BIOCHEMICAL AND STRUCTURAL COMPARISON OF HUMAN AND MOUSE CATHEPSIN B COMPLEXES WITH DARPINS

Miki Zarić

Doctoral Dissertation
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia

Supervisor: Prof. Dr. Olga Vasiljeva, Jožef Stefan Institute, Ljubljana, Slovenia
Co-Supervisor: Prof. DDr. Boris Turk, Jožef Stefan Institute, Ljubljana, Slovenia

Evaluation Board:
Prof. Dr. Veronika Stoka, Chair, Jožef Stefan Institute, Ljubljana, Slovenia
Assoc. Prof. Dr. Toni Petan, Member, Jožef Stefan Institute, Ljubljana, Slovenia
Assist. Prof. Dr. Gregor Gunčar, Member, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia

MEDNARODNA PODIPLOMSKA ŠOLA JOŽEFA STEFANA
JOŽEF STEFAN INTERNATIONAL POSTGRADUATE SCHOOL



Miki Zarić

BIOCHEMICAL AND STRUCTURAL COMPARISON OF HUMAN AND MOUSE CATHEPSIN B COMPLEXES WITH DARPINS

Doctoral Dissertation

BIOKEMIJSKA IN STRUKTURNA PRIMERJAVA KOMPLEKSOV ČLOVEŠKEGA IN MIŠJEGA KATEPSINA B Z DARPIN-I

Doktorska disertacija

Supervisor: Prof. Dr. Olga Vasiljeva

Co-Supervisor: Prof. DDr. Boris Turk

Ljubljana, Slovenia, May 2026

Acknowledgments

I would like to thank the Slovenian Research and Innovation Agency (ARIS) for funding my doctoral studies and the Jožef Stefan Institute, where I was able to conduct my work for my thesis.

I am grateful to my supervisor Prof. Dr. Olga Vasiljeva, for her guidance and supervision of my dissertation, advice regarding my work as well as for reviewing this thesis.

I would also like to thank my co-supervisor Prof. DDr. Boris Turk for accepting me into his research group and ensuring good working conditions despite everything, as well as for advice during the work.

I thank the members of the evaluation board Prof. Dr. Veronika Stoka, Assoc. Prof. Dr. Toni Petan, and Assist. Prof. Dr. Gregor Gunčar for lending their expertise and time in the revision of my dissertation.

I thank Prof. Dr. Dušan Turk for his help during the structural part of the work and for help with the interpretation of structural and cleavage comparisons.

I am grateful to Assist. Prof. Dr Livija Tušar for resolving the structure and interpretation of the mass spectrometry data and valuable discussions surrounding that.

I would like to thank Dr. Aleksandra Usenik for help with protein crystallisation experiments.

I am grateful to Ivica Štefe B.A. for her technical assistance and willingness to help in the lab.

I thank Urban Javoršek for help and discussions related to the field and all the colleagues from B1 department for their help, advice and for creating a pleasant and lively work environment.

Zahvaljujem se tudi mami Ani, očetu Radmilu, sestrama Merini in Martini ter dekletu Oliviji za vso podporo in spodbudo.

Abstract

Cathepsin B is a lysosomal cysteine protease implicated in various pathological conditions and therefore considered a potential therapeutic target. This study investigates the selectivity and binding of a designed ankyrin repeat protein (DARPin) 4m3 towards mouse cathepsin B (mCatB). It aims to explore the species-specific interaction of cathepsin B with DARPin molecules to gain insights into the differences with human cathepsin B (hCatB), including by comparing their cleavage site sequences with other human cathepsins using clustering and a support vector machine (SVM) algorithm. The mCatB-DARPin 4m3 complex was confirmed by size exclusion chromatography (SEC), nano-differential scanning fluorimetry (nano-DSF), and surface plasmon resonance (SPR) assay. The affinity of 4m3 for mCatB was determined by SPR (K_D = 65.7 nM), while the inhibition constant (K_i = 26.7 nM) and mode of inhibition were determined by kinetic analysis, which showed it to be hyperbolic mixed-type with a predominantly competitive inhibition. The crystal structure of the complex of mouse cathepsin B with DARPin 4m3 was determined for the first time at a resolution of 1.67 Å and compared with previously published structures of hCatB with DARPin 81 and DARPin 8h6, providing insight into the interactions that may be responsible for this selectivity. Structural analysis of the complexes revealed that DARPin 4m3 binds to the structurally equivalent region on mCatB that corresponds to the region targeted by DARPins 8h6 and 81 on hCatB. This region is less conserved among cathepsins across different species, and candidate amino acid interactions that may contribute to the selectivity of DARPin 4m3 were identified. DARPin 4m3 showed no binding to hCatB, which was confirmed by SEC, nano-DSF, SPR and inhibition assays. SVM-based analysis of cleavage-site sequences (P3–P4′) showed greater similarity between mCatB and hCatB (AUC = 0.73) than between mCatB and human CatK, CatL, CatV, CatS, or CatF (AUC = 0.94–0.97). The applied clustering and SVM-based comparison of cleavage-site sequences provide a framework that could be extended to other proteases and may help in the identification of selective substrates and protease-targeting tools. This study supports the close functional relationship between mouse and human CatB while demonstrating that species-specific differences can be exploited for selective targeting. The findings highlight the potential of DARPins as species-selective CatB inhibitors and provide structural and substrate-specificity insights relevant to the development of protease-targeting tools.

Povzetek

Katepsin B je lizosomska cisteinska proteaza, ki ima pomembno vlogo pri različnih patoloških procesih in zato predstavlja potencialno terapevtsko tarča. To delo obravnava selektivnost in vezavo sintetičnega mimetika protiteles na osnovi ankirinskih ponovitev (DARPin) 4m3 na mišji katepsin B (mCatB). Namen je raziskati vrstno specifične interakcije katepsina B z molekulami DARPin ter pridobiti vpogled v razlike v primerjavi s človeškim katepsinom B (hCatB), vključno s primerjavo njunih cepitvenih sekvenc z ostalimi človeškimi katepsini z uporabo gručenja in *support vector machine* (SVM) algoritma. Kompleks mCatB–DARPin 4m3 smo potrdili z gelsko izključitveno kromatografijo (SEC), nano-diferencialno vrstično fluorimetrijo (nano-DSF) in metodo površinske plazmonske resonance (SPR). Afiniteto 4m3 do mCatB smo določili z metodo SPR (K_D = 65,7 nM), medtem ko smo inhibicijsko konstanto (K_i = 26,7 nM) in način inhibicije določili s kinetično analizo, ki je pokazala hiperbolično mešano inhibicijo s pretežno kompetitivnim značajem. V tem delu smo tudi prvič določili kristalno strukturo mCatB v kompleksu z DARPinom 4m3 z ločljivostjo 1,67 Å. Strukturo kompleksa smo primerjali s prej objavljenimi strukturami hCatB v kompleksu z DARPinoma 81 in 8h6, kar je omogočilo vpogled v interakcije, ki bi lahko bile odgovorne za selektivnost. Strukturna analiza kompleksov je namreč pokazala, da se DARPin 4m3 veže na strukturno ekvivalentno regijo na mCatB, ki ustreza regiji, na katero se na hCatB vežeta DARPina 8h6 in 81. Ta regija je med katepsini različnih vrst manj ohranjena, pri čemer se je dalo identificirati aminokislinske interakcije, ki bi lahko prispevale k selektivnosti DARPina 4m3. DARPin 4m3 ni pokazal vezave na hCatB, kar smo potrdili z metodami SEC, nano-DSF, SPR in inhibicijskimi testi. Analiza cepitvenih zaporedij (P3–P4′) mCatB ter človeških katepsinov B, K, V, L, S in F z uporabo SVM je pokazala večjo podobnost med mCatB in hCatB (AUC = 0,73) kot med mCatB in ostalimi človeškimi katepsini (AUC = 0,94 – 0,97). Uporabljeno gručenje substratov mišjega in človeškega katepsina B ter njihova primerjava z algoritmom SVM predstavljata okvir, ki bi ga bilo mogoče uporabiti tudi pri drugih proteazah za identifikacijo selektivnih substratov in orodij za selektivno ciljanje. Rezultati podpirajo funkcionalno podobnost mišjega in človeškega katepsina B, hkrati pa kažejo, da je mogoče vrstno specifične razlike izkoristiti za selektivno ciljanje. Rezultati poudarjajo potencial proteinov DARPin kot vrstno selektivnih inhibitorjev katepsina B ter prispevajo k razumevanju strukturnih osnov interakcij in specifičnosti substratov, ki so pomembni za razvoj orodij za selektivno ciljanje proteaz.

Contents

List of Figures **xiii**

List of Tables **xv**

Abbreviations **xvii**

Symbols **xix**

Amino Acid Abbreviations **xxi**

1 Introduction **1**
1.1 Cysteine Cathepsins ... 2
1.2 Cysteine Cathepsins in Health ... 5
1.2.1 Intracellular role ... 5
1.2.2 Extracellular role ... 6
1.3 Cysteine Cathepsin B in Disease ... 6
1.3.1 Arthritis ... 7
1.3.2 Heart and lung diseases ... 7
1.3.3 Neurodegenerative diseases ... 7
1.3.4 Infections ... 8
1.3.5 Immune disorders and inflammation ... 8
1.3.6 Cancer ... 9
1.4 DARPins ... 10
1.4.1 DARPins in therapeutics and diagnostics ... 11
1.4.2 DARPins as research tools ... 12

2 Aims and Hypothesis **15**

3 Materials and Methods **17**
3.1 Materials ... 17
3.1.1 Laboratory equipment ... 17
3.1.2 Chemicals, commercial kits, and reagents ... 18
3.1.3 Buffers and growth mediums ... 19
3.1.4 Oligonucleotides ... 20
3.1.5 Vectors ... 20
3.1.6 Antibodies and enzymes ... 21
3.1.7 Bacterial and mammalian cell cultures ... 21
3.2 Methods ... 21
3.2.1 Selection of mouse cathepsin B-specific inhibitory DARPin ... 21
3.2.2 Molecular cloning of DARPin 4m3 sequence ... 21
3.2.3 Expression and purification of cathepsins ... 22
3.2.4 Expression and purification of DARPin 4m3 ... 23

3.2.5 Determination of protein concentration 23
3.2.6 Sodium dodecyl sulfate polyacrylamide electrophoresis 24
3.2.7 Western transfer and immunoblotting 24
3.2.8 Analytical size-exclusion chromatography 24
3.2.9 Inhibition assays 24
3.2.10 Nano-differential scanning fluorimetry 25
3.2.11 Surface plasmon resonance 25
3.2.12 Cellular thermal shift assay 25
3.2.13 Structure determination 25
3.2.14 Alignment and comparison of cathepsin B sequences from different animal species 26
3.2.15 Generation of datasets of human cathepsins K, V, L, S, F, and mouse cathepsin B substrates 26
3.2.16 Docking 26
3.2.17 Evaluation of the differences between the substrates of mouse cathepsin B and human cathepsins B, K, L, V, S, and F by support vector machine algorithm and clustering 27
3.2.18 Data availability 27

4 Results **29**
4.1 DARPin 4m3 Selectively Binds Mouse, but not Human, Cathepsin B 29
4.2 Binding Affinity and Mode of Inhibition by DARPin 4m3, Which is Retained in a Complex Sample 30
4.3 Crystal Structure Representation of the Complex of Mouse Cathepsin B with DARPin 4m3 35
4.4 Structural Comparison of the Cathepsins B, DARPins, and Their Complexes 37
4.5 Substrates as Indicators of Differences Between Mouse and Human CatB 45

5 Discussion **49**

6 Conclusions **53**

Appendix **55**
A.1 Appendix Figures 55
A.2 Appendix Tables 59

References **67**

Bibliography **83**

Biography **85**

List of Figures

Figure 1.1: Schematic representation of Schechter-Berger nomenclature of protease-substrate interaction with visualised exosite....2
Figure 1.2: Visual representation of a DARPin molecule....10
Figure 4.1: Interaction of purified DARPin 4m3 with mCatB and hCatB. 30
Figure 4.2: SPR and inhibition analysis of DARPin 4m3 binding and inhibition of mCatB and hCatB 31
Figure 4.3: Kinetics of inhibition of mCatB with DARPin 4m3. 33
Figure 4.4: CETSA melting curves (fitted with Boltzmann sigmoid function) for mCatB in RAW 264.7 cell lysates....34
Figure 4.5: Crystal structure of mouse cathepsin B in complex with DARPin 4m3....35
Figure 4.6: Crystal structure of mouse cathepsin B and DARPin 4m3 complex (molecules B and B2, shown in red) superimposed with DARPin 81 (cyan) and DARPin 8h6 (blue) 38
Figure 4.7: Interacting residues of mCatB/hCatB and DARPins....39
Figure 4.8: Structure-based alignment of mCatB, hCatB, and DARPins....40
Figure 4.9: Alignment of amino acid sequences of mature CatB from various animals ...41
Figure 4.10:3D visualisation of the interacting residues between mCatB/hCatB and DARPins 44
Figure 4.11:3D visualisation of DARPin molecules with cathepsins B, with surface coloured according to lipophilicity 45
Figure 4.12: Normality of substrate residue distributions 46
Figure 4.13: Calculated AUC values for pairwise comparisons of cleavage sites from different cathepsins....47
Figure 4.14: Functional enrichment analysis....48
Figure 6.1: Structure-based alignments of mCatB and hCatB and DARPins with MAIN with coloured differing residues....55
Figure 6.2: Structure-based alignments of mCatB and hCatB and DARPins with MAIN, with coloured differing residues according to chemical properties 56
Figure 6.3: ROC curves and calculated AUC values for the positions from P4 to P4′ for combinations of substrates....57
Figure 6.4: The positions S3–S4′ for the binding of substrates are shown in mCatB: prime sites S1′–S4′ (blue shading) and non-prime sites S1–S3 (green shading)....57
Figure 6.5: Docked complex of mutated mCatB (I65/Q66→S65/M66 substitution) and DARPin 4m3 58

List of Tables

Table 4.1: Kinetic and equilibrium constants for the interaction between mCatB and DARPin 4m3 32
Table 4.2: CETSA results showing temperatures at which mCatB reached 50% of its maximum band intensity 35
Table 4.3: Structural data of the mCatB and DARPin 4m3 complex 36
Table 4.4: Summary of the results of pairwise structure alignment 37
Table 4.5: Interaction energies were calculated for the mCatB–DARPin 4m3 complex (wild-type) and for docked models of the wild-type and two mutated complexes (I65/Q66→S65/M66) 42
Table 6.1: Comparison of the combinations of pairs of data sets of the substrates of human cysteine cathepsins K, V, B, L, S, F and mouse cathepsin B 59
Table 6.2: Separate clusters (336, 339, 389, 396) with only mCatB substrates 60
Table 6.3: Cluster scores (interaction energies between the two molecules) generated by the ClusPro server 61
Table 6.4: Calculated hydrogen bonds between mCatB and DARPin 4m3 are shown for the four top-ranked docking complexes generated by the ClusPro server 63
Table 6.5: Hydrogen bonds were calculated for the mutated mCatB–DARPin 4m3 complex (I65/Q66→S65/M66) using the two highest-ranked docked models 65

Abbreviations

AMC	...	7-amino-4-methylcoumarine
APS	...	Ammonium persulfate
ATP	...	Adenosine triphosphate
AUC	...	Area under the curve
Bis-Tris	...	2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol
BSA	...	Bovine serum albumin
CA074	...	[L-3-trans-(Propylcarbamoyl)oxirane-2-carbonyl]-L-isoleucyl-L-proline
DARPin	...	Designed ankyrin repeat protein
DMSO	...	Dimethyl sulfoxide
DNA	...	Deoxyribonucleic acid
dNTP	...	Deoxyribonucleotide triphosphate
DTT	...	Dithiothreitol
E. coli	...	*Escherichia coli*
E64	...	trans-Epoxysuccinyl-L-leucylamido(4-guanidino)butane
EDTA	...	Ethylenediaminetetraacetic acid
FBS-HI	...	Heat-inactivated fetal bovine serum
HEPES	...	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	...	Horseradish peroxidase
IEX	...	Ion exchange chromatography
IPTG	...	Isopropyl β-D-1-thiogalactopyranoside
LB	...	Lysogeny broth
MES	...	2-(N-morpholino)ethanesulfonic acid
MWCO	...	Molecular weight cut-off
Ni-NTA	...	Nickel-nitrilotriacetic acid
PBS	...	Phosphate-buffered saline
PCR	...	Polymerase chain reaction
PEG	...	Polyethylene glycol
SAPS-ESI	...	Statistical approach to peptidyl substrate-enzyme specific interactions
SDS	...	Sodium dodecyl sulfate
SDS-PAGE	...	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	...	Size-exclusion chromatography
SPR	...	Surface plasmon resonance
SVM	...	Support vector machine
TB	...	Terrific broth
TBS(T)	...	Tris-buffered saline (Tween-20 added)
TEMED	...	Tetramethylethylenediamine
Tris	...	Tris(hydroxymethyl)aminomethane
z	...	Benzyl chlorocarbonate

Symbols

Da	...	Dalton
E	...	Enzyme concentration
k_a	...	Association rate constant
k_{cat}	...	Catalytic constant
k_d	...	Dissociation rate constant
K_D	...	Dissociation constant
K_i	...	Inhibition constant
K_m	...	Michaelis-Menten constant
K_S	...	Equilibrium dissociation constants of the enzyme-substrate complex
RU	...	Resonance unit
S	...	Substrate concentration
U	...	Unit for enzyme activity
v_i	...	Velocity in presence of the inhibitor
v_o	...	Velocity of the uninhibited enzyme
α	...	Ratio between competitive and noncompetitive inhibition constant
β	...	Turnover rate of the inhibited enzyme compared to the non-inhibited enzyme (hyperbolic inhibition parameter)
σ	...	Ratio of substrate to Michaelis-Menten constant

Amino Acid Abbreviations

Amino acid	Three-letter abbreviation	One-letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamate	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1

Introduction

In the human genome, proteases occupy around 3 % of encoding regions, numbering over 600 members, which are divided into five different classes based on the amino acid in their active site that catalyses the cleavage of the peptide bond: metallo-, aspartyl-, serine-, threonine- and cysteine peptidases (Rawlings, Waller, Barrett, & Bateman, 2014). There are also representatives of two additional classes that are yet to be found in mammals, and those are glutamic peptidases and asparagine peptide lyases (Fujinaga, Cherney, Oyama, Oda, & James, 2004; Rawlings, Barrett, & Bateman, 2011). The latter ones differ from other proteolytic enzymes since their mechanism of cleavage does not use a water molecule as a nucleophile to attack the carbonyl, but rather an asparagine residue, which results in intramolecular cyclisation in the form of succinimide that is later resolved by cleaving away the other part of the polypeptide. Most proteases have molecular masses in the range of 20-200 kDa, but they can also form multi-domain complexes, such as the proteasome, with a molecular mass of 2 MDa. They can either cleave proteins within a sequence (endopeptidases) or at the termini (exopeptidases). Based on which end they process, exopeptidases can be further divided into carboxypeptidases (process C-termini) or aminopeptidases (process N-termini). Exopeptidases can cleave away one (monopeptidyl), two (dipeptidyl) or three amino acids (tripeptidyl). There is also a special case of oligopeptidases, which are endopeptidases that can cleave peptides but not proteins, due to the structure of the enzyme, which contains the active site at the end of a narrow cavity, so that only peptides short enough can reach.

Peptide bonds are very stable at neutral pH and have half-lives of centuries, which also enabled analysis of collagen from the mummified skin of a famous mummy, Oetzi, that was found in the Italian Alps (Janko, Zink, Gigler, Heckl, & Stark, 2010). In the presence of appropriate proteases and conditions that allow biochemical reaction, they are rapidly hydrolysed. Proteases bind to substrates through hydrophobic and electrostatic interactions with substrate amino acid side chains in so-called substrate binding pockets that define protease specificity. Additionally, they also bind the peptide backbone through hydrogen bonding, which furthermore stabilises the interaction.

A scissile bond is the bond where hydrolysis of the peptide bond happens. Amino acids around the scissile bond determine protease specificity since they are the ones responsible for binding to the active site area. In order to annotate positions around the scissile bond, Schechter-Berger nomenclature was developed (Figure 1.1). Subsites on the protease onto which the substrate binds are numbered S1-S_n towards the N-terminus of the substrate and S1'-S_n' towards the C-terminus of the substrate, while substrate residues that bind to subsites on the protease are numbered P1-P_n and P1'-P_n' in the same fashion (Schechter & Berger, 1967). Additionally, proteases can contain exosites, which are domains that enhance enzyme-substrate interactions but are not at the active site (Izaguirre, 2025). They

contribute to substrate recognition, stabilise the interaction, and enable subsequent cleavage reactions. These three-dimensional properties of the protease dictate specificity.

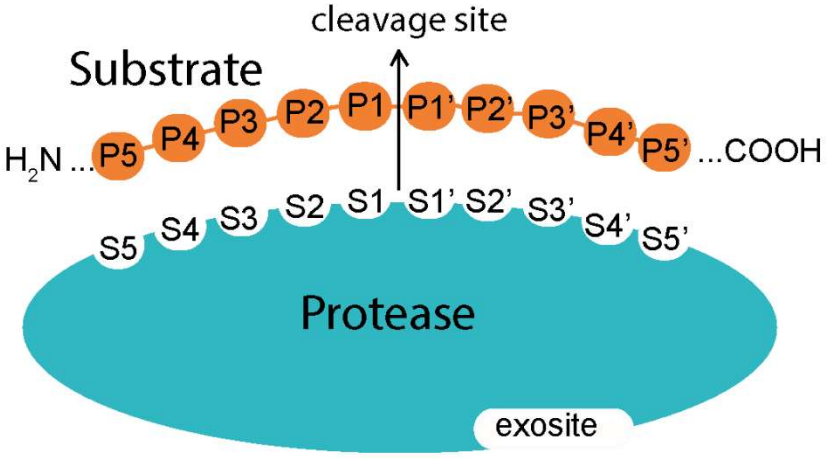


Figure 1.1: Schematic representation of Schechter-Berger nomenclature of protease-substrate interaction with visualised exosite. Substrate amino acids are notated with P1, P2 … P_n from cleavage site to N-terminus (also called non-prime sites) and P1', P2'… P_n' from cleavage site towards C-terminus (prime sites). Protease substrate-binding sites are marked with corresponding S labels.

Some proteases have very narrow specificity, while others, like lysosomal cathepsins that are responsible for general protein degradation and clearance, have broad specificity. Several factors govern the efficiency of cleavage *in vivo*: there has to be sufficient concentration of proteinase and its substrate in the same cellular environment under favourable conditions for proteinase activity, like pH or redox state. Simultaneously, cleavage must not be negatively influenced by post-translational modifications, such as methylation, phosphorylation, glycosylation, etc., which can either mark the substrate for degradation or protect it, usually by enabling or preventing ubiquitination (Lee, Hammarén, Savitski, & Baek, 2023). Likewise, the presence of endogenous inhibitors and allosteric effectors can affect cleavage success (Vizovišek, Vidmar, Fonović, & Turk, 2016). These mechanisms serve as means to finely regulate protease activity since even small quantities can induce physiological or pathological processes. Thus, *in vivo* protease activity is regulated on multiple levels, such as transcription (expression level), zymogen activation (if expressed as inactive zymogens, which is in most cases), localisation (e.g., separation from substrate in different compartments), and degradation by other proteases (half-life).

1.1 Cysteine Cathepsins

Cysteine proteases use the sulfhydryl group on the cysteine in their active site for nucleophilic attack on the carbonyl carbon of the peptide bond. Due to the proneness of this side chain group to oxidation, cysteine proteases are mostly active in reducing and/or a slightly acidic pH range of 4-6.5 (V. Turk *et al.*, 2012). In animals, the most well-known members are cytosolic calpains and lysosomal cathepsins, although the most studied one is papain from Carica papaya, which is also considered a typical representative of this family. Its active site is a catalytic dyad of cysteine and histidine, but cysteine proteases can also

contain a catalytic triad. However, when compared to serine proteases, this third catalytic residue (usually asparagine) is not as important for proteolysis as it is in serine proteases, despite its similar role of stabilising imidazole on the histidine side chain during catalysis (Klein, Eckhard, Dufour, Solis, & Overall, 2018).

Cysteine cathepsins (Cat) that belong to the papain superfamily are among the better-studied groups of cysteine proteases. In humans, there are eleven of them known: B, C, H, F, K, L, O, S, V, X and W. Cathepsins B and H possess endo- and exoprotease activity, which they owe to the presence of a chain near the active site that modifies it in such a way that only the substrate terminus can reach it. While CatB functions as a carboxypeptidase, CatH functions as an aminopeptidase. They are expressed as preproenzymes, and their N-terminal signal sequence is cleaved concurrently with N-linked glycosylation after their progression to the endoplasmic reticulum. The propeptide region then facilitates proper folding of the enzyme and targeting to the endosomal/lysosomal system by utilising the mannose-6-phosphate receptor (M6PR) pathway. Another role of the propeptide is to prevent any proteolytic activity of the zymogen. M6PR dissociates after zymogen entry to slightly acidic endosomes, where zymogens mature to active enzymes by cleavage of the propeptide region (V. Turk *et al.*, 2012). The propeptide region was shown to fold on the surface of the enzyme, running through the active site cleft, which blocks access of the substrate to the catalytic site (Coulombe *et al.*, 1996; Sivaraman, Lalumière1, Ménard, & Cygler, 1999; D. Turk, Podobnik, Kuhelj, Dolinar, & Turk, 1996). Usually, propeptides are bound to mature enzymes via noncovalent interactions, such as salt bridges and hydrogen bonds, except for CatX, where the propeptide is bound covalently via a disulfide bridge to the catalytic cysteine on mature cathepsin (Sivaraman, Nägler, Zhang, Ménard, & Cygler, 2000).

The propeptide thus has to be proteolytically removed, which can occur in several ways. This can be done by other proteases, such as the aspartic protease CatD that activates CatL (Nishimura, Kawabata, Furuno, & Kato, 1989), or autocatalytically under acidic conditions. This autoactivation process was shown to be a combination of unimolecular and bimolecular processes (Pungerčar *et al.*, 2009; Rozman, Stojan, Kuhelj, Turk, & Turk, 1999). In the case of procathepsin B, it was revealed that the zymogen form possesses a low catalytic activity, which is adequate to initiate the activation process. The propeptide region first dissociates from the active site cleft. This unimolecular process is then followed by bimolecular proteolytic processing of the propeptide region on another zymogen, which leads to a chain reaction. This kind of autocatalytic activation is present in endopeptidase cathepsins B, H, L, S, and K.

Typically, cathepsins are composed of the left and right domains. While the left domain consists of three alpha helices, the right contains a beta barrel with an alpha helix at the bottom. The active site is located between these domains, and each domain contributes one active site residue of the catalytic dyad – cysteine and histidine. Cysteine comes from the left domain at the N-terminus of the longest alpha helix, and histidine from the beta barrel of the right domain (Musil *et al.*, 1991; D. Turk, Gunčar, Podobnik, & Turk, 1998). In the case of endopeptidases, the active site cleft ranges along the whole length of the interface of the left and right domains, while in exopeptidases, there are additional features that limit binding to the termini of the substrate. In the case of aminopeptidases (CatH and CatC), they employ carboxylic groups from their propeptide parts to pack the active site on the non-primed side. CatH uses the carboxylic group from an eight-residue-long part, termed the mini-chain (Gunčar *et al.*, 1998), while CatC (strict aminopeptidase) provides the carboxylic group from the Asp1 side chain at the N-terminus of the exclusion domain (D. Turk, 2001). CatB and CatX contain occluding (Illy *et al.*, 1997; Nägler *et al.*, 1997) and mini-loop (Gunčar *et al.*, 2000), respectively. Specifically, histidine residues within these loops govern the docking of the C-terminal part of the substrate. Their binding

pockets are not really defined, which goes hand in hand with their most known role of general proteases, but CatK, L, and S show preference at the S2 site for hydrophobic amino acids, such as leucine, valine and isoleucine, with some differences. CatK is the only one of the three that prefers proline at this site as well, while CatL showed higher preference for aromatic residues when compared to the other two, and CatS prefers lysine at S2 (Vizovišek *et al.*, 2015). CatB has glutamate (Glu245) forming its S2 pocket, which is a unique residue among cathepsins. Mutating to glutamine demonstrated its importance for processing the substrate with an arginine in the P2 position (Hasnain, Hirama, Huber, Mason, & Mort, 1993). As mentioned, CatB possesses both endo- and exopeptidase activity. It is synthesised as a 339 amino acid long preproenzyme. Compared to other cathepsins, it has a relatively short proregion of 62 amino acids.

While cathepsins can be regulated via activation, they are also regulated via compartmentalisation within organelles, pH, and endogenous protein inhibitors. Their endogenous inhibitors consist of cystatins, thyropins and serpins, which can be divided into emergency and regulatory inhibitors. Emergency inhibitors act mostly on proteases that find their way into atypical compartments or on proteases that originate from pathogens (B. Turk, Turk, & Salvesen, 2002). Regulatory inhibitors possess a modulating ability in addition to their inhibitory activity. They are usually co-localised with their target protease and can be distinguished into three types: a) threshold-type, b) buffer-type and c) delay-type inhibitors (B. Turk *et al.*, 2002). Threshold inhibitors serve as a safety net in the event of accidental protease activation. Delay-type inhibitors are slow-binding, irreversible inhibitors that allow the activated protease to function for a limited amount of time, sufficient for maintaining normal homeostasis. Lastly, buffer-type inhibitors are reversible inhibitors that exhibit rapid association and dissociation, and they play a role in inhibiting proteases when no substrate is present.

Protein inhibitors of cysteine proteases can be divided into four groups: a) stefins – type 1 cystatins, b) cystatins – type 2 cystatins, c) kininogens – type 3 cystatins, and d) thyropins. Although there are also some serpins, which are usually inhibitors of serine proteases, that can inhibit cysteine cathepsins, such as squamous cell carcinoma antigen-1 (SCCA1) and hurpin (Schick *et al.*, 1998; Welss *et al.*, 2003). The first three groups are structurally related, whereas thyropins show homology to thyroglobulin type 1 domains (Lenarčič & Bevec, 1998). Stefins are primarily intracellular single-chain proteins, lacking carbohydrates, and consist of approximately 100 amino acids with a molecular mass of 11-12 kDa (V. Turk & Bode, 1991; V. Turk, Turk, Dolenc, & Stoka, 2019). The absence of a signal sequence and disulfide bonds distinguishes them from the other groups. Cystatins are slightly larger proteins (13-14 kDa) containing 2 disulfide bonds and a signal peptide, which enables their secretion to the extracellular milieu (V. Turk, Stoka, & Turk, 2008). Kininogens get their name from being precursors of the kinins, which are found in blood plasma. They are single-chain proteins with multiple domains. Compared to stefins and cystatins, they are present only in vertebrates, making them evolutionarily younger (Kordiš & Turk, 2009). Thyropins were first discovered as a fragment of p41 invariant chain and equistatin protein from sea anemone with the ability to inhibit CatL, papain, and CatD (Bevec, Stoka, Pungerčič, Dolenc, & Turk, 1996; Lenarčič, Ritonja, Štrukelj, Turk, & Turk, 1997; Lenarčič & Turk, 1999). Their sequences showed no homology to the other three groups, but they did to the thyroglobulin type-1 domain, from which they derived their name. It is worth mentioning that thyroglobulin type-1 domains are also present in other functionally unrelated proteins. Besides these, there are some serpins with the ability to inhibit cysteine proteases despite being primarily inhibitors of serine proteases (B. Turk *et al.*, 2002). As such, SCCA1 was found to be a potent inhibitor of CatK, CatL, and CatS (Schick *et al.*, 1998), whereas hurpin can selectively inhibit CatL (Welss *et al.*, 2003).

1.2 Cysteine Cathepsins in Health

1.2.1 Intracellular role

The main function of cysteine cathepsins is general intracellular protein degradation. These proteins are delivered by endocytosis, phagocytosis, and autophagy. Cathepsins play a crucial role in recycling and degradation, although they also have other important functions. During autophagy, CatB was shown to cleave the lysosomal MCOLN1/TRPML1 calcium channel, which reduces calcium efflux from lysosomes and leads to reduced activation of transcription factor EB (TFEB), a master transcription regulator of autophagy and lysosome biogenesis, showcasing CatB's role in suppression of these processes during homeostasis (Man & Kanneganti, 2016; Qi *et al.*, 2016). In addition, within the framework of innate immunity, cathepsins B, L, S, and F were shown to cleave Toll-like receptor 9 (TLR9) and ensure proper recognition of non-self DNA by TLR9 inside the endolysosomal system, which can continue the TLR signalling pathway, leading to a pro-inflammatory cytokine response (Matsumoto *et al.*, 2008; Park *et al.*, 2008). When it comes to adaptive immunity, cathepsins also play a role. Major histocompatibility complex (MHC) class II molecules in antigen-presenting cells present peptides derived from endolysosomal compartments. These protein antigens originate from the extracellular space and are transported into cells via phagocytosis or endocytosis. Inside these endolysosomal spaces, protein antigens are processed to peptides by lysosomal proteases. However, for these peptides to be presented, MHC class II molecules need CatS to process an invariant chain (Ii), a glycoprotein, which is needed for proper folding beforehand (Nakagawa *et al.*, 1999; Shi *et al.*, 1999). Similar to CatS, the same function was observed for CatL in mouse thymic cortical epithelial cells, where CatS is not expressed (Nakagawa *et al.*, 1998). In humans, CatL is replaced by CatV (Tolosa *et al.*, 2003).

