

Univerzita Karlova

1. lékařská fakulta

Studijní program: Biomedicína

Studijní obor: Biochemie a patobiochemie



UNIVERZITA KARLOVA
1. lékařská fakulta

Mgr. Jakub Benýšek

Regulace aktivity katepsinu K pomocí reaktivních inhibičních molekul využitelných
v biomedicině

Regulation of cathepsin K activity by reactive inhibitory molecules applicable in
biomedicine

Disertační práce

Školitel: RNDr. Michael Mareš, CSc.

Praha, 2022

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem řádně uvedl a citoval všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

V Praze, 1. 1. 2022

Jakub Benýšek

Identifikační záznam:

BENÝŠEK, Jakub, Regulace aktivity katepsinu K pomocí reaktivních inhibičních molekul využitelných v biomedicíně [Regulation of cathepsin K activity by reactive inhibitory molecules applicable in biomedicine]. Praha 2022. 162 stran, 3 přílohy. Disertační práce, Univerzita Karlova, 1. lékařská fakulta, Ústav organické chemie a biochemie AV ČR v. v. i. Vedoucí práce: RNDr. Michael Mareš, CSc

Obsah:

SEZNAM ZKRATEK.....	1
ABSTRAKT	2
ABSTRACT	3
1. ÚVOD	4
2. LITERÁRNÍ ÚVOD A PŘEHLED DANÉ PROBLEMATIKY	5
2.1 Proteolytické enzymy	5
2.1.1 Klasifikace proteas.....	6
2.1.2 Mechanismus katalytického působení cysteinových proteas.....	7
2.2 Lidské cysteinové proteasy	8
2.2.1 Kaspasy (klan CD, rodina C14).....	9
2.2.2 Legumainy a GPI-transamidasa (klan CD, rodina C13).....	10
2.2.3 Kalpainy (klan CA, rodina C2).....	11
2.2.4 Cysteinové katepsiny (klan CA, rodina C1)	12
2.3 Aktivita a struktura katepsinu K	14
2.3.1 Fyziologická úloha katepsinu K v lidském organismu	16
2.3.2 Regulace tvorby a resorpce kostní tkáně	17
2.3.3 Katepsin K a jeho role v patologických procesech.....	19
2.4 Regulace aktivity katepsinu K	23
2.4.1 Přírodní inhibitory a modulátory katepsinu K.....	25
2.4.2 Syntetické inhibitory katepsinu K.....	27
2.4.2.1 Ireverzibilní kovalentní syntetické inhibitory	28
2.4.2.2 Reverzibilní kovalentní syntetické inhibitory	32
2.4.3 Nekovalentní inhibitory katepsinu K syntetického a přírodního původu	35
3. CÍLE PRÁCE	36
4. MATERIÁL A METODY	37
4.1 Materiál a laboratorní výbava	37
4.2 Metody	38
4.2.1 Příprava rekombinantního katepsinu K	38
4.2.2 Biochemické metody	38
4.2.3 Enzymologické metody	39
4.2.4 Metody rentgenové krystalografické analýzy.....	39
5. VÝSLEDKY	41

5.1 Přehled výstupů dizertační práce	41
5.2 Příprava rekombinantního katepsinu K.....	42
5.3 Interakce katepsinu K s 1,2,4-thiadiazolovými inhibitory	43
5.3.1 Inhibiční vlastnosti sloučenin s 1,2,4-thiadiazolovým cyklem.....	43
5.4 Interakce katepsinu K s vinylketonovými inhibitory	46
5.4.1 Selektivita vinylketonových inhibitorů katepsinu K	46
5.4.2 Krystalografická analýza komplexů vinylketonových inhibitorů a určení vazebného módu	48
5.5 Interakce katepsinu K s cyanohydrazidovými inhibitory.....	54
5.5.1 Cyanohydrazidy jako účinné a selektivní inhibitory aktivity KatK a procesu aktivace zymogenu KatK	54
5.5.2 Krystalografická analýza komplexů cyanohydrazidových inhibitorů s katepsinem K	57
5.5.3 Analýza vazebného módu cyanhydrazidových inhibitorů s maturní formou katepsinu K.....	59
5.5.3.1 Interakce Gü1303 v aktivním místě maturního katepsinu K.....	60
5.5.3.2 Interakce Gü2602 v aktivním místě maturního katepsinu K.....	60
5.5.3.3 Interakce Gü1303 v aktivním místě aktivačního meziproduktu katepsinu K	63
5.5.3.4 Interakce Gü2602 v aktivním místě aktivačního meziproduktu katepsinu K	63
5.5.4 Analýza konformační flexibility inhibitorů Gü2602 a Gü1303 v aktivním místě. 65	
6. DISKUSE.....	68
7. ZÁVĚR	75
8. LITERÁRNÍ ZDROJE	76

Seznam zkratek

Å	Ångström, 10^{-10} m
AMC	7-amino-4-mehtylkumarin
BMM	indukční medium (z angl. „Buffered Minimal Methanol“)
cDNA	komplementární DNA, která vznikne přepisem z mRNA
DMSO	dimethylsulfoxid
DTT	dithiothreitol
Cbz	benzyloxykarbonyl
EDTA	kyselina ethylendiamintetraoctová
IC ₅₀	koncentrace inhibitoru potřebná k dosažení 50 % inhibice aktivity enzymu
KatK	katepsin K
iKatK	aktivační intermediát (meziprodukt) katepsinu K
mKatK	maturní forma katepsinu K
pKatK	proenzym katepsinu K
kDa	kilodalton
M-CSF	růstový hormon pro makrofágy (z angl. „Macrophage Colony Stimulating Factor“)
MHC	hlavní histokompatibilitní komplex (z angl. „Major Histocompatibility Complex“)
mRNA	mediatorová ribonukleová kyselina (z angl. „messenger RNA“)
MS	hmotnostní spektrometrie (z angl. „Mass Spectrometry“)
NMR	nukleární magnetická rezonance
PDB	databáze proteinových struktur (z angl. „Protein Data Bank“)
RANK	receptor prekurzorů osteoklastů (z angl. „Receptor Activator of Nuclear Factor $\kappa\beta$ “)
RANKL	substrát pro RANK receptor (z angl. „Receptor Activator of Nuclear Factor $\kappa\beta$ Ligand“)
RMSD	střední kvadratická odchylka (z angl. „Root Mean Square Deviation“)
RP-HPLC	vysokotlaká chromatografie na reverzní fázi (z angl. „Reverse Phase High Performace Liquid Chromatography“)
SDS-PAGE	elektroforéza na polyakrylamidovém gelu v přítomnosti dodecylsíranu sodného
YPD	kultivační medium pro kvasinky (z angl. „Yeast extract, Peptone, Dextrose“)

Abstrakt

Lidský katepsin K (KatK) je lysozomální cysteinová proteasa exprimovaná především v osteoklastech. Fyziologická funkce spočívá v degradaci extracelulární kostní matrix, kde je neúčinnějším enzymem pro štěpení kolagenu. Zvýšená enzymatická aktivita KatK je spojena s osteoporózou a dále revmatoidní artritidou a osteoartritidou. Proto je KatK cílovou molekulou pro léčbu uvedených patologií a k jeho regulaci jsou vyvíjena chemoterapeutika na bázi proteasových inhibitorů. Tato práce se zabývá reaktivními peptidomimetickými a nízkomolekulárními inhibitory KatK typu thiadiazolů, vinylketonů a cyanohydrazidů. Soustředí se zejména na určení vazebného módu selektivních inhibitorů a charakterizaci jejich interakcí s aktivním místem KatK. Přístupy rentgenostrukturní analýzy jsou kombinovány s výpočetní chemií, enzymologickou analýzou a buněčnými testy k objasnění vztahu mezi strukturou a biochemickou aktivitou těchto inhibitorů. Získané výsledky přinášejí významné informace pro navrhování a optimalizaci nových vysoce účinných a selektivních inhibitorů KatK určených pro vývoj potenciálních léčiv a diagnostických sond.

Klíčová slova: katepsin K, proteasa, inhibitor, 3D struktura, osteoporóza

Abstract

Human cathepsin K (KatK) is a lysosomal cysteine protease expressed predominantly in osteoclasts. It is the most effective enzyme for collagen breakdown and its physiological function lies in the degradation of the extracellular bone matrix. Increased enzymatic activity of KatK is associated with osteoporosis, rheumatoid arthritis, and osteoarthritis. This makes KatK a target for the treatment of pathologies, and chemotherapeutics based on protease inhibitors are being developed for its regulation. This work deals with reactive peptidomimetic and low molecular weight inhibitors of KatK namely thiadiazoles, vinyl ketones, and cyanohydrazides. It focuses mainly on the determination of the binding mode of selective inhibitors and characterization of their key interactions with the active site of KatK. Crystallographic structural analysis was combined with approaches of computational chemistry, enzymological analysis, and cell-based assays to elucidate the relationship between structure and biochemical activity of the investigated inhibitors. The obtained results provide important information for the design and optimization of new highly effective and selective KatK inhibitors as potential drugs and diagnostic probes.

Keywords: cathepsin K, protease, inhibitor, 3D structure, osteoporosis

1. Úvod

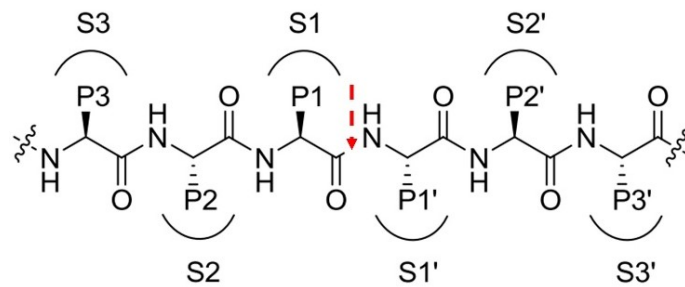
Katepsin K je lysozomální proteasa z rodiny papainu, která se řadí do třídy cysteinových proteas. Jedná se o enzym exprimovaný především v osteoklastických buňkách a je klíčový při procesu resorpce kostní tkáně a remodelace složek extracelulární matrix. Katepsin K je schopen jako jediná savčí proteasa degradovat kolagen typu I, který je hlavním strukturálním elementem kostní tkáně. Dysbalance enzymové aktivity této proteázy vede k celé řadě patologií. Příkladem je rozvoj nádorových onemocnění, zánět či neadekvátní imunitní odpověď. Proteolytickou degradací extracelulární matrix přispívá k rozvoji osteoporózy, která je jedním z nejčastějších onemocnění pohybového aparátu. Osteoporóza se vyznačuje řídnutím a zvýšenou křehkostí kostní tkáně a postihuje převážně ženy v postmenopauzálním věku. Dále katepsin K přispívá k rozvoji fibrózy srdeční a plicní tkáně či zánětlivých onemocnění osteoartrózy a revmatoidní artritidy. Zvýšená hladina exprese katepsinu K byla též prokázána u celé řady nádorových onemocnění, jakými jsou například rakovina prsu, plic, prostaty a melanomu, kde podporuje migraci nádorových buněk s následnou tvorbou metastáz do kostní tkáně. Z těchto důvodů je katepsin K atraktivní cílovou molekulou pro vývoj nových chemoterapeutik, která umožní snížení jeho enzymové aktivity. V současné době není komerčně dostupný žádný terapeutický prostředek tohoto typu a vyvíjené látky několika farmaceutických firem neprošly úspěšně klinickými studiemi. Disertační práce je zaměřena na katepsin K a regulaci jeho enzymové aktivity s využitím nových specifických peptidomimetických a nízkomolekulárních inhibitorů a dále na fluorescenčně značené aktivní sondy, určené pro diagnostiku patologií spojených s katepsinem K.

2. Literární úvod a přehled dané problematiky

2.1 Proteolytické enzymy

Proteasy (peptidasy) jsou definovány jako proteolytické enzymy, patřící do skupiny hydrolas (EC 3.4), které katalyzují štěpení proteinů nebo peptidů hydrolýzou peptidové vazby. Dle katalytického mechanismu jsou proteasy rozděleny na tyto čtyři hlavní třídy: serinové, cysteinové, aspartátové, metaloproteasy a minoritně zastoupené: asparaginové, threoninové a glutamátové proteasy (Rawlings 2020). V případě serinových proteas funguje hydroxylová skupina v postranním řetězci zbytku Ser v aktivním místě enzymu jako nukleofilní činidlo při hydrolýze peptidové vazby. Stejnou funkci zastává i thiolová skupina katalytického cysteinu v aktivním místě cysteinových proteas. Termín hydrolýza naznačuje fakt, že při štěpení peptidového řetězce je využita molekula vody. Ta je nukleofilním činidlem v případě aspartátových proteas, kdy karboxylová skupina asparagové kyseliny molekulu vody aktivuje a stabilizuje ji v aktivním místě. U metaloproteas je v aktivním místě enzymu iont kovu (Bond 2019).

Při rozpoznání a štěpení substrátu proteasou je aminokyselinová sekvence substrátu pojmenována podle modelu zavedeném Schechterem a Bergerem (Schechter and Berger 1967). Aminokyseliny jsou zde označeny jako P1 až P3 směrem od štěpné vazby k N-konci substrátu a P1' až P3' směrem k C-konci substrátu. Aminokyselina, která poskytuje svoji karboxylovou skupinu peptidové vazbě, má označení P1 a aminokyselina, která poskytuje aminoskupinu, nese označení P1'. Obdobným způsobem jsou pojmenována i vazebná podmísta pro tyto aminokyselinové zbytky na samotné protease jako S1 až S3 a S1' až S3' (Obr. 1, str. 6). U konkrétních proteas je počet funkčních vazebných podmíst (a rozpoznávaných P pozic substrátu) odlišný. Preference vazebných podmíst na protease vůči aminokyselinovým zbytkům substrátu se značně liší a tudíž stejný substrát může být štěpen různými proteasami za vzniku různých štěpných produktů.



Obrázek 1: *Pozice substrátu a vazebných podmíst v aktivním místě enzymu dle modelu Schechtera a Bergera. Jednotlivé pozice aminokyselin jsou značeny P1 až P3 směrem od štěpené vazby k N-konci substrátu a P1' až P3' směrem k C-konci substrátu. Odpovídající vazebná podmísta proteasy pro tyto aminokyselinové zbytky jsou značena S1 až S3 a S1' až S3'. Ke štěpení peptidové vazby substrátu (vyznačené červenou šipkou) dochází mezi aminokyselinami v pozicích P1 a P1'. Převzato a upraveno dle (Ali et al. 2010).*

Proteasy byly zpočátku považovány za pouhé nástroje buněčného katabolismu bílkovin na základní stavební bloky - aminokyseliny. S odstupem času se však nahromadilo velké množství důkazů, že proteasy hrají řadu dalších velmi významných fyziologických úloh. Řízená proteolýza je nezbytná pro štěpení větších prekurzorových bílkovin za uvolnění biologicky aktivních fragmentů (např. peptidových hormonů), při maturaci enzymů a buněčných receptorů, které jsou syntetizovány jako prekurzory (Kryczka and Boncela 2017). Proteolýzou jsou regulovány procesy jako je buněčný cyklus (degradací cyklinů), regulovány signalizační dráhy či komplexní děje při krevním srážení a imunitní odpovědi (López-Otín and Bond 2008). Dále se řízená proteolýza uplatňuje u apoptózy, autofagie či vnitrobuněčného transportu (odstraněním signální sekvence), Proteolýzu využívají i některé viry, bakterie a paraziti k průniku do hostitelského organismu či jeho buněk a k modulaci imunitní odpovědi hostitele (Rawlings 2013).

2.1.1 Klasifikace proteas

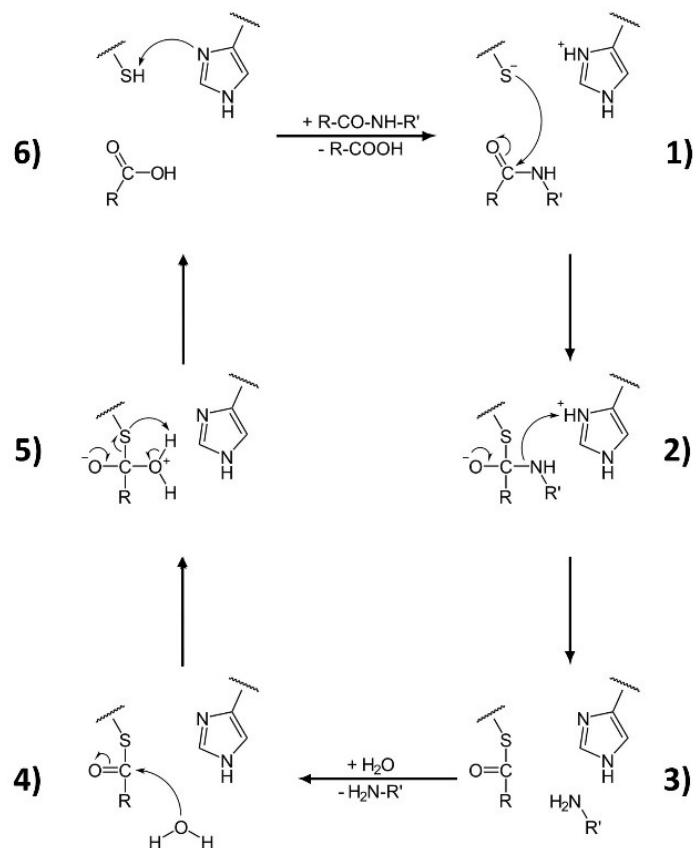
U lidí bylo doposud popsáno více jak 550 proteas, které se vzájemně liší svou strukturou, substrátovou specifitou, katalytickým mechanismem i molekulovou hmotností od 20 kDa po 6 MDa (Bond 2019). Tři největší skupiny představují proteasy cysteinové, serinové a metaloproteasy. Minoritně jsou pak zastoupeny proteasy threoninového, aspartátového a glutamátového typu (Kappelhoff et al. 2017).

Kromě klasifikace dle katalytického mechanismu se proteasy dělí na základě místa štěpení peptidového řetězce. To může probíhat uvnitř substrátu v případě endopeptidas či na konci substrátového řetězce, které katalyzují exopeptidasy. Exopeptidasy se dále dělí na aminopeptidasy štěpící na N-konci aminokyselinového řetězce substrátu a karboxypeptidasy

štěpící od C-konce řetězce substrátu. V případě štěpení po definovaných oligipeptidových fragmentech se používá označení dipeptidylpeptidasy, odštěpující dipeptid z N-konce a peptidyl-dipeptidasy, odštěpující z C-konce substrátu. Na základě evoluční příbuznosti a stavby katalytických center se jednotlivé třídy proteas dále člení do klanů a ty se dle primární aminokyselinové sekvence následně dělí do jednotlivých rodin. Aktuální klasifikaci lze nalézt v databázi proteas Merops (Rawlings et al. 2018). Následující kapitoly budou věnovány proteolytickým enzymům z třídy cysteinových proteas.

2.1.2 Mechanismus katalytického působení cysteinových proteas

Aktivní místo cysteinových proteas obsahuje katalytickou dyádu, tvořenou thiolovou skupinou cysteinu a imidazolovým kruhem histidinu, který přijímá proton a zvyšuje nukleofilitu thiolové skupiny přes imidazolium-thiolátový systém. Ke stabilizaci meziprojektu navázaného thioesterovou vazbou na katalytický cystein dochází v takzvané oxyaniontové jámě prostřednictvím postranního řetězce aminokyseliny glutaminu. Mechanismus katalýzy lze popsat v šesti krocích (Obr. 2, str. 8): 1) Zahájení reakce nukleofilním atakem thiolové skupiny katalytického cysteinu na karbonylový uhlík štěpené peptidové vazby. 2) Vznik nestabilního tetraedrického intermediátu, kde je uhlíkový atom štěpené vazby prostřednictvím thioesterové vazby kovalentně propojen s enzymem. 3) Donace protonu z kladně nabitého katalytického histidinu na aminoskupinu a jeho regenerace za současného odštěpení prvního produktu s koncovou aminoskupinou. 4) Nukleofilní atak molekuly vody na thioesterově vázaný uhlíkový atom. 5) Vznik druhého štěpeného produktu s koncovou karboxylovou skupinou. 6) Proton odštěpený z molekuly vody se stane součástí imidazolium-thiolátového systému enzymu a nastává regenerace aktivního místa a celý mechanismus se znovu opakuje (Cstorer and Ménard 1994).



Obrázek 2: **Schéma katalytického mechanismu cysteinových proteas.** Mechanismus probíhá v šesti následných krocích: 1) Nukleofilní atakem thiolové skupiny na karbonylový uhlík peptidové vazby. 2) Tvorba thioesterově vázaného intermediátu. 3) Přesun protonu z imidazolového kruhu histidinu na aminoskupinu substrátu, regenerace imidazolového zbytku a odštěpení prvního produktu reakce. 4) Nukleofilní atak molekuly vody na thioesterově vázaný atom uhlíku. 5) Uvolnění druhého reakčního produktu. 6) Regenerace aktivního místa enzymu. Převzato a upraveno z (Rzychon 2004).

2.2 Lidské cysteinové proteasy

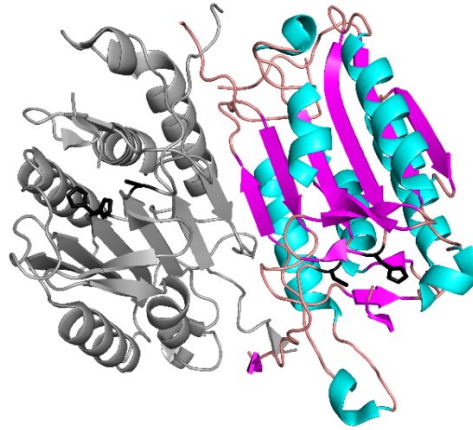
Cysteinové proteasy jsou rozděleny na 11 klanů a více jak 90 rodin. Klany se vzájemně odlišují vzájemnou pozicí katalytických zbytků v primární sekvenci enzymu. Nejpočetnější je klan CA, který čítá přibližně 40 rodin a nejvíce je prostudována rodina C1 (rodina papainu) a C2 (rodina kalpainu). Druhým nejvíce prostudovaným klanem je klan CD, který obsahuje 7 rodin. Nejznámější jsou rodiny C13 (asparaginyl endopeptidasy neboli legumainy) a C14 (kaspasy) (Rawlings et al. 2018). Některé klany vyžadují pouze katalytickou dyádu (CD, CL) ve svém aktivním místě, tvořenou aminokyselinovými zbytky cysteinu a histidinu. Výjimkou je klan CG, který má katalytickou dyádu tvořenou dvěma zbytky cysteinu.

U jiných klanů (např. CA, CE a CF) je dále nutná i přítomnost dalších katalytických zbytků v podobě sekvenčně konzervované kyseliny glutamové či asparagové, které napomáhají orientaci imidazolového kruhu histidinu a tvorbě tzv. oxyaniontové díry pro stabilizaci reakčního meziprojektu (Barrett 1996). V následujících odstavcích budou stručně popsány výše zmíněné rodiny cysteinových proteas.

2.2.1 Kaspasy (klan CD, rodina C14)

Tyto proteasy štěpí specificky za kyselinou asparagovou (Asp, odtud pochází jejich název) v aminokyselinové sekvenci substrátu v pozici P1. Lidský genom jich kóduje 12 a jsou rozděleny na tři skupiny. První skupinu tvoří prozánětlivé kaspasy (1, 4, 5 a 12), druhá skupina se označuje jako iniciační kaspasy (2, 8, 9 a 10) a poslední skupinu tvoří exekutivní kaspasy (3, 6 a 7). Kaspasa 14 je nezařazena. Kaspasy jsou syntetizovány v buňce ve formě proenzymů (prokaspas) a ke své katalytické funkci musí být nejprve aktivovány na maturní enzymy, za následné tvorby dimerních či oligomerních forem. Struktura kaspasy 3 je znázorněna na (Obr. 3, str. 10) (Van Opdenbosch and Lamkanfi 2019).

Kaspasy jsou velmi důležitými regulátory apoptózy. Jedná se o komplexní a vysoce regulovaný proces, během kterého buňka prochází morfologickými změnami a je řízeně destruována za minimálního poškození okolních buněk. Apoptózou je například řízena selekce B-lymfocytů v kostní dřeni a T-lymfocytů v brzlíku (Burger et al. 2014). Odlišným procesem je nekróza, kdy dochází k poruše celistvosti buněčné membrány a buněčný obsah je uvolněn do okolí. Specifickým procesem je pyroptóza, prozánětlivá programovaná buněčná smrt jako odpověď při napadení buňky intracelulárním patogenem, která je zahájena aktivitou kaspasy 1 (Fink and Cookson 2005).



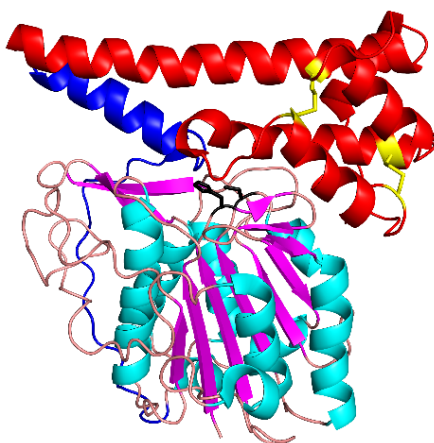
Obrázek 3: *Lidská kaspasa 3 ve formě dimeru (PDB: 5IAB)*. Barevně jsou v jedné monomerní jednotce zvýrazněny sekundární struktury (α -helix: růžově, β -skládaný list: tyrkysově) a katalytické zbytky (Cys163 a His121: černě). Monomer se skládá z charakteristických šesti segmentů β -skládaného listu a pěti α -helixů (Vega-Tejido et al. 2014).

2.2.2 Legumainy a GPI-transamidasa (klan CD, rodina C13)

Název legumain pochází z anglického „legume“ (luštěnina), protože byl poprvé izolován z fazolu *Phaseolus vulgaris*. Tyto proteasy jsou tkáňově specifické a jejich fyziologická úloha u člověka spočívá ve štěpení antigenů, které jsou následně prezentovány profesionální antigen prezentující buňkou na MHC glykoproteinu II (Wolk et al. 2005). Spolu s katepsiny proteolyticky aktivují toll-like receptory, které zahajují nespecifickou imunitní odpověď organismu. Dále inhibují proces tvorby osteoklastů v kostní tkáni.

Legumainy se svou substrátovou specifitou řadí mezi asparaginyl endopeptidasy. Jsou syntetizovány ve formě proenzymů (Obr. 4, str. 11) a v kyselém pH jsou autokatalyticky konvertovány na maturní proteasy (Dall and Brandstetter 2016). Během aktivace je odštěpena N-terminalní C-terminální propeptidová doména. C-doménu následně rozpoznává receptor diferenciovaných prekurzorových osteoklastů, což vede k zastavení přeměny na zralé osteoklasty; proto je tato doména (Obr. 4, str. 11 vyobrazena červeně) označována jako „osteoclast inhibitory peptide 2“ (Choi et al. 2001).

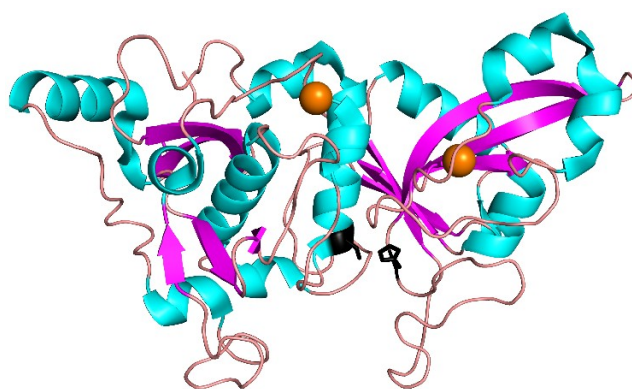
Do rodiny C13 je řazena též GPI-transamidasa, která je součástí multienzymového komplexu v lumen endoplazmatického retikula. Zde katalyzuje přenos glykosylfosfatidylinositolového zbytku, označovaného jako GPI kotva na C-konce proteinů, které rozeznává na základě specifického sekvenčního motivu. Takto posttranslačně modifikované proteiny zůstávají následně připojeny k plazmatické membráně (Kinoshita 2020).



Obrázek 4: **Lidský legumain ve formě proenzymu** (PDB: 4FGU). Molekula se skládá z N-koncového propeptidu (aktivačního peptidu) a C-koncového propeptidu (modře resp. červeně), které blokují aktivní místo enzymu (černě jsou vyznačeny katalytické zbytky Cys189 a His148); stabilizující disulfidické můstky jsou žlutě. Struktura katalytické domény legumainu vykazuje podobnou architekturu jako je tomu u kaspas s šesti segmenty β -skládaného listu a pěti α -helixy. Maturní forma je ve formě monomeru. Upraveno dle (Dall and Brandstetter 2013).

2.2.3 Kalpainy (klan CA, rodina C2)

Název kalpain, vychází ze závislosti enzymu na vápníkových iontech, které jsou nezbytné pro aktivaci této proteasy. V lidském genomu bylo doposud identifikováno 15 izoform těchto cytosolárních proteas. Kalpainy jsou regulační proteasy, které se spolu s katepsiny podílejí na specifické aktivaci biologicky aktivních proteinů např. během regulace buněčného cyklu a buněčné proliferace (Kovacs and Su 2014) či apoptózy, kde mezi jejich substráty spadá exekutivní kaspasa 3, antiapoptotické Bcl-2 a Bcl-xl proteiny a proapoptotické proteiny Bid a Bax (Li et al. 2009). Vyjma izoform kalpainu 1 (Obr. 5, str. 12) a kalpainu 2 se jedná o tkáňově specifické proteasy s enzymovou aktivitou v neutrální oblasti pH (Hosseini et al 2018).



Obrázek 5: **Katalytická doména lidského kalpainu 1** (PDB: 2ARY). Molekula je tvořena ze dvou domén s aktivním místem (katalyticky aktivní Cys115 a His272 jsou černě) uloženým mezi nimi. Levá doména s katalytickým Cys115 je tvořena převážně α -helixy (tyrkysově), kdežto pravou doménu s katalytickým His272 tvoří hlavně β -struktury (fialově). Dva vápníkové ionty (oranžově), charakteristické pro kalpains jsou nezbytné pro aktivaci a enzymovou aktivitu. Upraveno dle (Davis et al. 2007).

2.2.4 Cysteinové katepsiny (klan CA, rodina C1)

Název katepsinů je odvozen z řečtiny a odkazuje na význam „trávit“. Krom cysteinových proteas z rodiny C1 jsou mezi katepsiny řazeny ještě dvě serinové proteasy (katepsin A a G) a dvě proteasy aspartátové (katepsin D a E). Lidský genom kóduje celkem 11 cysteinových katepsinů, které mají různé fyziologické funkce (Turk et al. 2012). Katepsiny se obecně podílejí na lyzozomálním či endosomálním odbourávání proteinů a degradaci komponent extracelulární matrix (Bohley and Seglen 1992). Vedle této nespecifické degradace proteinových substrátů svoji aktivitou specificky ovlivňují řadu fyziologických procesů, jako jsou například: cytotoxická imunitní odpověď (katepsin C) (Zavašnik-Bergant and Turk 2006), regulace krevního tlaku přes renin angiotensinový systém (katepsin B) (Patel et al. 2017), štěpení antigenů a prezentace na MHC glykoproteinu II (katepsiny V, L, S, B, F) (Villadangos and Ploegh 2000), regulace buněčného cyklu a buněčná diferenciaci proteolýzou histonu H3 v jádře (katepsin L) (Duncan et al. 2008) či remodelace a resorpce kostní tkáně (katepsin K) (Brömme et al. 1996).

Vnitrobuněčný transport obstarává N-glykosidicky vázaná manosa-6-fosfátová značka, která je umístěna během posttranslační modifikace na molekulu katepsinu a zajišťuje transport do lysozomu prostřednictvím manosa-6-fosfátového receptoru (Brix et al. 2008). Lokalizace nezahrnuje pouze lysozomy, katepsiny jsou v rámci této dráhy transportovány do endozomů či fagozomů. Katepsin H se například nalézá v časném endozomu makrofágů, kdežto katepsin S je lokalizován až v endozomu pozdním

(Claus et al. 1998). Ne všechny katepsiny jsou závislé jen na manosa-6-fosfátové cestě; například katepsin W, ač má vázanou manosa-6-fosfátovou značku, setrvává v endoplazmatickém retikulu (Bühling et al. 2002). Katepsiny jsou biosyntetizovány ve formě proenzymů a své pH optimum mají v kyselé oblasti; výjimkou je katepsinu S s pH optimumem v neutrální oblasti (Shi et al. 1992; Vasiljeva et al. 2005).

Lysozomy poskytují cysteinovým katepsinům ideální prostředí (s nízkou hodnotou pH 4 až 4,5), kde jsou autokatalyticky aktivovány na maturní formu. Většina katepsinů rapidně ztrácí svoji aktivitu při neutrálním pH. Buněčná lokalizace a regulace enzymové aktivity na základě pH tedy poskytuje tkáni primární ochranu před nežádoucí aktivitou těchto proteas po jejich náhodném úniku z těchto organel (Turk et al. 2012). Tkáňová distribuce cysteinových katepsinů je převážně nespecifická; více vyhraněná je u několika členů, zejména katepsinu W, který se nachází převážně v cytotoxických NK buňkách, katepsinu S v dendritických buňkách a lymfatických tkáních, katepsinu V v epiteliálních buňkách brzlíku a varlat či katepsinu K v osteoklastech (Brix et al. 2008; Turk et al. 2012).

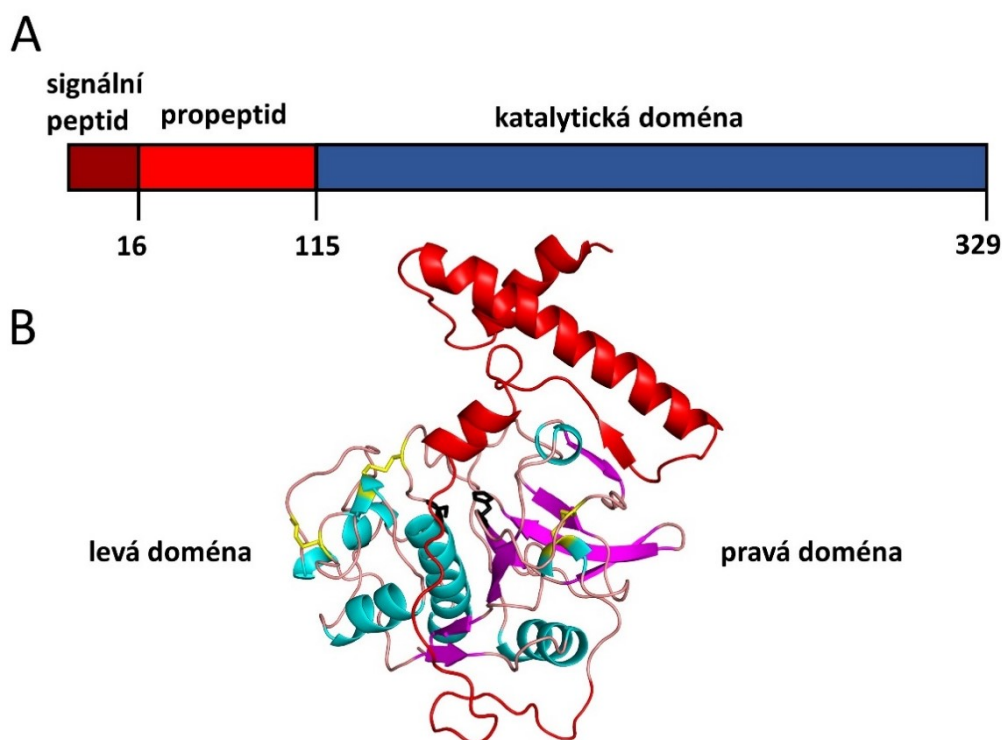
Zvýšená exprese katepsinů může poukazovat na patologický stav a z tohoto důvodu jsou některé katepsiny využívány jako prognostické markery, zejména tumorové. Například katepsiny B a L jsou spojeny s rakovinou slinivky či plic s tvorbou metastáz (Gocheva et al. 2006). Katepsin S slouží jako prognostický marker při karcinomu tlustého střeva (Gormley et al. 2011), katepsiny X a H při rakovině prsu (Decock et al. 2008) a katepsiny K a S při nádoru prostaty s následnou tvorbou metastáz (Brubaker et al. 2003; Wilkinson et al. 2015). Následující kapitoly budou pojednávat o katepsinu K, jeho fyziologické roli při remodelaci kostní tkáně, jeho patologickém působení a o cílené inhibici pro terapeutický zásah.

2.3 Aktivita a struktura Katepsinu K

V roce 1994 byl při diferenčním mapování králičí cDNA knihovny, pocházející ze slinivky a osteoklastů objeven gen, který byl podobný genu pro katepsiny L a S a jeho genový produkt dostal označení OC-2 (Tezuka et al. 1994). Lidský homolog OC-2 byl posléze úspěšně klonován několika vědeckými skupinami nezávisle na sobě a později přejmenován na katepsin K (KatK) (EC: 3.4.22.38) (Shi et al. 1995; Brömme and Okamoto 1995; Li et al. 1995; Inaoka et al. 1995). Gen pro lidský KatK (CTSK) se nalézá na prvním chromosomu, konkrétně v pozici 1q21.3 a obsahuje celkem 7 intronů a 8 exonů (Arman et al. 2014).

Produktem tohoto genu je proteasa, která je exprimována ve formě preproenzymu. Po odstranění signální sekvence (15 aminokyselin) v endoplazmatickém retikulu následuje v dalším proteolytickém kroku jeho kyselá autoaktivace v lysozomu. Vlivem změny pH dochází ke konformační změně propeptidové části o 99 aminokyselinách (Obr. 6B, str. 15) a jejímu následnému štěpení v pozicích Glu19, Ser98 a Glu110 (číslování sekvence je dle preproenzymu). Tímto způsobem se postupně tvoří aktivační meziprodukty a výsledně vzniká po odstranění celého propeptidu plně aktivovaný maturní enzym o délce 215 aminokyselin (Obr. 6A, str. 15) (McQueney et al. 1997).

Maturní forma KatK, stejně jako ostatní proteasy papainové rodiny, má charakteristickou prostorovou architekturu dvou domén označených jako levá L-doména se třemi α -helixy a pravá R-doména obsahuje β -barel, který je uzavřen α -helixem. Celá struktura (Obr. 6B, str. 15) je stabilizována třemi disulfidickými můstky. V mezeře mezi levou a pravou doménou je aktivní místo, tvořené katalytickou dyádou Cys25 a His162 a dále pomocným zbytkem Asp182, který napomáhá správné orientaci His162 (Zhao et al. 1997).



Obrázek 6: *Stavba molekuly katepsinu K.* A) Schéma hlavních segmentů v aminokyselinové sekvenci KatK ukazuje signální propeptid (zbytky 1-16, hnědě), propeptidovou část (16-115, červeně) a katalytickou doménu (115-329, modře) (číslování je podle preproenzymu). B) 3D struktura zymogenu lidského KatK (PDB: 1BY8). Propeptidová část tvořená třemi α -helixy (červeně) překrývá aktivní místo enzymu s katalytickými zbytky Cys25 a His162 (černě) (číslování je podle maturní formy). Celková katalytická doména odpovídající maturnímu proteinu (α -helixy jsou tyrkysově, β -struktury fialově, smyčky růžově) je tvořena dvěma dílčími doménami. Levá doména je tvořena převážně α -helixy s katalytickým Cys25 a je stabilizována dvěma disulfidickými můstky (žlutě). Pravá doména je skládána do β -barelu, který je uzavřen α -helixem, stabilizována disulfidickým můstkem a nese katalytický His162 (LaLonde et al. 1999).

KatK je klasifikován jako endopeptidasa se substrátovou specifitou pro hydrofobní aminokyseliny v pozici P1 (valin, leucin a izoleucin) a aminokyselinu prolin v pozici P2. Preference pro prolin v P2 pozici je pro KatK charakteristická (díky stavbě vazebné kapsy S2, která je tvořena Tyr67 a Leu205) a významně ho odlišuje od jeho strukturních homologů, jako jsou katepsiny L a S, se kterými má až 60 % strukturní podobnost. Ostatní vazebná podmísta mohou přijmout většinu zbytků aminokyselin, nicméně arginin, histidin a lysin s nabitými postranními řetězci jsou vázány přednostně (Lecaille et al. 2003; Choe et al. 2006; Brömme 2013).

2.3.1 Fyziologická úloha katepsinu K v lidském organismu

Na základě dat z *in situ* hybridizačních experimentů bylo zjištěno, že mRNA pro KatK se nachází ve více tkáních: v srdeční, plicní, kosterního svalu, vaječnících, folikulech štítné žlázy a placenty. Nicméně nejvyšší hladina exprese mRNA byla zjištěna v kostní tkáni (konkrétně v osteoklastech), které jsou zodpovědné za resorpci kostní tkáně. Dále pak byla nalezena v epiteliálních buňkách plicní tkáně, v buňkách synoviálních fibroblastů, chondrocytů a u aktivovaných makrofágů (Lecaille et al. 2008). Exprese KatK v lidských osteoklastech představuje více jak 90 % z celkové exprese všech produkovaných katepsinů v tomto typu buněk. Imunolokalizační pokusy odhalily, že se KatK nachází v lysozomech specificky lokalizovaných primárně poblíž zvlněného (tzv. kartáčového) lemu osteoklastů (Littlewood-Evans et al. 1997b). Buňka osteoklastu má polarizovaný charakter, kdy bazální strana svým kartáčovým lemem uzavírá prohlubeň v místě resorpce kostní tkáně, označovanou jako resorpční lakuna. Aktivitou vakuolární V-ATPasové pumpy, karbonické anhydrasy II a Na^+/H^+ antiportéru dochází ke snížení pH v oblasti resorpční lakuny (Rousselle and Heymann 2002), což zvyšuje rozpustnost anorganické složky kostní tkáně (hydroxyapatitu). Nízká hodnota pH je zároveň důležitá pro aktivitu hydrolytických enzymů, které se podílejí na odbourávání kostní matrix.

Kromě anorganické složky je kostní tkáň tvořena i organickou složkou, kterou je převážně kolagen typu I. Struktura kolagenu I se sestává ze dvou identických $\alpha 1$ řetězců a jednoho $\alpha 2$ řetězce, který se liší svou primární aminokyselinovou sekvencí. Tyto řetězce jsou bohaté na aminokyseliny prolin a jeho derivát 4-hydroxyprolin, lysin a jeho derivát 5-hydroxylysin a dále glycin. Tyto řetězce jsou vzájemně uspořádány do levotočivé šroubovice a stabilizovány za pomoci vodíkových můstků. Jednotlivá kolagenová vlákna jsou enzymaticky navzájem propojena působením enzymu lysyl oxidasy, která katalyzuje tvorbu pyridinolinů a deoxypyridinolinů ze zbytku lysinu, resp. 5-hydroxylysinu (Shoulders and Raines 2009).

KatK je jako jediná dosud známá savčí proteasa schopna štěpit kolagen ve formě rigidní trojitě šroubovice (tropokolagen) na více místech lokalizovaných poblíž N-konce kolagenového vlákna (Garnero et al. 1998). Pro kolagenasovou aktivitu KatK je nezbytně nutná komplexace této proteasy s chondroitin-4-sulfátem, který je součástí kostní extracelulární matrix (Li et al. 2002; Sharma et al. 2015). Vyjma KatK jsou schopny omezeně štěpit kolagen ještě proteasy ze třídy matrixových metaloproteas

(MMP-1, 8, 9, 13 a 14) s hemopexinovou doménou, která částečně rozvolňuje kolagenovou šroubovici, dále neutrofilní elastasa (u rozvolněné formy kolagenu) a minoritně též katepsiny S, B, D a L (Novinec and Lenarčič 2013). KatK vykazuje nejvyšší kolagenasovou aktivitu na rigidním tropokolagenu za tvorby pyridinolin–deoxypyridinolinových telopeptidů (markery kostní resorpce) (Drake et al. 2017). Tento fakt potvrzují studie na myších modelech, kdy byl gen pro tuto proteasu vypnut (Saftig 1998; Gowen et al. 1999) či u pacientů s pyknodystózou (Gelb et al. 1996). Dalším substrátem KatK je glykoprotein osteonektin, což je po kolagenu typu I druhý nejvíce zastoupený protein kostní extracelulární matrix, který je nezbytný pro mineralizaci kostní tkáně vazbou iontů vápníku. Dále pak proteoglykan agrekan, který je součástí chrupavčité tkáně a skleroprotein elastin, který je součástí pojivové tkáně (Brömme and Wilson 2011).

KatK má svou úlohu i v procesu hojení a udržování homeostáze kožní extracelulární matrix, kde je produkován kožními fibroblasty. Uplatňuje se též při přestavbě cévní stěny degradací elastinu (nejvyšší elastolytickou aktivitu nicméně vykazuje katepsin V) (Andrault et al. 2019). Za zmínku stojí ještě účast KatK v uvolňování hormonu thyroxinu (T4) štítné žlázy z komplexní struktury glykoproteinu thyreoglobulinu (Dauth et al. 2011).

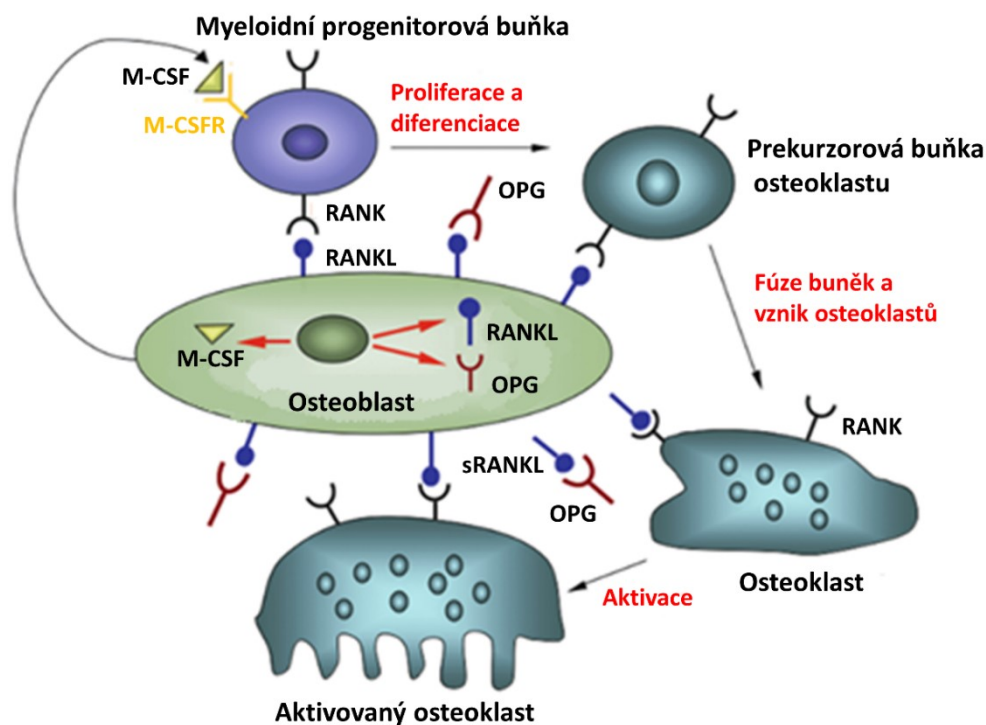
Výše uvedené studie potvrzují důležitou úlohu katepsinu K při procesu přestavby kostní tkáně a přestavby extracelulární matrix. V následující kapitole bude krátce popsána regulace přestavby kostní tkáně, kde je KatK nezbytným proteolytickým enzymem.

2.3.2 Regulace tvorby a resorpce kostní tkáně

Přestavba kostní tkáně je dynamický proces, který spočívá v cyklické produkci a resorpci kostní hmoty. Má za úkol adaptaci kostní tkáně na mechanické zatížení a reparaci drobných poškození. Odhaduje se, že po 25. roce života je posunuta rovnováha mezi tvorbou a resorpcí kostní tkáně na stranu její resorpce. Proces tvorby kostní hmoty zajišťují osteoblasty, diferenciované buňky odvozené z mezenchymálních kmenových buněk. Tyto buňky jsou zodpovědné za syntézu kostní matrix, která je tvořena převážně z kolagenu typu I, proteoglykanů, glykosaminoglykanů atp. Osteoblasty jsou po čase obklopeny kostní tkání a dochází k jejich přeměně na osteocyty, které tvoří většinu kostní tkáně a již se nepodílí na tvorbě nové extracelulární matrix. Oproti tomu odbourávání kostní tkáně zastávají mnohояaderné buňky osteoklasty, odvozené z hematopoetických buněk označovaných jako monocyto-makrofágová řada (Lin et al. 2020). Pro diferenciaci na zralé osteoklasty je třeba působení cytokinů jakými jsou M-CSF („Macrophage Colony-Stimulating Factor“)

a membránově vázaný RANKL („Receptor Activator of Nuclear Factor κ B Ligand“), produkovaný osteoblasty či jeho rozpustná a volná forma sRANKL, která vzniká proteolyticky za účasti metaloproteas. Vazbou RANKL či sRANKL na RANK receptor u osteoklastů se spustí vnitřní buněčná signalizace a následná produkce transkripčních faktorů, což vede k diferenciaci prekurzorových buněk na buňky plně diferenciované a expresi adhezivních receptorů z třídy integrinů, nutných pro adhezi osteoklastů (Li et al. 2000).

Adheze osteoklastů přes integrinové receptory je též závislá na aktivitě KatK, který štěpením kolagenu typu I v extracelulární matrix odhaluje sekvenční tripeptidové RGD motivy u vzniklých fragmentů kolagenu, které rozeznávají integriny na povrchu osteoklastů a dochází k jejich adhezi (Brömme and Wilson 2011). Osteoblasty produkují i rozpustný cytokinový receptor zvaný osteoprotegrin, který vychytává cytokin RANKL a nedochází tudíž k diferenciaci na zralé osteoklasty (Wada et al. 2006). Interakcemi RANKL/RANK/osteoprotegrin (Obr. 7, str. 19) osteoblasty ovlivňují nejen diferenciaci progenitorových buněk na zralé osteoklasty, ale i následnou adhezi. Regulaci exprese RANKL dále pozitivně ovlivňují hormony parathormon a kalcitriol či prozánětlivé cytokiny (IL-1 a TNF α) a negativně působí hormon estrogen a cytokin TGF β (Walsh and Choi 2014).



Obrázek 7: **Schéma diferenciacie osteoklastů.** Buňky osteoblastů produkují růstový faktor M-CSF, který rozeznávají myeloidní progenitorové buňky prostřednictvím svého receptoru (M-CSFR). Dochází k diferenciaci, proliferaci a vzniku prekurzorových osteoklastických buněk. Fúzí prekurzorových buněk vznikají mnohояderné osteoklasty, které rozeznávají cytokin RANKL či jeho rozpustnou formu sRANKL produkovaný osteoblasty prostřednictvím RANK receptoru na svém povrchu. To má za následek spuštění vnitřní buněčné signalizační kaskády s produkcí transkripčních faktorů a nastává aktivace osteoklastů a vznik plně diferenciováných osteoklastů. Osteoblasty produkují i rozpustný cytokinový receptor OPG (osteoprotegrin), který vychytává RANKL či jeho rozpustnou formu sRANKL a tím brání diferenciaci na zralé osteoklasty. Převzato a upraveno z (Alliston et al. 2008).

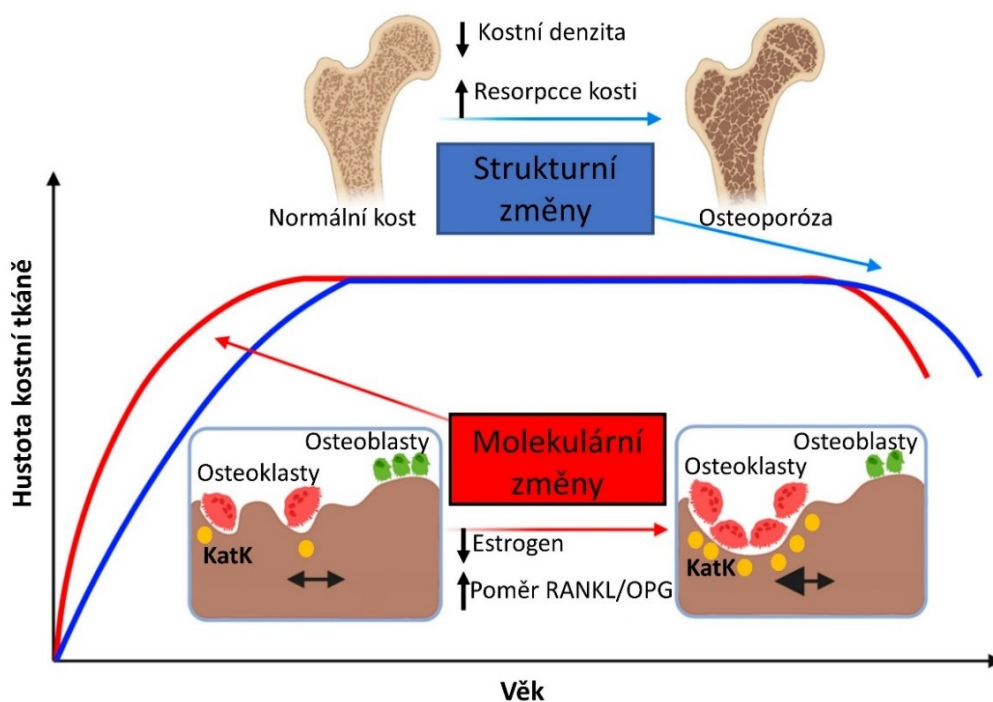
2.3.3 Katepsin K a jeho role v patologických procesech

V této podkapitole jsou uvedena onemocnění, u kterých je prokázáno přímé či zprostředkované spojení s KatK. Tyto patologické děje velmi často úzce souvisí s chybnou regulací a následně zvýšenou či sníženou aktivitou této proteasy v tkáni.

Mutace genu CTSK kódujícího KatK, obvykle v propeptidové či signální části sekvence, má za následek poruchu vnitrobuněčného transportu proteasy či dochází k jejímu špatnému sbalení a degradaci. Výsledkem je snížená či kompletní ztráta enzymové aktivity. Toto vzácné, autosomálně recesivní onemocnění je nazýváno pyknodistóza či Toulouse-Lautrec syndrom, podle francouzského malíře, který tímto onemocněním trpěl (Kumar 1988). Pacienti s tímto postižením trpí osteosklerózou, kdy je kostní tkáň kompaktnější oproti normálnímu stavu vlivem snížené resorpce kostní tkáně osteoklasty a kost je tudíž

křehčí a náchylnější k únavovým frakturám. Mezi další klinické projevy patří trpasličí vzrůst (dwarfismus), deformace lebky a zubů (Gelb et al 1996; Schmidt et al. 2020).

Pochopení molekulární podstaty pyknodistózy otevřelo nové možnosti pro klinickou terapii chorob, kde je naopak aktivita KatK zvýšená. V případě, kdy dochází k trvale převažující kostní resorpci nad tvorbou nové kostní tkáně, hovoříme o osteoporóze. Je důležité podotknout, že na rozdíl od osteomalacie, kde dochází k úbytku anorganické složky kostní tkáně, je u osteoporózy zachován poměr mezi anorganickou složkou a osteoidem. Osteoporóza se projevuje více u starších osob, zvláště u žen po menopauze (postmenopauzální osteoporóza) (Obr. 8, str. 21). Klesá hladina hormonu estrogeneru, čímž se zvyšuje tvorba proteinu RANKL, který stimuluje diferenciaci myeloidních prekurzorových buněk na zralé osteoklasty (Streicher et al. 2017). Studie naznačují, že redukce kostní tkáně se projevuje i u mužů se sníženou hladinou testosteronu a je častěji spojována s poruchou funkce pohlavních žláz (hypogonadismus) (Golds et al. 2017). Senilní osteoporóza (typ II), postihující pacienty nad 70 let je způsobena převládající diferenciací kmenových buněk kostní dřeně v adipocyty, oproti osteoblastům (Qadir et al. 2020). Osteoporóza může být vyvolána i v případě výživových poruch, po chirurgickém odstranění vaječnicků, při onemocnění ledvin či nadměrné aktivitě štítné žlázy (hyperthyreózy). Jedná se o velmi časté onemocnění, a tudíž je KatK atraktivním terapeutickým cílem (Drake et al. 2017).



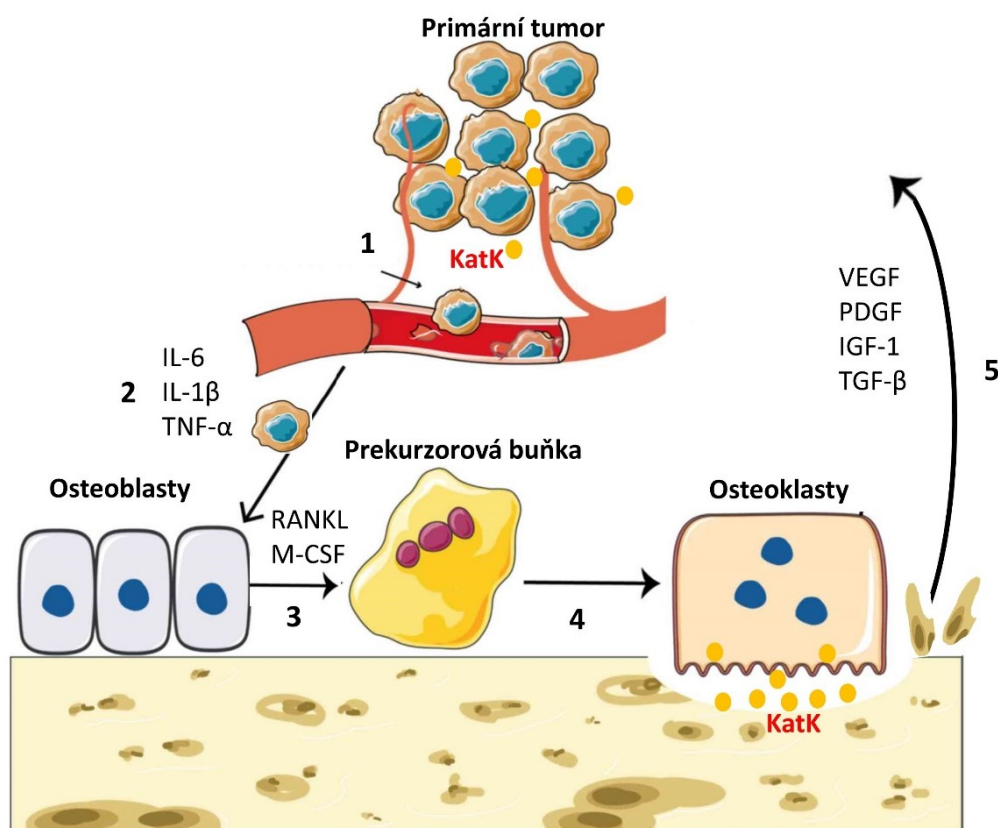
Obrázek 8: **Postmenopauzální osteoporóza.** S rostoucím věkem u žen klesá hladina hormonu estrogenu. Následkem je porušení poměru koncentrace proteinů RANKL/OPG systému ve prospěch RANKL a stimulace tvorby nových osteoklastů. Výsledkem je posun rovnováhy v prospěch kostní resorpce zprostředkované osteoklasty, které uvolňují do resorpční lakuny KatK, nad tvorbou nové kostní tkáně vzniklé aktivitou osteoblastů. Dochází ke snížení kompaktnosti kostní tkáně a kost se stává náchylnější k frakturám. Převzato a upraveno z (Park et al. 2021).

Schopnost KatK štěpit kolagen typu II a proteoglykan agrekan, který se nachází v chrupavčité tkáni je spojena s dalšími typy zánětlivých onemocnění pohybového aparátu, jakými jsou **osteoartróza** a **revmatoidní artritida**. Za normálních okolností je homeostáze chrupavčité tkáně zprostředkována chondrocyty v klidovém stádiu, které produkují jen omezené množství proteas (zejména matrixové metaloproteasy). Ty štěpí kolagen typu II charakteristickým způsobem poblíž jeho C-konce a generují zkrácené fragmenty. Naproti tomu KatK štěpí kolagen typu II v místě nehelikální C-koncové části a současně i v helikální oblasti N-konce (Dejica et al. 2012). Štěpením proteoglykanu agrekanu vznikají volné glykosaminoglykany, které následně stabilizují KatK v extracelulárním pH a současně alostericky ovlivňují jeho aktivitu. Při narušení rovnováhy homeostáze vlivem senescence, mechanickým poškození chrupavčité tkáně, nadváhou či hormonálně dochází ke změně fenotypu chondrocytů na metabolicky aktivní buňky, jejich shlukování a hypertrofické diferenciaci (Houard et al. 2013). Aktivované chondrocyty vlivem stresu produkují současně prozánětlivé cytokiny a proteolytické enzymy v mnohem vyšší míře. Bylo prokázáno, že oproti expresi ostatních cysteinových katepsinů je vlivem cytokinu IL-1 β zvýšena pouze

sekrece KatK, kterému je připisována hlavní proteolytická aktivita, podobně jako v případě osteoklastické eroze kostní tkáně (Hou et al. 2002). Degradací extracelulární matrix se generují signalizační stresové molekuly, které chondrocyty rozpoznávají svými povrchovými receptory a opětovnou sekrecí cytokinů a proteas amplifikují výslednou destrukci, osifikaci, tvorbu zánětu, apoptózu a nekrózu chrupavčité tkáně. Kromě chondrocytů jsou schopny KatK sekretovat i synoviální fibroblasty (Kozawa et al. 2016) a aktivované CD68⁺ makrofágy (Dodds et al. 1999; Bühling et al. 2001).

Při patologicky zvýšené aktivitě proteas dochází ke štěpení signálních a adhezivních molekul na povrchu buněk či degradaci složek extracelulární matrix za uvolnění biologicky aktivních biomolekul. Tyto děje vedou k narušení signálních vnitrobuněčných kaskád s následnou proliferací, diferenciací, ztrátě adheze buněk a jejich migraci po těle za tvorby metastáz z primárního tumoru. Zvýšená hladina exprese KatK byla zjištěna například u **karcinomu prsu** (Littlewood-Evans 1997a), **plic** (Yang et al. 2020) či **prostaty** (Brubaker et al. 2003) a KatK zde představuje současně prognostický marker těchto onemocnění. U karcinomu prsu, prostaty a plic je rovněž nejčastějším místem migrace těchto buněk kostní tkáň (Liang et al. 2019; Vashum et al. 2021). V případě primárního **nádoru kosti (osteosarkomu)** (Husmann et al. 2008) je naopak pozorována migrace nádorových buněk především do plic a mozku a v terminálním stádiu nemoci dochází též k zasažení kostní tkáně v jiných částech těla.

Remodelací extracelulárního prostředí a invazi nádorových buněk do kostní tkáně (Obr. 9, str. 23) zajišťují proteasy sekretované těmito buňkami. Aktivitou těchto proteas se uvolňují růstové faktory, které dále podporují angiogenezi, chemotaxi nádorových buněk do kostní tkáně, proliferaci a expresi adhezivních molekul na povrchu invazivních nádorových buněk (Verbovsek et al. 2015; Vashum et al. 2020). KatK a Katepsiny typu B, L a S rovněž štěpí endogenní inhibitory matrixových metaloproteas, což zvyšuje degradaci složek extracelulární matrix. Současně nadměrná metabolická aktivita nádorových buněk tvoří mikroprostředí se sníženým pH a obsahem kyslíku, které ještě umocňuje proteolytickou aktivitu cysteinových katepsinů (Fasanya and Siemann 2020).



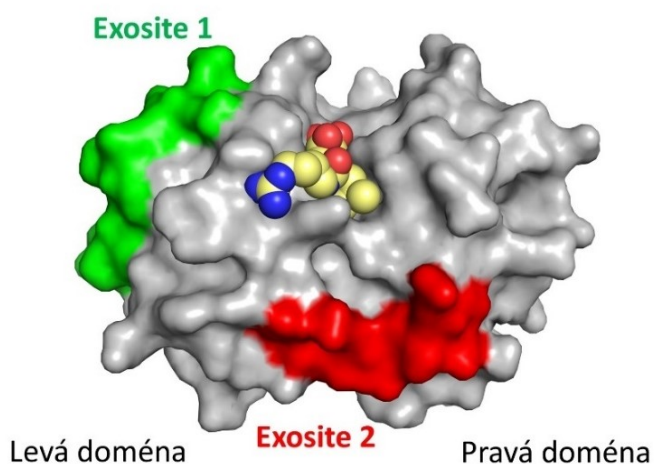
Obrázek 9: **Role katepsinu K při tvorbě metastáz do kostní tkáně.** 1) Buňky primárního tumoru sekretují KatK, který degraduje extracelulární matrix a adhezivní molekuly v okolí nádorových buněk. Uvolněné buňky vstupují do krevního řečiště a putují do kostní tkáně. 2) Tumorové buňky produkují škálu cytokinů a růstových faktorů (IL-6, IL-1 β , TNF- α). 3) Osteoblasty exprimují stimulační a růstové faktory (RANKL, M-CSF), které dávají povel prekurzorovým buňkám pro diferenciaci na zralé osteoklasty. 4) Plně diferenciované osteoklasty zahajují resorpci kostní tkáně a degradací složek extracelulární matrix se uvolňují stimulační růstové faktory (5), které podporují růst nádorových buněk (primárních a metastatických buněk odvozených z primárního nádoru). Převzato a upraveno z (Li et al. 2019).

2.4 Regulace aktivity katepsinu K

Modulace aktivity KatK probíhá na několika úrovních. 1) Regulace na úrovni genové exprese této proteasy (Corisdeo et al. 2001) 2) Regulace aktivace pomocí hodnoty pH, kdy vlivem kyselého pH v lysozomech dochází k reverzibilní autoaktivaci proenzymu a konverzi na proteolyticky aktivní maturní formu (McQueney et al. 1997). Naopak dlouhodobým působením kyselého pH na maturní formu má za následek její nevratnou inaktivaci, kdy dochází ke konformačním změnám. Stejný efekt nastává i při vystavení proteasy vlivu neutrálního či bazického pH (Brömme et al. 1996). 3) Inhibice aktivity maturní formy KatK interakcí s vlastním propeptidem. Ten představuje přirozený ligand aktivního místa, který po odštěpení ze struktury proenzymu a uvolnění je schopen opětovné interakce, kdy funguje

jako efektivní přirozený inhibitor s nanomolární hodnotou inhibiční konstanty K_i (Billington et al. 2000). 4) Alosterická modulace interakcí určitých glykosaminoglykanů s KatK. V tomto případě vazbou například chondroitin-4-sulfátu či keratan sulfátu do tzv. „exosite“ neboli exo-místa (je mimo aktivní místo) na povrchu KatK vznikají oligomerní struktury, kdy dochází ke stabilizaci KatKv neutrální oblasti pH a posílení jeho elastinolytické a kolagenolytické aktivity (Novinec et al. 2010). Opačný, inhibiční efekt má pak vazba dermatan sulfátu a heparan sulfátu (Li et al. 2004; Wilson et al. 2009). V molekule KatK byla identifikována dvě exo-místa (Sharma et al. 2015). „Exosite 1“ se nachází na L-doméně a je tvořeno segmentem Tyr67 až Gly102, „exosite 2“ je mezi L-doménou a R-doménou a je tvořeno segmentem Gly109 až Glu118 (Obr. 10, str. 24).

Bylo prokázáno, že specifickým blokováním těchto exo-míst nedochází k ovlivnění enzymatické aktivity KatK vůči malým oligopeptidovým substrátům, ale je potlačena pouze jeho elastinolytická a kolagenolytická aktivita, protože je znemožněna interakce s těmito makromolekulárními substráty. Exo-místo 1 pravděpodobně interaguje a účastní se degradace rozpustné i nerozpustné formy kolagenu, kdežto exo-místo 2 je určeno pro interakci s elastinem (Sharma et al. 2015; Panwar et al. 2018).



Obrázek 10: **Interakční místa na povrchu Katepsinu K.** Struktura lidského KatK s inhibitorem E-64 (kalotová reprezentace) v aktivním místě (PDB: 1ATK) s vyznačenými místy pro interakci s makromolekulárními proteinovými substráty. „Exosite 1“ interaguje s rozpustnou i nerozpustnou formou kolagenu, zatímco „exosite 2“ je specifické pro interakci s elastinem (Sharma et al. 2015).

2.4.1 Přírozené inhibitory a modulátory katepsinu K

V této kapitole jsou uvedeny základní rodiny přírozených (endogenních) inhibitorů proteinového charakteru (klasifikovány dle Merops databáze), které se účastní inhibiční regulace KatK a příbuzných cysteinových katepsinů.

Cystatiny (rodina I25, klan IH): cystatiny představují superrodinu inhibitorů cysteinových proteas, která zahrnuje tři hlavní rodiny: stefiny (cystatiny typu I), cystatiny typu II a kininogeny (cystatiny typu III). Jedná se o malé proteiny tvořené 100 až 120 aminokyselinovými zbytky (nebo fúzní konstrukty z těchto proteinů). Strukturně jsou tvořeny pěti antiparalelními β -strukturami a jedním α -helixem, orientovaným kolmo k β -skládanému listu, což vede k velmi stabilní struktuře i při vysokých teplotách a hraničních hodnotách pH. Typ inhibice je kompetitivní, reverzibilní s tzv. „tight-binding“ kinetikou (Lecaille et al 2008).

Stefiny (cystatiny typu I) (Turk et al. 1986; Turk et al 2008; Tusar et al. 2021): jsou tvořené jedním polypeptidovým řetězcem zhruba o 100 aminokyselinách a postrádají disulfidické můstky. Jsou převážně cytosolární a hrají protektivní úlohu při poruše integrity membrány lysozomu, kdy brání úniku aktivních lysozomálních proteas do cytosolu. Mezi zástupce stefinů patří lidské stefiny A a B. Reaktivní místo obsahuje N-konec proteinu spolu se dvěma β -vláskami, které tvoří klínovitou strukturu pro interakci s aktivním místem cílové papainové proteasy. Analogická stavba reaktivního místa je i u ostatních rodin cystatinů (Stubbs et al. 1990; Renko et al. 2010).

Cystatiny typu II: jsou na rozdíl od stefinů stabilizovány dvěma disulfidickými můstky v C-koncové polovině polypeptidového řetězce zhruba o 150 aminokyselinách. Jsou lokalizovány převážně extracelulárně, napříč všemi typy extracelulárních tělních tekutin, přičemž nejvíce zastoupeny jsou cystatin C a F. Snížená či naopak zvýšená hladina cystatinu C je spojována s celou řadou patologických jevů či nádorových onemocnění (Jones et al. 2017; Leto et al. 2018). Cystatin F je exprimován především buňkami imunitního systému například dendritickými buňkami či NK „natural killer“ buňkami ve formě neaktivního dimeru, který štěpí katepsin V na aktivní monomerní formu v lysozomech, kam je dopraven přes manosa-6-fosfátovou cestu. Mezi další zástupce cystatinů typu II patří neglykosylované cystatiny D, S, SA a SN či glykosylované cystatiny typu E a M. (Lecaille et al. 2008; Prunk et al. 2017).

Kininogeny (cystatiny typu III): souhrnně označují extracelulární polypeptidové, glykosylované prekurzory o více doménách, které proteolýzou (katalyzovanou serinovou proteasou kalikreinem) poskytují biologicky aktivní peptidy označované jako kininy. Ty se účastní procesů krevního srážení, vazodilatace či regulace kardiovaskulárního systému. Kininogeny jsou rozděleny dle molekulové hmotnosti na tři typy: nízkomolekulární (LK), vysokomolekulární (HK) a T-kininogeny, které byly nalezeny pouze u křesy. Na svém N-konci obsahují tři inhibiční domény (D1-D3), strukturně odvozené od cystatinů typu II, nicméně první doména (D1) evolučně ztratila schopnost inhibice a slouží jako transportér vápenatých iontů. Kininogeny jsou obecně produktem pouze jednoho genu a rozmanitost výsledných produktů je dána až jeho alternativním sestřihem (Lecaille et al. 2008).

Tyropiny (rodina I31, klan IX): rodina přirozených inhibitorů cysteinových proteas („thyroglobulin type-1 domain proteinase inhibitors“) obsahuje doménu Tg1, sekvenčně podobnou s thyreoglobulinovou doménou typu I. Tato doména má evolučně konzervovaný sekvenční motiv, bohatý na aminokyselinu cysteinu v primární sekvenci a podle jejich rozmístění jsou tyropiny rozděleny na podtyp 1A se šesti cysteiny a podtyp 1B se čtyřmi cysteiny. Typ inhibice je založen na reverzibilní „tight-binding“ kinetice (Mihelič and Turk 2007).

Ne všechny proteiny obsahující Tg1 doménu jsou účinnými inhibitory cysteinových proteas a KatK. Glykoprotein thyreoglobulin je produktem štítné žlázy a účastní se syntézy jejích hormonů, kde je substrátem pro KatK. Obdobným způsobem KatK štěpí Tg1 doménu testecanu-1, což je inhibitor katepsinu L, a Tg1 doménu nidogenu. Na rozdíl od cystatinů je inhibiční schopnost dosud známých tyropinů vůči KatK spíše omezená, zřejmě z důvodu sterických zábran objemné Tg1 domény (Meh et al. 2005; Novinec et al. 2007).

Serpiny (rodina I4, klan ID): superrodina serpinů je tvořena inhibitory přibližně o 400 aminokyselinách v primární sekvenci a charakteristickým sbalením do osmi α -helixů a tří β -skládaných listů, které mají v blízkosti C-konce tzv. reaktivní smyčku. Mechanismus inhibice je zde zcela odlišný od výše uvedených nekovalentních inhibitorů. Štěpení sekvence (podobné substrátu) na reaktivní smyčce působením cílové proteasy vede ke konformační změně celé molekuly serpinu a tvorbě kovalentního komplexu s inhibovanou proteasou. Název serpin označuje fakt, že se jedná převážně o inhibitory serinových proteas, ovšem jsou známy i tzv. „cross-class“ promiskuitní serpiny, které inhibují cysteinové, aspartátové proteasy či metaloproteasy. Typickými zástupci serpinů se schopností inaktivace KatK je

SCCA-1 (serpinB3) (Schick et al. 1998), serpin B13 (headpin) (Jayakumar et al. 2003) nebo serpin MENT (Irving et al. 2002).

α 2-makroglobulin (rodina I39, klan II): je plazmatický protein produkovaný makrofágy, fibroblasty či hepatocyty. Strukturně je složen ze čtyř identických podjednotek, navzájem spojených prostřednictvím disulfidických můstků. Je schopen kromě cysteinových proteas inhibovat i proteasy ostatních tříd. Mechanismus inhibice je založen na oblasti citlivé k proteolýze, která je štěpena cílovou proteasou, což spouští konformační změnu způsobující kovalentní zachycení a uvěznění proteasy ve struktuře α 2-makroglobulinu. Tím je stericky omezen přístup substrátu k aktivnímu místu inhibované proteasy (Barrett and Starkey 1973; Mason 1989; Zhu et al. 2021).

2.4.2 Syntetické inhibitory katepsinu K

Inhibitor preferovaný pro klinické použití by měl mít nízkou molekulovou váhu, vysokou selektivitu a specifitu, reverzibilní povahu inhibice a lipofilní charakter. Současně je žádoucí, aby splňoval farmakokinetická kritéria: jako je nízká toxicita, snadná prostupnost biologickými membránami či vysoký biologický poločas eliminace. Oproti stále hojně používaným léčivům na bázi bisfosfonátů, které inhibují aktivitu enzymu farnesylpyrofosfát synthasy (tím je narušena biosyntetická dráha izoprenoidů a osteoklasty podléhají apoptóze) (Drake et al. 2008). Doposud vyvíjené a experimentálně testované inhibitory cílené proti KatK zachovávají i částečnou tvorbu nové kostní tkáně.

Většina současných vyvíjených inhibitorů proti KatK a dalších cysteinových katepsinů je navržena tak, že část inhibitoru napodobuje sekvenci ideálního substrátu a část nese reaktivní skupinu, která následně reaguje s katalytickým cysteinovým zbytkem proteasy. Takové inhibitory se pak označují jako reaktivní peptidomimetické inhibitory (Leung et al 2000) a podle chemické povahy reaktivní skupiny jsou děleny na ketony, epoxidy, vinylsulfony, nitrily atp.

Interakcí katalytického místa enzymu s inhibitorem vzniká komplex enzym-inhibitor (EI). Na základě stability takto vzniklého komplexu rozlišujeme dále inhibitory na ireverzibilní a reverzibilní. Ireverzibilní inhibitory trvale modifikují cílový enzym prostřednictvím stabilní kovalentní vazby, zatímco reverzibilní inhibitory interagují s aktivním místem enzymu obvykle prostřednictvím nekovalentních interakcí (Strelow 2017). Nicméně určité reaktivní skupiny inhibitoru jsou hydrolyticky labilní (nitrily, azanitrily či aldehydy) a dochází k opětovné regeneraci aktivního místa

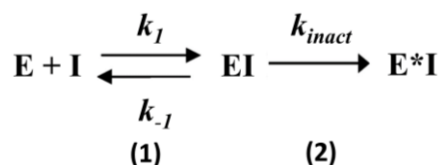
(Tusar et al. 2021). Méně zastoupeným typem inhibitorů KatK jsou pak nekovalentní inhibitory, které interagují s aktivním místem enzymu skrze vodíkové můstky, dipól-dipólové interakce či van der Waalsovy interakce. Tyto inhibitory jsou též schopny cílit místa na enzymu, která jsou lokalizována mimo aktivní centrum a jsou označována jako exo-místa (Sharma et al. 2015).

Cílené inhibici katK bylo za posledních 20 let věnováno více jak 200 publikací a více jak desítky inhibitorů vstoupila do různých fází klinických studií, což dokazuje intenzivní úsilí o nalezení specifického a účinného terapeutického prostředku pro léčbu patologií, kterých se KatK účastní (Lu et al. 2018). V následujících kapitolách je uveden stručný přehled hlavních skupin syntetických inhibitorů KatK.

2.4.2.1 Ireverzibilní kovalentní syntetické inhibitory

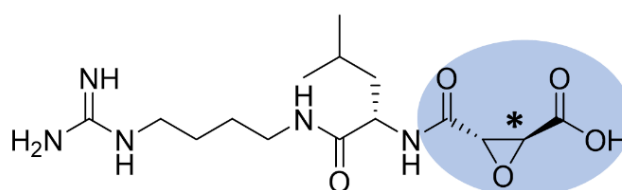
Ireverzibilní inhibitory jsou obvykle vybaveny elektrofilní reaktivní skupinou. Reakcí s nukleofilní skupinou v aktivním místě enzymu vzniká stabilní kovalentní adukt. Při dlouhodobém působení takového typu inhibitoru dojde k úplné inhibici cílových enzymů v místě působení (Ghosh et al. 2019).

Z kinetického hlediska probíhá ireverzibilní inhibice ve dvou krocích (Obr. 11, str. 28), kdy nejprve reaguje inhibitor s enzymem za tvorby přechodného nekovalentního komplexu (EI) a následuje chemická reakce za vzniku kovalentně modifikovaného enzymu s navázaným inhibitorem (E*I). Účinnost prvního kroku je dána poměrem mezi vazebnou rychlostní konstantou k_1 a disociační rychlostní konstantou k_{-1} , které popisují rychlost vzniku a rozpadu přechodného komplexu (EI). Druhý krok reakce popisuje inhibiční rychlostní konstanta k_{inact} , která je závislá na koncentraci přechodného komplexu (EI) a vyjadřuje maximální rychlost inhibice enzymu při jeho nasycení inhibitorem (Strelow 2017).



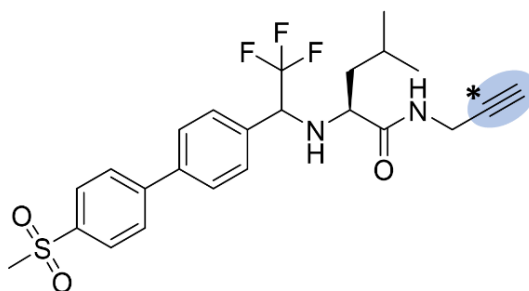
Obrázek 11: **Mechanismus inhibice ireverzibilního typu inhibice.** Reakce probíhá ve dvou následných krocích, kdy (1) inhibitor reaguje s volným enzymem za tvorby přechodného nekovalentního komplexu EI a (2) chemickou reakcí inhibitoru s katalytickým centrem enzymu dochází ke kovalentní modifikaci enzymu za vzniku komplexu E*I. Parametry k_1 , k_{-1} a k_{inact} jsou rychlostní konstanty příslušných kroků reakce. E je volný enzym a I je volný inhibitor (Strelow 2017).

Epoxidy: prvním objeveným inhibitorem, který ve své struktuře obsahoval epoxidový kruh, byla sloučenina přírodního původu. Jednalo se o L-trans-epoxysukcinyly-leucylamido(4-guanidino)butan, který byl izolován z plísně *Aspergillus japonicus* a byl pojmenován E-64 (Obr. 12, str. 29) (Hanada et al. 1978). E-64 je inhibitorem všech známých cysteinových katepsinů. Mechanismus inhibice je na bázi kovalentní alkylace katalytického cysteinu. Modifikací základního skeletu inhibitoru byl získán derivát E-64d (aloxistatin) schopný prostupovat skrze biologické membrány. E-64 se široce využívá v enzymologii při stanovení koncentrace cílových proteas metodou titrace aktivního místa či jako jejich identifikační marker v aktivitních testech (Matsumoto et al. 1999).



