

Univerzita Karlova
1. lékařská fakulta

Doktorský studijní program
Biochemie a patobiochemie



UNIVERZITA KARLOVA
1. lékařská fakulta

Mgr. Anežka Tichá

Vývoj inhibitorů proteas z rodiny rhomboidů jako nástrojů pro studium jejich biologických funkcí

Development of inhibitors of rhomboid proteases as tools for the study of their biological functions

Disertační práce

Vedoucí závěrečné práce/Školitel: Ing. Kvido Strišovský, PhD.

Praha, 2019

Prohlášení:

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

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

V Praze, 6. 3. 2019

Anežka Tichá

Podpis

Identifikační záznam:

TICHÁ, Anežka. *Vývoj inhibitorů proteas z rodiny rhomboidů jako nástrojů pro studium jejich biologických funkcí (Development of inhibitors of rhomboid proteases as tools for the study of their biological functions)*. 83 stran, 4 přílohy. Praha, 2019. Disertační práce. Univerzita Karlova, 1. lékařská fakulta, Ústav organické chemie a biochemie AV ČR, v. v. i. Vedoucí práce: Ing. Kvido Stříšovský, PhD.

Identification record:

TICHÁ, Anežka. *Development of inhibitors of rhomboid proteases as tools for the study of their biological functions (Vývoj inhibitorů proteas z rodiny rhomboidů jako nástrojů pro studium jejich biologických funkcí)*. 83 pages, 4 appendices. Praha, 2019. PhD thesis. Charles University, First Faculty of Medicine, Institute of organic chemistry and biochemistry of the CAS. Supervisor: Ing. Kvido Stříšovský, PhD.

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my supervisor Kvido Stříšovský. For giving me the opportunity to work in his lab, to be part of his amazing team. For his inspiring enthusiasm for science, for all the support and valuable advice he gave me.

I would also like to thank my awesome colleagues. Thanks for all the fun with you, the never-ending chats about the “sometimes-excruciating” parts of a life-as-a-PhD-student, our inevitable scientific ups and downs and for listening to all my (also never-ending) complaints. You are such an inspiration and thanks to you the long hours spent in the lab felt almost like being at home.

I would also like to acknowledge my colleagues from IOCB who helped me with the project. Mainly Stancho Stanchev who synthesised all of the substrates and inhibitors I describe in my thesis. Honza Škerle and Kuba Began for their help with the experiments and much helpful advice. Peťa Rampírová for looking after our lab supplies and making all of our lives so much easier. Radko Souček for amino acid analysis of the peptides and the team of mass-spectrometry for the MS measurements. And Blanka Collis for her support, optimism and useful feedback on my thesis.

Velký dík patří také mým přátelům mimo ÚOCHB, především těm z mého sboru. Díky, že jste mi věřili a vždycky mě ve všem podporovali, vážně to pro mě hodně znamená!

Na závěr bych chtěla poděkovat své rodině, především mým skvělým rodičům. Děkuju, že jste mi umožnili moje studia, díky, že se mám pořád kam vracet, že u vás vždycky najdu podporu (tak za jak dlouho budeš končit? ;-)), zázemí, lásku... prostě domov.



ABSTRACT

Rhomboids are intramembrane serine proteases that belong to the evolutionarily widespread rhomboid superfamily. Rhomboids developed a slightly different catalytic mechanism compared to classical serine proteases; they utilise a catalytic dyad (Ser/His) instead of the common triad (Ser/His/Asp), and the rhomboid active site is buried in the membrane. This, coupled with their hydrophobicity, makes them quite difficult to study. Therefore, even though they are known to be involved in several important biological processes it is still not clear how exactly most of them are involved in the regulation of or in the pathologies of diseases related to these processes (such as malaria, Parkinson's disease or cancer). Our understanding is hindered by the lack of tools for their characterisation both *in vitro* and *in vivo*. In my thesis I present new fluorogenic substrates based on the LacYTM2 sequence, which is hydrolysed by several different rhomboid proteases. Using Förster resonance energy transfer (FRET)-based methods, these substrates are suitable for continuous monitoring of rhomboid activity *in vitro*. Modifications in the P5-P1 residues can improve selectivity for a specific rhomboid, the choice of FRET pair of fluorophores that absorbs light of longer wavelengths makes them suitable for high throughput screening (HTS).

Selective and potent inhibitors are a valuable tool for studying the molecular mechanisms underlying enzyme function. However, such inhibitors have been lacking for rhomboid proteases. The inhibitors developed in my thesis are non-toxic and easily synthetically accessible and modifiable. The inhibitors based on the N-methylen saccharin or the benzoxazin-4-one scaffold are not potent enough for direct cell biological use, but further derivatisation could lead to improvement in potency. The peptidyl ketoamides are based on the structural understanding of rhomboid specificity and mechanism and they are the most promising class. They combine a substrate-derived peptidyl part with the electrophilic reactive group of the ketoamide, extended by a hydrophobic substituent. The resulting inhibitor scaffold is potent, selective, covalent and reversible. With low nanomolar potency *in vivo*, peptidyl ketoamides are by far the most effective rhomboid inhibitors available. Modifications of the peptidyl part and the C-terminal hydrophobic substituent will enable tuning of inhibitor selectivity to diverse rhomboid proteases.

Key words: intramembrane proteolysis, rhomboid, rhomboid inhibitors, peptidyl ketoamides, fluorogenic substrates

ABSTRAKT

Rhomboidy jsou serinové membránové proteasy, které patří do evolučně rozšířené rodiny rhomboidních proteinů. Rhomboidy si ve srovnání s klasickými serinovými proteasami vyvinuly odlišný katalytický mechanismus; místo běžně triády (Ser/His/Asp) k proteolýze využívají pouze katalytickou dyádu (Ser/His) a jejich aktivní místo se nachází uvnitř lipidové membrány, což — společně s faktem, že jsou hydrofobní — komplikuje jejich studium. Přestože je známé jejich zapojení v mnoha důležitých biologických procesech, stále u většiny rhomboidů není jasný konkrétní mechanismus, kterým přispívají k regulaci těchto procesů nebo k patologickým projevům souvisejících onemocnění (jako např. malárie, Parkinsonova choroba nebo rakovina). Pochopení těchto dějů je ztíženo nedostatkem nástrojů k jejich studiu jak *in vitro*, tak *in vivo*. Ve své práci uvádím nové fluorogenní substráty založené na Försterově rezonančním přenosu energie (FRET) a odvozené od sekvence LacYTM2, které umožňují široké využití pro kontinuální sledování aktivity mnoha rhomboidů *in vitro*. Modifikacemi v P5-P1 pozicích je možné zvýšit specifitu substrátu k cílovému rhomboidu, zatímco výběr FRET páru fluoroforů absorbujícího v červené oblasti viditelného světla umožňuje jejich využití pro rychlé testování knihoven molekul.

Selektivní a účinné inhibitory jsou cenným nástrojem ke studiu molekulárních mechanismů enzymů, avšak pro rhomboidy zatím žádné takové inhibitory nejsou známy. Inhibitory vyvinuté v rámci této práce jsou netoxické, jednoduše synteticky dostupné a modifikovatelné. Inhibitory odvozené od N-methylen saccharinu nebo benzoxazin-4-onu sice nejsou dostatečně aktivní pro biologické aplikace, ale další modifikace by mohly vést ke zvýšení účinnosti. Peptidyl ketoamidy jsou inhibitory založené na mechanismu interakce rhomboidu se substrátem a jsou zatím nejslibnější skupinou inhibitorů rhomboidů. Spojují peptidovou část odvozenou od substrátu s ketoamidovou elektrofilní reaktivní skupinou rozšířenou hydrofobním substituentem. Výsledné inhibitory jsou účinné, selektivní, kovalentní a reversibilní. Vzhledem k jejich aktivitě *in vivo* při nízkých nanomolárních koncentracích jsou peptidyl ketoamidy v současnosti zdaleka nejefektivnějšími inhibitory rhomboidů. Optimalizace peptidové části a C-koncového hydrofobního substituentu umožní návrh selektivních inhibitorů dalších členů této rodiny.

Klíčová slova: intramembránová proteolýza, rhomboid, inhibitory rhomboidů, peptidyl ketoamidy, fluorogenní substráty

TABLE OF CONTENTS

List of abbreviations	9
1 Introduction	13
1.1 THE RHOMBOID SUPERFAMILY	13
1.2 MEDICAL RELEVANCE OF RHOMBOID-LIKE PROTEINS	17
1.2.1 Protozoal rhomboid proteases	17
1.2.2 Mammalian rhomboid proteases	18
1.2.3 Rhomboid pseudoproteases	21
1.3 RHOMBOID PROTEASE STRUCTURE.....	25
1.4 MECHANISM OF RHOMBOID PROTEASE CATALYSIS	28
1.4.1 Substrate recognition	29
1.4.2 Catalytic mechanism.....	34
1.5 INHIBITORS OF RHOMBOID PROTEASES	35
1.5.1 Small heterocyclic inhibitors.....	36
1.5.2 Substrate-derived inhibitors	38
1.6 RHOMBOID PROTEASE ACTIVITY ASSAYS	40
1.6.1 <i>In vivo</i> assays.....	40
1.6.2 <i>In vitro</i> assays	41
2 Aims of the study.....	44
3 Methods.....	45
4 Results	46
4.1 PUBLICATION 1	46
4.2 PUBLICATION 2	49
4.3 PUBLICATION 3	52
4.4 PUBLICATION 4	54
5 Discussion	57
6 Summary	65
7 References	66
8 List of publications.....	82
9 Appendices	84

LIST OF ABBREVIATIONS

AarA	rhomboid protease AarA
ABP	activity-based probe
AcIATA-cmk	peptidyl chloromethyl ketone inhibitor (acetylated at the N-terminus)
AcVRMA-CHO	peptidyl aldehyde inhibitor (acetylated at the N-terminus)
ADAM10, 17	disintegrin and metalloproteinase domain-containing protein 10, 17
APP	amyloid precursor protein
A20	tumor necrosis factor alpha-induced protein 3
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
<i>B. thetaiotaomicron</i>	<i>Bacteroides thetaiotaomicron</i>
BACE1	beta-secretase 1, also called β -site APP cleaving enzyme 1
Bcl-3	B-cell lymphoma 3 protein
BIK	Bcl-2-interacting killer, apoptotic inducer
BSc5195	(1,1-Dioxido-3-oxobenzod[<i>d</i>]isothiazol-2(3 <i>H</i>)-yl)methyl 2-chlorobenzoate, saccharin derived inhibitor
BtioR3	rhomboid 3 from <i>Bacteroides thetaiotaomicron</i>
CAPF	Cbz-Ala ^P (O- <i>i</i> Pr)F, Cbz – carboxybenzyl
Cdc48	cell division control protein 48
CFP	cyan fluorescent protein
CMK	chloromethyl ketone
<i>D. melanogaster</i>	<i>Drosophila melanogaster</i>
DCI	3,4-dichloroisocoumarin
DDM	n-dodecyl β -D-maltoside
Der1	degradation in the endoplasmic reticulum protein 1
Dfm1	DER1-like family member protein 1
DFP	diisopropyl fluorophosphate
DMSO	dimethyl sulfoxide
<i>E. coli</i>	<i>Escherichia coli</i>
EGF(R)	epidermal growth factor (receptor)
EnPlex	a high-throughput flow cytometry platform to assess potency and selectivity of small molecule enzyme inhibitors
ER	endoplasmic reticulum
ERAD-L, ERAD-M	endoplasmic reticulum-associated degradation of misfolded luminal (-L) or membrane (-M) ER proteins
FITC	fluorescein isothiocyanate
FLAG	protein tag; sequence DYKDDDDK

FRET	Förster resonance energy transfer
FRMD8/iTAP	FERM domain-containing protein 8, iRhom-tail associated protein
GA	Golgi apparatus
GFP	green fluorescent protein
GlpG	rhomboid protease GlpG
Glut1	glucose transporter 1
HLE	human leukocyte elastase
Hrd1	ERAD-associated E3 ubiquitin-protein ligase
Hrd3	ERAD-associated E3 ubiquitin-protein ligase component that stabilizes Hrd1
HTS	high-throughput screening
IC₅₀	half maximal inhibitory concentration
IC 16	7-Benzoylamino-3-(3-butynoxy)-4-chloro-isocoumarin, isocoumarin rhomboid inhibitor
IκB	endogenous NF-κB inhibitor
IKK	IκB kinase complex
IKKγ	regulatory subunit of IKK
IL-6	interleukin 6
IRHD	iRhom homology domain
iRhom	inactive rhomboid protein
JLK 6	7-Amino-4-chloro-3-methoxy-1 <i>H</i> -2-benzopyran, isocoumarin rhomboid inhibitor
Jun	transcription factor AP-1, protein encoded by jun gene
k_{cat}	turnover number
K_i	inhibitory constant
K_m	Michaelis constant
KSp35	fluorogenic substrate based on LacYTM2 sequence, Edans-Dabcyl FRET fluorescence pair
KSp64	fluorogenic substrate based on LacYTM2 sequence with RVRHA in P5-P1, Tamra-QXL610 FRET fluorescence pair
KSp76	fluorogenic substrate based on LacYTM2 sequence, Tamra-QXL610 FRET fluorescence pair
LacYTM2	lactose permease, transmembrane domain 2
L1/2 loop	loop corresponding to the L1 loop of the rhomboid core, because PARL-like rhomboids have an extra N-terminal TMD it is actually between TMD2 and 3
λ_{ex}, λ_{em}	excitation/emission wavelength
LG	leaving group
L29	phenyl 2-oxo-4-phenylazetidine-1-carboxylate, β-lactam rhomboid inhibitor
L41	1-((4'-chloro-[1,1'-biphenyl]-4-yl)sulfonyl)-4-phenylazetidin-2-one,

	β -lactam rhomboid inhibitor
L62	cyclopentyl 2-oxo-4-phenylazetidine-1-carboxylate, β -lactam rhomboid inhibitor
MALDI	matrix assisted laser desorption/ionization
MAPK	mitogen-activated protein kinases
MBP	maltose binding protein
MPP	matrix processing peptidase
NF-κB	nuclear factor κ B
NR698	<i>Escherichia coli</i> MC4100 <i>imp4213</i>
<i>P. falciparum</i>	<i>Plasmodium falciparum</i>
<i>P. stuartii</i>	<i>Providencia stuartii</i>
PARL	Presenilin associated rhomboid like protease, also PINK1/PGAM5 associated rhomboid like protease
PD	Parkinson's disease
PfROM1, 4	Rhomboid proteases ROM1 and ROM4 from <i>Plasmodium falciparum</i>
PGAM5	Serine/threonine-protein phosphatase PGAM5, also Phosphoglycerate mutase family member 5
PINK1	Serine/threonine-protein kinase PINK1, also PTEN-induced putative kinase protein 1
Rbd2	Rhomboid protein 2 from <i>Saccharomyces cerevisiae</i>
Rce1	CAAX prenyl protease 2, from gene name Ras converting CAAX endopeptidase 1
RHBDD2, 3	Rhomboid domain-containing proteins 2 and 3
RHBDL1-4	Rhomboid-related protein 1-4
SAPE	streptavidin R-phycoerythrin conjugate
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
S006	7-amino-3-(but-3-yn-1-yloxy)-4-chloro-1 <i>H</i> -isochromen-1-one, isocoumarin rhomboid inhibitor
S016	7-amino-4-chloro-3-phenethoxy-1 <i>H</i> -isochromen-1-one, isocoumarin rhomboid inhibitor
<i>T. gondii</i>	<i>Toxoplasma gondii</i>
TatA	Sec-independent protein translocase protein, part of the twin-arginine translocation system
TGFα	transforming growth factor α
TgROM5	rhomboid protease ROM5 from <i>Toxoplasma gondii</i>
TLCK	N- α -tosyl-L-lysine chloromethyl ketone
TLR	toll-like receptor
TMD	transmembrane domain
TMD2/3	TMD corresponding to the TMD2 of the rhomboid core, but PARL-like rhomboids have extra N-terminal TMD, therefore it is actually TMD3

TMEM115	transmembrane protein 115
TNF	tumor necrosis factor
TOM/TIM	translocase of the outer membrane of mitochondria/ translocase of the inner membrane of mitochondria
TPCK	N- α -tosyl-L-phenylalanyl chloromethyl ketone
Trx	thyroxin domain
TSAP6	endosomal metalloredutase (also STEAP3)
T2DM	type 2 diabetes mellitus
Ub	ubiquitin
UBA	ubiquitin-assosiated domain
UBAC2	ubiquitin -associated domain-containing protein 2
Ubx2	UBX domain-containing protein 2
VBM	VCP (p97)-binding motif
WB	western blot
WR	conserved Trp-Arg motif in rhomboid L1 loop
YFP	yellow fluoescent protein
YqgP	rhomboid protease YqgP (also called GluP) from <i>Bacillus subtilis</i>
Yos9	protein OS-9 homologue

1 INTRODUCTION

Hydrolysis of proteins is a major regulatory mechanism in organisms. Numerous proteases are known to be essential for homeostasis and health. Their function is to cleave peptide bond(s) in the protein target, a process which requires the presence of water molecules. Unsurprisingly, most proteases are water-soluble proteins. It is well known that water is excluded from cellular lipid membranes so it was quite a surprise when transmembrane proteases whose catalytic residues resided within lipid membranes were identified.

Intramembrane proteases are present in all kingdoms of life and have many different functions, such as transcriptional control, cellular signalling, parasite invasion, bacterial protein translocation or control of mitochondrial membrane remodelling (Brown M. S. *et al.*, 2000; Urban S. and Freeman M., 2002). In analogy to soluble proteases, intramembrane proteases are divided into several groups according to their catalytic mechanism. The first group are the metalloproteases, which are exemplified by site-2-proteases (Rawson R. B. *et al.*, 1997). Then there are the aspartyl proteases with presenilin as the most prominent member of the group (Esler W. P. *et al.*, 2000; Wolfe M. S. *et al.*, 1999). Signal peptide peptidases are aspartyl proteases as well but they have opposite orientation compared to the presenilin-like proteins. The most recently identified group are the glutamate intramembrane proteases with the Rce1 protease as the best characterised member of this class (Manolaridis I. *et al.*, 2013). The last group are serine intramembrane proteases, also called rhomboids (Urban S. *et al.*, 2001), which are the focus of my thesis and are discussed further below.

1.1 The Rhomboid superfamily

Rhomboids were first discovered in the fruit fly *Drosophila melanogaster*, where mutations in *rhomboid* locus resulted in pointed head skeleton phenotype in embryos (Jurgens G. *et al.*, 1984). The encoded protein (named Rhomboid 1) was shown to be important in early embryonal differentiation (Bier E. *et al.*, 1990) through activation of the epidermal growth factor receptor/mitogen-activated protein kinases (EGFR/MAPK) pathway (Guichard A. *et al.*, 1999) although the exact mechanism was not clear. Later it was shown that Rhomboid 1 is actually a novel intramembrane serine protease that activates EGFR by shedding the membrane-bound TGF α -like growth factor Spitz. Thus, Rhomboid 1

plays a critical role in the regulation of the EGFR signalling pathway in *Drosophila* (Lee J. R. *et al.*, 2001; Urban S. *et al.*, 2001) (**Figure 1**).

