Univerzita Karlova

Přírodovědecká fakulta

Studijní program: Biochemie

Studijní obor: Biochemie

Mgr. Jaroslav Srp

Regulace aktivity aspartátových a serinových proteas pomocí selektivních přirozených inhibitorů

Regulation of the activity of aspartic and serine proteases by selective natural inhibitors

Disertační práce

Školitel: RNDr. Michael Mareš, CSc.

Praha, 2018

Prohlášení:

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

V Praze, dne

……..

Podpis

Poděkování

Rád bych na tomto místě poděkoval všem, kteří se podíleli na realizaci mé disertační práce.

Zejména děkuji svému školiteli Michaeli Marešovi za jeho cenné rady, vstřícnost, ochotu a čas, který mi věnoval, Martinu Hornovi a Martině Nussbaumerové za praktické rady v počátcích mých laboratorních experimentů a Ireně Pražákové a Milaně Štajfové za technickou pomoc v laboratoři. Rovněž děkuji i ostatním členům naší skupiny Katepsinové proteasy v patologii za všestrannou pomoc a vytvoření příjemného pracovního prostředí. Ze servisních pracovišť ÚOCHB AV ČR bych rád poděkoval především Dr. Petru Pachlovi a Dr. Pavlíně Řezáčové za pomoc při měření a zpracování difrakčních dat pro proteinovou krystalografii, dále pak Ing. Zdeňku Voburkovi za určení N koncové sekvence proteinů, členům proteomické sekce skupiny hmotnostní spektrometrie (Dr. Šandovi, Dr. Demianové a Dr. Hubálkovi) za určení proteinových sekvencí a Mirce Blechové a Dr. Hradilkovi za zakázkovou syntézu peptidů.

Závěrem bych chtěl poděkovat své rodině a přátelům za podporu a trpělivost v průběhu celého studia.

Obsah

Seznam zkratek 1
Abstrakt (ČJ) 3
Abstract (EN) 4
1. Literární přehled 5
1.1. Proteolytické enzymy a jejich klasifikace 5
1.1.1. Cysteinové proteasy C1 rodiny papainu a C13 rodiny legumainu 5
1.2. Regulace aktivity proteolytických enzymů 6
1.3. Aspartátové proteasy z A1 rodiny pepsinu 7
1.3.1. Prostorová struktura aspartátových proteas z A1 rodiny pepsinu 8
1.3.2. Popis aktivního místa a mechanismus katalýzy u aspartátových proteas z A1 rodiny pepsinu 9
1.3.3. Katepsin D 9
1.3.3.1. Prostorová struktura a aktivita katepsinu D 9
1.3.3.2. Fyziologie a patofyziologie spojené s katepsinem D 10
1.3.4. Medicinálně významné lidské aspartátové proteasy z A1 rodiny pepsinu 11
1.3.5. Aspartátové proteasy z A1 rodiny pepsinu u bezobratlých 12
1.4. Serinové proteasy z S1 rodiny chymotrypsinu 12
1.4.1. Prostorová struktura serinových proteas z S1 rodiny chymotrypsinu 13
1.4.2. Popis aktivního místa a mechanismus katalýzy u serinových proteas z S1 chymotrypsinové rodiny 14
1.4.3. Trypsin a chymotrypsin 14
1.4.4. Medicinálně významné serinové proteasy z S1 rodiny chymotrypsinu 15
1.4.5. Serinové proteasy z S1 rodiny chymotrypsinu u bezobratlých 16
1.5. Přirozené a syntetické inhibitory proteas 16
1.5.1. Peptidové inhibitory aspartátových a serinových proteas 17
1.5.2. Proteinové inhibitory proteas 19
1.5.2.1. Proteinové inhibitory aspartátových proteas 19
1.5.2.2. Proteinové inhibitory serinových proteas 20
1.5.2.3. Rostlinné proteasové inhibitory z Kunitzovy rodiny 21
2. Cíle práce 24
3. Materiál a metodika 25
3.1. Materiál a laboratorní vybavení 25
3.2. Metodika 25
3.2.1. Příprava rekombinantních proteas 25
3.2.2. Biochemické metody 26
3.2.3. Enzymologické metody 26
3.2.4. Krystalografické metody 26
4. Výsledky 27
4.1. Publikace č. 1: Digestive proteolysis in the Colorado potato beetle, *Leptinotarsa decemlineata*: Activity-based profiling and imaging of a multipeptidase network. 30
4.1.1. Proteolytická aktivita ve střevě mandelinky bramborové 30
4.1.2. Analýza trávicích aspartátových proteas 30
4.1.3. Analýza trávicích cysteinových proteas 31
4.1.4. Analýza trávicích serinových proteas 32
4.1.5. Závěr 32

4.2. Nepublikované výsledky 52
4.2.1. Strukturní analýza inhibitorů z Kunitzovy rodiny a jejich reaktivních center 52
4.2.1.1. Prostorová struktura inhibitoru PSPI v komplexu s trypsinem a chymotrypsinem 52
4.2.1.2. Prostorová struktura inhibitoru PCDI v komplexu s trypsinem 57
4.2.1.3. Prostorová struktura inhibitoru PCDI v komplexu s aspartátovou proteasou 58
4.2.1.4. Topologie reaktivních center ve struktuře inhibitorů PSPI a PCDI 59
4.2.2. Strukturní analýza aspartátové proteasy LdCD 61
4.3. Publikace č. 2: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids. 64
4.4. Publikace č. 3: Single- and double-headed chemical probes for detection of active cathepsin D in a cancer cell proteome. 78
4.5. Publikace č. 4: Profiling system for skin kallikrein proteolysis applied in gene deficient mouse models. 85
4.6. Publikace č. 5: Crystal structures of the complex of kallikrein inhibitor BbKI with trypsin and modeling of kallikrein complexes. 94
4.7. Publikace č. 6: Kunitzovy proteasové inhibitory z rostlin: strukturní a funkční diverzita. 133
5. Diskuze 143
5.1. Aspartátové proteasy a jejich regulace přirozenými inhibitory 143
5.2. Serinové proteasy a jejich regulace přirozenými inhibitory 145
5.3. Identifikace nových reaktivních center na rostlinných proteasových inhibitorech z Kunitzovy rodiny 146
6. Závěry 149
Prohlášení spoluautorů 151
7. Seznam použité literatury 153

Seznam zkratek

Abz	kyselina aminobenzoová
Ac	acetylový zbytek
Acc	7-amino-4-karbamoylmethylkumarin
AE	asparaginylendopeptidasa
Amc	7-amino-4-methylkumarin
API-A	inhibitor trypsinu z Kunitzovy rodiny ze šípatky (Arrowhead protease inhibitors)
APP	amyloidní prekursorový protein
BbKI	inhibitor serinových proteas z bauhinie (*Bauhinia bauhinioides* Kallikrein Inhibitor)
BPTI	inhibitor trypsinu z hovězí sleziny (Bovine Pankreas Trypsin Inhibitor)
BSA	hovězí sérový albumin (Bovine Serum Albumin)
CA074	L-3-trans-(propylkarbamoyl)oxiran-2-karbonyl)-L-isoleucyl-L-prolin
Cbz	benzyloxykarbonyl
Cmk	chloromethyl keton
E64	N-[N-(L-3-trans-karboxyoxirin-2-karbonyl)-L-leucin]-agmatin
FPLC	Fast Protein Liquid Chromatography
FRET	Fluroescence Resonance Energy Transfer
HIV	virus lidské imunodeficience (Human Immunodeficiency Virus)
HPLC	High Performance Liquid Chromatography
IA3	inhibitor aspartátové proteasy proteinasy A z kvasinky (*Saccharomyces cerevisiae*)
IC_{50}	koncentrace inhibitoru potřebná k dosažení 50% inhibice enzyme
IrCD1	katepsin D z klíštěte obecného (*Ixodes ricinus*)
k_{cat}	číslo přeměny
k_{cat}/Km	katalytická účinnost enzymu
kDa	kilodalton
K_i	inhibiční konstanta
KLK	serinová proteasa kalikrein
K_M	Michaelisova konstanta
LdCD	katepsin D z mandelinky bramborové (*Leptinotarsa decemlineata*)
MWCO	membrána s póry o definované velikosti (Molecular Weight Cut-Off)
NTA	nitrilotrioctová kyselina
PCDI	inhibitor katepsinu D z brambor (Potato Cathepsin D Inhibitor)
PDB	databáze proteinových struktur Protein Data Bank

PEG	polyethylenglykol
PI3	inhibitor aspartátové proteasy pepsinu ze škrkavky (*Ascaris suum*)
PSPI	inhibitor serinových proteas z brambor (Potato Serine Protease Inhibitor)
PVDF	polyvinyldifluorid
SDS	dodecylsíran sodný
SDS-PAGE	elektroforéza na polyakrylamidovém gelu v přítomnosti SDS
serpiny	inhibitor serinových proteas (Serine Protease Inhibitors)
SQAPI	inhibitor aspartátových proteas z tykve (*Cucurbita maxima*) Squash Aspartic Proteinase Inhibitor
STI	sójový inhibitor trypsinu z Kunitzovy rodiny (Soybean Trypsin Inhibitor)
Tos	tosyl
Tris	tris(hydroxymethyl)aminomehtan
WSCP	protein Kunitzovy rodiny schopný vázat chlorofyl (Water-Soluble Chlorophyll binding Protein)
YPD	médium pro expresi v kvasinkách *Pichia pastoris* (Yeast extract Peptone Dextrose)

Nestandartní aminokyselinové zbytky

NPh	4-nitrofenylalanin
Iva	kyselina izovalerová
Sta	statin

Abstrakt (ČJ)

Proteasy se účastní řady významných fyziologických procesů a jejich nedostatečná kontrola je spojena s řadou patologií. Přirozené inhibitory představují účinný nástroj pro regulaci jejich aktivity. Disertační práce se zaměřuje na živočišné a rostlinné inhibitory aspartátových a serinových proteas a přináší identifikaci, biochemickou charakterizaci a strukturní popis jejich mechanismu inhibice.

Rostlinné inhibitory z Kunitzovy rodiny jsou produkované jako obranné proteiny s unikátní schopností blokovat široké spektrum proteas. V této práci byl s využitím metod funkční proteomiky popsán trávicí proteolytický systém mandelinky bramborové, herbivorního škůdce lilkovitých rostlin. Byl podán přímý důkaz, že trávicí aspartátové a serinové proteasy tohoto herbivora jsou efektivně blokovány pomocí dvou Kunitzových inhibitorů (nazývaných PCDI a PSPI), které jsou produkované v lilku brambor. Pomocí strukturní analýzy byly na molekulách PCDI a PSPI identifikovány nové typy reaktivních center pro inhibici aspartátových proteas typu katepsinu D a serinových proteas typu trypsinu a chymotrypsinu. Popis reaktivního centra na PCDI inhibitoru společně s krystalovou strukturou trávicí proteasy typu katepsinu D mandelinky bramborové umožnilo vysvětlit mechanismus jejich vzájemné interakce.

Pro lidský katepsin D byly sfingolipidy identifikovány jako první specifické endogenní inhibitory. Tyto bioaktivní molekuly jsou známé svou protinádorovou aktivitou a lze předpokládat, že jedním z mechanismů jejich působení je regulace katepsinu D, který se tvorby nádorů účastní. Dále byla na bázi peptidomimetického inhibitoru aktivního místa vyvinuta proteomická sonda pro detekci katepsinu D jako prognostického nádorového markeru.

Pro serinové proteasy typu kalikreinů byl identifikován dosud nejúčinnější přirozený rostlinný inhibitor BbKI z Kunitzovy rodiny, který je selektivní pro nádorový marker kalikrein 4. Jejich interakce byla popsána pomocí strukturního modelu. Pro epidermální kalikreiny byl sestaven proteomický detekční systém selektivních substrátů a inhibitorů pro monitorování aktivit, který byl testován na geneticky modifikovaných myších.

Disertační práce přináší významné informace o specifitě a mechanismech inhibice aspartátových a serinových proteas, které lze využít pro racionální vývoj nových inhibičních molekul s uplatněním v biomedicíně a zemědělských biotechnologiích.

Abstract (EN)

Proteases are involved in many physiological processes and their dysregulation is associated with various pathologies. Protease activity is effectively controlled by natural inhibitors. This PhD thesis is focused on the inhibitors of aspartic and serine proteases of animal and plant origin and provides the identification, biochemical characterization and structural description of their inhibition mechanisms.

Plant Kunitz inhibitors are produced as defensive proteins, and they are able to block activities of a broad spectrum of proteases. In this thesis, the digestive proteolytic system of the Colorado potato beetle, a herbivore pest of potato plants, was described with the help of functional proteomics. It was shown that aspartic and serine proteases from this herbivore are effectively blocked by two potato Kunitz inhibitors (namely PCDI, PSPI). Using structural analysis, novel types of reactive centers were identified on PCDI and PSPI molecules for the inhibition of aspartic protease cathepsin D and the serine proteases trypsin and chymotrypsin. The analysis of the reactive center on a PCDI with the crystal structure of digestive cathepsin D from the Colorado potato beetle explained the mechanism of their interaction.

Sphingolipids were identified as the first endogenous inhibitors of human cathepsin D. Sphingolipids are bioactive molecules with anti-cancer activity, and the mechanism of their action may include the regulation of cathepsin D, which is associated with cancer proliferation. Moreover, a proteomic probe based on a peptidomimetic inhibitor was developed for the detection of human cathepsin D, which is a prognostic marker in cancer.

The plant Kunitz inhibitor BbKI was identified as the most potent natural inhibitor of kallikrein-type serine proteases. BbKI is selective for the cancer marker kallikrein 4, and their interaction was described using a structural model. Furthermore, a proteomic tool for the detection of epidermal kallikreins was designed combining selective substrates and inhibitors, and evaluated on mice with genetically knock-out kallikreins.

To conclude, this PhD thesis provides important information about the specificity and inhibition mechanisms of aspartic and serine proteases, which can be used for the rational design of new inhibitors with biomedical and agricultural applications.

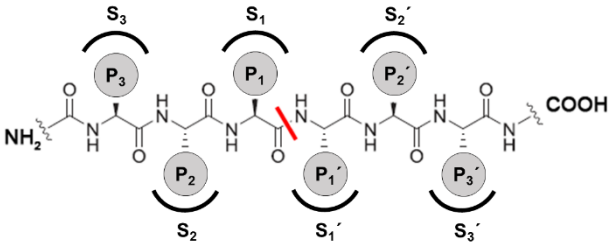
1. Literární přehled

Proteasy nebo také peptidasy jsou proteolytické enzymy, které hydrolyzují peptidovou vazbu v proteinech a peptidech. Vyskytují se ve všech živých organismech (archeobakterie, bakterie, rostliny a živočichové) i virech a jsou zapojeny do mnoha důležitých fyziologických procesů, jako je např. trávení, lysosomální degradace proteinů, apoptóza nebo krevní srážení.

1.1. Proteolytické enzymy a jejich klasifikace

Proteasy lze na základě místa štěpení polypeptidového substrátu dělit na endopeptidasy, které štěpí substrát uvnitř polypeptidového řetězce, a na exopeptidasy, které odštěpují aminokyselinové zbytky nebo krátké peptidy z N konce (aminopeptidasy) nebo C konce (karboxypeptidasy) substrátu. Podle vzájemné příbuznosti na základě homologie aminokyselinové sekvence je možné proteolytické enzymy klasifikovat do rodin a dále do klanů podle podobnosti prostorových struktur (Rawlings a Barrett 1993). V závislosti na složení katalytického centra a mechanismu katalýzy se proteasy dělí do tříd na serinové, cysteinové, aspartátové, asparaginové, glutamové, threoninové a metaloproteasy (Rawlings a Salvesen 2013).

Pro popis vazby peptidového substrátu do aktivního místa proteasy se používá konvenční terminologie (Obr. 1), kdy jednotlivé aminokyselinové zbytky substrátu jsou označeny od N konce substrátu ke štěpené vazbě jako P_n-P_1 a od štěpené vazby k C konci substrátu jako P_1‘-P_n‘. Analogicky se označují příslušná podmísta v aktivním místě enzymu, která vážou substrát jako S_n-S_1 od N konce substrátu ke štěpené vazbě a S_1‘-S_n‘ od štěpené vazby k C konci substrátu. Štěpená peptidová vazba se nachází mezi zbytky P_1 a P_1‘ substrátu, které se vážou do podmíst S_1 a S_1‘ proteasy (Motwani et al. 2015). Stejný způsobem se popisuje vazba inhibitoru do aktivního místa proteasy.



Obrázek 1: Schématické znázornění vazby peptidového substrátu do aktivního místa proteasy. *Zobrazeny jsou pozice jednotlivých aminokyselinových zbytků peptidového substrátu (P_3-P_3´), které se vážou do odpovídajících podmíst enzymu (S_3-S_3´). Štěpená peptidová vazba mezi zbytky P_1 a P_1´ je znázorněna červeně. Převzato a upraveno podle (Motwani et al. 2015).*

1.1.1. Cysteinové proteasy C1 rodiny papainu a C13 rodiny legumainu

Cysteinové proteasy vyžadují pro svou hydrolytickou aktivitu přítomnost nukleofilní thiolové skupiny cysteinu v aktivním místě enzymu. Podle homologie primárních struktur se cysteinové proteasy dělí do zhruba 70 rodin. V této práci byly studovány zástupci C1 rodiny

papainu, která je nejpočetnější a nejlépe charakterizována, a C13 rodiny legumainu (asparaginylendopeptidasy) (Rawlings a Barrett 2010). Do rodiny papainu patří také velká skupina savčích cysteinových katepsinů, které lze rozdělit na 11 zástupců (katepsin B, C, F, H, K, L, O, S, V, X a W). Tyto enzymy se vyskytují ve všech tkáních a buňkách a podílejí se zejména na lysosomální degradaci proteinů, presentaci antigenu, aktivaci hormonů a přestavbě extracelulární matrix (Novinec a Lenarcic 2013). Mezi medicinálně významné cysteinové proteasy patří také parazitární katepsiny, které zajišťují trávení přijatých proteinů a další biologické funkce parazitů. Proto jsou vhodnými cílovými molekulami pro vývoj nových antiparazitik (Jílková et al. 2011). Ukázkovým příkladem jsou kruzipain z *Trypanosoma cruzi*, falcipain z *Plasmodium falciparum* a SmCB1 ze *Schistosoma mansoni*, jako cíle potenciálních léčiv proti Chagasově chorobě, malárii a schistosomóze (McKerrow et al. 2009; Marco a Coteron 2012). Zástupci cysteinových katepsinů byly popsány jako trávicí enzymy u mnoha členovců včetně řádu brouků (Coleoptera), např. katepsiny L z *Diabrotica virgifera*, *Tenebrio molitor* nebo *Acyrthosiphon pisum* (Beton et al. 2012; Sainsbury et al. 2012; Terra a Ferreira 2012).

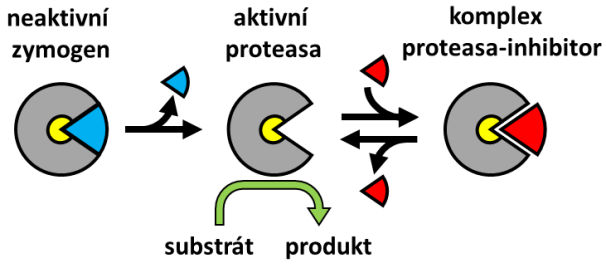
Asparaginylendopeptidasa (AEP) patřící do C13 rodiny cysteinových proteas byla původně popsána jako rostlinný protein v luštěninách, proto je také nazývaná legumain. Vyskytuje se v rostlinné i živočišné říši, zajímavé je, že u hmyzu řádu brouků (Coleoptera) AEP zcela chybí (Santamaria et al. 2012). Jedná se o cysteinovou endopeptidasu, která specificky štěpí za asparaginylovým zbytkem. Kromě endopeptidasové aktivity může v závislosti na pH a dalších podmínkách vykazovat i karboxypeptidasovou aktivitu (Dall a Brandstetter 2016). Rostlinné AEP se vyskytují v semenech, kde při procesu klíčení štěpí zásobní proteiny, a v rostlinných pletivech, kde jsou zapojeny do programované buněčné smrti a nahrazují ortologní kaspasy, které se u rostlin nevyskytují (Bonneau et al. 2008). Rostlinné i živočišné AEP jsou známé taky jako účinné peptidové ligasy a cyklasy (Dall et al. 2015), kdy přeměňují lineární polypeptidy na aktivní cyklické sloučeniny s různými biologickými funkcemi. Příkladem je cyklický trypsinový inhibitor SFTI ze slunečnice (Zauner et al. 2018).

Serinové a aspartátové proteasy, o kterých je převážná část předkládané disertační práce jsou detailně popsány v následujících kapitolách. Ostatními třídami proteas se tato práce nezabývá, proto zde nejsou popsány. Detaily o těchto proteasách jsou v přehledné knize o proteolytických enzymech „Handbook of proteolytic enzymes“ (Rawlings a Salvesen 2013).

1.2. Regulace aktivity proteolytických enzymů

Proteasy se uplatňují v řadě fyziologicky klíčových dějů, a proto musí být jejich aktivita přesně kontrolována na více úrovních. Množství proteas je v organizmu regulováno na úrovni genové

exprese, translace do proteinu a dalšími buněčnými regulačními mechanismy. Klíčovým mechanismem zabránění nechtěné proteolýzy je biosyntéza proteas ve formě neaktivních prekurzorů - proenzymů (zymogenů). K aktivaci zymogenu na aktivní enzym dochází při proteolytickém odštěpení tzv. aktivačního peptidu (propeptidu), který blokuje aktivní místo proteasy a také napomáhá při správném sbalení proteinu a zvyšuje jeho stabilitu (Obr. 2). Aktivace proběhne až v případě správné lokalizace proteasy buď v rámci buňky (např. katepsin D v lysosomu) nebo po sekreci mimo buňku (např. pepsin v žaludku). Propeptid má u zymogenů proteas funkci tzv. intramolekulárního inhibitoru, který zabraňuje vazbě substrátu do aktivního místa (Wiederanders, Kaulmann, a Schilling 2003). Bylo prokázáno, že syntetické peptidy odvozené od propeptidu jsou schopné působit jako intramolekulární inhibitory a blokovat zdrojové proteasy (Hánová et al. 2018; Máša et al. 2006; Horn et al. 2011).



Obrázek 2: Schématické znázornění příkladu dvouúrovňové regulace aktivity proteas.

Proteasy jsou syntetizovány jako neaktivní zymogeny, ve kterých je aktivní místo (žlutě) blokováno propeptidem (modře). Aktivace zymogenu probíhá proteolytickým odštěpením propeptidu za vzniku aktivní proteasy, která je schopna hydrolyzovat peptidovou vazbu substrátu (zeleně). Specifický inhibitor (červeně) je schopen se navázat do aktivního místa proteasy a vytvořit tak neaktivní komplex.

Mezi další významné regulátory aktivity proteas patří pH, iontová síla a hlavně specifické endogenní inhibitory peptidového nebo proteinového charakteru (Obr. 2). Tyto molekuly jsou hojně studovány, protože na základě jejich struktury, inhibiční specificity a mechanismu interakce s proteasami lze navrhovat nové bioaktivní látky, které lze využít např. při vývoji léčiv. Přirozeným inhibitorům proteas se věnuje kapitola 1.5.

1.3. Aspartátové proteasy z A1 rodiny pepsinu

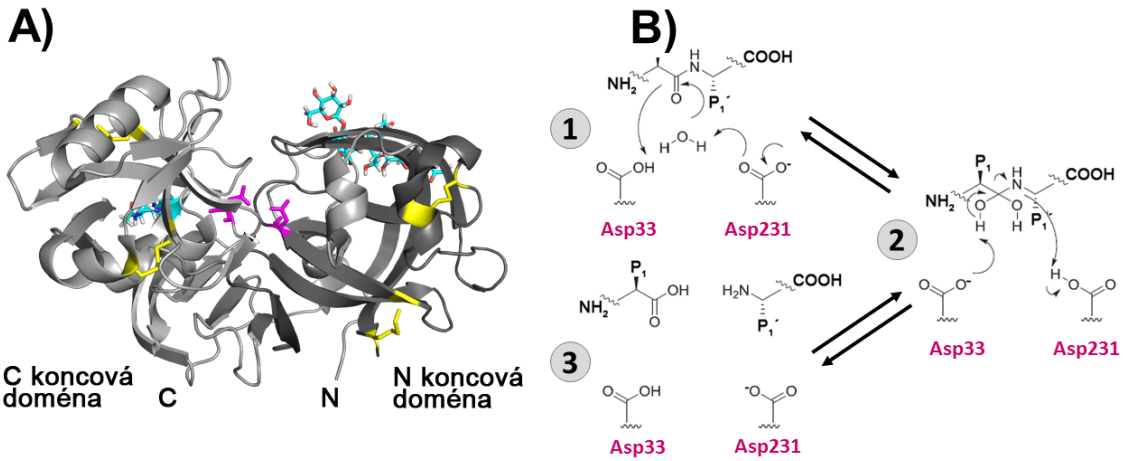
Aspartátové proteasy jsou široce rozšířené; jejich zástupci jsou popsány pro viry, bakterie a prvoky, v rostlinné i v živočišné říši. U savců jsou aspartátové proteasy zodpovědné za degradaci proteinů v trávicím traktu, dále se podílejí na katabolismu proteinů a peptidů na buněčné úrovni a také na regulaci homeostáze organismu (Rawlings a Salvesen 2013). Jako trávicí enzymy používají aspartátové proteasy někteří endoparazité, jako jsou zástupci prvoků (např. rody *Plasmodium, Trypanosoma*) a helmintů - hlístic (*Nematoda*) a motolic *(Trematoda*), nebo exoparazitiční členovci, např. klíšťata (*Ixodida*). Ti se živí krevními proteiny hostitele a aspartátové proteasy jsou důležitou složkou enzymové výbavy jejich trávicího traktu (Sojka et al. 2016). Rostlinné aspartátové proteasy se

podílejí na hydrolýze proteinů při procesu klíčení semen, opylení a interagují i s patogeny. Masožravé rostliny sekretují tyto enzymy a využívají je k trávení lapeného hmyzu (Kadek et al. 2014).

Třída aspartátových proteas se podle struktury dále dělí do pěti klanů, mezi nejznámější patří zástupci klanů AA a AD. Klan AA obsahuje rodinu A1 (pepsinová rodina), do které patří pepsin, katepsin D, renin a parazitární proteasy, a dále rodinu retrovirálních aspartátových proteas A2, jako je např. HIV-1 proteasa. Klan AD obsahuje presenilin a signální peptidasu, které se řadí mezi intramembránové proteasy (Rawlings a Salvesen 2013). Aspartátové proteasy tvoří nejméně početnou třídu proteas v lidském genomu, ale díky své nezastupitelné funkci jsou detailně studovány a považovány za perspektivní cílové proteiny pro vývoj potenciálních léčiv u řady onemocnění.

1.3.1. Prostorová struktura aspartátových proteas z A1 rodiny pepsinu

Aspartátové proteasy pepsinové rodiny jsou syntetizovány jako neaktivní zymogeny s propeptidem na N konci proteinu, který je odštěpen během procesu aktivace a tvorby zralého aktivního enzymu. Někteří zástupci jsou N glykosylováni (např. lidský katepsin D). Architektura proteinu je stabilizována pomocí tří konzervovaných disulfidových můstků, (Rawlings a Salvesen 2013). Molekula obsahuje dvě podjednotky, které se označují jako N a C terminální domény; vykazují sekvenční homologii a je tedy pravděpodobné, že vznikly duplikací jednoho genu (Obr. 3 A). Aktivní místo se v prostorové struktuře proteinu nachází mezi oběma podjednotkami a každá z nich nese jeden katalytický zbytek kyseliny asparagové. V současné době je v PDB databázi řada prostorových struktur aspartátových proteas pepsinové rodiny ve formě zymogenů, zralých enzymů nebo jejich komplexů s inhibitory.



Obrázek 3: Prostorová struktura a katalytický mechanismus katepsinu D.

A) Celkový pohled na strukturu lidského katepsinu D ve stužkovém modelu (PDB kód: 1lyb) tvořenou dvěma řetězci (světle resp. tmavě šedě), vyznačeny jsou N a C konce proteinu a členění na strukturní domény. Fialově jsou vyznačené katalytické zbytky (Asp33 a Asp231), žlutě cysteinové zbytky tvořící disulfidové můstky a modře N glykosylace. B) Schématické znázornění mechanismu štěpení peptidového substrátu u aspartátových proteas. V aktivním místě jsou dva katalytické zbytky kyseliny

asparagové, které vážou molekulu vody. (1) Aktivovaná molekula vody nukleofilně atakuje peptidovou vazbu navázaného substrátu, (2) dochází k přesunu elektronu a protonů v aktivním místě proteasy a štěpení vazby, (3) produkty jsou uvolněny z aktivního místa a enzym je regenerován vazbou nové molekuly vody. Převzato a upraveno podle (Coates et al. 2008).

1.3.2. Popis aktivního místa a mechanismus katalýzy u aspartátových proteas z A1 rodiny pepsinu

Všechny aspartátové proteasy pepsinové rodiny patří mezi endopeptidasy, které štěpí polypeptidový řetězec uvnitř sekvence substrátu. Aktivní místo je tvořeno podélnou dutinou a je schopno interagovat s peptidem o délce 6-10 aminokyselin po obou stranách štěpené vazby. Na dně aktivního místa jsou dva zbytky kyseliny asparagové (Asp33 a Asp231, číslování dle lidského katepsinu D), které v nativním stavu enzymu pevně vážou molekulu vody. Jeden ze zbytků tuto molekulu deprotonuje a aktivuje ji pro nukleofilní atak peptidové vazby substrátu, který je vázán do aktivního místa enzymu, za vzniku tetrahedrálního intermediátu. Intermediát je stabilizovaný vodíkovou vazbou k druhému aspartátovému zbytku. Štěpení peptidové vazby je doprovázeno přesunem protonu k amino skupině a hydroxylu ke karbonylu. Poté dochází k uvolnění produktů z aktivního místa a proteasa je regenerována vazbou nové molekuly vody (Obr. 3 B, str. 8) (Coates et al. 2008). Většina aspartátových proteas pepsinové rodiny je aktivní v kyselém pH, příkladem mohou být savčí trávicí enzymy, který patří mezi hlavní proteasy gastrointestinálního traktu.

1.3.3. Katepsin D

Katepsin D z pepsinové rodiny aspartátových proteas je exprimován ve všech typech savčích buněk, kde působí intracelulárně v lysosomech a pozdních endosomech. V těchto buněčných organelách patří mezi hlavní proteolytické enzymy a podílí se zde na katabolické degradaci proteinů a proteolytické inaktivaci např. inhibitorů a peptidových hormonů. Účastní se také tvorby peptidových fragmentů, které se na MHC II molekulách presentují jako antigeny. Další specifické role byly pro katepsin D popsány i v cytosolu, buněčném jádře a při extracelulární lokalizaci (Beneš, Větvička, a Fusek 2008). Přirozené endogenní inhibitory katepsinu D nebyly doposud popsány.

1.3.3.1. Prostorová struktura a aktivita katepsinu D

Lidský katepsin D je syntetizován na ribosomech drsného endoplasmatického retikula v neaktivní formě tzv. preproenzymu s N koncovou signální sekvencí, která je odštěpena v lumen endoplasmatického retikula. Za touto sekvencí pokračuje 44 aminokyselin dlouhý propeptid, který blokuje aktivní místo ve struktuře zymogenu. Katepsin D je N glykosylován v pozicích Asn70 a Asn199 oligosacharidovými řetězci vysokomannosového typu (Obr. 3 A, str. 8). Mannosový cukerný zbytek je v cis Golgi aparátu fosforylován a rozpoznán manosa-6-fosfátovým receptorem, který transportuje neaktivní zymogen katepsinu D do endosomu nebo lysosomu. V těchto buněčných

organelách se katepsin D působením nízkého pH uvolní z transportního receptoru (von Figura a Hasilik 1986).

V kyselém prostředí cílových buněčných kompartmentů dochází k aktivaci zymogenu katepsinu D. Propeptid se v kyselém pH vyvazuje z aktivního místa a stává se tak přístupný k proteolytickému odštěpení, které může proběhnout autokatalyticky nebo působením lysosomálních cysteinových katepsinů (Máša et al. 2006). Takto vzniká jednořetězcová forma katepsinu D o molekulové hmotnosti 44 kDa. Lidský katepsin D je dále proteolyticky štěpen za vzniku dvouřetězcové formy, která se skládá z N koncového lehkého řetězce (13 kDa) a C koncového těžkého řetězce (31 kDa) (Gieselmann, Hasilik, a von Figura 1985). Prostorová architektura katepsinu D (Obr. 3 A, str. 8) je podobná jiným aspartátovým proteasám z rodiny pepsinu, hlavní rozdíly jsou v rozdělení molekuly na dva řetězce a přítomnost čtvrtého disulfidového můstku (Baldwin et al. 1993).

Podlouhlé aktivní místo katepsinu D umožňuje vázat až osm aminokyselin dlouhý substrát v pozicích P5-P3´ a preferuje štěpení mezi hydrofobními zbytky. Katepsin D má poměrně široký rozsah pH v kyselé oblasti, kde vykazuje vysokou enzymovou aktivitu (pH 3,5-5,0). K nejlepším substrátům podle hodnoty k_{cat}/K_M patří syntetický oktapeptid Lys-Pro-Phe-Leu-Phe↓Nph-Arg-Leu, často používaný k měření aktivity katepsinu D (šipka znázorňuje místo štěpení, Nph znamená p-nitro-L-fenylalanin) (Dunn a Hung 2000; Sun et al. 2013).

1.3.3.2. Fyziologie a patofyziologie spojené s katepsinem D

Katepsin D hraje významnou roli při apoptose, kdy se z lysosomu uvolňuje do cytosolu a zapojuje do hydrolýzy buněčných substrátů a komponent kaspasových signálních drah (Benes, Vetvicka, a Fusek 2008). Z patofyziologického hlediska je katepsin D spojován s řadou onemocnění, zejména Alzheimerovou chorobou, aterosklerosou a nádorovými onemocněními.

Alzheimerova choroba patří mezi neurodegenerativní onemocnění mozku, kdy dochází k ukládání β amyloidních peptidů a vzniku plaků na povrchu neuronů. β amyloidní peptidy jsou výsledkem chybného proteolytického štěpení amyloidního prekursorového proteinu (APP) působením aspartátové proteasy rodiny pepsinu β sekretasou a následně γ sekretasou. Z degenerujících neuronů se uvolňuje katepsin D (Cataldo a Nixon 1990), který v *in-vitro* testech simuluje štěpení APP oběma sekretasami (Sadik et al. 1999).

Přítomnost neaktivního zymogenu katepsinu D v extracelulárním prostoru má silný proliferační efekt na rakovinné buněčné linie. Zymogen přidaný do média nádorových buněk stimuluje jejich proliferaci a naopak redukce množství zymogenu pomocí např. specifických protilátek proliferaci potlačuje. Extracelulárně sekretovaný zymogen katepsinu D, který má mitogenní aktivitu pro nádorové buňky interaguje s doposud nepopsaným membránovým receptorem (Větvička et al. 1997). Funkce zymogenu jako autokrinního růstového faktoru byla potvrzena u nádorových buněk

prsu, prostaty, vaječníků a plic, ale předpokládá se i u dalších typů nádorových buněk. Kromě zymogenu přispívá k rakovinnému bujení také aktivovaná forma katepsinu D, která je schopná degradovat extracelulární matrix a aktivovat růstové faktory (Fusek, Větvičková, a Větvička 2007).

1.3.4. Medicinálně významné lidské aspartátové proteasy z A1 rodiny pepsinu

Kromě katepsinu D patří k medicinálně nejvýznamnějším zástupcům aspartátových proteas rodiny pepsinu v lidském proteomu β sekretasa, renin a katepsin E.

β sekretasa

Transmembránová β sekretasa (z angl. β site amyloid precursor protein-cleaving enzyme 1) se účastní několika fyziologických procesů, jako je diferenciace a smrt buněk či regulace imunity. Změna aktivity β sekretasy vede k patologiím spojeným s nervovou soustavou. Nejlépe prozkoumána je její role u Alzheimerovy choroby, kde se β sekretasa účastní štěpení amyloidního prekursorového proteinu na β amyloidní peptidy (De Strooper, Vassar, a Golde 2010). Molekula β sekretasy obsahuje kromě katalytické domény ještě transmembránovou a cytosolární doménu na C konci proteinu. Posttranslační modifikace a dimerizace zvyšují aktivitu β sekretasy. V současné době se výzkum zaměřuje na vývoj nových specifických inhibitorů β sekretasy, které by blokovaly tvorbu β amyloidního peptidu, ale nebránily fyziologickým funkcím (Barao et al. 2016).

Renin

Regulace krevního tlaku a rovnováhy extracelulárních tekutin v lidském organismu je založena na kaskádě proteolytických kroků, které jsou součástí tzv. renin-angiotensinového systému. Dysfunkce tohoto systému je spojována s rozvojem vysokého krevního tlaku a srdečním selháním. Aspartátová proteasa rodiny pepsinu renin je zodpovědná za první aktivační krok v tomto systému, kdy specificky odštěpí deset aminokyselinových zbytků z N konce angiotensinogenu a spouští tvorbu angiotensinových hormonů. Hlavním zdrojem reninu v krevním řečišti jsou ledviny, ve kterých je exprimován a sekretován. Hladina reninu v krvi je přísně regulována a řízena krevním tlakem a koncentrací soli v plasmě (Sparks et al. 2014).

Katepsin E

Katepsin E má podobnou substrátovou specifitu jako katepsin D; jejich inhibiční specifita se liší a lze ji rozlišit podle citlivosti katepsinu E k PI3 proteinovému inhibitoru z hlístice (*Ascaris suum*) (Keilová a Tomášek 1972) nebo synteticky připravenému peptidovému inhibitoru (Kitamura et al. 2009). Na rozdíl od katepsinu D tvoří katepsinu E kovalentní dimer s celkovou Mr 84 kDa, kdy jsou dvě aktivní monomerní jednotky spojeny disulfidovým můstkem. Katepsin E je produkován v řadě tkání, např. v gastrointestinálním traktu a v buňkách imunitního systému (Rawlings a Salvesen 2013). Funkce katepsinu E není plně objasněna a je spojována s nádorovými onemocněními nebo atopickým ekzémem.

1.3.5. Aspartátové proteasy z A1 rodiny pepsinu u bezobratlých

Aspartátové proteasy pepsinové rodiny byly popsány u zástupců hmyzích řádů. Společně s cysteinovými proteasami se zde podílejí na degradaci proteinů v trávicím traktu. Specifická úloha v komplexním trávicím systému byla popsána u larev zrnokaze skvrnitého (*Callosobruchus maculatus*). Ty reagují na přítomnost proteasových inhibitorů pro cysteinové proteasy v potravě tak, že začnou produkovat aspartátovou proteasu typu katepsin D, která tyto inhibitory degraduje a zvyšuje tak aktivitu trávicích cysteinových proteas (Ahn a Zhu-Salzman 2009). Aktivita aspartátových proteas byla nalezena v řadě čeledí řádu brouků (Terra a Ferreira 2012), včetně mandelinky bramborové (*Leptinotarsa decemlineata*) (Brunelle et al. 1999). U mouchy domácí byly nalezeny tři rozdílné aspartátové proteasy typu katepsin D s odlišnou tkáňovou distribucí a byl studován mechanismus sekrece do trávicího traktu (Padilha et al. 2009).

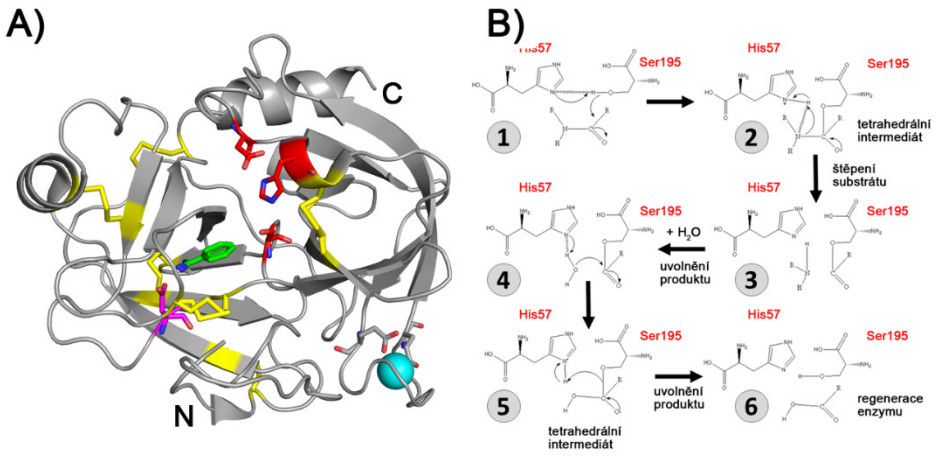
Aspartátové proteasy parazitů hrají důležitou roli při interakci parazita s hostitelem. Jsou součástí multienzymové kaskády degradace hemoglobinu u parazitů živících se krví, které mají trávení založené na proteasách funkčních v kyselé oblasti pH. Zde kooperují cysteinové a aspartátové proteasy, které iniciují štěpení krevních proteinů a zpřístupňují je dalším proteasám (Horn et al. 2009). Kromě trávicí funkce se tyto proteasy také mohou podílet na pronikání parazita do tkání nebo modulaci imunitního systému hostitele. Z tohoto důvodu jsou aspartátové proteasy slibné cílové molekuly pro léčbu parazitárních onemocnění. Mezi nejvíce studované aspartátové proteasy parazitických prvoků a helmintů patří plasmepsiny ze zimničky tropické (*Plasmodium falciparum*) způsobujícího malárii a katepsin D z krevničky střevní (*Schistosoma mansoni*) způsobující schistosomósu (Sojka et al. 2016).

1.4. Serinové proteasy z S1 rodiny chymotrypsinu

Mezi serinové proteasy patří enzymy, jejichž katalytický mechanismus využívá hydroxylovou skupinu serinového zbytku, který při hydrolýze peptidové vazby funguje jako nukleofil. Tuto třídu proteas lze rozdělit na základě prostorových struktur do 14 klanů a podle podobnosti v aminokyselinové sekvenci do přibližně 60 různých rodin. Nejznámější serinové proteasy jsou trypsin a chymotrypsin, které patří do S1 rodiny (označované jako rodina chymotrypsinu) z klanu PA. Jedná se o největší rodinu proteas co do počtu sekvencí. Zástupci této rodiny jsou typicky endopeptidasy s konzervovanou katalytickou triádou His, Asp, Ser a vyhraněnou substrátovou specifitou, která je určena vazbou do S_1 podmísa enzymu. Serinové proteasy S1 rodiny mají signální sekvenci pro sekreci proteinu z buňky a jsou produkované ve formě neaktivního zymogenu. Délka N koncového propeptidu se výrazně liší od dvou aminokyselinových zbytků u katepsinu G až po stovky zbytků u proteas z kaskády srážení krve nebo komplementu (Rawlings a Salvesen 2013). Aktivace je řízena odštěpením propeptidu, po kterém dochází ke konformační změně struktury a tvorbě aktivního enzymu.

1.4.1. Prostorová struktura serinových proteas z S1 rodiny chymotrypsinu

Serinové proteasy chymotrypsinové rodiny mají konzervovanou molekulární architekturu, která je tvořena dvěma doménami se šesti β strukturami (Obr. 4 A), mezi nimiž jsou umístěny zbytky katalytické triády. Rozdílná substrátová specifita mezi trypsinem, který preferuje bazické aminokyseliny v pozici P_1, a chymotrypsinem s preferencí pro objemné hydrofobní zbytky je způsobena záměnou aminokyseliny v pozici 189 na dně S_1 podmísta enzymu (Perona a Craik 1995). Zbytek Asp189 u trypsinu zvyšuje polaritu vazebného podmísta S_1 a kompenzuje náboj bazických zbytků (argininu nebo lysinu). Tvoří solný můstek s argininovým zbytkem substrátu, zatímco interakce s lysinem je zprostředkována přes molekulu vody. Záměna Asp189 mutagenezí za Ser189 u trypsinu vedla k poklesu hodnot katalytické účinnosti enzymu o pět řádů (Craik et al. 1987). U chymotrypsinu je dno S_1 vazebného podmísta tvořeno serinovým zbytkem, který je méně polární a umožnuje vazbu hydrofobním zbytkům v pozici P_1 (zejména Tyr, Phe, Trp a Leu). Při vazbě substrátu jsou z S_1 podmísta chymotrypsinu vytěsněny tři molekuly vody (Jelinek et al. 2004).



Obrázek 4: Prostorová struktura a katalytický mechanismus trypsinu.

A) Celkový pohled na prostorovou strukturu hovězího trypsinu ve stužkovém modelu (PDB kód: 4i8h), vyznačeny jsou N a C konce proteinu, červeně katalytické zbytky (His57, Asp102, Ser195) a žlutě cysteinové zbytky tvořící disulfidové můstky. Zelený tyčkový model znázorňuje molekulu inhibitoru benzamidinu, který je navázán v S_1 podmístě enzymu a tvoří vodíkové vazby s Asp189 (fialově), který je na dně S_1 podmísta. Znázorněny jsou dva glutamátové zbytky (Glu70 a Glu80), které váží iont vápníku (modře). B) Schématické znázornění štěpení peptidového substrátu u serinových proteas. V aktivním místě trypsinu jsou katalytické zbytky His57 a Ser195. (1) Substrát se váže do aktivního místa tak, aby karbonyl štěpené vazby byl v blízkosti Ser195, (2) hydroxylová skupina Ser195 nukleofilně atakuje karbonylový uhlík peptidové vazby za vzniku tetrahedrálního intermediátu, (3) dochází ke štěpení vazby, uvolnění produktu s nově utvořeným N koncem a vzniku kovalentního intermediátu acyl-enzymu, (4) molekula vody vstupuje do reakce (5) za vzniku druhého

tetrahedrálního intermediátu, (6) v posledním kroku dochází k uvolnění produktu s nově utvořeným C koncem a regeneraci enzymu. Převzato a upraveno podle (Hedstrom 2002).

1.4.2. Popis aktivního místa a mechanismus katalýzy u serinových proteas z S1 rodiny chymotrypsinu

Aktivní místo serinových proteas chymotrypsinové rodiny je tvořeno katalytickou triádou His57, Asp102 a Ser195. Serinový zbytek slouží jako nukleofil, který atakuje štěpenou peptidovou vazbu, imidazolový kruh histidinu jako donor protonů a aspartátový zbytek pomáhá udržovat správnou orientaci imidazolového kruhu. Ve struktuře těchto proteas je vysoce konzervovaný zbytek Ser214, který vytváří polární prostředí v aktivním místě enzymu a vodíkovou vazbu stabilizuje náboj na katalytickém Asp102 (McGrath et al. 1992). Štěpení peptidové vazby (Obr. 4 B, str. 13) probíhá tak, že při správné orientaci substrátu v aktivním místě dochází k nukleofilnímu ataku hydroxylové skupiny katalytického Ser195 na karbonyl štěpené peptidové vazby za vzniku tetrahedrálního intermediátu. V dalším kroku je peptidová vazba štěpena, uvolňuje se první produkt s nově utvořeným N koncem a vzniká kovalentní acyl-enzymový intermediát. Následně je tento intermediát hydrolyzován molekulou vody, která vstupuje do reakce. Na závěr se uvolňuje produkt s nově utvořeným C koncem a proteasa je regenerována (Hedstrom 2002). Pro malé syntetické peptidové substráty je krok určující rychlost reakce tvorba acyl-enzym intermediátu a pro velké proteinové substráty to je vazba substrátu do aktivního místa (Perona a Craik 1995).

1.4.3. Trypsin a chymotrypsin

Savčí trypsin i chymotrypsin jsou produkovány jako trávicí enzymy v buňkách slinivky břišní a skladovány v sekrečních váčcích. Jsou exprimovány ve formě neaktivního zymogenu (trypsinogenu a chymotrypsinogenu). Po uvolnění do trávicího traktu je trypsinogen proteolyticky aktivován enterokinasou a aktivní trypsin dále aktivuje chymotrypsinogen. Po odštěpení propeptidu následuje reorganizace aktivního místa enzymu, při které dochází ke správnému sterickému uspořádaní zbytků katalytické triády. U chymotrypsinu je odštěpený propeptid ukotven ke zbytku enzymu disulfidovým můstkem mezi Cys1 a Cys122. Jeho úlohou je bránit nespecifické proteolytické aktivaci, ale nemá vliv na aktivitu nebo substrátovou specifitu enzymu (Venekei, Graf, a Rutter 1996).

Trypsin a chymotrypsin jsou inhibované obecnými inhibitory serinových proteas, jako jsou např. diisopropyl fluorofosfát, Pefabloc, bakteriální leupeptin, proteasové inhibitory proteinového charakteru BPTI (z angl. Bovine Pancreatic Trypsin Inhibitor) ze slinivky břišní a STI (z angl. Soybean Trypsin Inhibitor) ze sóji. Specifickým syntetickým inhibitorem pro trypsin je Tos-Lys-CH_2Cl a benzamidin, a chymotrypsin je specificky blokován Tos-Phe-CH_2Cl a mikrobiálním chymostatinem. Oba enzymy mají nejvyšší aktivitu v oblasti kolem pH 8,0, ale tato hodnota se může mírně lišit pro konkrétní izoenzymy.

Trypsin a chymotrypsin vykazují 50% identitu v aminokyselinové sekvenci. Molekula trypsinu je jednořetězcová, u chymotrypsinu je rozdělena na tři řetězce, které jsou spojeny disulfidovými můstky (Rawlings a Salvesen 2013; Wang, Bode, a Huber 1985). Trypsin má ve své struktuře místo pro specifickou vazbu vápenatého iontu (Obr. 4 A, str. 14), které je tvořeno dvěma zbytky glutamátu (Glu70 a Glu80). Vazba vápníku zvyšuje aktivitu trypsinu a stabilizuje enzym proti autodegradaci (Bode a Schwager 1975).

1.4.4. Medicinálně významné serinové proteasy z S1 rodiny chymotrypsinu

Trombin

Kardiovaskulární onemocnění patří celosvětově k hlavním příčinám úmrtí a chemoterapeutika cílená proti srážení krve jsou nejčastěji předepisované léky. Výsledkem proteolytických dějů v koagulační kaskádě je aktivace zymogenu protrombinu na aktivní trombin. Ten je klíčovým enzymem koagulace, protože štěpí fibrinogen na fibrin, který polymeruje a umožňuje tvorbu krevní sraženiny. V prostorové struktuře trombinu jsou dvě alosterická místa, která jsou v závislosti na navázaném ligandu schopná specificky regulovat aktivitu enzymu (Huntington 2012). Velice důležitá je regulace trombinu endogenními inhibitory, mezi které patří zejména serpiny (z angl. Serine Protease Inhibitors) (Huntington 2014). Výzkum nových chemoterapeutik cílených na trombin se zaměřuje především na ligandy, které interagují s exo-místy mimo aktivní místo enzymu (Mehta, Jin, a Desai 2014).

Kalikreiny

Serinové proteasy kalikreinového typu (KLK) tvoří rodinu vysoce homologických sekretovaných proteas (KLK 1-KLK 15). Zástupci tkáňových KLK tvoří největší souvislý klastr proteas v lidském genomu (Yousef a Diamandis 2001) a jsou produkovány v různých tkáních. Vzdálenější příbuznost vykazuje plasmový KLK. Na publikovaných prostorových strukturách KLK jsou patrné typické strukturní rysy serinových proteas rodiny chymotrypsinu (Kalinska et al. 2016) a povrchové smyčky v okolí aktivního místa přispívající k drobným změnám v substrátové specifitě těchto enzymů. Významnou roli hraje tzv. „kalikreinová smyčka“ u KLK 1, 2, 3 a její interakce se substrátem (Skala et al. 2014). KLK proteasy lze podle substrátové specifity rozdělit na zástupce trypsinového typu, které štěpí peptidové substráty za bazickou aminokyselinou (Arg nebo Lys) a na KLK chymotrypsinového typu, které preferují hydrofobní aminokyseliny (Phe nebo Tyr); někteří zástupci mají duální specifitu (Kalinska et al. 2016).

KLK 4 se podílí na tvorbě zubní skloviny, když degraduje proteiny extracelulární matrix. Osm zástupců KLK proteas bylo detekováno v nejsvrchnější části kůže (*stratum corneum*), kde jsou zapojeny ve fyziologických procesech jako je deskvamace. KLK jsou spojovány s řadou patologií, jako jsou např. onemocnění dýchacích cest a neurodegenerativní a nádorová onemocnění (Kryza et al.

2016). K jejich přesnému fungovaní je nutná regulace aktivity pomocí endogenních inhibitorů (Masurier et al. 2018). Mezi endogenní inhibitory KLK proteas v krevní plasmě patří serpiny (Luo a Jiang 2006), některé KLK jsou inhibovány α_2 makroglobuliny a především LEKTI inhibitory z Kazalovy rodiny. LEKTI je multidoménový protein kódovaný genem SPINK5 (Goettig, Magdolen, a Brandstetter 2010), u kterého bylo prokázáno, že jeho mutace vede např. k rozvoji atopického ekzému. Aktivita KLK proteas je dále ovlivněna pH nebo zinkovými ionty, které interagují s aktivním místem některých zástupců.

1.4.5. Serinové proteasy z S1 rodiny chymotrypsinu u bezobratlých

U hmyzu byla identifikována řada sekvencí patřících k serinovým proteasám S1 chymotrypsinové rodiny. Mají společné rysy se savčími homology, včetně katalytické triády a definovaného S_1 podmísta, které určuje substrátovou specifitu. Někteří zástupci ale nemají vazebné místo pro vápník, který stabilizuje molekulu a na rozdíl od lidského trypsinu a chymotrypsinu nejsou stabilní v kyselém pH. Trypsinové nebo chymotrypsinové proteasy byly popsány u polokřídlých (Hemiptera), dvoukřídlých (Diptera), motýlu (Lepidoptera), brouku (Coleoptera) a blech (Siphonaptera). Variabilita trávicích izoenzymů hmyzích serinových proteas má vliv na jejich sensitivitu k rostlinným proteasovým inhibitorům produkovaným jako obranné molekuly (Terra a Ferreira 2012).

Serinové proteasy zprostředkovávají interakci parazitických helmintů a prvoků s hostitelem (Dvořák a Horn 2018). U motolic (*Trematoda*) byly prokázány serinové proteasy rodiny chymotrypsinu, např. u krevničky rodu *Schistosoma*, kde bylo identifikováno 16 unikátních sekvencí těchto proteas. Jsou zapojeny do vývoje parazitů, nutričního příjmu, invaze parazita do tkání i buněk hostitele, brání srážení krve a napomáhají úniku parazita před imunitním systémem hostitele (Yang et al. 2015). Tyto proteasy mohou sloužit k prevenci a léčbě parazitárních onemocnění a představuji perspektivní molekulární cíle pro vývoji nových vakcín a chemoterapeutik (Pearson et al. 2015; Tallima et al. 2017).

1.5. Přirozené a syntetické inhibitory proteas

Přirozené inhibitory proteas lze klasifikovat dle molekulové hmotnosti na nízkomolekulární (často peptidové) inhibitory a vysokomolekulární inhibitory proteinového charakteru. Mezi nízkomolekulární inhibitory proteas patří řada molekul izolovaných z nižších organismů, jako jsou například inhibitor cysteinových, serinových a threoninových proteas leupeptin z bakterií rodu *Streptomyces* (Hozumi et al. 1972), selektivní inhibitor cysteinových proteas papainového typu nazývaný E64 (trans-epoxysukcinyl-L-leucylamido-4-guanidinobutan) z plísní rodu *Aspergillus* (Hashida et al. 1980) nebo specifický inhibitor aspartátových proteas pepsinové rodiny pepstatin izolovaný z bakterií rodu *Actinomyces* (Knight a Barrett 1976). Proteinové inhibitory jsou velice rozšířené molekuly, které se uplatňují v řadě fyziologických i patologických dějů. Regulují endogenní

proteolytické procesy, ale jsou také sekretované z organismu, kde mohou sloužit jako obranné molekuly proti různým patogenům nebo se mohou naopak účastnit interakce patogena s hostitelem a regulovat proteolýzu hostitele (Chmelar et al. 2011; Tarasuk et al. 2009). V současné době je známo více než 80 rodin proteinových inhibitorů proteas (Rawlings a Barrett 2010). Většina rodin inhibitoru je cílena jen proti určité rodině proteas, ale některé rodiny mají širší selektivitu a regulují více skupin proteasových rodin. Příkladem jsou serpiny proti serinovým a cysteinovým proteasám, tyropiny proti cysteinovým a aspartátovým proteasám nebo inhibitory Kunitzovy rodiny, které blokují aktivitu serinových, cysteinových i aspartátových proteas (Mihelic a Turk 2007; Bendre, Ramasamy, a Suresh 2018; Law et al. 2006).

Syntetické inhibitory proteas nejčastěji vycházejí ze sekvencí substrátů nebo přirozených inhibitorů daných enzymů, které se modifikují tak, aby vznikly inhibitory s vyšší potencí a lepší specifitou. Jedním z přístupů racionálního navrhovaní je protekce štěpené vazby tak, aby nebyla proteolyticky hydrolyzovatelná. Příkladem je redukovaná amidová vazba (-CH_2-NH-) nebo substituce aminokyselinových zbytků kolem štěpené vazby z přirozené L formy na D formu (Motwani et al. 2015). Dále je možné ke specifické peptidové sekvenci, která se váže do aktivního místa proteasy, připojit reaktivní skupinu, která vytvoří kovalentní vazbu s katalytickým zbytkem proteasy. Mezi reaktivní skupiny pro serinové a cysteinové proteasy patří např. chlormetylketony, dimetylketony, nitrily nebo vinylsulfony (Powers et al. 2002). Na základě specifických inhibitorů lze konstruovat selektivní aktivitní sondy, které jsou schopné detekovat a vizualizovat aktivní proteasy *in vitro* a *in vivo* (Sanman a Bogyo 2014).

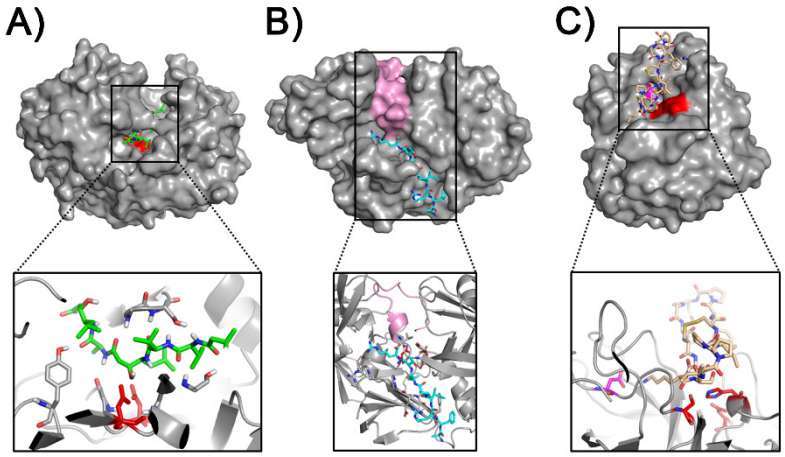
1.5.1. Peptidové inhibitory aspartátových a serinových proteas

Univerzálním inhibitorem aspartátových proteas pepsinové rodiny je hexapeptid pepstatin (Obr. 5 A, str. 18) izolovaný z bakterií rodu *Actinomyces* (Knight a Barrett 1976), který je účinný v pikomolárních koncentracích. Sekvence pepstatinu (Iva-Val-Val-Sta-Ala-Sta) obsahuje modifikaci kyselinou izovalerovou (Iva) a nestandardní aminokyselinu statin (Sta). Zbytek Sta4 zaujímá v aktivním místě enzymu pozici mezi podmísty P_1 a P_1' a jeho hydroxylová skupina přímo interaguje s karboxylovými skupinami katalytických aspartátů, odkud vytěsňuje molekulu vody, která je nezbytná pro hydrolýzu peptidové vazby. Statinový zbytek tak napodobuje transitní stav aktivního enzymu. Molekula pepstatinu byla cíleně minimalizována a optimalizována pro účinnou inhibici katepsinu D (Majer et al. 1997). Na základě takto upravené molekuly pepstatinu byly v rámci disertační práce připraveny aktivitní sondy pro vizualizaci aspartátových proteas (Nussbaumerova et al. 2010).

Další inhibitory aspartátových proteas na peptidové bázi byly odvozeny ze struktury propeptidu, který v neaktivním zymogenu blokuje aktivní místo proteasy. Detailně byly studovány propeptidy lidského katepsinu D (Mášaa et al. 2006) a jeho homologa z klíštěte obecného (*Ixodes*

ricinus) (Hánová et al. 2018). Byly připraveny různě dlouhé a mutované formy propeptidu v syntetické podobě, jejichž analýza ukázala, že za inhibiční funkci je zodpovědná N koncová část propeptidu. Pro odvozený syntetický oligopeptid byl identifikován strukturní mechanismus inhibice, který spočívá ve vazbě do exo-místa na povrchu enzymu (Obr. 5 B, str. 18). Tato interakce indukuje konformační změnu v N konci zralého enzymu, jenž následně blokuje přístup substrátu do aktivního místa enzymu (Hánová et al. 2018).

SFTI (z angl. Sunflower Trypsin Inhibitor) izolovaný ze slunečnice roční (*Helianthus annuus*) je nejmenší inhibitor patřící do rodiny Bowman-Birk proteasových inhibitorů (Luckett et al. 1999). SFTI je bicyklický 14 aminokyselin dlouhý peptid, který je stabilizován intramolekulárním disulfidovým můstkem. Jedná se o inhibitor serinových proteas trypsinového typu, který vazbou do aktivního místa enzymu napodobuje substrát a do S_1 podmísťa umisťuje Lys5 (Obr. 5 C). Unikátní strukturní a funkční vlastnosti dělají ze SFTI vhodný templát pro navrhování inhibičních derivátů se změněnou selektivitou (de Veer et al. 2015).



Obrázek 5: Prostorové struktury aspartátové a serinové proteasy v komplexu s peptidovými inhibitory.

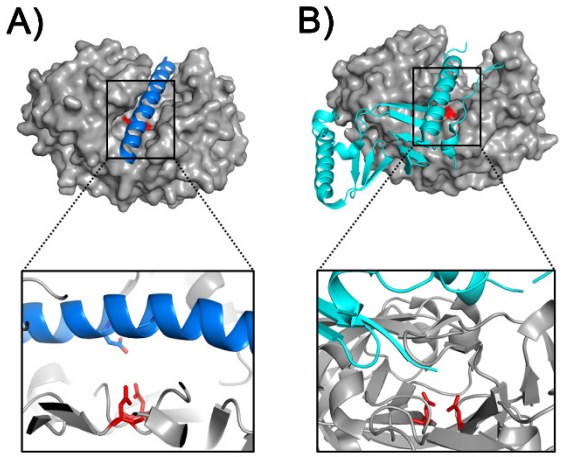
A) Lidský katepsin D s inhibitorem pepstatinem v aktivním místě (PDB kód: 1lyb), B) klíštěcí katepsin D s peptidovým inhibitorem v exo-místě (PDB kód: 5n70) a C) hovězí trypsin se SFTI inhibitorem (PDB kód: 1sfi). V horní části obrázku je celkový pohled na jednotlivé komplexy, kde jsou enzymy znázorněny jako šedý povrchový model, katalytické zbytky jsou červeně a inhibitory v tyčinkovém modelu. N konec katepsinu D, který alostericky blokuje aktivní místo, je růžově. V dolní části je detailní pohled na vazebné místo enzymu (ve stužkovém modelu) s inhibitorem. Tyčkovým modelem jsou červeně zvýrazněny katalytické zbytky enzymu. Důležité interakční zbytky jsou šedivě u katepsinu D a fialově u trypsinu (Asp189 v S_1 podmístě).

1.5.2. Proteinové inhibitory aspartátových a serinových proteas

1.5.2.1. Proteinové inhibitory aspartátových proteas

Přirozené inhibitory aspartátových proteas jsou vzácné a mechanismus interakce s enzymem byl popsán jen pro několik z nich. V PDB databázi jsou dostupné dvě prostorové struktury inhibitorů aspartátových proteas proteinového charakteru navázané na cílovou proteasu. Jedná se o inhibitor IA3 izolovaný z kvasinky *Saccharomyces cerevisiae* v komplexu s kvasinkovou proteinasou A (Li et al. 2000) a pepsinový inhibitor PI3 z hlístice *Ascaris suum* v komplexu s pepsinem (Ng et al. 2000). Další známé inhibitory aspartátových proteas jsou SQAPI z dýně (z angl. Squash Aspartic Protease Inhibitor, Christeller et al. 1998), doména 2 equistatinu ze sasanky mořské, který patří do tyropinové rodiny (Lenarcic a Turk 1999) a inhibitory z Kunitzovy rodiny izolované z lilkovitých rostlin (Mareš et al. 1989; Cater et al. 2002). U těchto inhibitorů dosud nebyl strukturně popsán mechanismus inhibice.

Prostorová struktura komplexu proteinasy A s inhibitorem IA3 ukázala (Obr. 6 A), že inhibitor, který v roztoku nemá žádnou sekundární strukturu, se při vazbě na proteasu stává strukturovaný a tvoří α šroubovici v aktivním místě enzymu. Jedná se o unikátní indukci sekundární struktury vazbou na enzym. IA3 tvoří s proteinasou A řadu kontaktů, které stabilizují vzniklý komplex. S katalytickým centrem enzymu interaguje zbytek Lys18 inhibitoru (Li et al. 2000). Pro inhibitor PI3 se podařilo vyřešit prostorovou strukturu jak volného inhibitoru, tak i inhibitoru v komplexu s pepsinem (Obr. 6 B). PI3 je tvořen dvěma doménami a interaguje s proteasou N koncovou částí. Tři aminokyselinové zbytky z N konce inhibitoru (Gln-Phe-Leu) obsazují pozice S_3'-S_1' v aktivním místu pepsinu a blokují vstup substrátu (Ng et al. 2000).



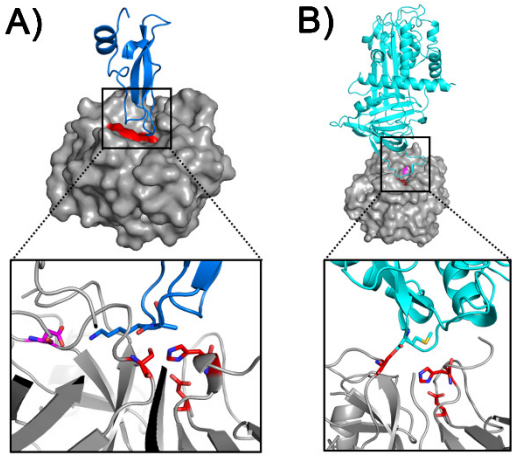
Obrázek 6: Prostorové struktury aspartátových proteas v komplexu s proteinovými inhibitory. *A) Kvasinková proteinasa A s inhibitorem IA3 ze Saccharomyces (PDB kód: 1dpj), B) vepřový pepsin s inhibitorem PI3 z Ascaris (PDB kód: 1f34). V horní části obrázku je celkový pohled na jednotlivé komplexy, kde jsou enzymy znázorněny jako šedý povrchový model, katalytické zbytky jsou červeně a inhibitory jako stužkový model (modře). V dolní části je detailní pohled do aktivního místa enzymu ve*

stužkovém modelu. Tyčkovým modelem jsou červeně zvýrazněny katalytické zbytky enzymů. IA3 interaguje pomocí Lys18 s katalytickým centrem proteinasy A; PI3 nemá přímý kontakt s katalytickými zbytky pepsinu, ale interaguje s vazebnými podmísty S_1'-S_3'.

1.5.2.2. Proteinové inhibitory serinových proteas

Proteinové inhibitory serinových proteas tvoří největší a nejrozmanitější skupinu proteasových inhibitorů, rozdíly najdeme v molekulové hmotnosti, struktuře a také mechanismu interakce s cílovým enzymem. Do této skupiny patří inhibitory, které (1) interaguji kovalentní vazbou s proteasami (rodina serpinu), (2) tvoří stabilní protein-proteinové komplexy na základě slabých interakcí (např. Bowman-Birkova, Kazalova nebo Kunitzova rodina) nebo (3) mají zvláštní mechanismus vazby α_2 makroglobulinů, které uvězní aktivní proteasu ve struktuře inhibitoru a zabraňují přístupu substrátu k enzymu. Jedná se o biochemicky a strukturně nejlépe charakterizované rodiny proteinových inhibitorů.

Inhibitor BPTI ze slinivky, známý také pod jménem aprotinin je typický zástupce rodiny inhibitoru I2 a jeden z nejlépe popsaných globulárních proteinů. BPTI má relativně širokou inhibiční specifitu, kdy je schopen blokovat zástupce serinových proteas s trypsinovou, chymotrypsinovou i elastasovou substrátovou specifitou. Se serinovými proteasami tvoří stabilní nekovalentní protein-proteinové komplexy. Jedná se o jednořetězcovou molekulu s M_r okolo 6,5 kDa s konzervovanou architekturou, kde reaktivní centrum je na smyčce exponované do roztoku a složené z Arg/Lys v P_1 a Ala/Gly v P_1'pozici (Obr. 7 A). Pomocí BPTI a jeho mutantních forem byl detailně studován kanonický (Laskowského) mechanismus interakce reaktivního centra tvořeného jednou smyčkou a vztah jeho struktury a funkčních vlastností (Helland et al. 1999; Krowarsch et al. 2005).



Obrázek 7: Prostorové struktury serinových proteas v komplexu s proteinovými inhibitory. *A) Hovězí trypsin s inhibitorem BPTI (z I2 rodiny, PDB kód: 3btk), B) hovězí trypsin s α1-antitrypsinovým inhibitorem (z I4 rodiny serpinů, PDB kód: 1ezx). V horní části obrázku je celkový pohled na jednotlivé komplexy, kde jsou enzymy znázorněny jako šedý povrchový model, katalytické zbytky jsou červeně a inhibitory jako stužkový model (modře). V dolní části je detailní pohled do*

aktivního místa enzymu ve stužkovém modelu. Tyčkovým modelem jsou červeně výrazněny katalytické zbytky enzymu a fialově Asp189 v S_1 podmístě. BPTI interaguje pomocí Arg13 s S_1 podmístem trypsinu; zbytek Met382 serpinu tvoří kovalentní vazbu s katalytickým Ser195 trypsinu.

Velkou rodinu inhibitorů tvoří serpiny (dle MEROPS databáze I4 rodina), které jsou schopné blokovat aktivitu především serinových proteas, ale i cysteinových proteas (např. kaspas nebo cysteinových katepsinů) (Law et al. 2006). Serpiny s M_r kolem 40-50 kDa mají dlouhou reaktivní smyčku, která je konformačně flexibilní. Tato smyčka je proteolyticky štěpena serinovou proteasou za vzniku kovalentního komplexu, kdy metioninový zbytek v P_1 pozici reaktivní smyčky serpinu se kovalentně váže na katalytický serinový zbytek proteasy. Tvoří se acyl-enzymový intermediát, který je vysoce stabilní, protože inhibitor zabraňuje přístupu vody do aktivního místa proteasy (Huntington, Read, a Carrell 2000) (Obr. 7 B, str. 20).

α_2 Makroglobuliny (dle MEROPS databáze rodina I39) patří mezi živočišné proteinové inhibitory endopeptidas s unikátním mechanismem interakce, který je popisován jako „mucholapka“ nebo „pastička na myši“. Při proteolytickém štěpení inhibitoru proteasou dochází ke konformačnímu přeskupení ve struktuře α_2 makroglobulinu a zachycení proteasy, která zůstává aktivní, ale má zamezený přístup substrátu do aktivního místa.