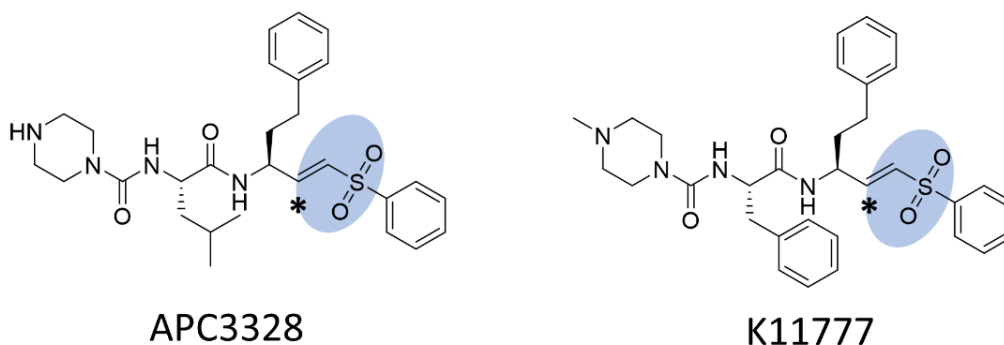
Obrázek 12: **Epoxidový inhibitor E-64.** Kostru inhibitoru tvoří modifikovaný dipeptid a pozice P1 obsahuje cyklickou reaktivní epoxidovou skupinu (vyznačená modře). Reakcí s katalytickou thiolovou skupinou enzymu v aktivním místě dochází k otevření epoxidového kruhu (* označuje uhlíkový atom, který se váže na thiolovou skupinu) a tvorbě ireverzibilního aduktu (Hanada et al. 1978).

Alkyny: alkyny byly dlouho dobu považovány za chemicky bioinertní reakční činidlo, a proto byly využívány při tzv. *in vivo* bioortogonálních studiích, využívající „click“ chemii na bázi rychlé, jednokrokové reakce bez interference s jinými biochemickými procesy. Alkynové inhibitory vznikly izoelektrickým nahrazením reaktivní nitrilové skupiny (viz dále, nitrilové inhibitory) methinovou skupinou. Svou povahou jsou alkyny sice klasifikovány jako velmi slabé (latentní) elektrofil, nicméně při dostatečném přiblížení během interakce specifického inhibitoru může nastat reakce mezi alkynovou reakční skupinou inhibitoru a thiolovým zbytkem cysteinu. Příkladem je struktura KatK (PDB: 6QBS) s modelovým inhibitorem tohoto typu (Obr. 13, str. 30) (Mons et al. 2019). Interakcí katalytického cysteinu s alkynovou reaktivní skupinou inhibitoru vzniká thiovinylový adukt.



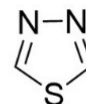
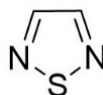
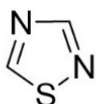
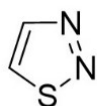
Obrázek 13: **Inhibitor s alkyňovou reaktivní skupinou.** Pozice P1 obsahuje terminální alkyňovou reaktivní skupinu (vyznačená modře). Reakcí s katalytickou thiolovou skupinou enzymu v aktivním místě dochází k tvorbě ireverzibilního thiovinylového aduktu (* označuje uhlík, který kovalentně reaguje s thiolovou skupinou) a tvorbě ireverzibilního aduktu (Mons et al. 2019).

Michaelovy akceptory: Michaelovy akceptory představují skupinu kovalentních ireverzibilních inhibitorů, které reagují s enzymem mechanismem tzv. Michaelovy reakce, kdy α,β -nenasycená karbonylová sloučenina reaguje nukleofilní adicí s thiolovou skupinou cysteinu v aktivním místě enzymu za tvorby thioetherového aduktu (Santos and Moreira 2007). Stejným způsobem reagují i α,β -nenasycené vinyl sulfony. Prvním vinylsulfonovým inhibitorem s popsáním mechanismem inhibice KatK na základě rentgenové strukturní analýzy byl APC3328 (Obr. 14, str. 30) (PDB: 1MEM). Nejznámějším vinylsulfonem je inhibitor s kódovým označením K11777 (APC3316) (Obr. 14, str. 30); jde o relativně širokospektrální inhibitor, který je slibným terapeutickým prostředkem pro léčbu parazitární Chagasovy choroby (inhibuje cysteinovou proteasu cruzipain z prvoka *Trypanosoma cruzi*) a současně je silným inhibitorem i pro lidský KatK v nanomolární koncentraci (McGrath et al. 1997; Kerr et al. 2009; Ndao et al. 2014).



Obrázek 14: **Inhibitory s vinylsulfonovou reaktivní skupinou.** Peptidomimetické inhibitory APC3328 a K11777 obsahují reaktivní vinylsulfonovou skupinou (vyznačena modře). Reakcí thiolové skupiny s katalytickou thiolovou skupinou cysteinu v aktivním místě enzymu s vinylsulfonovou skupinou vzniká thioetherový adukt. Hvězdičkou je označen uhlík, který je kovalentně spojen s atomem síry cysteinu (McGrath et al. 1997; Kerr et al. 2009).

Thiadiazoly: z chemického hlediska se jedná o syntetické pětičtené heterocyklické sloučeniny, které obsahují jeden atom síry a dva atomy dusíku. Vzájemnou polohou heteroatomů vznikají čtyři izomerní formy (Obr. 15, str. 31). Sloučeniny s thiadiazolovým heterocyklem reagují jako tzv. „thiol trapping“ inhibitory, kdy reakcí atomu síry s cílovou thiolovou skupinou dochází k otevření heterocyklu a vzniká kovalentní ireverzibilní disulfidický adukt (Tam et al 2005). U 1,2,4 thiadiazolů uhlíky v pozicích C3 a C5 mohou být dále substituovány. Substituce v pozici C3 (u izomeru 1,2,4 a 1,2,5) zvyšuje či snižuje reaktivitu heterocyklu a naopak substituce v pozici C5 (u izomerů 1,2,3, 1,2,4 a 1,3,4) vhodným peptidomimetickým postranním řetězcem zvyšuje selektivitu. Doposud byla detailně popsána a studována inhibice na lidských katepsinech B, H a S (Leung-Toung et al. 2003; Vega-Tejjido et al. 2014).

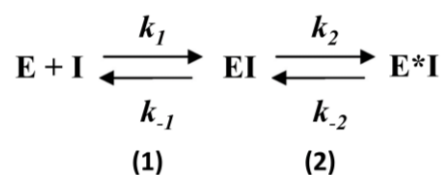


1,2,3-thiadiazol 1,2,4-thiadiazol 1,2,5-thiadiazol 1,3,4-thiadiazol

Obrázek 15: Vzorce čtyř izomerních forem thiadiazolů. Vzájemnou polohou heteroatomů vznikají čtyři izomerní formy, kdy číslování vychází od heteroatomu atomu síry. Při inhibiční reakci s cysteinovými proteasami se tvoří kovalentní disulfidická vazba mezi atom síry thiadiazolu s thiolovou skupinou katalytického cysteinu v aktivním místě enzymu (Tam et al. 2005).

2.4.2.2 Reverzibilní kovalentní syntetické inhibitory

Z farmakologického hlediska je často při dlouhodobém podávání léčiva preferována reverzibilní inhibice cílového enzymu, kdy nedochází k trvalé a dlouhodobé inhibici cílové molekuly a je sníženo riziko vzniku nežádoucích vedlejších například imunologických účinků, které mohou stabilní komplexy vyvolat. Při reverzibilním typu inhibice tvoří inhibitor s enzymem nestabilní komplex obvykle pouze na základě nekovalentní interakce. Tato kapitola se ale zaměřuje na peptidomimetické inhibitory cysteinových katepsinů, které vykazují kovalentně reverzibilní charakter. Tyto inhibitory jsou vybaveny několika typy elektrofilních reaktivních skupin (například nitrily, ketony či azanitrily), které kovalentně reagují s thiolovou skupinou cysteinu v aktivním místě, ale následně dochází k opětovné regeneraci tohoto katalytického zbytku (Singh et al. 2011). Mechanismus probíhá dvoukrokově a je znázorněn na (Obr. 16, str. 32). Úvodní krok je totožný jako v případě ireverzibilní inhibice, kdy se tvoří nekovalentně vázaný komplex (EI). V následném kroku vzniká kovalentně vázaný komplex (E*I), u kterého může docházet k jeho rozpadu. Tento jev nastává v případě, že je rychlostní konstanta k_2 tvorby komplexu E*I přibližně rovna rychlostní konstantě jeho rozpadu k_{-2} (Silva et al. 2020).

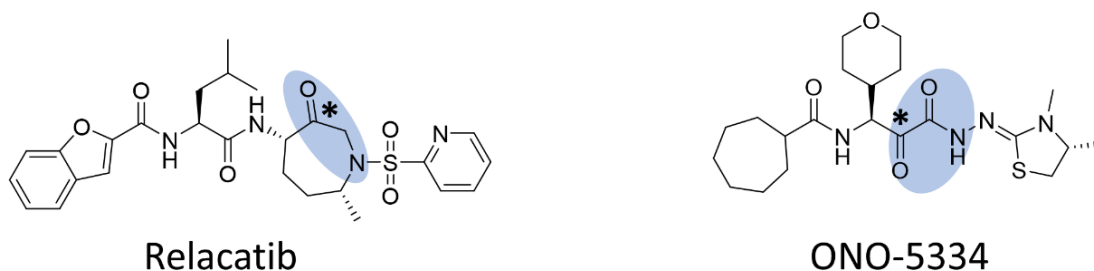


Obrázek 16: **Mechanismus reverzibilní kovalentní inhibice** Reakce probíhá ve dvou následných krocích, kdy (1) inhibitor reaguje s volným enzymem za tvorby přechodného nekovalentního komplexu EI a (2) chemickou reakcí inhibitoru s katalytickým centrem enzymu dochází ke kovalentní modifikaci enzymu za vzniku komplexu E*I. Parametry k_1 , k_{-1} , k_2 a k_{-2} jsou rychlostní konstanty příslušných kroků reakce. E je volný enzym a I je volný inhibitor (Silva et al. 2020).

Ketony: reakcí ketonů s katalytickým cysteinem enzymu vzniká reverzibilní thiohemiketalový adukt. Z inhibitorů na bázi symetrických ketonů a jejich cyklizací na tzv. azepanony vznikl relacatib (SB-462795), který byl vyvinut firmou GlaxoSmithKline a byl schopný inhibovat KatK v nanomolární koncentraci (Obr. 17, str. 33). Nicméně vykazoval nízkou selektivitu a jeho vývoj byl ukončen po I. fázi klinických studií (Kumar et al. 2007).

Firma Ono Pharmaceutical připravila vysoce selektivní inhibitor ONO-5334 (Obr. 17, str. 33), jehož vývoj byl ukončen ve III. fázi klinických studií pro své vedlejší účinky na kardiovaskulární systém. MIV-711 (L-006235) je selektivním inhibitorem od

firmy Medvir a v současné době se nachází v II. fázi klinických studií, jeho struktura nebyla zveřejněna (Wishart et al. 2006)

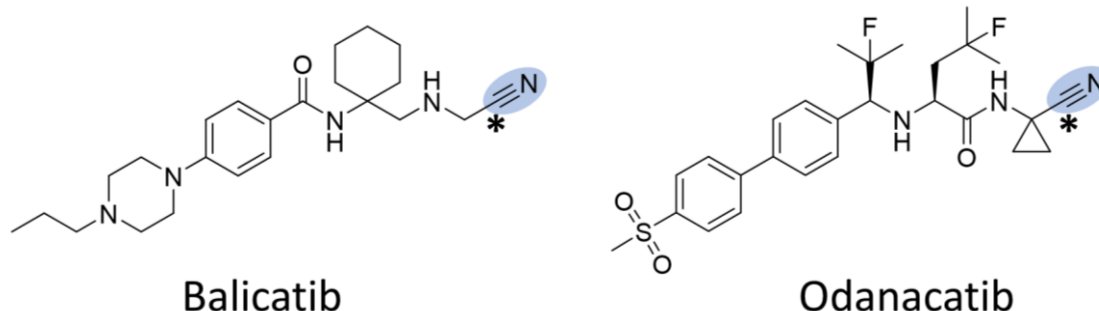


Obrázek 17: **Inhibitory Katepsinu K na bázi ketonů.** Uvedeny jsou dva příklady firemního vývoje: Relacatib obsahuje cyklický keton s heteroatomem dusíku (azepanon) jako reaktivní skupinu, ONO-5334 je cyklický α -ketoamidový inhibitor (Lu et al. 2018). Hvězdičkou je označen atom uhlíku tvořící kovalentní vazbu s atomem síry katalytického cysteinu enzymu.

Nitrily: reakcí nitrilové reaktivní skupiny inhibitoru s katalytickým cysteinem enzymu vzniká reverzibilní thioimidátový adukt. Jako první byla popsána reakce acetamidoacetonitrilu a benzamidoacetonitrilu s papainem (Lucas and Williams 1969). Systematickou analýzou substrátové specifity KatK byly určeny aminokyselinové zbytky, které se váží preferenčně do jednotlivých podmíst této proteasy.

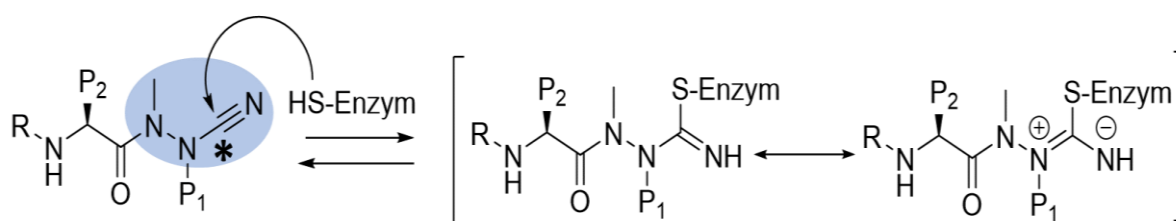
Na základě těchto poznatků vznikl peptidomimetický inhibitor balicatib (AAE581) firmy Novartis. Homocyklroleucin v pozici P2 inhibitoru a piperazinový zbytek v pozici P3 (Obr. 18, str. 40) zajišťoval selektivitu (Lu et al. 2018). Nicméně jeho vývoj byl ukončen v II. fázi klinických studií z důvodu jeho bazického a lipofilního charakteru, který způsoboval akumulaci v lysozomech (Falgueyret et al. 2005). To vedlo ke křížové inhibici katepsinů L a B s vedlejšími účinky v podobě vyrážky a změn morfologie kožní tkáně.

Odanacatib (MK-0822) firmy Merck byl metabolicky velmi stabilní peptidomimetický inhibitor, který v P1 pozici obsahoval cyklopropanový kruh, v P2 pozici fluorovaný leucinový zbytek a relativně objemný biarylový kruh v pozici P3 (Obr. 18, str. 34). Vzhledem k tomu, že nevykazoval bazický charakter, nedocházelo k jeho kumulaci v lysozomu. Vývoj tohoto inhibitoru byl ukončen ve III. fázi klinických studií pro podezření na spojitost s výskytem mozkové mrtvice u postmenopauzálních pacientek, kterým byl podáván (Drake et al. 2017; Lu et al. 2018).



Obrázek 18: **Inhibitory katepsinu K na bázi nitrilů.** Uvedeny jsou dva příklady firemního vývoje balicatib a odanacatib (Lu et al. 2018). Reaktivní nitrilová skupina je vyznačena modře a hvězdičkou je označen atom uhlíku tvořící kovalentní vazbu s atomem síry katalytického cysteinu enzymu. Balicatib obsahuje v pozici P2 cyklohexylový zbytek a v pozici P3 je piperazinový zbytek zajišťující selektivitu. Odanacatib reaktivní nitrilová skupina v pozici P1 rovněž obsahuje cyklopropanový kruh, který brání proteolytickému štěpení. V P2 pozici je fluorovaný leucinový derivát a P3 pozice je zastoupena biarylovým kruhem.

Azanitrily: izoelektrickým nahrazením CaH skupiny u těchto peptidomimetických nitrilových inhibitorů za atom dusíku vznikla nová skupina inhibitorů s tzv. azanitrilovou reaktivní skupinou (N-NH-C≡N). Ve srovnání s nitrily vykazovaly vyšší inhibici v pikomolárních koncentracích vůči katepsinům K, L a S. To je vysvětlováno reakčním mechanismem, kdy vzniká reverzibilní isothiosemikarbazidový adukt, který je rezonančně stabilizován, viz schéma (Obr. 19, str. 34) (Frizler et al. 2010; Schmitz et al. 2014). Azanitrilové peptidomimetické inhibitory vykazují obecně výborné farmakokinetické vlastnosti s prodlouženou dobou účinku a metabolickou stabilitou a patří mezi nejpokročilejší inhibitory cílené proti KatK (Proulx et al. 2011).



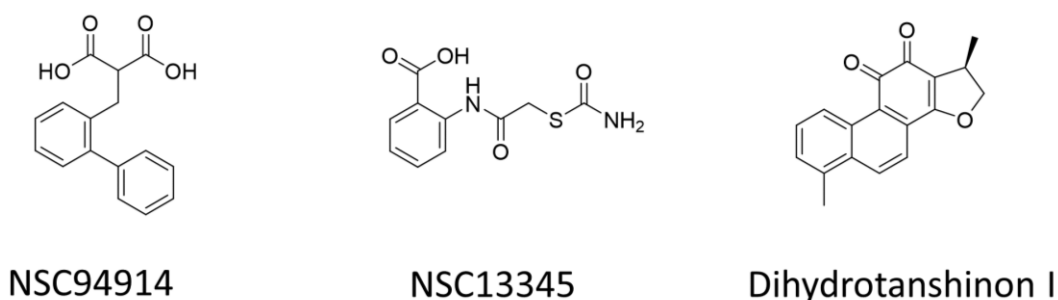
Obrázek 19: **Reakční mechanismus azanitrilových inhibitorů.** Uvedena je obecná struktura peptidomimetických inhibitorů s reaktivní azanitrilovou skupinou a zbytky v pozicích P1 a P2. Modře je zabarvena azanitrilová skupina a hvězdičkou je vyznačen atom uhlíku tvořící kovalentní vazbu s atomem síry katalytického cysteinu enzymu. Nukleofilní atakou thiolové skupiny vzniká rezonančně stabilizovaný isothiosemikarbazidový adukt. Převzato a upraveno z (Schmitz et al. 2014).

2.4.3 Nekovalentní inhibitory katepsinu K syntetického a přírodního původu

Tato kapitola se zaměřuje na nekovalentní typ inhibice s vazbou inhibitoru mimo aktivní místo enzymu. Přitom inhibiční efektor interaguje prostřednictvím slabých interakcí, jakými jsou vodíkové můstky, hydrofobní interakce či iontová vazba (Novinec et al. 2016). U KatK byla identifikována tato vazebná místa, která jsou často označována jako *exo*-místa („*exo-site*“) či alosterická místa (Novinec et al. 2010; Sharma et al. 2015). Prostřednictvím *exo*-míst KatK interaguje s některými glykosaminoglykany, čímž dochází ke tvorbě oligomerních komplexů vykazujících vysokou kolagenolytickou a elastololytickou aktivitu (Novinec et al. 2014; Lemaire et al. 2014).

Blokací *exo*-míst na povrchu KatK je následně zabráněno interakci s glykosaminoglykany, což má za následek selektivní snížení kolagenolytické a elastinolytické aktivity, za současného zachování proteolytické aktivity k jiným substrátům vstupujících do aktivního místa enzymu (Novinec et al. 2016; Rocha et al. 2020). To je zásadní rozdíl oproti inhibitorům cílených na aktivní místo, které plošně vypínají proteolytickou aktivitu enzymu, což může být v případě KatK spojeno s vedlejšími účinky (Li et al. 2002; Sharma et al. 2015; Panwar et al. 2018)

Příklady inhibitorů (Obr. 20, str. 35) cílených na *exo*-místa jsou syntetické nízkomolekulární látky jako NSC13345 a NSC94914 (Novinec et al. 2016) či lipofilní látky přírodního původu izolované z kořene byliny šalvěje červenokořenné (*Salvia miltiorrhiza*) používané v čínské medicíně na nemoci pohybového aparátu a označované jako tanshinony (Panwar et al. 2018).



Obrázek 20: Nekovalentní inhibitory Katepsinu K cílené na *exo*-místa. Synteticky připravené nízkomolekulární látky NSC13345 a NSC94914 interagují s vazebnými *exo*-místy na povrchu molekuly KatK na základě slabých interakcí s aminokyselinami argininem, lysinem a tyrosinem (PDB: 5J94). Tanshinony jsou lipofilní diterpeny přírodního původu izolované z kořene šalvěje červenokořenné, uveden je zástupce dihydrotanshinon I (Panwar et al. 2018).

3. Cíle práce

Cysteinová proteasa katepsin K (KatK) je cílovou molekulou pro léčbu osteoporózy a artritidy pomocí chemoterapeutik na bázi proteasových inhibitorů. Práce je zaměřena na identifikaci nových inhibitorů KatK a zejména na určení vazebného módu selektivních inhibitorů cílených na aktivní místo KatK. Cílem je přinést nové poznatky o vztahu mezi strukturou a biochemickou aktivitou inhibitorů, které umožní design účinných selektivních inhibitorů KatK jako potenciálních léčiv a molekulárních diagnostických sond pro detekci KatK.

Dílčí cíle práce jsou následující:

1) Připravit rekombinantní KatK a jeho jednotlivé aktivační formy. Testovat inhibiční vlastnosti studovaných inhibitorů za pomoci enzymologických metod.

2) Připravit kovalentní komplexy KatK s inhibitory, krystalizovat je a popsat rentgenostrukturální analýzou.

3) Určit vazebný mód inhibitorů podle 3D struktury komplexů. Integrovat strukturální data s výsledky získanými pomocí výpočetních metod, inhibičních analýz a buněčných testů.

4) Výše uvedeným způsobem studovat tři skupiny kovalentních inhibitorů se zaměřením na: (i) funkční interakci 1,2,4 thiadiazolů s KatK, (ii) využití vinylketonů pro konstrukci fluorescenčních sond k detekci KatK, (iii) vztah struktury a funkce cyanohydrazidů, patřících mezi nejlepší známe inhibitory KatK.

4. Materiál a metody

Obsahem této kapitoly je poskytnout stručný přehled o použitém materiálu, vybavení a metodice práce, které byly nezbytné pro vypracování této disertační práce. Detailní informace jsou uvedeny v jednotlivých přiložených publikacích.

4.1 Materiál a laboratorní vybavení

Disertační práce vznikla s využitím přístrojového vybavení laboratoří Ústavu organické chemie a biochemie AV ČR (ÚOCHB AV ČR, v. v. i) a ve spolupráci se zahraničními výzkumnými institucemi.

Plazmid pUC57 obsahující úsek DNA a kódující zymogen lidského KatK (Uniprot P43235) byl syntetizován firmou (GenScript, USA).

Sady inhibitorů poskytli M. Gütschow, Universität Bonn (Německo) a P. Majer (skupina medicínální chemie), ÚOCHB AV ČR, v. v. i. Syntetické inhibitory E-64a E-64d (Sigma-Aldrich, ČR) a K-1777 (Fisher-Scientific, USA) byly zakoupeny.

Fluorescenčně značený substrát Cbz-Gly-Pro-Arg-AMC byl zakoupen od firmy (Bachem, Švýcarsko).

Myší monoklonální protilátky proti lidskému KatK byly zakoupeny od firmy (Santa Cruz Biotechnology, USA) a sekundární kozi značené protilátky anti-mouse IgG (Sigma-Aldrich, ČR)

Buněčná linie lidského osteosarkomu (U2-OS) poskytla H. Mertlíková-Kaiserová (skupina biochemické farmakologie), ÚOCHB AV ČR, v. v. i.

Sady komerčních krystalizačních roztoků byly zakoupeny od firem Molecular Dimensions (UK) a Jena Bioscience (Německo)

Sběr difrakčních dat pro rentgenovou strukturní analýzu byl proveden na rentgenové difrakční stanici MicroMax-007 HF Microfocus (Rigaku, Japonsko) na ÚOCHB AV ČR, v. v. i. a na synchrotronu Bessy II electron storage ring (Helmholtz Zentrum, Německo).

4.2 Metody

4.2.1 Příprava rekombinantního katepsinu K

Gen kódující zymogen lidského KatK byl z plazmidu pUC57 do expresního vektoru pPICZ α A za využití dvou restrikčních míst (*XhoI* restrikční místo na 5' konci a *NotI* pro 3' konec). Expresní vektor byl linearizován restrikční endonukleázou *Sac I* a poté transformován elektroporací do kvasinky *Pichia pastoris* (kmen X-33). Pozitivní selekce klonů, byla provedena na sadě agarových ploten, s obsahem 50 μ g/ml a 100 μ g /ml antibiotika Zeocin. Selektce klonů s nejvyšší expresí rekombinantního proteinu do media byla provedena inokulací 10 kolonií do 10 ml YPD media a kultivací přes noc při 30 °C. Poté byly buňky centrifugovány (500 g, 10 min) a resuspendovány do 50 ml indukčního BMM media. Indukce exprese byla prováděna každý den přidáním metanolu do finální koncentrace 2 %, exprese probíhala jeden týden a z každého dne exprese byl odebrán vzorek media. Odebrané medium bylo analyzováno pomocí SDS-PAGE a kinetickým měřením proteolytické aktivity na přítomnost KatK. Identifikován byl klon s nejvyšší expresí KatK a optimální délka exprese (72 h). Identita rekombinantního KatK byla ověřena N-koncovým sekvenováním a mapováním peptidových fragmentů hmotnostní spektrometrií. Obdobným způsobem byla následně provedena velkoobjemová produkce rekombinantního KatK s výtěžkem přibližně 5 mg KatK na 1 l expresního média.

Expresní medium bylo zfiltrováno, lyofilizováno a rozpuštěno v 20 mM acetátovém pufru, pH 5,5 (na 10 % původního objemu). Následně byla provedena gelová filtrace na koloně Sephadex G-25 (500 ml) ekvilibrované stejným puftrem. Purifikační krok KatK byl proveden za využití ionexové chromatografie na koloně Mono S 5/50 GL (GE Healthcare Life) ekvilibrované 50 mM acetátovým puftrem, pH 5,5. Eluce byla provedena lineárním gradientem 2 M NaCl. Frakce obsahující zymogenní formu KatK byly spojeny, zakonzentrovány na koncentrátoru Amicon Ultracel-10K (Millipore) a skladovány při -80 °C. Konverze zymogenu na maturní formu KatK byla provedena autoaktivací v prostředí 100 mM acetátového pufru; 150 mM NaCl, pH 4 po dobu 1 h. Proces autoaktivece byl monitorován pomocí SDS-PAGE a kinetickým měřením proteolytické aktivity.

4.2.2 Biochemické metody

V disertační práci byly použity následující hlavní biochemické metody: elektroforetická separace proteinů pomocí SDS-PAGE, detekce KatK na polyvinylidfluoridové membráně metodou western blot pomocí specifických monoklonálních protilátek, analýza N-koncové

sekvence proteinů modifikovanou metodou Edmanova odbourávání s následnou RP-HPLC separací, stanovení koncentrace proteinu, peptidových substrátů a inhibitorů aminokyselinovou analýzou.

4.2.3 Enzymologické metody

V disertační práci byly použity následující enzymologické metody: měření aktivity KatK pomocí fluorogenního peptidového substrátu Cbz-Gly-Pro-Arg-AMC na fluorescenční čtečce (Tecan). Stanovení kinetických parametrů pro substrát (K_m) a pro inhibitory (IC_{50} , K_i , k_{on} , k_{off}). Kinetické parametry byly získány nelineární regresí v programech GraFit či GraphPad. Koncentrace aktivního KatK byla stanovena metodou titrace aktivního místa (Horn et al. 2009) ireverzibilním inhibitorem E-64 (Barrett et al. 1982).

4.2.4 Metody rentgenové krystalografické analýzy

Komplexy KatK s inhibitory byly pro krystalizační pokusy připraveny následujícím způsobem: purifikovaný KatK byl inkubován s nejméně pěti molárním nadbytkem příslušného inhibitoru v prostředí 100 mM acetátového pufru; 2,5 mM DTT; 300 mM NaCl; 1 mM EDTA; pH 4 po dobu 3 h při 4 °C. Z důvodů možné oxidace thiolové skupiny katalytického cysteinu KatK byla inhibiční reakce provedena za redukčních podmínek (2,5 mM DTT) a v inertní atmosféře argonu. Získané komplexy enzym-inhibitor byly následně a převedeny do pufru o složení: 20 mM acetátový pufr; 2,5 mM DTT; 300 mM NaCl, pH4 a zakoncentrovány na 2,5-6,0 mg/ml na koncentrátoru Amicon Ultracel-10K (Millipore). Primární krystalizační podmínky byly určeny metodou sedící kapky v 96 jamkové destičce s využitím sady komerčních krystalizačních roztoků a automatizovaného nanášecího systému Oryx (DouglasInstruments). Primární krystalizační podmínky byly následně optimalizovány metodou visící kapky, která poskytla krystaly pro rentgenovou strukturní analýzu.

Difrakční data byla sbírána na synchrotronu BESSY II (Helmholtz Zentrum, Německo) a na rentgenové difrakční stanici MicroMax-007 HF Microfocus (Rigaku). Difrakční data byla procesována programem XDSGUI (Kabsch 2010) a pro vyřešení 3D struktury komplexů byly použity programy z balíčku CCP4 (Winn et al. 2011). Knihovna geometrických a sterických charakteristik pro ligandy byla konstruována programem AceDRG (Long et al. 2017). Vyřešené 3D struktury komplexů byly graficky editovány a prezentovány s využitím programu PyMol (DeLano 2002). Analýza vazebného módu

inhibitorů a charakterizace interagujících zbytků enzymu byla provedena programy CONTACT a PLIP (Winn et al. 2011; Salentin et al. 2015).

5. Výsledky

5.1 Přehled výstupů dizertační práce

Publikace č. 1

1,2,4-Thiadiazole acyclic nucleoside phosphonates as inhibitors of cysteine dependent enzymes cathepsin K and GSK-3 β

Pomeislová A., Otmar M., Rubešová P., **Benýšek J.**, Matoušová M., Mertlíková-Kaiserová H., Pohl R., Poštová Slavětínská L., Pomeisl K., Krečmerová M., (2021) *Bioorganic and Medicinal Chemistry*, 32, 115998, DOI: 10.1016/j.bmc.2021.115998

Předmětem publikace je organická syntéza a biochemická aktivita nových sloučenin s reaktivní 1,2,4-thiadiazolovou skupinou zahrnující inhibici enzymové aktivity KatK.

Můj podíl na publikaci spočíval v přípravě a optimalizaci inhibičního testu pro KatK, určení inhibičního parametru IC₅₀ pro sadu derivátů obsahujících 1,2,4-thiadiazolovou skupinu a interpretaci strukturně-funkčních vztahů.

Publikace č. 2

An Activity-Based Probe for Cathepsin K Imaging with Excellent Potency and Selectivity

Lemke C., **Benýšek J.**, Brajtenbach D., Breuer Ch., Jílková A., Horn M., Buša M., Ulrychová L., Illies A., Kubatzky K. F., Bartz U., Mareš M., Gütschow M. (2021) *Journal of Medicinal Chemistry*, 64, 13793–13806. DOI: org/10.1021/acs.jmedchem.1c01178

Publikace popisuje přípravu unikátních fluorescenčních inhibičních sond na bázi peptidomimetických vinyl ketonů, umožňujících selektivní značení KatK. Analyzován byl strukturní mechanismus interakce sond a vztah mezi jejich strukturou a biochemickou aktivitou. Funkční vlastnosti sond byly testovány *in vitro* a na buněčné linii lidského osteosarkomu. Publikace přináší nové poznatky pro vývoj molekulárních sond detekujících KatK a jejich biomedicínskou aplikaci.

Můj podíl na publikaci spočíval v přípravě komplexů KaK s inhibičními sondami, jejich krystalizaci a řešení 3D struktur rentgenostrukturní analýzou. Dále jsem analyzoval vazebný mód sond a strukturně funkční vztahy a podílel se na přípravě manuskriptu a grafické prezentaci dat.

Publikace č. 3

Highly potent inhibitors of cathepsin K with a differently positioned cyanohydrazide warhead: structural analysis of binding mode to mature and zymogen-like enzymes

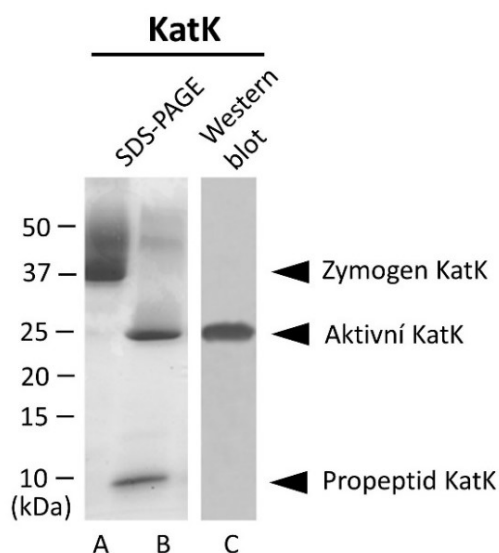
Benýšek J., Buša M., Rubešová P., Fanfrlík J., Lepšík M., Brynda J., Matoušková Z., Bartz U., Horn M., Gütschow M., Mareš M. (2022) *Journal of Enzyme Inhibition and Medicinal Chemistry* (accepted a momentálně v tisku). DOI:10.1080/14756366.2021.2024527

Publikace se zabývá strukturní analýzou interakce KatK se dvěma odlišnými typy vysoce účinných a selektivních peptidomimetických inhibitorů s reaktivní cyanohydrazidovou skupinou. Poprvé byla studována inhibice KatK na dvou různých úrovních. Jednak aktivního maturního enzymu a jednak aktivačního intermediátu, ze kterého vzniká maturní enzym. Za využití rentgenostrukturní analýzy, enzymových kinetik a výpočetních metod byl popsán vazebný mód těchto inhibitorů a objasněn výjimečný inhibiční potenciál, který z nich činí jedny z nejúčinnějších, dosud známých inhibitorů KatK.

Můj podíl na publikaci zahrnoval přípravu komplexů KatK s inhibitory, jejich krystalizaci a řešení 3D struktur rentgenostrukturní analýzou. Dále jsem analyzoval vazebný mód inhibitorů, jejich biochemické funkční vlastnosti a strukturně funkční vztahy. Jako první autor jsem se vypracoval manuskript a grafickou prezentaci dat.

5.2 Příprava rekombinantního katepsinu K

Rekombinantní lidský KatK byl připraven pomocí expresního systému v kvasinkách *P. pastoris* a izolován s využitím ionexové chromatografie s výtěžkem přibližně 5 mg/l expresního média. Rekombinantní protein byl vizualizován pomocí SDS-PAGE (Obr. 21, dráha A, str. 43), kde je patrný pás proteinu ve formě zymogenu (37 kDa). Byl optimalizován proces autoaktivace v kyselém prostředí; při pH 4 je během 45 min neaktivní zymogen autokatalytickou proteolýzou konvertován na svoji aktivní maturní formu (23,5 kDa) a současně lze pozorovat odštěpený propeptid (11 kDa) (Obr. 21, dráha B, str. 43). Dále byla provedena vizualizace metodou Western blot s použitím myši monoklonální protilátky rozpoznávající maturní formu KatK (Obr. 21, dráha C, str. 43). Detaily přípravy preparátu rekombinantního KatK použitého pro další experimentální práci jsou podrobně uvedeny v příložené autorské **publikaci č. 3** (sekce materiál a metody).



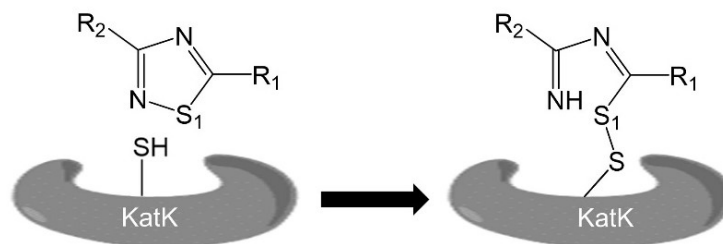
Obrázek 21: **Příprava rekombinantního katepsinu K.** Dráha A znázorňuje separaci purifikovaného produktu rekombinantní exprese pomocí SDS-PAGE, kterým je zymogen KatK (37 kDa). V dráze B je výsledek kyselé autoaktivace zymogenu, kdy vzniká maturní forma KatK (23,5 kDa) odštěpením propeptidové části (11 kDa). V dráze C je k vizualizaci autoaktivační směsi použita myšší monoklonální protilátka rozpoznávající maturní formu KatK.

5.3 Interakce katepsinu K s 1,2,4-thiadiazolovými inhibitory

V této kapitole jsou prezentovány výsledky autora, které jsou součástí **publikace č. 1**, která jako první popisuje inhibiční vlastností sloučenin s 1,2,4-thiadiazolovým heterocyklem vůči lidskému KatK.

5.3.1 Inhibiční vlastnosti sloučenin s 1,2,4-thiadiazolovým cyklem

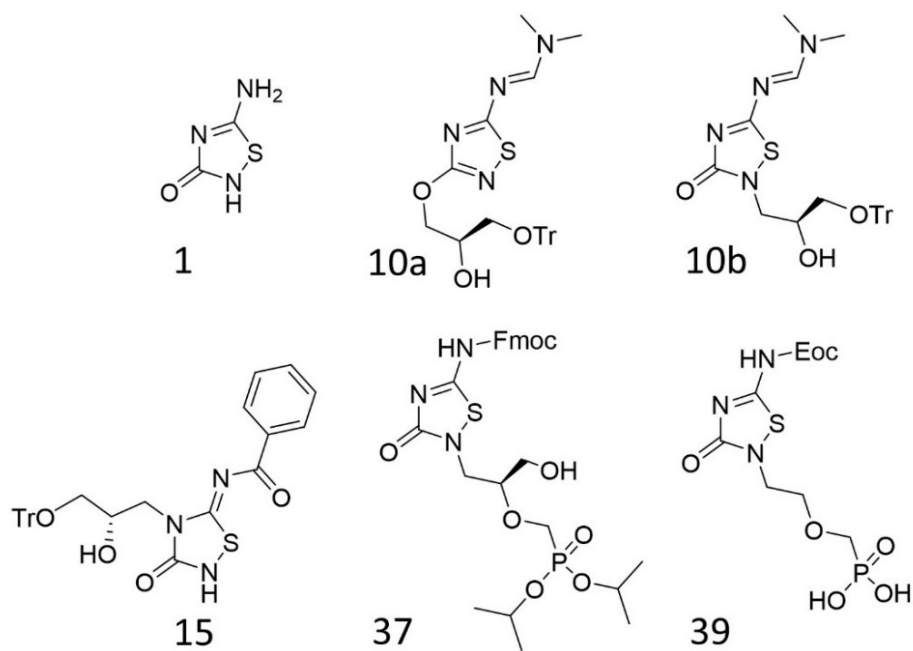
V **publikaci č. 1** byla poprvé popsána inhibiční schopnost látek obsahujících 1,2,4-thiadiazolovou skupinu vůči lidskému KatK. Navržený mechanismus inhibice spočívá v nukleofilním ataku thiolové skupiny katalytického cysteinu KatK na reaktivní heterocyklus inhibitoru. Následuje otevření thiadiazolového kruhu a tvorba kovalentního a ireverzibilního disulfidického aduktu podobně jak bylo postulováno pro jiné cysteinové katepsiny (Vega-Tejido et al. 2014) (Obr. 22, str. 44).



Obrázek 22: *Mechanismus inhibice katepsinu K thiadiazolovými deriváty.* Thiolová skupina katalytického cysteinu KatK tvoří kovalentní ireverzibilní adukt s reaktivní thiadiazolovou skupinou inhibitoru. Otevřením thiadiazolového kruhu vzniká disulfidové spojení mezi inhibitorem a enzymem tzv. „thiol trapping“ mechanismem.

Z vybrané sady 21 sloučenin (Tab. 1, str. 46) bylo v primárním optimalizovaném kinetickém inhibičním testu s KatK a fluorogenním substrátem identifikováno šest sloučenin (označeny červeně) s potenciálem inhibice pod hranici 50 % zbytkové enzymatické aktivity. Tyto deriváty vykazovaly hodnotu IC_{50} v mikromolárních koncentracích. Ze strukturního hlediska se dají rozdělit na 3 strukturní podtypy: a) Nesubstituovaná sloučenina 5-amino-1,2,4-thiadiazol-3(2H)-on s kódovým označením **1** ($IC_{50} = 6,34 \pm 0,63 \mu M$) b) Substituované sloučeniny na dusíku N5 dimethylaminomethylenem **10a** ($IC_{50} = 2,63 \pm 0,31 \mu M$), **10b** ($IC_{50} = 10,08 \pm 0,95 \mu M$) s rozdílnou polohou postranní hydroxy-3-trityloxypropylové skupiny spolu s látkou **15** ($IC_{50} = 14,9 \pm 0,75 \mu M$) c) Sloučeniny s chránicí skupinou na dusíku N5 a substituované na N2, zbytkem kyseliny fosfonové a jejích derivátů kódované **37** ($IC_{50} = 9,25 \pm 1,46$) a **39** ($IC_{50} = 10,15 \pm 1,48$). Vzorce šesti nejlepších thiadiazolových derivátů jsou znázorněny na (Obr. 23, str. 45)

Obecné vztahy, vyvozené z inhibiční studie ukazují, že látky substituované v pozici C2 derivátem kyseliny fosfonové (deriváty kódované **8**, **32** a **37**) jsou lepšími inhibitory oproti sloučeninám, které nejsou takto esterifikovány (**9**, **38** a **43**). Dále platí, že substituce v pozici C5 objemnější chránicí skupinou Fmoc (fluorenylmethoxykarbonyl) je pro inhibici lepší (**37**, **40** a **43**) oproti méně objemné fenylové skupině (**38**, **35** a **41**). Veškeré strukturní vzorce inhibitorů jsou k dispozici v příložené publikaci č. 1.



Obrázek 23: **Struktura thiadiazolových derivátů s nejvyšším inhibičním potenciálem.** Nesubstituovaná základní sloučenina 5-amino-1,2,4-thiadiazol-3(2H)-on (**1**), Substituované sloučeniny na dusíku N5 a rozdílnou polohou postranní hydroxy-3-trityloxypropylové skupiny (**10a**, **10b** a **15**), Sloučeniny s chránicí skupinou na N5 a substitucí zbytkem kyseliny fosfonové a jejich derivátů na N2 (**37**, **39**). Fmoc-fluorenylmethoxykarbonyl, Eoc-ethoxykarbonyl, OTr-trytyloxy.

Tabulka 1: *Inhibice aktivity lidského KatK sadou sloučenin se substituovanou 1,2,4-thiadiazolovou skupinou při 10 μ M koncentraci inhibitorů. Inhibice je vyjádřena relativní zbytkovou aktivitou (%) oproti neinhibované kontrole (100 %). Inhibitory jsou seřazeny v pořadí od nejúčinnějších s nejvyšším inhibičním potenciálem. Aktivita KatK byla měřena pomocí peptidového fluorogenního substrátu Cbz-Gly-Pro-Arg-AMC.*

Kódové označení derivátů	Zbytková aktivita (%)
10a	5,3 \pm 1,5
1	16,9 \pm 1,0
10b	28,3 \pm 1,7
37	34,1 \pm 2,4
15	34,2 \pm 3,0
39	39,1 \pm 1,0
7	52,9 \pm 3,8
6	67,0 \pm 5,8
40	67,0 \pm 0,6
14a	76,6 \pm 3,8
8	78,3 \pm 2,6
32	80,6 \pm 5,0
36	83,4 \pm 8,6
14b	85,0 \pm 3,5
43	85,5 \pm 5,3
33	86,8 \pm 2,7
9	90,0 \pm 3,5
38	90,3 \pm 6,0
41	92,3 \pm 4,5
35	92,9 \pm 2,8
42	97,5 \pm 2,8

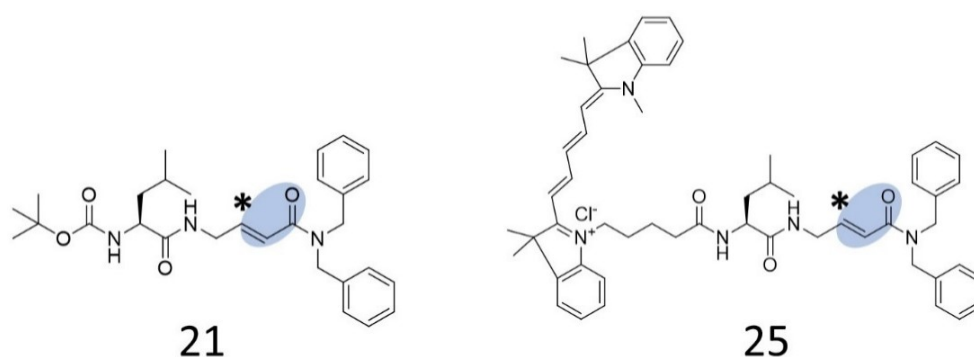
5.4 Interakce katepsinu K s vinylketonovými inhibitory

V této kapitole jsou prezentovány výsledky autora, které jsou součástí **publikace č. 2**, která popisuje přípravu, vazebný mód a použití fluorescenčních inhibičních sond na bázi peptidomimetických vinyl ketonů:

5.4.1 Selektivita vinylketonových inhibitorů katepsinu K

Navrženy a syntetizovány byly dva peptidomimetické vinylketonové inhibitory na bázi Michaelových akceptorů a ze struktury těchto látek byly připojením dvou typu fluorescenčních značek odvozeny inhibiční sondy klasifikované jako ABP („activity-based

probes“). Vzorce inhibitorů a sond jsou uvedeny v publikační příloze (**publikace č. 2**, str. 109). Na základě *in vitro* testování inhibiční aktivity proti čtyřem modelovým lidským katepsinům byly vybrány sloučeniny **21** a **25** (Tab. 2, str. 48) pro následnou rentgenostrukturní analýzu a identifikaci vazebného módu. Z (Tab. 2, str. 48) je patrné, že prekurzorová sloučenina **21** vykazuje dostatečnou selektivitu vůči KatK. Katepsin S je inhibován 36x nižší rychlostní inhibiční konstantou druhého řádu (k_{inact}/K_i) oproti KatK a inhibice katepsinu L a B je velmi nízká. Zavedením fluorescenční značky do struktury **21** v podobě pigmentu cyaninu 5 ($\lambda = 600$ nm) byla odvozena fluorescenční sonda **25** (Obr. 24, str. 47). Sonda **25** vykazovala ve funkčních testech ještě vyšší inhibici KatK než její prekurzor **21** při zachování obdobné selektivity.



Obrázek 24: **Struktury vinylketonových inhibitorů katepsinu K.** Prekurzorová látka s označením **21** je selektivním inhibitorem KatK a látka s označením **25** je fluorescenčně značená sonda z ní odvozená. Oba ligandy mají v P1' pozici dibenzylaminovou skupinu následovanou reaktivní vinylketonovou skupinou v P1 pozici (vyznačena modře). Hvězdičkou je označen atom uhlíku tvořící kovalentní vazbu s atomem síry katalytického cysteinu enzymu. Prekurzorová látka **21** obsahuje v P3 pozici chránící tert-butyloxykarbonylovou skupinu, zatímco sonda **25** obsahuje v tomto místě fluorofor cyanin 5.

Tabulka 2: **Inhibice lidských modelových katepsinů vinylketonovými inhibitory.** Z prekurzorového inhibitoru **21** byla odvozena fluorescenčně značená inhibiční sonda **25**. Inhibice je vyjádřena pomocí rychlostní inhibiční konstanty druhého řádu a relativních zbytkových aktivit enzymu.

Látka	k_{inact}/K_i ($M^{-1}s^{-1}$) ^a a zbytková aktivita enzymu (%) ^b			
	katepsin K	katepsin S	katepsin L	katepsin B
21	10 600 $M^{-1}s^{-1}$	291 $M^{-1}s^{-1}$	83%	91%
25	35 300 $M^{-1}s^{-1}$	505 $M^{-1}s^{-1}$	89%	96%

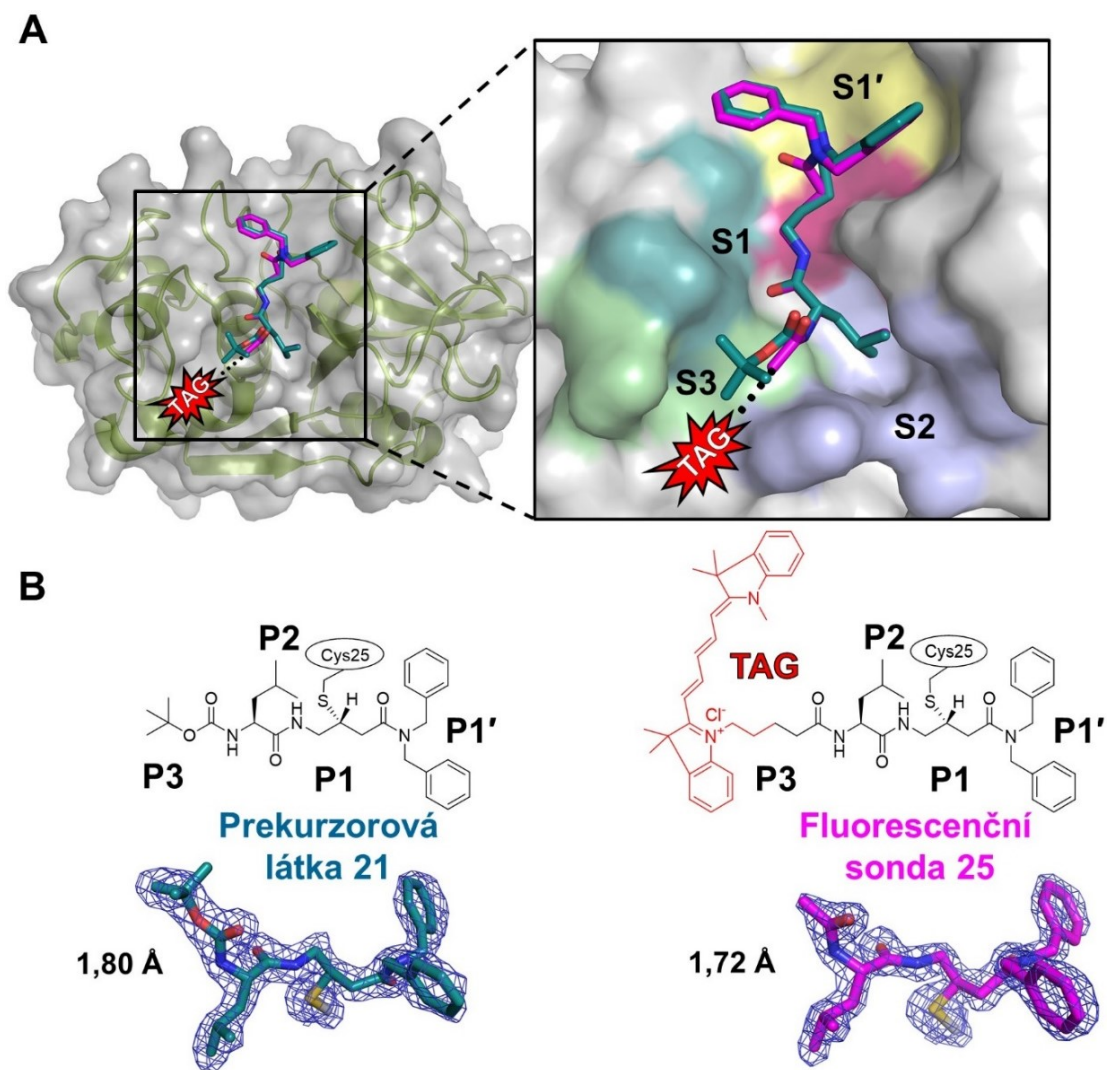
^a rychlostní inhibiční konstanta druhého řádu odpovídá poměru rychlostní konstanty inaktivace k_{inact} a inhibiční konstanty K_i . ^b Relativní hodnota zbytkové enzymové aktivity při koncentraci inhibitoru 2,5 μM (neinhibovaná kontrola odpovídá 100 %).

5.4.2 Krystalografická analýza komplexů vinylketonových inhibitorů a určení vazebného módu

Studovány byly dva vinylketonové inhibitory: prekurzorová sloučenina **21** s chránící skupinou tert-butyloxykarbonylem v pozici P3 a odvozená fluorescenčně značená sonda **25** se substitucí fluoroforem cyaninem 5 v pozici P3 (Obr. 25B, str. 50). Oba inhibiční ligandy byly vykrystalizovány v komplexu s rekombinantně připravenou maturní formou lidského KatK (komplexy KatK-**21** a KatK-**25**). Optimalizací podmínek byly získány krystaly s ortorombickou symetrií a prostorovou grupou P2₂1₂1. Difrakční data byla zpracována programem XDSGUI a pro vyřešení 3D struktury komplexů byly použity programy z balíčku CCP4 metodou molekulárního nahrazení. Finální struktury jsou deponované v databázi PDB s přístupovými kódy 7NXL pro strukturu KatK-**21** v rozlišení 1,80 Å a 7NXM pro komplex KatK-**25** s rozlišením 1,72 Å. Vyřešené 3D struktury obsahovaly maturní formu KatK o 215 aminokyselinových zbytcích (číslovány dle maturní formy Ala1 až Met215) s aminokyselinou argininem (Arg0) na N-konci polypeptidového řetězce navíc, která pochází z propeptidové části enzymu.

Analýza vazebného módu inhibitorů a charakterizace interagujících postranních zbytků enzymu s ligandem byla provedena programy CONTACT a online programem PLIP. Krystalografická analýza komplexů prokázala kovalentní typ vazby mezi katalytickým Cys25 a β -atomem karbonylu reaktivní skupiny ligandu (Obr. 25B, str. 50), což je v souladu s reakčním mechanismem na principu Michaelovy nukleofilní adice.

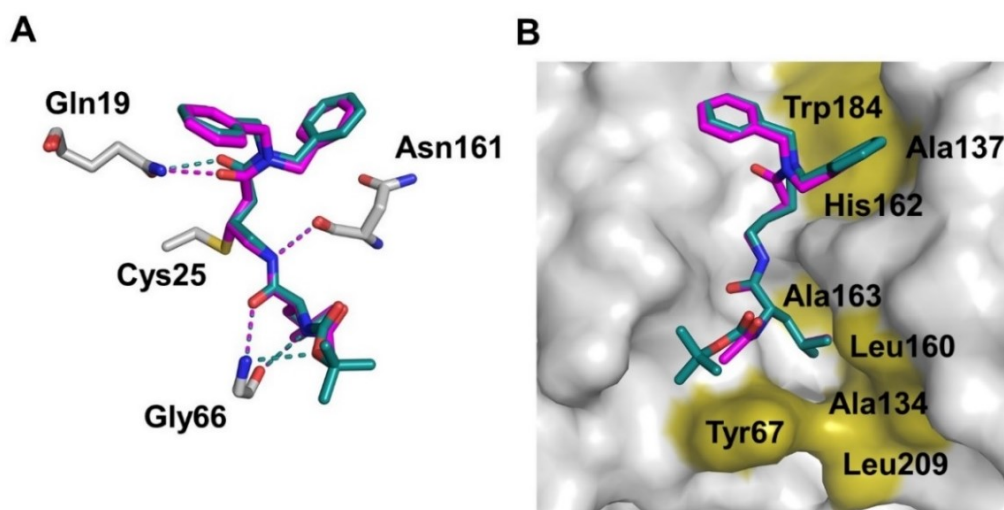
S výjimkou nízké hodnoty hustoty elektronové mapy pro oblast fluoroforu cyaninu 5 u sondy **25**, spojeného zbytkem kyseliny valerové, byly všechny ostatní strukturní elementy obou inhibitorů jasně zřetelné (Obr. 25B, str. 57). Z tohoto důvodu nebyl fluorofor modelován ve finální 3D struktuře a současně toto zjištění nasvědčuje zvýšenou flexibilitu této části s orientací ven z aktivního místa.



Obrázek 25: **Superpozice vinylketonových ligandů 21 a 25 v aktivním místě katepsinu K.** A) Levá část obrázku znázorňuje KatK ve stužkovém modelu s transparentním povrchem. Inhibiční ligandy jsou v tyčinkovém modelu, kdy uhlíkové atomy prekursoru 21 jsou znázorněny tyrkysově a u fluorescenční sondy 25 purpurově. Heteroatomy jsou vyznačeny ve standardním barevném kódování (O: červená, N: modrá, S: žlutá). Součástí sondy 25 je fluorofor cyanin 5; mapa elektronových hustot v této části vykazovala sníženou kvalitu, tudíž tento segment nebyl modelován (jeho pozice je vyznačena pomocí „tag“). Ligandy jsou navázány v orientaci jako u substrátu enzymu a vyplňují vazebná podmístia S3 až S1'. Povrchy těchto podmíst jsou barevně odlišeny (katalytické zbytky enzymu Cys25 a His162 jsou barveny růžově). B) Strukturní vzorce inhibičních ligandů (ve vázané formě na katalytický Cys25) a finální rozlišení krystalografických struktur (Å) s mapou elektronových hustot ligandu $2F_o - F_c$ konturovaných na 1σ .

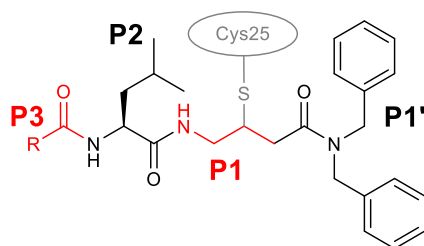
Aktivní místo KatK obsahuje katalytickou dyádu Cys25 a His162 s pomocnou aminokyselinou Asp182. Oba inhibiční ligandy se vážou do aktivního místa v orientaci analogické jako u substrátu a vyplňují jeho vazebná podmísta S3 až S1' (Obr. 25A, str. 50). Dibenzylaminová skupina ligandu v pozici P1' je umístěna obecně v podmístě S1', nicméně jeden z benzylových kruhů je namířen opačným směrem do podmísta S2'. Stabilizaci ligandů v aktivním místě enzymu zprostředkovává síť vodíkových můstků tvořená mezi kostrou ligandu a postranními řetězci aminokyselin v aktivním místě enzymu (Obr. 26A str. 52). Hlavní interakce jsou následující: zbytek Glu19 má klíčovou úlohu při stabilizaci meziprojektu enzymové katalýzy prostřednictvím tzv. oxyaniontové jámy a u obou ligandů tvoří vodíkový můstek s atomem kyslíku reaktivní skupiny, která je v sousedství s dibenzylaminovou skupinou ligandu. Zbytek Asp161 vstupuje do vodíkového můstku atomem kyslíku z karbonylu peptidové vazby a tvoří jej s amidickým dusíkem ligandu v oblasti pozice P1. Zbytek Gly66 přispívá k vodíkovým můstkům na jedné straně svým atomem dusíku a tvoří interakci s kyslíkem karbonylové skupiny leucinu v pozici P2 ligandu a zároveň tvoří interakci s atomem kyslíku karbamátu chránící skupiny (označen hvězdičkou -NH-(C=O)-O*-) na prekurzoru **21** v pozici P3. Gly66 je schopen tvořit i další vodíkový můstek svým atomem kyslíku s dusíkovým atomem karbamátu (-N*H-(C=O)-O-) v pozici P3. V celkové bilanci Gly66 tvoří tři vodíkové můstky s prekurzorem **21** a dva vodíkové můstky se sondou **25**.

Analýza interakcí obou ligandů s KatK prokázala dále i přítomnost hydrofobních interakcí (Obr. 26B, str. 59) které představují další typ nekovalentní interakce. Ty jsou definovány jako interakce nepolárních postranních řetězců vedlejších aminokyselinových zbytků enzymu s nepolárními atomy ligandu do vzdálenosti 4,1 Å. Účastní se jich celkem osm zbytků z aktivního místa KatK. Veškeré interakce ligandů **21** a **25** s KatK jsou shrnuty a uvedeny v přehledné (Tab. 3, str. 53)



Obrázek 26: Síť interakcí tvořených mezi vinylketonovými ligandy 21 a 25 a katepsinem K. Ligandy v superpozici jsou v tyčinkovém modelu a barevné kódování je analogické viz (Obr. 25, str. 50.) A) Přerušovanou čarou jsou znázorněny vodíkové můstky mezi ligandem a aminokyselinovými zbytky v aktivním místě KatK (tyčinkový model je barven šedě se standardním barevným kódováním pro heteroatomy). B) Hydrofobní interakce mezi ligandem a aminokyselinovými zbytky v aktivním místě KatK. Povrchový model enzymu je barven šedě a interagující zbytky jsou označeny žlutě.

Tabulka 3: Souhrn všech kontaktu tvořených mezi aminokyselinovými zbytky KatK a vinylketonovými ligandy 21 a 25. Analýza interakcí byla provedena v programu CONTACT. Struktury ligandů byly rozděleny na jednotlivé pozice s označením P3 až P1'; pro každou pozici jsou uvedeny příslušné kontaktní zbytky enzymu. Parametr s označením C představuje součet všech kontaktů daného zbytku enzymu s ligandem (včetně vodíkových můstků). Počet vodíkových můstků udává parametr Hb (zbytky enzymu, které tvoří vodíkové můstky, jsou uvedeny tučně). Pro určení kontaktů byly nastaveny vzdálenosti do 4,1 Å a pro vodíkové můstky vzdálenosti do 3,3 Å (Hb).



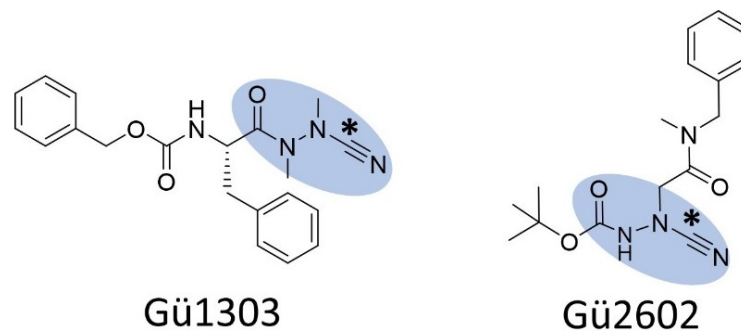
Pozice	Prekurzorová látka 21			Fluorescenční sonda 25		
	Aminokyselina	C	Hb	Aminokyselina	C	Hb
P1'	Gln 19	6	1	Gln19	5	1
	Gly 20	1		Gly 20	1	
	Gln 21	1		Gln 21	2	
	Cys 22	8		Cys 22	10	
	Gly 23	4		Gly 23	7	
	Cys 25	4		Cys 25	3	
	Ala 137	4		Ala 137	5	
	Ser 138	4		Ser 138	3	
	Gln 143	2		Gln 143	2	
	Asn 161	3		Asn161	6	
	His162	8		His162	4	
	Trp 184	4		Trp 184	3	
P1	Gln19	1		Gln 19	1	
	Gly 23	6		Gly 23	6	
	Cys 25	6		Cys 25	7	
	Asn 161	4	1	Asn 161	4	1
	His 162	1		His 162	1	
P2	Gly23	1		Gly23	1	
	Cys25	2		Cys25	3	
	Trp 26	4		Trp 26	4	
	Gly 65	3		Gly 65	3	
	Gly 66	12	2	Gly 66	13	2
	Tyr 67	1				
	Met 68	2		Met 68	2	
	Ala 134	2		Ala 134	2	
	Leu 160	3		Leu 160	3	
	Asn 161	5		Asn 161	4	
	His 162	1		His 162	1	
	Ala 163	1		Ala 163	1	
Leu 209	1		Leu 209	1		
P3	Asp 61	2				
	Gly 64	2				
	Gly 65	5				
	Gly 66	7	1	Gly 66	4	
	Tyr 67	2		Tyr 67	3	

5.5 Interakce katepsinu K s cyanohydrazidovými inhibitory

V této kapitole jsou prezentovány výsledky autora, které jsou součástí **publikace č. 3**, která popisuje strukturní analýzu interakcí KatK se dvěma vzájemně odlišnými chemotypy peptidomimetických inhibitorů s reaktivní cyanohydrazidovou skupinou. Poprvé zde byla studována inhibice KatK u plně aktivní maturní formy enzymu a současně i u jeho aktivačního intermediátu.

5.5.1 Cyanohydrazidy jako účinné a selektivní inhibitory aktivity KatK a procesu aktivace zymogenu KatK

Cyanohydrazidové inhibitory mají své chemické označení podle společného segmentu struktury, který nese reaktivní skupinu, a patří obecně mezi látky s triviálním označením azanitrily. Tato kapitola se zabývá dvěma konkrétními chemotypy cyanohydrazidových peptidomimetik, které jsou nazývány azadipeptidové nitrily a 3-cyano-3-aza- β -aminokyseliny (Obr. 27, str. 55). Představují nový typ kovalentních reverzibilních inhibitorů, které mají vysokou účinnost a selektivitu a vykazují dobré farmakologické vlastnosti (Löser et al. 2008; Frizler et al. 2011; Schmitz et al. 2014; Chingle et al. 2017). Pro tuto studii byli vybráni zástupci obou chemotypů cyanohydrazidů, které mají rozdílnou polohou reaktivní skupiny. Derivát **Gü1303** je azadipeptidový nitril s terminální cyanohydrazidovou skupinou (Frizler et al. 2011) a derivát **Gü2602** náležící mezi 3-cyano-3-aza- β -aminokyseliny má reaktivní skupinu umístěnou v molekule centrálně (Schmitz et al. 2014).



Obrázek 27: *Peptidomimetické inhibitory s cyanohydrazidovou reaktivní skupinou. Inhibitor Gü1303 patří mezi azadipeptidové nitrily a Gü2602 mezi 3-cyano-3-aza-β-aminokyseliny. Modře je zabarvená reaktivní cyanohydrazidová skupina a hvězdičkou je vyznačen atom uhlíku tvořící kovalentní isothiosemikarbazidový adukt s atomem síry katalytického cysteinu KatK. Pozice reaktivní skupiny je u Gü1303 v terminální poloze, zatímco u Gü2602 v centrální části molekuly.*

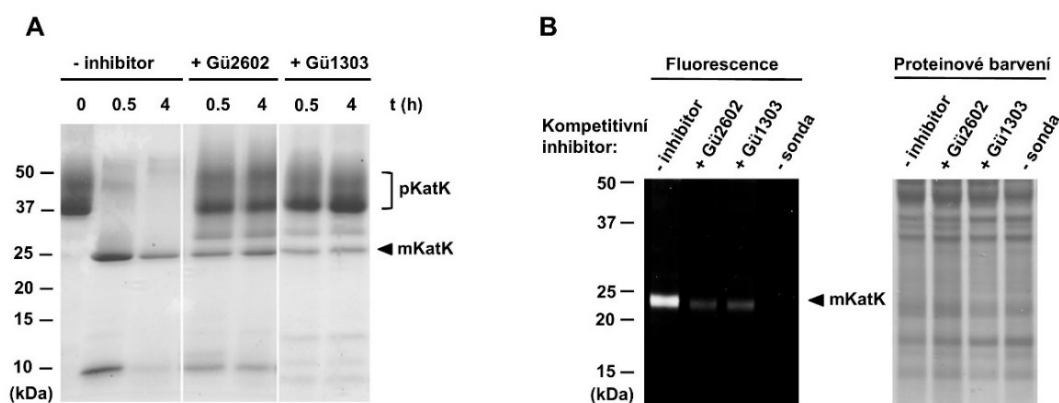
Kinetická inhibiční měření s fluorogenním substrátem (Cbz-Gly-Pro-Arg-AMC) a maturní formou KatK (v této kapitole bude označována mKatK) prokázala výjimečnou inhibiční účinnost studovaných cyanohydrazidových inhibitorů. Pro **Gü1303** byla určena hodnotou inhibiční konstanty $K_i = 0,91$ nM a téměř o dva řády nižší hodnotu $K_i = 0,013$ nM měl **Gü2602**, což z něj činí jeden z nejsilnějších doposud známých inhibitorů KatK (Tab. č. 4, str. 55). Dále analýza kinetických dat ukázala, že v případě **Gü1303** je patrná časová závislost typická pro tzv. pomalu vázající inhibitory (“slow-binding”) s asociační rychlostní konstantou $k_{on} = 527 \cdot 10^3 \text{ M}^{-1} \text{ s}^{-1}$, zatímco v případě **Gü2602** se jedná o rychle vázající inhibitor (“fast-binding”), kde hodnotu k_{on} nelze z principu stanovit.

Tabulka 4: *Inhibice aktivity maturního KatK (mKatK) pomocí cyanohydrazidů Gü2602 a Gü1303. Inhibice byla měřena v aktivním testu s fluorogenním substrátem Cbz-Gly-Pro-Arg-AMC při pH 5,5 a hodnoty kinetických parametrů byly určeny z grafů generovaných programem GraFit. Technické detaily jsou uvedeny v příložené autorské publikaci č. 2. n.d. – hodnotu parametru nelze stanovit.*

Inhibitor	Inhibice mKatK		
	K_i (nM)	k_{on} ($10^3 \text{ M}^{-1} \text{ s}^{-1}$)	k_{off} (10^{-3} s^{-1})
Gü2602	$0,013 \pm 0,008$	n.d.	n.d.
Gü1303	$0,91 \pm 0,04$	527 ± 8	$0,480 \pm 0,028$

Účinnost inhibitorů **Gü2602** a **Gü1303** při interakci s KatK byla dále testována na buněčném modelu pomocí kultury buněk lidského osteosarkomu (linie U-2 OS), které KatK exprimují. KatK byl monitorován s využitím fluorescenčně značené aktivní sondy **25** (z publikace č. 2), která se váže do aktivního místa KatK, a vizualizován po separaci na SDS-PAGE. Bylo prokázáno, že oba cyanohydrazidové inhibitory jsou schopny proniknout přes buněčnou membránu až do lysozomálního/endozomálního systému a cíleně se vázat do aktivního místa KatK, čímž kompetují se sondou **25** (Obr. 28B, str. 64).

Dále byl studován vliv inhibitoru **Gü2602** a **Gü1303** na proces autokatalytické aktivace zymogenu KatK při nízkém pH, při kterém je generován maturní KatK. Bylo pozorováno, že v přítomnosti cyanohydrazidových inhibitorů dochází k výraznému potlačení procesu aktivace zymogenu oproti neinhibované kontrole, která za stejných podmínek vedla k plné konverzi na maturní enzym (24 kDa) za současného oddělení propeptidové domény (10 kDa) (Obr. 28A, str. 57). V průběhu autoaktivační reakce cysteinových katepsinů jsou oslabeny interakce mezi propeptidem a katalytickou doménou, dochází ke konformační změně propeptidu a postupnému odhalení aktivního místa (McQueney et al. 1997; Pungerčar et al. 2009). Lze předpokládat, že za této situace dochází k efektivní interakci s reverzibilním cyanohydrazidovým inhibitorem, který proces autokatalytické konverze zpomalí.



Obrázek 28: **Funkční interakce inhibitorů Gü1303 a Gü2602 s katepsinem K v buňkách a při aktivaci zymogenu.** A) Proenzym KatK (pKatK) byl inkubován *in vitro* s inhibitory o finální koncentraci 10 μM při pH 4,0 a byla sledována jeho autokatalytická konverze na aktivní maturní formu (mKatK). V uvedených časových intervalech byly odebrány vzorky reakční směsi, které byly separovány pomocí SDS-PAGE. V přítomnosti inhibitoru dochází k potlačení procesu tvorby mKatK. B) Buněčná linie osteosarkomu (U-2 OS) byla inkubována v přítomnosti inhibitorů o finální koncentraci 1 μM po dobu 3 hodin. Následně byly buňky inkubovány s fluorescenční sondou 25 rozpoznávající aktivní místo KatK po dobu 24 hodin. Buněčný lyzát byl separován na SDS-PAGE a vizualizován fluorescenčně (levá část) a barvení proteinu pomocí Coomassie Brilliant Blue G (pravá část). Zhášení fluorescenční sondy 25 je důkazem kompetice s inhibitory o vazbu do aktivního místa KatK.

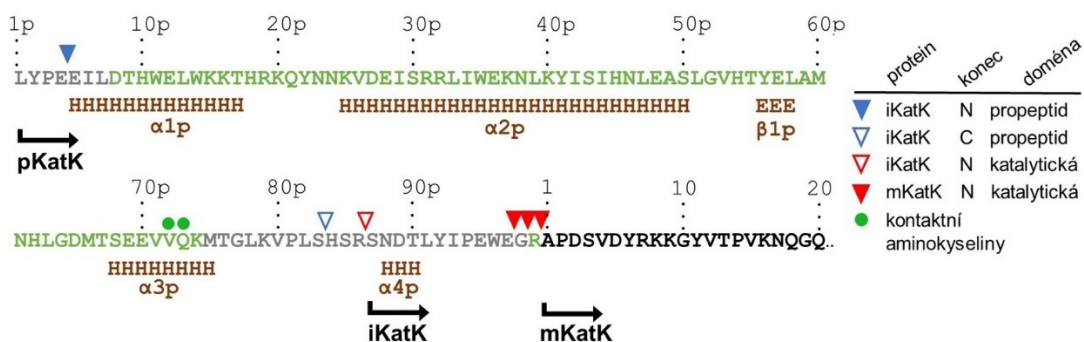
5.5.2 Krystalografická analýza komplexů cyanohydrazidových inhibitorů s katepsinem K

V rámci této studie byly připraveny komplexy inhibitorů **Gü1303** a **Gü2602** se dvěma molekulárními formami KatK. S maturním enzymem mKatK a s aktivačním meziproductem iKatK, který vzniká během konverze zymogenu KatK na maturní formu enzymu. Pro řešení komplexů mKatK byla zvolena metoda molekulárního nahrazení a v případě řešení komplexů iKatK byla použita kombinace molekulárního nahrazení a automatizovaného modelování propeptidové domény pomocí programu Buccaneer.

mKatK komplexy obsahovaly jednu molekulu v asymetrické jednotce s ortorombickou symetrií krystalové mříže a bodovou grupou $P2_12_12$. V případě komplexu mKatK-**Gü1303** bylo dosaženo finálního rozlišení na 1,55 Å a pro komplex mKatK-**Gü2602** bylo po ořezu datasetu ponecháno rozlišení na 2,00 Å.

iKatK komplexy obsahovaly též po jedné molekule v asymetrické jednotce, s tetragonální symetrií krystalové mříže a bodovou grupou $P4_32_12$. U komplexu iKatK-**Gü1303** bylo finální rozlišení 1,90 Å a pro komplex iKatK-**Gü2602** bylo docíleno rozlišení 1,88 Å.

Kombinací metod N-koncového sekvenování a hmotnostní spektrometrie byly určeny N- a C-konce polypeptidových řetězců tvořících mKatK a iKatK. Molekula mKatK obsahuje katalytickou doménu s polypeptidovým řetězcem Ala1 až Met215 (číslováno podle maturního enzymu) s heterogenitou na N-konci, která je způsobena prodloužením o dvě další aminokyseliny (Gly98p a Arg99p, písmeno p značí číslování propeptidu) zbylými z odštěpeného propeptidu (Obr. 29, str. 59). Ve vyřešeném 3D modelu komplexu mKatK-**Gü2602** je v mapách elektronových hustot patrná pouze část Ala1 až Met215. V případě komplexu mKatK-**Gü1303** je viditelná i N-koncová extenze. Komplexy s iKatK obsahovaly kromě katalytické domény též neúplný a nekovalentně vázaný propeptidový řetězec Glu5p až Ser83p. Katalytická doména iKatK byla na N-konci prodloužena o zbytkovou část třinácti aminokyselin z C-koncové části propeptidové domény a její celkový rozsah byl v rozmezí Ser87p až Met215 (Obr. 29, str. 59). Na základě viditelných map elektronových hustot byl modelován u komplexu iKatK-**Gü1303** polypeptidový řetězec reziduálního propeptidu od Asp8p po Gln73p a řetězec katalytické domény od Pro2 po Met215. Výsledný model iKatK-**Gü2602** měl viditelný reziduální propeptid v rozsahu Asp8p po Lys74p a katalytickou doménu Arg99p až Met215.



Obrázek 29: Aminokyselinová sekvence N-koncové oblasti zralého katepsinu K a aktivačního meziprojektu. Znárodněna je oblast sekvence KatK od N-konce propeptidové domény (zeleně a šedě) po prvních 20 aminokyselin katalytické domény (černě). Části propeptidu, které jsou viditelné ve finálních krystalografických modelech jsou zeleně, ostatní šedě. Pod sekvencí je uvedena lokalizace sekundárních struktur (H: α -helix, E: β -struktura; odvozeno z 3D struktury mKatK a iKatK, chybějící segmenty jsou odvozeny ze struktury zymogenu KatK PDB: 1BY8). Trojúhelníkové značky uvedené nad sekvencí označují N- a C-konce (uvedeno v legendě) polypeptidových řetězců určené N-koncovým sekvenováním a hmotnostní spektrometrií. Zeleným kruhem jsou označeny aminokyseliny z reziduálního propeptidu, které tvoří kontakt s inhibítozem Gü2602 v komplexu iKatK-Gü2602. Černé lomené šipky určují pozici N-konce u proenzymu pKatK, katalytické domény aktivačního meziprojektu iKatK a plně aktivované zralé formy mKatK.

5.5.3 Analýza vazebného módu cyanhydrazidových inhibítozů s zralou formou katepsinu K

Rentgenostrukturní analýza komplexů cyanhydrazidových inhibítozů s mKatK prokázala kovalentní vazbu inhibítozů v aktivním místě. Thiolová skupina katalytického Cys25 tvoří s uhlíkovým atomem inhibítozu (označen hvězdičkou na Obr. 27, str. 55) reverzibilní isothiosemikarbazidový adukt ve vzdálenosti 1,8 Å. Peptidomimetické inhibítozy jsou navázány v orientaci napodobující substrát, kdy pozice P1 až P3 ve struktuře inhibítozů vyplňují odpovídající vazebná podmišta S1 až S3 enzymu. Oba inhibítozy interagují s aktivním místem enzymu pomocí sítě vodíkových můsteků a nepolárních interakcí, které jsou shrnuty v (Tab. 5, str. 67). V následujících kapitolách budou tyto interakce detailně popsány.

5.5.3.1 Interakce Gü1303 v aktivním místě maturního katepsinu K

Nepolární interakce mezi **Gü1303** a mKatK jsou uvedeny na (Obr. 30B, str. 62) Pozice P1 inhibitoru obsahuje azaalanin a přestože tvoří početné kontakty s okolními aminokyselinami v S1 vazebném podmístě (Tab. 5, str. 67), žádná z těchto interakcí není definována jako striktně nepolární. Pozice P2 obsahuje aminokyselinu fenylalanin, která tvoří interakce s aminokyselinovými zbytky (Tyr67, Ala134, Leu160, Ala163 a Leu209) ve vazebném podmístě S2. Pozice P3 s terminální skupinou benzyloxykarbonylu interaguje s aminokyselinami (Asp61 a Tyr67) v S3 vazebném podmístě. 4-hydroxyfenylová skupina Tyr67 dále utváří stabilizační π - π interakci s fenylovým kruhem terminální skupiny.

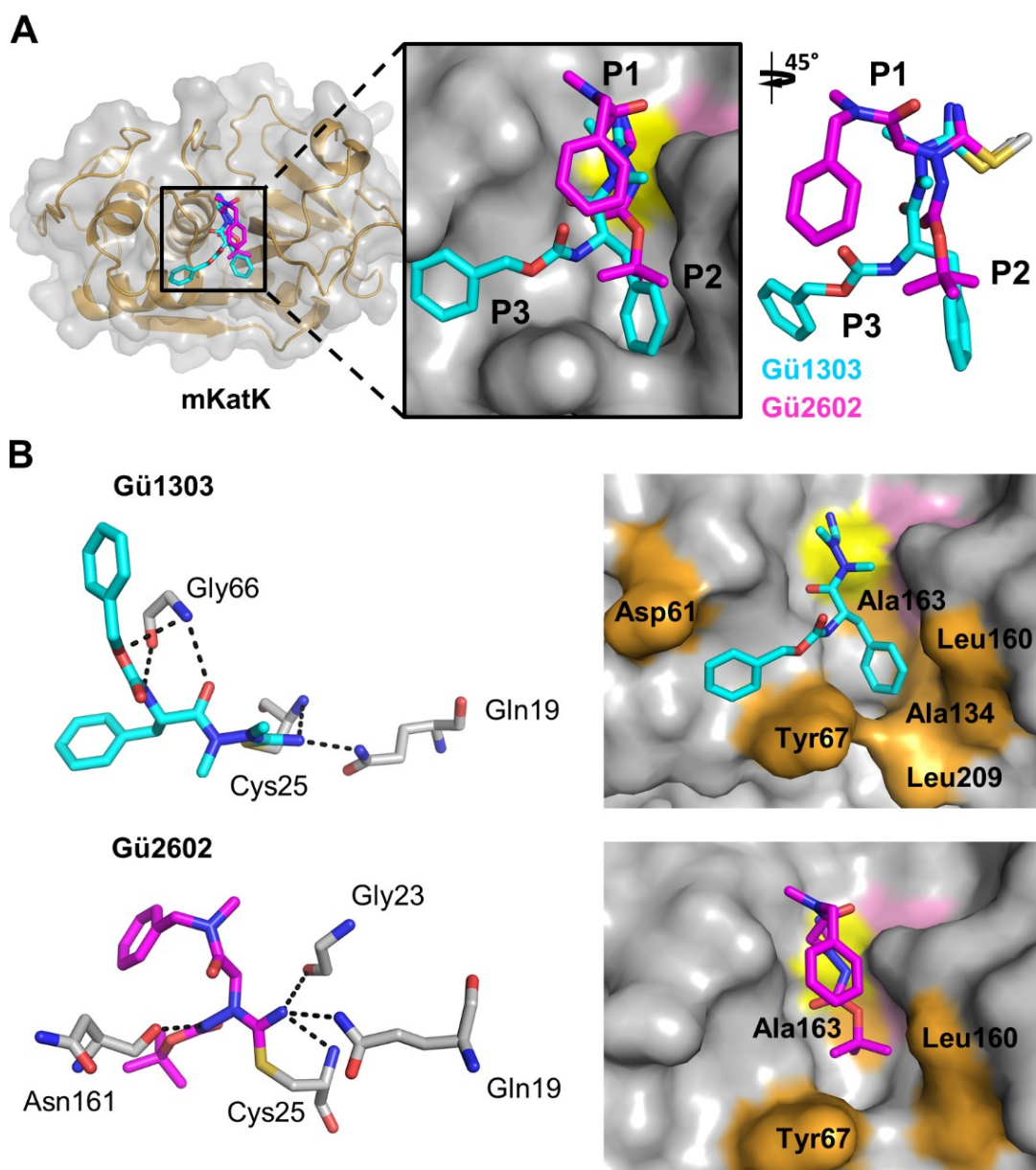
Inhibitor **Gü1303** je v aktivním místě mKatK stabilizován sítí vodíkových můstků uvedených na (Obr. 30B, str. 62). Atom dusíku z isothiosemikarbazidového aduktu (který vznikl reakcí nitrilové skupiny inhibitoru s thiolovou skupinou Cys25) interaguje vodíkovou vazbou s amidovou skupinou Cys25 a současně s amidovou skupinou z postranního řetězce Gln19. Další vodíkové můstky tvoří Gly66, který přispívá jednak svou amidovou skupinou k interakci s atomem kyslíku karbonylu na zbytku fenylalaninu v pozici P2 inhibitoru a jednak s atomem kyslíku karbamátu terminální skupiny (označen hvězdičkou -NH-(C=O)-O*) v pozici P3 inhibitoru. Atom kyslíku karbonylu Gly66 ještě tvoří vodíkový můstek s dusíkovým atomem karbamátu (-N*H-(C=O)-O-) inhibitoru.