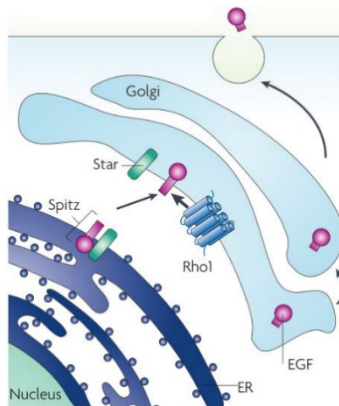


Figure 1: Involvement of Rhomboid 1 in EGFR signalling in *Drosophila melanogaster*

Membrane-bound precursor of EGFR ligand Spitz (*Drosophila* homologue of EGF) is translocated from the endoplasmic reticulum (ER) to the Golgi apparatus (GA) by protein Star. In the GA the Spitz is cleaved by Rhomboid 1 and subsequently released in its active form. Adopted from (Urban S., 2009).

Today it is clear that the rhomboid-like proteins make up a superfamily that consists both of active proteases as well as proteolytically inactive members – so called pseudoproteases (**Figure 2**). Rhomboid-like proteins occur in all kingdoms of life and are probably among the most widely distributed membrane proteins in nature. The rhomboid superfamily consists of numerous paralogue groups with low overall sequence similarity (Lemberg M. K. and Freeman M., 2007b), but all superfamily members share a typical rhomboid core formed by six transmembrane helices (Daley D. O. *et al.*, 2005; Maegawa S. *et al.*, 2005; Urban S. *et al.*, 2001) and conserved sequence features. This transmembrane core is sometimes extended by an extra N- or C-terminal transmembrane helix or by an extramembrane domain (Koonin E. V. *et al.*, 2003). Taking into account the six transmembrane domain (TMD) rhomboid core, conserved residues and by combining several statistical approaches four topological classes of rhomboid proteins were proposed (**Figure 2**).

Despite being evolutionarily widespread, based on phylogenetic analyses some authors asserted that it is not probable that the rhomboid family arose from the last universal common ancestor (Koonin E. V. *et al.*, 2003; Urban S. *et al.*, 2001). Instead it was proposed that the family appeared in some bacterial lineage, then spread by horizontal gene transfer, and archaea and eukaryotes obtained rhomboids on several independent occasions, giving rise to “secretase-type” and PARL-type rhomboids (Koonin E. V. *et al.*, 2003). Other authors speculate that rhomboids did evolve from one ancestor, supporting their claims by the wide

distribution of RHBDL4-like rhomboids and the fact that some features of secretase-type A class appear also in the secretase-type B class enzymes (Lemberg M. K. and Freeman M., 2007b). Probably a deeper analysis is needed to resolve this issue. It would perhaps be relevant now when many more genomes have been sequenced and the understanding of the superfamily has also increased. Nevertheless, the fact that internal gene duplication is more probable compared to the addition of extra non-homologous TMDs supports the first scenario (Liu Y. *et al.*, 2004a; von Heijne G., 2006). Based on transmembrane topology and sequence features, the rhomboid superfamily can be formally divided into several prominent architectural types (**Figure 2**).

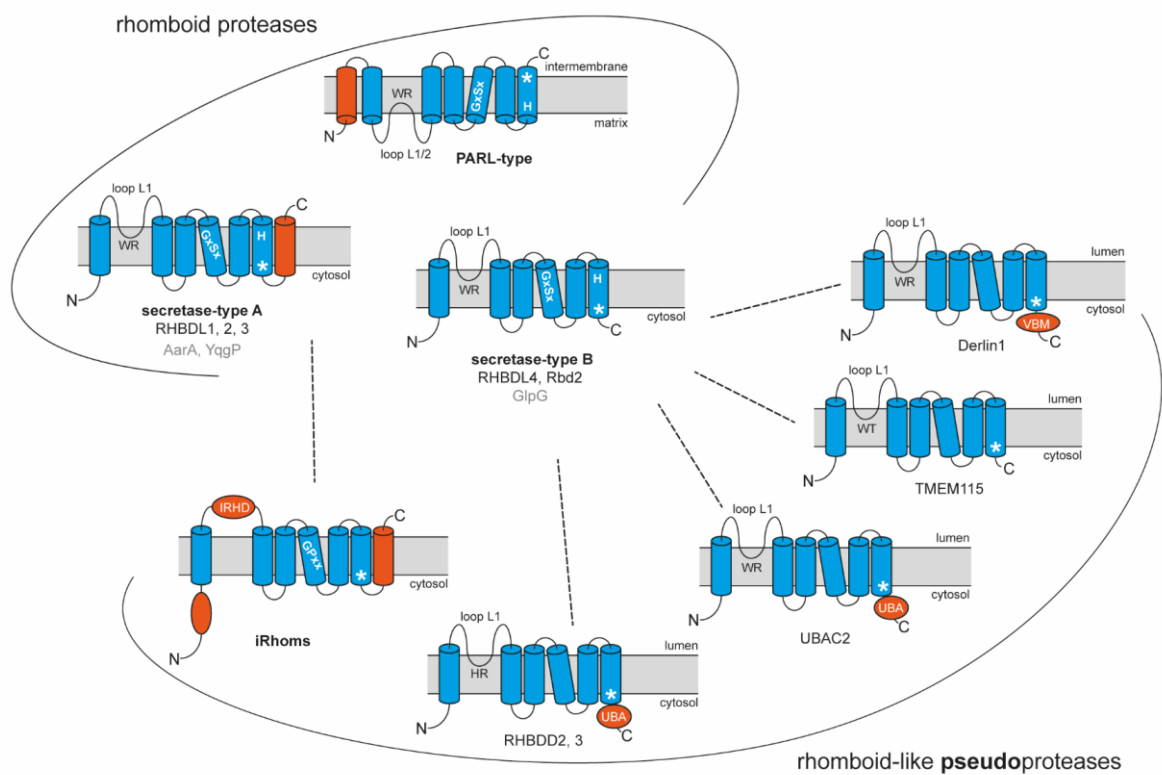


Figure 2: Subgroups of the rhomboid superfamily and their architectures

The rhomboid superfamily consists of both active proteases and catalytically inactive pseudoproteases that share the rhomboid core formed by 6 transmembrane helices (shown in blue) and several conserved elements (such as loop L1, structurally important residues). The subclasses often contain a special feature, such as extra helices or cytoplasmic domains (red). Representatives of each subclass are shown with prokaryotic members in grey. Probable evolutionary relationships are depicted by the dashed lines. The asterisk stands for GxxxG, helix dimerization motif. Abbreviations: IHRD, iRhom homology domain; UBA, ubiquitin-associated domain; VBM, VCP (p97)-binding motif. Adopted from (Ticha A. *et al.*, 2018).

Secretase-type A rhomboids have an extra C-terminal TMD, denoted 6+1 TMD. Previously it was thought that this topology is unique only for eukaryotes (Koonin E. V. *et al.*, 2003), but a later study suggests that a number of bacterial rhomboids are also predicted to have this structure (Lemberg M. K. and Freeman M., 2007b). Other typical features are

the highly conserved residues tryptophan and arginine in the L1 loop, the GxSxGVYA sequence motif around the catalytic centre and the GxxxG motif in TMD 6 (Lemberg M. K. and Freeman M., 2007b). The best characterised member of this group is *Drosophila* Rhomboid 1 (Urban S. *et al.*, 2001).

The secretase-type B class, also called the RHBDL4-like class, consists of rhomboids with six transmembrane helices. The best characterised member of this class is *E. coli* GlpG, but this topology is also seen in some eukaryotic rhomboids, for example *Saccharomyces cerevisiae* Rbd2 (Daley D. O. *et al.*, 2005; Kim H. *et al.*, 2006; Maegawa S. *et al.*, 2005) or RHBDL4 (Fleig L. *et al.*, 2012; Lemberg M. K. and Freeman M., 2007b). Unlike in secretase-type A, here the catalytic histidine is stabilised by phenylalanine (GxSxxxF motif) and most members retain only arginine as a conserved residue in the L1 loop (Lemberg M. K. and Freeman M., 2007b).

The PARL-subfamily is predicted to have an extra TMD at the N-terminus of the rhomboid core (1+6 TMD) and a seemingly opposite active site orientation compared to the secretase-type rhomboids (Lemberg M. K. and Freeman M., 2007a), i.e. facing the inside of the organelle where it resides (matrix of mitochondria). Another characteristic of this class are the long N-terminal extensions (Jeyaraju D. V. *et al.*, 2006) whereas the secretase-A and secretase-B rhomboids might have special soluble domains at either C- or N-terminus (Lemberg M. K. and Freeman M., 2007b). Although PARL-type rhomboids are missing several key parts typical for most rhomboid proteases (such as the arginine in loop L1/2 or a glutamate residue in the TMD2/3), the significant features of a rhomboid catalytic serine-histidine dyad are conserved. A high degree of sequence homology, identical topology of the members and their mitochondrial localization supports their common evolutionary origin and similar biological function (Lemberg M. K. and Freeman M., 2007b).

A remarkable number of rhomboid-like proteins appears to lack catalytic activity and most of these proteins form the rhomboid class called the iRhoms (Lemberg M. K. and Freeman M., 2007b; Zettl M. *et al.*, 2011). The characteristic feature of this class is a large globular domain between TMD1 and TMD2 (**Figure 2**) called the IRHD (for iRhom homology domain) and cytosolic N-terminal domain, which make them significantly larger than other members of the rhomboid-like superfamily. Different members of this group lack different catalytic residues but surprisingly there are also cases where both the catalytic serine and histidine are present. The catalytic (in)activity is thus probably also influenced by

the proline residue, which is present in all iRhoms in the first x-position of the conserved GxSx catalytic motif (Lemberg M. K. and Freeman M., 2007b), and was shown to disrupt activity of rhomboid proteases (Zettl M. *et al.*, 2011).

1.2 Medical relevance of rhomboid-like proteins

Several rhomboid-like proteins, including both the proteases and pseudoproteases, are known to play important roles in diseases such as colon cancer, Parkinson's disease, malaria, or inflammatory diseases (reviewed in (Dusterhoft S. *et al.*, 2017)), but in many cases the underlying mechanisms are not known. The emerging possibility that rhomboid-like proteins could be novel drug targets elevates the general interest in their research, after the first rhomboid protein was characterised nearly twenty years ago (Urban S. *et al.*, 2001). In this section, I will discuss several of the best understood rhomboid-like proteins associated with some human diseases.

1.2.1 Protozoal rhomboid proteases

The *Apicomplexa* phylum comprises pathogens that can cause severe diseases. They are obligate intracellular parasites, therefore invasion to the host cell is necessary for their survival and replication (Sibley L. D., 2004; Soldati D. *et al.*, 2004). For example, *Plasmodium spp.* cause malaria that leads to 438 000 deaths a year (World Health Organization) and up to a third of the human population is affected by *T. gondii* (Montoya J. G. and Liesenfeld O., 2004). Toxoplasmosis usually does not have obvious symptoms in adults, but it can be life threatening to immunocompromised patients because it can lead to brain or eye damage (Jones J. L. *et al.*, 2001).

Invasion of the host cell is a highly regulated process. When the host cell is recognised, the parasite attaches to it, which leads to the formation of a moving junction and to penetration (Carruthers V. B. and Boothroyd J. C., 2007). Adhesins are transmembrane proteins that interact with the receptors on the host cell and they are released on the parasite surface from the micronemes and rhoptries, the apical secretory organelles. In order to successfully invade the host cell, adhesins have to be shed which is ensured by proteolytic cleavage within their transmembrane domain (TMD) (**Figure 3**). Mutations in their TMD impair parasite invasion, showing that the shedding of adhesins is indispensable for the pathogen's life cycle (Brossier F. *et al.*, 2003; Ejigiri I. *et al.*, 2012; O'Donnell R. A. *et al.*, 2006; Olivieri A. *et al.*, 2011; Opitz C. *et al.*, 2002; Parussini F. *et al.*, 2012). In *T. gondii*

this cleavage is performed by rhomboid protease TgROM5 (Brossier F. *et al.*, 2005). In contrast, *Plasmodium* utilises two rhomboid proteases with different substrate specificity, PfROM1 and PfROM4, for adhesion-receptor complex cleavage (Baker R. P. *et al.*, 2006) (Harris P. K. *et al.*, 2005). Four classes of adhesins are needed during all stages of parasite invasion and PfROM1 and PfROM4 in combination are able to cleave all of them. Blocking rhomboid activity might prevent parasite invasion and therefore apicomplexan rhomboids present a potential drug target (Baker R. P. *et al.*, 2006).

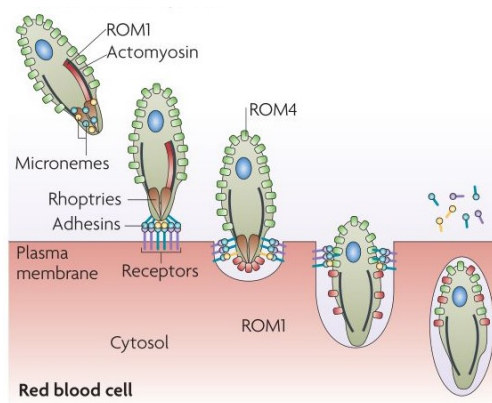


Figure 3: Rhomboid function in *Plasmodium falciparum*

Apicomplexan parasite *P. falciparum* use two rhomboids for its invasion into the host cell. Parasite adhesins interact with host cell receptors and form a moving junction that is broken down by ROM1 (red) and ROM4 (green). Adopted from (Urban S., 2009).

1.2.2 Mammalian rhomboid proteases

Mammalian genomes encode five different rhomboid proteases. RHBDL1-4 are so-called secretases as they are found in the secretory pathway (Lemberg M. K. and Freeman M., 2007b; Lohi O. *et al.*, 2004) whereas PARL is the mitochondrial rhomboid protease (McQuibban G. A. *et al.*, 2003). *Drosophila* rhomboids are involved in the EGFR signalling pathway (Lee J. R. *et al.*, 2001; Urban S. *et al.*, 2001) which is a very important signalling pathway also in mammals, both for homeostasis and cancers. However, it is well known that in mammalian cells EGFR ligands are shed by metalloproteases, mainly ADAM17 and ADAM10 (a disintegrin and metalloprotease) (Peschon J. J. *et al.*, 1998; Sahin U. *et al.*, 2004), suggesting that mammalian rhomboids may have different functions from their *Drosophila* homologues.

1.2.2.1 RHBDL2

RHBDL2 is probably the best understood mammalian rhomboid protease. It is located on the plasma membrane and expressed mainly in epithelia (Adrain C. *et al.*, 2011; Lohi O. *et al.*, 2004). Unlike other mammalian rhomboids it has similar substrate preferences to *Drosophila* Rhomboids 1-3 as it can cleave their substrates Spitz (Lohi O. *et al.*, 2004) and Gurken (Lemberg M. K. *et al.*, 2005). This suggested that RHBDL2 could also be involved in the EGFR pathway but it became clear that of the mammalian EGFR ligands RHBDL2 can only cleave EGF in the absence of metalloproteases activity (Adrain C. *et al.*, 2011). The physiological relevance of RHBDL2 in EGFR signaling is unclear.

Several other substrates of RHBDL2 were later discovered such as thrombomodulin (Lohi O. *et al.*, 2004), B-type ephrins (Pascall J. C. and Brown K. D., 2004), and EGFR (Liao H.-J. and Carpenter G., 2012). RHBDL2 was also proposed to play a role in cell migration and angiogenesis (Noy P. J. *et al.*, 2016), cell proliferation, wound healing, tumour metastasis (Cheng T. L. *et al.*, 2011) or epithelial homeostasis (Johnson N. *et al.*, 2017), but in none of these cases convincing animal data exist.

1.2.2.2 RHBDL4

RHBDL4 is a mammalian secretase located to the endoplasmic reticulum (ER) (Fleig L. *et al.*, 2012; Wunderle L. *et al.*, 2016). It was proposed to participate in the ER-associated degradation (ERAD), as its expression increases upon ER stress, and it can cleave membrane proteins that have unstable TMDs that are further degraded by ERAD (Fleig L. *et al.*, 2012). Apart from ER stress, RHBDL4 upregulation has also been connected to tumour growth (Han J. Y. *et al.*, 2015; Liu X. N. *et al.*, 2013; Song W. *et al.*, 2015). RHBDL4 was proposed to reduce apoptosis by cleaving the pro-apoptotic protein BIK (Wang Y. *et al.*, 2008) or by upregulating c-Jun and its downstream target, Bcl-3 (B-cell lymphoma 3 protein) (Ren X. *et al.*, 2013). The polytopic membrane protein TSAP6, which is involved in unconventional exosomal secretion (Keller S. *et al.*, 2006; Nickel W. and Rabouille C., 2009), was also shown to be cleaved by RHBDL4. Another RHBDL4 substrate might be proTGF α (Song W. *et al.*, 2015) but here it has been suggested that RHBDL4 instead promotes the transport of proTGF α through the secretory pathway, followed by its exosomal secretion (Wunderle L. *et al.*, 2016). RHBDL4 can also cleave the amyloid precursor protein (APP) (Paschkowsky S. *et al.*, 2016), a cell surface protein otherwise cleaved by BACE1 (β -site APP cleaving enzyme 1) (Li Q. and Sudhof T. C., 2004) and γ -secretase (De Strooper B. *et al.*, 1998) to produce deleterious A β peptides which are thought to be a molecular cause of Alzheimer's

disease (Viola K. L. and Klein W. L., 2015). The potential purpose of this might be the reduction of APP at the cell surface and subsequent A β generation (Paschkowsky S. *et al.*, 2016). Interestingly, the RHBDL4 activity towards APP appears to be regulated by membrane cholesterol levels; the lower the cholesterol levels, the higher the RHBDL4 proteolytic activity (Paschkowsky S. *et al.*, 2018). However, these findings again lack *in vivo* evidence and the precise roles of RHBDL4 in organisms remain unknown.

1.2.2.3 PARL

PARL is localised to the inner mitochondrial membrane and ubiquitous in eukaryotes (McQuibban G. A. *et al.*, 2003). PARL, abbreviated form of ‘Presenilin associated rhomboid-like protease’, was discovered as an interactor of Presenilin 2 (Pellegrini L. *et al.*, 2001), but later the connection was shown to be an artefact (McQuibban G. A. *et al.*, 2003). However, the acronym remained but with an alternative explanation proposed: ‘PGAM5/PINK1 associated rhomboid like protease’, as both proteins are PARL substrates (Spinazzi M. and De Strooper B., 2016). The substrate repertoire links PARL to mitochondrial homeostasis and mitochondrial dysfunction, mainly Parkinson’s disease (PD) and diabetes mellitus type 2 (Spinazzi M. and De Strooper B., 2016).

A mutation of PARL was identified in two patients with PD (Shi Guang *et al.*, 2011) although its pathogenicity was not confirmed (Heinitz S. *et al.*, 2011; Wust R. *et al.*, 2016). PARL cleaves PINK1 protein, a serine/threonine kinase (Deas E. *et al.*, 2011b; Greene A. W. *et al.*, 2012; Jin S. M. *et al.*, 2010; Lu W. *et al.*, 2014; Meissner C. *et al.*, 2011; Sekine S. *et al.*, 2012; Shi G. *et al.*, 2011; Shi Guang *et al.*, 2011). PINK1 monitors mitochondrial health and drives mitochondria to mitophagy upon mitochondrial function collapse (**Figure 4**) (Nguyen T. N. *et al.*, 2016), which enables a cell to eliminate defective mitochondria (Deas E. *et al.*, 2011a). PD is connected with excessive mitochondrial damage (Schapira A. H., 2008), hence impaired PINK1 induced mitophagy was proposed as a potential cause of autosomal recessive form of PD (Youle R. J. and Narendra D. P., 2011) where PINK1 is often mutated (Rogaeva E. *et al.*, 2004). Thus, PARL is at least indirectly involved in PD pathology, although the precise mechanism how PARL-PINK1 activity contributes to PD development is still not clear (Spinazzi M. and De Strooper B., 2016).