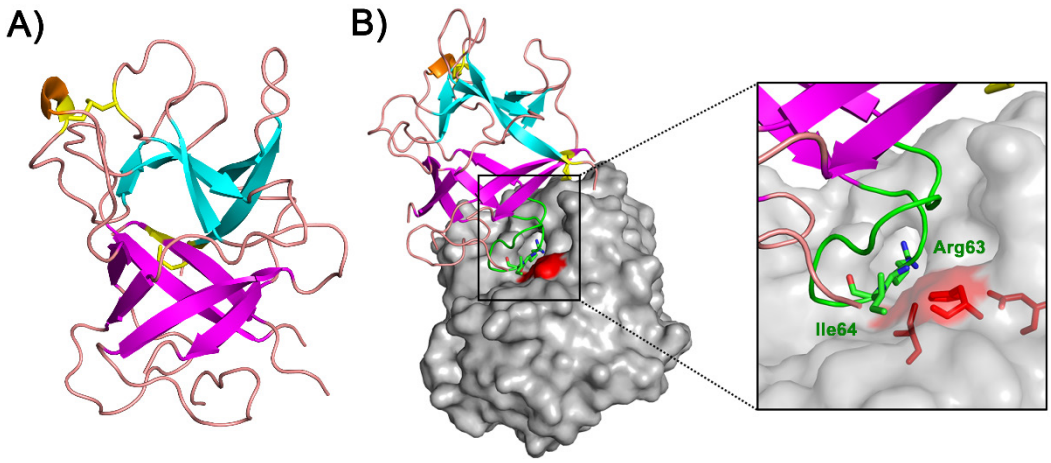
1.5.2.3. Rostlinné proteasové inhibitory z Kunitzovy rodiny

Proteasové inhibitory z Kunitzovy rodiny (dle MEROPS databáze rodina I3) jsou široce rozšířeny v rostlinné říši. Tyto proteiny s M_r okolo 20 kDa se vyskytují ve všech rostlinných pletivech, kde mohou mít odlišné funkce. Slouží např. jako obranné molekuly proti herbivorům v listech a v zásobních pletivech, dále jako regulátory endogenních proteas např. při klíčení semen a také se předpokládá jejich funkce jako zásobních proteinů. Inhibitory z Kunitzovy rodiny mají zcela unikátní rozsah inhibiční selektivity, protože jsou schopné blokovat aktivitu serinových, cysteinových i aspartátových proteas (Bendre, Ramasamy, a Suresh 2018). U některých zástupců se vyvinula multifunkčnost, kdy na jedné molekule inhibitoru se objevují reaktivní centra proti proteasám ze dvou tříd. Například inhibitory PCDI (z angl. Potato Cathepsin D Inhibitor) a NID (z angl. Novel Inhibitor of cathepsin D) z lilku brambor jsou schopné inhibovat serinové a aspartátové proteasy (Mareš et al. 1989; Ritonja et al. 1990); pro inhibitor BbCI z bauhinie byla popsána inhibice cysteinových proteas katepsinu L a kruzipainu a serinové proteasy elastasy (Hansen et al. 2007). Pro některé zástupce byla také identifikována aktivita proti medicinálně významným proteasám rodiny chymotrypsinu jako jsou např. proteasy krevní koagulace a fibrinolýzy nebo kalikreiny (D Zhou et al. 2015; Patil et al. 2012; Ferreira et al. 2013).

Proteiny z Kunitzovy rodiny fungují i mimo oblast proteolýzy. Příkladem je inhibice enzymu α amylasy endogenním inhibitorem BASI z ječmene setého (Vallée et al. 1998); protein mirakulin z plodů keře *Synsepalum dulcificum* moduluje receptory sladké chuti (Koizumi et al. 2011) a proteiny

WSCP (z angl. Water-Soluble Chlorophyll binding Protein) jsou v tetramerním uspořádání schopny vázat čtyři molekuly chlorofylu a chránit je tak před fotooxidací (Horigome et al. 2007).

Architektura proteinů z Kunitzovy rodiny (Obr. 8 A) označovaná jako β trojlístek (angl. β trefoil) je tvořena dvanácti segmenty β struktury, které jsou propojeny smyčkami (Renko, Sabotic, a Turk 2012). Struktura β segmentů je konzervovaná a naopak smyčky, které jsou exponované do roztoku, jsou variabilní v sekvenci i délce. Prostorové uspořádání molekuly je stabilizováno až třemi disulfidovými můstky, u některých zástupců však tyto vazby zcela chybí (Oliva et al. 2010). Mechanismus interakce Kunitzových inhibitorů se serinovými proteasami rodiny chymotrypsinu je ukázán na obrázku 8 B (str. 22) na komplexu inhibitoru STI s trypsinem (Song a Suh 1998). Inhibitor využívá kanonický Laskowského mechanismus interakce (Bode a Huber 1992; Krowarsch et al. 2003), kdy se reaktivní smyčka váže do aktivního místa enzymu a umisťuje klíčový bazický aminokyselinový zbytek do S_1 vazebného podmísta trypsinu, které je zodpovědné za substrátovou specifitu. Přestože je reaktivní smyčka v těsné blízkosti katalytického místa proteasy (Obr. 8 B), není efektivně štěpena díky celkové stabilizaci pomocí sítě protein-proteinových interakcí (Renko, Sabotic, a Turk 2012). Pouze za určitých podmínek, jako jsou extrémní hodnoty pH, dochází ke štěpení reaktivní smyčky, proto je tento mechanismus interakce někdy nazýván jako pomalý substrát (Krowarsch et al. 2003). Vysoká variabilita smyček Kunitzových proteinů umožnila v evoluci vývoj bifunkčních inhibitorů, které nesou dvě jednosmyčková reaktivní centra, jako např. u inhibitoru API-A (z angl. Arrowhead protease inhibitors), pro který byla vyřešena prostorová struktura ternárního komplexu (Bao et al. 2009). Kromě kanonických reaktivních center pro serinové proteasy chymotrypsinové rodiny byla u Kunitzových proteinů strukturně charakterizovaná reaktívní centra pro proteasu subtilisin,α amylasy a vazbu molekuly chlorofylu, která mají komplexnější stavbu (Vallée et al. 1998; Micheelsen et al. 2008; Horigome et al. 2007).



Obrázek 8: Prostorová struktura inhibitoru STI a jeho komplexu s trypsinem.

A) Struktura volného STI z rodiny Kunitzových inhibitorů (PDB kód: 1avw), který je znázorněn ve stužkovém modelu. Fialově a modře jsou znázorněny β struktury dvou domén, α šroubovice je oranžově, disulfidové můstky jsou žlutě. B) Struktura komplexu STI s hovězím trypsinem, který je

znázorněn jako šedý povrchový model, katalytické zbytky jsou zvýrazněny červeně. STI má stejné barevné kódování jako v A), reaktivní smyčka inhibitoru je zeleně. Detailní výřez ukazuje kanonické reaktivní centrum STI s kritickými zbytky Arg63 a Ile64 (zeleně).

V současné době jsou inhibitory z Kunitzovy rodiny intenzivně studovány pro jejich obranné funkce v rostlinách s možným využitím při přípravě transgenních rostlin, které by byly odolné proti hmyzím herbivorním škůdcům. Další směr výzkumu je zaměřen na enormní funkční variabilitu Kunitzových inhibitorů s cílem identifikovat nové typy reaktivních center a interakčních mechanismů, které jsou využitelné pro návrh inhibitorů medicinálně zajímavých proteas (Ferreira et al. 2013; Zani a Moreau 2010).

2. Cíle disertační práce

Disertační práce se zabývá aspartátovými a serinovými proteasami a regulací jejich aktivity pomocí přirozených inhibitorů. Zástupci těchto proteas hrají klíčovou roli v řadě fyziologicky významných procesů a jejich nedostatečná kontrola je spojena s mnoha patologickými stavy. Z tohoto důvodu jsou detailně studovány na biochemické a strukturní úrovni a představují cílové molekuly pro vývoj potenciálních chemoterapeutik. Jedním z možných přístupů při navrhování nových inhibitorů proteas je využití strukturních a funkčních vlastností přirozených inhibitorů, mezi které patří rostlinné proteiny z Kunitzovy rodiny, na něž se práce zaměřuje.

Práce má tyto dílčí cíle:

1) Pomocí moderních proteomických metod detailně popsat trávicí proteasy mandelinky bramborové, herbivorního škůdce lilku brambor. Jednotlivé proteasy klasifikovat a charakterizovat s využitím analýzy substrátové a inhibiční specifity a vizualizovat je pomocí aktivitních sond. Studovat interakci proteolytického systému tohoto škůdce s přirozenými inhibitory z Kunitzovy rodiny z lilku brambor.

2) S využitím proteinové krystalografie analyzovat interakci mezi inhibitory z Kunitzovy rodiny a jejich cílovými enzymy z třídy serinových a aspartátových proteas se zaměřením na identifikaci nových reaktivních center pro regulaci proteas. Dále pak vyřešit prostorovou strukturu trávicí aspartátové proteasy mandelinky bramborové a popsat mechanismus její inhibice Kunitzovým inhibitorem z lilku brambor.

3) Lidská aspartátová proteasa katepsin D je spojena s nádorovým onemocněním prsu. Cílem práce bylo identifikovat přirozené inhibiční mechanismy regulace aktivity tohoto enzymu se zaměřením na interakci se sfingolipidy. Dále pak připravit a testovat aktivitní sondy pro detekci katepsinu D jako prognostického nádorového markeru.

4) Lidské serinové proteasy kalikreinového typu jsou spojené s patologií řady tkání. Cílem práce bylo analyzovat substrátovou a inhibiční specifitu vybraných kalikreinů z epidermis. Dále pak určit inhibiční specifitu nového inhibitoru serinových proteas z Kunitzovy rodiny s kalikreinovými proteasami a vysvětlit jejich interakci na strukturní úrovni.

3. Materiál a metodika

Tato kapitola podává stručný přehled o metodice, vybavení a materiálu, které byly použity pro vypracování disertační práce, podrobné informace jsou uvedeny v jednotlivých publikacích. Detailně jsou popsány metody z dosud nepublikovaných výsledků, které jsou uvedeny v komentáři jedné kapitoly výsledků.

3.1. Materiál a laboratorní vybavení

Většina výsledků byla získána s využitím přístrojového vybavení laboratoří Ústavu organické chemie a biochemie AV ČR (ÚOCHB AV ČR). Sběr difrakčních dat pro rentgenostrukturní analýzu proteinových krystalů byl proveden na synchrotronu Bessy II electron storage ring (Helmholtz-Zentrum Berlin, Německo) a krystalové struktury byly řešeny ve spolupracující laboratoři Strukturní biologie na ÚOCHB AV ČR.

Substráty Abz-KPAEFF*AL (Máša et al. 2006), Abz-FRVF* a Abz-FRPF* (Jílková et al. 2011), Ac-WAVR-Acc a KHLY-Acc (Horn et al. 2018) pro měření proteolytické aktivity byly syntetizovány na ÚOCHB AV ČR ve skupině Medicinální chemie, ostatní substráty a inhibitory byly od firmy Bachem nebo Sigma-Aldrich. Enzymy trypsin a chymotrypsin byly od firmy Sigma-Aldrich, kalikreiny od firmy R&D Systems a lidský katepsin D byl izolován z placent (Fusek, Baudyš, a Metcalf 1992). Sady komerčních krystalizačních roztoků byly od firem Molecular Dimensions, Jena Bioscience a Hampton Research.

3.2. Metodika

3.2.1. Příprava rekombinantních proteas

IrCD1 byl připraven rekombinantní expresí v bakteriálním systému *Escherichia coli* a purifikován kombinací chromatografických kroků, jak bylo popsáno v (Sojka et al. 2012; Hánová et al. 2018). Pro rekombinantní expresi LdCD byl použit systém v kvasinkách *Pichia pastoris* (kmen X-33). Plasmid pGAPZα obsahující nukleotidovou sekvenci pro LdCD byl transformován do kvasinek elektroporací a pozitivní klony obsahující správný plasmid byly vybrány pomocí jejich rezistence na antibiotikum Zeocin. Tři vybrané kolonie byly zaočkovány do 10 ml YPD média a kultivovány přes noc, poté byly buňky převedeny do 0,5 L média a kultivovány po dobu dobu 1 týdne. LdCD byl konstitutivně produkován do média, ze kterého byl purifikován sadou chromatografických kroků s využitím FPLC systému zahrnující: (1) přípravu proteinové frakce na koloně s nosičem G-25, (2) afinitní chromatografii na koloně Ni-NTA a (3) gelovou chromatografii na koloně Superdex 75. Z 1 litru média bylo takto získáno 0,8 mg čistého aktivního enzymu, jehož identita byla potvrzena pomocí N koncového sekvenování a peptidového mapování hmotnostní spektrometrií. Proteasové inhibitory

z Kunitzovy rodiny byly purifikovány z proteinové frakce hlíz brambor pomocí iontoměničových kolon na systému FPLC dle postupu popsaného v článku (Mareš et al. 1989).

3.2.2. Biochemické metody

Mezi biochemické metody, které byly využity v této disertační práci patří: (1) separace proteinů na SDS-PAGE elektroforéze, (2) přenos proteinů na PVDF membránu metodou Western blot, (3) značení proteas v homogenátu s využitím aktivitních sond s fluorescenční nebo biotinovou značkou, označené proteasy byly separovány na SDS PAGE a vizualizovány pomocí fluorescenčního skeneru Typhoon nebo po převedení proteinů na PVDF membránu detekovány pomocí chemiluminiscence, (4) určení sekvence proteinu pomocí peptidového mapování hmotnostní spektrometrií a N koncového sekvenování z PVDF membrány, (5) určení koncentrace proteinů, peptidových substrátů a inhibitorů pomocí aminokyselinové analýzy, (6) purifikace proteas a proteasových inhibitorů na kolonách pro ionexovou, gelovou a afinitní chromatografii s využitím systému FPLC a (7) separace peptidů na C18 koloně s využitím HPLC systému.

3.2.3. Enzymologické metody

Měření enzymových aktivit pomocí fluorogenních, FRET (Fluroescence Resonance Energy Transfer) a chromogenních peptidových substrátů bylo provedeno na fluorescenční čtečce (Tecan), ze získaných dat byly stanoveny kinetické parametry proteas pro substráty (K_M, k_{cat}) a inhibiční konstanty pro inhibitory (IC_{50}, K_i) nelineární regresí v programu Grafit. Disociační konstanta K_D pro interakci LdCD s PCDI byly změřeny pomocí termoforetického efektu na přístroji Monolith NT.115 (NanoTemper).

3.2.4. Krystalografické metody

Pro krystalizační experimenty bylo připraveno celkem pět různých komplexů proteasa-inhibitor (Tab. 1, str. 27) inkubací dané proteasy s 1,5 molárním nadbytkem příslušného inhibitoru. Po 30 minutové inkubaci při laboratorní teplotě byla tato směs přečištěna pomocí gelové chromatografie na koloně Superdex 75. Absorpční vrchol obsahující komplex inhibitoru s proteasou byl zahuštěn pomocí koncentrátoru Amnicon (MWCO 10 kDa) na 2-7 mg/ml. Primární krystalizační podmínky pro jednotlivé komplexy byly identifikovány pomocí sady komerčních krystalizačních roztoků v 96 jamkové destičce metodou sedící kapky. Úspěšné krystalizační podmínky byly dále optimalizovány ve visící kapce k získání kvalitnějších a větších krystalů pro rentgenostrukturní analýzu. Přesné složení krystalizačních roztoků a koncentrace proteinových komplexů jsou uvedeny v tabulce 1 na straně 27. Difrakční data proteinových krystalů byla sebrána na synchrotronu BESSY II v Berlíně. Výsledné vyřešené prostorové struktury komplexů byly graficky upraveny a interpretovány s využitím programu PyMol.

Tabulka 1: Přehled složení krystalizačních roztoků pro jednotlivé komplexy proteasa-inhibitor. Uvedena je koncentrace proteinů, složení krystalizačních roztoků a roztoků pro mražení proteinových krystalů, které obsahovaly původní precipitační roztok a kryoprotektant.

Komplex proteasa-inhibitor	Koncentrace proteinu	Složení precipitačního roztoku	Kryoprotektant
LdCD-pepstatin	2,0 mg/mL	0.2 M NH_4-sulfát, 0.1 M Bis-Tris (pH 5,5), 25% PEG 3350	25% glycerol
IrCD1-PCDI	7,0 mg/mL	0,08 M Na-acetát (pH 5,0), 12% PEG 6000	30% PEG 400
trypsin-PCDI	5,0 mg/mL	0.1 M Tris-HCl (pH 8,0), 15% PEG 6000	25% glycerol
trypsin-PSPI	5,2 mg/mL	10% PEG 8000, 20% ethylenglykol 0.1 M Bicine-Tris (pH 8,5) 0.3 M Na-nitrát, 0.3 M Na_2H-fosfát, 0.3 M NH_4-sulfát	Nebylo použito
smíšený komplex trypsin/chymotrypsin-PSPI	5,5 mg/mL	12% PEG 8000, 24% ethylenglykol 0.1 M Bicine-Tris (pH 8,5) 0.6 M Na-nitrát, 0.6 M Na_2H-fosfát, 0.6 M NH_4-sulfát	Nebylo použito

4. Výsledky

Tato kapitola obsahuje jednotlivé publikace autora disertační práce a stručně se v ní diskutují hlavní výsledky společně s tím, jak se autor této disertační práce na rukopisu podílel.

Výsledky disertační práce jsou shrnuty v celkem pěti publikovaných článcích (kapitolách 4.1, 4.3-4.5 a 4.7) a v jednom rukopisu (kapitola 4.6), který je v recenzním řízení v oborovém impaktovaném časopise. Komentář první publikace je rozšířen o nepublikované výsledky, na jejichž základě jsou v současné době připravovány dva další rukopisy. Disertační práce se týká problematiky biochemie aspartátových a serinových proteas a jejich regulace pomocí proteasových inhibitorů. Předložené články a dílčí výsledky práce zahrnují (1) detailní popis systému trávicích proteolytických enzymů herbivorního škůdce mandelinky bramborové pomocí moderních proteomických nástrojů, (2) analýzu regulace aktivity lidského katepsinu D pomocí bioaktivních sfingolipidů, (3) vizualizaci lidského katepsinu D v extraktu buněk nádorového onemocnění prsu pomocí nově navržených a připravených aktivitních sond, (4) analýzu substrátové a inhibiční specifity hlavních kalikreinových proteas pokožky a jejich detekci v homogenátech epidermu geneticky modifikovaných myší, (5) určení inhibiční specifity kalikreinového inhibitoru z bauhínie (patřícího do Kunitzovy rodiny rostlinných inhibitorů) a její strukturní vysvětlení pomocí krystalografických dat tohoto inhibitoru s trypsinem a (6) přehledný referát o rostlinných proteasových inhibitorech z Kunitzovy rodiny se zaměřením na jejich strukturní a funkční diverzitu.

Nepublikovaná data zahrnují identifikaci a strukturní popis nových reaktivních center dvou proteasových inhibitorů Kunitzovy rodiny z brambor založený na krystalografické analýze prostorových struktur čtyř inhibičních komplexů s modelovými serinovými a aspartátovými proteasami. Dále je v této sekci uveřejněna krystalová struktura trávicí aspartátové proteasy mandelinky bramborové a navržen model přirozené regulace této proteasy pomocí interakce s Kunitzovým inhibitorem z brambor.

Seznam publikací

Publikace č. 1: Digestive proteolysis in the Colorado potato beetle, *Leptinotarsa decemlineata*: Activity-based profiling and imaging of a multipeptidase network.

Srp J, Nussbaumerová M, Horn M, Mareš M.

Insect Biochem Mol Biol, (78), 1-11 (2016)

Publikace č. 2: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids.

Žebrakovská I, Máša M, **Srp J**, Horn M, Vávrová K, Mareš M.

Biochim Biophys Acta, (12), 1097-1104 (2011)

Publikace č. 3: Single- and double-headed chemical probes for detection of active cathepsin D in a cancer cell proteome.

Nussbaumerová M, **Srp J**, Máša M, Hradilek M, Šanda M, Reiniš M, Horn M, Mareš M.

Chembiochem, (11), 1538-1541 (2010)

Publikace č. 4: Profiling system for skin kallikrein proteolysis applied in gene deficient mouse models

Horn M, Zbodáková O, Kašpárek P, **Srp J**, Hanečková R, Hradilek M, Mareš M, Sedláček R

Biol Chem, (9), 1085-1089 (2018)

Publikace č. 5: Crystal structures of the complex of kallikrein inhibitor BbKI with trypsin and modeling of kallikrein complexes

Li M, **Srp J**, Gustchina A, Dauter Z, Mareš M, Wlodawer A

Rukopis je v recenzním řízení v časopise *Acta Crystallographica Section D*

Publikace č. 6: Kunitzovy proteasové inhibitory z rostlin: strukturní a funkční diverzita

Srp J, Mareš M

Chemické listy, (110), 761-768 (2016)

4.1. Publikace č. 1: Digestive proteolysis in the Colorado potato beetle, *Leptinotarsa decemlineata*: Activity-based profiling and imaging of a multipeptidase network.

Souhrn

Mandelinka bramborová (*Leptinotarsa decemlineata*) je jeden z nejvýznamnějších herbivorních škůdců lilkovitých rostlin (*Solanaceae*), mezi které patří hospodářské plodiny (lilek brambor a rajče jedlé). Tento brouk pocházející původně ze střední Ameriky je v současné době rozšířen téměř po celém světě a je odolný na většinu hlavních insekticidů. Trávicí enzymy mandelinky byly zkoumány již dříve, ale až tento projekt přinesl detailní a systematický popis trávicího proteolytického aparátu s využitím moderní funkční proteomiky.

4.1.1. Proteolytická aktivita ve střevě mandelinky bramborové

Pro analýzu proteolytické aktivity ve střevech dospělců a larev mandelinky bramborové byly použity modelové proteinové substráty a bylo prokázáno, že proteolytický systém je schopen efektivně degradovat proteiny v širokém rozsahu pH od kyselého do mírně bazického. Tyto výsledky korelují s publikovanými hodnotami pH, které byly naměřeny ve střevě mandelinky bramborové v rozsahu pH 5,5-7,0 (Krishnan et al. 2007). Pomocí selektivních inhibitorů bylo prokázáno, že mezi hlavní proteolytické enzymy mandelinky patří cysteinové a aspartátové proteasy. V neutrálním pH se projevuje aktivita především cysteinových proteas, zatímco v mírně kyselém pH mají na celkové proteolýze kromě cysteinových proteas významný podíl aspartátové proteasy. Tato distribuce aktivity proteas je stejná u dospělců i larev. Zástupci cysteinových a aspartátových proteas jsou schopné zahájit štěpení proteinových substrátů a degradovat je až na malé peptidové fragmenty.

V dalším kroku byly využity specifické fluorogenní a chromogenní peptidové substráty společně se selektivními proteasovými inhibitory k detailnímu studiu jednotlivých proteolytických aktivit ve střevě mandelinky bramborové. Tyto aktivity byly dále rozděleny na sekretované, tedy přítomné v obsahu střev, a nesekretované. Mezi hlavní identifikované sekretované proteolytické enzymy patří aspartátová proteasa typu katepsinu D, cysteinové proteasy rodiny papainu a serinové proteasy - trypsin a karboxypeptidasy (serinová karboxypeptidáza a prolylkarboxypeptidasa). Jako nesekretované byly nalezeny aktivity aminopeptidasy patřící mezi metaloproteasy a prolyloligopeptidasy z třídy serinových proteas.

4.1.2. Analýza trávicích aspartátových proteas

Specifický peptidový FRET substrát (Abz-KPAEFF*AL) a selektivní inhibitor pepstatin byly použity k identifikaci sekretované aspartátové proteasy s pH optimem v kyselé oblasti (pH 4,5). Pomocí fluorescenční sondy FAP09, která je popsána v publikaci č. 3, byla tato proteasa specificky

označena v extraktu střev mandelinky bramborové a po separaci na SDS-PAGE vizualizována s M_r okolo 45 kDa. Analýza inhibiční specifity prokázala příbuznost s aspartátovou proteasou typu katepsinem D, a proto je dále používáno označení LdCD (*Leptinotarsa decemlineata* Cathepsin D). LdCD je účinně inhibován v nanomolárních koncentracích inhibitorem PCDI z lilku brambor, který patří do Kunitzovy rodiny. Tato interakce je fyziologicky významná a umožňuje klasifikovat PCDI jako rostlinný obranný protein, který cílí proti trávicímu systému hmyzího herbivora. V sekci nepublikovaných výsledků této disertační práce je uvedeno vyřešení prostorových krystalových struktur jak LdCD (kap. 4.2.2.), tak PCDI v komplexu s homologním katepsinem D (kap. 4.2.1.3) a je navržen model interakce obou proteinů.

4.1.3. Analýza trávicích cysteinových proteas

Cysteinové proteasy rodiny papainu (citlivé na inhibitor E64), které byly detekovány ve střevních extraktech mandelinky bramborové (tzv. intestainy), bylo možné s využitím specifických peptidových substrátů a selektivních inhibitorů klasifikovat jako proteasy, které vykazují (1) endopeptidasovou aktivitu katepsinu L, (2) endopeptidasovou a karboxydipeptidasovou aktivitu katepsinu B a (3) aminopeptidasovou aktivitu katepsinu H. Pro jednotlivé typy aktivit byly určeny hodnoty pH optima, které byly rozloženy v mírně kyselé až neutrální oblasti (pH 5,0-7,0).

Pro vizualizaci papainových proteas v extraktu střev byly využity tři fluorescenční sondy s rozdílnou specifitou, z toho dvě sondy (BoRC a BoFC) byly navrženy a připraveny v rámci této publikace. Tyto sondy umožnily identifikovat sadu minimálně osmi různých intestainů citlivých na inhibitor E64, s M_r v rozmezí 22-35 kDa. Tyto izoenzymy měly podobnou distribuci u dospělců a larev, kde jsou sekretovány do trávicího traktu. Kompetitivní značení fluorescenční sondou v přítomnosti sady selektivních inhibitorů cysteinových proteas dále ukázalo, že jednotlivé intestainy vykazují různou inhibiční specifitu. Trávicí trakt mandelinky bramborové tak obsahuje spektrum papainových proteas s výrazně odlišnými funkčními vlastnostmi a specifitou, což bylo potvrzeno i v aktivitním testu se sadou fluorogenních substrátů.

Analýza substrátové specifity prokázala, že některé z intestainů jsou schopné štěpit substrát s Asn zbytkem v P_1 pozici. Tento typ specifity je typický pro asparaginyl endopeptidasu (legumain), která patří mezi cysteinové proteasy rodiny C13. Asparaginyl endopeptidasy jsou časté proteasy trávicích systémů u bezobratlých, včetně členovců, ale překvapivě nejsou kódovány v genomech některých taxonů hmyzu jako je řád Coleoptera. Na základě této analýzy byla nyní formulována hypotéza, že cysteinové proteasy rodiny papainu mohou u hmyzu zastupovat funkční úlohu asparaginyl endopeptidas.

4.1.4. Analýza trávicích serinových proteas

Byly identifikovány dvě skupiny serinových proteas, které jsou sekretované ve střevě mandelinky bramborové: endopeptidasy z rodiny S1 a dva typy karboxypeptidas z rodin S10 a S28. V případě rodiny S1 se jedná o dva izoenzymy proteas se substrátovou specifitou trypsinového typu, pH optimem aktivity v alkalické oblasti a M_r 20 a 25 kDa. Tyto trypsinové proteasy vykazovaly odlišnou selektivitu k přirozeným inhibitorům z Kunitzovy rodiny: inhibitor PSPI (z angl. Potato Serine Protease Inhibitor) z lilku brambor byl účinným inhibitorem, na rozdíl od homologického inhibitoru STI ze sóji luštinaté. Tento výsledek naznačuje úlohu PSPI jako obranného proteinu proti mandelince bramborové. Interakce PSPI s trypsinovými proteasami byla charakterizována pomocí krystalové struktury komplexu, která je popsána v kapitole 4.2.1.1 a 4.1.1.2.

4.1.5. Závěr

Na základě detekovaných aktivit proteolytických enzymů bylo možné navrhnout schéma degradace proteinů v trávicím traktu mandelinky bramborové. V prvním kroku jsou proteiny v lumen střeva štěpeny na velké polypeptidové fragmenty působením endopeptidas ze tří proteasových tříd: (1) aspartátové proteasy typu katepsin D, (2) cysteinových proteas typu katepsinu L a B a (3) serinovou proteasou trypsinem. Vzniklé fragmenty jsou následně degradovány na vstřebatelné peptidy a aminokyseliny za účasti sekretovaných exopeptidas: (1) cysteinových proteas typu katepsinu B a H a (2) serinových proteas typu serinové karboxypeptidasy a prolylkarboxypeptidasy. Ve střevních buňkách je trávení dokončeno metaloaminopeptidasou a prolyloligopeptidasou.

Významným výsledkem byl přímý biochemický důkaz, že přirozené inhibitory z lilku brambor, které patří do Kunitzovy rodiny jsou schopné účinně inhibovat cílové trávicí enzymy mandelinky bramborové. To je umožňuje klasifikovat jako obranné proteiny produkované v rostlinných pletivech k potlačení trávicího systému herbivora. Na základě těchto výsledků byly provedeny detailní strukturní studie, které vysvětlují molekulární mechanismus interakcí mezi uvedenými proteasami a inhibitory, které jsou popsané v následující kapitole 4.2 a v současné době se připravují k publikování.

Můj podíl na práci zahrnoval (1) přípravu a charakterizaci tkáňových extraktů a analýzu jejich proteolytické aktivity s využitím enzymologických a biochemických přístupů, (2) identifikaci a klasifikaci proteas pomocí měření jejich enzymových aktivit a inhibice, (4) vizualizaci proteolytických enzymů v extraktech pomocí fluorescenčních sond, (5) bioinformatickou analýzu identifikovaných proteas v genomových a transkriptomových databázích, (6) přípravu obrázků a účast na psaní manuskriptu.

Publikace č. 1

Srp J., Nussbaumerová M., Horn M., Mareš M.

Digestive proteolysis in the Colorado potato beetle*, Leptinotarsa decemlineata*: Activity-based profiling and imaging of a multipeptidase network

Insect Biochemistry and Molecular Biology, (78), 1-11 (2016)

Insect Biochemistry and Molecular Biology 78 (2016) 1–11

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Digestive proteolysis in the Colorado potato beetle, *Leptinotarsa decemlineata*: Activity-based profiling and imaging of a multipeptidase network

Jaroslav Srp [a, b], Martina Nussbaumerová [a], Martin Horn [a], Michael Mareš [a, *]

[a] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, Prague 166 10, Czech Republic
[b] Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, Prague 128 40, Czech Republic

ARTICLE INFO

Article history:
Received 10 June 2016
Received in revised form
11 August 2016
Accepted 12 August 2016
Available online 15 August 2016

Keywords:
Colorado potato beetle
Peptidase
Cathepsin
Activity-based probe
Digestive system
Multienzyme network

ABSTRACT

The Colorado potato beetle (CPB), *Leptinotarsa decemlineata*, is a major pest of potato plants, and its digestive system is a promising target for development of pest control strategies. This work focuses on functional proteomic analysis of the digestive proteolytic enzymes expressed in the CPB gut. We identified a set of peptidases using imaging with specific activity-based probes and activity profiling with selective substrates and inhibitors. The secreted luminal peptidases were classified as: (i) endopeptidases of cathepsin D, cathepsin L, and trypsin types and (ii) exopeptidases with aminopeptidase (cathepsin H), carboxypeptidase (serine carboxypeptidase, prolyl carboxypeptidase), and carboxydipeptidase (cathepsin B) activities. The proteolytic arsenal also includes non-luminal peptidases with prolyl oligopeptidase and metalloaminopeptidase activities. Our results indicate that the CPB gut employs a multienzyme network of peptidases with complementary specificities to efficiently degrade ingested proteins. This proteolytic system functions in both CPB larvae and adults and is controlled mainly by cysteine and aspartic peptidases and supported by serine and metallopeptidases. The component enzymes identified here are potential targets for inhibitors with tailored specificities that could be engineered into potato plants to confer resistance to CPB.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The Colorado potato beetle (CPB), *Leptinotarsa decemlineata*, is the most devastating defoliator of potato in the northern hemisphere. Both larvae and adults feed on potato plants, causing large losses in crop yield. Several factors have contributed to the great success of CPB as a pest species, including adaptation to its potato host, high fecundity, a flexible life cycle with adult diapause, and the capability to evolve resistance to a wide variety of conventional insecticides. Therefore, there are pressing reasons to develop alternative strategies for CPB management. One promising approach for pest control is to disrupt protein digestion with dietary peptidase inhibitors, causing reduction in survival, growth, and fecundity of the insects. This has been illustrated in feeding trials with various species, including CPB (e.g., Bolter and LatoszekGreen, 1997; Wolfson and Murdock, 1987).

Plant peptidase inhibitors have been proposed as protein insecticides that could be engineered into crops to confer resistance to insect pests. However, the effectiveness of such transgenic plants is limited due to the adaptation response of insects, which counteract the plant defense proteins. CPB has exhibited several adaptation mechanisms to peptidase inhibitors, including compensation through upregulation of digestive peptidase genes, expression of inhibitor-insensitive peptidases, and proteolytic inactivation of peptidase inhibitors (Bolter and Jongsma, 1995; Cloutier et al., 2000; Gruden et al., 2003, 2004). To prevent CPB adaptation, current research concentrates on peptidase inhibitors of non-plant

Abbreviations: ABP, activity-based probe; Abz, 2-aminobenzoic acid; Ald, aldehyde; Amc, 7-amino-4-methylcoumarin; Bodipy, Bodipy TMR-X; BoFC, Bodipy-AAF-Cmk; BoRC, Bodipy-FPR-Cmk; BPTI, bovine pancreatic trypsin inhibitor; BSA, bovine serum albumin; Bz, benzoyl; Cbz, benzyloxycarbonyl; CPB, Colorado potato beetle; Cmk, chloromethyl ketone; dF, D-phenylalanine; DFP, diisopropyl fluorophosphate; Dmk, dimethylketone; DTT, dithiothreitol; E64, trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane; EDTA, ethylenediaminetetraacetic acid; F*, 4-nitrophenylalanine; Fmk, fluoromethyl ketone; FRET, fluorescence resonance energy transfer; nL, norleucine; PCDI, potato cathepsin D inhibitor; PCPI, potato carboxypeptidase inhibitor; PMSF, phenylmethanesulfonyl fluoride; PSPI, potato serine protease inhibitor; Suc, 3-carboxy-propionyl; STI, soybean trypsin inhibitor.

* Corresponding author.
E-mail address: mares@uochb.cas.cz (M. Mareš).

http://dx.doi.org/10.1016/j.ibmb.2016.08.004
0965-1748/© 2016 Elsevier Ltd. All rights reserved.

Abbreviations

ABP	activity-based probe
Abz	2-aminobenzoic acid
Ald	aldehyde
Amc	7-amino-4-methylcoumarin
Bodipy	Bodipy TMR-X
BoFC	Bodipy-AAF-Cmk
BoRC	Bodipy-FPR-Cmk
BPTI	bovine pancreatic trypsin inhibitor
BSA	bovine serum albumin
Bz	benzoyl
Cbz	benzyloxycarbonyl
CPB	Colorado potato beetle
Cmk	chloromethyl ketone
dF	D-phenylalanine
DFP	diisopropyl fluorophosphate
Dmk	dimethylketone
DTT	dithiothreitol
E64	trans-epoxysuccinyl-L-leucylamido(4-guanidino) butane
EDTA	ethylenediaminetetraacetic acid
F*	4-nitrophenylalanine
Fmk	fluoromethyl ketone
FRET	fluorescence resonance energy transfer
nL	norleucine
PCDI	potato cathepsin D inhibitor
PCPI	potato carboxypeptidase inhibitor
PMSF	phenylmethanesulfonyl fluoride
PSPI	potato serine protease inhibitor
Suc	3-carboxy-propionyl
STI	soybean trypsin inhibitor

origin and engineered inhibitors with improved potency, specificity, and stability (Brunelle et al., 2005; Goulet et al., 2008; Gruden et al., 1998). For selection and development of these inhibitors, more information is needed about the target digestive proteolytic system of CPB, its component enzymes, and their functional relationships.

Our understanding of the proteolytic arsenal in the CPB gut is fragmented. Previous studies focused on individual enzymes have identified a group of cysteine peptidases from the papain family called intestains and analyzed their interactions with proteinaceous inhibitors (Gruden et al., 2003, 2004; Petek et al., 2012; Sainsbury et al., 2012; Vorster et al., 2015). Little information is available about aspartic peptidases and serine peptidases (Brunelle et al., 1999; Petek et al., 2012). In the 1990s, researchers applied a systematic enzymological approach to obtain initial information about the digestive peptidases in the CPB gut (Brunelle et al., 1999; Novillo et al., 1997).

The present work provides novel biochemical insight into the molecular proteolytic machinery for plant protein degradation in the CPB gut by using a functional proteomic platform. This has become feasible only recently, with the help of chemical tools such as activity-based probes (ABPs) for selectively imaging target peptidases (for review, see e.g. Sanman and Bogyo, 2014) and new selective inhibitors and substrates. Here, we uncover how plant proteins are proteolytically digested in the CPB gut. We characterize a multienzyme network of peptidases that orchestrates luminal proteolysis, and we define the contributions of individual peptidases to this process. Based on these findings, we discuss the molecular evolution of digestive proteolysis in CPB and identify critical digestive peptidases as novel targets for pest control strategies.

2. Materials and methods

2.1. Biological materials and protein extracts

A laboratory colony of CPB was obtained from the Federal Biological Research Centre for Agriculture and Forestry, Braunschweig, Germany. Larvae and adults were reared on young potted potato plants (*Solanum tuberosum* cv. Magda) at 22 ± 2 °C, 60–70% relative humidity, and a 16/8 h (light/dark) photoperiod in a greenhouse chamber. The 4th instar larvae or adult beetle were removed from food for 2 h, immobilized on ice, and guts were dissected and stored at −80 °C. Guts of starved 4th instar larvae were prepared after 2 days of starvation. Regurgitant was collected from 4th instar larvae and stored at −80 °C. Protein extracts (5–10 mg protein/ml) from pooled guts were prepared by homogenization in 50 mM sodium acetate (pH 4.5) or 50 mM MES (pH 6.5) containing 1% CHAPS, followed by 30 min incubation on ice and sonication (Horn et al., 2009). The extracts were cleared by centrifugation (16,000g, 10 min, 4 °C), filtered with Ultrafree-MC 0.22 μm (Millipore), and stored at −80 °C. The protein extract from potato leaves was prepared from pooled leaves damaged by larvae feeding as previously described (Horn et al., 2005) using the gut extraction buffer (pH 4.5).

2.2. Proteomic tools: substrates, inhibitors, and activity-based probes

Fluorescence resonance energy transfer (FRET) substrates containing 2-aminobenzoic acid (Abz) and 4-nitrophenylalanine (F*) were synthesized as previously described (Jilkova et al., 2011; Masa et al., 2006). The fluorogenic substrates with a 7-amino-4-methylcoumarin (Amc) leaving group; the substrates Cbz-FL (Cbz, benzyloxycarbonyl) and Bz-GF (Bz, benzoyl) with a free C-terminal carboxyl group; and the inhibitors Cbz-PP-Ald (Ald, aldehyde), dFPR-Cmk (dF, D-phenylalanine, Cmk, chloromethyl ketone), and AAF-Cmk were purchased from Bachem. The inhibitors BPTI, STI, DFP, chymostatin, Pefabloc, PMSF, bestatin, 1,10-phenanthroline, and PCPI were from Sigma. The Kunitz-family inhibitors PCDI and PSPI were purified as previously described (Mares et al., 1989; Valueva et al., 2000). The inhibitor Cbz-AP-Cmk was synthesized as described (Fajtova et al., 2015), and AzaN5s (Gotz et al., 2008) was kindly donated by J. C. Powers. The activity-based probes BoRC and BoFC were prepared from the inhibitors dFPR-Cmk and AAF-Cmk, respectively, by modification of the N-terminal amino group with the NHS ester-activated fluorescent dye Bodipy TMR-X (Molecular Probes). Inhibitor (0.2 mg in 10 μl dry dimethylformamide) was mixed with reactive dye (0.2 mg in 10 μl dry dimethylformamide) and 1 μl N-ethyldiisopropylamine and incubated for 24 h. The probes were purified by RP-HPLC using a trifluoroacetic acid/acetonitrile system and characterized by mass spectrometry. The probe FAP09 was synthesized as described (Nussbaumerova et al., 2010). The probes Green-DCG-04 (Greenbaum et al., 2002) and Fhex-PD-AOMK (Sexton et al., 2007) were kindly donated by M. Bogyo.

2.3. Activity profiling of CPB peptidases

Proteolytic activities were identified and characterized by hydrolysis of the substrates listed in the Supplementary materials and methods; Abz- and Amc-containing substrates were applied at

25 μM concentrations and carboxypeptidase substrates Cbz-FL and Bz-GF at 1 mM concentrations. Measurements were performed at 26 °C in a total volume of 100 μl using an aliquot of the protein extract or regurgitant (0.05–15 μg protein) and the following 0.1 M assay buffers: sodium acetate (pH 4.0–5.0), sodium phosphate (pH 5.0–8.0), Tris (pH 7.0–9.0), and glycine-NaOH (pH 9.0–13.0), containing 2.5 mM DTT (for cysteine peptidases) and 25 mM NaCl (for cathepsin C). For activity assays in the presence of peptidase inhibitors, the extract was preincubated (15 min at 26 °C) in the assay buffer with the inhibitors listed in the Supplementary materials and methods. Proteolytic activity was continuously measured after addition of substrate in a Tecan Infinite M1000 fluorescence reader at 320 nm excitation and 420 nm emission wavelengths (for Abz-containing substrate) or at 360 nm excitation and 465 nm emission wavelengths (for Amc-containing substrates). Carboxypeptidase activities (Cbz-FL and Bz-GF substrates) were monitored after derivatization of the liberated product with fluorescamine at 375 nm excitation and 480 nm emission wavelengths (Sorgine et al., 2000). Assays of serine and aspartic peptidases were performed in the presence of 25 μM E64 to prevent possible interference of non-target peptidases. All measurements were performed in triplicate. The specific activity was expressed as pmoles of substrate hydrolyzed per second per μg of protein in the extract; the released fluorescent product was quantified by a calibration curve. Specific activities of protein extracts prepared at pH 4.5 and 6.5 were compared, and the more active one was used for further analysis.

2.4. Active-site labeling of CPB peptidases

For labeling of cysteine peptidases, the gut extract or regurgitant (10 μg protein) was incubated (1 h at 26 °C) with 1 μM activity-based probe (Green-DCG-04 (Greenbaum et al., 2002), BoFC, or BoRC) in 0.1 M sodium citrate (pH 5.5) containing 2.5 mM DTT. For labeling of trypsin peptidases, the gut extract (30 μg protein) was treated (15 min at 26 °C) with 50 μM E64 and incubated (1 h at 26 °C) with 1 μM BoRC in 0.1 M Tris (pH 9.0). For labeling of asparaginyl endopeptidases, the gut extract or leaf extract (30 μg protein) was treated with 25 μM E64 and 1 μM pepstatin (15 min at 26 °C) and incubated (1 h at 26 °C) with 1 μM Fhex-PD-AOMK in 0.1 M sodium acetate (pH 5.0) containing 2.5 mM DTT and 2 mM EDTA. For labeling of aspartic peptidases, the gut extract (50 μg protein) was treated with 25 μM E64 (15 min at 26 °C) and incubated (15 min at 26 °C) with 1 μM FAP09 in 0.1 M sodium acetate (pH 4.0). Afterward, the reaction mixture was irradiated for 20 min on ice with a 125 W high-pressure mercury-vapor lamp to allow for photoactivated cross-linking. The competitive labeling reactions were performed after treatment (15 min at 26 °C) of the extract or regurgitant with the indicated inhibitor. The labeled proteins were acetone precipitated and resolved by 15% Laemmli-SDS-PAGE. The labeled peptidases were visualized in-gel using a Typhoon 9410 Imager (GE Healthcare) as previously described (Horn et al., 2009; Nussbaumerova et al., 2010).

2.5. Digestion of protein substrates by CPB peptidases

For the azocasein assay, the gut extract or regurgitant (50 μg protein) was incubated for 2.5 h at 26 °C with 2% azocasein (Sigma) in 0.1 M Britton–Robinson buffer (pH 4.0–8.5) or 0.1 M sodium phosphate (pH 5.0–7.0) in a total volume of 100 μl. All buffers contained 1 M urea, 1 mM EDTA, 2.5 mM DTT, 50 mM NaCl, and 0.15% PEG 6000. The reaction was stopped by addition of 300 μl of 6% trichloroacetic acid and incubated for 30 min on ice. The precipitated substrate was removed by centrifugation (13,000*g*, 10 min, 4 °C), and the product was quantified by measuring the absorbance of the supernatant at 360 nm; one activity unit yields $A_{360} = 1$ in 1 h under assay conditions. For assay in the presence of inhibitors, the extract was treated (15 min, 26 °C) with the indicated inhibitors. All measurements were performed in triplicate.

For the bovine serum albumin (BSA) assay, the gut extract (0.5 μg protein) was incubated at 26 °C for 0.5–16 h with BSA (10 or 25 μg) in 0.1 M sodium citrate (pH 4.0–6.0) in the presence/

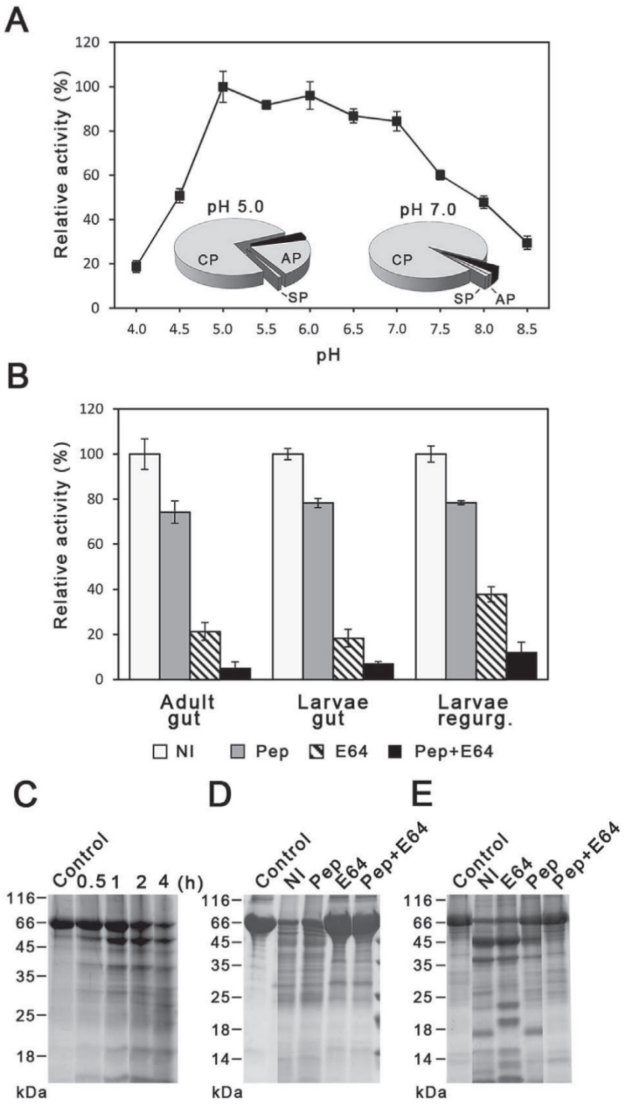


Fig. 1. Proteolytic potency and peptidase distribution in the CPB gut. (A) pH profile of proteolytic activity in the gut extract of CPB larvae. The chromogenic protein substrate azocasein was digested at various pH values. The proteolytic activities (mean values ± SE) are expressed relative to the maximum value (100%). Pie charts: The relative contributions of the activity of aspartic peptidases (AP), cysteine peptidases (CP), and serine peptidases (SP) are indicated for pH 5.0 and 7.0. The component activities were determined by sensitivity to selective peptidase inhibitors: pepstatin (10 μM) for AP, E64 (10 μM) for papain-family CP, and Pefabloc + BPTI (1 mM + 0.5 μM) for SP; the inhibitor insensitive activity is in black. (B) Inhibition of proteolytic activity in the gut extracts of CPB adults and larvae and the regurgitant of CPB larvae. Azocasein was digested at pH 6.0 in the presence of 10 μM pepstatin (Pep) and/or E64; NI indicates that no inhibitor was added. The inhibition of digestion is expressed as remaining proteolytic activity relative to the uninhibited control (100%). The mean values ± SE are given. (C–E) Protein cleavage pattern of peptidases from the CPB gut. The BSA substrate was digested by the gut extract of CPB larvae, and the reaction mixture was resolved by SDS-PAGE and visualized by protein staining. The time course of BSA degradation (C) was performed at pH 5.5. The contribution of peptidases to BSA degradation was analyzed at pH 6.5 (D) and 4.0 (E) in the presence of 10 μM pepstatin (Pep) and/or E64; NI indicates that no inhibitor was added. The controls contain non-digested BSA substrate: 10 μg (C, E) or 25 μg (D).

absence of 2.5 mM DTT in a final volume of 100 µl. For assay in the presence of inhibitors, the extract was treated (15 min at 26 °C) with the indicated inhibitors. The reaction was stopped by acetone precipitation. The reaction mixture was resolved by 15% Laemmli-SDS-PAGE, and gels were protein-stained with Coomassie Brilliant Blue.

3. Results

3.1. Protein degradation in the CPB gut is critically controlled by cysteine and aspartic peptidases

We investigated the proteolytic capability of the gut extract of CPB larvae *in vitro* using the non-specific protein substrates azocasein and BSA. Quantitative analysis of azocasein degradation showed that proteolysis is efficient at mildly acidic and neutral pH with a broad pH optimum ranging from approximately 5 to 7 (Fig. 1A). The fragmentation pattern of BSA was analyzed by SDS-PAGE. Time-course experiments demonstrated gradual hydrolysis of the protein substrate into initial large fragments of >35 kDa, and then into a wide spectrum of smaller fragments (Fig. 1C).

The contribution of individual peptidase activities to protein degradation was evaluated by analyzing the *in vitro* impact of selective peptidase inhibitors. An azocasein assay at pH 6 showed that individual application of E64 (which targets papain-family cysteine peptidases) and pepstatin (which targets aspartic peptidases) reduced azocasein digestion to roughly 20% and 80%, respectively (Fig. 1B). The relative contribution of aspartic peptidases *versus* papain-family cysteine peptidases increased at acidic pH (pH 5) and greatly decreased at neutral pH (Fig. 1A). The combination of both inhibitors led to suppression of degradation to less than 5% (Fig. 1B). In contrast, inhibitors of serine endopeptidases (Pefabloc, BPTI) had a marginal effect on degradation (Fig. 1A), and inhibition of metallopeptidases (EDTA) had no measurable effect (data not shown).

The BSA digest was performed at two different pH levels that favor different peptidase classes. Around neutral pH, BSA was efficiently digested into very small fragments (below the 10 kDa limit). Treatment with E64 resulted in near total inhibition of degradation, while pepstatin had little effect (Fig. 1D). On the other hand, at low pH, the substrate was converted into a rather limited set of distinct fragments. Treatment with pepstatin led to effective inhibition of degradation, while E64 exhibited a weak effect (Fig. 1E). These data demonstrated that aspartic peptidases acting mainly at low pH predominantly generated large- to medium-sized polypeptides, while cysteine peptidases favoring mildly acidic or neutral pH were capable of deeply degrading the protein substrate.

We compared the proteolytic activity of the gut extract of CPB larvae with that of the gut extract of CPB adults and regurgitant (oral secretions with gut contents) from CPB larvae; the specific activity to azocasein was 47, 59, and 57 U/mg protein, respectively. The inhibition pattern demonstrated the critical contribution of cysteine and aspartic peptidases to proteolysis in all samples tested (Fig. 1B). Furthermore, the results suggested that peptidases from these two classes are secreted into the gut lumen, where they act cooperatively in digestive proteolysis.

3.2. Functional profiling of peptidase activities in the CPB gut

We investigated the distribution of peptidases in the gut extract of CPB larvae by activity profiling with a grid of specific substrates and selective inhibitors used as diagnostic tools (Table 1). The major activities identified (higher than 10 pmol.s^{-1}.µg^{-1}) were analyzed in detail, including pH activity profiles and inhibitor specificity. Furthermore, these activities were screened in regurgitants from CPB larvae to determine their secretion status: those found in both the gut extract and regurgitant were classified as secreted and those found solely in the gut extract as non-secreted (Table 1).

3.3. Multiple cysteine and aspartic peptidases are secreted into the gut lumen

Activities of three types of papain-family cysteine peptidases

Table 1
Profiling of peptidase activities in the gut of CPB larvae. Peptidase activities in the gut extract of CPB larvae were screened against specific peptide substrates and selective inhibitors in a fluorescence assay (for details, see Materials and methods). All activities were tested at pH 4, 6, and 8. The pH values with substantial activities are listed; nsa indicates that no significant activity was detected (with up to 15 µg protein of the gut extract used in the assay). The specific activities were determined at the pH indicated in parentheses (for the secreted activities, pH optima were used). The effective inhibitors are listed. Peptidase activities were classified as secreted based on their presence in larvae regurgitant. To prevent the possible interference of non-target activities, the measurements were performed in the presence of 25 µM E64 (except for assay of cysteine peptidases); cathepsin H-like and asparaginyl endopeptidase-like intestain activities were measured in the presence of 100 µM bestatin and 1 µM AzaN5s, respectively. The specific activity was expressed as pmoles of substrate hydrolyzed per second per µg of protein of the extract.

Peptidase	Class	Substrate	Activity		Inhibitor
			pH	Specific activity (pmol.s^{-1}.µg^{-1})	
Secreted activity					
cathepsin D	aspartic	Abz-KPAEFF*AL	4.0	1337 (4.5)	pepstatin, PCDI
cathepsin L-like intestain	cysteine	Cbz-FR-Amc	4.0, 6.0, 8.0	82,097 (6.5)	E64, Cbz-FF-Dmk
cathepsin B-like intestain	cysteine	Cbz-RR-Amc	6.0, 8.0	17,563 (6.5)	E64, Cbz-FF-Dmk
		Abz-FRVF*	4.0, 6.0, 8.0	3767 (5.5)	E64, CA074
cathepsin H-like intestain	cysteine	R-Amc	6.0, 8.0	971 (7.0)	E64, L-Cmk
asparaginyl endopeptidase-like intestain	cysteine	Cbz-AAN-Amc	6.0	98 (6.0)	E64
trypsin	serine	Cbz-RR-Amc	6.0, 8.0	1790 (9.0)	Pefabloc, BPTI, PSPI
prolyl carboxypeptidase	serine	Abz-FRPF*	4.0, 6.0	1463 (5.0)	Cbz-PP-Ald
carboxypeptidase	serine	Cbz-FL	4.0, 6.0	230 (4.0)	DFP
Non-secreted activity					
aminopeptidase	metallo	L-Amc	6.0, 8.0	628 (8.0)	bestatin
prolyl oligopeptidase	serine	Cbz-GP-Amc	6.0, 8.0	258 (6.0)	Cbz-AP-Cmk, Cbz-PP-Ald
Weak/no activity					
carboxypeptidase	metallo	Bz-GF	6.0, 8.0	<10	1,10-Phenanthroline
chymotrypsin	serine	Suc-AAPF-Amc	8.0	<10	Chymostatin, PMSF
subtilisin	serine	Cbz-GGL-Amc	nsa		
elastase	serine	MeOSuc-AAPV-Amc	nsa		
cathepsin C	cysteine	GR-Amc	nsa		
caspase	cysteine	Ac-DEVD-Amc	nsa		

(sensitive to E64) called intestains were identified in the gut extract of CPB larvae (Table 1). We attributed the major activity that cleaves Cbz-FR-Amc (a substrate of cathepsins B and L) to cathepsin L-like intestains based on high sensitivity to Cbz-FF-Dmk, a preferential inhibitor of cathepsin L (Fig. 2A). Cbz-RR-Amc (a substrate of cathepsin B) was cleaved with ~5-fold lower efficiency than Cbz-FR-Amc (Table 1), and this activity had low sensitivity to the cathepsin B inhibitor CA074 (Fig. 2A). However, we found cathepsin B-like activity capable of hydrolyzing the carboxydipeptidase substrate Abz-FRVF* with high sensitivity to CA074 (Fig. 2B). This substrate is cleaved by cathepsins B, which can act as endopeptidases (cleaving Cbz-FR/RR-Amc) as well as exopeptidases, specifically carboxydipeptidases (Cezari et al., 2002; Jilkova et al., 2011). Further, intestains exhibited cathepsin H-like aminopeptidase activity, which was detected with the substrate R-Amc and was blocked by the inhibitor L-Cmk (Fig. 2B). Cathepsin D-like activity of aspartic peptidases was monitored with the substrate Abz-KPAEFF*AL, and it was fully inhibited by the class-selective inhibitor pepstatin as well as by PCDI (potato cathepsin D inhibitor), a specific inhibitor of cathepsin D-type peptidases (Fig. 2A). The gut-associated intestain and cathepsin D activities were also present in the regurgitant of CPB larvae, indicating that these enzymes are secreted into the gut lumen (Table 1).

The pH-activity profiles of the investigated peptidases show overlap in the region between pH 5 and 6 (Fig. 2C). The aspartic peptidases exhibited the greatest activity in the pH range from 3.5 to 5.5, while intestains had optimal endopeptidase and cathepsin H-like exopeptidase activities at pH levels ranging from 5.5 to 7.5. The cathepsin B-like exopeptidase activity had a broad pH optimum between pH 4 and 7, bridging the pH activity ranges of other intestains and aspartic peptidases.

The peptidases identified in the gut extract of CPB larvae by activity profiling were visualized using activity-based probes (ABPs, see Fig. 3). For imaging of aspartic peptidases (Fig. 3A), we used the photoreactive probe FAP09 (Nussbaumerova et al., 2010). This probe selectively labeled a species of ~45 kDa, and the finding was confirmed by quenching of the signal when the labeling was performed in competition with pepstatin, a specific active-site inhibitor of aspartic peptidases. Three ABPs were used for imaging of papain-family cysteine peptidases (Fig. 3B): Green-DCG-04 (Greenbaum et al., 2002) and two ABPs denoted BoFC and BoRC, which were designed for the purpose of this study (see Materials and methods and Fig. S2). With these probes, we visualized a set of at least five bands of intestains in the range from approximately 22 to 33 kDa, with a distinct pattern for each ABP. Labeling was quenched in the presence of E64, a specific active-site inhibitor of papain-family peptidases.

We analyzed the inhibitor specificity of intestains in detail using

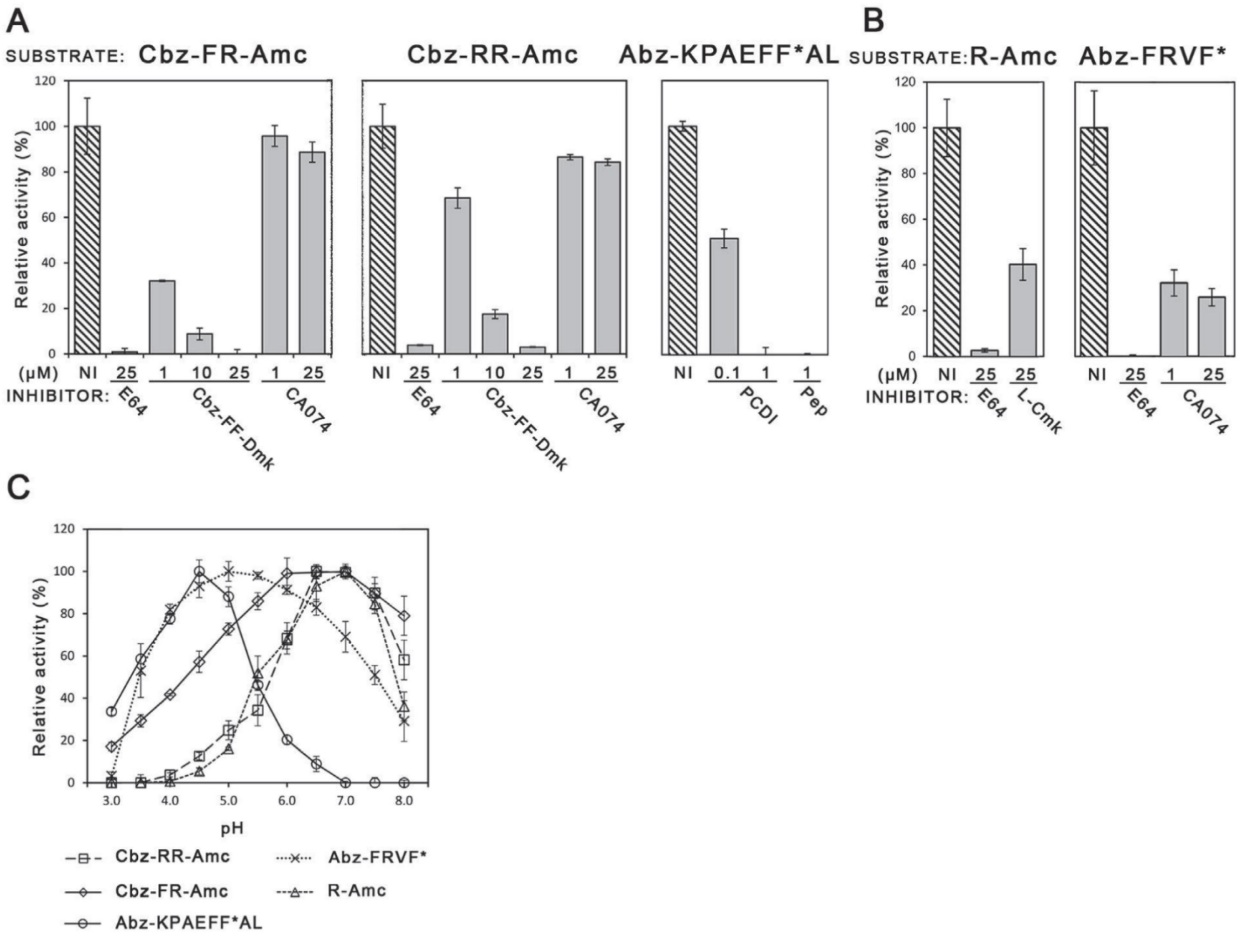


Fig. 2. Activity and inhibitor specificity of aspartic and papain-family cysteine peptidases in the CPB gut. (A–B) Peptidase activities in the gut extract of CPB larvae were measured with specific fluorogenic substrates in the presence of relevant selective inhibitors (concentrations are indicated). (A) Endopeptidase activities of papain-family cysteine peptidases attributed to cathepsin L-like and B-like intestains (Cbz-FR-Amc and Cbz-RR-Amc substrates measured at pH 6.5) and cathepsin D-like aspartic peptidases (Abz-KPAEFF*AL, pH 4.5). (B) Exopeptidase activities of papain-family cysteine peptidases attributed to cathepsin H-like (R-Amc, pH 7.0) and cathepsin B-like (Abz-FRVF*, pH 5.5) intestains. The inhibition is expressed as remaining proteolytic activity relative to the uninhibited control (NI, 100%); the mean values ± SE are given. (C) pH profiles of peptidase activities in the gut extract of CPB larvae were measured with specific fluorogenic substrates at various pH values using the same substrates and assayed peptidases as in panels (A) and (B). The activities (mean values ± SE) are expressed relative to the maximum values (100%). Cathepsin H-like intestain activity was measured in the presence of 100 μM bestatin to prevent interference from metalloaminopeptidases.

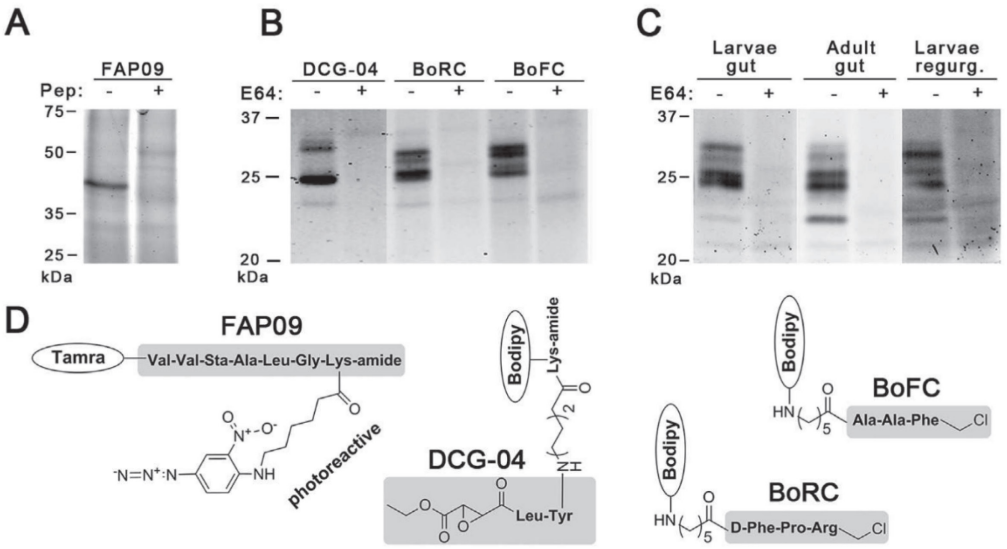


Fig. 3. Imaging of aspartic peptidases and papain-family cysteine peptidases with activity-based probes (ABPs). The gut extract of CPB larvae was treated with a panel of selective fluorescent ABPs that interact with the active site of the target peptidases. The reaction mixture was resolved by SDS-PAGE and the labeled peptidases were visualized using a fluorescence imager. The molecular mass standards are indicated in kDa. (A) Visualization of aspartic peptidases with the FAP09 ABP. The competitive labeling was performed in the presence of 10 μM pepstatin (Pep). (B) Visualization of papain-family peptidases (intestains) with Green-DCG-04, BoRC, and BoFC. The competitive labeling was performed in the presence of 10 μM E64. (C) Imaging of the gut extract of CPB larvae and adults and regurgitant of CPB larvae with BoRC. The competitive labeling was performed in the presence of 10 μM E64. (D) ABP structures: BoRC, BoFC, and Green-DCG-04 contain an irreversible active-site ligand that functions as a chemically reactive warhead (gray box); FAP09 contains a reversible active-site ligand (gray box), which is cross-linked to the protein target by a photoreactive group. Open ovals represent fluorescent tags serving as reporter groups.

BoRC, which labels the majority of isoenzyme species. Individual intestains exhibited different sensitivities to a panel of covalent peptidomimetic inhibitors of papain-family peptidases that were applied as quenching ligands (Fig. 4A). In general, endopeptidase inhibitors displayed a wide spectrum of binding potencies, including strong inhibition, while the efficiency of exopeptidase inhibitors (L-Cmk and CA074) was rather low. We also applied a highly effective inhibitor, Cbz-FF-Dmk, at various concentrations, which further demonstrated differences in the sensitivities of individual intestains to the inhibitors (Fig. 4B). Testing the cleavage specificity of intestains showed that peptide substrates with Arg/Lys residues in the P1 position were cleaved with different efficiencies, and Pro was a highly unfavorable residue in the P2 position (Fig. S1A). Importantly, the activities against individual substrates displayed different inhibitor specificity patterns (see Fig. 4).

These results indicate that the gut of CPB larvae contains a set of intestains that differ in molecular mass and in substrate and inhibitor specificities. The isoenzyme pattern was compared with that visualized for larvae regurgitant and the gut extract of CPB adults (Fig. 3C). This suggested that both larvae and adults produce very similar sets of intestain isoenzymes and that these peptidases are secreted into the gut lumen.

3.4. Dual origin of asparaginyl endopeptidase-like activity in the gut lumen

In the gut extract of CPB larvae, we measured the hydrolytic

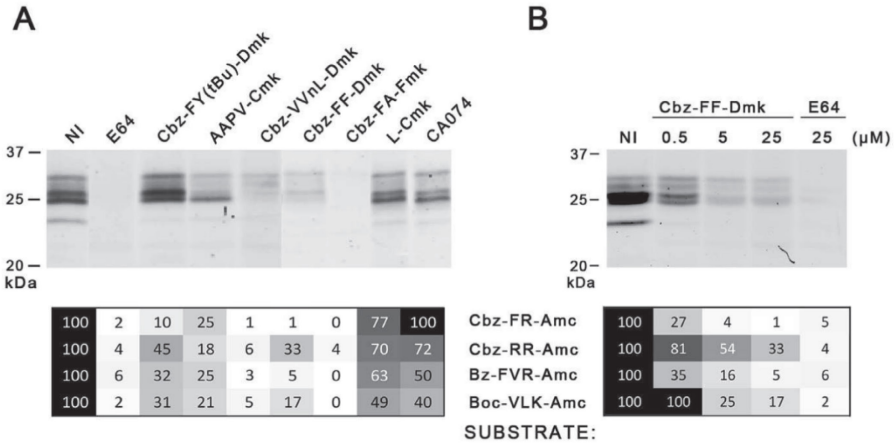


Fig. 4. Probing the inhibitor specificity of multiple papain-family cysteine peptidases in the CPB gut. The sensitivity of target peptidases in the gut extract of CPB larvae to inhibitors was analyzed using (A) a set of relevant inhibitors (applied at 10 μM concentration) and (B) various concentrations of the representative inhibitor Cbz-FF-Dmk. E64, a general inhibitor of papain-family cysteine peptidases, is shown as a positive control. Upper panels: The target peptidases were labeled with BoRC in the presence of the indicated competitive inhibitors. The reaction mixture was resolved by SDS-PAGE and visualized using a fluorescence imager. The molecular mass standards are shown in kDa. NI indicates that no inhibitor was added to the labeling reaction. Lower panels: Peptidase activities were measured at pH 6.5 with various fluorogenic substrates in the presence of inhibitors as in the upper panels. Activities are expressed as remaining proteolytic activity relative to the uninhibited control (NI, 100%) and displayed as a heat map (grayscale).

activity against Cbz-AAN-Amc (Table 1), a selective substrate used for detection of asparaginyl endopeptidases (legumains) of the CD clan of cysteine peptidases (Kembhavi et al., 1993). However, legumains should not be naturally present in the CPB, as they are not encoded in Coleoptera genomes (see Discussion). This activity was inhibited in part by AzaN5s, which specifically targets legumains, and in part by E64, which targets papain-family cysteine peptidases of the CA clan including intestains (Fig. 5A, control larvae). We found that the true legumain activity in the CPB gut extract is derived from ingested potato leaves that contain substantial legumain (but not intestain-like) activity. This was demonstrated by comparing legumain activity in both tissues (Fig. 5A). We found the same pH activity profile with a pH optimum of 6 (Fig. 5B) and the same imaging pattern; the legumain-specific probe Fhex-PD-AOMK labeled a species of ~37 kDa, and the signal was quenched by AzaN5s (Fig. 5C). Our conclusion was further supported by testing hydrolysis of the Cbz-AAN-Amc substrate in the gut extract of starved CPB larvae, where legumain activity was greatly reduced (Fig. 5A). To summarize, the gut of CPB larvae produces intestains with a cleavage preference for Asn in the P1 position, which mimics legumain substrate specificity.

3.5. Secreted peptidases of the serine class: trypsin and carboxypeptidases

Three types of serine peptidases were identified in the gut extract of CPB larvae, including trypsin-type peptidases (S1 family), serine carboxypeptidases (S10 family), and prolyl carboxypeptidases (S28 family). These activities were also detected in larvae regurgitant, indicating their secretion into the gut lumen (Table 1).

Trypsin-like activity was screened against a panel of fluorogenic substrates in the presence of E64 to prevent interference from intestains. Substrates with Arg/Lys residues in the P1 position were effectively cleaved; the highest efficiency was measured for extended substrates spanning the P4 to P1 positions (Fig. 6C). Pro was a favorable residue in the P2 position, in contrast to our findings for intestains (Fig. S1). High activity was exhibited in the pH range from 6 to 12, with a maximum at pH 9 (Fig. 6B). The activity was not inducible by calcium ions and was fully sensitive to general inhibitors of serine peptidases, such as Pefabloc and BPTI (Fig. 6A). Interestingly, CPB trypsins were able to discriminate between two plant Kunitz-type inhibitors of mammalian trypsins and chymotrypsins: pronounced inhibition was observed in the presence of PSPI from potato, while STI from soybean was ineffective. We visualized the gut extract of CPB larvae using the ABP BoRC (Fig. 6D), which is selective for trypsin-type peptidases under the conditions used and in the presence of E64 (Fig. S2). The labeling was authenticated by quenching the signal in a competition experiment with Pefabloc. The probe bound to two species of 20 and 24 kDa, indicating that different trypsin isoenzymes are produced in the CPB gut.