5.5.3.2 Interakce Gü2602 v aktivním místě maturního katepsinu K

Reaktivní cyanohydrazidová skupina je u **Gü2602** oproti inhibitoru **Gü1303** umístěna v centrální části molekuly. Inhibitor **Gü2602** byl navržen tak, aby byl schopný cílit i na indexovaná podmísta (S1' a S2') enzymu (Schmitz et al. 2014). Nepolární interakce mezi **Gü2602** a mKatK jsou uvedeny na (Obr. 30B, str. 62). Pozice P1 inhibitoru obsahuje segment *N*-benzyl-*N*-methylacetamidu, který interaguje s aminokyselinami (Gly23, Gly64 a Gly65) v podmístě S1. Mapy elektronových hustot finálního krystalografického modelu ukazují, že benzylový zbytek tohoto segmentu není směřován do původně zamýšlených indexovaných podmíst, ale je spíše orientována nad aktivní místo v oblasti S1 podmísta (Obr. 30A, str. 62). Tato skutečnost bude podrobněji komentována v následující kapitole (5.5.4). V pozici P2 inhibitoru se nachází terminální skupina Boc, která tvoří nepolární interakce se zbytky (Tyr67, Leu160 a Ala163) v podmístě S2.

I v případě inhibitoru **Gü2602** se při vazbě do aktivního místa vytváří stabilizační síť vodíkových můstků (Obr. 30B, str. 62). Dusíkový atom aduktu interaguje s amidem

na kostře katalytického Cys25, amidem postranního řetězce Gln19 a karbonylem na kostře Gly23. Oproti **Gü1303** nebyly v terminální oblasti pozice P2 pozorovány vodíkové interakce s Gly66, ale vznikla interakce s kostrou zbytku Asn161 tvořená amidickým dusíkem inhibitoru **Gü2602**.



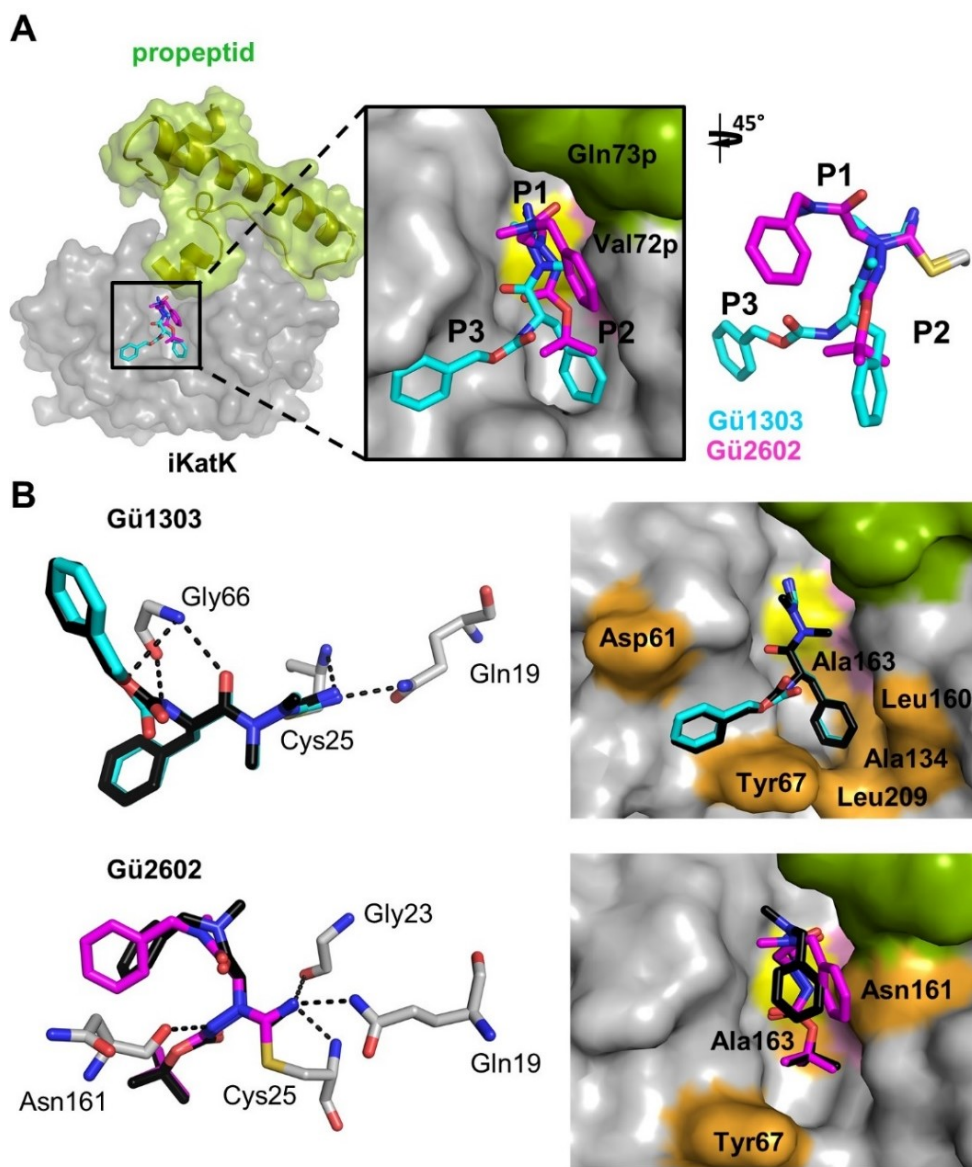
Obrázek 30: Vazebný mód cyanohydrazidových inhibitorů v aktivním místě maturního katepsinu K. A) Superpozice inhibitorů v aktivním místě mKatK a výřez s vyznačenými pozicemi inhibitorů P1 až P3 a odpovídajícími podmísty enzymu S1 až S3. Maturní forma enzymu mKatK je zobrazena v transparentním povrchovém modelu (šedě) a stužkovém modelu (hnědě). Katalytické místo enzymu ve výřezu je označeno pro zbytky Cys25 (žlutá) a His162 (růžová). Inhibitory jsou v tyčinkovém modelu a jejich uhlíkové atomy jsou barevně odlišeny tyrkysovou pro **Gü1303** a růžovou pro **Gü2602**. Heteroatomy mají standartní barevné kódování (O: červená, N: modrá, S: žlutá). B) Levé panely s tyčinkovými modely znázorňují síť vodíkových interakcí (čárkovaně) mezi inhibitory (barevně značení je stejné jako v sekci A) a aminokyselinovými zbytky mKatK (šedě). Pravé panely popisují nepolární interakce mezi inhibitory a aktivním místem mKatK v povrchové reprezentaci. Interagující zbytky enzymu jsou oranžově, ostatní šedě. Barevně jsou vyznačeny katalytické zbytky jako v panelu A.

5.5.3.3 Interakce Gü1303 v aktivním místě aktivačního meziproduktu katepsinu K

Aktivační meziprodukt (iKatK) vznikl limitovanou autokatalytickou proteolýzou zymogenu a obsahuje nekovalentně navázaný reziduální propeptid (Obr. 31A, str. 64) a podrobněji kapitola 5.5.2). V krystalových strukturách komplexu iKatK s oběma cyanohydrazidovými inhibitory je patrné, že aktivní místo je ve struktuře iKatK otevřené pro vstup ligandu s výjimkou částečně blokových indexovaných podmíst (konkrétně S1'). Interakce inhibitoru **Gü1303** s iKatK jsou analogické s výše popsanými pro jeho komplex s maturní formou enzymu (Tab. 5, str. 67) a nebyly nalezeny interakce s propeptidovým segmentem. Superpozice finálních modelů iKatK-**Gü1303** a mKatK-**Gü1303** (Obr. 31B, str. 64) prokázala téměř identickou konformaci inhibitorů v katalytickém centru enzymu s odchylkou pouhých 0,19 Å (RMSD). Pouze v P3 pozici je patrná rozdílná orientace karbonylové skupiny a mírná odchylka terminální skupiny, která má za následek ztrátu π - π interakce se zbytkem Tyr67.

5.5.3.4 Interakce Gü2602 v aktivním místě aktivačního meziproduktu katepsinu K

Inhibitor **Gü2602** vytváří v aktivním místě iKatK analogické interakce jako v komplexu s mKatK (Obr. 30B, str. 62). Nicméně superpozice finálních modelů iKatK-**Gü2602** a mKatK-**Gü2602** (Obr. 31B, str. 64) ukazuje rozdílnou orientaci segmentu *N*-benzyl-*N*-methylacetamidu v P1 pozici inhibitoru, který je natočený směrem k propeptidu, takže inhibitor vytváří kontakty s propeptidovými zbytky Val72p a Gln73p (Obr. 31A, str. 64). Hodnota odchylky mezi konformacemi **Gü2602** v komplexech s iKatK a mKatK činila z tohoto důvodu 1,23 Å (RMSD).

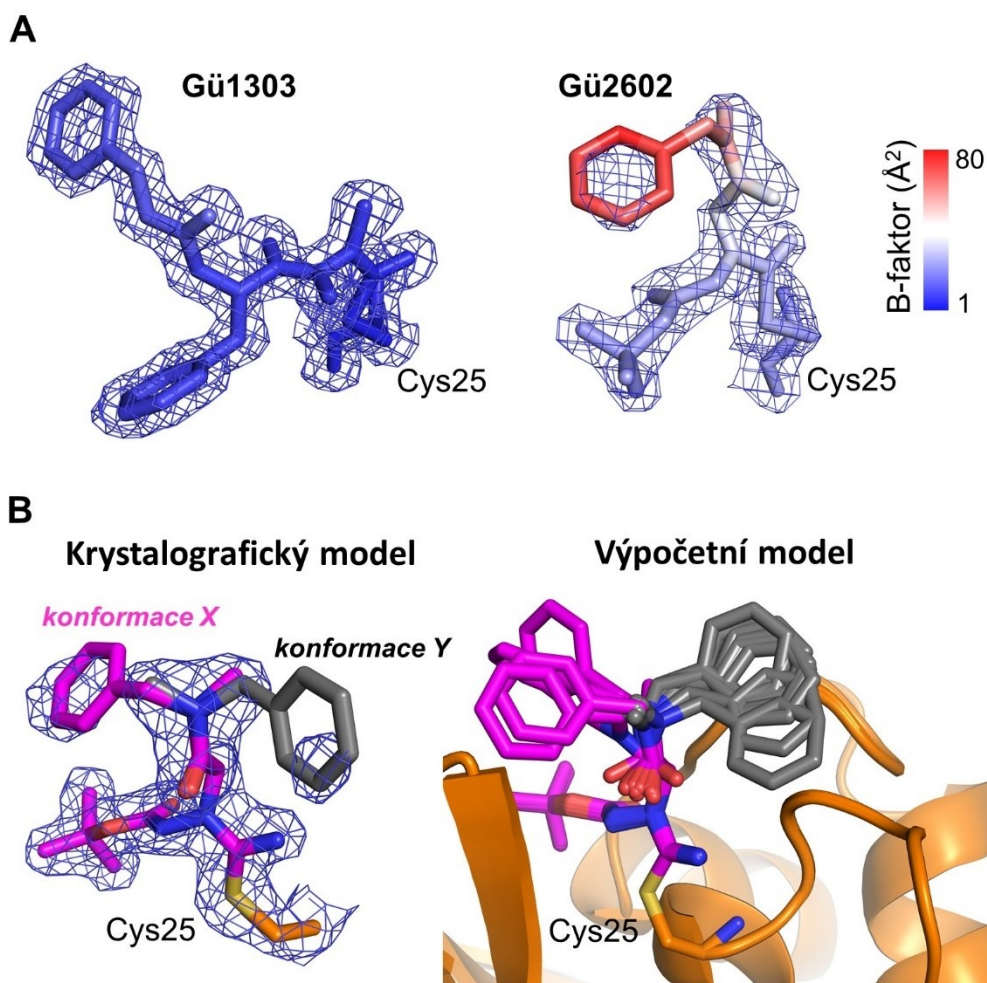


Obrázek 31: Vazebný mód cyanohydrázidových inhibitorů v aktivním místě aktivačního meziprojektu *katepsinu K*. A) Superpozice inhibitorů v aktivním místě iKatK a výřez s vyznačenými pozicemi inhibitorů P1 až P3 a odpovídajícími podmísty enzymu S1 až S3. Katalytická doména aktivačního meziprojektu iKatK je zobrazena v povrchovém modelu (šedě) a asociovaný reziduální propeptid je ve stužkovém modelu s transparentním povrchem (zeleně). Katalytické místo enzymu ve výřezu je označeno pro zbytky Cys25 (žlutá) a His162 (růžová). Gln73p a Val72p jsou zbytky propeptidu vytvářející kontakty s inhibitorem **Gü2602**. Inhibitory jsou znázorněny v tyčinkovém modelu a jejich uhlíkové atomy jsou barevně odlišeny tyrkysovou pro **Gü1303** a růžovou pro **Gü2602**; heteroatomy mají standartní barevné kódování (O: červená, N: modrá, S: žlutá). B) Levé panely s tyčinkovými modely znázorňují síť vodíkových interakcí (čárkovaně) mezi inhibitory v superpozici (Inhibitory z iKatK komplexů mají stejné barevné kódování jako v sekci A a inhibitory z mKatK komplexů mají uhlíkové atomy černě) a aminokyselinovými zbytky mKatK (šedě). Pravé panely popisují nepolární interakce mezi inhibitory a aktivním místem mKatK v povrchové reprezentaci. Interagující zbytky enzymu jsou oranžové, ostatní šedě. Barevně jsou vyznačeny katalytické zbytky jako v panelu A.

5.5.4 Analýza konformační flexibility inhibitorů **Gü2602** a **Gü1303** v aktivním místě

Krystalografické mapy elektronových hustot použité pro řešení inhibitorů **Gü1303** a **Gü2602** v aktivním místě mKatK (a iKatK) byly obecně ve vysoké kvalitě s výjimkou slabého signálu pro benzylovou skupinu **Gü2602** (Obr. 32A, str. 66). Proto byla dále analyzována distribuce krystalografických B-faktorů (teplotních faktorů). Jejich hodnoty jsou obecně nízké, ale zvyšují se pro N-benzyl-N-methylacetamidový segment v pozici P1 inhibitoru **Gü2602** s nevyššími hodnotami v jeho benzylové skupině (Obr. 32A, str. 66). To dokazuje zvýšenou flexibilitu a dynamickou neuspořádanost této krystalografické konformace benzylové skupiny (viz konformace X na Obr. 32B, str. 66), která je umístěna v oblasti podmísta S1 a orientována směrem ven z aktivního místa.

Vzhledem k uvedené flexibilitě P1 pozice a skutečnosti, že inhibitor **Gü2602** byl původně navržen tak, aby jeho P1 skupina mohla potenciálně interagovat s indexovanými podmísty enzymu (Schmitz et al. 2014), byly hledány důkazy pro další možné konformace. Detailní inspekce krystalografických map elektronových hustot odhalila slabý signál, který lze přiřadit druhé, alternativní konformaci P1 segmentu inhibitoru **Gü2602** (viz konformace Y na Obr. 32B, str. 73). Zde je benzylová skupina namířena směrem do S2' podmísta enzymu. Pro další analýzu této alternativní konformace byly použity přístupy výpočetní chemie, pomocí kterých byla generována sada energeticky výhodných poloh benzylové skupiny a celého N-benzyl-N-methylacetamidového segmentu. Ty obecně patří ke dvěma základním typům konformací, které jsou orientovány jednak do neindexované (podmísto S1), anebo indexované (podmísto S2') oblasti aktivního místa enzymu.



Obrázek 32: **Konformační flexibilita cyanohydrazidových inhibitorů Gü1303 a Gü2602 v aktivním centru maurního katepsinu K** A) Znáznorněny jsou inhibitory v tyčinkovém modelu, které jsou kovalentně navázány na katalytický Cys25. Mapy elektronových hustot ($2F_o - F_c$) inhibitorů byly konturovány na 1σ pro **Gü1303** a $1,5\sigma$ pro **Gü2602**. Inhibitory jsou vybarveny dle hodnot B-faktorů atomů v gradientu od modré (nízké hodnoty) po červenou (vysoké hodnoty). Nejvyšší hodnoty B-faktoru u benzylové skupiny **Gü2602** indikují vysokou flexibilitu. B) Levý panel ukazuje krystalografické řešení dvou konformací N-benzyl-N-methylacetamidového segmentu inhibitoru **Gü2602**. Pozorována majoritní konformace X je růžově a predikovaná alternativní konformace Y je šedě. Mapy $2F_o - F_c$ jsou zobrazeny pro 1σ . Katalytický Cys25 je oranžově; heteroatomy mají obvyklé barevné kódování (O: červená, N: modrá a S: žlutá). Pravý panel ukazuje sadu energeticky výhodných konformací odvozených z krystalografických konformací X a Y (levý panel), které byly získány metodami výpočetní chemie (detaily viz autorská publikace č. 3). Barevné provedení je totožné s levým panelem.

Tabulka 5: Souhrn všech kontaktů tvořených mezi aminokyselinovými zbytky mKatK/iKatK a cyanohydrazidovými ligandy **Gü1303** a **Gü2602**. Analýza interakcí byla provedena v programu CONTACT. Struktury ligandů byly rozděleny na jednotlivé pozice s označením P1 až P3, pro každou pozici jsou uvedeny příslušné kontaktní zbytky enzymu. Parametr s označením C představuje součet všech kontaktů daného zbytku enzymu s ligandem (včetně vodíkových můstků). Počet vodíkových můstků udává parametr Hb (zbytky enzymu, které tvoří vodíkové můstky, jsou uvedeny tučně). Pro určení kontaktů byly použity vzdálenosti do 4,2 Å a pro vodíkové můstky vzdálenosti do 3,3 Å.

Pozice	mKatK-Gü2602		mKatK-Gü1303		iKatK-Gü2602		iKatK-Gü1303			
	C	Hb	C	Hb	C	Hb	C	Hb		
P1					Val72p	10				
					Gln73p	2				
	Gln19	2	1	Gln19	3	1	Gln19	3	1	
	Gly23	10	1	Gly23	10		Gly23	10		
	Ser24	2		Ser24	3		Ser24	3		
	Cys25	13	1	Cys25	14	1	Cys25	14	1	
	Trp26	2		Trp26	3		Trp26	3		
	Cys63	1					Cys63	1		
	Gly64	7					Gly64	4		
	Gly65	2					Gly65	4		
	Asn161	10	1	Asn161	2		Asn161	6	1	
His162	3		His162	2		His162	3			
P2	Cys25	1	Cys25	2		Cys25	2	Cys25	2	
	Gly65	2	Gly65	6		Gly65	2	Gly65	6	
	Gly66	5	Gly66	11	2	Gly66	5	Gly66	11	2
	Tyr67	2	Tyr67	3		Tyr67	2	Tyr67	3	
			Met68	2				Met68	2	
			Ala134	3				Ala134	3	
	Leu160	3	Leu160	4		Leu160	3	Leu160	4	
	Asn161	6				Asn161	9			
	His162	3				His162	3			
	Ala163	1	Ala163	1		Ala163	1	Ala163	1	
			Leu209	2				Leu209	2	
P3			Glu59	3			Glu59	3		
			Asn60	3			Asn60	3		
			Asp61	5			Asp61	5		
			Gly65	3			Gly65	3		
			Gly66	15	1		Gly66	15	1	
			Tyr67	11			Tyr67	11		

6. Diskuse

Disertační práce se zabývá cysteinovou proteasou katepsinem K (KatK), který je cílovou molekulou pro léčbu několika patologií. Nejvíce studované je spojení KatK s osteoporózou, dále s artritidou a v neposlední řadě s metastázemi do kostní tkáně. Tato práce se soustředí na nové mechanismy regulace aktivity KatK pomocí proteasových inhibitorů, které představují potenciální chemoterapeutika.

V současné době (2021) není na trhu žádné komerční léčivo a bázi cílené inhibice KatK. Vývoj několika inhibičních preparátů jako je odanacatib (Merck), balicatid (Novartis) nebo ONO-5334 (Ono Pharmaceutical) byl ukončen v II. nebo III. fázi klinických studií vzhledem k vedlejším účinkům jako jsou zvýšené riziko cévní mozkové příhody, kardiovaskulárních komplikací nebo kožní defekty (Makras et al. 2015; Drake et al. 2017; Lu et al. 2018). Další inhibitory KatK zejména pro léčbu osteoporózy jsou v současné době ve vývoji a testování farmaceutickými firmami (Qiu et al. 2016). Dle statistických údajů v České republice osteoporózou trpí přibližně 5 % obyvatel (Kanis et al. 2021) a postihuje zejména ženy po menopauze. Nejčastějším projevem tohoto onemocnění je zlomenina krčku stehenní kosti a více jak 20 % nemocných zemře následně na komplikace spojené s tímto úrazem, nejčastěji v podobě zápalu plic. Klinická praxe je zaměřena na režimová opatření a suplementaci, která spočívají v podávání vápníku a vitamínu D. Dále v antiresorpční léčbě pomocí přípravků jako jsou estrogény, modulátor estrogenových receptorů Raloxifen, bisfosfonáty nebo protilátka proti cytokinu RANKL Denosumab, ale u tohoto přístupu je také známá celá škála vedlejších účinků (Zhang et al. 2011; Tu et al. 2018).

Na příkladu výše uvedených vyvíjených léčiv na bázi inhibice KatK, u kterých byly prokázány vedlejší účinky, je patrná komplexnost problému vyžadující nové podrobnější informace o KatK. Jedním z důležitých aspektů je otázka selektivity inhibice enzymové aktivity KatK. V proteomu je redundance cysteinových proteas katepsinového typu a jejich účasti v buněčných signalizačních a regulačních kaskádách, kdy dochází k částečnému překryvání a zastupitelnosti rolí vzhledem k blízkému funkčnímu profilu jednotlivých příbuzných enzymů. To klade zvýšené nároky na úzké cílení inhibice. Proto je kladen důraz na vývoj selektivnějších inhibitorů KatK nebo dokonce na inhibici pouze některých funkcí KatK. Příkladem jsou studie zabývající se inhibicí kolagenolytické aktivity KatK blokováním exo-míst bez ovlivnění aktivity KatK k jiným typům substrátů (Panwar et al. 2018). Dalším problémem, s nímž se setkal vývoj inhibitorů KatK je vhodný

testovací biologický model. Během výzkumu jsou využívány především studie na hlodavcích (myš, krysa a králík). Nicméně homologa KatK těchto hlodavců vykazují sekvenční identitu 86 % až 96 % a liší se stavbou aktivního místa, především substitucemi v podmístě S2, oproti lidskému KatK. To má za následek významné rozdíly v inhibičních parametrech testovaných inhibitorů (Desmarais et al. 2008) jako příklad lze uvést sníženou inhibiční účinnost vyvíjených inhibitorů balicatib a odanacatib, což pravděpodobně vedlo k tomu, že vedlejší účinky terapie nebyly odhaleny už v úvodních fázích testování (Law et al. 2017).

Pro design vysoce biologicky aktivních a selektivních inhibitorů KatK je nezbytné detailně popsat strukturní mechanismy inhibice a zmapovat interakce mezi inhibičními ligandy a cílovým enzymem. Tato dizertační práce si kladla za cíl testovat účinnost a analyzovat vazebný mód pro nové peptidomimetické a nízkomolekulární inhibitory KatK a přinést relevantní poznatky pro pokročilý vývoj potenciálních léčiv.

Regulace enzymové aktivity KatK pomocí nových nízkomolekulárních a peptidomimentických inhibitorů.

V autorské **publikaci č. 1** (Pomeislová et al. 2021) byly poprvé experimentálně testovány a charakterizovány inhibiční vlastnosti syntetických nízkomolekulárních sloučenin s 1,2,4-thiadiazolovou reaktivní skupinou vůči lidskému KatK. Tyto sloučeniny byly klasifikovány jako irverezibilní kovalentní inhibitory, které tvoří disulfidický adukt s thiolovou skupinou proteasy. Sada 21 sloučenin obsahovala šest sloučenin s výraznou účinností inhibice KatK pod hranici 50 % zbytkové enzymatické aktivity a tyto látky vykazovaly hodnotu IC_{50} v mikromolárních koncentracích. Nejúčinnější derivát **10a** substituovaný na dusíku N5 měl hodnotu $IC_{50} = 2,63 \pm 0,31 \mu M$. Publikace identifikovala substituované sloučeniny s 1,2,4-thiadiazolovým heterocyklem jako účinné inhibitory KatK a rozšiřuje tak spektrum potenciálního využití této skupiny látek. Deriváty byly již dříve popsány jako inhibitory cysteinového katepsinu B (Leung-Toung et al. 2003), který je studován v souvislosti s léčbou Alzheimerovy choroby (Bernstein and Keilhoff 2018) nebo jako regulátory medicíně zajímavých enzymů s vazebnými místy obsahujícími cysteinové zbytky jako je GSK-3 β (Martinez 2002; Pomeislová et al. 2021). Reaktivní skupina 1,2,4-thiadiazolu a popsané závislosti mezi strukturou a aktivitou studovaným derivátů otevírají cestu pro další design odvozených inhibitorů KatK s optimalizovanými funkčními vlastnostmi. Pro další racionální design zatím scházejí informace na úrovni

3D struktur, které zatím nebyly publikovány pro relevantní thiadiazolové ligandy kromě teoretických simulací (Vega-Tejido et al. 2014).

V autorské **publikaci č. 3** (Benýšek et al 2022) byl poprvé popsán strukturní mechanismus interakce nové skupiny vysoce selektivních peptidomimetických inhibitorů KatK vybavených cyanohydrazidovou reaktivní skupinou, které vykazují vynikající inhibiční účinnost s hodnotami K_i v až pikomolární oblasti koncentrací. Cyanohydrazidy představují zřejmě dosud nejpokročilejší design reverzibilních kovalentních inhibitorů KatK a tato práce se zabývá dvěma hlavními chemotypy označovanými jako azadipeptidové nitrily a 3-cyano-3-aza- β -aminokyseliny. Vznikly nahrazením C α H skupiny konvenčně používaných karbanitrilů (jako jsou např. dříve vyvíjená léčiva odanacatib a balicatib) atomem dusíku v P1 pozici inhibitorů a další stabilizací této skupiny substitucemi. V případě azadipeptidových nitrilů je cyanohydrazidová reaktivní skupina umístěna terminálně a představitelem těchto látek je inhibitor **Gü1303** s hodnotou $K_i = 0,91$ nM (Frizler et al. 2011). Pro dosažení vyšší selektivity vůči KatK byla reaktivní skupina umístěna centrálně do molekuly inhibitoru, aby byla možná interakce i s indexovanými podmísty enzymu, a tak byly odvozeny inhibitory typu 3-cyano-3-aza- β -aminokyselin. Jejich představitelem je inhibitor **Gü2602** s hodnotou $K_i = 0,013$ nM (Schmitz et al. 2014). Disertační práce přináší první strukturní analýzu vazebného módu těchto dvou prototypových inhibitorů.

Inhibitory s cyanohydrazidovou skupinou mohou být obecně atropochirální. Atropoisomery jsou stereoisomery, které se liší energií jednoduché substituované vazby s omezenou rotací kolem osy symetrie. V případě cyanohydrazidové skupiny jde o substituovanou vazbu N–N. Inhibitor **Gü1303** má dusíkové atomy NMe–NMe methylované a sousední vazba CO–NMe je u nenávaného inhibitoru v *E*-konfiguraci (Jílková et al. 2021). Po navázání do aktivního místa enzymu dochází ke konformační změně na *Z*-konfiguraci. Stejný efekt byl pozorován i v případě interakce azadipeptidového nitrilu s katepsinem B z motolice *Schistosoma mansoni* (SmCB1), který je cílovou molekulou pro léčbu schistosomózy (Jílková et al. 2021). V případě **Gü2602** je vazba NH–NMe pouze částečně methylovaná, a tak nepředstavuje atropochirální centrum, jak bylo prokázáno pomocí NMR dat (Ottersbach et al, 2012). Na sousední vazbě CO–NH proto nedochází ke konverzi konfigurace během vazby. Rozdíl v konformaci navázaného a nenávaného inhibitoru se u **Gü1303** a **Gü2602** odráží i v jejich odlišném kinetickém chování (viz příloha k **publikaci č. 3**). V případě **Gü1303** byla určena kinetika typická pro tzv. pomalu vázající

inhibitory, kdy dochází k pomalému ustálení rovnováhy mezi enzymem, inhibitorem a výsledným komplexem z důvodu konformační změny inhibitoru. Tato změna byla prokázána jako krok kontrolující kinetiku a získané výsledky jsou tak v souladu s dříve publikovaným článkem o kinetice inhibice SmCB1 pomocí azadipeptidového nitrilu (Jílková et al. 2021). Oproti tomu inhibitor **Gü2602** vykazuje kinetiku tzv. rychle vázajících inhibitorů, což odpovídá absenci časově náročné konformační změny během vazby. Tak se tento inhibitor podobá v kinetice a konfiguraci konvenčním karbanitrilovým inhibitorům (Jílková et al. 2021).

Rozdílné umístění reaktivní skupiny mezi inhibitory **Gü1303** a **Gü2602** se odráží i v rozdílném obsazení vazebných podmíst na KatK. **Gü1303** s terminálně umístěnou reaktivní skupinou cílí vazebná podmísta S1 až S3 a na základě analýzy B-faktorů (teplotních faktorů) se jedná o rigidní inhibitor. Naproti tomu **Gü2602** s centrálně umístěnou reaktivní skupinou obsazuje primárně vazebné podmísto S2 a svým flexibilním segmentem *N*-benzyl-*N*-methylacetamidu (zvláště jeho benzylovou skupinou) vytváří dvě alternativní konformace. Ty jsou směřovány buď do podmísta S1 nebo S2'. Konformace v podmístě S1 je ta hlavní, která byla pozorovaná v krystalové struktuře, kde je zřejmě upřednostněna stabilizací pomocí přídatných intramolekulárních kontaktů a krystalografických kontaktů se sousedními molekulami. Pro konformaci v podmístě S2' jsou důkazy v mapě elektronových hustot a zejména byla prokázána metodami výpočetní chemie. Přístup molekulárního modelování obecně odhalil sadu možných variací v orientaci pro obě základní konformace a potvrdil vysokou flexibilitu této části inhibitoru **Gü2602**. V této souvislosti je také zajímavé zmínit autorskou **publikaci č. 2** (Lemke et al. 2021), kde bylo popsáno umístění flexibilní benzylové skupiny do S2' podmísta KatK a její funkční význam u jiného typu inhibitoru.

Zvýšená flexibilita inhibitoru **Gü2602** je možným vysvětlením pro jeho výjimečné inhibiční vlastnosti s účinností v pikomolární oblasti. Lze předpokládat, že konformační flexibilita (na hranici mezi indexovanými a neindexovanými podmísty) poskytuje entropickou výhodu v celkové energetice vazby na KatK. Je známo, že ztráta konformační entropie po navázání na enzym nepříznivě přispívá k volné vazebné energii (Chia-en et al. 2007). Dalším významným faktorem pro vazbu **Gü2602** je stabilizace komplexu sítí vodíkových interakcí, kde klíčovou roli hraje interakce amidového dusíku NH na reaktivní skupině inhibitoru s Asn161 v podmístě S2. Deriváty **Gü2602**, které mají tento

dusík methylovaný a tudíž netvoří interakci, vykazují dramatický pokles inhibiční účinnosti o několik řádů (Schmitz et al. 2014).

Kromě komplexů inhibitorů **Gü1303** a **Gü2602** s maturní formou KatK (mKatK) byly také strukturně charakterizovány komplexy s aktivačním meziproduktem (iKatK). Ten vzniká v průběhu autokatalytické konverze nekativního zymogenu na plně aktivovaný maturní enzym a tento proces je spojen s proteolytickým odstraněním kovalentně spojené propeptidové domény. Nyní byla poprvé popsána struktura iKatK, která obsahuje vedle hlavní katalytické domény nekovalentně asociovaný fragment reziduálního propeptidu (v rozsahu Glu5p až Ser83p). Na rozdíl od struktury zymogenu (Sivaraman et al. 1999) u iKatK zbytkový propeptid pouze částečně překrývá aktivní místo enzymu a blokuje převážně podmísto S1'. Katalytické centrum iKatK se tak již stává přístupné pro ligand. Zde je vidět rozdíl oproti aktivačnímu meziprojektu katepsinu SmCB1, ve kterém je katalytické centrum stále blokováno reziduálním propeptidem a proteasa v této formě je neaktivní (Jílková et al. 2014).

Bylo prokázáno, že iKatK váže **Gü2602** a **Gü1303** analogickým způsobem, jaký byl pozorován u jejich komplexů s mKatK. To ukazuje, že aktivní místo iKatK je pro inhibitory plně přístupné ve všech potřebných vazebných podměstech a může být jimi regulováno. Tento závěr dále podporují výsledky testování **Gü2602** a **Gü1303** v autoaktivačním testu *in vitro* se zymogenní formou KatK. Ten prokázal, že oba inhibitory silně potlačují autokatalytickou konverzi zymogenu na maturní formu, která probíhá jako bimolekulární reakce katalyzovaná funkčními formami mKatK/iKatK. Tento závěr je v souladu s předchozími studiemi o mechanismu autoaktivace cysteinových katepsinů (Vernet et al. 1991; Pungercar et al. 2009).

Závěrem byla prokázána schopnost inhibitorů **Gü2602** a **Gü1303** regulovat KatK na buněčném modelu s využitím buněčné linie osteosarkomu. Pro vizualizaci KatK byla použita fluorescenčně značená sonda **25** z autorské **publikace č. 2** (Lemke et al. 2021) která se selektivně váže do aktivního centra KatK a kompetuje o vazbu s inhibitory. Oba inhibitory silně potlačily značení KatK, což dokazuje dobrou permeabilitu inhibitorů pro vstup do buňky a efektivní interakci s cílovým enzymem. Všechna získaná data ukazují inhibitory **Gü2602** a **Gü1303** jako účinné regulátory aktivního KatK a to jak přímou interakcí, tak cílením na aktivační proces, který enzym generuje.

Využití selektivních inhibitorů pro konstrukci fluorescenčních sond detekujících aktivní KatK

Selektivní vizualizace aktivních forem jednotlivých proteas v rámci komplexního proteomu má zásadní význam pro pochopení regulace proteolytických systémů a jejich dysbalance během patologií. Z terapeutického hlediska má důležité využití v diagnostice, zejména pro lokalizaci tkání s patologicky zvýšenými aktivitami proteasových markerů nebo pro monitorování redukováných aktivit cílových proteas po podání léčiv na bázi proteasových inhibitorů. Jednoduchým molekulárním nástrojem pro vizualizaci aktivity proteas jsou selektivní fluorogenní substráty, kde ovšem může dojít ke zkreslení výsledků při lokalizaci signálu vlivem difuze značeného štěpeného produktu (Kasperkiewicz 2021). Alternativním přístupem je použití značených sond na bázi selektivních proteasových inhibitorů, které se váží do aktivního místa proteasy a umožňují tak specifickou detekci aktivní formy enzymu („activity-based probes“, ABP) (Bogyo 2010; Sanman and Bogyo 2014). Tyto sondy jsou v současné době předmětem intenzivního výzkumu ve spojitosti s diagnostikou patologií. Příkladem jsou aktivní sondy pro detekci katepsinu L u rakoviny prsu (Poreba et al. 2018) nebo pro analýzu působení katepsinu B u Alzheimerovy choroby (Poreba et al. 2019).

Příspěvkem k autorské **publikaci č. 2** bylo objasnění vazebného módu první peptidomimetické aktivní sondy pro selektivní detekci KatK. Tato sonda je odvozena ze struktury inhibitoru vybaveného vinylketonovou reaktivní skupinou, která se kovalentně ireverzibilně váže do katalytického centra KatK mechanismem Michaelovy nukleofilní adice. Na základě dat z rentgenové strukturní analýzy komplexu sondy s KatK byly objasněny klíčové interakce v aktivním místě KatK. Sonda obsahovala v P1' pozici zásadní dibenzylaminový segment, který cílil benzylové skupiny na dvě podmísta S1' a S2', a byl zodpovědný za vynikající účinnost a selektivitu. O interakcích inhibitorů v těchto indexovaných vazebných podmínkách KatK je z krystalografických studií známo doposud málo. Dřívější studie se substráty ukazují, že jsou zde akceptovány aromatické aminokyseliny (Alves et al. 2003).

Na opačné straně sondy je umístěn fluorofor cyanin 5, který pro nízkou kvalitu map elektronových hustot nemohl být strukturně popsán. To naznačuje jeho flexibilitu a předpokládanou orientaci ven z aktivního místa. Pro posouzení, zdali fluorescenční značka neovlivňuje celkovou vazbu sondy v aktivním místě, byl analyzován i komplex KatK

s prekurzorovým inhibítozem bez fluoroforu. Výsledná superpozice obou struktur a distribuce kontaktů potvrdila, že síť klíčových interakcí je analogická. Získaná data umožňují konstrukci dalších derivátů například s jinými fluorofory či detekčními značkami v této poloze bez ztráty vazebných vlastností modifikovaných sond.

Funkční vlastnosti aktivní sondy byly úspěšně testovány v kontextu komplexního proteomu pomocí značení buněčných lyzátů a značení buněk lidského osteosarkomu. Fluorescenční mikroskopie a analýza buněk po separaci na SDS-PAGE ukazuje, že navržená sonda vstupuje do buňky až do lyzozomo/endozomálního systému, kde značí aktivní formu KatK. Tato sonda představuje tak účinný nástroj pro citlivou a selektivní detekci KatK.

Závěrem lze shrnout, že disertační práce přinesla důležité poznatky o medicíně relevantních inhibítozem KatK, zejména identifikaci nové třídy nízkomolekulárních thiazolových inhibítozem KatK, vysvětlení strukturního mechanismu interakce cyanohydrazidových inhibítozem jako dosud nejúčinnějších regulátorů KatK a první fluorescenční sondu pro selektivní vizualizaci KatK s potenciálním využitím v diagnostice.

7. Závěr

Disertační práce se zabývala cysteinovou proteasou katepsinem K (KatK), který je cílovou molekulou pro léčbu osteoporózy a artritidy pomocí chemoterapeutik na bázi proteasových inhibitorů. V rámci práce byly studovány tři skupiny syntetických kovalentních inhibitorů KatK nízkomolekulárního a peptidomimetického charakteru. Zjištěny byly nové poznatky o mechanismu inhibiční regulace KatK a vztahu mezi strukturou a aktivitou inhibitorů této proteasy. Získané výsledky mají využití při vývoji účinných selektivních chemoterapeutik a molekulárních diagnostických sond. Výsledky jsou shrnuty ve třech publikacích, které byly publikovány v mezinárodních odborných časopisech.

Zadané cíle práce byly splněny a hlavní získané výsledky jsou následující:

1) Byl připraven rekombinantní KatK a jeho jednotlivé aktivační formy (zymogen, aktivační intermediát a maturní enzym).

2) Testováním knihovny nízkomolekulárních inhibitorů na bázi 1,2,4 thiazolidů byla identifikována nová třída inhibitorů KatK.

3) Pro sadu peptidomimetických inhibitorů na bázi vinylketonů a cyanohydrazidů byly připraveny komplexy s KatK, nalezeny jejich krystalizační podmínky a byla určena 3D struktura celkem šesti komplexů pomocí rentgenostrukturní analýzy.

4) Strukturní analýzou byl určen vazebný mód peptidomimetických vinylketonů a cyanohydrazidů a byly popsány klíčové interakce mezi inhibitorem a aktivním místem KatK. Poprvé byla určena struktura aktivačního meziprojektu KatK a jeho inhibiční interakce. Mezi charakterizovanými inhibitory je jeden z nejlepších známých inhibitorů KatK a unikátní inhibiční sonda pro detekci KatK.

5) Byly odvozeny vztahy struktury a funkce studovaných inhibitorů na základě integrace strukturních dat s výsledky získanými pomocí výpočetních metod, inhibičních analýz a buněčných testů.

8. Literární zdroje

- Ali, A., R. M. Bandaranayake, Y. Cai, N. M. King, M. Kolli, S. Mittal, J. F. Murzycki, M. N. Nalam, E. A. Nalivaika, A. Ozen, M. M. Prabu-Jeyabalan, K. Thayer, and C. A. Schiffer. 2010. 'Molecular Basis for Drug Resistance in HIV-1 Protease', *Viruses*, 2: 2509-35.
- Alliston, T., E. Piek and R. Derynck, 2008. 'TGF-beta Family Signaling in Skeletal Development, Maintenance, and Disease', *Cold Spring Harbor M*, 50: 667-24.
- Alves, M. F., L. Puzer, S. S. Cotrin, M. A. Juliano, L. Juliano, D. Brömme, and A. K. Carmona. 2003. 'S3 to S3' subsite specificity of recombinant human cathepsin K and development of selective internally quenched fluorescent substrates', *Biochem J*, 373: 981-86.
- Andrault, P. M., P. Panwar, N. C. W. Mackenzie, and D. Brömme. 2019. 'Elastolytic activity of cysteine cathepsins K, S, and V promotes vascular calcification', *Sci Rep*, 9: 1-13
- Arman, A., A. Bereket, A. Coker, P. Ö. S. Kiper, T. Güran, B. Özkan, and S. Turan. 2014. 'Cathepsin K analysis in a pycnodysostosis cohort: demographic, genotypic and phenotypic features', *Orphanet J Rare Dis*, 9: 1-8.
- Barrett, A. J., and N. D. Rawlings. 1996. 'Families and clans of cysteine peptidases', *Perspect drug discov*, 6: 1-11.
- Barrett, A. J., and P. M. Starkey. 1973. 'The interaction of α 2-macroglobulin with proteinases. Characteristics and specificity of the reaction, and a hypothesis concerning its molecular mechanism', *Biochem J*, 133: 709-24.
- Barrett, A. J., A. A. Kembhavi, M. A. Brown, H. Kirschke, C. G. Knight, M. Tamai, K. Hanada. 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: 189-98.
- Bernstein, H. G., and G. Keilhoff. 2018. 'Putative roles of cathepsin B in Alzheimer's disease pathology: The good, the bad, and the ugly in one?', *Neural Regen Res*, 13: 2100-01.
- Billington, C. J., P. Mason, M. C. Magny, and J. S. Mort. 2000. 'The slow-binding inhibition of cathepsin K by its propeptide', *Biochem Biophys Res Commun*, 276: 924-9.
- Bogyo, M. 2010. 'Finding enzymes that are actively involved in cancer', *P Natl Acad Sci USA*, 107: 2379-80.
- Bohley, P., and P. O. Seglen. 1992. 'Proteases and proteolysis in the lysosome', *Experientia*, 48: 151-57.
- Bond, J. S. 2019. 'Proteases: History, discovery, and roles in health and disease', *J Biol Chem*, 294: 1643-51.
- Brix, K., A. Dunkhorst, K. Mayer and S. Jordans. 2008. 'Cysteine cathepsins: cellular roadmap to different functions', *Biochimie*, 90: 194-07.
- Brömme, D. and K. Okamoto. 1995. 'Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution', *Biol Chem Hoppe Seyler*, 376: 379-84.
- Brömme, D., K. Okamoto, B. B. Wang, and S. Biroc. 1996. 'Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O2 in *Spodoptera frugiperda* and characterization of the enzyme', *J Biol Chem*, 271: 2126-32.
- Brömme, D. 2013. 'Cathepsin K.' in, *Handbook of Proteolytic Enzymes*. (Third Edition), Academic Press, 1801-07
- Brömme, D., and S. Wilson. 2011. 'Role of Cysteine Cathepsins in Extracellular Proteolysis.' in, *Extracellular Matrix Degradation*, (Springer), 23-51

- Brubaker K. D., R. L. Vessella, L. D. True, R. Thomas and E. Corey. 2003. 'Cathepsin K mRNA and protein expression in prostate cancer progression', *J Bone Miner Res*, 18: 222-30.
- Bühling, F., A. Reisenauer, A. Gerber, S. Krüger, E. Weber, D. Brömme and C. H. Röcken. 2001. 'Cathepsin K a marker of macrophage differentiation?', *J Pathol*, 195: 375-82.
- Bühling, F., A. Fengler, W. Brandt, T. Welte, S. Ansorge, and D. K. Nagler. 2002. 'Novel cysteine proteases of the papain family.' in, *Cellular Peptidases in Immune Functions and Diseases 2* (Springer), 477: 241-54
- Burger, M. L., K. K. Leung, M. J. Bennett, and A. Winoto. 2014. 'T cell-specific inhibition of multiple apoptotic pathways blocks negative selection and causes autoimmunity', *eLife*, 3: e03468
- Claus, V., A. Jahraus, T. Tjelle, T. Berg, H. Kirschke, H. Faulstich, and G. Griffiths. 1998. 'Lysosomal enzyme trafficking between phagosomes, endosomes, and lysosomes in J774 macrophages enrichment of cathepsin H in early endosomes', *J Biol Chem*, 273: 9842-51.
- Corisdeo, S., M. Gyda, M. Zaidi, B. S. Moonga, and B. R. Troen. 2001. 'New insights into the regulation of cathepsin K gene expression by osteoprotegerin ligand', *Biochem Biophys Res Commun*, 285: 335-9.
- Cstorner, A., and R. Ménard. 1994. 'Catalytic mechanism in papain family of cysteine peptidases.' in, *Methods in enzymology* (Elsevier), 244: 1831-34
- Dall, E., and H. Brandstetter. 2016. 'Structure and function of legumain in health and disease', *Biochimie*, 122: 126-50.
- Dall, E. and H. Brandstetter. 2013. 'Mechanistic and structural studies on legumain explain its zymogenicity, distinct activation pathways, and regulation', *Proc Natl Acad Sci USA*, 110: 10940-5.
- Dauth, S., M. Arampatzidou, M. Rehders, D. M. T. Yu, D. Führer, and K. Brix. 2011. 'Thyroid Cathepsin K: Roles in Physiology and Thyroid Disease', *Clin Rev Bone Miner Metab*, 9: 94-106.
- Davis, T. L., J. R. Walker, P. J. Finerty, Jr., F. Mackenzie, E. M. Newman, and S. Dhe-Paganon. 2007. 'The crystal structures of human calpains 1 and 9 imply diverse mechanisms of action and auto-inhibition', *J Mol Biol*, 366: 216-29.
- Decock J, N. Obermajer, S. Vozelj, W. Hendrickx, R. Paridaens and J. Kos. 2008. 'Cathepsin B, Cathepsin H, Cathepsin X and Cystatin C in Sera of Patients with Early-Stage and Inflammatory Breast Cancer.', *Int J Biol Markers*, 28: 161-68.
- Dejica, V. M., J. S. Mort, S. Laverty, J. Antoniou, D. J. Zukor, M. Tanzer and A. R. Poole. 2012. 'Increased type II collagen cleavage by cathepsin K and collagenase activities with aging and osteoarthritis in human articular cartilage', *Arthritis Res Ther*, 14: 1-9.
- DeLano, W. L. 2002. 'Pymol: An open-source molecular graphics tool.', *CCP4 Newsletter on protein crystallography*, 40: 82-92.
- Desmarais, S., W. C. Black, R. Oballa, S. Lamontagne, D. Riendeau, P. Tawa, L. T. Duong, M. Pickarski, and M. D. Percival. 2008. 'Effect of cathepsin k inhibitor basicity on in vivo off-target activities', *Mol Pharmacol*, 73: 147-56.
- Dodds, R. A., J. R. Connor, F. H. Drake and M. Gowen. 1999. 'Expression of cathepsin K messenger RNA in giant cells and their precursors in human osteoarthritic synovial tissues', *Arthritis Rheum*, 42: 1588-93.
- Drake, M. T., B. L. Clarke, M. J. Oursler, and S. Khosla. 2017. 'Cathepsin K Inhibitors for Osteoporosis: Biology, Potential Clinical Utility, and Lessons Learned', *Endocr Rev*, 38: 325-50.

- Drake, M. T., B. L. Clarke, and S. Khosla. 2008. 'Bisphosphonates: mechanism of action and role in clinical practice', *Mayo Clinic proceedings*, 83: 1032-45.
- Duncan, E. M., T. L. Muratore-Schroeder, R. G. Cook, B. A. Garcia, J. Shabanowitz, D. F. Hunt, and C. D. Allis. 2008. 'Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation', *Cell*, 135: 284-94.
- Falgueyret, J. P., S. Desmarais, R. Oballa, W. C. Black, W. Cromlish, K. Khougaz, M. D. Percival. 2005. 'Lysosomotropism of basic cathepsin K inhibitors contributes to increased cellular potencies against off-target cathepsins and reduced functional selectivity', *J Med Chem*, 48: 7535-43.
- Fasanya, H. O., and D. W. Siemann. 2020. 'The Role of Cathepsins in the Growth of Primary and Secondary Neoplasia in the Bone', *Osteology*, 1: 3-28.
- Fink, S. L., and B. T. Cookson. 2005. 'Apoptosis, pyroptosis, and necrosis: mechanistic description of dead and dying eukaryotic cells', *Infect Immun*, 73: 1907-16.
- Frizler, M., F. Lohr, N. Furtmann, J. Klas, and M. Gütschow. 2011. 'Structural optimization of azadipeptide nitriles strongly increases association rates and allows the development of selective cathepsin inhibitors', *J Med Chem*, 54: 396-400.
- Frizler, M., M. Stirnberg, M. T. Sisay, and M. Gütschow. 2010. 'Development of nitrile-based peptidic inhibitors of cysteine cathepsins', *Curr Top Med Chem*, 10: 294-322.
- Garnero, P., O. Borel, I. Byrjalsen, M. Ferreras, F. H. Drake, M. S. McQueney, N. T. Foged, P. D. Delmas, and J. M. Delaisse. 1998. 'The collagenolytic activity of cathepsin K is unique among mammalian proteinases', *J Biol Chem*, 273: 32347-52.
- Gelb, B. D., G. P. Shi, H. A. Chapman, and R. J. Desnick. 1996. 'Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency', *Science*, 273: 1236-38.
- Ghosh, A. K., I. Samanta, A. Mondal, and W. R. Liu. 2019. 'Covalent Inhibition in Drug Discovery', *Chem Med Chem*, 14: 889-906.
- Gocheva, V., W. Zeng, D. Ke, D. Klimstra, T. Reinheckel, C. Peters, D. Hanahan, and J. A. Joyce. 2006. 'Distinct roles for cysteine cathepsin genes in multistage tumorigenesis', *Genes Dev*, 20: 543-56.
- Golds, G., D. Houdek, and T. Arnason. 2017. 'Male Hypogonadism and Osteoporosis: The Effects, Clinical Consequences, and Treatment of Testosterone Deficiency in Bone Health', *Int J Endocrinol*, 2017: 4602129.
- Gormley, J. A., S. M. Hegarty, A. O'Grady, M. R. Stevenson, R. E. Burden, H. L. Barrett, C. J. Scott, J. A. Johnston, R. H. Wilson, E. W. Kay, P. G. Johnston, and S. A. Olwill. 2011. 'The role of Cathepsin S as a marker of prognosis and predictor of chemotherapy benefit in adjuvant CRC: a pilot study', *Br J Cancer*, 105: 1487-94.
- Gowen, M., F. Lazner, R. Dodds, R. Kapadia, J. Feild, M. Tavarria, and I. Kola. 1999. 'Cathepsin K knockout mice develop osteopetrosis due to a deficit in matrix degradation but not demineralization', *J Bone Miner Res*, 14: 1654-63.
- Hanada, K., M. Tamai, M. Yamagishi, S. Ohmura, J. Sawada, and I. Tanaka. 1978. 'Isolation and Characterization of E-64, a New Thiol Protease Inhibitor', *Agric Biol Chem*, 42: 523-28.
- Horn, M., M. Nussbaumerová, M. Sanda, Z. Kovářová, J. Srba, Z. Franta, D. Sojka, M. Bogyo, C. R. Caffrey, P. Kopáček, and M. Mareš. 2009. 'Hemoglobin digestion in blood-feeding ticks: mapping a multi-peptidase pathway by functional proteomics', *Chem Biol*, 16: 1053-63.
- Hou, W. S., W. Li, G. Keyszer, E. Weber, R. Levy, M. J. Klein, E. M. Gravallese, S. R. Goldring, and D. Brömme. 2002. 'Comparison of cathepsins K and S expression within the rheumatoid and osteoarthritic synovium', *Arthritis Rheum*, 46: 663-74.
- Houard, X., M. B. Goldring, and F. Berenbaum. 2013. 'Homeostatic mechanisms in articular cartilage and role of inflammation in osteoarthritis', *Curr Rheumatol Rep*, 15: 375.

- Husmann, K., R. Muff, M. E. Bolander, G. Sarkar, W. Born, and B. Fuchs. 2008. 'Cathepsins and osteosarcoma: Expression analysis identifies cathepsin K as an indicator of metastasis', *Mol Carcinog*, 47: 66-73.
- Chia-en, A. C., W. Chen, and M. K. Gilson. 2007. 'Ligand configurational entropy and protein binding', *Proc Natl Acad Sci*, 104: 1534-39.
- Chingle, R., C. Proulx, and W. D. Lubell. 2017. 'Azapeptide Synthesis Methods for Expanding Side-Chain Diversity for Biomedical Applications', *Acc Chem Res*, 50: 1541-56.
- Choe, Y., F. Leonetti, D. C. Greenbaum, F. Lecaille, M. Bogyo, D. Brömme, J. A. Ellman, and C. S. Craik. 2006. 'Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities', *J Biol Chem*, 281: 12824-32.
- Inaoka, T., Bilbe, G., Ishibashi, O., K. I. Tezuka, M. Kumegawa, and T. Kokubo. 1995. 'Molecular cloning of human cDNA for cathepsin K: novel cysteine proteinase predominantly expressed in bone', *Biochem Biophys Res Commun*, 206: 89-96.
- Irving, J. A., S. S. Shushanov, R. N. Pike, E. Y. Popova, D. Brömme, T. H. Coetzer, S. P. Bottomley, I. A. Boulyenko, S. A. Grigoryev, and J. C. Whisstock. 2002. 'Inhibitory activity of a heterochromatin-associated serpin (MENT) against papain-like cysteine proteinases affects chromatin structure and blocks cell proliferation', *J Biol Chem*, 277: 13192-201.
- Jayakumar, A., Y. A. Kang, M. J. Frederick, S. C. Pak, Y. Henderson, P. R. Holton, and G. L. Clayman. 2003. 'Inhibition of the cysteine proteinases cathepsins K and L by the serpin headpin (SERPINB13): a kinetic analysis', *Arch Biochem Biophys*, 409: 367-74.
- Jílková, A., M. Horn, J. Fanfrlík, J. Kuppers, P. Páchl, P. Řezáčová, M. Lepšík, P. Fajtová, P. Rubešová, M. Chanová, C. R. Caffrey, M. Gütschow, and M. Mareš. 2021. 'Azanitrile Inhibitors of the SmCB1 Protease Target Are Lethal to *Schistosoma mansoni*: Structural and Mechanistic Insights into Chemotype Reactivity', *ACS Infect Dis*, 7: 189-201.
- Jílková, A., M. Horn, P. Řezáčová, L. Marešová, P. Fajtová, J. Brynda, J. Vondrášek, J. H. McKerrow, C. R. Caffrey, and M. Mareš. 2014. 'Activation route of the *Schistosoma mansoni* cathepsin B1 drug target: structural map with a glycosaminoglycan switch', *Structure*, 22: 1786-98.
- Jones, M., S. Denieffe, C. Griffin, W. Tinago, and M. C. Fitzgibbon. 2017. 'Evaluation of cystatin C in malignancy and comparability of estimates of GFR in oncology patients', *Pract Lab Med*, 8: 95-104.
- Kabsch, W. 2010. 'Integration, scaling, space-group assignment and post-refinement', *Acta Crystallogr D Biol Crystallogr*, 66: 133-44.
- Kanis, J. A., N. Norton, N. C. Harvey, T. Jacobson, H. Johansson, M. Lorentzon, E. V. McCloskey, C. Willers, and F. Borgstrom. 2021. 'SCOPE 2021: a new scorecard for osteoporosis in Europe', *Arch Osteoporos*, 16: 82.
- Kappelhoff, R., X. S. Puente, C. H. Wilson, A. Seth, C. López-Otín, and C. M. Overall. 2017. 'Overview of transcriptomic analysis of all human proteases, non-proteolytic homologs and inhibitors: Organ, tissue and ovarian cancer cell line expression profiling of the human protease degradome by the CLIP-CHIP DNA microarray', *Biochim Biophys Acta Mol Cell Res*, 1864: 2210-19.
- Kasperkiewicz, P. 2021. 'Peptidyl Activity-Based Probes for Imaging Serine Proteases', *Front Chem*, 9: 639410.
- Kerr, I. D., J. H. Lee, C. J. Farady, R. Marion, M. Rickert, M. Sajid, K. C. Pandey, C. R. Caffrey, J. Legac, E. Hansell, J. H. McKerrow, C. S. Craik, P. J. Rosenthal, and L.

- S. Brinen. 2009. 'Vinyl sulfones as antiparasitic agents and a structural basis for drug design', *J Biol Chem*, 284: 25697-703.
- Kinoshita, T. 2020. 'Biosynthesis and biology of mammalian GPI-anchored proteins', *Open Biol*, 10: 190290.
- Kozawa, E., X. W. Cheng, H. Urakawa, E. Arai, Y. Yamada, S. Kitamura, K. Sato, M. Kuzuya, N. Ishiguro, and Y. Nishida. 2016. 'Increased expression and activation of cathepsin K in human osteoarthritic cartilage and synovial tissues', *J Orthop Res*, 34: 127-34.
- Kryczka, J., and J. Boncela. 2017. 'Proteases Revisited: Roles and Therapeutic Implications in Fibrosis', *Mediators Inflamm*, 2017: 2570154.
- Kumar, A. 1988. 'Pycnodysostosis', *Int Orthop*, 12: 261-63.
- Kumar, S., L. Dare, J. A. Vasko-Moser, I. E. James, S. M. Blake, D. J. Rickard, S. M. Hwang, T. Tomaszek, D. S. Yamashita, R. W. Marquis, H. Oh, J. U. Jeong, D. F. Veber, M. Gowen, M. W. Lark, and G. Stroup. 2007. 'A highly potent inhibitor of cathepsin K (relacatib) reduces biomarkers of bone resorption both in vitro and in an acute model of elevated bone turnover in vivo in monkeys', *Bone*, 40: 122-31.
- LaLonde, J. M., B. Zhao, C. A. Janson, K. J. D'Alessio, M. S. McQueney, M. J. Orsini, and W. W. Smith. 1999. 'The crystal structure of human procathepsin K', *Biochemistry*, 38: 862-69.
- Kovacs, L. Su, Y. 2014. 'The Critical Role of Calpain in Cell Proliferation', *J Biomol Res Ther*, 03: 1000112
- Law, S., P. M. Andrault, A. H. Aguda, N. T. Nguyen, N. Kruglyak, G. D. Brayer, and D. Brömme. 2017. 'Identification of mouse cathepsin K structural elements that regulate the potency of odanacatib', *Biochem J*, 474: 851-64.
- Lecaille, F., D. Brömme, and G. Lalmanach. 2008. 'Biochemical properties and regulation of cathepsin K activity', *Biochimie*, 90: 208-26.
- Lecaille, F., E. Weidauer, M. A. Juliano, D. Brömme, and G. Lalmanach 2003. 'Probing cathepsin K activity with a selective substrate spanning its active site', *Biochem J*, 375: 307-12.
- Lemaire, P. A., L. Huang, Y. Zhuo, J. Lu, C. Bahnck, S. J. Stachel, S. S. Carroll, and L. T. Duong. 2014. 'Chondroitin sulfate promotes activation of cathepsin K', *J Biol Chem*, 289: 21562-72.
- Lemke, C., J. Benýšek, D. Brajtenbach, C. Breuer, A. Jílková, M. Horn, M. Buša, L. Ulrychova, A. Illies, K. F. Kubatzky, U. Bartz, M. Mareš, and M. Gütschow. 2021. 'An Activity-Based Probe for Cathepsin K Imaging with Excellent Potency and Selectivity', *J Med Chem*, 64: 13793-806.
- Leto, G., M. Crescimanno, and C. Flandina. 2018. 'On the role of cystatin C in cancer progression', *Life Sci*, 202: 152-60.
- Leung-Toung, R., J. Wodzinska, W. Li, J. Lowrie, R. Kukreja, D. Desilets, K. Karimian, and T. F. Tam. 2003. '1,2,4-thiadiazole: a novel Cathepsin B inhibitor', *Bioorg Med Chem*, 11: 5529-37.
- Leung, D., G. Abbenante, and D. P. Fairlie. 2000. 'Protease inhibitors: current status and future prospects', *J Med Chem*. 43: 305-41
- Li, F. X., J. J. Liu, F. Xu, X. Lin, J. Y. Zhong, F. Wu, and L. Q. Yuan. 2019. 'Role of tumor-derived exosomes in bone metastasis', *Oncol Lett*, 18: 3935-45.
- Li, J., I. Sarosi, X. Q. Yan, S. Morony, C. Capparelli, H. L. Tan, and W. J. 2000. 'RANK is the intrinsic hematopoietic cell surface receptor that controls osteoclastogenesis and regulation of bone mass and calcium metabolism', *Proc Natl Acad Sci*, 97: 1566-71.
- Li, Y., Z. H. Gong, L. Sheng, Y. T. Gong, X. Y. Tan, W. M. Li, D. L. Dong, B. F. Yang, S. B. Fu, and H. J. Xue. 2009. 'Anti-apoptotic effects of a calpain inhibitor on

- cardiomyocytes in a canine rapid atrial fibrillation model', *Cardiovasc Drugs Ther*, 23: 361-8.
- Li, Y. P., M. Alexander, A. L. Wucherpfennig, P. Yelick, W. Chen, and P. Stashenko. 1995. 'Cloning and complete coding sequence of a novel human cathepsin expressed in giant cells of osteoclastomas', *J Bone Miner Res*, 10: 1197-202.
- Li, Z., W. S. Hou, C. R. Escalante-Torres, B. D. Gelb, and D. Brömme. 2002. 'Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate', *J Biol Chem*, 277: 28669-76.
- Li, Z., Y. Yasuda, W. Li, M. Bogyo, N. Katz, R. E. Gordon, G. B. Fields, and D. Brömme. 2004. 'Regulation of collagenase activities of human cathepsins by glycosaminoglycans', *J Biol Chem*, 279: 5470-9.
- Liang, W., F. Wang, Q. Chen, J. Dai, J. Escara-Wilke, E. T. Keller, J. Zimmermann, N. Hong, Y. Lu, and J. Zhang. 2019. 'Targeting cathepsin K diminishes prostate cancer establishment and growth in murine bone', *J Cancer Res Clin Oncol*, 145: 1999-2012.
- Lin, X., S. Patil, Y. G. Gao, and A. Qian. 2020. 'The Bone Extracellular Matrix in Bone Formation and Regeneration', *Front Pharmacol*, 11: 757.
- Littlewood-Evans, A., J. Bilbe, G. Bowler, W. B. Farley, D. Wlodarski, B. Kokubo, and J. A. Gallagher. 1997a. 'The osteoclast-associated protease cathepsin K is expressed in human breast carcinoma', *Cancer Res*, 57: 5386-90.
- Littlewood-Evans, A., T. Kokubo, O. Ishibashi, T. Inaoka, B. Wlodarski, J. A. Gallagher, and G. Bilbe. 1997b. 'Localization of cathepsin K in human osteoclasts by in situ hybridization and immunohistochemistry', *Bone*, 20: 81-86.
- Long, F., R. A. Nicholls, P. Emsley, S. Graažulis, A. Merkys, A. Vaitkus, and G. N. Murshudov. 2017. 'AceDRG: a stereochemical description generator for ligands', *Acta Crystallogr D Struct Biol*, 73: 112-22.
- López-Otín, C., and J. S. Bond. 2008. 'Proteases: multifunctional enzymes in life and disease', *J Biol Chem*, 283: 30433-7.
- Löser, R., M. Frizler, K. Schilling, and M. Gütschow. 2008. 'Azadipeptide nitriles: highly potent and proteolytically stable inhibitors of papain-like cysteine proteases', *Angew Chem Int Ed Engl*, 47: 4331-4.
- Lu, J., M. Wang, Z. Wang, Z. Fu, A. Lu, and G. Zhang. 2018. 'Advances in the discovery of cathepsin K inhibitors on bone resorption', *J Enzyme Inhib Med Chem*, 33: 890-04.
- Lucas, E. C., and A. Williams. 1969. 'pH dependencies of individual rate constants in papain-catalyzed reactions', *Biochemistry*, 8: 5125-35.
- Makras, P., S. Delaroudis, and A. D. Anastasilakis. 2015. 'Novel therapies for osteoporosis', *Metabolism*, 64: 1199-214.
- Martinez, A., M. Alonso, A. Castro, C. Pérez, and F. J. Moreno. 2002. 'First non-ATP competitive glycogen synthase kinase 3 β (GSK-3 β) inhibitors: thiadiazolidinones (TDZD) as potential drugs for the treatment of Alzheimer's disease', *J Med Chem*, 45: 1292-99.
- Mason, R. W. 1989. 'Interaction of lysosomal cysteine proteinases with α 2-macroglobulin: Conclusive evidence for the endopeptidase activities of cathepsins B and H', *Arch Biochem Biophys*, 273: 367-74.
- Masoumeh H., H. Najmabadi, K. Kahrizi. 2018. 'Calpains: Diverse Functions but Enigmatic', *Arch Iran Med*, 21: 170-79.
- Matsumoto, K., K. Mizoue, K. Kitamura, W. C. Tse, C. P. Huber, and T. Ishida. 1999. 'Structural basis of inhibition of cysteine proteases by E-64 and its derivatives', *Peptide Sci*, 51: 99-107.

- McGrath, M. E., J. L. Klaus, M. G. Barnes, and D. Brömme. 1997. 'Crystal structure of human cathepsin K complexed with a potent inhibitor', *Nature structural biology*, 4: 105-09.
- McQueney, M. S., B. Y. Amegadzie, K. D'Alessio, C. R. Hanning, M. M. McLaughlin, D. McNulty, S. A. Carr, C. Ijames, J. Kurdyla, and C. S. Jones. 1997. 'Autocatalytic activation of human cathepsin K', *J Biol Chem*, 272: 13955-60.
- Meh, P., M. Pavšič, V. Turk, A. Baici, and B. Lenarčič. 2005. 'Dual concentration-dependent activity of thyroglobulin type-1 domain of testican: specific inhibitor and substrate of cathepsin L', *Biol Chem*, 386: 75-83.
- Mihelič, M., and D. Turk. 2007. 'Two decades of thyroglobulin type-1 domain research', *Biol Chem*, 388: 1123-30.
- Mons, E., I. D. C. Jansen, J. Loboda, B. R. van Doodewaerd, J. Hermans, M. Verdoes, C. A. A. van Boeckel, P. A. van Veelen, B. Turk, D. Turk, and H. Ovaa. 2019. 'The Alkyne Moiety as a Latent Electrophile in Irreversible Covalent Small Molecule Inhibitors of Cathepsin K', *J Am Chem Soc*, 141: 3507-14.
- Ndao, M., C. Beaulieu, W. C. Black, E. Isabel, F. Vasquez-Camargo, M. Nath-Chowdhury, F. Massé, C. Mellon, N. Methot, and D. A. Nicoll-Griffith. 2014. 'Reversible cysteine protease inhibitors show promise for a Chagas disease cure', *Antimicrob Agents Chemother*, 58: 1167-78.
- Novinec, M., R. N. Grass, W. J. Stark, V. Turk, A. Baici, and B. Lenarčič. 2007. 'Interaction between human cathepsins K, L, and S and elastins: mechanism of elastinolysis and inhibition by macromolecular inhibitors', *J Biol Chem*, 282: 7893-902.
- Novinec, M., L. Kovačič, B. Lenarčič, and A. Baici. 2010. 'Conformational flexibility and allosteric regulation of cathepsin K', *Biochem J*, 429: 379-89.
- Novinec, M., and B. Lenarčič. 2013. 'Cathepsin K: a unique collagenolytic cysteine peptidase', *Biol Chem*, 394: 1163-79.
- Novinec, M., B. Lenarčič, and B. Turk. 2014. 'Cysteine cathepsin activity regulation by glycosaminoglycans', *Biomed Res Int*, 2014: 309718.
- Novinec, M., M. Rebernik, and B. Lenarčič. 2016. 'An allosteric site enables fine-tuning of cathepsin K by diverse effectors', *FEBS Lett*, 590: 4507-18.
- Ottersbach, P. A., G. Schnakenburg, and M. Gütschow. 2012. 'Induction of chirality: experimental evidence of atropisomerism in azapeptides', *Chem Commun*, 48: 5772-4.
- Panwar, P., S. Law, A. Jamroz, P. Azizi, D. Zhang, M. Ciufolini, and D. Brömme. 2018. 'Tanshinones that selectively block the collagenase activity of cathepsin K provide a novel class of ectosteric antiresorptive agents for bone', *Br J Pharmacol*, 175: 902-23.
- Park, P. S. U., W. Y. Raynor, Y. Sun, T. J. Werner, C. S. Rajapakse, and A. Alavi. 2021. '(18)F-Sodium Fluoride PET as a Diagnostic Modality for Metabolic, Autoimmune, and Osteogenic Bone Disorders: Cellular Mechanisms and Clinical Applications', *Int J Mol Sci*, 22: 6504
- Patel, S., A. Rauf, H. Khan, and T. Abu-Izneid. 2017. 'Renin-angiotensin-aldosterone (RAAS): The ubiquitous system for homeostasis and pathologies', *Biomed Pharmacother*, 94: 317-25.
- Pomeislová, A., M. Otmar, P. Rubešová, J. Benýšek, M. Matoušová, H. Mertlíková-Kaiserová, R. Pohl, L. Poštová Slavětínská, K. Pomeisl, and M. Krečmerová. 2021. '1,2,4-Thiadiazole acyclic nucleoside phosphonates as inhibitors of cysteine dependent enzymes cathepsin K and GSK-3beta', *Bioorg Med Chem*, 32: 115998.

- Poreba, M., K. Groborz, M. Vizovisek, M. Maruggi, D. Turk, B. Turk, G. Powis, M. Drag, and G. S. Salvesen. 2019. 'Fluorescent probes towards selective cathepsin B detection and visualization in cancer cells and patient samples', *Chem Sci*, 10: 8461-77.
- Poreba, M., W. Rut, M. Vizovisek, K. Groborz, P. Kasperkiewicz, D. Finlay, K. Vuori, D. Turk, B. Turk, G. S. Salvesen, and M. Drag. 2018. 'Selective imaging of cathepsin L in breast cancer by fluorescent activity-based probes', *Chem Sci*, 9: 2113-29.
- Proulx, C., Sabatino, D., Hopewell, R., Spiegel, J., García Ramos, Y., and Lubell, W. D. 2011. 'Azapeptides and their therapeutic potential', *Future Med Chem*, 3: 1139-64.
- Prunk, M., M. P. Nanut, J. Sabotič, and Janko Kos. 2017. 'Cystatins, cysteine peptidase inhibitors, as regulators of immune cell cytotoxicity', *Period Biol*, 118: 353-2
- Pungerčar, J. R., D. Caglič, M. Sajid, M. Dolinar, O. Vasiljeva, U. Požgan, D. Turk, M. Bogyo, V. Turk, and B. Turk. 2009. 'Autocatalytic processing of procathepsin B is triggered by proenzyme activity', *FEBS J*, 276: 660-8.
- Qadir, A., S. Liang, Z. Wu, Z. Chen, L. Hu, and A. Qian. 2020. 'Senile Osteoporosis: The Involvement of Differentiation and Senescence of Bone Marrow Stromal Cells', *Int J Mol Sci*, 21: 349
- Qiu, Z. C., X. L. Dong, Y. Dai, G. K. Xiao, X. L. Wang, K. C. Wong, M. S. Wong, and X. S. Yao. 2016. 'Discovery of a New Class of Cathepsin K Inhibitors in Rhizoma Drynariae as Potential Candidates for the Treatment of Osteoporosis', *Int J Mol Sci*, 17: 2116
- Rawlings, N. D. 2020. 'Twenty-five years of nomenclature and classification of proteolytic enzymes', *Biochim Biophys Acta Proteins Proteom*, 1868: 140345.
- Rawlings, N. D., A. J. Barrett, P. D. Thomas, X. Huang, A. Bateman, and R. D. Finn. 2018. 'The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database', *Nucleic Acids Res*, 46: 624-32.
- Rawlings, N. D. 2013. *Proteases: Structure and Function* (First edition, Springer). 1-36
- Renko, M., U. Požgan, D. Majera, and D. Turk. 2010. 'Stefin A displaces the occluding loop of cathepsin B only by as much as required to bind to the active site cleft', *FEBS J*, 277: 4338-45.
- Rocha, G. V., L. S. Bastos, and M. G. S. Costa. 2020. 'Identification of potential allosteric binding sites in cathepsin K based on intramolecular communication', *Proteins*, 88: 1675-87.
- Rousselle, A. V., and Heymann, D. 2002. 'Osteoclastic acidification pathways during bone resorption', *Bone*, 30: 533-40.
- Rzychon, M., D. Chmiel, and J. Steć-Niemczyk. 2004. 'Modes of inhibition of cysteine proteases', *Acta Biochim Pol*, 51: 861-73.
- Saftig, P., Hunziker, E., Wehmeyer, O., Jones, S., Boyde, A., Rommerskirch, W., and Von Figura, K. 1998. 'Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice', *Proc Natl Acad Sci*, 95: 13453-58.
- Salentin, S., S. Schreiber, V. J. Haupt, M. F. Adasme, and M. Schroeder. 2015. 'PLIP: fully automated protein-ligand interaction profiler', *Nucleic Acids Res*, 43: 443-7.
- Sanman, L. E., and M. Bogyo. 2014. 'Activity-based profiling of proteases', *Annu Rev Biochem*, 83: 249-73.
- Santos, M. M., and Rui Moreira. 2007. 'Michael acceptors as cysteine protease inhibitors', *Mini-Rev Med Chem*, 7: 1040-50.
- Sharma, V., P. Panwar, A. J. O'Donoghue, H. Cui, R. V. Guido, C. S. Craik, and D. Brömme. 2015. 'Structural requirements for the collagenase and elastase activity of cathepsin K and its selective inhibition by an exosite inhibitor', *Biochem J*, 465: 163-73.

- Shi, G. P., Chapman, H. A., Bhairi, S. M., DeLeeuw, C., Reddy, V. Y., and Weiss, S. J. 1995. 'Molecular cloning of human cathepsin O, a novel endoproteinase and homologue of rabbit OC2', *FEBS Lett*, 357: 129-34.
- Shi, G. P., J. S. Munger, J. P. Meara, D. H. Rich, and H. A. Chapman. 1992. 'Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease', *J Biol Chem*, 267: 7258-62.
- Shoulders, M. D., and R. T. Raines. 2009. 'Collagen structure and stability', *Annu Rev Biochem*, 78: 929-58.
- Schechter, I., and A. Berger. 1967. 'On the size of the active site in proteases. I. Papain', *Biochem Biophys Res Commun*, 27: 157-62.
- Schick, C., P. A. Pemberton, G. P. Shi, Y. Kamachi, S. Çataltepe, A. J. Bartuski, G. A. Silverman. 1998. 'Cross-class inhibition of the cysteine proteinases cathepsins K, L, and S by the serpin squamous cell carcinoma antigen 1: a kinetic analysis', *Biochemistry*, 37: 5258-66.
- Schmidt, G. S., J. P. Schacht, T. S. Knee, M. K. M. Shakir, and T. D. Hoang. 2020. 'Pyknodysostosis (Osteopetrosis Acro-Osteolytica)', *AACE Clin Case Rep*, 6: 257-61.
- Schmitz, J., A. M. Beckmann, A. Dudic, T. Li, R. Sellier, U. Bartz, and M. Gütschow. 2014. '3-Cyano-3-aza-beta-amino Acid Derivatives as Inhibitors of Human Cysteine Cathepsins', *ACS Med Chem Lett*, 5: 1076-81.
- Silva, J. R. A., L. Cianni, D. Araujo, P. H. J. Batista, D. deVita, F. Rosini, A. Leitao, J. Lameira, and C. A. Montanari. 2020. 'Assessment of the Cruzain Cysteine Protease Reversible and Irreversible Covalent Inhibition Mechanism', *J Chem Inf Model*, 60: 1666-77.
- Singh, J., R. C. Petter, T. A. Baillie, and A. Whitty. 2011. 'The resurgence of covalent drugs', *Nat Rev Drug Discov*, 10: 307-17.
- Sivaraman, J., Lalumière, M., Ménard, R., Cygler, M. 1999. 'Crystal structure of wild-type human procathepsin K', *Protein Sci*, 8: 283-90.
- Streicher, C., A. Heyny, O. Andrukhova, B. Haigl, S. Slavic, C. Schüler, K. Kollmann, I. Kantner, V. Sexl, M. Kleiter, L. C. Hofbauer, P. J. Kostenuik, and R. G. Erben. 2017. 'Estrogen Regulates Bone Turnover by Targeting RANKL Expression in Bone Lining Cells', *Sci Rep*, 7: 6460.
- Strelow, J. M. 2017. 'A Perspective on the Kinetics of Covalent and Irreversible Inhibition', *SLAS Discov*, 22: 3-20.
- Stubbs, M. T., B. Laber, W. Bode, R. Huber, R. Jerala, B. Lenarčič, and V. Turk. 1990. 'The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction', *EMBO J*, 9: 1939-47.
- Choi, S. J., N. Kurihara, Y. Oba, and G. D. Roodman. 2001. 'Osteoclast Inhibitory Peptide 2 Inhibits Osteoclast Formation via Its C-Terminal Fragment', *J Bone Miner Res*, 16: 1804-11.
- Tam, T. F., Leung-Toung, R., Li, W., Spino, M., K. Karimian. 2005. 'Medicinal chemistry and properties of 1, 2, 4-thiadiazoles', *Mini-Rev Med Chem*, 5: 367-79.
- Tezuka, K., Y. Tezuka, A. Maejima, T. Sato, K. Nemoto, H. Kamioka, Y. Hakeda, and M. Kumegawa. 1994. 'Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts', *J Biol Chem*, 269: 1106-09.
- Tu, K. N., Lie, J. D., Wan, C. K. V., Cameron, M., Austel, A. G., Nguyen, J. K., and D. Hyun. 2018. 'Osteoporosis: a review of treatment options', *Pharm Ther*, 43: 92-104.
- Turk, V., Brzin, J., Lenarčič, B., Šali, A., W. Machleidt. 1986. *Human stefins and cystatins: their properties and structural relationships* (de Gruyter: Yugoslavia), 429-42.

