.* Amyloid fibril formation by human stefins: Structure, mechanism & putative functions. *Biochimie* **92**, 1597-1607, doi:10.1016/j.biochi.2010.05.012 (2010).
- 145 CDC. *Three-Host Ixodid Tick Life Cycle*, <<https://www.cdc.gov/dpdx/ticks/index.html>> (2023).
- 146 Nuttall, P. A. Pathogen-tick-host interactions: *Borrelia burgdorferi* and TBE virus. *Zbl Bakt-Int J Med M* **289**, 492-505, doi:Doi 10.1016/S0934-8840(99)80002-4 (1999).
- 147 Venclikova, K. *et al.* Neglected tick-borne pathogens in the Czech Republic, 2011-2014. *Ticks Tick Borne Dis* **7**, 107-112, doi:10.1016/j.ttbdis.2015.09.004 (2016).
- 148 Sonenshine, D. E. *et al.* 3-16 (ODU Digital Commons, 2014).
- 149 Starck, J. M. *et al.* Morphological responses to feeding in ticks (*Ixodes ricinus*). *Zoological Lett* **4**, 20, doi:10.1186/s40851-018-0104-0 (2018).
- 150 Lara, F. A. *et al.* Tracing heme in a living cell: hemoglobin degradation and heme traffic in digest cells of the cattle tick *Boophilus microplus*. *J Exp Biol* **208**, 3093-3101, doi:10.1242/jeb.01749 (2005).

- 151 Sojka, D. *et al.* Multienzyme degradation of host serum albumin in ticks. *Ticks Tick Borne Dis* **7**, 604-613, doi:10.1016/j.ttbdis.2015.12.014 (2016).
- 152 Lara, F. A. *et al.* 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, doi:10.1242/jeb.00334 (2003).
- 153 Sojka, D. *et al.* New insights into the machinery of blood digestion by ticks. *Trends Parasitol* **29**, 276-285, doi:10.1016/j.pt.2013.04.002 (2013).
- 154 Kotal, J. *et al.* Mialostatin, a Novel Midgut Cystatin from *Ixodes ricinus* Ticks: Crystal Structure and Regulation of Host Blood Digestion. *Int J Mol Sci* **22**, 5371, doi:10.3390/ijms22105371 (2021).
- 155 Xavier, M. A. *et al.* Blood anticlotting activity of a *Rhipicephalus microplus* cathepsin L-like enzyme. *Biochimie* **163**, 12-20, doi:doi.org/10.1016/j.biochi.2019.04.025 (2019).
- 156 Porter, L. M. *et al.* 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).
- 157 Parizi, L. F. *et al.* Peptidase inhibitors in tick physiology. *Med Vet Entomol* **32**, 129-144, doi:10.1111/mve.12276 (2018).
- 158 Chmelar, J. *et al.* Tick salivary secretion as a source of antihemostatics. *J Proteomics* **75**, 3842-3854, doi:10.1016/j.jprot.2012.04.026 (2012).
- 159 Lu, S. *et al.* A novel type 1 cystatin involved in the regulation of *Rhipicephalus microplus* midgut cysteine proteases. *Ticks Tick Borne Dis* **11**, 101374, doi:doi.org/10.1016/j.ttbdis.2020.101374 (2020).
- 160 Yamaji, K. *et al.* Hlcyst-1 and Hlcyst-2 are potential inhibitors of HICPL-A in the midgut of the ixodid tick *Haemaphysalis longicornis*. *J Vet Med Sci* **72**, 599-604, doi:10.1292/jvms.09-0561 (2010).
- 161 Anderson, J. M. *et al.* Exploring the mialome of ticks: an annotated catalogue of midgut transcripts from the hard tick, *Dermacentor variabilis* (Acari: Ixodidae). *BMC Genomics* **9**, 552, doi:10.1186/1471-2164-9-552 (2008).
- 162 Lima, C. A. *et al.* Bmcystatin, a cysteine proteinase inhibitor characterized from the tick *Boophilus microplus*. *Biochem Biophys Res Commun* **347**, 44-50, doi:doi.org/10.1016/j.bbrc.2006.06.018 (2006).
- 163 Wang, Y. *et al.* Functional characterization of a cystatin from the tick *Rhipicephalus haemaphysaloides*. *Parasites & Vectors* **8**, 140, doi:10.1186/s13071-015-0725-5 (2015).
- 164 Schwarz, A. *et al.* 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).
- 165 Wang, Y. *et al.* Functional characterization of a cystatin from the tick *Rhipicephalus haemaphysaloides*. *Parasit Vectors* **8**, 140, doi:10.1186/s13071-015-0725-5 (2015).
- 166 Chmelar, J. *et al.* Protease Inhibitors in Tick Saliva: The Role of Serpins and Cystatins in Tick-host-Pathogen Interaction. *Front Cell Infect Microbiol* **7**, 216, doi:10.3389/fcimb.2017.00216 (2017).
- 167 Schwarz, A. *et al.* 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).
- 168 Kotsyfakis, M. *et al.* Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick *Ixodes scapularis*. *J Biol Chem* **281**, 26298-26307, doi:10.1074/jbc.M513010200 (2006).