Serine carboxypeptidases from the S10 family remove C-terminal residues, preferentially hydrophobic amino acids. They were identified in the CPB gut extract based on hydrolysis of the

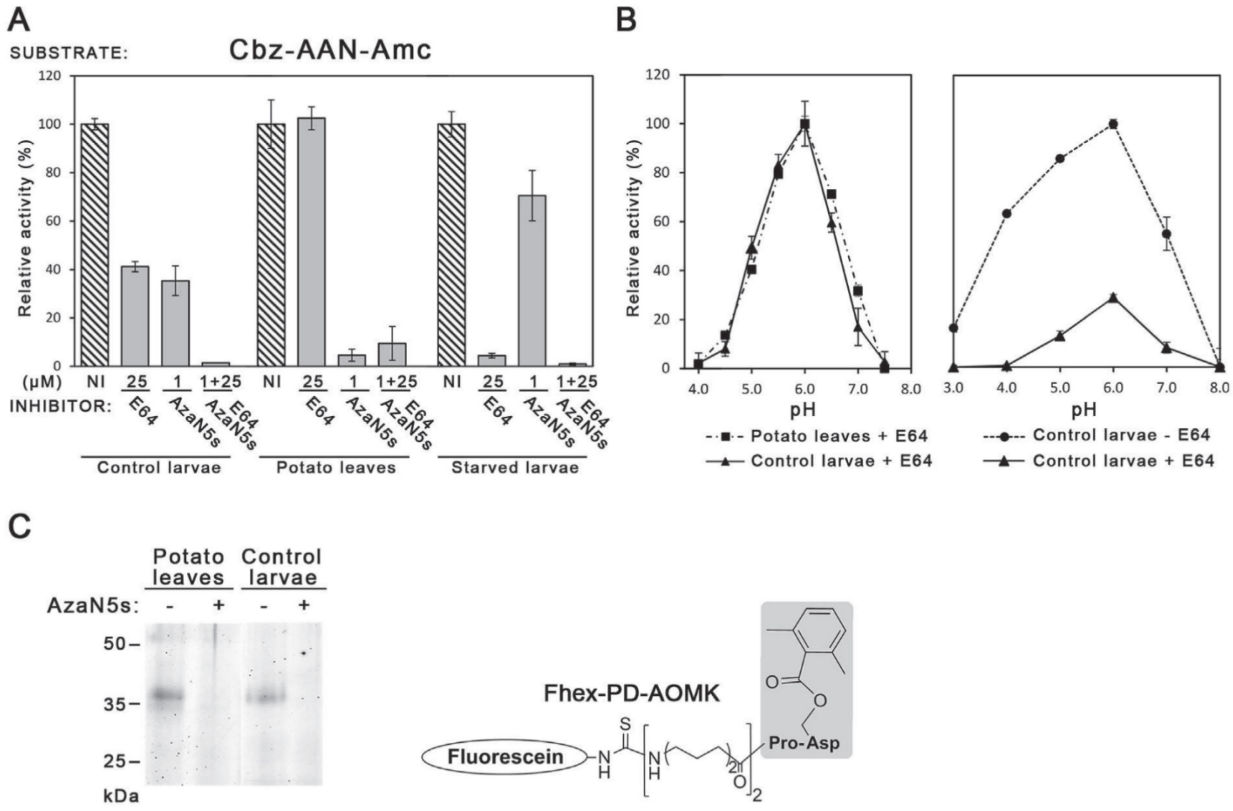


Fig. 5. Characterization of asparaginyl endopeptidase-like activity in the CPB gut. (A) Inhibitor sensitivity of asparaginyl endopeptidase-like activity in the gut extracts of control and starved CPB larvae and in potato leaf extract. The activity was measured with the fluorogenic substrate Cbz-AAN-Amc at pH 6.0 in the presence of inhibitors E64 and AzaN5s (concentrations are indicated), which are selective for papain-family and legumain-family cysteine peptidases, respectively. The inhibition is expressed as remaining proteolytic activity relative to the uninhibited control (NI, 100%); the mean values ± SE are given. (B) pH profiles of legumain activity in the gut extract of control larvae and potato leaf extract were measured with Cbz-AAN-Amc at various pH values. Treatment with E64 eliminates the effect of papain-family peptidases. The activities (mean values ± SE) are expressed relative to the maximum values (100%). (C) Imaging of legumains in the gut extract of control larvae and potato leaf extract with the selective fluorescent activity-based probe Fhex-PD-AOMK. The reaction mixture was resolved by SDS-PAGE and visualized using a fluorescence imager. The competitive labeling was performed in the presence of 1 μM AzaN5s. The molecular mass standards are indicated in kDa.

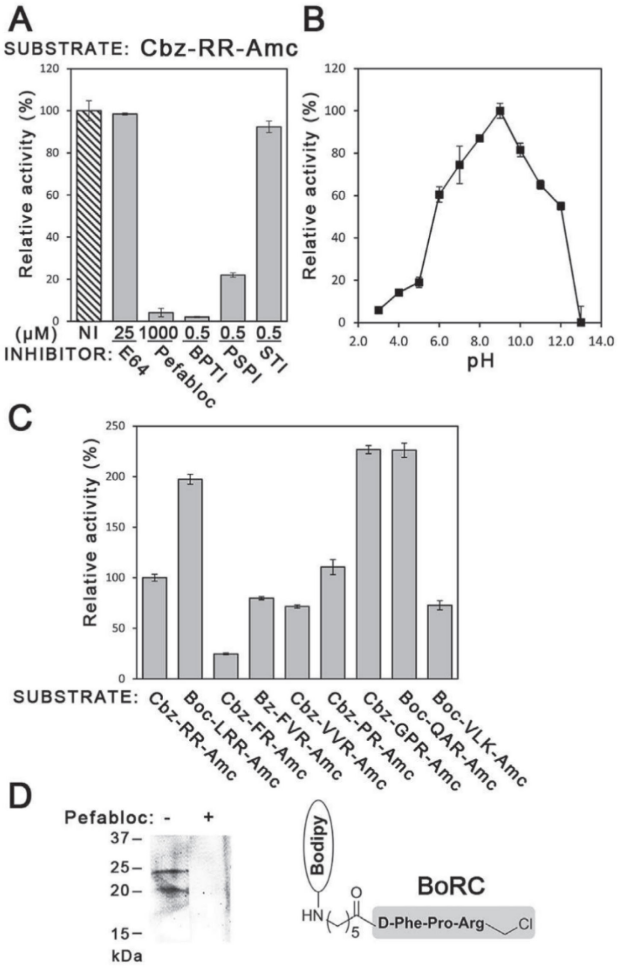


Fig. 6. Characterization of trypsin-type serine peptidases in the CPB gut. The analyzed gut extract of CPB larvae was treated with 50 μM E64 to eliminate the interference of papain-family cysteine peptidases. (A) Inhibitor sensitivity of trypsin-like activity in the gut extract was measured with the fluorogenic substrate Cbz-RR-Amc at pH 9.0 in the presence of various selective inhibitors (concentrations are indicated). The inhibition is expressed as remaining proteolytic activity relative to the uninhibited control (NI, 100%); the mean values ± SE are given. (B) The pH profile of trypsin-like activity was measured as in panel (A) at various pH values. The activities (mean values ± SE) are expressed relative to the maximum value (100%). (C) The substrate specificity of trypsin-like activity was probed with various fluorogenic substrates (25 μM) at pH 9.0. The activities are expressed relative to that with the substrate Cbz-RR-Amc (100%); the mean values ± SE are given. (D) Imaging of trypsin peptidases in the gut extract with BoRC, a fluorescent ABP that interacts with the trypsin active site. The reaction mixture was resolved by SDS-PAGE and visualized using a fluorescence imager. The competitive labeling was performed in the presence of 2 mM Pefabloc, an inhibitor of serine peptidases. The molecular mass standards are indicated in kDa.

substrate Cbz-FL and sensitivity to DFP, a general inhibitor of serine peptidases (Ferreira et al., 2015). Inhibitors of metallopeptidases, including EDTA, 1,10-phenanthroline, and PCPI (potato carboxypeptidase inhibitor), and the serine peptidase inhibitor PMSF were not effective (Fig. 7B). The enzyme was active below pH 7, with maximal activity at pH 4 (Fig. 7A).

Prolyl carboxypeptidases from the S28 family hydrolyze the C-terminal residue following the penultimate Pro residue. This activity was measured using the selective substrate Abz-FRPF* below pH 8 with a maximum at pH 5 (Fig. 7A). The enzyme was sensitive to the serine peptidase inhibitor Pefabloc and prolinal (Cbz-PP-Ald), an inhibitor of post-proline cleaving enzymes (Goptar et al., 2013).

3.6. Non-secreted proteolytic activities in the gut tissue

Significant activities of two types of peptidases were detected in the gut extract of CPB larvae but not in larvae regurgitant, indicating that they most likely act in epithelial cells of the CPB gut (Table 1). First, aminopeptidases of the metallopeptidase class were identified based on hydrolysis around neutral pH of the substrate L-Amc and sensitivity to the inhibitor bestatin. These properties and localization are consistent with the characteristics of leucine aminopeptidase and aminopeptidase N, which have been characterized in insects (Bozic et al., 2008; Cristofoletti and Terra, 1999). Second, a prolyl oligopeptidase from the SC clan of serine peptidases was identified based on hydrolysis at pH 6 of the substrate Cbz-GP-Amc and sensitivity to the inhibitors Cbz-AP-Cmk and prolinal (Cbz-PP-Ald), which is in line with the properties of this enzyme in other insects (Goptar et al., 2008).

4. Discussion

In the present work, we identify and unravel a proteolytic system functioning in protein digestion in the gut of CPB larvae. First, we biochemically dissected a number of peptidase activities in the gut extract. For activity profiling, we used a battery of specific substrates and inhibitors. The peptidases were further authenticated by imaging using a unique panel of selective activity-based probes. This combined approach enabled classification of the main component peptidase activities: (i) cysteine peptidases from the C1 (papain)-family of the CA clan, called intestains, with cathepsin L-like, B-like, and H-like activities; (ii) cathepsin D-like aspartic peptidases from the A1 (pepsin) family of the AA clan; (iii) serine peptidases of the trypsin type from the S1 family of the PA clan and serine carboxypeptidases (S10 family) and prolyl carboxypeptidases (S28 family) of the SC clan; and (iv) metalloaminopeptidases from the MA/MF clans and prolyl oligopeptidases from the SC clan. Peptidases (i) to (iii) are secreted enzymes that accumulate in the gut lumen and are present in the regurgitant, while peptidases (iv) are not secreted. Our classification of individual CPB peptidases is further supported by their identification in the transcriptome of CPB larvae (Kumar et al., 2014), as presented in Table S1.

Second, we simulated the degradation of model proteins *in vitro* with gut extract. Hydrolysis is optimal at mildly acidic and neutral pH, with a broad pH optimum ranging from approximately 5 to 7. This corresponds to pH environments in the gut lumen of CPB larvae, which range from 5.4 to 7.0 (Krishnan et al., 2007). We used selective inhibitors to dissect the contributions of major peptidase activities to the protein substrate degradation. The subsequent data on degradation potency and the fragment profile enabled us to trace the degradation pathway and propose a model of proteolytic digestion (Fig. 8).

In this model, endopeptidases perform the initial cleavage of ingested plant proteins. Intestains with cathepsin L-like activities are the most critical because of their high turnover efficiency and capability to deeply degrade the protein substrate; the role of cathepsin B-like intestains is less important at this stage. Cathepsin D has lower turnover efficiency and generates predominantly large initial fragments. Trypsin is least effective regarding the extent of fragmentation efficiency. Hence, its role may be supportive in digestion and/or it may play another specific role, such as in the adaptation process to plant defensive protease inhibitors. Exopeptidases are dominant in the next phases of proteolysis, which break down large and medium fragments into dipeptides and amino acids. The following enzymes catalyze final decomposition through N- and C-terminal trimming of fragments: intestains with aminopeptidase (cathepsin H-like) and carboxydipeptidase (cathepsin B-like) activities and serine and prolyl carboxypeptidases. We expect that the *in vivo* pathway might also incorporate peptide processing through the action of non-secreted peptidases

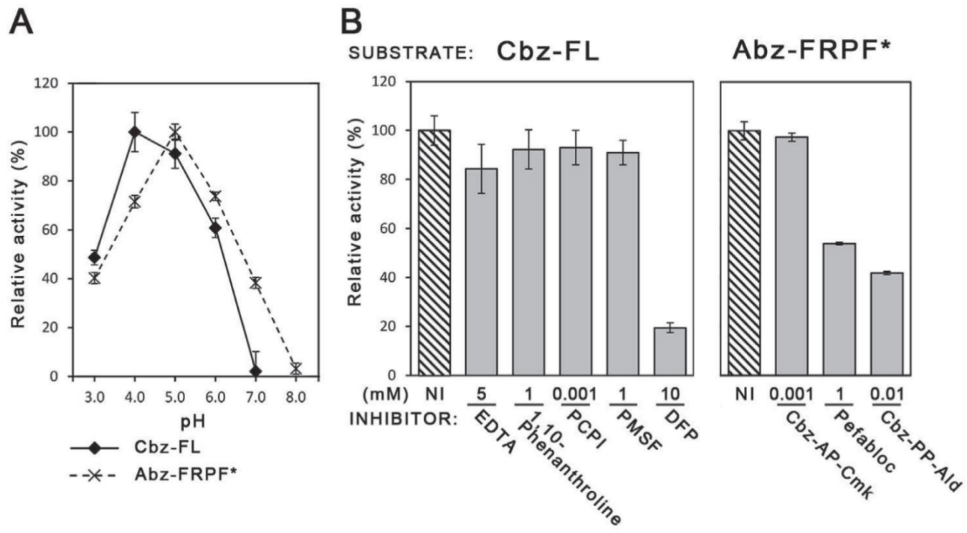


Fig. 7. Activity and inhibitor specificity of carboxypeptidases of the serine class in the CPB gut. Exopeptidase activity of S10-family serine carboxypeptidase (SCP) and S28-family prolyl carboxypeptidase (PCP) was analyzed in the gut extract of CPB larvae. (A) The pH activity profile was measured with the selective substrates Cbz-FL and Abz-FRPF* for SCP-like and PCP-like activities, respectively, at various pH values. The activities (mean values ± SE) are expressed relative to the maximum values (100%). (B) Inhibitor sensitivity of SCP-like and PCP-like activities was measured as in panel (A) at pH 4.0 and 5.0, respectively, in the presence of various selective inhibitors (concentrations are indicated). The inhibition is expressed as the remaining proteolytic activity relative to the uninhibited control (NI, 100%); the mean values ± SE are given.

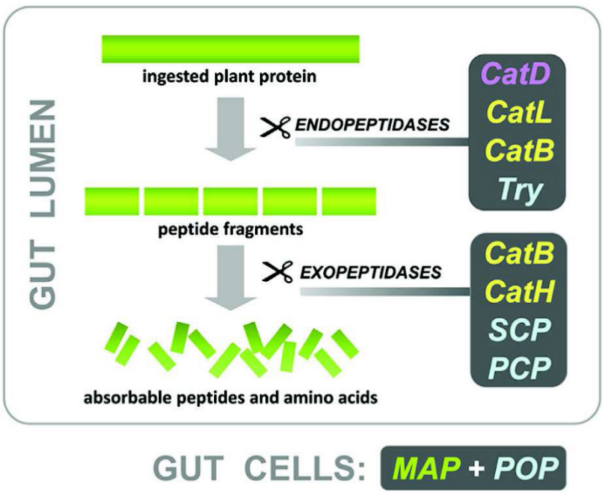


Fig. 8. Mechanistic model of the proteolytic pathway for degradation of plant proteins in the gut of CPB larvae. The luminal endopeptidases, including cathepsin D (CatD), a group of intestains with cathepsins L-like and B-like activities (CatL and CatB, respectively) and trypsin (Try) are responsible for primary cleavage of the protein substrate. Exopeptidases act on the peptides released by the action of the endopeptidases. This includes the aminopeptidase activity of cathepsin H-like intestain (CatH), carboxypeptidase activity of serine and prolyl carboxypeptidases (SCP and PCP, respectively), and carboxydipeptidase activity of cathepsin B-like intestain (CatB). In gut cells, prolyl oligopeptidase (POP) and metalloaminopeptidase (MAP) might participate in peptide degradation and liberation of free amino acids. The enzymes are color-coded according to class membership: aspartic peptidases (pink), papain-family cysteine peptidases (yellow), serine peptidases (blue), and metallopeptidases (green).

associated with the gut epithelium, including metalloaminopeptidases (leucine aminopeptidase and/or aminopeptidase N) and prolyl oligopeptidase.

Detailed biochemical characterization of the component peptidases provided new insights into the molecular evolution of the digestive proteolytic network in CPB. Intestains from the papain-family cysteine peptidases were found to be the major digestive peptidases in CPB larvae. More than thirty intestain sequences categorized into six phylogenetic groups are currently available (Gruden et al., 2004; Sainsbury et al., 2012); however, these enzymes have yet not been functionally analyzed in detail. Our work shows that previously used discrimination between cathepsin L-like and B-like intestains based on proteolysis of Cbz-FR-Amc and Cbz-RR-Amc endopeptidase substrates needs to be combined with a more accurate analysis of sensitivity to specific inhibitors such as CA074 targeting cathepsins B. We show that Cbz-RR-Amc, a substrate of mammalian cathepsins B (but not cathepsins L), is readily cleaved by cathepsin L-like intestains; this finding is in line with the promiscuous cleavage specificity reported for cathepsins L in two coleopteran species (Beton et al., 2012; Bown et al., 2004). The presence of cathepsins B in CPB was, however, clearly demonstrated by their carboxydipeptidase activity (detected with Abz-FRVF* substrate) sensitive to CA074. A dual mode of action, combining carboxydipeptidase and endopeptidase activities, is typical for cathepsins B (Jilkova et al., 2011). The enzyme identity was further supported by mining of the CPB transcriptome that revealed cathepsin B-type sequences with a characteristic pattern of the occluding loop, which controls the carboxydipeptidase activity (Table S1).

Cathepsin H-like intestains with aminopeptidase activity (and distinct inhibition specificity) have been detected in the CPB gut by us and others (e.g., Gruden et al., 1998; Novillo et al., 1997). Surprisingly, the CPB transcriptome mining did not identify any intestain sequence containing the mini-chain segment, a hallmark of mammalian cathepsins H that confers its aminopeptidase specificity (Table S1). Moreover, this structural feature was also absent in the sequences of all papain-family peptidases in available coleopteran genomes. Based on this, we propose that cathepsin H-like intestain(s) evolved a different structural element for steric hindrance of the active site that results in the aminopeptidase mode of action.

We demonstrated that at least some intestains display a post-asparagine cleaving activity, which is clearly distinguished from legumain activity present in ingested potato tissue (Outchkourov et al., 2003). Legumains are frequent component enzymes of the acidic digestive systems of invertebrates including arthropods, but surprisingly are not encoded in the genomes of Coleoptera and other Endopterygota insects (Horn et al., 2009; Santamaria et al., 2012). Therefore, we hypothesize that in these taxons, papain-family peptidases such as intestains with post-asparagine

cleaving activity may substitute for legumains.

Pepsin-family aspartic peptidases were identified as an important component of the CPB digestive network. For the first time, we provide evidence of their efficient inhibition by PCDI, a Kunitz protein produced in potato tissues and proposed to be involved in plant defense (Kreft et al., 1997; Mares et al., 1989). PCDI is a specific inhibitor of cathepsin D-related enzymes from the pepsin family (Cater et al., 2002; Sojka et al., 2012), which allows us to classify the targeted CPB peptidases. It remains to be determined whether one or several cathepsin D isoenzymes participate in digestive proteolysis in CPB (Table S1).

Serine peptidases secreted in the CPB gut include endopeptidases from the S1 family and two carboxypeptidases from S10 and S28 families. We detected two isoenzymes of trypsin-like peptidases from the S1 family but only a weak chymotrypsin-like activity. This prevalence of trypsins in CPB contrasts with the distribution of serine peptidases in granivorous tenebrionid beetles, which possess both trypsin-like and chymotrypsin-like activities (Elpidina et al., 2005; Vinokurov et al., 2009). The CPB trypsins were found to be sensitive to PSPI, a Kunitz protein from potato that inhibits mammalian S1 peptidases (Valueva et al., 2000), suggesting its role in plant defense against insect herbivores. The acidic carboxypeptidases secreted in the CPB gut include a serine carboxypeptidase and prolyl carboxypeptidase reported recently as digestive enzymes in Coleoptera (Ferreira et al., 2015; Goptar et al., 2013).

In conclusion, our study has identified and dissected a network of peptidases that orchestrate proteolytic digestion in the CPB gut. The process is partially ordered in that some peptidases act to initially cleave the protein substrate, and these and others then cooperate (with some redundancy) to complete substrate degradation. The luminal digestive proteolytic system of CPB is reminiscent of the acidic digestive proteolysis based on aspartic and cysteine peptidases that is evolutionarily old and occurs in some protozoa, helminths, and ticks (Horn et al., 2009; Sojka et al., 2013). However, CPB has also acquired enzyme tools of alkaline digestive proteolysis that rely on serine peptidases, as in lepidopteran insects (Terra and Ferreira, 2012). Finally, our results improve understanding of the molecular biology of CPB and its interactions with plants and provide a basis for pest management strategies. The major digestive peptidases identified in this study can be exploited to search for novel peptidase inhibitors with optimized functional properties that could be engineered as protein insecticides in potato plants to confer resistance to CPB.

Acknowledgements

We thank M. Bogyo (Stanford University School of Medicine, Stanford, CA) for providing the Fhex-PD-AOMK and Green-DCG-04 probes, J. C. Powers (Georgia Institute of Technology, Atlanta, GA) for the AzaN5s inhibitor, A. Metke (Federal Biological Research Centre for Agriculture and Forestry, Braunschweig, Germany) for the insect colony, E. Hausvater (Potato Research Institute Havlíčkův Brod, Ltd., Czech Republic) for potato plants, J. Illner and R. Maláková for preparation of initial biochemical and biological materials, and H. Hoffman for critical reading of the manuscript. This work was supported by project InterBioMed LO1302 (Ministry of Education of the Czech Republic), grant 15-18929S (Czech Science Foundation), grant 81122/2014 (Charles University Grant Agency, Czech Republic), and institutional project RVO61388963.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ibmb.2016.08.004.

References

Beton, D., Guzzo, C.R., Ribeiro, A.F., Farah, C.S., Terra, W.R., 2012. The 3D structure and function of digestive cathepsin L-like proteinases of *Tenebrio molitor* larval midgut. Insect Biochem. Mol. Biol. 42, 655–664.

Bolter, C.J., Jongsma, M.A., 1995. Colorado potato beetles (*Leptinotarsa-decemlineata*) adapt to proteinase-inhibitors induced in potato leaves by methyl jasmonate. J. Insect Physiol. 41, 1071–1078.

Bolter, C.J., LatoszekGreen, M., 1997. Effect of chronic ingestion of the cysteine proteinase inhibitor, E-64, on Colorado potato beetle gut proteinases. Entomol. Exp. Appl. 83, 295–303.

Bown, D.P., Wilkinson, H.S., Jongsma, M.A., Gatehouse, J.A., 2004. Characterisation of cysteine proteinases responsible for digestive proteolysis in guts of larval western corn rootworm (*Diabrotica virgifera*) by expression in the yeast *Pichia pastoris*. Insect Biochem. Mol. Biol. 34, 305–320.

Bozic, N., Ivanovic, J., Nenadovic, V., Bergstrom, J., Larsson, T., Vujcic, Z., 2008. Purification and properties of major midgut leucyl aminopeptidase of *Morimus funereus* (Coleoptera, *Cerambycidae*) larvae. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 149, 454–462.

Brunelle, F., Girard, C., Cloutier, C., Michaud, D., 2005. A hybrid, broad-spectrum inhibitor of Colorado potato beetle aspartate and cysteine digestive proteinases. Arch. Insect Biochem. Physiol. 60, 20–31.

Brunelle, F., Nguyen-Quoc, B., Cloutier, C., Michaud, D., 1999. Protein hydrolysis by Colorado potato beetle, *Leptinotarsa decemlineata*, digestive proteases: the catalytic role of cathepsin D. Arch. Insect Biochem. Physiol. 42, 88–98.

Cater, S.A., Lees, W.E., Hill, J., Brzin, J., Kay, J., Phylip, L.H., 2002. Aspartic proteinase inhibitors from tomato and potato are more potent against yeast proteinase A than cathepsin D. Biochim. Biophys. Acta 1596, 76–82.

Cezari, M.H., Puzer, L., Juliano, M.A., Carmona, A.K., Juliano, L., 2002. Cathepsin B carboxydipeptidase specificity analysis using internally quenched fluorescent peptides. Biochem. J. 368, 365–369.

Cloutier, C., Jean, C., Fournier, M., Yelle, S., Michaud, D., 2000. Adult Colorado potato beetles, *Leptinotarsa decemlineata* compensate for nutritional stress on oryzacystatin I-transgenic potato plants by hypertrophic behavior and overproduction of insensitive proteases. Arch. Insect Biochem. Physiol. 44, 69–81.

Cristofoletti, P.T., Terra, W.R., 1999. Specificity, anchoring, and subsites in the active center of a microvillar aminopeptidase purified from *Tenebrio molitor* (Coleoptera) midgut cells. Insect Biochem. Mol. 29, 807–819.

Elpidina, E.N., Tsybina, T.A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P., Oppert, B., 2005. A chymotrypsin-like proteinase from the midgut of *Tenebrio molitor* larvae. Biochimie 87, 771–779.

Fajtova, P., Stefanic, S., Hradilek, M., Dvorak, J., Vondrasek, J., Jilkova, A., Ulrychova, L., McKerrow, J.H., Caffrey, C.R., Mares, M., Horn, M., 2015. Prolyl oligopeptidase from the blood fluke *Schistosoma mansoni*: from functional analysis to anti-schistosomal inhibitors. PLoS Negl. Trop. Dis. 9, e0003827.

Ferreira, C., Rebola, K.G., Cardoso, C., Bragatto, I., Ribeiro, A.F., Terra, W.R., 2015. Insect midgut carboxypeptidases with emphasis on S10 hemipteran and M14 lepidopteran carboxypeptidases. Insect Mol. Biol. 24, 222–239.

Goptar, I.A., Filippova, I.Y., Lysogorskaya, E.N., Oksenoit, E.S., Vinokurov, K.S., Zhuzhikov, D.P., Bulushova, N.V., Zalunin, I.A., Dunaevsky, Y.E., Belozersky, M.A., Oppert, B., Elpidina, E.N., 2008. Localization of post-proline cleaving peptidases in *Tenebrio molitor* larval midgut. Biochimie 90, 508–514.

Goptar, I.A., Shagin, D.A., Shagina, I.A., Mudrik, E.S., Smirnova, Y.A., Zhuzhikov, D.P., Belozersky, M.A., Dunaevsky, Y.E., Oppert, B., Filippova, I.Y., Elpidina, E.N., 2013. A digestive prolyl carboxypeptidase in *Tenebrio molitor* larvae. Insect Biochem. Mol. Biol. 43, 501–509.

Gotz, M.G., James, K.E., Hansell, E., Dvorak, J., Seshaadri, A., Sojka, D., Kopacek, P., McKerrow, J.H., Caffrey, C.R., Powers, J.C., 2008. Aza-peptidyl Michael acceptors. A new class of potent and selective inhibitors of asparaginyl endopeptidases (legumains) from evolutionarily diverse pathogens. J. Med. Chem. 51, 2816–2832.

Goulet, M.C., Dallaire, C., Vaillancourt, L.P., Khalf, M., Badri, A.M., Preradov, A., Duceppe, M.O., Goulet, C., Cloutier, C., Michaud, D., 2008. Tailoring the specificity of a plant cystatin toward herbivorous insect digestive cysteine proteases by single mutations at positively selected amino acid sites. Plant Physiol. 146, 1010–1019.

Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K.F., Bogyo, M., 2002. Chemical approaches for functionally probing the proteome. Mol. Cell. Proteom. 1, 60–68.

Gruden, K., Kuipers, A.G., Guncar, G., Slapar, N., Strukelj, B., Jongsma, M.A., 2004. Molecular basis of Colorado potato beetle adaptation to potato plant defence at the level of digestive cysteine proteinases. Insect Biochem. Mol. Biol. 34, 365–375.

Gruden, K., Popovic, T., Cimerman, N., Krizaj, I., Strukelj, B., 2003. Diverse enzymatic specificities of digestive proteases, 'intestains', enable Colorado potato beetle larvae to counteract the potato defence mechanism. Biol. Chem. 384, 305–310.

Gruden, K., Strukelj, B., Popovic, T., Lenarcic, B., Bevec, T., Brzin, J., Kregar, I., Herzog-Velikonja, J., Stiekema, W.J., Bosch, D., Jongsma, M.A., 1998. The cysteine protease activity of Colorado potato beetle (*Leptinotarsa decemlineata* Say) guts, which is insensitive to potato protease inhibitors, is inhibited by thyroglobulin type-1 domain inhibitors. Insect Biochem. Mol. Biol. 28, 549–560.

Horn, M., Nussbaumerova, M., Sanda, M., Kovarova, Z., Srba, J., Franta, Z., Sojka, D., Bogyo, M., Caffrey, C.R., Kopacek, P., Mares, M., 2009. Hemoglobin digestion in blood-feeding ticks: mapping a multipeptidase pathway by functional

proteomics. Chem. Biol. 16, 1053–1063.
Horn, M., Patankar, A.G., Zavala, J.A., Wu, J., Doleckova-Maresova, L., Vujtechova, M., Mares, M., Baldwin, I.T., 2005. Differential elicitation of two processing proteases controls the processing pattern of the trypsin proteinase inhibitor precursor in *Nicotiana attenuata*. Plant Physiol. 139, 375–388.
Jilkova, A., Rezacova, P., Lepsik, M., Horn, M., Vachova, J., Fanfrlik, J., Brynda, J., McKerrow, J.H., Caffrey, C.R., Mares, M., 2011. Structural basis for inhibition of cathepsin B drug target from the human blood fluke, *Schistosoma mansoni*. J. Biol. Chem. 286, 35770–35781.
Kembhavi, A.A., Buttle, D.J., Knight, C.G., Barrett, A.J., 1993. The two cysteine endopeptidases of legume seeds: purification and characterization by use of specific fluorometric assays. Arch. Biochem. Biophys. 303, 208–213.
Kreft, S., Ravnikar, M., Mesko, P., Pungercar, J., Umek, A., Kregar, I., Strukelj, B., 1997. Jasmonic acid inducible aspartic proteinase inhibitors from potato. Phytochemistry 44, 1001–1006.
Krishnan, N., Kodrik, D., Turanli, F., Sehnal, F., 2007. Stage-specific distribution of oxidative radicals and antioxidant enzymes in the midgut of *Leptinotarsa decemlineata*. J. Insect Physiol. 53, 67–74.
Kumar, A., Congiu, L., Lindstrom, L., Piiroinen, S., Vidotto, M., Grapputo, A., 2014. Sequencing, De Novo assembly and annotation of the Colorado potato beetle, *Leptinotarsa decemlineata*, transcriptome. PLoS One 9, e86012.
Mares, M., Meloun, B., Pavlik, M., Kostka, V., Baudys, M., 1989. Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. FEBS Lett. 251, 94–98.
Masa, M., Maresova, L., Vondrasek, J., Horn, M., Jezek, J., Mares, M., 2006. Cathepsin D propeptide: mechanism and regulation of its interaction with the catalytic core. Biochemistry 45, 15474–15482.
Novillo, C., Castanera, P., Ortego, F., 1997. Characterization and distribution of chymotrypsin-like and other digestive proteases in Colorado potato beetle larvae. Arch. Insect Biochem. 36, 181–201.
Nussbaumerova, M., Srp, J., Masa, M., Hradilek, M., Sanda, M., Reinis, M., Horn, M., Mares, M., 2010. Single- and double-headed chemical probes for detection of active cathepsin D in a cancer cell proteome. Chembiochem 11, 1538–1541.
Outchkourov, N.S., Rogelj, B., Strukelj, B., Jongsma, M.A., 2003. Expression of sea anemone equistatin in potato. Effects of plant proteases on heterologous protein production. Plant Physiol. 133, 379–390.
Petek, M., Turnsek, N., Gasparic, M.B., Novak, M.P., Gruden, K., Slapar, N., Popovic, T., Strukelj, B., Gruden, K., Strukelj, B., Jongsma, M.A., 2012. A complex of genes involved in adaptation of *Leptinotarsa decemlineata* larvae to induced potato defense. Arch. Insect Biochem. Physiol. 79, 153–181.
Sainsbury, F., Rheaume, A.J., Goulet, M.C., Vorster, J., Michaud, D., 2012. Discrimination of differentially inhibited cysteine proteases by activity-based profiling using cystatin variants with tailored specificities. J. Proteome Res. 11, 5983–5993.
Sanman, L.E., Bogyo, M., 2014. Activity-based profiling of proteases. Annu. Rev. Biochem. 83, 249–273.
Santamaria, M.E., Hernandez-Crespo, P., Ortego, F., Grbic, V., Grbic, M., Diaz, I., Martinez, M., 2012. Cysteine peptidases and their inhibitors in *Tetranychus urticae*: a comparative genomic approach. BMC Genom. 13, 307.
Sexton, K.B., Witte, M.D., Blum, G., Bogyo, M., 2007. Design of cell-permeable, fluorescent activity-based probes for the lysosomal cysteine protease asparaginyl endopeptidase (AEP)/legumain. Bioorg. Med. Chem. Lett. 17, 649–653.
Sojka, D., Franta, Z., Frantova, H., Bartosova, P., Horn, M., Vachova, J., O'Donoghue, A.J., Eroy-Reveles, A.A., Craik, C.S., Knudsen, G.M., Caffrey, C.R., McKerrow, J.H., Mares, M., Kopacek, P., 2012. Characterization of gut-associated cathepsin D hemoglobinase from tick *Ixodes ricinus* (IrCD1). J. Biol. Chem. 287, 21152–21163.
Sojka, D., Franta, Z., Horn, M., Caffrey, C.R., Mares, M., Kopacek, P., 2013. New insights into the machinery of blood digestion by ticks. Trends Parasitol. 29, 276–285.
Sorgine, M.H., Logullo, C., Zingali, R.B., Paiva-Silva, G.O., Juliano, L., Oliveira, P.L., 2000. A heme-binding aspartic proteinase from the eggs of the hard tick *Boophilus microplus*. J. Biol. Chem. 275, 28659–28665.
Terra, W.R., Ferreira, C., 2012. Biochemistry and molecular biology of digestion. In: Insect Molecular Biology and Biochemistry, first ed. Elsevier B.V., Gilbert, L.I, pp. 365–418.
Valueva, T.A., Revina, T.A., Mosolov, V.V., Mentele, R., 2000. Primary structure of potato kunitz-type serine proteinase inhibitor. Biol. Chem. 381, 1215–1221.
Vinokurov, K.S., Elpidina, E.N., Zhuzhikov, D.P., Oppert, B., Kodrik, D., Sehnal, F., 2009. Digestive proteolysis organization in two closely related tenebrionid beetles: red flour beetle (*Tribolium castaneum*) and confused flour beetle (*Tribolium confusum*). Arch. Insect Biochem. Physiol. 70, 254–279.
Vorster, J., Rasoolizadeh, A., Goulet, M.C., Cloutier, C., Sainsbury, F., Michaud, D., 2015. Positive selection of digestive Cys proteases in herbivorous Coleoptera. Insect Biochem. Mol. Biol. 65, 10–19.
Wolfson, J.L., Murdock, L.L., 1987. Suppression of larval Colorado potato beetle growth and development by digestive proteinase-inhibitors. Entomol. Exp. Appl. 44, 235–240.

Supplementary materials and methods

Profiling of CPB peptidases with substrates and inhibitors

Proteolytic activities were identified and characterized by hydrolysis of the following substrates: 25 μM Cbz-RR-Amc for cathepsin B or trypsin (Barrett and Kirschke, 1981; Choe et al., 2006); 25 μM Cbz-FR-Amc for cathepsin L/B or trypsin (Barrett and Kirschke, 1981; Choe et al., 2006; Terra and Ferreira, 2012); 25 μM Cbz-VVR-Amc (Bromme et al., 1993; Choe et al., 2006), Cbz-GPR-Amc (Choe et al., 2006; Stack et al., 2008), Cbz-PR-Amc (Choe et al., 2006; Stack et al., 2008), Boc-VLK-Amc (Corvo et al., 2009; Choe et al., 2006), Boc-QAR-Amc (Choe et al., 2006; Schulz et al., 1998), Boc-LRR-Amc (Choe et al., 2006), and Bz-FVR-Amc (Bromme et al., 1993; Choe et al., 2006) for intestains or trypsin; 25 μM Cbz-AAN-Amc for legumain (Kembhavi et al., 1993); 25 μM Cbz-GP-Amc for prolyl oligopeptidase (Fajtova et al., 2015); 25 μM Cbz-AAPF-Amc for chymotrypsin (Elpidina et al., 2005); 25 μM Cbz-GGL-Amc for subtilisin (Kanaoka et al., 1985); MeOSuc-AAPV-Amc for elastase (Terra and Ferreira, 2012); 25 μM Ac-DEVD-Amc for caspase (Garcia-Calvo et al., 1999); 25 μM GR-Amc for cathepsin C (McGuire et al., 1997); 25 μM R-Amc for cathepsin H (Vasiljeva et al., 2003); 25 μM L-Amc for leucine aminopeptidase (Horn et al., 2009); 25 μM Abz-KPAEFF*AL for cathepsin D (Masa et al., 2006); 25 μM Abz-FRVF* and Abz-FRPF* for cathepsin B and prolyl carboxypeptidase, respectively (Jilkova et al., 2011); 1 mM Cbz-FL for serine carboxypeptidase (Horn et al., 2009); and 1 mM Bz-GF for metallocarboxypeptidase (Ferreira et al., 2015).

For activity assays in the presence of peptidase inhibitors, an aliquot of the extract was preincubated (15 min at 26 °C) in assay buffer with the following inhibitors: 10-50 μM E64 for intestains (Barrett et al., 1982); 10 μM CA074 for cathepsin B (Murata et al., 1991); 0.5-25 μM Cbz-FF-Dmk for cathepsin L (Caffrey and Ruppel, 1997); 10 μM Cbz-FY(tBu)-Dmk, AAPV-Cmk, Cbz-VVnL-Dmk, and Cbz-FA-Fmk for intestains (Powers et al., 2002); 25 μM L-Cmk for cathepsin H (Singh and Kalnitsky, 1980); 1 μM AzaN5s for legumain (Gotz et al., 2008); 1-10 μM pepstatin for aspartic peptidases (Knight and Barrett, 1976); 0.1-1 μM PCDI for cathepsin D (Mares et al., 1989); 1 mM Pefabloc, 1 mM PMSF, and 10 mM DFP for serine peptidases (Powers et al., 2002); 0.5 μM BPTI (Castro and Anderson, 1996), 0.5 μM PSPI (Valueva et al., 2000), and 0.5 μM STI (De Vonis Bidlingmeyer et al., 1972) for trypsin and chymotrypsin; 0.1 mM chymostatin for chymotrypsin (Galpin et al., 1984); 1 μM Cbz-AP-Cmk and 10 μM Cbz-PP-Ald for prolyl oligopeptidase and prolyl carboxypeptidase (Fajtova et al., 2015; Goptar et al., 2008); 5 mM EDTA and 1 mM 1,10 phenathroline for metallopeptidases (Auld, 1988; Novillo et al., 1997); 0.1 mM bestatin for leucine aminopeptidase (Cristofoletti and Terra, 2000); and 1 μM PCPI for metallocarboxypeptidase (Ryan et al., 1974; Terra and Ferreira, 2012).

References

Auld, D.S., 1988. Use of chelating agents to inhibit enzymes. Methods Enzymol 158, 110-114.

Barrett, A.J., Kembhavi, A.A., Brown, M.A., Kirschke, H., Knight, C.G., Tamai, M., Hanada, K., 1982. L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. Biochem J 201, 189-198.

Barrett, A.J., Kirschke, H., 1981. Cathepsin B, Cathepsin H, and cathepsin L. Methods Enzymol 80 Pt C, 535-561.

Bromme, D., Bonneau, P.R., Lachance, P., Wiederanders, B., Kirschke, H., Peters, C., Thomas, D.Y., Storer, A.C., Vernet, T., 1993. Functional expression of human cathepsin S in *Saccharomyces cerevisiae*. Purification and characterization of the recombinant enzyme. J Biol Chem 268, 4832-4838.

Caffrey, C.R., Ruppel, A., 1997. Cathepsin B-like activity predominates over cathepsin L-like activity in adult *Schistosoma mansoni* and *S. japonicum*. Parasitol Res 83, 632-635.

Castro, M.J., Anderson, S., 1996. Alanine point-mutations in the reactive region of bovine pancreatic trypsin inhibitor: effects on the kinetics and thermodynamics of binding to beta-trypsin and alpha-chymotrypsin. Biochemistry 35, 11435-11446.

Corvo, I., Cancela, M., Cappetta, M., Pi-Denis, N., Tort, J.F., 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.

Cristofoletti, P.T., Terra, W.R., 2000. The role of amino acid residues in the active site of a midgut microvillar aminopeptidase from the beetle *Tenebrio molitor*. Biochim Biophys Acta 1479, 185-195.

De Vonis Bidlingmeyer, U., Leary, T.R., Laskowski, M., Jr., 1972. Identity of the tryptic and alpha-chymotryptic reactive sites on soybean trypsin inhibitor (Kunitz). Biochemistry 11, 3303-3310.

Elpidina, E.N., Tsybina, T.A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P., Oppert, B., 2005. A chymotrypsin-like proteinase from the midgut of *Tenebrio molitor* larvae. Biochimie 87, 771-779.

Fajtova, P., Stefanic, S., Hradilek, M., Dvorak, J., Vondrasek, J., Jilkova, A., Ulrychova, L., McKerrow, J.H., Caffrey, C.R., Mares, M., Horn, M., 2015. Prolyl Oligopeptidase from the Blood Fluke *Schistosoma mansoni*: From Functional Analysis to Anti-schistosomal Inhibitors. PLoS Negl Trop Dis 9, e0003827.

Ferreira, C., Rebola, K.G., Cardoso, C., Bragatto, I., Ribeiro, A.F., Terra, W.R., 2015. Insect midgut carboxypeptidases with emphasis on S10 hemipteran and M14 lepidopteran carboxypeptidases. Insect Mol Biol 24, 222-239.

Galpin, I.J., Wilby, A.H., Place, G.A., Beynon, R.J., 1984. Synthetic analogues of the proteinase inhibitor: chymostatin. Int J Pept Protein Res 23, 477-486.

Garcia-Calvo, M., Peterson, E.P., Rasper, D.M., Vaillancourt, J.P., Zamboni, R., Nicholson, D.W., Thornberry, N.A., 1999. Purification and catalytic properties of human caspase family members. Cell Death Differ 6, 362-369.

Goptar, I.A., Filippova, I.Y., Lysogorskaya, E.N., Oksenoit, E.S., Vinokurov, K.S., Zhuzhikov, D.P., Bulushova, N.V., Zalunin, I.A., Dunaevsky, Y.E., Belozersky, M.A., Oppert, B., Elpidina, E.N., 2008. Localization of post-proline cleaving peptidases in *Tenebrio molitor* larval midgut. Biochimie 90, 508-514.

Gotz, M.G., James, K.E., Hansell, E., Dvorak, J., Seshaadri, A., Sojka, D., Kopacek, P., McKerrow, J.H., Caffrey, C.R., Powers, J.C., 2008. Aza-peptidyl Michael acceptors. A new class of potent and selective inhibitors of asparaginyl endopeptidases (legumains) from evolutionarily diverse pathogens. J Med Chem 51, 2816-2832.

Horn, M., Nussbaumerova, M., Sanda, M., Kovarova, Z., Srba, J., Franta, Z., Sojka, D., Bogyo, M., Caffrey, C.R., Kopacek, P., Mares, M., 2009. Hemoglobin digestion in blood-feeding ticks: mapping a multipeptidase pathway by functional proteomics. Chem Biol 16, 1053-1063.

Choe, Y., Leonetti, F., Greenbaum, D.C., Lecaille, F., Bogyo, M., Bromme, D., Ellman, J.A., Craik, C.S., 2006. Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. J Biol Chem 281, 12824-12832.

Jilkova, A., Rezacova, P., Lepsik, M., Horn, M., Vachova, J., Fanfrlik, J., Brynda, J., McKerrow, J.H., Caffrey, C.R., Mares, M., 2011. Structural basis for inhibition of cathepsin B drug target from the human blood fluke, *Schistosoma mansoni*. J Biol Chem 286, 35770-35781.

Kanaoka, Y., Takahashi, T., Nakayama, H., Tanizawa, K., 1985. New fluorogenic substrates for subtilisin. Chem Pharm Bull (Tokyo) 33, 1721-1724.

Kembhavi, A.A., Buttle, D.J., Knight, C.G., Barrett, A.J., 1993. The two cysteine endopeptidases of legume seeds: purification and characterization by use of specific fluorometric assays. Arch Biochem Biophys 303, 208-213.

Knight, C.G., Barrett, A.J., 1976. Interaction of human cathepsin D with the inhibitor pepstatin. Biochem J 155, 117-125.

Mares, M., Meloun, B., Pavlik, M., Kostka, V., Baudys, M., 1989. Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. FEBS Lett 251, 94-98.

Masa, M., Maresova, L., Vondrasek, J., Horn, M., Jezek, J., Mares, M., 2006. Cathepsin D propeptide: mechanism and regulation of its interaction with the catalytic core. Biochemistry 45, 15474-15482.

McGuire, M.J., Lipsky, P.E., Thiele, D.L., 1997. Cloning and characterization of the cDNA encoding mouse dipeptidyl peptidase I (cathepsin C). Biochim Biophys Acta 1351, 267-273.

Murata, M., Miyashita, S., Yokoo, C., Tamai, M., Hanada, K., Hatayama, K., Towatari, T., Nikawa, T., Katunuma, N., 1991. Novel epoxysuccinyl peptides. Selective inhibitors of cathepsin B, in vitro. FEBS Lett 280, 307-310.

Novillo, C., Castanera, P., Ortego, F., 1997. Characterization and distribution of chymotrypsin-like and other digestive proteases in Colorado potato beetle larvae. Arch Insect Biochem 36, 181-201.

Powers, J.C., Asgian, J.L., Ekici, O.D., James, K.E., 2002. Irreversible inhibitors of serine, cysteine, and threonine proteases. Chem Rev 102, 4639-4750.

Ryan, C.A., Hass, G.M., Kuhn, R.W., 1974. Purification and properties of a carboxypeptidase inhibitor from potatoes. J Biol Chem 249, 5495-5499.

Schulz, O., Sewell, H.F., Shakib, F., 1998. A sensitive fluorescent assay for measuring the cysteine protease activity of Der p 1, a major allergen from the dust mite *Dermatophagoides pteronyssinus*. Mol Pathol 51, 222-224.

Singh, H., Kalnitsky, G., 1980. alpha-N-benzoylarginine-beta-naphthylamide hydrolase, an aminoendopeptidase from rabbit lung. J Biol Chem 255, 369-374.

Stack, C.M., Caffrey, C.R., Donnelly, S.M., Seshaadri, 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., 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.

Terra, W.R., Ferreira, C., 2012. Biochemistry and Molecular Biology of Digestion, in: Gilbert, L.I. (Eds.), Insect Molecular Biology and Biochemistry, Elsevier, London, pp. 365-418.

Valueva, T.A., Revina, T.A., Mosolov, V.V., Mentele, R., 2000. Primary structure of potato kunitz-type serine proteinase inhibitor. Biol Chem 381, 1215-1221.

Vasiljeva, O., Dolinar, M., Turk, V., Turk, B., 2003. Recombinant human cathepsin H lacking the mini chain is an endopeptidase. Biochemistry 42, 13522-13528.

Supplementary table S1

Potential digestive peptidases in the CPB transcriptome. The CPB transcriptome (Kumar et al., 2014) was searched for the major types of gut peptidases identified and classified by functional proteomics analysis (see Table 1). The representative sequences of human homologs were used as a query for searching the transcriptome dataset using the BioEdit program (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Functional residues and regions characteristic of the individual peptidase types are listed, and their presence in the transcripts is indicated.

Peptidase	Sequence codes from the CPB transcriptome	Sequence hallmarks and comments
cathepsin D	Ld_rep_c32853 Ld_rep_c45248 Ld_rep_c32797 Ld_rep_c43651 Ld_rep_c43055	• Intact catalytic residues Asp97 and Asp295 (numbering according to human cathepsin D, UniProt code P07339)
cathepsin L: intestain A intestain B intestain C intestain D intestain E intestain F	 Ld_rep_c24167 Ld_c23771 Ld_rep_c24380 Ld_rep_c32804 Ld_c22604 Ld_c23771	• Intact catalytic residues Cys138, His276, and Asn300 (numbering according to human cathepsin L, UniProt code P07711) • Intestains divided into groups A-F based on (Sainsbury et al., 2012); only representative sequences are listed
cathepsin B	Ld_rep_c24418 Ld_rep_c33422 Ld_rep_c24152	• Intact catalytic residues Cys108, His278, and Asn298 • Intact occluding loop with two His110 and His111 residues (Musil et al., 1991) (numbering according to human cathepsin B, UniProt code P07858)
cathepsin H	not found	• Intact catalytic residues Cys141, His281, and Asn301 • Absence of characteristic mini chain (residues Glu98-Thr105) containing Cys102, which forms a disulfide with Cys327 in the enzyme core (Guncar et al., 1998) (numbering according to human cathepsin H, UniProt code P09668)
aspariginyl endopeptidase	not found	• Intact catalytic residues His148 and Cys189 (numbering according to human asparaginyl endopeptidase, UniProt code Q99538)
trypsin	Ld_rep_c38119 Ld_c7982 Ld_rep_c39034 Ld_c2341 Ld_rep_c33124 Ld_c4395	• Intact catalytic residues His57, Asp102, and Ser195 • S1 subsite residue Asp189 critical for specificity (Hedstrom, 2002) (numbering according to human trypsin, UniProt Code P07477)

prolyl carboxypeptidase	Ld_rep_c24488	• Intact catalytic residues Ser179, Asp430, and His455 (numbering according to human prolyl carboxypeptidase, UniProt code P42785)
serine carboxypeptidase	Ld_rep_c24194 Ld_rep_c41480 Ld_rep_c24300	• Intact catalytic residues Ser178, Asp400, His457 (numbering according to human cathepsin A, UniProt code P10619)
leucine aminopeptidase	Ld_c540	• Intact metal binding residues Lys282, Asp287, Asp305, Asp364, Glu366 and active site residues Lys294, Arg368 (numbering according to human leucyl aminopeptidase, UniProt code P28838)
aminopeptidase N	Ld_c66 Ld_c4002	• Intact metal binding residues His388, His392, and Glu411 and active site residues Glu389, Tyr477 (numbering according to human aminopeptidase N, UniProt code P15144)
prolyl oligopeptidase	Ld_rep_c24336	• Intact catalytic residues Ser554, Asp641, and His680 (numbering according to human prolyl endopeptidase, UniProt code P48147)

References

Guncar, G., Podobnik, M., Pungercar, J., Strukelj, B., Turk, V., Turk, D., 1998. Crystal structure of porcine cathepsin H determined at 2.1 A resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. Structure 6, 51-61.

Hedstrom, L., 2002. Serine protease mechanism and specificity. Chem Rev 102, 4501-4524.

Kumar, A., Congiu, L., Lindstrom, L., Piiroinen, S., Vidotto, M., Grapputo, A., 2014. Sequencing, De Novo assembly and annotation of the Colorado Potato Beetle, *Leptinotarsa decemlineata*, Transcriptome. PLoS One 9, e86012.

Musil, D., Zucic, D., Turk, D., Engh, R.A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T., Katunuma, N., et al., 1991. The refined 2.15 A X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. EMBO J 10, 2321-2330.

Sainsbury, F., Rheaume, A.J., Goulet, M.C., Vorster, J., Michaud, D., 2012. Discrimination of differentially inhibited cysteine proteases by activity-based profiling using cystatin variants with tailored specificities. J Proteome Res 11, 5983-5993.

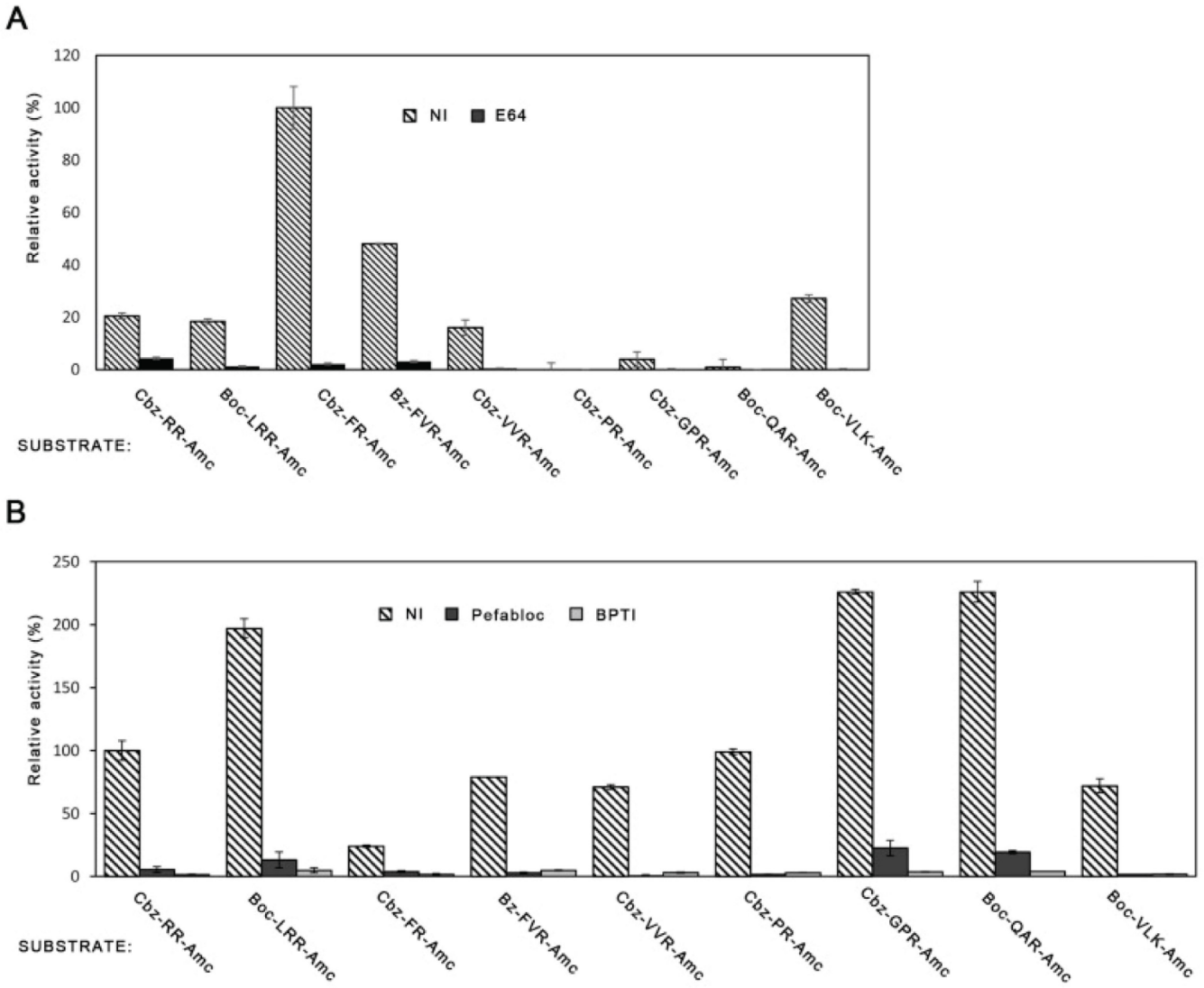


Fig. S1.

Substrate specificity of papain-family peptidases and trypsin peptidases in the CPB gut. Profiling of peptidase activities in the gut extract of CPB larvae using a kinetic assay with various fluorogenic substrates. (A) Papain-family peptidase activities were measured at pH 6.5 in the presence or absence of E64 (25 μM), a specific inhibitor of papain-family peptidases. The activities are expressed relative to that for Cbz-FR-Amc substrate (100%) presented in Table 1; the mean values ± SE are given. (B) Trypsin activities were measured at pH 9.0 in the presence or absence of Pefabloc (1 mM) or BPTI (0.5 μM), specific inhibitors of serine peptidases. The activities are expressed relative to that for Cbz-RR-Amc substrate (100%) presented in Table 1; the mean values ± SE are given. The extract was treated with 50 μM E64 to eliminate the interference of papain-family peptidases.

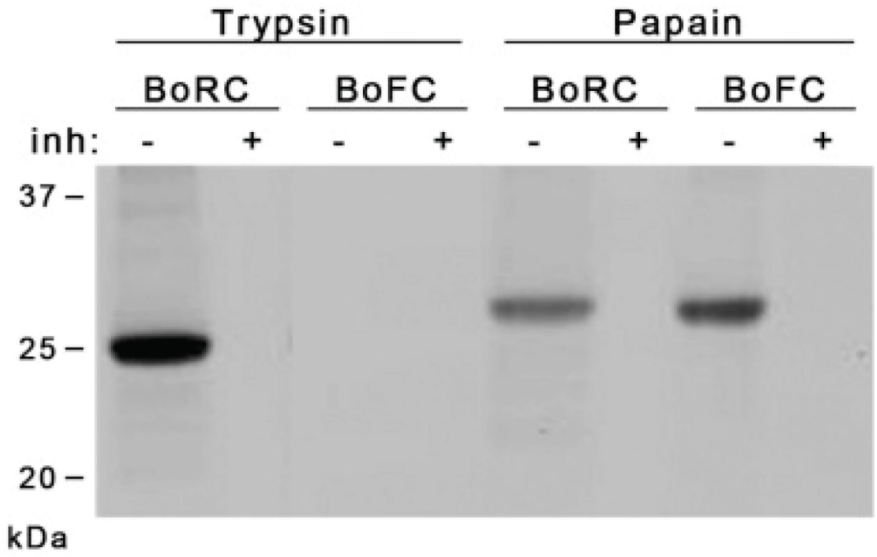


Fig. S2.

Labeling specificity of novel activity-based probes with a chloromethyl ketone-based reactive warhead. The “archetypal” peptidases trypsin and papain (as representative members of trypsin-type serine peptidases and papain-family cysteine peptidases) were treated with the activity-based probes BoRC and BoFC (for chemical structures, see Fig. 3). The competitive labeling was performed in the presence of selective active-site inhibitors (inh): 1 mM Pefabloc for trypsin and 10 μM E64 for papain. The reaction was performed at pH 8 (for trypsin) or 6 (for papain) with 2 μg of peptidases. The reaction mixture was resolved by SDS-PAGE, and the labeled peptidases were visualized using a fluorescence imager. The molecular mass standards are indicated in kDa.

4.2. Nepublikované výsledky

V této kapitole jsou presentovány výsledky o prostorových strukturách proteinů studovaných v publikaci č. 1, jedná se o proteasové inhibitory z Kunitzovy rodiny z lilku brambor a také trávicí aspartátové proteasy z mandelinky bramborové. Tato prozatím nepublikovaná data budou součástí dvou připravovaných publikací s mým prvoautorským podílem.

4.2.1. Strukturní analýza inhibitorů z Kunitzovy rodiny a jejich reaktivních center

V publikaci č. 1 bylo prokázáno, že aktivity aspartátových a serinových proteas v trávicím traktu mandelinky bramborové jsou inhibované inhibitory PCDI a PSPI, které patří do Kunitzovy rodiny z lilku brambor. Tyto proteiny s M_r kolem 20 kDa jsou široce rozšířené v rostlinné říši a vyznačují se konzervovaným uspořádáním segmentů β struktur, které jsou propojené variabilními smyčkami (Bendre, Ramasamy, a Suresh 2018; Renko, Sabotic, a Turk 2012). Tato práce se věnuje dvěma homologním inhibitorům označovaným PCDI a PSPI (aminokyselinové sekvence jsou na Obr. 13 B, str. 60). Tyto inhibitory byly testovány s modelovými zástupci proteas a byla nalezena efektivní inhibice následujících proteas: PSPI inhibuje vybrané serinové proteasy rodiny S1 (např. trypsin, chymotrypsin, elastasu a některé kalikreiny) a PCDI inhibuje stejně jako PSPI serinové proteasy a také vybrané aspartátové proteasy rodiny A1 (např. pepsin, katepsin D a proteinasu A). PCDI je tedy bifunkční inhibitor s duální inhibiční specifitou. Naměřené inhibiční konstanty byly v nanomolární oblasti. Prostorové struktury obou inhibitorů byly v nedávné době vyřešeny (Guo et al. 2015; Guerra et al. 2016; Meulenbroek et al. 2012), ale jejich reaktivní centra a interakční mechanismus nebyl doposud popsán. Proto se projekt zaměřil na vyřešení prostorových struktur těchto inhibitorů v komplexu s modelovými serinovými a aspartátovými proteasami. Konkrétně se jednalo o komplexy PSPI s trypsinem a chymotrypsinem, dále pak o komplexy PCDI s trypsinem a katepsinem D.

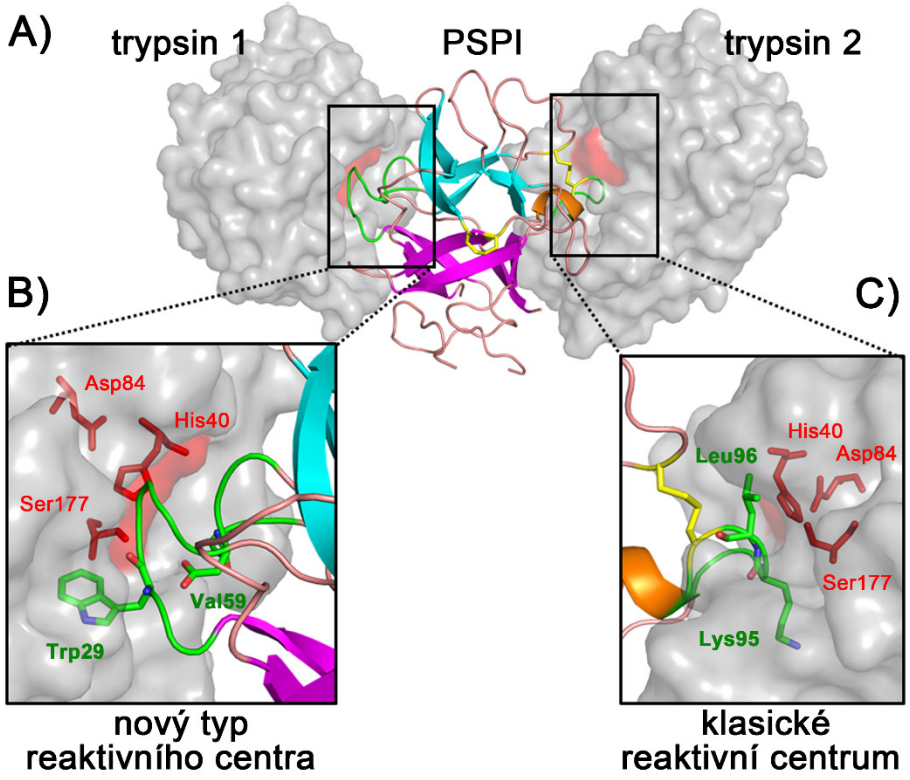
4.2.1.1. Prostorová struktura inhibitoru PSPI v komplexu s trypsinem a chymotrypsinem

Pro inhibitor PSPI byly vyřešeny dvě prostorové struktury ternárních komplexů se serinovými proteasami, obě ve vysokém rozlišení 1,7 Å. Jedná se o komplex PSPI se dvěma molekulami trypsinu (Obr. 9 A, str. 53) a smíšený komplex PSPI s trypsinem a chymotrypsinem (Obr. 10 A, str. 55). Tyto struktury umožnily popsat dvě nezávislá reaktivní centra PSPI s odlišnou stavbou.

První reaktivní centrum PSPI bylo v obou komplexech obsazeno trypsinem a je tvořeno jednou hlavní smyčkou; v sekvenci inhibitoru na obrázku 13 B (str. 60) se jedná o smyčku S3. Tato povrchová smyčka se váže do aktivního místa enzymu (Obr. 9 a 10 C, str. 53 a 55) a tvoří zde síť kontaktů (Tabulka 2 a 3, str. 54 a 56), které stabilizují komplex. Jde o již dříve popsaný jednosmyčkový mechanismus interakce nazývaný Laskowského nebo také kanonický mechanismus interakce, který je využíván většinou rodin inhibitorů serinových proteas. Kritické zbytky pro interakci reaktivní smyčky jsou Lys95 a Leu96 (Obr. 9 a 10, str. 53 a 55), které jsou umístěny v blízkosti

katalytické triády enzymu. Lys95 se váže do S_1 podmísta trypsinu, což je v souladu se substrátovou specifitou tohoto enzymu.

Druhé reaktivní centrum PSPI je schopné vázat kromě trypsinu (Obr. 9 A) také strukturně homologickou serinovou proteasu chymotrypsin (Obr. 10 A, str. 55), která má ale odlišnou substrátovou specifitu. Toto centrum je tvořené dvěma nezávislými smyčkami S1 a S2 na obrázku 13 (str. 59), které se zapojují do interakce s enzymem v aktivním centru a jeho okolí (Tabulka 2 a 3, str. 54 a 56). Jedná se o zcela nový, doposud nepopsaný typ reaktivního centra pro inhibici serinových proteas. Kritický aminokyselinový zbytek reaktivního centra, který vstupuje do S_1 podmísta enzymu, je Trp29 (Obr. 9 B a 10 B, str. 55). Tato aromatická aminokyselina odpovídá substrátové specifitě chymotrypsinu, který preferuje velké hydrofobní aminokyseliny v S_1 podmístě. Schopnost dvousmyčkového reaktivního centra vázat i trypsin s preferencí pro bazické zbytky v S_1 podmístě naznačuje, že toto centrum je velice flexibilní a zajišťuje širokou inhibiční specifitu.



Obrázek 9: Prostorová struktura ternárního komplexu inhibitoru PSPI se dvěma molekulami trypsinu.

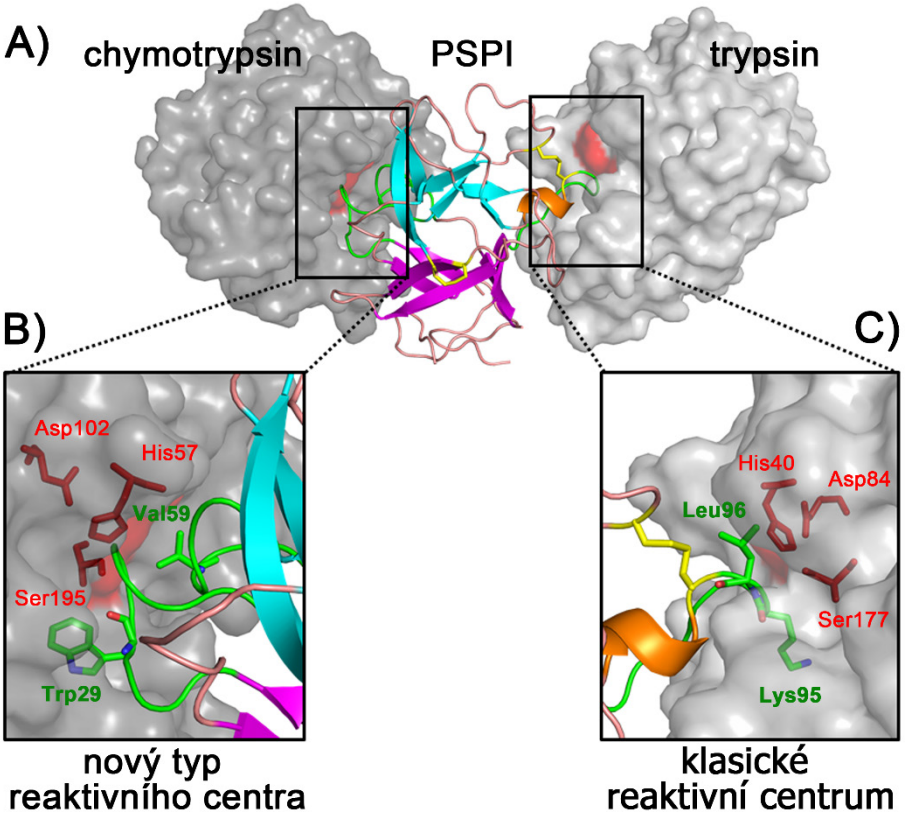
A) Celkový pohled na prostorovou strukturu komplexu. Obě molekuly trypsinu (trypsin 1 a 2) jsou znázorněny šedě jako povrchový model, katalytické zbytky enzymu jsou zvýrazněny červeně. Inhibitor PSPI je znázorněn ve stužkovém modelu, kde fialové a modré jsou β struktury dvou hlavních domén inhibitoru, α šroubovice je oranžová. Žlutě jsou vyznačeny cysteinové zbytky tvořící disulfidové můstky a zeleně dvě interakční smyčky reaktivních center. B) Detailní pohled na nový typ reaktivního centra PSPI se dvěma interakčními smyčkami v katalytickém místě trypsinu 1. Zvýrazněny jsou postranní

řetězce katalytických zbytků (červeně) a zeleně kritické zbytky interakčních smyček inhibitoru (Trp29 a Val59). C) Detailní pohled na klasické reaktivní centrum PSPI s jednou smyčkou v katalytickém místě trypsinu 2. Zvýrazněny jsou postranní řetězce katalytických zbytků (červeně) a zeleně kritické zbytky interakčních smyček inhibitoru (Lys95 a Leu96).

Tabulka 2: Seznam interakcí mezi inhibitorem PSPI a dvěma molekulami trypsinu.

Uvedeny jsou aminokyselinové zbytky trypsinu a PSPI, které tvoří vodíkové vazby, a vzdálenosti mezi příslušnými atomy.

	PSPI		**trypsin**		**vzdálenost**
	zbytek	**atom**	**zbytek**	**atom**	**(Å)**
trypsin 1	Thr27	O	Gly211	N	2,9
	Trp29	N	Gly209	O	3,1
	Trp29	NE1	Gly211	O	2,8
	Ala31	N	Ser207	O	2,9
	Asn56	OD1	Ser93	OG	3,2
	Asp58	OD1	Gln189	NE2	3,0
	Asp58	O	Gln189	NE2	3,2
	Val59	O	Lys58	NZ	2,8
	Arg175	NH2	Ser93	OG	3,1
	Arg175	NH2	Ser93	O	2,8
trypsin 2	Ser92	OG	Gln170	NE2	3,1
	Thr93	N	Gly209	O	3,1
	Thr93	O	Gly209	O	3,1
	Thr93	O	Gly209	N	3,2
	Lys95	N	Ser207	O	2,9
	Lys95	NZ	Ser187	OG	3,2
	Lys95	NZ	Ser187	O	3,1
	Ser99	N	Gln189	OE1	2,7
	Ser99	OG	Gln189	NE2	3,0
	Gln125	OE1	Thr144	N	3,1



Obrázek 10: Prostorová struktura ternárního komplexu inhibitoru PSPI s molekulou trypsinu a chymotrypsinu.