PARL is also connected to type 2 diabetes mellitus (T2DM), although its role there is even less well understood than in PD. Indirect data suggest decreased oxidative phosphorylation and lipid phosphorylation in T2DM and ageing, connecting T2DM with

mitochondrial damage (Kelley D. E. *et al.*, 2002; Morino K. *et al.*, 2005; Petersen K. F. *et al.*, 2003; Ukropcova B. *et al.*, 2007). Reduced PARL was proposed to be a risk factor for T2DM as its expression is lowered in the skeletal muscle of diabetic rodent *Psamomys obesus* (Walder K. *et al.*, 2005) and PARL downregulation leads to mitochondrial abnormalities similar to those that are observed in T2DM (Civitarese A. E. *et al.*, 2010). Although T2DM was putatively linked to previously proposed PARL-mediated mechanisms (cristae remodelling) (Civitarese A. E. *et al.*, 2010), there are still large gaps in our knowledge of the precise underlying mechanism.

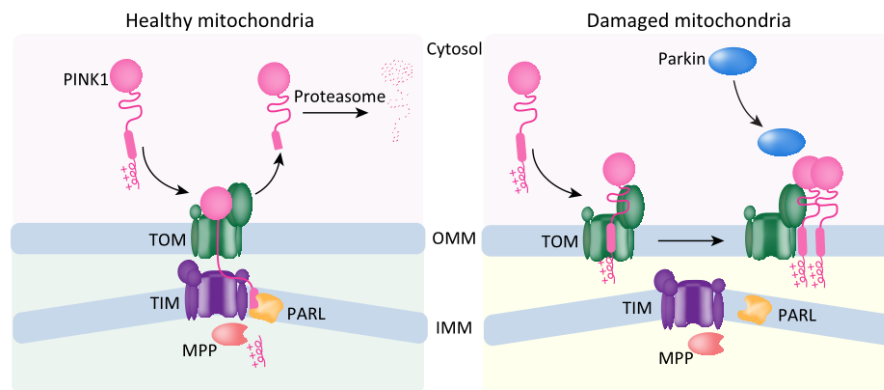


Figure 4: PINK1 mediated mitophagy

Under conditions with normal membrane potential (healthy mitochondria) PINK1 is cleaved by PARL. PINK1 is incorporated to the inner mitochondrial membrane (IMM) through the TOM/TIM complex. First, matrix processing peptidase (MPP) cleaves the N-terminal mitochondrial targeting sequence (Greene A. W. *et al.*, 2012) and subsequently the PINK1 TMD is cleaved by PARL (Deas E. *et al.*, 2011b; Jin S. M. *et al.*, 2010; Meissner C. *et al.*, 2011). This generates a fragment with an N-terminal destabilizing amino acid which drives PINK1 translocation back to the cytosol where it is degraded by the ubiquitin-proteasome system through the N-end rule pathway (Yamano K. and Youle R. J., 2013). In damaged mitochondria (with dissipated membrane potential) PINK1 is not fully imported and stays in the outer mitochondrial membrane (OMM) bound to the TOM complex (Jin S. M. *et al.*, 2010). Thus PINK1 cannot be cleaved by MPP and PARL, and is instead autophosphorylated, which is necessary to activate its kinase activity and for recruitment of the E3 ubiquitin ligase Parkin which leads to mitophagy (Okatsu K. *et al.*, 2012). Adopted from (Nguyen T. N. *et al.*, 2016).

1.2.3 Rhomboid pseudoproteases

The subfamilies of catalytically inactive rhomboid-like proteins were discovered by sequence comparisons (Lemberg M. K. *et al.*, 2005), gene annotation by hidden Markov models and BLAST searches (Greenblatt E. J. *et al.*, 2011; Hempel F. *et al.*, 2009; Kirst M. E. *et al.*, 2005; Lemberg M. K. and Freeman M., 2007b). Their wide distribution among eukaryotes points to a conserved function rather than random loss-of-function mutations (Adrain C. and Freeman M., 2012; Lemberg M. K., 2013). In fact, catalytically dead members are present in most enzyme families (Adrain C. and Freeman M., 2012). The rhomboid pseudoproteases probably resulted from several gene duplications of a number of rhomboid proteases, followed by a loss of activity of one of these, giving rise to several types

of rhomboid pseudoproteases (Adrain C. and Freeman M., 2012) (**Figure 2**). The loss-of-function is the result of either the mutation of the catalytic serine and histidine or the disturbance of the active site geometry by introducing a proline before the catalytic serine (Lemberg M. K. and Freeman M., 2007b; Zettl M. *et al.*, 2011). Rhomboid pseudoproteases have distinct functions, such as control of protein trafficking through the secretory pathway or protein dislocation in the ERAD (Lemberg M. K. and Adrain C., 2016).

1.2.3.1 iRhoms

The mammalian genome encodes two iRhoms, iRhom1 and iRhom2, which have at least partially redundant function in translocation and maturation of an important cell surface metalloprotease (shedase) ADAM17. The expression of the iRhoms correlates with ADAM17; the tissues with strong phenotype in *ADAM17*^{-/-} mouse show higher expression of both iRhom1 and iRhom2 (lung, heart, skin (Jackson L. F. *et al.*, 2003; Peschon J. J. *et al.*, 1998)). iRhom1 is widely expressed at higher levels Throughout a number of tissue types whereas iRhom2 is specifically elevated in macrophages (Christova Y. *et al.*, 2013). iRhoms are located in the ER where they bind to immature ADAM17 (Adrain C. *et al.*, 2012) and translocate it to the Golgi apparatus (GA). ADAM17 is glycosylated in the GA and its propeptide is cleaved by a furin-type proprotein convertase, generating mature ADAM17 that is further translocated to the plasma membrane (Schlondorff J. *et al.*, 2000). At the cell surface, the ADAM17-iRhom complex is stabilised by the interaction with FRMD8 (also called iTAP), which prevents its degradation through the endolysosomal pathway (**Figure 5**) (Kunzel U. *et al.*, 2018; Oikonomidi I. *et al.*, 2018).

Specific mutations in iRhom2 lead to ADAM17 hyperactivation, resulting in increased shedding of EGFR ligands (Maney S. K. *et al.*, 2015). EGFR signalling is a key factor for proliferation and differentiation in keratinocytes, therefore increased activation of EGFR leads to tylosis characterised by keratinocyte hyperproliferation that can result in oesophageal cancer (Blaydon D. C. *et al.*, 2012). ADAM17 is the main shedase of EGFR ligands, but it also catalyses the release of other signalling molecules, such as the primary inflammatory cytokine tumour necrosis factor (TNF) (Black R. A. *et al.*, 1997). Macrophages are the main source of TNF (Adrain C. *et al.*, 2012) and also the cells with the highest iRhom2 expression (Christova Y. *et al.*, 2013). Indeed, macrophages from *iRhom2*^{-/-} mice did not secrete TNF upon lipopolysaccharide stimulation (Adrain C. *et al.*, 2012; McIlwain D. R. *et al.*, 2012) and iRhom2 was shown to be a critical pathogenic mediator of inflammatory arthritis (Issuree P. D. A. *et al.*, 2013). As iRhom1 is not expressed in

hematopoietic cells, iRhom2 presents a potential therapeutic target (Issuree P. D. A. *et al.*, 2013).

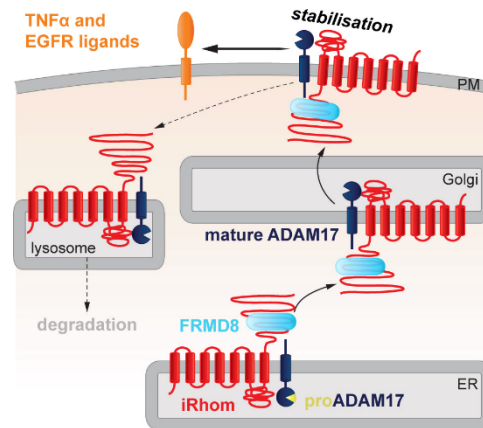


Figure 5: Model of ADAM17-mediated substrate shedding.

The complex of proADAM17-iRhom-FRMD8 is established in the ER. It is then translocated to the GA where ADAM17 is glycosylated and activated by cleavage of its pro-domain by furin protease. Afterwards the whole sheddase complex traffics to the plasma membrane where ADAM17 cleaves its substrates, e.g. membrane bound precursors of TNF, EGFR ligands. iRhom-FRMD8 interaction protects iRhom and ADAM17 from endolysosomal degradation. Adopted from (Kunzel U. *et al.*, 2018).

1.2.3.2 RHBDD3

iRhom2 is not the only rhomboid pseudoprotease involved in innate immunity. RHBDD3 provides negative regulation for a systemic inflammatory response. IL-6 is one of the pro-inflammatory cytokines which is released by dendritic cells upon microbial activation triggered by Toll-like receptors (TLRs) and its overexpression causes autoimmune diseases, e.g. multiple sclerosis and rheumatoid arthritis (Kimura A. and Kishimoto T., 2010). The pathological effect is caused mainly through activation of the NF- κ B pathway by IKK complex (I κ B kinase α and β). Normally, NF- κ B is regulated by one of the I κ B proteins (inhibitors of NF- κ B), which is phosphorylated by the IKK complex, thus activating NF- κ B (Gilmore T. D., 2006). The IKK complex is regulated by IKK γ (also called NEMO). Although the mechanism is poorly understood, it is known that ubiquitinated IKK γ is necessary for NF- κ B activation. RHBDD3 is localised to the endosome where it binds to K27-linked polyubiquitinated chains of IKK γ through its ubiquitin binding cytoplasmic domain (UBD). Afterwards, RHBDD3 is itself ubiquitinated by an unknown E3 ligase, recognises a deubiquitinase A20 and serves as a scaffold for IKK γ -A20 interaction. IKK γ is deubiquitinated by A20 (Ye Y. H. *et al.*, 2004) and NF- κ B signalling (IL-6 production) is blocked (**Figure 6**) (Liu J. *et al.*, 2014).

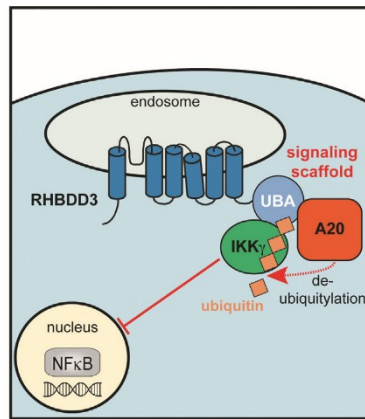


Figure 6: Negative feedback mediated by RHBDD3 on NF-κB signalling

RHBDD3 serves as a scaffold for interaction between regulatory protein IKK γ and a deubiquitinase A20. This pathway thus presents a negative feedback loop for NF- κ B signalling as the deubiquitinated IKK γ is not able to activate NF- κ B. Adopted from (Lemberg M. K. and Adrain C., 2016).

1.2.3.3 Derlins

Derlins (for “degradation in the ER”) were discovered in the budding yeast *S. cerevisiae* where they are required for the degradation of proteins connected to the ER (Knop M. *et al.*, 1996); their mammalian homologues were identified some years later (Lilley B. N. and Ploegh H. L., 2004; Oda Y. *et al.*, 2006; Ye Y. H. *et al.*, 2004). Derlins are involved in ERAD where they interact with misfolded proteins, enabling their delivery to the cytosol by retrotranslocation where these misfolded proteins are ubiquitinated by E3 ligases and targeted for proteasomal degradation (Greenblatt E. J. *et al.*, 2011; Mehnert M. *et al.*, 2014; Neal S. *et al.*, 2018). The two most well characterised derlin homologues are yeast Der1 and Dfm1. They take part in the ERAD-L (for misfolded luminal proteins) (Neal S. *et al.*, 2018) and ERAD-M (misfolded membrane proteins) pathways (Greenblatt E. J. *et al.*, 2011; Mehnert M. *et al.*, 2014), respectively (**Figure 7**). Mammalian genomes encode three derlin homologues, Derlin 1, Derlin 2 and Derlin 3, all of which play a role in ERAD as well (Lilley B. N. and Ploegh H. L., 2004; Oda Y. *et al.*, 2006; Ye Y. H. *et al.*, 2004), but differ in their expression patterns. Derlin1 and 2 are expressed widely whereas Derlin 3 is found only in the spleen, small intestine, placenta and pancreas (Oda Y. *et al.*, 2006). Although there is no direct evidence of derlin involvement in any disease, being involved in ERAD brings a possible association between derlins and cancer. Indeed, expression of all ERAD components was increased in HER2-positive breast cancer (Singh N. *et al.*, 2015) while derlin 3 was found to be silenced in some types of colorectal cancer (Lopez-Serra P. *et al.*, 2014). One of its substrates is the glucose transporter 1 (Glut1), Derlin 3 silencing thus

causes that Glut1 avoids ERAD, which might enable the cancer cell to switch from oxidative metabolism to glycolysis (Lopez-Serra P. *et al.*, 2014).

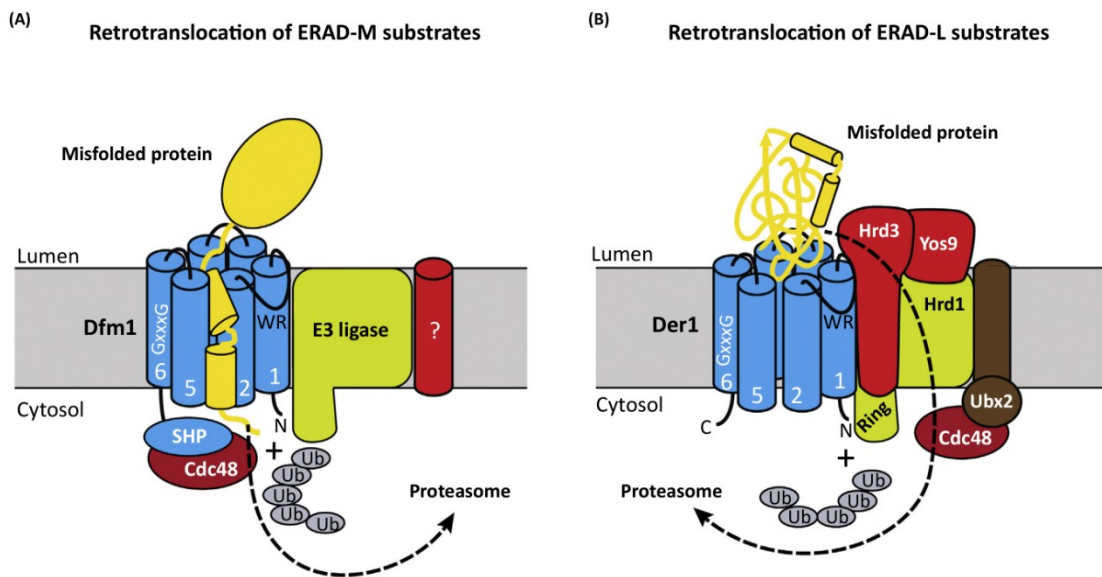


Figure 7: Involvement of derlins in the retrotranslocation of ERAD substrates in yeast

Derlins interact with other proteins to translocate damaged protein across the ER membrane for degradation by the ubiquitin-proteasome system. **A)** The ERAD-M pathway serves for translocation of proteins with defects in their TMDs. Dfm1 interactors are not known yet. **B)** Luminal misfolded proteins are retrotranslocated through the ERAD-L pathway. The complex is formed by Der1, E3 ligase Hrd1 which was proposed to work also as a retrotranslocation channel. Translocated proteins are subsequently ubiquitinated and degraded by the proteasome. Adopted from (Ticha A. *et al.*, 2018).

1.3 Rhomboid protease structure

The current understanding of rhomboid protease mechanism is largely based on enzymological and structural characterization of the *E. coli* rhomboid protease GlpG. The structure of GlpG (solved first time in 2006 as the first structure of any intramembrane protease) confirmed the proposed rhomboid core formed by the 6TMD bundle (Wang Y. *et al.*, 2006). Since then, several groups studied the crystal structures of *E. coli* GlpG and related GlpG from *Haemophilus influenzae* (Ben-Shem A. *et al.*, 2007; Lemieux M. J. *et al.*, 2007; Vinothkumar K. R., 2011; Vinothkumar K. R. *et al.*, 2010; Wu Z. *et al.*, 2006). The structures substantiate the serine-histidine catalytic dyad (Lemberg M. K. *et al.*, 2005) formed by Ser201 and His254 (in *E. coli* GlpG). The catalytic serine lies on the N-terminus of TMD4 which lies about 10 Å below the membrane surface and is encircled by the remaining five TMDs and the amphipatic structure of the L1 loop, which is half-immersed in the membrane as a ‘floatation device’. His254 is located on TMD6 below the level of the membrane. The catalytic dyad lies at the bottom of an aqueous internal cavity open to the

bulk solvent. This transmembrane architecture of rhomboid and its use of catalytic dyad instead of a classical Ser/His/Asp triad clearly indicate evolution independent from that of the soluble serine proteases (Kinch L. N. and Grishin N. V., 2013).

The overall stability of the rhomboid fold is maintained by several distinguishable structural elements (**Figure 8**) (Baker R. P. and Urban S., 2012). Firstly, the ‘bottom’ part of the rhomboid, which faces the cytoplasm, is stabilised by four hydrogen bonds between Glu166 (TMD2) and its neighbouring residues to assemble TMDs 1, 2 and 3. Hydrogen bonds also stabilise the hairpin conformation of the L1 loop, which is important for the overall protein stability (Wang Y. *et al.*, 2006; Wang Y. C. *et al.*, 2007). Mutations in specific sites of the L1 loop lead to a dramatic decrease in GlpG activity (Baker R. P. *et al.*, 2007) which explains the importance of the conserved WR motif (Koonin E. V. *et al.*, 2003). Association of TMD4 and TMD6 is ensured by the GxxxG dimerization motif, which is a hallmark of rhomboid superfamily (Lemberg M. K. and Freeman M., 2007b), and helps identify the 6TMD rhomboid core. Interaction of TMD4 with loop L3 is mediated by side chain packing of large residues. The only parts of the rhomboid fold that do not contribute to the protein stability are TMD5 and its adjacent loops L4 and L5 (Baker R. P. and Urban S., 2012).

Apart from the rhomboid core, most rhomboids possess extramembrane domains, extra helices (PARL-type, secretase-type A rhomboids) or the IRHD (iRhoms). These are dispensable for catalytic activity (Koonin E. V. *et al.*, 2003; Urban S. *et al.*, 2001) but they probably constitute additional interaction sites. For example, the C-terminal cytoplasmic domain of RHBDL4 contains a ubiquitin interacting motif in the cytosolic part that recognises ubiquitinated substrates (Fleig L. *et al.*, 2012). It was also proposed that extramembrane domains might be involved in GlpG dimerisation (Ghasriani H. *et al.*, 2014; Lazareno-Saez C. *et al.*, 2013), but later it was shown that dimerisation was an artefact caused by the detergents used and that the physiologically relevant rhomboid form is monomeric (Kreutzberger A. J. B. and Urban S., 2018). A recent study shows that the cytoplasmic domain (at least for GlpG) is responsible for the protein orientation in the membrane, which alters the speed of diffusion within the lipid bilayer (Kreutzberger A. J. B. *et al.*, 2019).