Cathepsins also play a role in growth and development-related processes by processing hormones. Insulin-like growth factor I (IGF-1) is necessary for the proper formation and development of the skeletal system, and CatB is responsible for endosomal processing of IGF-1 by cleaving the C-terminal part (Authier, Kouach, & Briand, 2005) and enabling binding to IGF-1 and insulin receptors. Additionally, CatB and CatL play a role in neuronal development, since mice deficient in these cathepsins showed signs of brain atrophy in the cerebral cortex and the cerebellar Purkinje and granule cell layers (Felbor *et al.*, 2002), whereas in another study, mice with knocked out gene for CatB had significantly reduced neuronal activity in prefrontal cortex, neuronal migration, axonal growth, and displayed depressive behaviour. Notably, knockdown of CatB in adult wild-type mice did not induce similar effects (Xie *et al.*, 2025), showcasing an important role of CatB in embryonic and early development.

Cysteine cathepsins act as important mediators of apoptosis, where they can cleave and activate the proapoptotic Bcl-2 family member Bid and degrade the antiapoptotic members Bcl-2, Bcl-xL, or Mcl-1 after release of lysosomal cysteine cathepsins into the cytosol. In the same study, *in vitro* experiments showed that CatB, CatL, CatS, CatK, and CatH can all carry out these cleavages (Droga-Mazovec *et al.*, 2008). Cathepsins B, L, C, S and X also affect NOD-like receptor containing a pyrin domain 3 (NLPR3)-mediated IL-1β activation, where CatX plays a nonredundant role in NLPR3 activation as a response to sterile particles (Orlowski *et al.*, 2015). Furthermore, cysteine cathepsins can also exert their effect in the nucleus. As such, CatL was recognised to localise to the nucleus and regulate the cell cycle by cleaving the CDP/Cux transcription factor, which produces the CDP/Cux p110 isoform, which is important for the G1-S transition (Goulet *et al.*, 2004). Additionally, CatL can cleave the N-terminal tail of the H3 histone during mouse

embryonic stem cell differentiation, affecting gene regulation (Duncan *et al.*, 2008). Cathepsins B, K, L, and S can localise to the nuclear membrane, where they can regulate the TGF-β signalling pathway by interacting with the importing transporter protein importin-β and the exporting RanBP3, which as a consequence, influences the translocation of Smad proteins, the main signal transducers of receptors for TGF-β proteins (Zhang *et al.*, 2019).

1.2.2 Extracellular role

Despite mainly functioning inside the cells, cysteine cathepsins are not limited to the intracellular space. Their secretion into the extracellular matrix (ECM) is observed during some physiological processes, such as prohormone processing, bone remodelling and wound healing. Thyroid function relies on proper processing of prohormone thyroglobulin (Tg). CatB, K, L, and S can solubilise Tg and release thyroid hormones thyroxine (T4) and triiodothyronine (T3) via proteolytic action (Friedrichs *et al.*, 2003; Jordans *et al.*, 2009). Double knockout $CatK^{-/-}/L^{-/-}$ mice showed reduced levels of free thyroxine, indicating hypothyroidism (Friedrichs *et al.*, 2003). CatK is also the most predominant cysteine cathepsin found in the extracellular space due to its role in bone resorption. Collagenolytic activity of CatK depends on its association with glycosaminoglycans (GAGs), most notably chondroitin sulfate (Aguda *et al.*, 2014). It is activated in the acidic environment of resorption lacunae, which are located between osteoclasts and the bone surface, leading to bone matrix degradation and is crucial for normal homeostasis and the development of skeletal tissue (Vizovišek, Fonović, & Turk, 2019). Additionally, it can also cleave osteonectin, a calcium-binding glycoprotein (Bossard *et al.*, 1996). CatB, L, H, and S were shown to cleave osteocalcin in vitro, a bone protein that also functions as a hormone and is often used as a biomarker for bone degradation in osteoporosis (Wheater, Elshahaly, Tuck, Datta, & van Laar, 2013).

CatK was also shown to affect wound healing in skin, where it was strongly expressed in young surgical scars, indicating CatK's role in the homeostasis of dermal ECM. It acts in equilibrium between matrix synthesis and proteolytic degradation during scar formation, preventing matrix deposition and the development of hypertrophic scars (Rünger, Quintanilla-Dieck, & Bhawan, 2007). CatB plays a role during tissue regeneration and wound healing of the skin. It removes type IV collagen in the ECM, enabling migration of keratinocytes (Büth *et al.*, 2007). CatX processes the β_2 subunit of integrin receptors, enhancing adhesion of monocytes/macrophages to fibrinogen and also affecting the proliferation of T lymphocytes, thereby modulating the immune response. (Obermajer *et al.*, 2008).

In the brain ECM, CatB and CatL can cleave perlecan on the C-terminal end, producing the LG3 peptide, which possesses a neuroprotective role in brain ischemia (Saini & Bix, 2012). CatV participates in producing neuropeptide neurotransmitters in secretory vesicles, enkephalin and neuropeptide Y, which are known to be involved in pain sensation and feeding behaviour (Funkelstein *et al.*, 2012).

1.3 Cysteine Cathepsin B in Disease

Cysteine cathepsins are key contributors to many pathologies, where they are secreted or mislocalised to the extracellular space, facilitating disease progression (Olson & Joyce, 2015; V. Turk *et al.*, 2012; Vasiljeva *et al.*, 2007). Among them, CatB is of particular interest as it is the most ubiquitously expressed cysteine cathepsin, active both inside and outside the cell, and has been implicated in diverse pathological processes ranging from

cancer invasion and neurodegeneration to cardiovascular disease and arthritis (Biasizzo, Javoršek, Vidak, Zarić, & Turk, 2022). Extracellular CatB cleaves structural proteins of the ECM, activates other proteases, such as matrix metalloproteinases (MMPs) through degradation of tissue inhibitors of matrix metalloproteinases (TIMPs), and directly activates urokinase plasminogen activator (uPA), and modulates signalling molecules, including cytokines and chemokines (Kobayashi *et al.*, 1993; Kostoulas, Lang, Nagase, & Baici, 1999; Orlowski *et al.*, 2015; Repnik, Starr, Overall, & Turk, 2015). Because these activities support both normal tissue turnover and destructive remodelling, the regulation of CatB by endogenous inhibitors such as cystatins is critical. Dysregulated CatB activity, however, is consistently associated with pathological tissue breakdown and disease progression.

1.3.1 Arthritis

The extracellular matrix is a dynamic structure continuously remodelled by proteases during development, repair, and homeostasis. Cysteine cathepsins contribute substantially to this remodelling by degrading collagens, elastin, laminin, and fibronectin (Vizovišek *et al.*, 2019). Although CatK is the dominant collagenase in osteoclast-mediated bone resorption, CatB also participates in matrix turnover, especially in pathological conditions. In arthritis, excessive proteolysis of cartilage and bone leads to irreversible joint damage. CatB, secreted by synovial fibroblasts, amplifies cartilage destruction (Ben-Aderet *et al.*, 2015; Požgan *et al.*, 2010) and its inhibition reduces the severity of disease in animal models (Esser *et al.*, 1994).

1.3.2 Heart and lung diseases

Extracellular cysteine cathepsins are important mediators of vascular and pulmonary remodelling. While CatS and CatK are traditionally highlighted for their elastolytic activity in arteries and lungs, CatB is also a significant contributor to disease progression. In atherosclerosis, CatB is secreted by macrophages within plaques, where it propagates inflammation and promotes plaque instability. Elevated extracellular CatB activity has been linked to unstable lesions and inflammatory cell death in vascular smooth cells (Cai, Xu, & Liu, 2024). CatB also cleaves apolipoprotein α-I at the C-terminus, which reduces its solubility and ability to promote cholesterol efflux by ATP-binding cassette transporter A1, thereby affecting antiatherogenic capacity (Dinnes *et al.*, 2016). Elevated CatB levels are also present in acute and chronic myocarditis, where it is linked to activation of inflammasomes and promotion of pyroptosis (Wang *et al.*, 2018). ECM remodelling by CatB is also observed in animal models of lung fibrosis, where high CatB levels were detected in bronchoalveolar lavage fluid and lung homogenates of the mouse model of bleomycin-induced fibrosis (Kasabova *et al.*, 2016).

1.3.3 Neurodegenerative diseases

Cathepsin B has been widely studied in the central nervous system, where it plays a dual role in both normal neuronal homeostasis and neurodegeneration. In Alzheimer's disease, traumatic brain injury and related brain disorders, CatB is secreted via lysosomal leakage into the cytosol, where it activates cell death and inflammatory pathways which participate in these pathologies (Hook *et al.*, 2020; Nakanishi, 2020). On the other hand, CatB was shown to possess antiamyloidogenic and neuroprotective function in Alzheimer's disease, where it reduces amyloid-β (Aβ) peptides (Mueller-Steiner *et al.*, 2006), but another study indicated that localisation of CatB might explain this finding since non-lysosomal CatB

was responsible for the production of Aβ peptides in astrocytes while lysosomal CatB mediated degradation of Aβ peptides in the glioma-derived H4 cells (Oberstein *et al.*, 2021). As mentioned before (see 1.2.2), CatB can, along with CatL, also cleave perlecan and generate a neuroprotective C-terminal fragment during brain ischemia (Saini & Bix, 2012). Additionally, CatB can also degrade axon-inhibitory chondroitin sulfate proteoglycans, which promotes axon regeneration and neuroplasticity following spinal cord injury (Tran, Sundar, Yu, Lang, & Silver, 2018). In amyotrophic lateral sclerosis (ALS), CatB participates in motor neuron degradation and elevated levels were found in mouse models of ALS (Kikuchi *et al.*, 2003; Offen *et al.*, 2009). Moreover, CatB knockout mice exhibited reduced motor neuron death after nerve injury in adult mice (Sun *et al.*, 2010). This dual action in neurodegenerative pathologies shows there is a necessity for further studies to resolve the CatB role in them.

1.3.4 Infections

Several pathogens exploit host proteases for invasion, and CatB is no exception. In the case of the Ebola virus, CatB and CatL are used to activate envelope glycoproteins needed for viral entry. CatB/L inhibitors reduced cell infection by Ebola glycoproteins in 294T and macrophage cell lines (Gnirß *et al.*, 2012). CatB's role in SARS-CoV-2 infection was also indicated by *in silico* analysis, which identified CatB, K, L, S, and V as possible proteases of SARS-CoV-2 S-protein (Bollavaram *et al.*, 2021). Moreover, dual inhibition of transmembrane serine protease 2 and CatB prevented SARS-CoV-2 infection in induced pluripotent stem cells, while inhibition of only one of these proteases had a lesser impact (Hashimoto *et al.*, 2021). In the context of bacterial infections, CatB inhibition was shown to increase survival of *Mycobacterium tuberculosis* in human macrophages, indicating its protective role against phagosome-residing bacteria (Pires *et al.*, 2016). A contrary effect was observed in the cytosolic-residing bacteria *Francisella novicida* infection, where genetic deletion or inhibition of CatB improved survival of infected mice and macrophage cell lines since this prevented cleavage of lysosomal calcium channel TRPML1, which led to increased lysosomal biogenesis and activation of autophagy genes, resulting in higher clearance of the bacteria (Qi *et al.*, 2016). It was observed that CatB-deficient mice, infected with parasitic protozoan *Leishmania*, resolved major lesions much faster than wild-type mice, although the mechanism is not known other than the fact that there is no involvement of CatB in TLR9-dependent or independent cytokine responses of dendritic cells and macrophages or in the innate immune response (Rasid *et al.*, 2016).

1.3.5 Immune disorders and inflammation

Cysteine cathepsins are key mediators of immune regulation, and their dysregulation contributes to autoimmune and inflammatory pathologies. Among them, CatB has emerged as an important factor in immune disorders. Genomic studies have linked increased plasma levels of CatB to systemic lupus erythematosus (SLE) (Mo *et al.*, 2020), suggesting it may contribute to disease susceptibility or progression. Complementarily with CatX, CatB has also been shown to promote delayed-type hypersensitivity reactions, which are characteristic of inflammatory skin conditions such as contact dermatitis and psoriasis (Schwenck *et al.*, 2019). CatB is involved in the onset of acute pancreatitis as well via intracellular trypsinogen activation, which leads to necrosis of pancreatic tissue (Halangk *et al.*, 2000). It was shown that it also initiates apoptosis in the acinar cells during pancreatitis (Sendler *et al.*, 2016) and is greatly activated by CatD in the latter, inflammatory cell–dependent phase of the disease (Aghdassi *et al.*, 2018). Additionally, CatB serves as a reliable biomarker in hepatic fibrosis in patients with chronic hepatitis-C

infection, where it was strongly correlated with levels of plasma kallikrein, another biomarker of hepatic fibrosis (Zanelatto *et al.*, 2022). Through these mechanisms, CatB appears to support both systemic and localised inflammatory processes.

1.3.6 Cancer

The role of cysteine cathepsins in cancer has been studied extensively, with CatB emerging as one of the most consistently implicated proteases in tumour progression. Tumour cells, stromal fibroblasts, and infiltrating immune cells all secrete CatB into the extracellular space, where it remodels the tumour microenvironment and supports invasion and metastasis (Olson & Joyce, 2015; Sloane *et al.*, 2005; Vasiljeva *et al.*, 2006; Vasiljeva & Turk, 2008).

CatB promotes cancer invasion primarily by degrading basement membrane components such as type IV collagen, laminin, and fibronectin, thereby clearing physical barriers to migration and inducing angiogenesis (Chang *et al.*, 2009; Guinec, Dalet-Fumeron, & Pagano, 1993; Kruszewski *et al.*, 2004). It further amplifies proteolytic cascades by activating uPA or degrading tissue inhibitors of matrix metalloproteinases (Kobayashi *et al.*, 1993; Kostoulas *et al.*, 1999). These activities promote not only invasion and angiogenesis but also pro-tumorigenic inflammation, as CatB processes cytokines and chemokines that shape the tumour microenvironment (Greten & Grivennikov, 2019; Repnik *et al.*, 2015; Sloane *et al.*, 2005; Yin *et al.*, 2012)

In addition, it can regulate cancer cell proliferation through the regulation of extracellular signal-regulated kinase (ERK)/mitogen-activated protein kinase (MAPK) in pancreatic cancer or the cell cycle regulator $p27^{Kip1}$ in colorectal cancer, which accelerates cancer cell proliferation. The loss of CatB activity results in reduced cancer cell proliferation (Bian *et al.*, 2016; Gopinathan *et al.*, 2012). Clinical sample analyses reinforce the importance of CatB: its expression and extracellular activity correlate with poor prognosis in breast, colorectal, glioblastoma, gastric, hepatic and pancreatic cancers (Chan *et al.*, 2010; Ebert *et al.*, 2005; Fujimoto *et al.*, 2020; Leto *et al.*, 1997; Ma *et al.*, 2022; Nouh *et al.*, 2011). CatB also regulates cell adhesion and motility by cleaving E-cadherin, which is associated with disaggregation of the tumour (Gocheva *et al.*, 2006). It cleaves neuropilin 1 as well, an adhesion molecule, in the MDA-MB-231 cancer cell line (Sobotič *et al.*, 2015).

Experimental models underscore CatB's tumour-promoting potential: its loss reduces proliferation, invasion, and angiogenesis in pancreatic neuroendocrine and mammary tumour models (Gocheva *et al.*, 2006; Vasiljeva *et al.*, 2006). Together, these findings highlight CatB not only as a biomarker of aggressive disease but also as a therapeutic target.

Although many cysteine cathepsins contribute to extracellular proteolysis, cathepsin B stands out as one of the most versatile and widely implicated in different diseases. Its ability to degrade ECM proteins, activate protease cascades, and process signalling molecules makes it an important driver of pathological remodelling. In arthritis, CatB accelerates cartilage destruction; in cancer, it fuels invasion and metastasis; in cardiovascular and pulmonary disease, it destabilises tissue structure; in neurodegeneration, it contributes to both protein aggregation and inflammation; and in infection, it is co-opted by pathogens.

These diverse functions highlight CatB as both a biomarker of disease progression and a promising therapeutic target. Targeting extracellular CatB selectively—without disrupting its essential intracellular roles—remains a challenge but also an opportunity for next-generation therapies in oncology, neurology, and cardiovascular medicine.

1.4 DARPins

Designed ankyrin repeat proteins (DARPins) are antibody mimetic proteins that were developed using a scaffold from ankyrin repeat (AR) proteins, which mediate the attachment of membrane proteins to the cytoskeleton. They were built from 33 amino acid residue AR modules with seven randomised positions (Binz, Stumpp, Forrer, Amstutz, & Plückthun, 2003). They usually consist of 4 to 6 AR modules, where N-terminal and C-terminal capping repeats serve to protect the inner hydrophobic AR proteins, and their size is in the 14-18 kDa range, much smaller when compared to conventional antibodies. Consensus sequence was obtained from the SMART database (Letunic *et al.*, 2002) and DNA libraries were assembled using trinucleotide phosphoramidites (Virnekas *et al.*, 1994). In 2013, a second-generation DARPin library with reduced hydrophobicity was developed, which gives fewer nonspecific hydrophobic binders (Seeger *et al.*, 2013). This is also the version of the library used in this work. Every inner AR module consists of a β-turn that connects two antiparallel α-helices.

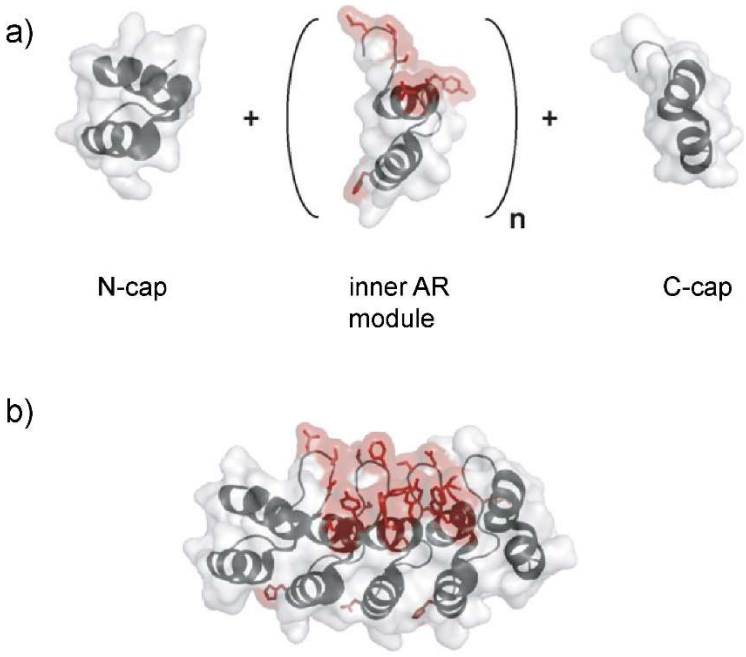


Figure 1.2: Visual representation of a DARPin molecule. **a**) Composition of the DARPin molecule. Every DARPin molecule consists of N- and C-cap with a varying number of inner ankyrin repeat (AR) modules (n). **b**) A three-dimensional representation of the DARPin molecule with three inner AR modules. Adapted from Stumpp, Binz, & Amstutz (2008).

Utilising ribosome and phage display, it is possible to select DARPins against a specific target. Ribosome display couples mRNA encoding DARPin and DARPin itself via the ribosome. This complex (mRNA-ribosome-DARPin) can then be displayed to the immobilised target, and it is possible to present up to 10^{12} candidates (Plückthun, 2012). Non-binding complexes are washed away, and the sequences of binding DARPins are recovered using reverse transcription PCR to obtain cDNA, which is then *in vitro* translated to mRNA to use for another round of the ribosome display. This cycle can then be repeated multiple times to further select those with the highest affinities. This approach can yield highly specific binders with an affinity as low as pM to the low nM range due to

randomised positions in the inner AR library modules, which form an interaction surface to bind the epitope of the target protein surface.

1.4.1 DARPins in therapeutics and diagnostics

DARPins can be used as diagnostic and therapeutic agents. Since they possess antibody functionality, they can also serve as binders for quantitative immunohistochemistry. Due to their robustness, it is possible to add various protein tags, fusion proteins, or conjugate them to diagnostic markers, such as fluorescent dyes or radiometal chelators (Plückthun, 2015). As a proof of principle, DARPin G3 with a K_D of 90 pM for epidermal growth factor receptor 2 (HER2), a predictor of poor prognosis in breast cancer, was used to successfully detect HER2 in breast cancer tissue microarrays. Signal was obtained by detecting biotin on the genetically fused AviTag via streptavidin-coupled horseradish peroxidase using 3,3'-diaminobenzidine as substrate. Its detecting ability was compared to an FDA-approved rabbit monoclonal antibody (clone 4B5; Ventana Medical Systems) and was found to detect HER2 with similar sensitivity but higher specificity, showcasing its potential as an alternative to the conventional antibody approach for diagnostics (Theurillat *et al.*, 2010). Another example of diagnostic use was displayed in our lab on the mouse model, where a bi-specific DARPin for both mouse and human CatB with a K_D of 35 pM was conjugated with fluorophore cyanine5.5 via an engineered C-terminal cysteine residue and used to image tumours *in vivo* in PyMT and 4T1 mouse models. Moreover, it selectively binds only to the mature form of CatB, which is clinically more relevant, since CatB contributes to cancer progression via proteolysis in the ECM (Kramer, Renko, *et al.*, 2017). Radiolabelled DARPins have been developed for diagnostic and imaging purposes, where their high specificity and affinity enable precise targeting, and some have entered clinical trials. Due to their small size (14-18 kDa), they also have fast clearance, affecting the target-to-background signal ratio, shorter workflow and therefore reduced side effects, making them promising tools for diagnostics. Two clinical trials have been done with radiolabelled DARPins targeting HER2. In the first one, ^{99m}Tc-(HE)3-G3 has been assessed as a tool for distinguishing between positive and negative HER2 primary breast cancer (NCT04277338) and in another one (NCT05376644), it was compared to another protein-based imaging agent, ADAPT6 (albumin-binding domain-derived affinity protein 6). While G3 was twice the size of ADAPT6 (14 kDa vs. 5-7 kDa), it also had a higher affinity for its target (90 pM vs. 1 nM) and showed higher contrast to background but lower accumulation. The third clinical trial utilising radiolabelled DARPin was done by targeting epithelial cell adhesion molecule (EpCAM; NCT05620472) by Ec1 DARPin in lung cancer patients. In all three trials, DARPins had no adverse effects and were effective as a tool for assessing EpCAM- and HER2-expressing tumours (reviewed in Stumpp, Kirkin, Sanderson, Dawson, & Legenne, 2025).

Many DARPins have entered clinical trials as therapeutic agents, mainly for cancer but also for COVID-19 infections. Two DARPin-based drugs targeting cancer conditions reached at least the initial part of phase 2, and three COVID-19-targeting DARPins reached phase 2 (Stumpp *et al.*, 2025). All of the therapeutic DARPins have been multi-specific, meaning that they target multiple targets, which is achieved by linking individual DARPins together. MP0250 is a DARPin that is designed for dual inhibition of vascular endothelial growth factor A (VEGF-A) and hepatocyte growth factor (HGF), which are proteins that many tumours rely on to support their growth, survival, invasiveness, and drug resistance (Saltarella *et al.*, 2023). Two human serum albumin (HSA)-targeting DARPin domains were also added to the construct to prolong the half-life of the construct, bringing the number of functional domains to four (2× anti-HSA, 1× anti-VEGF-A, 1× anti-HGF). In the phase 1 study (NCT02194426), where it was tested on 45 individuals

with advanced solid tumours, it showed suitable tolerability and appropriate pharmacokinetic properties for further development. The most common adverse effects were hypertension (69 % of patients), proteinuria (51 %), diarrhoea and nausea (both 36 %), and hypoalbuminemia (24 %), which are side effects consistent with inhibition of VEGF-A and HGF pathways (Baird *et al.*, 2021). This prompted the advancement to a phase 1b/2 study (NCT03136653), where 33 patients with relapsed or refractory multiple myeloma received at least one dose of MP0250, together with anti-cancer medications bortezomib and dexamethasone (Jongen-Lavrencic *et al.*, 2024). Despite recording an objective response in 32 % of patients and acceptable safety levels, the development was stopped due to the advent of novel effective drugs against multiple myeloma, mainly against CD38, SLAM7, and B-cell maturation antigen (Lin *et al.*, 2024). Another DARPin-based drug that reached the clinical phase 1/2a is MP0533, which is a tetra-specific T cell engager for patients with acute myeloid leukaemia (AML). It targets three tumour-associated antigens (CD33, CD123, CD70) expressed by leukemic stem cells and leukemic blasts, and CD3 for activation of T cells. Similar to MP0250, MP0533 contains two adjacent HSA-targeting DARPin domains on the N-terminus to increase systemic half-life (Bianchi *et al.*, 2024). In early 2023, a phase 1/2a study was launched, which is currently still in the recruitment process and is estimated to finish in 2029 (NCT05673057). Interim data obtained from 37 patients with relapsed or refractory AML, or myelodysplastic syndrome, who were treated with MP0533 showed clinical response in 4 cases (morphologically leukaemia-free state, MLFS (n=3); complete remission, CR (n=1)), and 8 of 30 evaluable patients displayed ≥50 % blast reduction in the bone marrow. There was cytokine release syndrome present, which usually occurred within 24 hours following MP0533 administration in treatment cycle 1 (28 days), and two dose-limiting toxicities occurred at dose-level-7 (Jongen-Lavrencic *et al.*, 2024). Nonetheless, the study concluded that MP0533 shows an acceptable safety profile and encouraging antitumour activity for future dose escalation.

1.4.2 DARPins as research tools

A lot of biochemical research is based on the identification of proteins by antibodies, and this is where DARPins can be used as a potent biosensor. Due to their selectivity, they can bind to specific forms of a chosen protein. Extracellular signal-regulated kinase 2 (ERK2) is a member of the mitogen-activated protein (MAP) kinase family, which is involved in a variety of cellular processes such as cell growth, adhesion, survival and differentiation through the regulation of transcription, translation and cytoskeletal rearrangements. ERK2 needs to be phosphorylated by upstream kinases to become active. The DARPin platform was used to select DARPins E40 and pE59 that can bind to either the non-phosphorylated (inactive) or phosphorylated (active) form, respectively. This enables either inhibiting activation or monitoring the level of activation in live cells (Kummer *et al.*, 2012), making it a valuable research tool for investigating protein dynamics. In a follow-up study, it was possible to selectively visualise active endogenous ERK in mouse embryoblasts, showing higher activation in the nucleus, perinuclear space, and especially in nucleoli (Kummer *et al.*, 2013).

Additionally, the same DARPins were used to obtain crystal structures of both forms of ERK2, and analysis showed that both DARPins bind to the same region but can recognise conformational changes which happen after phosphorylation (Kummer *et al.*, 2012), showcasing another use of DARPin, and that is chaperone-assisted crystallography. Their rigid nature allows them to readily crystallise with their targets (Sennhauser & Grütter, 2008), such as caspase-2 (Schweizer *et al.*, 2007), the efflux pump of *E. coli* AcrB (Sennhauser, Amstutz, Briand, Storchenegger, & Grütter, 2006) and others (reviewed in

Mittl, Ernst, & Plückthun, 2020). Similar work was done with DARPin 8h6 and 81, which can both bind to both mouse and human cathepsin B. Adding the cysteine tag at the N-terminus and labelling it via maleimide chemistry with fluorophore Cy5.5 made it possible to visualise tumours in PyMT and 4T1 mouse tumour models. Crystal structures of the complexes were also resolved, which showed that DARPins bind to the less structurally conserved region of cathepsins (Kramer, Renko, *et al.*, 2017).

Chapter 2

Aims and Hypothesis

CatB is involved in many diseases, such as various cancers, arthritis, lung disorders, Alzheimer's, and cardiovascular diseases. We aim to research mouse CatB and compare it to human CatB in interactions with protein inhibitors and the processing of physiological substrates. Understanding these differences and similarities will contribute to the evaluation of the validity of CatB-related mouse model conditions.

Our hypotheses were:

1. DARPin 4m3 shows true selectivity in its binding and inhibition, and it is possible to determine interactions responsible for this selectivity.
2. We expect mCatB and hCatB to showcase higher similarity in processing of the physiological substrate among themselves than to any other human cathepsin.

The study's goals are the characterisation of DARPin 4m3 interaction with mCatB and mCatB itself when it comes to substrate processing. The main goals of this work were:

1. Development of DARPin 4m3 and its use as a research tool
    a. Characterisation of DARPin 4m3 binding to mCatB
    b. Expression and purification of mCatB and DARPin 4m3 in large enough quantities to use for crystallisation purposes
2. Comparative analysis of substrate specificity between cathepsins with emphasis on mCatB and hCatB
    a. Identification of cleavage sites from physiological substrates of mammalian cell models.
    b. Using statistical and machine-learning methods to quantify the similarity of substrate processing.

Chapter 3

Materials and Methods

3.1 Materials

3.1.1 Laboratory equipment

- 96-well microtitre black plates (Greiner)
- agarose gel electrophoresis system Owl Easycast (Thermo Fisher Scientific)
- Amicon Ultra centrifugal filters 0.5, 4, 15 mL (Merck Millipore)
- centrifuges (Eppendorf)
- cell culture incubator (Binder)
- chemiluminescence fluorescence gel imaging system G:Box (Syngene)
- chromatography column HiPrep S75 120ml (GE Healthcare)
- chromatography column HiTrap IMAC Sepharose FF (GE Healthcare)
- chromatography column S75 25 mL (GE Healthcare)
- chromatography system Akta FPLC System (GE Healthcare)
- chromatography system Akta Pure (GE Healthcare)
- dispenser Gryphon for protein crystallography (Art Robbins)
- electrophoresis power supply PowerPac™ Basic Power Supply (Bio-Rad)
- gel casting apparatus Mini-Protean (Bio-Rad)
- laminar flow hood
- microplate reader Tecan Infinite M1000 (Tecan)
- multichannel pipettes (Eppendorf)
- Nanodrop microvolume spectrophotometer ND-1000 (Babtech International)
- nitrocellulose membrane
- PCR machine MiniAmp Plus (Thermo Fisher Scientific)
- pH meter SevenDirect SD20 (Mettler Toledo)
- polyacrylamide gel electrophoresis system Mini-Protean Tetra Cell (Bio-Rad)
- polystyrene Petri dishes 10 and 15 cm
- single-channel pipettes (Eppendorf)
- software PR.Stability Analysis, v1.1 (NanoTemper)
- software ChimeraX 1.11 (UCSF)
- sonicator SFX 550 (Branson)
- surface plasmon resonance system Biacore T200 (Cytiva)
- sensor chip CM5 (Cytiva)
- synchrotron Elettra Trieste S.C.p.A
- Trans-Blot Turbo Transfer system (Bio-Rad)

3.1.2 Chemicals, commercial kits, and reagents

- 2-(N-morpholino)ethanesulfonic acid (MES) (Serva)
- Amersham ELC Western Blotting Detection Reagent (Cytiva)
- ampicillin
- APS
- Bradford reagent (Bio-Rad)
- BSA (Sigma-Aldrich)
- CA074 (Bachem)
- casein (Sigma-Aldrich)
- chloramphenicol
- DMSO (Merck)
- dNTP
- DTT (Sigma-Aldrich)
- E-64 (Peptide Institute)
- EDTA (Serva)
- FBS-HI (Sigma-Aldrich)
- Glutamax (Gibco-Invitrogen)
- HEPES (Fluka Organics)
- imidazole (Acros Organics)
- IPTG (Sigma-Aldrich)
- kanamycin
- $MgCl_2$ (Serva)
- $MgSO_4$ (Sigma-Aldrich)
- mouse cathepsin B antibodies, goat-derived, AF965 (R&D Systems)
- $Na_2HPO_4{\cdot}2H_2O$ (Fluka Organics)
- NaCl (Fischer Chemical)
- NaH_2PO_4 (Acros Organics)
- Ni_2SO_4 (Sigma-Aldrich)
- PageRuler Prestained Protein Ladder, 10 to 250 kDa (Thermo Fisher Scientific)
- PCR buffer for KOD HS DNA polymerase (10X) (Merck Millipore)
- PEG_{6000} (Sigma-Aldrich))
- penicillin/streptomycin (Gibco-Invitrogen)
- rCutSmart buffer (NEB)
- SDS (Thermo Fisher Scientific)
- secondary antibody anti-goat (HRP), AB6885, Abcam
- T4 DNA Ligase Buffer (10X) (Thermo Fisher Scientific)
- TEMED
- tetracycline
- Trans-Blot Turbo 5x Transfer Buffer (Bio-Rad)
- Tris (Serva)
- tryptone (Becton Dickinson)
- Tween-20 (Sigma-Aldrich)
- Wizard® Plus SV Minipreps DNA Purification Systems kit (Promega)
- Wizard® SV Gel and PCR Clean-Up System kit (Promega)
- yeast extract (Merck Millipore)
- z-Arg-Arg-AMC (Bachem)
- pH indicator strips (Cytiva)

3.1.3 Buffers and growth mediums

- phosphate buffer
  11.54 mM $Na_2HPO_4 \cdot 2H_2O$
  8.46 mM NaH_2PO_4
  150 mM NaCl
  pH 7.0

- buffer A for Ni-NTA affinity chromatography
  20 mM HEPES
  300 mM NaCl
  30mM imidazole
  pH 7.4

- buffer B for Ni-NTA affinity chromatography
  20 mM HEPES
  300 mM NaCl
  250 mM imidazole
  pH 7.4

- buffer C following Ni-NTA affinity chromatography
  20 mM HEPES
  300 mM NaCl
  pH 7.4

- cathepsin activity buffer
  12 mM Na2HPO4·2H2O
  88 mM NaH2PO4
  1 mM EDTA
  0.1 % PEG 6000
  pH 6.0
  5 mM DTT

- LB growth medium with antibiotics
  1 % casein
  0.5 % yeast extract
  1 % NaCl
  pH 7.2
  100 µg·mL- ampicillin
  50 µg·mL- kanamycin
  25 µg·mL- chloramphenicol
  12.5 µg·mL- tetracycline

- TB growth medium with antibiotics
  17 mM
  72 mM
  1.2 % tryptone
  2.4 % yeast extract
  0.4 % glycerol
  pH 7.2
  100 µg·mL- ampicillin

50 µg·mL- kanamycin
25 µg·mL- chloramphenicol
12.5 µg·mL- tetracycline

- SDS-PAGE separating gel
  3.65 mL dH2O
  2.5 mL 1.5 M Tris, pH 8.8
  3.75 mL 40 % acrylamide/bisacrylamide, 37.5:1
  1 ‰ SDS
  2 ‰ TEMED
  0.4 ‰ APS

- SDS-PAGE stacking gel
  1.54 mL dH2O
  0.625 mL 0.5 M Tris, pH 6.8
  0.313 mL 40 % acrylamide/bisacrylamide, 37.5:1
  1 ‰ SDS
  2 ‰ TEMED
  0.4 ‰ APS

- SDS-PAGE running buffer
  25 mM Tris
  192 mM glycine
  1 ‰ SDS

- TBST buffer
  25 mM Tris
  500 mM NaCl
  0.05 % Tween-20

- TBST-BSA buffer
  25 mM Tris
  500 mM NaCl
  0.05 % Tween-20
  3% BSA

3.1.4 Oligonucleotides

- primer NdeI-FOR-pet22b4m3:
  5'-GCGCATATGGACCTGGGTAAGAAACTGCTGG-3'
- primer NotI-REV-pet22b4m3_3C:
  5'-TTAGCGGCCGCCGCCGCTTTTTGCAGCA-3'

3.1.5 Vectors

- pET22b+3CHis (for expression of the DARPin 4m3 with 3C cleavage site and His tag on C-terminus)
- pMCSG7 (for expression of the mouse and human cathepsin B)
- pQE30 (originating vector of DARPin 4m3)

3.1.6 Antibodies and enzymes

- KOD Hot Start DNA polymerase
- mouse cathepsin B antibody AF965
- NdeI restriction endonuclease
- NotI-HF® restriction endonuclease
- secondary donkey anti-goat IgG antibodies AB6885
- T4 DNA ligase

3.1.7 Bacterial and mammalian cell cultures

- Escherichia coli JM109
- Escherichia coli Rosetta Gami B(DE3)pLysS
- mouse macrophage cell line RAW 264.7

3.2 Methods

The same methods also appear in Zarić *et al.* (2025) except ones mentioned in 3.2.2, 3.2.3, 3.2.5, 3.2.6, and 3.2.7.