A) Celkový pohled na prostorovou strukturu komplexu. Molekula trypsinu je znázorněna šedě a molekula chymotrypsinu tmavě šedě, obě jako povrchový model, aktivní místa enzymů jsou zvýrazněny červeně. Inhibitor PSPI je znázorněn ve stužkovém modelu, kde fialové a modré jsou znázorněny β struktury dvou hlavních domén inhibitoru, α šroubovice je oranžová. Žlutě jsou vyznačeny cysteinové zbytky tvořící disulfidové můstky a zeleně dvě interakční smyčky reaktivních míst (kanonické a nekanonické). B) Detailní pohled na nový typ reaktivního centra PSPI se dvěma interakčními smyčkami v katalytickém místě chymotrypsinu. Zvýrazněny jsou postranní řetězce katalytických zbytků (červeně) a zeleně kritické zbytky interakčních smyček inhibitoru (Trp29 a Val59). C) Detailní pohled na klasické reaktivní centrum PSPI s jednou smyčkou v katalytickém místě trypsinu. Zvýrazněny jsou postranní řetězce katalytických zbytků (červeně) a zeleně kritické zbytky interakčních smyček inhibitoru (Lys95 a Leu96).

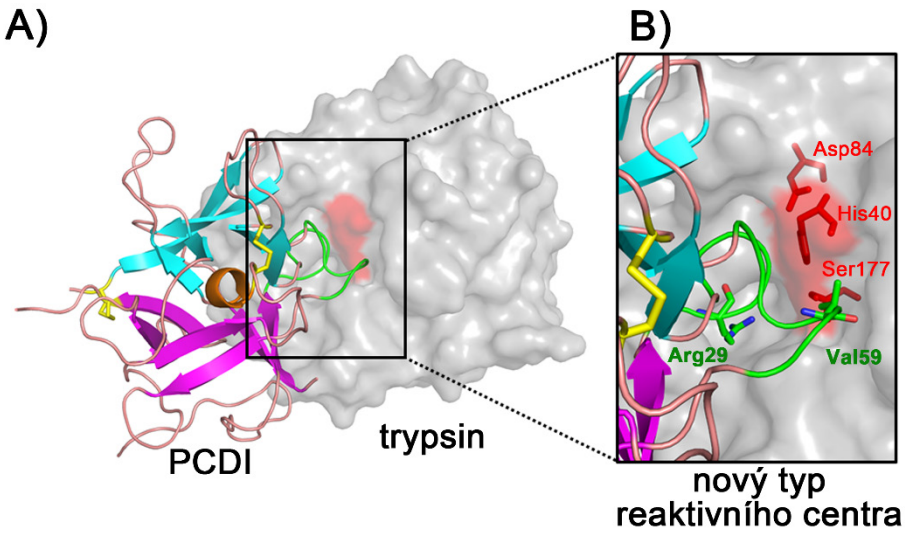
Tabulka 3: Seznam interakcí mezi inhibitorem PSPI a trypsinem a chymotrypsinem.

Uvedeny jsou aminokyselinové zbytky PSPI, trypsinu a chymotrypsinu, které tvoří vodíkové vazby, a vzdálenosti mezi příslušnými atomy.

	PSPI		**chymotrypsin / trypsin**		**vzdálenost**
	zbytek	**atom**	**zbytek**	**atom**	**(Å)**
chymotrypsin	Thr27	O	Ser218	OG	3,2
	Trp29	N	Gly216	O	2,9
	Trp29	NE1	Gly216	O	3,0
	Trp29	NE1	Ser217	O	3,0
	Ala31	N	Ser214	O	3,2
	Asn56	OD1	Ser96	OG	2,8
	Ser57	OG	His57	O	3,1
	His170	NE2	Ser96	OG	3,2
	Arg175	NE	Ser96	O	3,1
	Arg175	NH2	Ser96	O	2,9
trypsin	Ile91	O	Ser195	OG	3,0
	Ser92	OG	Gln155	NE2	2,7
	Thr93	N	Gly194	O	3,0
	Thr93	O	Gly194	N	3,3
	Thr93	O	Gly194	O	3,1
	Ser94	OG	His40	NE2	2,7
	Lys95	N	Ser192	O	2,9
	Ile123	O	Gly128	N	3,2
	Gln125	OE1	Thr129	N	2,5
	Gln125	OE1	Thr129	O	3,1
	Gln125	NE2	Thr129	OG1	3,2

4.2.1.2. Prostorová struktura inhibitoru PCDI v komplexu s trypsinem

Prostorová struktura komplexu PCDI s trypsinem (Obr. 11 A) byla vyřešena na rozlišení 1,9 Å a ukázala, že PCDI interaguje s enzymem obdobně jako PSPI s využitím dvou nezávislých smyček S1 a S2 (Obr. 13 B, str. 60), které se vážou do aktivního centra trypsinu a jeho okolí. Na rozdíl od inhibitoru PSPI, kde je kritický P_1 zbytek aromatický tryptofan, u PCDI se jedná o bazickou aminokyselinu arginin (Arg29, Obr. 11 B), která odpovídá trypsinové substrátové specifitě. Tento komplex je stabilizovaný sítí vodíkových vazeb mezi inhibitorem PCDI a trypsinem (Tab. 4, str. 58), žádná z nich se ale netvoří s katalytickými zbytky enzymu.



Obrázek 11: Prostorová struktura komplexu inhibitoru PCDI se serinovou proteasou trypsinem.
A) Celkový pohled na prostorovou strukturu komplexu. Molekula trypsinu je znázorněna šedě jako povrchový model, katalytické zbytky enzymu jsou zvýrazněny červeně. Inhibitor PCDI je znázorněn ve stužkovém modelu, kde fialové a modré jsou znázorněny β struktury dvou hlavních domén inhibitoru, α šroubovice je oranžová. Žlutě jsou vyznačeny cysteinové zbytky tvořící dva disulfidové můstky a zeleně dvě interakční smyčky. B) Detailní pohled na nový typ reaktivního centra PCDI se dvěma interakčními smyčkami v katalytickém místě trypsinu. Zvýrazněny jsou postranní řetězce katalytických zbytků (červeně) a zeleně kritické zbytky interakčních smyček inhibitoru (Arg29 a Val59).

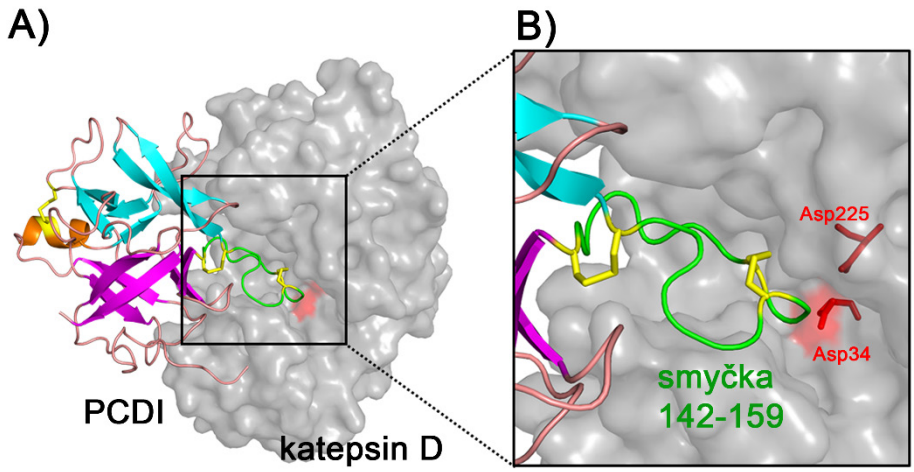
Tabulka 4: Seznam interakcí mezi inhibitorem PCDI a trypsinem.

Uvedeny jsou aminokyselinové zbytky trypsinu a PCDI, které mezi sebou tvoří vodíkové vazby, a vzdálenosti mezi příslušnými atomy.

PCDI		trypsin		vzdálenost (Å)
zbytek	**atom**	**zbytek**	**atom**	
Ile27	O	Gly211	N	2,9
Arg29	N	Gly209	O	2,9
Arg29	NE	Gly211	O	3,0
Arg29	NH2	Ser187	OG	3,1
Arg29	O	Trp208	CB	3,3
Ala31	N	Ser207	O	3,0
Asp58	OD1	Gln189	NE2	3,2
Val59	O	Lys58	NZ	2,9
Arg171	NH2	Ser93	O	3,0

4.2.1.3. Prostorová struktura inhibitoru PCDI v komplexu s aspartátovou proteasou

Prostorová struktura komplexu PCDI s modelovou aspartátovou proteasou typu katepsinu D (IrCD1 z klíštěte obecného) (Obr. 12 A, str. 59) byla vyřešena na vysoké rozlišení 1,7 Å. Jedná se o první strukturní důkaz interakce mezi aspartátovou proteasou a rostlinným inhibitorem z Kunitzovy rodiny. Tato analýza ukázala, že inhibitor PCDI interaguje s aspartátovými proteasami pomocí jediné velké smyčky (S4 na Obr. 13 B, str. 59) o délce 19 aminokyselin (zbytky 142-159). Tato smyčka je stabilizována dvěma disulfidovými můstky a váže se do aktivního centra enzymu (Obr. 12 B, str. 59), ale netvoří žádné přímé kontakty s katalytickými zbytky. Komplex je stabilizován rozsáhlou sítí vodíkových vazeb (Tabulka 5, str. 59) a řadu z nich tvoří smyčka S4. Proto lze předpokládat, že na základě struktury S4 smyčky by bylo možné v budoucnu navrhovat cyklické mimetické inhibitory aspartátových proteas. Intaktní smyčka S4 chybí ve struktuře homologního inhibitoru PSPI; tato pozice je u nativního PSPI posttranslačně proteolyticky štěpena, což může být umožněno její vyšší flexibilitou kvůli absenci jednoho z disulfidů (Obr. 13 B, str. 60).



Obrázek 12: Prostorová struktura inhibitoru PCDI v komplexu s aspartátovou proteasou katepsinem D.

A) Celkový pohled na prostorovou strukturu komplexu. Molekula katepsinu D (IrCD1) je znázorněna šedě jako povrchový model, katalytické zbytky enzymu jsou zvýrazněny červeně. Inhibitor PCDI je ve stužkovém modelu, kde fialově a modře jsou znázorněny β struktury dvou hlavních domén inhibitoru, α šroubovice je oranžová. Žlutě jsou vyznačeny cysteinové zbytky tvořící disulfidové můstky a zeleně smyčka, která interaguje s aktivním místem enzymu (označená jako smyčka 142-159). B) Detailní pohled na interakční smyčku 142-159 PCDI. Červeně jsou zvýrazněny postranní řetězce katalytických zbytků.

Tabulka 5: Seznam interakcí mezi inhibitorem PCDI a katepsinem D (IrCD1).

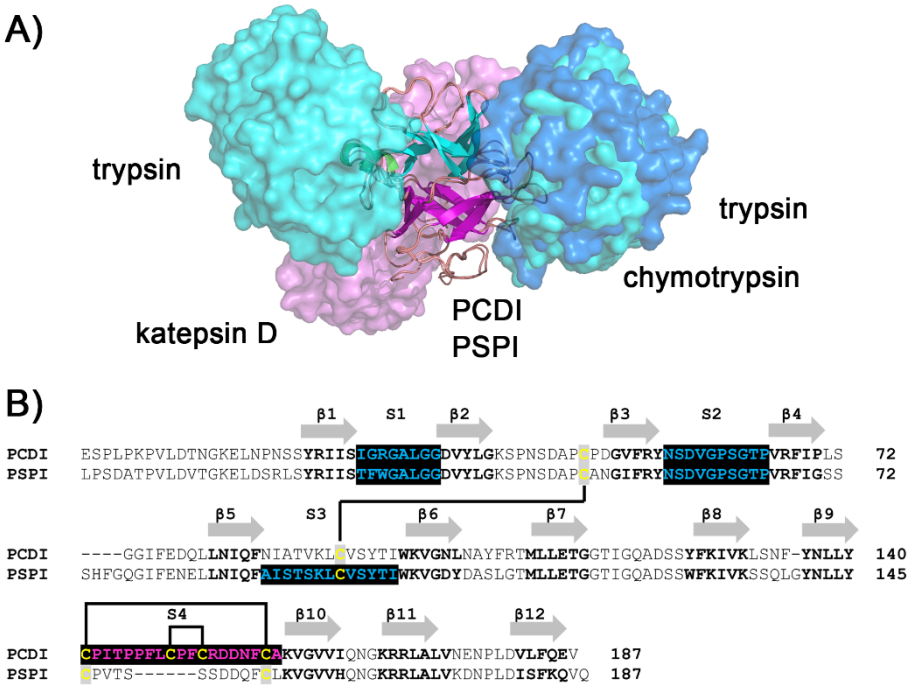
Uvedeny jsou aminokyselinové zbytky IrCD1 a PCDI, které mezi sebou tvoří vodíkové vazby, a vzdálenosti mezi příslušnými atomy.

PCDI		katepsin D		vzdálenost
zbytek	**atom**	**zbytek**	**atom**	**(Å)**
Arg109	O	Thr250	OG1	2,6
Lys128	NZ	Glu112	OE1	2,6
Leu149	N	Ser229	OG	2,7
Leu149	O	Ser229	N	3,1
Pro151	O	Gly78	N	3,1
Arg154	NH1	Gly75	O	2,6
Asn177	ND2	Lys249	O	2,5

4.2.1.4. Topologie reaktivních center ve struktuře inhibitorů PSPI a PCDI

V předchozích kapitolách byla strukturně popsána reaktivní centra pro serinové a aspartátové proteasy u dvou homologních Kunitzových inhibitorů z lilku brambor. Výsledná superpozice prostorových struktur čtyř komplexů inhibitorů s proteasami a porovnání aminokyselinových sekvencí inhibitorů jsou na obrázku 13 B (str. 59) a umožňují komentovat aspekty molekulární evoluce této rodiny inhibičních proteinů. Dvě reaktivní centra pro serinové proteasy (jednosmyčkové v pozici S3 a nové, dvousmyčkové kombinující pozice S1 a S2) se vyvinuly na prostorově opačných stranách inhibitoru PSPI. Nedochází tak ke sterickému bránění mezi oběma enzymy a tento „dvouhlavý“

inhibitor může tvořit ternární komplexy k efektivní regulaci cílových proteas. Dvousmyčkové reaktivní centrum pro serinové proteasy na PCDI má stejnou topologii jako u PSPI inhibitoru (Obr. 13 B). Nový typ reaktivního centra PCDI pro aspartátové proteasy je tvořený jedinou smyčkou S4, která je v jiné pozici než popsaná reaktivní centra pro serinové proteasy a stericky nebrání simultánní interakci obou tříd proteas (Obr. 13). Ze získaných výsledků vyplývá, že u Kunitzových inhibitorů z lilku brambor se vytvořila nová reaktivní centra pro protein-proteinové interakce, která vznikla díky sekvenční variabilitě smyček v různých pozicích umožňující získání nových funkčních vlastností pro komplexní regulaci proteolytických systémů, např. v trávicím traktu mandelinky bramborové.



Obrázek 13: Rozdílná topologie reaktivních center proti proteasám na bramborových Kunitzových inhibitorech

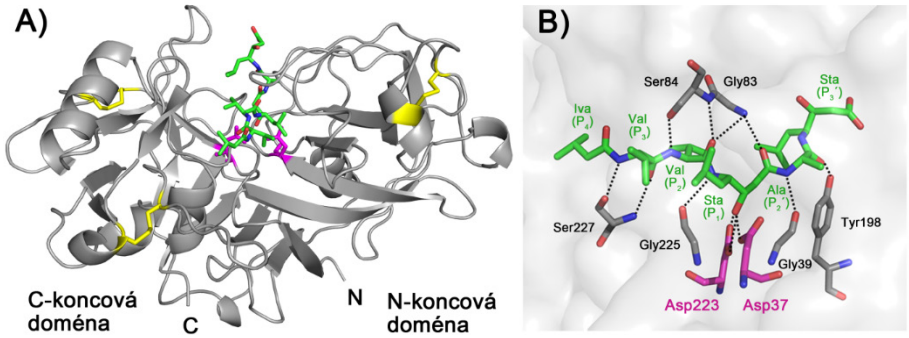
A) Superpozice prostorových struktur PCDI a PSPI v komplexu se serinovými proteasami a katepsinem D. Centrálně jsou umístěny molekuly PCDI a PSPI znázorněné ve stužkovém modelu s fialově-modrým barevným kódováním pro dvě hlavní domény inhibitorů. Molekuly proteas jsou znázorněny jako povrchový model, serinové proteasy jsou znázorněny modře (trypsin světlá, chymotrypsin tmavá) a aspartátová proteasa katepsin D má růžovou barvu. B) Přiřazení aminokyselinových sekvencí Kunitzových inhibitorů PCDI a PSPI. Šedé šipky nad aminokyselinovými sekvencemi znázorňují pozici dvanácti konzervovaných β struktur (β1-β12), párování cysteinových zbytků (žlutě na šedém pozadí) do disulfidových můstků je vyznačeno svorkou. Černým podbarvením jsou znázorněny aminokyseliny, které tvoří interakční smyčky se serinovými proteasami (modré písmo) a s aspartátovou proteasou (růžové písmo). Nový typ reaktivního centra se dvěma interakčními smyčkami je tvořeno smyčkami S1 a S2, klasické reaktivní centrum je tvořeno jednou interakční smyčkou S3 a interakční smyčka 142-159 proti aspratátovým proteasám je označena jako S4.

Přiřazení sekvencí bylo provedeno v programu ClustalW a dále upraveno podle superpozice prostorových struktur PCDI a PSPI v programu PyMol.

4.2.2. Strukturní analýza aspartátové proteasy LdCD

V publikaci č. 1 byla v trávicím traktu mandelinky bramborové identifikována aspartátová proteasa typu katepsinu D označená LdCD. Funkční analýza prokázala, že tento enzym hraje důležitou roli při trávení proteinů a zahajuje jejich degradační kaskádu. V dalším kroku byl připraven tento enzym v kvasinkovém expresním systému *Pichia pastoris* pro jeho strukturní analýzu. Prostorová struktura byla vyřešena pro komplex LdCD s pepstatinem (Obr. 14 A), inhibičním peptidovým analogem substrátu, v rozlišení 1,95 Å. LdCD má typickou architekturu aspartátových proteas pepsinové rodiny (Obr. 3 A, str. 8). Skládá se ze dvou domén (N koncové a C koncové), mezi kterými je umístěno aktivní místo enzymu se dvěma katalytickými zbytky kyseliny asparagové. Na rozdíl od lidského katepsinu D, který je posttranslačně štěpen na lehký a těžký řetězec (Baldwin et al. 1993), je LdCD jednořetězcová molekula, která je stabilizována třemi disulfidovými můstky.

Molekula pepstatinu se váže do aktivního místa LdCD obdobně jako to bylo popsáno u prostorových struktur jiných aspartátových proteas s tímto inhibitorem (Hánová et al. 2018), tj. od S_4 do S_3'podmísta, kterým odpovídají pozice P_4 až P_3' na molekule pepstatinu (Obr. 14 B). Klíčový aminokyselinový zbytek statinu interaguje v S_1 podmístě LdCD, kde tvoří přímé kontakty s katalytickými aspartáty a jeho hydroxylová skupina nahrazuje molekulu vody, která se u aktivního enzymu účastní katalytického mechanismu a štěpení peptidové vazby (Coates et al. 2008) (Obr. 3 B, str. 8). Další interakce mezi inhibitorem a aktivním centrem LdCD jsou uvedeny v tabulce 6 (str. 62), celkem je tento komplex stabilizován dvanácti vodíkovými vazbami.



Obrázek 14: Prostorová struktura aspartátové proteasy LdCD v komplexu s inhibitorem pepstatinem.

A) Celkový pohled na prostorovou strukturu LdCD ve stužkovém modelu, vyznačen je N a C konec proteinu a členění na domény. Fialově jsou vyznačené katalytické zbytky (Asp37 a Asp223), žlutě cysteinové zbytky tvořící tři disulfidové můstky. Zelený tyčkový model znázorňuje molekulu inhibitoru pepstatinu, který je navázán v aktivním místě LdCD jako analog transitního stavu substrátu. B) Interakce mezi pepstatinem a LdCD je znázorněná pomocí vodíkových vazeb (tečkovanou čárou). Ve

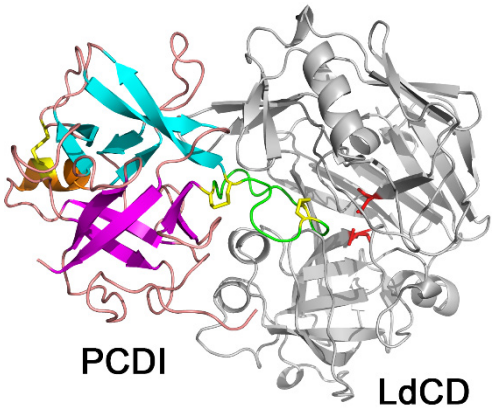
struktuře pepstatinu (zelená) jsou popsány jednotlivé pozice inhibitoru P4 až P3´(Iva označuje zbytek kyseliny izovalerové a Sta zbytek statinu), které se vážou do odpovídajících vazebných podmíst enzymu. Interagující aminokyselinové zbytky enzymu jsou označeny šedě a fialově (katalytické zbytky).

Tabulka 6: Seznam interakcí mezi pepstatinem a aktivním centrem LdCD.

Uvedeny jsou aminokyselinové zbytky v aktivním centru LdCD, které tvoří vodíkové vazby s inhibitorem (pepstatin), a vzdálenosti mezi příslušnými atomy tvořícími vodíkovou vazbu.

Pepstatin-A		**LdCD**		**vzdálenost (Å)**
zbytek	**Atom**	**zbytek**	**Atom**	
Val2	N	Ser227	OG	2,9
Val2	O	Ser227	N	3,0
Val3	N	Ser84	OG	2,9
Val3	O	Gly83	N	3,1
Val3	O	Ser84	N	2,9
Sta4	N	Gly225	O	3,1
Sta4	OH	Asp223	OD1	3,0
Sta4	OH	Asp37	OD1	2,6
Sta4	OH	Asp223	OD2	2,5
Sta4	O	Gly83	N	3,0
Ala5	N	Gly39	O	2,9
Ala5	O	Tyr198	OH	2,7

Analýza inhibiční specifity LdCD ukázala, že LdCD je efektivně inhibován Kunitzovým inhibitorem PCDI z lilku brambor (publikace č. 1). Pomocí dvou přístupů byla dále analyzována tato interakce a stanovena inhibiční hodnota IC_{50} = 0,11 µM v kinetickém aktivitním testu a hodnota K_D= 0,29 µM při měření pomocí MST termoforézy. Strukturní model interakce LdCD a PCDI byl navržen pomocí molekulárního modelování na základě prostorové struktury LdCD a prostorové struktury komplexu homologního enzymu IrCD1 s PCDI (kap. 4.2.1.3.) a je uveden na obrázku 15.



Obrázek 15: Model komplexu inhibitoru PCDI s aspartátovou proteasou LdCD.

Celkový pohled na model komplexu ve stužkovém znázornění. Molekula LdCD je znázorněna šedě, katalytické zbytky enzymu jsou zvýrazněny červeně. β struktury dvou hlavních domén inhibitoru PCDI

jsou barveny fialově a modře, a šroubovice je oranžově. Žlutě jsou vyznačeny cysteinové zbytky tvořící disulfidové můstky a zeleně smyčka, která interaguje s aktivním místem enzymu (označená jako smyčka 142-159). Model byl vytvořen na základě prostorových struktur enzymu LdCD a komplexu inhibitoru PCDI s homologní aspartátovou proteasou IrCD1 s využitím programu PyMol.

Na předložených dosud nepublikovaných výsledcích jsem se podílel (1) přípravou dvou rekombinantních katepsinů D v čisté podobě - IrCD1 a LdCD, (2) purifikací inhibitorů PCDI a PSPI z lilku brambor, (3) přípravou všech uvedených komplexů proteasa-inhibitor (LdCD-pepstatin, katepsin D-PCDI, trypsin-PCDI, trypsin-PSPI, trypsin/chymotrypsin-PSPI), (4) krystalizací komplexů proteasa-inhibitor metodou sedící a visící kapky, (5) grafickým zpracováním a interpretací dat z vyřešených prostorových struktur těchto komplexů.

4.3. Publikace č. 2: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids.

Souhrn

Působení bioaktivních sfingolipidů na proteolytické enzymy nebylo doposud detailně studováno. V této práci byla identifikována komplexní modulace enzymové aktivity lidského katepsinu D těmito molekulami. Jde o vysoce specifický typ regulační interakce, který nebyl nalezen u jiných modelových zástupců tříd proteolytických enzymů, včetně např. pepsinu, typického představitele A1 rodiny aspartátových proteas, tedy rodiny kam katepsin D patří.

Pomocí panelu přirozených a syntetických derivátů sfingolipidů byly charakterizovány vztahy mezi jejich strukturou a vlivem na aktivitu katepsinu D. Bylo zjištěno, že určité typy sfingosinu a ceramidu jsou účinné inhibitory katepsinu D v sub-mikromolární oblasti koncentrací. Pro nejlepší inhibitor C8 ceramin byla v aktivitním testu změřena hodnota IC_{50} = 0,14 μM. Pro fosforylované deriváty sfingolipidů typu sfingosin-1-fosfátu a ceramid-1-fosfátu byla nalezena naopak pozitivní modulace vedoucí k až několikanásobnému nárůstu aktivity katepsinu D.

Inhibiční i aktivační modulační efekty sfingolipidů jsou závislé na pH a nejvýrazněji se projevují v mírně kyselé oblasti. Přítomnost fosfosfingolipidů také způsobuje významný posun pH optima aktivity katepsinu D o jednu jednotku výše. Mechanismus inhibice pomocí sfingolipidů byl studován několika přístupy a prokázal přímou interakci s katepsinem D. Společné vazebné místo pro pro sfingosiny a ceramidy vykazuje kompetitivní typ inhibice. Lze předpokládat, že lipidové ligandy se budou alespoň částečně vázat do aktivního místa katepsinu D, jehož kritická vazebná podmísta preferují hydrofobní aminokyselinové zbytky substrátu.

Modulační efekty sfingolipidů na katepsin D testované *in vitro* se projevovaly za koncentrací, které jsou fyziologicky relevantní, a to i v přítomnosti liposomů simulujících prostředí biomembrán. Získané výsledky prokazují, že sfingolipidy jsou přirozené regulátory katepsinu D a jejich interakce může ovlivňovat procesy, jako jsou apoptóza a mitogeneze, jichž se tyto bioaktivní lipidy i proteasa účastní.

Můj podíl na práci zahrnoval (1) měření aktivit katepsinu D v přítomnosti vybraných sfingolipidů v aktivitním testu, (2) přípravu a analýzu vzorků pro HPLC analýzu modulace aktivity katepsinu D a (3) grafické zpracování dat pro manuskript.

Publikace č. 2

Žebrakovská I., Máša M., **Srp J.**, Horn M., Vávrová K., Mareš M.

Complex modulation of peptidolytic activity of cathepsin D by sphingolipids

Biochimica et Biophysica Acta, (12), 1097-1104 (2011)

Biochimica et Biophysica Acta 1811 (2011) 1097–1104

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbalip

Complex modulation of peptidolytic activity of cathepsin D by sphingolipids

Iva Žebrakovská [a,b], Martin Máša [a], Jaroslav Srp [a,b], Martin Horn [a], Kateřina Vávrová [c], Michael Mareš [a,*]

[a] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, 16610 Prague, Czech Republic
[b] Department of Biochemistry, Faculty of Science, Charles University, Albertov 6, 12843 Prague, Czech Republic
[c] Department of Inorganic and Organic Chemistry, Faculty of Pharmacy, Charles University, Heyrovskeho 1203, 50005 Hradec Kralove, Czech Republic

ARTICLE INFO

Article history:
Received 5 May 2011
Received in revised form 22 August 2011
Accepted 12 September 2011
Available online 23 September 2011

Keywords:
Sphingolipid
Phospholipid
Inhibition
Activation
Cathepsin D
Enzyme regulation

ABSTRACT

Cathepsin D is an aspartic peptidase involved in cellular processes including proliferation and apoptosis and implicated in human pathologies such as cancer and neurodegeneration. Our knowledge about the relationship between proteolysis and bioactive sphingolipids is still very limited. Here, we describe a complex pattern of modulation of the peptidolytic activity of cathepsin D by sphingolipids. A panel of sphingolipid derivatives was screened in a FRET-based assay; these molecules demonstrated negative or positive modulation of cathepsin D peptidolytic activity, depending on the sphingolipid structure. Certain sphingosines and ceramides inhibited cathepsin D in the submicromolar range, and structural requirements for this inhibitory effect were evaluated. The interaction of cathepsin D with sphingolipids was also demonstrated by fluorescence polarization measurements and determined to follow a competitive inhibition mode. In contrast, monoester phosphosphingolipids, especially ceramide-1-phosphate, were identified as activators of cathepsin D peptidolytic activity at submicromolar concentrations. Thus, sphingolipids and phosphosphingolipids, known to be antagonistic in their cell-signaling functions, displayed opposite modulation of cathepsin D. Sphingolipid-based modulators of cathepsin D are potentially involved in the control of cathepsin D-dependent processes and might serve as a scaffold for the development of novel regulators of this therapeutic target.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Sphingolipids (SLs) are potent signaling molecules capable of regulating vital cellular functions including cell growth, differentiation, senescence, apoptosis, and intracellular trafficking [1,2]. SLs are critically involved in tumor development and cancer cell death; *e.g.* ceramide has been recognized as a "tumor-suppressor lipid" based on its ability to block tumor initiation and metastasis [3,4]. Thus, SLs and their analogs represent a promising tool for manipulating the SL network in anticancer therapy. Despite recent advances in our understanding of SL metabolism and their roles in modulating cell responses, much less is known about SL-protein interactions and, in particular, about how SLs modulate the biochemical activity of target enzymes [5]. Our work focuses on cathepsin D (CatD), the only proteolytic enzyme that directly binds SLs reported to date [6].

CatD is an aspartic peptidase of general importance for cell growth and tissue homeostasis, as indicated by gene knockout studies [7]. It is localized in lysosomes and participates in the turnover of cellular proteins as well as in the selective processing of antigens, hormones, and growth factors. In cancer cells, CatD is overexpressed and secreted in the tumor microenvironment; it is associated with cancer development and progression through its complex proliferative and proteolytic action [8,9]. CatD is also an independent marker of poor prognosis in breast cancer [10,11]. During apoptosis, CatD is released into the cytosol, where it can interact with pro-apoptotic, anti-apoptotic, or nuclear proteins [8,9,12].

Evidence accumulated in the past decade indicates that there are multiple links between the physiological actions of CatD and bioactive SLs. Intracellular CatD zymogen interacts with prosaposin, a precursor of the sphingolipid activator proteins (saposins A-D), and this interaction supports CatD targeting and activation as well as prosaposin processing [13]. Extracellular CatD zymogen can be directed to lipid rafts [14], cell membrane structures that are generally regulated by SLs [15], by the CatD-binding protein LRP1. There is a functional link between production of intracellular ceramide by endolysosomal acid sphingomyelinase and CatD maturation [6,16]. Ceramide levels have been shown to regulate CatD maturation and transport in apoptosis models [16,17]. Heinrich et al. reported that SLs can induce CatD-mediated proteolysis [6,16]. Importantly, experiments using SL-affinity chromatography and SL-based photoreactive probes determined that CatD directly binds SLs such as sphingosine and ceramide [6].

In the present work, we performed the first detailed enzymological analysis of the interaction between SLs and CatD with regard to its impact on the peptidolytic activity of CatD. We show that specific SLs and

Abbreviations: CatD, cathepsin D; FRET, fluorescence resonance energy transfer; PSL, phosphosphingolipid; SL, sphingolipid

* Corresponding author at: Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, 16610 Prague 6, Czech Republic. Tel.: +420 220183358; fax: +420 220183578.
E-mail addresses: zebrakovska@uochb.cas.cz (I. Žebrakovská), mosha@atlas.cz (M. Máša), srp@uochb.cas.cz (J. Srp), horn@uochb.cas.cz (M. Horn), katerina.vavrova@faf.cuni.cz (K. Vávrová), mares@uochb.cas.cz (M. Mareš).

1388-1981/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbalip.2011.09.005

phosphosphingolipids (PSLs) are complex modulators of CatD capable of inhibiting and activating CatD, respectively. No endogenous mammalian CatD inhibitors are known and SLs therefore represent unique inhibitory regulators of this enzyme. We suppose that the SL-PSL network modulating CatD might be relevant in a physiological context.

2. Materials and methods

2.1. Materials

Human CatD was purified to homogeneity from placenta [18]. Pepsin and papain were from Sigma-Aldrich (MO, USA), and trypsin was from Serva (Germany). Specific peptidic substrates for CatD, namely the FRET (fluorescence resonance energy transfer)-based substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu (Abz,1-aminobenzoic acid; Nph, 4-nitrophenylalanine) and the CatD propeptide-derived substrate [15p-36p], were designed and synthesized as previously described [19]. Z-Phe-Arg-AMC substrate (AMC, 7-amino-4-methylcoumarin) was from Bachem (Switzerland). Lipids (see compound numbers in Tables 1 and S1) were obtained from the following sources: Sigma-Aldrich: **1**, **3**, **4**, **6**, **8**, **12**, **13**, **16**–**19**, **21**, **26**–**29**, **36**–**38**; Cayman Chemical Company (MI, USA): **2**, **11**, **33**–**35**; Avanti Polar Lipids (AL, USA): **10**, **22**; Invitrogen (CA, USA): **20**; and Matreya (PA, USA): L-*threo* stereoisomer of **12**. Compounds **5**, **7**, **9**, **14**, **15**, **23**–**25**, **30**–**32**, and L-*threo*-stereoisomer of **1** were prepared as previously described [20–22]. Heparin (18 kDa) and pepstatin A were from Sigma-Aldrich. Unilamellar liposomes (100 nm in diameter) were prepared from 1,2-dimyristoyl-phosphocholine (Avanti Polar Lipids) in water according to the manufacturer's protocol.

2.2. FRET-based assay of CatD activity

The peptidolytic activity of CatD and its modulation by lipids was assayed with the FRET-based optimized CatD substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu [19]. The reaction was performed in microplate format in a total assay volume of 100 µl. The reaction mixture typically contained 2 nM CatD and 0–200 µM lipid in 0.1 M sodium acetate, pH 4.2 (or pH 3.0–6.0 where indicated), containing 0.1% polyethylene glycol 1500. Where indicated, heparin or salts were added to the reaction mixture to give a final concentration of 750 µg heparin/ml or 1–250 mM NaCl, $CaCl_2$, or $MgCl_2$. In liposome experiments, 100 nm liposomes were added to the reaction mixture to give a final concentration of 100 µM. The mixture was incubated at 37 °C for 10 min, followed by the addition of substrate to a final concentration of 20 µM. The kinetics of product release were continuously monitored in an Infinite M200 microplate reader (Tecan) at excitation and emission wavelengths of 330 and 410 nm, respectively. Lipid stock solutions (10 mM) were prepared in dimethyl sulfoxide (alternatively in acetonitrile, methanol, or ethanol); the solvent concentration in the assay was 2%. IC_{50} values were determined from the residual velocities using the dose–response plot (v_i/v_0 vs [I]) analyzed by Grafit (Erithacus software, UK). The kinetic parameters (K_m and k_{cat}) and inhibition mode were determined using the same activity assay with substrate concentrations up to 400 µM. The initial velocities of product release were analyzed by non-linear regression (using Grafit) and by a Lineweaver-Burk plot. The reversibility of CatD inhibition by sphingosine was tested in a dilution experiment. CatD (20 nM) was preincubated (10 min at 37 °C) in 0.1 M sodium acetate, pH 4.2, with 0.4 or 4 µM sphingosine (corresponding to ~IC_{50} and 10 × IC_{50}, respectively). The mixture was then diluted 10-fold with the same buffer containing 0.4 µM sphingosine (for IC_{50} preincubation concentration) or without inhibitor (for 10 × IC_{50} preincubation concentration) and incubated for 10 min at 37 °C. The final activity was then measured with the FRET substrate as described above. All activity measurements were performed in triplicate.

Table 1
Modulation effects of sphingolipids and related compounds on the peptidolytic activity of CatD[a].

Comp. no.	Abbreviation	Inhibition	Activation
		IC_{50} (µM) ± SE	Fold increase ± SE
	Sphingosine derivatives		
1	Sphingosine	0.39 ± 0.04	n.e.
2	*N,N*-dimethylsphingosine	0.45 ± 0.11	n.e.
3	*N,N,N*-trimethylsphingosine	0.16 ± 0.01	n.e.
4	Azidosphingosine	n.e.	n.e.
5	C12-sphingosine	>200	n.e.
6	Dihydrosphingosine (sphinganine)	0.41 ± 0.03	n.e.
7	C20-sphingosine	0.22 ± 0.02	n.e.
8	Sphingosine-1-phosphate	n.e.	2.0 ± 0.1
9	(*2S,3R*)-2-aminoicos-4-yne-1,3-diol (triple bond C20-sphingosine)	0.27 ± 0.03	n.e.
10	Phytosphingosine	0.30 ± 0.04	n.e.
11	C8-ceramine (*N*-octylsphingosine)	0.14 ± 0.03	n.e.
	Ceramide derivatives		
12	C6-ceramide (C18 base)	1.3 ± 0.1	n.e.
13	C6-dihydroceramide (C18 base)	n.e.	n.e.
14	C2-ceramide (C12 base)	n.e.	n.e.
15	C6-ceramide (C12 base)	69.1 ± 8.5	n.e.
16	C2-ceramide (C18 base)	0.31 ± 0.03	n.e.
17	C8-ceramide-1-phosphate	n.e.	5.0 ± 0.1
18	1-*O*- β-D-galactosyl-ceramide	61.7 ± 8.1	n.e.
19	Sphingomyelin	n.e.	n.e.
20	BODIPY FL C5-ceramide	11.7 ± 1.9	n.e.
21	C18-ceramide-1-phosphate	n.e.	3.4 ± 0.9
22	C6-pyridinium-ceramide	0.41 ± 0.06	n.e.
23	C24-NBD-ceramide	n.e.	n.e.
24	C12-NBD-ceramide	4.3 ± 0.7	n.e.
25	C6-NBD-ceramide	6.9 ± 0.7	n.e.
26	3-*O*-benzoyl-C16-ceramide	n.e.	n.e.
	Others		
27	1,2-dipalmitoyl-glycerol	n.e.	n.e.
28	(*1R,2R*)-2-amino-1-phenyl-1,3-propanediol	n.e.	n.e.
29	(*1S,2S*)-2-amino-1-phenyl-1,3-propanediol	n.e.	n.e.
30	*N*-dodecanoylglycine dodecyl ester	n.e.	n.e.
31	*N*-dodecanoyl-L-serine dodecyl ester	n.e.	n.e.
32	*N*-oleoyl-L-serine dodecyl ester	n.e.	n.e.
33	Prostaglandin E1	n.e.	n.e.
34	Myriocin	n.e.	n.e.
35	(*1R,2R*)-2-myristylamino-1-(4-nitrophenyl)-1,3-propandiol	n.e.	n.e.
36	Phosphatidic acid	n.e.	4.9 ± 0.2
37	Phosphatidylcholine	n.e.	n.e.
38	Phosphatidylethanolamine	n.e.	n.e.

[a] CatD activity was measured with a FRET-based octapeptide substrate at pH 4.2 in the presence of the tested compounds. Structures of the compounds are given in Supplemental Table S1. The inhibitory effect was evaluated by determining the IC_{50} value. The activation effect was expressed as relative fold increase of CatD activity induced by 10 µM compound compared to the non-activated control. n.e. denotes no modulation effect with 200 µM compound. Mean values ± SE are given.

2.3. Activity assays for other proteases

The peptidolytic activities of trypsin and papain were assayed with the fluorogenic substrate Z-Phe-Arg-AMC. The reaction mixture contained 60 pM trypsin or 10 pM papain in 0.1 M Tris–HCl, pH 7.0, containing 0.1% polyethylene glycol 1500, or 0.1 M bis-Tris–HCl, pH 6.0, containing 2.5 mM dithiothreitol and 0.1% polyethylene glycol 1500, respectively. Where indicated, lipids were added to the reaction mixture to give a final concentration of 1 µM or 10 µM. The mixture was incubated at 37 °C for 10 min, followed by the addition of

50 μM Z-Phe-Arg-AMC. The instrumentation and measurement were as detailed in Section 2.2, except that the excitation and emission wavelengths were 360 and 465 nm, respectively. The peptidolytic activity of pepsin was determined as described for CatD (Section 2.2) with the following modifications: the reaction mixture contained 0.2 nM pepsin and was buffered with 0.1 M sodium acetate, pH 4.0, containing 0.1% polyethylene glycol 1500.

2.4. Fragmentation of peptide [15p-36p] by CatD

The reaction mixtures, composed of 50 μM peptide [15p-36p] and 1.25 nM CatD with 0–10 μM C8-ceramide-1-phosphate or 3.75 nM CatD with 0–10 μM sphingosine, were incubated in 0.1 M sodium acetate, pH 4.2, at 37 °C for 12 h. The reactions were stopped by adding trifluoroacetic acid (TFA) to a final concentration of 0.5% (v/v). The mixtures were separated by RP-HPLC using an Agilent 1200 Series system (Agilent Technologies) on a Kromasil C18 column (250 × 4.6 mm, 5 μm; Alltech) equilibrated in 0.1% (v/v) TFA and eluted with a linear 1.5%/min gradient of 99% (v/v) acetonitrile in 0.1% (v/v) TFA. The collected peak fractions were analyzed by ESI mass spectrometry on an LCQ Classic Finnigan Mat (Thermo Finnigan).

2.5. Fluorescence polarization measurements

Fluorescence polarization measurements were performed with an Aminco-Bowman Series 2 luminescence spectrophotometer (Thermo Electron) at 37 °C using excitation and emission wavelengths of 505 and 514 nm, respectively. Anisotropy titration was carried out in competition experiments with 0.001–40 μM sphingosine, 2 μM BODIPY FL C5-ceramide (Invitrogen) as a fluorescent "reporter" probe, and 50 nM CatD in 0.1 M sodium acetate, pH 5.5. The binding mixture was incubated for 45 min at 37 °C to reach equilibrium. The fluorescence of the buffer alone was subtracted from each measurement, and the anisotropy was calculated according to [23]. G-factor correction was performed to eliminate instrumental bias.

The anisotropy data were converted to the fraction of the probe bound using the following equation:

$$f_b = \frac{r_{obs} - r_0}{(r_{obs} - r_0) + R(r_{\max} - r_{obs})}$$

where f_b is the fraction of the probe bound and r_{obs}, r_0, and r_{max} are the observed anisotropy, anisotropy of the free probe, and anisotropy of the probe with CatD, respectively. R is the fluorescence intensity of the probe with CatD divided by the intensity of the free probe. The K_d of the binding of sphingosine to CatD was determined by nonlinear least-squares fit to a single-site binding model using Grafit software.

2.6. Analytical methods

The stock concentration of CatD was determined by titration with pepstatin A using the FRET-based assay. The stock solutions of peptide substrates were quantified by amino acid analysis. The stock solutions of SLs were quantified according to [24]; the acylated SLs were subjected to acidic hydrolysis prior to analysis.

3. Results

3.1. Structure-activity profiling of SLs reveals complex modulation of CatD activity

A panel of SL derivatives and structurally related compounds was screened for the ability to influence the peptidolytic activity of human CatD (Tables 1 and S1). An *in vitro* FRET-based assay with the fluorogenic peptidic substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu [19] was employed to measure CatD activity in the presence of SLs. The results presented in Table 1 show that these compounds display a complex pattern of modulation of CatD activity, including inhibition or activation; some derivatives had no significant effect.

The inhibitory compounds were evaluated by determining their IC_{50} values, which generally ranged from submicromolar to micromolar values (Table 1). D-*erythro*-sphingosine and C6-ceramide (compounds **1** and **12**, Table 1) were used as model SL structures; these compounds were compared with other derivatives to assess the structural requirements for inhibitory interaction. The naturally occurring D-*erythro*-sphingosine with C18 chain length (**1**) inhibited CatD with an IC_{50} of 0.39 μM. All the other sphingoid bases, including sphingosines with various chain lengths (**5**, **7**), *N*-alkylsphingosines (**2**, **3**, **11**), dihydrosphingosine (**6**), phytosphingosine (**10**), and a triple bond-containing analog (**9**) were able to inhibit CatD to different extents, with the only exception being an azidosphingosine analog (**4**). Thus, the inhibition was independent of the presence of the C4 *trans*-double bond (compare **1**, **6**, **10,** and **9**). The gradual shortening of the sphingosine chain length from C20 (**7**) to C12 (**5**) dramatically reduced the CatD inhibition (IC_{50} from 0.22 to >200 μM). The importance of the presence of an alkyl chain of a certain length was further confirmed by the lack of activity of **28** and **29**, which possess a phenyl ring instead of the pentadecenyl chain.

In the ceramide panel, C6-ceramide (with natural C18 sphingosine base, **12**) inhibited CatD with an IC_{50} of 1.4 μM, and the effect was abrogated after saturation of the C4 *trans*-double bond (**13**). Thus, an intact allylic hydroxyl in ceramides is critical for inhibition of CatD. This was also confirmed by the inactivity of **26**, which has the allylic hydroxyl masked with benzoyl group, and of **31** and **32**, in which the allylic hydroxyl moiety was replaced by an ester bond. Shortening of the acyl chain to C2-ceramide (**16**) markedly increased its inhibition potency (IC_{50} = 0.31 μM). On the other hand, the same change in the ceramide with the base chain truncated to C12 (**15** and **14**) resulted in a loss of inhibitory potency. Introduction of a bulky fluorescent label at the end of the acyl chain, *e.g.* in Bodipy-C5-ceramide (**20**) or the NBD-ceramides **23**–**25**, generally decreased the inhibitory effect towards CatD. Nevertheless, a similar acyl chain modification with pyridinium salt (highly soluble compound **22**) yielded a more effective inhibitor compared to the unmodified ceramide **12**.

The stereospecificity of SL-mediated inhibition of CatD was probed using unnatural L-*threo* stereoisomers of the natural D-*erythro* structures of sphingosine (**1**) and C6-ceramide (**12**). The IC_{50} values were rather similar for the sphingosine isomers (0.48 ± 0.06 μM for L-*threo* compared with 0.39 ± 0.04 μM for D-*erythro*), but differed to a

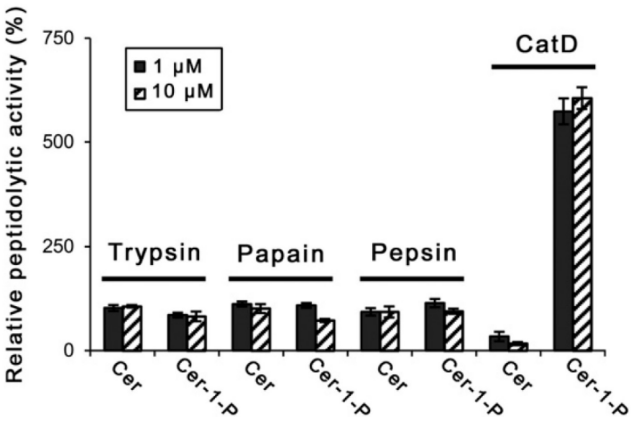


Fig. 1. Comparison of modulatory effects of sphingolipids on CatD and representative proteases. Activities of trypsin, papain, pepsin, and CatD were determined in the presence of 1 μM and 10 μM concentrations of C6-ceramide (Cer) and C8-ceramide-1-phosphate (Cer-1-P). The peptidolytic activities of pepsin and CatD were measured with a FRET-peptide substrate at pH 5.0 and 4.0, respectively, and those of papain and trypsin with a fluorogenic AMC-peptide substrate at pH 6.0 and 7.0, respectively. For each protease, mean values ± SD are expressed relative to the control experiment without lipid (100%).

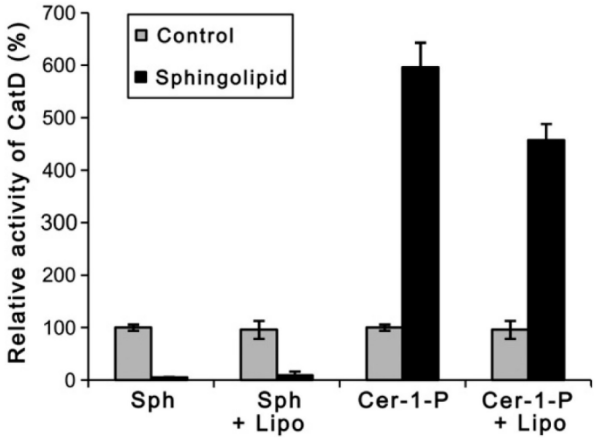


Fig. 2. Modulatory effects of sphingolipids on CatD activity in the presence of liposomes. The peptidolytic activity of CatD was measured in a FRET-based assay at pH 5.0 in the presence of 10 μM sphingosine (Sph) or 10 μM C8-ceramide-1-phosphate (Cer-1-P). The experiment was performed in the presence or absence of 100 μM liposomes (Lipo). The sphingosine-mediated inhibition and C8-ceramide-1-phosphate-mediated activation are expressed relative to the untreated control without sphingolipids and liposomes (100%). Mean values ± SD are given.

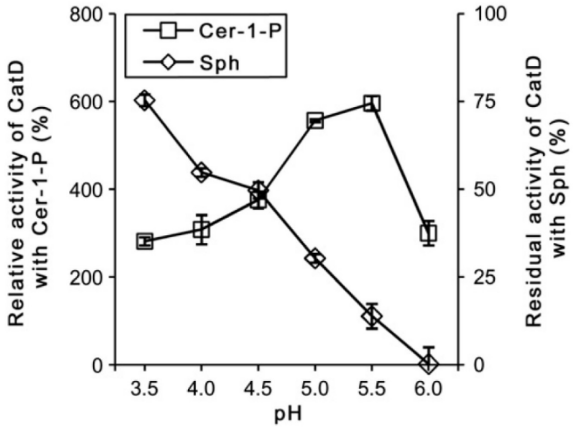


Fig. 4. Modulatory effects of sphingolipids on CatD activity are pH dependent. The peptidolytic activity of CatD was measured in a FRET-based assay in the presence of 0.4 μM sphingosine (Sph) or 1 μM C8-ceramide-1-phosphate (Cer-1-P) and compared with the untreated controls. For C8-ceramide-1-phosphate-mediated activation (left axis), the values are expressed relative to the untreated controls at the same pH (100%). For sphingosine-mediated inhibition (right axis), the values are expressed as the remaining activity relative to the untreated controls at the same pH (100%). Mean values ± SD are given.

greater extent for the C6-ceramide isomers (3.3 ± 0.3 μM for L-*threo* compared with 1.3 ± 0.1 μM for D-*erythro*).

SLs that potentiated CatD activity were evaluated by determining their activation effect at a 10 μM concentration (Table 1). A high effect was observed for short-chain C8-ceramide-1-phosphate (**17**) and naturally occurring long-chain ceramide-1-phosphate (with prevailing C18 acyl, **21**). Compounds **17** and **21** induced the activity of CatD by about 5-fold and 3.5-fold, respectively. In contrast, only ~2-fold induction was found for sphingosine-1-phosphate (**8**). The phosphate monoester group of these PSLs was essential for the activation effect; sphingomyelin **19** (with choline linked to the phosphate) was inactive, and cerebroside **18** (with the phosphate replaced with a monosaccharide) was weakly inhibitory. Interestingly, the activation effect was also found for phosphatidic acid (**36**) but not the phosphate diesters phosphatidylcholine (**37**) and phosphatidylethanolamine (**38**).

The modulation of CatD activity by SLs and PSLs was compared with the action of the lipids on other peptidases. The selected enzymes pepsin, papain, and trypsin are typical representatives of the aspartic, cysteine, and serine classes of peptidases, respectively. Their peptidolytic activities were measured in the presence of C6-ceramide (a CatD inhibitor) and C8-ceramide-1-phosphate (a CatD activator). Fig. 1 shows that these lipids do not have pronounced effects on the tested peptidases, in contrast to their modulatory effects on CatD. This demonstrates the specificity of the SL/PSL-mediated modulation of CatD activity.

Finally, we investigated the modulatory effects of SLs and PSLs in the presence of liposomes to simulate the environment of biological membranes. Fig. 2 shows that CatD was efficiently inhibited by sphingosine and was activated by ceramide-1-phosphate at a 10-fold molar excess of unilamellar liposomes over lipid. This experiment strongly suggests that CatD is capable of functional interaction with SLs/PSLs in membrane-containing systems.

3.2. Inhibitory interaction of SLs with CatD and its mechanism

The inhibitory interaction of SLs with CatD was characterized in detail. The inhibition curves of the peptidolytic activity of CatD measured with the FRET octapeptide substrate are given in Fig. 3A; sphingosine (**1**) and C6-ceramide (**12**) exerted concentration-dependent inhibition of CatD in the submicromolar and low micromolar range, respectively (Table 1). Moreover, their inhibition potency was influenced by pH. The pH profile of the most inhibitory natural SL, sphingosine, demonstrated that the extent of its CatD inhibition substantially

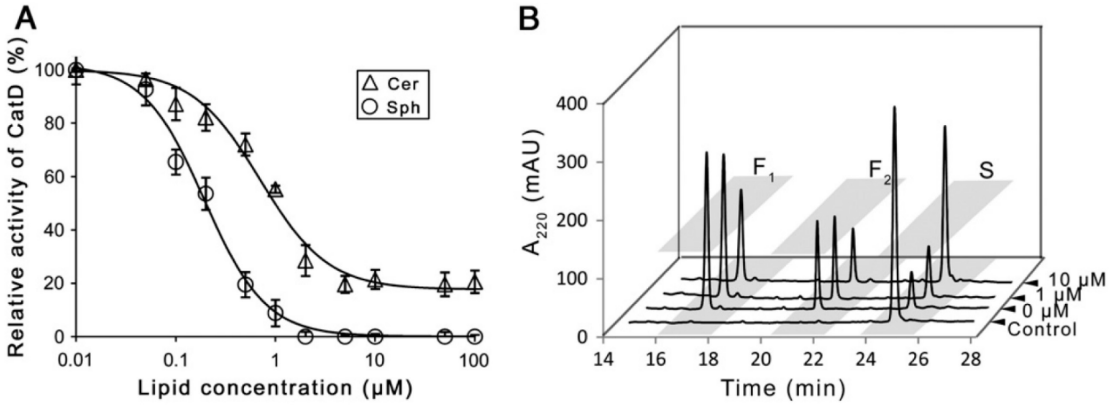


Fig. 3. Inhibition of the peptidolytic activity of CatD by sphingolipids. (A) Inhibition of CatD with varied concentrations of sphingosine (Sph; $IC_{50} = 0.19$ μM) and C6-ceramide **12** (Cer; $IC_{50} = 1.0$ μM). CatD activity was measured with a FRET-based octapeptide substrate at pH 5.5. Mean values ± SD are expressed as the remaining activity relative to the uninhibited control (100%). (B) The cleavage efficiency of CatD was inhibited in the presence of sphingosine. Synthetic peptide [15p-36p] (derived from the CatD propeptide) was digested by CatD at pH 4.2, and the reaction mixture was separated by RP-HPLC. The RP-HPLC elution profile was monitored by measuring the absorbance at 220 nm; the positions of substrate (*S*: [15p-36p]) and its cleavage products (fragments F_1: [27p-36p] and F_2: [15p-26p]) are indicated. The profiles are marked with the sphingosine concentration (1 μM; 10 μM; 0 μM, digest in the absence of sphingosine; Control, uncleaved substrate).

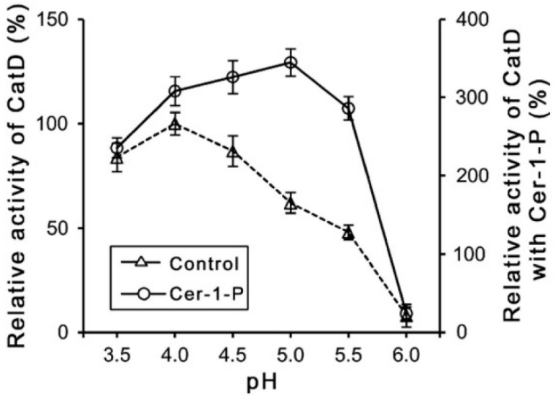


Fig. 5. pH profile of CatD determined in the absence (Control, left axis) and presence of 1 μM C8-ceramide-1-phosphate (Cer-1-P, right axis). The peptidolytic activity of CatD was measured in a FRET-based assay. Mean values ± SD are given normalized to the maximum value measured without C8-ceramide-1-phosphate (100%).

increased at higher pH values (Fig. 4). For comparison, the pH profile of CatD with a pH optimum at 4.2 is provided in Fig. 5. The reversibility of the sphingosine inhibition was examined by a dilution experiment, which showed that the activity of the mixture of CatD and sphingosine depends on the final concentration of sphingosine regardless of its preincubation concentration (see Materials and methods).

In addition to the FRET assay, we introduced an assay system based on an unmodified peptide substrate for testing the modulation of peptidolytic activity of CatD by SLs. The substrate was designed based on the structure of the CatD propeptide that is autocatalytically cleaved at the Leu26p-Ile27p bond during CatD activation [19]. The synthetic 22-mer peptide fragment [15p-36p] spanning the cleavage site was specifically hydrolyzed by CatD at the native position, and the reaction was monitored by separation of the cleavage products by HPLC and their identification by mass spectrometry (Fig. 3B). Using this assay, we demonstrated that the incubation of CatD in the presence of various concentrations of sphingosine inhibited cleavage of the peptide [15p-36p] in a SL concentration-dependent manner (Fig. 3B).

Kinetic analysis was performed to determine the mode of inhibition of CatD by SLs. The hydrolysis of the FRET peptidic substrate followed standard Michaelis-Menten kinetics with a K_m of 59 μM. The initial rates of product formation in the presence of C8-ceramine (**11**), one of the most inhibitory SLs tested, are shown in a Lineweaver-Burk plot (Fig. 6A). The family of straight lines intersects on the 1/v axis, indicating competitive inhibition (K_i = 0.07 μM).

Furthermore, a fluorescence polarization experiment was used to analyze the interaction of SLs with CatD; this approach allows for measurement of the inhibitor affinity independent of enzyme activity. As a fluorescent reporter probe, we employed C5-ceramide labeled with a BODIPY moiety (**20**), which inhibits CatD with an IC_{50} of 11.7 μM (Table 1). The probe was used as a primary CatD ligand in a binding experiment monitored by fluorescence anisotropy. The measurement was performed as a competition assay between the probe and sphingosine to yield an equilibrium dissociation constant (K_d) of 2.0 μM for the sphingosine-CatD complex (Fig. 6B). This experiment directly demonstrated the binding of SLs to CatD and indicated that sphingosine and ceramide-type inhibitors share the same binding site on the CatD molecule.

Finally, we investigated how the SL-mediated inhibition of CatD is influenced by a PSL activator. We compared the inhibition curves of sphingosine in the presence and absence of 10 μM C8-ceramide-1-phosphate at pH 5.0 using the FRET assay. The determined IC_{50} values were 0.32 ± 0.7 μM for sphingosine alone and 2.1 ± 0.7 μM for sphingosine combined with C8-ceramide-1-phosphate. This demonstrated that the apparent inhibitory effect was significantly reduced by the activator, indicating that SLs and PSLs are able to concurrently control the net activity of CatD.

3.3. CatD is activated by PSLs and this interaction is modulated by pH, ions, and heparin

PSLs without further substitution of the phosphate group induced the peptidolytic activity of CatD (Table 1). The activation curves of CatD measured with the FRET octapeptide substrate are shown in Fig. 7A. The induction of CatD activity increased with increasing PSL concentration of as shown for C8-ceramide-1-phosphate (**17**) and sphingosine-1-phosphate (**8**), and reached a maximum in the 100 micromolar range. The half-maximal induction was obtained with submicromolar PSL concentrations. Interestingly, the activation changed the pH profile of CatD peptidolytic activity (Fig. 4). In the presence of C8-ceramide-1-phosphate, the most activating PSL, the pH optimum of CatD was shifted about 1 pH unit higher to pH 5.0 (Fig. 5). Furthermore, an HPLC assay with the [15p-36p] peptidic substrate was employed to analyze the activation in an alternative assay system. It clearly demonstrated that incubation in the presence of C8-ceramide-1-phosphate augmented the cleavage efficiency of CatD in a PSL concentration-dependent manner (Fig. 7B).

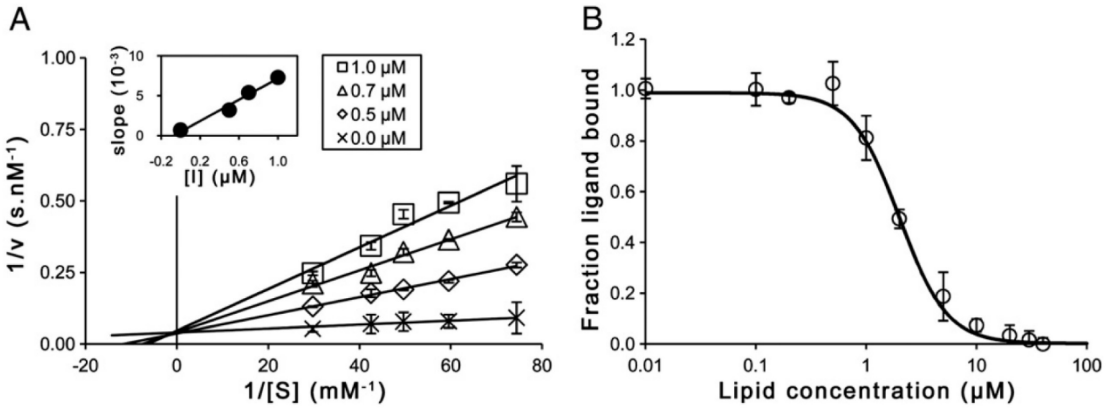


Fig. 6. Mechanism of the interaction of inhibitory sphingolipids with CatD. (A) The mode of inhibition of the peptidolytic activity of CatD by C8-ceramine was determined to be competitive using a FRET-based assay at pH 4.2. CatD activity was measured with varied concentrations of substrate (S) and inhibitor (I, concentrations are indicated). Lineweaver-Burk plot is presented together with a secondary plot of the same data (insert); K_i = 0.07 μM. Mean values ± SD are given. (B) Fluorescence polarization assay of the binding of sphingosine to CatD at pH 5.5 plotted as a fraction of the ligand bound. Competitive displacement of the fluorescent probe BODIPY FL C5-ceramide (2 μM) complexed with CatD (50 nM) by sphingosine (concentration indicated) revealed the optimized parameter K_d = 2.0 μM for the sphingosine-CatD complex.

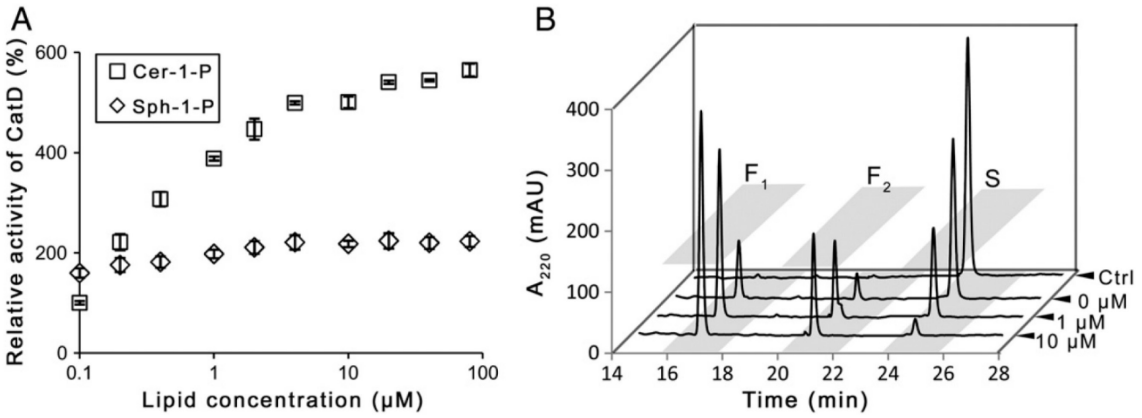


Fig. 7. Phosphosphingolipids induce the peptidolytic activity of CatD. (A) Activation of CatD with varied concentrations of sphingosine-1-phosphate (Sph-1-P) and C8-ceramide-1-phosphate (Cer-1-P). CatD activity was measured in a fluorescence assay with a specific FRET octapeptide substrate at pH 4.2. Mean values ± SD are expressed relative to the non-activated control (100%). (B) The cleavage efficiency of CatD was induced in the presence of C8-ceramide-1-phosphate. Synthetic peptide [15p-36p] (derived from the CatD pro-peptide) was digested by CatD at pH 4.2, and the reaction mixture was separated by RP-HPLC. The RP-HPLC elution profile was monitored by measuring absorbance at 220 nm; the positions of substrate (S: [15p-36p]) and its cleavage products (fragments F_1: [27p-36p] and F_2: [15p-26p]) are indicated. The profiles are marked with the C8-ceramide-1-phosphate concentration (1 μM; 10 μM; 0 μM, digest in the absence of C8-ceramide-1-phosphate; Control, uncleaved substrate).

To describe the activation mechanism in enzymological terms, we determined the kinetic parameters of CatD with the FRET substrate for the non-induced system and system induced with 10 μM C8-ceramide-1-phosphate. We obtained the following values: $K_m = 59 \pm 9$ μM, $k_{cat} = 54 \pm 6\ s^{-1}$, $k_{cat}/K_m = 910\ mM^{-1}.s^{-1}$ (non-induced), and $K_m = 4 \pm 0.7$ μM, $k_{cat} = 108 \pm 10\ s^{-1}$, $k_{cat}/K_m = 26967\ mM^{-1}.s^{-1}$ (induced) at pH 4.2. These data indicated that PSL increased the catalytic efficiency (k_{cat}/K_m) of CatD about 30-fold due to one order of magnitude better substrate-binding affinity (K_m) and, to a lesser extent, a faster turnover rate (k_{cat}).

Finally, we investigated the interaction between CatD and PSLs in the presence of ions and heparin. In Fig. 8A, CatD was activated by C8-ceramide-1-phosphate, and the effect of various concentrations of NaCl, $CaCl_2$, and $MgCl_2$ was compared. The PSL-mediated activation of CatD was suppressed by Ca^{2+} and Mg^{2+} by about 60% of the induced activity (Fig. 8A), although the tested salt concentrations did not significantly influence CatD activity in the absence of PSL (data not shown). The suppression was concentration-dependent, and Mg^{2+} exerted its effect at lower concentrations than Ca^{2+}. Furthermore, the combined action of C8-ceramide-1-phosphate and heparin on CatD was analyzed (Fig. 8B). Heparin is a natural sulfated glycosaminoglycan that activates CatD [25]. Here, it was applied at a saturating concentration that produced ~3-fold activation of CatD. The experiment showed that the induction effects mediated by both activators are additive at lower PSL concentrations and non-additive at saturating PSL concentrations.

4. Discussion

In this work, we discovered that SLs are complex modulators of the peptidolytic activity of CatD. SLs are capable of inhibiting or activating CatD, depending on their structure. First, a panel of SL derivatives was analyzed to identify the structural variables that determine their modulatory effects on CatD activity. Second, we investigated the mechanism of the interaction of SLs with CatD and factors that control this process. The effect of SLs was independently demonstrated in two different CatD activity assay systems using specific peptidic substrates applied in FRET-based and HPLC-based assays.

Sphingosine- and ceramide-type SLs exert inhibitory effects on CatD activity. The inhibitory activity of the sphingosine backbone was found to be relatively tolerant to structural modification. Of particular interest is the apparent unimportance of the allylic hydroxyl moiety; similar inhibitory activity was observed in sphingoid bases with this *trans*-double bond either saturated or substituted by a triple

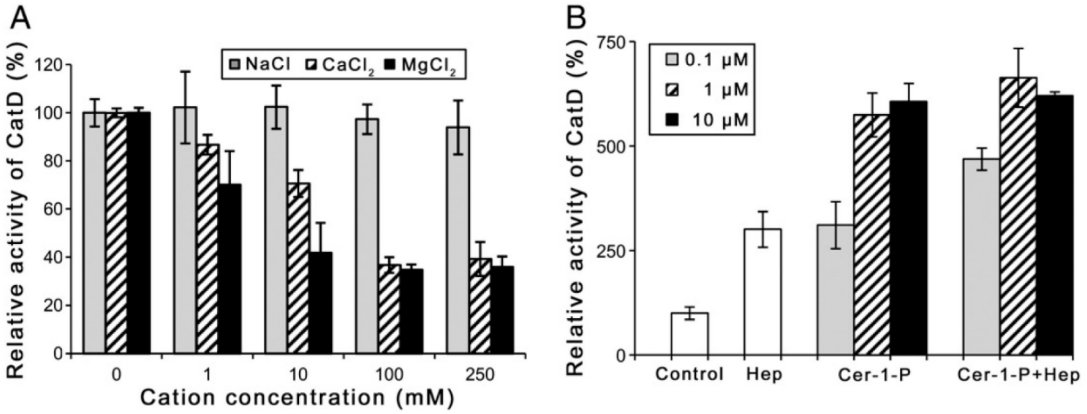


Fig. 8. Induction of CatD by C8-ceramide-1-phosphate in the presence of ions and heparin. The peptidolytic activity of CatD was measured in a FRET-based assay at pH 5.0; mean values ± SD are given. (A) CatD activity was induced by 1 μM C8-ceramide-1-phosphate in the presence of varying concentrations of NaCl, $CaCl_2$, and $MgCl_2$. The activity values are expressed relative to the controls induced in the absence of salts (100%). (B) CatD activity was induced by 0.1–10 μM C8-ceramide-1-phosphate (Cer-1-P) or 750 μg/ml heparin (Hep) or by the combined action of both compounds (Cer-1-P + Hep). The activity values are expressed relative to the non-induced control (100%).

bond. Interestingly, similar structural tolerance has been observed in the anticancer effects of the sphingoid bases [26]. On the other hand, ceramides, *i.e. N*-acylsphingosines, were more sensitive to structural (including sterochemical) alterations than the sphingoid bases. In particular, the allylic hydroxyl was found essential for the inhibitory effect of ceramides on CatD activity. This structural requirement of ceramides has also been reported for their ability to induce apoptosis [27]. Furthermore, the results revealed that the base chain length and the acyl chain length are important factors determining the ability of sphingosines and ceramides to inhibit CatD. Our data indicated that certain optimized chain lengths are necessary for effective binding to CatD, suggesting the importance of hydrophobic interactions in this process. Finally, we hypothesize that the (partial) positive charge on the nitrogen linked to the sphingoid C2 atom is important for the interaction of sphingolipids with CatD; this is suggested by the inactivity of an azido analog (**4**) compared to sphingosines and ceramides containing (alkyl)amino and amido moieties.

To study the mechanism of the inhibitory interaction of SLs with CatD, several approaches were applied. Fluorescence polarization measurements demonstrated that the SL inhibitors directly bind to the CatD molecule and that there is a common binding site for sphingosines and ceramides. Dilution experiments showed the reversibility of SL binding, and kinetic analysis indicated a competitive mode of inhibition of CatD by SLs. Using SDS PAGE analysis, we also found that the effect of SL-mediated inhibition is not associated with an autoproteolytic degradation of CatD under assay conditions (data not shown). Taken together, our results suggested that the SL inhibitors bind at the active site of CatD and compete with a peptidic substrate for its binding site. However, based on the obtained data and the inspection of the CatD crystallographic model (PDB code 1LYB), the exact position of this binding site cannot be unambiguously predicted. The active site of CatD contains subsites that accommodate hydrophobic parts of substrates, which is reflected in the enzyme cleavage specificity [28]. This was also demonstrated by the peptidic substrates of CatD used in this study, which were cleaved between the hydrophobic amino acids at the bonds Phe-Nph and Leu-Ile (in the FRET and [15p-36p] substrates, respectively). Therefore, we may assume that SLs as amphiphilic ligands can occupy, at least in part, such binding subsites. Nevertheless, to fully understand the binding mode of SLs to CatD, the crystal structure of a SL-CatD complex will be required. To that end, we are pursuing the co-crystallization of CatD with the most potent SL inhibitors.