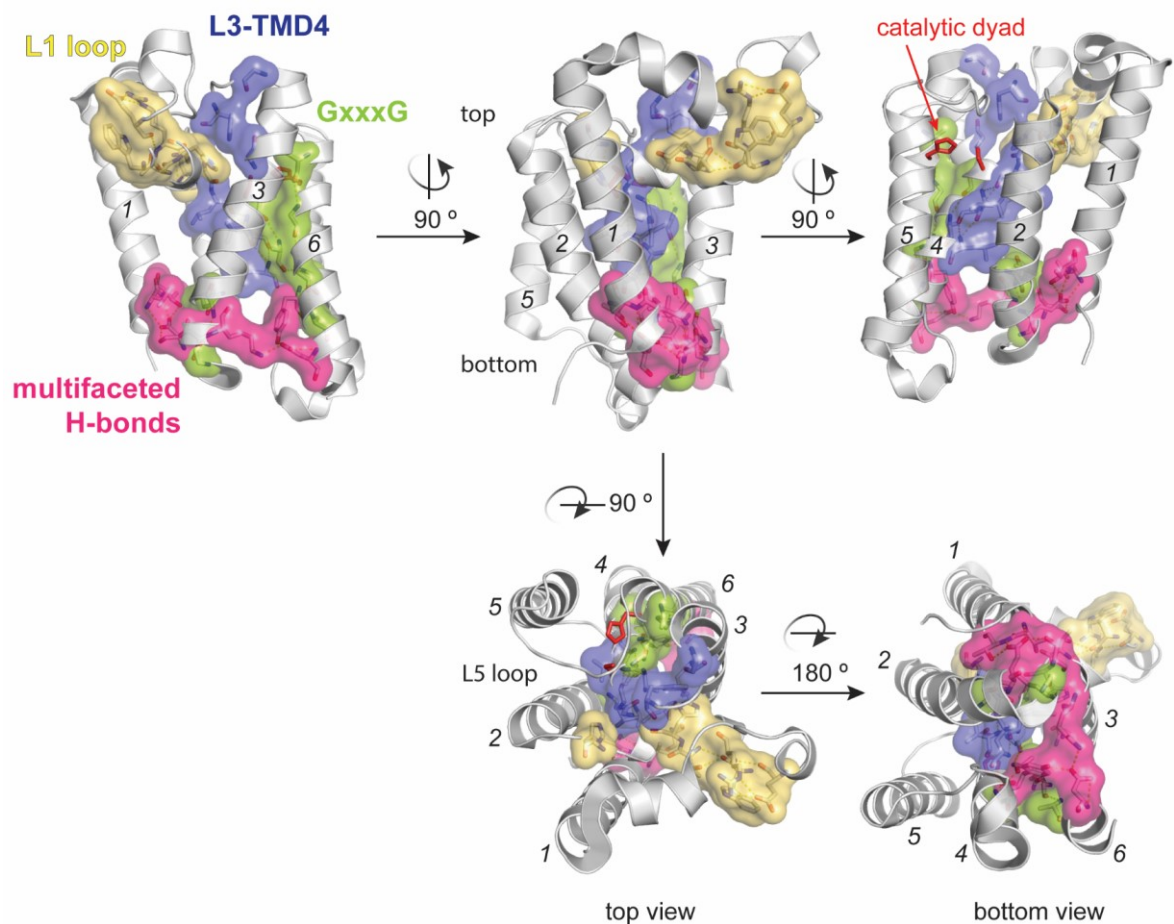


Figure 8: Structure of GlpG with stabilizing residues

GlpG is organised into six transmembrane helices with TMD4 in the middle of the helix bundle. This structure is maintained by several regions. Residues depicted in yellow form hydrogen bonds which stabilise the loop L1 and mediate the interactions with the top residues of TMD2 and TMD3. At the bottom of the helix bundle a cluster of H-bonds keeps the rhomboid core intact (magenta). The central TMD4 interacts with TMD6 through the GxxxG motif (green), and side-chain packing of large residues mediates its interaction with loop L3. The catalytic serine (Ser201) lies on the ‘top’ of TMD4 whereas the catalytic histidine (His254) is located on the TMD6. TMD5 and the corresponding loop L5 are not involved in the stabilising interactions, which supports its dynamic function in substrate recognition. Adopted from (Ticha A. *et al.*, 2018).

The crystal structures also helped to shed light on the source of water required for proteolysis. Water molecules seemed readily available for proteolysis as they were present near the catalytic residues in the active site in all of the crystal structures (Ben-Shem A. *et al.*, 2007; Lemieux M. J. *et al.*, 2007; Wang Y. *et al.*, 2006; Wu Z. *et al.*, 2006). In addition, molecular dynamics simulations showed that most of the water molecules in the active site were rapidly lost over time but for a few that remained within a small area in proximity to the catalytic serine. Specifically, three water molecules occupied a site lined by His141, Ser181, Ser185, and Gln189, leading to the surface of the GlpG molecule, to the bulk solvent. Residues Gln189 and Ser185 were important for rhomboid enzymatic activity as their mutation abrogated cleavage of rhomboid substrate Spitz while the protein stability remained

unaffected (Zhou Y. *et al.*, 2012), therefore concluding that that site serves to facilitate transfer of bulk water into the interior of GlpG and their stabilization in a so-called ‘water retention site’. Structural data of GlpG in complex with peptidyl chloromethyl ketones, substrate derived inhibitors, showed that the water retention site is formed by a cavity contiguous with the GlpG S1 subsite (the site that interacts with the substrate P1 residue, **Figure 9** (Schechter I. and Berger A., 1967)) (Zoll S. *et al.*, 2014).

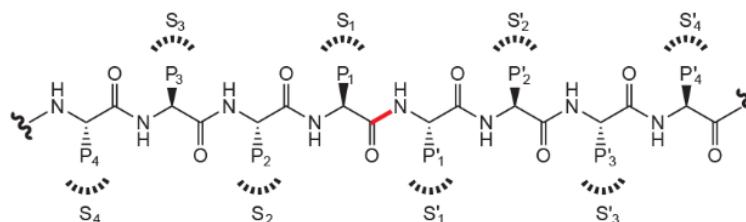


Figure 9: Subsite nomenclature for substrate and enzyme

The hydrolysed bond is depicted in red, the residues on its C-terminus are marked P1, P2 etc, the residues on the N-terminal side P1', P2' etc. Similarly, enzyme residues interacting with P1 are assigned S1 and so on. Adopted from (Strisovsky K., 2013).

1.4 Mechanism of rhomboid protease catalysis

The biological function of an enzyme is closely linked with its mechanism of action and catalysis, whose understanding in turn enables creation of chemical tools that can specifically interfere with the biological function, in order to study or perturb it therapeutically. In the case of rhomboids, our knowledge is based largely on biochemical assays using purified bacterial rhomboids, such as *E. coli* GlpG or AarA from *P. stuartii*, on crystal structures and on molecular dynamics studies. Rhomboid substrates were initially thought to be typically type I or type III transmembrane proteins (single TMD, N-terminus on the extracellular/luminal side, with or without signal peptide, respectively (Higy M. *et al.*, 2004)), which are cleaved near the N-terminus of their TMDs (Urban S. *et al.*, 2002). Surprisingly, rhomboids were shown to cleave also type II proteins (single TMD with cytosolic N-terminus) (Tsruya R. *et al.*, 2007) and the cleavage site can also be localised to the juxtamembrane region close to the membrane (Adrain C. *et al.*, 2011; Fleig L. *et al.*, 2012) or in some cases more distant from the transmembrane domain (Fleig L. *et al.*, 2012; Paschkowsky S. *et al.*, 2016).

Early advances in understanding the catalytic mechanism were made by comparing the ability of different rhomboids to cleave known rhomboid substrates. It was proposed that, to some extent, rhomboids share a common mechanism of substrate recognition because

bacterial rhomboids could cleave some of the eukaryotic substrates even though the sequence homology between the rhomboids that were compared was low (Urban S. *et al.*, 2002). This was later confirmed by other groups (Clemmer K. M. *et al.*, 2006; Gallio M. *et al.*, 2002; Kateete D. P. *et al.*, 2012). Together with the fact that rhomboid seemed to form two distinct groups based on their substrate preferences (Urban S. and Freeman M., 2002), this raised the obvious question of how rhomboids determine their substrate specificity. Another, perhaps less obvious, question was posed by the crystal structures. From those, it was clear that the rhomboid active site lies about 10 Å deep in the membrane which suggests that it is the membrane part of the substrate which is hydrolysed (Wang Y. *et al.*, 2006). In addition, most transmembrane domains are α -helical (Killian J. A. and von Heijne G., 2000; White S. H. and Wimley W. C., 1999), but proteases are expected to bind substrates in extended, β -strand conformation (Hubbard S. J., 1998; Tyndall J. D. A. *et al.*, 2005). So how do substrate recognition and substrate cleavage occur?

1.4.1 Substrate recognition

1.4.1.1 Helix destabilization

Initial theories about substrate recognition assumed that specificity is governed by helix-breaking residues present in the substrate TMD (Urban S. and Freeman M., 2003). Later studies confirmed the presence of helix-destabilising residues (mainly glycines) near the cleavage site (Maegawa S. *et al.*, 2005), but also in the TMD of an artificial LacY^{TM2} chimeric-substrate (the Gln-Pro sequence) (Akiyama Y. and Maegawa S., 2007). All of these residues were previously shown to have a helix-destabilising effect in a non-polar environment (Liu L.-P. and Deber C. M., 1998). Even though these motifs undoubtedly affected the ability of investigated proteins to be a rhomboid substrate, they did not explain the selection of the cleavage site because they were often quite distant from it. Additionally, the mere presence of TMD destabilising residues was insufficient to predict rhomboid substrates. Further research brought the nowadays consensus that helix destabilisation might be important for efficient substrate targeting into the active site (Akiyama Y. and Maegawa S., 2007) and it is necessary only when the cleavage site is located within or very near the TMD, not when it is outside the membrane (Strisovsky K. *et al.*, 2009).

1.4.1.2 Recognition motif

Interactions between the protease and its substrate underlie protease specificity and determine the cleavage site (Perona J. J. and Craik C. S., 1997). Rhomboid proteases do not

recognise a unique primary sequence in a substrate (unlike e.g. trypsin) but there are nevertheless some discernible preferences for specific amino acids around the cleavage site (also sometimes called a 'recognition motif'). Broadly, a small amino acid is required in the P1 position, whereas large residues are preferred in the P4 and P2' positions by a number of diverse rhomboid proteases (Akiyama Y. and Maegawa S., 2007; Strisovsky K. *et al.*, 2009), with some further variations (Zoll S. *et al.*, 2014). Some of these preferences were rationalised by crystal structures of *E. coli* GlpG in complex with peptidyl chloromethyl ketone and β -lactam inhibitors, which revealed binding pockets for the P1, P4 and P2' residues (Vinothkumar K. R. *et al.*, 2013; Zoll S. *et al.*, 2014).

Despite being widespread among rhomboid proteases, this recognition motif cannot be expected to be universal. Not all rhomboids cleave the same substrates (Baker R. P. *et al.*, 2006; Urban S. and Freeman M., 2003) so this might be the defining feature for just one group of them. Mitochondrial rhomboid substrates seem to lack some of the elements of the described motif (Herlan M. *et al.*, 2003; Tatsuta T. *et al.*, 2007), RHBDL2 mediated thrombomodulin cleavage depends on its cytoplasmic domain (Lohi O. *et al.*, 2004) and not all rhomboids can act interchangeably (Kateete D. P. *et al.*, 2012). On the other hand, all rhomboids are probably mechanistically related (Lemberg M. K. and Freeman M., 2007a; Lemberg M. K. and Freeman M., 2007b), which implies the existence of shared principles to determine rhomboid specificity (Strisovsky K. *et al.*, 2009).

1.4.1.3 Lateral gate

One of the big questions in the field has for a long time been the route of substrate access to rhomboid protease. In the absence of any structures of the rhomboid-substrate complex, the initial keys were provided by crystal structures of the rhomboid enzyme alone. These showed the rhomboid active site in the internal cavity sheltered from the environment by the surrounding helices and by loops L1 and L5, leading to the proposal of a gating function for the L1 loop (Wang Y. *et al.*, 2006). Displacement of the loop would open a space for substrate access into the active site from the area between TMD1 and TMD3 (Wang Y. *et al.*, 2006). This was questioned by subsequent structures where the TMD5 and loop L5 were widely displaced (Wu Z. *et al.*, 2006). These two types of structures, were then considered to capture rhomboid protease in a 'closed' and an 'open' state, leading to a proposal of the gating mechanism by TMD5-L5 (Baker R. P. *et al.*, 2007; Wu Z. *et al.*, 2006). This is supported by the fact that whilst the L1 loop is important for the overall protein stability, the TMD5-L5 interface is not involved in the stabilising interactions (Baker R. P.

and Urban S., 2012; Baker R. P. *et al.*, 2007), suggesting its movement might not disturb protein conformation and it could act as a substrate gate. Indeed, mutations increasing TMD5 flexibility increased enzyme activity (Baker R. P. *et al.*, 2007; Urban S. and Baker R. P., 2008).

Molecular dynamics showed that the transition between the closed and the open states in a phospholipid bilayer is slow and the gating could be the rate-limiting step (Zhou Y. *et al.*, 2012). While this model seems to cohere with other observations, there was one study where crosslinking the proposed gate shut by covalently linking TMD2 and TMD5 did not affect enzyme activity (Xue Y. and Ha Y., 2013), which would argue that mobility of TMD5 is not important for rhomboid mechanism. An alternative model was proposed where loop L5 alone could act as a gate, which would involve only subtle conformational changes in the L5 loop and the top of TMD5 (Wang Y. and Ha Y., 2007). Indeed, crystal structures of GlpG-inhibitor complexes confirmed a movement of the L5 loop upon inhibitor binding, but without any movement of TMD5 (Vinothkumar K. R. *et al.*, 2013; Vosyka O. *et al.*, 2013; Xue Y. and Ha Y., 2012; Xue Y. *et al.*, 2012; Zoll S. *et al.*, 2014), even with peptidyl inhibitors that are faithfully capturing half of the bound substrate. Resolving which of the proposed mechanisms is correct will require structural analysis of the complex of a full transmembrane substrate and rhomboid protease. For illustration, comparison of the solved crystal structures is shown in **Figure 10**.

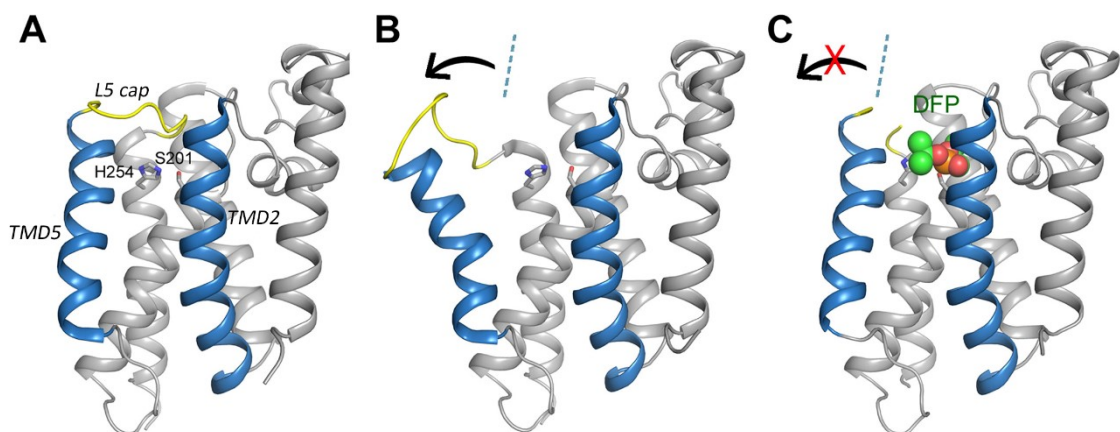


Figure 10: Lateral movement of TMD5

Differences in the positions of TMD5 and L5 loop in crystal structures of GlpG. TMD2 and 5 (blue) were proposed to act as a substrate gate. A) Structure of GlpG in closed conformation (PDB: 2IC8), B) GlpG in the open conformation, the TMD5 is tilted away from the rhomboid core $\sim 35^\circ$ accompanied by L5 loop displacement (PDB: 2NRF), C) Binding of diisopropyl fluorophosphate (DFP) into the rhomboid active site has no effect on the TMD5 conformation, but the L5 loop (yellow) was displaced in the crystal structure (PDB: 4H1D). Adopted from (Xue Y. and Ha Y., 2013).

Rhomboids have low affinity towards their substrates and their proteolysis was postulated to be driven kinetically (Dickey S. W. *et al.*, 2013), where substrate TMD instability (Moin S. M. and Urban S., 2012) and substrate sequence (Strisovsky K. *et al.*, 2009) together contribute to substrate cleavage kinetics (Strisovsky K., 2013). It was even proposed that GlpG patrols the membrane, recognizes the unfolded TMDs and acts as a repair mechanism, although without direct experimental evidence (Dickey S. W. *et al.*, 2013).

1.4.1.4 Effect of the lipid bilayer

The idea that lipids can affect the activity of membrane proteins is not surprising given that lipids form the natural environment of membrane proteins. Membrane composition could change rhomboid activity either by altering membrane properties such as its fluidity and thickness, or lipids can have an impact through direct interaction either with an enzyme or its substrate (reviewed (Lee A. G., 2004). Early experiments showed that GlpG cleaves Spitz when both are overexpressed in mammalian cells (Urban S. and Freeman M., 2003) but not when they are expressed in *E. coli* (Maegawa S. *et al.*, 2005) which was interpreted as an effect of its lipid environment, because the composition differs between mammalian and prokaryotic cells. This was confirmed when the activity and cleavage properties of purified GlpG and several other rhomboids were measured *in vitro* together with the effect of different lipids on its cleavage properties showing that different rhomboids had different requirements for lipids to enhance their activity (Urban S. and Wolfe M. S., 2005). Other experiments showed that the lipid bilayer in close proximity to GlpG is thinner compared to the bulk membrane (Bondar A. N. *et al.*, 2009; Reddy T. and Rainey J. K., 2012; Wang Y. C. *et al.*, 2007; Zhou Y. *et al.*, 2012), which could promote substrate binding (Wang Y. C. *et al.*, 2007). Furthermore, non-substrates can be turned into substrates by changing the membrane properties (Moin S. M. and Urban S., 2012; Urban S. and Moin S. M., 2014). RHBDL4 is hypothesised to contain several binding sites for cholesterol, the binding of which affects the ability of RHBDL4 to cleave APP (Paschkowsky S. *et al.*, 2018). This suggests that lipids can directly affect rhomboid activity but our understanding of this phenomenon, including a possible common mechanism of how lipids affect rhomboid catalysis, is rudimentary. It is also possible that membrane composition only helps to orient the substrates in a way where they can reach the rhomboid active site.

Recently, a fascinating study of membrane involvement in the rhomboid mechanism was published. *Kreutzberger et al.* observed that the rate limiting step in the rhomboid-

mediated catalysis is actually speed of the protease diffusion through the membrane. Their model rationalises some of the so far rather mysterious rhomboid features, such as its irregular shape, importance of the conserved residues in the L1 loop or the function of the cytoplasmic domain. Together these aspects contribute to the distortion of the surrounding lipids, resulting in a thinner hydrophobic belt around the protease that enhances speed of rhomboid diffusion through the membrane. Rhomboid migrates faster than presumed according to the viscosity-imposed speed limit of membrane diffusion, which suggests that it is an adaptation to overcome the catalytic rate limitation in the membrane environment (Kreutzberger A. J. B. *et al.*, 2019).