3.2.1 Selection of mouse cathepsin B-specific inhibitory DARPin

DARPins were obtained using ribosome display as previously described (Kramer, Renko, *et al.*, 2017). A second-generation DARPin library was used (Seeger *et al.*, 2013), which has reduced hydrophobicity.

Briefly, DARPin 4m3 was selected by panning the DARPin library four times against mCatB, with each panning selecting the most potent binders of the cycle. Each selection round also included a pre-panning step with biotinylated maltose-binding protein to remove nonspecific DARPin candidates.

3.2.2 Molecular cloning of DARPin 4m3 sequence

Due to problematic large-scale expression where DARPin 4m3 was being cleaved during expression, we cloned the DARPin sequence from the pQE30 plasmid to pET22b+ containing the HRV 3C site and 6×His-tag on the C-terminus using restriction digest and ligation protocol. Forward and reverse primers, containing NdeI and NotI restriction sites, for polymerase chain reaction (PCR) were obtained from Sigma-Aldrich. PCR of the DARPin 4m3 sequence was mixed as:

- vector pQE30/DARPin 4m3 (1 ng/µl)	1 µl
- dNTP (2 mM)	5 µL
- forward primer NdeI-FOR-pet22b4m3	1.5 µl
- reverse primer NotI-REV-pet22b4m3_3C	1.5 µl
- $MgSO_4$ (25 mM)	3 µl
- 10X PCR buffer for KOD HS DNA polymerase	5 µl
- KOD Hot Start DNA polymerase (1 U/µl)	1 µl
- dH_2O	32 µl
- final volume	50 µl

We used the following PCR program:

2 min 95°C

20 sec	95°C	⎫
10 sec	56°C	⎬ 40×
15 sec	70°C	⎭
2 min	70°C	
	4°C	

The PCR product was purified using the Wizard® SV Gel and PCR Clean-Up System kit. The PCR product containing the DARPin 4m3 sequence and the pET22b+ plasmid were then digested with the restriction endonucleases NdeI and NotI-HF and ligated using T4 ligase. Restriction reactions were mixed according to the following scheme:

- DNA material for restriction digestion (100 ng/µl) 20 µl
  (pET22b+3CHis or PCR product of the DARPin 4m3 sequence)
- rCutSmart buffer 5 µl
- NdeI restriction endonuclease 1 µl
- NotI-HF restriction endonuclease 1 µl
- dH_2O 23 µl
- final volume 50 µl

The mixture was incubated at 37°C for 1 h and then inactivated for 20 min at 80°C. Products were purified using the Wizard® SV Gel and PCR Clean-Up System kit again, followed by ligation reaction of restriction digestion products:

- pET22b+3CHiS (5.8 ng/µl) 8.8 µl
- PCR of the DARPin 4m3 sequence (35.4 ng/µl) 1.3 µl
- T4 DNA ligase buffer (10X) 2 µl
- T4 DNA ligase (5 U/µl) 1 µl
- dH_2O 6.9 µl
- final volume 20 µl

The ligation reaction was then incubated at 37°C for 1 hour. 10 µl was then used to transform JM109 competent cells, which were then spread on LB plates with ampicillin and incubated at 37°C overnight. The picked colony was then inoculated into a 10 mL liquid LB growth medium with ampicillin and incubated overnight with shaking. The new plasmid was then isolated from the culture using the Wizard® Plus SV Minipreps DNA Purification Systems kit.

3.2.3 Expression and purification of cathepsins

Human and mouse CatB were expressed as soluble proteins in Escherichia coli strain Rosetta gami B(DE3)pLysS, as previously described (Novinec, Pavšič, & Lenarčič, 2012). Human cathepsins K, V, L, and S were expressed in *Pichia pastoris* as previously described (Bromme, Nallaseth, & Turk, 2004; Mihelič, Doberšek, Gunčar, & Turk, 2008), and hCatF was expressed in baculovirus-infected cells (Fonovič, Brömme, Turk, & Turk, 2004). For cathepsins B, cells were transformed with pMCSG7 containing the sequence of the corresponding cathepsin. Expression cultures were grown at 37°C until OD600 reached 0.8 – 1.0. Flasks were then brought to 4°C for 15 minutes, and expression was induced by adding 1 mM IPTG. The cultures were incubated at 18°C and 300 rpm for 20 h. Cells were then pelleted by centrifugation at 4000 g for 10 minutes at 4°C. The cell pellet was stored at -20°C.

The cell pellet was thawed and resuspended in buffer A and sonicated on ice using a Branson Digital Sonifier SFX 550, 30 × 10 s at 25 % amplitude with a 30 s pause between pulses. The sonicated cells were then centrifuged at 16000 g for 45 minutes at 4°C. The supernatant was collected and applied to the Ni-NTA chromatography column. The bound proteins were eluted with buffer B with 250 mM imidazole. The eluted fractions were collected, and the imidazole was removed using Amicon filtration columns of 15 mL and buffer C that contains no imidazole. The imidazole-free samples were then concentrated and stored at -80°C.

To activate procathepsin B, the pH was lowered to approximately pH 5.0 by adding 10 % acetic acid and DTT to 5 mM. The sample was then incubated at room temperature for 2 h and centrifuged at 20000 g for 20 minutes at 4°C to remove aggregated and precipitated material. The supernatant was then further purified on the Superdex 75 size-exclusion chromatography column in 20 mM sodium acetate, 200 mM NaCl, 1 mM EDTA, pH 4.5. Fractions containing activated cathepsin were collected and concentrated using Amicon columns with a cut-off of 3 kDa.

The activity of cathepsins was then checked using the substrate z-Arg-Arg-AMC (z-RR-AMC) by titration against the cysteine protease inhibitor E64 in cathepsin activity buffer, and purity was checked by SDS-PAGE.

3.2.4 Expression and purification of DARPin 4m3

The DARPin 4m3 sequence was inserted into the pET22b+ vector by restriction digestion and ligation approach. The Rosetta gami B(DE3)pLysS strain was transformed using this construct. Cultures were grown in TB medium supplemented with ampicillin, kanamycin, chloramphenicol, and tetracycline at 37°C until the OD_{600} reached approximately 1.5. Flasks were then transferred to 4°C for 15 minutes, and expression was induced by the addition of 1 mM IPTG. The expression was continued at 18°C for 20h and 300 rpm.

Sample preparation for Ni-NTA was the same as for cathepsin. Ni-NTA chromatography was followed by size-exclusion chromatography on a Superdex 75 column in 20 mM HEPES, 200 mM NaCl, pH 7.4.

3.2.5 Determination of protein concentration

Protein concentrations were determined using a Nanodrop microvolume spectrometer by measuring absorbance at 280 (A_{280}) nm and calculated using the molar absorption coefficient ε_{280} according to the formula $c = A_{280} / (\varepsilon_{280} \cdot l)$, where l represents optical path length.

For cell lysate samples, concentrations were determined using the modified Bradford assay. BSA was used as an external standard in the concentration range of 1 to 10 µg/mL in the Bradford assay, where the absorption maximum of Coomassie Brilliant Blue G-250 shifts from 465 nm to 595 nm after binding to aromatic amino acid residues. A_{595} was measured using the Tecan plate reader Infinite M1000, and a standard curve was plotted using the BSA standard's measurements. As an external reference, bovine serum albumin (BSA) was utilised at concentrations between 1 and 10 µg/mL. A transparent 96-well plate was filled with the appropriate amounts of diluted samples and BSA stock solution, then dH2O was added to the final volume of 160 µL. The wells were filled with 40 µL of a 5×BioRad Reagent solution and pipetted. The microplate reader Tecan Infinite M1000 was used to measure A_{595}. The BSA standard measurements were used to plot a standard curve, from which the sample concentrations were computed.

3.2.6 Sodium dodecyl sulfate polyacrylamide electrophoresis

8 mL of SDS-PAGE separating gel solution was poured between glass plates in a casting frame on a casting stand. After pouring, 0.2 mL of isopropanol was poured on top to get a flat top line of separating gel. After 40 minutes, when the separating gel polymerised, isopropanol was removed, and the stacking gel was poured on top. A comb with 10 wells was inserted to form sample wells, which was removed after 30 minutes. The gel was transferred to a polyacrylamide gel electrophoresis system, and the system was filled with SDS-PAGE running buffer. Prestained protein ladder and samples were loaded into the wells, and electrophoresis was run at a constant current of 35 mA per gel for 1 h or until the markers of the prestained protein ladder were separated.

3.2.7 Western transfer and immunoblotting

A piece of nitrocellulose membrane and several pieces of filter paper of the same size as the polyacrylamide gel were soaked in the Trans-Blot Turbo 1x Transfer Buffer with 20 % ethanol. The unit for Western transfer was assembled as follows: the nitrocellulose membrane was placed on five filter papers; the polyacrylamide gel was placed over the membrane, and five more filter papers were placed over it. Any air bubbles and excess buffer were carefully squeezed out using the roller, and the unit was placed into the cassette of the Trans-Blot Turbo system. The cassette was closed and inserted into the system, which ran for 10 minutes at the current 1.3 A per gel.

After the transfer, the membrane was incubated in the TBST-BSA solution for 30 min at room temperature, in order to block the non-specific protein binding sites. After washing with TBST four times for 10 minutes, the primary antibody solution in TBST-BSA (1:4000) was added to the membrane and incubated overnight at 4 °C. After that, the membrane was washed for 1 h in TBST, which was changed four times. The membrane was then incubated with secondary antibody solution (1:12000) in TBST-BSA for 1 h and washed again for 1 h with TBST. For detection, the commercial kit for chemiluminescence Amersham ECL Western Blotting Detection Reagents was used following the manufacturer's instructions. 1 mL of the reagent was placed on the membrane, and after 1 min, the membrane was imaged in the gel imaging system G:Box (Syngene).

3.2.8 Analytical size-exclusion chromatography

We used analytical SEC to check the interaction between DARPin 4m3 and CatB. Samples were diluted in 20 mM sodium phosphate, 150 mM NaCl, pH 7.2. DARPin 4m3 (20 µM) and CatB were mixed in a 2:1 molar ratio, then incubated at room temperature for 10 minutes and centrifuged at 20000 g for 10 minutes. Samples were run at a flow rate of 1 mL/min at room temperature.

3.2.9 Inhibition assays

Before the enzyme activity experiments, a titration of the active site of CatB was performed with the general cysteine protease inhibitor E64. The substrate used was fluorogenic z-RR-AMC.

We measured the activity of human and murine CatB toward the substrate z-RR-AMC in the presence of DARPin 4m3. 10 nM concentrations of murine and human CatB were mixed with DARPin 4m3 (present at 1- to 1000-fold concentration of CatB) in cathepsin activity buffer and incubated at 37°C for 15 minutes before cathepsin activity was measured in a microplate reader.

The inhibition type of DARPin 4m3 for mCatB was determined using the general modifier scheme (Baici, 1981). The Km of z-RR-AMC for mCatB was determined to be 547.7 µM. The assay's enzyme concentration (mCatB) was 1 nM, while the substrate and inhibitor concentrations varied. Substrate concentrations (z-RR-AMC) ranged from 135 µM to 1080 µM, while inhibitor (DARPin 4m3) concentrations ranged from 53.7 nM to 1448.9 nM.

To determine the inhibitor constant (K_i) for DARPin 4m3, 1 nM mCatB was incubated with various concentrations of DARPin 4m3 (1.8 nM to 294 nM) for 15 minutes at 37°C in cathepsin activity buffer. The initial rates of enzyme activity against 700 µM z-RR-AMC were then measured and nonlinearly fitted as described by Szedlacsek, Ostafe, Serban, & Vlad (1988) for the hyperbolic mode of inhibition. However, this fit can be used as a standard for any mechanism, even in the absence of tight binding conditions. The fits and graphical analyses were performed using GraphPad Prism software.

3.2.10 Nano-differential scanning fluorimetry

Protein material was diluted to a concentration of 0.5 mg/mL in 100 mM phosphate buffer (pH 6 and 7). Measurements were made in the range of 20 to 95°C with a ramp rate of 1°C/min by monitoring the intrinsic tryptophan fluorescence absorbances at 350 and 330 nm on Nanotemper Prometheus. The manufacturer's software calculated the fluorescence intensity ratio and its first derivative (PR.Stability Analysis, v1.1).

3.2.11 Surface plasmon resonance

SPR was performed with Biacore T200 on CM5 sensor chips. Mouse and human CatB (100 nM) were immobilised on the chip at 30 µl/min until it reached 700 response units (RU). After immobilisation, seven different concentrations of DARPin 4m3 were then run over the chip. Results were analysed using the Biacore T200 Kinetics Summary software (v3.2). Assays were performed in phosphate buffers at pH 6 or 7 and 25°C.

3.2.12 Cellular thermal shift assay

CETSA was performed as previously described (Jafari *et al.*, 2014). Briefly, cell cultures of RAW 264.7 were grown in DMEM medium with 10 % FBS-HI, 1 % penicillin/streptomycin, and 1 % Glutamax. Cells were pelleted and then resuspended in a cathepsin activity buffer when they reached confluence. The cell suspensions were freeze-thawed three times in liquid nitrogen and centrifuged at 32000 RCF for 20 min at 4°C. The protein content of the supernatant was measured using the Bradford assay. 10 µM of a given analyte was added to the aliquots, followed by incubation at room temperature for 10 minutes. Samples were then exposed to a specified temperature in a preheated PCR machine for 3 minutes, then at room temperature for 3 minutes, and then snap-frozen. Samples were thawed at room temperature and centrifuged again at 32000 RCF for 20 minutes at 4°C. Supernatants were transferred to new tubes and analysed by Western blot using antibodies against mouse CatB (AF965, R&D Systems). Band intensities were then analysed using ImageJ software (Schneider, Rasband, & Eliceiri, 2012). To obtain apparent aggregation temperatures (T_{agg}), band intensities were fitted in GraphPad Prism using the Boltzmann sigmoidal equation.

3.2.13 Structure determination

The complex of DARPin 4m3 and mouse CatB was formed by mixing at a molar ratio of 3:2 in 20 mM MES, 30 mM NaCl, pH 6, and incubated for 15 minutes at room temperature.

The complex was then purified by size-exclusion chromatography on Superdex 75pg at 4°C and a flow rate of 1 mL/min. The sample was then concentrated to approximately 36 mg/mL using Amicon Ultra-15 Centrifugal Filter Units with a cut-off of 10 kDa. Crystals were grown by the sitting drop method in 0.1 M Bis-Tris, 20 % PEG3350, 0.2 M sodium acetate, and pH 6.5. The ratio of protein solution to mother liquor was 2:3 (5µl drops). The crystals grew for one week. They were then soaked in mother liquor containing 10 % glycerol and frozen in liquid nitrogen until data collection. Diffraction data were obtained at a synchrotron in Trieste.

The X-ray diffraction data collection, processing, structure solution, model building, refinement, and validation of the mCatB–DARPin 4m3 crystal structure were performed by Dr. Livija Tušar, as described in Zarić et al. (2025). The structure was solved to a resolution of 1.69 Å and deposited in the Protein Data Bank under accession code 9S60. A detailed description of the crystallographic methodology is available in the original publication.

Visualisation of surface 3D models of cathepsins B and DARPins, which are coloured according to their lipophilicity, was done in ChimeraX 1.11 (Pettersen *et al.*, 2021). The lipophilicity range was set to -15 to +15, where lipophilicity is represented by a more positive value.

3.2.14 Alignment and comparison of cathepsin B sequences from different animal species

The sequences of CatB from different animal species were compared using the Clustal Omega tool (Sievers & Higgins, 2021) and coloured in Jalview (Waterhouse, Procter, Martin, Clamp, & Barton, 2009) according to the percent identity with the consensus sequence of 9 animal species (rat, mouse, human, chicken, pig, cattle, sheep, orangutan, and macaque).

3.2.15 Generation of datasets of human cathepsins K, V, L, S, F, and mouse cathepsin B substrates

The proteomic experiments and acquisition of the cleavage-site datasets were performed as described in Zarić et al. (2025). Briefly, recombinant cathepsins were incubated with human SH-SY5Y or mouse embryonic fibroblast (MEF) cell lysates. MEF lysates were used for analysis of mCatB cleavage activity to provide a cellular proteomic background from the same species as the protease, based on the observation that proteases and their substrates can undergo co-evolution (Timmer et al., 2009). Cleavage sites were identified using N-terminal COFRADIC coupled to mass spectrometry. SILAC-based metabolic labelling was used to distinguish cathepsin-dependent cleavage from endogenous proteolytic activity.

3.2.16 Docking

Protein–protein docking simulations were performed by Dr. Livija Tušar using the ClusPro 2.0 server (Kozakov *et al.*, 2017) as described in Zarić et al. (2025). Both wild-type and mutant mCatB complexes were docked under identical conditions, with and without a specified interaction region. The resulting docking models were subsequently analysed in the present work for interaction energies and hydrogen-bond networks using MAIN software (D. Turk, 2013). The mutant complex contained the I65/Q66→S65/M66 substitution.

The ClusPro cluster scores and docking results are presented in Table 6.3 and Table 6.4, respectively. Docking IDs 1357923 and 1358636 correspond to the mutated mCatB

complex with DARPin 4m3, with and without the specified interaction region, respectively, while IDs 1357926 and 1358637 correspond to the wild-type complex, with and without the specified interaction region. The calculated interaction energies for the experimental complex and the selected wild-type and mutant docking models are presented in Table 4.5. Figure 6.5 shows the electrostatic model of the 002_00 docking model for the mutant mCatB–DARPin 4m3 complex containing the I65/Q66→S65/M66 substitution. Detailed docking conditions and analyses are provided in Zarić et al. (2025).

3.2.17 Evaluation of the differences between the substrates of mouse cathepsin B and human cathepsins B, K, L, V, S, and F by support vector machine algorithm and clustering

The statistical and computational analyses were performed by Dr. Livija Tušar and included cleavage-site distribution analysis, SVM-based classification, pairwise sequence similarity analysis, and clustering, as described in detail in Zarić *et al.* (2025). The cleavage-site distributions and positions contributing to differences between cathepsins are shown in Figure 4.12. Positions with non-normal or heterogeneous residue distributions are indicated by small red circles ($p \leq 0.05$), light grey circles ($0.05 < p \leq 0.08$), or dark grey circles ($p > 0.08$). The SVM approach, previously used by Tušar *et al.* (2023) to predict cathepsin cleavage sites, was adapted in Zarić *et al.* (2025) to classify cleavage sites between pairs of cathepsins based on AUC-ROC values, where higher AUC values indicate better discrimination between the two groups. Classification performance for the P4–P4′ region is presented in Figure 6.3, while AUC values for the P3–P4′, P4–P4′, and P15–P15′ regions are summarised in Table 6.1. The resulting clusters are provided in the Supplementary Materials of Zarić *et al.* (2025).

Substrates belonging to clusters with exclusively mCatB or hCatB were used for functional enrichment analysis using g:Profiler. In addition to Gene Ontology terms, pathways from KEGG, Reactome, and WikiPathways were included as well. Only terms that satisfied the significance threshold of 0.05, according to the g:SCS algorithm, were used for analysis. From these, only the top, non-redundant terms were selected for plotting in the functional enrichment analysis (Kolberg *et al.*, 2023).

3.2.18 Data availability

The CatB–DARPin 4m3 complex was deposited in the Protein Data Bank (PDB) and was assigned the entry 9S60. Data sets of substrates of human CatK, V, B, L, S, F, and mCatB are available in the Supplementary materials of Zarić *et al.* (2025).

Chapter 4

Results

The results and their interpretation also appear in Zarić *et al.* (2025), except the results of functional enrichment analysis in Chapter 4.5.

4.1 DARPin 4m3 Selectively Binds Mouse, but not Human, Cathepsin B

DARPin 4m3 was identified using ELISA screening of clones isolated after four successive rounds of ribosome display selection using the second-generation DARPin library (Seeger *et al.*, 2013) against mouse recombinant CatB (Figure 4.1).

Size-exclusion chromatography revealed complex formation between DARPin 4m3 and active mCatB, with an elution peak around 44 kDa (Figure 4.1b), but no complex formation with hCatB or the inactive mCatB zymogen (mproCatB; Figure 4.1b). NanoDSF analysis corroborated binding, showing increased thermal stability of mCatB in the presence of DARPin 4m3 (Figure 4.1c). At pH 7, mCatB alone aggregated at 37 °C (T_{agg}), increasing to 53 °C with DARPin 4m3. At pH 6, T_{agg} increased from 48 °C (mCatB alone) to over 60 °C (complex). Notably, mCatB exhibited greater intrinsic stability at pH 6 than at pH 7, regardless of complexation.

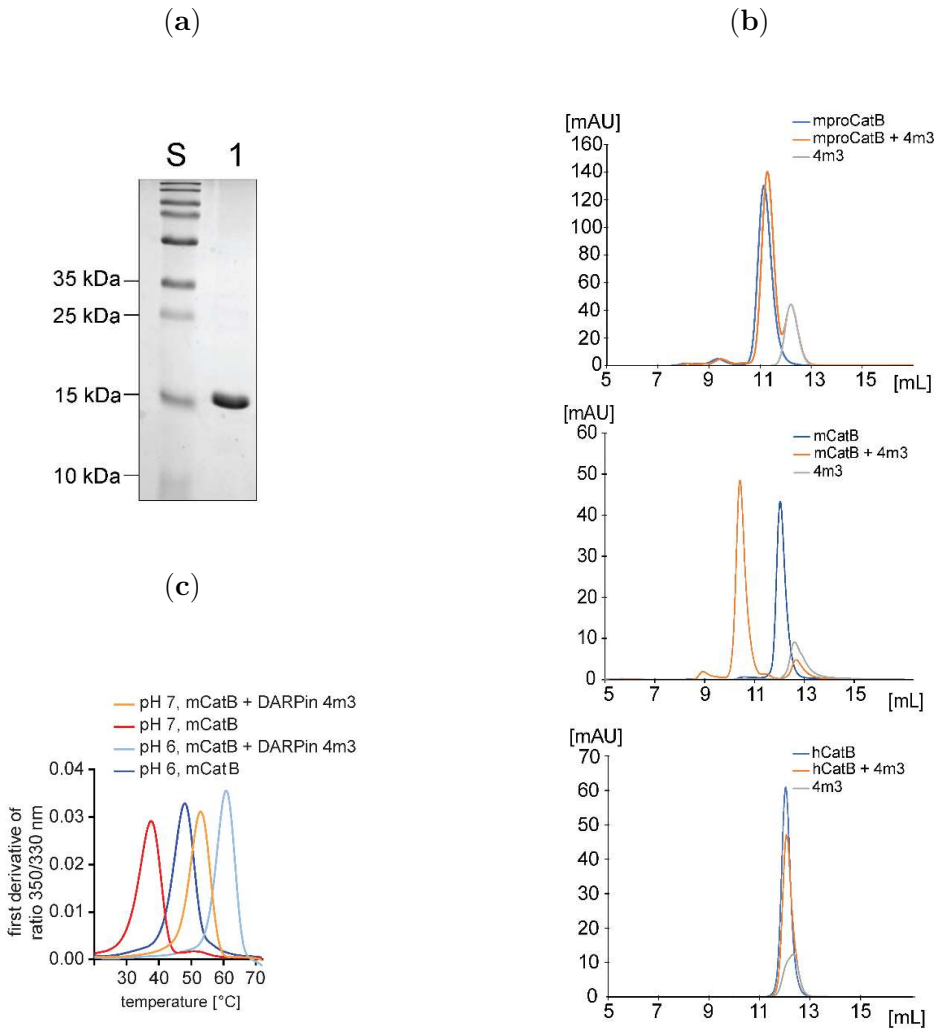


Figure 4.1: Interaction of purified DARPin 4m3 with mCatB and hCatB.(**a**) SDS-PAGE of the purified DARPin 4m3. Lane S contains molecular weight marker PageRuler Plus Prestained Protein Ladder. In lane 1, there is DARPin 4m3. (**b**) Analytical size-exclusion chromatography to assess complex formation between DARPin 4m3 and CatB, where the molar ratio of DARPin 4m3 to CatB is 2:1. The blue line represents CatB (mproCatB, mCatB, or hCatB). The orange line represents the sample in which DARPin 4m3 and CatB were mixed and incubated. The grey line represents DARPin 4m3 on SEC. (**c**) Nano-differential scan fluorometry analysis of mCatB alone and in complex with DARPin 4m3 at pH 6 or pH 7. The presence of 4m3 thermostabilises mCatB.

4.2 Binding Affinity and Mode of Inhibition by DARPin 4m3, Which is Retained in a Complex Sample

Surface plasmon resonance (SPR) analysis was used to determine the binding affinity of DARPin 4m3 for mCatB. DARPin 4m3 bound to immobilised mCatB at both pH 6 and pH 7; however, it showed no interaction with immobilised hCatB (Figure 4.2). Affinity was slightly higher at pH 6 than at pH 7 (K_D = 65.7 nM vs. 108.3 nM) (Table 4.1; Figure 4.2),

in agreement with enhanced complex thermostability at pH 6 observed by nano-DSF (Figure 4.1c).

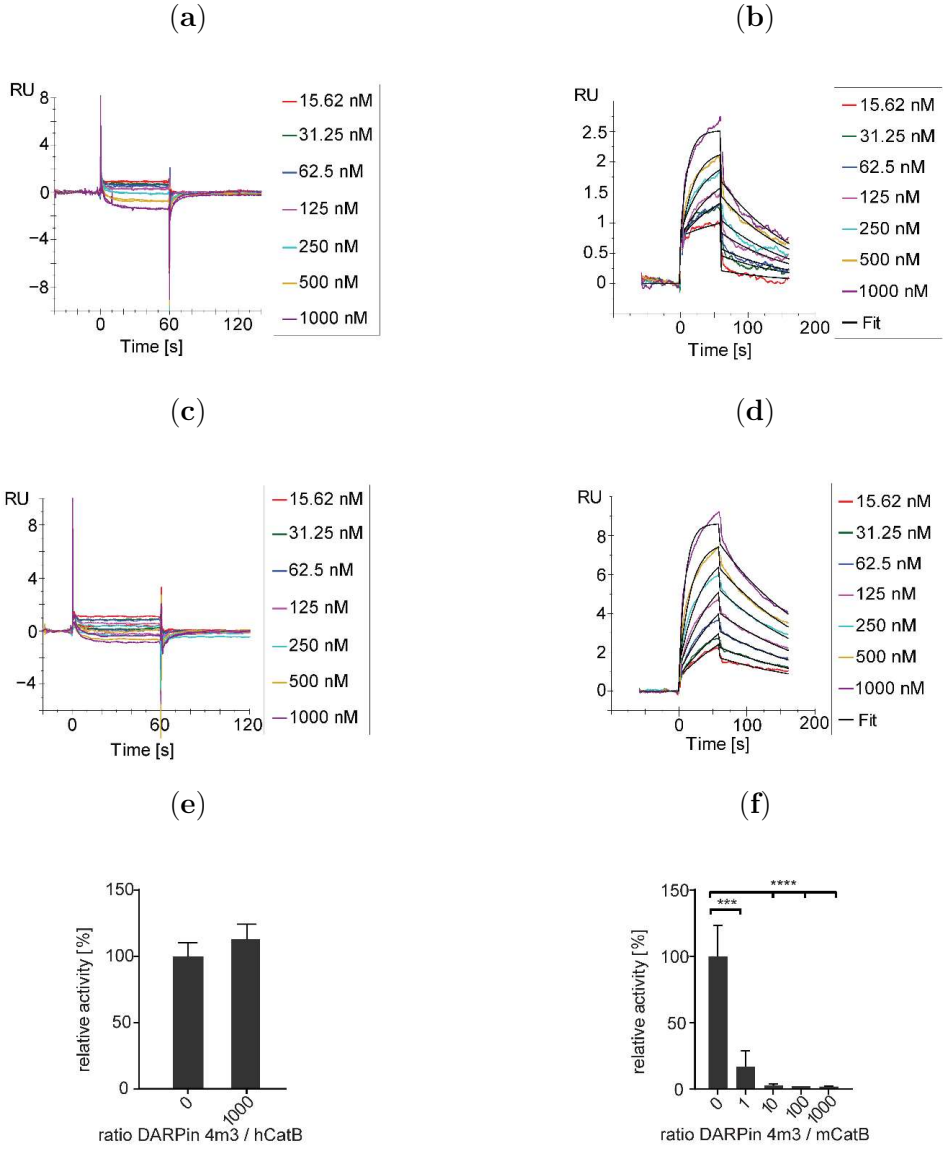


Figure 4.2: SPR and inhibition analysis of DARPin 4m3 binding and inhibition of mCatB and hCatB.(**a**) SPR of DARPin 4m3 binding to immobilised hCatB at pH 7. (**b**) DARPin 4m3 binding to immobilised mCatB at pH 7. (**c**) DARPin 4m3 binding to immobilised hCatB at pH 6. (**d**) DARPin 4m3 binding to immobilised mCatB at pH 6. DARPin 4m3 concentration used is shown for panels (**a**–**d**). (**e**) Inhibition assay of DARPin 4m3 against hCatB. Even at a 1000-fold excess, 4m3 did not inhibit hCatB activity. (**f**) Inhibition assay of DARPin 4m3 against mCatB. Reactions were performed at various molar ratios of 4m3 to mCatB. *** ($p < 0.001$) and **** ($p \leq 0.0001$) denote significant differences between groups according to one-way ANOVA/Tukey's multiple comparisons tests.

Table 4.1: Kinetic and equilibrium constants for the interaction between mCatB and DARPin 4m3. Experimental values were obtained from the SPR data shown in Figure 4.2.