In contrast to the SLs, specific PSLs were found to activate CatD. In particular, phosphorylation of the primary hydroxyl of both sphingosine and ceramide resulted in induction of the peptidolytic activity of CatD. However, the activation effect was also found for phosphatidic acid, which suggested that the induction of CatD activity could be attributed more generally to phospholipids with an unsubstituted monoester phosphate moiety. This is in line with an earlier finding that acidic lipids can enhance CatD proteolysis [29]. Using enzyme kinetics, we analyzed the activation effect for C8-ceramide-1-phosphate, the most activating compound, and found that the activation is based predominantly on the improved substrate-binding affinity. Furthermore, the activation mediated by C8-ceramide-1-phosphate was about two-fold higher than that mediated by heparin, a sulfated polysaccharide that has been reported to activate CatD [25]. The analysis of the combined action of C8-ceramide-1-phosphate and heparin suggested that these negatively-charged activators might interact with different binding sites on the CatD molecule. The exact binding site for heparin is unknown, but this interaction most likely occurs through positively-charged surface region(s) of CatD, as similarly reported for cysteine cathepsins [30,31]. Indeed, the CatD crystallographic model shows that there are four basic patches in the vicinity of the active site, and we hypothesize that some of these candidate sites are utilized for interaction with PSLs. This interaction can be potentially multivalent due to the macromolecular character of heparin and the micelle-forming ability of PSLs. The importance of the negatively-charged phosphate group of PSLs for the interaction with CatD was shown by the specific effect of Mg^{2+}/Ca^{2+}, which substantially suppressed the induction effect. This can be explained by the effective coordination of the divalent cations to the phosphate group, thus masking the phosphate functionality of PSLs and preventing their binding to CatD. Our results do not exclude the possibility that the activating PSLs might interact with CatD through a different mode of recognition than the inhibitory SLs, although their binding sites are not functionally independent, as suggested by the experiment combining SP and PSL action.

To evaluate the SL/PSL-mediated modulation of CatD in a physiological context, future work will be needed. We suppose that the modulation is physiologically relevant since CatD and SLs/PSLs can co-localize based on their broad intracellular and extracellular distribution. Furthermore, SLs/PSLs were effective *in vitro* at physiological concentrations (generally up to the low micromolar level). In addition to an inhibition/activation, we also demonstrated that the presence of PSLs causes a shift in the CatD pH optimum to a higher pH. This shift may be important for CatD activity when the enzyme is translocated outside lysosomes, *e.g.* during pathophysiologic conditions [8,9,14]. SLs/PSLs can be also involved in the modulation of autocatalytic activation of the CatD zymogen, as suggested by their effect on the cleavage of the synthetic processing region from the CatD propeptide. The polarity of the CatD response to SLs as inhibitors and the role of PSLs as activators deserve special attention. Bioactive SLs such as ceramide and sphingosine have been implicated in promoting apoptosis and cell cycle arrest, while PSLs (ceramide-1-phosphate and sphingosine-1-phosphate) are antagonistic molecules with generally antiapoptotic and mitogenic properties [3,4,12,32]. CatD is involved in both apoptosis and mitogenesis [8,9,12]. Thus, it is intriguing to speculate that the role of CatD in these processes is under complex control of the regulatory SL-PSL system and that CatD is evolutionarily designed to sense these modulators. In conclusion, we discovered that CatD is effectively modulated by naturally occurring SLs/PSLs as well as their specific synthetic derivatives. Based on this, we propose that SLs/PSLs might serve as a novel scaffold for the development of selective regulators of CatD-dependent processes with potential therapeutic applications.

Supplementary materials related to this article can be found online at doi:10.1016/j.bbalip.2011.09.005.

Acknowledgements

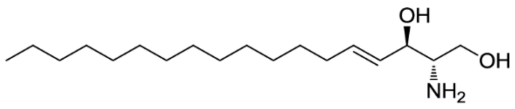
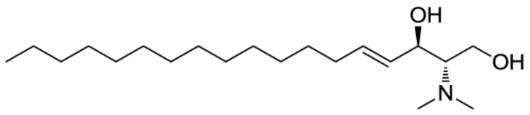
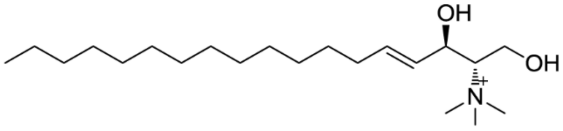
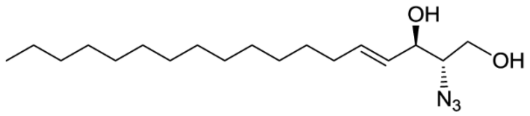
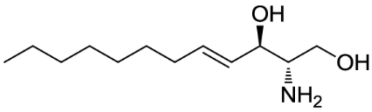
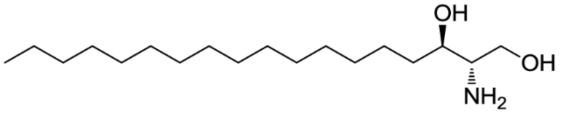
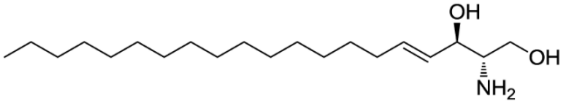
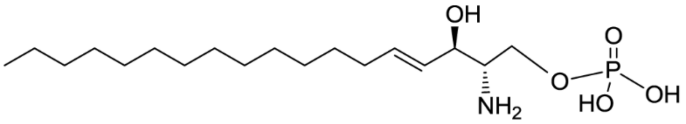
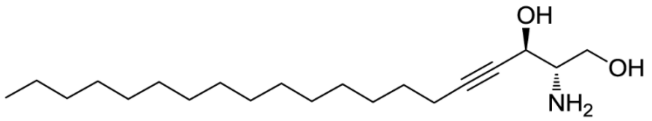
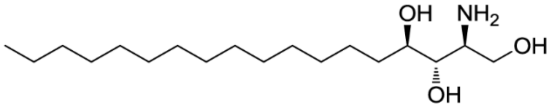
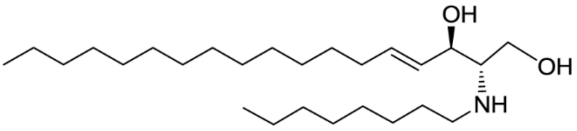
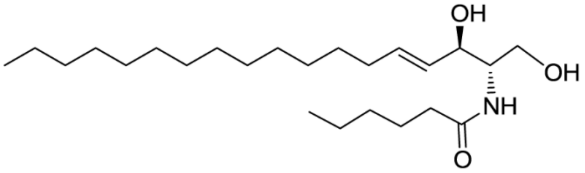
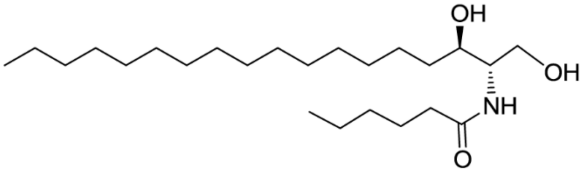
This work was supported by grant no. IAA400550705 from the Grant Agency of the Academy of Sciences of the Czech Republic and by research project Z40550506. The authors thank Hillary Hoffman for critical reading of the manuscript.

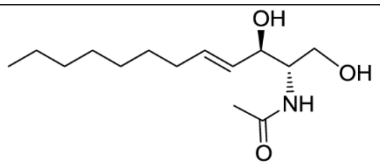
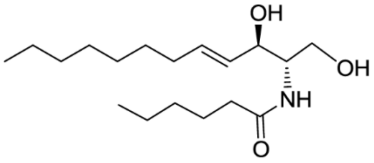
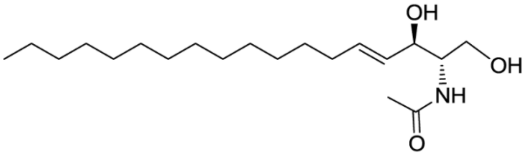
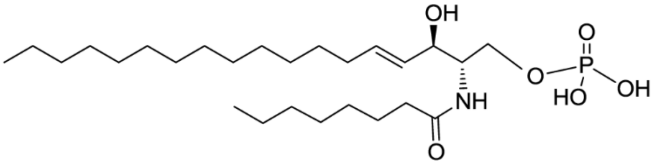
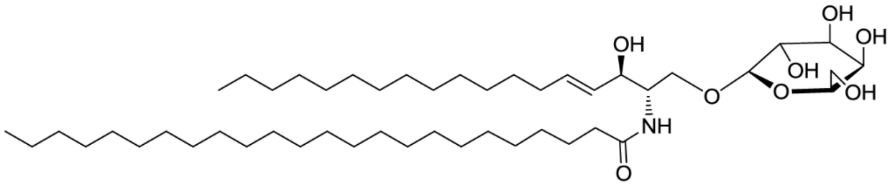
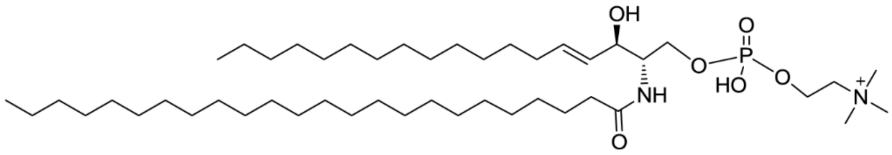
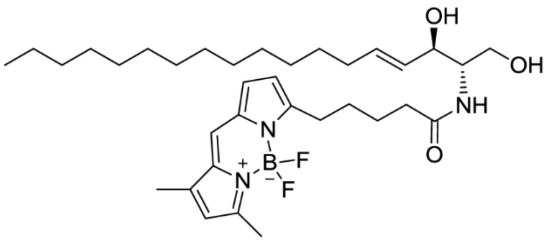
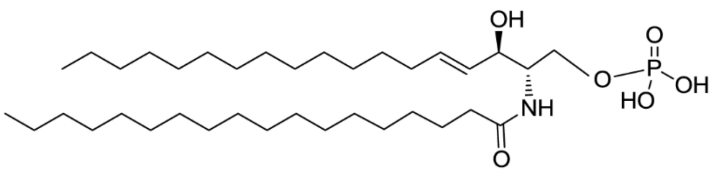
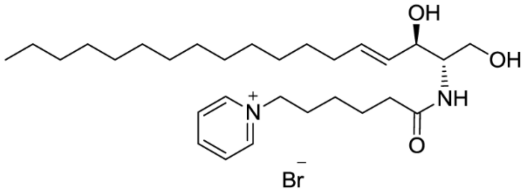
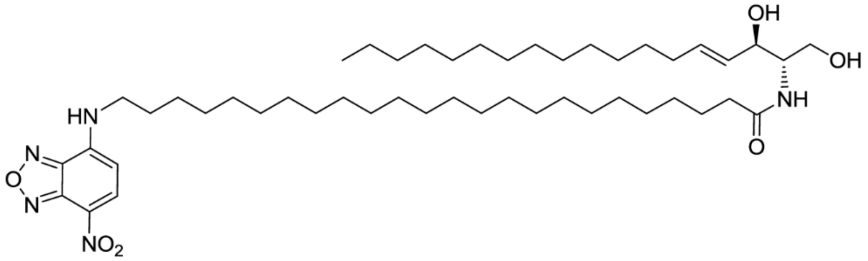
References

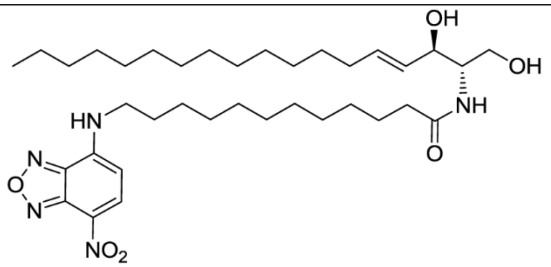
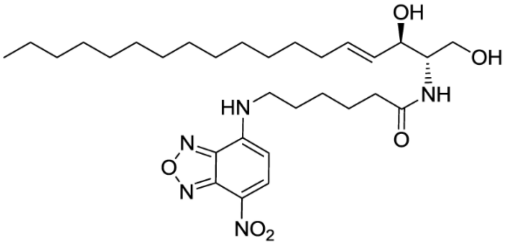
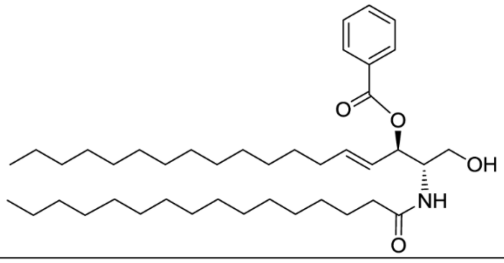
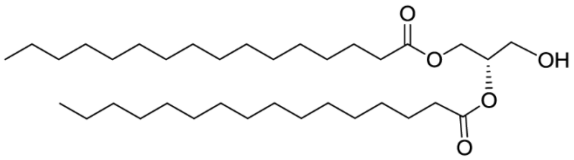
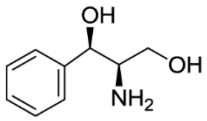
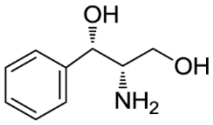
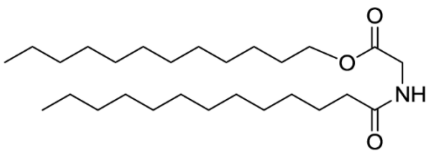
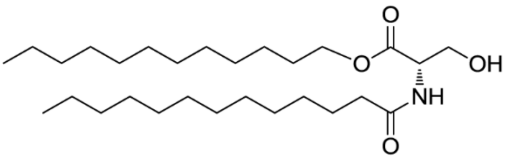
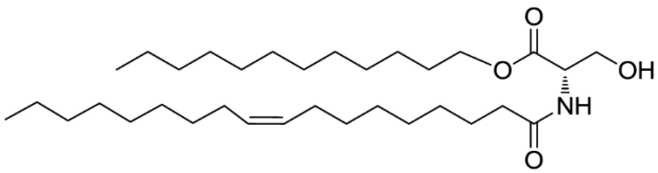
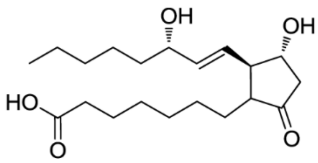
[1] N. Bartke, Y.A. Hannun, Bioactive sphingolipids: metabolism and function, J. Lipid Res. 50 (Suppl) (2009) S91–S96.
[2] Y.A. Hannun, L.M. Obeid, Principles of bioactive lipid signalling: lessons from sphingolipids, Nat. Rev. Mol. Cell Biol. 9 (2008) 139–150.
[3] B. Ogretmen, Sphingolipids in cancer: regulation of pathogenesis and therapy, FEBS Lett. 580 (2006) 5467–5476.
[4] A. Morales, J.C. Fernandez-Checa, Pharmacological modulation of sphingolipids and role in disease and cancer cell biology, Mini Rev. Med. Chem. 7 (2007) 371–382.
[5] C.F. Snook, J.A. Jones, Y.A. Hannun, Sphingolipid-binding proteins, Biochim. Biophys. Acta 1761 (2006) 927–946.
[6] M. Heinrich, M. Wickel, W. Schneider-Brachert, C. Sandberg, J. Gahr, R. Schwandner, T. Weber, P. Saftig, C. Peters, J. Brunner, M. Kronke, S. Schutze, Cathepsin D targeted by acid sphingomyelinase-derived ceramide, EMBO J. 18 (1999) 5252–5263.
[7] P. Saftig, M. Hetman, W. Schmahl, K. Weber, L. Heine, H. Mossmann, A. Koster, B. Hess, M. Evers, K. von Figura, C. Peters, Mice deficient for the lysosomal proteinase cathepsin D exhibit progressive atrophy of the intestinal mucosa and profound destruction of lymphoid cells, EMBO J. 14 (1995) 3599–3608.
[8] O. Masson, A.S. Bach, D. Derocq, C. Prebois, V. Laurent-Matha, S. Pattingre, E. Liaudet-Coopman, Pathophysiological functions of cathepsin D: targeting its catalytic activity versus its protein binding activity? Biochimie 92 (2010) 1635–1643.

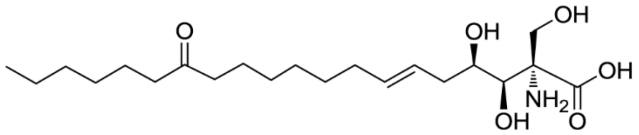
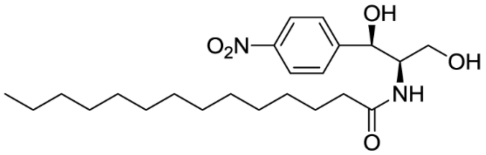
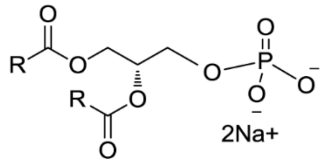
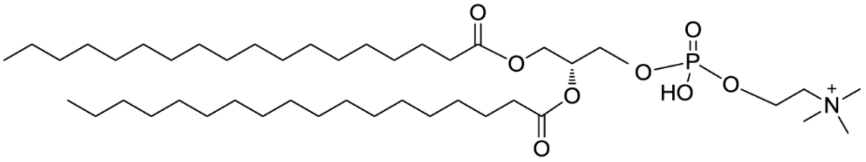
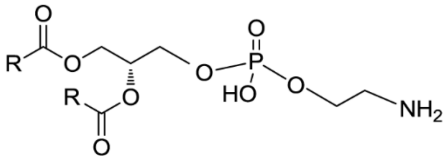
[9] P. Benes, V. Vetvicka, M. Fusek, Cathepsin D—many functions of one aspartic protease, Crit. Rev. Oncol. Hematol. 68 (2008) 12–28.
[10] J. Rodriguez, J. Vazquez, M.D. Corte, M. Lamelas, M. Bongera, M.G. Corte, A. Alvarez, M. Allende, L. Gonzalez, M. Sanchez, M. Vijande, M.J. Garcia, F. Vizoso, Clinical significance of cathepsin D concentration in tumor cytosol of primary breast cancer, Int. J. Biol. Markers 20 (2005) 103–111.
[11] N. Bossard, F. Descotes, A.G. Bremond, Y. Bobin, H.P. De Saint, F. Golfier, A. Awada, P.M. Mathevet, L. Berrerd, Y. Barbier, J. Esteve, Keeping data continuous when analyzing the prognostic impact of a tumor marker: an example with cathepsin D in breast cancer, Breast Cancer Res. Treat. 82 (2003) 47–59.
[12] E. Liaudet-Coopman, M. Beaujouin, D. Derocq, M. Garcia, M. Glondu-Lassis, V. Laurent-Matha, C. Prebois, H. Rochefort, F. Vignon, Cathepsin D: newly discovered functions of a long-standing aspartic protease in cancer and apoptosis, Cancer Lett. 237 (2006) 167–179.
[13] M.M. Gopalakrishnan, H.W. Grosch, S. Locatelli-Hoops, N. Werth, E. Smolenova, M. Nettersheim, K. Sandhoff, A. Hasilik, Purified recombinant human prosaposin forms oligomers that bind procathepsin D and affect its autoactivation, Biochem. J. 383 (2004) 507–515.
[14] M. Beaujouin, C. Prebois, D. Derocq, V. Laurent-Matha, O. Masson, S. Pattingre, P. Coopman, N. Bettache, J. Grossfield, R.E. Hollingsworth, H. Zhang, Z. Yao, B.T. Hyman, P. van der Geer, G.K. Smith, E. Liaudet-Coopman, Pro-cathepsin D interacts with the extracellular domain of the beta chain of LRP1 and promotes LRP1-dependent fibroblast outgrowth, J. Cell Sci. 123 (2010) 3336–3346.
[15] Y. Zhang, X. Li, K.A. Becker, E. Gulbins, Ceramide-enriched membrane domains—structure and function, Biochim. Biophys. Acta 1788 (2009) 178–183.
[16] M. Heinrich, J. Neumeyer, M. Jakob, C. Hallas, V. Tchikov, S. Winoto-Morbach, M. Wickel, W. Schneider-Brachert, A. Trauzold, A. Hethke, S. Schutze, Cathepsin D links TNF-induced acid sphingomyelinase to Bid-mediated caspase-9 and −3 activation, Cell Death Differ. 11 (2004) 550–563.
[17] D. De Stefanis, P. Reffo, G. Bonelli, F.M. Baccino, G. Sala, R. Ghidoni, P. Codogno, C. Isidoro, Increase in ceramide level alters the lysosomal targeting of cathepsin D prior to onset of apoptosis in HT-29 colon cancer cells, Biol. Chem. 383 (2002) 989.
[18] M. Fusek, M. Baudys, P. Metcalf, Purification and crystallization of human cathepsin D, J. Mol. Biol. 226 (1992) 555–557.
[19] M. Masa, L. Maresova, J. Vondrasek, M. Horn, J. Jezek, M. Mares, Cathepsin D propeptide: mechanism and regulation of its interaction with the catalytic core, Biochemistry 45 (2006) 15474–15482.
[20] J. Novotny, K. Pospechova, A. Hrabalek, R. Cap, K. Vavrova, Synthesis of fluorescent C24-ceramide: evidence for acyl chain length dependent differences in penetration of exogenous NBD-ceramides into human skin, Bioorg. Med. Chem. Lett. 19 (2009) 6975–6977.
[21] J. Novotny, B. Janusova, M. Novotny, A. Hrabalek, K. Vavrova, Short-chain ceramides decrease skin barrier properties, Skin Pharmacol. Physiol. 22 (2009) 22–30.
[22] K. Vavrova, A. Hrabalek, P. Dolezal, L. Samalova, K. Palat, J. Zbytovska, T. Holas, J. Klimentova, Synthetic ceramide analogues as skin permeation enhancers: structure-activity relationships, Bioorg. Med. Chem. 11 (2003) 5381–5390.
[23] J.R. Lakowics, Principles of Fluorescence Spectroscopy, second ed., Springer, New York, 1999.
[24] J. Ouveley, O.W. Thiele, A note on the spectrophotometric determination of long chain bases in lipids, Lipids 15 (1979) 194–195.
[25] M. Beckman, C. Freeman, C.R. Parish, D.H. Small, Activation of cathepsin D by glycosaminoglycans, FEBS J. 276 (2009) 7343–7352.
[26] J.M. Padron, Sphingolipids in anticancer therapy, Curr. Med. Chem. 13 (2006) 755.
[27] N.S. Radin, Designing anticancer drugs via the achilles heel: ceramide, allylic ketones, and mitochondria, Bioorg. Med. Chem. 11 (2003) 2123–2142.
[28] B.M. Dunn, S. Hung, The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases, Biochim. Biophys. Acta 1477 (2000) 231–240.
[29] K.R. Williams, N.D. Williams, W. Konigsberg, R.K. Yu, Acidic lipids enhance cathepsin D cleavage of the myelin basic protein, J. Neurosci. Res. 15 (1986) 137–145.
[30] M. Fairhead, S.M. Kelly, C.F. van der Walle, A heparin binding motif on the prodomain of human procathepsin L mediates zymogen destabilization and activation, Biochem. Biophys. Res. Commun. 366 (2008) 862–867.
[31] D. Caglic, J.R. Pungercar, G. Pejler, V. Turk, B. Turk, Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions, J. Biol. Chem. 282 (2007) 33076–33085.
[32] A. Gomez-Munoz, Ceramide 1-phosphate/ceramide, a switch between life and death, Biochim. Biophys. Acta 1758 (2006) 2049–2056.

Supplemental Table S1
Structures of the sphingolipids and related compounds presented in Table 1.

Comp. no.	Abbreviation	Structure
	Sphingosine derivatives	
1	sphingosine	
2	*N,N*-dimethylsphingosine	
3	*N,N,N*-trimethylsphingosine	
4	azidosphingosine	
5	C12-sphingosine	
6	dihydrosphingosine (sphinganine)	
7	C20-sphingosine	
8	sphingosine-1-phosphate	
9	(*2S,3R*)-2-aminoicos-4-yne-1,3-diol (triple bond C20-sphingosine)	
10	phytosphingosine	
11	C8-ceramine (*N*-octylsphingosine)	
	Ceramide derivatives	
12	C6-ceramide (C18 base)	
13	C6-dihydroceramide (C18 base)	

14	C2-ceramide (C12 base)	
15	C6-ceramide (C12 base)	
16	C2-ceramide (C18 base)	
17	C8-ceramide-1-phosphate	
18	1-*O*- β-D-galactosyl-ceramide[a]	
19	sphingomyelin	
20	BODIPY FL C5-ceramide	
21	C18-ceramide-1-phosphate[a]	
22	C6-pyridinium-ceramide	
23	C24-NBD-ceramide	

24	C12-NBD-ceramide	
25	C6-NBD-ceramide	
26	3-*O*-benzoyl-C16-ceramide	
	Others	
27	1,2-dipalmitoyl-glycerol	
28	(*1R,2R*)-2-amino-1-phenyl-1,3-propanediol	
29	(*1S,2S*)-2-amino-1-phenyl-1,3-propanediol	
30	*N*-dodecanoylglycine dodecyl ester	
31	*N*-dodecanoyl-L-serine dodecyl ester	
32	*N*-oleoyl-L-serine dodecyl ester	
33	prostaglandin E1	

34	myriocin	
35	(*1R,2R*)-2-myristylamino-1-(4-nitrophenyl)-1,3-propandiol	
36	phosphatidic acid[b]	
37	phosphatidylcholine	
38	phosphatidylethanolamine[b]	

[a]Major acyl component of natural mixtures from bovine brain is presented.

[b]Natural mixtures from egg yolk.

4.4. Publikace č. 3: Single- and double-headed chemical probes for detection of active cathepsin D in a cancer cell proteome.

Souhrn

Cílem této práce bylo navrhnout, připravit a otestovat aktivitní sondy (ABP, z angl. Activity Based Probe), které by specificky značily aspartátové proteasy pepsinové rodiny. ABP byly dosud intenzivně studovány jako nástroj pro zobrazování cysteinových a serinových proteas s mnoha aplikacemi *in vitro* a *in vivo* (Sanman a Bogyo 2014). V případě aspartátových proteas byla ale snaha o vývoj účinných a specifických ABP komplikována nedostatkem vhodných reaktivních chemických skupin pro kovalentní vazbu do katalytického místa těchto enzymů. V publikaci č. 3 byl využit přístup založený na konstrukci ABP kombinující nekovalentní specifický ligand aktivního místa s fotoreaktivní skupinou, jejíž aktivace kovalentně připojí ABP k enzymu.

Byly připraveny a testovány tři nové aktivitní sondy pro vizualizaci aspartátových proteas (FAP-09, sBAP-09 a dBAP-09), které se liší konstrukčními detaily umožňujícími odlišné aplikace. Tyto sondy se skládají ze tří funkčních modulů: (1) peptidomimetického hexapeptidu, který se váže do aktivního místa enzymu a který byl odvozen ze sekvence specifického reverzibilního inhibitoru pepstatinu (Majer et al. 1997), (2) fotoreaktivní skupiny, která při aktivaci UV zářením kovalentně váže sondu na proteasu a (3) detekční skupiny pro vizualizaci (fluorescenční skupina, biotin). Značené aspartátové proteasy lze vizualizovat z SDS-PAGE gelu pomocí fluorescenčního skeneru nebo po přenesení na PVDF membránu pomocí např. chemiluminiscence.

Pro připravené ABP byly v aktivitním testu zjištěny inhibiční konstanty s lidským katepsinem D. Naměřené hodnoty K_i byly v sub-nanomolární oblasti, tedy obdobné hodnoty, jaké vykazuje inhibitor pepstatin, ze kterého byly odvozeny struktury těchto sond a který jim uděluje vysokou citlivost. Na dvou modelových zástupcích aspartátových proteas (pepsinu a lidském katepsinu D) bylo prokázáno, že připravené ABP jsou schopné s oběma proteasami interagovat a že lze tyto enzymy vizualizovat. Detekční limit ABP byl stanoven na 2,5-5 ng katepsinu D v reakční směsi.

V dalším kroku byly použity připravené ABP pro značení katepsinu D v homogenátu buněk rakoviny prsu (MCF-7B), které produkují katepsin D. V tomto experimentu byly specificky vizualizovány dva pásy, které odpovídají těžkému (31 kDa) a lehkému řetězci (13-14 kDa) zralého aktivního lidského katepsinu D a jejich identita byla potvrzena pomocí hmotnostní spektrometrie. V práci je diskutována aplikace vyvinutých sond pro detekci katepsinu D jakožto markeru při nádorovém onemocnění prsu a v proteomice dalších typů medicinálně významných aspartátových proteas.

Můj podíl na práci zahrnoval (1) přípravu a charakterizaci extraktu z nádorových buněk, (2) značení katepsinu D v extraktu pomocí aktivitních sond a (3) analýzu a identifikaci vzorků vizualizovaných ve fluorescenční a chemiluminiscenční aplikaci.

Publikace č. 3

Nussbaumerová M., **Srp J.**, Máša M., Hradilek M., Šanda M., Reiniš M., Horn M., Mareš M.

Single- and double-headed chemical probes for detection of active cathepsin D in a cancer cell proteome

Chembiochem, (11), 1538-1541 (2010)

DOI: 10.1002/cbic.201000182

Single- and Double-Headed Chemical Probes for Detection of Active Cathepsin D in a Cancer Cell Proteome

Martina Nussbaumerová,[a] Jaroslav Srp,[a, b] Martin Máša,[a] Martin Hradilek,[a] Miloslav Šanda,[a] Milan Reiniš,[c] Martin Horn,[a] and Michael Mareš*[a]

Activity-based proteomics employing active-site-directed chemical probes for enzymatic activity profiling in complex proteomes has greatly accelerated the functional annotation of proteins.[1,2] In protease research, new generations of activity-based probes have led to tremendous progress in our understanding of the biochemistry and physiology of cysteine peptidases over the last decade.[3,4] These probes were designed for in vitro applications as well as for in vivo monitoring in living cells and whole animals, and has directed research towards biomarker discovery and drug screening. However, the current availability of analogous chemical tools selective for other classes of peptidases is limited. The situation is especially critical in the case of aspartic peptidases, despite the fact that these enzymes play an important role in a number of devastating human pathologies, such as AIDS, Alzheimer's disease, and cancer. A major reason for the limited number of available probes for aspartic peptidases is the lack of suitable covalent inhibitors that can be used as starting templates for probe derivation. There are only a few examples of recent advances in this field, including probes for human γ-secretase and malaria plasmepsins that were constructed based on the concept of photoaffinity labeling.[5,6] In this work, we focus on cathepsin D, an aspartic peptidase that is critically associated with cancer development and progression through its complex proteolytic and proliferative action and is a prognostic marker in breast cancer (for a recent review, see ref. [7]). There is still insufficient knowledge regarding the interaction partners and activity regulation mechanisms of this therapeutic target, and the introduction of novel imaging tools is required for these studies. Here, we present a convenient synthesis of novel proteomic-activity-based probes for cathepsin D that are suitable for applications in activity profiling and analysis of post-translational activation processing.

The design of three probes, sBAP-09, dBAP-09, and FAP-09, is shown in Figure 1. Their structures consist of three functionalities: 1) a peptidic binding core, 2) a photoreactive group as a probe "warhead", and 3) a fluorescent group or biotin as a reporter tag. The binding core was constructed by using pepstatin, a reversible peptidomimetic inhibitor of aspartic peptidases, as a template. The pepstatin molecule was structurally minimized and optimized for cathepsin D inhibition[8] to yield the binding core sequence shown in Figure 1. The statin residue in the peptidic core interacts with the enzyme's catalytic residues and acts as a transition-state mimic. The binding core scaffold was modified at its N and/or C terminus to incorporate functionalities 2) and 3). The crystal structure of human cathepsin D in complex with pepstatin (PDB ID: 1LYB) suggests that bulky substituents in the terminal positions will protrude from the enzyme active site and should not interfere with the interaction of ligand and enzyme. In the sBAP-09 and dBAP-09 probes, the photoreactive warhead and reporter tag were

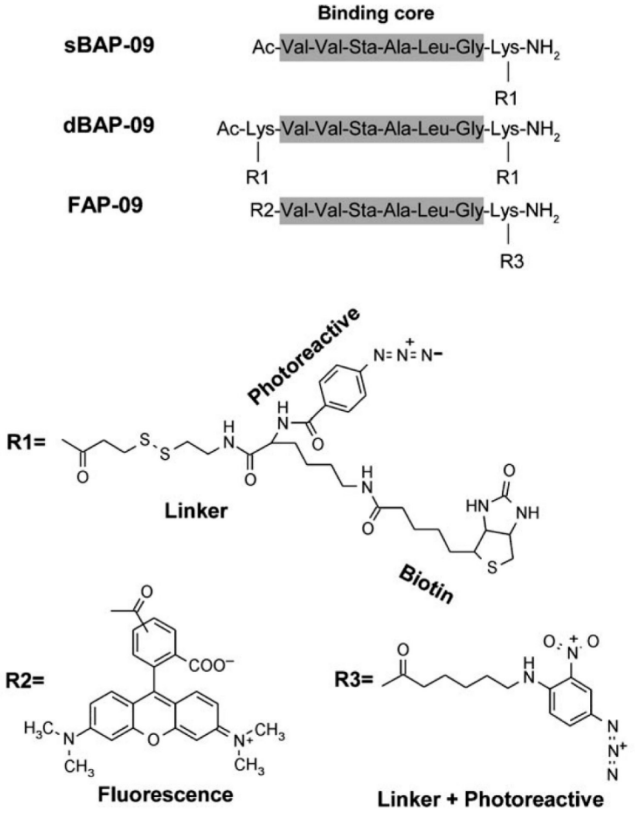


Figure 1. Design of activity-based probes. The structure of the probes consists of a binding core, a photoreactive group, and a reporter tag. The binding core (gray box) is a reversible active-site ligand containing a statin (Sta) residue. The photoreactive aryl azide group that crosslinks the probe to the target enzyme is connected through a flexible linker. A fluorescent TAMRA dye moiety or biotin is incorporated as a reporter group for visualization of the labeled enzyme. Lysine (Lys) spacers extending the binding core are substituted on the side-chain amino group.

[a] *Dr. M. Nussbaumerová, J. Srp, Dr. M. Máša, Dr. M. Hradilek, Dr. M. Šanda, Dr. M. Horn, Dr. M. Mareš*
Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic
Flemingovo n. 2, 16610 Prague (Czech Republic)
Fax: (+420) 220183578
E-mail: mares@uochb.cas.cz

[b] *J. Srp*
Department of Biochemistry, Faculty of Science
Charles University in Prague
Hlavova 8, 12843 Prague (Czech Republic)

[c] *M. Reiniš*
Department of Tumour Immunology
Institute of Molecular Genetics Academy of Sciences of the Czech Republic
Vídeňská 1083, 14220 Prague (Czech Republic)

WILEY InterScience DISCOVER SOMETHING GREAT
© 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

linked to a combined module that was attached to one terminus only (single-headed probe sBAP-09) or to both termini (double-headed probe dBAP-09). The role of the photoreactive warhead is to covalently crosslink the binding core to the enzyme upon irradiation with UV light. We used commercially available building blocks with an aryl azide photoreactive group as warheads. For the synthesis of the probes, we developed a simple strategy based on automated Fmoc solid-phase chemistry and one-step post-synthetic modification as described in the Experimental Section.

The binding properties of the probes were examined with an enzyme inhibition assay. Table 1 shows that the probes efficiently blocked (without the photoactivation step) the enzymatic activity of human cathepsin D. The measured K_i values were in the sub-nanomolar range and of the same order of magnitude as the K_i of pepstatin ($K_i = 0.25 \pm 0.02$ nM). This demonstrated that the N-/C-terminal modification of the binding core did not influence the affinity of the probes for the target enzyme.

Photoaffinity labeling of the target peptidases with individual probes is presented in Figure 2. Following the labeling reaction, the proteins were electrophoretically separated and visualized. The method of visualization depended on the nature of the tag: 1) for probe FAP-09, the SDS-PAGE gel was directly visualized by using a fluorescence scanner and 2) for s/dBAP-09, a blot was visualized by using a chemiluminescent biotin–avidin detection system. In the first step, we demonstrated that all probes covalently bind (in the presence of an excess of BSA) to purified human cathepsin D as well as to porcine pepsin, an "archetypal" aspartic peptidase (Figure 2A and B). Interestingly, the labeling signal of the double-headed probe dBAP-09 was significantly stronger than that of the single-headed probe sBAP-09, which most likely reflects a higher crosslinking efficiency of dBAP-09 as well as better visualization due to the double tag. In contrast, no labeling was found for purified papain and trypsin, which were used as an archetypal cysteine and serine peptidase, respectively (data not shown). In the next step, the protein extract from MCF-7B human breast cancer cells was labeled; this cell line efficiently expresses cathepsin D as demonstrated by immunoblot analysis (Figure 2C). Imaging of the cell extract clearly revealed two bands at molecular weights of 31 kDa and 13–14 kDa, which correspond to the theoretical molecular weight of the two polypeptide chains of mature cathepsin D; this labeling pattern was identical to that obtained for purified cathepsin D (Figure 2B). This indicated that the probes interact with the enzymatically active two-chain form of cathepsin D that is generated intracellularly by proteolytic processing from the inactive single-chain zymogen.[9] The labeling was specifically blocked when the labeling reaction was performed in the presence of an

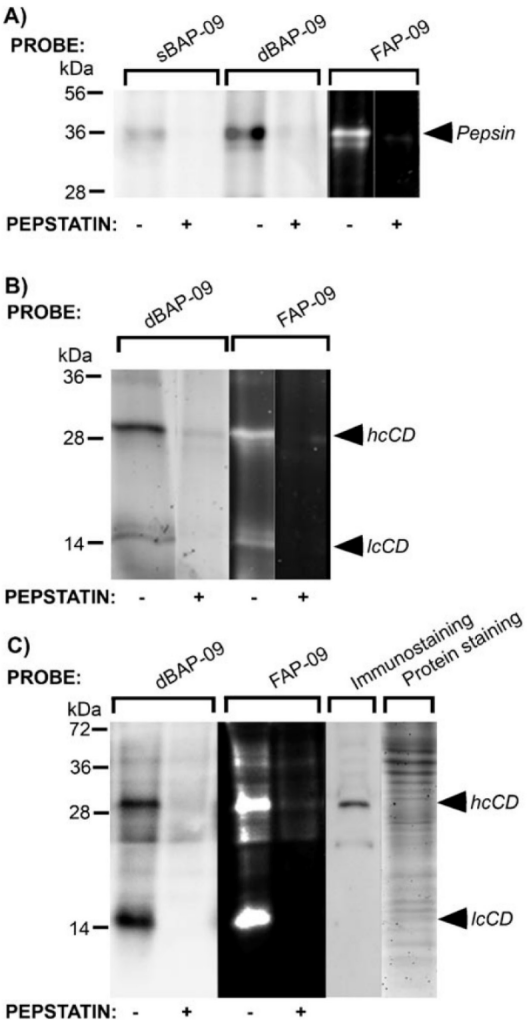


Figure 2. Imaging of cathepsin D and pepsin with activity-based probes. A) Labeling of pepsin (0.2 μg), an "archetypal" aspartic peptidase. B) Labeling of purified cathepsin D (0.1 μg), and C) cathepsin D in the protein extract (100 μg of protein) from MCF-7B human breast cancer cells. The position of the heavy and light chain (hcCD and lcCD, respectively) of mature cathepsin D is indicated (note: in the immunoblot only hcCD is recognized by anti-cathepsin D Ab). After the labeling reaction, the proteins were resolved by SDS-PAGE, and the labeled peptidases were visualized in the gel (fluorescence signal of FAP-09) by using a fluorescence scanner or on a chemiluminescent avidin blot (biotin signal of sBAP-09 or dBAP-09) by using a CCD camera. The competitive labeling was performed in the presence of the selective active-site inhibitor pepstatin. Molecular mass standards are shown on the left.

Table 1. Properties and proteomic applications of activity-based probes.

Name	M_W	Design	Inhibition K_i [nM][a]	Applications
FAP-09	1442.4	single-headed, fluorescence-tagged	0.84 ± 0.07	in-gel visualization, mass spectrometry
sBAP-09	1445.3	single-headed, biotin-tagged	0.50 ± 0.05	on-blot visualization, microsequencing
dBAP-09	2236.0	double-headed, biotin-tagged	0.77 ± 0.04	on-blot visualization, microsequencing

[a] Inhibition of the enzymatic activity of human cathepsin D by the probes was assayed with a FRET-based peptidic substrate; mean values of the inhibition constant ± standard errors are given.

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

excess of pepstatin, which competes with the probe for binding to the enzyme active site (Figure 2). Furthermore, by using these probes, we were able to visualize cathepsin D-like peptidases in two parasite proteomes rich in cathepsin D activity,[10,11] namely in extracts from the digestive tract of the hard tick *Ixodes ricinus* and the blood fluke *Schistosoma mansoni* (data not shown). In summary, these results clearly demonstrated the selectivity of the probes for cathepsin D and related aspartic peptidases from the A1 (pepsin) family and their ability to label the target in a complex protein mixture with a low background signal. The detection limit of the probes, based on the labeling of purified cathepsin D and pepsin with dBAP-09/FAP-09, is 2.5–5 ng of target enzyme.

In addition to the application of the probes as a visualization tool, we pursued further characterization of the labeled cathepsin D. The fluorescent probe FAP-09 was used for analysis of cancer-cell-derived cathepsin D (Figure 2C). The labeled bands on the SDS PAGE gel were subjected to enzymatic digestion followed by LC–MS/MS analysis, which provided ~32% peptide coverage of the human cathepsin D sequence in the heavy and light chains. Cathepsin D labeled with the biotinylated probe dBAP-09 and detected on the blot was subjected to N-terminal analysis (Figure 2B). The visualization protocol described in the Experimental Section was compatible with subsequent Edman microsequencing, and this analysis revealed an N-terminal sequence of LGGVK for the heavy chain and (E)GPIPE (trimmed forms in about a 1:1 ratio) for the light chain. These mass spectrometry and protein-sequencing experiments highlight the suitability of the probes for various analytical applications using 1D/2D gels and blots.

In conclusion, we present the design of specific active-site directed probes for cathepsins D and related aspartic peptidases, which enable the effective labeling of these peptidases in a complex proteome (summarized in Figure 3). The applicability of this system to activity-based proteomics was demonstrated for cancer-cell-derived cathepsin D, a therapeutic target in several human tumors. The modular assembly renders the probe synthesis convenient and widely available. Due to their versatile fluorescent and biotin tags, the probes can be used as a powerful tool for a broad range of "in-gel" and "on-blot" proteomic applications in profiling medically important cathepsin D-type peptidases.

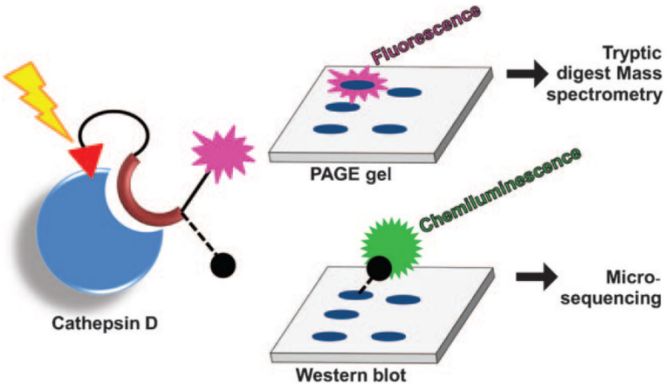


Figure 3. From labeling reaction to proteomic analysis. Left) In a complex protein mixture, the activity-based probe specifically interacts with the active site of cathepsin D through the binding core (brown). Then, the probe is covalently crosslinked to the target peptidase through its photoreactive group (red triangle) upon irradiation with UV light. Right) The proteins are separated on a 1D/2D PAGE gel and in some cases subsequently blotted onto a PVDF membrane. The fluorescent-tagged probe (magenta) allows for direct "in-gel" visualization of the labeled peptidase by using a fluorescence scanner, which can be followed by analysis of the band/spot by mass spectrometry. The target labeled with the biotin-tagged probe (black ball) is detected on the blot with a biotin–avidin chemiluminescent detection system that is compatible with subsequent N-terminal microsequencing.

Experimental Section

Synthesis of the probes: The peptidic components of sBAP-09, dBAP-09, and FAP-09 (Ac-Val-Val-Sta-Ala-Leu-Gly-Lys-NH_2, Ac-Lys-Val-Val-Sta-Ala-Leu-Gly-Lys-NH_2 and H-Val-Val-Sta-Ala-Leu-Gly-Lys-NH_2, respectively) were synthesized in the form of peptidyl amides containing a free or acetylated N terminus by standard Fmoc solid-phase chemistry on an ABI 433A peptide synthesizer (Applied Biosystems) by using Fmoc-protected L-amino acids and Rink amide resin (IRIS Biotech, Marktredwitz, Germany). 5(6)-Carboxytetramethylrhodamine (TAMRA; Novabiochem) was coupled to the N-terminal Val residue of the FAP-09 binding core in the last step of the automated synthesis. The peptides were cleaved from the solid support and deprotected by the TFA procedure, TBME precipitated, and RP-HPLC purified.[12] In the next step, the ε-amino group of lysine in the binding core peptides was modified by reaction with an amino-reactive (NHS-ester) building block. The following compounds were used as building blocks containing a photoreactive group: 1) sulfo-*N*-succinimidyl-2-(6-[biotinamido]-2-(*p*-azido benzamido)-hexanoamido) ethyl-1,3′-dithioproprionate (Sulfo-SBED Reagent; Pierce) and 2) *N*-succinimidyl-6-[4′-azido-2′-nitrophenylamino]hexanoate (SANPAH Reagent; Pierce). The modification reaction was performed by using the following procedure: Sulfo-SBED reagent (0.3 mg and 0.7 mg for sBAP-09 and dBAP-09, respectively) or SANPAH reagent (0.15 mg for FAP-09) in dry DMF (15 μL) was mixed with the peptide solution (0.3 mg in 10 μL of dry DMF) and iPr_2EtN (1 μL). The reaction mixture was incubated in a Pyrex test tube under an argon atmosphere at 37 °C overnight and then lyophilized. The probes were purified by RP-HPLC by using an Agilent 1200 Series system (Agilent Technologies) on a Kromasil C18 column (250×4.6 mm, 5 μm; Alltech) equilibrated in 1% HOAc at a flow rate of 0.8 mL min^{-1}. The probes were eluted with a 1% min^{-1} gradient of MeOH, and the eluate was monitored at 220/260 nm. The purified probes were characterized by ESI mass spectrometry on an LCQ Classic Finnigan Mat (Thermo Finnigan); sBAP-09 showed an $[M-H]^-$ ion at *m/z* 1444.3, dBAP-09 showed an $[M+2Na]^{2+}$ ion at *m/z* 1141.0, and FAP-09 showed an $[M+2H]^{2+}$ ion at *m/z* 722.2. The composition was verified by isotope cluster analysis and amino acid analysis.

Cathepsin D inhibition assay: Inhibition of cathepsin D activity by synthetic probes was assayed with the mature human enzyme purified from placenta and the fluorogenic substrate Abz-Lys-Pro-Ala-Glu-Phe-Nph-Ala-Leu (Abz, aminobenzoic acid; Nph, 4-nitrophenylalanine) in 50 mM NaOAc (pH 4.5) as previously described.[13] The reaction was performed in a microplate format in a total assay volume of 100 μL and was continuously monitored on a GENios microplate reader (TECAN) at excitation and emission wavelengths of 330 and 410 nm, respectively. All measurements were performed in triplicate.

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Breast cancer cell extract: MCF-7B human breast cancer cells were cultured in RPMI 1640 medium (supplemented with 10% fetal calf serum, L-glutamine, and antibiotics) to confluence and harvested by trypsinization. The PBS-washed cells were subjected to three freeze-thaw cycles and homogenized in 50 mM NaOAc (pH 4.5), 1% CHAPS, and 10 μM E-64 on ice. The extract (1 mg of protein per mL) was cleared by centrifugation (10 min, 4 °C, 21 000 *g*), filtered with Ultrafree-MC 0.22 mm (Millipore), and stored at −80 °C.

Labeling reaction: The labeling was performed in a microcentrifuge tube in a total volume of 100 μL for 1) the breast cancer cell extract (100 μg of protein) and 2) purified enzymes: human cathepsin D (0.1 μg) or porcine pepsin (0.2 μg) supplemented with BSA (2 μg). Protein aliquots were incubated with 0.5 μM probe (100 μM stock in DMSO) in 50 mM NaOAc (pH 4.5) for 20 min at 26 °C. Afterward, the reaction mixture was irradiated in an open tube for 10 min on ice with a 125 W high-pressure mercury-vapor lamp (at a distance of 20 cm) to allow for photoactivated crosslinking. Competitive labeling was performed after preincubation (15 min at 26 °C) with 2 μM pepstatin. The reaction mixture was precipitated with four volumes of acetone and boiled in reducing Laemmli sample buffer.

In-gel and on-blot visualization and analysis of peptidases: The labeled proteins were separated by Laemmli-SDS-PAGE (15% gel), and in the case of s/dBAP-09 labeling, transferred to a PVDF membrane by electroblotting. FAP-09-labeled peptidases were visualized directly in the gel by using a Typhoon 9400 Fluorescence Imager (GE Healthcare) with excitation at 532 nm (green laser) and emission at 580 nm (bp 30 nm). The fluorescent bands were excised and subjected to tryptic and chymotryptic digestion. The digests were analyzed by LC-MS/MS on an LTQ Orbitrap XL mass spectrometer (Thermo, Waltham, U.S.A.) coupled to a 2D capillary chromatography Rheos 2000 (Flux Instruments, Reinach, Switzerland). The first dimension column was a Symmetry C18 (180 μm × 20 mm × 3 μm, Waters) and the second dimension column was a PepMap C18 (75 μm × 150 mm × 3 μm, LC Packings, Amsterdam, Netherlands). The LC–MS/MS data were processed with Bioworks software (Thermo).

The blot with s/dBAP-09-labeled peptidases was blocked (30 min at 26 °C) with blocking buffer (10 mM Tris pH 7.4, 150 mM NaCl, 0.05% Tween, 1% polyvinylpyrrolidone, 4% BSA) or protein-free blocking buffer compatible with protein sequencing (10 mM Tris pH 7.4, 150 mM NaCl, 1% Tween, 5% polyvinylpyrrolidone). The biotin tag was detected with horseradish peroxidase–avidin D conjugate (Vector Laboratories; 1:10 000, 30 min at 26 °C) and Super-Signal West Pico chemiluminescent substrate (Pierce), and the blot was visualized with a LAS-4000 luminescent image analyzer (Fuji-film). N-terminal sequencing of the labeled bands was performed by using a Procise 494 cLC protein sequencer (Applied Biosystems). Cathepsin D in the cell extract (50 μg of total protein) was visualized on an immunoblot by using rabbit polyclonal anti-cathepsin D antiserum (raised against human enzyme from placenta; 1:5000), goat anti-rabbit IgG conjugated to horseradish peroxidase (Pierce; 1:40 000), and the chemiluminescent substrate described above.

Acknowledgements

This work was supported by grant no. IAA400550705 from the Grant Agency of the Academy of Sciences of the Czech Republic, and by research projects Z40550506 and AVOZ50520514. We thank Zdeněk Voburka and Miroslava Blechová for technical assistance, and Hillary Hoffman for critical reading of the manuscript.

Keywords: activity-based probes · breast cancer cells · cathepsins · photoaffinity labeling · proteomics

[1] K. T. Barglow, B. F. Cravatt, *Nat. Methods* **2007**, *4*, 822 – 827.
[2] A. E. Speers, B. F. Cravatt, *ChemBioChem* **2004**, *5*, 41 – 47.
[3] M. Fonovic, M. Bogyo, *Expert Rev. Proteomics* **2008**, *5*, 721 – 730.
[4] S. H. Verhelst, M. Bogyo, *Biotechniques* **2005**, *38*, 175 – 177.
[5] C. C. Shelton, L. Zhu, D. Chau, L. Yang, R. Wang, H. Djaballah, H. Zheng, Y. M. Li, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 20228 – 20233.
[6] K. Liu, H. Shi, H. Xiao, A. G. Chong, X. Bi, Y. T. Chang, K. S. Tan, R. Y. Yada, S. Q. Yao, *Angew. Chem.* **2009**, *121*, 8443 – 8447; *Angew. Chem. Int. Ed.* **2009**, *48*, 8293 – 8297.
[7] P. Benes, V. Vetvicka, M. Fusek, *Crit. Rev. Oncol. Hematol.* **2008**, *68*, 12 – 28.
[8] P. Majer, J. R. Collins, S. V. Gulnik, J. W. Erickson, *Protein Sci.* **1997**, *6*, 1458 – 1466.
[9] G. R. Richo, G. E. Conner, *J. Biol. Chem.* **1994**, *269*, 14 806 – 14 812.
[10] M. Horn, M. Nussbaumerová, M. Šanda, Z. Kovářová, J. Srba, Z. Franta, D. Sojka, M. Bogyo, C. R. Caffrey, P. Kopáček, M. Mareš, *Chem. Biol.* **2009**, *16*, 1053 – 1063.
[11] M. Delcroix, M. Sajid, C. R. Caffrey, K. C. Lim, J. Dvořák, I. Hsieh, M. Bahgat, C. Dissous, J. H. McKerrow, *J. Biol. Chem.* **2006**, *281*, 39316 – 39329.
[12] L. Dolečková-Marešová, M. Pavlík, M. Horn, M. Mareš, *Chem. Biol.* **2005**, *12*, 1349 – 1357.
[13] M. Máša, L. Marešová, J. Vondrášek, M. Horn, J. Ježek, M. Mareš, *Biochemistry* **2006**, *45*, 15 474 – 15 482.

Received: March 19, 2010
Published online on June 16, 2010

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

4.5. Publikace č. 4: Profiling system for skin kallikrein proteolysis applied in gene deficient mouse models.

Souhrn

Serinové proteasy kalikreinového typu (KLK) jsou exprimovány v různých orgánech a jejich 15 zástupců se podílí na řadě biologických procesů. Překrývající se substrátová specifita těchto proteas znemožňuje nezávisle monitorovat aktivitu jednotlivých kalikreinů v komplexních směsích. Tato práce se zaměřuje na analýzu substrátové a inhibiční specifity hlavních KLK z tkáně pokožky a na jejich detekci v homogenátech epidermu geneticky modifikovaných myší.

Zástupci myších a lidských kalikreinů s trypsinovou (KLK 5 a KLK 14) a chymotrypsinovou (KLK 7) substrátovou specifitou byly testovány na citlivost k sadě relevantních syntetických fluorogenních substrátů a reaktivních inhibitorů. Kromě komerčně dostupných substrátů byly také připraveny a analyzovány nové fluorogenní substráty Ac-WAVR-Acc a KHLY-Acc navržené podle substrátové specificity studovaných KLK. Pro jednotlivé substráty byly s rekombinantními KLK určeny kinetické parametry k_{cat} a K_M a na základě katalytické účinnosti (k_{cat}/K_M) byly vybrány selektivní substráty pro specifickou detekci KLK proteas. Nejvyšší míru selektivity vykazoval substrát Ac-WAVR-Acc pro KLK 14 a substrát KHLY-Acc pro KLK 7. Analýza inhibiční specifity rekombinantních KLK v aktivitním testu umožnila identifikovat inhibitor dFPR-Cmk jako nástroj pro rozlišení aktivity trypsinových KLK (KLK 5 a KLK 14) a chymostatin jako selektivní citlivý inhibitor KLK 7 s chymotrypsinovou specifitou.

Byly připraveny geneticky modifikované myši, které měly deficience v genech pro studované KLK a také pro přirozený kalikreinový inhibitor LEKTI, který slouží jako fyziologický regulátor KLK proteas. Deficience LEKTI umožňuje stanovení nemodulovaných aktivit KLK a představuje unikátní model pro studium KLK v tkáni pokožky. Z epidermis geneticky modifikovaných myší byly připraveny extrakty, které byly analyzovány pomocí kombinace identifikovaných selektivních substrátů a inhibitorů pro určení distribuce jednotlivých KLK proteas a vztahů mezi nimi.

Tato práce přináší nové molekulární nástroje pro detekci aktivit kožních KLK a jejich kvantifikaci při patologických podmínkách. Výsledky pomohou objasnit roli KLK proteas v komplexním proteolytickém systému v epidermální tkáni, do kterého přispívá několik typů enzymů.

Můj podíl na práci zahrnoval měření inhibiční specifity pro jednotlivé rekombinantní KLK se sadou syntetických inhibitorů.

Publikace č. 4

Horn M., Zbodáková O., Kašpárek P., **Srp J.**, Hanečková R., Hradilek M., Mareš M., Sedláček R.

Profiling system for skin kallikrein proteolysis applied in gene deficient mouse models

Biological Chemistry, (9), 1085-1089 (2018)

Short Communication

Martin Horn[a], Olga Zbodakova[a], Petr Kasparek*, Jaroslav Srp, Radka Haneckova, Martin Hradilek, Michael Mares* and Radislav Sedlacek

Profiling system for skin kallikrein proteolysis applied in gene-deficient mouse models

https://doi.org/10.1515/hsz-2018-0116
Received January 12, 2018; accepted May 14, 2018; previously published online June 9, 2018

Abstract: Kallikrein-related proteases (KLKs) play a critical role in epidermis physiology and have been implicated in skin pathologies such as Netherton syndrome. The contribution of individual KLKs to skin proteolysis is poorly understood. Monitoring of their activities in skin proteome is hampered by overlapping substrate specificities, and there is a need for novel assays. Here, we present a platform of selective and sensitive fluorogenic substrates and inhibitors for profiling KLK5, KLK7 and KLK14. These chemical tools were evaluated using recombinant KLKs and tissue from a unique set of mice deficient in eight combinations of KLKs and their natural regulator LEKTI.

Keywords: inhibitors; kallikreins; mouse model; proteolysis; skin; substrates.

[a]**Martin Horn and Olga Zbodakova:** These authors contributed equally to this work.
***Corresponding authors: Petr Kasparek,** Laboratory of Transgenic Models of Diseases, Institute of Molecular Genetics of the ASCR, Prumyslova 595, 252 50 Vestec, Czech Republic; and Czech Centre for Phenogenomics, Institute of Molecular Genetics of the ASCR, Prumyslova 595, 252 50 Vestec, Czech Republic, e-mail: petr.kasparek@img.cas.cz; and **Michael Mares,** Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Flemingovo n. 2, 166 10 Prague, Czech Republic, e-mail: mares@uochb.cas.cz
Martin Horn and Martin Hradilek: Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Flemingovo n. 2, 166 10 Prague, Czech Republic
Olga Zbodakova and Radka Haneckova: Laboratory of Transgenic Models of Diseases, Institute of Molecular Genetics of the ASCR, Prumyslova 595, 252 50 Vestec, Czech Republic
Jaroslav Srp: Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Flemingovo n. 2, 166 10 Prague, Czech Republic; and Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, 128 00 Prague, Czech Republic
Radislav Sedlacek: Laboratory of Transgenic Models of Diseases, Institute of Molecular Genetics of the ASCR, Prumyslova 595, 252 50 Vestec, Czech Republic; and Czech Centre for Phenogenomics, Institute of Molecular Genetics of the ASCR, Prumyslova 595, 252 50 Vestec, Czech Republic

The family of kallikrein-related proteases (KLKs) consists of 15 serine proteases expressed in many organs and involved in a number of diverse biological processes (Chao et al., 2006; Tamura et al., 2006; Lundwall and Brattsand, 2008; Stopkova et al., 2017). Typically, multiple KLKs are expressed in one tissue compartment along with other proteases (Beaufort et al., 2010; de Veer et al., 2014). Monitoring of activities of individual KLKs is hampered by their overlapping substrate specificities, and moreover, interference by the proteolytic activity of other non-KLK proteases. Thus, there is a need for novel assays and chemical tools to profile members of KLK family in complex biological systems and understand their biological roles. This study focuses on three KLKs that are expressed in skin and were predicted to play important roles in epidermal physiology and pathology – KLK5, KLK7 and KLK14 (Caubet et al., 2004; Descargues et al., 2005; Borgono et al., 2007; Kasparek et al., 2017). We developed a platform for sensitive analysis of their activities based on selective fluorogenic substrates and inhibitors identified by *in vitro* screening of a set of recombinant skin KLKs of human and mouse origin. This assay platform was further evaluated *in vivo* using mouse models deficient in several *Klk* genes that we recently reported ($Klk5^{-/-}$, $Klk7^{-/-}$, $Klk14^{-/-}$) (Kasparek et al., 2017) and one newly prepared strain $Klk5^{-/-}Klk14^{-/-}$ (Supplementary Figure 1) Kasparek et al. (2014). In addition, we analysed animals combining KLK deficiency with null-mutation in the gene *Spink5*, which encodes the natural KLK inhibitor LEKTI ($Klk5^{-/-}Sp5^{-/-}$, $Klk7^{-/-}Sp5^{-/-}$, $Klk14^{-/-}Sp5^{-/-}$, $Klk5^{-/-}Klk14^{-/-}Sp5^{-/-}$). Such genetic combinations enable evaluation of those KLK activities that would otherwise be undetectable in physiological conditions.

In total, all these mutant strains allow for a KLK substrate validation in the context of all physiologically relevant proteases, lacking only specific genetically defined components.

The activities of a set of five recombinant KLKs including mouse and human KLK5, mouse and human KLK7 and human KLK14 were probed with fluorogenic peptide substrates. A panel of substrates with a fluorogenic leaving group 7-amino-4-methylcoumarin (amc) comprised four substrates for trypsin-like proteases (Boc-VPR-amc, Tos-GPR-amc, Boc-FSR-amc and Boc-QAR-amc; Boc, t-Butyloxycarbonyl; Tos, 4-toluenesulfonyl) and three substrates for chymotrypsin-like proteases (Suc-LLVY-amc, Suc-LY-amc and Suc-AAPF-amc; Suc, succinyl). In addition, two substrates, Ac-WAVR-acc and KHLY-acc, containing a fluorogenic leaving group 7-amino-4-carbamoylmethylcoumarin (acc) were designed. Their structure is based on peptide binding sequences for KLK14 and KLK7, respectively, that were previously identified by positional scanning libraries (de Veer et al., 2012, 2017) and converted to new substrate derivatives for sensitive fluorescence detection.

Michaelis-Menten kinetic parameters (k_{cat} and K_M) of substrates were determined with recombinant KLKs using a kinetic activity assay in microplate format (Table 1). According to catalytic efficiency values (k_{cat}/K_M), the substrate Boc-VPR-amc is the most effective substrate for KLK5 and KLK14 (Table 1A). The other substrates for trypsin-like kallikreins displayed one to three orders of magnitude lower k_{cat}/K_M values. In terms of selectivity, the substrate Ac-WAVR-acc was found to have much better catalytic efficiency for KLK14 (k_{cat}/K_M of 43 313 M^{-1} s^{-1}) compared to both KLK5 enzymes (k_{cat}/K_M of 139 and 79 M^{-1} s^{-1}), which gives selectivity ratio of k_{cat}/K_M values for KLK14 to KLK5 enzymes of 331 (human KLK14/mouse KLK5) and 588 (human KLK14/human KLK5). This selectivity parameter is one to two orders of magnitude better than that for other substrates for trypsin-like kallikreins.

The activity of KLK7 enzymes was probed with substrates preferred by chymotrypsin-like proteases (Table 1B). No effective cleavage was detected for the substrates Suc-LY-amc and Suc-AAPF-amc, and catalytic efficiency with Suc-LLVY-amc was low (k_{cat}/K_M of 144 and 89 M^{-1} s^{-1}). Meanwhile, the substrate KHLY-acc was a good substrate for both KLK7 enzymes, displaying two orders of magnitude higher k_{cat}/K_M values (10 732 and 28 373 M^{-1} s^{-1}).

In conclusion, we identified efficient and selective fluorogenic substrates for high-throughput screening of KLK activities, including (i) Boc-VPR-amc for KLK5 and KLK14, (ii) Ac-WAVR-acc for KLK14 and (iii) KHLY-acc for KLK7.

Inhibitory specificity of KLK enzymes was probed with synthetic peptide inhibitors of proteases. A panel

Table 1: Profiling of recombinant KLK activities with fluorogenic substrates for trypsin-like (A) and chymotrypsin-like (B) proteases.

Substrate	mKLK5			hKLK5			hKLK14			Selectivity (k_{cat}/K_M to k_{cat}/K_M)	
	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})	hKLK14 to mKLK5	hKLK14 to hKLK5
(A)											
Boc-VPR-amc	1.48	48.3	31 947	8.63	101.1	23 482	3.94	29	135 873	4.3	5.8
Tos-GPR-amc	0.26	39.0	6905	0.65	153.1	2336	2.33	87.7	26 583	3.8	11.4
Boc-FSR-amc	0.61	99.2	6421	1.01	100.2	5544	2.04	34.5	58 907	9.2	10.6
Boc-QAR-amc	1.04	139.4	7832	0.41	204.1	1104	7.72	206.4	37 386	4.8	33.8
Ac-WAVR-acc	0.01	82.2	139	1.52	127.6	79	2.88	62.1	46 313	331.2	586.5

	mKLK7			hKLK7		
	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})
(B)						
Suc-LLVY-amc	0.01	64.4	144	0.01	104.1	89
Suc-LY-amc	nd	nd	nd	nd	nd	nd
Suc-AAPF-amc	nd	nd	nd	nd	nd	nd
KHLY-acc	1.09	101.5	10 732	2.81	99.1	28 373

The Michaelis-Menten kinetic parameters were determined by measuring the rate of hydrolysis of the substrate (0–600 μM) using a kinetic activity assay with recombinant KLKs (1–2 nM) in 0.1 M Tris-HCl pH 8.0, 150 mM NaCl. K_M and k_{cat} values were obtained by nonlinear regression using GraFit software. Mean values are given for triplicate measurements (SE values are within 10% of the mean). Selectivity is expressed as a ratio of k_{cat}/K_M values for KLK14 to KLK5. nd, No cleavage detected. Note that panel-A substrates are not cleaved by KLK7, and panel-B substrates are not cleaved by KLK5 and KLK14. All KLKs (R&D Systems) were produced in mouse myeloma cell lines. Substrates Ac-WAVR-acc and KHLY-acc were synthesised as described previously (Fajtova et al., 2015), all other substrates were purchased from Bachem.

of inhibitors included compounds with chloromethylketone (cmk), fluoromethylketone (fmk) or aldehyde reactive warheads: (i) dFPR-cmk, FFR-cmk and Tos-K-cmk for trypsin-like proteases and (ii) AAPV-cmk, Z-FA-fmk (Z, benzyloxycarbonyl), AAPF-cmk and chymostatin aldehyde for chymotrypsin-like proteases.

Inhibitory potency against recombinant KLKs was tested in a kinetic fluorescence assay with three selective fluorogenic substrates (as specified in previous chapter). Among inhibitors with a basic residue in the P1 position targeting preferentially trypsin-like proteases, Tos-K-cmk was found ineffective, and dFPR-cmk and FFR-cmk completely inhibited both KLK5 and KLK14 at 1 μM concentrations. However, at 0.1 μM concentrations dFPR-cmk strongly inhibited KLK14 but not KLK5 (IC_{50} values of 2.6 and 79 nM, respectively), and FFR-cmk was a weak inhibitor of both of enzymes (Figure 1). The inhibitor dFPR-cmk can be applied at 10-nM concentrations as selective inhibitor of KLK14 providing about 19-fold better inhibition of KLK14 than KLK5 (Figure 1). Among inhibitors with a hydrophobic residue in the P1 position targeting preferentially chymotrypsin-like proteases, only chymostatin was found to be a potent inhibitor of KLK7 enzymes with IC_{50} values of 366 and 162 nM (Figure 1).

To conclude, we identified two inhibitors, dFPR-cmk and chymostatin, as selective for KLK14 and KLK7, respectively. They represent a novel tool for discriminating among proteolytic activities of skin kallikreins.

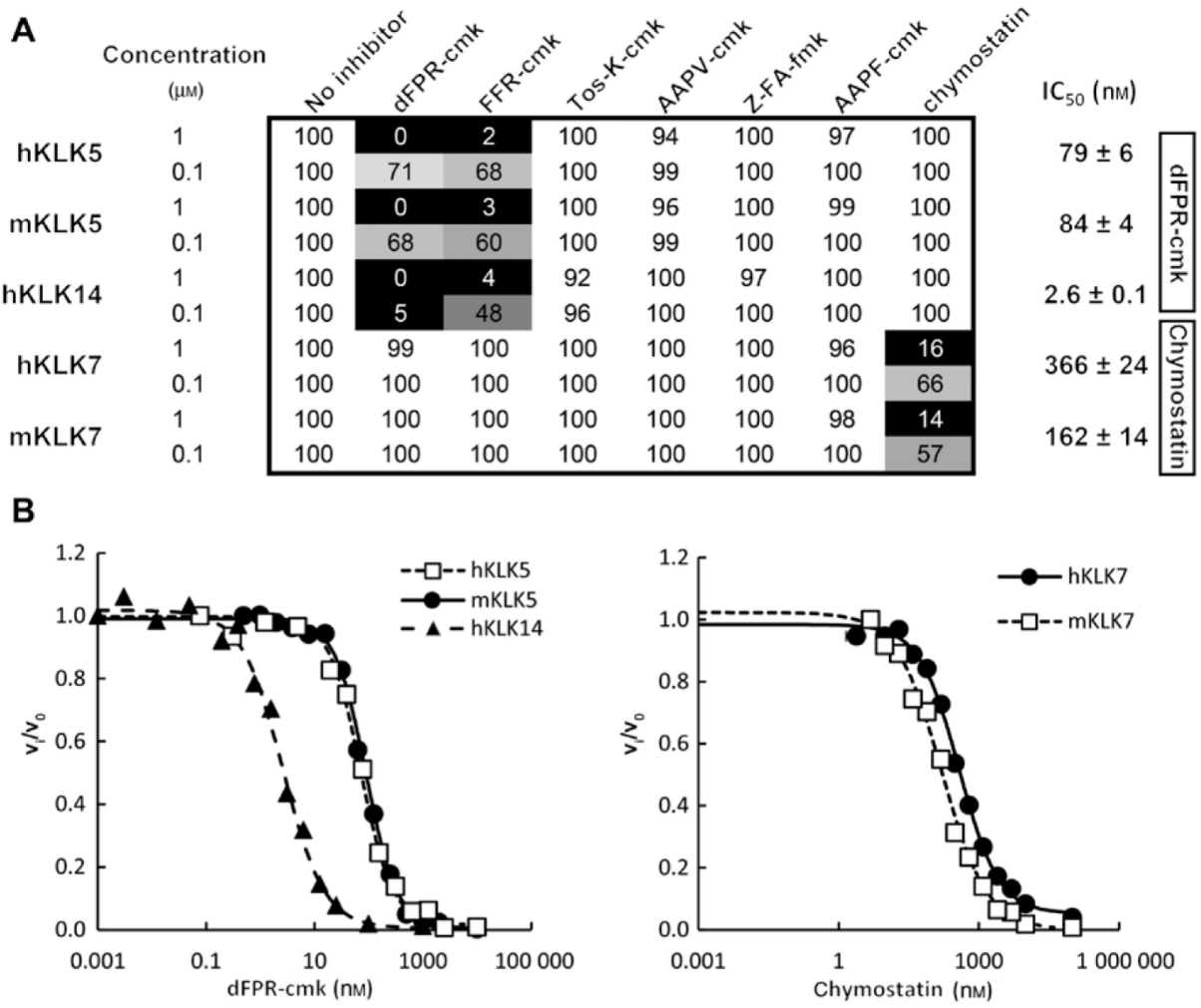


Figure 1: Profiling of recombinant KLKs with synthetic reactive peptide inhibitors of trypsin-like and chymotrypsin-like proteases.
KLK activities were measured with specific fluorogenic substrates (Boc-VPR-amc for KLK5, Ac-WAVR-acc for KLK14, and KHLY-acc for KLK7) in 0.1 M Tris-HCl pH 8.0, 150 mM NaCl in the presence of inhibitors (concentrations are indicated). (A) Activities are expressed as remaining proteolytic activity relative to the uninhibited control (100%) and displayed as a grayscale heat map. (B) For IC_{50} determination, KLKs were incubated with inhibitors (0–10 μM final concentrations) for 10 min and the reaction was initiated by the addition of substrate. Activities are expressed as ratio of proteolytic activity in the presence of inhibitor (v_i) to activity of the uninhibited control (v_0). The mean values of three replicates are shown (SE values are within 10% of the mean). IC_{50} values were determined by nonlinear regression using the GraFit software (Erithacus Software).