1.4.1.5 Concluding remarks on substrate recognition

Despite extensive research effort, substrate binding is still poorly understood. According to our current knowledge it probably happens in two steps (Dickey S. W. *et al.*, 2013; Cho S. *et al.*, 2016; Strisovsky K. *et al.*, 2009). There is presumably another interaction site in addition to the active site, the so called “exosite”, where substrate binds in the first instance and this is called “interrogation” (Dickey S. W. *et al.*, 2013) or substrate “docking” (Strisovsky K. *et al.*, 2009). In the second step the substrate part presenting the recognition motif unwinds and bind to the rhomboid active site (Shokhen M. and Albeck A., 2017; Strisovsky K. *et al.*, 2009). The mechanism is summarized in **Figure 11**. Again, the final proof will require structural analysis of rhomboid-substrate complex.

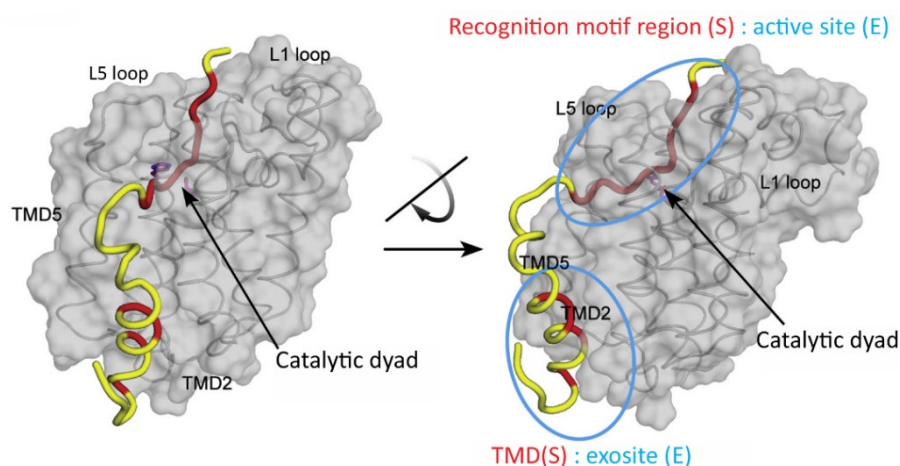


Figure 11: Mechanism of substrate recognition

A model of the rhomboid-substrate interaction created using molecular dynamics simulation (Shokhen M. and Albeck A., 2017). GlpG (grey) interacts with a Gurken-based substrate (yellow). The substrate residues that show highest interaction energy with GlpG are depicted in red. Note that the most intense interactions are concentrated into two regions – the substrate recognition motif interacts with the rhomboid active site and the substrate TMD with the proposed exosite. Adopted from (Ticha A. *et al.*, 2018).

1.4.2 Catalytic mechanism

As mentioned above, most of the soluble serine proteases use a Ser-His-Asp catalytic triad to cleave their substrates. Ser acts as a nucleophile, His is a general base, and Asp is a general acid, which is important for the right orientation of His and also to compensate for the charge that is generated on His in the transition state (Ekici O. D. *et al.*, 2008; Hedstrom, 2002). Rhomboids, on the other hand, only use a catalytic Ser-His dyad, missing the Asp (Lemberg M. K. *et al.*, 2005; Wang Y. *et al.*, 2006). Molecular dynamics simulations and quantum mechanics calculations on *E. coli* GlpG proposed that the absence of a catalytic Asp makes the catalytic histidine (His254) more acidic than the serine (Ser201) hydroxyl group which results in the catalytic serine being protonated in the unbound enzyme state (Uritsky N. *et al.*, 2012). Ligand binding causes lower exposure of the active site to water and its gradual desolvation which leads to an increase of His254 pK_a. This results in proton transfer from Ser201 to His254 and a nucleophilic attack in a concerted mechanism. Similar experiments showed that for classical serine proteases (represented by trypsin) this mechanism is stepwise; the single-step mechanism could help explain the low catalytic efficiency of rhomboids (Uritsky N. *et al.*, 2016).

Crystal structures of GlpG in complex with peptidyl aldehyde inhibitors showed that the negatively charged oxygen of the covalent tetrahedral intermediate is stabilised by an oxyanion hole, formed by hydrogen bonds to the side chains of His150 and Asn154 and to the amide backbone of Ser201 (Cho S. *et al.*, 2016). The tetrahedral intermediate is hydrolysed to produce a covalently bound acyl-enzyme intermediate, which is further hydrolysed to finish the catalytic cycle (**Figure 12**) (Strisovsky K., 2016). GlpG structures also showed that rhomboids attack the scissile bond from the *si*-face (Wang Y. C. *et al.*, 2007) instead of the more commonly used *re*-face (James M. N. G., 1993). In addition to this, the S1 cavity, which combines the water retention site (Zhou Y. *et al.*, 2012) and the P1 interacting site, seems to be able to accommodate residues larger than just alanine or serine, and these could interfere with the water transfer to the active site (Zoll S. *et al.*, 2014). If understood better, these phenomena could facilitate design of potent and selective rhomboid inhibitors.

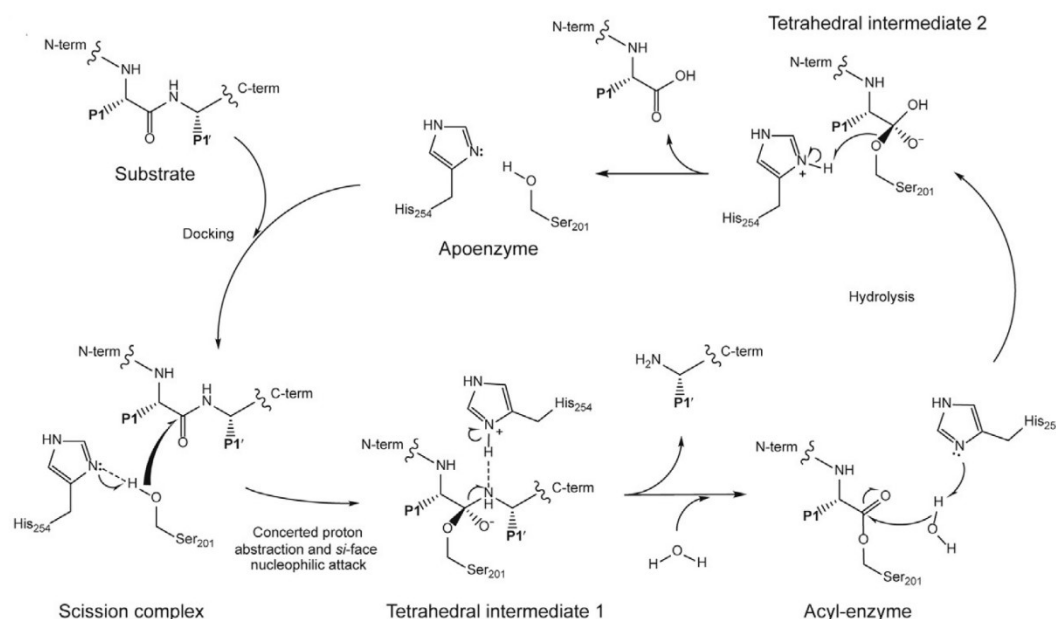


Figure 12: Reaction cycle of rhomboid mediated proteolysis

The apoenzyme interacts with the substrate and upon a local unwinding of the substrate TMD they form a scission complex. This interaction leads to a proton abstraction from the Ser201 by His254 and results in nucleophilic attack from the si-face. The resulting tetrahedral intermediate 1 is hydrolytically cleaved and produces an acyl-enzyme complex which is further hydrolytically cleaved to regenerate the apoenzyme and release the second product. Adopted from (Strisovsky K., 2016).

1.5 Inhibitors of rhomboid proteases

During the last two decades the medical importance of many rhomboids has been confirmed several times (chapter 1.2). Beyond these examples, in general, the investigation of the cell biology of rhomboid proteases would greatly benefit from specific and potent rhomboid inhibitors. As rhomboid activity is necessary for several pathological states such as invasion of malaria parasite into the host cell, such rhomboid inhibitors could be also evaluated for therapeutic use. Although there were several attempts to identify rhomboid inhibitors prior to the current study, only few rhomboid inhibitors were described, none of them fulfilling the above mentioned requirements. They were useful for better understanding of rhomboid-substrate interaction and mechanism (Vinothkumar K. R. *et al.*, 2010; Xue Y. and Ha Y., 2012), but unsuitable for cell biology due to low potency, low selectivity and high toxicity, etc. The current rhomboid inhibitors can be conceptually divided into two groups; i) small non-peptidic, usually heterocyclic compounds and ii) substrate derived inhibitors that combine a peptidyl part with a reactive electrophilic group. The overview of the currently known inhibitors is in the **Figure 13**.

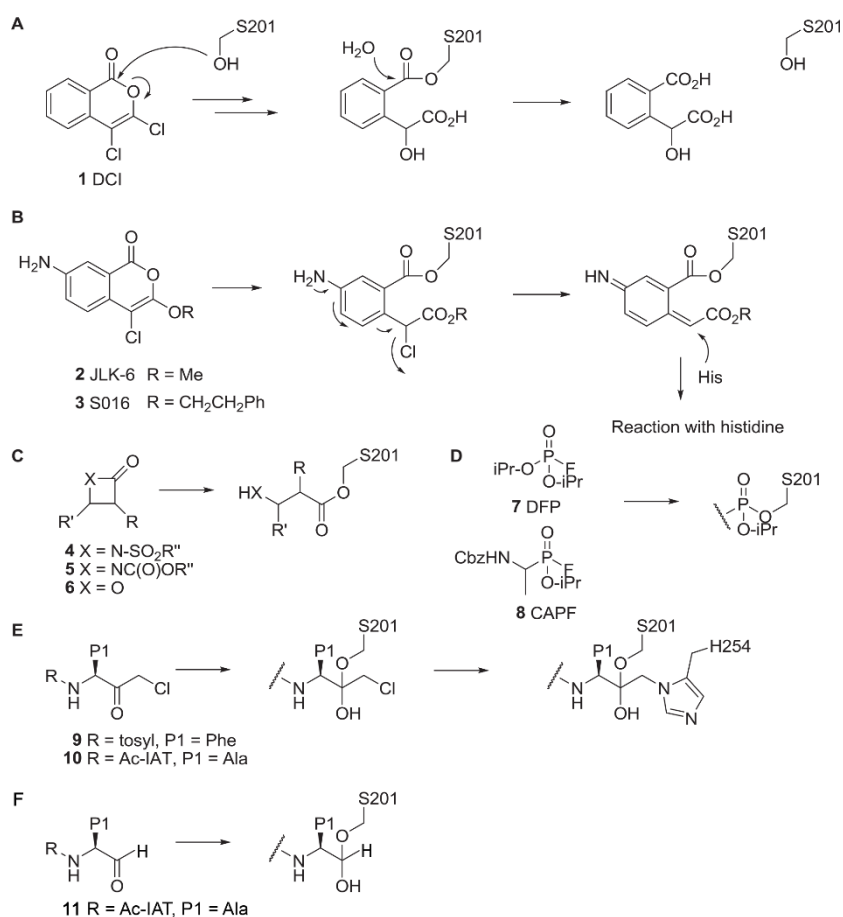


Figure 13: Structure and mechanism of rhomboid inhibitors.

A) 3,4-dichloroisocoumarin (DCI) forms a covalent adduct with catalytic serine that undergoes slow hydrolysis. **B)** Inhibitors based on the 4-chloro-isocoumarin scaffold with a substituent at position 7 form bind to the catalytic serine which leads to a formation of an electrophilic intermediate that further reacts with the catalytic histidine and forms a cross-link to the active site. **C)** Reaction of β -lactam and β -lactones with the catalytic serine results in the opening of the heterocyclic ring. **D)** Diisopropyl fluorophosphate (DFP) and Cbz-AlaP(O-iPr)F (CAPF; Cbz – carboxybenzyl) contain a reactive phosphoryl atom that acts as a leaving group upon reaction with the catalytic serine. **E)** The reaction of chloromethyl ketones with the active site serine results in the formation of hemi-ketal, similar to the tetrahedral intermediate substrate-rhomboid complex, which further alkylates the catalytic histidine. **F)** The nucleophilic attack by the catalytic serine on the peptidyl aldehyde inhibitor results in the tetrahedral complex. Adopted from (Wolf E. V. and Verhelst S. H., 2016).

1.5.1 Small heterocyclic inhibitors

Two broad spectrum serine protease inhibitors, 3,4-dichloroisocoumarin (DCI) and tosyl phenylalanyl chloromethyl ketone (TPCK), were able to inhibit Rhomboid 1 dependent cleavage of Spitz (Urban S. *et al.*, 2001) and indicated that rhomboids were serine proteases. Both of them have low affinity towards rhomboids and as expected DCI was shown to be a “pan rhomboid” inhibitor, being able to inhibit not only Rhomboid 1 but also GlpG, AarA, PARL and many others (Wolf E. V. *et al.*, 2015a). The isocoumarin scaffold was further developed and soon isocoumarins with better potency were described. Two inhibitors, JLK 6 and IC 16, were shown to covalently bind GlpG through two covalent bonds, each of them

using a different mechanism (Vinothkumar K. R. *et al.*, 2010; Vosyka O. *et al.*, 2013). JLK 6 binds to the catalytic Ser201 and afterwards makes another covalent bond to the second catalytic residue, His254, and the rhomboid-inhibitor structure represents the acyl enzyme complex. This confirms that despite independent evolution rhomboid proteases developed a similar mechanism to classical serine proteases (Vinothkumar K. R. *et al.*, 2010). IC 16 also forms two covalent bonds with GlpG, but instead of His254 it binds to His150, suggesting a further role for this residue besides being part of an oxyanion hole (Vosyka O. *et al.*, 2013). IC 16 has higher potency than the other isocoumarins investigated but it lacks selectivity as it inhibits chymotrypsin similarly well as GlpG (Vosyka O. *et al.*, 2013).

A high-through put screening study identified a new class of rhomboid inhibitors, β -lactams, as exemplified by L62 (Fig. 14) (Pierrat O. A. *et al.*, 2011). They are slowly reversible (Pierrat O. A. *et al.*, 2011), covalently bound to the catalytic Ser and occupy the proposed S2' subsite (Vinothkumar K. R. *et al.*, 2013). Their advantage is lower cytotoxicity compared to isocoumarins. The fact that the β -lactam scaffold is already used in the pharmaceutical industry qualifies them as good candidates for potential drug development. Unfortunately, the best candidates do not show high selectivity for rhomboids over chymotrypsin (the only non-rhomboid serine protease tested) and have low efficiency *in vivo* (Pierrat O. A. *et al.*, 2011). Comparison of binding modes of representative isocoumarin and β -lactam is shown in **Figure 14**.

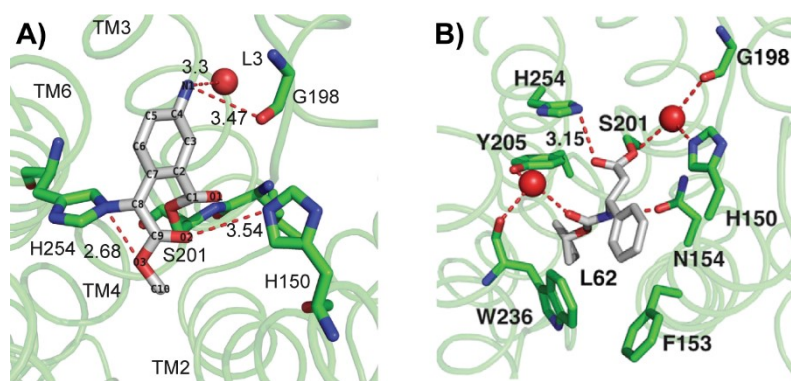


Figure 14: Different binding modes of A) isocoumarin JLK 6 and B) β -lactam L62.

JLK6 and L62 interact with different parts of the rhomboid active site. JLK6 binds to the non-prime site whereas L62 binds to the prime site of the rhomboid. Adopted from (Vinothkumar K. R. *et al.*, 2013; Vinothkumar K. R. *et al.*, 2010).

More recently, several other classes of rhomboid inhibitors were discovered. For example, diisopropylfluorophosphate (DFP), another typical inhibitor of the serine

hydrolases, was shown to inhibit GlpG. The structure of its complex with GlpG is a model for a transition tetrahedral state which precedes the acyl enzyme complex, represented by the GlpG-JLK 6 complex. Based on the differences between these two structures authors hypothesise about the conformational changes that both substrate and enzyme have to undergo (Xue Y. and Ha Y., 2012). The complex of GlpG with CAPF (Cbz-Ala^P(O-*i*Pr)F, Cbz – carboxybenzyl), another organophosphate inhibitor that spans the S' side, confirms that the changes are rather small. This is most interesting for the TMD5 position whose tilt angle changes by about 7.5 ° compared to the unbound state (Xue Y. *et al.*, 2012). This creates doubts as to its function as a lateral gate (chapter 1.5.1.3), at least to the proposed extent, as in the “open” state the top of TMD5 is tilted 35 ° away from the enzyme core compared to the “closed” state (Wu Z. *et al.*, 2006).

β-lactones are the most recent group of heterocyclic rhomboid inhibitors. They bind covalently and irreversibly to the catalytic serine but their potency to inhibit GlpG is rather low (IC₅₀ ~ 40 μM). They are similar to β-lactams although β-lactam electrophilic group is more reactive. They are not specific to rhomboids but their scaffold offers scope for further optimisation (Wolf E. V. *et al.*, 2013).

1.5.2 Substrate-derived inhibitors

The most straightforward approach towards inhibitor design is probably based on mimicking the substrate-enzyme interaction. It usually combines a substrate-like peptidyl part with a reactive electrophilic group. The peptidyl element should ensure inhibitor potency and ideally selectivity whereas the electrophile interacts with the catalytic serine (Drag M. and Salvesen G. S., 2010). Only two groups of substrate-derived inhibitors are known so far, peptidyl chloromethyl ketones (Zoll S. *et al.*, 2014) and peptidyl aldehydes which were published a few years after this project started (Cho S. *et al.*, 2016).

As mentioned above, chloromethyl ketones (CMKs) are known to be rhomboid inhibitors since the identification of Dm Rhomboid 1 as a protease. TPCK and TLCK (N-α-tosyl-L-lysine chloromethylketone), two commercially available CMKs, are weak inhibitors of Dm Rhomboid 1 and YqgP (Urban S. and Wolfe M. S., 2005; Urban S. *et al.*, 2001). The poor inhibitory potency is probably caused by mismatched residues in the P1 position (Phe and Lys) as the S1 subsite requires small residues in P1 position (Strisovsky K. *et al.*, 2009; Vinothkumar K. R. *et al.*, 2010). In 2014, Zoll *et al.* presented the first substrate-derived peptide mimicking rhomboid inhibitors. They are also based on the

chloromethyl ketone reactive warhead but the reactive group is attached to the P4-P1 residues (IATA) from TatA (Zoll S. *et al.*, 2014). TatA is part of a twin-arginine translocase system (Sargent F. *et al.*, 1998) and is a physiological substrate of *P. stuartii* AarA (Stevenson L. G. *et al.*, 2007). IATA-cmk binds to GlpG in a substrate-like manner as mutations in defined positions impaired both substrate cleavage and inhibitory potency of the CMK. Thus, the structure of the GlpG-IATA-cmk complex provides significant information about the rhomboid-substrate interaction as it was the first rhomboid-substrate structure published. It shows which enzyme residues are in contact with the P4-P1 part of a substrate. On the other hand, CMKs form a double-bonded covalent irreversible adduct with the catalytic dyad, hence crosslinking TMD4 and TMD6 (Zoll S. *et al.*, 2014), which could distort the resulting oxyanion hole. This was later confirmed when the structure of GlpG in complex with peptidyl aldehydes was published (Cho S. *et al.*, 2016).