- Turk, V., V. Stoka and D. Turk. 2008. 'Cystatins: biochemical and structural properties, and medical relevance', *Front Biosci*, 13: 5780-86.
- Turk, V., V. Stoka, O. Vasiljeva, M. Renko, T. Sun, B. Turk, and D. Turk. 2012. 'Cysteine cathepsins: from structure, function and regulation to new frontiers', *Biochim Biophys Acta*, 1824: 68-88.
- Tusar, L., A. Usenik, B. Turk, and D. Turk. 2021. 'Mechanisms Applied by Protein Inhibitors to Inhibit Cysteine Proteases', *Int J Mol Sci*, 22: 997
- Van Opdenbosch, N., and M. Lamkanfi. 2019. 'Caspases in Cell Death, Inflammation, and Disease', *Immunity*, 50: 1352-64.
- Vashum, Y., R. Premasingh, A. Kottaiswamy, M. Soma, A. Padmanaban, P. Kalaiselvan, and S. Samuel. 2021. 'Inhibitory effect of cathepsin K inhibitor (ODN-MK-0822) on invasion, migration and adhesion of human breast cancer cells in vitro', *Mol Biol Rep*, 48: 105-16.
- Vashum, Y., Z. Khashim, and F. B. M. Sheriff. 2020. 'Cathepsin K and Its Specific Target in Mediating Breast Cancer Skeletal Metastasis', *Canc Oncol Res*, 6: 36-46.
- Vasiljeva, O., M. Dolinar, J. R. Pungercar, V. Turk, and B. Turk. 2005. 'Recombinant human procathepsin S is capable of autocatalytic processing at neutral pH in the presence of glycosaminoglycans', *FEBS Lett*, 579: 1285-90.
- Vega-Tejido, M. A., S. E. Maluf, C. R. Bonturi, J. R. Sambrano, and O. N. Ventura. 2014. 'Theoretical insight into the mechanism for the inhibition of the cysteine protease cathepsin B by 1,2,4-thiadiazole derivatives', *J Mol Model*, 20: 2254.
- Verbovsek, U., C. J. Van Noorden, and T. T. Lah. 2015. 'Complexity of cancer protease biology: Cathepsin K expression and function in cancer progression', *Semin Cancer Biol*, 35: 71-84.
- Vernet, T., H. E. Khouri, P. Laflamme, D. C. Tessier, R. Musil, B. J. Gour-Salin, A. C. Storer, and D. Y. Thomas. 1991. 'Processing of the papain precursor. Purification of the zymogen and characterization of its mechanism of processing', *J Biol Chem*, 266: 21451-57.
- Villadangos, J. A., and H. L. Ploegh. 2000. 'Proteolysis in MHC class II antigen presentation: who's in charge?', *Immunity*, 12: 233-39.
- Wada, T., T. Nakashima, N. Hiroshi, and J. M. Penninger. 2006. 'RANKL-RANK signaling in osteoclastogenesis and bone disease', *Trends Mol Med*, 12: 17-25.
- Walsh, M. C., and Y. Choi. 2014. 'Biology of the RANKL-RANK-OPG System in Immunity, Bone, and Beyond', *Front Immunol*, 5: 511.
- Wilkinson, R. D., R. Williams, C. J. Scott, and R. E. Burden. 2015. 'Cathepsin S: therapeutic, diagnostic, and prognostic potential', *Biol Chem*, 396: 867-2.
- Wilson, S., S. L. Hashamiyan, P. Clarke, P. Saftig, J. Mort, V. M. Dejica, and D. Brömme. 2009. 'Glycosaminoglycan-mediated loss of cathepsin K collagenolytic activity in MPS I contributes to osteoclast and growth plate abnormalities', *Am J Pathol*, 175: 2053-62.
- Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson. 2011. 'Overview of the CCP4 suite and current developments', *Acta Crystallogr D Biol Crystallogr*, 67: 235-42.
- Wishart, D. S., C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, and J. Woolsey. 2006. 'DrugBank: a comprehensive resource for in silico drug discovery and exploration', *Nucleic Acids Res*, 34: 668-72.

- Wolk, K., G. Grutz, K. Witte, H. D. Volk, and R. Sabat. 2005. 'The expression of legumain, an asparaginyl endopeptidase that controls antigen processing, is reduced in endotoxin-tolerant monocytes', *Genes Immun*, 6: 452-6.
- Yang, H., J. Heyer, H. Zhao, S. Liang, R. Guo, and L. Zhong. 2020. 'The Potential Role of Cathepsin K in Non-Small Cell Lung Cancer', *Molecules*, 25: 4136
- Zavašnik-Bergant, T. and B. Turk. 2006. 'Cysteine cathepsins in the immune response', *HLA*, 67: 349-55.
- Zhang, J., K. G. Saag, and J. R. Curtis. 2011. 'Long-term safety concerns of antiresorptive therapy', *Rheum Dis Clin North Am*, 37: 387-40.
- Zhao, B., C. A. Janson, B. Y. Amegadzie, K. D'Alessio, C. Griffin, C. R. Hanning, and S. S. Abdel-Meguid. 1997. 'Crystal structure of human osteoclast cathepsin K complex with E-64', *Nat Struct Biol*, 4: 109-11.
- Zhu, M., B. Zhao, L. Wei, and S. Wang. 2021. 'alpha-2-Macroglobulin, a Native and Powerful Proteinase Inhibitor, Prevents Cartilage Degeneration Disease by Inhibiting Majority of Catabolic Enzymes and Cytokines', *Curr Mol Biol Rep*, 7: 1-7.

Příloha – publikací in extenso, které jsou podkladem disertace, včetně příloh

Publikace č. 1

1,2,4-Thiadiazole acyclic nucleoside phosphonates as inhibitors of cysteine dependent enzymes cathepsin K and GSK-3 β

Pomeislová A., Otmar M., Rubešová P., **Benýšek J.**, Matoušová M., Mertlíková-Kaiserová H., Pohl R., Poštová Slavětínská L., Pomeisl K., Krečmerová M., (2021) *Bioorganic and Medicinal Chemistry*, 32, 115998. DOI: 10.1016/j.bmc.2021.115998



Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc1,2,4-Thiadiazole acyclic nucleoside phosphonates as inhibitors of cysteine dependent enzymes cathepsin K and GSK-3 β

Alice Pomeislová^{a,b}, Miroslav Otmar^c, Petra Rubešová^b, Jakub Benýšek^b, Marika Matoušová^b, Helena Mertlíková-Kaiserová^b, Radek Pohl^b, Lenka Poštová Slavětínská^b, Karel Pomeisl^d, Marcela Krečmerová^{b,*}

^a Department of Organic Chemistry, Faculty of Science, Charles University, Hlavova 8, 128 43 Prague 2, Czech Republic

^b Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic

^c Institute of Macromolecular Chemistry, Czech Academy of Sciences, Heyrovského nám. 2, 162 06 Prague 6, Czech Republic

^d Institute of Physics, Czech Academy of Sciences, Na Slovance 1999/2, 182 21 Prague 8, Czech Republic

ARTICLE INFO

Keywords:

Thiadiazoles
5-Amino-1,2,4-thiadiazol-3-(2H)-one
Acyclic nucleoside phosphonates
Epigenetic
Cathepsin K, Glycogen synthase kinase 3 β

ABSTRACT

In analogy to antiviral acyclic nucleoside phosphonates, a series of 5-amino-3-oxo-1,2,4-thiadiazol-3(2H)-ones bearing a 2-phosphonomethoxyethyl (PME) or 3-hydroxy-2-(phosphonomethoxy)propyl (HPMP) group at the position 2 of the heterocyclic moiety has been synthesized. Diisopropyl esters of PME- and HPMP-amines have been converted to the *N*-substituted ureas and then reacted with benzoyl, ethoxycarbonyl, and Fmoc thiocyanates to give the corresponding thiobiotins, which were oxidatively cyclized to diisopropyl esters of 5-amino-3-oxo-2-PME- or 2-HPMP- 1,2,4-thiadiazol-3(2H)-ones. The phosphonate ester groups were cleaved with bromotrimethylsilane, yielding *N*²-protected phosphonic acids. The subsequent attempts to remove the protecting group from *N*² under alkaline conditions resulted in the cleavage of the 1,2,4-thiadiazole ring. Similarly, compounds with a previously unprotected 5-amino-1,2,4-thiadiazolone base moiety were stable only in the form of phosphonate esters. The series of twenty-one newly prepared 1,2,4-thiadiazol-3(2H)-ones were explored as potential inhibitors of cysteine-dependent enzymes – human cathepsin K (CatK) and glycogen synthase kinase 3 β (GSK-3 β). Several compounds exhibited an inhibitory activity toward both enzymes in the low micromolar range. The inhibitory potency of some of them toward GSK-3 β was similar to that of the thiadiazole GSK-3 β inhibitor tideglusib, whereas others exhibited more favorable toxicity profile while retaining good inhibitory activity.

1. Introduction

It has been reported that 1,2,4-thiadiazole compounds are capable of inhibiting cysteine proteases, which contain a catalytic cysteine residue in the active site, as demonstrated for a selective inhibitor of cathepsin B acting in the micromolar range.¹ Among cathepsin-type cysteine proteases, human cathepsin K (CatK) has been one of the most investigated medically important members in the last decade. This lysosomal enzyme is predominantly expressed in osteoclasts and plays a critical role in the bone resorption process. It degrades the protein components of the bone matrix, including its major constituent, type I collagen. The hyperactivity of osteoclasts and an imbalance between bone formation and resorption lead to the loss bone tissue, which is associated with osteoporosis.² CatK has been evaluated as the main drug target for

osteoporosis treatment, and a number of CatK inhibitors are being developed or tested in clinical trials (for reviews, see e.g.²⁻⁴). Most of these inhibitors contain an electrophilic warhead, such as ketone or nitrile groups, for covalent binding to the catalytic cysteine of CatK.⁵⁻⁸ This work describes for the first time CatK inhibitors based on 1,2,4-thiadiazole employed as a thiol-reactive warhead targeting the catalytic site. The interaction of 1,2,4-thiadiazole derivatives with the thiol group of the cysteine residue of the enzyme target is outlined in Fig. 1.

The nucleophilic attack of the cysteine thiol results in the formation of a covalent adduct and the loss of enzymatic activity. Human cathepsin K (CatK) is proposed to be inhibited by the modification of catalytic Cys25 in the enzyme active site, while glycogen synthase kinase 3 β (GSK3 β) is modified on Cys199, located in the ATP cofactor binding site.

A great effort has been made in the investigation of 1,2,4-

* Corresponding author.

E-mail address: marcela.krecmerova@uochb.cas.cz (M. Krečmerová).

<https://doi.org/10.1016/j.bmc.2021.115998>

Received 7 December 2020; Accepted 31 December 2020

Available online 6 January 2021

0968-0896/© 2021 Elsevier Ltd. All rights reserved.

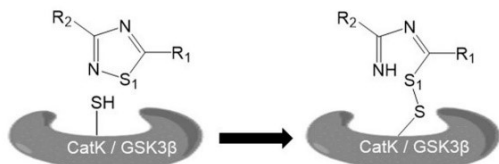


Fig. 1. Scheme of the interaction of 1,2,4-thiadiazole-based inhibitors with the thiol group of the functionally important cysteine residue of the enzyme target.

thiadiazolidine-3,5-diones (TDZD) as inhibitors of glycogen synthase kinase-3 β (GSK-3), a key enzyme of glycogen metabolism.^{9,10} In cancer cells, aberrant nuclear accumulation of GSK-3 positively regulates the binding of the transcription factor NF- κ B to its target gene promoters and its further transcriptional activity. In this way, the NF- κ B activation promotes the progression of human cancer and also causes resistance to chemotherapeutic drugs.^{9,10} So far, the most promising inhibitor of GSK-3 is 4-benzyl-2-(1-naphthyl)-1,2,4-thiadiazolidine-3,5-dione (tideglusib). Regarding its ability to influence gene expression, its function can be considered epigenetic. GSK-3 also plays a crucial role in Alzheimer's disease and other neurodegenerative processes. Clinical studies of tideglusib for the treatment of Alzheimer's disease^{11–13} and myotonic muscular dystrophy¹⁴ are currently being conducted in Phase II. In dental medicine, tideglusib is a topic of extensive research aimed at natural tooth repair by GSK3 antagonists. The ability of the promotion of dentine reinforcement has been successfully proven *in vivo*.¹⁵

A completely different mechanism of action has been reported in a series of neuroprotective 1,2,4-thiadiazole derivatives influencing ionotropic transmembrane receptors for glutamate (NMDA and AMPA receptors). These receptors are important for controlling synaptic plasticity and memory function and mediate fast synaptic transmission in the central nervous system.^{16–18} A general structural motif of these derivatives is outlined in Fig. 2.

Besides, there are many 1,2,4-thiadiazol structures whose biological properties still deserve further exploration. Our team of nucleoside chemistry has focused on 5-amino-1,2,4-thiadiazol-3(2H)-one, which could be understood as a sulfur-containing mimic of cytosine in which the $-\text{CH}=\text{CH}-$ group between the positions 5 and 6 of the pyrimidine ring is replaced by a divalent sulfur ($-\text{S}-$) as reported in several papers.^{19,20} 5-Amino-1,2,4-thiadiazol-3(2H)-one is considered a bioisostere of cytosine because of the similar electronic and spatial properties of the $-\text{CH}=\text{CH}-$ group in cytosine and the divalent sulfur.^{19,20} Due to the proposed bioisosterism, there has been an effort to

use 5-amino-1,2,4-thiadiazol-3(2H)-one as a cytosine surrogate in the synthesis of nucleosides and their acyclic analogs.^{19,20} 5-Amino-1,2,4-thiadiazol-3(2H)-one itself has exhibited slight antibacterial activities. The fact that the antibacterial activity of 5-amino-1,2,4-thiadiazol-3(2H)-one is not reversed by the presence of natural pyrimidines implies that in this case, it does not behave as a pyrimidine antimetabolite but rather interferes with bacterial proteins and thus inhibits the growth of bacteria.²¹ In order to identify the active form for biological interactions, a systematic study of the tautomerism in 5-amino-1,2,4-thiadiazol-3(2H)-one has been performed. In aqueous solutions, a significant stabilization of the amino tautomer has been found, with the equilibrium shifted to favor the amino form.²²

The present study focuses on the synthesis of compounds in which the 5-amino-1,2,4-thiadiazol-3(2H)-one moiety (as the cytosine mimic) is attached to a 2-(phosphonomethoxy)ethyl (PME) or 3-hydroxy-2-(phosphonomethoxy)propyl (HPMP) group analogously to acyclic nucleoside phosphonates (ANPs). ANPs are known for their large spectrum of biological activities, especially antiviral. Three of them have already become successful commercial drugs: cidofovir, approved for the treatment of human cytomegalovirus retinitis in AIDS patients but used off-label also against many other DNA viruses, adefovir for the treatment of hepatitis B (HBV), and tenofovir for the treatment of HIV and HBV infections.^{23–25} The possibility that 5-amino-1,2,4-thiadiazol-3(2H)-one could really work as a cytosine mimic together with the ability of the reactive N–S bond to attack strong nucleophiles such as the cysteine thiol group of enzymes have provided the impetus to synthesize 5-amino-1,2,4-thiadiazol-3(2H)-one ANPs and to subject them to a detailed biological activity screening.

2. Chemistry

This work is mainly focused on the synthesis of 5-amino-1,2,4-thiadiazol-3(2H)-one derivatives with the HPMP arrangement, compounds structurally analogous to the antiviral agents cidofovir (HPMP-C)^{23–25} and its 5-azacytosine analog (HPMP-azaC), both depicted in Fig. 3.^{26,27} The HPMP series was also extended to the PME derivatives with respect to the antiviral activity of PMEC and PME-azaC (see also Fig. 3), which, even if it is lower, can be increased *via* transformation to appropriate prodrugs.²⁸

5-Amino-1,2,4-thiadiazol-3(2H)-one (**1**) was prepared by oxidative cyclization of thiobiuret under basic conditions¹⁹ (Scheme 1). The direct alkylation of **1** with the synthon $i\text{Pr}_2\text{PME-Cl}$ (**2**) or (*S*)-glycidyl trityl ether (**4**) led to a mixture of compounds, whose TLC analysis profile did not correspond to the desired products. Therefore, the 5-amino group of

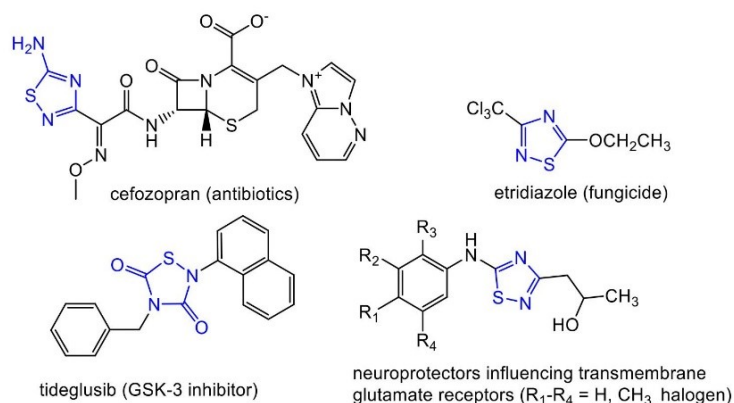


Fig. 2. Examples of 1,2,4-thiadiazoles with diverse types of biological activities.

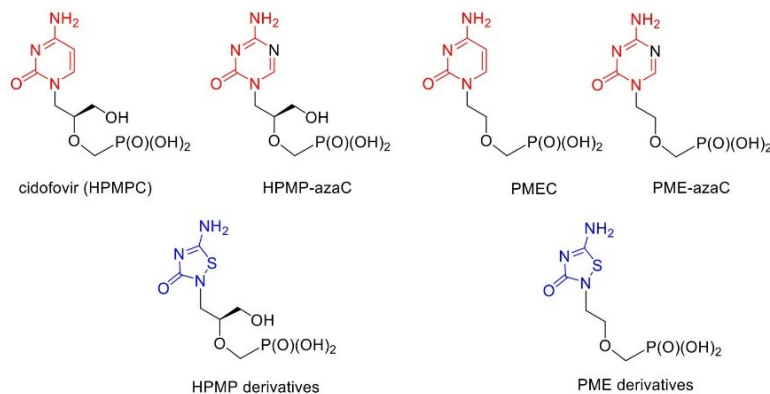
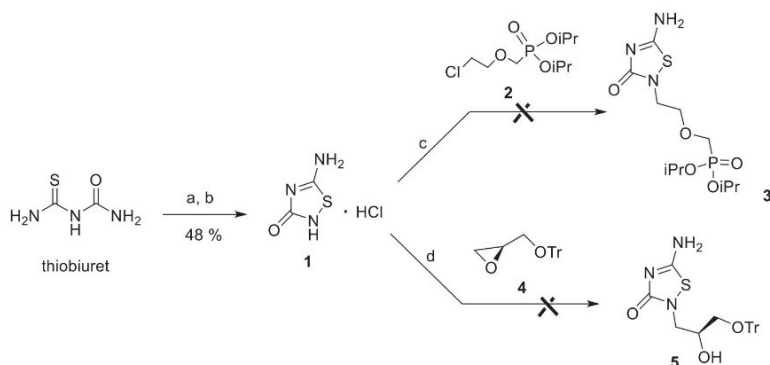


Fig. 3. Acyclic nucleoside phosphonates with cytosine and 5-azacytosine base moieties and their 5-amino-1,2,4-thiadiazol-3(2H)-one analogues.



Scheme 1. The synthesis of 5-amino-1,2,4-thiadiazol-3(2H)-one (**1**). The reagents and conditions: (a) 2 M NaOH, H₂O₂, MeOH, 0 °C, 2 h; (b) HCl; (c) Cs₂CO₃, DMF, reflux, 4.5 h; (d) Cs₂CO₃, DMF, reflux, 24 h.

1 was protected by dimethylaminomethylene (DMA) to form **6**, which was subsequently alkylated with the synthons (**Scheme 2**). However, in the case of the *i*Pr₂-PME-Cl (**2**), only the product of the *O*³-alkylation of **7** was formed, and with the synthon **4**, a 4:1 mixture of the *O*³-**10a** and the *N*2-alkylated **10b** product was obtained. The deprotection of **7** by hydrazine and the subsequent removal of phosphonate ester groups from the resulting compound **8** by bromotrimethylsilane afforded the free phosphonate **9**. Like in the case of the *i*Pr₂-PME-Cl (**2**), the alkylation of the isomers **10a** and **10b** with the tosylate **11** did not provide the desired HPMP derivatives **12a** and **12b**.

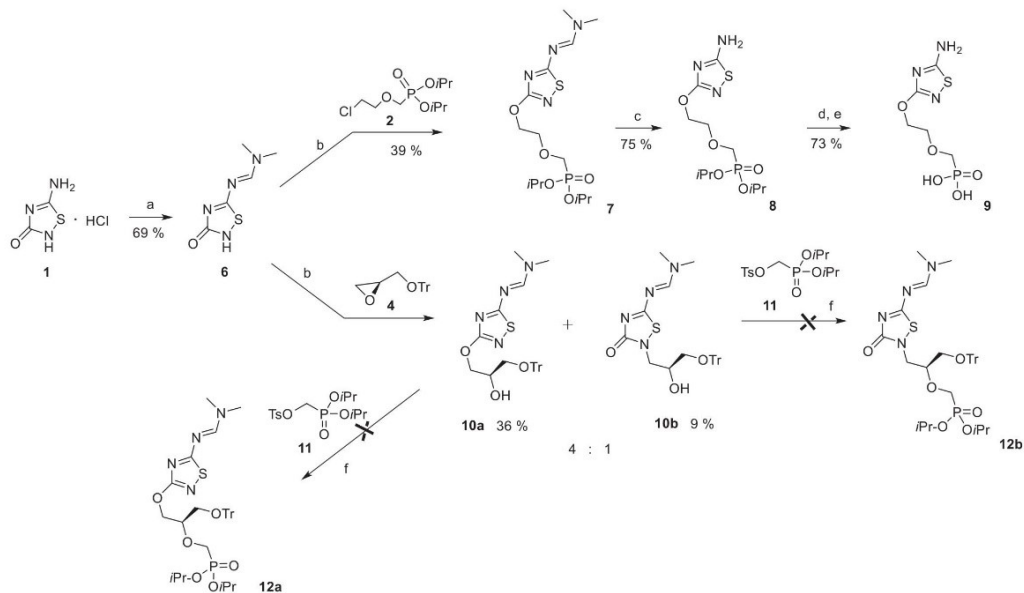
In a further attempt to achieve *N*2 alkylation of **1** with the synthons **2** and **4**, the DMA protecting group was replaced by acyl (**Scheme 3**). The *N*⁵-acetyl- (**14a**) and *N*⁵-benzoyl- (**14b**) derivatives were thus prepared by the reaction of acetyl or benzoyl isothiocyanate with urea, yielding the formation of the corresponding 1-acetyl- (**13a**) and 1-benzoyl-2-thiobiuret (**13b**), followed by oxidative cyclization with bromine or NBS. The *N*⁵-benzoyl derivative **14b** was alkylated with **4** to yield only the undesired *N*4-substituted product **15**. Therefore, this synthetic strategy based on the *N*⁵-acyl protection was abandoned.

In order to obtain the desired *N*²-substituted derivatives of **1**, a stepwise construction of the 5-amino-1,2,4-thiadiazol-3(2H)-one ring on PME and HPMP scaffolds was performed. The starting synthetic precursors were *i*Pr₂-PME-NH₂ (**17**) and *i*Pr₂-Tr-HPMP-NH₂ (**23**). Compound **17** was synthesized according to **Scheme 4** by two different

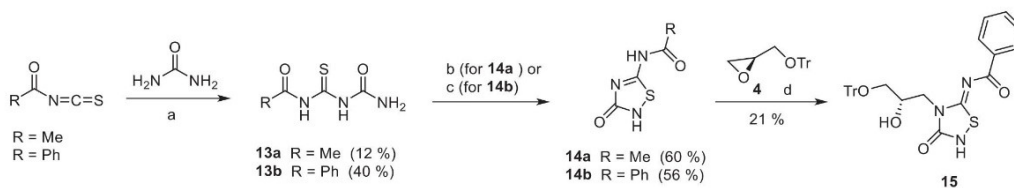
methods. In the first one, *i*Pr₂-PME-Cl (**2**) was converted to **16** by reaction with potassium phthalimide, and the subsequent removal of the phthalic moiety by hydrazine released *i*Pr₂-PME-NH₂ (**17**). In the second method, 2-aminoethanol was *N*-tritylated to **18**, then *O*-alkylated with the tosylated synthon **11** in the presence of magnesium *tert*-butoxide in *N*-methyl-2-pyrrolidone²⁹ to form **19**, and after that deprotected to obtain *i*Pr₂-PME-NH₂ (**17**).

The synthon *i*Pr₂-Tr-HPMP-NH₂ (**23**) was prepared in three steps from (*S*)-glycidyl trityl ether (**4**) (**Scheme 5**). In the first step, the oxirane ring of **4** was opened with sodium azide at 45 °C to form **20**. To minimize the formation of undesired side products, it was necessary to increase the temperature only slightly and to use a 40-fold excess of sodium azide in the reaction. Compound **20** was alkylated with the bromomethylphosphonate **21** to obtain the corresponding ether **22**. The azide function in **22** was then reduced to an amino group by the Staudinger reaction with triphenylphosphine, yielding *i*Pr₂-Tr-HPMP-NH₂ (**23**).

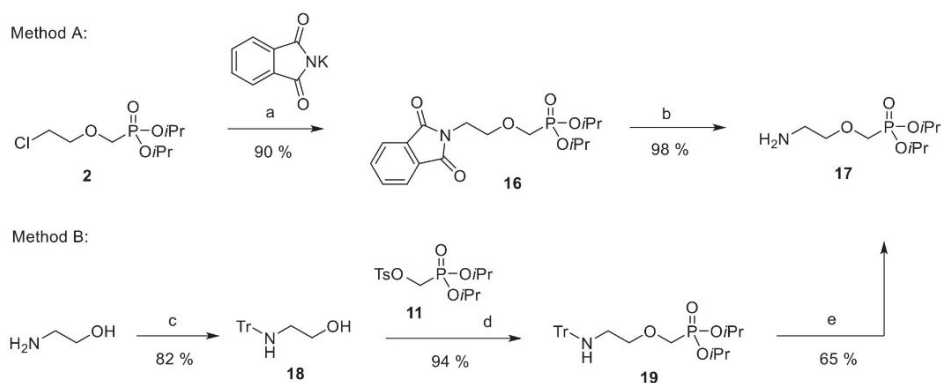
The starting synthons, *i*Pr₂-PME-NH₂ (**17**) and *i*Pr₂-Tr-HPMP-NH₂ (**23**), on which the stepwise construction of the 5-amino-1,2,4-thiadiazol-3(2H)-one moiety, mimicking the cytosine nucleobase, was performed, were converted to the corresponding *N*-substituted ureas **24** and **25** (**Scheme 6**). The urea **24**, with the *i*Pr₂-PME substituent, was prepared by the reaction of the amine **17** with potassium cyanate under acidic conditions. In the case of the amine **23**, bearing the *i*Pr₂-Tr-HPMP substituent, the acidic treatment could not be used, because the trityl



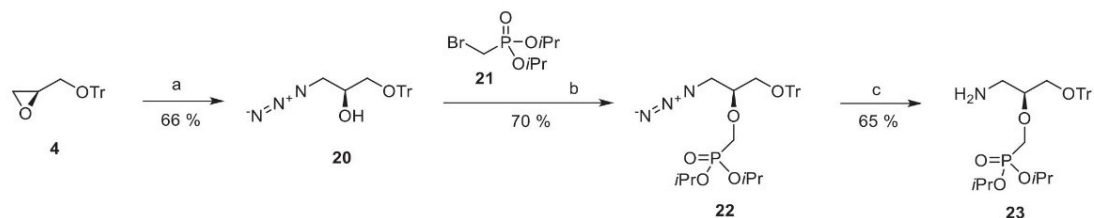
Scheme 2. Alkylation of the DMA protected 5-amino-1,2,4-thiadiazol-3(2H)-ones with PME a HPMP synthons. The reagents and conditions: (a) DMF/DMA, acetonitrile, 24 h; (b) Cs_2CO_3 , DMF, 95 °C, 4.5 h; (c) NH_2NH_2 , MeOH, 1 h; (d) TMSBr, CH_3CN , 24 h; (e) H_2O ; (f) NaH, THF, 24 h.



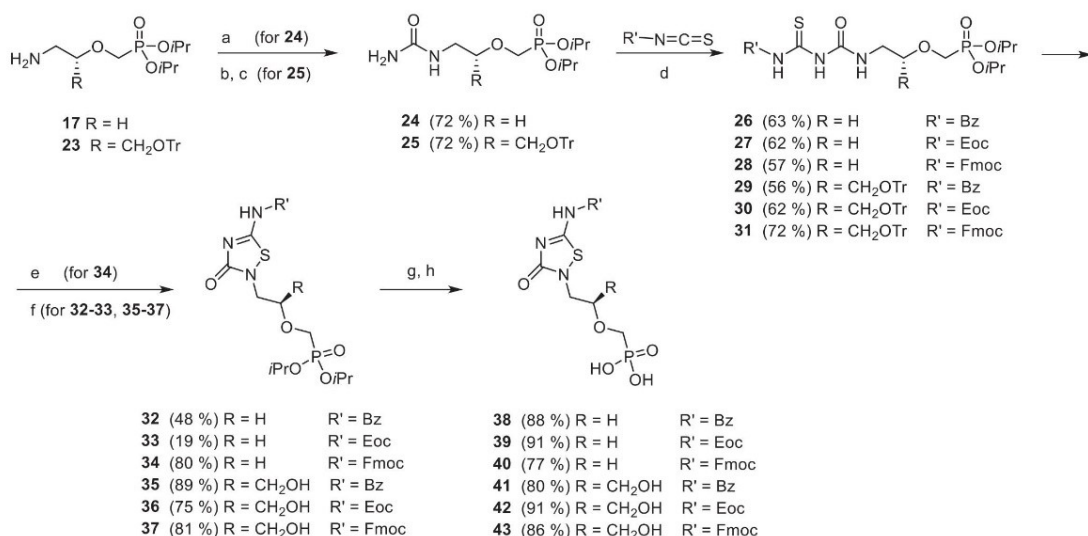
Scheme 3. The synthesis of N^5 -acyl-5-amino-1,2,4-thiadiazol-3(2H)-ones and their N^4 -alkylation. The reagents and conditions: (a) acetone, 60 °C, 5 h; (b) $\text{Br}_2/\text{CHCl}_3$ (1 M), EtOH, 35 °C, 20 min; (c) NBS, MeOH, reflux, 1 h; (d) Cs_2CO_3 , DMF, reflux, 24 h.



Scheme 4. The synthesis of $i\text{Pr}_2\text{-PME-NH}_2$ (17). The reagents and conditions: (a) DMF, 120 °C, 3.5 h; (b) 1 M NH_2NH_2 /THF, MeOH, 2 h; (c) TrCl, Et_3N , CH_2Cl_2 , 2 h; (d) $\text{Mg}(\text{O}^t\text{Bu})_2$, NMP, 75 °C, 2 h; (e) 35% HCl (aq.), dioxane, 1 h.



Scheme 5. The synthesis of iPr_2 -Tr-HPMP-NH₂ (**23**). The reagents and conditions: (a) NaN₃, DMF, 45 °C, 1 week; (b) NaH, DMF, 60 °C, 24 h; (c) Ph₃P, H₂O, THF, 0 °C to RT, 24 h.

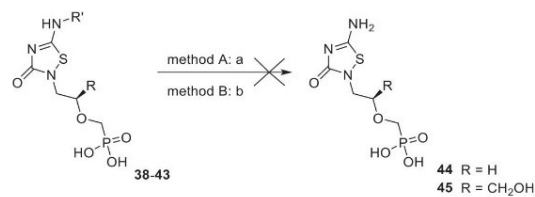


Scheme 6. The synthesis of N^5 -protected 5-amino-1,2,4-thiadiazol-3(2H)-ones substituted on N2 with the PME and HPMP moiety. The reagents and conditions: (a) KOcN, 35% HCl, 70 °C, 24 h; (b) CDI, CH₂Cl₂, 3 h; (c) NH₃/MeOH, 24 h; (d) acetone, 60 °C, 24 h; (e) 0.5 M Br₂/EtOAc, 4 °C, 24 h; (f) NBS, MeOH, 80 °C, 30 min; (g) TMSBr, CH₃CN, 24 h; (h) H₂O.

protecting group would be removed. Instead, the amine **23** was treated with carbonyldiimidazole (CDI) and then with methanolic ammonia to form the urea **25** with an intact trityl group. In the next step, the ureas **24** and **25** were reacted with benzoyl, ethoxycarbonyl (Eoc), and Fmoc isothiocyanate with the formation of the corresponding thiobiurets **26–31**.

The oxidative cyclization of compounds **26–28** with bromine or NBS led to the formation of the N^5 -protected 5-amino-1,2,4-thiadiazol-3(2H)-ones **32–34** alkylated on the N2-position with the iPr_2 -PME synthon. In the compounds **29–31** the oxidative cyclization also caused the removal of the trityl protecting group, yielding the iPr -HPMP derivatives **35–37**. In the next step, the phosphonate ester groups of the compounds **32–37** were cleaved with trimethylsilyl bromide to give the N^5 -protected phosphonic acids **38–43**.

The attempts to remove the protecting groups with sodium hydroxide or hydrochloric acid at elevated temperature were unsuccessful (Scheme 7). The use of pyrrolidine and ethanolamine for the deprotection of the Fmoc derivatives **40** and **43** according to the literature³⁰ only led to the open-ring phosphonates **46** and **47** (Scheme 8). Similarly, the removal of Fmoc from the PME derivative **40** with hydrochloric acid in harsh conditions under microwave irradiation resulted in the formation of the open-ring phosphonate **46**. The same product was obtained by the deprotection of **40** with sodium azide.³¹ The attempt for the oxidative recyclization of the compound **46** to the 1,2,4-thiadiazol-3

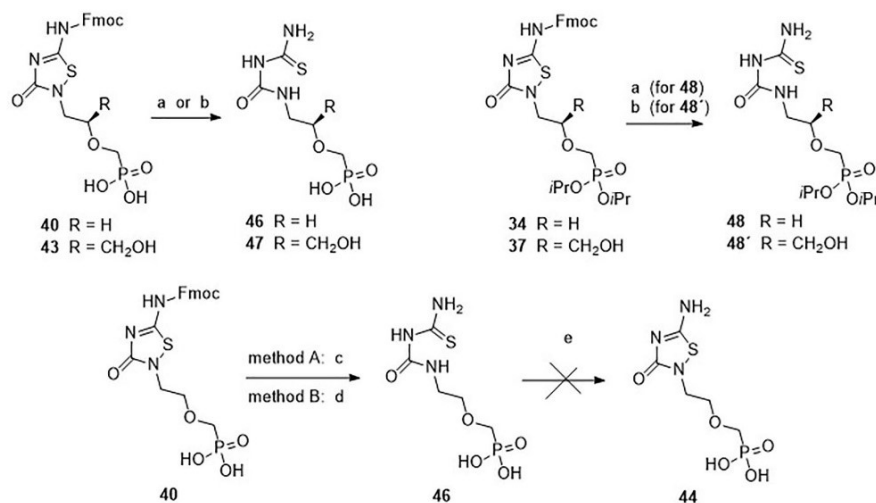


Scheme 7. Attempts for the removal of protecting groups from the amino group in position 5. The reagents and conditions: (a) 35% aq. HCl/dioxane 1:1, reflux, 8 h; (b) 2 M KOH, reflux, 8 h.

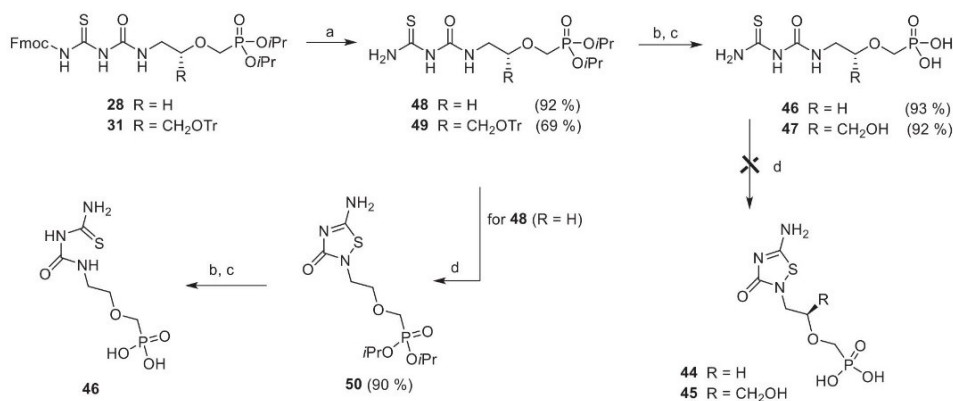
(2H)-one **44** with bromine was unsuccessful.

The remaining synthetic path to obtain the desired N2-substituted derivatives of **1** consists in the oxidative cyclization of the deprotected 2-thiobiurets bearing at the N5 position the PME or HPMP group with or without ester groups (**46–49**, Scheme 9).

The compounds **28** and **31** were thus deprotected with piperidine, yielding the 5-substituted biurets **48** and **49**. The isopropyl ester and trityl groups were then cleaved with trimethylsilyl bromide to give the free phosphonic acids **46** and **47**. The attempts to cyclize the compounds **46** and **47** oxidatively with bromine to the 1,2,4-thiadiazol-3(2H)-ones **44** and **45** were unsuccessful, only recovering the starting material.



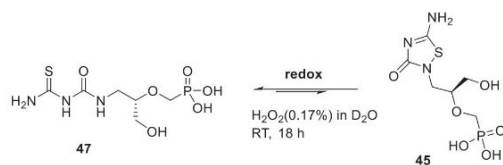
Scheme 8. The formation of open-ring derivatives during the deprotection of the 5-amino group. The reagents and conditions: (a) pyrrolidine, CH_2Cl_2 , 24 h; (b) piperidine, CH_2Cl_2 , 24 h; (c) 0.5 M HCl, M-W, 150 °C, 16 min; (d) NaN_3 , DMF, 50 °C, 3.5 h; (e) Br_2/EtOAc .



Scheme 9. Cyclization and ring-opening of 5-amino-1,2,4-thiadiazol-3(2H)-ones caused by a redox character of S–N bond. The reagents and conditions: (a) piperidine, CH_2Cl_2 , 24 h; (b) TMSBr, CH_3CN , 24 h; (c) H_2O , 1 h; (d) 0.5 M Br_2/EtOAc , H_2O , 4 °C.

Interestingly, the phosphonate diester **48** was successfully cyclized with bromine to the 1,2,4-thiadiazol-3(2H)-one **50**. However, the subsequent ester cleavage with trimethylsilyl bromide gave only the open-ring product **46** instead, which, as mentioned above, cannot be recycled.

The redox character of the S–N bond was studied in the reaction of



Scheme 10. The cyclization of **47** by hydrogen peroxide; the experiment was performed in a NMR tube in D_2O .

47 with 0.17% hydrogen peroxide in D_2O . The process was monitored at one-hour intervals by ^{31}P NMR (Scheme 10). Hydrogen peroxide was chosen for its easy decomposition to water not disturbing the measurement. The reaction exhibited equilibrium between the linear compound **47** and the cyclization product **45**. The ratio of the peaks in ^{31}P NMR spectra for **47** and **45** was 1:1. The addition of another equivalent of hydrogen peroxide did not lead to a change in the equilibrium.

This equilibrium illustrates the above-mentioned tendency to ring opening of 5-amino-1,2,4-thiadiazol-3(2H)-ones bearing at the N^2 position a PME or HPMP substituent with a free phosphonate function and the impossibility to recycle the resulted open-ring derivatives successfully. If the amino group at position 5 of the 1,2,4-thiadiazol-3(2H)-ones is protected and thus the tautomerism is prevented, the stability of the ring is much higher. In order to explain the simultaneous presence of the unprotected 1,2,4-thiadiazol-3(2H)-one ring and the free phosphonate function is not a stable arrangement, it is necessary to look

for an analogy with the formation of cyclic HPMP (51) under the cleavage of HPMP diisopropyl ester (52) (Scheme 11).³² The spontaneous origination of cyclic HPMP (51) is conditioned by the possibility of the formation of the six-membered ring by the nucleophilic attack of the primary hydroxy group on the intermediate reactive trimethylsilyl phosphonate ester. A similar situation occurred in the compound 50, where even the more nucleophilic N2 nitrogen atom is in an analogous position (enabling the formation of a six-membered ring) to that of the primary hydroxy group of HPMP. Contrary to the cyclic HPMP (51), the cyclic phosphonoamide 53 is more easily cleavable to the open-ring phosphonate 46 due to the inherent reactivity of the P–N bond than the phosphonate cyclic ester 51.

To summarize the results of the above-described syntheses, only the N⁵-substituted 1,2,4-thiadiazol-3(2H)-one analogs of ANPs were sufficiently stable and isolable in the form of free phosphonic acids. Compounds with the unprotected amino group of the 5-amino-1,2,4-thiadiazol-3(2H)-one base moiety were stable only in the form of phosphonate esters. These findings correspond to our previous observations, when the esterification of the phosphonic acid residue increased the stability of a labile base moiety, e.g. the 5-azacytosine ring in 5-azacytosine ANPs.^{33,34}

Most of the newly prepared 5-amino-1,2,4-thiadiazol-3(2H)-ones have exhibited interesting biological activities concerning inhibitory effects to cysteine-dependent enzymes.

3. Results and Discussion

3.1. The inhibition of cathepsin K by 5-amino-1,2,4-thiadiazol-3(2H)-ones

Twenty-one derivatives of 5-amino-1,2,4-thiadiazol-3(2H)-one derivatives have been explored as potential inhibitors of human CatK. The compounds were screened *in vitro*, at the concentration of 10 μM, against the recombinant enzyme using a kinetic inhibition assay with the fluorogenic substrate Z-Gly-Pro-Arg-AMC. They exhibited a wide range of inhibitory activities, with the six most potent inhibitors showing higher inhibition than 50% (residual enzyme activity below 50%, Table 1). The determined IC₅₀ values of the effective compounds ranged from 2.6 to 14.9 μM. According to structural features, the active compounds can be divided into three categories: (i) the parental unsubstituted 5-amino-

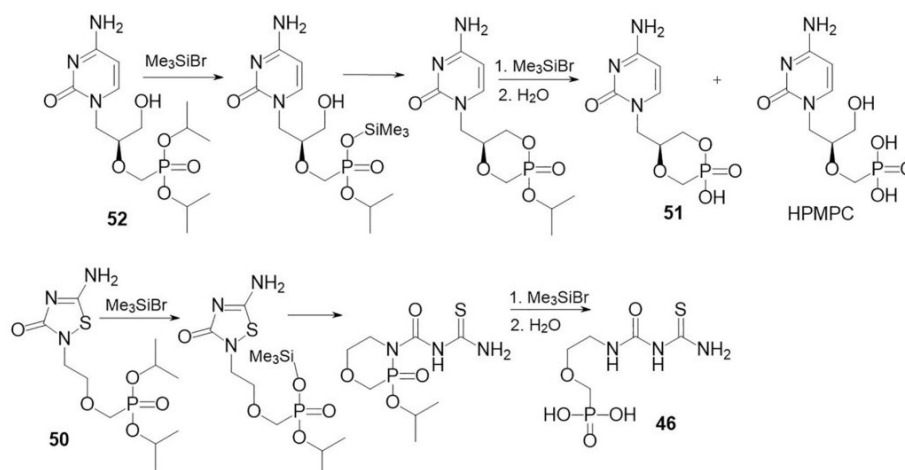
Table 1

The inhibition of human cathepsin K and GSK-3β by 5-amino-1,2,4-thiadiazol-3(2H)-ones applied at the concentration of 10 μM. Residual enzyme activities were measured as described in the Experimental section and expressed relative to the uninhibited control (100%). The compounds are ordered according to decreasing inhibitory potency toward CatK.

Compound name	CatK residual activity (%)	GSK3 residual activity (%)
10a	5.3 ± 1.5	72 ± 16.7
1	16.9 ± 1.0	78 ± 7.3
10b	28.3 ± 1.7	64 ± 14.1
37	34.1 ± 2.4	76 ± 2.4
15	34.2 ± 3.0	59 ± 13.6
39	39.1 ± 1.0	87 ± 4.1
7	52.9 ± 3.8	102 ± 8.8
6	67.0 ± 5.8	97 ± 12.3
40	67.0 ± 0.6	85 ± 4.8
14a	76.6 ± 3.8	82 ± 9.7
8	78.3 ± 2.6	113 ± 15.5
32	80.6 ± 5.0	87 ± 25.0
36	83.4 ± 8.6	110 ± 9.1
14b	85.0 ± 3.5	85 ± 18.5
43	85.5 ± 5.3	98 ± 12.5
33	86.8 ± 2.7	154 ± 0.5
9	90.0 ± 3.5	106 ± 6.2
38	90.3 ± 6.0	75 ± 7.4
41	92.3 ± 4.5	106 ± 12.3
35	92.9 ± 2.8	81 ± 3.5
42	97.5 ± 2.8	101 ± 10.0
Tideglusib	N/A*	70 ± 7.2
CHIR99021	N/A	10 ± 5.2

* not applicable for CatK

1,2,4-thiadiazol-3(2H)-one (1) (IC₅₀ of 6.34 ± 0.63), (ii) derivatives protected on N⁵ bearing 2-hydroxy-3-trityloxypropyl group: 10a (2.63 ± 0.31 μM), 10b (10.08 ± 0.95 μM), and 15 (14.9 ± 0.75 μM), and (iii) derivatives protected on N⁵ bearing a phosphonate group on HPMP or PME aliphatic chain, with: 37 (9.25 ± 1.46) or without: 39 (10.15 ± 1.48) esterification. The whole series tested revealed several general trends in SAR. The esterification of the phosphonate function resulted in more potent inhibitors of CatK in comparison with free phosphonic acids (8, 32, and 37 versus 9, 38, and 43, respectively). The presence of the polar hydroxymethyl group (HPMP) decreased inhibition in contrast to its absence (PME) (35, 41, 43, and 42 versus 32, 38, 40, and 39,



Scheme 11. An explanation of the tendency to ring opening of 5-amino-1,2,4-thiadiazol-3(2H)-ones bearing at the N2 position a PME or HPMP substituent with the free phosphonate function.

respectively). At the amino group in position 5, the Fmoc protecting group effects a higher inhibition than that of benzoyl (**40**, **37**, and **43** versus **38**, **35**, and **41**, respectively). In conclusion, the screened series of 1,2,4-thiadiazoles provided several inhibitors of CatK effective in the low micromolar range.

3.2. The inhibition of GSK-3 β and the cytotoxicity of 5-amino-1,2,4-thiadiazol-3(2H)-ones

The prepared 5-amino-1,2,4-thiadiazol-3(2H)-ones have also been screened for their ability to inhibit GSK-3 β (human, recombinant). Interestingly, the inhibition by 10- μ M 1,2,4-thiadiazole derivatives (including the reference inhibitor tideglusib) was relatively weak in comparison with previously published data,³⁵ with the IC₅₀ values > 10 μ M, which was not the case for another non-thiadiazole inhibitor, CHIR99021 (IC₅₀ = 9.9 \pm 0.65 nM). Since we have employed a validated commercial GSK-3 β kit (Promega), there might be some unknown interferences of assay components with the 1,2,4-thiadiazole moiety, resulting in the apparently low inhibition effects. Therefore, we have set tideglusib as a benchmark compound when evaluating the activity of the compounds toward GSK-3 β . Most compounds that have exhibited activity similar to tideglusib (i.e. the inhibition of approx. 30%) under the conditions of the assay fall into the cluster of compounds that have also been identified as CatK inhibitors (**10a**, **1**, **10b**, **37**, **15**) (Table 1) and therefore probably follow a similar SAR pattern to that described above.

The cytotoxicity of the compounds intended for use in the treatment of human diseases is, of course, of concern. This paper shows that the hereby presented thiadiazole derivatives are generally less toxic than tideglusib, which exhibited cytotoxicity in two leukemic cell lines (CCRF-CEM, HL-60) but, more importantly, also to normal human dermal fibroblasts (NHDF) at 10 μ M (Table 2). With the exception of compounds **10a**, **10b** and **15**, which all include the triphenylmethyl moiety, and unsubstituted 5-amino-1,2,4-thiadiazol-3(2H)-ones (**1**), the compounds have proven to be safe at concentrations up to 100 μ M. The presence of the trityl group increased the inhibitory activity of the compound toward Cys-containing enzymes CatK and GSK-3 β , but at the expense of undesired toxicity.

Table 2

The cytotoxicity of 5-amino-1,2,4-thiadiazol-3(2H)-ones to tumor and non-tumor cell lines following the 72-h incubation of the cells with the compounds (10 and 100 μ M). The data are expressed as the percentage of viable cells vs. untreated control.

Compound name	Cell viability [%]											
	CCRF-CEM		HL-60		HeLa S3		HepG2		SH-SY5Y		NHDF	
	10 μ M	100 μ M	10 μ M	100 μ M	10 μ M	100 μ M	10 μ M	100 μ M	10 μ M	100 μ M	10 μ M	100 μ M
1	120	4	98	95	100	87	87	54	99	56	101	49
6	122	84	102	105	97	112	97	85	100	75	105	107
7	106	93	97	98	101	89	99	90	101	87	101	99
8	114	103	103	108	97	98	100	99	98	55	104	114
9	113	112	103	100	101	102	97	93	101	99	105	96
10a	101	2	86	1	91	2	92	1	99	2	97	1
10b	15	1	64	2	80	1	87	1	67	3	77	1
14a	104	102	105	108	97	102	103	88	102	99	99	105
14b	102	96	101	110	99	101	99	77	103	90	104	100
15	81	5	93	9	45	3	91	37	96	34	98	79
32	100	94	104	115	101	100	103	101	104	101	103	101
33	92	98	102	111	100	101	104	100	105	104	101	98
35	86	97	106	118	97	105	109	99	100	97	103	104
36	67	45	97	101	90	87	101	93	104	104	104	100
37	82	82	103	110	99	100	102	92	102	85	98	100
38	98	100	102	113	101	101	100	100	104	105	103	102
39	74	77	96	100	94	96	98	97	103	105	94	97
40	85	90	102	106	97	102	105	103	102	99	93	84
41	72	61	103	106	97	101	101	100	101	103	105	98
42	94	114	95	99	97	94	99	89	100	99	101	99
43	88	119	93	100	98	99	106	88	99	93	98	95
Tideglusib	30	n.a.*	34	n.a.	69	n.a.	92	n.a.	80	n.a.	57	n.a.
CHIR99021	28	n.a.*	119	n.a.	67	n.a.	86	n.a.	77	n.a.	53	n.a.

* n.a. – not available.

4. Conclusions

In this work, we describe for the first time CatK and GSK-3 β inhibitors based on the 1,2,4-thiadiazole scaffold, employed as a thiol-reactive warhead targeting the thiol group located in the catalytic (cofactor binding) site of these cysteine-dependent enzymes. We have screened a set of 5-amino-1,2,4-thiadiazol-3(2H)-one compounds; among them, we have identified several inhibitors of CatK effective in the low micromolar range. The inhibitors are comparable in efficacy to previously published cathepsin B inhibitors.¹ These compounds have also been shown to inhibit GSK-3 β with similar potency as the established 1,2,4-thiadiazole GSK-3 β inhibitor tideglusib, with some of them (e.g. **37**) exhibiting even a more favorable toxicity profile while retaining good inhibitory activity.

The future research of 1,2,4-thiadiazole inhibitors should take more into account the interior of the enzyme recognition site and design the new molecules accordingly, similarly as it was published in the case of peptidomimetic inhibitors of CatK with ketone or nitrile warheads.⁵⁻⁸

5. Experimental

5.1. General

The melting points were determined on a Kofler melting point apparatus with a microscope KBT 300 (Santiago) and are uncorrected. The analytical TLC was performed on silica gel pre-coated aluminum plates with a fluorescent indicator (Merck 5554, 60 F₂₅₄). The spots were visualized with UV light (254 nm) or by spraying with ninhydrine (1% solution in ethanol) followed by short heating to 300–400 °C. The column chromatography was performed on silica gel (Sigma S-0507, 40–63 μ m). Mass spectra were measured on a Q-ToF micro (Waters) and the HR MS were taken on a LTQ Orbitrap XL (Thermo Fisher Scientific) spectrometer. ¹H and ¹³C NMR spectra were recorded on a Bruker Avance 400 (¹H at 400.0 MHz, ¹³C at 100.6 MHz and ³¹P at 162.0 MHz), Bruker Avance 500 (¹H at 500.0 MHz, ¹³C at 125.7 MHz and ³¹P at 202.4 MHz), and Bruker Avance 600 (¹H at 600.1 MHz and ¹³C at 150.9 MHz) in CDCl₃ (referenced to TMS as an internal standard), in DMSO-d₆.

(referenced to the solvent signal $\delta(^1\text{H}) = 2.50$ ppm and $\delta(^{13}\text{C}) = 39.70$ ppm), in CD_3OD (referenced to the solvent signal $\delta(^1\text{H}) = 3.31$ ppm and $\delta(^{13}\text{C}) = 49.0$ ppm), or in D_2O (referenced to 1,4-dioxane as an internal standard, $\delta(^1\text{H}) = 3.75$ ppm and $\delta(^{13}\text{C}) = 69.3$ ppm). The complete assignment was based on H,H-COSY, H,C-HSQC and H,C-HMBC correlation experiments. The chemical shifts (δ) are given in ppm and the coupling constants (J) in Hz. The IR spectra were obtained on an FT IR NICOLET 6700 spectrometer in CCl_4 or KBr pellets. Optical rotations were measured on an Autopol IV polarimeter (Rudolph Research Analytical, U.S.A.) at 20 °C, $[\alpha]_D$ values are given in 10^{-1} deg $\text{cm}^2 \text{g}^{-1}$. The concentrations for specific rotation measurements are in the units of g/100 mL.

The structure numbering for NMR assignment and the NMR spectra of all compounds are available in the Supporting Information.

5.2. The synthesis of compounds

5.2.1. 5-Amino-1,2,4-thiadiazol-3(2H)-one hydrochloride (1)

30% Hydrogen peroxide (12 mL) was added dropwise to a 0 °C cold solution of thiobiotret (10.0 g; 83.9 mmol) in 2 M NaOH (12 mL) and the mixture was stirred at 0 °C for 3 h. Upon successful conversion, the mixture was acidified to pH 4.5 by the addition of 35% HCl. The resulting precipitate was filtered off and recrystallized from boiling water. Yield: 6.19 g (48%) of white crystals, mp 219–220 °C. The spectral data are in accordance with Ref.¹⁹

5.2.2. (E)-N,N-Dimethyl-N'-(3-oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)formimidamide (6)

A solution of **1** (5.450 g; 35.5 mmol) and *N,N*-dimethylformamide dimethyl acetal (12.4 mL; 93.3 mmol) in acetonitrile (50 mL) was stirred at RT overnight. An excess of diethyl ether was added, the resulting precipitate was filtered off and washed with diethyl ether. Yield: 4.7 g (77%) of white crystals, mp 197–198 °C. MS (ESI) m/z (%): 195 (100, $\text{M}+\text{Na}^+$), 173 (15, $\text{M}+\text{H}^+$). HRMS (ESI) calcd. for $\text{C}_5\text{H}_9\text{ON}_4\text{S}$ ($\text{M}+\text{H}^+$): 173.0492, found: 173.0492. ^1H NMR (DMSO- d_6): δ 10.56 (bs, 1H, NH), 8.45 (s, 1H, CH-7), 3.17 (s, 3H, 9b), 3.02 (s, 3H, 9a). ^{13}C NMR (DMSO- d_6): δ 187.02 (C-5), 168.22 (C-3), 157.98 (CH-7), 40.95 (CH₃-9b), 35.04 (CH₃-9a).

5.2.3. Diisopropyl (E)-((2-((5-(((dimethylamino)methylene)amino)-1,2,4-thiadiazol-3-yl)oxy)ethoxy)methyl)phosphonate (7)

Diisopropyl 2-(phosphonmethoxy)ethyl chloride (6.28 g; 24.3 mmol) and caesium carbonate (11.89 g; 36.5 mmol) were added to a solution of **6** (4.59 g; 26.7 mmol) in DMF (150 mL) and the mixture was stirred at 95 °C for 5 h. After cooling to RT, the mixture was diluted with EtOAc (300 mL), washed with aqueous NaHCO_3 (2x 250 mL) and then with brine (1x 250 mL). An organic layer was dried over MgSO_4 and evaporated. The residue was chromatographed on a silica gel column in a gradient of 0–5% MeOH in CHCl_3 . Yield: 4.15 g (39%) as yellow oil. MS (ESI) m/z (%): 417 (90, $\text{M}+\text{Na}^+$), 395 (100, $\text{M}+\text{H}^+$). HRMS (ESI) calcd. for $\text{C}_{14}\text{H}_{27}\text{O}_5\text{N}_4\text{NaPS}$ ($\text{M}+\text{Na}^+$): 417.1332, found: 417.1337; HRMS (ESI) calcd. for $\text{C}_{14}\text{H}_{26}\text{O}_5\text{N}_4\text{PS}$ ($\text{M}+\text{H}^+$): 395.1513, found: 395.1511. ^1H NMR (DMSO- d_6): δ 8.40 (m, 1H, CH-7), 4.59 (dsept, 2H, J (H,P) = 7.8 Hz, $J_{\text{vic}} = 6.2$ Hz, CH-*i*-Pr), 4.37 (m, 2H, H-1'), 3.81 (m, 2H, H-2'), 3.80 (d, 2H, J (3',P) = 8.3 Hz, H-3'), 3.16 (d, 3H, J (9b,7) = 0.5 Hz, 9b), 3.02 (d, 3H, J (9a,7) = 0.8 Hz, 9a), 1.24 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.23 (d, 6H, $J_{\text{vic}} = 6.1$ Hz, CH₃-*i*-Pr). ^{13}C NMR (DMSO- d_6): δ 191.16 (C-5), 168.05 (C-3), 157.93 (CH-7), 70.71 (d, J (2',P) = 12.0 Hz, C-2'), 70.35 (d, J (C,P) = 6.3 Hz, CH-*i*-Pr), 67.59 (C-1'), 65.08 (d, J (3',P) = 164.6 Hz, C-3'), 40.83 (9b), 34.99 (9a), 24.01 (d, J (C,P) = 3.8 Hz, CH₃-*i*-Pr), 23.87 (d, J (C,P) = 4.6 Hz, CH₃-*i*-Pr). ^{31}P $\{^1\text{H}\}$ NMR (DMSO- d_6): δ 20.40. FTIR (KBr, cm^{-1}): ν : 2980, 2881, 2820, 1705, 1624, 1468, 1375, 1253, 1193, 1107, 1060, 1010, 985, 950, 889.

5.2.4. Diisopropyl ((2-((5-amino-1,2,4-thiadiazol-3-yl)oxy)ethoxy)methyl)phosphonate (8)

Hydrazine (8 mL; 35% aqueous solution) was added to a solution of **7** (4.10 g; 10.4 mmol) in MeOH (15 mL) and the mixture was stirred at RT for 1 h. The solvent was evaporated and the residue was chromatographed on a silica gel column in a gradient of 0–5% MeOH in CHCl_3 . Yield: 2.7 g (75%) of a colorless amorphous solid. MS (ESI) m/z (%): 362 (100, $\text{M}+\text{Na}^+$), 340 (23, $\text{M}+\text{H}^+$). HRMS (ESI) calcd. for $\text{C}_{11}\text{H}_{22}\text{O}_5\text{N}_3\text{PS}$ ($\text{M}+\text{Na}^+$): 362.0910, found: 362.0909. ^1H NMR (DMSO- d_6): δ 7.94 (s, 2H, NH₂), 4.59 (dsept, 2H, J (H,P) = 7.8 Hz, $J_{\text{vic}} = 6.2$ Hz, CH-*i*-Pr), 4.30 (m, 2H, H-1'), 3.79 (d, 2H, J (3',P) = 8.3 Hz, H-3'), 3.78 (m, 2H, H-2'), 1.24 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.23 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr). ^{13}C NMR (DMSO- d_6): δ 182.76 (C-5), 167.37 (C-3), 70.72 (d, J (2',P) = 12.1 Hz, CH₂-2'), 70.38 (d, J (C,P) = 6.3 Hz, CH-*i*-Pr), 67.47 (CH₂-1'), 65.00 (d, J (3',P) = 164.4 Hz, CH₂-3'), 24.03 (d, J (C,P) = 3.8 Hz, CH₃-*i*-Pr), 23.89 (d, J (C,P) = 4.5 Hz, CH₃-*i*-Pr). ^{31}P $\{^1\text{H}\}$ NMR (DMSO- d_6): δ 20.38. FTIR (KBr, cm^{-1}): ν_{max} : 3298, 3162, 2980, 2935, 2882, 1631, 1548, 1507, 1465, 1454, 1409, 1387, 1375, 1337, 1237, 1222, 1178, 1142, 1105, 1013, 995, 891.

5.2.5. ((2-((5-Amino-1,2,4-thiadiazol-3-yl)oxy)ethoxy)methyl)phosphonic acid (9)

Bromotrimethylsilane (21.6 mL; 0.164 mol) was added to a solution of **8** (2.78 g; 8.19 mmol) in acetonitrile (50 mL) and the mixture was stirred at RT overnight. The volatiles were removed *in vacuo*, the residue was co-evaporated with acetonitrile (3x) and then with 25% aqueous ammonia. The crude product was purified on a column of Dowex 50 × 8 (H^+ form). Impurities and salts were eluted first with water and the product was then eluted with 2.5% aqueous ammonia. The compound was crystallized from the MeOH/diethyl ether mixture. Yield: 1.52 g (73%) of a white crystalline solid, mp 161–163 °C. MS (ESI) m/z (%): 276 (30, $\text{M}-2\text{H}^++\text{Na}^+$), 254 (100, $\text{M}-\text{H}^+$). HRMS (ESI) calcd. for $\text{C}_5\text{H}_9\text{O}_5\text{N}_3\text{PS}$ ($\text{M}-\text{H}^+$): 254.0006, found: 254.0006. ^1H NMR (DMSO- d_6): δ 8.00 (s, 2H, NH₂), 7.47 (vbs, NH₂), 4.26 (m, 2H, H-1'), 3.71 (m, 2H, H-2'), 3.37 (d, 2H, J (3',P) = 8.7 Hz, H-3'). ^{13}C NMR (DMSO- d_6): δ 182.74 (C-5), 167.44 (C-3), 70.07 (d, J (2',P) = 10.0 Hz, CH₂-2'), 69.08 (d, J (C,P) = 155.0 Hz, CH₂-3'), 67.88 (CH₂-1'). ^{31}P $\{^1\text{H}\}$ NMR (DMSO- d_6): δ 12.88. FTIR (KBr, cm^{-1}): ν_{max} : 3290, 3179, 2855, 1635, 1547, 1513, 1455, 1414, 1390, 1339, 1105, 1063, 975, 922.

5.2.6. The reaction of compound 6 with (S)-glycidyl trityl ether

Compound **6** (2.50 g; 14.5 mmol), (S)-glycidyl trityl ether (4.18 g; 13.2 mmol) and caesium carbonate (0.860 g; 2.64 mmol) were dissolved in DMF (50 mL) and the mixture was stirred at 95 °C for 4 h. Diethyl ether (200 mL) was added and the mixture was washed with an aqueous solution of NaHCO_3 (2x 150 mL) and then with brine (150 mL). The organic phase was dried over MgSO_4 , evaporated, and the residue was purified on a silica gel column in a gradient of 0–2% MeOH in CHCl_3 . Compound **10a** was eluted first, followed by compound **10b**.

5.2.6.1. (*R,E*)-N'-(3-(2-Hydroxy-3-(trityloxy)propoxy)-1,2,4-thiadiazol-5-yl)-N,N-dimethylformimidamide (10a). Yield: 2.5 g (35%) as a white crystalline solid, mp 72–74 °C. $[\alpha]_D - 2.5$ (c 0.360 g/100 mL, MeOH). MS (ESI) m/z (%): 511 (100, $\text{M}+\text{Na}^+$), 999 (18, $2\text{M}+\text{Na}^+$). HRMS (ESI) calcd. for $\text{C}_{27}\text{H}_{28}\text{O}_3\text{N}_4\text{NaS}$ ($\text{M}+\text{Na}^+$): 511.1774, found: 511.1773. ^1H NMR (DMSO- d_6): δ 8.39 (m, 1H, H-7), 7.37–7.42 (m, 6H, H-*o*-Tr), 7.29–7.35 (m, 6H, H-*m*-Tr), 7.23–7.28 (m, 3H, H-*p*-Tr), 5.23 (d, 1H, J (OH,2') = 5.6 Hz, OH), 4.33 (dd, 1H, $J_{\text{gem}} = 10.8$ Hz, J (1'b,2') = 4.4 Hz, H-1'b), 4.26 (dd, 1H, $J_{\text{gem}} = 10.8$ Hz, J (1'a,2') = 5.8 Hz, H-1'a), 3.97 (pentd, 1H, J (2',1'a) = J (2',3'a) = J (2',3'b) = J (2',OH) = 5.7 Hz, J (2',1'b) = 4.4 Hz, H-2'), 3.15 (bs, 3H, 9b), 3.04 (dd, 1H, $J_{\text{gem}} = 9.2$ Hz, J (3'b,2') = 5.6 Hz, H-3'b), 3.02 (d, 3H, J (9a,7) = 0.8 Hz, 9a), 3.00 (dd, 1H, $J_{\text{gem}} = 9.2$ Hz, J (3'a,2') = 5.7 Hz, H-3'a). ^{13}C NMR (DMSO- d_6): δ 191.07 (C-5, determined by HSQC and HMBC), 168.24 (C-3), 157.87 (CH-7), 143.98 (C-*i*-Tr), 128.47 (CH-*o*-Tr), 128.19 (CH-*m*-Tr), 127.19

(CH-*p*-Tr), 86.00 (C-*Tr*), 70.32 (C-1'), 67.91 (C-2'), 64.99 (C-3'), 40.83 (9b), 35.00 (9a). FTIR (KBr, cm^{-1}) ν_{max} : 3415, 3086, 3057, 3032, 3023, 3001, 2928, 2815, 1750, 1685, 1623, 1597, 1583, 1523, 1497, 1449, 1337, 1323, 1154, 1107, 1076, 1033, 1002, 946, 938, 765, 707, 700, 633, 316.

5.2.6.2. (*S,E*)-*N'*-(2-(2-Hydroxy-3-(trityloxy)propyl)-3-oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)-*N,N*-dimethylformimidamide (**10b**). Yield: 0.62 g (9%) of a white crystalline solid, mp 200–202 °C. $[\alpha]_{\text{D}} - 3.4$ (c 0.295 g/100 mL, MeOH). MS (ESI) m/z (%): 511 (100, M+Na⁺), 999 (19, 2M+Na⁺). HRMS (ESI) calcd. for C₂₇H₂₈O₃N₄Na₄S (M+Na⁺): 511.1774, found: 511.1773. ¹H NMR (DMSO-*d*₆): δ 8.55 (s, 1H, H-7), 7.38–7.42 (m, 6H, H-*o*-Tr), 7.31–7.36 (m, 6H, H-*m*-Tr), 7.23–7.28 (m, 3H, H-*p*-Tr), 5.41 (d, 1H, J(OH,2') = 5.4 Hz, OH), 3.86 (dd, 1H, $J_{\text{gem}} = 13.8$ Hz, J(1'b,2') = 3.4 Hz, H-1'b), 3.81 (m, 1H, H-2'), 3.52 (dd, 1H, $J_{\text{gem}} = 13.8$ Hz, J(1'a,2') = 7.2 Hz, H-1'a), 3.20 (s, 3H, 9b), 3.03 (s, 3H, 9a), 2.97 (dd, 1H, $J_{\text{gem}} = 9.3$ Hz, J(3'b,2') = 5.2 Hz, H-3'b), 2.88 (dd, 1H, $J_{\text{gem}} = 9.4$ Hz, J(3'a,2') = 6.2 Hz, H-3'a). ¹³C NMR (DMSO-*d*₆): δ 181.53 (C-5), 165.94 (C-3), 158.01 (CH-7), 143.88 (C-*i*-Tr), 128.48 (CH-*o*-Tr), 128.07 (CH-*m*-Tr), 127.20 (CH-*p*-Tr), 86.11 (C-*Tr*), 68.67 (C-2'), 65.34 (C-3'), 47.11 (C-1'), 41.12 (9b), 35.17 (9a). FTIR (KBr, cm^{-1}) ν_{max} : 3288, 3085, 3057, 3032, 3022, 3001, 2926, 2810, 1678, 1623, 1596, 1585, 1491, 1448, 1418, 1339, 1307, 1153, 1113, 1077, 1033, 1002, 950, 935, 762, 709, 699, 633, 317.

5.2.7. The general procedure for the synthesis of *N*¹-substituted thiouretes
Potassium thiocyanate (48.0 g; 0.49 mol) was dissolved in acetone (400 mL) and the appropriate acyl chloride (0.41 mol) was added dropwise. The mixture was stirred at 50 °C for 3.5 h. After cooling to RT, the precipitated potassium chloride was filtered off over Celite and the volume of the filtrate was reduced *in vacuo*. Urea (24.6 g; 0.41 mol) was added to the solution and the mixture stirred at 65 °C for 5 h. After cooling to RT, the flask was cooled in an ice bath to initiate crystallization and the solution was then left at RT overnight. Crystals were collected by filtration and recrystallized from methanol.

5.2.7.1. *N*-(Carbamoylcarbamothioyl)acetamide (**13a**). Yield: 9.4 g (14%) of yellowish crystals, mp 196–198 °C. The spectral data are in agreement with Ref.³⁶

5.2.7.2. *N*-(Carbamoylcarbamothioyl)benzamide (**13b**). Yield: 36.18 g (40%) of yellow crystals, mp 174–175 °C. The spectral data are in agreement with Ref.¹⁹

5.2.8. *N*-(3-Oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)acetamide (**14a**)

1 M Br₂ in CHCl₃ (43.8 mL; 43.8 mmol) was added dropwise to a solution of **13a** (5.00 g; 31.0 mmol) in EtOH (40 mL) at 35 °C and the mixture was stirred at the same temperature for 20 min. After cooling to RT, the crystals formed were filtered off, washed with diethyl ether and recrystallized from water. Yield: 2.95 g (60%) as white crystals, mp 236–237 °C. The spectral data are in agreement with Ref.³⁶

5.2.9. *N*-(3-Oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)benzamide (**14b**)

Compound **13b** (7.46 g, 33.4 mmol) and *N*-bromosuccinimide (7.43 g; 41.8 mmol) were dissolved in MeOH (200 mL) and the mixture was stirred at 70 °C for 6 h. After cooling to RT, the crystals formed were filtered off and washed with diethyl ether. Yield: 4.15 g (56%) as white crystals. The spectral data are in agreement with Ref.^{19,37}

5.2.10. (*S,Z*)-*N*-(4-(2-Hydroxy-3-(trityloxy)propyl)-3-oxo-1,2,4-thiadiazolidin-5-ylidene)benzamide (**15**)

Compound **14b** (2.00 g; 9.04 mmol), (*S*)-glycidyl trityl ether (2.62 g; 8.28 mmol) and caesium carbonate (0.553 g; 1.70 mmol) were dissolved in DMF (70 mL), the mixture was stirred at 110 °C for 24 h and evaporated. The residue was chromatographed on a silica gel column in a

gradient of 0–2% MeOH in CHCl₃ and the product was recrystallized from ethyl acetate/hexane. Yield: 1.02 g (21%) as white crystals, mp 176–177 °C. $[\alpha]_{\text{D}} - 23.3$ (c 0.313 g/100 mL, MeOH). MS (ESI) m/z (%): 560 (100, M+Na⁺). HRMS (ESI) calcd. for C₃₁H₂₇O₄N₃NaS (M+Na⁺): 560.1615, found: 560.1609. ¹H NMR (DMSO-*d*₆): δ 10.24 (bs, 1H, NH, determined by HMBC), 8.16 (m, 2H, H-*o*-Ph), 7.63 (m, 1H, H-*p*-Ph), 7.46 (m, 2H, H-*m*-Ph), 7.35–7.41 (m, 6H, H-*o*-Tr), 7.24–7.30 (m, 6H, H-*m*-Tr), 7.19–7.24 (m, 3H, H-*p*-Tr), 5.30 (d, 1H, J(OH,2') = 5.8 Hz, OH), 4.31 (bdpent, 1H, J(2',1'a) = 7.8 Hz, J(2',1'b) = J(2',3') = J(2',OH) = 5.7 Hz, H-2'), 4.09 (dd, 1H, $J_{\text{gem}} = 13.2$ Hz, J(1'b,2') = 5.5 Hz, H-1'b), 4.02 (dd, 1H, $J_{\text{gem}} = 13.2$ Hz, J(1'a,2') = 7.9 Hz, H-1'a), 3.12 (dd, 1H, $J_{\text{gem}} = 9.5$ Hz, J(3'b,2') = 5.1 Hz, H-3'b), 2.98 (dd, 1H, $J_{\text{gem}} = 9.5$ Hz, J(3'a,2') = 6.3 Hz, H-3'a). ¹³C NMR (DMSO-*d*₆): δ 176.97 (C-7), 171.52 (C-5), 153.51 (C-3), 143.91 (C-*i*-Tr), 133.97 (C-*i*-Ph), 133.23 (CH-*p*-Ph), 129.25 (CH-*o*-Ph), 128.80 (CH-*m*-Ph), 128.34 (CH-*o*-Tr), 128.02 (CH-*m*-Tr), 127.17 (CH-*p*-Tr), 86.27 (C-*Tr*), 66.70 (CH₂-3'), 66.33 (CH-2'), 48.34 (CH₂-1').

5.2.11. Diisopropyl ((2-(*N*-phthalimido)ethoxy)methyl)phosphonate (**16**)

A solution of diisopropyl 2-(phosphonomethoxy)ethyl chloride (20.28 g; 78.4 mmol) and potassium phthalimide (14.52 g; 78.4 mmol) in DMF (200 mL) was stirred at 120 °C for 3.5 h and then evaporated. The residue was co-evaporated with toluene and purified on a silica gel column in a gradient of 0–5% MeOH in CHCl₃. Yield: 24.7 g (85%) of a white amorphous solid. MS (ESI) m/z (%): 392 (100, M+Na⁺). HRMS (ESI) calcd. for C₁₇H₂₄O₆NNaP (M+Na⁺): 392.1234, found: 392.1235. ¹H NMR (CDCl₃): δ 7.77 (m, 2H, CH-4,7-Ar), 7.67 (m, 2H, CH-5,6-Ar), 4.63 (m, 2H, CH-*i*-Pr), 3.85 (m, 2H, H-1'), 3.80 (m, 2H, H-2'), 3.71 (d, 2H, J(3',P) = 8.3 Hz, H-3'), 1.19–1.22 (m, 12H, CH₃-*i*-Pr). ¹³C NMR (CDCl₃): δ 167.95 (C-2, C-9), 133.85 (C_A-3,4), 131.93 (C_A-1,6), 123.09 (CH-4,7-Ar), 70.96 (d, J(C,P) = 6.6 Hz, CH-*i*-Pr), 69.55 (d, J(2',P) = 10.8 Hz, CH₂-2'), 65.28 (d, J(3',P) = 167.2 Hz, CH₂-3'), 36.90 (CH₂-1'), 23.87 (d, J(C,P) = 3.8 Hz, CH₃-*i*-Pr), 23.76 (d, J(C,P) = 4.6 Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (CDCl₃): δ 18.84. FTIR (CCL₄, cm^{-1}) ν_{max} : 3088, 3061, 2981, 2938, 2875, 1778, 1720, 1617, 1469, 1393, 1375, 1241, 1179, 1142, 1107, 1009, 990, 889, 718, 529.

5.2.12. 2-(Tritylamino)ethan-1-ol (**18**)

Triethylamine (60 mL; 0.431 mol) was added to a solution of 2-aminoethan-1-ol (13.0 g; 0.21 mol) in CH₂Cl₂ (250 mL). Trityl chloride (59.1 g; 0.212 mol) was added under cooling in an ice bath and the mixture was stirred at RT for 2 h. The solvent was evaporated and the residue was partitioned between ethyl acetate (300 mL) and aqueous NaHCO₃ (2x 250 mL) and brine (250 mL). The organic layer was dried over MgSO₄ and evaporated. The residue was crystallized from the mixture of ethyl acetate/hexane. Yield: 52.4 g (82%) as a white crystalline solid. The spectral data are in agreement with Ref.³⁸

5.2.13. Diisopropyl ((2-(tritylamino)ethoxy)methyl)phosphonate (**19**)

The synthetic methodology described in Ref.²⁴ was utilized. Magnesium di-*tert*-butoxide (5.63 g; 33.0 mmol) and diisopropyl tosyloxymethanephosphonate (6.94 g; 19.8 mmol) were added to a solution of **18** (5.0 g; 16.5 mmol) in *N*-methylpyrrolidone (30 mL). The mixture was stirred at 75–78 °C for 4.5 h. After cooling to RT, water (30 mL) was added and the mixture was extracted with diethyl ether (50 mL). An organic layer was washed with aqueous NaHCO₃ (2x 50 mL), then with brine (1x 50 mL), and it was dried over MgSO₄. The solvent was evaporated and the residue was chromatographed on a silica gel column in the CHCl₃/hexane system (1:9 to 1:1). Yield: 7.2 g (91%) of a colorless oil. MS (ESI) m/z (%): 504 (90, M+Na). HRMS (ESI) calcd. for C₂₈H₃₆O₄NNaP (M+Na⁺): 504.2274, found: 504.2273; HRMS (ESI) calcd. for C₂₈H₃₇O₄NP (M+H⁺): 482.2455, found: 482.2455. ¹H NMR (DMSO-*d*₆): δ 7.41 (m, 6H, H-*o*-Tr), 7.29 (m, 6H, H-*m*-Tr), 7.19 (m, 3H, H-*p*-Tr), 4.59 (dsept, 2H, J(H,P) = 7.8 Hz, J_{vic} = 6.2 Hz, CH-*i*-Pr), 3.68 (d, 2H, J(3,P) = 8.4 Hz, H-3), 3.62 (t, 2H, J(2,1) = 5.6 Hz, H-2), 2.59 (t, 1H, J(NH,1) = 7.9 Hz, NH), 2.16 (dt, 2H, J(1,NH) = 7.8 Hz, J(1,2) = 5.6

H_z, H-1), 1.23 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr), 1.21 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr). ¹³C NMR (DMSO-*d*₆): δ 146.09 (C-*i*-Tr), 128.47 (CH-*o*-Tr), 127.88 (CH-*m*-Tr), 126.29 (CH-*p*-Tr), 72.56 (d, $J(2, \text{P}) = 12.1$ Hz, CH₂-2), 70.27 (C-*Tr*), 70.22 (d, $J(\text{C}, \text{P}) = 6.5$ Hz, CH-*i*-Pr), 64.96 (d, $J(3, \text{P}) = 165.0$ Hz, CH₂-3), 43.09 (CH₂-1), 23.96 (d, $J(\text{C}, \text{P}) = 3.9$ Hz, CH₃-i-Pr), 23.88 (d, $J(\text{C}, \text{P}) = 4.5$ Hz, CH₃-i-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 21.81. FTIR (CCl₄, cm⁻¹) ν_{max} : 3335, 3085, 3062, 3033, 2980, 2934, 2873, 1597, 1489, 1466, 1449, 1386, 1375, 1318, 1242, 1216, 1191, 1180, 1142, 1107, 1033, 1008, 990, 940, 902, 706, 698, 648, 625.

5.2.14. Diisopropyl ((2-aminoethoxy)methyl)phosphonate (17)

Method A. Hydrazine (1 M in THF; 140 mL, 0.141 mol) was added to a solution of **16** (5.2 g; 14.1 mmol) in EtOH (50 mL) and the mixture was stirred at RT for 2 h. The precipitated phthallhydrazide was filtered off. The filtrate was evaporated and the residue was purified on a silica gel column under basic conditions in the system of 2–10% methanolic ammonia (2.35 M NH₃) in CHCl₃. Yield: 3.07 g (91%) of a colorless oil.

Method B. 35% Hydrochloric acid (30 mL) was added to a 0 °C cold solution of **19** (7.0 g; 4.5 mmol) in dioxane (40 mL). The solution was then stirred at RT for 1.5 h, washed with diethyl ether, and the aqueous phase was concentrated *in vacuo*. The residue was purified on a silica gel column under basic conditions (5% TEA in toluene/CHCl₃ 2:1 to TEA/CHCl₃ 5:95). Yield: 2.26 g (65%) of a colorless oil. MS (ESI) *m/z* (%): 262 (27, M+Na⁺), 240 (100, M+H⁺). HRMS (ESI) calcd. for C₉H₂₂O₄NNaP (M+Na⁺): 262.1179, found: 262.1179; HRMS (ESI) calcd. for C₉H₂₃O₄NP (M+H⁺): 240.1359, found: 240.1360. ¹H NMR (CDCl₃): δ 4.73 (dsept, 2H, J(H,P) = 7.7 Hz, $J_{\text{vic}} = 6.2$ Hz, CH-*i*-Pr), 3.73 (d, 2H, J(3,P) = 8.5 Hz, H-3), 3.68 (t, 2H, J(2,1) = 5.1 Hz, H-2), 2.97 (t, 2H, J(1,2) = 4.9 Hz, H-1), 1.33 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr), 1.32 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr). ¹³C NMR (CDCl₃): δ 75.34 (d, J(C,P) = 11.2 Hz, CH₂-2), 70.89 (d, J(C,P) = 6.7 Hz, CH-*i*-Pr), 65.68 (d, J(3,P) = 168.2 Hz, CH₂-3), 41.38 (CH₂-1), 24.02 (d, J(C,P) = 3.9 Hz, CH₃-i-Pr), 23.95 (d, J(C,P) = 4.6 Hz, CH₃-i-Pr). ³¹P {¹H} NMR (CDCl₃): δ 20.15. FTIR (CCl₄, cm⁻¹) ν_{max} : 3334, 2981, 2934, 2876, 1467, 1387, 1376, 1241, 1179, 1142, 1106, 1010, 993, 889.

5.2.15. (S)-1-Azido-3-(trityloxy)propan-2-ol (20)

Sodium azide (41.1 g; 632 mmol) was added to a solution of (S)-glycidyl trityl ether (10.0 g; 31.6 mmol) in DMF (500 mL) and the mixture was stirred at 45 °C for 7 days. Water (5.0 mL) was added and the reaction mixture was stirred for 10 min. The precipitated salts were filtered off on a Büchner funnel and the filtrate was evaporated. The residue was partitioned between water (350 mL) and ethyl acetate (2x 350 mL). An organic layer was dried over Na₂SO₄, evaporated, and the residue was purified on a silica gel column in CHCl₃/hexane (15:85). Yield: 7.53 g (66%) of a colorless oil. $[\alpha]_{\text{D}}^{20} = -17.6$ (c 0.250 g/100 mL, MeOH). MS (ESI) *m/z* (%): 382 (100, M+Na⁺). HRMS (ESI) calcd. for C₂₂H₂₁O₂N₃Na (M+Na⁺): 382.1526, found: 382.1523. ¹H NMR (DMSO-*d*₆): δ 7.36–7.42 (m, 6H, H-*o*-Tr), 7.30–7.38 (m, 6H, H-*m*-Tr), 7.23–7.30 (m, 3H, H-*p*-Tr), 5.35 (d, 1H, J(OH,2) = 5.3 Hz, OH), 3.83 (tt, 1H, J(2,1b) = J(2,3b) = 6.5 Hz, J(2,OH) = J(2,3a) = 5.3 Hz, J(2,1a) = 3.6 Hz, H-2), 3.36 (dd, 1H, $J_{\text{gem}} = 12.6$ Hz, J(1a,2) = 3.6 Hz, H-1a), 3.30 (dd, 1H, $J_{\text{gem}} = 12.6$ Hz, J(1b,2) = 6.5 Hz, H-1b), 3.01 (dd, 1H, $J_{\text{gem}} = 9.2$ Hz, J(3a,2) = 5.3 Hz, H-3a), 2.89 (dd, 1H, $J_{\text{gem}} = 9.2$ Hz, J(3b,2) = 6.5 Hz, H-3b). ¹³C NMR (DMSO-*d*₆): δ 143.90 (C-*i*-Tr), 128.45 (CH-*o*-Tr), 128.08 (CH-*m*-Tr), 127.22 (CH-*p*-Tr), 86.11 (C-*Tr*), 69.14 (CH-2), 65.27 (CH₂-3), 53.80 (CH₂-1). FTIR (CCl₄, cm⁻¹) ν_{max} : 3600, 2928, 2858, 2105, 1737, 1467, 1393, 1087, 1077, 529.