Three substrates (Boc-VPR-amc, Ac-WAVR-acc, KHLY-acc) identified as selective for individual KLKs were employed for analysis of KLK activities in skin homogenates. Tissue samples were prepared from eight gene-deficient mice strains, including mice deficient for KLK 5, 7 and 14 ($Klk5^{-/-}$, $Klk7^{-/-}$, $Klk14^{-/-}$), and KLK5/KLK14 double-deficient mice ($Klk5^{-/-}Klk14^{-/-}$) (Kasparek et al., 2017). All strains were crossed to a strain with null mutation in KLK-inhibitor LEKTI ($Sp5^{-/-}$), which results in non-modulated activities of epidermal KLKs and thus provides an excellent model system for their analysis. Because some of the strains survive only several hours after birth (Kasparek et al., 2017), all samples were taken at postnatal day 0.

KLK7 chymotrypsin-like activities measured with selective substrate KHLY-acc displayed dramatic changes corresponding to mouse genotype (Figure 2). $Klk7^{-/-}$ mice had reduced activity to about 40% compared to wild type (wt), and in $Sp5^{-/-}$, activity was highly increased to over 200%. Deficiency of trypsin-like KLKs (KLK5 and KLK5/KLK14) had no effect on KLK7 activity at this time point, indicating selectivity of the used substrate.

Trypsin-like KLK5 and KLK14 activities measured with Boc-VPR-amc substrate substantially increased signal determined in $Sp5^{-/-}$, which reached 145% of wt (Figure 2). KLK14 contributed to this total activity very mildly as $Klk14^{-/-}Sp5^{-/-}$ showed only 10% decrease compared to $Sp5^{-/-}$. In contrast, most of the activity came from KLK5: (a) in $Klk5^{-/-}Sp5^{-/-}$, the signal was reduced by about one third compared to $Sp5^{-/-}$, reaching the same level as $Klk5^{-/-}$ single KO. $Klk5^{-/-}Klk14^{-/-}Sp5^{-/-}$ triple-deficient mice showed additive reduction to about one half compared to $Sp5^{-/-}$, again to the level in $Klk5^{-/-}Klk14^{-/-}$. Deficiency in chymotrypsin-like KLK7 had no effect in line with selectivity of the used substrate. The results also show that there is substantial background of activity detected by this substrate originating from other proteases than KLK5 and KLK14 (see a final level of signal in $Klk5^{-/-}$ and $Klk5^{-/-}Klk14^{-/-}$ samples). The main candidate could be a matriptase, which is not inhibited by LEKTI (Sales et al., 2010).

Activities of KLK14 specifically measured with Ac-WAVR-acc were elevated to about 160% in $Sp5^{-/-}$ animals over wt (Figure 2). This activity can be attributed to KLK14, because $Klk14^{-/-}Sp5^{-/-}$ reduced the signal below the wt level. Interestingly, $Klk5^{-/-}Sp5^{-/-}$ displayed 70% activity of $Sp5^{-/-}$. This is most likely caused by the fact that KLK5 is responsible for processing the enzymatically inactive KLK14 zymogen into active enzyme (Yoon et al., 2007). Thus, KLK14 is not effectively activated in the absence of KLK5. The substrate Ac-WAVR-acc, similarly to Boc-VPR-amc, produces considerable background of non-KLK activities.

In conclusion, we demonstrated that the evaluated selective substrates can serve as a reliable tool for profiling proteolytic activities of skin KLKs and their quantification

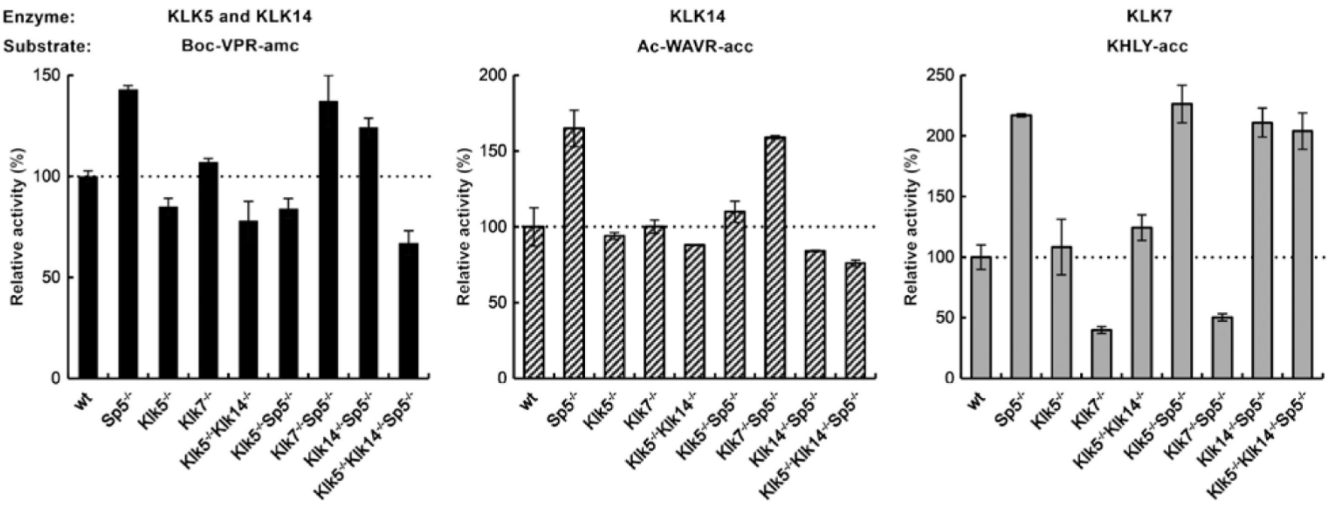


Figure 2: Profiling of KLK activities in the skin of gene-deficient mice.
KLK activities were measured with selective fluorogenic substrates targeting individual KLKs as indicated. Skin homogenates were prepared from mice deficient in one or more *Klk* genes and with null mutation in the gene *Spink5* ($Sp5^{-/-}$). Skin samples were crushed in liquid nitrogen, and soluble protein extracts were prepared by homogenisation in 0.1 M Tris-HCl, pH 8.0, 150 mM NaCl, 1% CHAPS, 1 mM EDTA and 1 μM cysteine protease inhibitor K11777. The extracts, cleared by centrifugation, were applied in the same protein concentration (0.2–0.5 mg/ml). Proteolytic activities were measured in a kinetic assay at pH 8.0 as described in Table 1, and serine protease activities sensitive to inhibition by 1 mM Pefabloc SC (targeting trypsin- and chymotrypsin-like proteases) are shown. The mean values ± SE of three replicates are normalised to the wild type animals (wt set as 100%).

in pathological conditions of gene-deficient mice. In addition, these results can further elucidate the roles of KLKs in epidermal proteolytic network that is contributed by several types of proteolytic enzymes.

Acknowledgements: The study was supported by projects InterBioMed LO1302 to M.H., J.S. and M.M., LM2015040 'Czech Centre for Phenogenomics' and LQ1604 'National Sustainability Program II (NPUII)' to R.S. funded by the Ministry of Education, Youth and Sports and CZ.1.05/2.1.00/19.0395 grant for 'Higher quality and capacity for transgenic models' to R.S. funded by the Ministry of Education, Youth and Sports and the European Regional Development Fund., by Benevolent fund for support of science – Neuron to P.K. and also by Academy of Sciences of the Czech Republic (Funder Id: 10.13039/501100004240) (RVO 61388963 and RVO 68378050).

References

Beaufort, N., Plaza, K., Utzschneider, D., Schwarz, A., Burkhart, J.M., Creutzburg, S., Debela, M., Schmitt, M., Ries, C., and Magdolen, V. (2010). Interdependence of kallikrein-related peptidases in proteolytic networks. Biol. Chem. *391*, 581–587.

Borgono, C.A., Michael, I.P., Komatsu, N., Jayakumar, A., Kapadia, R., Clayman, G.L., Sotiropoulou, G., and Diamandis, E.P. (2007). A potential role for multiple tissue kallikrein serine proteases in epidermal desquamation. J. Biol. Chem. *282*, 3640–3652.

Caubet, C., Jonca, N., Brattsand, M., Guerrin, M., Bernard, D., Schmidt, R., Egelrud, T., Simon, M., and Serre, G. (2004). Degradation of corneodesmosome proteins by two serine proteases of the kallikrein family, SCTE/KLK5/hK5 and SCCE/KLK7/hK7. J. Invest. Dermatol. *122*, 1235–1244.

Chao, J., Bledsoe, G., Yin, H., and Chao, L. (2006). The tissue kallikrein-kinin system protects against cardiovascular and renal diseases and ischemic stroke independently of blood pressure reduction. Biol. Chem. *387*, 665–675.

de Veer, S.J., Swedberg, J.E., Parker, E.A., and Harris, J.M. (2012). Non-combinatorial library screening reveals subsite cooperativity and identifies new high-efficiency substrates for kallikrein-related peptidase 14. Biol. Chem. *393*, 331–341.

de Veer, S.J., Furio, L., Harris, J.M., and Hovnanian, A. (2014). Proteases: common culprits in human skin disorders. Trends Mol. Med. *20*, 166–178.

de Veer, S.J., Furio, L., Swedberg, J.E., Munro, C.A., Brattsand, M., Clements, J.A., Hovnanian, A., and Harris, J.M. (2017). Selective substrates and inhibitors for kallikrein-related peptidase 7 (KLK7) shed light on KLK proteolytic activity in the stratum corneum. J. Invest. Dermatol. *137*, 430–439.

Descargues, P., Deraison, C., Bonnart, C., Kreft, M., Kishibe, M., Ishida-Yamamoto, A., Elias, P., Barrandon, Y., Zambruno, G., Sonnenberg, A., et al. (2005). Spink5-deficient mice mimic Netherton syndrome through degradation of desmoglein 1 by epidermal protease hyperactivity. Nat. Genet. *37*, 56–65.

Fajtova, P., Stefanic, S., Hradilek, M., Dvorak, J., Vondrasek, J., Jilkova, A., Ulrychova, L., McKerrow, J.H., Caffrey, C.R., Mares, M., et al. (2015). Prolyl oligopeptidase from the blood fluke *Schistosoma mansoni*: from functional analysis to anti-schistosomal inhibitors. PLoS Negl. Trop. Dis. *9*, e0003827.

Kasparek, P., Krausova, M., Haneckova, R., Kriz, V., Zbodakova, O., Korinek, V., and Sedlacek, R. (2014). Efficient gene targeting of the Rosa26 locus in mouse zygotes using TALE nucleases. FEBS Lett. *588*, 3982–3988.

Kasparek, P., Ileninova, Z., Zbodakova, O., Kanchev, I., Benada, O., Chalupsky, K., Brattsand, M., Beck, I.M., and Sedlacek, R. (2017). KLK5 and KLK7 ablation fully rescues lethality of Netherton syndrome-like phenotype. PLoS Genet. *13*, e1006566.

Lundwall, A. and Brattsand, M. (2008). Kallikrein-related peptidases. Cell Mol. Life Sci. *65*, 2019–2038.

Sales, K.U., Masedunskas, A., Bey, A.L., Rasmussen, A.L., Weigert, R., List, K., Szabo, R., Overbeek, P.A., and Bugge, T.H. (2010). Matriptase initiates activation of epidermal pro-kallikrein and disease onset in a mouse model of Netherton syndrome. Nat. Genet. *42*, 676–683.

Stopkova, R., Klempt, P., Kuntova, B., and Stopka, P. (2017). On the tear proteome of the house mouse (*Mus musculus musculus*) in relation to chemical signalling. Peer J. *5*, e3541.

Tamura, H., Ishikawa, Y., Hino, N., Maeda, M., Yoshida, S., Kaku, S., and Shiosaka, S. (2006). Neuropsin is essential for early processes of memory acquisition and Schaffer collateral long-term potentiation in adult mouse hippocampus *in vivo*. J. Physiol. *570*, 541–551.

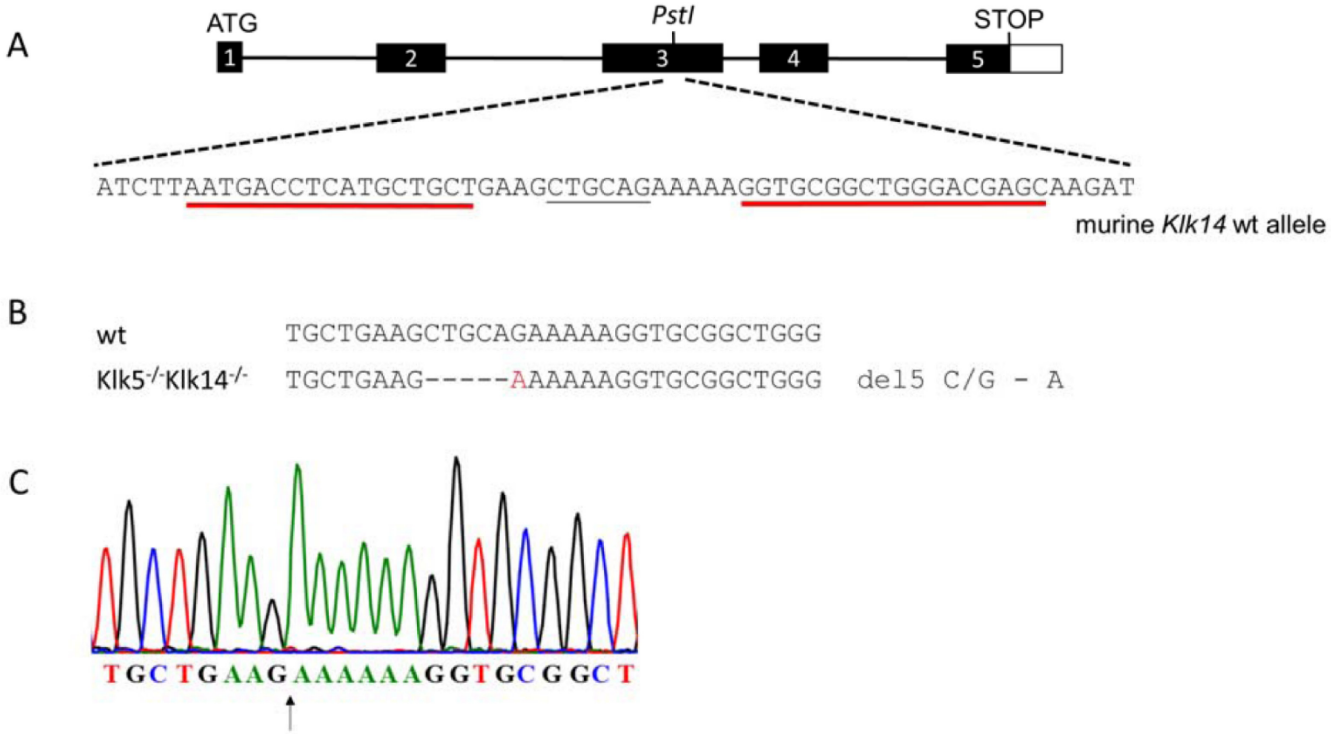
Yoon, H., Laxmikanthan, G., Lee, J., Blaber, S.I., Rodriguez, A., Kogot, J.M., Scarisbrick, I.A., and Blaber, M. (2007). Activation profiles and regulatory cascades of the human kallikrein-related peptidases. J. Biol. Chem. *282*, 31852–31864.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/hsz-2018-0116).

Profiling system for skin kallikrein proteolysis applied in gene deficient mouse models

Martin Horn, Olga Zbodakova, Petr Kasparek, Jaroslav Srp, Radka Haneckova, Martin Hradilek, Michael Mares and Radislav Sedlacek

Supplementary material



Supplementary Figure 1 Generation of Klk5$^{-/-}$Klk14$^{-/-}$ mutant mice.

(A) Schematic representation of *Klk14* gene structure. To generate Klk5$^{-/-}$Klk14$^{-/-}$ mutant mice, TALE nucleases targeting exon3 of *Klk14* gene were designed (TALEN-binding sites marked by red underline). Talen assembly and preparation of mRNA was performed as described previously (Kasparek et al., 2014). Klk14-TALEN encoding mRNA was microinjected into zygotes obtained from Klk5$^{-/-}$ deficient mutant line (Kasparek et al., 2017). Mice carrying mutations in Klk14 coding sequence were analyzed by PCR using primers Klk14-g-F: 5´ACAAGAGGCCACAAAATGCT 3´ and Klk14-g-R: 5´ ACCCAGGAGTCCAACATCAG 3´ and subsequent digestion using *PstI* enzyme (PstI sequence is marked by black underline) (B) Sequence of the target locus in wt mouse and in Klk5/Klk14 mutant founder, which was selected for establishing Klk5$^{-/-}$Klk14$^{-/-}$ mutant mouse line. Resulting mutation consisted of 5bp deletion and one nucleotide substitution, causing frame shift and premature STOP codon in *Klk14* coding sequence. (C) Chromatogram showing DNA sequence of *Klk14* target site, which was obtained by sequencing of cDNA isolated from Klk5$^{-/-}$Klk14$^{-/-}$ mutant mouse.

Position of del5 mutation is marked by black arrow. PCR product was amplified using primers Klk14-cDNA-F: 5´-CCTGGGCAAGCACAACAT-3´ and Klk14-cDNA-R: 5´-ACTGCAGAGCAGTGGGGTAC-3´.

Reference

Kasparek P., Krausova M., Haneckova R., Kriz V., Zbodakova O., Korinek V. and Sedlacek R. (2014). Efficient gene targeting of the Rosa26 locus in mouse zygotes using TALE nucleases. FEBS Lett. 588, 3982-3988.

4.6. Publikace č. 5: Crystal structures of the complex of kallikrein inhibitor BbKI with trypsin and modeling of kallikrein complexes.

Souhrn

Rukopis je v recenzním řízení v časopise Acta Crystallographica Section D.

Proteasový inhibitor BbKI izolovaný ze semen bauhinií (*Bauhinia bauhinioides*) patří mezi rostlinné inhibitory z Kunitzovy rodiny. Již dříve byla publikována jeho schopnost inhibovat aktivitu serinových proteas trypsinu, plasminu a plasmového kalikreinu (Oliva a Sampaio 2008) a také prostorová struktura volného BbKI (Zhou et al. 2015).

V předkládaném manuskriptu byl inhibitor BbKI připraven jako rekombinantní protein v bakteriálním expresním systému ve dvou formách, nemutovaný a mutovaný v pozici 54 (L54R). Analýza inhibiční specifity inhibitoru BbKI se zástupci lidských a myších serinových proteas kalikreinového typu (KLK) a s modelovými proteasami trypsinem a chymotrypsinem prokázala výrazné rozdíly v inhibiční specifitě a afinitě k BbKI. Nejlepší inhibiční konstanty s hodnotami K_i v sub-nanomolární oblasti byly stanoveny pro medicinálně významné proteasy KLK 4, KLK 7 a plasmový KLK. Z této analýzy vyplývá, že BbKI je nejvíce efektivní inhibitor lidských KLK proteas z dosud známých přirozených inhibitorů proteinového charakteru.

Byla vyřešena krystalová struktura komplexu BbKI s trypsinem na rozlišení 2,0 Å, která ukázala, že reaktivní centrum BbKI se váže do aktivního místa enzymu pomocí jedné interakční smyčky. Tento inhibiční mechanismus byl již dříve popsán u jiných komplexů rostlinných inhibitorů z Kunitzovy rodiny, jedná se o tzv. kanonický nebo Laskowského mechanismus interakce. Klíčová interakce v S_1 podmístě enzymu je zprostředkována mezi Arg64 na straně BbKI a Asp189 na straně trypsinu. Řada dalších kontaktů uvnitř i mimo tuto interakční smyčku stabilizují celý komplex.

Na základě vyřešené prostorové struktury komplexu BbKI s modelovou serinovou proteasou trypsinem byla studována interakce BbKI s KLK proteasami. S využitím tohoto templátu a prostorových struktur KLK 4, KLK 7 a plasmového KLK (dostupných v PDB databázi) byly modelovány jednotlivé komplexy a analyzována interakce s BbKI. Byla nalezena konformační flexibilita reaktivního centra BbKI a přídatné determinanty jeho okolí umožňující zvýšený počet kontaktů s enzymem, což vysvětluje vysokou afinitu BbKI ke studovaným KLK proteasám.

Můj podíl na práci zahrnoval měření inhibičních konstant K_i pro BbKI s devíti zástupci serinových proteas KLK typu a modelových serinových proteas trypsinu a chymotrypsinu a dále analýzu naměřených enzymologických dat.

Publikace č. 5

Li M., **Srp J.**, Gustchina A., Dauter Z., Mareš M., Wlodawer A.

Crystal structures of the complex of kallikrein inhibitor BbKI with trypsin and modeling of kallikrein complexes

Rukopis je v recenzním řízení v časopise *Acta Crystallographica Section D*, (2018)

Crystal structures of the complex of kallikrein inhibitor BbKI with trypsin and modeling of kallikrein complexes

Mi Li[1,2], Jaroslav Srp[3,4], Alla Gustchina[1], Zbigniew Dauter[5], Michael Mares[3] and Alexander Wlodawer[1*]

[1]Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

[2]Basic Science Program, Leidos Biomedical Research, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA

[3]Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Prague, Czech Republic

[4]Department of Biochemistry, Faculty of Science, Charles University, 12840 Prague, Czech Republic

[5]Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, NCI, Argonne National Laboratory, Argonne, IL 60439, USA

[*]Corresponding author:

E-mail: wlodawer@nih.gov

Synopsis

Crystal structures of the complex of a serine protease inhibitor from *Bauhinia bauhinioides* (BbKI) with bovine trypsin were determined in two crystal forms, serving as templates for modeling structures of the complexes with plasma and tissue kallikreins.

Abstract

Structures of a recombinant Kunitz-type serine protease inhibitor from *Bauhinia bauhinioides* (BbKI) complexed with bovine trypsin were determined in two crystal forms. Crystal structure utilizing the L55R mutant of BbKI was determined at 1.94 Å resolution in space group $P6_4$, and in the monoclinic space group $P2_1$ at 3.95 Å. An asymmetric unit of the latter crystals contains 44 independent complexes, thus representing one of the largest number of independent objects deposited in the Protein Data Bank. Additionally, the structure of the complex utilizing native BbKI was determined at 2.0 Å resolution in the $P6_4$ crystals isomorphous to those of the mutant. Since BbKI has been previously found to be a potent inhibitor of trypsin-like plasma kallikrein, it was tested against several tissue kallikreins as well. We have found that BbKI is a potent inhibitor of tissue kallikrein 4 (KLK4) and of the chymotrypsin-like human tissue kallikrein 7 (KLK7). We modeled structures of BbKI complexed with the catalytic domain of human plasma kallikrein, as well as KLK4 and KLK7 and analyzed them in order to identify the interactions responsible for its inhibitory potency.

1. Introduction

The Kunitz family of plant protease inhibitors shares its β-trefoil fold with a variety of proteins with different biological functions, including lectins, interleukins, fibroblast growth factors, as well as fragments of DNA-binding and actin-crosslinking proteins (Renko *et al.*, 2012). The inhibitor family was named after the discoverer of the soybean trypsin inhibitor (STI), which he isolated and crystallized (Kunitz, 1947). The inhibitors of the Kunitz-type are usually (but not always) single-chain proteins containing 160-180 amino acid residues, although they may also be domains of larger proteins. These inhibitors usually contain one or two disulfides, but some have no cysteine residues (Oliva *et al.*, 2010, Oliva & Sampaio, 2008). Structures of a number of Kunitz-type inhibitors have been determined to date (Renko *et al.*, 2012). The family members contain a very similar core with pseudo-threefold symmetry, whereas they differ quite significantly in the structure of the loops connecting the central β strands. The functional properties of these inhibitors are largely defined by the structure of their variable parts.

Inhibitory properties of the plant-derived Kunitz-type inhibitors have been studied very extensively. They were determined for the enzymes involved in digestive processes, such as trypsin and chymotrypsin, for a number of blood clotting enzymes, such as plasma kallikrein, factor XIIa, factor Xa, and thrombin, and for the enzymes involved in inflammatory processes, such as elastase (Batista *et al.*, 1996, Souza-Pinto *et al.*, 1996, Oliva & Sampaio, 2008, Vadivel *et al.*, 2014, Odei-Addo *et al.*, 2014).

BbKI is a Kunitz-type inhibitor originally isolated from the seeds of *Bauhinia bauhinioides*. This 18 kDa protein contains 165 amino acids with only one free cysteine, and forms no disulfides. Originally designated as BbTI-II (Oliva *et al.*, 1999), it was shown to be a potent inhibitor of trypsin, as well as plasmin and human and porcine plasma kallikrein, with no detectable inhibition of thrombin or Factor Xa (Oliva *et al.*, 1999, Oliva *et al.*, 2001).

BbKI, or its recombinant form rBbKI (hereafter both are referred to as BbKI), become attractive molecules for studying pathological models of the circulatory system, since this protein acts on enzymes involved in coagulation, fibrinolysis, and inflammation. One of the targets of BbKI is human plasma kallikrein, an enzyme that participates in the process of blood coagulation, platelet aggregation, and muscle contraction (Botos & Wlodawer, 2007, Pampalakis & Sotiropoulou,

2007, Turk, 2006). Indeed, BbKI was effective in inhibiting the viability of tumor cell lines (Nakahata *et al.*, 2013). Other important aspects of the inhibitor were reported by (Brito *et al.*, 2014), who found that BbKI may prolong formation of blood clots *in vitro* and that it exhibits antithrombotic activity in venous and arterial thrombosis *in vivo* models. To the best of our knowledge, BbKI is the only plant-derived inhibitor of both plasma and tissue kallikreins (Oliva *et al.*, 2001, Oliva *et al.*, 2010).

In this study we evaluated inhibitory properties of BbKI towards several kallikreins. We determined the high-resolution crystal structure of BbKI complexed with bovine trypsin that was subsequently used to build models of the inhibitor complexes with selected kallikreins. These models were analyzed in order to reveal the important interactions that might be responsible for the differences in the inhibitory potency of BbKI towards various kallikreins.

2. Materials and methods

2.1. Preparation and purification of BbKI and its complex with bovine trypsin

The ORF of BbKI was cloned from the previously used construct of BbKI (Araujo *et al.*, 2005) with primers (Operon) p300-BbKIf (5'-GAGAACCTGTACTTCCAGTCGGTCGTTGTCGACACCAATGG-3') and p301-BbKIr (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTATTACTCATCAGTTGCCTTCCTTATC-3'), using KOD Hot Start DNA polymerase (Novagen). The product was gel purified and used as template for 2nd PCR with primer PE277 (5' – GGGGACAAGTTTGTACAAAAAAGCAGGCTCGGAGAACCTGTACTTCCAG -3') and p301-BbKIr. This amplicon was inserted into pDONR201 using Gateway technology (Invitrogen) to generate entry clone pZD226. The entry clone was then recombined with the destination vector pCB1426 to construct His6-MBP-TEV-BbKI fusion protein expression vector pZD228. To facilitate protein purification, the plasmid pZD228 was modified using Quick Change with primer p303 (5'-GAGAACCTGTACTTCCAGGGTGGCGGTGGCGGTGGCTCGGTCGTTGTCGACACCAATGG-3') and p305 (5'-CCATTGGTGTCGACAACGACCGAGCCACCGCCACCGCCACCCTGGAAGTACAGGTTCTC-3'). This generated the construct His6-MBP-TEV-BbKI (pZD263) with six glycine residues

inserted before BbKI ORF. The plasmid was transformed into *E. coli* Rosetta 2 (DE3) cells for expression. The cells were grown in Luria–Bertani medium supplemented with 25 μg mL^{-1} chloramphenicol (Invitrogen) at 37 °C to an optical density $OD_{600\ nm}$ of 0.6, followed by induction of the expression of the fusion protein with isopropyl β-D-thiogalactopyranoside (Invitrogen). IPTG was added at a final concentration of 1 mM and the culture was grown overnight at 16 °C on an orbital shaker, at 180 RPM. Subsequently, the cells were harvested by centrifugation (6000 g, 20 min, 4 °C) (Thermo Scientific Evolution RC). The pellets were then resuspended in lysis buffer (0.05 M Tris/HCl, pH 8.0, 0.5 M NaCl, 0.5 mM MgCl, and 0.5 mM $CaCl_2$) and lysed by adding Novagen BugBuster at 1:50 volume ratio. After 40 min lysis at 4 °C, the lysis solution was centrifuged at 16000 RPM for 30 minutes at 4 °C (Sorvall RC). The supernatant was loaded onto Ni–NTA affinity chromatography column which was then eluted with an imidazole gradient (10–500 mM) in 20mM Tris/HCl, pH 7.5 buffer. Fractions containing the fusion protein were combined and dialyzed against 20 mM Tris/HCl, pH 8.0, and 0.2 M NaCl buffer overnight to remove the imidazole. TEV protease (1 mg for 8 g of cells) was added to dialyzed solution to cleave the MBP at 4 °C overnight. As a consequence of the introduced TEV cleavage site in the expression vector, the BbKI contains six extra glycine residues at the N terminus. The MBP (containing a His-tag) was removed by the second run on a Ni-NTA column. The flow-through containing BbKI was concentrated and further purified by size exclusion chromatography (Superdex 75 HR 10/30 column; GE Healthcare) in 0.02 M Tris/HCl, pH 7.5, and 0.2 M NaCl.

For preparation of the complex of BbKI with trypsin, equimolar amounts of BbKI and bovine pancreatic trypsin (Sigma 8003) were mixed and incubated on ice for 1 hour in 20 mM Tris pH 7.5, 0.200 M NaCl. The mixture was then applied onto a Superdex 75 column pre-equilibrated with the same buffer. Fractions of 1 ml volume were collected at a flow rate of 0.5 mL min^{-1}. The identity of samples pooled from different peaks was verified by SDS-PAGE. The fractions corresponding to the BbKI-trypsin complex were selected and concentrated to between 10 to 25 mg mL^{-1} and stored at -80 °C for crystallization trials.

2.2. BbKI mutagenesis

Mutation of Leu55 to arginine was introduced to pZD263 using Quick Change with primers p359 (5'-CTCACCACCGTCCCGGTCGTCCGGTTAGATTTGAATCC-3') and p260 (5'-GGATTCAAATCTAACCGGACGACCGGGACGGTGGTGAG-3') to generate pZD264. The

mutated gene of L55R BbKI was expressed and the resulting protein was purified in a procedure analogous to the one used for the native BbKI. The complex of BbKI L55R/trypsin was prepared and purified with the same protocol as for the native BbKI.

2.3. Enzymatic activity and inhibition assays

Proteolytic activities were measured with selective fluorogenic peptide substrates (**Table 1**) containing a fluorescent leaving group 7-amino-4-methylcoumarin (Amc) or 7-amino-4-carbamoylmethylcoumarin (Acc). The reaction mixture included enzyme and substrate at concentrations indicated in **Table 1**, and 0.1 M Tris-HCl pH 8.0, 0.1% PEG 1500, containing 0.15 M NaCl (for most KLKs), 1 M NaCl (for KLK3), or 10 mM $CaCl_2$ (for trypsin and chymotrypsin). The kinetics of product release was continuously monitored in an Infinite M1000 microplate reader (Tecan) at excitation wavelengths of 360 or 380 nm and emission wavelengths 465 or 460 nm for substrates with Amc and Acc groups, respectively. For inhibition measurements, the enzyme mixture was preincubated with the BbKI inhibitor (at up to 10 μM concentration) for 10 min followed by the addition of substrate. Reaction rates were obtained at various inhibitor concentrations, and IC_{50} values were determined by nonlinear regression using GraFit software (Erithacus Software). The K_i values were calculated using the equation for competitive inhibition $K_i = IC_{50} / (1 + ([S]/K_m))$; substrate concentrations [S] and K_m values are presented in **Table 1** (K_m values were determined by nonlinear regression of substrate-dependent velocity curves or obtained from literature data). Each measurement was performed in triplicate. The concentration of the BbKI inhibitors and substrates were determined by amino acid analysis. Kallikreins were purchased from R&D Systems and activated according to manufacturer's protocol, other enzymes were from Sigma. Substrates were purchased from Bachem; Ac-KHLY-Acc was synthesized as described in (Horn *et al.*, 2018).

2.4. Protein crystallization and X-ray data collection and processing

Two crystal forms of the BbKI/trypsin complex have been grown. Monoclinic crystals (space group $P2_1$) were grown by the hanging drop technique, starting from protein sample concentrated to 21 mg mL^{-1} in 20 mM Tris and 0.2 M NaCl, pH 7.5. The well solution contained 17.5% PEG3350 at pH 8.0. Each 4.0 μl hanging drop consisted of 2 μL sample and 2 μL well solution, and it was equilibrated against 500 μL well solution. Crystals of the L55R mutant of BbKI/trypsin

complex grew at the same condition as those of BbKI/trypsin crystals, the only difference being that the starting sample concentration was 13 mg mL^{-1}. Hexagonal crystals of the complex (space group $P6_4$) were grown by an analogous procedure, starting with a protein sample at 16.5 mg mL^{-1} concentration in the same buffer. However, the well solution contained 1.6 M ammonium sulfate pH 4.2. Each hanging drop was mixed using 4 µL of the sample and 2 µL of well solution, equilibrated against 500 µL well solution.

Diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID at the Advanced Photon Source, Argonne National Laboratory. Single crystals were transferred to a cryoprotectant solution (mother liquor supplemented with 25% glycerol) for approximately 2 min and then were flash frozen at 100 K in a stream of cold nitrogen gas. All three data sets were collected at the wavelength of 1.000 Å with a Rayonix 300HS detector. Diffraction data were indexed, integrated, and scaled with the program *HKL2000* (Otwinowski & Minor, 1997). Data processing statistics are shown in **Table 2**.

2.5. Structure determination and refinement

The structure of the L55R BbKI/Trypsin complex was solved first for the hexagonal crystals by molecular replacement using the program *PHASER* (McCoy *et al.*, 2007). The structures of BbKI (PDB ID 4zot; (Zhou *et al.*, 2015)) and of bovine trypsin (PDB ID 5eg4; (Hinkes *et al.*, 2016)) were used independently as search models to locate the six complexes present in the asymmetric unit. The structure was refined with *REFMAC5* (Murshudov *et al.*, 2011) in *CCP4* (CCP4, 1994)and rebuilt with *COOT* (Emsley & Cowtan, 2004). Crystals of the BbKI L55R-trypsin complex were isomorphous with those of the native BbKI, thus the former structure was used directly as a starting point for refinement of the latter one with *REFMAC5.* A single BbKI L55R-trypsin complex was used as a search model to solve the structure of the BbKI-trypsin complex in the monoclinic space group with a total of 44 complexes located within an asymmetric unit. The structure was subsequently refined with *PHENIX* (Adams *et al.*, 2010).

2.6. Modeling of kallikrein complexes

A model of a complex of BbKI with human plasma kallikrein was built by superimposing the coordinates of the protease domain of the enzyme (PDB ID 5tjx; (Li *et al.*, 2017)) on the trypsin molecule (A) in the L55R BbKI/trypsin complex. Complexes with human tissue kallikrein 4

(KLK4; PDB ID 4k8y; (Riley *et al.*, 2016)) and kallikrein 7 (KLK7; PDB ID 2qxi; (Debela *et al.*, 2007)) were prepared in an analogous manner. The coordinates of all three complexes were subjected to energy minimization of the contact region. For that, an all-atom version (including all the hydrogens) was created for each assembly using the *PSFGEN* plugin for *VMD* v1.9.1 (Humphrey *et al.*, 1996). All the charged residues were set in their default dissociation state at neutral pH. For the residues with two alternate locations in the original assemblies, only the second alternate conformation (i.e. conformation B) was included into the final structures for minimization. The disulfides present in the original structures were secured with explicit covalent bonds (C419-C435, C517-C584, C548-C563, C574-C602 cysteine crosslinks for 5tjx, and C22-C157, C42-C58, C128-C232, C136-C201, C168-C182, C191-C220 for 4k8y and 2qxi). Only a narrow interfacial region (10 Å layer centered at the interface) was allowed to move and adjust during the minimization- i.e. all the atoms at the interface within 5 Å cutoff from any atom of the other contact partner. The rest of the system was fixed.

The energy minimizations were done using *NAMD* package (Phillips *et al.*, 2005), with *CHARMM36* force field (MacKerell *et al.*, 1998, Klauda *et al.*, 2010) in vacuum, using conjugate energy gradient algorithm, with 12 Å cutoff for non-covalent interactions. We have performed 1000 steps of energy minimization, which was sufficient to resolve all the steric conflicts (thus making structures compatible for use in Molecular Dynamics simulations without further adjustments, but still avoiding unnatural longer-distance rearrangements of the polar contacts that occur at longer minimizations due to the absence of an explicit solvent).

3. Results and Discussion

3.1 Crystallization of the complex of BbKI with bovine trypsin

The recombinant form of BbKI used in determining the structure of free inhibitor (Zhou *et al.*, 2015) included two serine residues at the N terminus (residues 19 and 20 of the precursor; GenBank: AAR01967.1). Although Ser19 of the precursor is most likely part of the signal peptide and is not present in the native form of the inhibitor, it was retained in its recombinant form due to the way the protein was cloned. Since that residue was well ordered in the crystal structure, it was previously retained in the numbering scheme and thus the reported sequence of BbKI started

with Ser1-Ser2-Val3... This numbering scheme of residues is followed in this work and the "native BbKI" referred to below has an extra Ser1 at its N terminus.

Our initial crystallization experiments utilized the complex of native BbKI with trypsin at pH 7.5 and resulted in the growth of the monoclinic crystals with very large volume of a unit cell (see below). An initial estimate was that at least 30 molecules of the complex could be present in an asymmetric unit, and all efforts of determining the structure by molecular replacement using separately the coordinates of BbKI (PDB ID 4zot) and bovine trypsin (PDB ID 5eg4) were unsuccessful. In a search for a different crystal form we prepared the L55R mutant of BbKI and continued crystallization efforts using this version of the inhibitor.

3.2. Hexagonal crystals of the BbKI-trypsin complex

Hexagonal crystals of the BbKI-trypsin complex were initially grown at low pH for the L55R mutant, and subsequently, under the same conditions, also for the complex utilizing the native inhibitor. Crystals of the former complex grew in the hexagonal space group $P6_4$ with the unit cell parameters $a=b=207.30$ Å, $c=107.17$ Å and diffracted beyond 2 Å resolution. An analysis of the Matthews parameter suggested the presence of six complexes in an asymmetric unit. The structure was solved by molecular replacement (see Materials and Methods) and refined with *REFMAC5* (**Table 2**). The electron density is clear and continuous for all molecules in the unit cell and the geometry of the model is acceptable. The isomorphous structure of the complex with native BbKI was subsequently solved by direct refinement starting with the coordinates of the complex utilizing the L55R mutant. With the exception of the mutated residues, clearly visible in the electron density in both structures, no major differences between these two coordinate sets are present. Since the structure of the L55R complex was determined at a slightly higher resolution and the differences between the six complexes in an asymmetric unit are minor, the coordinates of the complex consisting of molecule A of trypsin and molecule G of L55R-BbKI (**Figure 1**) will be used in all comparisons reported below. The RMSD's for the other complexes compared to AG are 0.17 Å (BH), 0.10 Å (CI), 0.17 Å (DJ), 0.05 Å (EK), and 0.17 Å (FL), justifying this choice.

The structures of both components of the complex are very similar to their counterparts in the available structures of the individual proteins. Superposition of the bovine trypsin molecule of the complex on the highest-resolution structure of that enzyme (PDB ID 4i8h) yields RMSD of 0.55

Å for all 223 Cα atoms. Deviations exceeding 1 Å are found in only three loops – 76-79, 96-98, and 114-117 (chymotrypsin numbering is used throughout). In many published structures of bovine trypsin, including the one used in this comparison, loop 76-79 is stabilized by a calcium ion, but calcium was not present in the crystallization buffers used to obtain crystals of the complex and the ions are replaced by putative water molecules. The attempts to soak calcium into crystals of the complex were unsuccessful, and adding calcium chloride to crystallization buffer did not yield any crystals. The reasons for this behavior are not known. Loop 114-117 is located at the surface of the protein and its conformation may be influenced by crystal contacts, whereas loop 96-98 is found at the interface between the enzyme and the inhibitor (see below), and thus its conformation may be influenced by the presence of intermolecular interactions.

Superposition of the native BbKI on molecule G of the complex yields an RMSD of 0.78 Å for all 163 Cα atoms of the inhibitor. Significant differences are only present in several loops, with the surface loops 26-28, 36-38, 49-50, 106-107 and 140-144 exhibiting maximum deviations between the positions of Cα atoms, up to 3 Å. Two loops at the interface between the inhibitor and the enzyme (82-84 and 128-131) were also shifted by up to 2.8 Å.

3.3. Monoclinic crystals of the BbKI-trypsin complex

The cell dimensions of this crystal form, a=137.2, b=483.9, c=137.0 Å, β=116.8°, and systematic absences with k=2n+1 for 0,k,0 reflections point to the monoclinic space group $P2_1$, but the close similarity of the lengths of the a and c unit cell axes potentially could lead to the orthorhombic symmetry and $C222_1$ space group, or indicate the existence of pseudo-merohedral twinning. However, all symmetry and twinning tests performed with *XTRIAGE* excluded these possibilities and it was therefore assumed that, indeed, this crystal is monoclinic and not twinned.

The cell volume of this crystal form suggested the presence of a large number of independent molecules in the cell, but initially it was not clear exactly how many. Molecular replacement runs with *PHASER* (McCoy *et al.*, 2007), using the structure of the L55R BbKI/trypsin complex determined in this study (PDB ID 6dwf), first identified positions of 22 complexes, but inspection of the molecular packing suggested that this was not a complete structure. The successive runs of *PHASER* eventually completed the structure finding 44 complexes in the asymmetric unit of the

cell. The structure was refined with the TLS option, with each of the 88 individual protein chains assigned as a separate TLS domain. The final refinement statistics are summarized in **Table 2**.

The 44 heterodimers (trypsin:inhibitor complexes) are arranged in the cell in four columns, running along four independent 2_1 axes at 0,y,0; 0,y,½; ½,y,0; and ½,y,½. Within each pair of successive dimers along each column they are rotated with respect to each other by about 114° on average. This corresponds to a full turn rotation between the 1st and 23rd dimer, or 180° rotation between the 1st and 12th dimer along the column axis. Each column, therefore, is built by dimers arranged according to the approximate 22_7 screw axis, since 114° ~ 7×(360°/22). Two views of the contents of the asymmetric unit are shown in **Figure 2**.

A search of the PDB performed in January 2018 indicated that the number of biological units present in an asymmetric unit of these crystals is the largest among all structures deposited to date. Although some structures with as many as 60 molecules in an asymmetric unit are present in the PDB, they contain fewer biologically-relevant units. For example, crystals of lumazine synthase (PDB ID 3mk3; (Kumar *et al.*, 2011)) contain 60 protein molecules in an asymmetric unit, but only one biological unit, since in solution the protein forms icosahedral capsids. Similarly, several proteins with as many as 48 molecules in an asymmetric unit contain fewer biological units – for example DUTPase (PDB ID 4apz; (Garcia-Nafria *et al.*, 2013)) contains 16 trimers, and ethanolamine utilization microcompartment protein (PDB ID 3ia0; (Tanaka *et al.*, 2010)) contains 8 hexamers. However, we consider the properties of this particular crystal form of the BbKI/trypsin complex to be merely a crystallographic curiosity, indicating many ways that crystals might be formed. This structure was not analyzed further.

3.4. Interactions between the inhibitor and the enzyme

Three major epitopes on BbKI, each comprising several loops, are involved in contacts with bovine trypsin. The primary interactions involve the so-called “reactive loop” of the inhibitor (residues 61-66). The extensive interactions of Arg64, the key residue of the inhibitor that makes a strong ion pair with the carboxylate of Asp189 of trypsin, are of particular importance. The distance between NH1 of Arg64 and OD1 of Asp189 is 2.89 Å, whereas the distance between Arg64/NH2 and Asp189/OD2 is 2.71 Å. Other interactions of Arg64 include hydrogen bonds between Arg64/NH2 and Gly218/O (2.85 Å), as well as Arg64/NH1 and Ser190/OG (2.73 Å). Hydrogen bonds between Arg64/O and Gly193/N (2.74 Å), Arg64/N and Ser214/O (3.05 Å), Leu63/O and

Gln192/NE2 (2.94 Å), as well as Pro62/O and Gly216/N (3.12 Å) play a crucial role in positioning the main chain of the reactive loop in the active site of the enzyme and maximizing the interactions of the specificity-providing Arg64. Other interactions stabilizing the reactive loop involve putative hydrogen bonds between both OD1 and ND2 of Asn66 of BbKI and Tyr151/OH of trypsin.

A number of interactions outside of the area of the reactive loop help to stabilize the complex and provide additional specificity. The guanidinium group of Arg129 of BbKI interacts with the side chain of Gln175 of trypsin, as well as with Asn97/O. Lys69/NZ and Tyr72/OH of BbKI are hydrogen bonded to the carbonyl oxygens of Tyr59 and His57 of trypsin, respectively, with the former residue making hydrogen bonds to both carbonyls. Additional non-polar interactions involve a number of residues of both proteins and presumably contribute to the stability of the complex.

3.5. A comparison with the structures of trypsin complexes of related inhibitors

Crystal structures of trypsin complexes of four inhibitors related to BbKI have been previously published. They are soybean trypsin inhibitor STI (PDB ID 1avw; (Song & Suh, 1998)), tamarind Kunitz inhibitor TKI (PDB ID 4an7; (Patil *et al.*, 2012), *Enterolobium contortisiliquum* inhibitor EcTI (PDB ID 4j2y; (Zhou *et al.*, 2013), and an engineered version of the winged bean chymotrypsin inhibitor WCI, in which the specificity-determining Leu65 was mutated to arginine (PDB ID 3veq; (Majumder *et al.*, 2012)). STI and TKI were complexed with porcine trypsin, whereas EcTI and WCI were complexed with bovine trypsin, similarly to BbKI. (Parenthetically, trypsin sequence in 3veq has an error by identifying Ile118 as valine, unlike in the other structures involving cationic bovine trypsin – the electron density clearly supports the presence of isoleucine). However, the differences between the source of trypsin used in the structural studies seem to be immaterial, since the RMSDs after superposition of the trypsin molecules of the complexes were 0.52 Å, 0.53 Å, 0.55 Å, and 0.48 Å for all Cα atoms in the complexes with STI, TKI, EcTI, and WCI, respectively. The reactive loops of the inhibitors in all five complexes adopt very similar conformations and superimpose very well when trypsin is used as a reference, although such trypsin-based superposition leads to comparatively large shifts in most of the inhibitor molecules. That is happening despite the fact that the inhibitors superimpose quite well by themselves. The RMSDs are 1.44 Å for the superposition of 143 Cα atoms of STI (sequence identity 32.2%), 1.55 Å for the superposition of 143 Cα atoms of TKI (sequence identity 36%),

1.35 Å for the superposition of 145 Cα atoms of EcTI (sequence identity 34.5%), and 1.65 Å for the superposition of 148 Cα atoms of WCI (sequence identity 32.5%). In all these comparisons we find very accurate overall superposition of the β sheets of the inhibitors, whereas many loops, including the reactive loops, are not well superimposed. The only way to keep the reactive loops superimposed is by rigid body movement of the whole inhibitors in relation to the enzyme, thus affecting their interactions in the areas distant from the active site. However, although rigid-body rotation was previously invoked as a potential determinant of secondary specificity of interactions (Zhou *et al.*, 2013), these adjustments of the quaternary structures may also be due to crystal contacts, as suggested by 7.25° rotation of STI in the two high-resolution structures of the complex in different crystal forms (Song & Suh, 1998). A major difference between BbKI and the other four inhibitors is observed in the orientation of the loop 80-83, facing the enzyme in the former inhibitor and facing away in the latter ones.

3.6. Inhibitory specificity of BbKI

We investigated the inhibitory specificity of BbKI using a panel of nine kallikreins of human and mouse origin, in addition to trypsin and chymotrypsin, the latter enzymes representing archetypal serine proteases from the S1 family. BbKI was screened using a kinetic activity assay and the inhibition constants K_i were determined (**Table 3**). The results show that BbKI is a potent inhibitor of human kallikreins KLK4 (K_i ~0.045 nM), KLK7 (K_i ~0.18 nM), and plasma pKLK (K_i ~12.9 nM). In contrast, it displayed only weak inhibition of KLK5 and KLK14 (Ki > 0.5 and 5 μM, respectively) and no detectable inhibition of KLK2 and KLK3. Relative sensitivity of mouse KLK7 and KLK5 to BbKI inhibition was rather similar to that of their human orthologs. Furthermore, we compared the inhibitory profiles of both the native BbKI and its L55R mutant, indicating that this substitution did not influence the binding properties to serine proteases in any significant way (**Table 3**).

These results demonstrated that BbKI is able to effectively inhibit kallikreins with trypsin-like cleavage specificities (pKLK and KLK4), as well as chymotrypsin-like enzymes (KLK7). This is in line with BbKI inhibition of both archetypal proteases, trypsin and chymotrypsin (K_i ~0.2 and 33 nM, respectively). A versatile reactive site loop with a basic P1 residue that is responsible for inhibition of trypsin and also chymotrypsin has been proposed for the plant Kunitz inhibitor STI (De Vonis Bidlingmeyer *et al.*, 1972) and later demonstrated crystallographically in the animal

Kunitz inhibitor BPTI (Helland *et al.*, 1999, Scheidig *et al.*, 1997). We provide evidence that BbKI is a highly potent inhibitor of human KLK4, effective in the low picomolar range. To our knowledge, BbKI is one of the best natural proteinaceous inhibitors of KLK4 (and kallikreins in general) with the non-covalent, Laskowski mechanism of action. A similar inhibitory potency was reported for a rationally designed derivative of the sunflower trypsin inhibitor (SFTI) interacting with KLK4 (Swedberg *et al.*, 2011).

3.7. Modeling of the interactions between BbKI and selected kallikreins

A model of the complex between human plasma kallikrein (pKLK) and BbKI was built by superimposing the coordinates of pKLK refined at the 1.4 Å resolution (PDB ID 5tjx; (Li *et al.*, 2017)) on those of trypsin in the L55R BbKI/trypsin complex (RMSD 1.11 Å for 205 Cα pairs). Unlike in the previous model of such a complex which was based on the structure of free BbKI (Zhou *et al.*, 2015), in the new energy-minimized model the specificity-determining residue on the reactive loop of BbKI, Arg64, is making excellent contacts with Asp572 of kallikrein, creating an ion pair **(Figure 3A)**. This result is largely due to the repositioning of the reactive loop upon complex formation with trypsin in the experimentally-determined structure. Part of the primary binding pocket consisting of residues 216-222 (chymotrypsin numbering) has two different conformations in various kallikreins, due to the presence of the inserted Pro218A in some of them **(Figure 3A)**.

Secondary contributions to specific binding of BbKI to pKLK are provided by the interactions between Lys69/NZ and Tyr72/OH of BbKI with the carbonyl oxygens of His434 and Cys435 of pKLK, respectively. The interaction pattern for the residues that follow is unique to pKLK, due to an insertion in the main chain of this enzyme. This is an area where the structures of trypsin and pKLK diverge significantly **(Figure 3B)**. Several other specific interactions between BbKI and pKLK are seen in the modeled complex. Strong charge-charge interactions are formed between Arg129 of BbKI and Glu479 of pKLK. Also, Arg129/NH1 is hydrogen bonded to Tyr555/OH of pKLK **(Figure 3B)**. Ser81 and Ser82 in BbKI interact extensively with the side chain of Lys528 of pKLK **(Figure 3C)**. The hydroxyl group of Ser82 is hydrogen bonded to Lys528, whereas Cβ atoms of both Ser81 and Ser82 make hydrophobic contacts with that residue in pKLK. This loop does not make similar contacts with trypsin. Some other interactions include an ion pair between Lys106 of BbKI and Glu600 of pKLK in another contact area. Close to the specificity-defining

Arg64, Asn66 of BbKI is hydrogen bonded to the guanidinium group of Arg416 in pKLK, as well as to a carbonyl oxygen of Leu418. The latter interaction is unique to this complex, since the trypsin counterpart of Arg416 in pKLK is Tyr39.

Since the inhibitory properties of BbKI evaluated with a panel of kallikreins were shown to be the best for human KLK4, a model of a complex between KLK4 and BbKI was prepared in a manner analogous to the model of a complex of pKLK, by superimposing the published atomic resolution model of KLK4 (PDB ID 4k8y; (Riley *et al.*, 2016)) on the coordinates of trypsin in the BbKI/trypsin complex (RMSD 1.09 Å for 215 Cα pairs). Comparative analysis of the interactions of BbKI in the experimental structure of the complex with trypsin and in the model of the complex with KLK4 does not provide clear justification for the much higher inhibitory potency of BbKI towards KLK4. The interactions around the P1 Arg64 of the inhibitor seem to be very similar for KLK4 and several other kallikreins, such as KLK2, 3, 5, 7 and 14, all containing the same Pro218A insertion in the protein fragment 216-222 **(Figure 4).** This fragment interacts with the region that includes the specificity determining ion pair between Arg64 (BbKI) and Asp189 of the enzyme. Insertion of Pro218A changes the conformation of this fragment in a similar way, creating an identical pattern of BbKI-kallikrein interactions in all these enzymes **(Figures 5A,3A)**. Kallikreins that contain Pro218A insertion gain additional hydrophobic contacts with the loop in BbKI carrying Ser82.

Three other fragments of BbKI, containing Val3, residues Lys69 and Tyr72, as well as the loop that includes Arg129 and Gly130, consistently maintain contacts with the corresponding fragments comprising residues 56-60, 95-98, and 175 in trypsin, KLK4, and all other kallikreins. Only two hydrogen bonds, that include the one between Lys69/NZ (BbKI) and Tyr59/O in trypsin, as well as the other between hydroxyl of Tyr72/OH (BbKI) and carbonyl oxygen of His57, are conserved in all modeled kallikrein complexes **(Figure 5B)**. The other interactions vary in different complexes. The hydrogen bond between Arg129 (BbKI) and Gln175 in trypsin is not present in the complexes with KLK4 due to the substitution Q175L.

In the absence of an experimental structure very little can be said about inhibitor interactions with residues 96-98 in KLK4. The structure of KLK4, which was used for modeling (PDB ID 4k8y), includes bound inhibitor SFTI-1. Loop 96-98 in that complex adopts a different conformation than in trypsin, KLKs 5 and 7, and pKLK due to collision with the bulky side chain of Phe12 of SFTI-

1. As a result, the tip of the loop is shifted to a different position **(Figure 5B)**. Therefore, the contact area for this loop in the complex with BbKI could be properly modeled.

The BbKI loop containing residues Ser81-Ser82, as well as two loops that include Asn66 and Ala16, respectively, make almost no contacts with the fragment of KLK4 comprising residues 147-152, whereas a corresponding fragment in KLK5 makes several interactions with the same loops **(Figure 5C)**, although BbKI is a very weak inhibitor of KLK5 (**Table 3**).

Whereas pKLK and KLK4 are trypsin-like enzymes, KLK7 is chymotrypsin-like, thus the reasons for the potent inhibitory activity of BbKI against this enzyme were not obvious. A model of a complex between KLK7 and BbKI was prepared in a manner analogous to the model of a complex of pKLK, by superimposing the published atomic resolution model of KLK7 (PDB ID 2qxi; (Debela *et al.*, 2007)) on the coordinates of trypsin in the BbKI/trypsin complex (RMSD 1.14 Å for 218 Cα pairs). A crucial difference between the KLK7 on one hand, and trypsin and some other kallikreins on the other hand, is in the presence of Asn in place of Asp189 (or Ser189 in chymotrypsin and KLK3). Both carboxyl oxygens of Asp189 in trypsin and other kallikreins are hydrogen bonded to NH1 and NH2 of Arg64 of BbKI. In KLK7 only the Asn189/OD1 is an acceptor of a hydrogen bond from Arg64/NH1, in addition to an intramolecular hydrogen bond provided by Gly220/N. Asn189/ND2 is stabilized by an intramolecular hydrogen bond to Lys188/O. Other interactions of Arg64 observed in the structure of BbKI-trypsin complex include hydrogen bonds between Arg64/NH1 and Ser190/OG (2.73 Å) and Arg64/NH2 and Gly218/O (2.85 Å). Both interactions are not preserved in the KLK7 complex due to the presence of Ala190 in place of serine, as well as due to insertion of a proline residue at position 218A that changes the conformation of the loop around it **(Figure 6A)**. The same insertion is also found in KLK4, 5, 2 and 3, with Arg218A present in KLK14. The conformational changes induced by this insertion are very similar in all compared kallikreins, leading to the loss of a hydrogen bond between Arg64/NH2 and Gly218/O, while gaining hydrophobic interactions between the side chains of the inserted proline and Ser82 **(Figure 6A)**.

Substitution of Ser190 by alanine combined with a change of Asp189 to Asn reorients the side chain of Arg64, leading to dramatic changes in the network of interactions of the guanidinium group of Arg64 with KLK7 **(Figure 6A)**. An extensive network of interactions involving the NH1,

NH2, and NE atoms of Arg64 with different counterparts in the enzyme might contribute to the high inhibitory potency of BbKI against KLK7.

The fact that the changes of the orientation of Arg64 in the KLK7 complex are directly connected to the substitutions D189N and S190A, and not to the Pro218A insertion, is supported by a comparison with KLK4. While the same insertion of a proline residue is present in KLK4 as in KLK7, Asp189 and Ser190 are not replaced**. Figure 6A** shows very clearly the same orientation and very similar interaction pattern of Arg64 in KLK4 as in trypsin. The inhibition constants for these two enzymes are also very similar.

The interactions between the loop in BbKI containing Lys69 and Tyr72 does not change much in the complex with KLK7, whereas the longer side chain of Gln97 enables much better interaction pattern with Arg129 in KLK7 than in trypsin. Numerous contacts that include hydrophobic interactions, as well as hydrogen bonds, may compensate for the loss of contacts with Val3 and Leu175 in KLK7 **(Figure 6B)**.

Three loops in BbKI, containing Ser81-Ser82, Asn66, and Ala16, respectively, make more extensive interactions with KLK7 than with trypsin **(Figure 6C)**. Although a strong hydrogen bond between Asn66 and Tyr151 in the trypsin complex is not present in the KLK7 complex due to substitution by Phe151, the hydrophobic contacts between these residues still remain. An extra hydrogen bond is formed between the carbonyl oxygen of Pro147 and the hydroxyl of Ser81, and several hydrophobic contacts are made between Pro147 and Ser81-Ser82 **(Figure 6C)**. These interactions appear to be present in all kallikreins that simultaneously contain a proline insertion (Pro218A) and Pro147 in the corresponding loop. The loops that contain Pro147 and Pro218A are located just above the entrance to the specific P1 pocket of the enzyme. The interactions of the Ser81-Ser82 loop of the inhibitor with these two loops in KLK7 may reduce their mobility, and thus facilitate docking of the side chain of Arg64 in the primary recognition pocket **(Figure 7)**.

3.8. Structural basis for the lack of inhibition of several other kallikreins

As seen in **Table 3**, BbKI shows only very low (if any) inhibitory potency against KLK5 and KLK14, and does not inhibit KLK2 and KLK3 at all. To justify the lack of inhibition in structural terms, the available structure of human KLK5 (PDB ID 2psx) was superimposed on trypsin in the

complex with BbKI in order to examine the differences between the enzyme-inhibitor interactions compared to the trypsin complex.

In the complex with trypsin Arg64, the specificity-determining P1 basic residue on the reactive loop of BbKI, contacts Asp189 and Ser190, and is hydrogen bonded to the carbonyl oxygen of Gly218. The latter interaction is most likely not present in KLK5 and KLK14 due to the insertion of Pro218A in the former, and Arg218A in the latter, that changes the conformation of the main chain of the enzyme around this residue **(Figure 8A)**. However, an equivalent insertion is also present in KLK4 and KLK7. While the loss of the hydrogen bond between P1 Arg64 with O of Gly218 might explain the lower inhibitory activity of BbKI against KLK5 and KLK14, this effect is compensated by other interactions in KLK4 and KLK7. Two other interactions might be lost in KLK5 due to the substitution of Asn97 (trypsin) by Pro97 (a hydrogen bond between the carbonyl oxygen of Asn97 and NE of Arg129 in BbKI, as well as a hydrophobic contact between CG atom of Asn97 and the Cα atom of Gly130) **(Figure 8B)**. In addition, the bridging interactions between the side chain of Arg129 in BbKI and two loops, 174-175 and 96-97 in trypsin, most likely will be lost in the complex with KLK5 **(Figure 8B)**.

Close contacts were found between the loop containing Ser80-Ser81 in BbKI and the two loops in KLK5 containing Pro147 (less than 2 Å) and Tyr218 (~1.5 Å) **(Figure 9A)**, that may contribute to the lack of inhibition. In addition, strong hydrogen bond between Asn66 in BbKI and Tyr151 in trypsin (2.5 Å) cannot be present in the complex with KLK5 due to the Y151F substitution.

We attribute the complete inability of BbKI to inhibit KLK2 and KLK3 predominantly to the large insertion in the loop 94-100 in both enzymes (**Figure 9B)**, which leads to severe collisions and overlaps with several loops in BbKI (**Figure 9B)**. In particular, two points of collision are found in the vicinity of residues 62 and 71 in BbKI, part of the loop that includes the P1 Arg64.

The side chain of Glu218 in the loop that includes the Pro218A insertion in KLK2 and KLK3 collides with the loop containing Ser61 in BbKI in a manner similar to the Tyr218 in KLK5 described above **(Figure 9A)**. There appears to be fairly good correlation between the size of the side chain of residue 218 and the potency of the inhibitor **Table 3**).

4.0. Conclusions

BbKI is a plant derived inhibitor of various proteolytic enzymes. In particular, it is a potent inhibitor of plasma kallikrein and several tissue kallikreins, towards which it exhibits various levels of inhibitory activity. Whereas BbKI is indeed the most potent inhibitor of some enzymes from this family among all proteinaceous inhibitors studied to date, it is not particularly specific and its activity towards some tissue kallikreins is fairly low or totally absent. Although BbKI is called a “kallikrein inhibitor”, it is also active against both trypsin and chymotrypsin.

In this study we determined a high-resolution structure of the complex of BbKI and trypsin. The structure allowed to map three major interfaces between the enzyme and the inhibitor. Based on this structure we constructed models of the complexes of several kallikreins with BbKI, evaluating the structural basis of inhibition of these enzymes, as well as the lack of inhibition of some other members of the family. The models revealed some areas of the enzymes that are important determinants of the potency of inhibition. Substitutions in the enzymes in the specificity-determining pocket that accommodates the P1 Arg64 modify the network of interactions in this area. Another important interface maintains the interactions with the loop in BbKI carrying Ser81-Ser82. As an example, interactions in these two areas work in a concerted way in KLK7, enhancing inhibition potency of BbKI for this chymotrypsin-like enzyme.

Kallikreins have been found to be important in the progression and metastasis of various cancer types, and their activity is also associated with other diseases, making them important targets of investigations. Crystal structure of the BbKI complex with trypsin and modeling of its interactions with selected kallikreins for which experimental structures are available allowed us to analyze the structural basis for its inhibitory properties for this important family of serine proteases. Differences in the inhibitor-enzyme interactions in the corresponding sites can provide insights towards the specificity of individual kallikreins. Comparative analysis of such structural data combined with their inhibitory measurements might guide us in the design of a more potent and specific inhibitor targeting a particular kallikrein.

Acknowledgements

We would like to thank Dr. Andriy Anishkin (University of Maryland) for the calculations of the energy minimized structures of selected kallikrein complexes. These models are available from the corresponding author of this manuscript. We acknowledge the use of beamline 22-ID of the

Southeast Regional Collaborative Access Team (SER-CAT), located at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. This work was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research and with Federal funds from the National Cancer Institute, NIH, under Contract No. HHSN261200800001E (to ML). JS and MM were supported by project ChemBioDrug CZ.02.1.01/0.0/0.0/16_019/0000729 from European Regional Development Fund (OP RDE) and the institutional project RVO 61388963. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U. S. Government.

Figures

Fig. 1. A cartoon representation of one of the six complexes of BbKI-trypsin (consisting of molecules A (trypsin - green) and G (L55R-BbKI - yellow)) present in the monoclinic crystals. The side chain of the specificity-determining Arg64 in BbKI is shown in sticks.

Fig. 2. Packing of 44 independent BbKI-trypsin complexes in the monoclinic crystal. Each of the four helical columns of 11 complexes arranged along the four crystallographic 2_1 axes is presented in a different color, with the trypsin molecules in darker and the inhibitor molecules in paler shades. The unit cell is outlined in thin black lines. A) The view along the length of the helical columns, i.e. along the monoclinic 2_1 axis. B) The view perpendicular to the direction of helices.

Fig. 3. Modeling of the interactions between BbKI and pKLK. a) Interactions in the vicinity of the specificity-providing P1 Arg64 in trypsin and several kallikreins. b) Interactions of three loops in BbKI comprising residues Val3, Lys69 and Tyr72, and Arg129 and Gly130, respectively, with the corresponding fragments in trypsin and kallikreins. c) Interactions of three loops in BbKI which include Ser81-Ser82, Asn66, and Ala16, respectively, with trypsin and pKLK.

Fig. 4. Structure-based alignment of the amino acids sequences of trypsin, chymotrypsin, and selected kallikreins.

Fig. 5. Modeling of the interactions between BbKI and KLK4. a) Interactions in the vicinity of Arg64. Inserted Pro218A in KLK4 maintains additional hydrophobic contacts with the loop in BbKI carrying Ser82. b) Interactions of three loops in BbKI comprising residues: Val3, Lys69 and Tyr72, and Arg129 and Gly130, respectively, with trypsin and KLK4. Interactions of BbKI with residues 96-98 in KLK4 (shown as red ribbon) cannot be modeled. c) Open conformation of the loop 96-98 in KLK4 complexed with the Sunflower Trypsin Inhibitor (PDB ID 4k8y). d) Interactions of BbKI loop containing residues Ser81-Ser82, as well as two loops that include Asn66 and Ala16, respectively with KLK4 and trypsin.

Fig. 6. Modeling of the interactions between BbKI and KLK7. a) Differences in the interactions of Arg64 of BbKI in the primary pocket of KLK7 and their counterparts in trypsin and KLK4. b) Interactions of three loops in BbKI comprising residues: Val3, Lys69 and Tyr72, and Arg129 and Gly130, respectively, with trypsin and KLK7. c) Extensive interactions between three loops in BbKI which include Ser81-Ser82, Asn66, and Ala16, respectively, with KLK7 compared to trypsin.

Fig. 7. Hydrophobic interactions of Ser81-Ser82 of the inhibitor with two loops in kallikreins simultaneously containing Pro147 and Pro218A.

Fig. 8. Modeling of the interactions between BbKI and KLK5. a) Insertion of Pro218A in KLK5 leads to the loss of a hydrogen bond between Arg64 in BbKI and Gly218. b) Modified interactions between BbKI and KLK5 due to the amino acid substitution in a loop containing residue 97.

Fig. 9. Structural basis for the lack of inhibition of KLK2 and KLK3. a) Extensive clashes of Glu218 in KLK2 and KLK3 with the loop containing Ser61 in BbKI may contribute to the lack of inhibition, in a manner similar to the conflict between Tyr218 in KLK5 and the Ser81-Ser82 loop in BbKI. b) The large insertion in the loop 94-100 in both enzymes leads to collision with several loops in BbKI, leading to the inability of BbKI to inhibit KLK2 and KLK3.

Table 1 Enzyme activity assay conditions used for inhibition kinetics.

Enzyme	Enzyme concentration (nM)	Substrate	Substrate concentration (μM)	K_m (μM)
plasma hKLK	10	PFR-Amc	25	28.7
hKLK2	40	PFR-Amc	50	3.1†
hKLK3	40	Suc-AAPF-Amc	50	> 250
hKLK4	1	Boc-VPR-Amc	50	5.9
hKLK5	0.2	Boc-VPR-Amc	100	101.1‡
mKLK5	0.2	Boc-VPR-Amc	100	48.3 ‡
hKLK7	0.2	Ac-KHLY-Acc	25	99.1‡
mKLK7	2	Ac-KHLY-Acc	25	101.5 ‡
hKLK14	0.2	Boc-VPR-Amc	30	29.0 ‡
trypsin	0.6	Cbz-FR-Amc	25	67.8
chymotrypsin	0.024	Suc-AAPF-Amc	25	41.5

† (Fogaca *et al.*, 2004), ‡ (Horn *et al.*, 2018)

Table 2. Data collection and structure refinement

	BbKI/Trypsin complex	BbKI/Trypsin complex	L55R/Trypsin complex
Data Collection			
Space Group	$P2_1$	$P6_4$	$P6_4$
Cell Dimensions (Å)	137.2 483.9 137.0	207.81 207.81 107.17	207.30 207.30 107.17
angles (°)	90 116.8 90	90 90 120	90 90 120
Resolution (Å)	50.0-3.96 (4.03-3.96)*	50.0-2.0 (2.03-2.00)	50.0-1.94 (1.97-1.94)
<I>/<σI>	8.16 (3.05)	14.4 (0.66)	26.5 (0.68)
Completeness (%)	88.7 (37.3)	96.5 (71.6)	98.2 (74.1)
Redundancy	4.0 (2.1)	5.6 (1.8)	10.1 (2.3)
R_{merge} (%)	13.5 (27.9)	9.9 (110.5)	7.7 (115.1)
Refinement			
Resolution (Å)	49.23-3.95	49.91-2.0	49.79-1.94
No. Reflections	118575	140125	168232
R_{work}/R_{free} (%)	22.60/28.15	18.99/22.3	19.6/22.3
R.m.s. deviations			
Bond lengths (Å)	0.004	0.008	0.009
Bond angles (°)	0.77	1.32	1.44
No. Atoms			
Protein	127116	17407	17407
Ligand/ion	-	8	0
Water	-	810	1642
PDB ID	6dwu	6dwh	6dwf

*The highest resolution shell is shown in parentheses.

†$R_{merge} = \sum_h\sum_i|I_i-\langle I\rangle|/\sum_h\sum_i I_i$, where I_i is the observed intensity of the i-th measurement of reflection h, and ⟨I⟩ is the average intensity of that reflection obtained from multiple observations.

‡$R = \sum||F_o|-|F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. R_{free} was defined in Brünger (1992).

Table 3 Inhibitory potency of wild type BbKI and its mutant (BbKI-L55R) against various kallikreins and archetypal proteases from the S1 family.
n.s.i. indicates that no significant inhibition was observed with 10 μM inhibitor.

Enzyme	K_i (nM)	
	BbKI	**BbKI-L55R**
human kallikreins		
plasma hKLK	12.9 ± 3.2	14.4 ± 1.1
hKLK2	n. s. i.	n. s. i.
hKLK3	n. s. i.	n. s. i.
hKLK4	0.045 ± 0.002	0.036 ± 0.006
hKLK5	> 500	> 500
hKLK7	0.18 ± 0.03	2.2 ± 0.2
hKLK14	> 5000	> 5000
mouse kallikreins		
mKLK5	> 500	> 500
mKLK7	9.8 ± 1.6	15.9 ± 3.4
archetypal serine proteases		
trypsin	0.20 ± 0.01	1.3 ± 0.1
chymotrypsin	33.4 ± 2.3	50.6 ± 6.4

References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Crystallogr.* **D66**, 213-221.

Araujo, A. P., Hansen, D., Vieira, D. F., Oliveira, C., Santana, L. A., Beltramini, L. M., Sampaio, C. A., Sampaio, M. U. & Oliva, M. L. (2005). *Biol. Chem.* **386**, 561-568.