	$\mathbf{k_a}$ **[1/Ms]**	$\mathbf{k_d}$ **[1/s]**	$\mathbf{K_D}$ **[nM]**	$\boldsymbol{\chi}^2$ **[RU²]**
mCatB pH 6	9.80×10^4	6.44×10^{-3}	65.7	0.021
mCatB pH 7	8.55×10^4	9.25×10^{-3}	108.3	0.004

Activity measurements demonstrated that DARPin 4m3 inhibits mCatB, but not hCatB, even at 1000-fold molar excess (Figure 4.2e, f), aligning with size-exclusion chromatography data. To elucidate the inhibition mechanism, we used the general modifier scheme (Figure 4.3a; Baici, 1981) and determined the inhibition constant using the equation for hyperbolic mixed-type inhibition (4.1; Szedlacsek *et al.*, 1988):

$$v_i = \frac{v_0}{2} \cdot \left[1 - \frac{\beta(1+\sigma)}{\alpha+\sigma}\right] \cdot \left\{ \begin{array}{c} \sqrt{\left[\left(\frac{1+\sigma}{\alpha+\sigma} \cdot \frac{\alpha K_i}{[E]_t} + \frac{[I]_t}{[E]_t} - 1\right)^2 + 4 \cdot \frac{1+\sigma}{\alpha+\sigma} \cdot \frac{\alpha K_i}{[E]_t}\right]} \\ +\alpha+\sigma+\beta(1+\sigma) \\ \frac{}{\alpha+\sigma-\beta(1+\sigma)} - \frac{1+\sigma}{\alpha+\sigma} \cdot \frac{\alpha K_i}{[E]_t} - \frac{[I]_t}{[E]_t} \end{array} \right\} \tag{4.1}$$

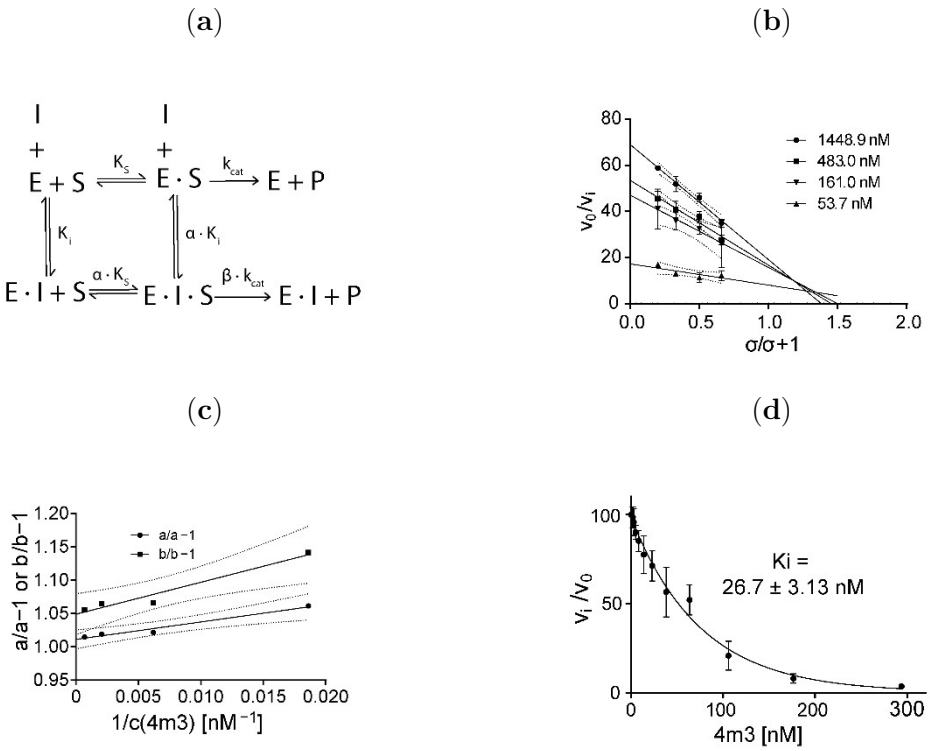
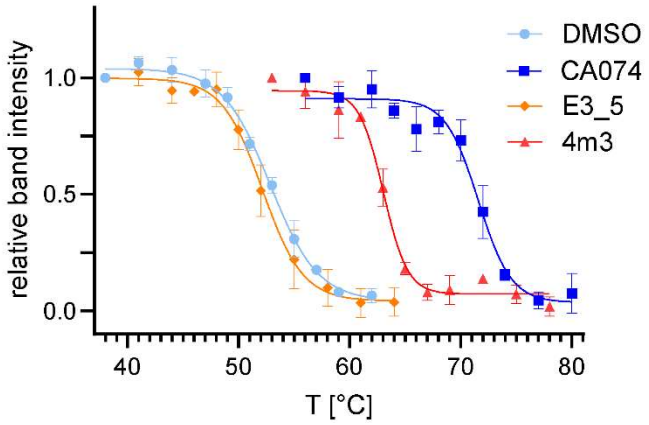


Figure 4.3: Kinetics of inhibition of mCatB with DARPin 4m3. (**a**) Enzyme inhibition (modulation) scheme according to the general modifier mechanism. (**b**) Primary specific velocity plot for mCatB in the presence of inhibitory DARPin 4m3. (**c**) Secondary specific plot for mCatB in the presence of inhibitory DARPin 4m3. (**d**) Inhibition assay of mCatB (1 nM) by DARPin 4m3 (1.80–294 nM) for determination of K_i values. The solid line represents a nonlinear fit using an equation from Szedlacsek *et al.* (1988). A K_i value of 26.7 nM was determined for the interaction.

Kinetic analysis revealed hyperbolic mixed-type inhibition ($\alpha = 1.77 \pm 0.042$; and $\beta = 0.047 \pm 0.0066$; Figure 4.3c). The parameter α, representing the ratio between competitive and noncompetitive inhibition constants, was greater than 1, indicating predominantly competitive inhibition. This confirms competition between DARPin 4m3 and the substrate z-Arg-Arg-AMC for the binding site. We calculated a competitive K_i of 26.7 ± 3.13 nM (Figure 4.3d) and, using α, a noncompetitive K_i of 47.25 nM. The parameter β represents the ratio of the turnover rate of the inhibited enzyme to that of the uninhibited enzyme. A β value greater than 0 indicates residual catalytic activity in the enzyme–inhibitor complex, albeit at reduced turnover rates.

Further, we performed cellular thermal shift analysis (CETSA) in RAW 264.7 cell lysates, a mouse macrophage cell line, which is known to express high levels of CatB, to evaluate DARPin 4m3 binding to endogenous mCatB. DARPin 4m3 successfully stabilised mCatB, increasing its apparent aggregation temperature (T_{agg}) from 52.1 ± 0.6 °C (negative control: DARPin E3_5, specific for maltose-binding protein (Kohl *et al.*, 2003) to 62.9 ± 0.5 °C (Figure 4.4a, Table 4.2). Western blot quantification using ImageJ software (Schneider *et al.*, 2012) confirmed stabilisation comparable to the positive control CA-074 (an irreversible CatB-selective E-64 analogue; $T_{agg} = 71.6 \pm 0.9$ °C), while DMSO (52.9 ± 0.6 °C; solvent control) and DARPin E3_5 showed no such effect (Figure 4.4b, Table 4.2).

(**a**)



(**b**)

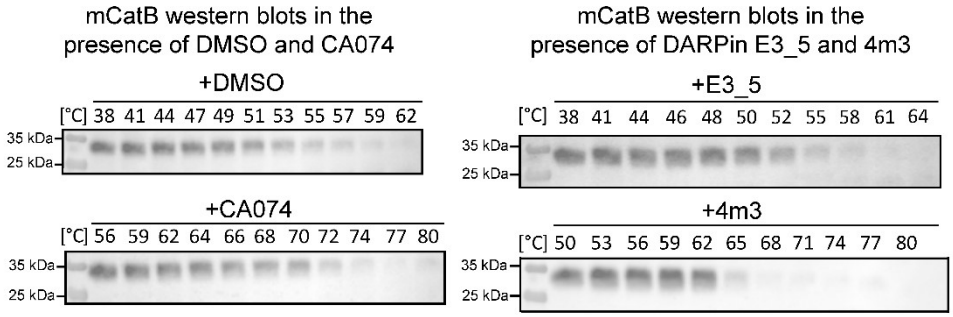


Figure 4.4: CETSA melting curves (fitted with Boltzmann sigmoid function) for mCatB in RAW 264.7 cell lysates in the presence of various analytes (DMSO, CA074, DARPin E3_5, and DARPin 4m3). Western blots of mCatB, from which band intensities were quantified, are shown below the melting curves. (**a**) Melting curves of mCatB in the presence of the mentioned molecules. Solid lines represent Boltzmann sigmoid fits of the relative mCatB band intensities from the Western blots, which were quantified by densitometry. The apparent T_{agg} of mCatB in the presence of the analytes was 52.2 ± 0.6 °C (DARPin E3_5), 62.9 ± 0.5 °C (DARPin 4m3), 52.9 ± 0.6 °C (DMSO), and 71.6 ± 0.9 °C (CA074). (**b**) On the left, there are examples of Western blots of mCatB after incubation with the selective CatB inhibitor CA074 (+) and its solvent DMSO (−) and on the right, of Western blots of mCatB in the presence of the generic DARPin E3_5 (−) and the mCatB-selective DARPin 4m3.

Table 4.2: CETSA results showing temperatures at which mCatB reached 50% of its maximum band intensity on Western blots in the presence of DARPin E3_5, DARPin 4m3, DMSO, and CA074.

	DARPin E3_5	**DMSO**	**DARPin 4m3**	**CA074**
T_{agg} (50% of mCatB maximum band intensity)	52.1 °C	52.9 °C	62.9 °C	71.6 °C
95% confidence interval	51.6 °C–52.8 °C	52.3 °C–53.5 °C	62.39 °C–63.35 °C	70.7 °C–72.4 °C

4.3 Crystal Structure Representation of the Complex of Mouse Cathepsin B with DARPin 4m3

To elucidate differences in DARPins binding between human and mouse cathepsin B, we determined the crystal structure of mCatB in complex with DARPin 4m3 at 1.67 Å resolution (Figure 4.5). The complex crystallised in the orthorhombic space group $P2_12_12_1$ with two complexes per asymmetric unit. The structures include residues L1–D254 of mCatB and D2–A160 of DARPin 4m3. Superimposition of Cα atoms yielded RMSDs of 0.37 Å (mCatB monomers), 0.39 Å (DARPin 4m3 molecules), and 0.65 Å (full complexes), indicating only minor crystal-packing effects. For comparison, hCatB monomers (PDB: 1HUC) superimpose with 0.41 Å. Pairwise RMSDs between mCatB and hCatB (0.57–0.64 Å; 83.0% sequence identity) reveal subtle structural differences. Table 4.3 shows the diffraction data and refinement statistics. Table 4.4 presents a summary of the results from pairwise structural alignment of complexes between cathepsins B and DARPins.

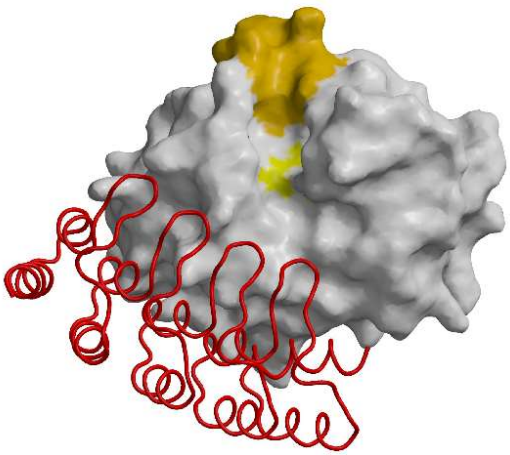


Figure 4.5: Crystal structure of mouse cathepsin B in complex with DARPin 4m3 viewed from the top of the active-site cleft. mCatB is shown as a white surface, with the catalytic residue coloured yellow and the occluding loop residues at the top in orange. DARPin 4m3 is shown as a red ribbon.

Table 4.3: Structural data of the mCatB and DARPin 4m3 complex: (**a**) the data collection and refinement statistics, and (**b**) the calculated RMSD between cathepsin B and DARPin molecules in the asymmetric unit and between human and mouse cathepsin B using superposition.

(**a**)

Data collection and refinement statistics.	
PDB entry	**9S60**
Data collection	
Space group	P21 21 21
Cell dimensions	
a, *b*, *c* (Å)	52.643, 109.921, 142.637
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	49.387 – 1.690 (1.67)
R_{means} (%)	6.0 (82.6)
I / σ*I*	22.11 (3.11)
Completeness (%)	99.9 (99.9)
Redundancy	12.44 (12.54)
Refinement	
Resolution (Å)	49.387 – 1.67
No. reflections	96,907
R_{work} / R_{free}	0.1931 / 0.2177
No. atoms	14,655
Protein	3,741 – 3,750
Inhibitor	2,262
Water	2,640
B-factors	37.79
Protein	36.77
Inhibitor	35.08
Water	49.34
R.m.s. deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.527

Data were collected from one crystal. Hydrogen atoms were excluded from the calculations.
Values in parentheses are for the highest resolution shell.

(**b**)

Molecule 1	Molecule 2	RMSD (Å)	Sequence Identity (%)
Mouse cathepsin B	Human cathepsin B	0.70	83.1
Mouse cathepsin B – molecule A	Mouse cathepsin B – molecule B	0.38	100
Mouse cathepsin B – molecule A and DARPin B	Mouse cathepsin B – molecule B and DARPin B2	0.65	100
DARPin B	DARPin B2	0.39	100

Table 4.4: Summary of the results of pairwise structure alignment.

Protein 3D Structure 1	Protein 3D Structure 2	Alignment Length	RMSD	Sequence Identities (%)	Number of Different Residues	Number of Homologs
mCatB	hCatB	253	0.68	83.0	23	20
4m3	81	158	0.65	86.8	14	6
4m3	8h6	158	0.85	88.1	7	11
8h6	81	159	0.68	89.3	6	3

4.4 Structural Comparison of the Cathepsins B, DARPins, and Their Complexes

To visualise binding similarities and differences, we superimposed all available DARPin–CatB complex structures (PDB: 5MBM; 5MBL) onto the first pair of molecules in the mCatB–DARPin 4m3 complex. The DARPin molecules revealed overall similar positioning with solvent-exposed catalytic sites, though subtle differences were evident (Figure 4.6). Specifically, the mCatB–DARPin 4m3 and hCatB–DARPin 8h6 complexes shared analogous orientations at the interdomain interface of the enzyme and L-domain surface, maintaining exposure of the Cys–His catalytic dyad (Figure 4.6).

To map interface differences, we created a 2D interaction table comparing DARPin binding across species. While 11 residues in mCatB and hCatB interacted with 11 residues in DARPin 4m3 and 8h6, only 6 hCatB residues interacted with 6 residues in DARPin 81. Colour coding revealed three interaction groups (Figure 4.7). Group 1 (shared between mCatB/hCatB and 4m3/8h6) featured key mCatB residues C63, I65, and Q66 binding DARPin 4m3 Y79, R37, and W45, respectively, while in hCatB, S65 and M66 interacted with DARPin 8h6 N88 and R48. Position 66 thus showed additional divergence (mCatB Q66 vs. hCatB M66). Group 2 (exclusive to hCatB with 8h6/81) involved hCatB N72 interacting with DARPin 8h6 R111 and DARPin 81 W73, R106, plus hCatB G73 and T125 binding DARPin 8h6 W78 and R111. Group 3 contained nearly identical residues conserved across all complexes.

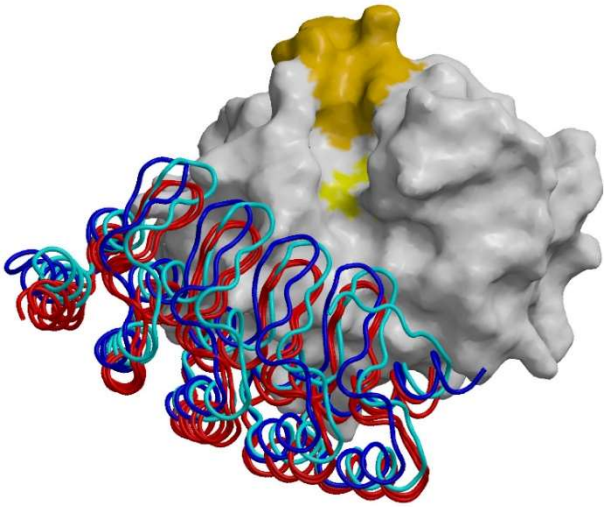


Figure 4.6: Crystal structure of mouse cathepsin B and DARPin 4m3 complex (molecules B and B2, shown in red) superimposed with DARPin 81 (cyan) and DARPin 8h6 (blue).

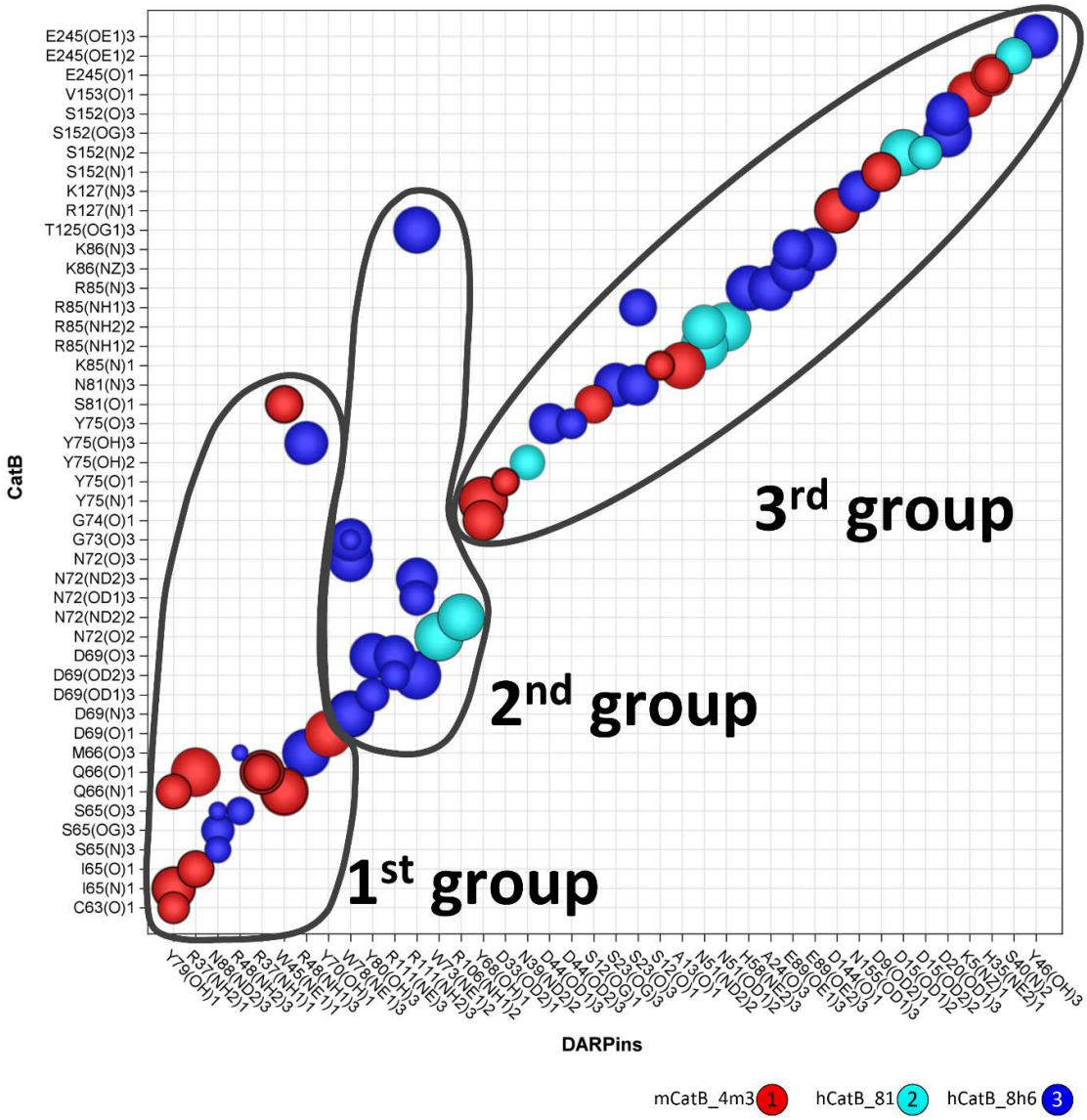


Figure 4.7: Interacting residues of mCatB/hCatB and DARPins. Each circle denotes interactions between cathepsin and DARPin residues on the vertical and horizontal axes. Circle colours (red, cyan, and blue) correspond to the mouse and human complexes, as indicated in the bottom right corner. The circle radius represents interaction distance: smaller circles indicate closer interactions, while larger circles indicate longer interactions.

To correlate interaction interfaces with sequence divergence, we mapped the interacting residues onto a sequence alignment of mouse and human CatB (Figure 4.8), colour-coding them by interaction group (shades of red). The primary contact region (residues S65 to K87 in hCatB and C63 to K86 in mCatB) coincided with the area of greatest sequence divergence between mCatB and hCatB. This is also one of the least conserved areas between species (Figure 4.9). Only minor sequence differences occurred outside this interface when comparing mCatB and hCatB. Consistent with the interaction differences analysis (Figure 4.7), the most significant variations occurred at positions 65 and 66: Ile and Gln in mCatB versus Ser and Met in hCatB. Notably, hCatB N72 formed contacts with DARPin 8h6 R111 and DARPin 81 residues W73 and R106. These interactions were absent in mCatB, where tyrosine (Y101) in DARPin 4m3, a smaller residue compared to

R111 in DARPin 8h6 or R106 in DARPin 81, occupies the structurally equivalent position (Figure 4.8).

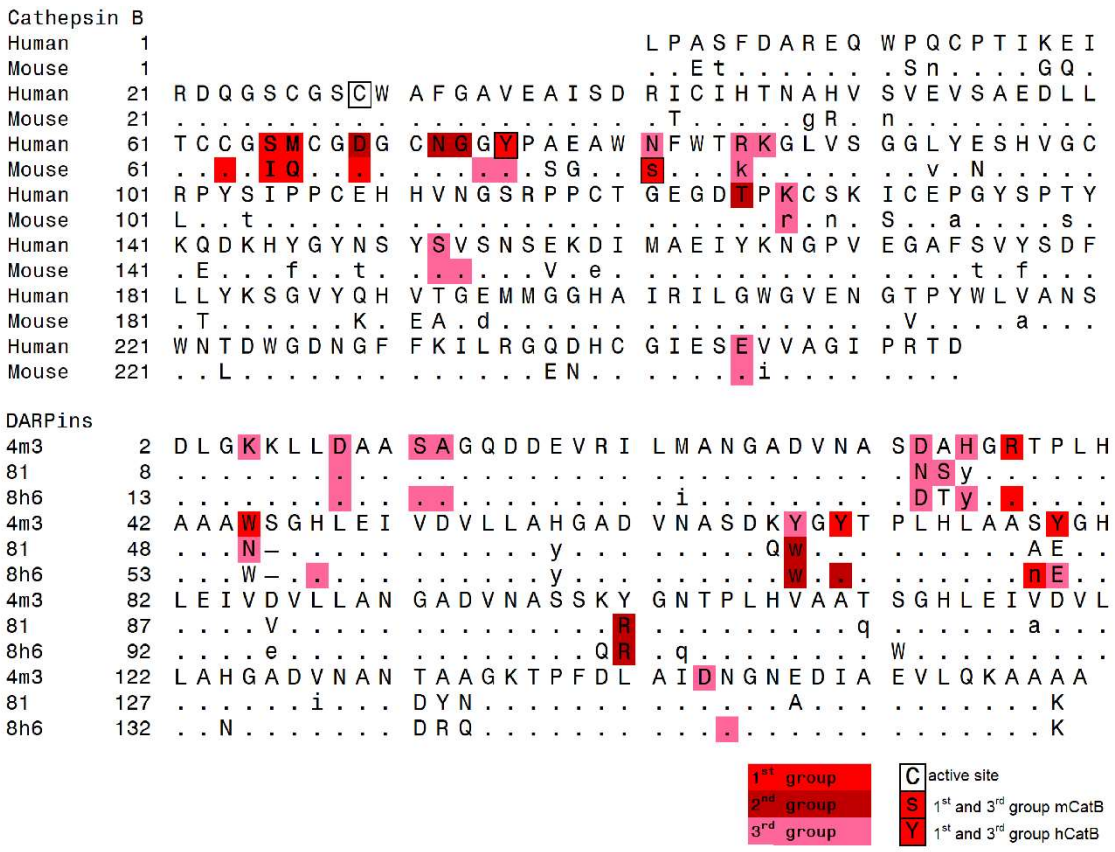


Figure 4.8: Structure-based alignment of mCatB, hCatB, and DARPins generated with FatCat (Li, Jaroszewski, Iyer, Sedova, & Godzik, 2020) and MAIN software. Identical residues are written as dots, homologous residues as lowercase letters, and different residues as uppercase letters. The active site C29 and hCatB Y65, and mCatB S81 are highlighted with frames. Residues forming interactions between cathepsin and DARPin are colored red, dark red, and Indian red, corresponding to the three interacting groups defined in Figure 4.7. Residues labelled Y and s indicate interactions that split between two groups.

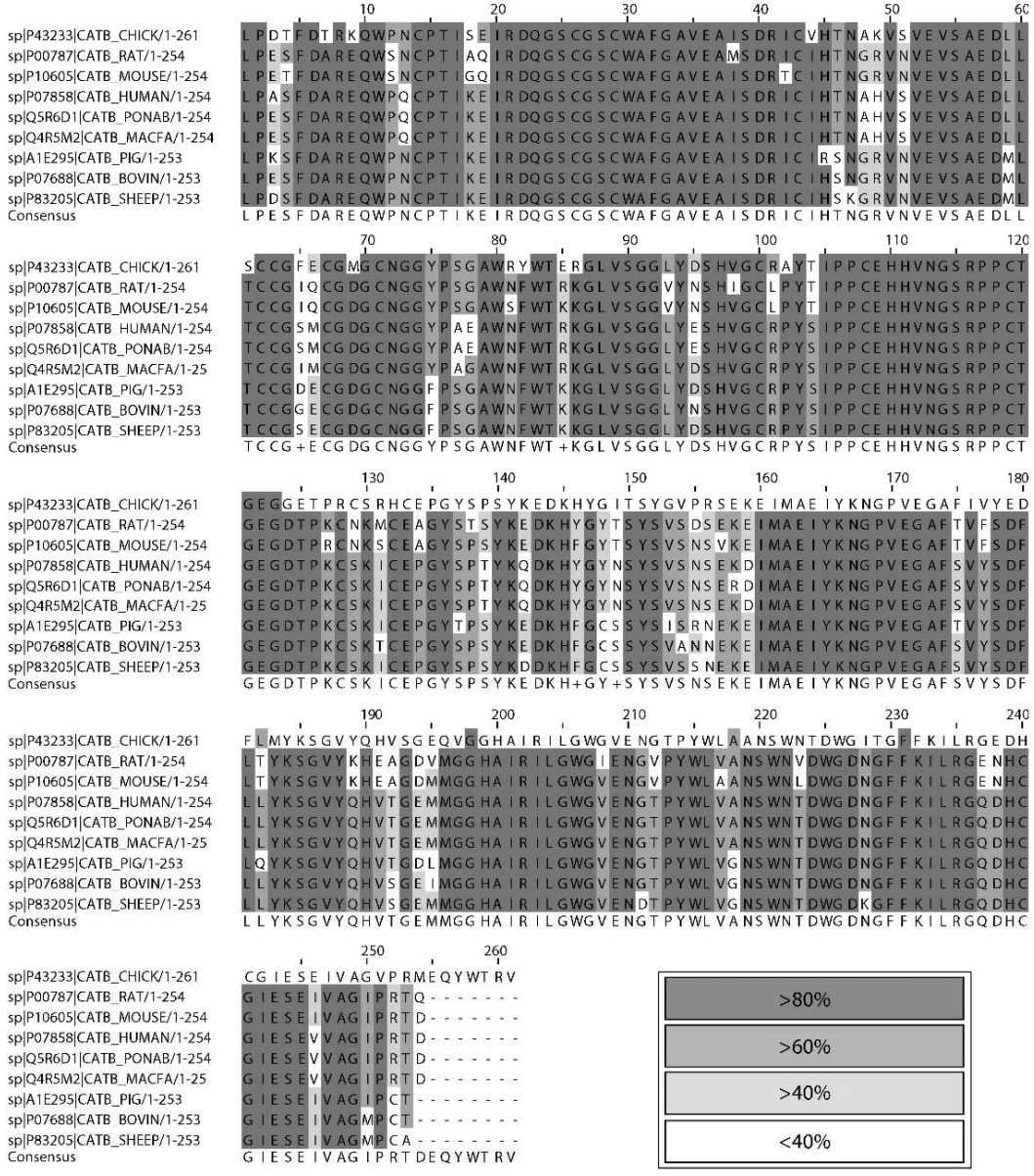


Figure 4.9: Alignment of amino acid sequences of mature CatB from various animals: chicken, rat, mouse, human, orangutan, macaque, pig, bovine, and sheep. A legend in the bottom right denotes percentage identity with the consensus sequence of these nine animal species.

To assess the impact of the I65/Q66→S65/M66 substitution on the mCatB-DARPin 4m3 interface, protein–protein docking simulations were performed using ClusPro 2.0 (Kozakov *et al.*, 2017). Both wild-type and mutant complexes were docked under identical conditions. Analysis of the top-ranked docked complexes revealed minimal changes in interfacial hydrogen bond networks upon mutation (Table 6.4 and Table 6.5). However, we observed slight differences in ClusPro docking scores, particularly when the interaction region was not specified (average difference 0.5, standard deviation 4.6; Table 6.3, IDs 1358636 & 1358637), with larger differences noted when the interaction region was specified (average difference 1235.8, standard deviation 52.8; Table 6.3).

Interaction energies, including van der Waals and electrostatic interactions, were also calculated for our complex and for the four top-ranked ClusPro docking models of both wild-type and mutated mCatB (Table 4.5) using MAIN (D. Turk, 2013). Our complex showed interaction energies of −425.3554 (molecules A/B) and −538.8692 (molecules A2/B2). For the mCatB wild-type complex, the least favourable interaction energy (−344.8188) was obtained with the 006_00 docking model (docking ID 1357926), representing a difference of −80.5366 and −194.0504 compared to the A/B and A2/B2 models, respectively. The most favourable interaction energy (−748.2675 energy units) was obtained with the 000_00 docking model. For the mutated mCatB complex, the least favourable interaction energy (−416.8979) was obtained with the 000_00 docking model (docking ID 1357923), while the 002_00 docking model yielded the most favourable interaction energy (−548.9533).

Table 4.5: Interaction energies were calculated for the mCatB–DARPin 4m3 complex (wild-type) and for docked models of the wild-type and two mutated complexes (I65/Q66→S65/M66) using the MAIN program. Both van der Waals interactions (Lennard–Jones potential) and electrostatic energies were evaluated. Interaction energies are presented for two molecular pairs derived from our complex: mCatB (molecule A) with DARPin 4m3 (molecule B), and mCatB (molecule A2) with DARPin 4m3 (molecule B2). For comparison, four top-ranked ClusPro-docked models were analysed for both the wild-type complex (docking ID 1357926) and the mutated complex (docking ID 1357923). ClusPro provides models generated under four scoring schemes: Balanced (000_00), Electrostatic-favoured (002_00), Hydrophobic-favoured (004_00), and van der Waals plus electrostatics (006_00).

Complex	**Van Der Waals Contribution**	**Electrostatic Contribution**	**Total**
mCatB and DARPin 4m3 (PDB code 9S60), molecules A (mCatB) and B (DARPin 4m3)	−85.6724	−339.6830	−425.3554
mCatB and DARPin 4m3 (PDB code 9S60), molecules A2 (mCatB) and B2 (DARPin 4m3)	−94.5953	−444.2739	−538.8692
mCatB and DARPin 4m3, docking model 000_00, ID 1357926	−28.4291	−719.8384	−748.2675
mCatB and DARPin 4m3, docking model 002_00, ID 1357926	−64.7532	−555.6167	−620.3699
mCatB and DARPin 4m3, docking model 004_00, ID 1357926	−41.1605	−529.7449	−570.9054
mCatB and DARPin 4m3, docking model 006_00, ID 1357926	−27.8825	−316.9363	−344.8188
Mutated mCatB and DARPin 4m3, docking model 000_00, ID 1357923	−65.0729	−351.8250	−416.8979
Mutated mCatB and DARPin 4m3, docking model 002_00, ID 1357923	−32.7603	−516.1930	−548.9533
Mutated mCatB and DARPin 4m3, docking model 004_00, ID 1357923	−32.6505	−426.5532	−459.2037
Mutated mCatB and DARPin 4m3, docking model 006_00, ID 1357923	−10.2918	−416.8580	−427.1498

These analyses suggest the I65/Q66→S65/M66 substitution induces subtle changes in the energetic landscape of the DARPin interaction (Figure 6.1 and Figure 6.2). The substitution of hydrophobic/polar residues to polar/hydrophobic residues may contribute to local structural divergence near the active site but does not appear to be a primary determinant of DARPin binding energetics. While docking simulations and MAIN calculations provide estimates of interaction energy and are not equivalent to experimental determination of interaction constants, these in silico analyses support the hypothesis that the I65/Q66/S65/M66 residues contribute to the energetic stability of the DARPin interaction.

Rotating the structures of the complex by ±90 ° with minor spatial adjustments visualises the interaction groups by mapping their colour codes from Figure 4.8 onto the molecular surfaces (Figure 4.10).