5.2.16. (Diisopropyl (S)-(((1-azido-3-(trityloxy)propan-2-yl)oxy)methyl)phosphonate (22)

Sodium hydride (60% suspension in mineral oil; 2.095 g; 52.3 mmol) was added to a solution of **20** (7.5 g; 20.9 mmol) in DMF (300 mL) and the mixture was stirred at RT for 15 min. Diisopropyl (bromomethyl)phosphonate **21** (8.12 g; 31.4 mmol) was added and the mixture stirred at 60 °C overnight. The reaction was quenched with water (30 mL), the

mixture was stirred for 5 min and evaporated. The residue was extracted between water (300 mL) and ethyl acetate (2x 450 mL). Combined organic layers were dried over Na₂SO₄ and evaporated. The residue was purified on a silica gel column, first in CHCl₃/hexane (40:60), followed by a gradient of MeOH in CHCl₃ (0–15%). Yield: 7.85 g (70%) of a colorless oil. $[\alpha]_{\text{D}}^{20} = -7.4$ (c 0.298 g/100 mL, MeOH). MS (ESI) *m/z* (%): 560 (100, M+Na⁺). HRMS (ESI) calcd. for C₂₉H₃₆O₅N₃NaP (M+Na⁺): 560.2285, found: 560.2284. ¹H NMR (DMSO-*d*₆): δ 7.36–7.43 (m, 6H, H-*o*-Tr), 7.31–7.38 (m, 6H, H-*m*-Tr), 7.25–7.29 (m, 3H, H-*p*-Tr), 4.53–4.64 (m, 2H, CH-*i*-Pr), 3.84 (dd, 1H, $J_{\text{gem}} = 13.9$ Hz, J(3a,P) = 9.0 Hz, H-3a), 3.81 (dd, 1H, $J_{\text{gem}} = 13.9$ Hz, J(3b,P) = 9.0 Hz, H-3b), 3.74 (dddd, 1H, J(2,1b) = 5.7 Hz, J(2,4b) = 5.3 Hz, J(2,4a) = 4.9 Hz, J(2,1a) = 4.1 Hz, H-2), 3.56 (dd, 1H, $J_{\text{gem}} = 13.1$ Hz, J(1a,2) = 4.1 Hz, H-1a), 3.45 (dd, 1H, $J_{\text{gem}} = 13.1$ Hz, J(1b,2) = 5.7 Hz, H-1b), 3.10 (dd, 1H, $J_{\text{gem}} = 10.2$ Hz, J(4a,2) = 4.9 Hz, H-4a), 3.08 (dd, 1H, $J_{\text{gem}} = 10.2$ Hz, J(4b,2) = 5.3 Hz, H-4b), 1.18–1.28 (m, 12H, CH₃-i-Pr). ¹³C NMR (DMSO-*d*₆): δ 143.70 (C-*i*-Tr), 128.44 (CH-*o*-Tr), 128.12 (CH-*m*-Tr), 127.30 (CH-*p*-Tr), 86.36 (C-*Tr*), 79.35 (d, J(2,P) = 12.2 Hz, CH-2), 70.38 (d, J(C,P) = 6.3 Hz, CH-*i*-Pr), 65.86 (CH₂-4), 64.07 (d, J(3,P) = 165.2 Hz, CH₂-3), 50.93 (CH₂-1), 23.72–24.06 (CH₃-i-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 18.82. FTIR (CCl₄, cm⁻¹) ν_{max} : 3088, 3062, 3035, 3024, 2981, 2934, 2874, 2103, 1598, 1491, 1467, 1449, 1386, 1375, 1310, 1260, 1244, 1178, 1142, 1159, 1153, 1107, 1088, 1080, 1035, 1008, 991, 936, 889, 705, 697, 633, 555.

5.2.17. Diisopropyl (S)-(((1-amino-3-(trityloxy)propan-2-yl)oxy)methyl)phosphonate (23)

Triphenylphosphine (4.19 g; 16.0 mmol) and water (2.0 mL) were added while stirring to a 0 °C cold solution of **22** (7.80 g; 14.5 mmol) in THF (200 mL) and the mixture left stirring overnight to reach RT. The solvent was evaporated and the residue was chromatographed on a silica gel column (a gradient of 0–5% MeOH in CHCl₃ containing 1% TEA). Yield: 4.8 g (65%) of a colorless oil. $[\alpha]_{\text{D}}^{20} = -8.8$ (c 0.656 g/100 mL, MeOH). MS (ESI) *m/z* (%): 1024 (88, 2M+H⁺), 512 (85, M+H⁺). HRMS (ESI) calcd. for C₂₉H₃₉O₅NP (M+H⁺): 512.2560, found: 512.2557. ¹H NMR (DMSO-*d*₆): δ 7.38–7.42 (m, 6H, H-*o*-Tr), 7.31–7.36 (m, 6H, H-*m*-Tr), 7.23–7.28 (m, 3H, H-*p*-Tr), 4.55–4.65 (m, 2H, CH-*i*-Pr), 3.85 (dd, 1H, $J_{\text{gem}} = 13.3$ Hz, J(3a,P) = 8.6 Hz, H-3a), 3.81 (dd, 1H, $J_{\text{gem}} = 13.3$ Hz, J(3b,P) = 8.6 Hz, H-3b), 3.50 (dddd, 1H, J(2,1b) = 6.1 Hz, J(2,4b) = 5.7 Hz, J(2,1a) = 5.2 Hz, J(2,4a) = 3.9 Hz, H-2), 3.11 (dd, 1H, $J_{\text{gem}} = 10.0$ Hz, J(4a,2) = 3.9 Hz, H-4a), 3.04 (dd, 1H, $J_{\text{gem}} = 10.0$ Hz, J(4b,2) = 5.7 Hz, H-4b), 2.66 (dd, 1H, $J_{\text{gem}} = 13.2$ Hz, J(1a,2) = 5.2 Hz, H-1a), 2.62 (dd, 1H, $J_{\text{gem}} = 13.2$ Hz, J(1b,2) = 6.1 Hz, H-1b), 1.24 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr), 1.23 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr), 1.22 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr), 1.21 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-i-Pr). ¹³C NMR (DMSO-*d*₆): δ 143.94 (C-*i*-Tr), 128.45 (CH-*o*-Tr), 128.06 (CH-*m*-Tr), 127.20 (CH-*p*-Tr), 86.15 (C-*Tr*), 82.78 (d, J(2,P) = 11.8 Hz, CH-2), 70.30 (d, J(C,P) = 6.2 Hz, CH-*i*-Pr), 70.26 (d, J(C,P) = 6.2 Hz, CH-*i*-Pr), 64.34 (d, J(3,P) = 165.3 Hz, CH₂-3), 63.83 (CH₂-4), 42.59 (CH₂-1), 24.02 (d, J(C,P) = 3.6 Hz, CH₃-i-Pr), 24.01 (d, J(C,P) = 3.6 Hz, CH₃-i-Pr), 23.93 (d, J(C,P) = 4.5 Hz, CH₃-i-Pr), 23.91 (d, J(C,P) = 4.5 Hz, CH₃-i-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 19.55. ³¹P {¹H} NMR (DMSO-*d*₆): δ 19.55. FTIR (CCl₄, cm⁻¹) ν_{max} : 3088, 3061, 3035, 3024, 2980, 2874, 1598, 1491, 1466, 1449, 1386, 1375, 1241, 1179, 1159, 1153, 1142, 1106, 1089, 1078, 992, 936, 890, 787, 764, 706, 633.

5.2.18. Diisopropyl ((2-ureidoethoxy)methyl)phosphonate (24)

35% hydrochloric acid (1.65 mL) was added dropwise to a 10 °C cold solution of **17** (5.12 g; 21.4 mmol) in H₂O (50 mL). Potassium cyanate (3.47 g; 42.8 mmol) was added and the reaction mixture was stirred at 70 °C overnight. The mixture was extracted with CHCl₃ (4x 70 mL), an organic layer was dried over MgSO₄ and concentrated *in vacuo*. The residue was purified on a silica gel column (0–5% MeOH in CHCl₃). Yield: 4.33 g (72%) of a colorless oil. MS (ESI) *m/z* (%): 305 (100, M+Na⁺). HRMS (ESI) calcd. for C₁₀H₂₃O₅N₂NaP (M+Na⁺): 305.1237, found: 305.1237. ¹H NMR (DMSO-*d*₆): δ 5.95 (bt, 1H, J(H,1') = 5.8 Hz, NH), 5.48 (s, 2H, NH₂), 4.60 (dsept, 2H, J(H,P) = 7.7 Hz, $J_{\text{vic}} = 6.2$ Hz,

CH-*i*-Pr), 3.72 (d, 2H, $J(3',P) = 8.2$ Hz, H-3'), 3.47 (t, 2H, $J(2',1') = 5.8$ Hz, H-2'), 3.12 (q, 2H, $J(1',2') = J(1',NH) = 5.8$ Hz, H-1'), 1.25 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.24 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 158.86 (C-1), 72.41 (d, $J(2',P) = 11.2$ Hz, CH₂-2'), 70.42 (d, $J(C,P) = 6.4$ Hz, CH-*i*-Pr), 64.94 (d, $J(C,P) = 164.4$ Hz, CH₂-3'), 39.05 (CH₂-1'), 24.09 (d, $J(C,P) = 3.7$ Hz, CH₃-*i*-Pr), 23.98 (d, $J(C,P) = 4.6$ Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 19.67. FTIR (CCl₄, cm⁻¹) ν_{max} : 3362, 3210, 2982, 2935, 2876, 1690, 1665, 1612, 1555, 1467, 1387, 1376, 1240, 1179, 1142, 1105, 1014, 995, 890.

5.2.19. Diisopropyl (*S*)-(((1-(*trityloxy*)-3-ureidopropan-2-yl)oxy)methyl)phosphonate (25)

Compound **23** (4.75 g; 9.29 mmol) and 1,1'-carbonyldiimidazole (1.81 g; 11.2 mmol) were dissolved in dichloromethane (200 mL) and the reaction mixture was stirred at RT for 3 h. Afterwards, 7 M methanolic ammonia (50 mL; 0.350 mol) was added and the mixture was stirred at RT overnight. The solvent was evaporated and the residue was purified on a silica gel column (0–5% MeOH in CHCl₃). Yield: 3.71 g (72%) of a colorless oil. $[\alpha]_D^{20} = -3.7$ (c 0.216 g/100 mL, MeOH). MS (ESI) *m/z* (%): 577 (100, M+Na⁺). HRMS (ESI) calcd. for C₃₀H₃₉O₆N₂NaP (M+Na⁺): 577.2438, found: 577.2439; HRMS (ESI) calcd. for C₃₀H₄₀O₆N₂P (M+H⁺): 555.2619, found: 555.2619. ¹H NMR (CDCl₃): δ 7.43 (m, 6H, H-*o*-Tr), 7.30 (m, 6H, H-*m*-Tr), 7.24 (m, 3H, H-*p*-Tr), 6.26 (bs, 1H, NH), 4.77 (dsept, 1H, $J(H,P) = 7.6$ Hz, $J_{vic} = 6.2$ Hz, CH-*i*-Pr), 4.72 (dsept, 1H, $J(H,P) = 7.6$ Hz, $J_{vic} = 6.2$ Hz, CH-*i*-Pr), 4.52 (bs, 2H, NH₂), 4.13 (dd, 1H, $J_{gem} = 14.1$ Hz, $J(3'a,P) = 6.4$ Hz, H-3'a), 3.76 (dd, 1H, $J_{gem} = 14.1$ Hz, $J(3'b,P) = 8.7$ Hz, H-3'b), 3.62 (m, 1H, H-2'), 3.50 (ddd, 1H, $J_{gem} = 13.8$ Hz, $J(1'a,2') = 7.8$ Hz, $J(1'a,NH) = 3.1$ Hz, H-1'a), 3.17 (dd, 1H, $J_{gem} = 10.2$ Hz, $J(4'a,2') = 6.1$ Hz, H-4'a), 3.13 (dd, 1H, $J_{gem} = 10.2$ Hz, $J(4'b,2') = 4.0$ Hz, H-4'b), 3.01 (ddd, 1H, $J_{gem} = 13.8$ Hz, $J(1'b,2') = 9.3$ Hz, $J(1'b,NH) = 3.0$ Hz, H-1'b), 1.356 (d, 3H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.349 (d, 3H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.340 (d, 3H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.336 (d, 3H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr). ¹³C NMR (CDCl₃): δ 159.02 (C-1), 143.65 (C-*i*-Tr), 128.61 (CH-*o*-Tr), 127.85 (CH-*m*-Tr), 127.10 (CH-*p*-Tr), 86.83 (C-*Tr*), 83.25 (d, $J(2',P) = 11.2$ Hz, CH₂-2'), 71.82 (d, $J(C,P) = 6.6$ Hz, CH-*i*-Pr), 71.24 (d, $J(C,P) = 7.1$ Hz, CH-*i*-Pr), 66.99 (d, $J(3',P) = 168.1$ Hz, CH₂-3'), 64.42 (CH₂-4'), 42.16 (CH₂-1'), 24.15 (d, $J(C,P) = 3.7$ Hz, CH₃-*i*-Pr), 24.09 (d, $J(C,P) = 4.6$ Hz, CH₃-*i*-Pr), 24.02 (d, $J(C,P) = 3.9$ Hz, CH₃-*i*-Pr), 23.93 (d, $J(C,P) = 4.5$ Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (CDCl₃): δ 21.09. FTIR (CHCl₃, cm⁻¹) ν_{max} : 3518, 3413, 3331, 3222, 3089, 3062, 3035, 2985, 2934, 2875, 1675, 1598, 1554, 1491, 1467, 1449, 1443, 1388, 1377, 1305, 1238, 1179, 1159, 1153, 1142, 1103, 1088, 1079, 1035, 1015, 1000, 890, 707, 701, 642, 633, 617.

5.2.20. *N*¹-Substituted thiobiurets with *N*⁵-[[diisopropoxy]phosphoryl]methoxyalkyl arrangement (26–31). The general procedure

Compound **24** or **25** was dissolved in acetone and the appropriate isothiocyanate was added dropwise. The solution was stirred at 60 °C until quantitative conversion (4–24 h). The solvent was removed *in vacuo* and the residue was purified on a silica gel column (a gradient of 40–100% EtOAc/hexane). For specific conditions, see below.

5.2.20.1. Diisopropyl (1,5-dioxo-1-phenyl-3-thioxo-9-oxa-2,4,6-triazadecan-10-yl)phosphonate (26). Prepared from **24** (7.88 g; 27.9 mmol) and benzoyl isothiocyanate (4.56 g; 27.9 mmol) in acetone (300 mL); reaction time: 4 h. Yield: 7.8 g (63%) of a yellowish oil. MS (ESI) *m/z* (%): 468 (100, M+Na⁺), 446 (15, M+H⁺). HRMS (ESI) calcd. for C₁₈H₂₈O₆N₃NaP (M+Na⁺): 468.1329, found: 468.1328; HRMS (ESI) calcd. for C₁₈H₂₉O₆N₃P (M+H⁺): 446.1509, found: 446.1509. ¹H NMR (DMSO-*d*₆): δ 7.90 (m, 2H, H-*o*-Ph), 7.68 (m, 1H, H-*p*-Ph), 7.58 (m, 2H, H-*m*-Ph), 4.60 (dsept, 2H, $J(H,P) = 7.7$ Hz, $J_{vic} = 6.2$ Hz, CH-*i*-Pr), 3.78 (d, 2H, $J(3',P) = 8.5$ Hz, H-3'), 3.61 (t, 2H, $J(2',1') = 5.3$ Hz, H-2'), 3.38 (m, 2H, H-1'), 1.24 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.23 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 178.96 (C-2), 168.03 (C-3),

153.60 (C-1; determined by HMBC), 133.56 (CH-*p*-Ph), 131.42 (C-*i*-Ph), 129.16 (CH-*m*-Ph), 128.09 (CH-*o*-Ph), 71.08 (d, $J(2',P) = 11.4$ Hz, CH₂-2'), 70.38 (d, $J(C,P) = 6.4$ Hz, CH-*i*-Pr), 64.97 (d, $J(3',P) = 164.5$ Hz, CH₂-3'), 39.20 (CH₂-1'), 24.03 (d, $J(C,P) = 3.8$ Hz, CH₃-*i*-Pr), 23.92 (d, $J(C,P) = 4.6$ Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 19.34. FTIR (CCl₄, cm⁻¹) ν_{max} : 3497, 3423, 3330, 3237, 3068, 3035, 2981, 2937, 2875, 1723, 1689, 1675, 1603, 1582, 1540, 1502, 1491, 1467, 1451, 1386, 1375, 1316, 1238, 1179, 1143, 1106, 1070, 1013, 994, 890, 704, 686, 644.

5.2.20.2. Diisopropyl (6,10-dioxo-8-thioxo-2,11-dioxo-5,7,9-triazatridecyl)phosphonate (27). Prepared from **24** (11.22 g; 39.7 mmol) and ethoxycarbonyl isothiocyanate (5.21 g; 39.7 mmol) in acetone (400 mL); reaction time: 4 h. Yield: 10.3 g (63%) of a yellowish oil. MS (ESI) *m/z* (%): 436 (100, M+Na⁺), 414 (47, M+H⁺). HRMS (ESI) calcd. for C₁₄H₂₈O₇N₃NaP (M+Na⁺): 436.1278, found: 436.1281; HRMS (ESI) calcd. for C₁₄H₂₉O₇N₃P (M+H⁺): 414.1458, found: 414.1451. ¹H NMR (DMSO-*d*₆): δ 12.21 (bs, 1H, NH), 11.05 (bs, 1H, NH), 8.34 (bs, 1H, NH), 4.55–4.65 (m, 2H, CH-*i*-Pr), 4.18 (q, 2H, $J(4,5) = 7.1$ Hz, H-4), 3.76 (d, 2H, $J(3',P) = 8.3$ Hz, H-3'), 3.59 (bt, 2H, $J(2',1') = 5.4$ Hz, H-2'), 3.35 (bq, 2H, $J(1',2') = 5.4$ Hz, H-1'), 1.24 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.23 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.229 (t, 3H, $J(5,4) = 7.1$ Hz, H-5). ¹³C NMR (DMSO-*d*₆): δ 178.93 (C-2), 153.13 (C-1), 152.00 (C-3), 71.04 (d, $J(2',P) = 12.1$ Hz, CH₂-2'), 70.36 (d, $J(C,P) = 6.4$ Hz, CH-*i*-Pr), 64.97 (d, $J(3',P) = 164.4$ Hz, CH₂-3'), 62.39 (CH₂-4), 39.22 (CH₂-1'), 24.00 (d, $J(C,P) = 3.7$ Hz, CH₃-*i*-Pr), 23.90 (d, $J(C,P) = 4.5$ Hz, CH₃-*i*-Pr), 14.18 (CH₃-5). ³¹P {¹H} NMR (DMSO-*d*₆): δ 20.34. FTIR (CCl₄, cm⁻¹) ν_{max} : 3252, 3186, 2980, 2935, 2875, 1705, 1537, 1508, 1468, 1454, 1386, 1374, 1337, 1261, 1227, 1177, 1142, 1104, 1012, 993, 890, 794.

5.2.20.3. Diisopropyl (1-(9H-fluoren-9-yl)-3,7-dioxo-5-thioxo-2,11-dioxo-4,6,8-triazadodecan-12-yl)phosphonate (28). Prepared from **24** (1.97 g; 6.96 mmol) and 9-fluorenylmethoxycarbonyl isothiocyanate (2.43 g; 6.96 mmol) in acetone (100 mL); reaction time: 6 h. Yield: 2.13 g (54%) of a white amorphous solid. MS (ESI) *m/z* (%): 586 (95, M+Na⁺), 564 (12, M+H⁺). HRMS (ESI) calcd. for C₂₆H₃₄O₆N₃NaP (M+Na⁺): 586.1747, found: 586.1748. ¹H NMR (DMSO-*d*₆): δ 12.37 (bs, 1H, NH), 11.13 (bs, 1H, NH), 8.52 (bs, 1H, NH), 7.90 (dt, 2H, $J(10a,9a+10b,9b) = 7.5$ Hz, $J(10a,8a+10b,8b) = 10(10a,7a+10b,7b) = 1.1$ Hz, H-10a,b), 7.76 (bd, 2H, $J(7a,8a+7b,8b) = 7.5$ Hz, H-7a,b), 7.43 (td, 2H, $J(9a,10a+9b,10b) = J(9a,8a+9b,8b) = 7.5$ Hz, $J(9a,7a+9b,7b) = 1.2$ Hz, $J(9a,5+9b,5) = 0.7$ Hz, H-9a,b), 7.35 (td, 2H, $J(8a,9a+8b,9b) = J(8a,7a+8b,7b) = 7.5$ Hz, $J(8a,10a+8b,10b) = 1.1$ Hz, H-8a,b), 4.60 (dsept, 2H, $J(H,P) = 7.7$ Hz, $J_{vic} = 6.2$ Hz, CH-*i*-Pr), 4.46 (d, 2H, $J(4,5) = 7.1$ Hz, H-4), 4.31 (bt, 1H, $J(5,4) = 7.1$ Hz, H-5), 3.77 (d, 2H, $J(3',P) = 8.2$ Hz, H-3'), 3.61 (t, 2H, $J(2',1') = 5.3$ Hz, H-2'), 3.37 (q, 2H, $J(1',2') = 5.3$ Hz, H-1'), 1.25 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr), 1.24 (d, 6H, $J_{vic} = 6.2$ Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 178.94 (C-2), 152.93 (C-1), 152.17 (C-3), 143.43 (C-6a,b), 140.94 (C-11a,b), 128.04 (CH-9a,b), 127.37 (CH-8a,b), 125.50 (CH-7a,b), 120.40 (CH-10a,b), 71.03 (d, $J(2',P) = 11.5$ Hz, CH₂-2'), 70.37 (d, $J(C,P) = 6.4$ Hz, CH-*i*-Pr), 67.66 (CH₂-4), 64.98 (d, $J(3',P) = 164.3$ Hz, CH₂-3'), 46.22 (CH-5), 39.30 (CH₂-1'), 24.02 (d, $J(C,P) = 3.7$ Hz, CH₃-*i*-Pr), 23.92 (d, $J(C,P) = 4.5$ Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 20.32. FTIR (KBr, cm⁻¹) ν_{max} : 3322, 3243, 3186, 3066, 3040, 2979, 1721, 1705, 1609, 1538, 1509, 1466, 1451, 1386, 1360, 1338, 1323, 1224, 1209, 1172, 1142, 1105, 1084, 994, 840, 760, 741, 726, 621.

5.2.20.4. Diisopropyl (*S*)-((1,5-dioxo-1-phenyl-3-thioxo-8-((trityloxy)methyl)-9-oxa-2,4,6-triazadecan-10-yl)phosphonate (29). Prepared from **25** (1.50 g; 2.71 mmol) and benzoyl isothiocyanate (0.53 g; 3.25 mmol) in acetone (100 mL); reaction time: 24 h. Yield: 0.59 g (30%) of a yellow amorphous solid. $[\alpha]_D^{20} = -3.4$ (c 0.205 g/100 mL, MeOH). MS (ESI) *m/z* (%): 740 (100, M+Na⁺). HRMS (ESI) calcd. for C₃₈H₄₄O₇N₃NaP (M+Na⁺): 740.2530, found: 740.2524. ¹H NMR (DMSO-*d*₆): δ 7.86–7.89

(m, 2H, H-o-Ph), 7.67 (m, 1H, H-p-Ph), 7.54–7.58 (m, 2H, H-m-Ph), 7.38–7.41 (m, 6H, H-o-Tr), 7.30–7.35 (m, 6H, H-m-Tr), 7.25 (m, 3H, H-p-Tr), 4.53–4.63 (m, 2H, CH-*i*-Pr), 3.84 (dd, 1H, $J_{\text{gem}} = 13.7$ Hz, $J(3'a,P) = 8.8$ Hz, H-3'a), 3.79 (dd, 1H, $J_{\text{gem}} = 13.7$ Hz, $J(3'b,P) = 9.3$ Hz, H-3'b), 3.75 (p, 1H, $J(2',1') = J(2',3') = 5.2$ Hz, H-2'), 3.42–3.47 (bm, 2H, H-1'), 3.09 (dd, 1H, $J_{\text{gem}} = 10.1$ Hz, $J(4'a,2') = 4.9$ Hz, H-4'a), 3.06 (dd, 1H, $J_{\text{gem}} = 10.1$ Hz, $J(4'b,2') = 5.3$ Hz, H-4'b), 1.22 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.21 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.20 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.19 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 168.05 (C-3), 143.72 (C-*i*-Tr), 133.47 (CH-*p*-Ph), 133.10 (C-*i*-Ph), 129.09 (CH-*m*-Ph), 128.44 (CH-*o*-Tr), 128.07 (CH-*o*-Ph, CH-*m*-Tr), 127.26 (CH-*p*-Tr), 86.33 (C-Tr), 78.91 (d, $J(C,P) = 12.0$ Hz, CH-2'), 70.39 (d, $J(C,P) = 6.3$ Hz, CH-*i*-Pr), 70.38 (d, $J(C,P) = 6.3$ Hz, CH-*i*-Pr), 64.02 (d, $J(3',P) = 165.2$ Hz, CH₂-3'), 62.89 (CH₂-4'), 39.87 (CH₂-1'), 24.00 (d, $J(C,P) = 3.7$ Hz, CH₃-*i*-Pr), 23.90 (d, $J(C,P) = 4.4$ Hz, CH₃-*i*-Pr), 23.88 (d, $J(C,P) = 4.4$ Hz, CH₃-*i*-Pr); C-1 and C-2 not detected. ³¹P {¹H} NMR (DMSO-*d*₆): δ 20.11. FTIR (CCl₄, cm⁻¹) ν_{max} : 3423, 3225, 3063, 3035, 2981, 2934, 1722, 1687, 1678, 1603, 1578, 1541, 1491, 1468, 1449, 1386, 1375, 1258, 1238, 1179, 1159, 1143, 1105, 1087, 1079, 1026, 1009, 993, 881, 706, 696, 648, 633.

5.2.20.5. Diisopropyl (S)-(6,10-dioxo-8-thioxo-3-(trityloxymethyl)-2,11-dioxo-5,7,9-triazatridecyl)phosphonate (30). Prepared from 25 (1.00 g; 1.80 mmol) and ethoxycarbonyl isothiocyanate (0.355 g; 2.70 mmol) in acetone (80 mL); reaction time: 4 h. Yield: 0.76 g (62%) of a yellowish crystalline solid, mp 61–62 °C. $[\alpha]_D - 0.7$ (c 0.430 g/100 mL, MeOH). MS (ESI) *m/z* (%): 708 (100, M+Na⁺). HRMS (ESI) calcd. for C₃₄H₄₄O₈N₃NaPS (M+Na⁺): 708.2479, found: 708.2477. ¹H NMR (DMSO-*d*₆): δ 12.18 (bs, 1H, NH), 11.04 (bs, 1H, NH), 8.24 (bs, 1H, NH), 7.37–7.41 (m, 6H, H-o-Tr), 7.30–7.35 (m, 6H, H-m-Tr), 7.26 (m, 3H, H-p-Tr), 4.52–4.63 (m, 2H, CH-*i*-Pr), 4.18 (q, 2H, $J(4,5) = 7.1$ Hz, H-4), 3.83 (dd, 1H, $J_{\text{gem}} = 13.7$ Hz, $J(3'a,P) = 8.8$ Hz, H-3'a), 3.78 (dd, 1H, $J_{\text{gem}} = 13.7$ Hz, $J(3'b,P) = 9.3$ Hz, H-3'b), 3.72 (p, 1H, $J(2',1') = J(2',4') = 5.0$, H-2'), 3.35–3.48 (bm, 2H, H-1'), 3.08 (dd, 1H, $J_{\text{gem}} = 10.4$ Hz, $J(4'a,2') = 4.8$ Hz, H-4'a), 3.05 (dd, 1H, $J_{\text{gem}} = 10.4$ Hz, $J(4'b,2') = 5.3$ Hz, H-4'b), 1.16–1.26 (m, 15H, CH₃-*i*-Pr, H-5). ¹³C NMR (DMSO-*d*₆): δ 178.86 (C-2), 153.13 (C-1), 152.03 (C-3), 143.72 (C-*i*-Tr), 128.43 (CH-*o*-Tr), 128.08 (CH-*m*-Tr), 127.26 (CH-*p*-Tr), 86.33 (C-Tr), 78.84 (d, $J(2',P) = 12.2$ Hz, CH-2'), 70.38 (d, $J(C,P) = 6.3$ Hz, CH-*i*-Pr), 64.01 (d, $J(3',P) = 165.1$ Hz, CH₂-3'), 62.90 (CH₂-4'), 62.40 (CH₂-4), 39.70 (CH₂-1'), 24.00 (d, $J(C,P) = 3.7$ Hz, CH₃-*i*-Pr), 23.90 (d, $J(C,P) = 4.4$ Hz, CH₃-*i*-Pr), 23.87 (d, $J(C,P) = 4.4$ Hz, CH₃-*i*-Pr), 14.19 (CH₃-5). ³¹P {¹H} NMR (DMSO-*d*₆): δ 18.91. FTIR (CCl₄, cm⁻¹) ν_{max} : 3413, 3255, 3201, 3088, 3062, 3035, 3024, 2981, 2934, 2874, 1777, 1742, 1717, 1707, 1597, 1535, 1503, 1478, 1467, 1449, 1442, 1386, 1374, 1304, 1223, 1174, 1170, 1142, 1121, 1105, 1096, 1077, 1035, 1011, 993, 936, 890, 706, 696, 642, 633.

5.2.20.6. Diisopropyl (S)-(1-(9H-fluoren-9-yl)-3,7-dioxo-5-thioxo-10-(trityloxymethyl)-2,11-dioxo-4,6,8-triazadodecan-12-yl)phosphonate (31). Prepared from 25 (1.50 g; 2.71 mmol) and 9-fluorenylmethoxycarbonyl isothiocyanate (0.92 g; 3.25 mmol) in acetone (100 mL); reaction time: 18 h. Yield: 1.14 g (51%) of a white crystalline solid, mp 89–90 °C. $[\alpha]_D - 2.8$ (c 0.109 g/100 mL, MeOH). MS (ESI) *m/z* (%): 859 (100, M+Na⁺). HRMS (ESI) calcd. for C₄₆H₅₀O₈N₃NaPS (M+Na⁺): 858.2948, found: 858.2949; HRMS (ESI) calcd. for C₄₆H₅₁O₈N₃PS (M+H⁺): 836.3129, found: 836.3134. ¹H NMR (DMSO-*d*₆): δ 12.36 (bs, 1H, NH), 11.11 (bs, 1H, NH), 8.48 (bs, 1H, NH), 7.90 (dt, 2H, $J(10a,9a\&10b,9b) = 7.5$ Hz, $J(10a,8a\&10b,8b) = J(10a,7a\&10b,7b) = 0.9$ Hz, H-10a,b), 7.76 (bd, 2H, $J(7a,8a\&7b,8b) = 7.5$ Hz, H-7a,b), 7.43 (bt, 2H, $J(9a,10a\&9b,10b) = J(9a,8a\&9b,8b) = 7.5$ Hz, H-9a,b), 7.39 (m, 6H, H-o-Tr), 7.30–7.35 (m, 8H, H-8a,b, H-m-Tr), 7.25 (m, 3H, H-p-Tr), 4.53–4.63 (m, 2H, CH-*i*-Pr), 4.47 (d, 2H, $J(4,5) = 7.1$ Hz, H-4), 4.31 (t, 1H, $J(5,4) = 7.1$ Hz, H-5), 3.83 (dd, 1H, $J_{\text{gem}} = 13.7$ Hz, $J(3'a,P) = 8.8$ Hz, H-3'a), 3.79 (dd, 1H, $J_{\text{gem}} = 13.7$ Hz, $J(3'b,P) = 9.3$ Hz, H-3'b),

3.74 (p, 1H, $J(2',1') = J(2',3') = 5.2$ Hz, H-2'), 3.39–3.51 (bm, 2H, H-1'), 3.08 (dd, 1H, $J_{\text{gem}} = 10.8$ Hz, $J(4'a,2') = 5.2$ Hz, H-4'a), 3.06 (dd, 1H, $J_{\text{gem}} = 10.8$ Hz, $J(4'b,2') = 5.2$ Hz, H-4'b), 1.22 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.21 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.19 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.18 (d, 3H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 178.90 (C-2), 152.94 (C-1), 152.18 (C-3), 143.74 (C-*i*-Tr), 143.43 (C-6a,b), 140.96 (C-11a,b), 128.46 (CH-*o*-Tr), 128.12 (CH-*m*-Tr), 128.08 (CH-9a,b), 127.39 (CH-8a,b), 127.29 (CH-*p*-Tr), 125.54 (CH-7a,b), 120.45 (CH-10a,b), 86.34 (C-Tr), 78.80 (d, $J(2',P) = 12.2$ Hz, CH-2'), 70.42 (d, $J(C,P) = 6.3$ Hz, CH-*i*-Pr), 67.70 (CH₂-4), 64.00 (d, $J(3',P) = 165.1$ Hz, CH₂-3'), 62.82 (CH₂-4'), 46.23 (CH-5), 39.70 (CH₂-1'), 24.05 (d, $J(C,P) = 3.7$ Hz, CH₃-*i*-Pr), 23.95 (d, $J(C,P) = 4.2$ Hz, CH₃-*i*-Pr), 23.92 (d, $J(C,P) = 4.2$ Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 20.11. FTIR (CHCl₃, cm⁻¹) ν_{max} : 3395, 3263, 3208, 3088, 3066, 2985, 1768, 1742, 1712, 1607, 1598, 1538, 1505, 1478, 1450, 1387, 1377, 1305, 1249, 1172, 1104, 1056, 999, 890, 707, 642, 545, 427.

5.2.21. N⁵-Protected thiadiazolones with N²-[(disopropoxy)phosphoryl]methoxyalkyl arrangement. The general procedure

An appropriate starting compound (26, 27, 29, 30 or 31) and *N*-bromosuccinimide (1.25 eq.) were dissolved in MeOH and stirred at 80 °C for 30 min. After cooling to RT, the solvent was evaporated and the residue was chromatographed on a silica gel column (0–3% MeOH/CHCl₃).

5.2.21.1. Diisopropyl ((2-(5-benzamido-3-oxo-1,2,4-thiadiazol-2(3H)-yl)ethoxy)methyl)phosphonate (32). Prepared from 26 (7.78 g; 17.5 mmol) and *N*-bromosuccinimide (3.90 g; 21.9 mmol) in MeOH (200 mL). Yield: 3.74 g (48%) of a white crystalline solid, mp 83–84 °C. MS (ESI) *m/z* (%): 466 (100, M+Na⁺), 444 (9, M+H⁺). HRMS (ESI) calcd. for C₁₈H₂₆O₆N₃NaPS (M+Na⁺): 466.1172, found: 466.1171; HRMS (ESI) calcd. for C₁₈H₂₇O₆N₃PS (M+H⁺): 444.1353, found: 444.1352. ¹H NMR (DMSO-*d*₆): δ 13.42 (vbs, 1H, NH), 8.12 (m, 2H, H-o-Ph), 7.66 (m, 1H, H-p-Ph), 7.56 (m, 2H, H-m-Ph), 4.59 (dsept, 2H, $J(H,P) = 7.7$ Hz, $J_{\text{vic}} = 6.2$ Hz, CH-*i*-Pr), 3.80 (d, 2H, $J(3',P) = 8.2$ Hz, H-3'), 3.71–3.79 (m, 4H, H-1', H-2'), 1.23 (bt, 12H, $J_{\text{vic}} = 5.9$ Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 174.52 (C-7, determined by HMBC), 170.86 (C-5, determined by HMBC), 161.53 (C-3), 133.57 (CH-*p*-Ph), 130.56 (C-*i*-Ph), 129.04 (CH-*m*-Ph), 128.94 (CH-*o*-Ph), 70.74 (d, $J(2',P) = 11.5$ Hz, CH₂-2'), 70.40 (d, $J(C,P) = 6.4$ Hz, CH-*i*-Pr), 64.98 (d, $J(3',P) = 163.9$ Hz, CH₂-3'), 42.47 (CH₂-1'), 23.98 (d, $J(C,P) = 3.8$ Hz, CH₃-*i*-Pr), 23.88 (d, $J(C,P) = 4.5$ Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 19.01. FTIR (CCl₄, cm⁻¹) ν_{max} : 3503, 3430, 3123, 3068, 3032, 2981, 2930, 2877, 1714, 1706, 1684, 1676, 1653, 1601, 1570, 1559, 1488, 1468, 1449, 1387, 1375, 1336, 1302, 1231, 1177, 1143, 1106, 1069, 1017, 999, 890, 687, 639, 529.

5.2.21.2. Ethyl (2-(2-((disopropoxyphosphoryl)methoxy)ethyl)-3-oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)carbamate (33). Prepared from 27 (10.20 g; 24.7 mmol) and *N*-bromosuccinimide (5.50 g; 30.9 mmol) in MeOH (200 mL). Yield: 2.42 g (19%) of a white crystalline solid, mp 84–86 °C. MS (ESI) *m/z* (%): 434 (100, M+Na⁺), 412 (20, M+H⁺). HRMS (ESI) calcd. for C₁₄H₂₆O₇N₃NaPS (M+Na⁺): 434.1121, found: 434.1121; HRMS (ESI) calcd. for C₁₄H₂₇O₇N₃PS (M+H⁺): 412.1302, found: 412.1302. ¹H NMR (DMSO-*d*₆): δ 12.79 (bs, 1H, NH), 4.59 (dsept, 2H, $J(H,P) = 7.7$ Hz, $J_{\text{vic}} = 6.2$ Hz, CH-*i*-Pr), 4.23 (q, 2H, $J(9,10) = 7.1$ Hz, CH₂-9), 3.77 (d, 2H, $J(3',P) = 8.2$ Hz, H-3'), 3.76 (m, 2H, H-1'), 3.69 (m, 2H, H-2'), 1.25 (t, 3H, $J(10,9) = 7.1$ Hz, H-10), 1.23 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr), 1.21 (d, 6H, $J_{\text{vic}} = 6.2$ Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 170.54 (C-7), 162.95 (C-3), 157.49 (C-5), 71.94 (d, $J(C,P) = 6.4$ Hz, CH-*i*-Pr), 71.76 (d, $J(2',P) = 11.2$ Hz, CH₂-2'), 65.66 (d, $J(3',P) = 164.7$ Hz, CH₂-3'), 64.53 (CH₂-9), 43.70 (CH₂-1'), 24.85 (d, $J(C,P) = 3.8$ Hz, CH₃-*i*-Pr), 24.75 (d, $J(C,P) = 4.6$ Hz, CH₃-*i*-Pr), 15.23 (CH₃-10). ³¹P {¹H} NMR (DMSO-*d*₆): δ 19.02. FTIR (CCl₄, cm⁻¹) ν_{max} : 3426, 3145, 3076, 2982, 2933, 2876, 1716, 1689, 1635, 1576, 1467, 1386, 1375, 1320, 1306, 1252, 1233, 1178, 1144, 1106, 1014, 996, 890.

5.2.21.3. **Diisopropyl (S)-(((1-(5-benzamido-3-oxo-1,2,4-thiadiazol-2(3H)-yl)-3-hydroxypropan-2-yl)oxy)methyl)phosphonate (35)**. Prepared from **29** (0.58 g; 0.802 mmol) and *N*-bromosuccinimide (0.18 g; 1.01 mmol) in MeOH (40 mL). Yield: 0.34 g (89%) of a white amorphous solid. $[\alpha]_D - 4.2$ (c 0.169 g/100 mL, MeOH). MS (ESI) m/z (%): 969 (50, 2M+Na⁺), 496 (100, M+Na⁺), 474 (44, M+H⁺). HRMS (ESI) calcd. for C₁₉H₂₈O₇N₃NaPS (M+Na⁺): 496.1278, found: 496.1275; HRMS (ESI) calcd. for C₁₉H₂₉O₇N₃PS (M+H⁺): 474.1458, found: 474.1458. ¹H NMR (DMSO-*d*₆): δ 13.40 (bs, 1H, NH), 8.10–8.14 (m, 2H, H-*o*-Ph), 7.66 (m, 1H, H-*p*-Ph), 7.54–7.58 (m, 2H, H-*m*-Ph), 4.86 (t, 1H, J(OH,4') = 5.3 Hz, OH), 4.54–4.65 (m, 2H, CH-*i*-Pr), 3.92 (dd, 1H, J_{gem} = 14.0 Hz, J(3'a,P) = 8.7 Hz, H-3'a), 3.84 (dd, 1H, J_{gem} = 14.0 Hz, J(3'b,P) = 8.7 Hz, H-3'b), 3.81 (dd, 1H, J_{gem} = 14.5 Hz, J(1'a,2') = 3.1 Hz, H-1'a), 3.67 (m, 1H, H-2'), 3.61 (dd, 1H, J_{gem} = 14.5 Hz, J(1'b,2') = 7.1 Hz, H-1'b), 3.45–3.54 (m, 2H, H-3'), 1.23 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.22 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.21 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.20 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 133.46 (CH-*p*-Ph), 129.01 (CH-*m*-Ph), 128.91 (CH-*o*-Ph), 80.53 (d, J(2',P) = 10.7 Hz, CH-2'), 70.41 (d, J(C,P) = 6.4 Hz, CH-*i*-Pr), 70.34 (d, J(C,P) = 6.4 Hz, CH-*i*-Pr), 63.86 (d, J(3',P) = 164.7 Hz, CH₂-3'), 60.22 (CH₂-4'), 43.75 (CH₂-1'), 23.97 (d, J(C,P) = 3.5 Hz, CH₃-*i*-Pr), 23.94 (d, J(C,P) = 3.5 Hz, CH₃-*i*-Pr), 23.88 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr), 23.85 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr); C-3, C-5, C-7, and C-*i*-Ph not detected. ³¹P {¹H} NMR (DMSO-*d*₆): δ 20.38. FTIR (KBr, cm⁻¹) ν_{max}: 3424, 3253, 2981, 1703, 1648, 1600, 1579, 1569, 1555, 1487, 1465, 1449, 1387, 1376, 1257, 1240, 1143, 1104, 1081, 1012, 1001, 994, 893, 690.

5.2.21.4. **Ethyl (S)-((2-((diisopropoxyphosphoryl)methoxy)-3-hydroxypropyl)-3-oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)carbamate (36)**. Prepared from **30** (0.70 g; 1.02 mmol) and *N*-bromosuccinimide (0.23 g; 1.28 mmol) in MeOH (50 mL). Yield: 0.34 g (75%) of a white amorphous solid. $[\alpha]_D - 5.6$ (c 0.197 g/100 mL, MeOH). MS (ESI) m/z (%): 464 (100, M+Na⁺), 442 (18, M+H⁺). HRMS (ESI) calcd. for C₁₅H₂₈O₈N₃NaPS (M+Na⁺): 464.1227, found: 464.1226. ¹H NMR (CD₃OD): δ 4.69–4.78 (m, 2H, CH-*i*-Pr), 4.33 (q, 2H, J(9,10) = 7.1 Hz, H-9), 4.02 (dd, 1H, J_{gem} = 13.9 Hz, J(3'a,P) = 9.0 Hz, H-3'a), 3.96 (dd, 1H, J_{gem} = 14.9 Hz, J(1'a,2') = 3.4 Hz, H-1'a), 3.93 (dd, 1H, J_{gem} = 13.9 Hz, J(3'b,P) = 9.3 Hz, H-3'b), 3.82 (dd, 1H, J_{gem} = 14.9 Hz, J(1'b,2') = 6.8 Hz, H-1'b), 3.73 (m, 1H, H-2'), 3.68 (dd, 1H, J_{gem} = 12.1 Hz, J(4'a,2') = 4.7 Hz, H-4'a), 3.62 (dd, 1H, J_{gem} = 12.1 Hz, J(4'b,2') = 4.9 Hz, H-4'b), 1.31–1.37 (m, 15H, CH₃-*i*-Pr, H-10). ¹³C NMR (CD₃OD): δ 172.10 (C-5), 166.20 (C-3), 157.19 (C-7), 82.50 (d, J(2',P) = 11.6 Hz, CH₂-2'), 73.34 (d, J(C,P) = 6.8 Hz, CH-*i*-Pr), 73.25 (d, J(C,P) = 6.8 Hz, CH-*i*-Pr), 65.28 (d, J(3',P) = 168.6 Hz, CH₂-3'), 64.84 (CH₂-9), 61.82 (CH₂-4'), 45.40 (CH₂-1'), 24.33 (d, J(C,P) = 3.6 Hz, CH₃-*i*-Pr), 24.32 (d, J(C,P) = 3.6 Hz, CH₃-*i*-Pr), 24.29 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr), 24.28 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr), 14.55 (CH₃-10). ³¹P {¹H} NMR (CD₃OD): δ 19.76. FTIR (CCL₄, cm⁻¹) ν_{max}: 3385, 2981, 2934, 2875, 2737, 1715, 1670, 1635, 1570, 1467, 1387, 1375, 1245, 1178, 1143, 1105, 1062, 1012, 996, 890, 690.

5.2.21.5. **(9H-Fluoren-9-yl)methyl (S)-((2-((diisopropoxyphosphoryl)methoxy)-3-hydroxypropyl)-3-oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)carbamate (37)**. Prepared from **31** (1.04 g; 1.25 mmol) and *N*-bromosuccinimide (0.28 g; 1.56 mmol) in MeOH (60 mL). Yield: 0.53 g (72%) of a white crystalline solid, mp 68–70 °C. $[\alpha]_D - 2.8$ (c 0.212 g/100 mL, MeOH). MS (ESI) m/z (%): 1181 (8, 2M-H⁺), 590 (100, M-H⁺). HRMS (ESI) calcd. for C₂₇H₃₃O₈N₃PS (M-H⁺): 590.1732, found: 590.1729. ¹H NMR (DMSO-*d*₆): δ 7.91 (ddd, 2H, J(15a,14a&15b,14b) = 7.5 Hz, J(15a,13a&15b,13b) = 1.2 Hz, J(15a,12a&15b,12b) = 0.7 Hz, H-15a,b), 7.71 (bd, 2H, J(12a,13a&12b,13b) = 7.5 Hz, H-12a,b), 7.43 (tdd, 2H, J(14a,13a&14b,13b) = J(14a,15a&14b,15b) = 7.5 Hz, J(14a,12a&14b,12b) = 1.2 Hz, J(14a,11a&14b,11b) = 0.7 Hz, H-14a,b), 7.34 (td, 2H, J(13a,12a&13b,12b) = J(13a,14a&13b,14b) = 7.5 Hz, J(13a,12a&13b,12b) = 1.2 Hz, H-13a,b), 4.84 (t, 1H, J(OH,4') = 5.4 Hz, OH), 4.52–4.63 (m, 4H, CH₂-9, CH-*i*-Pr), 4.34 (t, 1H, J(10,9) = 6.8 Hz,

H-10), 3.89 (dd, 1H, J_{gem} = 14.0 Hz, J(3'a,P) = 8.8 Hz, H-3'a), 3.78–3.85 (m, 2H, H-1'a, H-3'b), 3.58–3.65 (m, 2H, H-2', H-1'b), 3.40–3.52 (m, 2H, H-4'), 1.219 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.214 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.208 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.199 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 143.41 (C-11a,b), 140.99 (C-16a,b), 128.03 (CH-14a,b), 127.37 (CH-13a,b), 125.34 (CH-12a,b), 120.44 (CH-15a,b), 80.64 (d, J(2',P) = 10.5 Hz, CH₂-2'), 70.41 (d, J(C,P) = 6.4 Hz, CH-*i*-Pr), 70.34 (d, J(C,P) = 6.4 Hz, CH-*i*-Pr), 68.09 (CH₂-9), 63.85 (d, J(3',P) = 164.8 Hz, CH₂-3'), 60.16 (CH₂-4'), 46.36 (CH-10), 43.99 (CH₂-1'), 23.97 (d, J(C,P) = 3.5 Hz, CH₃-*i*-Pr), 23.94 (d, J(C,P) = 3.5 Hz, CH₃-*i*-Pr), 23.87 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr), 23.85 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr); C-3, C-5, and C-7 not detected. ³¹P {¹H} NMR (DMSO-*d*₆): δ 20.34. FTIR (CCL₄, cm⁻¹) ν_{max}: 3371, 3068, 3042, 3018, 2981, 2931, 2874, 2856, 2742, 1717, 1668, 1571, 1478, 1466, 1452, 1387, 1375, 1341, 1294, 1243, 1179, 1143, 1105, 1015, 997, 891, 729, 689, 621, 428.

5.2.22. **(9H-Fluoren-9-yl)methyl (2-((diisopropoxyphosphoryl)methoxy)ethyl)-3-oxo-2,3-dihydro-1,2,4-thiadiazol-5-yl)carbamate (34)**. 0.5 M bromine solution in ethyl acetate (19.0 mL; 9.5 mmol) was added at 0 °C to a solution of **28** (2.54 g; 4.51 mmol) in ethyl acetate (150 mL). The mixture was stirred at 4 °C for 17 h and evaporated. The residue was purified on a silica gel column (a gradient of 0–5% MeOH/CHCl₃). Yield: 2.03 g (80%) of a colorless oil. MS (ESI) m/z (%): 584 (100, M+Na⁺), 562 (7, M+H⁺). HRMS (ESI) calcd. for C₂₆H₃₂O₇N₃NaPS (M+Na⁺): 584.1591, found: 584.1592. ¹H NMR (DMSO-*d*₆): δ 7.90 (bd, 2H, J(15a,14a&15b,14b) = 7.5 Hz, H-15a,b), 7.71 (bd, 2H, J(12a,13a&12b,13b) = 7.5 Hz, H-12a,b), 7.43 (bt, 2H, J(14a,13a&14b,13b) = J(14a,15a&14b,15b) = 7.5 Hz, H-14a,b), 7.34 (td, 2H, J(13a,12a&13b,12b) = J(13a,14a&13b,14b) = 7.5 Hz, H-13a,b), 4.57 (dsep, 2H, J(H,P) = 7.7 Hz, J_{vic} = 6.2 Hz, CH-*i*-Pr), 4.56 (d, 2H, J(9,10) = 6.8 Hz, H-9), 4.33 (bt, 1H, J(10,9) = 6.8 Hz, H-10), 3.77 (d, 2H, J(3',P) = 8.1 Hz, H-3'), 3.76 (bt, 2H, J(1',2') = 4.8 Hz, H-1'), 3.69 (bt, 2H, J(2',1') = 4.8 Hz, H-2'), 1.21 (d, 6H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.20 (d, 6H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 169.90 (C-5), 160.40 (C-3), 157.55 (C-7), 143.39 (C-11a,b), 140.99 (C-16a,b), 128.03 (CH-14a,b), 127.37 (CH-13a,b), 125.34 (CH-12a,b), 120.44 (CH-15a,b), 70.84 (d, J(2',P) = 11.2 Hz, CH₂-2'), 70.39 (d, J(C,P) = 6.4 Hz, CH-*i*-Pr), 68.12 (CH-10), 64.94 (d, J(3',P) = 164.0 Hz, CH₂-3'), 46.35 (CH-10), 42.70 (CH₂-1'), 23.97 (d, J(C,P) = 3.7 Hz, CH₃-*i*-Pr), 23.86 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr). ³¹P NMR {¹H} (DMSO-*d*₆): δ 20.04. FTIR (KBr, cm⁻¹) ν_{max}: 3435, 3250, 2980, 2927, 1716, 1636, 1568, 1478, 1465, 1451, 1232, 1254, 1180, 1145, 1105, 1003, 761, 741, 621.

5.2.23. **N⁵-Substituted N²-[2-(phosphonomethoxy)ethyl]thiadiazolones (38–40)**. *The general procedure*

Bromotrimethylsilane (2.7 mL; 20 mmol) was added to the appropriate phosphonate ester **32–34** (1 mmol) dissolved in acetonitrile (30 mL) and the mixture was left at RT overnight. The volatiles were removed *in vacuo* and the residue was co-evaporated three times with acetonitrile, then with 25% aqueous ammonia and finally three times with ethanol. The product was recrystallized from ethanol.

5.2.23.1. **((2-(5-Benzamido-3-oxo-1,2,4-thiadiazol-2(3H)-yl)ethoxy)methyl)phosphonic acid (38)**. Prepared from the starting compound **32** (0.44 g; 1 mmol). Yield: 0.34 g (96%) of a white crystalline solid, mp 219–221 °C. MS (ESI) m/z (%): 380 (13, M-2H⁺+Na⁺), 358 (100, M-H⁺). HRMS (ESI) calcd. for C₁₂H₁₃O₆N₃PS (M-H⁺): 358.0268, found: 358.0268. ¹H NMR (DMSO-*d*₆): δ 8.09 (m, 2H, H-*o*-Ph), 7.40–7.48 (m, 3H, H-*m*-Ph, H-*p*-Ph), 6.97 (vbs, NH⁺), 3.68 (m, 2H, H-1'), 3.60 (m, 2H, H-2'), 3.36 (d, 2H, J(H,P) = 8.8 Hz, H-3'). ¹³C NMR (DMSO-*d*₆): δ 176.89 (C-5), 175.28 (C-7), 166.16 (C-3), 136.88 (C-*i*-Pr), 130.96 (CH-*p*-Ph), 128.42 (CH-*o*-Ph), 128.22 (CH-*m*-Ph), 71.13 (d, J(C,P) = 10.7 Hz, CH₂-2'), 69.17 (d, J(C,P) = 15.4 Hz, CH₂-3'), 42.10 (CH₂-1'). ³¹P {¹H} NMR (DMSO-*d*₆): δ 12.45. FTIR (KBr, cm⁻¹) ν_{max}: 3127,

3010, 2930, 1669, 1655, 1617, 1603, 1582, 1564, 1536, 1496, 1448, 1399, 1334, 1302, 1145, 1113, 1070, 1015, 995, 952, 811, 772, 704, 687, 641, 529.

5.2.23.2. ((2-(5-((Ethoxycarbonyl)amino)-3-oxo-1,2,4-thiadiazol-2(3H)-yl)ethoxy)methyl)phosphonic acid (39). Prepared from the starting compound **33** (0.41 g; 1 mmol). Yield: 0.30 g (91%) of a white crystalline solid, mp 184–186 °C. MS (ESI) m/z (%): 348 (26, M–2H⁺+Na⁺), 326 (100, M–H⁺). HRMS (ESI) calcd. for C₈H₁₃O₇N₃PS (M–H⁺): 326.0217, found: 326.0213. ¹H NMR (D₂O): δ 4.23 (q, 2H, J(9,10) = 7.1 Hz, H-9), 3.88 (m, 2H, H-1'), 3.82 (m, 2H, H-2'), 3.62 (d, 2H, J(3',P) = 8.5 Hz, H-3'), 1.29 (t, 3H, J(10,9) = 7.1 Hz, H-10). ¹³C NMR (D₂O): δ 177.80 (C-5), 166.86 (C-3), 163.26 (C-7), 71.39 (d, J(2',P) = 10.4 Hz, CH₂-2'), 68.32 (d, J(3',P) = 154.2 Hz, CH₂-3'), 63.94 (CH₂-9), 43.70 (CH₂-1'), 14.41 (CH₃-10). ³¹P NMR (D₂O): δ 15.56. FTIR (KBr, cm⁻¹) ν_{max} : 3138, 2978, 2934, 2884, 1706, 1665, 1500, 1483, 1403, 1339, 1312, 1218, 1107, 1027, 1011, 956, 937.

5.2.23.3. ((2-(5-(((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-oxo-1,2,4-thiadiazol-2(3H)-yl)ethoxy)methyl)phosphonic acid (40). Prepared from the starting compound **34** (0.56 g; 1 mmol). Yield: 0.39 g (82%) of a yellowish amorphous solid. MS (ESI) m/z (%): 476 (43, M–H⁺). HRMS (ESI) calcd. for C₂₀H₁₉O₇N₃PS (M–H⁺): 476.0684, found: 476.0684. ¹H NMR (MeOD-*d*₄): δ 7.78 (dt, 2H, J(15a,14a&15b,14b) = 7.5 Hz, J(15a,13a&15b,13b) = J(15a,12a&15b,12b) = 1.1 Hz, H-15a,b), 7.64 (dq, 2H, J(12a,13a&12b,13b) = 7.5 Hz, J(12a,14a&12b,14b) = J(12a,15a&12b,15b) = J(12a,10&12b,10) = 1.1 Hz, H-12a,b), 7.38 (tdd, 2H, J(14a,13a&14b,13b) = J(14a,15a&14b,15b) = 7.5 Hz, J(14a,12a&14b,12b) = 1.1 Hz, J(14a,11a&14b,11b) = 0.7 Hz, H-14a,b), 7.31 (td, 2H, J(13a,12a&13b,12b) = J(13a,14a&13b,14b) = 7.5 Hz, J(13a,15a&13b,15b) = 1.1 Hz, H-13a,b), 4.59 (d, 2H, J(9,10) = 6.5 Hz, H-9), 4.27 (t, 1H, J(10,9) = 6.5 Hz, H-10), 3.87 (m, 2H, H-1'), 3.79 (m, 2H, H-2'), 3.76 (d, 2H, J(3',P) = 8.3 Hz, H-3'). ¹³C NMR (MeOD-*d*₄): δ 171.87 (C-5), 165.70 (C-3), 156.99 (C-7), 144.47 (C-11a,b), 142.63 (C-16a,b), 129.04 (CH-14a,b), 128.29 (CH-13a,b), 126.04 (CH-12a,b), 121.08 (CH-15a,b), 72.34 (d, J(2',P) = 10.3 Hz, CH₂-2'), 69.91 (CH₂-9), 67.33 (d, J(3',P) = 162.9 Hz, CH₂-3'), 47.90 (CH-10), 44.44 (CH₂-1'). ³¹P NMR (MeOD-*d*₄): δ 19.15. FTIR (KBr, cm⁻¹) ν_{max} : 3431, 3065, 3039, 3015, 2925, 2720, 2320, 1731, 1713, 1643, 1566, 1478, 1450, 1295, 1256, 1230, 1107, 1033, 1005, 942, 761, 741, 621, 428.

5.2.24. N⁵-Substituted N²-[3-hydroxy-2-(phosphonomethoxy)propyl]thiadiazolones (41–43)

The general procedure. Bromotrimethylsilane (2.7 mL; 20 mmol) was added to a solution of the appropriate phosphonate ester **35–37** (1 mmol) in acetonitrile (30 mL) and the mixture was left at RT overnight. The volatiles were removed *in vacuo* and the residue was co-evaporated three times with acetonitrile. Water (5 mL) was added and the mixture was stirred at RT for 1 h. The mixture was evaporated, the residue was co-evaporated three times with ethanol and the crude product was purified on a reverse-phase (C₁₈) column in a gradient of 0–60% acetonitrile/water. Product-containing fractions were combined and lyophilized.

5.2.24.1. (S)-(((1-(5-Benzamido-3-oxo-1,2,4-thiadiazol-2(3H)-yl)-3-hydroxypropan-2-yl)oxy)methyl)phosphonic acid (41). Prepared from the starting compound **35** (0.47 g; 1 mmol). Yield: 0.31 g (80%) as a white lyophilizate. [α]_D – 1.2 (c 0.332 g/100 mL, H₂O). MS (ESI) m/z (%): 410 (32, M–2H⁺+Na⁺), 388 (M–H⁺). HRMS (ESI) calcd. for C₁₃H₁₅O₇N₃PS (M–H⁺): 388.0374, found: 388.0370. ¹H NMR (MeOD-*d*₄): δ 8.10–8.14 (m, 2H, H-*o*-Ph), 7.66 (m, 1H, H-*p*-Ph), 7.52–7.59 (m, 2H, H-*m*-Ph), 3.90 (dd, 1H, J_{gem} = 14.7 Hz, J(1'a,2') = 4.6 Hz, H-1'a), 3.78–3.86 (m, 2H, H-1'b, H-3'a), 3.76 (dd, 1H, J_{gem} = 12.2 Hz, J(4'a,2') = 3.6 Hz, H-4'a), 3.71 (m, 1H, H-2'), 3.67 (dd, 1H, J_{gem} = 12.8 Hz, J(3'b,P) = 9.9 Hz, H-3'b), 3.58 (dd, 1H, J_{gem} = 12.2 Hz, J(4'b,2') = 4.8 Hz, H-4'b). ¹³C NMR

(MeOD-*d*₄): δ 173.57 (C-7), 171.68 (C-5), 162.56 (C-3), 134.74 (CH-*p*-Ph), 132.54 (C-*i*-Ph), 129.94 (CH-*m*-Ph), 129.82 (CH-*o*-Ph), 82.41 (d, J(2',P) = 10.4 Hz, CH-2'), 68.28 (d, J(3',P) = 157.6 Hz, CH₂-3'), 62.21 (CH₂-4'), 44.96 (CH₂-1'). ³¹P NMR (MeOD-*d*₄): δ 16.50. FTIR (KBr, cm⁻¹) ν_{max} : 3390, 3200, 3136, 2893, 2516, 2270, 1753, 1705, 1687, 1652, 1600, 1589, 1547, 1487, 1472, 1449, 1303, 1261, 1185, 1157, 1107, 1057, 1027, 1013, 838, 709, 686, 632.

5.2.24.2. (S)-(((1-(5-((Ethoxycarbonyl)amino)-3-oxo-1,2,4-thiadiazol-2(3H)-yl)-3-hydroxypropan-2-yl)oxy)methyl)phosphonic acid (42). Prepared from the starting compound **36** (0.44 g; 1 mmol). Yield: 0.26 g (73%) as a white lyophilizate. [α]_D – 2.6 (c 0.272 g/100 mL, H₂O). MS (ESI) m/z (%): 735 (12, 2M–2H⁺+Na⁺), 713 (15, 2M–H⁺), 378 (5, M–2H⁺+Na⁺), 356 (100, M–H⁺). HRMS (ESI) calcd. for C₉H₁₅O₈N₃PS (M–H⁺): 356.0323, found: 356.0322. ¹H NMR (MeOD-*d*₄): δ 4.33 (q, 2H, J(9,10) = 7.1 Hz, H-9), 3.95 (dd, 1H, J_{gem} = 14.8 Hz, J(1'a,2') = 3.9 Hz, H-1'a), 3.78–3.90 (m, 3H, H-1'b, 2H-4'), 3.68–3.74 (m, 2H, H-2', H-3'a), 3.60 (dd, 1H, J_{gem} = 13.5 Hz, J(3'b,2') = 6.3 Hz, H-3'b), 1.35 (t, 3H, J(10,9) = 7.1 Hz, H-10). ¹³C NMR (MeOD-*d*₄): δ 172.09 (C-5), 166.32 (C-3), 156.97 (C-7), 82.50 (d, J(2',P) = 10.7 Hz, CH-2'), 66.81 (d, J(3',P) = 162.1 Hz, CH₂-3'), 64.86 (CH₂-9), 62.02 (CH₂-4'), 45.14 (CH₂-1'), 14.53 (CH₃-10). ³¹P NMR (MeOD-*d*₄): δ 19.45. FTIR (KBr, cm⁻¹) ν_{max} : 3259, 2281, 2137, 1758, 1714, 1689, 1563, 1513, 1370, 1363, 1323, 1276, 1257, 1231, 1034, 1010, 988, 942, 883, 701, 552.

5.2.24.3. (S)-(((1-(5-(((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-oxo-1,2,4-thiadiazol-2(3H)-yl)-3-hydroxypropan-2-yl)oxy)methyl)phosphonic acid (43). Prepared from the starting compound **37** (0.59 g; 1 mmol). Yield: 0.44 g (86%) as a white lyophilizate. [α]_D – 1.8 (c 0.221 g/100 mL, H₂O). MS (ESI) m/z (%): 528 (10, M–2H⁺+Na⁺), 506 (60, M–H⁺). HRMS (ESI) calcd. for C₂₁H₂₀O₈N₃NaPS (M–2H⁺+Na⁺): 528.0612, found: 528.0608; HRMS (ESI) calcd. for C₂₁H₂₁O₈N₃PS (M–H⁺): 506.0793, found: 506.0789. ¹H NMR (MeOD-*d*₄): δ 7.82 (ddd, 2H, J(15a,14a&15b,14b) = 7.6 Hz, J(15a,13a&15b,13b) = 1.3 Hz, J(15a,12a&15b,12b) = 0.7 Hz, H-15a,b), 7.67 (ddd, 2H, J(12a,13a&12b,13b) = 7.5 Hz, J(12a,14a&12b,14b) = 1.3 Hz, J(12a,13a&12b,13b) = J(12a,10&12b,10) = 0.7 Hz, H-12a,b), 7.41 (tdd, 2H, J(14a,13a&14b,13b) = J(14a,15a&14b,15b) = 7.5 Hz, J(14a,12a&14b,12b) = 1.3 Hz, J(14a,11a&14b,11b) = 0.7 Hz, H-14a,b), 7.33 (td, 2H, J(13a,12a&13b,12b) = J(13a,14a&13b,14b) = 7.5 Hz, J(13a,15a&13b,15b) = 1.3 Hz, H-13a,b), 4.65 (d, 2H, J(9,10) = 6.5 Hz, H-9), 4.32 (t, 1H, J(10,9) = 6.5 Hz, H-10), 3.94 (dd, 1H, J_{gem} = 14.9 Hz, J(1'a,2') = 3.8 Hz, H-1'a), 3.87 (dd, 1H, J_{gem} = 13.4 Hz, J(3'a,P) = 9.1 Hz, H-3'a), 3.83 (dd, 1H, J_{gem} = 13.4 Hz, J(3'b,P) = 9.5 Hz, H-3'b), 3.81 (dd, 1H, J_{gem} = 14.9 Hz, J(1'b,2') = 6.5 Hz, H-1'b), 3.65–3.72 (m, 2H, H-2', H-3'a), 3.60 (m, 1H, H-3'b). ¹³C NMR (MeOD-*d*₄): δ 172.11 (C-5), 166.11 (C-3), 157.00 (C-7), 144.53 (C-11a,b), 142.70 (C-16a,b), 129.07 (CH-14a,b), 128.32 (CH-13a,b), 126.07 (CH-12a,b), 121.10 (CH-15a,b), 82.52 (d, J(2',P) = 11.0 Hz, CH-2'), 69.93 (CH₂-9), 66.49 (d, J(3',P) = 163.0 Hz, CH₂-3'), 61.99 (CH₂-4'), 47.96 (CH-10), 45.16 (CH₂-1'). ³¹P NMR (MeOD-*d*₄): δ 20.15. FTIR (KBr, cm⁻¹) ν_{max} : 3427, 3203, 3069, 3000, 1706, 1626, 1568, 1545, 1496, 1477, 1449, 1247, 1223, 1116, 1057, 1024, 1009, 621, 430.

5.2.25. N⁵-[[Diisopropoxy]phosphoryl]methoxyalkylthioburets (48, 49)

The general procedure. Piperidine (0.15 mL; 1.5 mmol) was added to a solution of Fmoc derivative **28** or **31** (1 mmol) in dichloromethane (40 mL). The mixture was stirred at RT overnight, evaporated, and the residue was chromatographed on a silica gel column (0–3% MeOH/CHCl₃).

5.2.25.1. Diisopropyl ((2-(3-carbamothioylureido)ethoxy)methyl)phosphonate (48). Prepared from starting compound **28** (0.56 g; 1 mmol). Yield: 92% (0.31 g) of white crystals, mp 125–126 °C. MS (ESI) m/z (%):

705 (35, 2M+Na⁺), 364 (100, M+Na⁺), 342 (38, M+H⁺). HRMS (ESI) calcd. for C₁₁H₂₄O₅N₃NaPS (M+Na⁺): 364.1067, found: 364.1068; HRMS (ESI) calcd. for C₁₁H₂₅O₅N₃PS (M+H⁺): 342.1247, found: 342.1248. ¹H NMR (CDCl₃): δ 9.75 (bs, 1H, NH), 7.22 (bs, 1H, NH), 7.12 (bs, 2H, NH₂), 4.76–4.89 (m, 2H, CH-*i*-Pr), 3.79 (d, 2H, J(3',P) = 8.8 Hz, H-3'), 3.66 (m, 2H, H-2'), 3.39 (m, 2H, H-1'), 1.35 (d, 6H, J_{vic} = 6.3 Hz, CH₃-*i*-Pr), 1.33 (d, 6H, J_{vic} = 6.3 Hz, CH₃-*i*-Pr). ¹³C NMR (CDCl₃): δ 182.43 (C-2), 154.35 (C-1), 72.39 (d, J(2',P) = 12.1 Hz, CH₂-2'), 71.99 (d, J(C,P) = 6.8 Hz, CH-*i*-Pr), 66.23 (d, J(3',P) = 170.3 Hz, CH₂-3'), 39.54 (CH₂-1'), 24.07 (d, J(C,P) = 3.8 Hz, CH₃-*i*-Pr), 24.00 (d, J(C,P) = 4.7 Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (CDCl₃): δ 19.98. FTIR (CHCl₃, cm⁻¹) ν_{max}: 3488, 3340, 3296, 3270, 3190, 2984, 2939, 2876, 1704, 1580, 1537, 1466, 1415, 1388, 1377, 1232, 1180, 1143, 1117, 1103, 1018, 1002, 891.

5.2.25.2. Diisopropyl (S)-(((1-(3-carbamothioylureido)-3-(trityloxy)propan-2-yl)oxy)methyl)phosphonate (49). Prepared from the starting compound 31 (0.83 g; 1 mmol). Yield: 0.43 g (69%) of white crystals, mp 77–78 °C. [α]_D – 4.6 (c 0.240 g/100 mL, MeOH). MS (ESI) *m/z* (%): 1250 (14, 2M+Na⁺), 636 (100, M+Na⁺). HRMS (ESI) calcd. for C₃₁H₄₀O₆N₃NaPS (M+Na⁺): 636.2268, found: 636.2269. ¹H NMR (DMSO-*d*₆): δ 9.75 (d, 1H, ⁴J = 1.6 Hz, CSNH), 9.36 (d, 1H, J_{gem} = 3.3 Hz, NH₂), 8.95 (d, 1H, J_{gem} = 3.3 Hz, NH₂), 7.37–7.41 (m, 6H, H-*o*-Tr), 7.31–7.36 (m, 6H, H-*m*-Tr), 7.26 (m, 3H, H-*p*-Tr), 6.82 (bt, 1H, J(NH,1') = 5.3 Hz, CONH), 4.53–4.63 (m, 2H, CH-*i*-Pr), 3.81 (dd, 1H, J_{gem} = 13.7 Hz, J(3'a,P) = 8.7 Hz, H-3'a), 3.76 (dd, 1H, J_{gem} = 13.7 Hz, J(3'b,P) = 9.3 Hz, H-3'b), 3.65 (p, 1H, J(2',1') = J(2',4') = 5.1 Hz, H-2'), 3.23–3.35 (m, 2H, H-1'), 3.05 (dd, 1H, J_{gem} = 10.4 Hz, J(4'a,2') = 4.6 Hz, H-4'a), 3.02 (dd, 1H, J_{gem} = 10.4 Hz, J(4'b,2') = 5.4 Hz, H-4'b), 1.23 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.22 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.20 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.19 (d, 3H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 181.51 (C-2), 154.31 (C-1), 143.77 (C-*i*-Tr), 128.44 (CH-*o*-Tr), 128.09 (CH-*m*-Tr), 127.28 (CH-*p*-Tr), 86.32 (C-*r*-Tr), 79.17 (d, J(2',P) = 12.0 Hz, CH-2'), 70.42 (d, J(C,P) = 6.3 Hz, CH-*i*-Pr), 70.39 (d, J(C,P) = 6.3 Hz, CH-*i*-Pr), 64.00 (d, J(3',P) = 165.1 Hz, CH₂-3'), 63.01 (CH₂-4'), 39.48 (CH₂-1'), 24.02 (d, J(C,P) = 3.6 Hz, CH₃-*i*-Pr), 23.91 (d, J(C,P) = 4.4 Hz, CH₃-*i*-Pr), 23.87 (d, J(C,P) = 4.4 Hz, CH₃-*i*-Pr). ³¹P {¹H} NMR (DMSO-*d*₆): δ 18.99. FTIR (CHCl₃, cm⁻¹) ν_{max}: 3488, 3390, 3340, 3295, 3264, 3088, 3063, 3035, 2984, 2935, 2875, 1705, 1599, 1578, 1519, 1491, 1466, 1449, 1442, 1412, 1388, 1377, 1228, 1180, 1159, 1152, 1143, 1117, 1101, 1090, 1080, 1036, 1018, 1002, 891, 707, 700, 643, 633, 617.

5.2.26. N⁵-[(Phosphonomethoxy)alkyl]thiobiurets (46, 47). The general procedure

Bromotrimethylsilane (2 mL; 15 mmol) was added to the appropriate phosphonate ester 48 or 49 (1 mmol) in acetonitrile (30 mL) and the mixture was left at RT overnight. The volatiles were evaporated and the residue was co-evaporated with acetonitrile. After the addition of water, the mixture was stirred for 1 h and evaporated (for compound 46) or extracted with diethyl ether (for compound 47) in order to remove trityl alcohol) and then evaporated.

5.2.26.1. ((2-(3-Carbamothioylureido)ethoxy)methyl)phosphonic acid (46). Prepared from the starting compound 48 (0.34 g; 1 mmol). Yield: 0.24 g (93%) of a yellowish crystalline solid, mp 139–140 °C. MS (ESI) *m/z* (%): 513 (25, 2M–H⁺), 256 (100, M–H⁺). HRMS (ESI) calcd. for C₅H₁₁O₅N₃PS (M–H⁺): 256.0163, found: 256.0161. ¹H NMR (CD₃OD): δ 3.76 (d, 2H, J(3',P) = 9.0 Hz, H-3'), 3.64–3.67 (m, 2H, H-2'), 3.36–3.39 (m, 2H, H-1'). ¹³C NMR (CD₃OD): δ 183.67 (C-2), 155.73 (C-1), 72.84 (d, J(2',P) = 11.9 Hz, CH₂-2'), 67.08 (d, J(3',P) = 163.7 Hz, CH₂-3'), 40.16 (CH₂-1'). ³¹P {¹H} NMR (CD₃OD): δ 19.57. FTIR (KBr, cm⁻¹) ν_{max}: 3336, 3247, 3199, 3139, 2675, 2258, 1723, 1714, 1626, 1604, 1563, 1535, 1223, 1120, 1102.

5.2.26.2. (S)-(((1-(3-Carbamothioylureido)-3-hydroxypropan-2-yl)oxy)methyl)phosphonic acid (47). Prepared from the starting compound 49 (0.61 g; 1 mmol). Yield: 0.26 g (92%) of a colorless oil. [α]_D – 11.7 (c 0.214 g/100 mL, MeOH). MS (ESI) *m/z* (%): 595 (9, 2M–2H⁺+Na⁺), 308 (7, M–2H⁺+Na⁺), 286 (100, M–H⁺). HRMS (ESI) calcd. for C₆H₁₃O₆N₃PS (M–H⁺): 286.0268, found: 286.0267. ¹H NMR (CD₃OD): δ 3.85–3.91 (m, 2H, CH₂-3'), 3.65 (dd, 1H, J_{gem} = 11.8 Hz, J(4'a,2') = 4.5 Hz, H-4'a), 3.59 (dd, 1H, J_{gem} = 11.8 Hz, J(4'b,2') = 5.3 Hz, H-4'b), 3.55 (m, 1H, H-2'), 3.46 (dd, 1H, J_{gem} = 14.2 Hz, J(1'a,2') = 4.5 Hz, H-1'a), 3.27 (dd, 1H, J_{gem} = 14.2 Hz, J(1'b,2') = 6.1 Hz, H-1'b). ¹³C NMR (CD₃OD): δ 183.74 (C-2), 155.89 (C-1), 82.72 (d, J(2',P) = 11.5 Hz, CH-2'), 66.04 (d, J(3',P) = 164.2 Hz, CH₂-3'), 62.51 (CH₂-4'), 40.64 (CH₂-1'). ³¹P {¹H} NMR (CD₃OD): δ 21.13. FTIR (KBr, cm⁻¹) ν_{max}: 3375, 3280, 2945, 2306, 1730, 1711, 1643, 1599, 1546, 1437, 1222, 1120, 1040, 1002, 943.

5.2.27. Diisopropyl ((2-(5-amino-3-oxo-1,2,4-thiadiazol-2(3H)-yl)ethoxy)methyl)phosphonate (50)

0.5 M Bromine solution in ethyl acetate (2.5 mL; 12.5 mmol) was added to a 4 °C cold solution of 48 (0.20 g; 0.59 mmol) in ethyl acetate (10 mL). The mixture was stirred overnight at 4 °C for 16.5 h (until quantitative conversion). The solvent was removed *in vacuo* and the residue was purified on a preparative TLC plate (5% MeOH/CHCl₃). Yield: 0.18 g (90%) of a colorless oil. MS (ESI) *m/z* (%): 362 (100, M+Na⁺), 340 (M+H⁺). HRMS (ESI) calcd. for C₁₁H₂₂O₅N₃NaPS (M+Na⁺): 362.0910, found: 362.0910. ¹H NMR (DMSO-*d*₆): δ 8.33 (bs, 2H, NH₂), 4.59 (dsep, 2H, J(H,P) = 7.7 Hz, J_{vic} = 6.2 Hz, CH-*i*-Pr), 3.76 (d, 2H, J(3',P) = 8.2 Hz, H-3'), 3.66 (m, 2H, H-1'), 3.64 (m, 2H, H-2'), 1.24 (dd, 6H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr), 1.23 (dd, 6H, J_{vic} = 6.2 Hz, CH₃-*i*-Pr). ¹³C NMR (DMSO-*d*₆): δ 172.14 (C-5), 164.86 (C-3), 71.22 (d, J(2',P) = 11.3 Hz, CH₂-2'), 70.46 (d, J(C,P) = 6.4 Hz, CH-*i*-Pr), 64.95 (d, J(3',P) = 164.1 Hz, CH₂-3'), 43.16 (CH₂-1'), 24.07 (d, J(C,P) = 3.6 Hz, CH₃-*i*-Pr), 23.95 (d, J(C,P) = 4.5 Hz, CH₃-*i*-Pr). ³¹P NMR {¹H} (MeOD-*d*₄): δ 19.08. FTIR (CHCl₃, cm⁻¹) ν_{max}: 3464, 3398, 3274, 3126, 2983, 2877, 1668, 1630, 1608, 1544, 1388, 1377, 1237, 1179, 1143, 1103, 1000.

5.3. Biological assays

5.3.1. Recombinant expression, purification, and activation of CatK

The synthetic gene of human CatK (GenScript) was cloned into the expression vector pPICZα and expressed in the X-33 strain of the methylotrophic yeast *Pichia pastoris*.³⁹ The medium was lyophilized and desalted over a Sephadex G25 column. The recombinant protein was purified over a Mono S HR 5/5 column (GE Healthcare Life Sciences) equilibrated in 50 mM sodium acetate pH 5.5 and eluted using a linear gradient of 0–1 M NaCl. The purified CatK in the zymogen form was concentrated to 1 mg/mL by an Amicon Ultracel-10 k ultrafiltration device (Millipore) and activated to the mature form by autocatalytic processing at acidic pH as described previously.⁴⁰ The activation process was monitored by a CatK activity assay and SDS-PAGE. The concentration of mature CatK was determined by active site titration with E-64 inhibitor.

5.3.2. CatK inhibition assay

Inhibition measurements were performed in a microplate format (100-μL assay volume) at 30 °C. The reaction mixture contained 0.7 nM CatK, 20 μM fluorogenic substrate Z-Gly-Pro-Arg-AMC (Bachem), 0.1 M sodium acetate pH 5.5, 150 mM NaCl, 50 μM dithiothreitol, 0.1% PEG 6000, and 1 mM EDTA. CatK was preincubated with 0–100 μM inhibitor (stock solution in DMSO) at 30 °C for 10 min followed by the addition of substrate. The kinetics of the release of the fluorescent product were monitored continuously for 30 min in an Infinite M200 microplate reader (Tecan) at excitation and emission wavelengths of 360 and 465 nm, respectively. The IC₅₀ values were determined using nonlinear regression (GraFit software; Erithacus Software). The final concentration of DMSO did not exceed 5%. The assay concentration of

dithiothreitol was set as low as possible to maintain sufficient CatK activity and avoid interference with 1,2,4-thiadiazoles. The measurements were performed in triplicate.