Batista, I. F., Oliva, M. L., Araujo, M. S., Sampaio, M. U., Richardson, M., Fritz, H. & Sampaio, C. A. (1996). *Phytochemistry* **41**, 1017-1022.

Botos, I. & Wlodawer, A. (2007). *Curr. Opin. Struct. Biol* **17**, 1-8.

Brito, M. V., C., D. O., Salu, B. R., Andrade, S. A., Malloy, P. M., Sato, A. C., Vicente, C. P., Sampaio, M. U., Maffei, F. H. & Oliva, M. L. (2014). *Thromb. Res* **133**, 945-951.

CCP4 (1994). *Acta Crystallogr.* **D50**, 760-763.

De Vonis Bidlingmeyer, U., Leary, T. R. & Laskowski, M., Jr. (1972). *Biochemistry* **11**, 3303-3310.

Debela, M., Hess, P., Magdolen, V., Schechter, N. M., Steiner, T., Huber, R., Bode, W. & Goettig, P. (2007). *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16086-16091.

Emsley, P. & Cowtan, K. (2004). *Acta Crystallogr.* **D60**, 2126-2132.

Fogaca, S. E., Melo, R. L., Pimenta, D. C., Hosoi, K., Juliano, L. & Juliano, M. A. (2004). *Biochem. J.* **380**, 775-781.

Garcia-Nafria, J., Timm, J., Harrison, C., Turkenburg, J. P. & Wilson, K. S. (2013). *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1367-1380.

Helland, R., Otlewski, J., Sundheim, O., Dadlez, M. & Smalas, A. O. (1999). *J. Mol. Biol.* **287**, 923-942.

Hinkes, S., Wuttke, A., Saupe, S. M., Ivanova, T., Wagner, S., Knorlein, A., Heine, A., Klebe, G. & Steinmetzer, T. (2016). *J. Med. Chem.* **59**, 6370-6386.

Horn, M., Zbodakova, O., Kasparek, P., Srp, J., Haneckova, R., Hradilek, M., Mares, M. & Sedlacek, R. (2018). *Biol. Chem.*

Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph* **14**, 33-38.

Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A. D., Jr. & Pastor, R. W. (2010). *J. Phys. Chem. B* **114**, 7830-7843.

Kumar, P., Singh, M. & Karthikeyan, S. (2011). *Acta Crystallogr. D Biol. Crystallogr.* **67**, 131-139.

Kunitz, M. (1947). *J. Gen. Phys* **30**, 311-320.

Li, Z., Partridge, J., Silva-Garcia, A., Rademacher, P., Betz, A., Xu, Q., Sham, H., Hu, Y., Shan, Y., Liu, B., Zhang, Y., Shi, H., Xu, Q., Ma, X. & Zhang, L. (2017). *ACS Med. Chem. Lett.* **8**, 185-190.

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. & Karplus, M. (1998). *J. Phys. Chem. B* **102**, 3586-3616.

Majumder, S., Khamrui, S., Dasgupta, J., Dattagupta, J. K. & Sen, U. (2012). *Biochim. Biophys. Acta* **1824**, 882-890.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst* **40**, 658-674.

Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Crystallogr.* **D67**, 355-367.

Nakahata, A. M., Mayer, B., Neth, P., Hansen, D., Sampaio, M. U. & Oliva, M. L. (2013). *Planta Med.* **79**, 227-235.

Odei-Addo, F., Frost, C., Smith, N., Ogawa, T., Muramoto, K., Oliva, M. L., Graf, L. & Naude, R. (2014). *J. Enzyme Inhib. Med. Chem* **29**, 633-638.

Oliva, M. L., Mendes, C. R., Juliano, M. A., Chagas, J. R., Rosa, J. C., Greene, L. J., Sampaio, M. U. & Sampaio, C. A. (1999). *Immunopharmacology* **45**, 163-169.

Oliva, M. L., Mendes, C. R., Santomauro-Vaz, E. M., Juliano, M. A., Mentele, R., Auerswald, E. A., Sampaio, M. U. & Sampaio, C. A. (2001). *Curr. Med. Chem* **8**, 977-984.

Oliva, M. L. & Sampaio, U. M. (2008). *Biol. Chem* **389**, 1007-1013.

Oliva, M. L., Silva, M. C., Sallai, R. C., Brito, M. V. & Sampaio, M. U. (2010). *Biochimie* **92**, 1667-1673.

Pampalakis, G. & Sotiropoulou, G. (2007). *Biochim. Biophys. Acta* **1776**, 22-31.

Patil, D. N., Chaudhary, A., Sharma, A. K., Tomar, S. & Kumar, P. (2012). *FEBS J* **279**, 4547-4564.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L. & Schulten, K. (2005). *J. Comput. Chem.* **26**, 1781-1802.

Renko, M., Sabotic, J. & D., T. (2012). *Biol. Chem* **393**, 1043-1054.

Riley, B. T., Ilyichova, O., Costa, M. G., Porebski, B. T., de Veer, S. J., Swedberg, J. E., Kass, I., Harris, J. M., Hoke, D. E. & Buckle, A. M. (2016). *Sci. Rep.* **6**, 35385.

Scheidig, A. J., Hynes, T. R., Pelletier, L. A., Wells, J. A. & Kossiakoff, A. A. (1997). *Protein Sci.* **6**, 1806-1824.

Song, H. K. & Suh, S. W. (1998). *J. Mol. Biol* **275**, 347-363.

Souza-Pinto, J. C., Oliva, M. L., Sampaio, C. A., Damas, J., Auerswald, E. A., Limaos, E., Fritz, H. & Sampaio, M. U. (1996). *Immunopharmacology* **33**, 330-332.

Swedberg, J. E., de Veer, S. J., Sit, K. C., Reboul, C. F., Buckle, A. M. & Harris, J. M. (2011). *PLoS One* **6**, e19302.

Tanaka, S., Sawaya, M. R. & Yeates, T. O. (2010). *Science* **327**, 81-84.

Turk, B. (2006). *Nat. Rev. Drug Discov* **5**, 785-799.

Vadivel, K., Ponnuraj, S. M., Kumar, Y., Zaiss, A. K., Bunce, M. W., Camire, R. M., Wu, L., Evseenko, D., Herschman, H. R., Bajaj, M. S. & Bajaj, S. P. (2014). *J. Biol. Chem* **289**, 31647-31661.

Zhou, D., Hansen, D., Shabalin, I. G., Gustchina, A., Vieira, D. F., de Brito, M. V., Araujo, A. P., Oliva, M. L. & Wlodawer, A. (2015). *Acta Crystallogr. F. Struct. Biol. Commun* **71**, 1055-1062.

Zhou, D., Lobo, Y. A., Batista, I. F., Marques-Porto, R., Gustchina, A., Oliva, M. L. & Wlodawer, A. (2013). *PLoS. One* **8**, e62252.

Figure 1

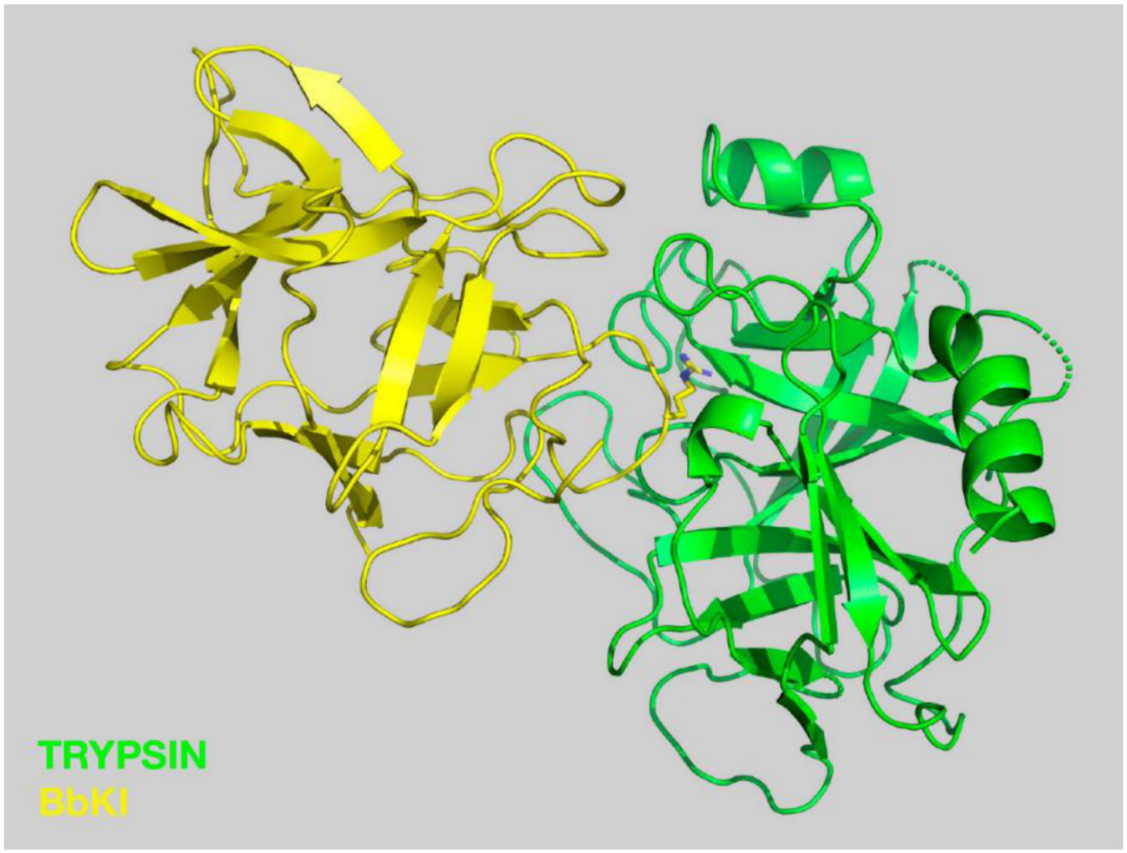


Figure 2

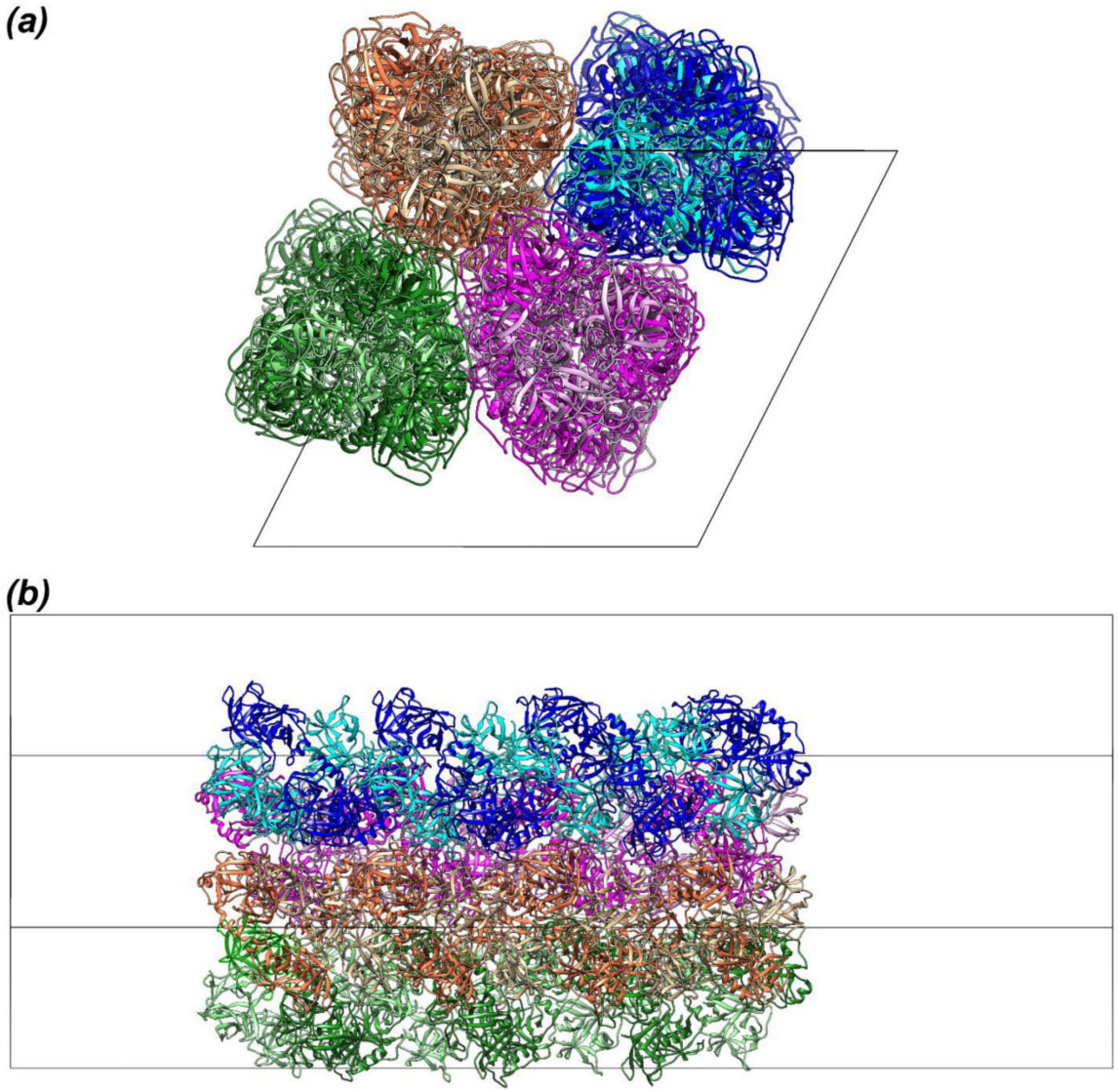


Figure 3

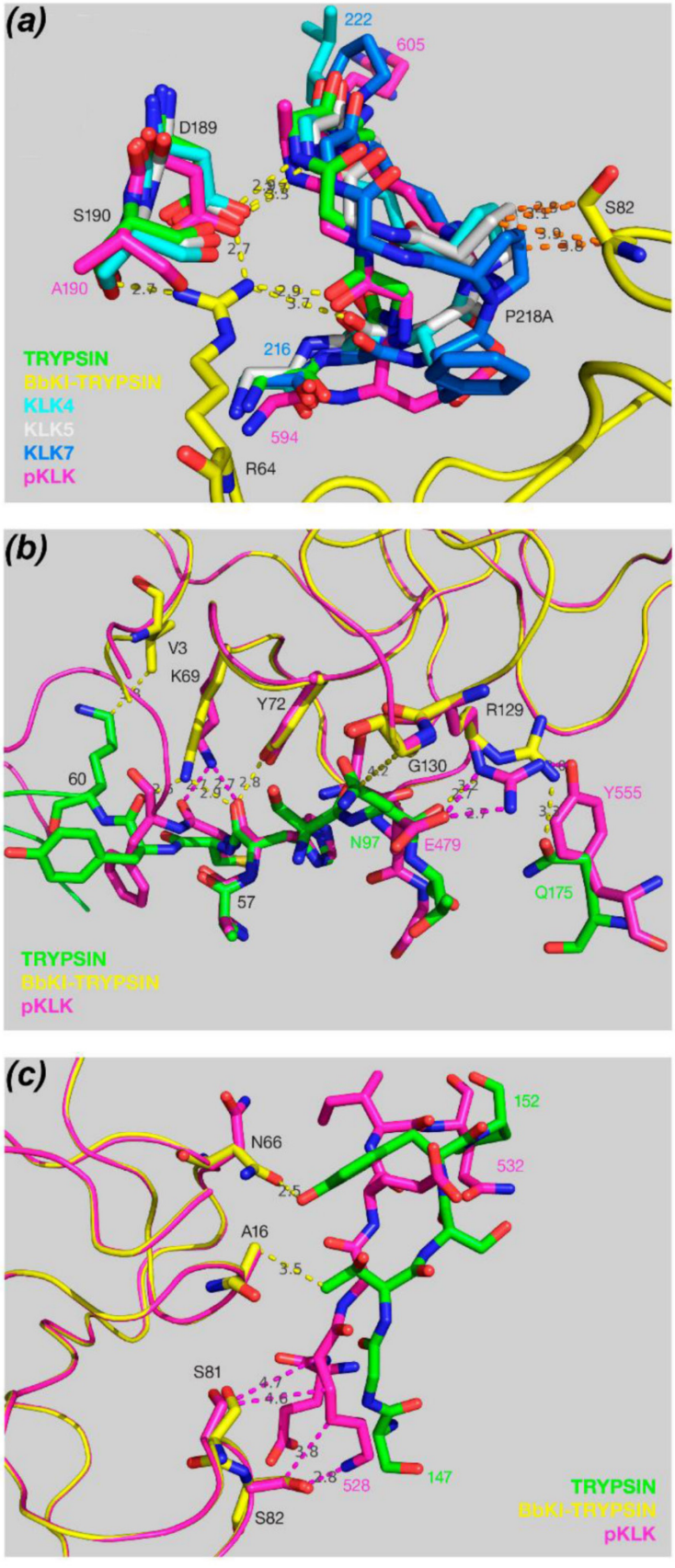

(a)
222
605
D189
S190
A190
S82
P218A
216
594
R64
TRYPSIN
BbKI-TRYPSIN
KLK4
KLK5
KLK7
pKLK
(b)
V3
K69
Y72
R129
G130
60
Y555
N97
E479
Q175
57
TRYPSIN
BbKI-TRYPSIN
pKLK
(c)
152
N66
532
A16
S81
147
528
S82
TRYPSIN
BbKI-TRYPSIN
pKLK

Figure 4

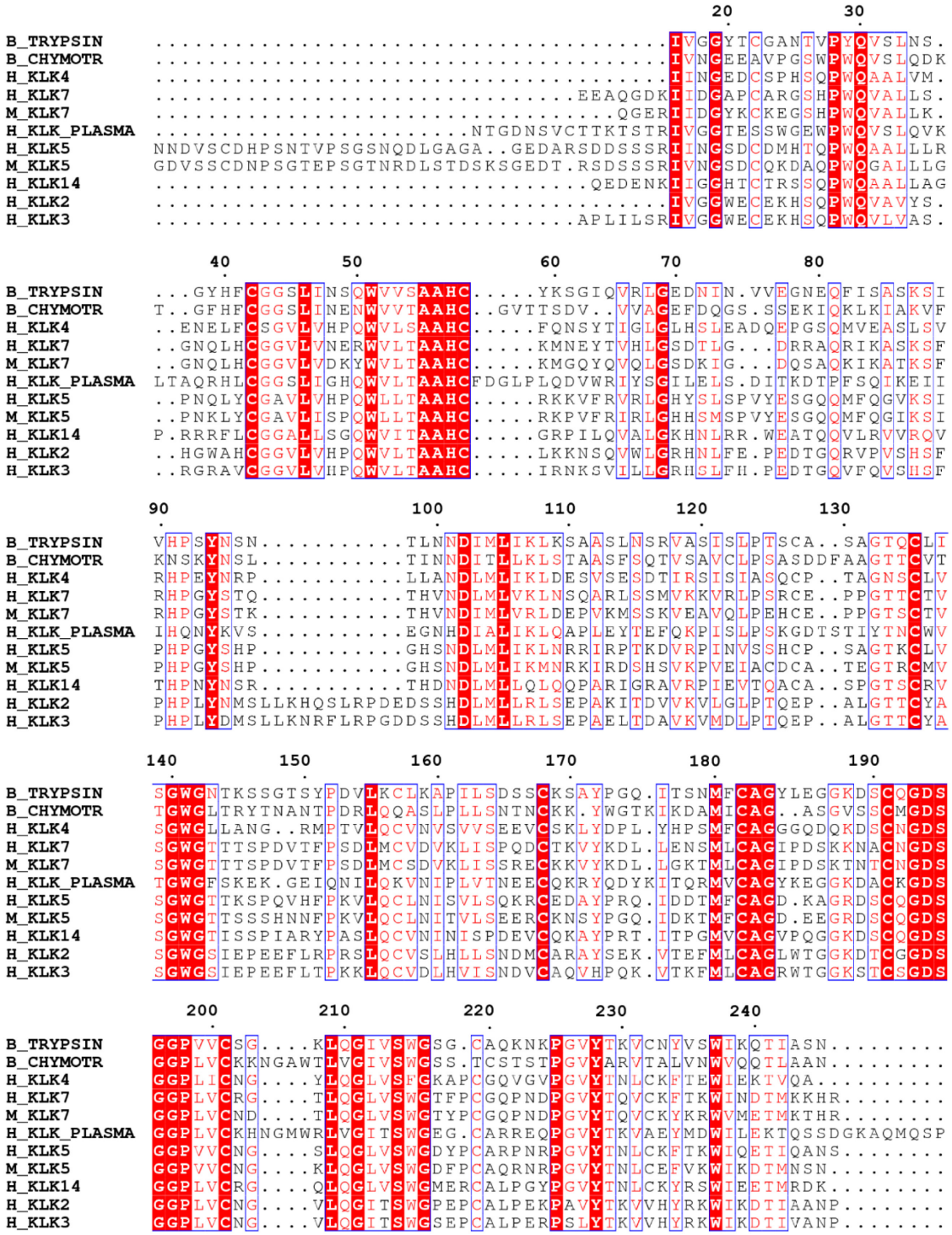

B_TRYPSIN
B_CHYMOTR
H_KLK4
H_KLK7
M_KLK7
H_KLK_PLASMA
H_KLK5
M_KLK5
H_KLK14
H_KLK2
H_KLK3

Figure 5

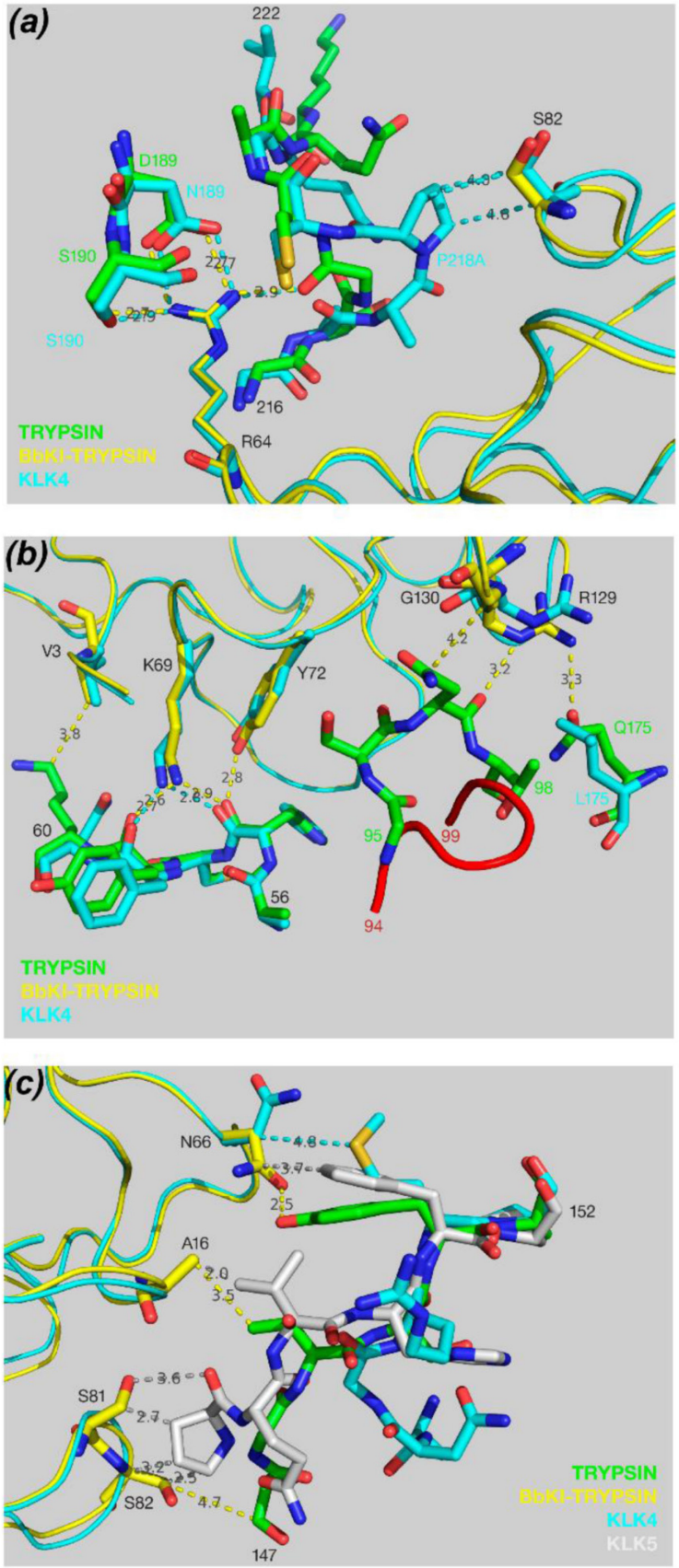

(a)
222
S82
D189
N189
S190
P218A
S190
216
R64
TRYPSIN
BbKI-TRYPSIN
KLK4
(b)
G130
R129
V3
K69
Y72
Q175
L175
98
60
95
99
56
94
TRYPSIN
BbKI-TRYPSIN
KLK4
(c)
N66
152
A16
S81
S82
147
TRYPSIN
BbKI-TRYPSIN
KLK4
KLK5

Figure 6

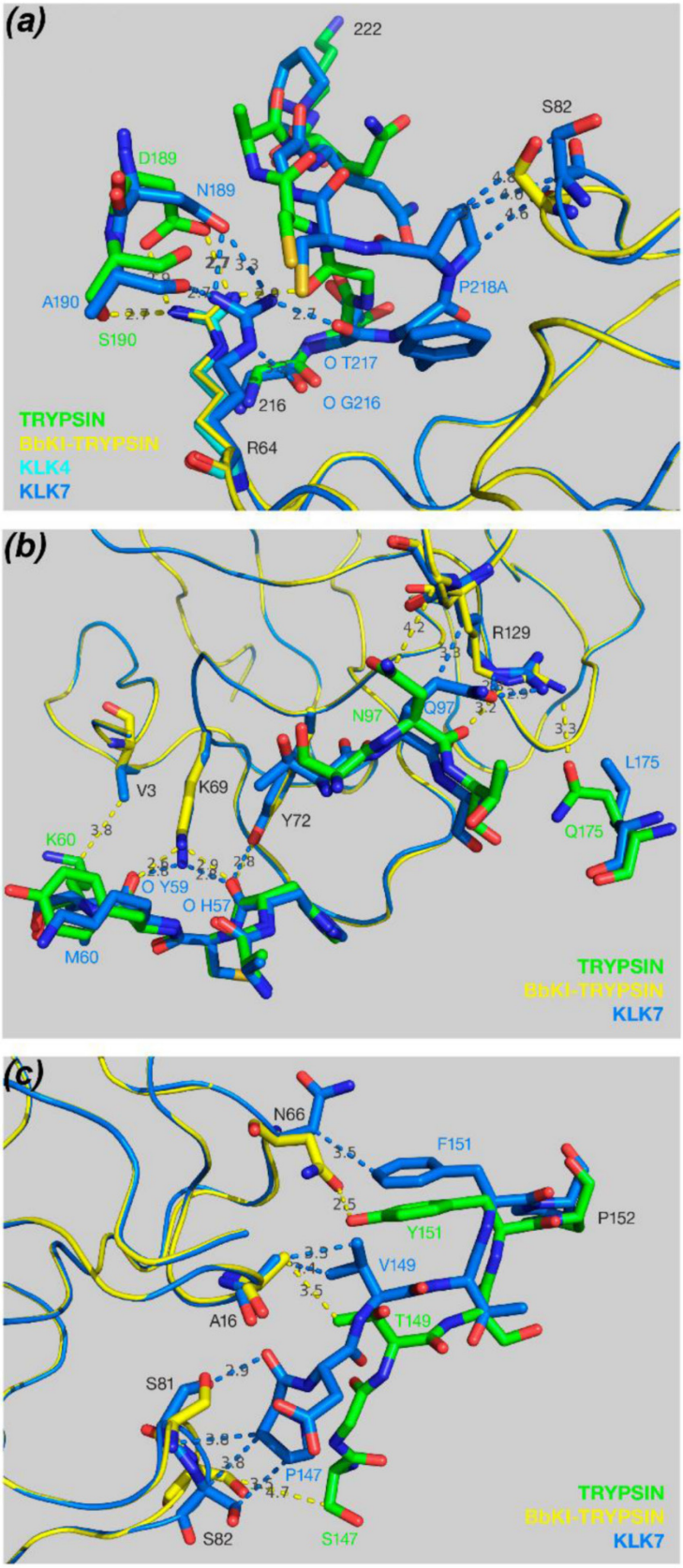

(a)
222
S82
D189
N189
A190
S190
P218A
O T217
216
O G216
R64
TRYPSIN
BbKI-TRYPSIN
KLK4
KLK7
(b)
R129
N97
Q97
L175
V3
K69
Y72
Q175
K60
O Y59
O H57
M60
TRYPSIN
BbKI-TRYPSIN
KLK7
(c)
N66
F151
P152
Y151
V149
A16
T149
S81
P147
S82
S147
TRYPSIN
BbKI-TRYPSIN
KLK7

Figure 7

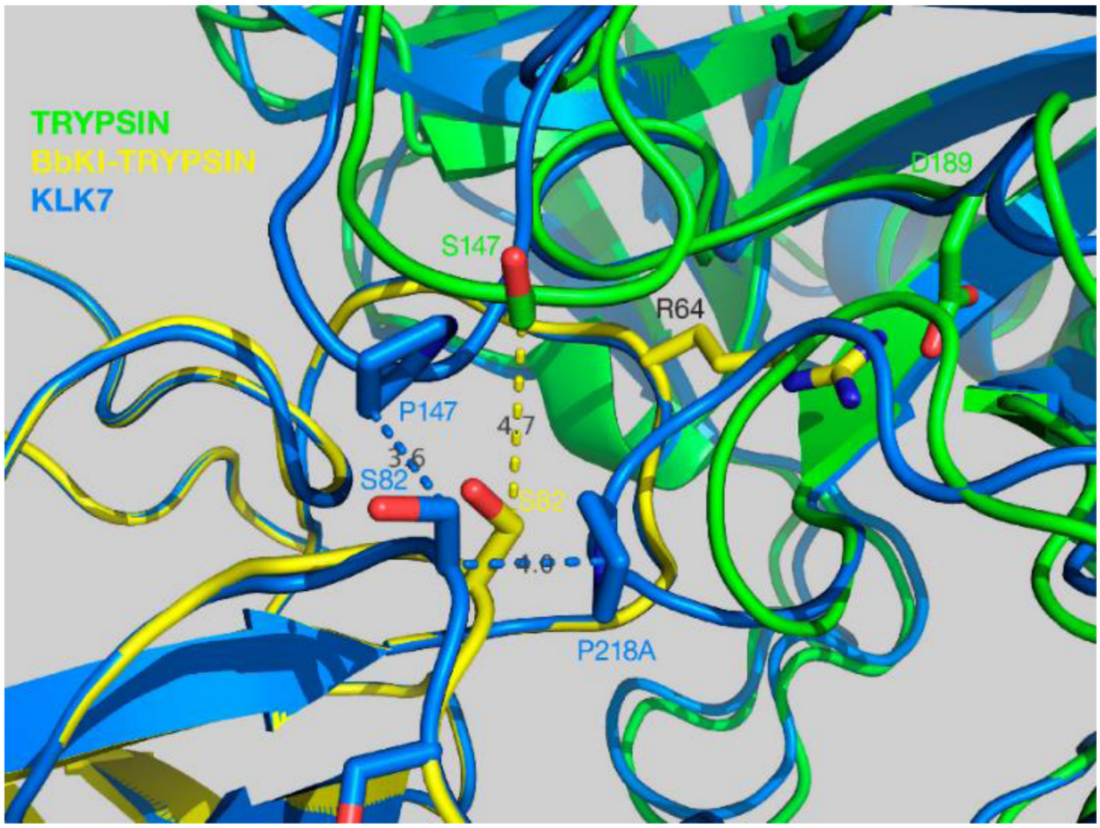

TRYPSIN
BbKI-TRYPSIN
KLK7
D189
S147
R64
P147
4.7
3.6
S82
S82
4.0
P218A

Figure 8

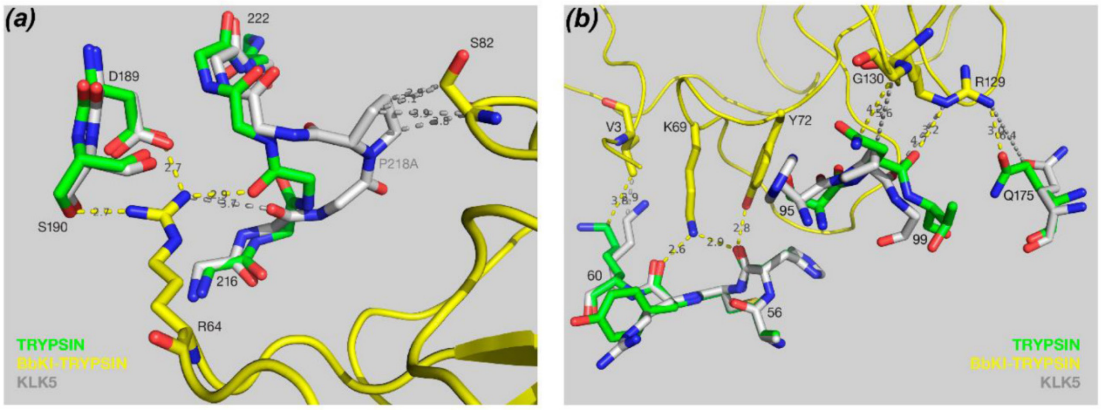
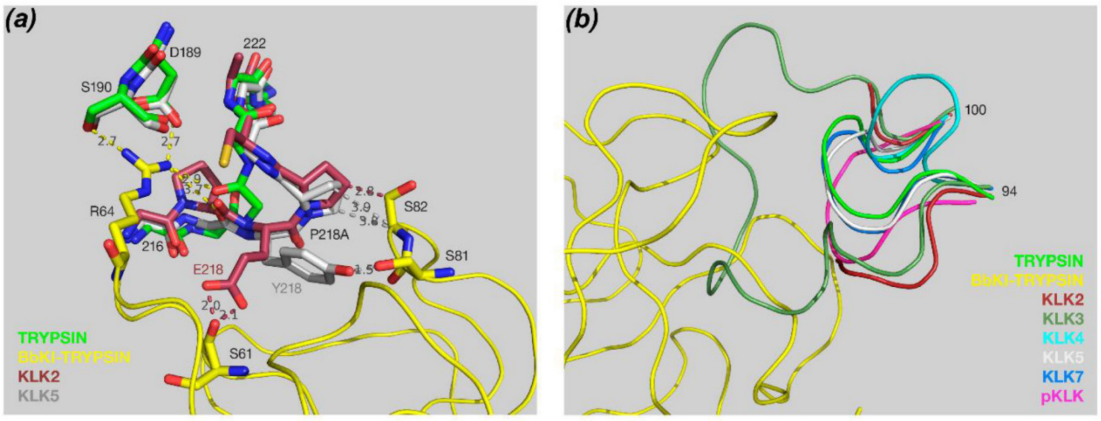

(a)
222
S82
D189
P218A
S190
216
R64
TRYPSIN
BgAI-TRYPSIN
KLK5
(b)
G130
R129
V3
K69
Y72
Q175
95
99
60
56
TRYPSIN
BgAI-TRYPSIN
KLK5

Figure 9


(a)
D189
222
S190
R64
216
E218
Y218
P218A
S82
S81
S61
TRYPSIN
BbKI-TRYPSIN
KLK2
KLK5
(b)
100
94
TRYPSIN
BbKI-TRYPSIN
KLK2
KLK3
KLK4
KLK5
KLK7
pKLK

4.7. Publikace č. 6: Kunitzovy proteasové inhibitory z rostlin: strukturní a funkční diverzita.

Souhrn

Cílem tohoto referátu bylo podat ucelený přehled o rostlinných proteasových inhibitorech z Kunitzovy rodiny, který v současné literatuře chybí. Zaměřuje se na jejich strukturní a funkční vlastnosti, zejména pak na diverzitu v inhibiční specifitě, která zahrnuje inhibici tří různých tříd proteas (aspartátových, cysteinových a serinových proteas) a je v kontextu rodin proteasových inhibitorů unikátní. V práci je popsána konzervovaná molekulární architektura Kunitzových inhibitorů označovaná jako β trojlístek (angl. β trefoil), která je tvořena dvanácti antiparalelními β řetězci propojených systémem variabilních smyček.

Pozornost je zaměřena na prostorové struktury komplexů Kunitzových inhibitorů se serinovými proteasami, které jsou v současné době jako jediné dostupné pro tyto rostlinné inhibitory v PDB databázi proteinových struktur. Diskutován je jednotný, tzv. Laskowského mechanismus interakce Kunitzových inhibitorů s modelovou serinovou proteasou trypsinem pomocí centrální reaktivní smyčky, která řídí klíčovou interakci v S_1 podmístě enzymu. Ta je zprostředkována převážně bazickými aminokyselinovými zbytky (Arg nebo Lys) na straně inhibitoru a zbytkem kyseliny asparagové v S_1 podmístě trypsinu. Ze superpozice prostorových struktur sady komplexů je patrné, že reaktivní centra proti serinovým proteasám se vyvinula nezávisle na několika různých smyčkách na povrchu molekuly inhibitoru. Tato rozdílná topologie tří doposud popsaných reaktivních center Laskowského typu je prezentována také pomocí přiřazení aminokyselinových sekvencí Kunitzových inhibitorů, které byly dosud strukturně charakterizovány ve formě komplexů. V práci je také nastíněno možné využití těchto rostlinných inhibitorů při konstrukci transgenních rostlin, které by byly odolné vůči herbivorním škůdcům a jiným patogenům.

Můj podíl na práci zahrnoval vyhledávání dostupné literatury a prostorových struktur v PDB databázi, analýzu prostorových struktur a sekvencí, přípravu obrázků pomocí molekulární grafiky a psaní manuskriptu.

Publikace č. 6

Srp J., Mareš M.

Kunitzovy proteasové inhibitory z rostlin: strukturní a funkční diverzita

Chemické listy, (110), 761-768 (2016)

KUNITZOVY PROTEASOVÉ INHIBITORY Z ROSTLIN: STRUKTURNÍ A FUNKČNÍ DIVERZITA

JAROSLAV SRP[a,b] a MICHAEL MAREŠ[a]

[a] *Ústav organické chemie a biochemie Akademie věd České republiky, Flemingovo nám. 2, 166 10 Praha,* [b] *Katedra biochemie, Přírodovědecká fakulta, Univerzita Karlova v Praze, Hlavova 2030/8, 128 40 Praha*
srp@uochb.cas.cz

Došlo 23.3.16, přijato 12.5.16.

Klíčová slova: inhibitory proteas, Kunitzovy inhibitory, proteolytické enzymy, prostorová struktura proteinů, obranné proteiny rostlin

Obsah

1. Úvod
2. Specifita a funkce proteinů z Kunitzovy rodiny
   2.1. Proteasové inhibitory z Kunitzovy rodiny
      2.1.1. Inhibitory serinových proteas
      2.1.2. Inhibitory aspartátových proteas
      2.1.3. Inhibitory cysteinových proteas
   2.2. Další funkce proteinů z Kunitzovy rodiny
3. Struktura inhibitorů Kunitzovy rodiny a jejich interakce s proteasami
   3.1. Architektura a strukturní variabilita inhibitorů
   3.2. Struktura komplexů inhibitorů s proteasami
4. Závěr

1. Úvod

Proteolýza patří mezi nejvýznamnější biochemické děje, které zahrnují na jedné straně nespecifickou degradaci proteinů při procesech jako je trávení a na druhé straně vysoce specifickou úpravu struktury proteinů a regulaci jejich funkce[1]. Vzhledem ke své důležitosti je proteolýza ve všech organismech pod přísnou kontrolou, která se odehrává na několika úrovních, a jednu z nich představují inhibitory proteas. Největší skupinou přírodních inhibitorů proteas jsou inhibitory proteinového charakteru, které se klasifikují podle struktury do 67 rodin uvedených v databázi proteas a jejich inhibitorů Merops (http://merops.sanger.ac.uk/inhibitors/).

Podle mechanismu interakce s proteasami lze inhibitory rozdělit na reverzibilní a ireverzibilní. Reverzibilní inhibitory tvoří nekovalentní protein-proteinové komplexy s proteasami, kde stericky blokují přístup substrátu do katalytického místa, zatímco ireverzibilní inhibitory jsou atakovány proteasami za vzniku kovalentních inhibičních komplexů[2,3]. U proteasových inhibitorů došlo během evoluce ke značné specializaci a jednotlivé rodiny inhibitorů se obvykle zaměřují na regulaci pouze jedné konkrétní třídy proteas. Několik rodin inhibitorů ale toto obecné pravidlo porušuje a je u nich zřejmý spíše opačný trend: jednotlivé inhibitory se strukturně a funkčně diverzifikovaly, takže rodina inhibitorů jako celek umožňuje regulaci dvou nebo dokonce tří proteasových tříd. V tomto kontextu jsou unikátní rostlinné inhibitory z Kunitzovy rodiny, jejíž zástupci jsou schopni inhibovat enzymy ze třídy serinových, cysteinových nebo aspartátových proteas. Dokonce se u některých inhibitorů této rodiny vyvinula multifunkčnost – více reaktivních center na jediné molekule cílených proti odlišným třídám proteas.

Tento přehledný článek je zaměřen na rostlinné proteasové inhibitory z Kunitzovy rodiny (rodina I3 dle Merops databáze, také nazývaná Kunitz-P), které jsou široce rozšířené u řady rostlinných taxonů[4]. Historie inhibitorů Kunitzovy rodiny je spojena se dvěma jmény – Kunitz a Laskowski. Prvně jmenovaný ve 40. letech 20. století izoloval a charakterizoval inhibitor serinové proteasy trypsinu ze semen sóje označovaný jako STI (soybean trypsin inhibitor)[5]. Laskowski v 80. letech popsal mechanismus interakce trypsinu s Kunitzovými inhibitory a zobecnil ho pro další rodiny reverzibilních inhibitorů interagující se serinovými proteasami[6]. Od 70. let jsou inhibitory Kunitzovy rodiny studovány na úrovni prostorové struktury a tyto informace spolu s detailní analýzou biochemických vlastností a molekulární fylogeneze z nich dělají jednu z nejvíce studovaných rodin proteasových inhibitorů. Mezi hlavní důvody současného zájmu o Kunitzovy inhibitory patří jejich funkce v rostlinách, kde hrají důležitou roli při obraně rostlin: tvoří významnou součást komplexního anti-proteolytického systému, kterým jsou blokovány trávicí proteasy herbivorního hmyzu a patogenů[7,8]. Z tohoto důvodu jsou geny Kunitzových inhibitorů atraktivní pro konstrukci transgenních plodin odolných proti škůdcům[9,10]. Další potenciální využití Kunitzových inhibitorů vyplývá z jejich unikátní strukturní architektury, která je základem enormní funkční variability. To umožňuje vyhledávat přírodní formy účinných inhibitorů specifických pro medicinálně zajímavé proteasy nebo takové inhibitory cíleně konstruovat mutacemi a racionální strukturní přestavbou molekuly[11,12]. Výzkum zaměřený tímto směrem vyžaduje nové informace zejména o struktuře reaktivních center Kunitzových inhibitorů a mechanismech jejich interakce se širokým spektrem relevantních typů proteas.

2. Specifita a funkce proteinů z Kunitzovy rodiny

Rostlinné inhibitory z Kunitzovy rodiny jsou proteiny nebo N-glykosylované glykoproteiny s molekulovou hmotností v rozmezí 18–22 kDa. Jde zpravidla o jednořetězcové molekuly, ale jsou známé i dvouřetězcové inhibitory, které vznikají proteolytickou fragmentací v oblasti povrchové smyčky[4]. Kunitzovy proteiny byly identifikovány a studovány v různých rostlinných pletivech, zejména v semenech, hlízách, listech a kůře. V listech hrají významnou roli jako obranné molekuly proti hmyzím herbivorům; jejich exprese je indukována při poškození listů požerem nebo spouštěna methyljasmonátem jako signální molekulou pro systémovou odpověď rostliny[13]. Vysoký obsah inhibitorů je v zásobních pletivech (semenech a hlízách), kde zřejmě vytvářejí permanentní obrannou bariéru proti proteasám patogenů, a předpokládá se, že by zde mohly hrát také roli jako zásobní proteiny a jako endogenní inhibitory, které kontrolují rostlinné proteasy např. během klíčení semen[4].

2.1. Proteasové inhibitory z Kunitzovy rodiny

2.1.1. Inhibitory serinových proteas

Výrazně největší skupinou inhibitorů z Kunitzovy rodiny jsou inhibitory serinových proteas (tab. I). Dokáží inhibovat především aktivitu proteas z rodiny S1 a menší počet inhibitorů interaguje s proteasami z rodiny S8. V prvním případě jde o inhibitory širokého spektra trávicích proteas trypsinového a chymotrypsinového typu, od enzymů hmyzích herbivorů po savčí enzymy. Při detailní analýze jejich inhibiční specifity byla také u řady inhibitorů nalezena účinná křížová inhibice medicinálně zajímavých proteas z rodiny S1 jako jsou proteasy krevní koagulace (faktory VIIa, XIa, Xa), fibrinolýzy (plasmin), kallikreiny nebo proteasy zánětu (elastasa a katepsin G)[4]. Přestože proteasy trypsinového a chymotrypsinového typu mají zcela odlišnou substrátovou specifitu, některé inhibitory Kunitzovy rodiny jako STI nebo PTI dokáží inhibovat oba typy těchto proteas (tab. I). Byly také konstruovány mutované formy původně chymotrypsinového inhibitoru WCI, které byly konvertovány na trypsinové inhibitory pomocí bodové mutace na reaktivní smyčce inhibitoru nebo transplantací reaktivní smyčky z jiného trypsinového inhibitoru[14,15]. Ačkoli interakce Kunitzových inhibitorů s trypsinovými proteasami byla intenzivně studována na strukturní úrovni, obdobné informace zatím chybí pro chymotrypsinové proteasy. Skupina Kunitzových inhibitorů ze semen obilnin má inhibiční specifitu k subtilisinům z rodiny S8 serinových proteas. Jde o bifunkční inhibitory, které kromě subtilisinu inhibují odlišným reaktivním centrem také endogenní α-amylasu. Nejlépe prostudovaným je inhibitor BASI z ječmene, který bude dále komentován v kap. 2.2.

2.1.2. Inhibitory aspartátových proteas

Inhibitory aspartátových proteas proteinového charakteru vznikly během evoluce pouze několikrát a představují velmi malou skupinu proteasových inhibitorů. Patří mezi ně několik inhibitorů z Kunitzovy rodiny, které byly izolovány z brambor jako dva izoinhibitory PDI (cit.[16]) a NID (cit.[17]) a z rajčat jako TDI (cit.[18]). Tyto proteiny jsou silné bifunkční inhibitory jednak aspartátových proteas typu katepsinu D a jednak serinových proteas trypsinového a chymotrypsinového typu. Jejich schopnost inhibovat aspartátové proteasy představuje evoluční odpověď na přítomnost tohoto typu proteas v trávicím traktu přirozených hmyzích herbivorů jako je mandelinka bramborová, typický škůdce těchto plodin[19]. Jejich obrannou úlohu v listech podporuje i inducibilní exprese methyljasmonátem[20]. PDI je také schopen inhibovat proteinasu A z kvasinek, která je blízkým homologem hmyzích a savčích katepsinů D, což naznačuje další možnou obrannou funkci inhibitorů proti fungálním patogenům[21]. Prostorová struktura PDI byla nedávno určena, nicméně unikátní reaktivní místo pro interakci s aspartátovými proteasami je neznámé a předmětem spekulací založených na molekulárním modelování[22,23].

2.1.3. Inhibitory cysteinových proteas

Bylo popsáno jen několik zástupců Kunitzovy rodiny z bobovitých a lilkovitých rostlin, kteří jsou schopni inhibovat aktivitu cysteinových proteas. Jedná se obecně o bifunkční inhibitory interagující s cysteinovými proteasami rodiny C1 (rodina papainu) a také se serinovými proteasami rodiny S1. Nejlépe charakterizován, včetně prostorové struktury, byl inhibitor BbCI z bauhínie funkční proti modelovým cysteinovým proteasám jako je katepsin L a cruzipain (a serinovým proteasam typu elastas)[24]. Pro skupinu izoinhibitorů PCPI z bramboru s inducibilní expresí v listech zatím chybí detailní biochemická analýza[25]. Interakční mechanismus těchto inhibitorů s cysteinovými proteasami není známý, ale je pravděpodobné, že reaktivní centrum inhibitorů bude na rozdíl od Kunitzových inhibitorů serinových proteas obsahovat více než jednu strukturní smyčku. Tato hypotéza je založena na inhibičním mechanismu clitocypinu (rodina I48), který má analogickou prostorovou architekturu jako proteiny z Kunitzovy rodiny a blokuje aktivní místo cysteinové proteasy pomocí dvou smyček[26].

2.2. Další funkce proteinů z Kunitzovy rodiny

U několika proteinů Kunitzovy rodiny byly podrobně popsány interakce s jinými skupinami proteinů, než jsou proteasy. Bifunkční inhibitor BASI je inhibitorem endogenní α-amylasy (dále také inhibuje serinovou proteasu subtilisin) a jeho úlohou je regulace štěpení zásobního škrobu v semenech ječmene při klíčení[27]. Příbuzné inhibitory WASI a RASI byly identifikovány v dalších obilninách[28]. Protein mirakulin z Kunitzovy rodiny byl izolován z plodů keře *Synsepalum dulcificum*. Ovlivňuje činnost chuťových pohárků tím, že se váže v neutrálním pH na

Tabulka I
Přehled rostlinných proteinů Kunitzovy rodiny s vyřešenou prostorovou strukturou. Uvedeno je označení proteinů a zdrojové rostliny. Typické modelové proteasy (senzitivní enzymy), které interagují s inhibitory, jsou klasifikovány podle tříd jako serinové (SerPR), cysteinové (CysPR) a aspartátové (AspPR) proteasy. Kód v PDB databázi se týká prostorových struktur samotných proteinů a jejich komplexů s uvedenými enzymy.

Inhibitor	Původ	Inhibiční specifita (senzitivní enzymy)	3D struktura (PDB kód)	
			inhibitor	komplex s enzymem
STI	*Glycine max*	SerPR (trypsin, chymotrypsin)	1BA7; 1AVU	trypsin 1AVW, 1AVX
DrTI	*Delonix regia*	SerPR (trypsin, kallikren)	1R8N	
WCI	*Psophocarpus tetragonolobus*	SerPR (chymotrypsin)	1WBC, 2WBC, 4WBC, 1EYL; mutované inhibitory[a]: 1FMZ, 1FNO, 3QYD, 3I2A, 3I2X, 1XG6, 2ET2, 2BEA, 2BEB, 2ESU, 4H9W, 4TLP, 4HA2	trypsin 3VEQ, 3I29, 2QYI
ETI	*Erythrina caffra*	SerPR (trypsin)	1TIE	
WBA	*Psophocarpus tetragonolobus*	nebyla nalezena	1WBA	
CTI	*Copaifera langsdorffii*	SerPR (trypsin)	1R8O	
BASI	*Hordeum vulgare*	SerPR (subtilisin) α-amylasa		subtilisin 3BX1 α-amylasa 1AVA
BbCI	*Bauhinia bauhinoides*	CysPR (cruzipain, katepsin L) SerPR (elastasa)	2GZB	
BbKI	*Bauhinia bauhinoides*	SerPR (kallikrein)	4ZOT	
API-A	*Sagittaria sagittifolia*	SerPR (trypsin)		2x trypsin[b] 3E8L
Mirakulin	*Murraya koenigii*	nebyla nalezena[c]	3IIR, 3ZC8	
CrataBL	*Crataeva tapia*	SerPR (trypsin, faktor Xa)	4IHZ	
EcTI	*Enterolobium contortisiliquum*	SerPR (trypsin)	4J2K	trypsin 4J2Y
TKI	*Tamarindus indica*	SerPR (trypsin, faktor Xa)	4AN6	trypsin 4AN7
PPI	*Carica papaya*	SerPR (trypsin, chymotrypsin, elastasa)	3S8K	
PTI	*Solanum tuberosum*	SerPR (trypsin, chymotrypsin)	3TC2	
PCPI	*Solanum tuberosum*	CysPR (katepsiny L, B) SerPR (trypsin, chymotrypsin)	5DVD, 5DVH, 5DWH[d]	
PDI	*Solanum tuberosum*	AspPR (katepsin D, proteinasa A) SerPR (trypsin, chymotrypsin)	5DZU	

[a] Rekombinantní inhibitory s mutacemi zavedenými v aminokyselinové sekvenci, [b] ternární komplex inhibitoru se dvěma molekulami trypsinu, [c] interaguje s chuťovými receptory, [d] nepublikované struktury

receptory sladké chuti jako antagonista a v kyselém pH se mění na agonistu; kyselý pokrm se tak zdá konzumentovi sladký[29]. Příbuzným s mirakulinem je protein CrataBL izolovaný z kůry stromu *Crataeva tapia*, který vykazuje insekticidní aktivitu. Jde o lektin vážící glykoproteiny a glykosaminoglykany. Kromě toho je také inhibitorem serinových proteas trypsinového typu[11].

3. Struktura inhibitorů Kunitzovy rodiny a jejich interakce s proteasami

3.1. Architektura a strukturní variabilita inhibitorů

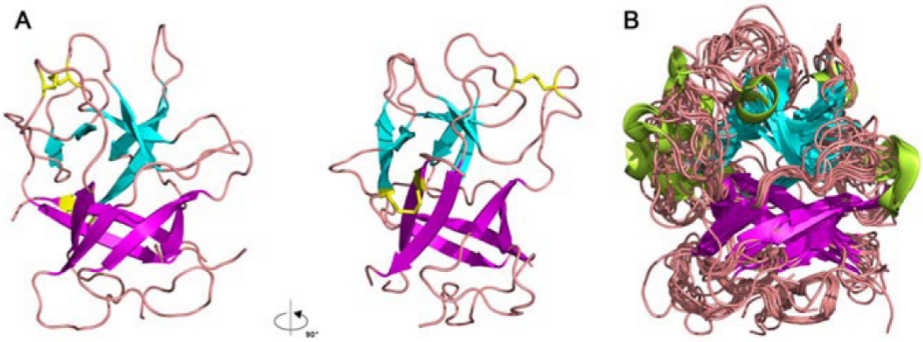
V současné době jsou v databázi PDB více než tři desítky prostorových struktur rostlinných proteinů z Kunitzovy rodiny a šest struktur jejich komplexů, kde je Kunitzův inhibitor navázán na hydrolytický enzym – proteasu nebo α-amylasu (tab. I). Tyto prostorové struktury mají jednotnou architekturu, která je charakteristická pro proteiny z Kunitzovy rodiny a je označována jako β-trojlístek (angl. β-trefoil). Obr. 1A ukazuje prostorovou strukturu molekuly trypsinového inhibitoru STI jako typického modelového zástupce Kunitzovy rodiny[30]. Architektura β-trojlístku je tvořena dvanácti segmenty se sekundární strukturou (β-strukturou), které jsou organizovány po dvojicích do vlásenek a jsou propojeny smyčkami. Tato uspořádání bývá popisováno jako strom, jehož kmen je tvořen šesti segmenty β-struktury vytvářejícími β-soudek a koruna se skládá ze tří párů β-struktur uspořádaných do trojúhelníku[31]. Sekvenční a strukturní analýza ukázala, že architekturu β-trojlístku využívá kromě proteinů Kunitzovy rodiny ještě nejméně pět dalších skupin proteinů, přičemž vzájemná sekvenční homologie těchto skupin je velmi malá[31,32].

Na obr. 1B je superpozice patnácti prostorových struktur Kunitzových proteinů dostupných v PDB databázi; analyzovány byly volné přirozené formy proteinů bez navázaného interakčního partnera. Z porovnání je patrné, že tyto Kunitzovy proteiny na jedné straně sdílejí společnou vnitřní architekturu β-trojlístku (topologii β-struktur) a na druhé straně jsou vysoce variabilní na povrchu v konformaci smyček exponovaných do roztoku. Některé smyčky mohou obsahovat segment se sekundární strukturou, konkrétně krátký motiv α-šroubovice, který se vyskytuje u řady z těchto molekul. Porovnání sekvencí pěti Kunitzových inhibitorů na obr. 2 názorně ukazuje variabilitu smyček v jejich délce i sekvenci.

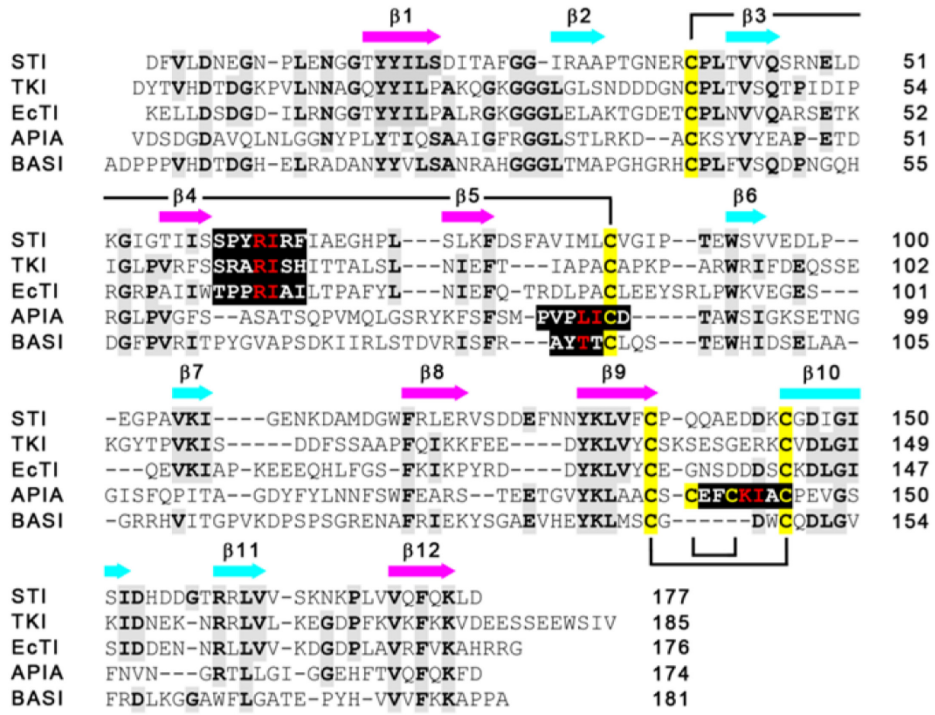
Architektura Kunitzových proteinů je stabilizována pomocí dvou konzervativních disulfidických můstků mezi cysteinovými zbytky Cys39-Cys86 a Cys136-Cys145 (číslování podle STI, obr. 1 a 2). Některé Kunitzovy proteiny obsahují další přídatný disulfid jako je tomu u inhibitoru serinových proteas API-A (cit.[33]) (obr. 2) a inhibitorů serinových/aspartátových proteas PDI (cit.[16]) a NDI (cit.[17]) nebo mají redukovaný počet disulfidů, případně disulfidy zcela chybí, jak bylo zjištěno u inhibitorů serinových a cysteinových proteas BbCI (cit.[34]) a BbKI (cit.[35]) (tab. I). Počet a rozmístění disulfidických vazeb na molekule Kunitzových proteinů byly také navrženy jako kritérium pro klasifikaci této proteinové rodiny do čtyř skupin[36].

3.2. Struktura komplexů inhibitorů s proteasami

Z šesti dostupných prostorových struktur komplexů Kunitzových inhibitorů s vazebnými partnery jsou čtyři pro komplex se serinovou proteasou trypsinem a další komplexy jsou se serinovou proteasou subtilisinem a s α-amylasou (tab. I).



Obr. 1. **Prostorová struktura molekul Kunitzových inhibitorů.** (A) Struktura reprezentativního inhibitoru STI (PDB kód: 1AVU). Molekula má architekturu β-trojlístku a je znázorněna pomocí stužkového modelu s barevně vyznačenými segmenty β-struktur. Šest fialových segmentů vytváří β-soudek označovaný jako "kmen" molekuly, šest modrých segmentů je umístěno v "koruně" molekuly. Smyčky spojující segmenty β-struktur mají lososovou barvu a dva disulfidové můstky mezi cysteinovými zbytky jsou znázorněny žlutě. (B) Superpozice prostorových struktur Kunitzových inhibitorů (15 struktur dostupných v PDB databázi). Barevné kódování je stejné jako v panelu (A); disulfidové můstky nejsou kvůli přehlednosti znázorněny. Zřetelné je konzervativní uspořádání segmentů β-struktur a konformační variabilita smyček. Zeleně jsou znázorněny krátké segmenty s konformací α-šroubovice umístěné v odlišných smyčkách některých inhibitorů. Molekulární grafika pro všechny obrázky byla připravena programem PyMol.



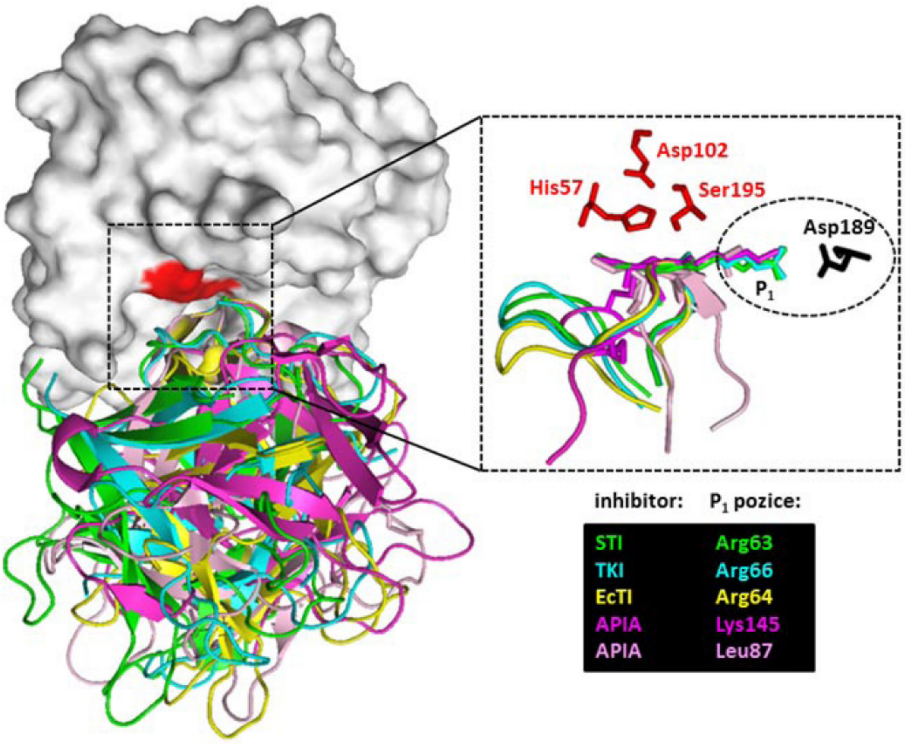
Obr. 2. **Přiřazení aminokyselinových sekvencí Kunitzových inhibitorů.** Analyzovány byly inhibitory, pro které je známá prostorová struktura komplexu s proteasou (tab. I), konkrétně s trypsinem (inhibitory STI, TKI a EcTI), dvěma molekulami trypsinu (API-A) a subtilisinem (BASI). Černým pozadím jsou zvýrazněny aminokyselinové zbytky, které tvoří reaktivní smyčku interagující s proteasou: červeně jsou uvedeny zbytky P_1'-P_1 , které se vážou v blízkosti katalytického místa trypsinu (v podmístech S_1' a S_1) a zbytek P_3, který se váže do S_3 podmísta subtilisinu. Párování cysteinových zbytků (žluté) do disulfidických můstků je vyznačeno svorkou. Šedým pozadím jsou zvýrazněny zbytky identické alespoň ve třech sekvencích. Barevné šipky udávají konzervativní pozici segmentů β-struktur (barevné kódování odpovídá topologii uvedené v obr. 1; nad šipkami je uvedeno značení segmentů. Přiřazení sekvencí bylo provedeno programem ClustalW (http://embnet.vital-it.ch/software/ClustalW.html) a upraveno podle superpozice prostorových struktur v programu PyMol.

Komplexy s trypsinem umožňují podrobnou srovnávací analýzu interakce. Obr. 2 dává na úrovni sekvencí inhibitorů základní informaci o poloze reaktivního místa, které zprostředkovává vazbu na trypsin. Reaktivní místo nemá stejné umístění v sekvenci jednotlivých inhibitorů, ale během molekulární evoluce se vytvořilo nezávisle na třech odlišných smyčkách umístěných mezi segmenty β-struktur. Je to smyčka β4-β5 u inhibitorů STI (cit.[30]), TKI (cit.[37]) a EcTI (cit.[38]) a dále smyčky β5-β6 a β9-β10 u inhibitoru API-A (cit.[33]), který je schopen interagovat se dvěma molekulami trypsinu (smyčky jsou značeny podle číslování okolních segmentů β-struktur v obr. 2).

Obr. 3 analyzuje prostorovou stavbu reaktivních smyček inhibitorů vázaných v aktivním místě trypsinu. Superpozice smyček ukazuje, že všechny Kunitzovy inhibitory využívají analogický mechanismus interakce s trypsinem. Ten umisťuje klíčový zbytek P_1 inhibitoru do vazebného podmísta S_1 trypsinu, které je v těsné blízkosti katalytických zbytků His57, Asp102 a Ser195. Podmísto S_1 je zodpovědné za substrátovou specifitu trypsinu s preferencí pro bazické aminokyseliny, proto Kunitzovy inhibitory využívají v pozici P_1 především zbytky Arg a Lys (vzácnější je nahrazení hydrofobní aminokyselinou jako je Leu87 v β5-β6 smyčce u inhibitoru API-A) (obr. 2 a 3). Tímto způsobem reaktivní smyčka napodobuje stavbu peptidového substrátu, který je trypsinem vázán a štěpen mezi pozicemi P_1 a P_1' (pozn. v konvenčním značení aminokyselinových zbytků substrátů NH_2 -P_x-P_2-P_1-P_1'-P_2'- P_x'-COOH je štěpena vazba P_1-P_1'). Uvedený mechanismus inhibice serinových proteas se nazývá Laskowského kanonický mechanismus[39,40].

Reaktivní smyčky Kunitzových inhibitorů, přestože jsou v těsné blízkosti katalytických zbytků trypsinu, nejsou efektivně štěpeny, protože jejich konformace je stabilizována. K tomu dochází pomocí interakce reaktivní smyčky s konzervovaným zbytkem Asn14 v architektuře inhibitoru, jako je tomu u inhibitorů STI, TKI a EcTI, nebo pomocí sítě vodíkových interakcí v rámci smyčky a pomocí disulfidických můstků, jako tomu je u inhibitoru API-A (cit.[33]) (obr. 2). Podrobně byla stabilizační role zbytku Asn14 studována u inhibitoru WCI s využitím mutantů, které byly funkčně a strukturně analyzovány[41,42] (tab. I).

Vznik reaktivních smyček s odlišným umístěním v architektuře Kunitzových inhibitorů umožnil evoluci



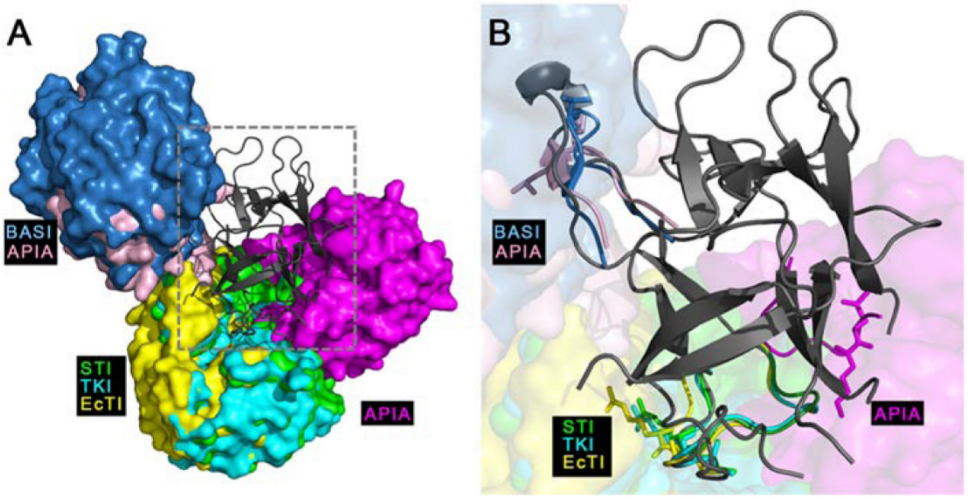
Obr. 3. **Jednotný mechanismus interakce trypsinu s reaktivními místy Kunitzových inhibitorů.** Molekula trypsinu je znázorněna šedě jako povrchový model, katalytické zbytky jsou zvýrazněny červeně. Jednotlivé inhibitory jsou znázorněny stužkovým modelem s barevným kódováním, které umožní identifikovat inhibitor v panelu s vysvětlivkami. V komplexu s trypsinem mají inhibitory rozdílnou orientaci, protože reaktivní místa inhibitorů mají odlišné umístění v architektuře molekuly inhibitoru. Zvětšený výřez obrázku vpravo ukazuje analogickou konformaci reaktivní smyčky inhibitorů, která se váže do katalytického místa trypsinu. Ve struktuře smyčky jsou zobrazeny postranní řetězce aminokyselinových zbytků, které se váží do S_1 a S_1' podmíst trypsinu. Ovál rámuje klíčovou interakci v S_1 podmístě mezi zbytkem Asp189 trypsinu a zbytkem v P_1 pozici inhibitoru (přehled P_1 zbytků a jejich barevné kódování jsou uvedeny v panelu s vysvětlivkami). Katalytická triáda trypsinu (His57, Asp102 a Ser195) je zobrazena červeně.

multifunkčních inhibitorů interagujících s více partnery. Na strukturní úrovni byl charakterizován ternární komplex inhibitoru API-A se dvěma molekulami trypsinu, jež se váží na nezávislé reaktivní smyčky umístěné na opačných stranách molekuly inhibitoru (obr. 4). Druhým bifunkčním inhibitorem je BASI, inhibitor serinové proteasy subtilisinu a α-amylasy. Byly určeny prostorové struktury obou komplexů, které ukazují rozdílnou stavbu a umístění příslušných reaktivních míst na molekule BASI.

V komplexu BASI-subtilisin je reaktivní místo inhibitoru tvořeno smyčkou β5-β6, která polohou odpovídá reaktivní smyčce pro trypsin na inhibitoru API-A (obr. 2 a 4). Inhibiční mechanismus pro subtilisin se ale liší od Laskowkého kanonického mechanismu, který využívají Kunitzovy inhibitory pro inhibici trypsinu. BASI interaguje se subtilisinem pouze v pozicích P_5-P_2, protože Cys90 v P_1 pozici tvoří disulfid a vychyluje tak reaktivní smyčku z katalytického centra subtilisinu, což ji chrání před štěpením (obr. 2). Pro inhibici je klíčový zbytek P3 (Thr88), který pomocí intramolekulárních interakcí stabilizuje reaktivní smyčku[43]. Vznik těchto odlišných inhibičních mechanismů lze vysvětlit strukturní odlišností subtilisinu a trypsinu, jež patří do jiných klanů serinových proteas, takže nejsou sekvenčně homologické, ačkoli mají analogický katalytický mechanismus (s katalytickou triádou zbytků Asp, His a Ser). Prostorová struktura komplexu BASI s α-amylasou ukazuje, že inhibitor využívá pro vazbu na enzym rozsáhlou interakční plochu (2355 Å^2), která zahrnuje šest smyček a dva segmenty β-struktury inhibitoru a je tak největší u dosud známých komplexů Kunitzových inhibitorů[27]. Na straně inhibitoru se interakce účastní celkem 26 aminokyselinových zbytků. Katalytické zbytky α-amylasy nejsou v přímém kontaktu s BASI a mezi oběma proteiny je specificky vázán iont vápníku.