In Figure 4.11, a visual representation of opened complexes of mouse and human cathepsin B and corresponding DARPins with coloured surfaces according to lipophilicity shows the difference between I65 of the mCatB and S65 of the hCatB. When mCatB forms a complex with DARPin 4m3, it uses a more lipophilic I65 to additionally stabilise the interaction, since the DARPin 4m3 has a tiny hydrophobic pocket formed by Y70, L75 and V108, where I65 is located in the complex. In the case of hCatB, these hydrophobic interactions do not occur due to the presence of S65 in place of I65.

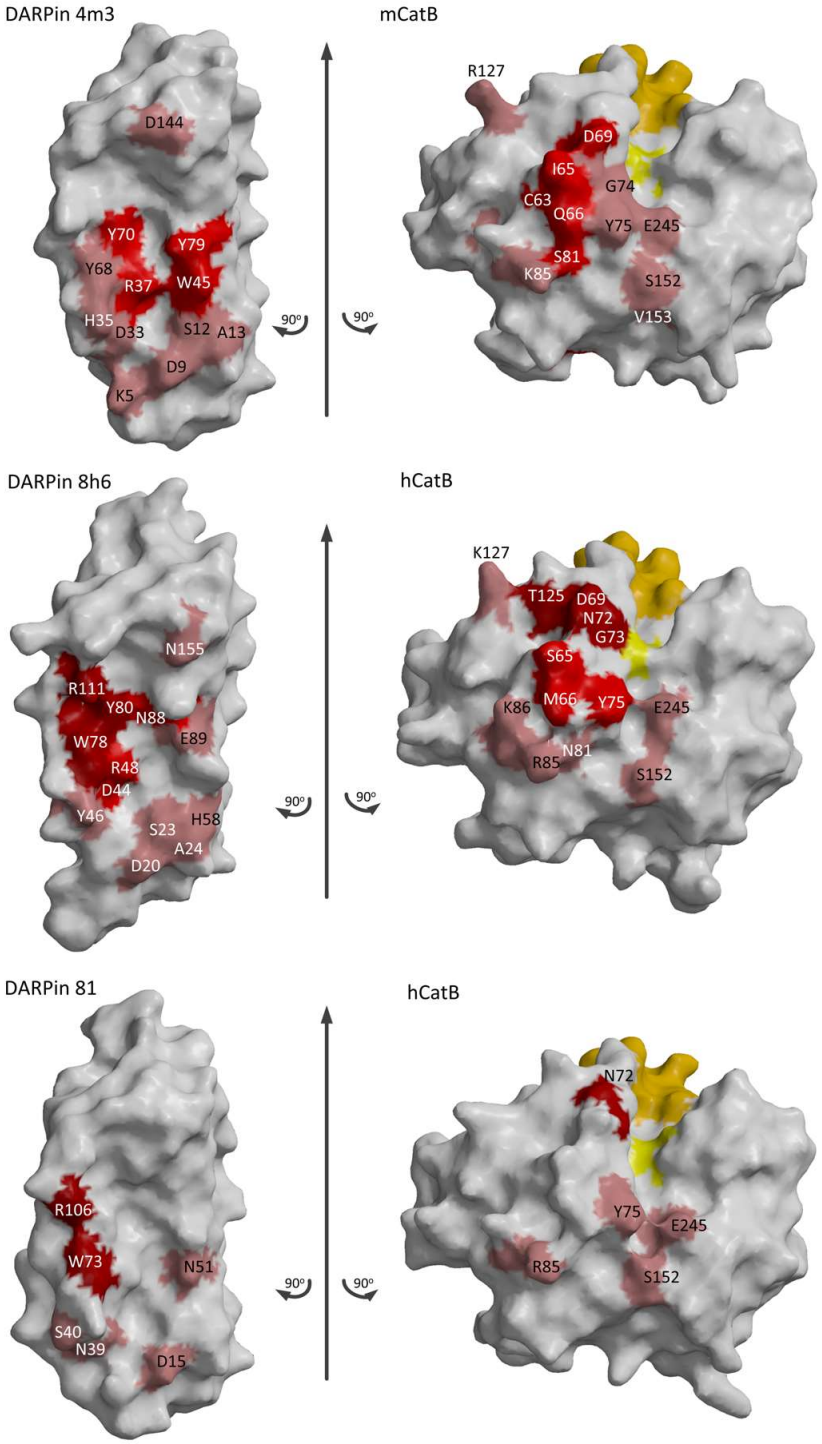


Figure 4.10: 3D visualisation of the interacting residues between mCatB/hCatB and DARPins, divided into three groups (red, dark red, and Indian red) as defined in Figure 7. The active site residue is shown in yellow, and the occluding loop residues are shown in orange. Protein structures were generated with MAIN (D. Turk, 2013) and rendered with Raster3D (Merritt & Bacon, 1997). Arrows indicate the direction of the rotation of the complex subunits.

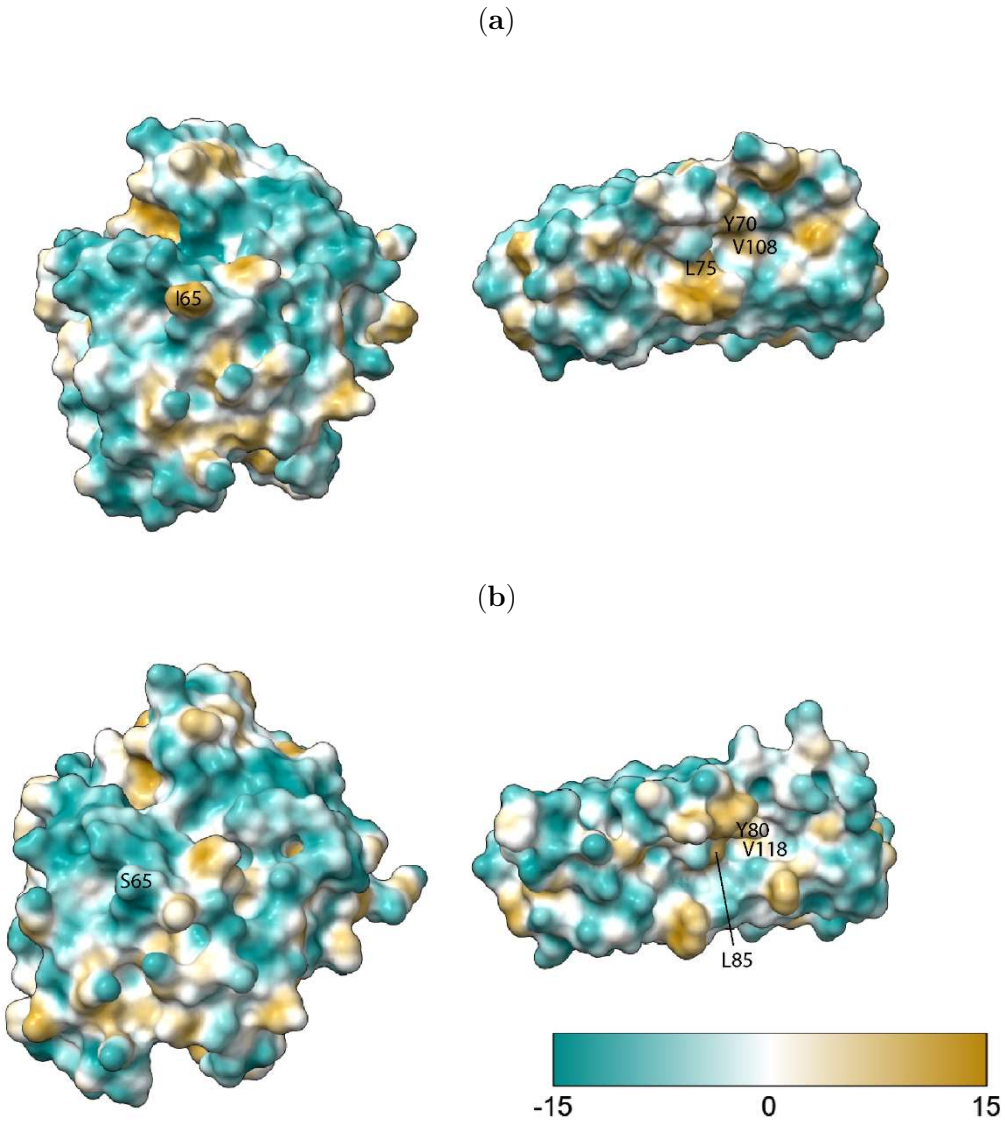


Figure 4.11: 3D visualisation of DARPin molecules with cathepsins B, with surface coloured according to lipophilicity. **a**) On the left, there is mCatB with marked I65, with DARPin 4m3 on the right, which has marked amino acids where I65 from mCatB ends up after formation of the complex. **b**) On the left, there is hCatB with marked S65, which structurally corresponds to I65 on mCatB. On the right, there is DARPin 8h6 with the same marked amino acids in the structurally equivalent positions as with DARPin 4m3 above. Bottom right is the legend for lipophilicity, where positive values represent higher lipophilicity and negative values represent lower.

4.5 Substrates as Indicators of Differences Between Mouse and Human CatB

To compare the specificity between mouse and human CatB, a mass spectrometry analysis of MEF cell lysates treated with mCatB was performed using the same approach as before.

The data for human cathepsins K, V, B, L, S, and F were also taken from the same study (Tušar *et al.*, 2023).

The peptides/substrates data were then analysed using Schechter-Berger nomenclature. The Anderson–Darling test revealed non-normal residue distributions (heterogeneous positions) at 7 substrate positions (P3–P4′) for mCatB (Figure 4.12; structurally visualised in Figure 6.4), compared to 5 positions for hCatB. This suggests mCatB engages more substrate-binding subsites during catalysis.

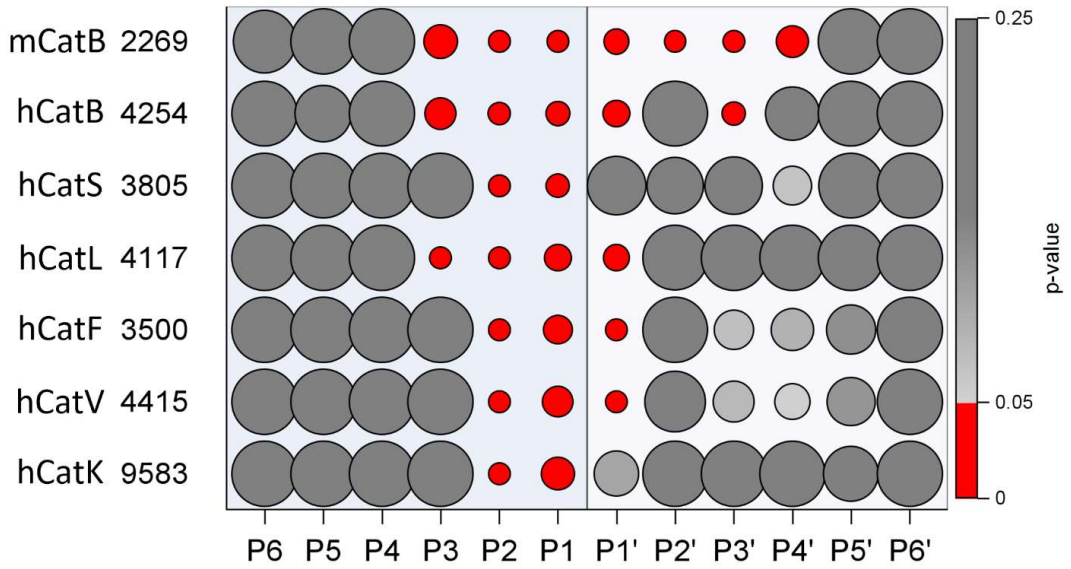


Figure 4.12: Normality of substrate residue distributions. The p-values for the normality of the distributions of the residues at the positions from P6 to P6′ (columns) for each cathepsin (rows) are indicated by the size of the circles. The red, light grey, and grey circles indicate p-values ≤ 0.05 (not normal), >0.05 and ≤ 0.08 (normal but close to the limit 0.05), and >0.08, respectively.

Previously, Support Vector Machine learning algorithms were used to predict cleavage sites by differentiating cleaved and non-cleaved sequences (Tušar *et al.*, 2023). Modelled performance was quantified using the area under the Receiver Operating Characteristic curve (AUC-ROC). Here, AUC was repurposed as a similarity metric for pairwise comparisons of cathepsin cleavage datasets for a different goal - discrimination of two datasets. The AUC values for pairwise comparisons of substrate cleavage datasets across cathepsins were thus calculated by selecting different ranges of positions: P3–P4′ (Figure 4.13, Table 6.1), P4–P4′ (Figure 6.3, Table 6.1), and P15–P15′ (Table 6.1). The differences between the peptides of mCatB vs. hCatB and human cathepsins K, V, L, S, and F are most obvious when positions P3–P4′ and P4–P4′ are considered (Table 6.1). The lower AUC values indicate greater similarity between two data sets (reduced discriminative power), while higher AUC values reflect greater dissimilarity (enhanced discriminative power). The colour gradient (light pink towards maroon red shades) represents increasing dissimilarity (higher AUC). Mouse and human CatB datasets showed strong similarity (AUC = 0.73), yet they remained distinguishable from each other and formed a distinct cluster relative to the datasets of other cathepsins (Figure 4.13). Notably, hCatV and hCatF data sets exhibited the highest similarity (lowest AUC, light pink).

The discrimination of mCatB and hCatB cleavage sites encouraged us to seek substrates specific to hCatB or mCatB. To identify species-specific substrates, the combined cleavage data sets were clustered. Using a maximum of 400 clusters, we obtained 22 clusters containing 227 substrates (complete data available in the Supplementary data of Zarić *et al.* (2025)). These included 18 clusters (218 substrates, belonging to 180 unique proteins) specific to hCatB and 4 clusters (9 substrates, 9 unique proteins) specific to mCatB (Table 6.2). The hCatB-specific substrates represented 161 unique cleavage sequences, which include actins, tubulins, kinases, ribosomal proteins, histones, and nucleolar proteins (Supplementary data of Zarić *et al.* (2025)). This further supports the idea that mouse and human cathepsin B do not have entirely identical physiological roles, although the differences in substrate list may in part reflect the different types of cell lines used in the proteomic analysis—mouse embryonic fibroblasts (MEFs) vs. human neuroblastoma cell line (SH-SY5Y).

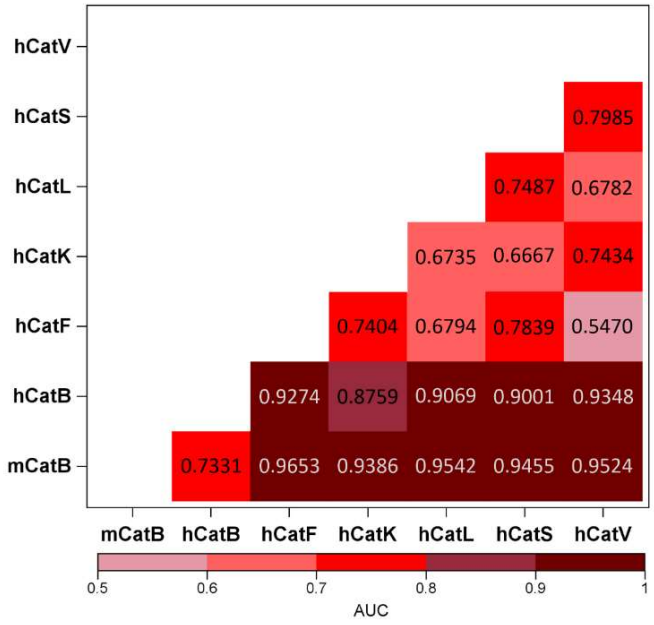
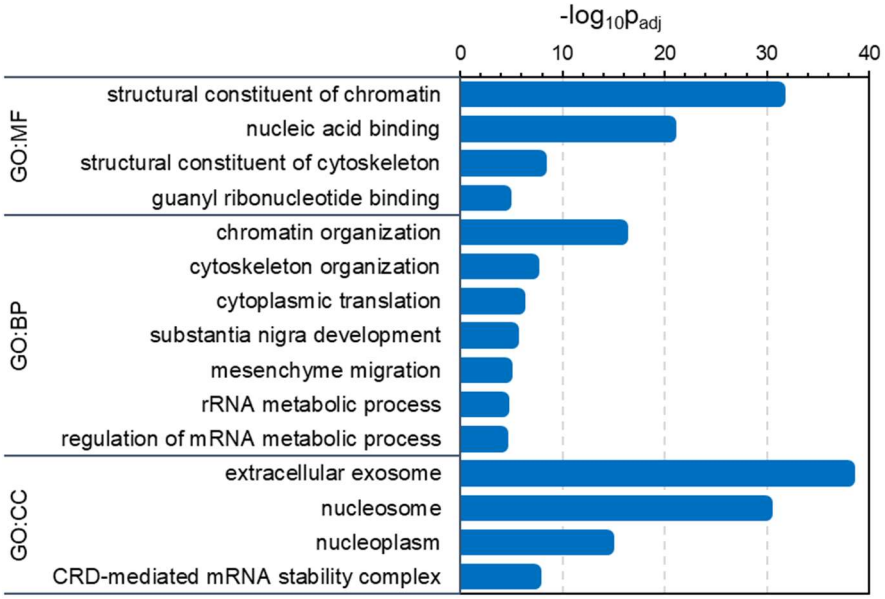


Figure 4.13: Calculated AUC values for pairwise comparisons of cleavage sites from different cathepsins for the P3–P4′ region, which proved most informative for distinguishing cathepsin cleavage site specificities. Values between 0.5 and 0.6 (light pink) indicate the most similar substrate groups; 0.6 to 0.7 (rose), 0.7 to 0.8 (red), 0.8 to 0.9 (deep pink), and 0.9 to 1.0 (maroon) represent progressively greater differences, with maroon indicating the most distinct groups.

Gene set enrichment analysis (Figure 4.14) was done by using g:Profiler (Kolberg *et al.*, 2023). Analysis was performed on substrates from clusters unique to mCatB or hCatB. In the case of hCatB, 180 unique proteins belonged to clusters exclusive to hCatB. On the other hand, there were only 9 unique proteins that belonged to clusters exclusive to mCatB, and these do not represent the complete set of mCatB substrates. In the case of hCatB (Figure 4.14a), the top terms were associated with the structure and regulation of chromatin, binding and metabolism of nucleic acids. In GO:CC, the top scores were extracellular exosome and nucleus-associated terms, such as nucleosome, nucleoplasm, and mRNA stability complex. Mouse CatB showed nucleus-related terms as top scores as well (Figure 4.14b), but focused more on chromatin/histone organisation and regulation.

However, because these results are based on only nine proteins, they should be interpreted with caution and be considered exploratory. More experiments would be needed to confirm if there are any biologically relevant differences in substrate processing which this analysis cannot really confirm. Additionally, because mCatB and hCatB were profiled using different cell lines (MEFs and SH-SY5Y), differences in identified substrates may reflect both intrinsic enzyme specificity and differences in the underlying cellular proteomes. Consequently, species-specific substrate assignments should be interpreted with caution.

(a) hCatB



(b) mCatB

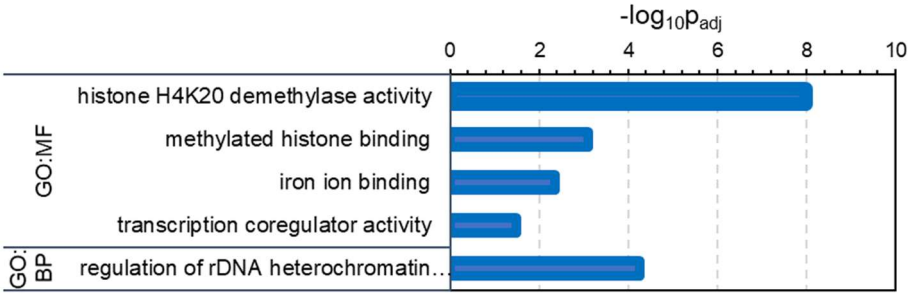


Figure 4.14: Functional enrichment analysis, where only top statistically significant and non-redundant terms are shown according to -$\log_{10}p_{adj}$. GO:MF stands for Gene Ontology terms related to molecular function. GO:BP and GO:CC stand for biological process and cellular compartment, respectively. **a**) Top results from enrichment analysis based on substrates from clusters which are processed exclusively by hCatB. **b**) Top results from enrichment analysis of substrates from clusters processed exclusively by mCatB. In both cases, we can see emphasis on interaction with nucleic acids.

Chapter 5

Discussion

Cathepsin B has been one of the most studied human cysteine cathepsins. It is the most abundant and has numerous roles in physiological and many pathological conditions, including cancer and arthritis. In addition to its direct therapeutic potential, the enzyme was widely targeted for prodrug and antibody-drug conjugate activation and evaluated for diagnostic imaging as well as for a target for targeted drug delivery in cancer animal models (Biasizzo *et al.*, 2022; Kramer, Renko, *et al.*, 2017; Kramer, Turk, & Turk, 2017; Mikhaylov *et al.*, 2014). However, despite a vast amount of information about human cathepsin B (Keppler & Sloane, 1996; Kos, Mitrović, & Mirković, 2014; Xie *et al.*, 2023), little is known about mouse cathepsin B (Caglič *et al.*, 2009), including a detailed comparison of inhibitory efficacy across species. In their mature active forms, human and mouse CatB (hCatB and mCatB, respectively) exhibit 83% amino acid identity and similar biochemical properties, such as K_i for interaction with chicken cystatin C and k_{on} for irreversible cathepsin B inhibitor CA-074 (Figure 4.9, Caglič *et al.* (2009)). Still, despite sequence similarities among cathepsins from different species, many inhibitors show different affinities for human and rodent homologs, as shown in studies by Marquis *et al.* (2001) and Desmarais *et al.* (2009). Inhibitors showing high affinities for mouse variants and lower affinities for human ones are problematic since there is poor translatability in the context of the development of a pharmaceutical drug. On the other hand, inhibitors that show great affinities for human but not for mouse variants are hard to validate in preclinical models, necessitating costly use of primate (Jerome, Missbach, & Gamse, 2011; Kumar *et al.*, 2007; Lindström *et al.*, 2018; Stroup *et al.*, 2001) and rabbit models (Pennypacker *et al.*, 2011; Pennypacker, Oballa, Levesque, Kimmel, & Duong, 2013; Takahashi *et al.*, 2009). Another approach is to use humanised mice or a mouse xenograft model; however, these models introduce additional considerations due to the need for specialised housing and handling of immunodeficient animals. Additionally, immunodeficient mice provide a less realistic tumour microenvironment due to the incomplete immune system, which plays an important role in tumour-stroma interactions, tumour progression, and therapeutic response (Kähkönen, Halleen, & Bernoulli, 2020; Richmond & Su, 2008).

To address this gap, we identified mCatB-selective DARPin 4m3, which has been used for biochemical, crystallographic, and kinetic studies of mCatB. 4m3 displays selective and strong binding in the nanomolar range, as shown by the SPR measurement (K_D = 65.7 nM) and the determination of K_i (26.7 nM). The discrepancy between the SPR measurement of K_D and the determination of K_i by the inhibition assay for 4m3 against mCatB could be due to the methodology. In the SPR experiment, the mCatB molecules are bound to the sensor chip via amines, and DARPin 4m3 may not have optimal access to the binding pocket, whereas in the solution-based inhibition assay, this is not an issue. A similar difference is observed in the case of DARPin AR_F8 against caspase-2 (Schweizer *et al.*, 2007). Additionally, SPR shows a higher affinity of DARPin 4m3 for mCatB at pH

6 than it does at pH 7, which is likely due to the greater stability of mCatB at lower pH values (Figure 4.1c). Resolving the structure and comparison to previously determined structures of DARPin 8h6/81 and hCatB shows that DARPin 4m3 binds in a similar orientation on the same region as in the case of hCatB, where it binds adjacent to the active site but not directly over it (Figure 4.6). This is also corroborated by kinetics analysis, which demonstrates the same type of inhibition, hyperbolic mixed-type inhibition (Figure 4.3), where the inhibitor binds in such a way that the enzyme still retains a partial substrate-processing ability, as shown by the parameter β, which is greater than 0. The same type of inhibition was observed with DARPin 8h6 and 81 against hCatB, and with DARPin AR_F8 against caspase-2 (Kramer, Renko, *et al.*, 2017; Schweizer *et al.*, 2007). It should be noted that the substrate used consists of only two amino acids (z-RR-AMC), whereas physiological substrates are much larger proteins.

Analysing DARPin 4m3 binding contacts, we see that they occur at regions that were most different between mCatB and hCatB, indicating that the selection process of DARPin 4m3 targeted the most significant areas of difference (Figure 4.8 and Figure 4.9), which are evolutionarily less conserved among species (Figure 4.9). Notably, three distinct interaction groups at the DARPin interface (Figure 4.7) may contribute to the high selectivity of 4m3, showcasing the potential of DARPin scaffolds to exploit subtle ortholog differences. Interactions between DARPins and CatB can be separated into three groups, suggesting that the residues I65, Q66 in mCatB and S65, M66 in hCatB are likely contributors to the differences between mCatB and hCatB when it comes to binding to DARPins. Additionally, 8h6 residues R111 and W78 form bonds with hCatB via N72 and T125 that are also present on mCatB. In DARPin 81, these interactions are also present, while in the case of 4m3, these interactions are not possible since R111 and W78 are replaced by tyrosine at both sites. However, in DARPin 4m3, Y79 forms contacts with I65, M66, and C63 on mCatB, whereas in DARPins 8h6 and 81, Y79 is replaced by glutamic acid in a structurally equivalent position. DARPin 4m3 can also bind S81 on mCatB via its W45, while in hCatB, S81 is replaced by N81. 8h6 possesses W45 (81 has asparagine instead) as well, and compared to DARPin 81 has stronger binding to mCatB (Kramer, Renko, *et al.*, 2017). Interaction between DARPin 4m3 and mCatB differs from DARPins 81 and 8h6 binding hCatB in another aspect, and this is its hydrophobic interactions between I65 on mCatB and Y70, L75, and V108 on DARPin 4m3 (Figure 4.11), stabilising the interaction further. But *in silico* interaction studies showed that replacing I65S and Q66M was not the most crucial factor for driving interaction. Nevertheless, analysis confirmed that these residues still affect the stability of a complex in a meaningful way. It has to be noted that *in silico* studies cannot replace experimental data, but can still offer valuable insight into the molecular dynamics.

While active site subsites show identical residues (Figure 6.1 and Figure 6.2), substrate specificity analysis using SVM-based machine learning reveals differences in cleavage preferences between mCatB and hCatB, although with greater mutual similarity (AUC = 0.73 compared to CatL or CatV (AUC = 0.95), Figure 4.13, Table 6.1). Together with the observation that only 22 out of 400 clusters were exclusive to either hCatB or mCatB, this indicates that these proteases recognise highly similar cleavage sequences, which aligns with their closer phylogenetic relationship (Reiser, Adair, & Reinheckel, 2010).

Clustering identified species-specific cleavage patterns (Supplementary materials of Zarić *et al.* (2025) providing a pathway to develop selective substrates. Additionally, clusters that were unique to either mCatB or hCatB showed that these substrates were linked with DNA and RNA homeostasis functionality (Figure 4.14). GO term enrichment analysis of unique substrates showed, in both cases, increased terms related to DNA metabolism and regulation, especially relating to histones/chromatin. This shows that these two enzymes may have overlapping biological processes, even in cases of unique sequences

of substrates. These findings are consistent with previous studies suggesting that CatB may be involved in nuclear processes related to transcription and translation processes. It must be said that in the case of mCatB, we managed to obtain only 4 clusters with mCatB-specific cleavages. These 4 clusters contained 7 unique cleavage sites belonging to 9 different proteins (Table 6.2), while there were 18 hCatB-specific clusters with 160 unique cleavage sites belonging to 180 unique proteins (Supplementary materials of Zarić *et al.* (2025)), reflecting a larger number of identified hCatB-specific cleavage events in this dataset. The interpretation of mCatB-specific substrates requires caution, as only four clusters containing seven unique cleavage sites from nine proteins were identified. In comparison, hCatB-specific clusters comprised 18 clusters containing 160 unique cleavage sites from 180 proteins. This difference may reflect genuine differences in substrate processing, but it may also be influenced by differences in dataset size, cellular background, and proteome composition. An additional limitation of the substrate comparison is that human CatB and mouse CatB were analysed using lysates from different cell types. SH-SY5Y and MEF cells have distinct proteomic compositions, which may influence the availability of potential substrates and therefore the observed cleavage profiles. Consequently, differences between mCatB and hCatB substrate datasets cannot be attributed exclusively to species-specific enzyme preferences. For a more reliable interpretation, additional datasets would need to be obtained to determine if these enrichments reflect genuine biological differences between mouse and human CatB. Also, our substrate data were gained from cell lysates, and this may not fully reflect substrate processing in intact cells, where compartmentalisation and regulatory mechanisms influence protease activity. Taken together, these limitations mean that the observed differences between the mCatB and hCatB substrate datasets should not be interpreted as definitive evidence of species-specific substrate preferences. Nevertheless, the overall similarity observed between mCatB and hCatB cleavage patterns compared with other human cathepsins supports their close functional relationship. Additionally, the broader substrate specificity analysis, based on cleavage-site patterns across larger datasets, demonstrated that mCatB and hCatB are more similar to each other than to other human cathepsins. This supports the close functional relationship regarding substrate recognition between the two orthologs, while suggesting that only a subset of substrates may be processed in a species-specific pattern.

The identification of nuclear-associated proteins among CatB substrates is consistent with previous studies that have reported the presence of CatB in the nucleus and have described its nuclear functions. One way could be translocation into the nucleus – bypassing the mannose-6-pathway – as a consequence of truncation, which leads to CatB missing a part of the signal sequence, as well as parts of the N-terminal proregion (Bestvater, Dallner, & Spiess, 2005). It was found to degrade sirtuins in the nucleus in senescent microglia, inducing neurodegeneration (Meng et al., 2020). It was also found to degrade histones H1 in thyroid carcinoma cells and H3 in genetic and pharmacological models of ferroptosis (Nagakannan, Islam, Conrad, & Eftekharpour, 2021). This kind of proteolytic activity in the nucleus has been proposed to influence transcriptional regulation and potentially drive disease progression and malignancy, as in the example of oncogene-induced senescence (Duarte *et al.*, 2014). Interestingly, in our experiment, mCatB and hCatB both demonstrated cleavages of members belonging to H1 and H2A families, while only hCatB also cleaved members of H2B, H3, and H4 families of histone proteins (see Supplementary materials in Zarić *et al.* (2025)). Although these findings should be interpreted cautiously, they may reflect differences in substrate processing that warrant further investigation.

Beyond the biological findings, the use of the SVM-based approach here to analyse cleavage sequences may have broader methodological applications. This approach could be extended to other proteases and provide a comparative value for proteases, both within and between species, providing a framework for identifying functional similarities and

differences between proteases. However, such applications would depend on the generation of sufficiently large and high-quality proteomics data sets. Furthermore, this approach could also be used to find selective cleavage sequences, which could be used for the development of selective substrates and protease-specific probes for research or diagnostic applications.

Collectively, mCatB and hCatB appear highly similar proteases with only limited functional differences, supporting the continued use of mouse models in CatB-related research, while also indicating that careful validation may be required to obtain highly selective inhibitors and probes. Our mCatB structure enables rational design of murine-specific probes and inhibitors (Katunuma, 2011), facilitating translational research. Species-selective inhibitors like DARPin 4m3 may also help to dissect host-derived CatB roles in xenograft tumour models (Olson & Joyce, 2015; Sloane *et al.*, 2005), distinguishing tumour-cell versus microenvironmental contributions (Gocheva *et al.*, 2010; Vasiljeva *et al.*, 2006). Thus, while mCatB remains valuable for human CatB research, its differences necessitate context-specific validation when interpreting species-selective interactions or developing highly selective CatB targeting strategies.

Chapter 6

Conclusions

We characterised and demonstrated the selectivity of the mouse CatB-selective DARPin 4m3 with a high-affinity binding K_D of 65.7 nM and potent inhibition K_i of 26.7 nM for mCatB, and determined it to be a hyperbolic mixed-type inhibition with predominantly competitive character according to the general modifier mechanism.

By determining the structure of mCatB in complex with the mCatB-selective DARPin 4m3, we show that observed species selectivity is associated with interactions involving less conserved surface regions adjacent to the active site, providing a structural rationale for the proposed mechanism of binding and selectivity. Additionally, the DARPin platform enabled the exploitation of subtle differences between orthologs, which are often inaccessible to small-molecule inhibitors. Such selective tools could provide the ability to gauge CatB function *in vivo* without cross-species effects, which could be used in mouse xenograft tumour models for host-derived CatB, helping to disentangle tumour cell and tumour microenvironment-derived CatB functions in such models.

Beyond inhibitor binding, SVM-based substrate analysis uncovered shared and potential species-specific differences in cleavage preferences. There is overall similarity, but there are distinct clusters of unique substrates, especially relating to DNA and RNA metabolism and regulation. These differences, including the enrichment of proteins associated with DNA and RNA metabolism, suggest potential functional differences, but these are preliminary and need further investigation, particularly given the limited number of mCatB-specific substrates and the different cellular background used for substrate profiling.

This work demonstrates that mouse CatB closely resembles human CatB in substrate specificity, but it may not be fully representative of hCatB in all contexts, and these differences should be carefully considered when designing and interpreting studies of hCatB-related conditions. The structure presented here enables rational design of mouse-specific probes and inhibitors that can be valuable for pre-clinical studies. Taken together, this work shows that mCatB is sufficiently similar to hCatB to serve as a relevant experimental model for hCatB-related research, while careful context-dependent validation remains essential to maximise translational relevance in CatB-targeted research and therapeutic development.