Complex of GlpG-aldehyde inhibitor presents the most probable way of oxyanion stabilisation as all the other structures are complexes with inhibitors that either do not mimic the substrate interaction (Vinothkumar K. R. *et al.*, 2013; Vinothkumar K. R. *et al.*, 2010; Xue Y. and Ha Y., 2012; Xue Y. *et al.*, 2012) or with inhibitors containing a warhead bound to both catalytic residues, thus distorting the catalytic centre (Zoll S. *et al.*, 2014). Peptidyl aldehydes represent reversible inhibitors (Cho S. *et al.*, 2016), which makes them suitable for biological assays (or potential drug design) as irreversible covalent protein modifications might increase their antigenicity and lower their specificity (Drag M. and Salvesen G. S., 2010). However, they are also only weak inhibitors (Cho S. *et al.*, 2016) which means that at the time of starting this project there were no potent and selective inhibitors of rhomboid proteases that could help to better understand their mechanism and function. The differences between interaction of CMK and CHO with GlpG are shown in **Figure 15**.

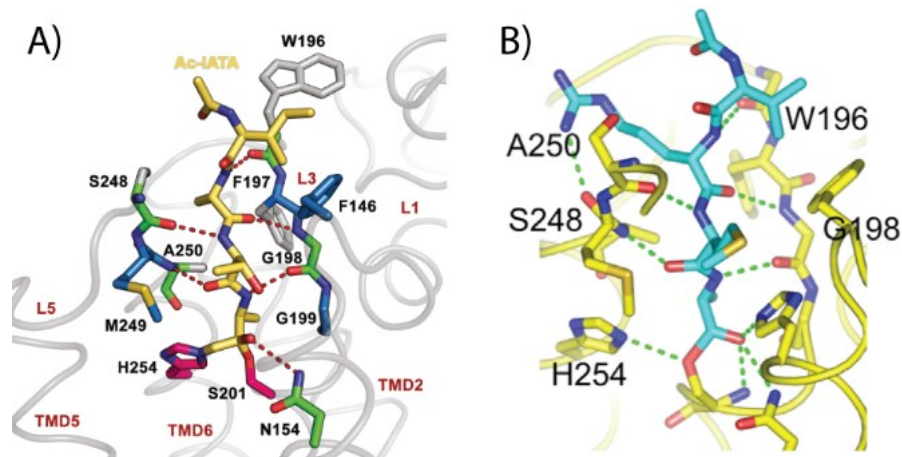


Figure 15: Comparison of interactions between GlpG and A) AcIATA-cmk and B) AcVRMA-CHO

Chloromethyl ketone crosslinks the rhomboid active site while peptidyl aldehyde binds covalently only to the catalytic serine and this complex mimics the structure of the tetrahedral intermediate. The peptidyl parts of both inhibitors make H-bonds to the corresponding GlpG residues. Adopted from (Zoll S. *et al.*, 2014) and (Cho S. *et al.*, 2016), resp.

1.6 Rhomboid protease activity assays

The previous chapters explain that most of our understanding of rhomboid mechanisms and function comes from qualitative experiments, such as genetic manipulations, rhomboid structures alone or in complex with an inhibitor, but our knowledge of the kinetic aspects of their function lags behind. It is mostly caused by the nature of rhomboids, since the fact they are intramembrane proteases makes *in vitro* studies extremely difficult. As both interactors (substrate, enzyme) are hydrophobic, constant presence of detergents and/or lipids is required to maintain their solubility. Thus, assays for determining rhomboid activity are performed either in cells *in vivo* or with purified recombinant proteins, *in vitro*. Here I present an overview of the most widely used methods for determining specific rhomboid activity.

1.6.1 *In vivo* assays

Rhomboid *in vivo* assays are carried out either in bacterial cells or in mammalian cells. It is important to choose the right cell type, since intracellular trafficking and perhaps lipid composition may affect rhomboid activity as demonstrated by different behaviour of GlpG in *E. coli* and in HEK293T cells (Maegawa S. *et al.*, 2005; Urban S. and Freeman M., 2003). In general, *in vivo* assays depend on overexpression of tagged substrate and tagged enzyme (or only tagged substrate if the activity of endogenous rhomboid is measured). For example, LacYTM2 substrate (Akiyama Y. and Maegawa S., 2007; Maegawa S. *et al.*, 2005), Spitz

TMD (Baker R. P. *et al.*, 2007; Moin S. M. and Urban S., 2012; Urban S. and Wolfe M. S., 2005; Urban S. and Baker R. P., 2008) or a chimeric substrate combining MBP (maltose binding protein), substrate TMD and a thioredoxin domain (Adrain C. *et al.*, 2011; Strisovsky K. *et al.*, 2009) have all been used. Eukaryotic rhomboids are usually difficult to prepare recombinantly, so their activity is frequently determined by transient co-transfection of both substrate and rhomboid into mammalian cells (Adrain C. *et al.*, 2011; Baker R. P. *et al.*, 2006; Lemberg M. K. *et al.*, 2005; Lohi O. *et al.*, 2004; Urban S. *et al.*, 2001). The main disadvantage of these steady state assays is their readout as they use SDS-PAGE or WB analysis. Neither of them allows continuous measurement and they are also not suitable for HTS of libraries to identify new inhibitors.

1.6.2 *In vitro* assays

In vitro assays can be classed into several groups, depending on the readout they use – either SDS-PAGE, fluorescence, analysis by mass spectrometry, or the fourth class which are activity-based probes (ABPs). The main disadvantage is the need for purified enzyme and substrate and the above-mentioned need for the presence of detergent or lipids. Solubilisation of membrane proteins with detergents might change their tertiary structure, whereas reactions in liposomes require joint reconstitution of enzyme and substrate.

In vitro assays using gel-based or MALDI readout are based on a similar principle to the previously discussed *in vivo* assays. They usually use the transmembrane part of a known substrate conjugated to several other domains such as the β -lactamase domain and MBP (Maegawa S. *et al.*, 2005) or to green fluorescent protein (GFP) (Urban S. and Wolfe M. S., 2005) or a chimeric substrate combining MBP (maltose binding protein), substrate TMD and a thioredoxin domain (Adrain C. *et al.*, 2011; Strisovsky K. *et al.*, 2009; Zoll S. *et al.*, 2014). The use of full-length natural substrates is limited, as mammalian substrates are usually poorly expressed and the only known bacterial substrate is TatA (Arutyunova E. *et al.*, 2017). The main disadvantage of both methods is that they are quite laborious and offer only end-point data, not continuous measurement of enzyme activity.

On the other hand, fluorescence-based assays allow continuous measurement and they can detect small amounts of product as they have higher sensitivity (Arutyunova E. *et al.*, 2017). Fluorescent substrates are frequently used for determining protease activity (Huang Z. J., 1991). Three types of fluorescent substrates have been employed in studies of rhomboids. Firstly, it is the TatA TMD bound to fluorescein isothiocyanate (FITC).

TatA-FITC can only be used in liposomes where the FITC fluorescence is quenched and is only restored upon rhomboid cleavage which releases the dye into the solvent (Dickey S. W. *et al.*, 2013). The other two substrate types both depend on the principle of Förster resonance energy transfer (FRET) between two fluorophores, the donor and the acceptor. The efficiency of FRET depends on the distance of the fluorophores (inversely proportional to the sixth power of the distance between them), and therefore the closer the fluorophores are, the more efficient FRET is. KSp21 is a peptidyl substrate that is derived from the Gurken TMD and uses the FRET pair Chromis-645/QSY21. It is suitable for HTS but has limited solubility which precludes determination of kinetic parameters of the reaction. It is also cleaved only very poorly by GlpG and other rhomboids, with the exception of AarA (Pierrat O. A. *et al.*, 2011). The last substrate group is similar to KSp21 but it uses GFP and its derivatives instead of small fluorophores, such as cyan or yellow fluorescent proteins (CFP and YFP, resp.) (Arutyunova E. *et al.*, 2014). In both cases one member of the FRET pair is located at either side of the cleavage site, and thus an increase in fluorescence is observed upon rhomboid cleavage (Arutyunova E. *et al.*, 2014; Pierrat O. A. *et al.*, 2011). A disadvantage of the third group of substrates is its use of fluorescent proteins, which makes it dependent on expression levels and limits their photochemical variability. As red-shifted fluorescence substrates are preferred for HTS, this makes them unsuitable for such purpose (Simeonov A. *et al.*, 2008).

Finally, ABPs are based on an electrophilic residue connected via a spacer to a reporter tag, usually a fluorophore or biotin. ABPs often target a large number of mechanistically related enzymes as they recognise conserved features in the active site. Several ABPs were described, which enable rhomboid activity to be measured, using fluorescence polarization, fluorescence microscopy or in gel-based assay as a readout (Sherratt A. R. *et al.*, 2012; Vosyka O. *et al.*, 2013; Wolf E. V. *et al.*, 2015b; Wolf E. V. *et al.*, 2013). As part of this study we established a collaboration with Daniel Bachovchin's research group (Memorial Sloan Kettering Cancer Centre, New York) to test the selectivity of our compounds. They use the EnPlex method for HTS of serine hydrolase inhibitors (**Figure 16**). Here, purified enzymes are coupled to differently-coloured Luminex beads and treated with the compound being tested. Afterwards the bead complexes are incubated with the ABP (biotin-labelled) and a streptavidin R-phycoerythrin conjugate (SAPE). The detection combines signals from two lasers, the first of which recognises the bead colour (the enzyme) and the second one the phycoerythrin (enzyme activity). A higher phycoerythrin signal means that the inhibitory

activity of the tested compound is lower, and vice versa. The screen contains about 100 human serine hydrolases (Bachovchin D. A. *et al.*, 2014).

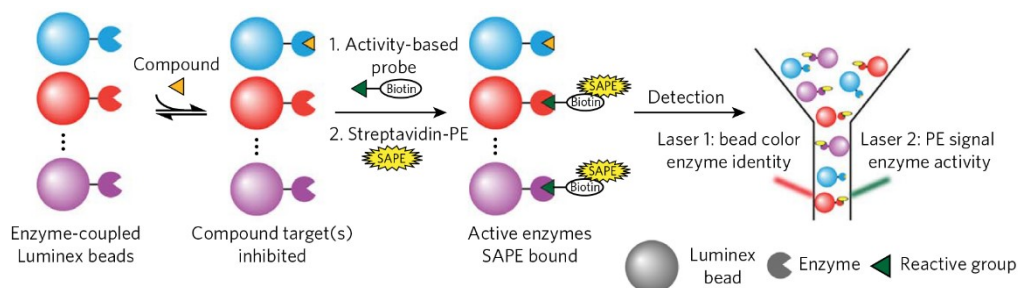


Figure 16: EnPlex method

The compound under study is incubated with purified enzyme coupled with the Luminex bead. The complex is subsequently treated with the ABP and SAPE to determine the compound activity by a two-laser system. The signal from the first laser identifies the enzyme (Luminex bead signal) and the signal from the second one the enzyme activity (phycoerythrin signal). Adopted from (Bachovchin D. A. *et al.*, 2014).

2 AIMS OF THE STUDY

In the Introduction of my thesis, I have highlighted the importance of the rhomboid superfamily, and the hitherto pronounced lack of chemical tools to understand their function in detail. Therefore, the general aim of my study was to expand the range of chemical tools and methods available for studying rhomboid proteases.

In vitro characterisation of the rhomboid protease mechanism is hampered by the lack of easily obtainable, wide-selectivity or bespoke rhomboid substrates, which would allow continuous measurement of reaction progress. Therefore, the initial aim of my project was the development of a new and better substrate and assay platform. As fluorescence intensity remains the most easily accessible and accurate readout, the substrate was to be based on FRET. The goal of this project was to design new fluorogenic rhomboid substrates that would be recognised by several rhomboids and could be used in detergent micelles as well as in proteoliposomes. To be suitable for HTS, the substrate needed to utilise red-shifted variants of fluorophores.

However, the main focus of my thesis is development of potent and selective rhomboid inhibitors. Prior to the start of this work (as summarised in chapter 1.5), no rhomboid inhibitors were available that would meet all the requirements for inhibitors suitable for cell biological assays. The main objective was to deliver a general strategy for rhomboid inhibitor design based on known rhomboid-substrate interactions. The second objective was the characterisation of novel classes of small-molecule rhomboid inhibitors.

3 METHODS

This chapter summarises the methods I have used in the publications presented. My contribution to each publication is specified at the end of the related chapter in the Results section. The work was carried out at the Institute of Organic Chemistry and Biochemistry of the CAS. As part of these projects we established collaborations with several research groups. First part of my thesis are projects that originated in and were led by our lab (Publication 1 and 2). Crystal structures of GlpG in complex with peptidyl α -ketoamides were solved partly in collaboration with Kutti R. Vinothkumar (Richard Henderson's group at Medical Research Council Laboratory of Molecular Biology, Cambridge, UK), the selectivity of inhibitors against a panel of rhomboids was tested in the laboratory of Steven Verhelst (University of Leuven, Leuven, Belgium) and the selectivity against human hydrolases using the EnPlex method in the lab of Daniel A. Bachovchin (Memorial Sloan Kettering Cancer Center, New York, USA). The second part of my thesis is the characterisation of the compounds developed by Sascha Weggen's group (Heinrich Heine University Düsseldorf, Düsseldorf, Germany). They synthesised all the saccharin and benzoxazinone inhibitors, did the docking studies and preliminary activity testing.

This is a brief list of the methods used in this thesis, detailed descriptions are within the publications presented below.

Molecular biology methods:

Transformation of plasmid into *E. coli* cells; expression of recombinant proteins in *E. coli*; isolation and purification of proteins of interest.

Characterisation of substrates and inhibitors:

Measurement of peptide solubility in different conditions; acquisition of fluorescent spectra; measurement of reaction rates and enzyme activity using i) a continuous fluorescence-based assay (following an increase in fluorescence intensity upon fluorescent substrate cleavage), and ii) an end-point assay (following the difference in the amount of substrate hydrolysed after the equivalent time under different conditions); sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE); immunoblot analysis (detection by chemiluminescence or fluorescence); characterisation of substrate and inhibitor properties (K_m , k_{cat} , IC_{50} , K_i , reversibility and modality of inhibition); inhibition of endogenous GlpG in *E. coli*.

4 RESULTS

During my PhD studies I co-authored six publications (three as a first author) describing tools to manipulate rhomboid activity, all of them in peer-reviewed journals. Here I present and discuss four of these; my contribution to *Yang et al.* is rather small and the final paper is a review of the different approaches to study rhomboid proteins that we were asked to write after publishing the paper about peptidyl ketoamides. Our work on peptidyl ketoamides also resulted in an international patent. All of the presented work is the result of joint efforts of several people and my contribution to the publications is summarised at the end of each chapter.

4.1 Publication 1

Sensitive versatile fluorogenic transmembrane peptide substrates for rhomboid intramembrane proteases

Anežka Tichá, Stancho Stanchev, Jan Škerle, Jakub Began, Marek Ingr, Kateřina Švehlová, Lucie Polovinkin, Martin Růžička, Lucie Bednářová, Romana Hadravová, Edita Poláchová, Petra Rampírová, Jana Březinová, Václav Kašička, Pavel Majer, Kvido Stříšovský

Previous assays to determine rhomboid protease activity suffer from several disadvantages discussed in chapter 1.6. Briefly, these are as follows: gel-based readout is unsuitable for HTS (Maegawa S. *et al.*, 2005; Urban S. and Wolfe M. S., 2005), and substrates have low affinity towards different rhomboids (Pierrat O. A. *et al.*, 2011), their use is limited to liposomes (Dickey S. W. *et al.*, 2013), or the presence of large fluorescent domains constrains their photochemical variability (Arutyunova E. *et al.*, 2014). Here we present new fluorogenic substrates based on LacYTM2 peptide with the Edans-Dabcyl or the Tamra-QXL610 fluorophore FRET pairs.

We have chosen LacYTM2 as we know it is a substrate for a range of different rhomboids. Almost the entire TMD is required for efficient cleavage by GlpG (**Figure 17**), so we did not truncate the peptide and we introduced Glu-Edans at the P5 and Lys-Dabcyl at the P4' sites, positions where mutations do not affect recognition by rhomboid (Strisovsky K. *et al.*, 2009; Zoll S. *et al.*, 2014). The resultant peptide is a new rhomboid fluorogenic substrate, KSp35. We have characterised the peptide in terms of its secondary structure, fluorescence spectra, cleavage site, pH optimum and solubility. For KSp35 to be soluble, the presence of detergent in buffer was required. Importantly, the cleavage rate also strongly

depended on detergent concentration (**Figure 18**), which might have been neglected in most reports to date.

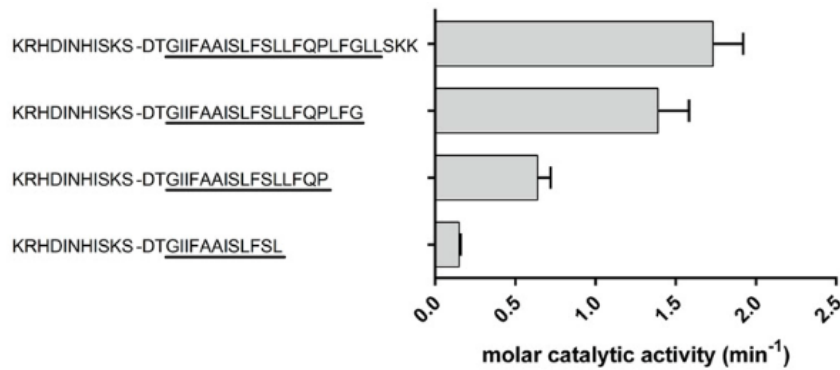


Figure 17: The substrate TMD is important for recognition and cleavage by GlpG

C-terminally truncated peptides were prepared and their cleavage by GlpG was measured. The reaction was performed in buffer containing 0.05 % (w/v) n-dodecyl β -D-maltoside (DDM); 250 μ M substrate was hydrolysed by 2.6 μ M GlpG, samples were taken every 15 min for 2 h, the amount of a formed product was followed by capillary electrophoresis. Removal of as few as 5 amino acids from their TMD caused a reduction in their processing.