4. Závěr

Rostlinné proteasové inhibitory z Kunitzovy rodiny jsou hojně zastoupené obranné molekuly rostlinných pletiv cílené proti proteolytickému systému hmyzích herbivorů a patogenů. Tyto inhibitory jsou unikátní svou širokou funkční diverzitou, která je umožněna molekulární architekturou inhibitorů se systémem variabilních smyček.



Obr. 4. **Reaktivní místa Kunitzových inhibitorů pro interakci s proteasami mají rozdílnou topologii.** (A) Superpozice prostorových struktur komplexů Kunitzových inhibitorů se serinovými proteasami. Centrálně je umístěn reprezentativní inhibitor znázorněný pomocí černého stužkového modelu. Molekuly serinových proteas, které interagují s inhibitory, jsou znázorněny pomocí povrchových modelů: barevné kódování proteas odpovídá popiskám, které uvádějí příslušný inhibitor tvořící komplex. Inhibitor BASI interaguje se subtilisinem, ostatní inhibitory s trypsinem (API-A se dvěma molekulami trypsinu). (B) Zvětšený výřez z panelu (A) ukazuje, jak jsou rozmístěna reaktivní centra jednotlivých inhibitorů v architektuře molekuly. Reaktivní smyčky inhibitorů jsou znázorněny ve stejném barevném kódování, jako mají příslušné interagující proteasy v panelu (A). Ve struktuře smyčky je zobrazen postranní řetězec aminokyselinových zbytků v pozicích P_1 a P_1', které se vážou do podmíst S_1 a S_1' trypsinu, a zbytku v pozici P_3, který se váže do podmísta S_3 subtilisinu.

V současné době v PDB databázi rychle přibývají prostorové struktury Kunitzových inhibitorů. Lze očekávat, že výzkum se nyní bude soustředit především na technicky náročnou strukturní analýzu komplexů inhibitorů s proteasami. Tím budou získány informace o reaktivních centrech inhibitorů a identifikovány nové inhibiční mechanismy pro jednotlivé typy proteas, zejména pro cysteinové a aspartátové proteasy. Na jejich základě bude možné konstruovat vysoce specifické inhibitory cílené proti různým typům proteas, které najdou uplatnění v zemědělství a biomedicíně.

Referát vznikl v rámci řešení grantového projektu InterBioMed LO1302 a výzkumného záměru RVO 61388963.

LITERATURA

1. Rawlings N. D., Salvesen G. (ed.): *Handbook of Proteolytic Enzymes*. 3. vyd. Elsevier, Amsterdam 2013.
2. Bode W., Huber R.: Biochim. Biophys. Acta *1477*, 241 (2000).
3. Law R. H. P., Zhang Q., McGowan S., Buckle A. M., Silverman G. A., Wong W., Rosado C. J., Langendorf C. G., Pike R. N., Bird P. I., Whisstock J. C.: Genome Biol. *7*, 216 (2006).
4. Oliva M. L., Ferreira R. S., Ferreira J. G., de Paula C. A. A., Salas C. E., Sampaio M. U.: Curr. Protein Pept. Sci. *12*, 348 (2011).
5. Kunitz M.: J. Gen. Physiol. *20*, 311 (1947).
6. Laskowski M. Jr., Qasim M. A.: Biochim. Biophys. Acta *1477*, 324 (2000).
7. Gatehouse J. A.: New Phytol. *156*, 145 (2002).
8. Jongsma M. A., Bolter C.: J. Insect Physiol. *43*, 885 (1997).
9. Habib H., Fazili K. M.: Biotechnol. Mol. Biol. Rev. *2*, 68 (2007).
10. Schlüter U., Benchabane M., Munger A., Kiggundu A., Vorster J., Goulet M. C., Cloutier C., Michaud D.: J. Exp. Bot. *61*, 4169 (2010).
11. Ferreira R. dS., Zhou D., Ferreira J. G., Silva M. C. C., Silva-Lucca R. A., Mentele R., Paredes-Gamero E. J., Bertolin T. C., Correia M. T. dS., Paiva P. M. G., Gustchina A., Wlodawer A., Oliva M. L.: PLoS One *8*, e64426 (2013).
12. Zani M. L., Moreau T.: Biochimie *92*, 1986 (2010).
13. Ishikawa A., Ohta S., Matsuoka K., Hattori T., Nakamura K.: Plant Cell Physiol. *35*, 303 (1994).
14. Khamrui S., Dasgupta J., Dattagupta J. K., Sen U.: Biochim. Biophys. Acta *1752*, 65 (2005).
15. Khamrui S., Majumder S., Dasgupta J., Dattagupta J. K., Sen U.: Protein Sci. *19*, 593 (2010).
16. Mareš M., Meloun B., Pavlík M., Kostka V., Baudyš M.: FEBS Lett. *251*, 94 (1989).
17. Ritonja A., Krizaj I., Mesko P., Kopitar M., Lucovnik P., Strukelj B., Pungercar J., Buttle D. J., Barrett A. J., Turk V.: FEBS Lett. *267*, 13 (1990).
18. Werner R., Guitton M. C., Mühlbach H. P.: Plant Physiol. *103*, 1473 (1993).
19. Brunelle F., Nguyen-Quoc B., Cloutier C., Michaud D.: Arch. Insect Biochem. Physiol. *42*, 88 (1999).

20. Kreft S., Ravnikar M., Mesko P., Pungercar J., Umek A., Kregar I., Strukelj B.: Phytochemistry *44*, 1001 (1997).
21. Cater S. A., Lees W. E., Hill J., Brzin J., Kay J., Phylip L. H.: Biochim. Biophys. Acta *1596*, 76 (2002).
22. Guo J., Erskine P. T., Coker A. R., Wood S. P., Cooper J. B., Mareš M., Baudyš M.: J. Struct. Biol. *192*, 554 (2015).
23. Guerra Y., Valiente P. A., Berry C., Pons T.: J. Mol. Model. *18*, 2673 (2012).
24. Hansen D., Macedo-Ribeiro S., Veríssimo P., Yoo Im S., Sampaio M. U., Oliva M. L.: Biochem. Biophys. Res. Commun. *360*, 735 (2007).
25. Krizaj I., Drobnic-Kosorok M., Brzin J., Jerala R., Turk V.: FEBS Lett. *333*, 15 (1993).
26. Renko M., Sabotic J., Mihelic M., Brzin J., Kos J., Turk D.: J. Biol. Chem. *285*, 308 (2010).
27. Vallée F., Kadziola A., Bourne Y., Juy M., Rodenburg K. W., Svensson B., Haser R.: Structure *6*, 649 (1998).
28. Yamasaki T., Deguchi M., Fujimoto T., Masumura T., Uno T., Kanamaru K., Yamagata H.: Biosci. Biotechnol. Biochem. *70*, 1200 (2006).
29. Koizumi A., Tsuchiya A., Nakajima K., Ito K., Terada T., Shimizu-Ibuka A., Briand L., Asakura T., Misaka T., Abe K.: Proc. Natl. Acad. Sci. U.S.A. *108*, 16819 (2011).
30. Song H. K., Suh S. W.: J. Mol. Biol. *275*, 347 (1998).
31. Renko M., Sabotič J., Turk D.: Biol. Chem. *393*, 1043 (2012).
32. Murzin A. G., Lesk A. M., Chothia C.: J. Mol. Biol. *223*, 531 (1992).
33. Bao R., Zhou C. Z., Jiang C., Lin S. X., Chi C. W., Chen Y.: J. Biol. Chem. *284*, 26676 (2009).
34. de Oliveira C., Santana L. A., Carmona A. K., Cezari M. H., Sampaio M. U., Sampaio C. A., Oliva M. L.: Biol. Chem. *382*, 847 (2001).
35. Zhou D., Hansen D., Shabalin I. G., Gustchina A., Vieira D. F., de Brito M. V., Araújo A. P., Oliva M. L., Wlodawer A.: Acta Crystallogr., Sect. F: Struct. Biol. Commun. *71*, 1055 (2015).
36. Oliva M. L., Silva M. C., Sallai R. C., Brito M. V., Sampaio M. U.: Biochimie *92*, 1667 (2010).
37. Patil D. N., Chaudhary A., Sharma A. K., Tomar S., Kumar P.: FEBS J. *279*, 4547 (2012).
38. Zhou D., Lobo Y. A., Batista I. F., Marques-Porto R., Gustchina A., Oliva M. L., Wlodawer A.: PLoS One *8*, e62252 (2013).
39. Laskowski M. Jr., Kato I.: Annu. Rev. Biochem. *49*, 593 (1980).
40. Bode W., Huber R.: Eur. J. Biochem. *204*, 433 (1992).
41. Dasgupta J., Khamrui S., Dattagupta J. K., Sen U.: Biochemistry *45*, 6783 (2006).
42. Ravichandran S., Dasgupta J., Chakrabarti C., Ghosh S., Singh M., Dattagupta J. K.: Protein Eng., Des. Sel. *14*, 349 (2001).
43. Micheelsen P. O., Vévodová J., De Maria L., Ostergaard P. R., Friis E. P., Wilson K., Skjøt M.: J. Mol. Biol. *380*, 681 (2008).

J. Srp[a,b] and M. Mareš[a] *([a] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., Prague, [b] Department of Biochemistry, Faculty of Science, Charles University, Prague):*
Plant Kunitz Protease Inhibitors: Structural and Functional Diversity

Protease inhibitors of the plant Kunitz family (I3 or Kunitz-P family) are widely distributed in the plant kingdom. They are critically involved in plant defense against insect herbivores and pathogens through the regulation of exogenous proteases. The majority of characterized members of the Kunitz family are inhibitors of proteases of the serine class (S1 and S8 families); inhibitors of aspartic proteases (A1 family) and cysteine proteases (C1 family) are less abundant. Some of the inhibitors are double-headed, with two independent reactive sites, or bifunctional, targeting different protease families. The high functional versatility stems from a system of multiple variable loops exposed on the conserved β-trefoil scaffold. Several reactive sites on the loops with different locations evolved independently during evolution as demonstrated by recent 3D structures of the complexes of Kunitz inhibitors with serine proteases. This review provides new insight into structure-function relationships in the Kunitz inhibitor family.

5. Diskuze

Proteasy se zásadním způsobem účastní řady fyziologicky významných dějů, jako jsou trávení, presentace antigenů, srážení krve, regulace tlaku krve a mnoha dalších. Nedostatečná regulace proteolytické aktivity těchto enzymů vede k závažným patologiím, např. nádorovým, neurodegenerativním a kardiovaskulárním onemocněním. Proto jsou specifické molekuly, které účinně inhibují aktivitu medicinálně významných proteas, považovány za perspektivní chemoterapeutika. K jejich racionálnímu vývoji jsou intenzivně studovány strukturní mechanismy umožňující kontrolu proteas a unikátní informace v tomto směru poskytuje analýza přirozených antiproteolytických molekul, které jsou součástí komplexních fyziologických systémů. Cílem disertační práce bylo popsat nové principy regulace vybraných zástupců aspartátových a serinových proteas pomocí interakce s přirozenými proteasovými inhibitory.

5.1. Aspartátové proteasy a jejich regulace přirozenými inhibitory

K výrazným úspěchům molekulární medicíny při využití poznatků z oblasti proteolýzy patří chemoterapie pomocí inhibitorů aspartátové proteasy viru HIV z rodiny A2 retrovirálních proteas. Tato proteasa je zodpovědná za štěpení gag-pol polyproteinů při tzv. zrání viru, kterým vznikají infekční virové částice. Díky úsilí akademické sféry a farmaceutického průmyslu je v současné době k dispozici desítka specifických inhibitorů HIV proteasy pro antivirovou terapii (Pokorná et al. 2009). Disertační práce se zabývala aspartátovými proteasami z pepsinové rodiny A1, do které patří medicinálně významné enzymy, jako jsou např. renin, β sekretasa nebo katepsin D. Renin je klíčový enzym renin-angiotensinového systému regulujícího krevní tlak a rovnováhu extracelulárních tekutin v lidském organismu. Aliskiren, specifický inhibitor reninu, je široce rozšířené léčivo při hypertenzi (Nadeem a Batisky 2014). Zvýšená aktivita β sekretasy vede ke štěpení amyloidního prekursorového proteinu za vzniku β amyloidních peptidů, které se ukládají v mozkové tkáni a způsobují Alzheimerovu chorobu. V současné době je v klinických testech několik molekul, které specificky inhibují β sekretasu (Ghosh, Brindisi, a Tang 2012; Volloch a Rits 2018). Katepsin D, který v lysosomu degraduje buněčné proteiny a generuje peptidy pro presentaci na MHC II molekulách, je součástí řady nádorových procesů, kde se podílí na mitotické signalizaci a také degradaci extracelulární matrix a aktivaci růstových faktorů (Fusek, Větvičková, a Větvička 2007). Pro katepsin D nejsou doposud známé žádné specifické inhibitory testované v klinických studiích. V této disertační práci byly popsány dva nové strukturní motivy pro inhibiční interakci s aspartátovými proteasami rodiny pepsinu, které mají odlišný charakter - proteinové domény a sfingolipidu.

V publikaci č. 2 byl prokázán specifický modulační efekt bioaktivních lipidů na peptidolytickou aktivitu lidského katepsinu D, kdy určité sfingolipidy odvozené ze struktury

sfingosinu a ceramidu potlačují jeho aktivitu a jejich fosforylované deriváty naopak aktivitu katepsinu D zvyšují.

Výsledky uvedené v publikaci č. 2 naznačují, že se sfingolipidy vážou do aktivního místa katepsinu D, které preferuje hydrofobní aminokyseliny. Nicméně pro přesný popis a plné porozumění mechanismu interakce mezi sfingolipidem a katepsinem D bude nutné vyřešit prostorovou strukturu jejich komplexu. Na jejím základě bude možné racionálním způsobem navrhovat modifikované syntetické molekuly lipidů a jejich mimetika s cílem zlepšit inhibiční potenciál a další funkční vlastnosti. Fosforylace primárního hydroxylu u sfingolipidů vede ke zvýšení peptidolytické aktivity katepsinu D, která je způsobena převážně zlepšením afinity substrátu k enzymu. Vazebné místo pro fosforylované lipidy také není známé, ale lze spekulovat, že se jedná o pozitivně nabitou povrchovou část katepsinu D v okolí jeho aktivního centra. Analogicky interagují glykosaminoglykany s pozitivně nabitými povrchovými oblastmi cysteinových proteas katepsinového typu (Fairhead, Kelly, a van der Walle 2008; Caglic et al. 2007). Popsaný posun pH optima katepsinu D v přítomnosti fosforylovaných sfingolipidů o jednotku výše by mohl významně zvyšovat aktivitu tohoto enzymu v situacích, kdy je transportován mimo lysosom, k čemuž dochází např. za patofyziologických podmínek (Masson et al. 2010; Beneš, Větvička, a Fusek 2008; Beaujouin et al. 2010).

V jiných pracích bylo prokázáno, že tyto bioaktivní lipidy mají přímou souvislost s apoptózou a jejich fosfoderiváty působí naopak jako mitotické molekuly (Ponnusamy et al. 2010), tedy s procesy, ve kterých je zapojen i katepsin D (Beneš, Větvička, a Fusek 2008). Proto lze předpokládat, že vliv těchto lipidů na popsané buněčné procesy bude souviset s aktivitou katepsinu D, který se účastní řady nádorových procesů. Sfingolipidové templáty mají tedy potenciální využití při vývoji selektivních molekul regulujících procesy, ve kterých je zapojen katepsin D v protinádorové terapii.

Rostlinný proteasový inhibitor PCDI z lilku brambor, který patří do Kunitzovy rodiny, je specifický inhibitor katepsinu D (Mareš et al. 1989). V rámci disertační práce bylo identifikováno a strukturně charakterizováno reaktivní centrum, které je zodpovědné za interakci PCDI s enzymem. Jedná se o 22 aminokyselin dlouhou smyčku, která je stabilizována dvěma disulfidovými můstky a váže se do aktivního místa katepsinu D. V genomu lilku brambor se vyskytuje i další protein z Kunitzovy rodiny označený jako NID, pro který byla popsána inhibice aspartátových proteas, ale reaktivní centrum není známé (Ritonja et al. 1990; Guerra et al. 2016). Porovnání aminokyselinových sekvencí PCDI a NID ukázalo, že oba proteiny obsahují motiv reaktivní smyčky se dvěma disulfidy. Ačkoli celkové aminokyselinové sekvence inhibitorů jsou vysoce homologické (94% identita), vykazují určené interakční smyčky výraznou variabilitu (56% identita). Rozdílnost ve složení interakční smyčky může být způsobena tím, že oba inhibitory cílí na rozdílné spektrum aspartátových proteas, což souvisí s předpokládanou fyziologickou rolí Kunitzových inhibitorů jako obranných proteinů v rostlinných pletivech; jejich funkční variabilita umožňuje inhibici odlišných izoenzymů trávicích proteas produkovaných herbivory (viz kap. 5.3 s diskuzí k publikaci č. 1). Strukturně

popsaná interakční smyčka PCDI je v kontextu jiných inhibitorů aspartátových proteas nový, zcela unikátní inhibiční motiv. Doposud byly analyzovány pouze čtyři strukturní motivy pro inhibici těchto proteas: (1) peptidomimetický inhibitor pepstatin se váže do aktivního centra enzymu a jeho statinový zbytek interaguje s katalytickými aspartáty (Baldwin et al. 1993) (Obr. 3 A, str. 8), (2) polypeptidový inhibitor IA3 při vazbě na proteinasu A tvoří α helix, který překrývá aktivní místo enzymu (Li et al. 2000) (Obr. 6 A, str. 19), (3) proteinový inhibitor PI3 blokuje pozice S_1'-S_3' aktivního místa pepsinu a brání tak vstupu substrátu (Obr. 6 B, str. 19) (Ng et al. 2000) a (4) oligopeptidový inhibitor vyštěpený z propeptidové domény zymogenu katepsinu D se váže do povrchového exo-místa na aktivní protease a způsobuje její alosterickou inhibici (Hánová et al. 2018) (Obr. 5 B, str. 18). Reaktivní centrum PCDI vytváří strukturně kompaktní a relativně nezávislou malou doménu na povrchu inhibitoru. Lze proto předpokládat, že bude možné navrhovat syntetická mimetika reaktivního centra v podobě cyklických peptidů a dále upravovat jejich funkční vlastnosti včetně inhibičního potenciálu a specifity pro jiné medicinálně významné aspartátové proteasy. Tento přístup byl již úspěšně otestován na jiných peptidových cyklických inhibitorech serinových proteas (Xu, Andreasen, a Huang 2017; Swedberg et al. 2011; de Veer et al. 2015).

V publikaci č. 3 byly na základě inhibitoru aspartátových proteas pepstatinu (Majer et al. 1997) navrženy a připraveny syntetické proteomické sondy, které jsou selektivní pro detekci proteolyticky aktivních forem aspartátových proteas rodiny pepsinu. Bylo prokázáno, že připravené sondy jsou schopné vizualizovat katepsin D v komplexních proteinových směsích z extraktů nádorových buněk. Představují tak nové účinné nástroje, které lze využít pro detekci katepsinu D jako prognostického markeru u nádorového onemocnění prsu a terapeutické cílové molekuly (Brouillet et al. 1997). Modulární složení připravených sond umožňuje jejich využití při různých proteomických aplikacích jako je analýza aspartátových proteas v proteolytických systémech, včetně jejich identifikace pomocí hmotnostní spektrometrie nebo N koncového sekvenování.

5.2. Serinové proteasy a jejich regulace přirozenými inhibitory

Serinové proteasy chymotrypsinové rodiny, mezi které patří modeloví zástupci trypsin a chymotrypsin lišící se svou substrátovou specifitou, jsou nejvíce zastoupené proteasy v lidském genomu. Do této rodiny patří i trombin a proteasové faktory, které se účastní koagulační kaskády srážení krve (Huntington 2014) nebo kalikreinové proteasy, které mají řadu specifických fyziologických funkcí a jsou spojovány s nádorovými onemocněními (Kalinska et al. 2016; Masurier et al. 2018). Proto jsou tyto proteolytické enzymy perspektivní cílové molekuly a jejich inhibitory představují potenciální farmaka. Jako příklad lze uvést Dabigatran, syntetický reversibilní inhibitor trombinu, který patří mezi hojně užívané léky proti infarktu myokardu nebo mozkové mrtvici (Ageno, Eikelboom, a Lip 2016). Přirozené proteinové inhibitory serinových proteas chymotrypsinové rodiny byly intenzivně studovány po biochemické i strukturní stránce. V PDB databázi je řada zástupců těchto inhibitorů z několika inhibičních rodin, včetně jejich komplexů. Jejich analýza ukazuje, že

inhibitory tvořící se serinovými proteasami nekovalentní protein-proteinové komplexy interagují s cílovou proteasou pomocí analogického reaktivního centra založeného na jedné interakční smyčce, která vstupuje do aktivního místa proteasy. Tento interakční mechanismus se nazývá kanonický nebo také Laskowského mechanismus (Krowarsch et al. 2003) a byl detailně studován pro živočišný inhibitor BPTI (Helland et al. 1999) a pro rostlinné inhibitory z Kunitzovy rodiny (Bendre, Ramasamy, a Suresh 2018).

Protein BbKI produkovaný rostlinou bauhinie patří do Kunitzovy rodiny a byla popsána jeho schopnost inhibovat trypsin a plasmový kalikrein (vykazující trypsinovou substrátovou specifitu) (Zhou et al. 2013). Vyřešená prostorová struktura komplexu BbKI s trypsinem v publikaci č. 5 ukázala, že BbKI využívá pro interakci se serinovými proteasami jednosmyčkové reaktivní centrum a umisťuje do kritického S_1 podmísta trypsinu argininový zbytek, jedná se tedy o kanonický mechanismus interakce. V této práci bylo prokázáno, že BbKI je schopen inhibovat i další zástupce kalikreinových proteas a jejich interakce byla detailně popsána pomocí molekulárních modelů. Pro kalikrein 4, který je spojován s rakovinou prostaty (Xi et al. 2004), je BbKI s hodnotou K_i ~45 pM nejlepší přirozený inhibitor, který byl doposud popsán. Kalikrein 7, který má chymotrypsinovou substrátovou specifitu, byl také účinně inhibován BbKI s K_i v subnanomolární oblasti.

V publikaci č. 4 byla studována specifita hlavních zástupců epidermálních kalikreinů (kalikreiny 5, 7 a 14), pro které byl sestaven detekční proteomický systém založený na syntetických selektivních substrátech a inhibitorech. Tento nástroj může být využit pro rozlišení a kvantifikaci aktivit jednotlivých kalikreinových proteas v komplexním proteomu, což bylo potvrzeno v testech s extrakty epidermis geneticky upravených myší, které měly defekty v kalikreinové proteolýze.

5.3. Identifikace nových reaktivních center na rostlinných proteasových inhibitorech z Kunitzovy rodiny

Proteasové inhibitory patří mezi přirozené obranné proteiny rostlin cílené proti trávicím proteasám škůdců, např. herbivorního hmyzu. Jejich interakce je komplexní díky dlouhotrvající společné evoluci. Hmyz se adaptoval na přítomnost rostlinných inhibitorů ve stravě a reaguje na ně zvýšením produkce trávicích enzymů, expresí izoenzymů, které nejsou senzitivní na určité inhibitory, nebo proteolytickou deaktivací těchto obranných proteinů (Wielkopolan a Obrepalska-Steplowska 2016; Zhu-Salzman a Zeng 2015). Disertační práce se zaměřila na analýzu modelového proteolytického systému mandelinky bramborové a jeho interakce s antiproteolytickým systémem inhibičních proteinů lilku brambor. Larvy i dospělci mandelinky bramborové způsobují významné ekonomické ztráty požerem listů lilkovitých rostlin. Jedním z možných přístupů pro regulaci tohoto škůdce je využiti přirozených inhibičních molekul blokující jeho trávicího systému, což vede k snížení schopnosti přežít, růstu a množení (Bolter a LatoszekGreen 1997; Wolfson a Murdock 1987).

V publikaci č. 1 byl s využitím nástrojů funkční proteomiky (selektivních substrátů, inhibitorů a aktivitních sond) popsán trávicí proteolytický systém mandelinky bramborové. Na základě zjištěných výsledků bylo možné sestavit schéma procesu trávení proteinů, kterého se účastní cysteinové, aspartátové a serinové proteasy. Pro klíčový enzym zahajující degradaci proteinů, aspartátovou proteasu typu katepsinu D (LdCD), byla v rámci disertační práce také vyřešena prostorová struktura (Obr. 14, str. 61). Dále bylo v publikaci č. 1 prokázáno, že inhibitory z Kunitzovy rodiny produkované v lilku brambor (PCDI a PSPI) jsou schopné inhibovat aspartátové a serinové proteasy mandelinky bramborové. Jedná se o první přímý důkaz interakce Kunitzova inhibitoru s konkrétní cílovou trávicí proteasou a identifikaci fyziologické role PCDI a PSPI.

Další krok se zaměřil na strukturní popis interakce PCDI a PSPI s modelovými proteasami s cílem popsat nové typy inhibičních mechanismů. V kapitole 5.1 je diskutovaná identifikace interakční smyčky na PCDI pro aspartátové proteasy (Obr. 12, str. 59), která společně s prostorovou strukturou cílové proteasy LdCD umožnila sestavit model interakce těchto biologicky relevantních partnerů (Obr. 15, str. 62). Dále byly pro PSPI a PCDI vyřešeny tři prostorové struktury se serinovými proteasami. PSPI tvoří ternární komplexy se dvěma molekulami serinových proteas, tzn. že PSPI má dvě nezávislá reaktivní centra, která se liší pozicí na molekule inhibitoru. První reaktivní centrum využívá kanonicky jednosmyčkový mechanismus při interakci s trypsinem (Obr. 9, str. 53). Má stejnou topologii a stavbu jako reaktivní centrum na Kunitzově inhibitoru API-A z šípatky (Bao et al. 2009), u kterého byla popsaná stabilizační role cysteinového zbytku (Cys89), který se nachází v těsné blízkosti reaktivního centra a tvoří disulfidový můstek s Cys43 z jiné smyčky. Druhé reaktivní centrum inhibitoru PSPI využívá nový, dosud nepopsaný typ interakčního mechanismu. Je tvořené dvěma nezávislými smyčkami pocházející ze sekvenčně vzdálených oblastí inhibitoru (Obr. 13 B, str. 60), které se v aktivním místě enzymu skládají do jednoho reaktivního centra (Obr. 9 a 10, str. 53 a 55). Pro inhibitor PSPI bylo pomocí dvou prostorových struktur prokázáno, že dvousmyčkové reaktivní centrum je schopno vázat trypsin (Obr. 9, str. 53) i chymotrypsin (Obr. 10, str. 55), tedy proteasy s odlišnou substrátovou specifitou. Interakční mechanismus založený na systému dvou smyček využívá i inhibitor PCDI při vazbě trypsinu, jak bylo potvrzeno na prostorové struktuře jejich společného komplexu (Obr. 11, str. 57). V literatuře je kanonické jednosmyčkové reaktivní centrum označováno jako „pomalý substrát“, protože v komplexu se serinovou proteasou podléhá za určitých podmínek degradaci, kdy je štěpeno v pozicích mezi P_1 a P_1' zbytky (Krowarsch et al. 2003). Reaktivní centrum složené ze dvou smyček je naopak stabilní, protože tyto kritické pozice jsou obsazeny zbytky, které pochází ze dvou nespojitých smyček.

V disertační práci byla identifikovaná na sérii prostorových struktur komplexů rostlinných inhibitorů z Kunitzovy rodiny tři reaktivní centra pro vazbu serinových proteas a jedno reaktivní centrum pro inhibici aspartátových proteas. Z hlediska umístění jednotlivých reaktivních center na architektuře molekuly inhibitoru je kanonické reaktivní centrum BbKI v pozici, která byla prokázána

na dalších Kunitzových inhibitorech (STI, TKI a EcTI) v komplexu s trypsinem (Song a Suh 1998; Patil et al. 2012; Dongwen Zhou et al. 2013). Jednosmyčkové kanonické reaktivní centrum na PSPI se shoduje svou topologií s jedním z center na API-A (Bao et al. 2009). Druhé reaktivní centrum pro serinové proteasy na API-A je ve stejné pozici jako centrum proti aspartátovým proteasám na PCDI molekule. Dvousmyčkové reaktivní centrum na inhibitorech PSPI a PCDI se vyskytuje na smyčkách, pro které doposud nebyla popsána interakční funkce pro inhibici proteas. Z této analýzy vyplývá, že rostlinné inhibitory Kunitzovy rodiny díky variabilitě v délce i sekvenci smyček byly schopné vyvinout reaktivní centra proti různým typům proteas v odlišných pozicích molekuly a tím dokázaly tyto obranné molekuly efektivně reagovat na evoluční tlak. Molekulární architektura Kunitzových proteinů představuje vysoce flexibilní templát pro racionální navrhování nových antiproteolytických funkcí (a obecně protein-proteinových interakcí) a jejich kombinování do multifunkčních proteinů. Popsaných unikátních vlastností inhibitorů z Kunitzovy rodiny je možné využít při konstrukci rezistentních transgenních plodin produkujících nové insekticidní inhibiční proteiny nebo při vývoji specifických proteasových inhibitorů pro biomedicínu.

6. Závěry

Disertační práce se zabývala regulací aktivity proteolytických enzymů ze třídy aspartátových a serinových proteas. Výsledky jsou shrnuty v pěti publikacích a jednom rukopisu, který je podáván do mezinárodního časopisu. Práce také obsahuje doposud nepublikované výsledky, které budou součástí dvou připravovaných publikací v impaktovaném časopise, kde bude autor disertace uveden jako první autor.

Obecně práce přináší první ucelený biochemický popis proteolytického trávicího systému mandelinky bramborové, herbivorního škůdce lilkovitých rostlin, a nové poznatky o jeho regulaci pomocí proteasových inhibitorů Kunitzovy rodiny z lilku brambor. Také podává strukturní popis nových reaktivních center pro serinové a aspartátové proteasy na těchto Kunitzových inhibitorech. Dále jsou prezentovány nové poznatky o přirozené regulaci aktivity aspartátové proteasy lidského katepsinu D pomocí biogenních sfingolipidů a vývoj aktivitních sond pro detekci tohoto enzymu. Pro serinové proteasy kalikreinového typu byl sestaven systém pro monitorování aktivity kalikreinů v tkáni epidermis a byl identifikován nejvíce efektivní přirozený inhibitor lidských kalikreinů, který patří mezi rostlinné Kunitzovy inhibitory.

V disertační práci byly splněny zadané cíle a získané výsledky jsou následující:

1) S využitím moderních proteomických přístupů byl popsán proteolytický trávicí systém mandelinky bramborové, který je založen na kombinovaném působení cysteinových, aspartátových a serinových proteas. Jednotlivé enzymy byly identifikovány a charakterizovány pomocí selektivních substrátů a inhibitorů a vizualizovány v extraktech střev mandelinky bramborové pomocí specifických aktivitních sond. Na základě získaných výsledků bylo možné sestavit schéma degradace rostlinných proteinů v procesu trávení tohoto herbivorního škůdce lilkovitých rostlin. Významným poznatkem bylo prokázání, že Kunitzovy inhibitory z lilku brambor efektivně blokují aktivitu serinových a aspartátových proteas mandelinky bramborové a fungují jako obranné proteiny cílené proti trávicímu traktu herbivora.

2) Podařilo se vyřešit unikátní sadu čtyř prostorových struktur protein-proteinových komplexů, které jsou tvořeny inhibitory z Kunitzovy rodiny v komplexu s cílovými proteasami. Na základě těchto struktur bylo možné popsat nová reaktivní centra pro interakci se serinovými a aspartátovými proteasami, která měla rozdílnou topologii na molekule inhibitoru. Konkrétně se jednalo o:

    a) Dvě reaktivní centra pro serinové proteasy: první je tvořené jedinou smyčkou s klasickým mechanismem interakce popsaným v literatuře a druhé se váže dosud neznámým mechanismem, který je založený na dvou smyčkách, jež se skládají do aktivního místa enzymu. Nový typ reaktivního centra je schopen interagovat se serinovými proteasami s odlišnou substrátovou specifitou - trypsinem a chymotrypsinem.

b) Unikátní reaktivní centrum pro aspartátové proteasy tvořené jednou smyčkou, která je stabilizována dvěma vnitřními disulfidovými můstky. Jeho mechanismus interakce byl identifikován vyřešením první prostorové struktury Kunitzova inhibitoru v komplexu s aspartátovou proteasou.

3) Byla vyřešena prostorová struktura trávicí aspartátové proteasy z mandelinky bramborové. Tato struktura společně s údaji o pozici a struktuře reaktivního centra pro aspartátové proteasy na molekule Kunitzova inhibitoru (viz 2b) umožnila navrhnout model interakce mezi biologicky relevantními partnery - trávicím enzymem mandelinky bramborové a obraným proteinem lilku brambor.

4) Sfingolipidy byly popsány jako první endogenní molekuly, které jsou schopné regulovat aktivitu aspartátové proteasy typu katepsinu D. V závislosti na struktuře byly sfingolipidy identifikovány jako účinné inhibitory katepsinu D, zatímco jejich fosforylované deriváty jako aktivátory tohoto enzymu. Jedná se o specifickou modulaci aktivity katepsinu D, která se potenciálně uplatňuje při regulaci tvorby nádorů.

5) Byly navrženy a připraveny syntetické aktivitní sondy pro aspartátové proteasy pepsinové rodiny, které byly odvozené ze sekvence přirozeného specifického inhibitoru pepstatinu. Bylo prokázáno, že tyto sondy je možné využít k detekci katepsinu D jako markeru v buněčných extraktech při prognóze nádorových onemocnění.

6) Byly nalezeny specifické fluorogenní substráty a selektivní inhibitory, které umožňují rozlišit aktivity jednotlivých epidermálních serinových proteas kalikreinového typu v biologických extraktech. Tento detekční systém byl využit pro analýzu příslušných kalikreinů u geneticky modifikovaných myší jako modelu pro studium kožních onemocnění.

7) Byla určena inhibiční specifita inhibitoru BbKI z Kunitzovy rodiny se serinovými proteasami. Tento inhibitor byl identifikován jako nejúčinnější přirozený proteinový inhibitor některých medicinálně významných lidských kalikreinů s preferencí pro kalikrein 4, který je spojovaný s nádorovými onemocněními. Mechanismus interakce BbKI s kalikreiny byl vysvětlen pomocí molekulárního modelování na základě vyřešené prostorové struktury komplexu BbKI s trypsinem jako templátu.

Prohlášení autorů:

Mgr. Jaroslav Srp se významně podílel na přípravě publikací uvedených v disertační práci.

Publikace č. 1: Digestive proteolysis in the Colorado potato beetle, *Leptinotarsa decemlineata*: Activity-based profiling and imaging of a multipeptidase network

(1) Příprava a charakterizace tkáňových extraktů a analýza jejich proteolytické aktivity s využitím enzymologických a biochemických přístupů, (2) identifikace a klasifikace proteas pomocí měření jejich enzymových aktivit a inhibice, (4) vizualizace proteolytických enzymů v extraktech pomocí fluorescenčních sond, (5) bioinformatická analýza identifikovaných proteas v genomových a transkriptomových databázích, (6) příprava obrázků a účast na psaní manuskriptu.

Nepublikovaná data:

(1) Příprava dvou rekombinantních katepsinů D v čisté podobě - IrCD1 a LdCD, (2) purifikace inhibitorů PCDI a PSPI z lilku brambor, (3) příprava komplexů LdCD-pepstatin, katepsin D-PCDI, trypsin-PCDI, trypsin-PSPI, trypsin/chymotrypsin-PSPI, (4) krystalizace těchto komplexů, (5) grafické zpracování a interpretace získaných prostorových struktur komplexů.

Publikace č. 2: Complex modulation of peptidolytic activity of cathepsin D by sphingolipids

(1) Měření aktivit katepsinu D v přítomnosti vybraných sfingolipidů v aktivitním testu, (2) příprava a HPLC analýza vzorků s modulovanou aktivitou katepsinu D, (3) příprava grafiky pro manuskript.

Publikace č. 3: Single- and double-headed chemical probes for detection of active cathepsin D in a cancer cell proteome

(1) Příprava a charakterizace extraktu z nádorových buněk, (2) značení katepsinu D v extraktu pomocí aktivitních sond, (3) analýza a identifikace vzorků vizualizovaných ve fluorescenční a chemiluminiscenční aplikaci.

Publikace č. 4: Profiling system for skin kallikrein proteolysis applied in gene deficient mouse models

(1) Měření inhibiční specifity pro jednotlivé rekombinantní KLK se sadou syntetických inhibitorů.

Publikace č. 5: Crystal structures of the complex of kallikrein inhibitor BbKI with trypsin and modeling of kallikrein complexes

(1) Měření inhibičních konstant K_i pro BbKI se zástupci serinových proteas kalikreinového typu a modelových serinových proteas trypsinu a chymotrypsinu a (2) analýza enzymologických dat.

Publikace č. 6: Kunitzovy proteasové inhibitory z rostlin: strukturní a funkční diverzita

(1) Vyhledání dostupné literatury a prostorových struktur v PDB databázi, (2) analýza prostorových struktur a sekvencí, (3) příprava molekulární grafiky manuskriptu.

……..

RNDr. Michael Mareš, CSc.

Školitel a spoluautor všech uvedených publikací

7. Seznam použité literatury

Ageno, W., J. Eikelboom, and G. Y H Lip. 2016. "Dabigatran in Clinical Practice: Contemporary Overview of the Evidence." *International Journal of Cardiology* 220: 417–28.

Ahn, J. E., and K. Zhu-Salzman. 2009. "CmCatD, a Cathepsin D-like Protease Has a Potential Role in Insect Defense against a Phytocystatin." *Journal of Insect Physiology* 55 (8): 678–85.

Baldwin, E. T., T. N. Bhat, S. Gulnik, M. V. Hosur, R. C. Sowder, R. E. Cachau, J. Collins, A. M. Silva, and J. W. Erickson. 1993. "Crystal Structures of Native and Inhibited Forms of Human Cathepsin D: Implications for Lysosomal Targeting and Drug Design." *Proceedings of the National Academy of Sciences of the United States of America* 90 (14): 6796–6800.

Bao, R., C. Zhou, C. Jiang, S. Lin, C. Chi, and Y. Chen. 2009. "The Ternary Structure of the Double-Headed Arrowhead Protease Inhibitor API-A Complexed with Two Trypsins Reveals a Novel Reactive Site Conformation." *The Journal of Biological Chemistry* 284 (39): 26676–84.

Barao, S., D. Moechars, S. F. Lichtenthaler, and B. De Strooper. 2016. "BACE1 Physiological Functions May Limit Its Use as Therapeutic Target for Alzheimer's Disease." *Trends in Neurosciences* 39 (3): 158–69.

Beaujouin, M., C. Prebois, D. Derocq, V. Laurent-Matha, O. Masson, S. Pattingre, P. Coopman, et al. 2010. "Pro-Cathepsin D Interacts with the Extracellular Domain of the Beta Chain of LRP1 and Promotes LRP1-Dependent Fibroblast Outgrowth." *Journal of Cell Science* 123 (19): 3336–46.

Bendre, A. D., S. Ramasamy, and C. G. Suresh. 2018. "Analysis of Kunitz Inhibitors from Plants for Comprehensive Structural and Functional Insights." *International Journal of Biological Macromolecules* 113: 933–43.

Beneš, P., V. Vetvicka, and M. Fusek. 2008. "Cathepsin D-Many Functions of One Aspartic Protease." *Critical Reviews in Oncology/Hematology* 68 (1): 12–28.

Beton, D. , C. R. Guzzo, A. F. Ribeiro, C. S. Farah, and W. R. Terra. 2012. "The 3D Structure and Function of Digestive Cathepsin L-like Proteinases of Tenebrio Molitor Larval Midgut." *Insect Biochemistry and Molecular Biology* 42 (9): 655–64.

Bode, W., and R. Huber. 1992. "Natural Protein Proteinase Inhibitors and Their Interaction with Proteinases." *European Journal of Biochemistry* 204 (2): 433–51.

Bode, W., and P. Schwager. 1975. "The Single Calcium-Binding Site of Crystallin Bovin Beta-Trypsin." *FEBS Letters* 56 (1): 139–43.

Bolter, C. J., and M. LatoszekGreen. 1997. "Effect of Chronic Ingestion of the Cysteine Proteinase Inhibitor, E-64, on Colorado Potato Beetle Gut Proteinases." *Entomologia Experimentalis Et Applicata* 83 (3): 295–303.

Bonneau, L., Y. Ge, G. E. Drury, and P. Gallois. 2008. "What Happened to Plant Caspases?" *Journal of Experimental Botany* 59 (3): 491–99.

Brouillet, J. P., F. Dufour, G. Lemamy, M. Garcia, N. Schlup, J. Grenier, J. C. Mani, and H. Rochefort. 1997. "Increased Cathepsin D Level in the Serum of Patients with Metastatic Breast Carcinoma Detected with a Specific Pro-Cathepsin D Immunoassay." *Cancer* 79 (11): 2132–36.

Brunelle, F., B. Nguyen-Quoc, C. Cloutier, and D. Michaud. 1999. "Protein Hydrolysis by Colorado Potato Beetle, *Leptinotarsa Decemlineata*, Digestive Proteases: The Catalytic Role of Cathepsin D." *Archives of Insect Biochemistry and Physiology* 42 (1): 88–98.

Caglic, D., J. Rozman Pungercar, G. Pejler, V. Turk, and B. Turk. 2007. "Glycosaminoglycans Facilitate Procathepsin B Activation through Disruption of Propeptide-Mature Enzyme Interactions." *The Journal of Biological Chemistry* 282 (45): 33076–85.

Cataldo, A. M., and R. A. Nixon. 1990. “Enzymatically Active Lysosomal Proteases Are Associated with Amyloid Deposits in Alzheimer Brain.” *Proceedings of the National Academy of Sciences of the United States of America* 87 (10): 3861–65.

Cater, S. A, W. E. Lees, J. Hill, J. Brzin, J. Kay, and L. H. Phylip. 2002. “Aspartic Proteinase Inhibitors from Tomato and Potato Are More Potent against Yeast Proteinase A than Cathepsin D.” *Biochimica et Biophysica Acta* 1596 (1): 76–82.

Chmelař, J., C. J. Oliveira, P. Řezáčová, I. M. B. Francischetti, Z. Kovářová, G. Pejler, P. Kopáček, et al. 2011. “A Tick Salivary Protein Targets Cathepsin G and Chymase and Inhibits Host Inflammation and Platelet Aggregation.” *Blood* 117 (2): 736–44.

Christeller, J. T., P. C. Farley, R. J. Ramsay, P. A. Sullivan, and W. A. Laing. 1998. “Purification, Characterization and Cloning of an Aspartic Proteinase Inhibitor from Squash Phloem Exudate.” *European Journal of Biochemistry* 254 (1): 160–67.

Coates, L., H. F. Tuan, S. Tomanicek, A. Kovalevsky, M. Mustyakimov, P. Erskine, and J. Cooper. 2008. “The Catalytic Mechanism of an Aspartic Proteinase Explored with Neutron and X-Ray Diffraction.” *Journal of the American Chemical Society* 130 (23): 7235–37.

Craik, C. S., S. Roczniak, C. Largman, and W. J. Rutter. 1987. “The Catalytic Role of the Active Site Aspartic Acid in Serine Proteases.” *Science* 237 (4817): 909–13.

Dall, E., and H. Brandstetter. 2016. “Structure and Function of Legumain in Health and Disease.” *Biochimie* 122: 126–50.

Dall, E., J. C. Fegg, P. Briza, and H. Brandstetter. 2015. “Structure and Mechanism of an Aspartimide-Dependent Peptide Ligase in Human Legumain.” *Angewandte Chemie (International Ed in English)* 54 (10): 2917–21.

Dunn, B. M., and S. Hung. 2000. “The Two Sides of Enzyme-Substrate Specificity: Lessons from the Aspartic Proteinases.” *Biochimica et Biophysica Acta* 1477 (1–2): 231–40.

Dvořák, J., and M. Horn. 2018. “Serine Proteases in Schistosomes and Other Trematodes.” *International Journal for Parasitology* 48 (5): 333–44.

Fairhead, M., S. M. Kelly, and C. F. van der Walle. 2008. “A Heparin Binding Motif on the Pro-Domain of Human Procathepsin L Mediates Zymogen Destabilization and Activation.” *Biochemical and Biophysical Research Communications* 366 (3): 862–67.

Ferreira, R., D. Zhou, J. G. Ferreira, M. C. C. Silva, R. A. Silva-Lucca, R. Mentele, E. J. Paredes-Gamero, et al. 2013. “Crystal Structure of Crataeva Tapia Bark Protein (CrataBL) and Its Effect in Human Prostate Cancer Cell Lines.” *PloS One* 8 (6): e64426.

Figura, K. von, and A. Hasilik. 1986. “Lysosomal Enzymes and Their Receptors.” *Annual Review of Biochemistry* 55: 167–93.

Fusek, M., M. Baudyš, and P. Metcalf. 1992. “Purification and Crystallization of Human Cathepsin D.” *Journal of Molecular Biology* 226 (2): 555–57.

Fusek, M., J. Větvičková, and V. Větvička. 2007. “Secretion of Cytokines in Breast Cancer Cells: The Molecular Mechanism of Procathepsin D Proliferative Effects.” *Journal of Interferon & Cytokine Research : The Official Journal of the International Society for Interferon and Cytokine Research* 27 (3): 191–99.

Ghosh, A. K, M. Brindisi, and J. Tang. 2012. “Developing Beta-Secretase Inhibitors for Treatment of Alzheimer’s Disease.” *Journal of Neurochemistry* 120 Suppl: 71–83.

Gieselmann, V., A. Hasilik, and K. von Figura. 1985. “Processing of Human Cathepsin D in Lysosomes in Vitro.” *The Journal of Biological Chemistry* 260 (5): 3215–20.

Goettig, P., V. Magdolen, and H. Brandstetter. 2010. “Natural and Synthetic Inhibitors of Kallikrein-Related Peptidases (KLKs).” *Biochimie* 92 (11): 1546–67.

Guerra, Y., P. A. Valiente, T. Pons, C. Berry, and E. Rudino-Pinera. 2016. “Structures of a Bi-Functional Kunitz-Type STI Family Inhibitor of Serine and Aspartic Proteases: Could the Aspartic Protease Inhibition Have Evolved from a Canonical Serine Protease-Binding Loop?” *Journal of Structural Biology* 195 (2): 259–71.

Guo, J., P. T. Erskine, A. R. Coker, S. P. Wood, and J. B. Cooper. 2015. “Structure of a Kunitz-Type Potato Cathepsin D Inhibitor.” *Journal of Structural Biology* 192 (3): 554–60.

Hánová, I., J. Brynda, R. Houštecká, N. Alam, D. Sojka, P. Kopáček, L. Marešová, et al. 2018. “Novel Structural Mechanism of Allosteric Regulation of Aspartic Peptidases via an Evolutionarily Conserved Exosite.” *Cell Chemical Biology* 25 (3): 318–329.e4.

Hansen, D., S. Macedo-Ribeiro, P. Verissimo, S. Y. Im, M. U. Sampaio, and M. L. V. Oliva. 2007. “Crystal Structure of a Novel Cysteinless Plant Kunitz-Type Protease Inhibitor.” *Biochemical and Biophysical Research Communications* 360 (4): 735–40.

Hashida, S., T. Towatari, E. Kominami, and N. Katunuma. 1980. “Inhibitions by E-64 Derivatives of Rat Liver Cathepsin B and Cathepsin L in Vitro and in Vivo.” *Journal of Biochemistry* 88 (6): 1805–11.

Hedstrom, L. 2002. “Serine Protease Mechanism and Specificity.” *Chemical Reviews* 102 (12): 4501–24.

Helland, R., J. Otlewski, O. Sundheim, M. Dadlez, and A. O Smalas. 1999. “The Crystal Structures of the Complexes between Bovine Beta-Trypsin and Ten P1 Variants of BPTI.” *Journal of Molecular Biology* 287 (5): 923–42.

Horigome, D., H. Satoh, N. Itoh, K. Mitsunaga, I. Oonishi, A. Nakagawa, and A. Uchida. 2007. “Structural Mechanism and Photoprotective Function of Water-Soluble Chlorophyll-Binding Protein.” *The Journal of Biological Chemistry* 282 (9): 6525–31.

Horn, M., A. Jílková, J. Vondrášek, L. Marešová, C. R. Caffrey, and M. Mareš. 2011. “Mapping the Pro-Peptide of the Schistosoma Mansoni Cathepsin B1 Drug Target: Modulation of Inhibition by Heparin and Design of Mimetic Inhibitors.” *ACS Chemical Biology* 6 (6): 609–17.

Horn, M., M. Nussbaumerová, M. Šanda, Z. Kovářová, J. Srba, Z. Franta, D. Sojka, et al. 2009. “Hemoglobin Digestion in Blood-Feeding Ticks: Mapping a Multipeptidase Pathway by Functional Proteomics.” *Cell Chemical Biology* 16 (10): 1053–63.

Horn, M., O. Zbodaková, P. Kašpárek, J. Srp, R. Hanečkova, M. Hradilek, M. Mareš, and R. Sedláček. 2018. “Profiling System for Skin Kallikrein Proteolysis Applied in Gene-Deficient Mouse Models.” *Biological Chemistry*. 399 (9): 1085–89

Hozumi, M., M. Ogawa, T. Sugimura, T. Takeuchi, and H. Umezawa. 1972. “Inhibition of Tumorigenesis in Mouse Skin by Leupeptin, a Protease Inhibitor from Actinomycetes.” *Cancer Research* 32 (8): 1725–28.

Huntington, J. A., R. J. Read, and R. W. Carrell. 2000. “Structure of a Serpin-Protease Complex Shows Inhibition by Deformation.” *Nature* 407 (6806): 923–26.

Huntington, J. A.. 2012. “Thrombin Plasticity.” *Biochimica et Biophysica Acta* 1824 (1): 246–52.

Huntington, J. A.. 2014. “Natural Inhibitors of Thrombin.” *Thrombosis and Haemostasis* 111 (4): 583–89.

Jelinek, B., J. Antal, I. Venekei, and L. Graf. 2004. “Ala226 to Gly and Ser189 to Asp Mutations Convert Rat Chymotrypsin B to a Trypsin-like Protease.” *Protein Engineering, Design and Selection* 17 (2): 127–31.

Jílková, A., P. Řezáčová, M. Lepšík, M. Horn, J. Váchová, J. Fanfrlík, J. Brynda, J. H. McKerrow, C. R. Caffrey, and M. Mareš. 2011. "Structural Basis for Inhibition of Cathepsin B Drug Target from the Human Blood Fluke, Schistosoma Mansoni." *The Journal of Biological Chemistry* 286 (41): 35770–81.

Kadek, A., V.Tretyachenko, H. Mrázek, Ljubina Ivanova, P. Halada, M. Rey, D. C. Schriemer, and P. Man. 2014. "Expression and Characterization of Plant Aspartic Protease Nepenthesin-1 from Nepenthes Gracilis." *Protein Expression and Purification* 95: 121–28.

Kalinska, M. , U. Meyer-Hoffert, T. Kantyka, and J. Potempa. 2016. "Kallikreins - The Melting Pot of Activity and Function." *Biochimie* 122: 270–82.

Keilová, H., V. Tomašek. 1972. "Effect of Pepsin Inhibitor from Ascaris Lumbricoides on Cathepsin D and E." *Biochimica et Biophysica Acta* 284 (2): 461–64.

Kitamura, K., C. Yoshida, Y. Kinoshita, T. Kadowaki, Y. Takahashi, T. Tayama, T. Kawakubo, et al. 2009. "Development of Systemic in Vitro Evolution and Its Application to Generation of Peptide-Aptamer-Based Inhibitors of Cathepsin E." *Journal of Molecular Biology* 387 (5): 1186–98.

Knight, C. G., and A. J. Barrett. 1976. "Interaction of Human Cathepsin D with the Inhibitor Pepstatin." *The Biochemical Journal* 155 (1): 117–25.

Koizumi, A., A. Tsuchiya, K. Nakajima, K. Ito, T. Terada, A. Shimizu-Ibuka, L. Briand, T. Asakura, T. Misaka, and K. Abe. 2011. "Human Sweet Taste Receptor Mediates Acid-Induced Sweetness of Miraculin." *Proceedings of the National Academy of Sciences of the United States of America* 108 (40): 16819–24.

Krishnan, N., D. Kodrik, F. Turanli, and F. Sehnal. 2007. "Stage-Specific Distribution of Oxidative Radicals and Antioxidant Enzymes in the Midgut of *Leptinotarsa Decemlineata.*" *Journal of Insect Physiology* 53 (1): 67–74.

Krowarsch, D., T. Cierpicki, F. Jelen, and J. Otlewski. 2003. "Canonical Protein Inhibitors of Serine Proteases." *Cellular and Molecular Life Sciences : CMLS* 60 (11): 2427–44.

Krowarsch, D., M. Zakrzewska, A. O. Smalas, and J. Otlewski. 2005. "Structure-Function Relationships in Serine Protease-Bovine Pancreatic Trypsin Inhibitor Interaction." *Protein and Peptide Letters* 12 (5): 403–7.

Kryza, T., M. L. Silva, D. Loessner, N. Heuze-Vourc'h, and J. A. Clements. 2016. "The Kallikrein-Related Peptidase Family: Dysregulation and Functions during Cancer Progression." *Biochimie* 122: 283–99.

Law, R. H. P., Q. Zhang, S. McGowan, A. M. Buckle, G. A. Silverman, W. Wong, C. J. Rosado, et al. 2006. "An Overview of the Serpin Superfamily." *Genome Biology* 7 (5): 216.

Lenarcic, B., and V. Turk. 1999. "Thyroglobulin Type-1 Domains in Equistatin Inhibit Both Papain-like Cysteine Proteinases and Cathepsin D." *The Journal of Biological Chemistry* 274 (2): 563–66.

Li, M., L. H. Phylip, W. E. Lees, J. R. Winther, B. M. Dunn, A. Wlodawer, J. Kay, and A. Gustchina. 2000. "The Aspartic Proteinase from Saccharomyces Cerevisiae Folds Its Own Inhibitor into a Helix." *Nature Structural Biology* 7 (2): 113–17.

Luckett, S., R. S. Garcia, J. J. Barker, A. V. Konarev, P. R. Shewry, A. R. Clarke, and R. L. Brady. 1999. "High-Resolution Structure of a Potent, Cyclic Proteinase Inhibitor from Sunflower Seeds." *Journal of Molecular Biology* 290 (2): 525–33.

Luo, L., and W. Jiang. 2006. "Inhibition Profiles of Human Tissue Kallikreins by Serine Protease Inhibitors." *Biological Chemistry* 387 (6): 813–16.

Majer, P., J. R. Collins, S. V. Gulnik, and J. W. Erickson. 1997. “Structure-Based Subsite Specificity Mapping of Human Cathepsin D Using Statine-Based Inhibitors.” *Protein Science* 6 (7): 1458–66.

Marco, M., and J. M. Coteron. 2012. “Falcipain Inhibition as a Promising Antimalarial Target.” *Current Topics in Medicinal Chemistry* 12 (5): 408–44.

Mareš, M., B. Meloun, M. Pavlík, V. Kostka, and M. Baudyš. 1989. “Primary Structure of Cathepsin D Inhibitor from Potatoes and Its Structure Relationship to Soybean Trypsin Inhibitor Family.” *FEBS Letters* 251 (1–2): 94–98.

Máša, M., L. Marešova, J. Vondrášek, M. Horn, J. Ježek, and M. Mareš. 2006. “Cathepsin D Propeptide: Mechanism and Regulation of Its Interaction with the Catalytic Core.” *Biochemistry* 45 (51): 15474–82.

Masson, O., A. Bach, D. Derocq, C. Prebois, V. Laurent-Matha, S. Pattingre, and E. Liaudet-Coopman. 2010. “Pathophysiological Functions of Cathepsin D: Targeting Its Catalytic Activity versus Its Protein Binding Activity?” *Biochimie* 92 (11): 1635–43.

Masurier, N., D. P. Arama, C. El Amri, and V. Lisowski. 2018. “Inhibitors of Kallikrein-Related Peptidases: An Overview.” *Medicinal Research Reviews* 38 (2): 655–83.

McGrath, M. E., J. R. Vasquez, C. S. Craik, A. S. Yang, B. Honig, and R. J. Fletterick. 1992. “Perturbing the Polar Environment of Asp102 in Trypsin: Consequences of Replacing Conserved Ser214.” *Biochemistry* 31 (12): 3059–64.

McKerrow, J. H., P. S. Doyle, J. C. Engel, L. M. Podust, S. A. Robertson, R. Ferreira, T. Saxton, et al. 2009. “Two Approaches to Discovering and Developing New Drugs for Chagas Disease.” *Memorias Do Instituto Oswaldo Cruz* 104 Suppl: 263–69.

Mehta, A. Y., Y. Jin, and U. R. Desai. 2014. “An Update on Recent Patents on Thrombin Inhibitors (2010 - 2013).” *Expert Opinion on Therapeutic Patents* 24 (1): 47–67.

Meulenbroek, E. M., E. A. Thomassen, L. Pouvreau, J. P. Abrahams, H. Gruppen, and N. S. Pannu. 2012. “Structure of a Post-Translationally Processed Heterodimeric Double-Headed Kunitz-Type Serine Protease Inhibitor from Potato.” *Acta Crystallographica Section D* 68 (Pt 7): 794–99.

Micheelsen, P. O., J. Vevodova, L. De Maria, P. R. Ostergaard, E. P. Friis, K. Wilson, and M. Skjot. 2008. “Structural and Mutational Analyses of the Interaction between the Barley Alpha-Amylase/Subtilisin Inhibitor and the Subtilisin Savinase Reveal a Novel Mode of Inhibition.” *Journal of Molecular Biology* 380 (4): 681–90.

Mihelic, M., and D. Turk. 2007. “Two Decades of Thyroglobulin Type-1 Domain Research.” *Biological Chemistry* 388 (11): 1123–30.

Motwani, H. V., M. De Rosa, L. R. Odell, A. Hallberg, and M. Larhed. 2015. “Aspartic Protease Inhibitors Containing Tertiary Alcohol Transition-State Mimics.” *European Journal of Medicinal Chemistry* 90: 462–90.

Nadeem, S., and D. L. Batisky. 2014. “Aliskiren, the First Direct Renin Inhibitor: Assessing a Role in Pediatric Hypertension and Kidney Diseases.” *Pediatric Nephrology (Berlin, Germany)* 29 (11): 2105–11.

Ng, K. K., J. F. Petersen, M. M. Cherney, C. Garen, J. J. Zalatoris, C. Rao-Naik, B. M. Dunn, M. R. Martzen, R. J. Peanasky, and M. N. James. 2000. “Structural Basis for the Inhibition of Porcine Pepsin by Ascaris Pepsin Inhibitor-3.” *Nature Structural Biology* 7 (8): 653–57.

Novinec, M., and B. Lenarcic. 2013. “Papain-like Peptidases: Structure, Function, and Evolution.” *Biomolecular Concepts* 4 (3): 287–308.

Nussbaumerová, M., J. Srp, M. Máša, M. Hradilek, M. Šanda, M. Reiniš, M. Horn, and M. Mareš.

2010. “Single- and Double-Headed Chemical Probes for Detection of Active Cathepsin D in a Cancer Cell Proteome.” *Chembiochem : A European Journal of Chemical Biology* 11 (11): 1538–41.

Oliva, M. L. V., M. C. C. Silva, R. C. Sallai, M. V. Brito, and M. U. Sampaio. 2010. “A Novel Subclassification for Kunitz Proteinase Inhibitors from Leguminous Seeds.” *Biochimie* 92 (11): 1667–73.

Oliva, M. L. V., and U. M. Sampaio. 2008. “Bauhinia Kunitz-Type Proteinase Inhibitors: Structural Characteristics and Biological Properties.” *Biological Chemistry* 389 (8): 1007–13.

Padilha, M. H., A. C. Pimentel, A. F. Ribeiro, and W. R. Terra. 2009. “Sequence and Function of Lysosomal and Digestive Cathepsin D-like Proteinases of Musca Domestica Midgut.” *Insect Biochemistry and Molecular Biology* 39 (11): 782–91.

Patil, D. N., A. Chaudhary, A. K. Sharma, S. Tomar, and P. Kumar. 2012. “Structural Basis for Dual Inhibitory Role of Tamarind Kunitz Inhibitor (TKI) against Factor Xa and Trypsin.” *The FEBS Journal* 279 (24): 4547–64.

Pearson, M. S., L. Becker, P. Driguez, N. D. Young, S. Gaze, T. Mendes, X. Li, et al. 2015. “Of Monkeys and Men: Immunomic Profiling of Sera from Humans and Non-Human Primates Resistant to Schistosomiasis Reveals Novel Potential Vaccine Candidates.” *Frontiers in Immunology* 6: 213.

Perona, J. J., and C. S. Craik. 1995. “Structural Basis of Substrate Specificity in the Serine Proteases.” *Protein Science* 4 (3): 337–60.

Pokorná, J., L. Machala, P. Řezáčová, and J. Konvalinka. 2009. “Current and Novel Inhibitors of HIV Protease.” *Viruses* 1 (3): 1209–39.

Ponnusamy, S., M. Meyers-Needham, C. E. Senkal, S. A. Saddoughi, D. Sentelle, S. P. Selvam, A. Salas, and B. Ogretmen. 2010. “Sphingolipids and Cancer: Ceramide and Sphingosine-1-Phosphate in the Regulation of Cell Death and Drug Resistance.” *Future Oncology* 6 (10): 1603–24.

Powers, J. C., J. L. Asgian, O. D. Ekici, and K. E. James. 2002. “Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases.” *Chemical Reviews* 102 (12): 4639–4750.

Rawlings, N. D., and A. J. Barrett. 1993. “Evolutionary Families of Peptidases.” *Biochemical Journal* 290 (Pt 1: 205–18.

Rawlings, N. D., and A. J. Barrett. 2010. “MEROPS: The Peptidase Database.” *Nucleic Acids Research* 38 (1): 325–31.

Rawlings, N. D., and G. Salvesen. 2013. *Handbook of Proteolytic Enzymes*. *Elsevier*.

Renko, M., J. Sabotic, and D. Turk. 2012. “Beta-Trefoil Inhibitors--from the Work of Kunitz Onward.” *Biological Chemistry* 393 (10): 1043–54.

Ritonja, A., I. Krizaj, P. Mesko, M. Kopitar, P. Lucovnik, B. Strukelj, J. Pungercar, D. J. Buttle, A. J. Barrett, and V. Turk. 1990. “The Amino Acid Sequence of a Novel Inhibitor of Cathepsin D from Potato.” *FEBS Letters* 267 (1): 13–15.

Sadik, G., H. Kaji, K. Takeda, F. Yamagata, Y. Kameoka, K. Hashimoto, K. Miyanaga, and T. Shinoda. 1999. “In Vitro Processing of Amyloid Precursor Protein by Cathepsin D.” *The International Journal of Biochemistry & Cell Biology* 31 (11): 1327–37.

Sainsbury, F., A. Rheaume, M. Goulet, J. Vorster, and D. Michaud. 2012. “Discrimination of Differentially Inhibited Cysteine Proteases by Activity-Based Profiling Using Cystatin Variants with Tailored Specificities.” *Journal of Proteome Research* 11 (12): 5983–93.

Sanman, L. E., and M. Bogyo. 2014. “Activity-Based Profiling of Proteases.” *Annual Review of Biochemistry* 83: 249–73.

Santamaria, M. E., P. Hernandez-Crespo, F. Ortego, V. Grbic, M. Grbic, I. Diaz, and M. Martinez. 2012. “Cysteine Peptidases and Their Inhibitors in Tetranychus Urticae: A Comparative Genomic Approach.” *BMC Genomics* 13: 307.

Skala, W., D. T. Utzschneider, V. Magdolen, M. Debela, S. Guo, C. S. Craik, H. Brandstetter, and P. Goettig. 2014. “Structure-Function Analyses of Human Kallikrein-Related Peptidase 2 Establish the 99-Loop as Master Regulator of Activity.” *The Journal of Biological Chemistry* 289 (49): 34267–83.

Sojka, D., Z. Franta, H. Frantová, P. Bartošová, M. Horn, J. Váchová, A. J. O’Donoghue, et al. 2012. “Characterization of Gut-Associated Cathepsin D Hemoglobinase from Tick Ixodes Ricinus (IrCD1).” *The Journal of Biological Chemistry* 287 (25): 21152–63.

Sojka, D., D. Hartmann, P. Bartošová-Sojková, and J. Dvořák. 2016. “Parasite Cathepsin D-Like Peptidases and Their Relevance as Therapeutic Targets.” *Trends in Parasitology* 32 (9): 708–23.

Song, H. K., and S. W. Suh. 1998. “Kunitz-Type Soybean Trypsin Inhibitor Revisited: Refined Structure of Its Complex with Porcine Trypsin Reveals an Insight into the Interaction between a Homologous Inhibitor from Erythrina Caffra and Tissue-Type Plasminogen Activator.” *Journal of Molecular Biology* 275 (2): 347–63.

Sparks, M. A., S. D. Crowley, S. B. Gurley, M. Mirotsou, and T. M. Coffman. 2014. “Classical Renin-Angiotensin System in Kidney Physiology.” *Comprehensive Physiology* 4 (3): 1201–28.

Strooper, B. De, R. Vassar, and T. Golde. 2010. “The Secretases: Enzymes with Therapeutic Potential in Alzheimer Disease.” *Nature Reviews Neurology* 6 (2): 99–107.

Sun, H., X. Lou, Q. Shan, J. Zhang, X. Zhu, J. Zhang, Y. Wang, Y. Xie, N. Xu, and S. Liu. 2013. “Proteolytic Characteristics of Cathepsin D Related to the Recognition and Cleavage of Its Target Proteins.” *PloS One* 8 (6): e65733.

Swedberg, J. E., S. J. de Veer, K. C. Sit, C. F. Reboul, A. M. Buckle, and J. M. Harris. 2011. “Mastering the Canonical Loop of Serine Protease Inhibitors: Enhancing Potency by Optimising the Internal Hydrogen Bond Network.” *PloS One* 6 (4): e19302.

Tallima, H., J. Dvorak, S. Kareem, M. A. El Dahab, N. A. Aziz, J. P. Dalton, and R. El Ridi. 2017. “Protective Immune Responses against Schistosoma Mansoni Infection by Immunization with Functionally Active Gut-Derived Cysteine Peptidases Alone and in Combination with Glyceraldehyde 3-Phosphate Dehydrogenase.” *PLoS Neglected Tropical Diseases* 11 (3): e0005443.

Tarasuk, M., S. V. Grams, V. Viyanant, and R. Grams. 2009. “Type I Cystatin (Stefin) Is a Major Component of Fasciola Gigantica Excretion/Secretion Product.” *Molecular and Biochemical Parasitology* 167 (1): 60–71.

Terra, W. R., and C. Ferreira. 2012. “Biochemistry and Molecular Biology of Digestion.” In *Insect Molecular Biology and Biochemistry*, edited by Lawrence I Gilbert, 1st ed., 365–418. Elsevier B.V.

Vallee, F., A. Kadziola, Y. Bourne, M. Juy, K. W. Rodenburg, B. Svensson, and R. Haser. 1998. “Barley Alpha-Amylase Bound to Its Endogenous Protein Inhibitor BASI: Crystal Structure of the Complex at 1.9 A Resolution.” *Structure* 6 (5): 649–59.

Veer, S. J. de, J. E. Swedberg, M. Akcan, K. J. Rosengren, M. Brattsand, D. J. Craik, and J. M. Harris. 2015. “Engineered Protease Inhibitors Based on Sunflower Trypsin Inhibitor-1 (SFTI-1) Provide Insights into the Role of Sequence and Conformation in Laskowski Mechanism Inhibition.” *The Biochemical Journal* 469 (2): 243–53.

Venekei, I., L. Graf, and W. J. Rutter. 1996. “Expression of Rat Chymotrypsinogen in Yeast: A Study on the Structural and Functional Significance of the Chymotrypsinogen Propeptide.” *FEBS Letters* 379 (2): 139–42.

Větvička, V., J. Větvičková, I. Hilgert, Z. Voburka, and M. Fusek. 1997. “Analysis of the Interaction of Procathepsin D Activation Peptide with Breast Cancer Cells.” *International Journal of Cancer* 73 (3): 403–9.

Volloch, V., and S. Rits. 2018. “Results of Beta Secretase-Inhibitor Clinical Trials Support Amyloid Precursor Protein-Independent Generation of Beta Amyloid in Sporadic Alzheimer’s Disease.” *Medical Sciences* 6 (2).

Wang, D., W. Bode, and R. Huber. 1985. “Bovine Chymotrypsinogen A X-Ray Crystal Structure Analysis and Refinement of a New Crystal Form at 1.8 A Resolution.” *Journal of Molecular Biology* 185 (3): 595–624.

Wiederanders, B., G. Kaulmann, and K. Schilling. 2003. “Functions of Propeptide Parts in Cysteine Proteases.” *Current Protein & Peptide Science* 4 (5): 309–26.

Wielkopolan, B., and A. Obrepalska-Steplowska. 2016. “Three-Way Interaction among Plants, Bacteria, and Coleopteran Insects.” *Planta* 244 (2): 313–32.

Wolfson, J. L., and L. L. Murdock. 1987. “Suppression of Larval Colorado Potato Beetle Growth and Development by Digestive Proteinase-Inhibitors.” *Entomologia Experimentalis Et Applicata* 44 (3): 235–40.

Xi, Z., J. Kaern, B. Davidson, T. I. Klokk, B. Risberg, C. Trope, and F. Saatcioglu. 2004. “Kallikrein 4 Is Associated with Paclitaxel Resistance in Ovarian Cancer.” *Gynecologic Oncology* 94 (1): 80–85.

Xu, P., P. A. Andreasen, and M. Huang. 2017. “Structural Principles in the Development of Cyclic Peptidic Enzyme Inhibitors.” *International Journal of Biological Sciences* 13 (10): 1222–33.

Yang, Y., Y. Wen, Y. N. Cai, I. Vallee, P. Boireau, M. Y. Liu, and S. P. Cheng. 2015. “Serine Proteases of Parasitic Helminths.” *The Korean Journal of Parasitology* 53 (1): 1–11.

Yousef, G. M., and E. P. Diamandis. 2001. “The New Human Tissue Kallikrein Gene Family: Structure, Function, and Association to Disease.” *Endocrine Reviews* 22 (2): 184–204.

Zani, M., and T. Moreau. 2010. “Phage Display as a Powerful Tool to Engineer Protease Inhibitors.” *Biochimie* 92 (11): 1689–1704.

Zauner, F. B., E. Dall, C. Regl, L. Grassi, C. G. Huber, C. Cabrele, and H. Brandstetter. 2018. “Crystal Structure of Plant Legumain Reveals a Unique Two-Chain State with PH-Dependent Activity Regulation.” *The Plant Cell* 30 (3): 686–99.

Zhou, D., D. Hansen, I. G. Shabalin, A. Gustchina, D. F. Vieira, M. V. de Brito, A. P. Araujo, M. L. V. Oliva, and A. Wlodawer. 2015. “Structure of BbKI, a Disulfide-Free Plasma Kallikrein Inhibitor.” *Acta Crystallographica Section F Structural Biology Communications* 71 (Pt 8): 1055–62.

Zhou, D., Y. A. Lobo, I. F. C. Batista, R. Marques-Porto, A. Gustchina, M. L. V. Oliva, and A. Wlodawer. 2013. “Crystal Structures of a Plant Trypsin Inhibitor from Enterolobium Contortisiliquum (EcTI) and of Its Complex with Bovine Trypsin.” *PloS One* 8 (4): e62252.

Zhu-Salzman, K., and R. Zeng. 2015. “Insect Response to Plant Defensive Protease Inhibitors.” *Annual Review of Entomology* 60: 233–52.