Appendix

Appendix figures and tables appear in Zarić *et al.* (2025) as well.

A.1 Appendix Figures

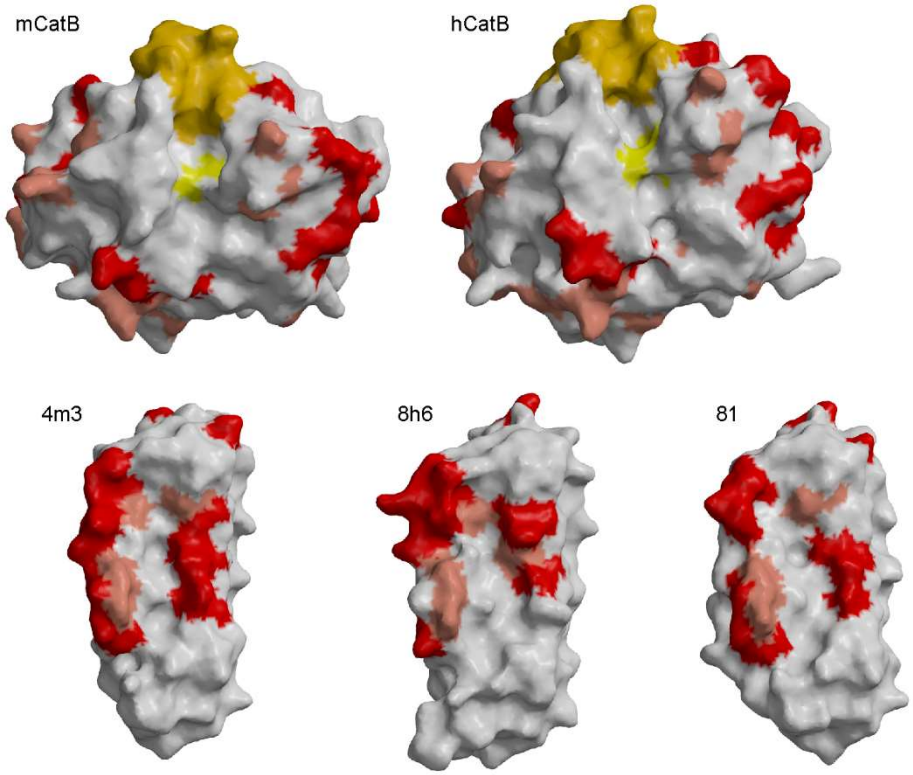


Figure 6.1: Structure-based alignments of mCatB and hCatB and DARPins with MAIN with coloured differing residues. Areas with different amino acids are shown in red, and homologs are shown in tomato red. The active site is coloured yellow, and the amino acid residues of the occlusion loop from 108 to 119 are coloured orange. Residues involved in the interaction between cathepsin B and DARPins are indicated by a one-letter code and a number. Protein structures were generated using MAIN software (D. Turk, 2013) and rendered using Raster3D (Merritt & Bacon, 1997)

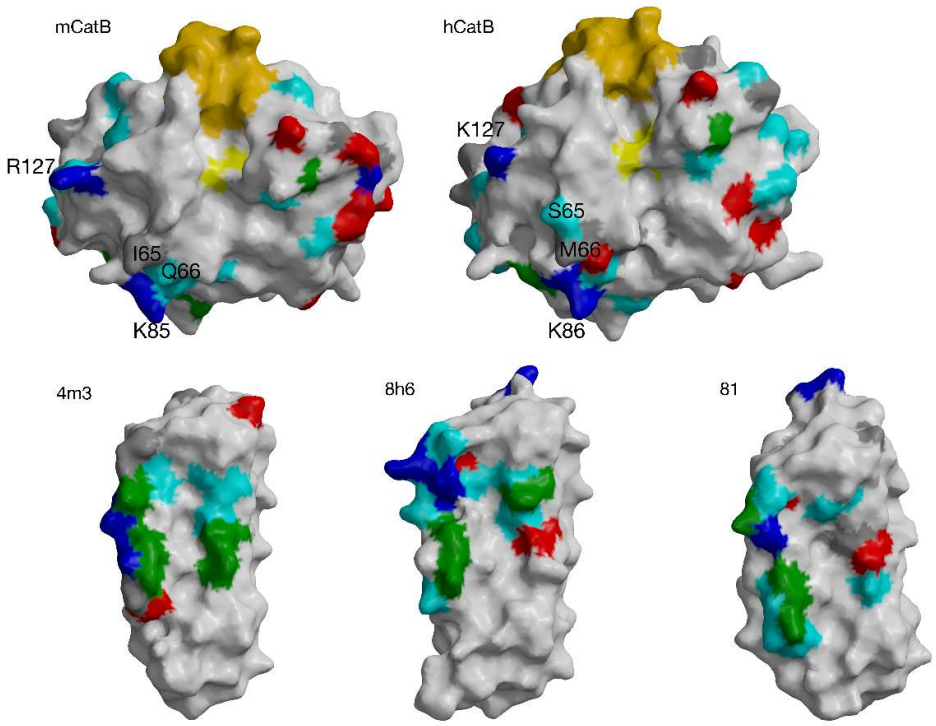


Figure 6.2: Structure-based alignments of mCatB and hCatB and DARPins with MAIN, with coloured differing residues according to chemical properties. The colours of the different residues and homologs represent their chemical properties, namely the hydrophobic A, G, I, V, L, P, M, and C (dark grey), the hydrophilic S, T, N, Q (cyan), the negatively charged D, E (red), the positively charged K, R (blue) and the aromatic hydrophobic Y, F, W, H (green). The CSO29C is colored yellow, and the amino acid residues of the occlusion loop from C108 to C119 are colored orange. Residues involved in the interaction between cathepsin B and DARPins are indicated by a one-letter code and a number. The protein structures were generated with MAIN and rendered using Raster3D.

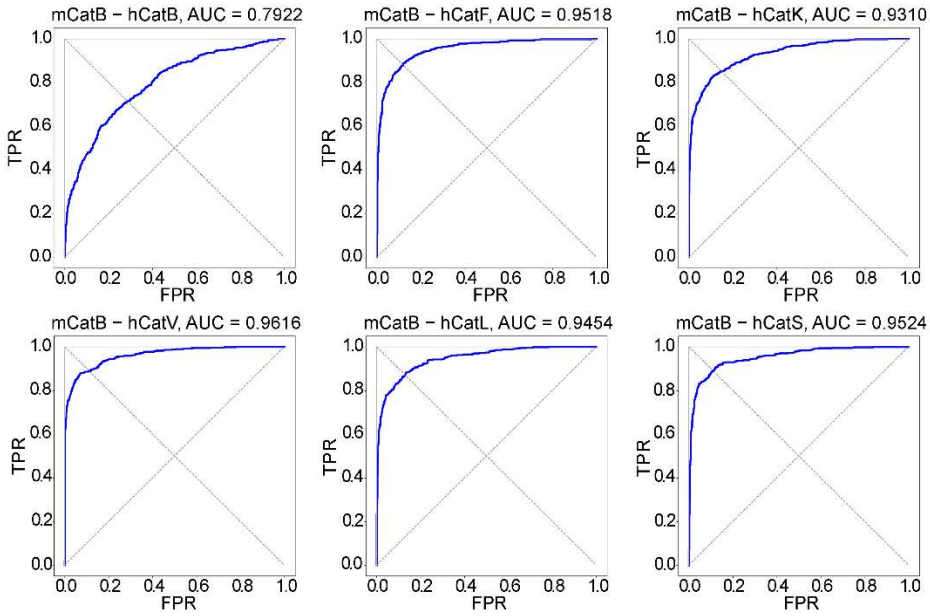


Figure 6.3: ROC curves and calculated AUC values for the positions from P4 to P4′ for combinations of substrates of mouse cathepsin B (mCatB) with separate substrates of human cathepsin B (hCatB), human cathepsin F (hCatF), human cathepsin K (hCatK), human cathepsin L (hCatL), human cathepsin V (hCatV), and human cathepsin S (hCatS) (see also Table 6.1).

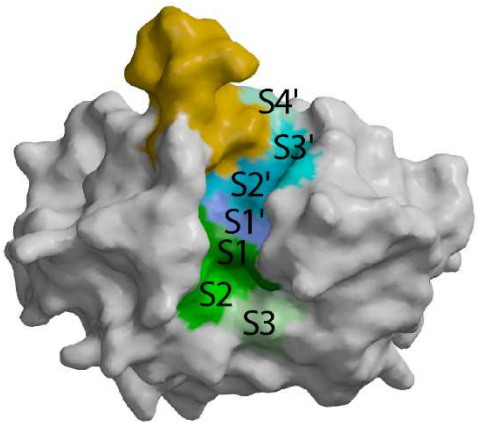


Figure 6.4: The positions S3–S4′ for the binding of substrates are shown in mCatB: prime sites S1′–S4′ (blue shading) and non-prime sites S1–S3 (green shading). Residues at the sites: S1′ 23, 176, 199; S2′ 24, 33, 221; S3′ 180, 181, 225, 231; S4′: 22, 224; S1 26, 27, 29, 71, 198; S2 30, 73, 74, 200, 245; S3 76, 173, 245. Positions were determined based on inhibitors bound to human cathepsin B.

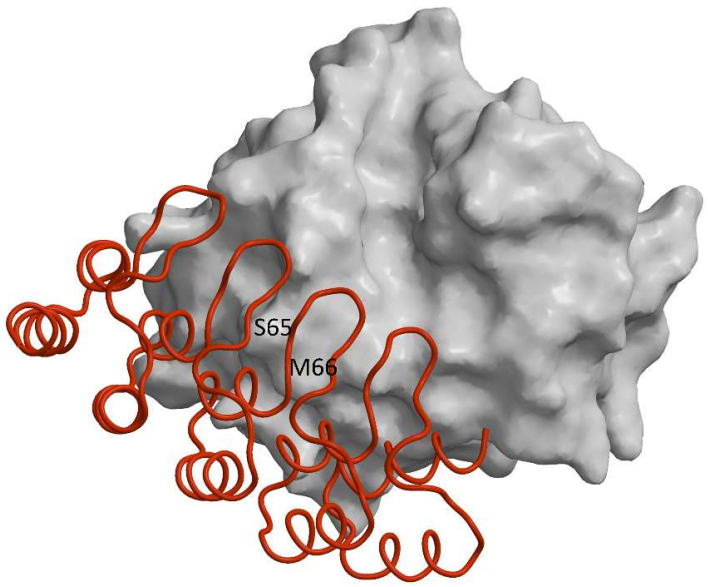


Figure 6.5: Docked complex of mutated mCatB (I65/Q66→S65/M66 substitution) and DARPin 4m3. The electrostatic model with code 002_00 (Table 6.5, red letters) is shown. Mutated mCatB is displayed as a white surface. DARPin 4m3 is shown as a red ribbon. The positions of the mutated residues in mCatB are labelled S65 and M66.

A.2 Appendix Tables

Table 6.1: Comparison of the combinations of pairs of data sets of the substrates of human cysteine cathepsins K, V, B, L, S, F and mouse cathepsin B by calculation of areas under the curve (AUC) of the ROC curve after using the Support Vector Machine (SVM) algorithm, where lower AUC denotes higher similarity.

No	**Cathepsin 1**	**Cathepsin 2**	**P3–P4′ (AUC)**	**P4–P4′ (AUC)**	**P15–P15′ (AUC)**
1	mCatB	hCatB	0.7331	0.7922	0.9668
2	mCatB	hCatK	0.9386	0.9310	0.9722
3	mCatB	hCatV	0.9524	0.9616	0.9743
4	mCatB	hCatL	0.9542	0.9454	0.9754
5	mCatB	hCatS	0.9455	0.9524	0.9788
6	mCatB	hCatF	0.9653	0.9518	0.9745
7	hCatF	hCatB	0.9274	0.9360	0.9384
8	hCatK	hCatB	0.8759	0.9270	0.9042
9	hCatV	hCatB	0.9348	0.8291	0.9573
10	hCatL	hCatB	0.9069	0.8980	0.8884
11	hCatS	hCatB	0.9001	0.9192	0.9105
12	hCatF	hCatK	0.7404	0.7324	0.7744
13	hCatF	hCatL	0.6794	0.7406	0.7423
14	hCatK	hCatL	0.6735	0.7338	0.7325
15	hCatK	hCatV	0.7434	0.7759	0.9267
16	hCatK	hCatS	0.6667	0.7920	0.8163
17	hCatF	hCatS	0.7839	0.8788	0.8655
18	hCatF	hCatV	0.5470	0.6393	0.8667
19	hCatL	hCatV	0.6782	0.6557	0.6694
20	hCatS	hCatL	0.7487	0.7931	0.8082
21	hCatS	hCatV	0.7985	0.8387	0.8371

Table 6.2: Separate clusters (336, 339, 389, 396) with only mCatB substrates. The columns are UniProt code, sequential number of clusters among 400 clusters, positions from P3 to P4′, cleavage position in the protein and the name of the protein.

UniProt	**Cluster**	**P3**	**P2**	**P1**	**P1′**	**P2′**	**P3′**	**P4′**	**Cleavage (P1 Position)**	**Protein Name**
Q3UWM4	336	K	K	G	M	A	T	A	916	Lysine-specific demethylase 7
Q80TJ7	336	K	K	G	L	A	T	A	1001	Histone lysine demethylase PHF8
Q9WTU0	336	K	K	G	M	A	T	A	1074	Lysine-specific demethylase PHF2
Q9R1Q8	339	Q	A	G	M	T	G	Y	188	Transgelin-3
Q9WVA4	339	Q	A	G	M	T	G	Y	188	Transgelin-2
P63325	389	K	K	A	E	A	G	A	140	40S ribosomal protein S10
Q61464	389	K	F	T	Q	A	G	A	395	Zinc finger protein 638
P39749	396	K	T	G	G	A	G	K	369	Flap endonuclease 1
P97386	396	K	R	G	T	A	G	C	102	DNA ligase 3

Table 6.3: Cluster scores (interaction energies between the two molecules) generated by the ClusPro serverare summarised for the docking runs performed with mutated mCatB (mut mCatB) and DARPin 4m3 (4m3) carrying the I65/Q66→S65/M66 substitution. Two docking jobs were used: one with (ID 1357923) and one without (ID 1358636) a specified interaction region. These results were validated against dockings of the wild-type mCatB–DARPin 4m3 complex performed under the same conditions, with (ID 1357926) and without (ID 1358637) the specified interaction region. For each docking, the coefficient weights for the interaction energy terms between the two proteins are listed. Here, Erep and Eattr represent the repulsive and attractive components of the van der Waals energy, respectively; Eelec denotes the electrostatic energy; and EDARS primarily reflects desolvation contributions. The parameter set "600Eelec" corresponds to the balanced weighting scheme, which has been shown to provide reliable performance for enzyme–inhibitor complexes (Kozakov *et al.*, 2017).

		ID 1357923		ID 1357926		ID 1358636		ID 1358637	
		mut mCatB+4m3		mCatB+4m3		mut mCatB+4m3		mCatB+4m3	
		Specified interaction region		Specified interaction region					
	Coefficient Weights*	E=0.40Erep+−0.40Eatt+600Eelec+1.00EDARS		E=0.40Erep+−0.10Eatt+600Eelec+0.00EDARS		E=0.40Erep+−0.40Eatt+600Eelec+1.00EDARS		E=0.40Erep+−0.40Eatt600Eelec+1.00EDARS	
Representative	Cluster	Weighted score	Number of cluster members	Weighted score	Number of cluster members	Weighted score	Number of cluster members	Weighted score	Number o cluster members
Center	0	-1927.8	227	-407.7	249	-603.3	89	-603.5	8
Lowest Energy		-1991.6		-569.9		-633.8		-633.8	
Center	1	-1733.2	96	-321.1	107	-736.4	68	-715.5	8
Lowest Energy		-1797.4		-362.6		-840.9		-762.6	
Center	2	-1412.9	83	-318.4	101	-690.2	66	-736.4	6
Lowest Energy		-1628.9		-496.6		-690.2		-840.9	
Center	3	-1444.7	64	-334.2	85	-687.0	59	-690.3	6
Lowest Energy		-1570.3		-447.2		-687.0		-690.3	
Center	4	-1508.3	62	-359.4	77	-770.8	52	-686.8	5
Lowest Energy		-1887.9		-404.4		-770.8		-686.8	
Center	5	-1528.5	47	-365.4	62	-633.6	49	-762.3	5
Lowest Energy		-1621.3		-365.4		-762.3		-762.3	
Center	6	-1422.7	44	-312.8	43	-661.5	47	-770.8	4
Lowest Energy		-1560.4		-383.4		-758.2		-770.8	
Center	7	-1467.2	33	-333.9	39	-693.6	43	-693.6	4
Lowest Energy		-1650.4		-350.1		-708.4		-708.4	
Center	8	-1497.1	33	-308.6	34	-615.5	43	-615.8	3
Lowest Energy		-1653.3		-409.5		-615.5		-615.8	
Center	9	-1446.4	31	-313.0	33	-628.5	41	-628.0	3
Lowest Energy		-1624.5		-384.5		-628.5		-628.0	
Center	10	-1415.3	30	-315.3	32	-580.5	39	-581.0	3
Lowest Energy		-1665.5		-397.4		-635.7		-636.1	
Center	11	-1531.8	29	-311.3	29	-605.6	36	-605.6	3
Lowest Energy		-1645.1		-376.4		-697.9		-697.8	
Center	12	-1464.0	27	-336.5	25	-590.7	34	-590.7	3
Lowest Energy		-1673.4		-340.1		-692.7		-692.7	
Center	13	-1559.0	23	-323.1	21	-613.5	30	-613.8	3
Lowest Energy		-1646.4		-344.3		-675.2		-675.6	
Center	14	-1472.0	23	-330.9	18	-597.8	26	-576.4	2
Lowest Energy		-1625.1		-341.3		-680.2		-680.2	
Center	15	-1446.2	22	-330.0	14	-614.0	25	-614.3	2
Lowest Energy		-1446.2		-342.3		-630.8		-630.9	
Center	16	-1410.8	21	-321.3	13	-631.9	22	-592.0	2
Lowest Energy		-1513.6		-343.8		-647.9		-647.8	
Center	17	-1404.4	19	-327.5	12	-600.5	21	-600.8	2
Lowest Energy		-1678.2		-343.8		-636.0		-636.2	
Center	18	-1448.3	17			-571.1	18	-571.4	1
Lowest Energy		-1559.4				-679.8		-679.9	
Center	19	-1480.9	16			-642.9	15	-649.2	1
Lowest Energy		-1574.2				-642.9		-649.2	
Center	20	-1415.9	13			-602.4	14	-642.6	1
Lowest Energy		-1525.9				-668.9		-642.6	
Center	21	-1550.4	7			-596.3	14	-596.2	1
Lowest Energy		-1590.9				-634.0		-633.9	
Center	22	-1475.9	6			-579.3	13	-602.7	1
Lowest Energy		-1475.9				-607.5		-669.1	
Center	23	-1432.7	5			-655.1	13	-579.3	1
Lowest Energy		-1645.6				-655.1		-607.5	
Center	24	-1435.2	5			-626.3	13	-577.4	1
Lowest Energy		-1435.5				-626.3		-613.6	
Center	25					-577.4	12	-655.2	1
Lowest Energy						-613.5		-655.2	
Center	26					-598.9	10	-625.8	1
Lowest Energy						-708.5		-625.8	
Center	27					-581.8	9	-581.8	
Lowest Energy						-604.5		-604.5	
Center	28					-568.6	5		
Lowest Energy						-647.8			

Table 6.4: Calculated hydrogen bonds between mCatB and DARPin 4m3 are shown for the four top-ranked docking complexes generated by the ClusPro server (docking ID 1357926) and are compared with the hydrogen bonds identified in the experimentally determined complex (PDB 9S60). These docking models were used to validate the mutated mCatB–DARPin 4m3 complex (I65/Q66→S65/M66) presented in Table 6.5. ClusPro provides four sets of docking models based on different scoring schemes: balanced (000_00), electrostatic-favoured (002_00), hydrophobic-favoured (004_00), and van der Waals plus electrostatics (006_00). The docking complexes most similar to our complex (PDB code 9S60), based on hydrogen-bond patterns, are highlighted in cyan, magenta, red, and green. If a docking model lacked a hydrogen bond corresponding to one present in our complex, the respective cell was shaded with the appropriate solid colour.

mouse cathepsin B				DARPin 4m3					000_00	006_00	002_00	004_00										000_00	006_00	002_00	004_00
Seq.1	Atom	Res.	Seq.2	Seq.3	Atom	Atom	Seq.4	9S60	Bond length (Å)	Bond length (Å)	Bond length (Å)	Bond length (Å)	Seq.1	Atom	Res.	Seq.2	Seq.3	Atom	Atom	Seq.4	9S60	Bond length (Å)	Bond length (Å)	Bond length (Å)	Bond length (Å)
896	O	CYS	62	4330	NE1	TRP	45						1219	NZ	LYS	86	5768	OD1	ASN	145					
906	O	CYS	63	3877	OG	SER	12						1219	NZ	LYS	86	5789	OD1	ASN	147					
906	O	CYS	63	4837	OH	TYR	79	2.77	2.6702	2.7861	2.6399	2.638	1723	OE1	GLU	122	5132	NZ	LYS	100					2.61
907	N	GLY	64	4330	NE1	TRP	45						1724	OE2	GLU	122	5132	NZ	LYS	100					2.778
914	N	ILE	65	4816	OG	SER	78						1724	OE2	GLU	122	5155	OH	TYR	101				2.9329	2.8396
914	N	ILE	65	4837	OH	TYR	79	3.22	2.8232	3.4548	2.8894	3.0181	1742	OD1	ASP	124	5769	ND2	ASN	145					
932	O	ILE	65	4219	NH2	ARG	37	2.98	2.8218	3.1575	3.2417	2.9154	1743	OD2	ASP	124	5769	ND2	ASN	145					
933	N	GLN	66	4837	OH	TYR	79	2.91	2.871	3.2043	2.7597	2.7713	1743	OD2	ASP	124	5155	OH	TYR	101					
944	OE1	GLN	66	3772	NZ	LYS	5						1759	O	THR	125	5769	ND2	ASN	145					
944	OE1	GLN	66	4216	NH1	ARG	37	3.34	2.6802	2.7271	2.6809	2.66	1746	N	THR	125	5155	OH	TYR	101					
944	OE1	GLN	66	4701	OH	TYR	70						1752	OG1	THR	125	5155	OH	TYR	101					
944	OE1	GLN	66	4788	O	LEU	75						1787	NE	ARG	127	5768	OD1	ASN	145					
944	OE1	GLN	66	4816	OG	SER	78						1787	NE	ARG	127	4819	O	SER	78					
945	NE2	GLN	66	3777	O	LYS	5						1787	NE	ARG	127	5610	O	ALA	134					
945	NE2	GLN	66	3838	N	ASP	9						1790	NH1	ARG	127	4840	O	TYR	79					
945	NE2	GLN	66	3847	OD2	ASP	9						1790	NH1	ARG	127	5327	NE2	HIS	114					
945	NE2	GLN	66	4330	NE1	TRP	45	3.54				2.969	1790	NH1	ARG	127	5756	OD1	ASP	144		3.4703	2.7657		
945	NE2	GLN	66	4837	OH	TYR	79						1790	NH1	ARG	127	5759	O	ASP	144	3.38	2.9402		2.8585	2.8483
949	O	GLN	66	4216	NH1	ARG	37	3	2.6756	2.6955	2.9473	2.8724	1790	NH1	ARG	127	5768	OD1	ASN	145				2.7478	2.6823
949	O	GLN	66	4219	NH2	ARG	37		3.3582	3.0714	3.1282	3.2709	1790	NH1	ARG	127	5773	O	ASN	145					
966	O	GLY	68	4216	NH1	ARG	37						1790	NH1	ARG	127	5780	O	GLY	146					
966	O	GLY	68	4701	OH	TYR	70						1793	NH2	ARG	127	5756	OD1	ASP	144		2.8136			
975	OD1	ASP	69	4216	NH1	ARG	37						1793	NH2	ARG	127	5757	OD2	ASP	144		2.8135			
975	OD1	ASP	69	4219	NH2	ARG	37						1793	NH2	ARG	127	5768	OD1	ASN	145					
975	OD1	ASP	69	4676	O	TYR	68		3.3878				1793	NH2	ARG	127	5773	O	ASN	145					
975	OD1	ASP	69	4701	OH	TYR	70				2.9	2.7721	1793	NH2	ARG	127	5780	O	GLY	146					
975	OD1	ASP	69	5112	OG	SER	99					3.2379	1793	NH2	ARG	127	5789	OD1	ASN	147					
975	OD1	ASP	69	5116	N	LYS	100						1793	NH2	ARG	127	5756	OD1	ASP	144			2.7293		
975	OD1	ASP	69	5132	NZ	LYS	100		2.7435	2.7872			1793	NH2	ARG	127	4819	O	SER	78					
975	OD1	ASP	69	5175	ND2	ASN	103				3.061	3.0298	1793	NH2	ARG	127	5305	O	SER	112					
976	OD2	ASP	69	4183	ND1	HIS	35						1793	NH2	ARG	127	5327	NE2	HIS	114					
976	OD2	ASP	69	4216	NH1	ARG	37						1793	NH2	ARG	127	5757	OD2	ASP	144				2.6142	
976	OD2	ASP	69	4650	NZ	LYS	67						1793	NH2	ARG	127	5759	O	ASP	144	3.32			2.8846	
976	OD2	ASP	69	4655	O	LYS	67						1793	NH2	ARG	127	5610	O	ALA	134					
976	OD2	ASP	69	4673	OH	TYR	68						1793	NH2	ARG	127	5617	O	GLY	135					
976	OD2	ASP	69	4701	OH	TYR	70				3.0169		1793	NH2	ARG	127	5759	O	ASP	144					2.9702
976	OD2	ASP	69	5112	OG	SER	99				2.8871	2.8617	1838	NZ	LYS	130	5756	OD1	ASP	144					
976	OD2	ASP	69	5116	N	LYS	100				3.4947		1838	NZ	LYS	130	5757	OD2	ASP	144					
976	OD2	ASP	69	5132	NZ	LYS	100		2.9037	2.8426			1838	NZ	LYS	130	4353	O	SER	46					
976	OD2	ASP	69	5138	N	TYR	101				3.0628	3.1699	1838	NZ	LYS	130	5759	O	ASP	144					
976	OD2	ASP	69	5175	ND2	ASN	103						1838	NZ	LYS	130	5768	OD1	ASN	145					
1005	OD1	ASN	72	5155	OH	TYR	101		2.7815	2.5237			1991	NZ	LYS	141	5759	O	ASP	144					
1005	OD1	ASN	72	5132	NZ	LYS	100						1991	NZ	LYS	141	5773	O	ASN	145					
1006	ND2	ASN	72	5155	OH	TYR	101						1991	NZ	LYS	141	5780	O	GLY	146					
1010	O	ASN	72	4650	NZ	LYS	67						2136	O	SER	150	3772	NZ	LYS	5					
1010	O	ASN	72	5132	NZ	LYS	100					3.1431	2158	N	SER	152	3846	OD1	ASP	9		2.8916	2.8338	2.9605	2.9852
1018	N	GLY	74	5155	OH	TYR	101						2158	N	SER	152	3847	OD2	ASP	9	3.05	3.0655		2.9639	3.1333
1018	N	GLY	74	4673	OH	TYR	68						2165	OG	SER	152	3846	OD1	ASP	9		2.8168	2.7938	2.9298	2.9162
1024	O	GLY	74	4183	ND1	HIS	35		2.6731	2.754			2165	OG	SER	152	3847	OD2	ASP	9				3.0765	
1024	O	GLY	74	4673	OH	TYR	68	3.14		2.6048	2.6226	2.6693	2165	OG	SER	152	3877	OG	SER	12					
1042	OH	TYR	75	4162	OD2	ASP	33	2.58	2.9256	2.8356	2.8501	2.8709	2165	OG	SER	152	3880	O	SER	12					
1042	OH	TYR	75	4192	O	HIS	35						2165	OG	SER	152	4162	OD2	ASP	33					
1042	OH	TYR	75	4330	NE1	TRP	45						2168	O	SER	152	3847	OD2	ASP	9			2.9045		
1042	OH	TYR	75	4630	OD1	ASP	66						2168	O	SER	152	3772	NZ	LYS	5					3.4027
1067	OG	SER	77	4162	OD2	ASP	33				3.0181		2184	O	VAL	153	3772	NZ	LYS	5	3.37		2.6812	2.9603	
1060	N	SER	77	4189	NE2	HIS	35					3.447	2185	N	SER	154	3772	NZ	LYS	5					3.3823
1067	OG	SER	77	4189	NE2	HIS	35					2.9973	2192	OG	SER	154	3794	NZ	LYS	6		2.688	2.561		
1077	O	GLY	78	4837	OH	TYR	79						2192	OG	SER	154	3910	NE2	GLN	15					
1119	OG	SER	81	3880	O	SER	12		3.427				2192	OG	SER	154	3772	NZ	LYS	5					2.8022
1119	OG	SER	81	4330	NE1	TRP	45	3.04		2.77			2204	OD1	ASN	155	3726	OD1	ASP	2					
1119	OG	SER	81	3847	OD2	ASP	9						2204	OD1	ASN	155	3794	NZ	LYS	6					
1119	OG	SER	81	3877	OG	SER	12	3.03			2.7744	2.7806	2205	ND2	ASN	155	3726	OD1	ASP	2					
1122	O	SER	81	4330	NE1	TRP	45		2.8007		3.1864		2205	ND2	ASN	155	3727	OD2	ASP	2					
1122	O	SER	81	4837	OH	TYR	79						2205	ND2	ASN	155	3847	OD2	ASP	9					
1123	N	PHE	82	4837	OH	TYR	79						2205	ND2	ASN	155	3727	OD2	ASP	2					
1197	NZ	LYS	85	3846	OD1	ASP	9						2205	ND2	ASN	155	3794	NZ	LYS	6					
1197	NZ	LYS	85	3849	O	ASP	9						2205	ND2	ASN	155	3847	OD2	ASP	9					
1197	NZ	LYS	85	3877	OG	SER	12						2209	O	ASN	155	3794	NZ	LYS	6					
1197	NZ	LYS	85	3880	O	SER	12	2.71	2.8243		2.8216	2.7342	3482	ND2	ASN	238	3726	OD1	ASP	2					
1197	NZ	LYS	85	3890	O	ALA	13	3.46	2.9464		2.8628	2.7085	3482	ND2	ASN	238	3727	OD2	ASP	2					
1197	NZ	LYS	85	3950	OE1	GLU	18						3482	ND2	ASN	238	3772	NZ	LYS	5					
1197	NZ	LYS	85	3951	OE2	GLU	18						3564	OG	SER	244	3772	NZ	LYS	5					
1197	NZ	LYS	85	4342	O	TRP	45			2.7365		2.7946	3567	O	SER	244	3772	NZ	LYS	5		2.6728	3.0517	3.4348	
1197	NZ	LYS	85	4353	O	SER	46		2.8315	3.0939	2.8378	2.959	3567	O	SER	244	3846	OD1	ASP	9					
1197	NZ	LYS	85	4840	O	TYR	79						3567	O	SER	244	3847	OD2	ASP	9					
1197	NZ	LYS	85	4862	NE2	HIS	81			2.9283			3579	OE1	GLU	245	3772	NZ	LYS	5		2.5268			
1197	NZ	LYS	85	5327	NE2	HIS	114						3579	OE1	GLU	245	4165	N	ALA	34		3.1376	2.7023	3.4453	
1219	NZ	LYS	86	3859	O	ALA	10						3579	OE1	GLU	245	4174	O	ALA	34				3.3716	
1219	NZ	LYS	86	3880	O	SER	12						3579	OE1	GLU	245	4175	N	HIS	35			3.0971	3.1293	
1219	NZ	LYS	86	3890	O	ALA	13						3579	OE1	GLU	245	4183	ND1	HIS	35					
1219	NZ	LYS	86	3897	O	GLY	14						3579	OE1	GLU	245	4189	NE2	HIS	35				3.4091	
1219	NZ	LYS	86	3909	OE1	GLN	15						3579	OE1	GLU	245	4650	NZ	LYS	67				2.807	2.5528
1219	NZ	LYS	86	4342	O	TRP	45						3580	OE2	GLU	245	4165	N	ALA	34		2.8597			
1219	NZ	LYS	86	4353	O	SER	46						3580	OE2	GLU	245	4175	N	HIS	35		2.762	3.3268		
1219	NZ	LYS	86	4840	O	TYR	79		2.9024				3580	OE2	GLU	245	4183	ND1	HIS	35					2.85
1219	NZ	LYS	86	5294	O	THR	111						3580	OE2	GLU	245	4189	NE2	HIS	35	2.9			3.3773	
1219	NZ	LYS	86	5305	O	SER	112						3580	OE2	GLU	245	4650	NZ	LYS	67				2.8116	3.1788
1219	NZ	LYS	86	5757	OD2	ASP	144						3580	OE2	GLU	245	4673	OH	TYR	68					