- 169 Lu, S. *et al.* Rmcystatin3, a cysteine protease inhibitor from *Rhipicephalus microplus* hemocytes involved in immune response. *Biochimie* **106**, 17-23, doi:10.1016/j.biochi.2014.07.012 (2014).

- 170 Kotsyfakis, M. *et al.* 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).
- 171 Lieskovska, J. *et al.* Tick salivary cystatin sialostatin L2 suppresses IFN responses in mouse dendritic cells. *Parasite Immunol* **37**, 70-78, doi:10.1111/pim.12162 (2015).
- 172 Lieskovská, J. *et al.* Tick sialostatins L and L2 differentially influence dendritic cell responses to *Borrelia spirochetes*. *Parasite Vector* **8**, 275, doi:10.1186/s13071-015-0887-1 (2015).
- 173 Jakos, T. *et al.* Cysteine Cathepsins in Tumor-Associated Immune Cells. *Front Immunol* **10**, doi:10.3389/fimmu.2019.02037 (2019).
- 174 Kotsyfakis, M. *et al.* Selective cysteine protease inhibition contributes to blood-feeding success of the tick *Ixodes scapularis*. *J Biol Chem* **282**, 29256-29263, doi:10.1074/jbc.M703143200 (2007).
- 175 Furst, T. *et al.* Global burden of human food-borne trematodiasis: a systematic review and meta-analysis. *Lancet Infect Dis* **12**, 210-221, doi:10.1016/S1473-3099(11)70294-8 (2012).
- 176 Caravedo, M. A. *et al.* Human Fascioliasis: Current Epidemiological Status and Strategies for Diagnosis, Treatment, and Control. *Res Rep Trop Med* **11**, 149-158, doi:10.2147/rrtm.S237461 (2020).
- 177 Fairweather, I. *et al.* Drug resistance in liver flukes. *Int J Parasitol Drugs Drug Resist* **12**, 39-59, doi:10.1016/j.ijpddr.2019.11.003 (2020).
- 178 CDC. *Fascioliasis*, <<https://www.cdc.gov/dpdx/fascioliasis>> (2023).
- 179 Mas-Coma, S. *et al.* Fascioliasis and other plant-borne trematode zoonoses. *Int J Parasitol* **35**, 1255-1278, doi:10.1016/j.ijpara.2005.07.010 (2005).
- 180 Arjona, R. *et al.* Fascioliasis in developed countries: a review of classic and aberrant forms of the disease. *Medicine (Baltimore)* **74**, 13-23, doi:10.1097/00005792-199501000-00002 (1995).
- 181 Saba, R. *et al.* Human fascioliasis. *Clin Microbiol Infect* **10**, 385-387, doi:doi.org/10.1111/j.1469-0691.2004.00820.x (2004).
- 182 Verissimo, C. D. *et al.* *Fasciola hepatica* is refractory to complement killing by preventing attachment of mannose binding lectin (MBL) and inhibiting MBL-associated serine proteases (MASPs) with serpins. *PLoS Path* **18**, doi:10.1371/journal.ppat.1010226 (2022).
- 183 Norbury, L. J. *et al.* Adult and juvenile *Fasciola* cathepsin L proteases: different enzymes for different roles. *Biochimie* **93**, 604-611, doi:10.1016/j.biochi.2010.12.004 (2011).
- 184 Acosta, D. *et al.* *Fasciola hepatica* leucine aminopeptidase, a promising candidate for vaccination against ruminant fasciolosis. *Mol Biochem Parasitol* **158**, 52-64, doi:10.1016/j.molbiopara.2007.11.011 (2008).
- 185 Corvo, I. *et al.* Dissecting the active site of the collagenolytic cathepsin L3 protease of the invasive stage of *Fasciola hepatica*. *PLoS Negl Trop Dis* **7**, e2269, doi:10.1371/journal.pntd.0002269 (2013).
- 186 McGonigle, L. *et al.* The silencing of cysteine proteases in *Fasciola hepatica* newly excysted juveniles using RNA interference reduces gut penetration. *Int J Parasitol* **38**, 149-155, doi:10.1016/j.ijpara.2007.10.007 (2008).
- 187 Cwiklinski, K. *et al.* The cathepsin-like cysteine peptidases of trematodes of the genus *Fasciola*. *Adv Parasitol* **104**, 113-164, doi:10.1016/bs.apar.2019.01.001 (2019).

- 188 Donnelly, S. *et al.* Helminth cysteine proteases inhibit TRIF-dependent activation of
macrophages via degradation of TLR3. *J Biol Chem* **285**, 3383-3392,
doi:10.1074/jbc.M109.060368 (2010).
- 189 Brady, M. T. *et al.* Fasciola hepatica suppresses a protective Th1 response against
Bordetella pertussis. *Infect Immun* **67**, 5372-5378, doi:10.1128/iai.67.10.5372-
5378.1999 (1999).
- 190 Cancela, M. *et al.* Functional characterization of single-domain cystatin-like cysteine
proteinase inhibitors expressed by the trematode Fasciola hepatica. *Parasitology* **144**,
1695-1707, doi:10.1017/S0031182017001093 (2017).