5.3.3. GSK-3 inhibition assay

Derivatives of 1,2,4-thiadiazole were screened for their potential inhibition effect toward GSK3 β at the concentration of 10 μ M using the GSK3 β kinase assay plus ADP-Glo[®] assay (Promega) according to the manufacturer's instructions with minor modifications. Briefly, the reaction was performed in 384-well white plates (NUNC) in a total volume of 5 μ L. The reaction contained commercial GSK3 β kinase buffer with 2 ng of GSK3 β kinase recombinant protein, 1 μ g of a GSK3 substrate peptide (GSM) and 10 μ M of the respective inhibitors added to the reaction using the ECHO 550 liquid handler. Following a 15 min pre-incubation, the reaction was started by 25 μ M ATP and incubated for 90 min at RT. The kinase reaction was stopped by adding 5 μ L of the Kinase Glo[®] reagent and incubated at RT for 40 min (ATP depletion). Luminescence was recorded 40 min after the addition of 10 μ L of the kinase detection reagent (ADP to ATP conversion) on a Cytation 3 imaging reader (BioTek). Tideglusib and CHIR99021 were used as reference GSK3 inhibitors, while staurosporin represented a non-specific pan-kinase inhibitor (all 1 μ M).

5.3.4. Cytotoxicity assay

Compound cytotoxicity was evaluated in five cancer cell lines (CCRF-CEM, HepG2, Hela, HL-60, SH-SY5Y), while non-tumor dermal fibroblasts (NHDF) represented normal tissue. All cell lines were from ATCC (Manassas, VA, USA). The cells were maintained in RPMI-1640 or DMEM culture medium containing 10% FBS and 1% GlutaMax without antibiotics. Cells were seeded in 384-well white plates (Thermo Fisher Scientific, Waltham, USA) at the concentrations of 2000–50,000 cells per well and left to rest overnight. The next day, varying concentrations of the test compounds were added, the cells were incubated at 37 °C, in an atmosphere containing 5% CO₂ for 72 h, after which CellTiter-Glo[®] detection reagent (Promega, Madison, USA) was added. The plate was left on a shaker (350 rpm) at RT for 20 min. Luminescence was measured by a multimode plate reader. The signal of the compound-treated cells was compared to the value of untreated control, which was set to 100% viability.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Subvention for Development of Research Organization RVO 61388963, by the MEYS, project Czech National Node to the European Infrastructure for Translational Medicine EATRIS-CZ, grant No. LM2018133 and by the project ChemBioDrug CZ.02.1.01/0.0/0.0/16.019/0000729 from the European Regional Development Fund (OP RDE).

References

- 1 Leung Toung R, Wodzinska J, Li W, et al. *Bioorg Med Chem*. 2003;11:5529.
- 2 Lu J, Wang M, Wang Z, Fu Z, Lu A, Zhang G. *J Enzyme Inhib Med Chem*. 2018;33:890.
- 3 Drake MT, Clarke BL, Onusler MJ, Khosla S. *Endocr Rev*. 2017;38:325.
- 4 Mulkherje K, Chattopadhyay N. *Biochem Pharmacol*. 2016;117:10.
- 5 Qutbell M, Benn A, Flinn N, et al. *Bioorg Med Chem*. 2004;12:5689.
- 6 Ochi Y, Yamada H, Mori H, et al. *Bone*. 2011;49:1351.
- 7 Robichaud J, Bayly C, Oballa R, et al. *Bioorg Med Chem Lett*. 2004;14:4291.
- 8 Rankovic Z, Cai J, Kerr J, et al. *Bioorg Med Chem Lett*. 2010;20:1524.
- 9 Walz A, Ugolnikov A, Chandra S, et al. *Clin Cancer Res*. 2017;23:1891.
- 10 Aguilar Morante D, Morales Garcia JA, Sanz SanCristobal M, Garcia Cabezas MA, Santos A, Perez-Castillo A. *PLoS ONE*. 2010;5:e13879.
- 11 del Ser T, Steinwachs KC, Gertz HJ, et al. *J Alzheimers Dis*. 2013;33:205.
- 12 Lovestone S, Boada M, Dubois B, et al. ARGO investigators. *J Alzheimers Dis*. 2015;45:75.
- 13 Martinez A, Alonso M, Castro A, Pérez C, Moreno FJ. *J Med Chem*. 2002;45:1292.
- 14 <https://clinicaltrials.gov/ct2/show/NCT03692312>.
- 15 Neves VCM, Babb R, Chandrasekaran D, Sharpe PT. *Sci Rep*. 2017;7:39654.
- 16 Bachurin SO, Bovina EV, Ustyugov AA. *Med Chem Res Rev*. 2017;37:1186.
- 17 Perlovich GL, Proshin AN, Volkova TV, Petrova LN, Bachurin SO. *Mol Pharm*. 2012;9:2156.
- 18 Volkova T, Chibunova E, Silyukov O, Proshin A, Terekhova I. *J Mol Liq*. 2017;247:64.
- 19 Párkányi C, Yuan HL, Cho NS, Jaw J-HJ, Woodhouse TE, Aung TL. *J Heterocyclic Chem*. 1989;26:1331.
- 20 Li Y, Geng J, Liu Y, Yu S, Zhao G. *ChemMedChem*. 2013;8:27.
- 21 Pískala A, Vacháková A, Masojdková M, et al. *Pharmazie*. 2004;59:756.
- 22 Encinas A, Castro A, Campillo NE, Eur PJA. *J Org Chem*. 2007;5603.
- 23 Holý A. *Curr Pharm Des*. 2003;9:2567.
- 24 De Clercq E, Holý A. *Nat Rev Drug Discov*. 2005;4:928.
- 25 De Clercq E. *Med Res Rev*. 2013;33:1278.
- 26 Krečmerová M, Holý A, Pískala A, et al. *J Med Chem*. 2007;50:1069.
- 27 Krečmerová M, Holý A, Pohl R, et al. *J Med Chem*. 2007;50:5765.
- 28 Krečmerová M, Dračinský M, Snoeck R, Balzarini J, Pomeisl K, Andrei G. *Bioorg Med Chem*. 2017;25:4637.
- 29 Cai ZR, Jabri SY, Jin H, et al. *PCT Int Appl WO2007014352A2 (20070201)*.
- 30 Chang CD, Waki M, Ahmad M, Meienhofer J, Lindell EO, Haug JD. *Peptide Protein Res*. 1980;15:59.
- 31 Chen C-C, Rajagopal B, Liu XY, Chen KL, Lin F, Lin P-C. *Amino Acids*. 2014;46:367.
- 32 Holý A, Rosenberg I, Dvorakova H. *Collect Czech Chem Commun*. 1989;54:2470.
- 33 Dračinský M, Krečmerová M, Holý A. *Bioorg Med Chem*. 2008;16:6778.
- 34 Krečmerová M, Pohl R, Masojdková M, Balzarini J, Snoeck R, Andrei G. *Bioorg Med Chem*. 2014;22:2896.
- 35 Domínguez JM, et al. *J Biol Chem*. 2012;287:893–990.
- 36 Ra DY, Cho NS, Moon JH, Kang SK. *J Heteroc Chem*. 1998;35:1435.
- 37 Castro A, Encinas A, Gil C, et al. *Bioorg Med Chem*. 2008;16:495.
- 38 Xie A, Yan J, Yue L, et al. *Mo Pharmacol*. 2011;80:965.
- 39 Limerers CJ, et al. *Protein Sci*. 1997;6:919.
- 40 Emmott AA, Mort JS. *Protein Express Purif*. 2013;91:37.

Publikace č. 2

An Activity-Based Probe for Cathepsin K Imaging with Excellent Potency and Selectivity

Lemke C., **Benýšek J.**, Brajtenbach D., Breuer Ch., Jílková A., Horn M., Buša M., Ulrychová L., Illies A., Kubatzky K. F., Bartz U., Mareš M., Gütschow M. (2021) *Journal of Medicinal Chemistry*, 64: 13793–13806. DOI: [org/10.1021/acs.jmedchem.1c01178](https://doi.org/10.1021/acs.jmedchem.1c01178)

An Activity-Based Probe for Cathepsin K Imaging with Excellent Potency and Selectivity

Carina Lemke, Jakub Benýšek, Dominik Brajtenbach, Christian Breuer, Adéla Jílková, Martin Horn, Michal Buša, Lenka Ulrychová, Annika Illies, Katharina F. Kubatzky, Ulrike Bartz, Michael Mareš,* and Michael Gütschow*

Cite This: *J. Med. Chem.* 2021, 64, 13793–13806

Read Online

ACCESS |



Metrics & More

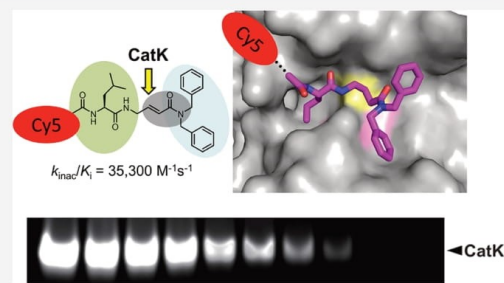


Article Recommendations



Supporting Information

ABSTRACT: The cysteine protease cathepsin K is a target for the treatment of diseases associated with high bone turnover. Cathepsin K is mainly expressed in osteoclasts and responsible for the destruction of the proteinaceous components of the bone matrix. We designed various fluorescent activity-based probes (ABPs) and their precursors that bind to and inactivate cathepsin K. ABP 25 exhibited extraordinary potency ($k_{\text{inac}}/K_i = 35,300 \text{ M}^{-1}\text{s}^{-1}$) and selectivity for human cathepsin K. Crystal structures of cathepsin K in complex with ABP 25 and its nonfluorescent precursor 21 were determined to characterize the binding mode of this new type of acrylamide-based Michael acceptor with the particular orientation of the dibenzylamine moiety to the primed subsite region. The cyanine-5 containing probe 25 allowed for sensitive detection of cathepsin K, selective visualization in complex proteomes, and live cell imaging of a human osteosarcoma cell line, underlining its applicability in a pathophysiological environment.



INTRODUCTION

Human cysteine cathepsins constitute a group of lysosomal proteases that exert a variety of physiological functions and are essential for cellular protein digestion, zymogen activation, and extracellular matrix remodeling. Alterations in their expression, localization, and activity are associated with manifold pathological disorders, rendering them important targets for synthetic bioactive molecules.¹ Cathepsin K is a bone matrix-degrading enzyme, tissue-specifically expressed in osteoclasts. In case of osteoporosis, the activity of these cells is out-of-equilibrium with that of osteoblasts, and the resulting imbalance between bone resorption and formation ultimately leads to deficient bone density and connectivity. Mature osteoclasts excavate characteristic pits and trails on bone, by pumping protons and secreting proteases, particularly cathepsin K, into the resorption lacuna. The acidified milieu is responsible for solubilizing hydroxyapatite.^{2,3} Cathepsin K expression is regulated by the master regulator of the osteoclastogenesis inducer, NFATc1, after RANKL stimulation. RANKL is a type II membrane cytokine belonging to the tumor necrosis factor family.² Cathepsin K is able to degrade the triple-helical and telopeptide regions of type I collagen, the main constituent of the organic bone, as well as other organic components of the extracellular environment, such as elastin, osteopontin, and osteonectin. Its unique collagenolytic activity

is modulated by glycosaminoglycans forming complexes with the enzyme.^{1,4}

Conceptually, selective cathepsin K inhibition would reduce the osteoclasts' resorptive activity without interrupting paracrine signaling to the osteoblasts.⁵ Hence, cathepsin K proved to be a promising therapeutic target for osteoporosis. Moreover, it is implicated in the pathophysiology of the two common forms of arthritis, osteo- and rheumatoid arthritis.² Drug development efforts culminated in the discovery of the cathepsin K inhibitor odanacatib (compound 1 in Figure 1), which efficiently increased bone mineral density in postmenopausal women; however, because of unforeseen adverse effects in clinical trials, the development was discontinued.^{1,5}

Odanacatib is a representative of dipeptide nitriles in which the scissile peptide bond has been replaced by the C-terminal nitrile warhead, susceptible to an attack by the active site cysteine nucleophile, leading to the reversible formation of a covalent thioimidate adduct.⁶ The P2 and P1 amino acid side chains allow for specific binding to the S2 and S1 pockets of

Received: July 1, 2021

Published: September 2, 2021



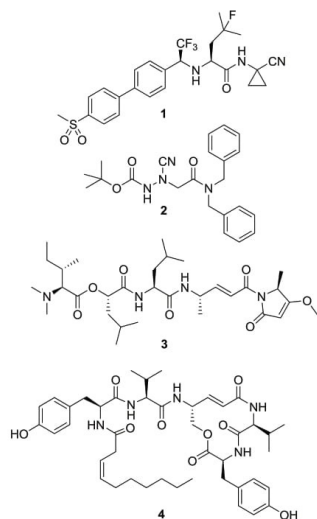


Figure 1. Chemical structures of the small-molecule inhibitors odanacatib (1) and a nitrile-based cathepsin K inhibitor (2), the natural product gallinamide A (3), and the cyclic depsipeptide *N*-desmethyl thalassospiramide C (4).

the active site. An exchange of the α -CH moiety of the P1 amino nitrile by a nitrogen atom led to potent protease-inhibiting azadipeptide nitriles.⁷ In the somewhat modified structure 2 (Figure 1), the cyano group was centrally arranged, and a selective and intriguingly potent cathepsin K inhibitor was achieved.⁸

Tracing the activity of cathepsin K can be realized with activity-based probes (ABPs). In contrast to the above-mentioned, covalent-reversible inhibitors, ABPs have to interact with the target enzyme in an irreversible manner. The inhibitory substructure is directly conjugated to a fluorophore or biotin tag to assemble the ABP ready for target protein labeling, detection, and visualization. Activity-based probing has been particularly successful to analyze cysteine and serine proteases as well as lipases, that is, enzymes which have a covalent mode of catalytic action and can be modified by electrophilic agents.^{9,10} The α,β -unsaturated Michael acceptor motif fulfills these requirements and serves as a standard warhead in ABPs for cysteine proteases. Its carbon in the β position undergoes a nucleophilic addition by the active site cysteine thiolate. Most commonly, vinyl sulfones,^{11–13} vinyl sulfonates,¹⁴ acrylates,¹⁵ or α,β -unsaturated ketones¹⁶ have been implemented in such types of inhibitors and in ABPs derived therefrom.

The acrylamide warhead has not been utilized for the design of ABPs so far. However, this moiety is part of some synthetic inhibitors,¹⁷ as well as remarkable natural compounds and their analogues. Gallinamide A (compound 3 in Figure 1), a marine cyanobacterial metabolite, combines an imide with two acrylamide groups. It acts as an irreversible inhibitor of cathepsin L¹⁰ and was highly potent against parasite cathepsin L-like proteases cruzain and falcipain.¹⁸ Cathepsin L inhibition by gallinamide A was causative for a reduced SARS-CoV-2 infection in vitro.¹⁹ Thalassospiramide C, produced by marine bacteria, contains the acrylamide within the 12-membered

depsipeptide cycle. An X-ray crystal structure of its *N*-desmethyl derivative (compound 4 in Figure 1; PDB ID: 6HGY) in the active site of cathepsin K corroborated the nucleophilic attack of the active-site cysteine on the β position of the Michael acceptor part.²⁰

Inspired by the structures of these inhibitory natural compounds and motivated by the pathophysiological relevance of human cathepsin K, this study was aimed at developing fluorescent, acrylamide-based ABPs. We performed in-depth characterization of the probe–target interaction and demonstrated the excellent applicability of these ABPs for selective imaging of cathepsin K.

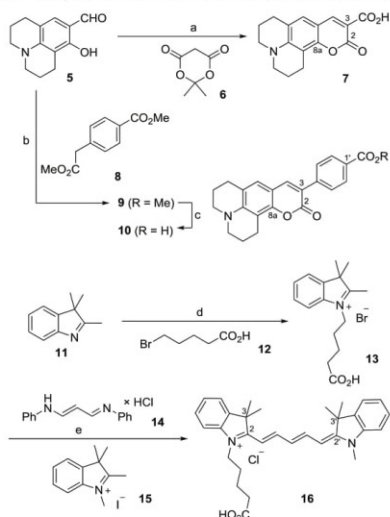
RESULTS AND DISCUSSION

Design. We conceptualized ABPs for cathepsin K resting upon the following considerations. The dibenzylamine portion was applied as the primed site binding element (PSBE) based on the structure of the nitrile inhibitor 2 (Figure 1). An acryloyl group as the electrophilic warhead was considered to enable this structural elongation toward the primed binding site. As in the highly potent cathepsin K inhibitor 2, two benzyl groups have been supposed to be oriented toward the primed binding region thus allowing for favorable interactions with local hydrophobic amino acid residues. Hence, we designed *N,N*-dibenzylacrylamide-based ABPs. This idea was also realized to revisit the acrylamide substructure of the aforementioned natural products 3 and 4.

The selectivity of cathepsins mainly relies on the S2 subsite specificity.²¹ In case of cathepsin K, mapping studies revealed a clear prevalence for hydrophobic and aliphatic amino acids at the P2 position, in particular for leucine.^{22,23} Moreover, affinity can be further increased by concomitant introduction of natural or unnatural amino acids at other positions, and norleucine was identified in a combinatorial substrate screening approach to be well tolerated at the P3 position.²⁴ Accordingly, a dipeptidic or tripeptidic moiety (Leu-Gly or Nle-Leu-Gly) was introduced at the P2-P1 or P3-P1 positions of our probes to interact with the corresponding binding pockets.

We selected two fluorophores, coumarin 343 and cyanine 5 (Cy5), for detection of the target protease. In coumarin 343, the amino group of 7-aminocoumarin is rigidized because of cyclization leading to increased quantum yields and high fluorescence in aqueous media.²⁵ In the present design, this commonly used fluorophore was either directly attached or connected via a phenylene linker to the dipeptide moiety. Cy5 is valued for its exceptional biocompatibility, a fluorescence window ranging from more than 600 nm to the near infrared, and a high signal-to-noise ratio because of minimal autofluorescence from biological samples in this spectral range.²⁶ In case of our Cy5 probes, a valeroyl linker combined the fluorophore and the inhibitor substructure.

Chemistry. The convergent syntheses of the envisaged probes are composed of two branches. The first is depicted in Scheme 1 showing the preparation of coumarins and Cy5. The tricyclic aldehyde 5 was reacted in a Knoevenagel condensation with the malonate 6 to obtain coumarin 343 (compound 7). By employing dimethyl homoterephthalate (8), the new coumarin ester 9 was produced whose saponification gave the building block 10. The preparation of the Cy5 dye 16 was performed by following a procedure to a similar cyanine.²⁷ The first step, an *N*-alkylation of 2,3,3-trimethyl-3*H*-indole (11) with δ -bromovaleric acid (12) in

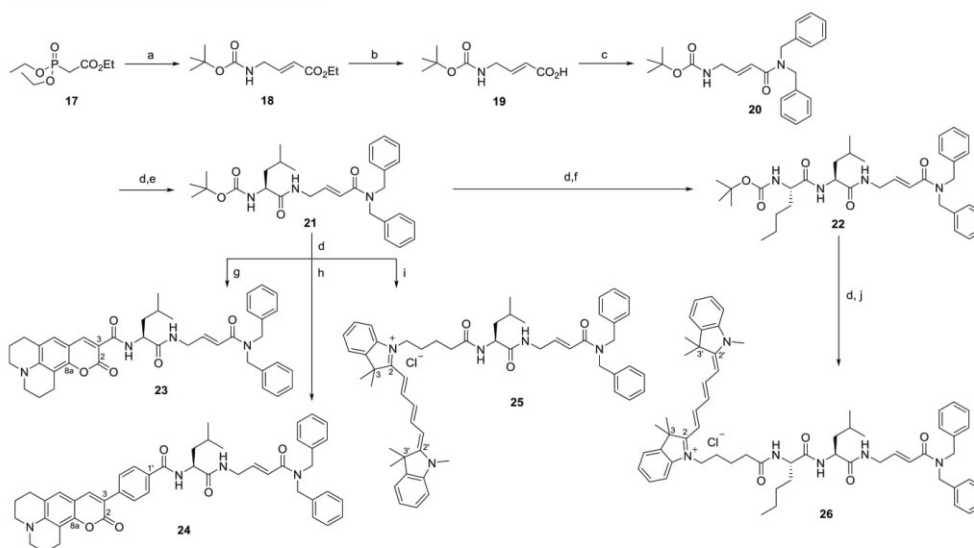
Scheme 1. Synthesis of Fluorescent Dyes 7, 10, and 16^{4f}

^aReagents and conditions: (a) piperidinium acetate, toluene, rt to reflux, 3 h; (b) piperidine toluene, reflux, 4 h; (c) LiOH × H₂O, THF/H₂O/EtOH 6 + 3 + 2, rt, 16 h; (d) MeNO₂, 80 °C, 6 h; (e) 1. Ac₂O, 120 °C, 0.5 h, 2. pyridine, rt, 70 h.

nitromethane, gave the indolium bromide **13** in low yields. Afterward, the condensation of **13** with malonaldehyde dianilide hydrochloride (**14**) to an intermediate hemicyanine was performed followed by a further, pyridine-promoted condensation with the tetramethylindolium salt **15** leading to the desired fluorescent component **16**. Its structural elucidation was in agreement with reported spectroscopic data.^{27–29}

The second branch of our synthesis started with a Horner–Wadsworth–Emmons reaction of the diethyl phosphonate **17** with Boc-protected glycinal in the presence of potassium *tert*-butylate (Scheme 2). After hydrolysis of **18** to the free acid **19**, this was converted to the dibenzylcarboxamide **20**. The next uronium-promoted peptide coupling of the corresponding deprotected amino component produced the dipeptidic building block **21** with leucine at the P2 position. In the converging step, deprotected **21** was linked via an amide bond with each of the three fluorescent components **7**, **10**, and **16**, respectively, to yield the probes **23–25**. Our fourth probe was structurally extended by norleucine as the P3 amino acid. For this purpose, the protected peptide **22** was prepared from **21**, prior to the completion of the Cy5-containing probe **26**.

Enzyme Kinetics. The four ABPs **23–26** together with their precursors **21** and **22** were fully characterized in *in vitro* kinetic assays with four human cathepsins. Next to cathepsin K, the closely related lysosomal cysteine proteases cathepsins S, L, and B were examined. Enzymatic activities were monitored for 60 min in the presence of a chromogenic or fluorogenic peptide substrate and five different inhibitor concentrations. The progress curves indicated time-dependent

Scheme 2. Synthesis of Probes 23–26^{4f}

^aReagents and conditions: (a) Boc-Gly-H, KO^tBu, THF, 0 °C to rt, 20 h; (b) LiOH × H₂O, THF/H₂O 2 + 1, 0 °C to rt, 5 h; (c) HATU, Bn₂NH, DIPEA, CH₂Cl₂, rt, 16 h; (d) TFA/CH₂Cl₂ 1 + 6, 0 °C to rt, 20 h; (e) Boc-Leu-OH × H₂O, HATU, DIPEA, CH₂Cl₂, rt, 16 h; (f) Boc-Nle-OH × DCHA, HATU, DIPEA, CH₂Cl₂, rt, 16 h; (g) **7**, HATU, DIPEA, CH₂Cl₂, rt, 18 h; (h) **10**, HBTU, DIPEA, CH₂Cl₂, rt, 16 h; (i) **16**, HATU, DIPEA, CH₂Cl₂, rt, 18 h; (j) **16**, HATU, DIPEA, CH₂Cl₂, rt, 68 h.

13795

<https://doi.org/10.1021/acs.jmedchem.1c01178>
J. Med. Chem. 2021, 64, 13793–13806

inhibition with rates of product formation approaching zero (Figure S1, Supporting Information). This was in accordance with the expected mode of enzyme–inhibitor interaction, comprising the irreversible nucleophilic attack of the active-site cysteine at the electrophilic β -position of the α,β -unsaturated Michael acceptor system present in 21–26. Nonlinear regression of the progress curves gave k_{obs} values, which were replotted against the inhibitor concentrations to obtain second-order rate constants of inactivation k_{inac}/K_i (Table 1).

Table 1. Inhibition of Human Cathepsins by Precursors 21 and 22 and Probes 23–26^a

cmpd.	k_{inac}/K_i ($\text{M}^{-1}\text{s}^{-1}$) ^b or residual activity (%) ^c			
	cathepsin K	cathepsin S	cathepsin L	cathepsin B
21	10,600 $\text{M}^{-1}\text{s}^{-1}$	291 $\text{M}^{-1}\text{s}^{-1}$	83%	91%
22	6440 $\text{M}^{-1}\text{s}^{-1}$	485 $\text{M}^{-1}\text{s}^{-1}$	76%	91%
23	9550 $\text{M}^{-1}\text{s}^{-1}$	80%	89%	96%
24	84,200 $\text{M}^{-1}\text{s}^{-1}$	1510 $\text{M}^{-1}\text{s}^{-1}$	2820 $\text{M}^{-1}\text{s}^{-1}$	365 $\text{M}^{-1}\text{s}^{-1}$
25	35,300 $\text{M}^{-1}\text{s}^{-1}$	505 $\text{M}^{-1}\text{s}^{-1}$	89%	96%
26	17,900 $\text{M}^{-1}\text{s}^{-1d}$	556 $\text{M}^{-1}\text{s}^{-1}$	91%	96%

^aKinetic inhibition assays were carried out with chromogenic peptide substrates. ^bSecond-order rate constants of inactivation were determined in duplicate measurements with five different inhibitor concentrations, when residual activities were less than 60%. ^cMean values from duplicate measurements at a single inhibitor concentration of 2.5 μM . ^dThe inhibition assay was carried out with a fluorogenic peptide substrate.

Overall, a striking selectivity for cathepsin K over the other investigated human cysteine proteases was achieved, thus indicating that the knowledge-driven conceptualization of the ABPs was successful. The two precursors 21 and 22 showed satisfactory activity against cathepsin K ($k_{\text{inac}}/K_i > 5000 \text{ M}^{-1}\text{s}^{-1}$, Table 1). Concomitantly, only the most closely related cathepsin S was inhibited, still yielding a 36- and 13-fold selectivity for 21 and 22, respectively. It is noteworthy that the elongation by a further amino acid, realized in the tripeptidomimetic 22, neither improved the inactivation capacity nor the selectivity profile. The ethyl acrylate counterpart of 21 showed less favorable selectivity indexes,¹³ indicating the relevance of the dibenzyl PSBE to achieve cathepsin K selectivity.

By introducing coumarin 343, the inhibitory profile of the resulting ABP 23 was not strongly influenced; selectivity and potency were roughly maintained. ABP 24 exhibited a greatly improved cathepsin K inhibition by 8-fold when compared to its precursor 21 and was the most potent compound out of the herein described series with a second-order rate constant k_{inac}/K_i of $84,200 \text{ M}^{-1}\text{s}^{-1}$. The design of 24 was guided by the structure of odanacatib (compound 1 in Figure 1) and several other dipeptidic cathepsin K inhibitors with a common biaryllic pattern reaching into the S3 pocket. In odanacatib, the biphenyl is attached to the amide-bioisosteric trifluoroethylamine group, and the $\text{C}_6\text{H}_4\text{—CHCF}_3\text{—NH}$ unit corresponds to the $\text{C}_6\text{H}_4\text{—CO—NH}$ unit of 24.

The CyS-containing ABPs 25 and 26 were found to efficiently inhibit cathepsin K. Again, the incorporation of norleucine reduced the potency nearly twofold (25 vs. 26). The highly active compound 25 ($k_{\text{inac}}/K_i = 35,300 \text{ M}^{-1}\text{s}^{-1}$) displayed an extraordinary selectivity for cathepsin K over cathepsins S, L, and B. This finding was preferable, but somewhat unexpected in the light of the similar substrate

specificities of human cysteine cathepsins.²³ The target preference of the probes could be attributed to the presence of the tailored PSBE, which we implemented based on the structure of cathepsin K inhibitor 2 (Figure 1). To unequivocally confirm the binding mode of our probes and the accommodation of the PSBE in the active site of human cathepsin K, X-ray crystal structure analyses were conducted.

Crystallographic Analysis and Identification of Binding Modes. Recombinant human cathepsin K was crystallized in complex with two inhibitory ligands, ABP 25 and its precursor 21 (Figure 2), confirming the covalent protease–inhibitor interaction. The crystal structures were determined by molecular replacement using the structure of human

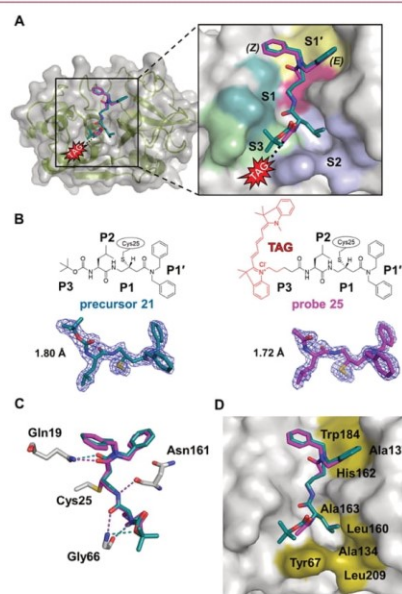


Figure 2. Binding mode of precursor 21 and probe 25 in the active site of human cathepsin K (PDB codes 7NXL and 7NXM). (A) The zoomed-in view of the active site shows a superposition of 21 and 25 bound to the S3 to S1' subsites. Cathepsin K is displayed in surface representation with the catalytic residues Cys25 and His162 highlighted in pink and individual subsites in different colors. Ligands are shown in stick representation with carbon atoms in dark cyan for 21 and magenta for 25. Heteroatoms have standard color-coding (O, red; N, blue; S, yellow) and (Z)- and (E)-positioned benzyl groups are marked. The fluorescence tag of the probe 25 is not included in the crystal structure, and its position is indicated (TAG). (B) P3 to P1' moieties of the ligands bind to corresponding S3 to S1' subsites. The active site Cys25 forms a covalent bond to the ligands, and the newly formed (R)-configured stereocenter is shown. The $2F_o - F_c$ electron density maps are contoured at 1 σ ; the side chain of the covalently linked Cys25 is depicted. (C) Hydrogen bond network formed by cathepsin K with ligands 21 and 25 (dark cyan and magenta dashed lines, respectively). Inhibitors are colored as in (A), and interacting enzyme residues are in gray. (D) Surface representation of the cathepsin K active site shows residues (highlighted in yellow) forming nonpolar interactions with the ligands (in stick representation).

cathepsin K as the template. The final crystallographic models contain mature cathepsin K residues 1–215 and an additional N-terminal Arg residue derived from the cathepsin K propeptide. Both complexes crystallized in the orthorhombic space group $P2_2_1$ with one molecule in the asymmetric unit and a solvent content of ~38% (Table S1; CatK-21, PDB ID 7NXL; CatK-25, PDB ID 7NXM). The structures were refined using data to resolutions of 1.80 and 1.72 Å for **21** and **25** complexes, respectively. The electron density applied for the modeling of the ligands was of high quality in both structures, with the exception of the valeryl-connected Cy5 fluorophore part of ABP **25**, for which the electron density was missing (Figure 2A,B). This suggested that the Cy5 tag protrudes from the enzyme. A comparison of cathepsin K in both complexes revealed a shared core scaffold with no significant differences in the overall protein structure (backbone RMSD value of 0.126 Å).

The active site cleft of cathepsin K contains the catalytic triad residues Cys25, His162, and Asn182. The binding mode of both ligands is presented in Figure 2. It shows the formation of a covalent bond between the active-site cysteine sulfur and the prochiral electrophilic β -carbon of the acrylamide warhead. The underlying reaction resulted in a (*R*)-chirality center at the sp^3 -hybridized carbon, implicating a nucleophilic attack of the Cys25 thiolate from the *si* face of the inhibitor (Figure 2B). Our data are concordant with the X-ray crystal structure of *N*-desmethyl thalassospiramide **4** in complex with human cathepsin K.²⁰ As predicted, both inhibitory ligands adopt a substrate-like orientation; their N-terminal parts bind into the nonprimed S3 to S1 region of the active site cleft, and the PSBEs are directed toward the S1' subsite (Figure 2). Although the N-terminal Cy5 tag of ABP **25** was not included in the crystallographic model, the binding mode was analogous to that of its precursor **21**, with a slight deviation of the acrylamide carbonyl groups only.

In both complexes, there is a common hydrogen bond network between the inhibitor backbone and the active site of cathepsin K (Figure 2C, Table S2). The Gln19 side chain plays a major role in cathepsin K catalysis by forming the transition-state stabilizing oxyanion hole.³⁰ In the CatK-**21** and CatK-**25** complexes, it donates a hydrogen bond to the inhibitors' acrylamide oxygen, adjacent to the PSBE. The respective interaction of Gln19, a component of the primed binding region of cathepsin K,^{31,32} with a sulfonyl oxygen of a cocrystallized vinyl sulfone inhibitor was reported.³³ The Asn161 backbone oxygen makes a hydrogen bond with the amide NH of both ligands. The NH of Gly66 acts as a bifurcated hydrogen bond donor for the carbonyl oxygen of the P2 leucine and the noncarbonyl carbamate oxygen of **21**. A further hydrogen bond is established between the Gly66 oxygen and the carbamate NH of **21**. As a consequence of the exchange of the Boc-capping group in **21** by the fluorophore-linker in **25** and the loss of the noncarbonyl carbamate oxygen, only two hydrogen bond interactions between Gly66 and the cocrystallized **25** emerged.

We further analyzed structural determinants of both ligands with respect to nonhydrogen bond interactions with components of the cathepsin K subsites. Selected nonpolar contacts are depicted in Figure 2D. The PSBEs of the ligands made the largest number of contacts with the protease (Table S2). The (*E*)-positioned benzyl group points toward Ala137, Ser138, and Gln143 and is flanked by Trp184 and the catalytically essential His162 (Figure 2D). These amino acid

residues are assumed to form the S1' pocket of cathepsin K.^{30,32–34} The (*Z*)-positioned benzyl group approaches the opposite site of the S1' pocket, a region close to the S2' pocket with residues Gly20, Gln21, Cys22, and Gly23.³¹ Electron density maps indicated several potential conformations of the benzyl groups in both complexes suggesting their static or dynamic disorder, and only the major conformation of each was modeled into the map.

The P2 leucine residues of **21** and **25** accommodate the S2 pocket of cathepsin K and are engaged in nonpolar interactions with the S2 residues Tyr67, Ala134, Leu160, Ala163, and Leu 209 (Figure 2D). The Boc-capping group of **21** resides in the narrow S3 groove of cathepsin K and makes contact with Asp61, Gly64, Gly65, Gly66, and Tyr67 (Table S2). In case of **25**, an apparent acetyl group derived from the linker was modeled into the electron density map interacting with Gly66 and Tyr67 of the S3 subsite.

In-Gel Detection of Cathepsin K. Having selected the two most active compounds out of the four ABPs (Table 1), as a further step of probe evaluation, we established in-gel detection of human cathepsin K with ABPs **24** and **25**. Recombinantly expressed, purified, and activated cathepsin K was labeled with 5 μ M of the dipeptidic Cy5-tagged probe **25**, subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), and visualized by fluorescence imaging (Figure 3). We observed a distinct fluorescent band at

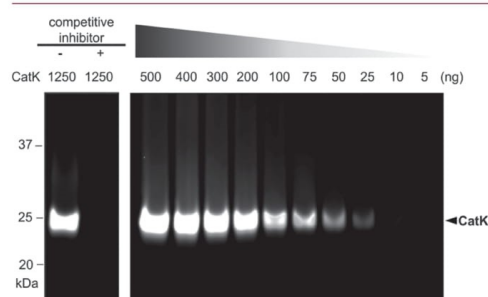


Figure 3. Imaging of human cathepsin K with ABP **25**. Left: Purified recombinant cathepsin K (CatK, 1.25 μ g) was incubated with **25** (5 μ M). In the control experiment, the enzyme was pretreated with a mixture of the competitive inhibitors E-64/K11777 (both 10 μ M) prior to incubation with **25** (5 μ M). Right: Cathepsin K (CatK, 500–5 ng) was labeled with **25** (5 μ M). Incubations were performed for 1 h at 25 °C and at pH 5.5. The reactions were resolved by SDS-PAGE and visualized with a Typhoon RGB imager.

~25 kDa which was completely abolished upon preincubation with a mixture of two competitive and irreversible inhibitors, that is, the pan-cathepsin inhibitors E-64 and K11777 (Figure 3, left). This finding confirmed the active site-directed mode of inactivation by **25**. Additionally, probe **25** was not interacting with nucleophilic residues at the protein's surface, which would have led to a residual fluorescence at ~25 kDa. The sensitivity of ABP **25** was assessed by applying decreasing amounts of purified cathepsin K, and a detection limit of 25 ng was determined (Figure 3, right).

For selectivity studies, **25** was applied in the complex protein mixture of non-cathepsin-K-expressing human embryonic kidney (HEK) cell lysate, either in the presence or absence

of additionally added cathepsin K. Bovine serum albumin (BSA) was added to the reaction mixtures to protect HEK lysate proteins from being excessively degraded by cathepsin K, thus maintaining the protein composition for the entire incubation time. The reactions were resolved by SDS-PAGE and visualized by fluorescence imaging followed by protein staining (Figure 4). Under these conditions of protein excess

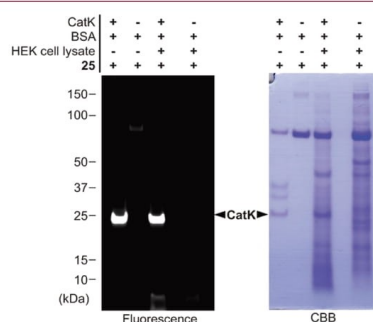


Figure 4. Imaging of human cathepsin K with ABP 25 in the presence of the cell lysate. Human recombinant cathepsin K (CatK, 300 ng) was treated with ABP 25 (5 μ M) in the presence of the HEK cell lysate (30 μ g) and BSA (500 ng) for 1 h at 25 $^{\circ}$ C and pH 5.5. In control experiments, cathepsin K (300 ng) and BSA (500 ng), or BSA (500 ng) alone, or HEK cell lysate (30 μ g) and BSA (500 ng) were treated with ABP 25 (5 μ M), separately. The reactions were resolved by SDS-PAGE and visualized with a Typhoon RGB imager (fluorescence, left) and subsequently stained with Coomassie Brilliant Blue (CBB, right).

(Figure 4, right, third lane), a single clear and bright fluorescent band was observed (Figure 4, left, third lane), corresponding to the fluorescence signal of cathepsin K in the reaction without the cell lysate (Figure 4, left, first lane). Furthermore, the incubation of 25 with the HEK cell lysate without spiked cathepsin K did not show any off-target labeling (Figure 4, left, fourth lane). Also, we confirmed the preference of 25 for cathepsin K over the highly related human cathepsins L, B, and S, which were not (cathepsin B) or only marginally (cathepsins L and S) labeled (Figure S2).

The cyanine fluorophore mainly contributed to the advantageous imaging properties of 25 in comparison to the kinetically more active, dipeptidic coumarin probe 24. The latter ABP was employed in the same set of experiments (Figures S3,S4). However, probe 24 exhibited lower fluorescence intensity, and the fluorescence signal reduction after incubation with the competing mixture of inhibitors was less pronounced than with probe 25. While we were able to detect as little as 25 ng of cathepsin K with ABP 25 (Figure 3), the detection limit of 24 ranged between 100 and 75 ng (Figure S3).

Imaging of Cathepsin K in Human Osteosarcoma Cells U-2 OS. Next, we demonstrated the cellular functionality of the superior ABP 25 by imaging live osteosarcoma cells, which show enhanced expression levels of cathepsin K.³⁵ Labeling of cathepsin K in U-2 OS cell lines was analyzed by in-gel fluorescence and live-cell microscopy. First, ABP 25 was incubated with U-2 OS cells for various times, and the harvested cells were subjected to SDS-PAGE and visualized by fluorescence imaging (Figure 5A). A fluorescent band at 25

kDa corresponding to cathepsin K was clearly detected after 4 h incubation, the maximum intensity was reached after 16 h, and the fluorescence was persistent after 24 h treatment. In contrast, noncathepsin K-expressing cells, that is, hepatocellular carcinoma HepG cells and cervical cancer HeLa cells, did not provide a fluorescent cathepsin K signal. Slight off-target effects of 25, presumably associated with other cysteine cathepsins, were observed with HepG cells, while the HeLa cell proteome remained unaffected (Figure S2).

The labeling of cathepsin K was quenched when the U-2 OS cells were preincubated with E-64d, a cell-permeable, irreversible, and active-site-directed inhibitor of cysteine cathepsins, prior to application of the ABP 25 (Figure 5A). Cathepsin K was authenticated by immunoblotting analysis using a specific anticathepsin K antibody, which detected a 25 kDa single band. The highly potent, cathepsin K-specific inhibitor odanacatib strongly diminished the labeling of cathepsin K originating from U-2 OS cells (Figure S2). The finding that the signal did not entirely disappear was attributed to the reversible nature of odanacatib.

Furthermore, the ABP 25 was used to localize cathepsin K in U-2 OS cells by confocal fluorescence microscopy. The cells were incubated with the probe for 16 h, and a staining pattern was observed, which colocalized with the signal of LysoTracker, a marker of late endosomes and lysosomes (Figure 5B). In a competition experiment, the cathepsin K labeling was completely prevented by preincubation with E-64d. These results demonstrated that ABP 25 was efficiently taken up into endolysosomal compartments of U-2 OS cells and labeled the active form of cathepsin K, providing specific detection of the target in live cells.

CONCLUSIONS

We designed an exceptional ABP for cathepsin K by incorporating a dibenzylamine structural motif for a specific interaction with the primed binding region of the target protease. Fusing this PSBE motif with a peptidomimetic Michael acceptor gave rise to the irreversible inhibitor 21, whose binding mode was elucidated by X-ray structure analysis of the covalent enzyme–inhibitor complex, refined at a 1.8 Å resolution. The introduction of a linker-connected fluorophore (in 25) did not implement changes in the alignment of the inhibitory substructure within the active site cleft. While one benzyl group of the PSBE adopted the substrate-like orientation of P1' residues, the other resided on the opposite side close to the S2' pocket. Little is known from crystallographic and mapping studies on this binding region in the active site cleft of cathepsin K.^{22,31} At least, as shown with internally quenched fluorescent peptide substrates, aromatic residues at P1' and P2' positions seemed to be acceptable.²² Nevertheless, the presence of both benzylic groups, when integrated into the appropriate overall structure of our inhibitors 21–26, accounted for their outstanding potency and cathepsin K selectivity. The superior, Cy5-containing fluorescent probe 25, when applied to SDS-PAGE, enabled for highly specific and sensitive in-gel detection even in the presence of a complex proteome. The cellular uptake of probe 25 and trafficking into the lysosomes was manifested. Finally, we exposed a human osteosarcoma cell line to ABP 25 and demonstrated its capability to detect cathepsin K in a pathophysiologically relevant context. Owing to its promising biological properties, ABP 25 may be advanced to animal

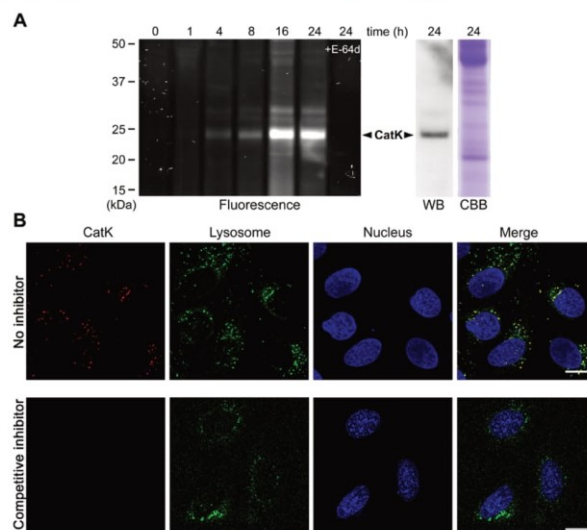


Figure 5. Imaging of cathepsin K with ABP 25 in human osteosarcoma cells U-2 OS. (A) Cells were incubated with 25 (1 μ M) for the indicated time and resolved by SDS-PAGE. The gel was visualized using a fluorescence scanner (left) and protein-stained with CBB (right) or analyzed by western blot with anticathepsin K antibody (WB, right). The position of cathepsin K is indicated (CatK). (B) U-2 OS cells were incubated for 24 h with 25 (1 μ M, CatK panels), stained with LysoTracker Green (50 nM, Lysosome) and Hoechst (Nucleus), and imaged live using confocal fluorescence microscopy. Scale bar: 20 μ m. In the control experiment, U-2 OS cells were pretreated with cathepsin K inhibitor E-64d (1 μ M, Competitive inhibitor) for 1 h prior to incubation with 25.

studies for investigating the in vivo detection of cathepsin K in murine models of osteoarthritis.

EXPERIMENTAL SECTION

Chemistry. General Information. Preparative column chromatography was performed using Merck silica gel 60 (63–200 mesh). Melting points were determined on a Büchi 510 oil bath apparatus and were uncorrected. ^1H NMR and ^{13}C NMR spectra were recorded on a Bruker Avance 500 MHz NMR spectrometer or on a Bruker Avance III 600 MHz NMR spectrometer, respectively. Chemical shifts are given in parts per million (ppm), coupling constants J are given in Hertz, and spin multiplicities are given as s (singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). High-resolution mass spectrometry (HRMS) spectra were recorded on a micrOTOF-Q mass spectrometer (Bruker) with an electrospray ionization (ESI)-source coupled with an HPLC Dionex Ultimate 3000 (Thermo Scientific). The purity of the compounds was determined by HPLC/DAD/MS obtained on an LC–MS instrument (Applied Biosystems API 2000 LC/MS/MS, HPLC Agilent 1100). An EC50/2 Nucleodur C18 Gravity 3 μ m column (Macherey-Nagel, Düren, Germany) was used. The mobile phase was a gradient of 90% H_2O to 100% MeOH containing 2 mM ammonium acetate in 10 min, then 100% MeOH containing 2 mM ammonium acetate to 20 min. Retention times (t_{R}) of all compounds are noted. The purity of all compounds was confirmed to be >95% purity by liquid chromatography–mass spectrometry (LC–MS). High-performance liquid chromatography (HPLC) traces of all in vitro test compounds are shown in the Supporting Information. All compounds passed the PAINS filter³⁶ with a false-positive remover.³⁷

Coumarin 343 (7).³⁸ 8-Hydroxyjulolidine-9-carboxaldehyde (5, 0.44 g, 2.00 mmol) was dissolved in absolute ethanol (20 mL). Piperidinium acetate (6 mg, 0.04 mmol) and isopropylidene malonate (6, 0.29 g, 2.00 mmol) were added, and the reaction mixture was stirred at rt for 20 min. Then, the reaction mixture was refluxed for

further 2 h. Cooling to rt and subsequent cooling on ice led to the crystallization of 7 as a bright orange solid (0.15 g, 0.53 mmol, 27%, mp 238–240 $^{\circ}\text{C}$, lit.³⁹ mp 253 $^{\circ}\text{C}$). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 1.83–1.91 (m, 4H, $\text{N}(\text{CH}_2\text{CH}_2)_2$), 2.67–2.74 (m, 4H, $\text{N}((\text{CH}_2)_2\text{CH}_2)_2$), 7.22 (s, 1H, 5-H), 8.44 (s, 1H, 4-H), the signal for $\text{N}(\text{CH}_2)_2$ is obscured by the H_2O peak; ^{13}C NMR (125 MHz, $\text{DMSO}-d_6$) δ 19.65 ($\text{N}(\text{CH}_2\text{CH}_2)_2$), 20.62, 26.90 ($\text{N}((\text{CH}_2)_2\text{CH}_2)_2$), 49.27, 49.80 ($\text{N}(\text{CH}_2)_2$), 104.91 (C-4a), 105.28 (C-8), 107.41 (C-6), 119.71 (C-3), 127.61 (C-4), 148.87 (C-5), 149.29 (C-7), 152.80 (C-8a), 160.79 (C-2), 164.77 (CO_2H). LC–MS (ESI), t_{R} 8.56 min, m/z calcd for $[\text{C}_{16}\text{H}_{15}\text{NO}_4 + \text{H}]^+$, 286.1, found, 285.9.

Methyl 4-(2,3,6,7-tetrahydro-11-oxo-1H,5H,11H-[1]-benzopyrano[6,7,8-ij]-10-quinoliziny)-benzoate (9). 8-Hydroxyjulolidine-9-carboxaldehyde (5, 0.44 g, 2.00 mmol), dimethyl homoterephthalate (8, 0.42 g, 2.00 mmol), and piperidine (0.55 g, 4.00 mmol) were dissolved in toluene (25 mL) and refluxed for 4 h. The orange precipitate was filtered off and purified by column chromatography using petroleum ether/EtOAc (3 + 1) as the eluent to obtain 9 as an orange solid (0.29 g, 0.77 mmol, 39%, mp 167–171 $^{\circ}\text{C}$, decomposition). ^1H NMR (600 MHz, CDCl_3) δ 1.93–2.00 (m, 4H, $\text{N}(\text{CH}_2\text{CH}_2)_2$), 2.76 (t , $J = 6.3$ Hz, 2H, $\text{N}(\text{CH}_2)_2\text{CH}_2$), 2.91 (t , $J = 6.5$ Hz, 2H, $\text{N}(\text{CH}_2)_2\text{CH}_2$), 3.26, 3.29 (both t , $J = 6.3$ Hz, 4H, $\text{N}(\text{CH}_2)_2$), 3.90 (s, 3H, OCH_3), 6.90 (s, 1H, 5-H), 7.67 (s, 1H, 4-H), 7.77 (d , $J = 8.4$ Hz, 2H, 3'-H), 8.03 (d , $J = 8.4$ Hz, 2H, 2'-H); ^{13}C NMR (150 MHz, CDCl_3) δ 20.28, 20.50 ($\text{N}(\text{CH}_2\text{CH}_2)_2$), 21.42, 27.52 ($\text{N}((\text{CH}_2)_2\text{CH}_2)_2$), 49.69, 50.05 ($\text{N}(\text{CH}_2)_2$), 52.07 (OCH_3), 106.21 (C-4a), 108.81 (C-8), 118.33, 118.74 (C-5, C-6), 125.37 (C-3), 128.71 (C-1'), 127.94, 129.55 (C-2', C-3'), 140.81 (C-4'), 141.71 (C-4), 146.26 (C-7), 151.51 (C-8a), 161.49 (C-2), 166.97 (CO_2Me). LC–MS (ESI), t_{R} 12.11 min, m/z calcd for $[\text{C}_{23}\text{H}_{21}\text{NO}_4 + \text{H}]^+$, 376.2, found, 376.2.

4-(2,3,6,7-Tetrahydro-11-oxo-1H,5H,11H-[1]-benzopyrano[6,7,8-ij]-10-quinoliziny)-benzoic acid (10). Methyl ester 9 (0.38 g, 1.00 mmol) was dissolved in THF (60 mL), H_2O (30 mL), and EtOH (20

mL). Upon addition of LiOH \times H₂O (0.42 g, 10.0 mmol), the reaction mixture was stirred for 16 h at rt. The organic solvent was evaporated, and the aqueous layer was acidified using 2 M HCl (6 mL). The resulting yellowish precipitate was filtered off and dried in vacuo (0.13 g, 0.36 mmol, 36%, mp 186–190 °C, decomposition). ¹H NMR (600 MHz, CDCl₃) δ 1.97–2.01 (m, 4H, N(CH₂CH₂)₂), 2.78 (t, ³J = 6.4 Hz, 2H, N(CH₂)₂CH₂), 2.93 (t, ³J = 6.5 Hz, 2H, N(CH₂)₂CH₂), 3.28, 3.31 (both t, ³J = 6.7 Hz, 4H, N(CH₂)₂), 6.93 (s, 1H, 5-H), 7.72 (s, 1H, 4-H), 7.83 (d, ³J = 8.4 Hz, 2H, 3'-H), 8.12 (d, ³J = 8.4 Hz, 2H, 2'-H); ¹³C NMR (150 MHz, CDCl₃) δ 20.27, 20.49 (N(CH₂CH₂)₂), 21.41, 27.52 (N((CH₂)₂CH₂)₂), 49.70, 50.07 (N(CH₂)₂), 106.21 (C-4a), 108.81 (C-8), 118.18 (C-6), 118.79 (C-5), 125.44 (C-3), 127.68 (C-1'), 128.03, 130.18 (C-2', C-3'), 141.88 (C-4'), 143.16 (C-4), 146.35 (C-7), 151.56 (C-8a), 161.49 (C-2), 170.23 (CO₂H). LC–MS (ESI), t_R 10.27 min, m/z calcd for [C₂₂H₁₉NO₄ + H]⁺, 362.1, found, 362.1.

1-(4-Carboxybutyl)-2,3,3-trimethyl-3H-indolium bromide (13). 2,3,3-Trimethyl-3H-indole (**11**, 3.17 g, 3.2 mmol, 20.00 mmol) was added to a solution of 5-bromovaleric acid (**12**, 3.70 g, 20.00 mmol) in nitromethane (12 mL) and stirred at 80 °C for 8 h. The reaction mixture was treated with diethyl ether (80 mL), and the precipitate was filtered off. Washing of the precipitate with dry acetone (25 mL) and recrystallization from MeOH/diethyl ether (**1** + **9**) led to **13** as a yellow solid (1.00 g, 3.00 mmol, 15%; mp 200–202 °C; lit.²⁸ mp 188–190 °C). ¹H NMR (500 MHz, DMSO-*d*₆) δ 1.54 (s, 6H, 3-CH₃), 1.62–1.68 (m, 2H, CH₂CH₂CO₂H), 1.81–1.91 (m, 2H, CH₂(CH₂)₂CO₂H), 2.31 (t, ³J = 7.3 Hz, 2H, CH₂CO₂H), 2.85 (s, 3H, 2-CH₃), 4.49 (t, ³J = 7.7 Hz, 2H, CH₂(CH₂)₂CO₂H), 7.54–7.74 (m, 2H, 4-H, 5-H), 7.80–7.89 (m, 1H, 6-H), 7.93–8.07 (m, 1H, 7-H), 12.09 (s, 1H, CO₂H); ¹³C NMR (125 MHz, DMSO) δ 13.96 (2-CH₃), 21.40 (CH₂CH₂CO₂H), 22.00 (3-CH₃), 26.63 (CH₂(CH₂)₂CO₂H), 32.97 (C-14, CH₂CO₂H), 47.24 (C-3), 54.15 (CH₂(CH₂)₂CO₂H), 115.40 (C-7), 123.46 (C-4), 128.88 (C-5), 129.38 (C-6), 141.08 (C-7a), 141.84 (C-3a), 174.00 (CO₂H), 196.57 (C-2). LC–MS (ESI), t_R 5.64 min, m/z calcd for [C₁₆H₂₂NO₂]⁺, 260.2, found, 259.9.

1-(4-Carboxybutyl)-3,3-dimethyl-2-((1E,3E)-5-(1,3,3-trimethylindolin-2-ylidene)penta-1,3-dienyl)-3H-indolium chloride (Cys, 16).²⁸ 1-(4-Carboxybutyl)-2,3,3-trimethyl-3H-indolium bromide (**13**, 0.50 g, 1.45 mmol) was added to a solution of malonaldehyde diamide hydrochloride (**14**, 0.46 mg, 1.75 mmol) in acetic anhydride (30 mL) and stirred for 30 min at 120 °C. A solution of 1,2,3,3-tetramethyl-3H-indolium iodide (**15**, 0.62 mg, 2.05 mmol) in pyridine (20 mL) was added to the prepared solution at rt, and the reaction mixture was stirred for 70 h. Solvents were removed, the crude product was dissolved in CH₂Cl₂ (10 mL), and *n*-hexane (100 mL) was added. The solution was decanted, and the precipitate was dissolved in CH₂Cl₂ (80 mL), washed with H₂O (3 \times 80 mL) and 0.1 M HCl (80 mL), and dried over Na₂SO₄. Purification by column chromatography (CH₂Cl₂/MeOH, 19 + 1 + 3% AcOH) led to **16** as a blue, red shining solid (0.27 g, 0.53 mmol, 36%, mp. 134–136 °C). ¹H NMR (600 MHz, DMSO-*d*₆) δ 1.20–1.28 (m, 2H, CH₂CH₂CO₂H), 1.56–1.62 (m, 2H, CH₂(CH₂)₂CO₂H), 1.67 (s, 12H, 3-CH₃, 3'-CH₃), 2.10 (t, ³J = 7.1 Hz, 2H, CH₂CO₂H), 3.59 (s, 3H, NCH₃), 4.09 (t, ³J = 7.2 Hz, 2H, CH₂(CH₂)₂CO₂H), 6.26 (d, ³J = 13.7 Hz, 1H, CH(CH₃)CH), 6.34 (d, ³J = 13.8 Hz, 1H, CH(CH₃)CH), 6.59 (t, ³J = 12.4 Hz, 1H, (CH)₂CH(CH₂)₂), 7.24 (t, ³J = 7.2 Hz, 2H, 5-H, 5'-H), 7.35–7.41 (m, 4H, 6-H, 7-H, 6'-H, 7'-H), 7.60 (dd, ³J = 7.4 Hz, ⁴J = 3.3 Hz, 2H, 4-H, 4'-H), 8.31 (t, ³J = 13.0 Hz, 2H, CHCH(CH₃)). ¹³C NMR (125 MHz, DMSO-*d*₆) δ 22.69 (CH₂CH₂CO₂H), 26.72 (CH₂(CH₂)₂CO₂H), 27.02, 27.12 (3-CH₃, 3'-CH₃), 31.00 (NCH₃), 35.85 (CH₂CO₂H), 43.33 (CH₂(CH₂)₂CO₂H), 48.74, 48.86 (C-3, C-3'), 103.17, 103.26 (CH(CH₃)CH), 110.91, 111.13 (C-7, C-7'), 122.23, 122.34 (C-4, C-4'), 124.57, 124.62 (C-5, C-5'), 125.49 ((CH)₂CH(CH₂)₂), 128.30, 128.40 (C-6, C-6'), 140.94, 141.09 (C-3a, C-3'a), 142.01, 142.78 (C-7a, C-7'a), 153.81, 154.05 (CHCH(CH₃)), 172.58, 172.99 (C-2, C-2'), 173.57 (CO₂H). LC–MS (ESI), t_R 7.90 min, m/z calcd for [C₃₁H₃₉N₂O₂]⁺, 469.3, found, 469.2.

(E)-Ethyl 4-(tert-butoxycarbonylamino)but-2-enoate (18).¹³ Potassium *tert*-butoxide (0.57 g, 5.00 mmol) was added to a solution of ethyl 2-(diethoxyphosphoryl)acetate (**17**, 1.15 g, 5.00 mmol) in THF (55 mL) at 0 °C and stirred for 30 min while it was allowed to reach room temperature. Boc-Gly-H (0.80 g, 5.00 mmol) was then added dropwise to the reaction mixture and stirred for 20 h at room temperature. The solvent was evaporated under reduced pressure, and the resulting oil was dissolved in diethyl ether (100 mL), washed with saturated NaHSO₄ solution (2 \times 100 mL), H₂O (100 mL), and 10% KHSO₄ solution (2 \times 100 mL), and dried over Na₂SO₄. Purification by column chromatography (petroleum ether/EtOAc, 3 + 1) led to **18** as a yellow oil (0.74 g, 3.20 mmol, 65%). ¹H NMR (600 MHz, DMSO-*d*₆) δ 1.19 (t, ³J = 7.1 Hz, 3H, CH₂CH₃), 1.37 (s, 9H, C(CH₃)₃), 3.71 (t, ³J = 4.0 Hz, 2H, CH₂CHCH), 4.10 (q, ³J = 7.1 Hz, 2H, CH₂CH₃), 5.82 (d, ³J = 15.7 Hz, 1H, CH₂CHCH), 6.79 (dt, ³J = 15.7 Hz, ³J = 4.7 Hz, 1H, CH₂CHCH), 7.14 (s, 1H, NH). ¹³C NMR (150 MHz, DMSO-*d*₆) δ 14.30 (CH₂CH₃), 28.36 (C(CH₃)₃), 40.87 (CH₂CHCH), 60.06 (CH₂CH₃), 78.21 (C(CH₃)₃), 120.19 (CH₂CHCH), 146.71 (CH₂CHCH), 155.71 (OCONH), 165.65 (CO₂CH₂). LC–MS (ESI), t_R 9.59 min, m/z calcd for [C₁₁H₂₀NO₄]⁺, 230.2, found, 230.1.

(E)-4-(tert-Butoxycarbonylamino)but-2-enoic acid (19). (E)-Ethyl 4-(tert-butoxycarbonylamino)but-2-enoate (**18**, 0.70 g, 3.00 mmol) was dissolved in THF (20 mL) and H₂O (10 mL) and stored at 0 °C. Upon addition of LiOH \times H₂O (0.38 g, 9.00 mmol), the reaction mixture was stirred for 30 min at 0 °C and for further 4.5 h at room temperature. After evaporation of THF, the aqueous phase was acidified to pH < 2 with 1 N HCl solution and extracted with EtOAc (3 \times 30 mL) and dried over Na₂SO₄ to yield a white solid (0.56 g, 2.79 mmol, 93%, mp 116–118 °C, lit.⁴¹ mp 115 °C). ¹H NMR (500 MHz, DMSO-*d*₆) δ 1.38 (s, 9H, C(CH₃)₃), 3.70 (t, ³J = 4.2 Hz, 2H, CH₂CHCH), 5.75 (dt, ³J = 15.5 Hz, ⁴J = 1.5 Hz, 1H, CH₂CHCH), 6.72 (dt, ³J = 15.7 Hz, ³J = 4.8 Hz, 1H, CH₂CHCH), 7.12 (s, 1H, OCONH), 12.22 (s, 1H, CO₂H). ¹³C NMR (125 MHz, DMSO-*d*₆) δ 28.34 (C(CH₃)₃), 40.79 (CH₂CHCH), 78.12 (C(CH₃)₃), 121.18 (CH₂CHCH), 145.93 (CH₂CHCH), 155.68 (OCONH), 167.04 (CO₂H). LC–MS (ESI), t_R 2.87 min, m/z calcd for [C₉H₁₉N₂O₄]⁺, 219.1, found, 218.9.

(E)-tert-Butyl 4-(dibenzylamino)-4-oxobut-2-enylcarbamate (20). HATU (1.25 g, 3.35 mmol) was added to a solution of (E)-4-(tert-butoxycarbonyl-amino)but-2-enoic acid (**19**, 0.47 g, 2.30 mmol) and DIPEA (0.4 mL, 0.30 g, 2.30 mmol) in dry CH₂Cl₂ (15 mL). The mixture was flooded with argon and stirred for 30 min. A solution of dibenzylamine (0.60 g, 3.00 mmol) and DIPEA (0.4 mL, 0.30 g, 2.30 mmol) in dry CH₂Cl₂ (5 mL) was prepared and added to the reaction mixture which was stirred for 16 h at room temperature. Solvents were removed, and the crude product was dissolved in EtOAc (15 mL). The organic phase was washed with 10% KHSO₄ solution (3 \times 50 mL) and dried over Na₂SO₄. Purification by column chromatography (petroleum ether/EtOAc 3 + 2) led to **20** as a white solid (0.58 g, 1.51 mmol, 66%, mp 114–116 °C). ¹H NMR (600 MHz, DMSO-*d*₆) δ 1.32 (s, 9H, C(CH₃)₃), 3.71 (s, 2H, CH₂CHCH), 4.52 (s, 2H, NCH₂), 4.56 (s, 2H, NCH₂), 6.49 (d, ³J = 15.1 Hz, 1H, CH₂CHCH), 6.74 (dt, ³J = 15.1 Hz, ³J = 4.9 Hz, 1H, CH₂CHCH), 7.09 (s, 1H, OCONH), 7.19 (d, ³J = 7.5 Hz, 2H, NCH₂Ph), 7.22 (d, ³J = 7.5 Hz, 2H, NCH₂Ph), 7.23–7.39 (m, 6H, NCH₂Ph). ¹³C NMR (125 MHz, DMSO-*d*₆) δ 28.13 (C(CH₃)₃), 40.98 (CH₂CHCH), 48.17 (NCH₂), 49.77 (NCH₂), 77.86 (C(CH₃)₃), 119.85 (CH₂CHCH), 126.65, 127.08, 127.35, 127.69, 128.44, 128.69, 137.26, 137.62 (NCH₂Ph), 143.42 (CH₂CHCH), 155.48 (OCONH), 165.89 (CON). LC–MS (ESI), t_R 11.69 min, m/z calcd for [C₂₃H₃₂N₃O₃]⁺, 398.2, found, 398.2.

(S,E)-tert-Butyl 1-(4-(dibenzylamino)-4-oxobut-2-enylamino)-4-methyl-1-oxopentane-2-ylcarbamate (21). (E)-tert-Butyl 4-(dibenzylamino)-4-oxobut-2-enylcarbamate (**20**, 0.35 g, 0.90 mmol) was added to a solution of TFA (1 mL) in CH₂Cl₂ (6 mL) and stirred for 20 h at room temperature. Solvents were removed, and the crude brownish resin was dissolved in dry CH₂Cl₂ (50 mL) with DIPEA (0.20 g, 1.55 mmol). A solution of Boc-Leu-OH \times H₂O (0.25 g, 1.00 mmol), HATU (0.52 g, 1.35 mmol), and DIPEA (0.15 g, 1.15 mmol)

in dry CH_2Cl_2 (100 mL) was prepared, flooded with argon, and stirred for 30 min. Both solutions were combined and stirred for 16 h at room temperature. Solvents were removed, the crude product was dissolved in EtOAc (20 mL) and washed with saturated NaHCO_3 solution (2×20 mL), brine (20 mL), and 10% KHSO_4 solution (20 mL), and the organic phase was dried over Na_2SO_4 . Purification by column chromatography (petroleum ether/EtOAc, 1 + 1) led to **21** as a colorless resin (0.39 g, 0.80 mmol, 87%). ^1H NMR (600 MHz, $\text{DMSO}-d_6$) δ 0.79 ($d, J = 6.6$ Hz, 6H, $\text{CH}(\text{CH}_3)_2$), 1.32 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.35–1.43 (m, 2H, NHCHCH_2), 1.51–1.56 (m, 1H, $\text{CH}(\text{CH}_3)_2$), 3.84–3.93 (m, 3H, CH_2CHCH , NHCHCH_2), 4.40 ($d, J = 15.0$ Hz, 1H, NCH_2), 4.49 ($d, J = 16.8$ Hz, 1H, NCH_2), 4.57 ($d, J = 15.2$ Hz, 1H, NCH_2), 4.66 ($d, J = 16.8$ Hz, 1H, NCH_2), 6.59 ($d, J = 15.1$ Hz, 1H, CH_2CHCH), 6.77 ($dt, J = 14.7$ Hz, $J = 4.8$ Hz, 1H, CH_2CHCH), 6.84 ($d, J = 7.9$ Hz, 1H, OCONH), 7.17–7.38 (m, 10H, NCH_2Ph), 8.12 (bs, 1H, CONHCH_2). ^{13}C NMR (150 MHz, $\text{DMSO}-d_6$) δ 22.86, 24.10, 24.25 ($\text{CH}(\text{CH}_3)_2$), 28.15 ($\text{C}(\text{CH}_3)_3$), 40.76 (NHCHCH_2), 47.85 (NCH_2), 49.71 (NCH_2), 52.87 (NHCHCH_2), 77.93 ($\text{C}(\text{CH}_3)_3$), 119.75 (CH_2CHCH), 126.91, 127.08, 127.41, 127.63, 128.43, 128.71, 137.26, 137.56 (NCH_2Ph), 143.20 (CH_2CHCH), 155.41 (OCONH), 165.78 (CON), 172.65 (CONHCH_2). One signal (CH_2CHCH) is obscured by the DMSO solvent signal. LC–MS (ESI), t_R 11.89 min, m/z calcd for $[\text{C}_{29}\text{H}_{40}\text{N}_4\text{O}_3]^+$, 494.3, found, 494.1. HRMS m/z calcd for $[\text{C}_{29}\text{H}_{40}\text{N}_4\text{O}_3]^+$, 494.3013, found, 494.2964.

tert-Butyl-(*S*)-1-(*S*)-1-(*E*)-4-(dibenzylamino)-4-oxobut-2-enylamino)-4-methyl-1-oxopent-2-ylamino)-1-oxohexan-2-ylcarbamate (**22**). (*S*)-*tert*-Butyl-1-(4-(dibenzylamino)-4-oxobut-2-enylamino)-4-methyl-1-oxopent-2-ylcarbamate (**21**, 0.21 g, 0.43 mmol) was added to a solution of TFA (2 mL) in CH_2Cl_2 (10 mL) and stirred for 20 h. Solvents were removed, and the crude brownish resin was dissolved in dry CH_2Cl_2 (8 mL) with DIPEA (0.08 g, 0.60 mmol). A solution of Boc-Nle-OH \times DCHA (0.18 g, 0.48 mmol), HATU (0.25 g, 0.65 mmol), and DIPEA (0.09 g, 0.70 mmol) in dry CH_2Cl_2 (8 mL) was prepared, flooded with argon, and stirred for 30 min. Both solutions were combined and stirred for 16 h at room temperature. Solvents were removed, the crude product was dissolved in EtOAc (25 mL) and washed with saturated NaHCO_3 solution (2×25 mL), brine (25 mL), and 10% KHSO_4 solution (25 mL), and the organic phase was dried over Na_2SO_4 . Purification by column chromatography ($2 \times \text{CH}_2\text{Cl}_2/\text{MeOH}$, 29 + 1) led to **22** as a colorless oil (0.19 g, 0.32 mmol, 74%). ^1H NMR (600 MHz, $\text{DMSO}-d_6$) δ 0.74–0.84 (m, 9H, $\text{CH}(\text{CH}_3)_2$, CH_2CH_3 , $\text{CH}(\text{CH}_3)_2$), 1.14–1.26 (m, 4H, $\text{CHCH}_2(\text{CH}_2)_2\text{CH}_3$), 1.30–1.38 (s + m, 11H, $\text{C}(\text{CH}_3)_3$, $\text{CHCH}_2(\text{CH}_2)_2\text{CH}_3$), 1.39–1.47 (m, 2H, $\text{CH}_2\text{CH}(\text{CH}_3)_2$), 1.48–1.58 (m, 1H, $\text{CH}_2\text{CH}(\text{CH}_3)_2$), 3.72–3.94 (m, 3H, NHCHCH_2CH , CH_2CHCH), 4.24 ($dt, J = 8.8$ Hz, $J = 4.7$ Hz, 1H, $\text{CH}(\text{CH}_3)_2$), 4.42–4.66 (m, 4H, NCH_2 , NCH_2), 6.57 ($dt, J = 15.0$ Hz, $J = 1.9$ Hz, 1H, CH_2CHCH), 6.74–6.80 (m, 2H, OCONH , CH_2CHCH), 7.19 ($t, J = 6.3$ Hz, 4H, NCH_2Ph), 7.23–7.38 (m, 6H, NCH_2Ph), 7.74 ($d, J = 8.0$ Hz, 1H, CONH), 8.14–8.21 (m, 1H, CONH). ^{13}C NMR (150 MHz, $\text{DMSO}-d_6$) δ 13.84 ($\text{CH}(\text{CH}_3)_2$), 21.45, 21.77, 22.87, 24.04 ($\text{CH}(\text{CH}_2)_2\text{CH}_2\text{CH}_3$, $\text{CH}(\text{CH}_3)_2$), 27.49 ($\text{CHCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$), 28.11 ($\text{C}(\text{CH}_3)_3$), 31.34 ($\text{CHCH}_2(\text{CH}_2)_2\text{CH}_3$), 38.21 (CH_2CHCH), 40.86 ($\text{CH}_2\text{CH}(\text{CH}_3)_2$), 47.88 (NCH_2), 49.64 (NCH_2), 50.99 (NHCHCH_2CH), 54.41 ($\text{CH}(\text{CH}_3)_2$), 78.08 ($\text{C}(\text{CH}_3)_3$), 119.82 (CH_2CHCH), 126.85, 127.03, 127.32, 127.62, 128.38, 128.62, 137.30, 137.57 (NCH_2Ph), 142.90 (CH_2CHCH), 155.39 (OCONH), 165.64, 171.87, 171.99 (CON , CONH). LC–MS (ESI), t_R 8.87 min, m/z calcd for $[\text{C}_{35}\text{H}_{51}\text{N}_4\text{O}_5]^+$, 607.4, found, 607.6. HRMS m/z calcd for $[\text{C}_{35}\text{H}_{51}\text{N}_4\text{O}_5]^+$, 607.3854, found, 607.3883.

(*S*)-2-(2,3,6,7-Tetrahydro-11-oxo-1H,5H,11H-[1]benzopyrano[6,7,8-ij]-10-quinolizinyloxy)carbonylamino)-N-(4-(dibenzylamino)-4-oxobut-2-enyl)-4-methylpentanamide (**23**). (*S*)-*tert*-Butyl-1-(4-(dibenzylamino)-4-oxobut-2-enylamino)-4-methyl-1-oxopent-2-ylcarbamate (**21**, 69 mg, 0.14 mmol) was added to a solution of TFA (2 mL) in CH_2Cl_2 (10 mL) and stirred for 20 h. Solvents were removed, and the crude brownish resin was dissolved in dry CH_2Cl_2 (5 mL) with DIPEA (19 mg, 0.29 mmol). Coumarin 343 (7, 44 mg, 0.15

mmol), HATU (80 mg, 0.21 mmol), and DIPEA (19 mg, 0.29 mmol) were dissolved in CH_2Cl_2 (10 mL) and stirred at room temperature under argon for 30 min. Both solutions were combined and stirred at room temperature for 18 h. The solvent was removed under reduced pressure, and the crude product was dissolved in EtOAc (15 mL), washed with saturated NaHCO_3 solution (20 mL), brine (20 mL), and 10% KHSO_4 solution (20 mL), and dried over Na_2SO_4 . Purification using column chromatography with EtOAc as the solvent yielded **23** as an orange solid (22 mg, 0.033 mmol, 22%, mp 118–120 °C). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 0.83 ($t, J = 6.6$ Hz, 6H, $\text{CH}(\text{CH}_3)_2$), 1.45–1.60 (m, 3H, NHCHCH_2 , $\text{CH}(\text{CH}_3)_2$), 1.83–1.91 (m, 4H, $\text{N}(\text{CH}_2\text{CH}_2)_2$), 2.71 ($q, J = 7.0$ Hz, 4H, $\text{N}(\text{CH}_2)_2\text{CH}_2$), 3.84–3.90 (m, 2H, CH_2CHCH), 4.40–4.62 (m, 5H, NCH_2Ph , NCH_2Ph , NHCHCH_2), 6.60 ($dt, J = 14.8$ Hz, $J = 1.9$ Hz, 1H, CH_2CHCH), 6.78 ($dt, J = 15.0$ Hz, $J = 4.5$ Hz, 1H, CH_2CHCH), 7.11 (s, 1H, 5-H), 7.13–7.19 (m, 4H, NCH_2Ph), 7.21–7.26 (m, 2H, NCH_2Ph), 7.28–7.31 (m, 4H, NCH_2Ph), 8.34 (s, 1H, 4-H), 8.43 ($t, J = 5.7$ Hz, 1H, CONH), 8.87 ($d, J = 7.8$ Hz, 1H, CONH). One signal ($\text{N}(\text{CH}_2\text{CH}_2)_2$) is obscured by the H_2O peak. ^{13}C NMR (125 MHz, $\text{DMSO}-d_6$) δ 19.71 ($\text{N}(\text{CH}_2\text{CH}_2)_2$), 20.66 ($\text{N}(\text{CH}_2)_2\text{CH}_2$), 21.99, 22.95, 24.64 ($\text{CH}(\text{CH}_3)_2$), 26.92 ($\text{N}(\text{CH}_2)_2\text{CH}_2$), 41.77 (NHCHCH_2), 47.96, 49.17, 49.70, 49.82 ($\text{N}(\text{CH}_2)_2\text{CH}_2$), NCH_2Ph , NCH_2Ph), 51.71 (NHCHCH_2), 104.76 (C-4a), 107.43 (C-8), 107.58 (C-6), 119.60 (C-3), 119.89 (CH_2CHCH), 127.12, 127.17, 127.27, 127.49, 127.75 (NCH_2Ph , C-4), 128.55, 128.74, 137.37, 137.70 (NCH_2Ph), 143.16 (CH_2CHCH), 147.58 (C-5), 148.27 (C-7), 152.23 (C-8a), 162.05, 162.36, 165.77, 171.83 (C-2, CON , CONH); one signal (CH_2CHCH) is obscured by the DMSO solvent signal. LC–MS (ESI), t_R 12.63 min, m/z calcd for $[\text{C}_{40}\text{H}_{48}\text{N}_4\text{O}_5]^+$, 661.3, found, 661.4. HRMS m/z calcd for $[\text{C}_{40}\text{H}_{48}\text{N}_4\text{O}_5]^+$, 661.3377, found, 661.3384.

(*S*)-2-(4-(2,3,6,7-Tetrahydro-11-oxo-1H,5H,11H-[1]benzopyrano[6,7,8-ij]-10-quinolizinyloxy)benzoylamino)-N-(4-(dibenzylamino)-4-oxobut-2-enyl)-4-methylpentanamide (**24**). Compound **21** (79 mg, 0.16 mmol) was dissolved in CH_2Cl_2 (12 mL), treated with TFA (2 mL), and stirred at room temperature for 20 h. After evaporation of the solvent, the ammonium trifluoroacetate was dissolved in CH_2Cl_2 (10 mL) and added to a solution of the coumarin **10** (72 mg, 0.20 mmol), HBTU (76 mg, 0.20 mmol), and DIPEA (129 mg, 1.00 mmol) in CH_2Cl_2 (40 mL). After stirring at room temperature for 16 h, the reaction mixture was washed with saturated NaHCO_3 (2×30 mL). The organic layer was evaporated, and the residue was purified by column chromatography using $\text{CH}_2\text{Cl}_2/\text{MeOH}$ (25 + 1) to obtain **24** as an orange solid (58 mg, 0.08 mmol, 49%, mp 96–98 °C). ^1H NMR (600 MHz, $\text{DMSO}-d_6$) δ 0.82 (d, $J = 6.4$ Hz, $\text{CH}(\text{CH}_3)_2$), 0.84 ($d, J = 6.5$ Hz, 3H, $\text{CH}(\text{CH}_3)_2$), 1.44–1.50 (m, 1H, $\text{CH}(\text{CH}_3)_2$), 1.60–1.71 (m, 2H, NHCHCH_2), 1.85–1.93 (m, 4H, $\text{N}(\text{CH}_2\text{CH}_2)_2$), 2.72 ($t, J = 6.2$ Hz, 2H, $\text{N}(\text{CH}_2)_2\text{CH}_2$), 2.76 ($t, J = 6.5$ Hz, 2H, $\text{N}(\text{CH}_2)_2\text{CH}_2$), 3.25–3.30 (m, 4H, $\text{N}(\text{CH}_2\text{CH}_2)_2$), 3.85–3.89 (m, 2H, CH_2CHCH), 4.41–4.46 (m, 1H, NHCHCH_2), 4.46–4.48 (m, 2H, NCH_2Ph), 4.56 ($d, J = 16.8$ Hz, 1H, NCH_2Ph), 4.61 ($d, J = 16.9$ Hz, 1H, NCH_2Ph), 6.62 ($dt, J = 15.0$ Hz, $J = 1.7$ Hz, 1H, CH_2CHCH), 6.79 ($dt, J = 15.0$ Hz, $J = 4.6$ Hz, 1H, CH_2CHCH), 7.12 (s, 1H, 5-H), 7.17 ($t, J = 7.8$ Hz, 4H, NCH_2Ph), 7.26–7.35 (m, 6H, NCH_2Ph), 7.73 ($d, J = 8.5$ Hz, 2H, 3'-H), 7.85 ($d, J = 8.5$ Hz, 2H, 2'-H), 8.03 (s, 1H, 4-H), 8.30 ($t, J = 5.7$ Hz, 1H, CONH), 8.44 ($d, J = 7.7$ Hz, 1H, CONH). ^{13}C NMR (150 MHz, $\text{DMSO}-d_6$) δ 20.23, 20.45 ($\text{N}(\text{CH}_2\text{CH}_2)_2$), 21.34 ($\text{N}(\text{CH}_2)_2\text{CH}_2$), 21.88, 23.41, 24.94 ($\text{CH}(\text{CH}_3)_2$), 27.36 ($\text{N}(\text{CH}_2)_2\text{CH}_2$), 40.53 (CH_2CHCH), 40.74 (NHCHCH_2), 48.30, 49.34, 49.83, 50.16 ($\text{N}(\text{CH}_2\text{CH}_2)_2$, NCH_2Ph , NCH_2Ph), 52.67 (NHCHCH_2), 105.47 (C-4a), 108.65 (C-8), 117.12 (C-6), 118.96 (C-5), 120.14 (CH_2CHCH), 126.21 (C-3), 126.88, 127.42, 127.81, 127.85, 128.10, 128.50, 128.91, 129.14 ($\text{N}(\text{CH}_2\text{Ph})_2$, C-2', C-3'), 132.82 (C-1'), 137.78, 138.06 ($\text{N}(\text{CH}_2\text{Ph})_2$), 139.08 (C-4'), 142.38 (C-4), 143.80 (CH_2CHCH), 146.52 (C-7), 151.38 (C-8a), 160.86 (C-2), 166.22, 166.55, 172.77 (CON , CONH). LC–MS (ESI), t_R 13.07 min, m/z calcd for $[\text{C}_{46}\text{H}_{48}\text{N}_4\text{O}_5]^+$, 737.4, found, 737.5. HRMS m/z calcd for $[\text{C}_{46}\text{H}_{48}\text{N}_4\text{O}_5]^+$, 737.3697, found, 737.3704.