We have also shown that KSp35 is hydrolysed by a panel of rhomboids and that it can be used in liposomes. Its affinity towards different rhomboids can be changed by mutation of the recognition motif. Mutations in P5-P1 (RVRHA) to amino acids preferred by GlpG (Zoll S. *et al.*, 2014) increased its catalytic efficiency about 23-times and had a striking effect on selectivity as the mutated substrate was cleaved only by GlpG (**Figure 19**). This suggests

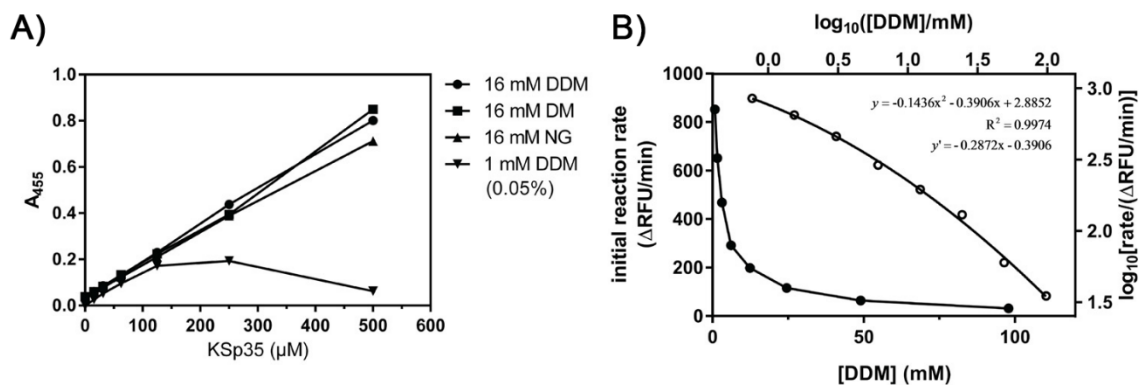


Figure 18: Effect of detergent concentration on substrate solubility and on the reaction rate

A) KSp35 solubility depends on the detergent concentration in the assay buffer. In 1 mM n-dodecyl β -D-maltoside (DDM) (standard assay conditions) it is soluble only to 100 μ M whereas in 16 mM detergent it is soluble up to 500 μ M (higher concentrations were not tested). The KSp35 peptide, at different concentrations, was dissolved in the buffer being tested from 10 mM stock solution in dimethyl sulfoxide (DMSO), incubated for 2 h at 37 $^{\circ}$ C, and centrifuged at $21,130 \times g$ for 20 min. Its solubility was evaluated by measuring the absorbance of the supernatant ($\lambda = 455$ nm). **B)** The effect of detergent concentration on the reaction rate. 10 μ M KSp35 was cleaved by 0.4 μ M GlpG in the reaction buffer with varying concentrations of DDM. The reaction rate was determined as an increase in fluorescence ($\lambda_{\text{ex}} = 335$ nm, $\lambda_{\text{em}} = 493$ nm). Each graph shows representative measurements of three independent experiments.

an easy route to increased substrate selectivity. The KSp35 peptide can also be turned into a red-shifted variant with the Tamra-QXL610 fluorescence pair which makes it suitable for HTS (the resultant substrate is KSp76 as the WT peptide, and KSp64 as the RVRHA mutant). From previous work, it is known that red-shifted fluorophores are suitable for characterisation of inhibitors that absorb in the UV region, such as isocoumarins (Vosyka O. *et al.*, 2013). We based our method on conditions similar to those previously published (Pierrat O. A. *et al.*, 2011; Vosyka O. *et al.*, 2013; Zoll S. *et al.*, 2014) using KSp76 and obtained a comparable IC_{50} for all inhibitors.

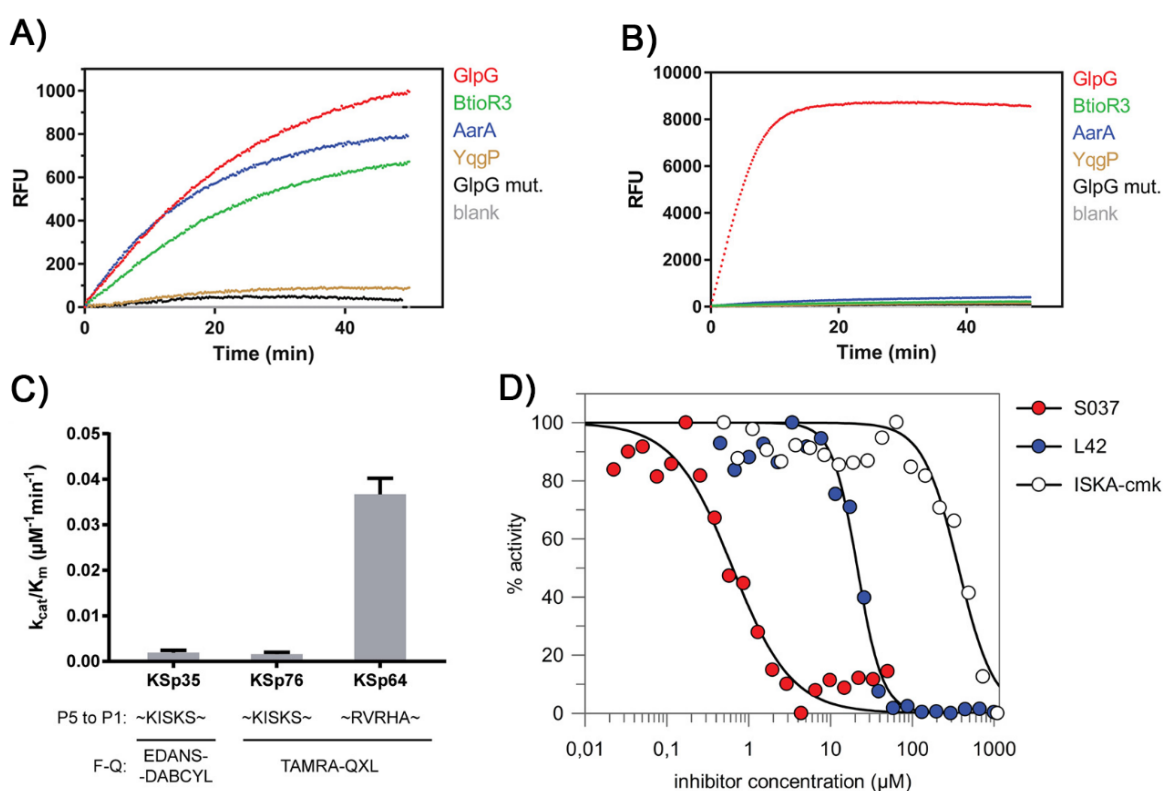


Figure 19: The recognition motif (P5-P1) is important for substrate turnover and specificity

A) Cleavage of KSp76, the red-shifted variant of KSp35 with wild type sequence in the P5-P1 region. KSp76 is cleaved by GlpG, BtioR3 and AarA but not by YqgP and an inactive GlpG mutant (S201A/H254A). **B)** Cleavage of KSp64, the red-shifted variant of KSp35 with the P5-P1 sequence optimised for GlpG cleavage (RVRHA). In contrast to KSp76, KSp64 is cleaved by GlpG only, thus the mutation drastically improved specificity towards GlpG. Moreover, the catalytic rate is higher than for KSp76 cleavage. In both cases 10 μM substrate was cleaved by 0.4 μM enzyme (0.04 μM in case of BtioR3). **C)** Catalytic efficiency (k_{cat}/K_m) for the hydrolysis of KSp35, KSp76 and KSp64 by GlpG. Substrates (0.5-20 μM) were incubated with 0.4 μM GlpG in the presence of 0.5 % (w/v) DDM. Initial reaction rates were plotted against substrate concentration to obtain k_{cat}/K_m . **D)** Inhibition of GlpG by three previously published inhibitors, using KSp76 as a substrate. Inhibitors were pre-incubated with GlpG for 60 min. All experiments were performed in 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05 % (w/v) DDM and 10 % (v/v) DMSO at 37 °C unless noted otherwise. The readout for reaction progress was an increase in fluorescence (KSp35: λ_{ex} = 335 nm, λ_{em} = 493 nm; KSp76 and KSp64: λ_{ex} = 553 nm, λ_{em} = 583 nm).

My contribution to this publication:

- expression and purification of GlpG (both WT and mutant) and chimeric protein substrate MBP-LacYTM2-thioredoxin-His₆
- sample preparation for acquisition of CD spectra and for MS analysis
- measurement of fluorescence spectra; peptide solubility studies; determination of the pH optimum and the effect of detergent concentration on the cleavage rate
- cleavage of the fluorogenic substrates by different rhomboids (fluorescence readout) and cleavage of the chimeric protein substrate (gel-based end point assay)
- determination of k_{cat}/K_m for cleavage of different substrates by GlpG.

4.2 Publication 2

General and modular strategy for designing potent, selective, and pharmacologically compliant inhibitors of rhomboid proteases

Anežka Tichá, Stancho Stanchev, Kutti R. Vinothkumar, David C. Mikles, Petr Pachel, Jakub Began, Jan Škerle, Kateřina Švehlová, Minh T.N. Nguyen, Steven H.L. Verhelst, Darren C. Johnson, Daniel A. Bachovchin, Martin Lepšík, Pavel Majer and Kvido Strišovský

Potent and selective rhomboid protease inhibitors are highly desirable to enable detailed characterisation of rhomboid function as well as for their potential therapeutic use (e.g. in malaria). However, none of the rhomboid inhibitor groups known so far (e.g. isocoumarins, β -lactams/lactons, peptidyl chloromethyl ketones, peptidyl aldehydes, summarised in chapter 1.5) achieve all the desired properties needed for such inhibitors – potency, specificity and non-toxicity. Therefore, based on our knowledge of substrate-rhomboid interactions (chapter 1.4), we designed a general strategy for rhomboid inhibitor design, demonstrating the principle using GlpG protease.

Based on the knowledge that interactions with both the recognition motif (Akiyama Y. and Maegawa S., 2007; Strisovsky K. *et al.*, 2009) and the TMD are important (Ticha A. *et al.*, 2017) we divided the design into three parts. To most efficiently mimic the substrate, we designed inhibitors that have a peptidyl part to mimic the recognition motif, then a reactive electrophilic group to interact with the catalytic serine (the so-called “warhead”), and thirdly a prime site hydrophobic substituent that could be extended further if needed. The mutations introduced into the fluorogenic substrate in Publication 1 (Ticha A. *et al.*, 2017) proved to be also effective in chloromethyl ketone inhibitors. We modified a parent compound AcIATA-cmk (Zoll S. *et al.*, 2014), a scaffold in which the individual mutations

in P5-P1 had additive effect. Thus, we used the RVRHA sequence to form the peptidyl part of the new inhibitor. Next, we screened several warheads and out of the six we tested we chose the α -ketoamide group. Other substituents can be attached to it through the ketoamide nitrogen (Chatterjee S. *et al.*, 1999; Liu Y. *et al.*, 2004b) and the α -ketoamide group is already used as a pharmacophore (Njoroge F. G. *et al.*, 2008). These features might thus be useful for improving potency and non-toxicity in inhibitors, respectively. The final step in the design was optimisation of the prime site hydrophobic substituent. We tested a panel of

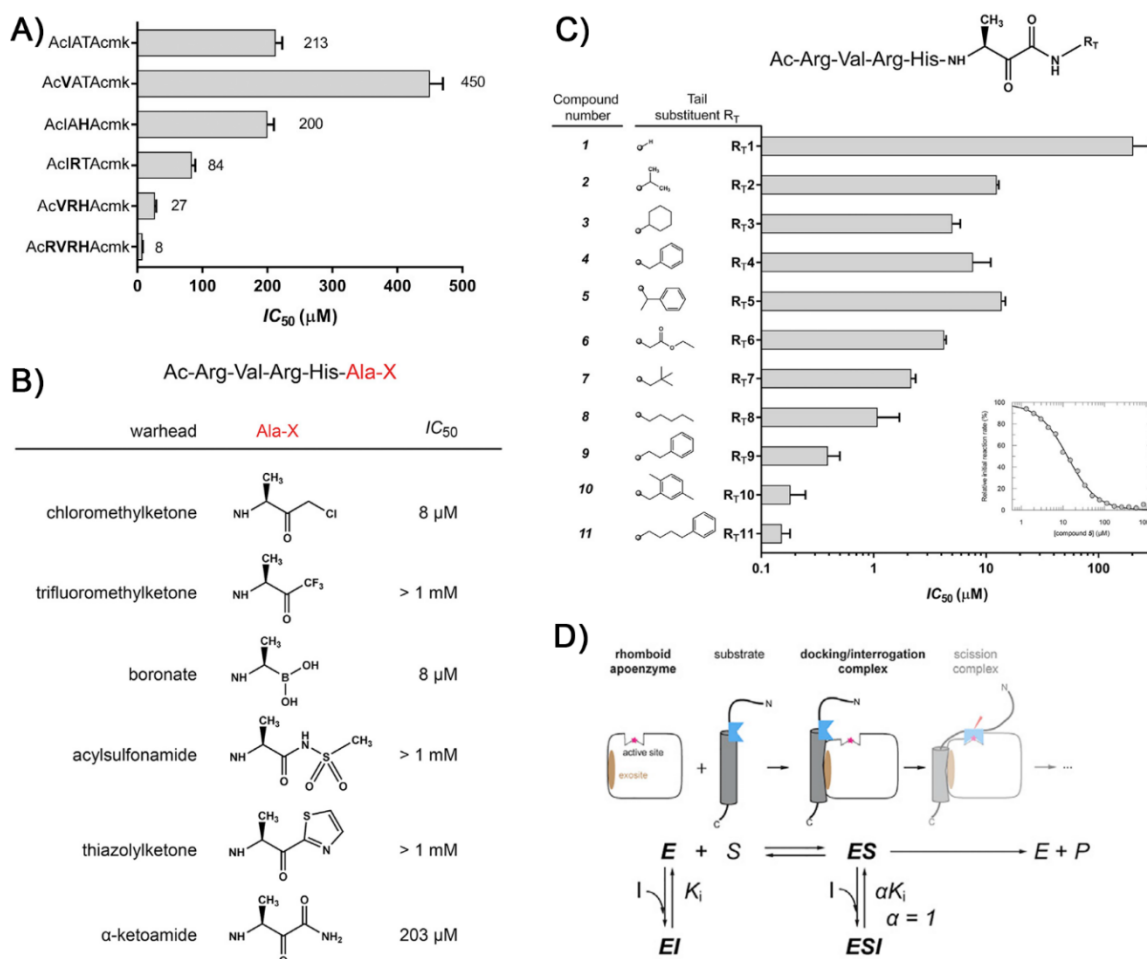


Figure 20: Design of novel rhomboid inhibitors

A) Mutations of the parent compound AcIATA-cmk have an additive effect on inhibitor potency. **B)** Screen of several warheads. α -ketoamide group was chosen as it showed the highest potency out of those that can be extended to the prime site. **C)** Hydrophobic prime site substituents have a dramatic effect on inhibitor potency, the optimal substituent improved IC_{50} 1000-fold. The enzyme concentration in the assay was 0.4 μM , so we reached the lower limit of the assay. All experiments were performed in 20 mM HEPES pH 7.4, 150 mM NaCl, 10 % (v/v) DMSO and 0.05 % (w/v) DDM at 37 °C. GlpG was pre-incubated with each inhibitor for 60 min and the reaction was started by addition of 10 μM KSp35. The readout for initial reaction rates was an increase in fluorescence ($\lambda_{ex} = 335$ nm, $\lambda_{em} = 493$ nm) during the first few minutes of measurement (below 5 % substrate conversion). IC_{50} values were determined from the dependence of initial rate on inhibitor concentration (the inset in C)). The IC_{50} values shown here are means \pm SD of best fits of 2-3 measurements. **D)** α -ketoamides act as non-competitive inhibitors. As such, they can bind to both free enzyme and enzyme-substrate complex (in this case the docking/interrogation complex), without affecting substrate affinity towards GlpG ($\alpha = 1$).

11 different substituents and the most potent one identified was the longest, phenylbutyl (**Figure 20**).

We determined the kinetic constants and the mechanism of inhibition for the three most potent compounds (9, 10 and 11). The K_i of the best inhibitor (cpd11) was ~ 50 nM. Its higher potency compared to the other compounds tested is probably the result of slow dissociation of enzyme-inhibitor complex ($k_{on} > k_{off}$, whereas for the others they are almost the same). All of them are reversible non-competitive inhibitors that covalently bind to GlpG in a substrate-like manner occupying the S4-S2' sites, and efficiently inhibit endogenous rhomboids in living cells with an IC_{50} value of 2.7 ± 0.1 nM for the best compound. We also tested their selectivity and ability to discriminate between several rhomboids using established methods (**Figure 21**) (Wolf E. V. *et al.*, 2015a), as well as their ability to inhibit other human enzymes using the EnPlex assay that includes about 100 human serine hydrolases (Bachovchin D. A. *et al.*, 2014). Our inhibitors showed only partial selectivity between tested rhomboids, but they do not inhibit most of the human hydrolyses tested.

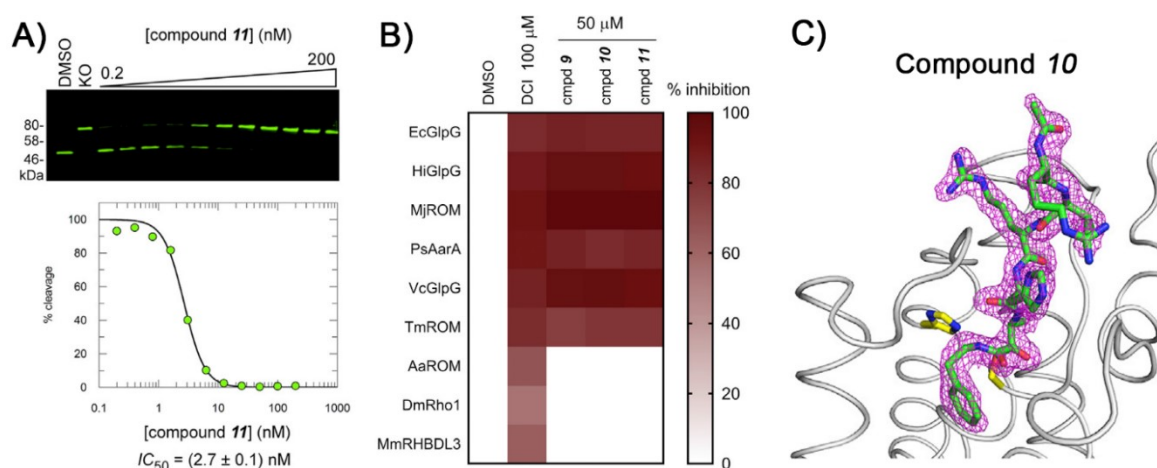


Figure 21: Characterisation of α -ketoamides as rhomboid inhibitors

A) Inhibition of endogenous GlpG mediated cleavage of MBP-FLAG-LacYTM2-Trx substrate (Strisovsky K. *et al.*, 2009) overexpressed in live *E. coli* NR698 with genetically permeabilised outer membrane (Ruiz N. *et al.*, 2005). Substrate cleavage was monitored in cell lysates and quantified by near-infrared fluorescence based immunoblotting (α -FLAG). A representative immunoblot and inhibition curve are shown. The IC_{50} is a mean \pm SD of best-fits from 2-3 independent measurements. KO is a negative control ($\Delta glpG$ NR698). B) Selectivity of three most potent compounds towards nine rhomboid proteases. DMSO and DCI are included as negative and positive controls, respectively. The assay is based on competition with an activity-based probe for the rhomboid active site (Wolf E. V. *et al.*, 2015a). C) Crystal structure of the complex GlpG-cpd10. GlpG crystals were soaked with cpd10, the structure of the complex was solved by X-ray diffraction. The inhibitor binds covalently to Ser201 and occupies the S4-S2' subsite of GlpG.