- 191 Sánchez Di Maggio, L. *et al.* Serpins in Fasciola hepatica: insights into host-parasite
interactions. *Int J Parasitol* **50**, 931-943, doi:10.1016/j.ijpara.2020.05.010 (2020).
- 192 De Marco Verissimo, C. *et al.* Fasciola hepatica serine protease inhibitor family
(serpins): Purposely crafted for regulating host proteases. *PLoS Negl Trop Dis* **14**,
e0008510, doi:10.1371/journal.pntd.0008510 (2020).
- 193 Smith, D. *et al.* Unexpected Activity of a Novel Kunitz-type Inhibitor: Inhibition of
Cysteine Proteases but Not Serine Proteases. *J Biol Chem* **291**, 19220-19234,
doi:10.1074/jbc.M116.724344 (2016).
- 194 Zhang, K. *et al.* Molecular Characteristics and Potent Immunomodulatory Activity of
Fasciola hepatica Cystatin. *Korean J Parasitol* **60**, 117-126,
doi:10.3347/kjp.2022.60.2.117 (2022).
- 195 Tarasuk, M. *et al.* Type I cystatin (stefin) is a major component of Fasciola gigantica
excretion/secretion product. *Mol Biochem Parasitol* **167**, 60-71,
doi:10.1016/j.molbiopara.2009.04.010 (2009).
- 196 Horn, M. *et al.* Free-thiol Cys331 exposed during activation process is critical for native
tetramer structure of cathepsin C (dipeptidyl peptidase I). *Protein Science* **11**, 933-943,
doi:10.1110/ps.2910102 (2002).
- 197 Horn, M. *et al.* Activation processing of cathepsin H impairs recognition by its
propeptide. *Biol Chem* **386**, 941-947, doi:10.1515/bc.2005.109 (2005).
- 198 Benýšek, J. *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* **37**, 515-526,
doi:10.1080/14756366.2021.2024527 (2022).
- 199 Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta
Crystallogr D Biol Crystallogr* **66**, 133-144, doi:10.1107/S0907444909047374 (2010).
- 200 Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons
and electrons: recent developments in Phenix. *Acta Crystallographica Section D-
Structural Biology* **75**, 861-877, doi:10.1107/S2059798319011471 (2019).
- 201 Somoza, J. R. *et al.* Crystal structure of human cathepsin V. *Biochemistry* **39**, 12543-
12551, doi:10.1021/bi000951p (2000).
- 202 Renko, M. *et al.* *Unreduced cathepsin V in complex with stefin A*, 2009).
- 203 Renko, M. *et al.* Versatile Loops in Mycocypins Inhibit Three Protease Families. *J Biol
Chem* **285**, 308-316, doi:10.1074/jbc.M109.043331 (2010).
- 204 Robinson, P. J. *et al.* Mechanisms of Disulfide Bond Formation in Nascent
Polypeptides Entering the Secretory Pathway. *Cells* **9**, doi:10.3390/cells9091994
(2020).
- 205 Olson, O. C. *et al.* Cysteine cathepsin proteases: regulators of cancer progression and
therapeutic response. *Nat Rev Cancer* **15**, 712-729, doi:10.1038/nrc4027 (2015).
- 206 Nanut, M. P. *et al.* The Role of Cysteine Peptidases in Hematopoietic Stem Cell
Differentiation and Modulation of Immune System Function. *Front Immunol* **12**,
doi:10.3389/fimmu.2021.680279 (2021).

- 207 Maizels, R. M. *et al.* Modulation of Host Immunity by Helminths: The Expanding Repertoire of Parasite Effector Molecules. *Immunity* **49**, 801-818, doi:10.1016/j.immuni.2018.10.016 (2018).
- 208 Salat, J. *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).
- 209 Turk, V. *et al.* Cystatins: biochemical and structural properties, and medical relevance. *Front Biosci* **13**, 5406-5420, doi:10.2741/3089 (2008).
- 210 Cribb, T. H. *et al.* Life cycle evolution in the digenea: a new perspective from phylogeny. *Adv Parasitol* **54**, 197-254, doi:10.1016/s0065-308x(03)54004-0 (2003).
- 211 Netea, M. G. *et al.* Innate and Adaptive Immune Memory: an Evolutionary Continuum in the Host's Response to Pathogens. *Cell Host Microbe* **25**, 13-26, doi:10.1016/j.chom.2018.12.006 (2019).
- 212 Grunclova, L. *et al.* 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).
- 213 Sojka, D. *et al.* New insights into the machinery of blood digestion by ticks. *Trends Parasitol* **29**, 276-285, doi:doi.org/10.1016/j.pt.2013.04.002 (2013).
- 214 Musil, D. *et al.* The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J* **10**, 2321-2330 (1991).
- 215 Renko, M. *et al.* Stefin A displaces the occluding loop of cathepsin B only by as much as required to bind to the active site cleft. *FEBS J* **277**, 4338-4345, doi:10.1111/j.1742-4658.2010.07824.x (2010).