1-(5-((S)-1-((E)-4-(Dibenzylamino)-4-oxobut-2-enylamino)-4-methyl-1-oxopentan-2-ylamino)-5-oxopentyl)-3,3-dimethyl-2-((1E,3E,5E)-5-(1,3,3-trimethylindolin-2-ylidene)penta-1,3-dienyl)-3H-indolium chloride (25). ((S,E)-tert-Butyl-1-(4-(dibenzylamino)-4-oxobut-2-enylamino)-4-methyl-1-oxopentan-2-ylcarbamate (21, 84 mg, 0.17 mmol) was added to a solution of TFA (1 mL) in CH₂Cl₂ (6 mL) and stirred for 20 h. Solvents were removed, and the crude brownish resin was dissolved in dry CH₂Cl₂ (2 mL) with DIPEA (0.08 g, 0.62 mmol). A solution of 16 (0.08 g, 0.16 mmol), HATU (0.10 g, 0.28 mmol), and DIPEA (0.06 g, 0.45 mmol) in dry CH₂Cl₂ (4 mL) was prepared, flooded with argon, and stirred for 10 min. Both solutions were combined and stirred for 18 h at room temperature. Solvents were removed, the crude product was dissolved in EtOAc (10 mL) and washed with saturated NaHCO₃ solution (2 × 20 mL), brine (20 mL), and 10% KHSO₄ solution (20 mL), and the organic phase was dried over Na₂SO₄. Purification by column chromatography (CH₂Cl₂/MeOH, 59 + 1 gradient to 19 + 1) led to 25 as a red shining resin (51 mg, 0.06 mmol, 39%). ¹H NMR (600 MHz, DMSO-*d*₆) δ 0.70 (d, ³J = 6.5 Hz, 3H, CHCH₃), 0.75 (d, ³J = 6.6 Hz, 3H, CHCH₃), 1.34–1.40 (m, 2H, CH₂CH₂CONH), 1.42–1.49 (m, 1H, CH(CH₂)₂), 1.49–1.58 (m, 2H, NHCHCH₂), 1.57–1.64 (m, 2H, CH₂(CH₂)₂CONH), 1.65–1.72 (m, 12H, 3-CH₃, 3'-CH₃), 2.02 (dt, ²J = 14.0 Hz, ³J = 6.9 Hz, 1H, CH₂CONH), 2.15 (dt, ²J = 14.2 Hz, ³J = 7.1 Hz, 1H, CH₂CONH), 3.59 (s, 3H, NCH₃), 3.76–3.90 (m, 2H, CH₂CHCH), 4.05 (t, ³J = 7.2 Hz, 2H, CH₂(CH₂)₂CONH), 4.16–4.23 (m, 1H, NHCHCH₂), 4.43–4.51 (m, 2H, NCH₂Ph), 4.53–4.61 (m, 2H, NCH₂Ph), 6.24 (d, ³J = 13.8 Hz, 1H, CH(CH₂)₂CH), 6.28 (d, ³J = 13.7 Hz, 1H, CH(CH₂)₂CH), 6.50–6.61 (m, 2H, CH₂CHCH, (CH₂)₂CH(CH₂)), 6.74–6.80 (m, 1H, CH₂CHCH), 7.17–7.20 (m, 4H, NCH₂Ph), 7.21–7.45 (m, 12H, 5-H, 6-H, 7-H, 5'-H, 6'-H, 7'-H, NCH₂Ph), 7.56–7.67 (m, 2H, 4-H, 4'-H), 7.92 (d, ³J = 7.9 Hz, 1H, CONH), 8.23 (dd, ²J = 8.8 Hz, ³J = 5.7 Hz, 1H, CONH), 8.28–8.36 (m, 2H, CHCH(CH₂)). ¹³C NMR (150 MHz, DMSO-*d*₆) δ 21.34, 22.34, 22.81, 24.20 (CH(CH₂)₂, CH₂CH₂CONH), 26.43 (CH₂(CH₂)₂CONH), 26.98, 27.14 (3-CH₃, 3'-CH₃), 31.04 (NCH₃), 34.63 (CH₂CONH), 40.78 (NHCHCH₂), 43.06 (CH₂(CH₂)₂CONH), 47.88, 48.83, 48.85, 49.64 (C-3, C-3', NCH₂Ph, NCH₂Ph), 51.04 (NHCHCH₂), 103.09, 103.21 (CH(CH₂)₂CH), 111.01 (C-7, C-7'), 119.71 (CH₂CHCH), 122.28, 122.39 (C-4, C-4'), 124.63, 124.73 (C-5, C-5'), 125.36 ((CH₂)₂CH(CH₂)), 126.88, 126.91, 127.06, 127.35, 127.64, 128.36, 128.42, 128.64 (NCH₂Ph, C-6, C-6'), 137.28, 137.56 (NCH₂Ph), 140.98, 141.05 (C-3a, C-3'a), 141.99 (CH₂CHCH), 142.75, 143.18 (C-7a, C-7'a), 153.95, 154.02 (CHCH(CH₂)), 165.66, 171.82, 172.19, 172.54, 173.22 (C-2, C-2', CON, CONH). One signal (CH₂CHCH) is obscured by the DMSO solvent signal. LC-MS (ESI), *t*_R 12.09 min, *m/z* calcd for [C₅₃H₆₆N₆O₃]⁺, 844.5, found, 844.9. HRMS *m/z* calcd for [C₅₃H₆₆N₆O₃]⁺, 844.5166, found, 844.5239.

1-(9S,12S,E)-2-Benzyl-12-butyl-9-isobutyl-3,8,11,14-tetraoxo-1-phenyl-2,7,10,13-tetraazooctadec-4-en-18-yl)-3,3-dimethyl-2-((1E,3E,5E)-5-(1,3,3-trimethylindolin-2-ylidene)penta-1,3-dienyl)-3H-indolium chloride (26). tert-Butyl ((S)-1-((S)-1-((E)-4-(dibenzylamino)-4-oxobut-2-enylamino)-4-methyl-1-oxopentan-2-ylamino)-1-oxohexan-2-ylcarbamate (22, 0.15 g, 0.25 mmol) was added to a solution of TFA (2 mL) in CH₂Cl₂ (10 mL) and stirred for 20 h. Solvents were removed, and the crude brownish resin was dissolved in dry CH₂Cl₂ (8 mL) with DIPEA (0.10 g, 0.75 mmol). A solution of 16 (0.13 g, 0.25 mmol), HATU (0.16 g, 0.40 mmol), and DIPEA (0.11 g, 0.85 mmol) in dry CH₂Cl₂ (8 mL) was prepared, flooded with argon, and stirred for 30 min. Both solutions were combined and stirred for 68 h at room temperature. Solvents were removed, the crude product was dissolved in EtOAc (25 mL) and washed with saturated NaHCO₃ solution (2 × 25 mL), brine (25 mL), and 10% KHSO₄ solution (25 mL), and the organic phase was dried over Na₂SO₄. Purification by column chromatography (CH₂Cl₂/MeOH 29 + 1) led to 26 as a blue resin (0.06 g, 0.06 mmol, 23%). ¹H NMR (600 MHz, DMSO-*d*₆) δ 0.72–0.80 (m, 9H, CH(CH₂)₂CH₃, CH(CH₂)₂), 1.12–1.20 (m, 5H, CHCH₂(CH₂)₂CH₃), 1.39–1.43 (m, 2H, CHCH₂(CH₂)₂CH₃, CH(CH₂)₂), 1.47–1.52 (m, 2H,

CH₂CH(CH₂)₂), 1.58–1.63 (m, 2H, CH₂(CH₂)₂CONH), 1.68 (s, 6H, 3-CH₃ or 3'-CH₃), 1.68 (s, 6H, 3-CH₃ or 3'-CH₃), 1.96–2.04 (m, 1H, CH₂CONH), 2.13–2.25 (m, 3H, CH₂CONH, CH₂CH₂CONH), 3.60 (s, 3H, NCH₃), 3.76–3.83 (m, 1H, CH₂CHCH), 3.84–3.91 (m, 1H, CH₂CHCH), 4.09 (t, ²J = 7.3 Hz, 2H, CH₂(CH₂)₂CONH), 4.12–4.18 (m, 1H, NHCHCH₂CH), 4.18–4.24 (m, 1H, CH₂(CH₂)₂), 4.43–4.62 (m, 4H, NCH₂Ph), 6.25 (d, ³J = 13.8 Hz, 1H, CH(CH₂)₂CH), 6.32 (d, ³J = 13.7 Hz, 1H, CH(CH₂)₂CH), 6.53–6.59 (m, 2H, (CH₂)₂CH(CH₂), CH₂CHCH), 6.75 (dt, ²J = 15.1 Hz, ³J = 4.9 Hz, 1H, CH₂CHCH), 7.17–7.36 (m, 12H, 5-H, 5'-H, NCH₂Ph), 7.37–4.41 (m, 4H, 6-H, 7-H, 6'-H, 7'-H), 7.61 (dd, ²J = 7.3, ³J = 2.8 Hz, 2H, 4-H, 4'-H), 7.83 (d, ³J = 8.0 Hz, 1H, CONH), 7.88 (d, ³J = 7.7 Hz, 1H, CONH), 8.14 (t, ²J = 5.7 Hz, 1H, CONH), 8.33 (t, ²J = 13.0 Hz, 2H, CHCH(CH₂)₂). ¹³C NMR (150 MHz, DMSO-*d*₆) δ 13.98 (CH(CH₂)₂CH₃), 21.63, 21.92, 22.22, 22.57, 22.99, 24.32 ((CH₂)₂CH₂CONH, CH(CH₂)₂CH₂CH, CH(CH₂)₂), 27.14, 27.31 (3-CH₃, 3'-CH₃), 27.58 (CHCH₂CH₂CH₂CH₃), 31.20 (NCH₃), 31.63 (CHCH₂(CH₂)₂CH₃), 34.77 (CH₂CONH), 40.90 (CH₂CH(CH₂)₂), 43.27 (CH₂(CH₂)₂CONH), 48.06 (C-3 or C-3'), 49.00 (NCH₂Ph), 49.83 (C-3 or C-3'), 51.24 (NHCHCH₂CH), 52.74 (CH(CH₂)₂), 103.22, 103.40 (CH(CH₂)₂CH), 111.18 (C-7, C-7'), 120.10 (CH₂CHCH), 122.44, 122.56 (C-4, C-4'), 124.78, 124.88 (C-5, C-5'), 125.52 ((CH₂)₂CH(CH₂)), 127.05, 127.21, 127.51, 127.79, 128.51, 128.55, 128.80 (NCH₂Ph, C-6, C-6'), 137.43, 137.72 (NCH₂Ph), 141.14, 141.21, 142.14, 142.91 (C-3a, C-3'a, C-7a, C-7'a), 143.09 (CH₂CHCH), 154.14, 154.21 (CHCH(CH₂)₂), 165.82, 171.87, 171.96, 172.11, 172.69, 173.40 (C-2, C-2', CON, CONH). One signal (CH₂CHCH) is obscured by the solvent signal. LC-MS (ESI), *t*_R 12.64 min, *m/z* calcd for [C₆₁H₇₇N₆O₄]⁺, 957.6, found, 958.1. HRMS *m/z* calcd for [C₆₁H₇₇N₆O₄]⁺, 957.6001, found, 957.6105.

Enzymatic Assays. Spectrophotometric assays were carried out on a Varian Cary 50 Bio or 100 Bio device (Agilent) in cuvettes. Fluorometric assays were performed on a FLUOstar Optima plate reader (BMG Labtech) in 96 well plates. Human cathepsin B was purchased from Calbiochem (Merck), and human cathepsins K, L, and S were from Enzo Life Sciences. All substrates were purchased from Bachem. The inhibitor stock solutions were prepared in DMSO. The residual activity was determined in duplicates after 1 h of measurements via endpoint determination and subsequent comparison to uninhibited reactions. Monitoring of enzymatic reactions was performed for 60 min in duplicate with six different inhibitor concentrations. Pseudofirst order rate constants *k*_{obs} were determined via nonlinear regression using the eq [P] = *v*_i × (1 - exp(-*k*_{obs} × *t*)) / *k*_{obs} + *d*, where [P] is the product concentration, *v*_i is the initial product formation rate, *t* is the time, and *d* is the offset. The mean of the *k*_{obs} values was plotted versus the inhibitor concentrations [I], and a second nonlinear regression was performed to obtain second-order rate constants *k*_{inac}/*K*_i using the equation *k*_{obs} = *k*_{inac} × [I] / (*K*_i + [I]), with *k*_{inac} as the first-order rate constant of inactivation and *K*_i as the inhibitory constant. *K*_i values were additionally corrected by means of the Cheng-Prusoff equation.

Cathepsin K. The enzyme stock solution (23 μg/mL in 50 mM sodium acetate buffer pH 5.5, 50 mM NaCl, and 0.5 mM EDTA) was diluted 1:100 with a 100 mM sodium citrate buffer pH 5.0 containing 100 mM NaCl, 1 mM EDTA, 0.01% (w/v) CHAPS, and 5 mM DTT, incubated at 37 °C for 30 min, and then stored on ice. A 10 mM solution of the substrate Z-Phe-Arg-pNA was prepared in DMSO. Into a cuvette containing 940 μL assay buffer (100 mM sodium phosphate buffer pH 6.0, 100 mM NaCl, 5 mM EDTA, and 0.01% (w/v) Brij 35), the chromogenic substrate (10 μL) and DMSO and/or inhibitor solution (10 μL) were pipetted. Upon addition of cathepsin K (40 μL), the measurement was started and followed at 37 °C for 60 min. The final enzyme concentration was 9.2 ng/mL, the final substrate concentration was 100 μM (= 0.85 × *K*_m), and the total DMSO concentration was 2% (v/v). The corresponding *K*_m value of 118 μM was determined in duplicate measurements with 10 different substrate concentrations (10–200 μM). The fluorometric assay was performed in 96-well plates as described elsewhere using Z-Leu-Arg-AMC (40 μM = 13.33 × *K*_m).¹²

Cathepsin S. The enzyme stock solution (70 $\mu\text{g}/\text{mL}$ in 100 mM MES buffer pH 6.5, 1 mM EDTA, 50 mM L-cysteine, 10 mM DTT, 0.5% (w/v) Triton X-100, and 30% (w/v) glycerol) was diluted 1:100 with a 100 mM sodium citrate buffer pH 6.0 containing 50 mM NaCl, 2 mM EDTA, 5 mM DTT, and 0.01% (w/v) Triton-X, incubated at 37 °C for 60 min, and then stored on ice. A 10 mM solution of the substrate Z-Phe-Val-Arg-pNA was prepared in DMSO. Into a cuvette containing 940 μL assay buffer (100 mM sodium phosphate buffer pH 6.0, 100 mM NaCl, 5 mM EDTA, and 0.01% (w/v) Brij 35), the chromogenic substrate (7 μL) and DMSO and/or inhibitor solution (13 μL) were pipetted. Upon addition of cathepsin S (40 μL), the measurement was started and followed at 37 °C for 60 min. The final enzyme concentration was 28 ng/mL, the final substrate concentration was 70 μM ($= 0.58 \times K_m$), and the total DMSO concentration was 2% (v/v). The corresponding K_m value of 121 μM was determined in duplicate measurements with eight different substrate concentrations (2–100 μM).

Cathepsin L.⁴² The enzyme stock solution (0.14 mg/mL in 20 mM malonate buffer pH 5.5, 400 mM NaCl, and 1 mM EDTA) was diluted 1:100 with assay buffer (100 mM sodium phosphate buffer pH 6.0, 100 mM NaCl, 5 mM EDTA, and 0.01% (w/v) Brij 35) containing 5 mM DTT, incubated at 37 °C for 30 min, and then stored on ice. A 10 mM solution of the substrate Z-Phe-Arg-pNA was prepared in DMSO. Into a cuvette containing 940 μL assay buffer, the chromogenic substrate (10 μL) and DMSO and/or inhibitor solution (10 μL) were pipetted. Upon addition of cathepsin L (40 μL), the measurement was started and followed at 37 °C for 60 min. The final enzyme concentration was 56 ng/mL, the final substrate concentration was 100 μM ($= 5.88 \times K_m$), and the total DMSO concentration was 2% (v/v).

Cathepsin B.⁴² The enzyme stock solution (0.47 mg/mL in 20 mM sodium acetate buffer pH 5.0 and 1 mM EDTA) was diluted 1:500 with assay buffer (100 mM sodium phosphate buffer pH 6.0, 100 mM NaCl, 5 mM EDTA, and 0.01% (w/v) Brij 35) containing 5 mM DTT, incubated at 37 °C for 30 min, and then stored on ice. A 100 mM solution of the substrate Z-Arg-Arg-pNA was prepared in DMSO. Into a cuvette containing 960 μL assay buffer, the chromogenic substrate (5 μL) and DMSO and/or inhibitor solution (15 μL) were pipetted. Upon addition of cathepsin B (20 μL), the measurement was started and followed at 37 °C for 60 min. The final enzyme concentration was 19 ng/mL, the final substrate concentration was 500 μM ($= 0.45 \times K_m$), and the total DMSO concentration was 2% (v/v).

Preparation of Cathepsin K-Ligand Complexes, Protein Crystallization, and Data Collection. Human cathepsin K zymogen (Uniprot accession number P43235) was expressed in the yeast *Pichia pastoris*, autoactivated, and purified as described previously.⁴³ The active mature cathepsin K (0.5 mg/mL) was incubated with a 5-fold molar excess of the ligand 21 or 25 in 100 mM sodium acetate, pH 4.0, 2.5 mM DTT, 30 mM NaCl, and 1 mM EDTA for 3 h at 4 °C under an argon atmosphere. The enzyme inhibition was monitored using a kinetic assay. The complexes were buffer-exchanged into 20 mM sodium acetate, pH 4.0, 2.5 mM DTT, 300 mM NaCl using an Amicon Ultracel-10 k centrifugal filter unit (Millipore) and concentrated to 2.3 mg/mL for cathepsin K-21 and 1.1 mg/mL for cathepsin K-25. Crystals were grown by vapor diffusion in hanging drops at 20 °C. The drops consisting of 1 μL of protein solution and 1 μL of reservoir solution were equilibrated over 350 μL of reservoir solution consisting of 100 mM MES/imidazole pH 6.5, 12.5% (w/v) PEG 1000, 12.5% (w/v) PEG 3350, 12.5% (v/v) 2-methyl-2,4-pentanediol, 30 mM magnesium chloride, and 30 mM calcium chloride. The crystals were flash-cooled by plunging into liquid nitrogen, and diffraction data were collected at 100 K on a MicroMax-007 HF Microfocus rotating anode X-ray generator equipped with a PILATUS 300 K detector (Rigaku). All diffraction data were processed using the XDS suite of programs.⁴⁴ Crystal parameters, data collection statistics, and final refinement statistics are given in Table S1.

Structure Determination, Refinement, and Analysis. The structures of cathepsin K-21 and cathepsin K-25 complexes were

determined by molecular replacement with program Molrep⁴⁵ from the CCP4 package⁴⁶ using the structure of human cathepsin K (PDB code 5TDI) as a search model. Model refinement was carried out using the program REFMAC 5.5⁴⁶ interspersed with manual adjustments using Coot.⁴⁷ The geometric restraints for ligands were constructed using the program AceDRG.⁴⁸ The structures were refined using data to a resolution of 1.80 and 1.72 Å for cathepsin K-21 and cathepsin K-25, respectively. The final crystallographic models contain cathepsin K residues Ala1 to Met215 (mature cathepsin K numbering) and the N-terminal Arg0 residue (corresponding to the P99 residue of the activation peptide). The quality of the final models was validated with Molprobtly.⁴⁹ The final refinement statistics are given in Table S1. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with the accession codes 7NXL and 7NXM for cathepsin K-21 and cathepsin K-25, respectively. Ligand interactions were analyzed using the server PLIP⁵⁰ and program CONTACT.⁴⁶ All figures showing structural presentations were prepared using the program PyMOL.⁵¹

Preparation of the HEK Cell Lysate. The HEK 293 cell lysate was obtained as described elsewhere.⁵² HEK cells were cultivated in Dulbecco's modified Eagle's medium containing fetal bovine serum (FBS) and penicillin/streptomycin at a concentration of 100 mg/mL each at 37 °C. When a ~70% confluence was reached, the culture medium was discarded. The cells were then washed twice with phosphate buffered saline (PBS). After further incubation at 37 °C in Minimal Essential Medium (Opti-MEM), the supernatant was discarded, and the cells were mechanically lysed in PBS. After a short centrifugation at 2000 g and 4 °C, the supernatant was collected, and the protein concentration was determined using Roti-Nanoquant Bradford solution (Carl Roth) and subsequently stored at –20 °C.

SDS-PAGE Analyses and Imaging of Cathepsin K with ABPs 24 and 25. Stock solutions of recombinant cathepsin K at a concentration of 2.5 mg/mL in 100 mM sodium acetate buffer pH 5.5 containing 150 mM NaCl were used and correspondingly diluted. Stock solutions (2 mM) of E-64 and K11777 and 0.5 mM solutions of ABP 24 and 25 were prepared in DMSO. All incubations (see below) were performed in a total volume of 13 μL in the assay buffer (100 mM sodium acetate buffer pH 5.5, 1 mM EDTA, and 2.5 mM DTT) with a probe concentration of 5 μM . The reactions were stopped by addition of SDS-PAGE loading buffer (7 μL) and heating at 75 °C for 10 min. The protein samples were separated on 4–12% bis-tris polyacrylamide gels with MES running buffer (Bolt series from Invitrogen). The gels were visualized with a Typhoon RGB imager (GE Healthcare Life Sciences) using excitation at 488 nm and emission at 515 nm (long pass filter) for ABP 24 and excitation at 635 nm and emission at 660 nm (long pass filter) for ABP 25, respectively, and subsequently stained with CBB.

Competition and Sensitivity Experiments. Cathepsin K (1,250 ng in 10 μL of assay buffer) was treated with 3 μL of the probe 24 or 25 (final concentration 5 μM) at 25 °C for 60 min in the dark. The competitive labeling was performed after preincubation of cathepsin K with E-64 and K11777 (final concentration 10 μM each) in 10 μL of the assay buffer at 37 °C for 15 min. For the estimation of the detection limit of probes 24 and 25, cathepsin K (5–1250 ng) was treated with probe 24 or 25 (final concentration 5 μM) at 25 °C for 60 min in the dark.

HEK Cell Spiking. Cathepsin K (300 ng) was incubated with the HEK293 cell lysate (30 μg), BSA (500 ng) and the probe 24 or 25 (final concentration 5 μM) at 25 °C for 60 min in the dark. In the control experiment, cathepsin K (300 ng), HEK293-cell lysate (30 μg), and BSA (500 ng) were separately treated with the probes 24 or 25 (final concentration: 5 μM).

Osteosarcoma Cell Imaging. Human bone osteosarcoma cells, U-2 OS, were cultivated in McCoy's 5A medium supplemented with 10% (v/v) FBS, 2 mM glutamax (L-alanyl-L-glutamine), 100 units/mL penicillin, and 100 $\mu\text{g}/\text{mL}$ streptomycin at 37 °C in 5% CO₂. Fifty thousand cells were seeded into a well of 12-well plate and allowed to attach. After 24 h, 1 μM probe 25 was added, and cells were incubated at 37 °C in 5% CO₂. At different time points (0 to 24 h), cells were detached using 0.25% (w/v) trypsin–EDTA solution

(Sigma-Aldrich, Merck), collected by centrifugation (500 g for 5 min), washed with PBS, resuspended in 100 μ L of SDS-PAGE loading buffer, and heated at 100 $^{\circ}$ C for 10 min. The competitive labeling was performed after preincubation of cells with 1 μ M E-64d for 1 h. The samples (30 μ L) were separated on 4–12% bis-tris polyacrylamide gels (Thermo Fisher Scientific). The gels were visualized using a Typhoon RGB imager (GE Healthcare Life Sciences) with excitation at 635 nm and emission at 660 nm (long pass filter) or transferred onto a PVDF membrane. The membrane was blocked for 1 h with 5% BSA in TTBS, washed with TTBS, and incubated for 1 h with anticathepsin K mouse monoclonal antibody (Santa Cruz Biotechnology) diluted 1:200 in TTBS. The membrane was incubated for 1 h with goat HRP-conjugated antimouse IgG antibody (Sigma-Aldrich, Merck) at a dilution of 1:10,000, developed with Luminata Forte Western HRP substrate (Merck), and imaged using an ImageQuant LAS 4000 biomolecular imager (GE Healthcare Life Sciences). For live cell microscopy, 33,000 cells were seeded in a 35 mm glass bottom dish and cultured overnight. The cells were then incubated with 1 μ M ABP 25 for 24 h. LysoTracker Green (50 nM, Invitrogen) was added to the cells for the last 2 h to label late endosomes and lysosomes, and 1 μ M Hoechst dye (Thermo Fisher Scientific) was added for the last 15 min to label nuclei. The competitive labeling was performed after preincubation of cells with 1 μ M E-64d for 1 h. Images were captured on a Zeiss LSM 980 Airyscan 2 confocal microscope using a 40 \times water objective (C-Apochromat 40x/1.20 W Corr). Hoechst dye, LysoTracker Green, and Cy5 signals were detected using lasers with excitation wavelengths of 405, 488, and 639 nm, respectively. In all cases, appropriate beamsplitters were used (MBS 405 for the 405 nm laser line and MBS 488/639 for 488 and 639 nm laser lines), and emission was collected from 300 to 720 nm. Channels were captured sequentially. All images were acquired using ZEN software and processed using Fiji software.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jmedchem.1c01178>.

Kinetic evaluation of precursor 22 with cathepsin K, X-ray data collection, and refinement statistics, list of contacts formed between cathepsin K and ligands, fluorescence detection of human cathepsin K with ABP 24 in the absence and presence of the cell lysate, selectivity investigations with ABP 25, and HPLC purity traces and NMR spectra of test compounds (PDF)

Molecular formula strings (CSV)

Accession Codes

Atomic coordinates and experimental structure factors have been deposited in the Protein Data Bank with accession codes 7NXL and 7NXM for cathepsin K-21 and cathepsin K-25 complexes, respectively.

■ AUTHOR INFORMATION

Corresponding Authors

Michael Mareš – *Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 16610, Czech Republic*; orcid.org/0000-0002-0847-5022; Phone: (+42) 220-183358; Email: mares@uochb.cas.cz

Michael Gütschow – *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, Bonn 53121, Germany*; orcid.org/0000-0002-9376-7897; Phone: (+49)228-732317; Email: guetschow@uni-bonn.de

Authors

Carina Lemke – *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, Bonn 53121, Germany*

Jakub Benýšek – *Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 16610, Czech Republic*; *First Faculty of Medicine, Charles University, Prague 12108, Czech Republic*

Dominik Brajtenbach – *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, Bonn 53121, Germany*

Christian Breuer – *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, Bonn 53121, Germany*; *Department of Natural Sciences, University of Applied Sciences Bonn-Rhein-Sieg, Rheinbach 53359, Germany*

Adéla Jilková – *Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 16610, Czech Republic*

Martin Horn – *Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 16610, Czech Republic*; orcid.org/0000-0001-9110-2018

Michal Buša – *Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 16610, Czech Republic*; *Department of Biochemistry, Faculty of Science, Charles University, Prague 12800, Czech Republic*

Lenka Ulrychová – *Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague 16610, Czech Republic*

Annika Illies – *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, Bonn 53121, Germany*

Katharina F. Kubatzky – *Department of Infectious Diseases, Medical Microbiology and Hygiene, Heidelberg University Hospital, Heidelberg 69120, Germany*

Ulrike Bartz – *Department of Natural Sciences, University of Applied Sciences Bonn-Rhein-Sieg, Rheinbach 53359, Germany*

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jmedchem.1c01178>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

C.L. thanks the Bonn International Graduate School of Drug Sciences (BIGS-DrugS)/BMBF for obtaining a grant (16GW0025). C.B. was supported by a grant from the Department of Natural Sciences, University of Applied Sciences Bonn-Rhein-Sieg. M.M., J.B., A.J., M.H., M.B., and L.U. were supported by the project ChemBioDrug CZ.02.1.01/0.0/0.0/16_019/0000729 from ERDF/OPRDE, Gilead Sciences & IOCB Research Center, and the institutional project RVO 61388963. The authors thank Marion Schneider, Sabine Terhart-Krabbe, Dr. Anna-Christina Schulz-Fincke, Michaela Hympánová, and Rabea Voget for support, Helena Mertlíková-Kaiserová for providing U-2 OS cells, and Rebecca May for support in artwork design.

■ ABBREVIATIONS USED

ABP, activity-based probe; BSA, bovine serum albumin; CatK, cathepsin K; CBB, Coomassie brilliant blue staining; Cy5,

cyanine 5; HEK cells, human embryonic kidney cells; NFATc1, nuclear factor-activated T cells c1; PSBE, primed-site binding element; RANKL, receptor activator of nuclear factor kappa-B ligand; U-2 OS cells, human bone osteosarcoma epithelial cells; WB, western blot.

REFERENCES

- (1) Kramer, L.; Turk, D.; Turk, B. The Future of Cysteine Cathepsins in Disease Management. *Trends Pharmacol. Sci.* **2017**, *38*, 873–898.
- (2) Dai, R.; Wu, Z.; Chu, H. Y.; Lu, J.; Lyu, A.; Liu, J.; Zhang, G. Cathepsin K: The Action in and Beyond Bone. *Front. Cell Dev. Biol.* **2020**, *8*, 433.
- (3) Gruber, R. Molecular and Cellular Basis of Bone Resorption. *Wien. Med. Wochenschr.* **2015**, *165*, 48–53.
- (4) Li, Z.; Kienetz, M.; Cherney, M. M.; James, M. N.; Brömme, D. The Crystal and Molecular Structures of a Cathepsin K:Chondroitin Sulfate Complex. *J. Mol. Biol.* **2008**, *383*, 78–91.
- (5) Khosla, S.; Hofbauer, L. C. Osteoporosis Treatment: Recent Developments and Ongoing Challenges. *Lancet Diabetes Endocrinol.* **2017**, *5*, 898–907.
- (6) (a) Frizler, M.; Stirnberg, M.; Sisay, M. T.; Gütschow, M. Development of Nitrile-based Peptidic Inhibitors of Cysteine Cathepsins. *Curr. Top. Med. Chem.* **2010**, *10*, 294–322. (b) Borišek, J.; Vizovišek, M.; Sosnowski, P.; Turk, B.; Turk, D.; Mohar, B.; Novič, M. Development of *N*-(Functionalized benzoyl)-homocycloleucylglycinitriles as Potent Cathepsin K Inhibitors. *J. Med. Chem.* **2015**, *58*, 6928–6937.
- (7) (a) Löser, R.; Frizler, M.; Schilling, K.; Gütschow, M. Azadipeptide Nitriles: Highly Potent and Proteolytically Stable Inhibitors of Papain-like Cysteine Proteases. *Angew. Chem. Int. Ed.* **2008**, *47*, 4331–4334. (b) Yang, P. Y.; Wang, M.; Li, L.; Wu, H.; He, C. Y.; Yao, S. Q. Design, Synthesis and Biological Evaluation of Potent Azadipeptide Nitrile Inhibitors and Activity-based Probes as Promising anti-*Trypanosoma brucei* Agents. *Chem. Eur. J.* **2012**, *18*, 6528–6541. (c) Breidenbach, J.; Lemke, C.; Pillaiyar, T.; Schäkel, L.; Al Hamwi, G.; Dieltz, M.; Gedtschold, R.; Geiger, N.; Lopez, V.; Mirza, S.; Namasivayam, V.; Schiedel, A. C.; Sylvester, K.; Thimm, D.; Vielmuth, C.; Vu, P. L.; Zylina, M.; Bodem, J.; Gütschow, M.; Müller, C. E. Targeting the Main Protease of SARS-CoV-2: From the Establishment of High Throughput Screening to the Design of Tailored Inhibitors. *Angew. Chem. Int. Ed.* **2021**, *60*, 10423–10429.
- (8) Schmitz, J.; Beckmann, A. M.; Dudic, A.; Li, T.; Sellier, R.; Bartz, U.; Gütschow, M. 3-Cyano-3-aza- β -amino Acid Derivatives as Inhibitors of Human Cysteine Cathepsins. *ACS Med. Chem. Lett.* **2014**, *5*, 1076–1081.
- (9) (a) Löser, R.; Pietzsch, J. Cysteine Cathepsins: Their Role in Tumor Progression and Recent Trends in the Development of Imaging Probes. *Front. Chem.* **2015**, *3*, 37. (b) Vizovišek, M.; Vidmar, R.; Drag, M.; Fonović, M.; Salvesen, G. S.; Turk, B. Protease Specificity: Towards *In Vivo* Imaging Applications and Biomarker Discovery. *Trends Biochem. Sci.* **2018**, *43*, 829–844.
- (10) Dana, D.; Pathak, S. K. A Review of Small Molecule Inhibitors and Functional Probes of Human Cathepsin L. *Molecules* **2020**, *25*, 698.
- (11) (a) Ettari, R.; Nizi, E.; Di Francesco, M. E.; Dude, M. A.; Pradel, G.; Vicik, R.; Schirmeister, T.; Micale, N.; Grasso, S.; Zappalà, M. Development of Peptidomimetics with a Vinyl Sulfone Warhead as Irreversible Falcipain-2 Inhibitors. *J. Med. Chem.* **2008**, *51*, 988–996. (b) Frizler, M.; Yampolsky, I. V.; Baranov, M. S.; Stirnberg, M.; Gütschow, M. Chemical Introduction of the Green Fluorescence: Imaging of Cysteine Cathepsins by an Irreversibly Locked GFP Fluorophore. *Org. Biomol. Chem.* **2013**, *11*, 5913–5921.
- (12) Mertens, M. D.; Schmitz, J.; Horn, M.; Furtmann, N.; Bajorath, J.; Mareš, M.; Gütschow, M. A Coumarin-labeled Vinyl Sulfone as Tripeptidomimetic Activity-based Probe for Cysteine Cathepsins. *ChemBioChem* **2014**, *15*, 955–959.
- (13) Breuer, C.; Lemke, C.; Schmitz, J.; Bartz, U.; Gütschow, M. Synthesis and Kinetic Evaluation of Ethyl Acrylate and Vinyl Sulfone Derived Inhibitors for Human Cysteine Cathepsins. *Bioorg. Med. Chem. Lett.* **2018**, *28*, 2008–2012.
- (14) Dana, D.; De, S.; Rathod, P.; Davalos, A. R.; Novoa, D. A.; Paroly, S.; Torres, V. M.; Afzal, N.; Lankalapalli, R. S.; Rotenberg, S. A.; Chang, E. J.; Subramaniam, G.; Kumar, S. Development of a Highly Potent, Selective, and Cell-Active Inhibitor of Cysteine Cathepsin L - A Hybrid Design Approach. *Chem. Commun.* **2014**, *50*, 10875–10878.
- (15) Ettari, R.; Micale, N.; Schirmeister, T.; Gelhaus, C.; Leippe, M.; Nizi, E.; Di Francesco, M. E.; Grasso, S.; Zappalà, M. Novel Peptidomimetics Containing a Vinyl Ester Moiety as Highly Potent and Selective Falcipain-2 Inhibitors. *J. Med. Chem.* **2009**, *52*, 2157–2160.
- (16) (a) Previti, S.; Ettari, R.; Cosconati, S.; Amendola, G.; Chouchene, K.; Wagner, A.; Hellmich, U. A.; Ulrich, K.; Krauth-Siegel, R. L.; Wich, P. R.; Schmid, I.; Schirmeister, T.; Gut, J.; Rosenthal, P. J.; Grasso, S.; Zappalà, M. Development of Novel Peptide-Based Michael Acceptors Targeting Rhodocytin and Falcipain-2 for the Treatment of Neglected Tropical Diseases (NTDs). *J. Med. Chem.* **2017**, *60*, 6911–6923. (b) Ettari, R.; Previti, S.; Maiorana, S.; Amendola, G.; Wagner, A.; Cosconati, S.; Schirmeister, T.; Hellmich, U. A.; Zappalà, M. Optimization Strategy of Novel Peptide-Based Michael Acceptors for the Treatment of Human African Trypanosomiasis. *J. Med. Chem.* **2019**, *62*, 10617–10629.
- (17) (a) Ma, Y.; Li, L.; He, S.; Shang, C.; Sun, Y.; Liu, N.; Meek, T. D.; Wang, Y.; Shang, L. Application of Dually Activated Michael Acceptor to the Rational Design of Reversible Covalent Inhibitor for Enterovirus 71 3C Protease. *J. Med. Chem.* **2019**, *62*, 6146–6162. (b) Chenna, B. C.; Li, L.; Mellott, D. M.; Zhai, X.; Siqueira-Neto, J. L.; Calvet Alvarez, C.; Bernatchez, J. A.; Desormeaux, E.; Alvarez Hernandez, E.; Gomez, J.; McKerrow, J. H.; Cruz-Reyes, J.; Meek, T. D. Peptidomimetic Vinyl Heterocyclic Inhibitors of Cruzain Effect Antitrypanosomal Activity. *J. Med. Chem.* **2020**, *63*, 3298–3316.
- (18) (a) Conroy, T.; Guo, J. T.; Elias, N.; Cergol, K. M.; Gut, J.; Legac, J.; Khatoon, L.; Liu, Y.; McGowan, S.; Rosenthal, P. J.; Hunt, N. H.; Payne, R. J. Synthesis of Gallinamide A Analogues as Potent Falcipain Inhibitors and Antimalarials. *J. Med. Chem.* **2014**, *57*, 10557–10563. (b) Boudreau, P. D.; Miller, B. W.; McCall, L.; Almaliti, J.; Reher, R.; Hirata, K.; Le, T.; Siqueira-Neto, J. L.; Hook, V.; Gerwick, W. H. Design of Gallinamide A Analogs as Potent Inhibitors of the Cysteine Proteases Human Cathepsin L and *Trypanosoma cruzi* Cruzain. *J. Med. Chem.* **2019**, *62*, 9026–9044.
- (19) Ashhurst, A. S.; Tang, A. H.; Fajtová, P.; Yoon, M.; Aggarwal, A.; Stoye, A.; Larence, M.; Beretta, L.; Drellich, A.; Skinner, D.; Li, L.; Meek, T. D.; McKerrow, J. H.; Hook, V.; Tseng, C. K.; Turville, S.; Gerwick, W. H.; O'Donoghue, A. J.; Payne, R. J. Potent *In Vitro* Anti-SARS-CoV-2 Activity by Gallinamide A and Analogues Via Inhibition of Cathepsin L. *bioRxiv* **2020**, No. 424111.
- (20) Fournier, J.; Chen, K.; Mailyan, A. K.; Jackson, J. J.; Buckman, B. O.; Emayan, K.; Yuan, S.; Rajagopalan, R.; Misialek, S.; Adler, M.; Blaesse, M.; Griessner, A.; Zakarian, A. Total Synthesis of Covalent Cysteine Protease Inhibitor *N*-Desmethyl Thalassospiramide C and Crystallographic Evidence for Its Mode of Action. *Org. Lett.* **2019**, *21*, 508–512.
- (21) Schmitz, J.; Gilberg, E.; Löser, R.; Bajorath, J.; Bartz, U.; Gütschow, M. Cathepsin B: Active Site Mapping with Peptidic Substrates and Inhibitors. *Bioorg. Med. Chem.* **2019**, *27*, 1–15.
- (22) Alves, M. F.; Puzer, L.; Cotrin, S. S.; Juliano, M. A.; Juliano, L.; Brömme, D.; Carmona, A. K. S3 to S3' Subsite Specificity of Recombinant Human Cathepsin K and Development of Selective Internally Quenched Fluorescent Substrates. *Biochem. J.* **2003**, *373*, 981–986.
- (23) Vizovišek, M.; Vidmar, R.; Van Quickenberghe, E.; Impens, F.; Andjelković, U.; Sobotić, B.; Stoka, V.; Gevaert, K.; Turk, B.; Fonović, M. Fast Profiling of Protease Specificity Reveals Similar Substrate Specificities for Cathepsins K, L and S. *Proteomics* **2015**, *15*, 2479–2490.

- (24) Choe, Y.; Leonetti, F.; Greenbaum, D. C.; Lecaille, F.; Bogoy, M.; Brömme, D.; Ellman, J. A.; Craik, C. S. Substrate Profiling of Cysteine Proteases Using a Combinatorial Peptide Library Identifies Functionally Unique Specificities. *J. Biol. Chem.* **2006**, *281*, 12824–12832.
- (25) Breidenbach, J.; Bartz, U.; Gütschow, M. Coumarin as a Structural Component of Substrates and Probes for Serine and Cysteine Proteases. *Biochim. Biophys. Acta Proteins Proteomics* **2020**, *1868*, No. 140445.
- (26) Li, Y.; Zhou, Y.; Yue, X.; Dai, Z. Cyanine Conjugate-Based Biomedical Imaging Probes. *Adv. Healthc. Mater.* **2020**, *9*, No. 2001327.
- (27) Kvach, M. V.; Ustinov, A. V.; Stepanova, I. A.; Malakhov, A. D.; Skorobogatyi, M. V.; Shmanai, V. V.; Korshun, V. A. A Convenient Synthesis of Cyanine dyes: Reagents for the Labeling of Biomolecules. *Eur. J. Org. Chem.* **2008**, 2107–2117.
- (28) Korbel, G. A.; Lalic, G.; Shair, M. D. Reaction Microarrays: A Method for Rapidly Determining the Enantiomeric Excess of Thousands of Samples. *J. Am. Chem. Soc.* **2001**, *123*, 361–362.
- (29) Peng, Z. H.; Qun, L.; Zhou, X. F.; Carroll, S.; Geise, H. J.; Peng, B.; Dommissie, R.; Carleer, R. Synthesis and Physical Performance of Indole and Benzimidazole Cyanine Dyes. *J. Mater. Chem.* **1996**, *6*, 559–565.
- (30) Ma, S.; Devi-Kesavan, L. S.; Gao, J. Molecular Dynamics Simulations of the Catalytic Pathway of a Cysteine Protease: A Combined QM/MM Study of Human Cathepsin K. *J. Am. Chem. Soc.* **2007**, *129*, 13633–13645.
- (31) Lecaille, F.; Weidauer, E.; Juliano, M. A.; Brömme, D.; Lalmanach, G. Probing Cathepsin K Activity with a Selective Substrate Spanning its Active Site. *Biochem. J.* **2003**, *375*, 307–312.
- (32) Lu, J.; Wang, M.; Wang, Z.; Fu, Z.; Lu, A.; Zhang, G. Advances in the Discovery of Cathepsin K Inhibitors on Bone Resorption. *J. Enzyme Inhib. Med. Chem.* **2018**, *33*, 890–904.
- (33) McGrath, M. E.; Klaus, J. L.; Barnes, M. G.; Brömme, D. Crystal Structure of Human Cathepsin K Complexed with a Potent Inhibitor. *Nat. Struct. Biol.* **1997**, *4*, 105–109.
- (34) Barrett, D. G.; Catalano, J. G.; Deaton, D. N.; Hassell, A. M.; Long, S. T.; Miller, A. B.; Miller, L. R.; Shewchuk, L. M.; Wells-Knecht, K. J.; Willard, D. H., Jr.; Wright, L. L. Potent and Selective P2-P3 Ketoamide Inhibitors of Cathepsin K with Good Pharmacokinetic Properties via Favorable P1', P1, and/or P3 Substitutions. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4897–4902.
- (35) Husmann, K.; Muff, R.; Bolander, M. E.; Sarkar, G.; Born, W.; Fuchs, B. Cathepsins and Osteosarcoma: Expression Analysis Identifies Cathepsin K as an Indicator of Metastasis. *Mol. Carcinog.* **2008**, *47*, 66–73.
- (36) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAIS) from Screening Libraries and for their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719–2740.
- (37) False Positive Remover. https://www.cbligand.org/PAIS/search_struct.php (accessed Jun 17, 2021).
- (38) Schulz-Fincke, A. C.; Tikhomirov, A. S.; Braune, A.; Girbl, T.; Gilberg, E.; Bajorath, J.; Blaut, M.; Nourshargh, S.; Gütschow, M. Design of an Activity-Based Probe for Human Neutrophil Elastase: Implementation of the Lossen Rearrangement to Induce Förster Resonance Energy Transfers. *Biochemistry* **2018**, *57*, 742–752.
- (39) Van Gompel, J.; Schuster, G. B. Chemiluminescence in Organic Peroxides: Intramolecular Electron-Exchange Luminescence from a Secondary Perester. *J. Org. Chem.* **1987**, *52*, 1465–1468.
- (40) Becht, J. M.; Meyer, O.; Helmchen, G. Enantioselective Syntheses of (–)-(R)-Rolipram, (–)-(R)-Baclofen and Other GABA Analogues via Rhodium-Catalyzed Conjugate Addition of Arylboronic Acids. *Synthesis* **2003**, 2805–2810.
- (41) Grinson, C.; Geneve, S.; Halbin, E.; Coutrot, P. Stereoselective Synthesis of Vinylogous Peptides. *Tetrahedron Lett.* **2001**, *57*, 4903–4923.
- (42) Frizler, M.; Lohr, F.; Lülsdorff, M.; Gütschow, M. Facing the gem-Dialkyl Effect in Enzyme Inhibitor Design: Preparation of Homocycloleucine-Based Azadipeptide Nitriles. *Chem. Eur. J.* **2011**, *17*, 11419–11423.
- (43) Pomeislová, A.; Otmar, M.; Rubešová, P.; Benýšek, J.; Matoušová, M.; Mertlíková-Kaiserová, H.; Pohl, R.; Poštová Slavětinská, L.; Pomeisl, K.; Krečmerová, M. 1,2,4-Thiadiazole Acyclic Nucleoside Phosphonates as Inhibitors of Cysteine Dependent Enzymes Cathepsin K and GSK-3 β . *Bioorg. Med. Chem.* **2021**, *32*, No. 115998.
- (44) Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132.
- (45) Vagin, A.; Teplyakov, A. An Approach to Multi-copy Search in Molecular Replacement. *Acta Crystallogr. D Biol. Crystallogr.* **2000**, *56*, 1622–1624.
- (46) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Ead, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242.
- (47) Emsley, P.; Cowtan, K. Coot: Model-building Tools for Molecular Graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- (48) Long, F.; Nicholls, R. A.; Emsley, P.; Graeulis, S.; Merks, A.; Vaitkus, A.; Murshudov, G. N. AceDRG: A Stereochemical Description Generator for Ligands. *Acta Crystallogr. D Struct. Biol.* **2017**, *73*, 112–122.
- (49) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. MolProbity: All-atom Contacts and Structure Validation for Proteins and Nucleic Acids. *Nucleic Acids Res.* **2007**, *35*, W375–W383.
- (50) Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: Fully Automated Protein-Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447.
- (51) DeLano, W. L. Pymol: An Open-source Molecular Graphics Tool. *CCP Newsl. Prot. Crystallogr.* **2002**, *40*, 82–92.
- (52) Mangold, M.; Gütschow, M.; Stimberg, M. A Short Peptide Inhibitor as an Activity-Based Probe for Matriptase-2. *Pharmaceuticals* **2018**, *11*, 49.

SUPPORTING INFORMATION

An Activity-Based Probe for Cathepsin K Imaging with Excellent Potency and Selectivity

Carina Lemke,[§] Jakub Benýšek,^{‡,ζ} Dominik Brajtenbach,[§] Christian Breuer,^{§,ψ} Adéla Jílková,[‡]
Martin Horn,[‡] Michal Buša,^{‡,¥} Lenka Ulrychová,[‡] Annika Illies,[§] Katharina F. Kubatzky,[♦]
Ulrike Bartz,^ψ Michael Mareš^{*,‡} and Michael Gütschow^{*,§}

[§] Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany.

[‡] Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n. 2, 16610 Prague, Czech Republic.

^ζ First Faculty of Medicine, Charles University, Kateřinská 32, 12108 Prague, Czech Republic.

^ψ Department of Natural Sciences, University of Applied Sciences Bonn-Rhein-Sieg, von-Liebig-Str. 20, 53359 Rheinbach, Germany.

[¥] Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, 12800 Prague, Czech Republic.

[♦] Department of Infectious Diseases, Medical Microbiology and Hygiene, Heidelberg University Hospital, Im Neuenheimer Feld 324, Heidelberg, Germany.

Corresponding Authors

* Email: mares@uochb.cas.cz (M.M.)

* Email: guetschow@uni-bonn.de (M.G.)

Table of Contents

Supporting Figures and Tables.....	S3
Figure S1. Kinetic evaluation of precursor 22 with cathepsin K.....	S3
Table S1. X-ray data collection and refinement statistics.....	S4
Table S2. List of contacts formed between cathepsin K and ligands.....	S5
Figure S2. Selectivity of ABP 25 . Experiments with related cysteine cathepsins and different cell lines.....	S6
Figure S3. Fluorescence detection of human cathepsin K with ABP 24	S7
Figure S4. Imaging of human cathepsin K with ABP 24 in the presence of cell lysate.....	S8
HPLC Purity Traces for Compounds with <i>in vitro</i> Data.....	S9
NMR Spectra of Compounds with <i>in vitro</i> Data.....	S13
References.....	S19

Supporting Figures and Tables

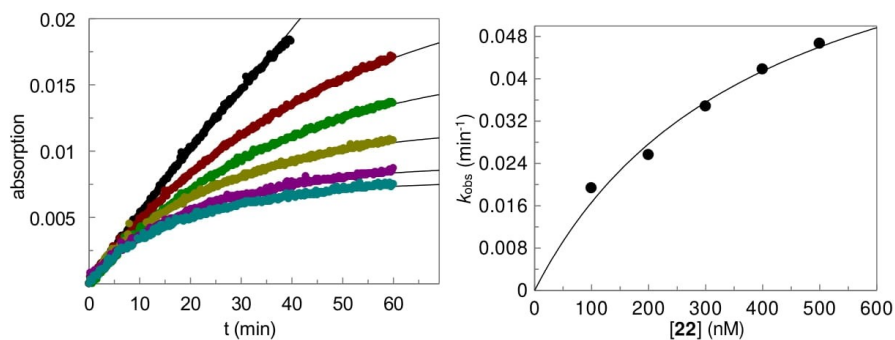


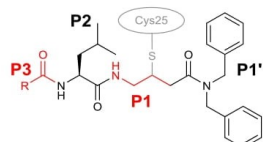
Figure S1. Kinetic evaluation of precursor **22** with cathepsin K. Left: Product formation is measured as absorption *versus* time. From top to bottom: 0 nM, 100 nM, 200 nM, 300 nM, 400 nM, 500 nM. From these progress curves, first-order rate constants k_{obs} were determined following the equation $[P] = v_i \times (1 - \exp(-k_{\text{obs}} \times t))/k_{\text{obs}} + d$, where $[P]$ is the product concentration, v_i is the initial product formation rate, t is the time, and d is the offset. Right: The mean of the k_{obs} values was plotted *versus* the inhibitor concentrations $[I]$. Non-linear regression was performed to obtain second-order rate constants k_{inac}/K_i using the equation $k_{\text{obs}} = k_{\text{inac}} \times [I]/(K_i + [I])$, with k_{inac} as the first-order rate constant of inactivation and K_i as the inhibitory constant. K_i values were additionally corrected by means of the Cheng-Prusoff equation.

Table S1. X-ray data collection and refinement statistics

Cathepsin K-ligand complex	Cathepsin K-21	Cathepsin K-25
Data collection statistics ^a		
Wavelength (Å)	1.541	1.541
Space group	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 22 ₁ 2 ₁
Temperature (K)	100	100
<i>a</i> , <i>b</i> , <i>c</i> (Å)	32.37, 71.82, 80.78	32.05, 71.68, 80.86
α , β , γ (deg)	90, 90, 90	90, 90, 90
Resolution (Å)	35.20–1.80 (1.865–1.80)	35.84–1.72 (1.782–1.72)
Number of total reflections	106,253 (6,827)	65,175 (968)
Number of unique reflections	18,053 (1,705)	17,404 (552)
R_{merge}^b	0.1468 (1.211)	0.0564 (0.3496)
Average <i>I</i> / σ (<i>I</i>)	9.60 (1.06)	15.85 (2.08)
Completeness (%)	99.57 (96.99)	85.38 (27.25)
Multiplicity	5.9 (4.0)	3.7 (1.8)
CC _{1/2} ^c (%)	99.6 (46.0)	99.8 (81.5)
Wilson B (Å ²)	20.13	16.20
Refinement statistics ^a		
Resolution range (Å)	35.20–1.80 (1.865–1.80)	35.84–1.72 (1.782–1.72)
Number of reflections in working set	18,050 (1,705)	17,510 (552)
Number of reflections in test set	910 (80)	895 (39)
<i>R</i> value ^d (%)	17.46 (31.25)	16.08 (26.04)
R_{free} value ^e (%)	20.94 (31.74)	19.08 (31.22)
Number of atoms in AU ^f protein/inhibitor/solvent	1,674/36/172	1,692/32/184
Average ADP ^g for protein/inhibitor/solvent (Å ²)	23.92/43.32/30.69	17.93/38.16/27.80
RMSD bond length (Å)	0.015	0.015
RMSD bond angle (deg)	1.87	1.80
Ramachandran plot statistics ^h		
favored regions (%)	96.3	96.7
allowed regions (%)	3.7	3.3
PDB code	7NXL	7NXM

^aNumbers in parentheses refer to the highest-resolution shell. ^b $R_{\text{merge}} = 100 \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is an individual intensity of the i^{th} observation of reflection hkl and $\langle I(hkl) \rangle$ is the average intensity of reflection hkl with summation over all data. ^cCC_{1/2} is the percentage of correlation between intensities from random half-datasets. ^d R value = $\|F_o\| - \|F_c\| / \|F_o\|$, where F_o and F_c are the observed and calculated structure factors, respectively. ^e R_{free} is equivalent to R value but is calculated for up to 5% of the reflections chosen at random and omitted from the refinement process. ^fAU, asymmetric unit. ^gADP, atomic displacement parameter, formally B-factor. ^hAs determined by Molprobity.³

S4

Table S2. List of contacts formed between cathepsin K and ligands.^a

Position	Precursor 21			Probe 25		
	Residue	C ^b	Hb ^b	Residue	C	Hb
P1'	Gln19	6	1	Gln19	5	1
	Gly20	1		Gly20	1	
	Gln21	1		Gln21	2	
	Cys22	8		Cys22	10	
	Gly23	4		Gly23	7	
	Cys25	4		Cys25	3	
	Ala137	4		Ala137	5	
	Ser138	4		Ser138	3	
	Gln143	2		Gln143	2	
	Asn161	3		Asn161	6	
	His162	8		His162	4	
Trp184	4		Trp184	3		
P1	Gln19	1		Gln19	1	
	Gly23	6		Gly23	6	
	Cys25	6		Cys25	7	
	Asn161	4	1	Asn161	4	1
	His162	1		His162	1	
P2	Gly23	1		Gly23	1	
	Cys25	2		Cys25	3	
	Trp26	4		Trp26	4	
	Gly65	3		Gly65	3	
	Gly66	12	2	Gly66	13	2
	Tyr67	1				
	Met68	2		Met68	2	
	Ala134	2		Ala134	2	
	Leu160	3		Leu160	3	
	Asn161	5		Asn161	4	
	His162	1		His162	1	
Ala163	1		Ala163	1		
Leu209	1		Leu209	1		
P3	Asp61	2				
	Gly64	2				
	Gly65	5				
	Gly66	7	1	Gly66	4	
	Tyr67	2		Tyr67	3	

^aAnalysis of protein-inhibitor contacts between the cathepsin K active site and precursor **21** or probe **25** was performed using the program CONTACT.⁴ The fragmentation of ligands into segments (in red and black) corresponding to the P3 to P1' positions is depicted in the scheme; R is a *tert*-butyloxy group for **21** and an apparent methyl group (derived from the fluorescent tag) for **25**. The distance cutoffs were set to 4.1 Å for all contacts and 3.2 Å for hydrogen bonds. The cathepsin K residues interacting with the individual ligand positions are specified. ^bFor each cathepsin K residue, total number of contacts (C) is listed, including hydrogen bonds (Hb). Residues forming hydrogen bonds are listed in bold.

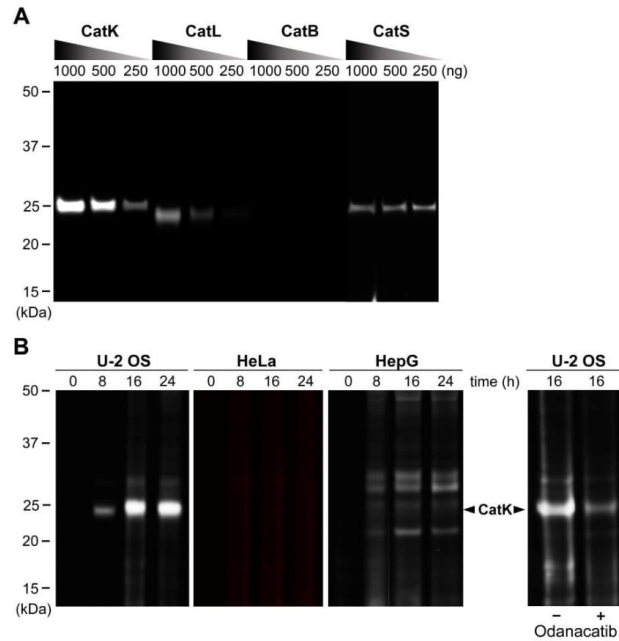


Figure S2. Selectivity of ABP **25**. Experiments with related cysteine cathepsins and different cell lines. **(A)** Indicated amounts of recombinant human cathepsin K (CatK), purified human liver cathepsins L (CatL, Sigma) and B (CatB, Calbiochem) as well as recombinant human cathepsin S (CatS, Enzo Life Sciences) were incubated with ABP **25** (5 μ M) at 25 °C for 1 h. The incubations were resolved by SDS-PAGE and visualized using a fluorescence scanner. **(B)** Imaging of CatK with ABP **25** in human osteosarcoma cells (U-2 OS), human epithelial cancer cells (HeLa), and human liver cancer cells (HepG). 50,000 cells each were seeded into a well of a 12-well plate and allowed to attach. The cells were incubated at 37 °C in 5% CO₂ with **25** (1 μ M) for the indicated time. After cell lysis with 100 μ L of SDS-PAGE loading buffer and heating at 100 °C for 10 min, the samples were resolved by SDS-PAGE and visualized using a fluorescence scanner. Additionally, U-2 OS cells were treated with the covalent-reversible cathepsin K-selective inhibitor odanacatib (10 μ M, 1 h) prior to treatment with ABP **25** (1 μ M).

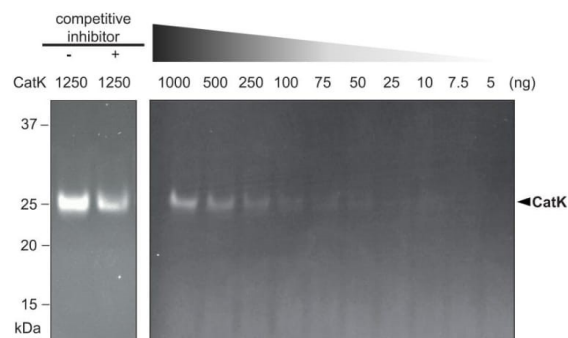


Figure S3. Imaging of human cathepsin K with ABP **24**. Left: Purified recombinant cathepsin K (CatK, 1.25 µg) was incubated with **24** (5 µM). In the control experiment, the enzyme was pretreated with a mixture of the competitive inhibitors E-64/K11777 (both 10 µM) prior to incubation with **24** (5 µM). Right: Cathepsin K (CatK, 1000-5 ng) was labeled with **24** (5 µM). Incubations were performed for 1 h at 25 °C and at pH 5.5. The reactions were resolved by SDS-PAGE and visualized with a Typhoon RGB imager using excitation at 488 nm and emission at 515 nm.

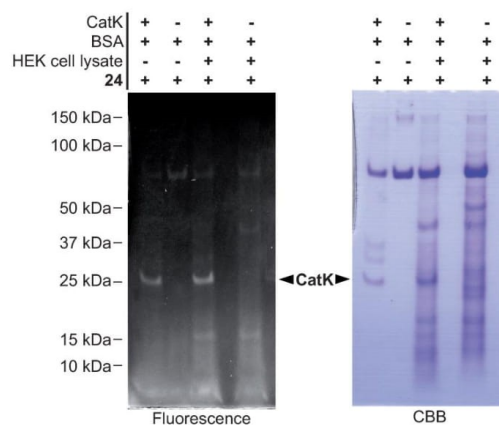
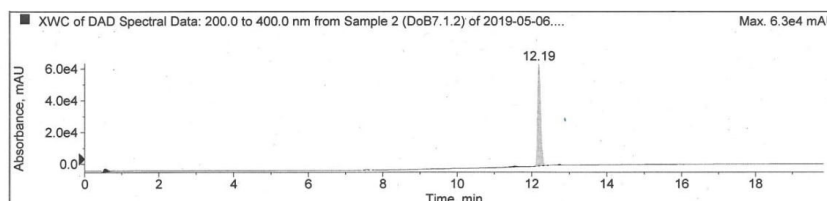


Figure S4. Imaging of human cathepsin K with ABP **24** in the presence of cell lysate. Human recombinant cathepsin K (CatK, 300 ng) was treated with ABP **24** (5 μ M) in the presence of HEK cell lysate (30 μ g) and BSA (500 ng) for 1 h at 25 $^{\circ}$ C and pH 5.5. In control experiments, cathepsin K (300 ng) and BSA (500 ng), or BSA (500 ng) alone, or HEK cell lysate (30 μ g) and BSA (500 ng) were treated with ABP **24** (5 μ M), separately. The reactions were resolved by SDS-PAGE and visualized with a Typhoon RGB imager using excitation at 488 nm and emission at 515 nm (Fluorescence, left) and subsequently stained with Coomassie Brilliant Blue (CBB, right).

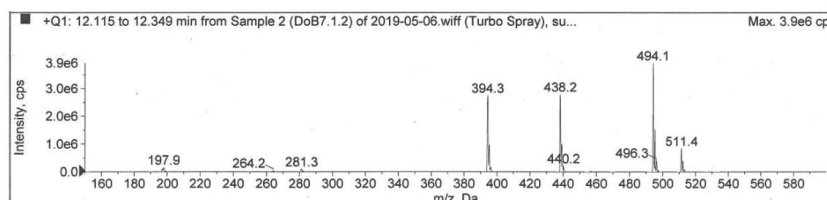
HPLC Purity Traces and MS Signals for Compounds with *in vitro* Data

HPLC traces were analyzed with HPLC/DAD/MS. The purity was determined based on area under the curve (AUC) values for absorption at 200-400 nm (**21**), 200-600 nm (**22**), 220-500 nm (**23**, **24**), 220-700 nm (**25**, **26**). Wavelengths ranges were selected with respect to compounds' spectroscopic properties. The sum of all AUCs except for that of the injection peak was set at 100%.

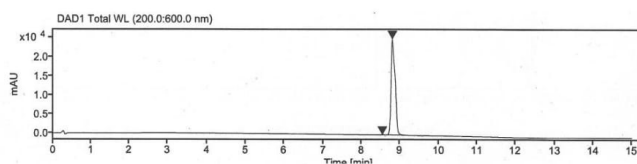
HPLC-DAD trace of **21** $t_R = 12.19$ min, purity = 96.24%



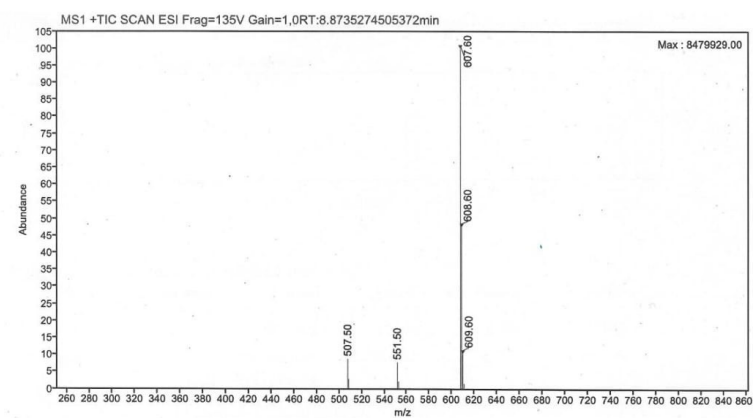
MS signal of **21**



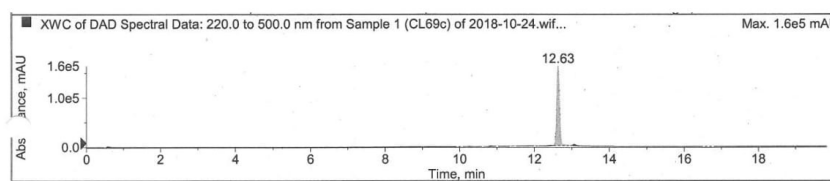
HPLC-DAD trace of **22** $t_R = 8.80$ min, purity = 99.75%



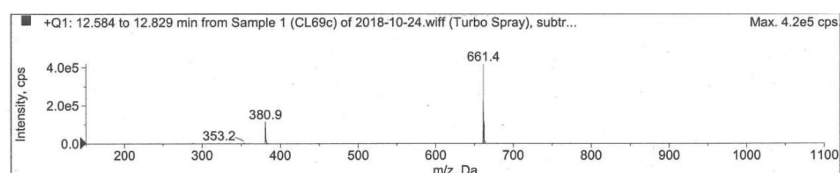
MS signal of **22**



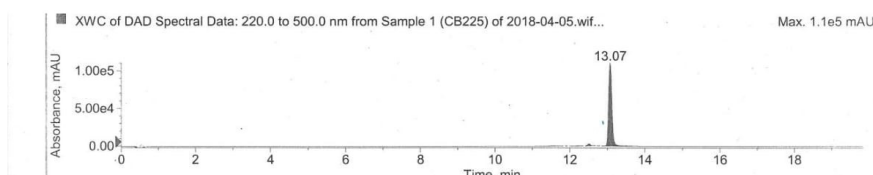
HPLC/DAD trace of **23** $t_R = 12.63$ min, purity = 96.17%



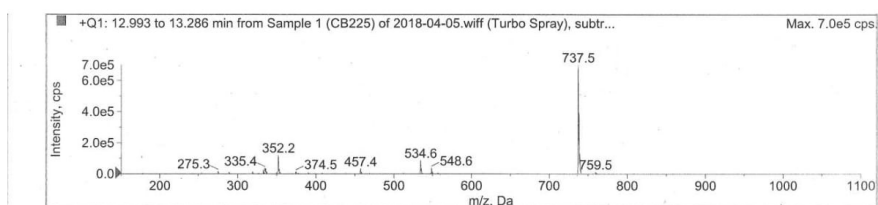
MS signal of **23**



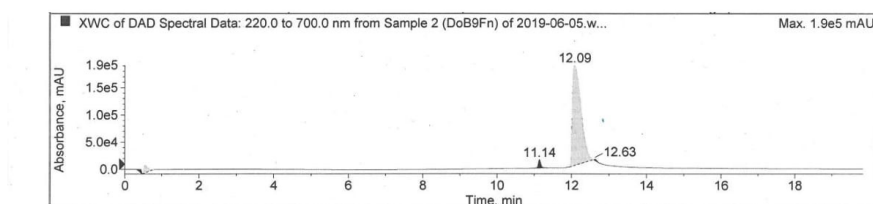
HPLC/DAD trace of **24** $t_R = 13.07$ min, purity = 95.39%



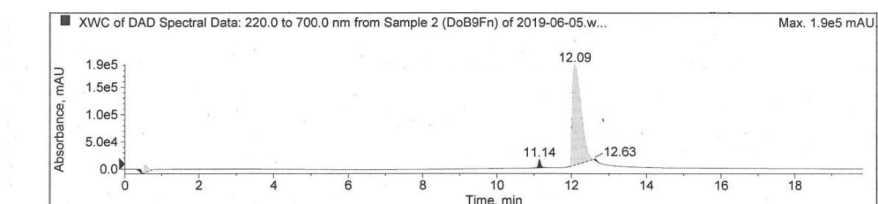
MS signal of **24**



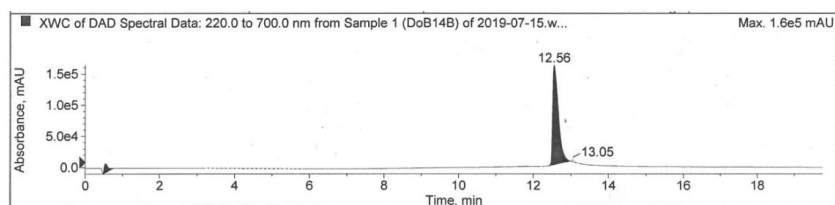
HPLC/DAD trace of **25** $t_R = 12.09$ min, purity = 97.16%



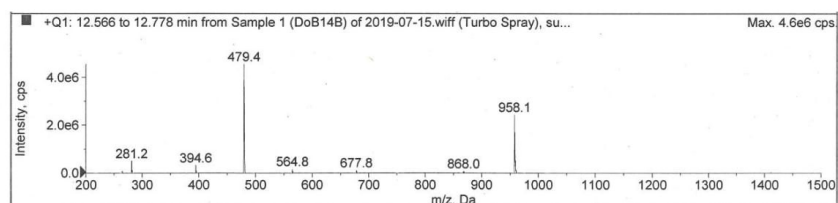
MS signal of **25**



HPLC/DAD trace of **26** $t_R = 12.56$ min, purity = 97.35%

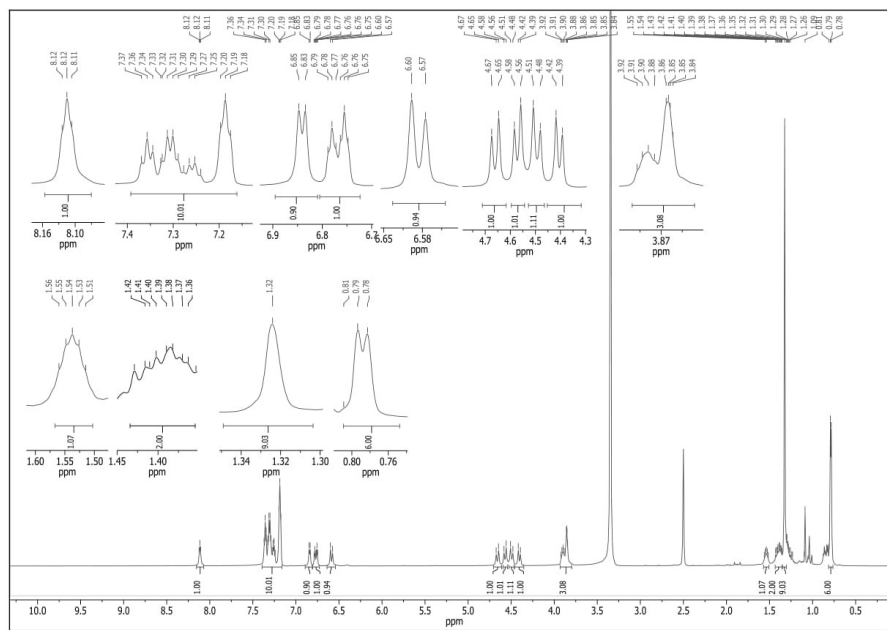


MS signal of **26**

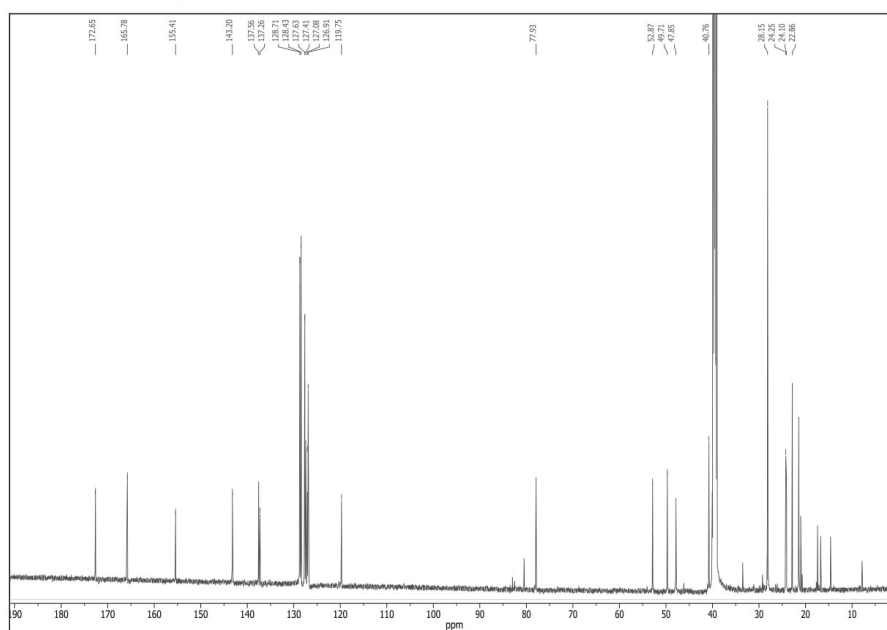


NMR Spectra of Compounds with *in vitro* Data

^1H NMR of compound **21**.

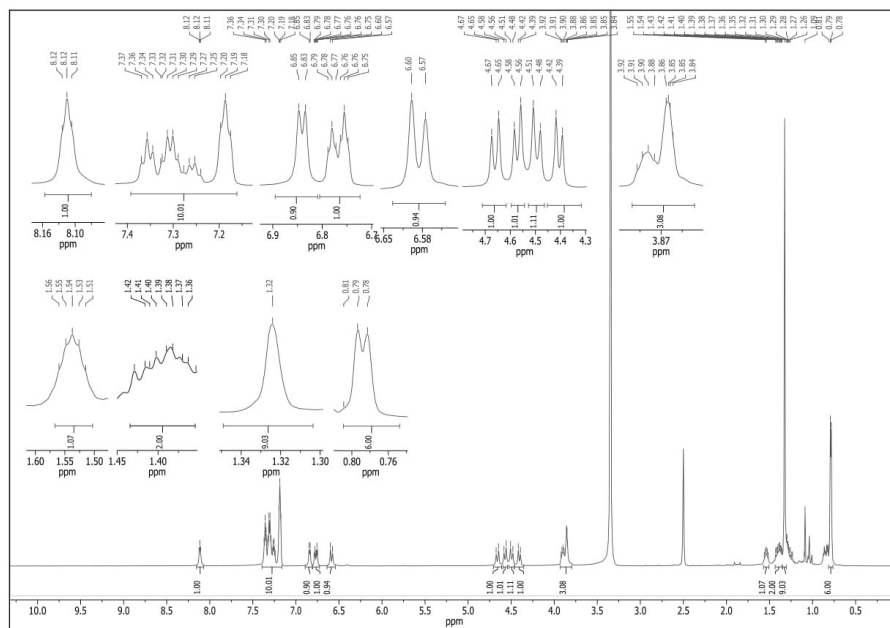


^{13}C NMR of compound **21**.

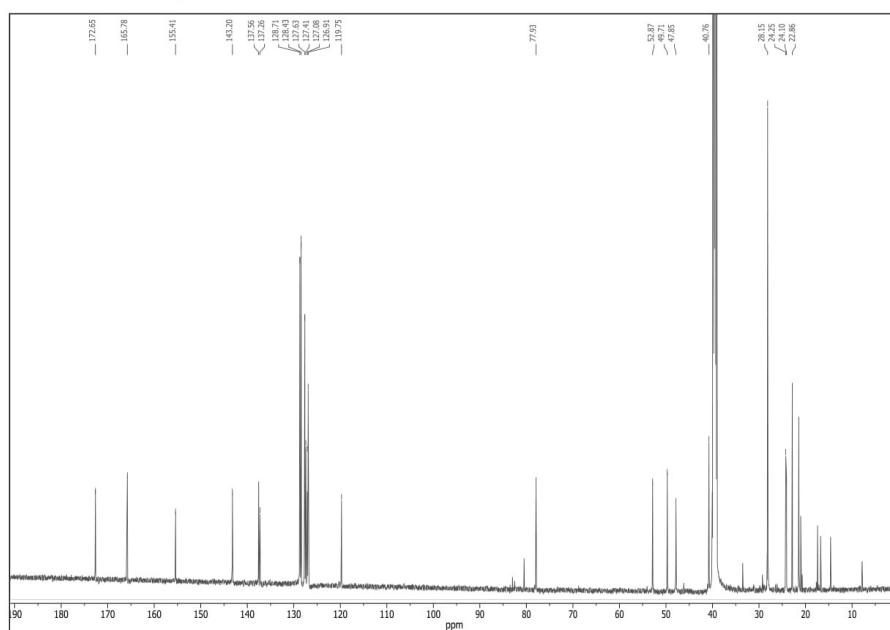


NMR Spectra of Compounds with *in vitro* Data

¹H NMR of compound **21**.

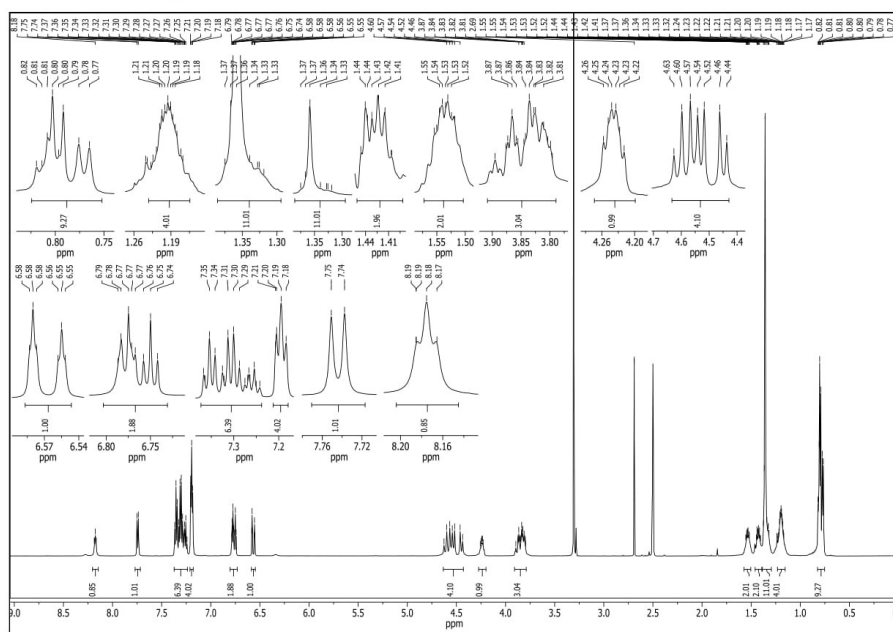


¹³C NMR of compound **21**.

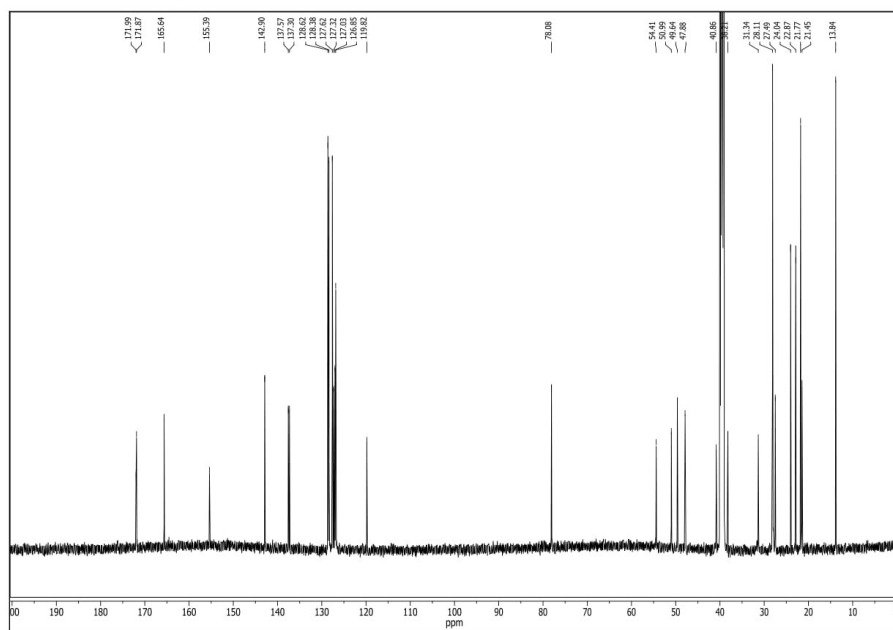


S13

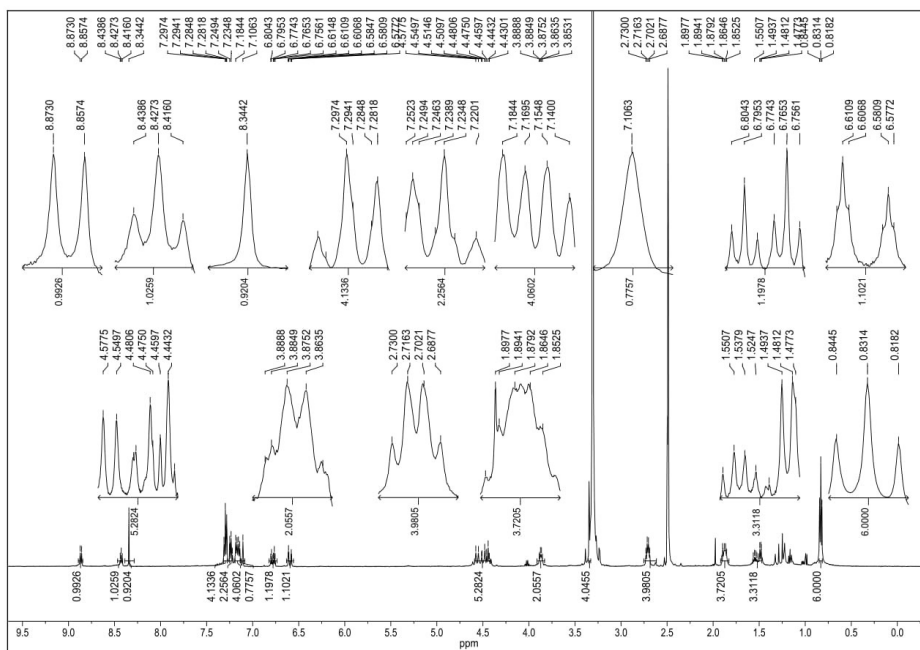
¹H NMR of compound 22.