Table 6.5: Hydrogen bonds were calculated for the mutated mCatB–DARPin 4m3 complex (I65/Q66→S65/M66) using the two highest-ranked docked models generated by the ClusPro server (docking ID 1357923). ClusPro produces four scoring-based model sets; in this case, only two were used: Balanced (000_00) and Electrostatic-favoured (002_00). Among these, the 002_00 model (highlighted in red) shows the greatest similarity in hydrogen-bond pattern to our complex (PDB code 9S60) and is therefore also illustrated in Figure 6.5.

mutated mCatB				DARPin 4m3				000_00	002_00									000_00	002_00
Seq.1	Atom	Res.	Seq.2	Seq.3	Atom	Atom	Seq.4	Bond length (Å)	Bond length (Å)	Seq.1	Atom	Res.	Seq.2	Seq.3	Atom	Atom	Seq.4	Bond length (Å)	Bond length (Å)
886	O	THR	61	4322	NE1	TRP	45			1211	NZ	LYS	86	5761	ND2	ASN	145		
906	O	CYS	63	4829	OH	TYR	79	2.5494	2.5218	1211	NZ	LYS	86	5781	OD1	ASN	147		
906	O	CYS	63	3869	OG	SER	12			1211	NZ	LYS	86	3901	OE1	GLN	15		
914	N	SER	65	4829	OH	TYR	79	3.023	2.9128	1211	NZ	LYS	86	3942	OE1	GLU	18		
914	N	SER	65	5167	ND2	ASN	103			1211	NZ	LYS	86	3943	OE2	GLU	18		
921	OG	SER	65	4280	O	HIS	41			1716	OE2	GLU	122	5147	OH	TYR	101	2.8632	
921	OG	SER	65	4281	N	ALA	42			1715	OE1	GLU	122	5124	NZ	LYS	100		
921	OG	SER	65	4761	O	HIS	74	2.8814		1716	OE2	GLU	122	5124	NZ	LYS	100		
921	OG	SER	65	4808	OG	SER	78	3.4602	2.8349	1715	OE1	GLU	122	5147	OH	TYR	101		
921	OG	SER	65	4829	OH	TYR	79		3.4503	1735	OD2	ASP	124	5147	OH	TYR	101		
921	OG	SER	65	4854	NE2	HIS	81			1734	OD1	ASP	124	5761	ND2	ASN	145		
921	OG	SER	65	5166	OD1	ASN	103			1735	OD2	ASP	124	5761	ND2	ASN	145		
921	OG	SER	65	5167	ND2	ASN	103			1735	OD2	ASP	124	5279	OG1	THR	111		
924	O	SER	65	4211	NH2	ARG	37	3.335	2.7094	1735	OD2	ASP	124	4829	OH	TYR	79		
925	N	MET	66	4829	OH	TYR	79	2.6753	2.697	1738	N	THR	125	5147	OH	TYR	101		
941	O	MET	66	4208	NH1	ARG	37	2.8147	2.788	1744	OG1	THR	125	5147	OH	TYR	101		
941	O	MET	66	4211	NH2	ARG	37	3.2862		1738	N	THR	125	4829	OH	TYR	79		
941	O	MET	66	4829	OH	TYR	79			1751	O	THR	125	4829	OH	TYR	79		
952	N	GLY	68	4211	NH2	ARG	37			1779	NE	ARG	127	5602	O	ALA	134		
958	O	GLY	68	4693	OH	TYR	70			1779	NE	ARG	127	5765	O	ASN	145		
958	O	GLY	68	4208	NH1	ARG	37			1782	NH1	ARG	127	5751	O	ASP	144	2.8501	2.8624
967	OD1	ASP	69	4208	NH1	ARG	37			1782	NH1	ARG	127	5760	OD1	ASN	145	2.7165	2.7153
967	OD1	ASP	69	4211	NH2	ARG	37			1782	NH1	ARG	127	5297	O	SER	112		
967	OD1	ASP	69	4693	OH	TYR	70	3.0486		1782	NH1	ARG	127	4832	O	TYR	79		
967	OD1	ASP	69	4808	OG	SER	78			1785	NH2	ARG	127	5748	OD1	ASP	144	2.6152	2.9704
967	OD1	ASP	69	5104	OG	SER	99			1785	NH2	ARG	127	5749	OD2	ASP	144	2.6152	2.6974
967	OD1	ASP	69	5167	ND2	ASN	103	3.098		1785	NH2	ARG	127	5751	O	ASP	144	2.8891	2.9265
967	OD1	ASP	69	5124	NZ	LYS	100		2.8428	1785	NH2	ARG	127	5602	O	ALA	134		
968	OD2	ASP	69	5124	NZ	LYS	100		3.1307	1785	NH2	ARG	127	5609	O	GLY	135		
968	OD2	ASP	69	4208	NH1	ARG	37			1785	NH2	ARG	127	5286	O	THR	111		
968	OD2	ASP	69	4211	NH2	ARG	37			1785	NH2	ARG	127	5765	O	ASN	145		
968	OD2	ASP	69	4693	OH	TYR	70	3.0127		1785	NH2	ARG	127	5781	OD1	ASN	147		
968	OD2	ASP	69	4808	OG	SER	78			1785	NH2	ARG	127	5319	NE2	HIS	114		
968	OD2	ASP	69	5104	OG	SER	99	2.8854		1830	NZ	LYS	130	5751	O	ASP	144		
968	OD2	ASP	69	5108	N	LYS	100			1830	NZ	LYS	130	5760	OD1	ASN	145		
968	OD2	ASP	69	5130	N	TYR	101	3.019		1830	NZ	LYS	130	3872	O	SER	12		
997	OD1	ASN	72	5124	NZ	LYS	100			1830	NZ	LYS	130	4345	O	SER	46		
997	OD1	ASN	72	5147	OH	TYR	101		2.8299	1983	NZ	LYS	141	5765	O	ASN	145		
998	ND2	ASN	72	5147	OH	TYR	101			1983	NZ	LYS	141	5772	O	GLY	146		
997	OD1	ASN	72	5167	ND2	ASN	103			1983	NZ	LYS	141	3882	O	ALA	13		
1002	O	ASN	72	4693	OH	TYR	70			1983	NZ	LYS	141	3889	O	GLY	14		
1009	O	GLY	73	4665	OH	TYR	68			2125	OG	SER	150	3764	NZ	LYS	5		
1016	O	GLY	74	4665	OH	TYR	68	2.6996	2.5513	2157	OG	SER	152	3764	NZ	LYS	5		
1016	O	GLY	74	4175	ND1	HIS	68		3.0273	2150	N	SER	152	3838	OD1	ASP	9	2.9512	2.8781
1010	N	GLY	74	4665	OH	TYR	68		3.4597	2157	OG	SER	152	3838	OD1	ASP	9	2.9279	2.8744
1034	OH	TYR	75	4154	OD2	ASP	33	2.8301	2.8683	2150	N	SER	152	3839	OD2	ASP	9	2.9574	2.8958
1034	OH	TYR	75	4153	OD1	ASP	33			2157	OG	SER	152	3839	OD2	ASP	9	3.1087	3.1892
1059	OG	SER	77	4154	OD2	ASP	33	3.0075		2157	OG	SER	152	3869	OG	SER	12		
1059	OG	SER	77	4153	OD1	ASP	33			2157	OG	SER	152	3872	O	SER	12		
1059	OG	SER	77	4175	ND1	HIS	35			2157	OG	SER	152	4154	OD2	ASP	33		
1069	O	GLY	78	4829	OH	TYR	79			2176	O	VAL	153	3764	NZ	LYS	5	2.9539	
1111	OG	SER	81	3869	OG	SER	12	2.8087	2.7825	2184	OG	SER	154	3902	NE2	GLN	15		
1114	O	SER	81	4322	NE1	TRP	45	2.9881	2.8261	2184	OG	SER	154	3786	NZ	LYS	6		2.6395
1111	OG	SER	81	3838	OD1	ASP	9			2197	ND2	ASN	155	3839	OD2	ASP	9		
1111	OG	SER	81	3764	NZ	LYS	5			2847	O	GLY	198	5124	NZ	LYS	100		
1111	OG	SER	81	3838	OD1	ASP	9			3559	O	SER	244	3764	NZ	LYS	5	3.426	
1111	OG	SER	81	3839	OD2	ASP	9			3556	OG	SER	244	3764	NZ	LYS	5		
1114	O	SER	81	3838	OD1	ASP	9			3559	O	SER	244	3764	NZ	LYS	5		2.7417
1115	N	PHE	82	3838	OD1	ASP	9			3559	O	SER	244	3838	OD1	ASP	9		
1189	NZ	LYS	85	3718	OD1	ASP	2			3571	OE1	GLU	245	4157	N	ALA	34	3.4619	2.876
1189	NZ	LYS	85	3719	OD2	ASP	2			3572	OE2	GLU	245	4157	N	ALA	34		
1189	NZ	LYS	85	3838	OD1	ASP	9			3571	OE1	GLU	245	4166	O	ALA	34	3.364	
1189	NZ	LYS	85	3839	OD2	ASP	9			3571	OE1	GLU	245	4167	N	HIS	35	3.0915	
1189	NZ	LYS	85	3841	O	ASP	9			3571	OE1	GLU	245	4181	NE2	HIS	35	3.4298	
1189	NZ	LYS	85	3872	O	SER	12	2.8225	2.7958	3571	OE1	GLU	245	4642	NZ	LYS	67	2.8035	
1189	NZ	LYS	85	3882	O	ALA	13	2.887	2.9848	3572	OE2	GLU	245	4181	NE2	HIS	35	3.3363	
1189	NZ	LYS	85	4334	O	TRP	45	3.2962	2.7322	3572	OE2	GLU	245	4642	NZ	LYS	67	2.7912	
1189	NZ	LYS	85	4345	O	SER	46	2.8047	2.7191	3571	OE1	GLU	245	3764	NZ	LYS	5		2.7712
1189	NZ	LYS	85	4832	O	TYR	79			3572	OE2	GLU	245	3764	NZ	LYS	5		2.8691
1211	NZ	LYS	86	4832	O	TYR	79		2.9751	3571	OE1	GLU	245	4642	NZ	LYS	67		
1211	NZ	LYS	86	5286	O	THR	111			3572	OE2	GLU	245	4181	NE2	HIS	35		
1211	NZ	LYS	86	5297	O	SER	112			3572	OE2	GLU	245	4642	NZ	LYS	67		
										3572	OE2	GLU	245	4665	OH	TYR	68		

References

Aghdassi, A. A., John, D. S., Sendler, M., Weiss, F. U., Reinheckel, T., Mayerle, J., & Lerch, M. M. (2018). Cathepsin D regulates cathepsin B activation and disease severity predominantly in inflammatory cells during experimental pancreatitis. *Journal of Biological Chemistry*, *293*(3), 1018–1029. https://doi.org/10.1074/jbc.M117.814772

Aguda, A. H., Panwar, P., Du, X., Nguyen, N. T., Brayer, G. D., & Brömme, D. (2014). Structural basis of collagen fiber degradation by cathepsin K. *Proceedings of the National Academy of Sciences*, *111*(49), 17474–17479. https://doi.org/10.1073/pnas.1414126111

Authier, F., Kouach, M., & Briand, G. (2005). Endosomal proteolysis of insulin-like growth factor-I at its C-terminal D-domain by cathepsin B. *FEBS Letters*, *579*(20), 4309–4316. https://doi.org/10.1016/j.febslet.2005.06.066

Baici, A. (1981). The specific velocity plot. A graphical method for determining inhibition parameters for both linear and hyperbolic enzyme inhibitors. *European Journal of Biochemistry*, *119*(1), 9–14. https://doi.org/10.1111/j.1432-1033.1981.tb05570.x

Baird, R. D., Linossi, C., Middleton, M., Lord, S., Harris, A., Rodón, J., … Omlin, A. (2021). First-in-Human phase I study of MP0250, a first-in-class DARPin drug candidate targeting VEGF and HGF, in patients with advanced solid tumors. *Journal of Clinical Oncology*, *39*(2), 145–154. https://doi.org/10.1200/JCO.20.00596

Ben-Aderet, L., Merquiol, E., Fahham, D., Kumar, A., Reich, E., Ben-Nun, Y., … Dvir-Ginzberg, M. (2015). Detecting cathepsin activity in human osteoarthritis via activity-based probes. *Arthritis Research & Therapy*, *17*(1), 69. https://doi.org/10.1186/s13075-015-0586-5

Bestvater, F., Dallner, C., & Spiess, E. (2005). The C-terminal subunit of artificially truncated human cathepsin B mediates its nuclear targeting and contributes to cell viability. *BMC Cell Biology*, *6*(1), 16. https://doi.org/10.1186/1471-2121-6-16

Bevec, T., Stoka, V., Pungerčič, G., Dolenc, I., & Turk, V. (1996). Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L. *The Journal of Experimental Medicine*, *183*(4), 1331–1338. https://doi.org/10.1084/jem.183.4.1331

Bian, B., Mongrain, S., Cagnol, S., Langlois, M.-J., Boulanger, J., Bernatchez, G., … Rivard, N. (2016). Cathepsin B promotes colorectal tumorigenesis, cell invasion, and metastasis. *Molecular Carcinogenesis*, *55*(5), 671–687. https://doi.org/10.1002/mc.22312

Bianchi, M., Reichen, C., Croset, A., Fischer, S., Eggenschwiler, A., Grübler, Y., …

Goubier, A. (2024). The CD33xCD123xCD70 multispecific CD3-engaging DARPin MP0533 induces selective T Cell–mediated killing of AML leukemic stem cells. *Cancer Immunology Research*, *12*(7), 921–943. https://doi.org/10.1158/2326-6066.CIR-23-0692

Biasizzo, M., Javoršek, U., Vidak, E., Zarić, M., & Turk, B. (2022). Cysteine cathepsins: a long and winding road towards clinics. *Molecular Aspects of Medicine*, *88*(October). https://doi.org/10.1016/j.mam.2022.101150

Binz, H. K., Stumpp, M. T., Forrer, P., Amstutz, P., & Plückthun, A. (2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *Journal of Molecular Biology*, *332*(2), 489–503. https://doi.org/10.1016/S0022-2836(03)00896-9

Bollavaram, K., Leeman, T. H., Lee, M. W., Kulkarni, A., Upshaw, S. G., Yang, J., … Platt, M. O. (2021). Multiple sites on SARS-CoV-2 spike protein are susceptible to proteolysis by cathepsins B, K, L, S, and V. *Protein Science*, *30*(6), 1131–1143. https://doi.org/10.1002/pro.4073

Bossard, M. J., Tomaszek, T. A., Thompson, S. K., Amegadzie, B. Y., Hanning, C. R., Jones, C., … Levy, M. A. (1996). Proteolytic activity of human osteoclast cathepsin K. *Journal of Biological Chemistry*, *271*(21), 12517–12524. https://doi.org/10.1074/jbc.271.21.12517

Bromme, D., Nallaseth, F. S., & Turk, B. (2004). Production and activation of recombinant papain-like cysteine proteases. *Methods*, *32*(2), 199–206. https://doi.org/10.1016/S1046-2023(03)00212-3

Büth, H., Luigi Buttigieg, P., Ostafe, R., Rehders, M., Dannenmann, S. R., Schaschke, N., … Brix, K. (2007). Cathepsin B is essential for regeneration of scratch-wounded normal human epidermal keratinocytes. *European Journal of Cell Biology*, *86*(11–12), 747–761. https://doi.org/10.1016/j.ejcb.2007.03.009

Caglič, D., Kosec, G., Bojič, L., Reinheckel, T., Turk, V., & Turk, B. (2009). Murine and human cathepsin B exhibit similar properties: possible implications for drug discovery. *Biological Chemistry*, *390*(2), 175–179. https://doi.org/10.1515/BC.2009.021

Cai, Z., Xu, S., & Liu, C. (2024). Cathepsin B in cardiovascular disease: Underlying mechanisms and therapeutic strategies. *Journal of Cellular and Molecular Medicine*, *28*(17). https://doi.org/10.1111/jcmm.70064

Chan, A. T., Baba, Y., Shima, K., Nosho, K., Chung, D. C., Hung, K. E., … Ogino, S. (2010). Cathepsin B expression and survival in colon cancer: implications for molecular detection of neoplasia. *Cancer Epidemiology, Biomarkers & Prevention*, *19*(11), 2777–2785. https://doi.org/10.1158/1055-9965.EPI-10-0529

Chang, S.-H., Kanasaki, K., Gocheva, V., Blum, G., Harper, J., Moses, M. A., … Dvorak, H. F. (2009). VEGF-A induces angiogenesis by perturbing the cathepsin cysteine protease inhibitor balance in venules, causing basement membrane degradation and mother vessel formation. *Cancer Research*, *69*(10), 4537–4544. https://doi.org/10.1158/0008-5472.CAN-08-4539

Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J. S., & Cygler, M. (1996). Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *The EMBO Journal*, *15*(20), 5492–5503. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/8896443

Desmarais, S., Massé, F., & Percival, M. D. (2009). Pharmacological inhibitors to identify roles of cathepsin K in cell-based studies: A comparison of available tools. *Biological Chemistry*, *390*(9), 941–948. https://doi.org/10.1515/BC.2009.092

Dinnes, D. L. M., White, M. Y., Kockx, M., Traini, M., Hsieh, V., Kim, M., … Kritharides, L. (2016). Human macrophage cathepsin β-mediated C-terminal cleavage of apolipoprotein α-I at Ser 228 severely impairs antiatherogenic capacity. *The FASEB Journal*, *30*(12), 4239–4255. https://doi.org/10.1096/fj.201600508R

Droga-Mazovec, G., Bojič, L., Petelin, A., Ivanova, S., Romih, R., Repnik, U., … Turk, B. (2008). Cysteine cathepsins trigger caspase-dependent cell death through cleavage of bid and antiapoptotic Bcl-2 homologues. *Journal of Biological Chemistry*, *283*(27), 19140–19150. https://doi.org/10.1074/jbc.M802513200

Duarte, L. F., Young, A. R. J., Wang, Z., Wu, H.-A., Panda, T., Kou, Y., … Bernstein, E. (2014). Histone H3.3 and its proteolytically processed form drive a cellular senescence programme. *Nature Communications*, *5*(1), 5210. https://doi.org/10.1038/ncomms6210

Duncan, E. M., Muratore-Schroeder, T. L., Cook, R. G., Garcia, B. A., Shabanowitz, J., Hunt, D. F., & Allis, C. D. (2008). Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. *Cell*, *135*(2), 284–294. https://doi.org/10.1016/j.cell.2008.09.055

Ebert, M. P. A., Krüger, S., Fogeron, M., Lamer, S., Chen, J., Pross, M., … Röcken, C. (2005). Overexpression of cathepsin B in gastric cancer identified by proteome analysis. *Proteomics*, *5*(6), 1693–1704. https://doi.org/10.1002/pmic.200401030

Esser, R. E., Angelo, R. A., Murphey, M. D., Watts, L. M., Thornburg, L. P., Palmer, J. T., … Smith, R. E. (1994). Cysteine proteinase inhibitors decrease articular cartilage and bone destruction in chronic inflammatory arthritis. *Arthritis & Rheumatism*, *37*(2), 236–247. https://doi.org/10.1002/art.1780370213

Felbor, U., Kessler, B., Mothes, W., Goebel, H. H., Ploegh, H. L., Bronson, R. T., & Olsen, B. R. (2002). Neuronal loss and brain atrophy in mice lacking cathepsins B and L. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 7883–7888. https://doi.org/10.1073/pnas.112632299

Fonovič, M., Brömme, D., Turk, V., & Turk, B. (2004). Human cathepsin F: expression in baculovirus system, characterization and inhibition by protein inhibitors. *Biological Chemistry*, *385*(6). https://doi.org/10.1515/BC.2004.059

Friedrichs, B., Tepel, C., Reinheckel, T., Deussing, J., Von Figura, K., Herzog, V., … Brix, K. (2003). Thyroid functions of mouse cathepsins B, K, and L. *Journal of Clinical Investigation*, *111*(11), 1733–1745. https://doi.org/10.1172/JCI15990

Fujimoto, T., Tsunedomi, R., Matsukuma, S., Yoshimura, K., Oga, A., Fujiwara, N., … Nagano, H. (2020). Cathepsin B is highly expressed in pancreatic cancer stem-like cells and is associated with patients' surgical outcomes. *Oncology Letters*, *21*(1), 1–1. https://doi.org/10.3892/ol.2020.12291

Fujinaga, M., Cherney, M. M., Oyama, H., Oda, K., & James, M. N. G. (2004). The molecular structure and catalytic mechanism of a novel carboxyl peptidase from Scytalidium lignicolum. *Proceedings of the National Academy of Sciences*, *101*(10),

3364–3369. https://doi.org/10.1073/pnas.0400246101

Funkelstein, L., Lu, W. D., Koch, B., Mosier, C., Toneff, T., Taupenot, L., … Hook, V. (2012). Human cathepsin V protease participates in production of enkephalin and NPY neuropeptide neurotransmitters. *Journal of Biological Chemistry*, *287*(19), 15232–15241. https://doi.org/10.1074/jbc.M111.310607

Gnirß, K., Kühl, A., Karsten, C., Glowacka, I., Bertram, S., Kaup, F., … Pöhlmann, S. (2012). Cathepsins B and L activate Ebola but not Marburg virus glycoproteins for efficient entry into cell lines and macrophages independent of TMPRSS2 expression. *Virology*, *424*(1), 3–10. https://doi.org/10.1016/j.virol.2011.11.031

Gocheva, V., Wang, H. W., Gadea, B. B., Shree, T., Hunter, K. E., Garfall, A. L., … Joyce, J. A. (2010). IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion. *Genes and Development*, *24*(3), 241–255. https://doi.org/10.1101/gad.1874010

Gocheva, V., Zeng, W., Ke, D., Klimstra, D., Reinheckel, T., Peters, C., … Joyce, J. A. (2006). Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. *Genes & Development*, *20*(5), 543–556. https://doi.org/10.1101/gad.1407406

Gopinathan, A., DeNicola, G. M., Frese, K. K., Cook, N., Karreth, F. A., Mayerle, J., … Tuveson, D. A. (2012). Cathepsin B promotes the progression of pancreatic ductal adenocarcinoma in mice. *Gut*, *61*(6), 877–884. https://doi.org/10.1136/gutjnl-2011-300850

Goulet, B., Baruch, A., Moon, N.-S., Poirier, M., Sansregret, L. L., Erickson, A., … Nepveu, A. (2004). A cathepsin L isoform that Is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. *Molecular Cell*, *14*(2), 207–219. https://doi.org/10.1016/S1097-2765(04)00209-6

Greten, F. R., & Grivennikov, S. I. (2019). Inflammation and cancer: triggers, mechanisms, and consequences. *Immunity*, *51*(1), 27–41. https://doi.org/10.1016/j.immuni.2019.06.025

Guinec, N., Dalet-Fumeron, V., & Pagano, M. (1993). "In vitro" study of basement membrane degradation by the cysteine proteinases, cathepsins B, B-like and L. Digestion of collagen IV, laminin, fibronectin, and release of gelatinase activities from basement membrane fibronectin. *Biological Chemistry*, *374*(12), 1135–1146. https://doi.org/10.1515/bchm3.1993.374.7-12.1135

Gunčar, G., Klemenčič, I., Turk, B., Turk, V., Karaoglanovic-Carmona, A., Juliano, L., & Turk, D. (2000). Crystal structure of cathepsin X: a flip–flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease. *Structure*, *8*(3), 305–313. https://doi.org/10.1016/S0969-2126(00)00108-8

Gunčar, G., Podobnik, M., Pungerčar, J., BorutŠtrukelj, Turk, V., & Turk, D. (1998). Crystal structure of porcine cathepsin H determined at 2.1 å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. *Structure*, *6*(1), 51–61. https://doi.org/10.1016/S0969-2126(98)00007-0

Halangk, W., Lerch, M. M., Brandt-Nedelev, B., Roth, W., Ruthenbuerger, M., Reinheckel, T., … Deussing, J. (2000). Role of cathepsin B in intracellular trypsinogen activation and the onset of acute pancreatitis. *Journal of Clinical Investigation*, *106*(6), 773–781. https://doi.org/10.1172/JCI9411

Hashimoto, R., Sakamoto, A., Deguchi, S., Yi, R., Sano, E., Hotta, A., … Takayama, K. (2021). Dual inhibition of TMPRSS2 and Cathepsin B prevents SARS-CoV-2 infection in iPS cells. *Molecular Therapy - Nucleic Acids*, *26*(December), 1107–1114. https://doi.org/10.1016/j.omtn.2021.10.016

Hasnain, S., Hirama, T., Huber, C. P., Mason, P., & Mort, J. S. (1993). Characterization of cathepsin B specificity by site-directed mutagenesis. Importance of Glu245 in the S2-P2 specificity for arginine and its role in transition state stabilization. *Journal of Biological Chemistry*, *268*(1), 235–240. https://doi.org/10.1016/S0021-9258(18)54140-5

Hook, V., Yoon, M., Mosier, C., Ito, G., Podvin, S., Head, B. P., … Hook, G. (2020). Cathepsin B in neurodegeneration of Alzheimer's disease, traumatic brain injury, and related brain disorders. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1868*(8), 140428. https://doi.org/10.1016/j.bbapap.2020.140428

Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T., & Mort, J. S. (1997). Role of the occluding loop in cathepsin B activity. *Journal of Biological Chemistry*, *272*(2), 1197–1202. https://doi.org/10.1074/jbc.272.2.1197

Izaguirre, G. (2025). Exosites: beyond the limitations of the protease active site. *The FEBS Journal*, *292*(8), 1884–1886. https://doi.org/10.1111/febs.70057

Jafari, R., Almqvist, H., Axelsson, H., Ignatushchenko, M., Lundbäck, T., Nordlund, P., & Molina, D. M. (2014). The cellular thermal shift assay for evaluating drug target interactions in cells. *Nature Protocols*, *9*(9), 2100–2122. https://doi.org/10.1038/nprot.2014.138

Janko, M., Zink, A., Gigler, A. M., Heckl, W. M., & Stark, R. W. (2010). Nanostructure and mechanics of mummified type I collagen from the 5300-year-old Tyrolean Iceman. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1692), 2301–2309. https://doi.org/10.1098/rspb.2010.0377

Jerome, C., Missbach, M., & Gamse, R. (2011). Balicatib, a cathepsin K inhibitor, stimulates periosteal bone formation in monkeys. *Osteoporosis International*, *22*(12), 3001–3011. https://doi.org/10.1007/s00198-011-1529-x

Jongen-Lavrencic, M., Pabst, T., Bories, P., Griškevičius, L., Huls, G., de Leeuw, D. C., … Subklewe, M. (2024). MP0533 (CD33 x CD123 x CD70 x CD3), a tetra-specific CD3-engaging Darpin for the treatment of patients with relapsed/refractory AML or MDS/AML: Results of an ongoing phase 1/2a study. *Blood*, *144*(Supplement 1), 2881–2881. https://doi.org/10.1182/blood-2024-199211

Jordans, S., Jenko-Kokalj, S., Kühl, N. M., Tedelind, S., Sendt, W., Brömme, D., … Brix, K. (2009). Monitoring compartment-specific substrate cleavage by cathepsins B, K, L, and S at physiological pH and redox conditions. *BMC Biochemistry*, *10*(1), 23. https://doi.org/10.1186/1471-2091-10-23

Kähkönen, T. E., Halleen, J. M., & Bernoulli, J. (2020). Immunotherapies and metastatic cancers: understanding utility and predictivity of human immune cell engrafted mice in preclinical drug development. *Cancers*, *12*(6), 1615. https://doi.org/10.3390/cancers12061615

Kasabova, M., Villeret, B., Gombault, A., Lecaille, F., Reinheckel, T., Marchand-Adam, S., … Lalmanach, G. (2016). Discordance in cathepsin B and cystatin C expressions

in bronchoalveolar fluids between murine bleomycin-induced fibrosis and human idiopathic fibrosis. *Respiratory Research*, *17*(1), 118. https://doi.org/10.1186/s12931-016-0432-6

Katunuma, N. (2011). Structure-based development of specific inhibitors for individual cathepsins and their medical applications. *Proceedings of the Japan Academy Series B: Physical and Biological Sciences*, *87*(2), 29–39. https://doi.org/10.2183/pjab.87.29

Keppler, D., & Sloane, B. F. (1996). Cathepsin B: multiple enzyme forms from a single gene and their relation to cancer. *Enzyme and Protein*, *49*(1–3), 94–105. https://doi.org/10.1159/000468619

Kikuchi, H., Yamada, T., Furuya, H., Doh-ura, K., Ohyagi, Y., Iwaki, T., & Kira, J. (2003). Involvement of cathepsin B in the motor neuron degeneration of amyotrophic lateral sclerosis. *Acta Neuropathologica*, *105*(5), 462–468. https://doi.org/10.1007/s00401-002-0667-9

Klein, T., Eckhard, U., Dufour, A., Solis, N., & Overall, C. M. (2018). Proteolytic cleavage - mechanisms, function, and "omic" approaches for a near-ubiquitous posttranslational modification. *Chemical Reviews*, *118*(3), 1137–1168. https://doi.org/10.1021/acs.chemrev.7b00120

Kobayashi, H., Moniwa, N., Sugimura, M., Shinohara, H., Ohi, H., & Terao, T. (1993). Effects of membrane-associated cathepsin B on the activation of receptor-bound prourokinase and subsequent invasion of reconstituted basement membranes. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1178*(1), 55–62. https://doi.org/10.1016/0167-4889(93)90109-3

Kohl, A., Binz, H. K., Forrer, P., Stumpp, M. T., Plückthun, A., & Grütter, M. G. (2003). Designed to be stable: crystal structure of a consensus ankyrin repeat protein. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 1700–1705. https://doi.org/10.1073/pnas.0337680100

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., & Peterson, H. (2023). g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*, *51*(W1), W207–W212. https://doi.org/10.1093/nar/gkad347

Kordiš, D., & Turk, V. (2009). Phylogenomic analysis of the cystatin superfamily in eukaryotes and prokaryotes. *BMC Evolutionary Biology*, *9*(1), 266. https://doi.org/10.1186/1471-2148-9-266

Kos, J., Mitrović, A., & Mirković, B. (2014). The current stage of cathepsin B inhibitors as potential anticancer agents. *Future Medicinal Chemistry*, *6*(11), 1355–1371. https://doi.org/10.4155/fmc.14.73

Kostoulas, G., Lang, A., Nagase, H., & Baici, A. (1999). Stimulation of angiogenesis through cathepsin B inactivation of the tissue inhibitors of matrix metalloproteinases. *FEBS Letters*, *455*(3), 286–290. https://doi.org/10.1016/S0014-5793(99)00897-2

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., … Vajda, S. (2017). The ClusPro web server for protein–protein docking. *Nature Protocols*, *12*(2), 255–278. https://doi.org/10.1038/nprot.2016.169

Kramer, L., Renko, M., Završnik, J., Turk, D., Seeger, M. A., Vasiljeva, O., … Turk, B. (2017). Non-invasive in vivo imaging of tumour-associated cathepsin B by a highly

selective inhibitory DARPin. *Theranostics*, *7*(11), 2806–2821. https://doi.org/10.7150/thno.19081

Kramer, L., Turk, D., & Turk, B. (2017). The future of cysteine cathepsins in disease management. *Trends in Pharmacological Sciences*, *38*(10), 873–898. https://doi.org/10.1016/j.tips.2017.06.003

Kruszewski, W. J., Rzepko, R., Wojtacki, J., Skokowski, J., Kopacz, A., Jaśkiewicz, K., & Drucis, K. (2004). Overexpression of cathepsin B correlates with angiogenesis in colon adenocarcinoma. *Neoplasma*, *51*(1), 38–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15004658

Kumar, S., Dare, L., Vasko-Moser, J. A., James, I. E., Blake, S. M., Rickard, D. J., … Stroup, G. (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*(1), 122–131. https://doi.org/10.1016/j.bone.2006.07.015

Kummer, L., Hsu, C.-W., Dagliyan, O., MacNevin, C., Kaufholz, M., Zimmermann, B., … Plückthun, A. (2013). Knowledge-based design of a biosensor to quantify localized ERK activation in living Cells. *Chemistry & Biology*, *20*(6), 847–856. https://doi.org/10.1016/j.chembiol.2013.04.016

Kummer, L., Parizek, P., Rube, P., Millgramm, B., Prinz, A., Mittl, P. R. E., … Plückthun, A. (2012). Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries. *Proceedings of the National Academy of Sciences*, *109*(34). https://doi.org/10.1073/pnas.1205399109

Lee, J. M., Hammarén, H. M., Savitski, M. M., & Baek, S. H. (2023). Control of protein stability by post-translational modifications. *Nature Communications*, *14*(1), 1–16. https://doi.org/10.1038/s41467-023-35795-8

Lenarčič, B., & Bevec, T. (1998). Thyropins--new structurally related proteinase inhibitors. *Biological Chemistry*, *379*(2), 105–111. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9524061

Lenarčič, B., Ritonja, A., Štrukelj, B., Turk, B., & Turk, V. (1997). Equistatin, a new inhibitor of cysteine proteinases from Actinia equina, is structurally related to thyroglobulin type-1 domain. *Journal of Biological Chemistry*, *272*(21), 13899–13903. https://doi.org/10.1074/jbc.272.21.13899

Lenarčič, B., & Turk, V. (1999). Thyroglobulin type-1 domains in equistatin inhibit both papain-like cysteine proteinases and cathepsin D. *Journal of Biological Chemistry*, *274*(2), 563–566. https://doi.org/10.1074/jbc.274.2.563