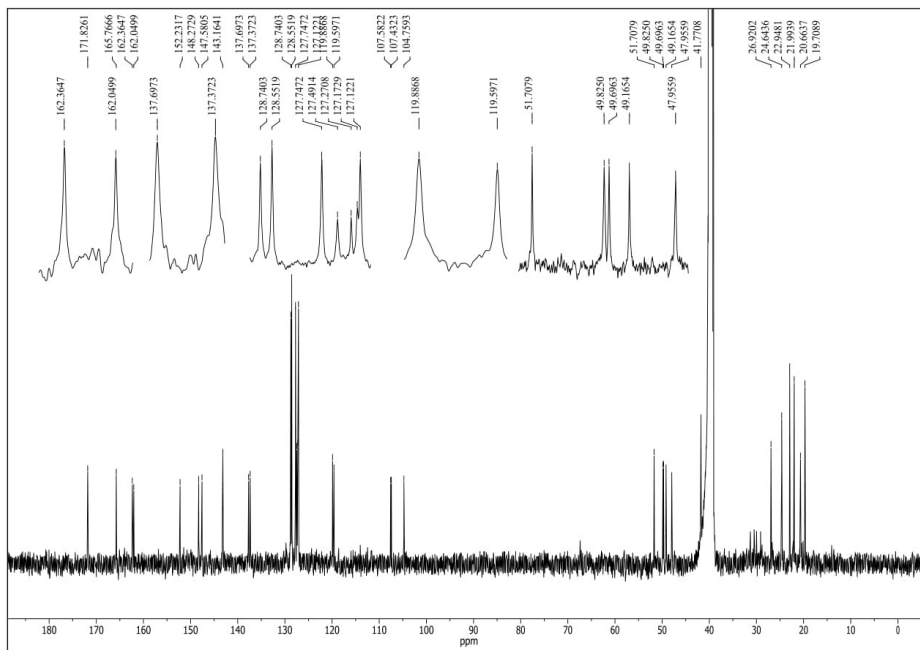
¹³C NMR of compound 22.



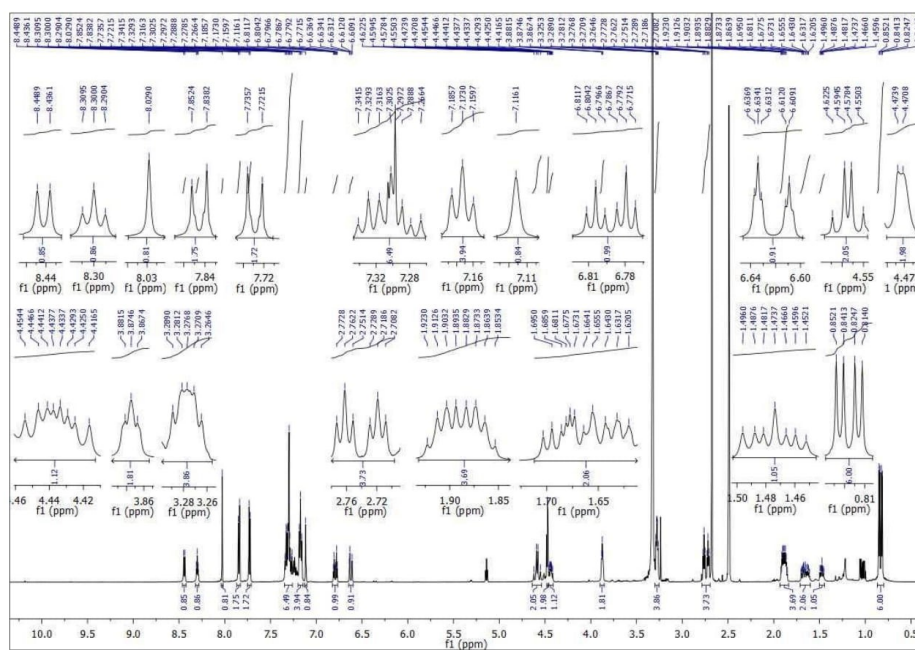
¹H NMR of compound **23**.



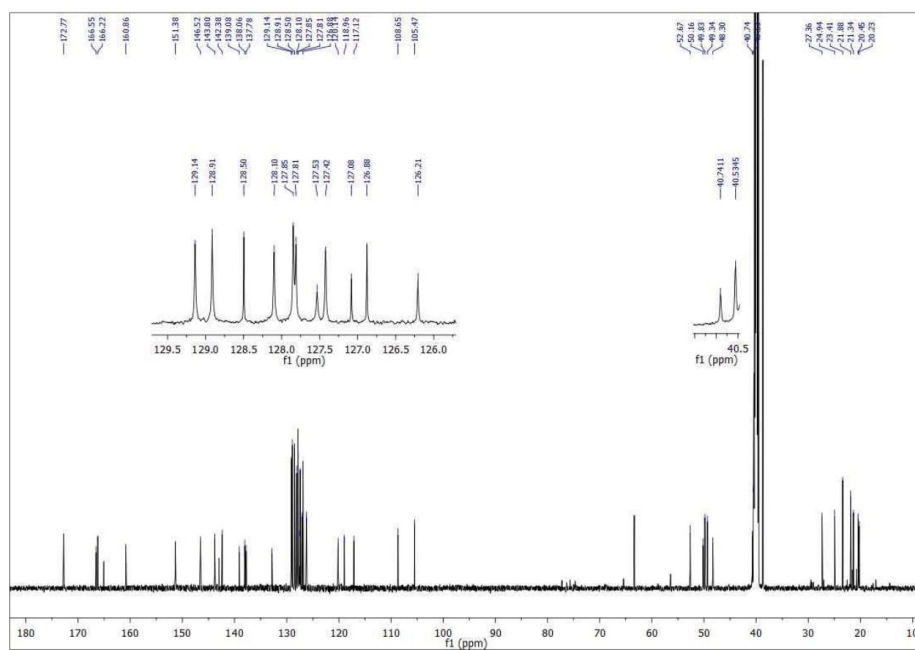
¹³C NMR of compound **23**.



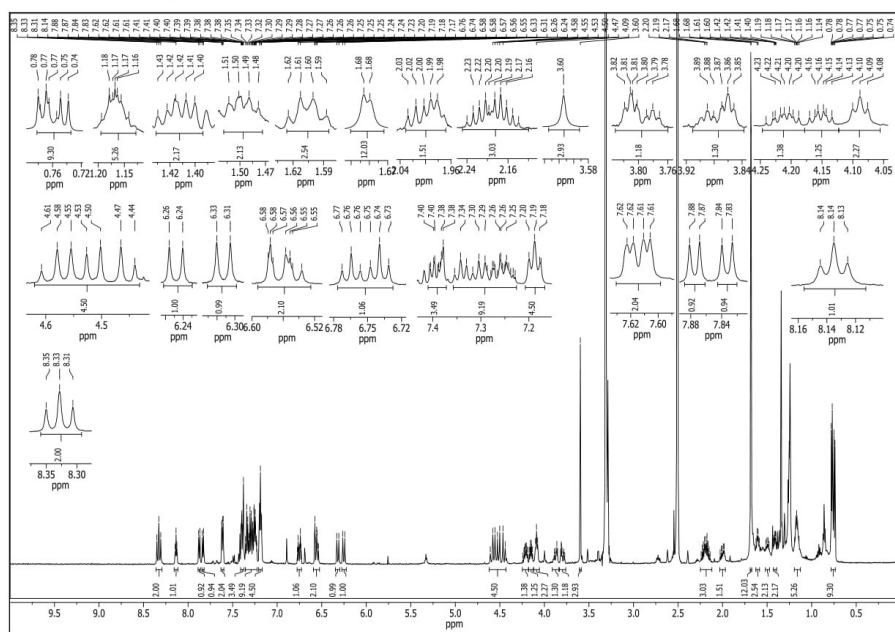
¹H NMR of compound 24.



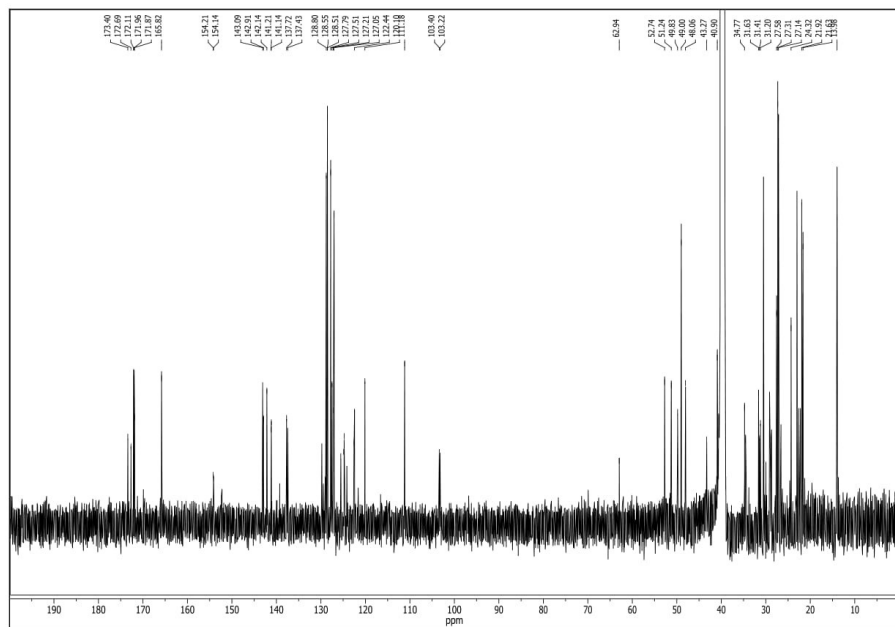
¹³C NMR of compound 24.



¹H NMR of compound 26.



¹³C NMR of compound 26.



References

- (1) Karplus, P. A.; Diederichs, K. Linking crystallographic model and data quality. *Science* **2012**, *336*, 1030–1033.
- (2) Brünger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **1992**, *355*, 472–475.
- (3) Lovell, S. C.; Davis, I. W.; Arendall, W. B., III; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure validation by $C\alpha$ geometry: ϕ , ψ and $C\beta$ deviation. *Proteins: Struct. Funct. Bioinform.* **2003**, *50*, 437–450.
- (4) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242.

Publikace č. 3

Highly potent inhibitors of cathepsin K with a differently positioned cyanohydrazide warhead: structural analysis of binding mode to mature and zymogen-like enzymes

Benýšek J., Buša M., Rubešová P., Fanfrlík J., Lepšík M., Brynda J., Matoušková Z., Bartz U., Horn M., Gütschow M., Mareš M. (2022) *Journal of Enzyme Inhibition and Medicinal Chemistry* (accepted a momentálně v tisku). DOI: 10.1080/14756366.2021.2024527

Highly potent inhibitors of cathepsin K with a differently positioned cyanohydrazone warhead: structural analysis of binding mode to mature and zymogen-like enzymes

Jakub Benýšek^{a,b}, Michal Buša^{a,c}, Petra Rubešová^a, Jindřich Fanfrlík^a, Martin Lepšík^a, Jiří Brynda^a, Zuzana Matoušková^a, Ulrike Bartz^d, Martin Horn^a, Michael Gütschow^e and Michael Mareš^a

^aInstitute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Prague, Czech Republic; ^bFirst Faculty of Medicine, Charles University, Prague, Czech Republic; ^cDepartment of Biochemistry, Faculty of Science, Charles University, Prague, Czech Republic; ^dDepartment of Natural Sciences, University of Applied Sciences Bonn-Rhein-Sieg, Rheinbach, Germany; ^ePharmaceutical & Medicinal Chemistry, Pharmaceutical Institute, University of Bonn, Bonn, Germany

ABSTRACT

Cathepsin K (CatK) is a target for the treatment of osteoporosis, arthritis, and bone metastasis. Peptidomimetics with a cyanohydrazone warhead represent a new class of highly potent CatK inhibitors; however, their binding mechanism is unknown. We investigated two model cyanohydrazone inhibitors with differently positioned warheads: an azadipeptide nitrile **Gü1303** and a 3-cyano-3-aza- β -amino acid **Gü2602**. Crystal structures of their covalent complexes were determined with mature CatK as well as a zymogen-like activation intermediate of CatK. Binding mode analysis, together with quantum chemical calculations, revealed that the extraordinary picomolar potency of **Gü2602** is entropically favoured by its conformational flexibility at the nonprimed-primed subsites boundary. Furthermore, we demonstrated by live cell imaging that cyanohydrins effectively target mature CatK in osteosarcoma cells. Cyanohydrins also suppressed the maturation of CatK by inhibiting the autoactivation of the CatK zymogen. Our results provide structural insights for the rational design of cyanohydrazone inhibitors of CatK as potential drugs.

ARTICLE HISTORY

Received 30 November 2021
Accepted 27 December 2021

KEYWORDS

Cathepsin K; protease inhibitor; cyanohydrazone warhead; azadipeptide nitrile; Crystal structure

1. Introduction



Cathepsin K (CatK) is one of the most investigated cysteine cathepsins, both by academia and pharmaceutical companies. It is expressed in high levels in osteoclasts, where it serves as the principal protease involved in bone remodelling. It has been validated as a therapeutic target for osteoporosis, an increasing health problem in the modern world^{1–4}. This disorder is caused by progressive loss of bone mass due to excessive activity of osteoclastic CatK. Several anti-remodelling inhibitors of CatK, such as odanacatib or balicatib^{5–7} (Figure 1), have been developed for the treatment of osteoporosis but have not yet been approved. CatK has also been implicated in the pathophysiology of two common forms of arthritis, osteoarthritis and rheumatoid arthritis^{8,9}. In bone and cartilage disorders, CatK functions as a potent collagen degrading enzyme with the unique ability to cleave the triple helix of collagen molecules at multiple locations, an activity that is unparalleled among human collagenases^{10,11}. This activity is induced by glycosaminoglycans, foremost chondroitin-4-sulphate, which mediates the formation of a complex between CatK and the collagen substrate^{10,12,13}.


Further, there is increasing evidence that CatK is a pro-tumorigenic protease that plays an important role in processes associated with tumour growth, invasion, and metastasis of cancer

cells and their interactions with the tumour microenvironment (for review see^{14,15}). Its complex action includes direct degradation of collagen and other extracellular matrix proteins, e.g. in bone metastases, and indirect affecting of the signalling pathways^{16–19}. In glioma, CatK can regulate cancer stem-like cell mobilisation and proteolytically modulate levels of chemokines and growth factors²⁰. Thus, clinical research is investigating CatK as a marker for diagnosis and survival prognosis in metastatic cancer and as a target for anticancer inhibitors^{21,22}.

At the protein level, CatK activity is regulated by endogenous protein inhibitors^{23,24} and by zymogen activation^{24–26}. CatK is synthesised as an inactive zymogen (procathepsin) in which the N-terminal propeptide blocks the active site^{27,28}. Activation to the mature, active form occurs upon proteolytic removal of the N-terminal propeptide (also termed the “activation peptide”). This process was shown to be autocatalytic and bimolecular; it is triggered by acidic pH and also enhanced by interaction with chondroitin-4-sulfate^{25,29}. The activation pathway of CatK includes an activation intermediate with a partially processed propeptide²⁵, which has not been studied in detail so far, and its spatial structure remains unknown.

Peptidomimetics with a reactive nitrile functionality have attracted particular attention as potent inhibitors of CatK and other cysteine cathepsins. The electrophilic nitrile warhead allows

CONTACT Michael Mareš  michael.mares@uochb.cas.cz  Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n. 2, 16610 Prague 6, Czech Republic

 Supplemental data for this article can be accessed [here](#).

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

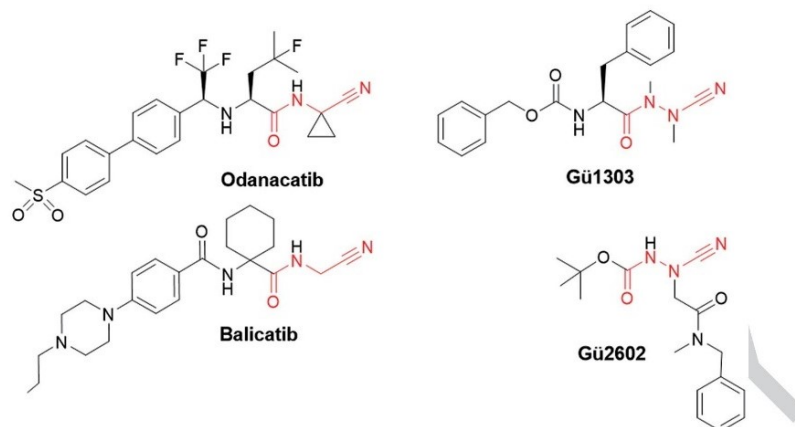


Figure 1. Chemical structures of peptidomimetic inhibitors of cathepsin K with a reactive nitrile functionality. The warhead is indicated in red. The dipeptide nitriles odanacatib and balicatib were developed as osteoporosis drugs. The azadipeptide nitrile Gü1303 and 3-cyano-3-aza- β -amino acid Gü2602 contain the cyanohydrate warhead.

for covalent interaction with the catalytic cysteine nucleophile, leading to the reversible formation of a covalent thioimide adduct. Representatives of such CatK inhibitors are dipeptide nitriles odanacatib and balicatib^{5–7} (Figure 1). An exchange of the α -CH moiety of the P1 amino nitrile by a nitrogen atom led to azadipeptide nitriles with the cyanohydrate warhead forming a stabilised isothiosemicarbazide adduct^{30–32}. They have been introduced as a class of efficient covalent-reversible inhibitors of human cysteine cathepsins, including CatK, and their homologs from parasites and pathogens^{30–34}. Compared to their parent carbapeptide analogs, bioactive azapeptides can possess improved potency, selectivity, and pharmacokinetics^{35–38}. The 3-cyano-3-aza- β -amino acid derivatives represent another scaffold bearing the cyanohydrate warhead³⁹. They were designed to position the warhead centrally in the peptidomimetic inhibitor molecule for extended interactions of inhibitor substructures with the non-primed and primed binding regions of the target enzymes. These compounds were found to be exceptionally potent, in particular towards CatK³⁹.

The exact binding mode of peptidomimetics with the cyanohydrate warhead to CatK had not been characterised so far. To do this, we investigated two model cyanohydrate inhibitors selective for CatK with high potency in the subnanomolar to picomolar range, namely the azadipeptide nitrile Gü1303^{30,31} and the 3-cyano-3-aza- β -amino acid derivative Gü2602³⁹ (Figure 1). The crystal structures of their complexes with mature CatK and the activation intermediate of CatK were determined and functional properties *in vitro* and in cells were described. The present structures and inhibitor interaction data provide a footing for the rational design of next generation cyanohydrate inhibitors of CatK as potential therapeutics.

2. Materials and methods

2.1. Materials

Inhibitors Gü1303 and Gü2602 and the CatK activity-based probe 25 were synthesised as described previously^{30,39,40}.

2.2. Expression and purification of the recombinant zymogen of cathepsin K

Human CatK (Uniprot accession number P43235) was expressed in the X-33 strain of the methylotrophic yeast *Pichia pastoris* (Thermo Fisher). A gene coding for the zymogen form of CatK was purchased from GenScript and recloned into expression plasmid pPICZ α A (Thermo Fisher) using XhoI and NotI restriction sites. Transformation of *P. pastoris* cells and protein expression were carried out as described previously^{41,42}. The yeast medium containing the recombinant CatK zymogen was centrifuged (2,500 g for 10 min), and the supernatant was lyophilised and dissolved in 20 mM MES pH 6.0 (to 10% of the original volume). The protein solution was then desalted over a Sephadex G-25 column equilibrated with the same buffer. The CatK zymogen was purified using chromatography on Mono S (HR 5/5 column) equilibrated with 50 mM sodium acetate pH 5.5, and eluted by a linear gradient of 2 M NaCl. The purified protein was concentrated to 2 mg/ml using an Amicon Ultracel-10k centrifugal filter device (Millipore).

2.3. Activation of the cathepsin K zymogen and preparation of inhibitor complexes

The purified CatK zymogen (75 μ M) was activated by incubation in 0.1 M sodium acetate pH 4.0 containing 2.5 mM DTT, 1 mM EDTA, and 0.3 M NaCl under an argon atmosphere at room temperature. The zymogen-like activation intermediate iCatK was obtained after 30 min of incubation, and fully activated mature enzyme mCatK after 75 min. Activation was terminated by the addition of 6-fold molar excess of the inhibitor Gü1303 or Gü2602, followed by incubation under argon atmosphere for 3 h at room temperature. The activation and inhibition were monitored with a kinetic activity assay using the fluorogenic substrate Cbz-Gly-Pro-Arg-AMC (Cbz, benzyloxycarbonyl; AMC, 7-amino-4-methylcoumarin) and Laemmli-SDS-PAGE. The processing sites were identified by N-terminal protein sequencing after electroblotting of Laemmli-SDS-PAGE gels to a PVDF membrane using a Procise 494 cLC protein sequencer (Applied Biosystems) and by peptide mapping using mass spectrometry (LC-MS/MS) on an LTQ Orbitrap XL mass

spectrometer (Thermo Scientific) coupled to a UHPLC system. The LC-MS/MS data were processed with Bioworks software (Thermo). The complexes were buffer-exchanged into 20 mM sodium acetate pH 5.5 containing 2.5 mM DTT and 0.25 M NaCl, and concentrated to 3.5 mg/ml for mCatK and 5 mg/ml for iCatK using an Amicon Ultracel-10k centrifugal filter device; the inhibitors were maintained during buffer exchange and concentration in a 6-fold molar excess to mCatK/iCatK in the mixture.

2.4. Autoactivation assay with the cathepsin K zymogen and inhibitors

The CatK zymogen (2.8 μ M) was incubated at room temperature in the presence or absence of the inhibitor (10 μ M **Gü1303** or **Gü2602**) in 150 μ L of 0.1 M sodium acetate pH 4.0 containing 2.5 mM DTT and 0.3 M NaCl for 0, 30 and 240 min. The reaction was terminated by the addition of E-64 (10 μ M final concentration), followed by acetone precipitation; the reaction mixture was separated by Laemmli-SDS-PAGE.

2.5. Protein crystallisation and data collection

Crystals were obtained by the vapour diffusion technique in hanging drops at 18 °C. Drops consisted of 1 μ L of the protein-inhibitor complex, 0.15 μ L 1:100 diluted seed stock and 1 μ L of the reservoir solution. The drops were equilibrated over the following reservoir solutions: (1) 10% PEG 8000, 20% ethylene glycol, 0.02 M sodium L-glutamate, 0.02 M DL-alanine, 0.02 M glycine, 0.02 M DL-lysine HCl, 0.02 M DL-serine, 0.1 M MES/imidazole pH 6.5 for iCatK-**Gü1303** complex; (2) 10% PEG 8000, 20% ethylene glycol, 0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M trisodium citrate, 0.02 M sodium potassium L-tartrate, 0.02 M sodium oxamate, 0.1 M MES/imidazole pH 6.5 for iCatK-**Gü2602** complex; (3) 12.5% PEG 1000, 12.5% PEG 3350, 12.5% MPD, 0.03 M sodium nitrate, 0.03 M disodium hydrogen phosphate, 0.03 M ammonium sulphate, 0.1 M MES/imidazole pH 6.5 for mCatK-**Gü1303** complex; and (4) 10% PEG 8000, 20% ethylene glycol, 0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M trisodium citrate, 0.02 M sodium potassium L-tartrate, 0.02 M sodium oxamate, 0.1 M MES/imidazole pH 6.5 for mCatK-**Gü2602** complex. Crystals were flash-cooled by plunging them into liquid nitrogen, and diffraction data from the crystals of mCatK complexes were collected at 100 K on a MicroMax-007 HF Microfocus rotating anode X-ray generator equipped with a PILATUS 300K detector (Rigaku). Data from the crystals of iCatK complexes were collected at 100 K on an MX 14.1 beamline operated by Helmholtz-Zentrum Berlin at the BESSY II electron storage ring in Berlin-Adlershof, Germany⁴³. All diffraction data was processed using the XDS suite of programs⁴⁴. Crystal parameters, data collection statistics, and final refinement statistics are in Table S1.

2.6. Structure determination, refinement and analysis

The structures of the mCatK-**Gü1303**/**Gü2602** complexes and the main domain of iCatK in complex with **Gü2602** were solved by molecular replacement with the program MolRep⁴⁵ from the CCP4 program suite⁴⁶ using the structure of human cathepsin K (PDB code 7NXM)⁴⁰ as a search model. The propeptide domain structure in iCatK-**Gü2602** was built using the *de novo* model building program Buccaneer⁴⁷. The structure of the iCatK-**Gü1303** complex was solved by molecular replacement using the structure of iCatK-**Gü2602** as a search model. Model refinement was carried out

using the program REFMAC 5.5, interspersed with manual adjustments using Coot. The geometric restraints for ligands were constructed by the program AceDRG⁴⁸. The quality of the final models was validated with MolProbity⁴⁹. The final refinement statistics are given in Table S1. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with the accession codes: 7QBL, 7QBN, 7QBM, and 7QBO for mCatK-**Gü2602**, mCatK-**Gü1303**, iCatK-**Gü2602**, and iCatK-**Gü1303**, respectively. Inhibitor interactions were analysed using the programs CONTACT⁴⁶ and PLIP⁵⁰. The distance cut-offs were set to 3.3 Å for hydrogen bonds and 4.2 Å for contacts. The nonpolar interactions represent contacts between two hydrophobic atoms defined as carbon atoms having carbon or hydrogen atoms as neighbours. All figures showing structural representations were prepared with the PyMOL Molecular Graphics System, version 1.40 (Schrodinger, LLC).

2.7. Molecular modelling

The X-ray structure of the mCatK-**Gü2602** complex was used for molecular modelling. Hydrogen atoms were added to the protein by the Reduce and Leap programs in AMBER 14⁵¹. The Asp, Glu, Lys, Arg, and His residues were charged, with the exception of Asp82 and His162. Hydrogen atoms of the inhibitor were added manually using the PyMOL. The ff14SB force field⁵¹ was used for the protein, while the GAFF force field⁵¹ and RESP charges at HF/6-31G* level were used for the ligand. The molecular dynamics/quenching (MD/Q) technique was used to search for possible conformations of **Gü2602** in the complex. The simulations were performed using AMBER 14. During the simulations, the *N*-benzyl-*N*-methylformamide segment of the inhibitor was relaxed; specifically, the C, O, and N heavy atoms in this segment and all H atoms of the inhibitor were relaxed while the rest of the system was frozen. The structures were collected every 1 ps in two independent runs with two orientations of the relaxed inhibitor segment (i.e. the heavy atoms as above). The simulations were 10 ps long at 600 K using 1 fs time step and a Berendsen thermostat. All the obtained structures were optimised (i.e. residues within 6 Å of the inhibitor were relaxed) by using the corrected semiempirical quantum mechanical (SQM) PM6-D3H4 method^{52,53}. The environment was described by the COSMO implicit solvent model^{54,55}. The SQM calculations were done by Cuby4⁵⁶ and MOPAC2016⁵⁷. Residues further than 6 Å from the inhibitor were frozen during the optimisation.

2.8. Cathepsin K activity and inhibition assays

Enzymatic activity of mCatK was measured using a kinetic continuous assay with the fluorogenic substrate Cbz-Gly-Pro-Arg-AMC (Bachem). The assay was performed in a 96-well microplate format in a total assay volume of 100 μ L at 37 °C. The assay mixture contained an aliquot of mCatK (e.g. from the autoactivation assay) and 20 μ M Cbz-Gly-Pro-Arg-AMC in 0.1 M sodium acetate pH 5.5 containing 2.5 mM DTT, 0.15 M NaCl, 0.1% PEG 6000, and 1 mM EDTA. The kinetics of the product release were continuously monitored in an Infinite M1000 microplate reader (Tecan) at excitation and emission wavelengths of 360 and 465 nm, respectively. The Michaelis-Menten kinetic parameters were determined by measuring the rate of hydrolysis of the substrate (0–100 μ M) using the same assay with mCatK (0.4 nM); the K_m value obtained by nonlinear regression using GraFit software was 17.4 μ M.

Inhibition measurements were performed analogously. mCatK (0.42 nM) was added to a mixture of the fluorogenic substrate

Cbz-Gly-Pro-Arg-AMC (20 μ M) and an inhibitor (0–100 nM) in 0.1 M sodium acetate pH 5.5 containing 2.5 mM DTT, 0.15 M NaCl, 0.1% PEG 6000, and 1 mM EDTA. The substrate hydrolysis was monitored for 40 min. For the slow-binding inhibitor **Gü1303**, an observed first-order rate constant k_{obs} was calculated at each inhibitor concentration by fitting the progress curve to the equation $P = v_s t + (v_i - v_s)(1 - \exp(-k_{\text{obs}} t)) / k_{\text{obs}} + d$, where P is the product formation, v_s is the steady-state velocity, t is the reaction time, v_i is the initial velocity, and d is offset. The apparent inhibition constant K_i' was determined by non-linear regression using equation $v_s/v_0 = 1/(1 + [I]/K_i')$. The true inhibition constants K_i were calculated using the Cheng Prusoff equation $K_i = K_i'/(1 + [S]/K_m)$, where $[S]$ is the substrate concentration and K_m is the Michaelis constant. The apparent second-order rate constant $k_{2\text{nd}}$ was determined by fitting to the linear equation $k_{\text{obs}} = k'_{2\text{nd}} [I] + k_{\text{off}}$, and the true constant $k_{2\text{nd}}$ was calculated by correction $k_{2\text{nd}} = k'_{2\text{nd}} (1 + [S]/K_m)$. The fast-binding inhibitor **Gü2602** showed linear progress curves, and the apparent inhibition constant K_i' was determined by non-linear regression using the Morrison equation for tight binding inhibition⁵⁸ with GraphPad Prism software. The concentration of mCatK was determined by active site titration as described previously⁵⁹ with E-64 used as the titrant⁶⁰. The final concentration of DMSO in the assay systems did not exceed 1.5%.

2.9. Osteosarcoma cell imaging

Imaging of cathepsin K in human bone osteosarcoma cells (U-2 OS) was done as described previously using the CatK activity-based probe **25**⁴⁰. U-2 OS cells were cultivated in McCoy's 5A medium supplemented with 10% FBS, 2 mM glutamax (L-alanyl-L-glutamine), 100 units/ml penicillin, and 100 μ g/ml streptomycin at 37 °C in 5% CO₂. Fifty thousand cells were seeded into a well of a 12-well plate and allowed to attach. After 24 h, 1 μ M probe was added, and cells were incubated for 16 h. Cells were detached using 0.25% trypsin-EDTA solution (Sigma-Aldrich, Merck), collected by centrifugation (500 g for 5 min), washed with PBS, resuspended in 100 μ l of loading buffer, and heated at 100 °C for 10 min. The competitive labelling was performed after the preincubation of cells with 1 μ M **Gü1303** or **Gü2602** for 3 h. The samples (30 μ l) were separated on 4–12% Bis-Tris polyacrylamide gels (Thermo Fisher Scientific). The gels were visualised using a Typhoon RGB imager (GE Healthcare Life Sciences) with excitation at 635 nm and emission at 660 nm (long pass filter).

3. Results

3.1. Cyanohydrazides are potent inhibitors of cathepsin K activity and zymogen activation both in vitro and in cells

Functional properties of the azadipeptide nitrile inhibitor **Gü1303**³⁰ and the 3-cyano-3-aza- β -amino acid inhibitor **Gü2602**²⁹ were characterised using an *in vitro* kinetic assay with the mature form of recombinant human cathepsin K (mCatK) and the fluorogenic substrate Cbz-Gly-Pro-Arg-AMC. Subnanomolar values of the inhibition constant K_i were determined: Notably, we found a K_i of 0.91 nM for **Gü1303** and an almost two orders of magnitude lower K_i of 0.013 nM for **Gü2602** (Table 1). A detailed analysis of the kinetic behaviour showed non-linear progress curves that indicated time-dependent inhibition for **Gü1303** typical for slow-binding inhibitors (Figure 2(A)). This allowed for the calculation of the second-order rate constant of inactivation k_{on} of $527 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$, thus demonstrating the extraordinary potency of **Gü1303** (Table 1). In contrast, linear progress curves were

Table 1. Inhibition of human mature cathepsin K.

Compound	mCatK inhibition ^a		
	K_i (nM)	k_{on} ($10^3 \text{ M}^{-1} \text{ s}^{-1}$)	k_{off} (10^{-3} s^{-1})
Gü2602	0.013 \pm 0.008	n.d. ^b	n.d. ^b
Gü1303	0.91 \pm 0.04	527 \pm 8	0.480 \pm 0.028

^aThe inhibition parameters were measured using a kinetic activity assay with the fluorogenic peptide substrate Cbz-Gly-Pro-Arg-AMC at pH 5.5. ^bn.d.: not determined for linear progress curves.

obtained for **Gü2602** that are characteristic of fast-binding inhibitors (Figure 2(B)).

The mature enzyme mCatK is generated during the acidic autocatalytic processing of an inactive precursor, the cathepsin K zymogen (pCatK), and this activation process is associated with the proteolytic removal of the propeptide domain from pCatK^{24–26}. We examined the effect of **Gü1303** and **Gü2602** on the autoactivation of recombinant pCatK induced by acidic pH using an *in vitro* assay with SDS-PAGE visualisation of the processing forms pCatK and mCatK (bands of ca 37 kDa and ca 24 kDa, respectively) (Figure 3(A)). Inhibitors were tested under conditions providing full conversion to mCatK and the cleaved propeptide (a band of ca 10 kDa). Both inhibitors substantially suppressed the autoactivation processing of pCatK that resulted in approximately a 10–15% conversion compared to uninhibited conditions. Also, both inhibitors were able to suppress the autodegradation of the generated mCatK observed under prolonged incubation (compare 0.5 and 4 h experiments, Figure 3(A)).

Finally, we investigated the interaction of the inhibitors with CatK using a cell-based assay with the human osteosarcoma cell line U-2 OS, which has an enhanced expression level of CatK⁶². For CatK imaging, we used a fluorescent activity-based probe that binds specifically and irreversibly to the active site of CatK⁴⁰. Competition of the probe and inhibitor was monitored by SDS-PAGE and in-gel fluorescence. As shown in Figure 3(B), both reversible inhibitors **Gü1303** and **Gü2602** strongly diminished the mCatK labelling, demonstrating that these cyanohydrazides are cell-permeable compounds that effectively interact with the active form of CatK in the lysosomal/endosomal system.

In conclusion, the cyanohydrazide inhibitors were demonstrated to inhibit mCatK with subnanomolar potency and different binding kinetics, to suppress the generation of mCatK from its zymogen, and to effectively target active mCatK in the cell context.

3.2. Crystallography of complexes of two cyanohydrazide inhibitors with mature cathepsin K and its zymogen-like activation intermediate

Human mature cathepsin K (mCatK) and a zymogen-like activation intermediate of cathepsin K (iCatK) were crystallised in complexes with two cyanohydrazide inhibitors, the azadipeptide nitrile **Gü1303** and the 3-cyano-3-aza- β -amino acid **Gü2602** (Figures 4 and 5). The crystal structures of mCatK complexes with **Gü1303** and **Gü2602** were determined by molecular replacement using the structure of uncomplexed mCatK as a template. Both complexes crystallised in the orthorhombic space group $P2_12_12$ with one molecule in the asymmetric unit (Table S1). The structure of the mCatK-**Gü1303** complex was refined using data to resolution 1.55 Å, and to 2.00 Å for mCatK-**Gü2602**. The final crystallographic models contained mCatK residues Ala1 to Met215; additional N-terminal residues derived from the propeptide, Gly-Arg (98p–99p, propeptide numbering), were visible in the mCatK-**Gü1303**

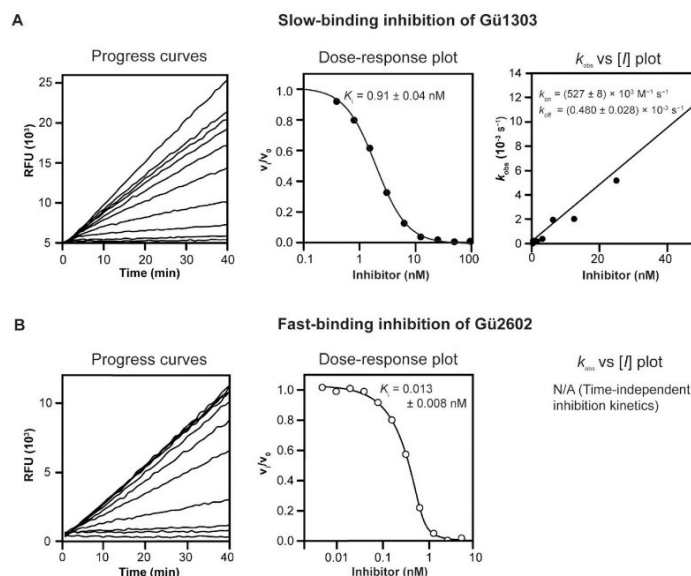


Figure 2. Different inhibition kinetics of mature cathepsin K with **Gü1303** and **Gü2602**. Progress curves show the hydrolysis of the fluorogenic substrate Cbz-Gly-Pro-Arg-AMC by mCatK at pH 5.5 in the presence of increasing inhibitor concentrations. (A) **Gü1303** exhibited a time-dependent inhibition characterised by non-linear progress curves typical of slow-binding kinetics. (B) Linear progress curves obtained for **Gü2602** are characteristic of fast-binding inhibitors. In dose-response plots, the derived steady-state reaction velocities were plotted against inhibitor concentration, and the inhibition constants K_i were obtained after correction by the Cheng-Prusoff and Morrison equations (see Materials and Methods). In the k_{obs} versus $[I]$ plot, the first-order rate constants k_{obs} from the time-dependent progress curves were plotted against inhibitor concentrations to show a linear dependence.

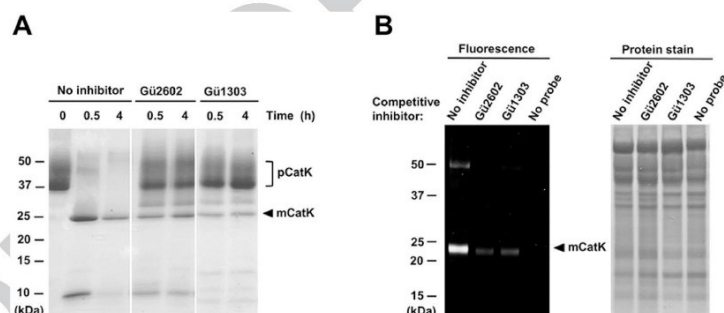


Figure 3. The inhibitors **Gü1303** and **Gü2602** suppress the autocatalytic activation of the cathepsin K zymogen and target cathepsin K in cells. (A) The zymogen of cathepsin K (pCatK) was incubated in the presence and absence of inhibitor (10 μ M) at pH 4.0, and the generation of mature cathepsin K (mCatK) was analysed at the indicated times. The reaction mixture was resolved by SDS-PAGE and visualised by protein staining. The positions of pCatK and mCatK are indicated; note mass heterogeneity of pCatK due to glycosylation⁵¹. (B) The U-2 OS cells were pre-treated with inhibitor (1 μ M) for 3 h, followed by 24 h incubation with a fluorescent activity-based probe specific for cathepsin K⁴⁰ (1 μ M); quenching of the labelling reaction by competitive inhibition was analysed. Cell lysates were resolved by SDS-PAGE and visualised by fluorescence imaging (left) and protein staining (right). The position of mCatK is indicated. In control experiments, the probe or inhibitor was omitted.

structure (Figure 5). A comparison of both mCatK complexes did not reveal any significant differences in protein structure (a backbone r.m.s.d. of 0.51 Å, a value within the range observed for different crystal structures of identical proteins).

iCatK was produced by limited autocatalytic processing of the zymogen pCatK at acidic pH. This resulted in fragmentation of the propeptide with cleavage sites identified after Glu4p, Ser83p, and Arg86p residues by Edman sequencing and mass spectrometry peptide mapping analysis (Figure 5(A)). The obtained iCatK comprised two non-covalently bound domains (chains), namely the

residual propeptide domain (Glu5p to Ser83p) and the main domain (Ser87p to Met215), including the C-terminal part of the propeptide and mCatK. The structure of the iCatK-**Gü2602** complex was solved by a combination of molecular replacement for the main domain (see previous paragraph) and automated model building by fragment-fitting technique for the residual propeptide. The structure of the iCatK-**Gü1303** complex was solved by molecular replacement based on the iCatK-**Gü2602** structure. Both iCatK complexes crystallised in the tetragonal space group $P4_32_12$ containing one molecule in the asymmetric unit (Table S1).

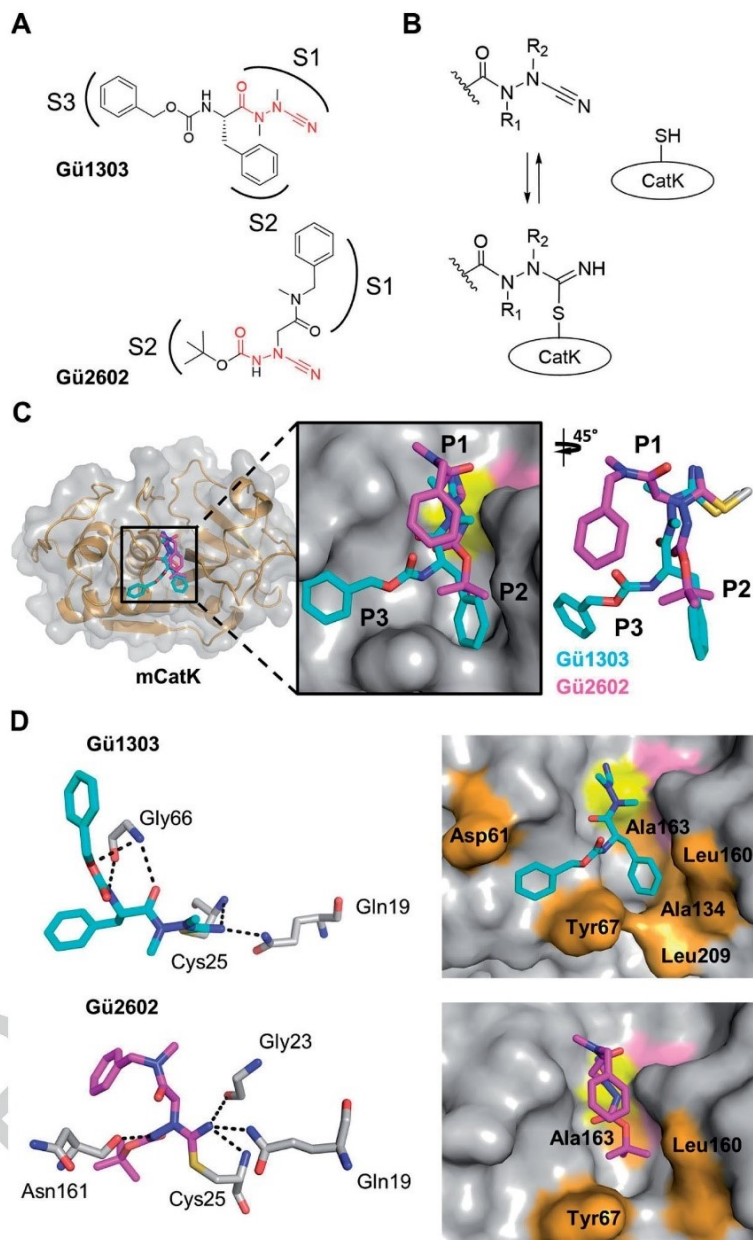


Figure 4. Binding mode of the cyanohydrazone inhibitors Gü1303 and Gü2602 in the active site of mature cathepsin K. (A) Chemical structure of the azadipeptide nitrile inhibitor Gü1303 and 3-cyano-3-aza- β -amino acid inhibitor Gü2602; the binding subsites (S) are marked, and the cyanohydrazone warheads are in red. (B) Reactive warheads form a covalent reversible bond with the thiol of the catalytic cysteine residue of the enzyme; R₁ and R₂ are substituents on the N atoms of the warheads. (C) The zoomed-in view of the mCatK active site shows a superposition of the inhibitors bound to the S1 to S3 subsites (corresponding inhibitor positions P1 to P3 are indicated). mCatK is displayed in surface representation (grey); highlighted are the catalytic residues Cys25 (yellow) and His162 (pink). Inhibitors are shown in stick representation with carbon atoms in cyan for Gü1303 and magenta for Gü2602; heteroatoms have standard colour coding (O, red; N, blue; S, yellow). (D) Interaction of the inhibitors with active site residues of mCatK. Left panels: the hydrogen bond network formed between inhibitors and mCatK residues (dashed black lines). Inhibitors are coloured as in (C), and interacting enzyme residues are in grey; the side chain of the covalently linked catalytic cysteine residue Cys25 is depicted. Right panels: the surface representation of the mCatK active site shows enzyme residues (highlighted in orange) forming nonpolar interactions with the inhibitors (in stick representation); both inhibitors are in the same orientation.

The final crystallographic model of the iCatK-**Gü1303** complex contained propeptide residues Asp8p to Gln73p and main domain residues Pro2 to Met215. The model of the iCatK-**Gü2602** complex contained residues Asp8p to Lys74p and Arg99p to Met215. The structures were refined using data to resolution 1.90 Å for iCatK-**Gü1303** and 1.88 Å for iCatK-**Gü2602**. A comparison of iCatK complexes did not reveal any significant differences in protein structure (a backbone r.m.s.d. of 0.28 Å).

The residual propeptide domain of iCatK is folded in a similar manner to the intact propeptide in the structure of the zymogen pCatK (PDB: 1BY8) and bound at the same position (a backbone r.m.s.d. of 0.81 and 0.48 Å for the propeptide domains and catalytic domains, respectively). However, we observed a slightly different orientation of the α 2p helix in iCatK complexes, which is rotated by approximately 10° compared to pCatK. The propeptide segment that blocks the active site in pCatK (downstream of the α 3p helix, Figure 5) is proteolytically removed or flexible in iCatK, and therefore the active site cleft of iCatK becomes accessible for inhibitors. However, the residual propeptide domain partially occludes the primed region of the active site, in particular the S1' subsite of iCatK is occupied.

3.3. Interaction of cyanohydrate inhibitors with the active site of cathepsin K

3.3.1. Binding mode of the azadipeptide nitrile **Gü1303** to mature and zymogen-like cathepsin K

The active site cleft of mCatK contains the catalytic triad residues Cys25, His162, and Asn182. **Gü1303** is bound in a substrate-like orientation, and its P1 to P3 residues occupy the S1 to S3 subsites of mCatK (Figure 4(A,C)). The cyanohydrate warhead reacts with the thiol group of the catalytic Cys25, forming a covalent isothiosemicarbazide adduct through the connection to the C-atom of the nitrile moiety (Figure 4(B)). Azadipeptides such as **Gü1303** are atropochiral molecules due to the restricted rotation around the methylated N–N axis, and, in the unbound state, they preferentially adopt the *E*-configuration of the respective CO–NMe bond^{63,64} (Figure 4(B)). However, a *Z*-configuration at the CO–NMe bond was observed in **Gü1303** bound to mCatK, suggesting an *E*-to-*Z* conformational change in **Gü1303** upon binding to the enzyme, most likely due to a "configurational selection" that we recently reported for an azadipeptide nitrile inhibitor of the protease SmCB1³².

Interactions between **Gü1303** and the mCatK active site are presented in Figure 4(D) (for details see Table S2). The inhibitor forms a network of hydrogen bonds with the active site residues. The nitrogen atom of the imidate moiety, derived from the warhead nitrile group, is stabilised by two hydrogen bonds to the backbone amide of the catalytic Cys25 and the side chain amide of Gln19 (Figure 4(D)). An analogous interaction pattern was observed for the warhead of an azadipeptide nitrile inhibitor reacted with the protease SmCB1 (PDB: 6YI7)³². The NH of Gly66 acts as a bifurcated hydrogen bond donor for the carbonyl oxygen of the P2 phenylalanine and the noncarbonyl carbamate oxygen in the P3 position of **Gü1303**. An additional hydrogen bond is formed between the Gly66 oxygen and the carbamate NH of **Gü1303**. A similar network of Gly66-mediated hydrogen bonds was also reported for a Boc-protected precursor of a reactive activity-based probe for mCatK (PDB 7NXL)⁴⁰. Nonpolar interactions of **Gü1303** with mCatK are depicted in Figure 4(D). They are absent in the S1 subsite, although the warhead containing the P1 azaalanine residue forms a number of contacts (Supporting information Table S2). At the P2 position, the phenylalanine residue

makes nonpolar interactions with Tyr67, Ala134, Leu160, Ala163, and Leu209. The P3 benzyloxycarbonyl capping group forms nonpolar interactions with Asp61 and Tyr67, and the phenyl moiety of this group is stabilised by a T-shaped π - π stacking interaction to the 4-hydroxyphenyl group of Tyr67.

The binding mode of **Gü1303** in the active site of iCatK is analogous to that in mCatK, and there is no interaction of the inhibitor with the residual propeptide domain that blocks a part of the primed region of the iCatK active site (Figure 5(B)). High conformational similarity of the inhibitor is indicated by a r.m.s.d. of 0.19 Å, with a certain change in the position of the P3 carbonyl oxygen. Also, **Gü1303** forms the same network of hydrogen bonds and nonpolar interactions in the S1 to S3 subsites of iCatK and mCatK (Figure 5(C), Table S2). However, no π - π stacking interaction with Tyr67 residue was observed, due to a slightly different orientation of the benzyloxycarbonyl capping group of **Gü1303**.

3.3.2. Binding mode of the 3-cyano-3-aza- β -amino acid **Gü2602** to mature and zymogen-like cathepsin K

Gü2602 binds to the active site of mCatK in a substrate-like orientation, and the major conformation of the inhibitor occupies the S1 and S2 subsites (Figure 4(A,C)). The cyanohydrate warhead positioned centrally in the inhibitor molecule forms a covalent isothiosemicarbazide adduct with a thiol group of the catalytic Cys25 (Figure 4(B)). There is a shift in the inhibitor backbone when comparing **Gü2602** and **Gü1303** that might be the result of nitrogen methylation in the warhead (CO–NH vs. CO–NMe in **Gü2602** and **Gü1303**, respectively) (Figure 4(C)). The CO–NH bond of **Gü2602** adopts the *Z*-configuration, similar to what has been observed for the CO–NMe bond of the enzyme bound inhibitor **Gü1303**.

The nitrogen atom of the imidate moiety of **Gü2602** is strongly stabilised by three hydrogen bonds to the backbone amide of the catalytic Cys25, carbonyl group of Gly23, and the side chain amide of Gln19 (Figure 4(D)). In contrast to **Gü1303**, there is no hydrogen bonding between **Gü2602** and the Gly66 residue. However, a new hydrogen bond is formed between the backbone oxygen of Asn161 and the amide NH of **Gü2602**. Analysis of the inhibitor-mCatK complexes available in the PDB shows that peptidomimetic inhibitors of mCatK frequently establish hydrogen bonding with Asn161 as well as Gly66 as important interaction determinants. The P1 and P2 residues of **Gü2602** are located in the non-primed subsites of the mCatK active site as follows (Figure 4(A,C)). The *N*-benzyl-*N*-methylacetamide substructure at P1 occupies the S1 subsite, making contacts with Gly23, Gly64, and Gly65 residues. Its terminal benzyl moiety is oriented out of the S1 subsite and towards the S2 pocket. The P2 Boc-capping group resides in the S2 subsite of mCatK and makes nonpolar interactions with Tyr67, Leu160, and Ala163 (Figure 4(D), Table S2).

In general, **Gü2602** binds to the S1 and S2 subsites of iCatK in a manner that is similar to what has been shown for the **Gü2602**-mCatK complex, with a r.m.s.d. of 1.23 Å; this value is substantially increased compared to **Gü1303** in its complexes, with a r.m.s.d. of 0.19 Å (Figure 5(B)). The major conformational change, however, is in the benzyl group of the P1 *N*-benzyl-*N*-methylacetamide substructure. It is rotated towards the residual propeptide domain of iCatK and forms new contacts with its Val72p and Gln73p residues (Figure 5(A,B), Table S2). The terminal Boc group accommodates the S2 pocket of iCatK analogously as in mCatK. The network of hydrogen bonds is identical, and the pattern of nonpolar interactions (Tyr67, Asn161, and Ala163 in iCatK) is similar for both **Gü2602** complexes (Figure 5(C)).

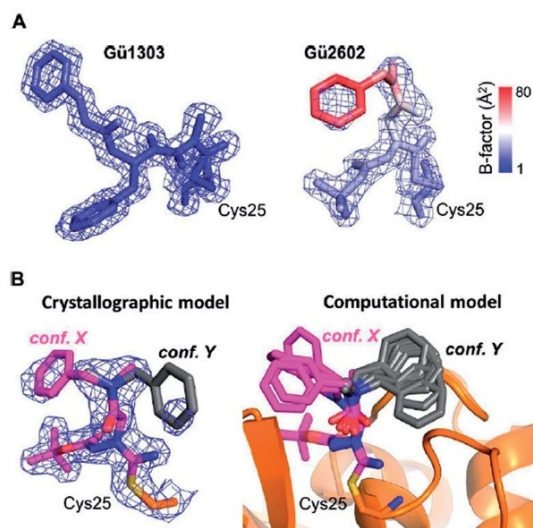


Figure 6. Analysis of conformational flexibility of **Gü1303** and **Gü2602** inhibitors in the active site of mature cathepsin K. (A) The inhibitors and the side chain of the covalently linked catalytic cysteine residue Cys25 are shown in stick representation. Their $2F_o - F_c$ electron density maps are contoured at 1σ and 1.5σ for **Gü1303** and **Gü2602**, respectively. Structures are coloured according to atomic B-factor values, from blue (low) to red (high). The highest B-factors indicating flexibility are observed for the benzyl moiety of **Gü2602**. (B) Conformational flexibility of the benzyl moiety of **Gü2602**. Left panel: Crystallographic model of **Gü2602** with the major observed conformation of the benzyl moiety in magenta (conf. X located in the non-primed area) and a predicted alternative minor conformation in grey (conf. Y located in the primed area). The side chain of the covalently linked catalytic Cys25 is in orange; heteroatoms have a standard colour coding (O, red; N, blue; S, yellow). The $2F_o - F_c$ electron density maps are contoured at 1σ . Right panel: Conformations of the benzyl moiety were searched by the molecular dynamics/quenching (MD/Q) technique and optimised by the semiempirical quantum mechanical (SQM) method. The lowest-energy conformations up to relative Gibbs "free" energy of 5 kcal/mol are shown for the general orientations X and Y (as defined by crystallography in left panel).

3.4. Conformational flexibility of cyanohydrzide inhibitors in the cathepsin K active site

The crystallographic electron density maps used for modelling **Gü1303** and **Gü2602** in the active site of mCatK were of high quality except for a weak electron density signal of the benzyl group of **Gü2602** (Figure 6(A)). This prompted us to analyse the B-factor distribution in both inhibitors. The B-factor values were generally in the low range, but increased values were only found for the P1 *N*-benzyl-*N*-methylacetamide part of **Gü2602**, with the highest value for its benzyl group (Figure 6(A)). An analogous B-factor pattern was also observed for inhibitors in iCatK complexes (Figure S1). This finding indicated an increased flexibility and dynamic disorder of the major crystallographic conformation of the benzyl group of **Gü2602**, which is located in the non-primed part of the active site and oriented out of the S1 subsite (Figure 5(C)).

A detailed inspection of electron density maps in the active site of the mCatK-**Gü2602** complex provided a further weak electron density signal of poor quality in the primed region of the mCatK active site that can be assigned to **Gü2602**. Based on this signal, we attempted to model an alternative conformation of the benzyl group of **Gü2602**, which is oriented towards the S2' subsite as presented in Figure 6(B) (see conformation Y). For this

purpose, we employed molecular modelling to examine conformational space of the flexible *N*-benzyl-*N*-methylacetamide substructure of **Gü2602**. The molecular dynamics/quenching (MD/Q) technique was utilised to generate accessible conformations of this segment in the active site, and semiempirical quantum mechanical optimisation yielded a set of conformations with two general orientations into the non-primed and primed areas (the orientations are marked X and Y, respectively, in Figure 6(B)).

In conclusion, we demonstrated that **Gü2602** is more flexible in the enzyme active site in contrast to the rigid ligand **Gü1303**. The former contains a highly flexible benzyl moiety that is capable of adopting to two different types of conformations that reside in the non-primed part or primed part of the active site.

4. Discussion and conclusions

Cathepsin K (CatK) is a target for the treatment of osteoporosis, arthritis, and bone metastasis, and its potent and selective inhibitors are being intensively pursued as chemotherapeutics^{6,7,24}. In this study, we investigated peptidomimetic inhibitors with a cyanohydrzide warhead, a class of highly efficient inhibitors of CatK that have been recently discovered^{30,31,39}, yet for which the interaction mechanism at the atomic level with CatK is unknown. For two representative cyanohydrzide compounds, **Gü1303** and **Gü2602**, we present crystallographic analysis of the binding mode to mature enzyme mCatK and its zymogen-like activation intermediate iCatK, as well as functional analysis *in vitro* and in cells.

The crystal structures of the complexes of **Gü1303** or **Gü2602** with mCatK showed that the inhibitors are bound in a substrate-like orientation, and their cyanohydrzide warhead reacted with a thiol group of the catalytic Cys25, forming a covalent isothiosemicarbazide adduct. The CO-NMe and CO-NH bond (in **Gü1303** and **Gü2602**, respectively) of the warhead adopted a *Z*-configuration in the mCatK active site. This is in line with our recent analysis of another cyanohydrzide inhibitor (with the azadipeptide nitrile scaffold) in the active site of a cysteine protease³² which also demonstrated that the warhead with the methylated *N*-*N* axis provides atropochirality, and the *E*-configuration of the unbound inhibitor is transformed to a *Z*-configuration upon binding³². Therefore, such an *E*- to *Z*-conformational change in the course of its interaction with CatK is proposed for **Gü1303**, containing the CO-NMe-NMe portion.

The cyanohydrzide warhead is differently positioned in the inhibitor scaffold of the azadipeptide nitrile **Gü1303** and 3-cyano-3-aza- β -amino acid **Gü2602**. This is reflected in the distribution of the binding subsites that are targeted by the inhibitors. **Gü1303** occupies the non-primed subsites S1 to S3, and B-factor distribution showed that **Gü1303** is rigid in the mCatK active site. In contrast, **Gü2602** primarily occupies the S2 subsite, and the *N*-benzyl-*N*-methylacetamide part (especially its benzyl moiety) is highly flexible and has two alternative types of conformations that reside in the non-primed part (the S1 subsite) or primed part (the S2' subsite) of the active site. The non-primed orientation is the major one observed in the crystal structure, where it might be stabilised by additional contacts, including crystal packing contacts (with a symmetry-related protein molecule) and intramolecular contacts. The primed orientation was clearly demonstrated by molecular modelling using MD/Q and SQM techniques and is further supported by the recent identification of a benzyl group oriented towards the S2' pocket in the activity-based probe for CatK⁴⁰. The computational approach showed a set of conformations for both types of orientations, further highlighting the high flexibility of this part of **Gü2602** in the enzyme active site.

We hypothesise that the conformational flexibility of the complexed ligand at the boundary of the non-primed and primed subsites provides an entropic advantage contributing to the extraordinary potency of **Gü2602** in the low picomolar range. It has been shown that the loss of configurational entropy upon ligand binding contributes significantly unfavourably to the binding free energy⁶⁵; this entropy loss is decreased for **Gü2602** making its binding more favourable. An analysis of active site interactions formed by **Gü2602** also identified an important role of hydrogen bonding between the amide NH of the warhead and Asn161 of the S2 subsite, an interaction that is absent in **Gü1303**. This is supported by analogues of **Gü2602** with methylated NH that exhibited a dramatic decrease in inhibitory potency by 3 orders of magnitude^{39,40}.

The differences in the binding mode between **Gü2602** and **Gü1303** are also reflected in their kinetic behaviour. The slow-binding of **Gü1303** is attributed to a conversion from an *E*- to *Z*-configuration upon binding that was recently described as the kinetic controlling step for a prototype azadipeptide nitrile with an atropochiral warhead³². In contrast, the 3-cyano-3-aza- β -amino acid scaffold of **Gü2602** bears a non-atropochiral warhead, as demonstrated by NMR studies showing the absence of diastereotopic methylene protons³⁹. Hence, the configuration conversion step does not occur, as is reflected by the fast-binding of **Gü2602** that resembles analogous kinetics and a reaction mechanism described for carbanitrile inhibitors³².

In addition to mCatK-inhibitor complexes, we also structurally characterised the complexes of **Gü2602** and **Gü1303** with the activation intermediate iCatK produced during autocatalytic processing of the inactive zymogen to mature enzyme that is associated with the removal of the propeptide. The iCatK structure is described here for the first time. It contains the main catalytic domain, which is non-covalently bound to the residual propeptide domain. Unlike the zymogen structure^{27,28}, the residual propeptide (residues 4p to 83p) of iCatK only partially occluded the active site cleft (mainly the S1' subsite), and thus the catalytic centre is accessible for ligands. This contrasts to the previously investigated activation intermediate of the cathepsin SmCB1, in which the catalytic centre is still blocked by the residual propeptide⁴¹.

iCatK was demonstrated to bind **Gü2602** and **Gü1303** in an analogous manner to that observed in the mCatK complexes. This indicated that the iCatK active site is competent to bind inhibitory ligands in the available subsites and can be regulated by them. This conclusion is supported by the results on testing **Gü2602** and **Gü1303** in an *in vitro* autoactivation assay with the zymogen pCatK. It showed that both inhibitors strongly suppress the formation of mCatK from the zymogen that proceeds as a bimolecular processing reaction catalysed by functional forms of iCatK/mCatK. Furthermore, we provided evidence that **Gü2602** and **Gü1303** are capable of effectively targeting mCatK as well as the mCatK-generating pathway in the pathophysiologically relevant context of osteosarcoma cells.

In conclusion, our work provides the first crystallographic, computational chemical, and functional insights into the binding mode of the cyanohydrizide inhibitors to the CatK target and will facilitate a further rational design of therapeutics against CatK-mediated pathologies.

Acknowledgements

The authors thank Helena Mertlíková-Kaiserová for providing U-2 OS cells.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the project ChemBioDrug CZ.02.1.01/0.0/0.0/16_019/0000729 from ERDF/OPRDE and the institutional project RVO 61388963. This work was also supported by Czech Academy of Sciences, Lumina quaeruntur.

Data availability statement

Diffraction data were collected on MX14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin. Atomic coordinates and experimental structure factors have been deposited in the Protein Data Bank with accession codes 7QBL, 7QBM, 7QBN, and 7QBO.

References

- Li H, Xiao Z, Quarles LD, Li W. Osteoporosis: mechanism, molecular target and current status on drug development. *Curr Med Chem* 2021;28:1489–507.
- Costa AG, Cusano NE, Silva BC, et al. Cathepsin K: its skeletal actions and role as a therapeutic target in osteoporosis. *Nat Rev Rheumatol* 2011;7:447–56.
- Makras P, Delaroudis S, Anastasilakis AD. Novel therapies for osteoporosis. *Metabolism* 2015;64:1199–214.
- Kramer L, Turk D, Turk B. The future of cysteine cathepsins in disease management. *Trends Pharmacol Sci* 2017;38:873–98.
- Drake MT, Clarke BL, Oursler MJ, Khosla S. Cathepsin K inhibitors for osteoporosis: biology, potential clinical utility, and lessons learned. *Endocr Rev* 2017;38:325–50.
- Brömme D, Lecaille F. Cathepsin K inhibitors for osteoporosis and potential off-target effects. *Expert Opin Investig Drugs* 2009;18:585–600.
- Lu J, Wang M, Wang Z, et al. Advances in the discovery of cathepsin K inhibitors on bone resorption. *J Enzyme Inhib Med Chem* 2018;33:890–904.
- Dejica VM, Mort JS, Laverty S, et al. Increased type II collagen cleavage by cathepsin K and collagenase activities with aging and osteoarthritis in human articular cartilage. *Arthritis Res Ther* 2012;14:R113–9.
- Yasuda Y, Kaleta J, Brömme D. The role of cathepsins in osteoporosis and arthritis: rationale for the design of new therapeutics. *Adv Drug Deliv Rev* 2005;57:973–93.
- Aguda AH, Panwar P, Du X, et al. Structural basis of collagen fiber degradation by cathepsin K. *Proc Natl Acad Sci USA* 2014;111:17474–9.
- Novinec M, Lenarčič B. Cathepsin K: a unique collagenolytic cysteine peptidase. *Biol Chem* 2013;394:1163–79.
- Novinec M, Kovacic L, Lenarčič B, Baici A. Conformational flexibility and allosteric regulation of cathepsin K. *Biochem J* 2010;429:379–89.
- Cherney MM, Lecaille F, Kienitz M, et al. Structure-activity analysis of cathepsin K/chondroitin 4-sulfate interactions. *J Biol Chem* 2011;286:8988–98.
- Verbovšek U, Van Noorden CJ, Lah TT. Complexity of cancer protease biology: cathepsin K expression and function in cancer progression. *Semin Cancer Biol* 2015;35:71–84.

15. Olson OC, Joyce JA. Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nat Rev Cancer* 2015;15:712–29.
16. Jiang H, Cheng XW, Shi GP, et al. Cathepsin K-mediated Notch1 activation contributes to neovascularization in response to hypoxia. *Nat Commun* 2014;5:3838.
17. Herroon MK, Rajagurubandara E, Rudy DL, et al. Macrophage cathepsin K promotes prostate tumor progression in bone. *Oncogene* 2013;32:1580–93.
18. Kamolmatyakul S, Chen W, Yang S, et al. IL-1 α stimulates cathepsin K expression in osteoclasts via the tyrosine kinase-NF- κ B pathway. *J Dent Res* 2004;83:791–6.
19. Asagiri M, Hirai T, Kunigami T, et al. Cathepsin K-dependent toll-like receptor 9 signaling revealed in experimental arthritis. *Science* 2008;319:624–7.
20. Hira VV, Ploegmakers KJ, Grevers F, et al. CD133+ and Nestin+ glioma stem-like cells reside around CD31+ arterioles in niches that express SDF-1 α , CXCR4, osteopontin and cathepsin K. *J Histochem Cytochem* 2015; 63: 481–93.
21. Calio A, Brunelli M, Gobbo S, et al. Cathepsin K: a novel diagnostic and predictive biomarker for renal tumors. *Cancers* 2021;13:2441.
22. Leusink FK, Koudounarakis E, Frank MH, et al. Cathepsin K associates with lymph node metastasis and poor prognosis in oral squamous cell carcinoma. *BMC Cancer* 2018;18:385.
23. Kos J, Lah TT. Cysteine proteinases and their endogenous inhibitors: target proteins for prognosis, diagnosis and therapy in cancer (review). *Oncol Rep* 1998;5:1349–410.
24. Lecaille F, Brömme D, Lalmanach G. Biochemical properties and regulation of cathepsin K activity. *Biochimie* 2008;90: 208–26.
25. McQueney MS, Amegadzie BY, D'Alessio K, et al. Autocatalytic activation of human cathepsin K. *J Biol Chem* 1997;272:13955–60.
26. Linnevers CJ, Mcgrath ME, Armstrong A, et al. Brömme, D, expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex. *Protein Sci* 1997;6:919–21.
27. Sivaraman J, Lalumière M, Ménard R, Cygler M. Crystal structure of wild-type human procathepsin K. *Protein Sci* 1999;8: 283–90.
28. LaLonde JM, Zhao B, Janson CA, et al. The crystal structure of human procathepsin K. *Biochemistry* 1999;38:862–9.
29. Lemaire PA, Huang L, Zhuo Y, et al. Chondroitin sulfate promotes activation of cathepsin K. *J Biol Chem* 2014;289: 21562–72.
30. Löser R, Frizler M, Schilling K, Gütschow M. Azadipeptide nitriles: highly potent and proteolytically stable inhibitors of papain-like cysteine proteases. *Angew Chem Int Ed Engl* 2008;47:4331–4.
31. Frizler M, Lohr F, Furtmann N, et al. Structural optimization of azadipeptide nitriles strongly increases association rates and allows the development of selective cathepsin inhibitors. *J Med Chem* 2011;54:396–400.
32. Jílková A, Horn M, Fanfrlík J, et al. Azanitrile inhibitors of the SmCB1 protease target are lethal to *Schistosoma mansoni*: structural and mechanistic insights into chemotype reactivity. *ACS Infect Dis* 2021;7:189–201.
33. Yang PY, Wang M, Li L, et al. Design, synthesis and biological evaluation of potent azadipeptide nitrile inhibitors and activity-based probes as promising anti-*Trypanosoma brucei* agents. *Chemistry* 2012;18:6528–41.
34. Breidenbach J, Lemke C, Pillaiyar T, et al. Targeting the main protease of SARS-CoV-2: from the establishment of high throughput screening to the design of tailored inhibitors. *Angew Chem Int Ed Engl* 2021;60:10423–9.
35. Chingle R, Proulx C, Lubell WD. Azapeptide synthesis methods for expanding side-chain diversity for biomedical applications. *Acc Chem Res* 2017;50:1541–56.
36. Proulx C, Sabatino D, Hopewell R, et al. Azapeptides and their therapeutic potential. *Future Med Chem* 2011;3: 1139–64.
37. Verhelst SH, Witte MD, Arastu-Kapur S, et al. Novel aza peptide inhibitors and active-site probes of papain-family cysteine proteases. *ChemBioChem* 2006;7:943–50.
38. Sexton KB, Kato D, Berger AB, et al. Specificity of aza-peptide electrophile activity-based probes of caspases. *Cell Death Differ* 2007;14:727–32.
39. Schmitz J, Beckmann AM, Dudic A, et al. 3-Cyano-3-aza- β -amino acid derivatives as inhibitors of human cysteine cathepsins. *ACS Med Chem Lett* 2014;5:1076–81.
40. Lemke C, Benýšek J, Brajtenbach D, et al. An activity-based probe for cathepsin K imaging with excellent potency and selectivity. *J Med Chem* 2021;64:13793–806.
41. Jílková A, Horn M, Řezáčová P, et al. Activation route of the *Schistosoma mansoni* cathepsin B1 drug target: structural map with a glycosaminoglycan switch. *Structure* 2014;22: 1786–98.
42. Horn M, Jílková A, Vondrášek J, et al. Mapping the pro-peptide of the *Schistosoma mansoni* cathepsin B1 drug target: modulation of inhibition by heparin and design of mimetic inhibitors. *ACS Chem Biol* 2011;6:609–17.
43. Mueller U, Förster R, Hellmig M, et al. The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *Eur Phys J Plus* 2015;130:141.
44. Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr* 2010;66: 125–32.
45. Vagin A, Teplyakov A. An approach to multi-copy search in molecular replacement. *Acta Crystallogr D Biol Crystallogr* 2000;56:1622–4.
46. Winn MD, Ballard CC, Cowtan KD, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 2011;67:235–42.
47. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* 2006;62:1002–11.
48. Long F, Nicholls RA, Emsley P, et al. AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol* 2017;73:112–22.
49. Davis IW, Leaver-Fay A, Chen VB, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 2007;35:W375–83.
50. Salentin S, Schreiber S, Haupt VJ, et al. PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res* 2015;43: W443–7.
51. Case DA, Babin V, Berryman JT, et al. AMBER 14. San Francisco (USA): University of California San Francisco; 2014.
52. Stewart JJ. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J Mol Model* 2007;13:1173–213.
53. Řezáč J, Hobza P. Advanced corrections of hydrogen bonding and dispersion for semiempirical quantum mechanical methods. *J Chem Theory Comput* 2012;8:141–51.

54. Klamt A, Schuurmann G. Cosmo - a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J Chem Soc Perk T 2* 1993;799–805.
55. Kříž K, Řezáč J. Reparametrization of the COSMO solvent model for semiempirical methods PM6 and PM7. *J Chem Inf Model* 2019;59:229–35.
56. Řezáč J. Cuby: an integrative framework for computational chemistry. *J Comput Chem* 2016;37:1230–7.
57. Stewart JJP. Optimization of parameters for semiempirical methods IV: extension of MNDO, AM1, and PM3 to more main group elements. *J Mol Model* 2004;10:155–64.
58. Kuzmic P, Sideris S, Cregar LM, et al. High-throughput screening of enzyme inhibitors: automatic determination of tight-binding inhibition constants. *Anal Biochem* 2000;281:62–7.
59. Horn M, Nussbaumerová M, Šanda M, et al. Hemoglobin digestion in blood-feeding ticks: mapping a multipetidase pathway by functional proteomics. *Chem Biol* 2009;16:1053–63.
60. Barrett AJ, Kembhavi AA, Brown MA, et al. L-trans-Epoxy succinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochem J* 1982;201:189–98.
61. Billington CJ, Mason P, Magny MC, Mort JS. The slow-binding inhibition of cathepsin K by its propeptide. *Biochem Biophys Res Commun* 2000;276:924–9.
62. Husmann K, Muff R, Bolander ME, et al. Cathepsins and osteosarcoma: expression analysis identifies cathepsin K as an indicator of metastasis. *Mol Carcinog* 2008;47:66–73.
63. Ottersbach PA, Schnakenburg G, Gütschow M. Induction of chirality: experimental evidence of atropisomerism in azapeptides. *Chem Commun* 2012;48:5772–4.
64. Ottersbach PA, Schnakenburg G, Gütschow M. Atropisomerism in azadipeptides: evaluation of N1-methylation and thioamide introduction. *Tetrahedron Lett* 2015;56:4889–91.
65. Chia-en AC, Chen W, Gilson MK. Ligand configurational entropy and protein binding. *Proc Natl Acad Sci USA* 2007;104:1534–9.

PROOF ONLY

Supplementary materials

Highly potent inhibitors of cathepsin K with a differently positioned cyanohydrazone warhead: structural analysis of binding mode to mature and zymogen-like enzyme

Jakub Benýšek^{a,b}, Michal Buša^{a,c}, Petra Rubešová^a, Jindřich Fanfrlík^a, Martin Lepšík^a, Jiří Brynda^a, Zuzana Matoušková^a, Ulrike Bartz^d, Martin Horn^a, Michael Gütschow^e, Michael Mareš^{*a}

^aInstitute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n. 2, 16610 Prague, Czech Republic

^bFirst Faculty of Medicine, Charles University, Kateřinská 1660/32, 12108 Prague, Czech Republic

^cDepartment of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, 12800 Prague, Czech Republic

^dDepartment of Natural Sciences, University of Applied Sciences Bonn-Rhein-Sieg, von-Liebig-Strasse 20, D-53359 Rheinbach, Germany

^ePharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany

*Corresponding author: mares@uochb.cas.cz

Table S1. X-ray data collection and refinement statistics

CatK-inhibitor complex ^a	mCatK-Gü2602	mCatK-Gü1303	iCatK-Gü2602	iCatK-Gü1303
Data Collection Statistics				
wavelength (Å)	1.542	1.542	0.918	0.918
temperature (K)	100	100	100	100
space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 4 ₃ 2 ₁ 2	<i>P</i> 4 ₃ 2 ₁ 2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	70.63, 78.52, 31.02	74.94, 76.12, 32.85	103.28, 103.28, 55.46	103.21, 103.21, 54.99
α , β , γ (deg)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
resolution (Å)	50.0–2.00 (2.12–2.00)	50.0–1.55 (1.64–1.55)	50.0–1.88 (1.92–1.88)	50.0–1.90 (2.02–1.90)
number of unique reflections	11 829 (1880)	23 646 (1843)	24 933 (1541)	23 911 (3763)
redundancy	4.1 (4.0)	6.5 (4.7)	11.1 (4.8)	7.9 (7.5)
completeness (%)	96.5 (97.8)	83.4 (41.0)	99.7 (96.4)	99.8 (99.3)
R_{merge}^b (%)	7.8 (59.2)	8.5 (38.7)	18.7 (214.1)	11.1 (63.0)
average I/σ (I)	12.4 (2.5)	19.4 (4.4)	7.9 (0.7)	13.3 (2.9)
CC _{1/2} ^c (%)	99.8 (88.9)	99.8 (89.1)	97.2 (31.5)	99.8 (85.3)
Wilson B (Å ²)	35.06	15.94	28.40	27.48
Refinement Statistics				
resolution range (Å)	28.90–2.0 (2.05–2.0)	37.50–1.55 (1.59–1.55)	48.91–1.88 (1.93–1.88)	46.20–1.90 (1.95–1.90)
number of reflections in working set	11 344 (836)	22 487 (439)	23 673 (1662)	22 715 (1633)
number of reflections in test set	598 (44)	1184 (23)	1213 (84)	1196 (86)
<i>R</i> value ^d (%)	22.1 (35.5)	14.4 (23.8)	19.0 (37.4)	16.4 (25.5)
R_{free} value ^e (%)	28.6 (34.5)	16.6 (26.0)	23.4 (37.5)	21.2 (29.1)
number of molecules in AU ^f	1	1	1	1
number of atoms in AU ^g protein/inhibitor/solvent	1657/23/47	1705/27/302	2226/23/235	2226/27/293
average ADP ^g for protein/inhibitor/solvent (Å ²)	38.0/43.2/33.0	12.9/14.8/24.3	36.9/49.2/40.5	24.0/29.7/33.2
RMSD bond length (Å)	0.014	0.015	0.014	0.014
RMSD bond angle (deg)	1.91	1.77	1.74	1.68
Ramachandran plot statistics ^h				
favoured regions (%)	95.8	97.7	95.3	97.9
allowed regions (%)	4.2	2.3	4.7	2.2
PDB code	7QBL	7QBN	7QBM	7QBO

^aNumbers in parentheses refer to the highest-resolution shell. ^b $R_{\text{merge}} = 100 \sum_{hkl} \sum_j |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_j I_i(hkl)$, where $I_i(hkl)$ is an individual intensity of the i^{th} observation of reflection hkl and $\langle I(hkl) \rangle$ is the average intensity of reflection hkl with summation over all data. ^cCC_{1/2} is the percentage of correlation between intensities from random half-datasets. ^d R -value = $|F_o| - |F_c| / |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. ^e R_{free} is equivalent to R value but is calculated for up to 5% of the reflections chosen at random and omitted from the refinement process. ^fAU, asymmetric unit. ^gADP, atomic displacement parameter, formally B-factor. ^hAs determined by Molprobitry.³

Table S2. List of contacts formed between the inhibitors **Gü1303** and **Gü2602** and mature cathepsin K (mCatK) or the activation intermediate of cathepsin K (iCatK). Analysis of protein-inhibitor contacts between the mCatK/iCatK active site and inhibitors was performed using program CONTACT.⁴ The distance cutoffs were set to 4.2 Å for all contacts and 3.3 Å for hydrogen bonds. The enzyme residues interacting in the individual inhibitor positions (P3 to P1) are specified. For each enzyme residue, total number of contacts (C) are listed, including hydrogen bonds (Hb) (residues forming hydrogen bonds are in bold).

Subsite	mCatK-Gü2602			mCatK-Gü1303			iCatK-Gü2602			iCatK-Gü1303		
	Residue	C	Hb	Residue	C	Hb	Residue	C	Hb	Residue	C	Hb
P1							Val72p	10				
							Gln73p	2				
	Gln19	2	1	Gln19	3	1	Gln19	2	1	Gln19	3	1
	Gly23	10	1	Gly23	10		Gly23	9	1	Gly23	10	
	Ser24	2		Ser24	3		Ser24	2		Ser24	3	
	Cys25	13	1	Cys25	14	1	Cys25	13	1	Cys25	14	1
	Trp26	2		Trp26	3		Trp26	1		Trp26	3	
	Cys63	1					Cys63	1				
	Gly64	7					Gly64	4				
	Gly65	2					Gly65	4				
Asn161	10	1	Asn161	2		Asn161	6	1	Asn161	2		
His162	3		His162	2		His162	3		His162	2		
P2	Cys25	1		Cys25	2		Cys25	2		Cys25	2	
	Gly65	2		Gly65	6		Gly65	2		Gly65	6	
	Gly66	5		Gly66	11	2	Gly66	5		Gly66	11	2
	Tyr67	2		Tyr67	3		Tyr67	2		Tyr67	3	
				Met68	2					Met68	2	
				Ala134	3					Ala134	3	
	Leu160	3		Leu160	4		Leu160	3		Leu160	4	
	Asn161	6					Asn161	9				
	His162	3					His162	3				
	Ala163	1		Ala163	1		Ala163	1		Ala163	1	
P3				Glu59	3					Glu59	3	
				Asn60	3					Asn60	3	
				Asp61	5					Asp61	5	
				Gly65	3					Gly65	3	
				Gly66	15	1				Gly66	15	1
				Tyr67	11					Tyr67	11	

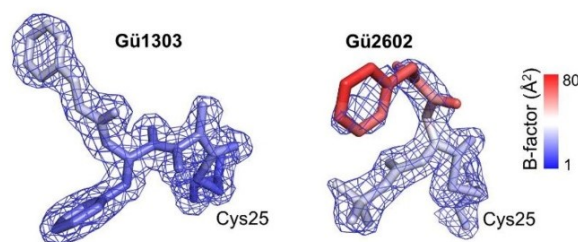


Figure S1. Analysis of conformational flexibility of the inhibitors **Gü1303** and **Gü2602** in the active site of the activation intermediate of cathepsin K. The inhibitors and the side chain of the covalently linked catalytic cysteine residue Cys25 are shown in stick representation. Their $2F_o - F_c$ electron density maps are contoured at 1σ and 1.5σ for **Gü1303** and **Gü2602**, respectively. Structures are coloured according to atomic B-factor values, from blue (low) to red (high). The highest B-factors indicating flexibility are observed for the benzyl moiety of **Gü2602**.

References

1. Karplus, P. A.; Diederichs, K., Linking crystallographic model and data quality. *Science* **2012**, *336* (6084), 1030-1033.
2. Brünger, A. T., Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **1992**, *355* (6359), 472-475.
3. Lovell, S. C.; Davis, I. W.; Arendall, W. B., 3rd; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C., Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins Struct Funct Bioinform* **2003**, *50* (3), 437-450.
4. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 4), 235-242.