Leto, G., Tumminello, F. M., Pizzolanti, G., Montalto, G., Soresi, M., & Gebbia, N. (1997). Lysosomal cathepsins B and L and Stefin A blood levels in patients with hepatocellular carcinoma and/or liver cirrhosis: potential clinical implications. *Oncology*, *54*(1), 79–83. https://doi.org/10.1159/000227666

Letunic, I., Goodstadt, L., Dickens, N. J., Doerks, T., Schultz, J., Mott, R., … Bork, P. (2002). Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Research*, *30*(1), 242–244. https://doi.org/10.1093/nar/30.1.242

Li, Z., Jaroszewski, L., Iyer, M., Sedova, M., & Godzik, A. (2020). FATCAT 2.0: towards

a better understanding of the structural diversity of proteins. *Nucleic Acids Research*, *48*(1), W60–W64. https://doi.org/10.1093/NAR/GKAA443

Lin, C. H.-T., Tariq, M. J., Ullah, F., Sannareddy, A., Khalid, F., Abbas, H., … Dima, D. (2024). Current novel targeted therapeutic strategies in multiple myeloma. *International Journal of Molecular Sciences*, *25*(11), 6192. https://doi.org/10.3390/ijms25116192

Lindström, E., Rizoska, B., Henderson, I., Terelius, Y., Jerling, M., Edenius, C., & Grabowska, U. (2018). Nonclinical and clinical pharmacological characterization of the potent and selective cathepsin K inhibitor MIV-711. *Journal of Translational Medicine*, *16*(1), 1–14. https://doi.org/10.1186/s12967-018-1497-4

Ma, K., Chen, X., Liu, W., Chen, S., Yang, C., & Yang, J. (2022). CTSB is a negative prognostic biomarker and therapeutic target associated with immune cells infiltration and immunosuppression in gliomas. *Scientific Reports*, *12*(1), 4295. https://doi.org/10.1038/s41598-022-08346-2

Man, S. M., & Kanneganti, T. D. (2016). Regulation of lysosomal dynamics and autophagy by CTSB/cathepsin B. *Autophagy*, *12*(12), 2504–2505. https://doi.org/10.1080/15548627.2016.1239679

Marquis, R. W., Ru, Y., LoCastro, S. M., Zeng, J., Yamashita, D. S., Oh, H. J., … Veber, D. F. (2001). Azepanone-based inhibitors of human and rat cathepsin K. *Journal of Medicinal Chemistry*, *44*(9), 1380–1395. https://doi.org/10.1021/jm000481x

Matsumoto, F., Saitoh, S., Fukui, R., Kobayashi, T., Tanimura, N., Konno, K., … Miyake, K. (2008). Cathepsins are required for Toll-like receptor 9 responses. *Biochemical and Biophysical Research Communications*, *367*(3), 693–699. https://doi.org/10.1016/j.bbrc.2007.12.130

Meng, J., Liu, Y., Xie, Z., Qing, H., Lei, P., & Ni, J. (2020). Nucleus distribution of cathepsin B in senescent microglia promotes brain aging through degradation of sirtuins. *Neurobiology of Aging*, *96*, 255–266. https://doi.org/10.1016/j.neurobiolaging.2020.09.001

Merritt, E. A., & Bacon, D. J. (1997). Raster3D: Photorealistic molecular graphics. In *Methods in Enzymology* (Vol. 277, pp. 505–524). https://doi.org/10.1016/S0076-6879(97)77028-9

Mihelič, M., Doberšek, A., Gunčar, G., & Turk, D. (2008). Inhibitory fragment from the p41 form of invariant chain can regulate activity of cysteine cathepsins in antigen presentation. *Journal of Biological Chemistry*, *283*(21), 14453–14460. https://doi.org/10.1074/jbc.M801283200

Mikhaylov, G., Klimpel, D., Schaschke, N., Mikac, U., Vizovišek, M., Fonović, M., … Vasiljeva, O. (2014). Selective targeting of tumor and stromal cells by a nanocarrier system displaying lipidated cathepsin B inhibitor. *Angewandte Chemie - International Edition*, *53*(38), 10077–10081. https://doi.org/10.1002/anie.201402305

Mittl, P. R. E., Ernst, P., & Plückthun, A. (2020). Chaperone-assisted structure elucidation with DARPins. *Current Opinion in Structural Biology*, *60*, 93–100. https://doi.org/10.1016/j.sbi.2019.12.009

Mo, X., Guo, Y., Qian, Q., Fu, M., Lei, S., Zhang, Y., & Zhang, H. (2020). Mendelian randomization analysis revealed potential causal factors for systemic lupus

erythematosus. *Immunology*, *159*(3), 279–288. https://doi.org/10.1111/imm.13144

Mueller-Steiner, S., Zhou, Y., Arai, H., Roberson, E. D., Sun, B., Chen, J., … Gan, L. (2006). Antiamyloidogenic and neuroprotective functions of cathepsin B: implications for Alzheimer's disease. *Neuron*, *51*(6), 703–714. https://doi.org/10.1016/j.neuron.2006.07.027

Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., … Katunuma, N. (1991). The refined 2.15 A X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *The EMBO Journal*, *10*(9), 2321–2330. https://doi.org/10.1002/j.1460-2075.1991.tb07771.x

Nagakannan, P., Islam, M. I., Conrad, M., & Eftekharpour, E. (2021). Cathepsin B is an executioner of ferroptosis. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1868*(3), 118928. https://doi.org/10.1016/j.bbamcr.2020.118928

Nägler, D. K., Storer, A. C., Portaro, F. C. V., Carmona, E., Juliano, L., & Ménard, R. (1997). Major increase in endopeptidase activity of human cathepsin B upon removal of occluding loop contacts. *Biochemistry*, *36*(41), 12608–12615. https://doi.org/10.1021/bi971264+

Nakagawa, T. Y., Brissette, W. H., Lira, P. D., Griffiths, R. J., Petrushova, N., Stock, J., … Rudensky, A. Y. (1999). Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. *Immunity*, *10*(2), 207–217. https://doi.org/10.1016/S1074-7613(00)80021-7

Nakagawa, T. Y., Roth, W., Wong, P., Nelson, A., Farr, A., Deussing, J., … Rudensky, A. Y. (1998). Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus. *Science*, *280*(5362), 450–453. https://doi.org/10.1126/science.280.5362.450

Nakanishi, H. (2020). Microglial cathepsin B as a key driver of inflammatory brain diseases and brain aging. *Neural Regeneration Research*, *15*(1), 25. https://doi.org/10.4103/1673-5374.264444

Nishimura, Y., Kawabata, T., Furuno, K., & Kato, K. (1989). Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. *Archives of Biochemistry and Biophysics*, *271*(2), 400–406. https://doi.org/10.1016/0003-9861(89)90289-0

Nouh, M. A., Mohamed, M. M., El-Shinawi, M., Shaalan, M. A., Cavallo-Medved, D., Khaled, H. M., & Sloane, B. F. (2011). Cathepsin B: a potential prognostic marker for inflammatory breast cancer. *Journal of Translational Medicine*, *9*(1), 1. https://doi.org/10.1186/1479-5876-9-1

Novinec, M., Pavšič, M., & Lenarčič, B. (2012). A simple and efficient protocol for the production of recombinant cathepsin v and other cysteine cathepsins in soluble form in Escherichia coli. *Protein Expression and Purification*, *82*(1), 1–5. https://doi.org/10.1016/j.pep.2011.11.002

Obermajer, N., Repnik, U., Jevnikar, Z., Turk, B., Kreft, M., & Kos, J. (2008). Cysteine protease cathepsin X modulates immune response via activation of β 2 integrins. *Immunology*, *124*(1), 76–88. https://doi.org/10.1111/j.1365-2567.2007.02740.x

Oberstein, T. J., Utz, J., Spitzer, P., Klafki, H. W., Wiltfang, J., Lewczuk, P., … Maler, J. M. (2021). The role of cathepsin B in the degradation of Aβ and in the production of Aβ peptides starting with Ala2 in cultured astrocytes. *Frontiers in Molecular*

Neuroscience, *13*, 615740. https://doi.org/10.3389/fnmol.2020.615740

Offen, D., Barhum, Y., Melamed, E., Embacher, N., Schindler, C., & Ransmayr, G. (2009). Spinal cord mRNA profile in patients with ALS: comparison with transgenic mice expressing the human SOD-1 mutant. *Journal of Molecular Neuroscience*, *38*(2), 85–93. https://doi.org/10.1007/s12031-007-9004-z

Olson, O. C., & Joyce, J. A. (2015). Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nature Reviews Cancer*, *15*(12), 712–729.

Orlowski, G. M., Colbert, J. D., Sharma, S., Bogyo, M., Robertson, S. A., & Rock, K. L. (2015). Multiple cathepsins promote pro–IL-1β synthesis and NLRP3-mediated IL-1β activation. *The Journal of Immunology*, *195*(4), 1685–1697. https://doi.org/10.4049/jimmunol.1500509

Park, B., Brinkmann, M. M., Spooner, E., Lee, C. C., Kim, Y.-M., & Ploegh, H. L. (2008). Proteolytic cleavage in an endolysosomal compartment is required for activation of Toll-like receptor 9. *Nature Immunology*, *9*(12), 1407–1414. https://doi.org/10.1038/ni.1669

Pennypacker, B. L., Duong, L. T., Cusick, T. E., Masarachia, P. J., Gentile, M. A., Gauthier, J. Y., … Kimmel, D. B. (2011). Cathepsin K inhibitors prevent bone loss in estrogen-deficient rabbits. *Journal of Bone and Mineral Research*, *26*(2), 252–262. https://doi.org/10.1002/jbmr.223

Pennypacker, B. L., Oballa, R. M., Levesque, S., Kimmel, D. B., & Duong, L. T. (2013). Cathepsin K inhibitors increase distal femoral bone mineral density in rapidly growing rabbits. *BMC Musculoskeletal Disorders*, *14*. https://doi.org/10.1186/1471-2474-14-344

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., … Ferrin, T. E. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Pires, D., Marques, J., Pombo, J. P., Carmo, N., Bettencourt, P., Neyrolles, O., … Anes, E. (2016). Role of cathepsins in Mycobacterium tuberculosis survival in human macrophages. *Scientific Reports*, *6*(1), 32247. https://doi.org/10.1038/srep32247

Plückthun, A. (2012). Ribosome display: a perspective. In J. Douthwaite & R. Jackson (Eds.), *Methods in Molecular Biology: Ribosome Display and Related Technologies* (pp. 3–28). Springer, New York, NY. https://doi.org/10.1007/978-1-61779-379-0_1

Plückthun, A. (2015). Designed ankyrin repeat proteins (DARPins): Binding proteins for research, diagnostics, and therapy. *Annual Review of Pharmacology and Toxicology*, *55*, 489–511. https://doi.org/10.1146/annurev-pharmtox-010611-134654

Požgan, U., Caglič, D., Rozman, B., Nagase, H., Turk, V., & Turk, B. (2010). Expression and activity profiling of selected cysteine cathepsins and matrix metalloproteinases in synovial fluids from patients with rheumatoid arthritis and osteoarthritis. *Biological Chemistry*, *391*(5), 571–579. https://doi.org/10.1515/BC.2010.035

Pungerčar, J. R., Caglič, D., Sajid, M., Dolinar, M., Vasiljeva, O., Požgan, U., … Turk, B. (2009). Autocatalytic processing of procathepsin B is triggered by proenzyme activity. *The FEBS Journal*, *276*(3), 660–668. https://doi.org/10.1111/j.1742-4658.2008.06815.x

Qi, X., Man, S. M., Malireddi, R. K. S., Karki, R., Lupfer, C., Gurung, P., … Kanneganti, T.-D. (2016). Cathepsin B modulates lysosomal biogenesis and host defense against Francisella novicida infection. *Journal of Experimental Medicine*, *213*(10), 2081–2097. https://doi.org/10.1084/jem.20151938

Rasid, O., Mériaux, V., Khan, E. M., Borde, C., Ciulean, I. S., Fitting, C., … Doyen, N. (2016). Cathepsin B-deficient mice resolve Leishmania major inflammation faster in a T Cell-dependent manner. *PLOS Neglected Tropical Diseases*, *10*(5), e0004716. https://doi.org/10.1371/journal.pntd.0004716

Rawlings, N. D., Barrett, A. J., & Bateman, A. (2011). Asparagine peptide lyases. *Journal of Biological Chemistry*, *286*(44), 38321–38328. https://doi.org/10.1074/jbc.M111.260026

Rawlings, N. D., Waller, M., Barrett, A. J., & Bateman, A. (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*, *42*(D1), D503–D509. https://doi.org/10.1093/nar/gkt953

Reiser, J., Adair, B., & Reinheckel, T. (2010). Specialized roles for cysteine cathepsins in health and disease. *Journal of Clinical Investigation*, *120*(10), 3421–3431. https://doi.org/10.1172/JCI42918

Repnik, U., Starr, A. E., Overall, C. M., & Turk, B. (2015). Cysteine cathepsins activate ELR chemokines and inactivate non-ELR chemokines. *Journal of Biological Chemistry*, *290*(22), 13800–13811. https://doi.org/10.1074/jbc.M115.638395

Richmond, A., & Su, Y. (2008). Mouse xenograft models vs GEM models for human cancer therapeutics. *Disease Models & Mechanisms*, *1*(2–3), 78–82. https://doi.org/10.1242/dmm.000976

Rozman, J., Stojan, J., Kuhelj, R., Turk, V., & Turk, B. (1999). Autocatalytic processing of recombinant human procathepsin B is a bimolecular process. *FEBS Letters*, *459*(3), 358–362. https://doi.org/10.1016/S0014-5793(99)01302-2

Rünger, T. M., Quintanilla-Dieck, M. J., & Bhawan, J. (2007). Role of cathepsin K in the turnover of the dermal extracellular matrix during scar formation. *Journal of Investigative Dermatology*, *127*(2), 293–297. https://doi.org/10.1038/sj.jid.5700535

Saini, M. G., & Bix, G. J. (2012). Oxygen–glucose deprivation (OGD) and interleukin-1 (IL-1) differentially modulate cathepsin B/L mediated generation of neuroprotective perlecan LG3 by neurons. *Brain Research*, *1438*, 65–74. https://doi.org/10.1016/j.brainres.2011.12.027

Saltarella, I., Altamura, C., Campanale, C., Laghetti, P., Vacca, A., Frassanito, M. A., & Desaphy, J.-F. (2023). Anti-angiogenic activity of drugs in multiple myeloma. *Cancers*, *15*(7), 1990. https://doi.org/10.3390/cancers15071990

SAS Institute, I. (2016). *Statistical analysis software (SAS) user's guide version 9.4.* Cary, NC: SAS Institute, Inc.

Schechter, I., & Berger, A. (1967). On the size of active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications*, *2*, 157–162.

Schick, C., Pemberton, P. A., Shi, G.-P., Kamachi, Y., Çataltepe, S., Bartuski, A. J., … Silverman, G. A. (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*(15), 5258–5266. https://doi.org/10.1021/bi972521d

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Schweizer, A., Roschitzki-Voser, H., Amstutz, P., Briand, C., Gulotti-Georgieva, M., Prenosil, E., … Grütter, M. G. (2007). Inhibition of caspase-2 by a designed ankyrin repeat protein: Specificity, structure, and inhibition mechanism. *Structure*, *15*(5), 625–636. https://doi.org/10.1016/j.str.2007.03.014

Schwenck, J., Maurer, A., Fehrenbacher, B., Mehling, R., Knopf, P., Mucha, N., … Kneilling, M. (2019). Cysteine-type cathepsins promote the effector phase of acute cutaneous delayed-type hypersensitivity reactions. *Theranostics*, *9*(13), 3903–3917. https://doi.org/10.7150/thno.31037

Seeger, M. A., Zbinden, R., Flütsch, A., Gutte, P. G. M., Engeler, S., Roschitzki-Voser, H., & Grütter, M. G. (2013). Design, construction, and characterization of a second-generation DARPin library with reduced hydrophobicity. *Protein Science*, *22*(9), 1239–1257. https://doi.org/10.1002/pro.2312

Sendler, M., Maertin, S., John, D., Persike, M., Weiss, F. U., Krüger, B., … Lerch, M. M. (2016). Cathepsin B activity initiates apoptosis via digestive protease activation in pancreatic acinar cells and experimental pancreatitis. *Journal of Biological Chemistry*, *291*(28), 14717–14731. https://doi.org/10.1074/jbc.M116.718999

Sennhauser, G., Amstutz, P., Briand, C., Storchenegger, O., & Grütter, M. G. (2006). Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS Biology*, *5*(1), e7. https://doi.org/10.1371/journal.pbio.0050007

Sennhauser, G., & Grütter, M. G. (2008). Chaperone-assisted crystallography with DARPins. *Structure*, *16*(10), 1443–1453. https://doi.org/10.1016/j.str.2008.08.010

Shi, G.-P., Villadangos, J. A., Dranoff, G., Small, C., Gu, L., Haley, K. J., … Chapman, H. A. (1999). Cathepsin S required for normal MHC class II peptide loading and germinal center development. *Immunity*, *10*(2), 197–206. https://doi.org/10.1016/S1074-7613(00)80020-5

Sievers, F., & Higgins, D. G. (2021). The Clustal Omega multiple alignment package. In K. Katoh (Ed.), *Multiple Sequence Alignment* (pp. 3–16). Humana New York, NY. https://doi.org/10.1007/978-1-0716-1036-7_1

Sivaraman, J., Lalumière1, M., Ménard, R., & Cygler, M. (1999). Crystal structure of wild-type human procathepsin K. *Protein Science*, *8*(2), 283–290. https://doi.org/10.1110/ps.8.2.283

Sivaraman, J., Nägler, D. K., Zhang, R., Ménard, R., & Cygler, M. (2000). Crystal structure of human procathepsin X: a cysteine protease with the proregion covalently linked to the active site cysteine. *Journal of Molecular Biology*, *295*(4), 939–951. https://doi.org/10.1006/jmbi.1999.3410

Sloane, B., Yan, S., Podgorski, I., Linebaugh, B. E., Cher, M. L., Mai, J., … Moin, K. (2005). Cathepsin B and tumor proteolysis: contribution of the tumor microenvironment. *Seminars in Cancer Biology*, *15*(2), 149–157. https://doi.org/10.1016/j.semcancer.2004.08.001

Sobotič, B., Vizovišek, M., Vidmar, R., Van Damme, P., Gocheva, V., Joyce, J. A., … Fonović, M. (2015). Proteomic identification of cysteine cathepsin substrates shed

from the surface of cancer cells. *Molecular & Cellular Proteomics*, *14*(8), 2213–2228. https://doi.org/10.1074/mcp.M114.044628

Stroup, G. B., Lark, M. W., Veber, D. F., Bhattacharyya, A., Blake, S., Dare, L. C., … Gowen, M. (2001). Potent and selective inhibition of human cathepsin K leads to inhibition of bone resorption in vivo in a nonhuman primate. *Journal of Bone and Mineral Research*, *16*(10), 1739–1746. https://doi.org/10.1359/jbmr.2001.16.10.1739

Stumpp, M. T., Binz, H. K., & Amstutz, P. (2008). DARPins: A new generation of protein therapeutics. *Drug Discovery Today*, *13*(15–16), 695–701. https://doi.org/10.1016/j.drudis.2008.04.013

Stumpp, M. T., Kirkin, V., Sanderson, M. P., Dawson, K. M., & Legenne, P. (2025, May 22). A decade of clinical experience with DARPins in oncology and virology: a systematic review. *MedRxiv*, pp. 1–34. https://doi.org/10.1101/2025.05.21.25328103

Sun, L., Wu, Z., Baba, M., Peters, C., Uchiyama, Y., & Nakanishi, H. (2010). Cathepsin B-dependent motor neuron death after nerve injury in the adult mouse. *Biochemical and Biophysical Research Communications*, *399*(3), 391–395. https://doi.org/10.1016/j.bbrc.2010.07.084

Szedlacsek, S. E., Ostafe, V., Serban, M., & Vlad, M. O. (1988). A re-evaluation of the kinetic equations for hyperbolic tight-binding inhibition. *Biochemical Journal*, *254*(1), 311–312. https://doi.org/10.1042/bj2540311

Takahashi, D., Iwasaki, N., Kon, S., Matsui, Y., Majima, T., Minami, A., & Uede, T. (2009). Down-regulation of cathepsin K in synovium leads to progression of osteoarthritis in rabbits. *Arthritis and Rheumatism*, *60*(8), 2372–2380. https://doi.org/10.1002/art.24718

Theurillat, J.-P., Dreier, B., Nagy-Davidescu, G., Seifert, B., Behnke, S., Zürrer-Härdi, U., … Moch, H. (2010). Designed ankyrin repeat proteins: a novel tool for testing epidermal growth factor receptor 2 expression in breast cancer. *Modern Pathology*, *23*(9), 1289–1297. https://doi.org/10.1038/modpathol.2010.103

Timmer, J. C., Zhu, W., Pop, C., Regan, T., Snipas, S. J., Eroshkin, A. M., … Salvesen, G. S. (2009). Structural and kinetic determinants of protease substrates. *Nature Structural and Molecular Biology*, *16*(10), 1101–1108. https://doi.org/10.1038/nsmb.1668

Tolosa, E., Li, W., Yasuda, Y., Wienhold, W., Denzin, L. K., Lautwein, A., … Brömme, D. (2003). Cathepsin V is involved in the degradation of invariant chain in human thymus and is overexpressed in myasthenia gravis. *Journal of Clinical Investigation*, *112*(4), 517–526. https://doi.org/10.1172/JCI18028

Tran, A. P., Sundar, S., Yu, M., Lang, B. T., & Silver, J. (2018). Modulation of receptor protein tyrosine phosphatase sigma increases chondroitin sulfate proteoglycan degradation through cathepsin B secretion to enhance axon outgrowth. *The Journal of Neuroscience*, *38*(23), 5399–5414. https://doi.org/10.1523/JNEUROSCI.3214-17.2018

Turk, B., Turk, D., & Salvesen, G. (2002). Regulating cysteine protease activity: essential role of protease inhibitors as guardians and regulators. *Current Pharmaceutical Design*, *8*(18), 1623–1637. https://doi.org/10.2174/1381612023394124

Turk, D. (2001). Structure of human dipeptidyl peptidase I (cathepsin C): exclusion

domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *The EMBO Journal*, *20*(23), 6570–6582. https://doi.org/10.1093/emboj/20.23.6570

Turk, D. (2013). MAIN software for density averaging, model building, structure refinement and validation. *Acta Crystallographica Section D: Biological Crystallography*, *69*(8), 1342–1357. https://doi.org/10.1107/S0907444913008408

Turk, D., Gunčar, G., Podobnik, M., & Turk, B. (1998). Revised definition of substrate binding sites of papain-like cysteine proteases. *Biological Chemistry*, *379*(2), 137–148. https://doi.org/10.1515/bchm.1998.379.2.137

Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M., & Turk, V. (1996). Crystal structures of human procathepsin B at 3.2 and 3.3 Å resolution reveal an interaction motif between a papain-like cysteine protease and its propeptide. *FEBS Letters*, *384*(3), 211–214. https://doi.org/10.1016/0014-5793(96)00309-2

Turk, V., & Bode, W. (1991). The cystatins: protein inhibitors of cysteine proteinases. *FEBS Letters*, *285*(2), 213–219. https://doi.org/10.1016/0014-5793(91)80804-C

Turk, V., Stoka, V., & Turk, D. (2008). Cystatins: biochemical and structural properties, and medical relevance. *Frontiers in Bioscience*, *Volume*(13), 5406–5420. https://doi.org/10.2741/3089

Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T., Turk, B., & Turk, D. (2012). Cysteine cathepsins: from structure, function and regulation to new frontiers. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1824*(1), 68–88. https://doi.org/10.1016/j.bbapap.2011.10.002

Turk, V., Turk, D., Dolenc, I., & Stoka, V. (2019). Characteristics, structure, and biological role of stefins (type-1 cystatins) of human, other mammals, and parasite origin. *Acta Chimica Slovenica*, *66*(1), 5–17. https://doi.org/10.17344/acsi.2018.4639

Tušar, L., Loboda, J., Impens, F., Sosnowski, P., Van Quickelberghe, E., Vidmar, R., … Turk, D. (2023). Proteomic data and structure analysis combined reveal interplay of structural rigidity and flexibility on selectivity of cysteine cathepsins. *Communications Biology*, *6*(1), 1–15. https://doi.org/10.1038/s42003-023-04772-8

Vasiljeva, O., Papazoglou, A., Krüger, A., Brodoefel, H., Korovin, M., Deussing, J., … Reinheckel, T. (2006). Tumor cell–derived and macrophage-derived cathepsin B promotes progression and lung metastasis of mammary cancer. *Cancer Research*, *66*(10), 5242–5250. https://doi.org/10.1158/0008-5472.CAN-05-4463

Vasiljeva, O., Reinheckel, T., Peters, C., Turk, D., Turk, V., & Turk, B. (2007). Emerging roles of cysteine cathepsins in disease and their potential as drug targets. *Current Pharmaceutical Design*, *13*(4), 387–403. https://doi.org/10.2174/138161207780162962

Vasiljeva, O., & Turk, B. (2008). Dual contrasting roles of cysteine cathepsins in cancer progression: Apoptosis versus tumour invasion. *Biochimie*, *90*(2), 380–386. https://doi.org/10.1016/j.biochi.2007.10.004

Virnekas, B., Ge, L., Plückthun, A., Schneider, K. C., Wellnhofer, G., & Moroney, S. E. (1994). Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acids Research*, *22*(25), 5600–5607. https://doi.org/10.1093/nar/22.25.5600

Vizovišek, M., Fonović, M., & Turk, B. (2019). Cysteine cathepsins in extracellular matrix

remodeling: extracellular matrix degradation and beyond. *Matrix Biology*, *75–76*, 141–159. https://doi.org/10.1016/j.matbio.2018.01.024

Vizovišek, M., Vidmar, R., Fonović, M., & Turk, B. (2016). Current trends and challenges in proteomic identification of protease substrates. *Biochimie*, *122*, 77–87. https://doi.org/10.1016/j.biochi.2015.10.017

Vizovišek, M., Vidmar, R., Van Quickelberghe, E., Impens, F., Andjelković, U., Sobotič, B., … Fonović, M. (2015). Fast profiling of protease specificity reveals similar substrate specificities for cathepsins K, L and S. *Proteomics*, *15*(14), 2479–2490. https://doi.org/10.1002/pmic.201400460

Wang, Y., Jia, L., Shen, J., Wang, Y., Fu, Z., Su, S., … Xiang, M. (2018). Cathepsin B aggravates coxsackievirus B3-induced myocarditis through activating the inflammasome and promoting pyroptosis. *PLOS Pathogens*, *14*(1), e1006872. https://doi.org/10.1371/journal.ppat.1006872

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Welss, T., Sun, J., Irving, J. A., Blum, R., Smith, A. I., Whisstock, J. C., … Abts, H. F. (2003). Hurpin is a selective inhibitor of lysosomal cathepsin L and protects keratinocytes from ultraviolet-induced apoptosis. *Biochemistry*, *42*(24), 7381–7389. https://doi.org/10.1021/bi027307q

Wheater, G., Elshahaly, M., Tuck, S. P., Datta, H. K., & van Laar, J. M. (2013). The clinical utility of bone marker measurements in osteoporosis. *Journal of Translational Medicine*, *11*(1), 201. https://doi.org/10.1186/1479-5876-11-201

Xie, Z., Yang, Q., Lan, F., Kong, W., Zhao, S., Sun, J., … Ni, J. (2025). Cathepsin B deficiency disrupts cortical development via PEG3, leading to depression-like behavior. *Communications Biology*, *8*(1), 1097. https://doi.org/10.1038/s42003-025-08508-8

Xie, Z., Zhao, M., Yan, C., Kong, W., Lan, F., Narengaowa, … Ni, J. (2023). Cathepsin B in programmed cell death machinery: mechanisms of execution and regulatory pathways. *Cell Death and Disease*, *14*(4). https://doi.org/10.1038/s41419-023-05786-0

Yin, M., Soikkeli, J., Jahkola, T., Virolainen, S., Saksela, O., & Hölttä, E. (2012). TGF-β signaling, activated stromal fibroblasts, and cysteine cathepsins B and L drive the invasive growth of human melanoma cells. *The American Journal of Pathology*, *181*(6), 2202–2216. https://doi.org/10.1016/j.ajpath.2012.08.027

Zanelatto, A. de C. O., Lacerda, G. de S., Accardo, C. de M., Rosário, N. F. do, Silva, A. A. da, Motta, G., … Xavier, A. R. (2022). Cathepsin B and plasma kallikrein are reliable biomarkers to discriminate clinically significant hepatic fibrosis in patients with chronic hepatitis-C infection. *Microorganisms*, *10*(9), 1769. https://doi.org/10.3390/microorganisms10091769

Zarić, M., Tušar, L., Kramer, L., Vasiljeva, O., Novak, M., Impens, F., … Turk, B. (2025). Structural and proteomic analysis of the mouse cathepsin B-DARPin 4m3 complex reveals species-specific binding determinants. *International Journal of Molecular Sciences*, *26*(24), 1–37. https://doi.org/10.3390/ijms262411910

Zhang, X., Zhou, Y., Yu, X., Huang, Q., Fang, W., Li, J., … Shi, G.-P. (2019). Differential roles of cysteinyl cathepsins in TGF-β signaling and tissue fibrosis. *IScience*, *19*, 607–622. https://doi.org/10.1016/j.isci.2019.08.014

Bibliography

Publications Related to the Thesis

Journal Articles

Zarić, M., Tušar, L., Kramer, L., Vasiljeva, O., Novak, M., Impens, F., Usenik, A., Gevaert, K., Turk, D., Turk, B. (2025). Structural and proteomic analysis of the mouse cathepsin B-DARPin 4m3 complex reveals species-specific binding determinants. *International Journal of Molecular Sciences*, *26*(24), 11910, str. 1-37. doi: 10.3390/ijms262411910

Biasizzo, M.; Javoršek, U.; Vidak, E.; Zarić, M.; Turk, B. (2022). Cysteine cathepsins: a long and winding road towards clinics. *Molecular Aspects of Medicine, 88*, 101150. doi: 10.1016/j.mam.2022.101150

Conference Contributions

Zarić, Miki, Tušar, L., Usenik, A., Kramer, L., Vasiljeva, O., Novak, M., Turk, D., Turk, B. Biochemical and structural comparison of human and mouse cathepsin B complexes with DARPins. In: STOKA, Veronika (ur.), *et al. Proteolysis: at the interface between health and disease: FEBS 2024 advanced course: Bled, Slovenia 14 - 18 September, 2024: programme and book of abstracts*. FEBS 2024 advanced course, Bled, Slovenia, 14 - 18 September, 2024.

Other Publications

Majc, B.; Sever, T.; Zarić, M.; Breznik, B.; Turk, B.; Lah, T. T. (2020). Epithelial-to-mesenchymal transition as the driver of changing carcinoma and glioblastoma microenvironment. *Biochimica et Biophysica Acta – Molecular Cell Research, 1867*(10), 118782. doi: 10.1016/j.bbamcr.2020.118782

Panevska, A., Hodnik, V., Skočaj, M. *et al.* Pore-forming protein complexes from *Pleurotus* mushrooms kill western corn rootworm and Colorado potato beetle through targeting membrane ceramide phosphoethanolamine. (2019). *Scientific Reports, 9*, 5073. doi: 10.1038/s41598-019-41450-4

Conference Contributions

Stantič, M., Novak, A., Zarić, M., Javoršek, U., Seeger, M. A., Turk, B. Generation of synthetic nanobodies for non-invasive detection of cathepsin B in cancer. In: STOKA, Veronika (ur.), *et al. Proteolysis: at the interface between health and disease: FEBS 2024 advanced course: Bled, Slovenia 14 - 18 September, 2024: programme and book of abstracts*. FEBS 2024 advanced course, Bled, Slovenia, 14 - 18 September, 2024.

Novak, A., Zarić, M., Stantič, M., Javoršek, U., Seeger, M. A., Turk, B. Generation of synthetic nanobodies for non-invasive detection of cathepsin B in cancer. In: SEPČIĆ, Kristina (ed.), PAVŠIČ, Miha (ed.). *15th Meeting of the Slovenian Biochemical Society with International Participation = 15. srečanje Slovenskega biokemijskega društva z mednarodno udeležbo : book of abstracts = zbornik povzetkov : Portorož, 20-23 September 2023 = [20.–23. september 2023, Portorož, Slovenija]*. Ljubljana: Slovenian Biochemical Society: = Slovensko biokemijsko društvo, 2023.

Biography

Miki Zarić completed his undergraduate studies at the Biotechnical Faculty, University of Ljubljana, in September 2015 and obtained the Bsc in Biology. He enrolled in a master's program at the same faculty, the Molecular and Functional Biology programme, which he completed in September 2018. For his master's thesis, under the supervision of Prof. Dr Kristina Sepčić and Dr. Jaka Razinger, he studied the toxicity of proteins from the oyster mushroom (*Pleurotus ostreatus*) against widely distributed insect agricultural pests. In October 2018, he enrolled in the Nanosciences and Nanotechnologies programme at the Jožef Stefan International Postgraduate School under the supervision of Prof. Dr. Olga Vasiljeva and co-supervisor Prof. DDr. Boris Turk at the Jožef Stefan Institute in the department B1 Biochemistry, Molecular and Structural Biology, where he worked on the comparison of mouse and human cathepsins and their interactions with synthetic antibody mimetics.