My contribution to this publication:

- all kinetics and inhibition measurements *in vitro*
- *in vivo* assay in *E. coli*
- expression and purification of GlpG

4.3 Publication 3

Discovery and biological evaluation of potent and selective N-methylene saccharin-derived inhibitors for rhomboid intramembrane proteases

Parul Goel, Thorsten Jumpertz, David C. Mikles, Anežka Tichá, Minh T. N. Nguyen, Steven Verhelst, Martin Hubalek, Darren C. Johnson, Daniel A. Bachovchin, Isabella Ogorek, Claus U. Pietrzik, Kvido Stříšovský, Boris Schmidt, and Sascha Weggen

An alternative approach for rhomboid inhibitor identification was chosen by Sascha Weggen's group (Heinrich Heine University Düsseldorf, Düsseldorf, Germany), a project on which we were invited to collaborate. They took N-methylene-substituted saccharin (1,2-benzisothiazol-3-one 1,1-dioxide) as a starting template for inhibitor development. Saccharins are known to i) act as suicide-inhibitors (Powers J. C. *et al.*, 2002; Walker B. and Lynas J. F., 2001), ii) be easily modifiable and iii) have safe toxicological profile (Hlasta D. J. *et al.*, 1995). The inhibitor design strategy was based on a computer-aided approach, which screened the candidate molecules before chemical synthesis and again during the *in vitro* screening (**Figure 22**).

N-acylsaccharins are known inhibitors of human leukocyte elastase (HLE) and chymotrypsin (Zimmerman M. *et al.*, 1980). The Weggen study functionalised saccharin inhibitors of a “next generation” (Groutas W. C. *et al.*, 1996) by attaching a range of different leaving groups (LG) to the saccharin scaffold through the nitrogen of the heterocyclic ring. Initial *in silico* docking studies tested 50 different LGs and selected the 10 most promising candidates for *in vitro* testing. The most promising molecules were saccharins with aryl carboxylic acid or aryl sulphides and sulfones LGs. Inhibitors with an aryl carboxylic acid LG showed a ~25x higher inhibitory potency than the remaining compounds, where inhibition was only observed at 10 μ M. These findings were used for a second round of docking studies, now with 103 saccharin derivatives out of which 20 were subsequently synthesised and tested *in vitro*. The best analogues inhibited GlpG with an IC_{50} of 200 nM *in vitro* and a similar effectivity *in vivo* (in *E. coli*) (**Figure 23**).

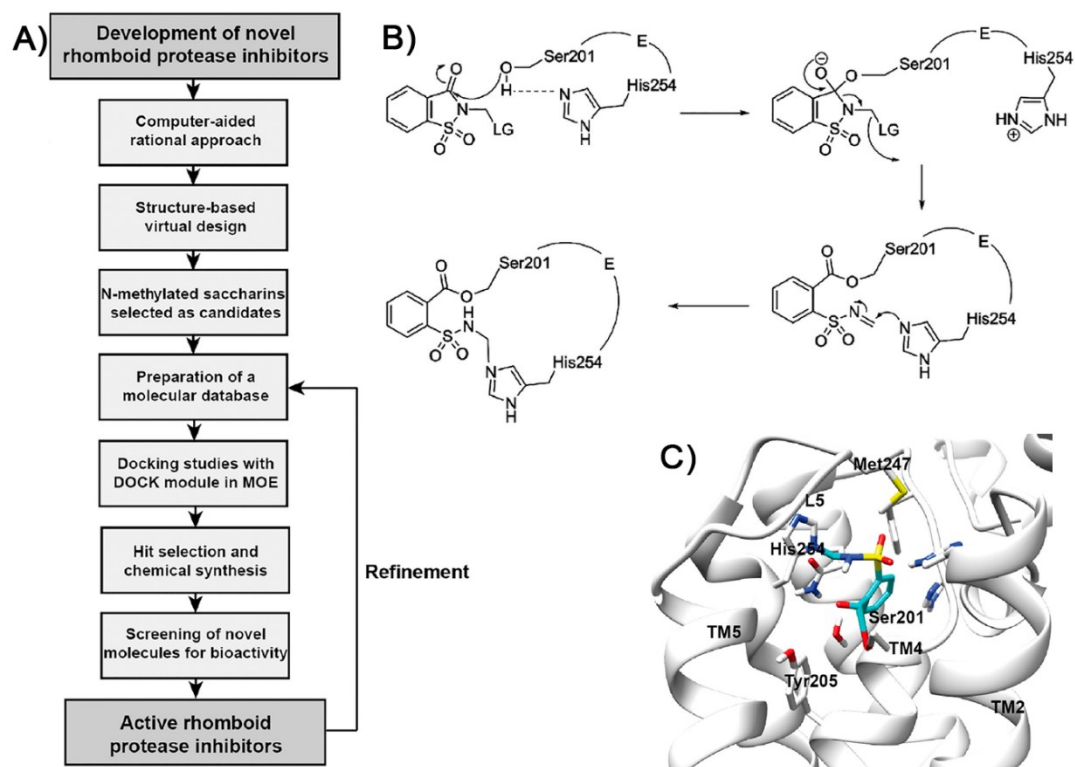


Figure 22: Next-generation saccharin inhibitors

A) A strategy for computer aided design of novel saccharin based rhomboid inhibitors. **B)** Mechanism of rhomboid inhibition by next-generation saccharin inhibitors. The catalytic serine (Ser201) attacks the carbonyl of the inhibitor. This leads to the prototropic shift and subsequent LG release and formation of a reactive intermediate that covalently cross-links the Ser201 with His254. **C)** Model of N-methylated saccharin inhibitor (blue) bound in the GlpG active site (grey). The inhibitor covalently cross-links to the rhomboid active site.

Mechanistic studies of classical serine proteases show that release of the acidic leaving group is followed by the formation of an irreversible covalent crosslink to the active site serine and histidine (Powers J. C. *et al.*, 2002). Crosslinking of the catalytic residues of GlpG causes a characteristic shift in mobility on SDS-PAGE electrophoresis (Vosyka O. *et al.*, 2013; Zoll S. *et al.*, 2014) which was not observed in the presence of saccharin inhibitor. Complexes of saccharin inhibitors with different GlpG mutants were analysed by mass spectrometry to determine which mutations prevent the formation of the characteristic adduct. However, no conclusive results were obtained from this experiment. Coupled to the fact that inhibition of GlpG by the saccharin inhibitor under study was slowly reversible (**Figure 23**), these data suggest that the mechanism of rhomboid inhibition by this class of inhibitors differs from the mechanism of inhibition of classical serine proteases.

The selectivity of the best compounds was assessed using the assays discussed in the section on Publication 2 (Bachovchin D. A. *et al.*, 2014; Wolf E. V. *et al.*, 2015a). The saccharins under study had no inhibitory activity towards eukaryotic rhomboids and

inhibited only few human serine hydrolases (out of 71). These results suggest that further optimisation of the saccharin scaffold could possibly lead to effective rhomboid inhibitors.

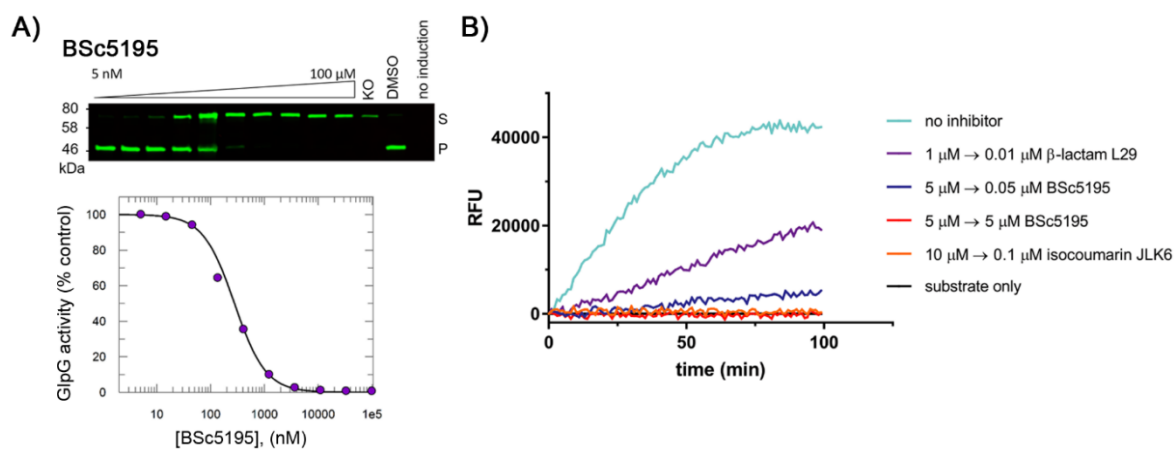


Figure 23: Characterisation of BSc5195, one of the most potent new saccharin inhibitors

A) Inhibition of endogenous GlpG in live NR698 *E. coli* strain transformed with the MBP-FLAG-LacYTM2-Trx substrate. Inhibition of substrate cleavage effected by different concentrations of inhibitor was detected in cell lysates by immunoblotting with fluorescence detection (α -FLAG). KO ($\Delta glpG$ NR698) serve as a negative control. A representative measurement (out of two) is shown. **B)** Next-generation saccharins are slowly reversible inhibitors. GlpG was pre-incubated with inhibitors for 1 h at 37 °C and rapidly diluted 100-fold into the reaction buffer or into inhibitor solution at equivalent concentration. The readout for reaction progress was cleavage of fluorogenic substrate KSp76 resulting in increased fluorescence ($\lambda_{ex} = 553$ nm, $\lambda_{em} = 583$ nm). For comparison, β -lactam L29 is a known reversible inhibitor whereas isocoumarin JLK6 binds GlpG irreversibly.

My contribution to this publication:

- *in vivo* assay in *E. coli*
- reversibility test

4.4 Publication 4

Discovery and validation of 2-styryl substituted benzoxazin-4-ones as a novel scaffold for rhomboid protease inhibitors

Parul Goel, Thorsten Jumpertz, Anežka Tichá, Isabella Ogorek, David C. Mikles, Martin Hubalek, Claus U. Pietrzik, Kvido Strišovský, Boris Schmidt, Sascha Weggen

This project was a follow up to the previous publication, again in collaboration with Sascha Weggen. An approach similar to that used for saccharin inhibitor identification was used for optimisation of 2-styryl substituted benzoxazin-4-ones. These are known inhibitors of HLE, chymotrypsin and cathepsin G (Gutschow M. and Neumann U., 1997; Gutschow M. *et al.*, 1999; Krantz A. *et al.*, 1990; Powers J. C. *et al.*, 2002; Teshima T. *et al.*, 1982).

The benzoxazinone scaffold consists of two fused aromatic rings which provide opportunities for modifications and improved binding to the target enzyme.

Thirteen 2-alkyl and 2-aryl benzoxazinone derivatives were tested *in silico* and *in vitro*. Docking studies of the most favourable molecule (2-styryl substituted benzoxazinone scaffold, compound 3) showed that the heterocyclic ring is oriented towards the S1 subsite whereas the substituent points towards the S2' subsite. The inhibitor also interacts with the adjacent His254 and Phe245. *In vitro*, compound 3 could inhibit Gurken hydrolysis by GlpG, thus confirming the results from the docking studies; compound 3 is the most potent analogue ($IC_{50} \sim 5 \mu\text{M}$). Surprisingly, compound 3 did not inhibit chymotrypsin whereas 50 % inhibition of HLE and trypsin was achieved with 50 μM inhibitor and 10 μM for cathepsin G, therefore showing at least some selectivity towards GlpG. Dilution experiments confirmed the expected reversible mechanism (**Figure 24**).

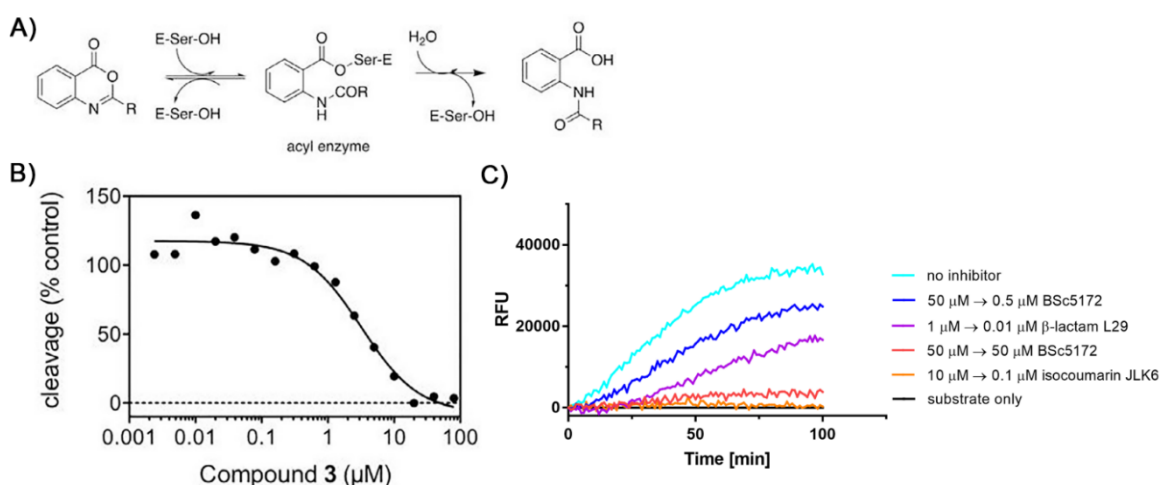


Figure 24: Characterisation of 2-substituted benzoxazin-4-ones as rhomboid inhibitors

A) Mechanism of 2-substituted benzoxazin-4-ones mediated inhibition of serine proteases. Attack by the catalytic serine on the C-4 benzoxazinone carbonyl leads to opening of the heterocyclic ring and the subsequent formation of the O-acyl enzyme intermediate (Powers J. C. *et al.*, 2002). **B)** Inhibition of GlpG by 2-aryl substituted benzoxazinones. GlpG was pre-incubated with the inhibitor for 1 h at 37 °C. The reaction was started by addition of the fluorogenic substrate KSp76. IC_{50} was determined only for the 2-aryl substituted inhibitors as 2-alkyl substituted benzoxazinones did not show sufficient inhibitory activity in a preliminary experiment. **C)** 2-aryl substituted benzoxazin-4-ones are reversible rhomboid inhibitors. The mixture of GlpG pre-incubated with a specific inhibitor for 1 h at 37 °C was diluted 100-fold into the reaction buffer or into inhibitor solution at concentration equivalent to that used during pre-incubation. The reaction was started by addition of KSp64, and the readout was an increase in fluorescence. Both IC_{50} and reversibility measurements were performed in triplicates. Here, the representatives of each experiment are presented. GlpG was pre-incubated with inhibitors for 1 h at 37 °C and rapidly 100× diluted into the reaction buffer or the same concentration of inhibitor. The reaction progress was followed as a cleavage of fluorogenic substrate KSp76 resulting in increase in fluorescence ($\lambda_{\text{ex}} = 553 \text{ nm}$, $\lambda_{\text{em}} = 583 \text{ nm}$). For a comparison, β -lactam L29 is a known reversible inhibitor whereas isocoumarin JLK 6 binds GlpG irreversibly.

My contribution to this publication:

- IC_{50} measurement
- reversibility test

5 DISCUSSION

In my thesis I focused on introducing new methodology to study rhomboid proteases. Proteins from the rhomboid superfamily, both the catalytically active proteases as well as the pseudoproteases, were several times shown to be involved in many important biological processes. New tools that I present in this thesis might help to better understand the biology of rhomboid proteases. Novel fluorescent substrates for rhomboids enable continuous measurement of reaction progress and provide a robust method for the assessment of rhomboid activity. Therefore, they were subsequently used for characterisation of newly developed inhibitors.

We based the fluorogenic substrates on the peptide sequence of the artificial rhomboid substrate LacYTM2 (Maegawa S. *et al.*, 2005). This, out of four different substrates (TatA from *P. stuartii*, Gurken and Spitz from *D. melanogaster* and *E. coli* LacYTM2), is best tolerated by the panel of rhomboids we tested (*E. coli* GlpG, *B. subtilis* YqgP, *P. stuartii* AarA and BtioR3 - rhomboid 3 from *B. thetaiotaomicron*). In order to keep the substrate as small as possible and to better understand substrate-rhomboid interactions, we prepared a series of C-terminally truncated versions of LacYTM2 and demonstrated that essentially the entire substrate TMD is required for cleavage to be unaffected. At the time we published this work, there was a general agreement that the recognition motif alters the catalytic turnover of the substrate (Dickey S. W. *et al.*, 2013; Strisovsky K. *et al.*, 2009), but no research to assess the importance of the substrate TMD had been published at that stage. Following the publication of our study a computational characterisation of rhomboid-substrate interaction was published showing high interaction energies for some of the residues deep in the substrate TMD, suggesting the importance of the substrate TMD for substrate-enzyme interaction (Shokhen M. and Albeck A., 2017). However, these data await experimental confirmation.

Consequently, the fluorogenic substrates are based on the entire TMD region of LacYTM2, and thus they are highly hydrophobic. For this reason, it is not surprising that their solubility depends on the presence of a suitable detergent in the reaction buffer. The effect of the detergent concentration on the rhomboid reaction rate was determined so far only for the recombinant YqgP in fusion with MBP. However, MBP is known to be a solubilising agent (Kapust R. B. and Waugh D. S., 1999) and indeed, MBP-YqgP is soluble. Probably due to the hydrophobic nature of YqgP, it also forms multimeric

complexes and addition of n-dodecyl β -D-maltoside (DDM) reduces its activity, albeit all concentrations have the same effect (Lei X. *et al.*, 2008). Our recombinant GlpG only contains a His₆-tag on its C-terminus and precipitates in the absence of any detergent (and probably mimics better the native protein) therefore it is not surprising that we observed different results.

We see a clear dependence of the reaction rate on detergent concentration. Increased detergent concentration did not have any impact on the secondary structure of any of the reactants, hence, the decrease in the reaction rate is not caused by altered substrate/enzyme conformation that could prevent recognition by rhomboid. Adding more detergent into the solution leads to higher concentration of empty micelles in the reaction, thus diluting the system, as the reaction between substrate and enzyme takes place only in the detergent micelles (both are hydrophobic). Decreased reaction rate with increased detergent concentration is therefore expectable. In our model we assume that the substrate distributes equally into the detergent micelles upon dilution from DMSO stock. Based on the critical micellar concentration and aggregation number we assume that under our reaction conditions (10 μ M substrate and 0.4 μ M GlpG in 0.05 % DDM), each micelle contains about '1.5' molecules of the substrate and only about 4 % of all micelles are occupied by the enzyme. Increasing DDM concentration lowers the probability of a substrate and an enzyme molecule being on the same micelle. Kinetic parameters of rhomboid catalysis were measured multiple times to date. However, the effect of the hydrophobic environment was taken into account only with the reactions in liposomes when the impact of detergent on kinetic constants had not been considered yet (Arutyunova E. *et al.*, 2014; Dickey S. W. *et al.*, 2013). For the *in vitro* assays it is important to keep this in mind as purification and concentration of both enzyme and substrate can increase detergent concentration in the stock solutions. To reach the same conditions in all experiments it is not sufficient to measure just the protein content, but the detergent concentration should be determined as well.

One of the main goals of this work was to prepare a substrate useful for HTS. Screened libraries often contain molecules that absorb in the UV region (Simeonov A. *et al.*, 2008), which disqualifies KSp35 with its Edans-Dabcyl FRET pair that absorbs light at 300-400 nm. We therefore synthesised KSp76, a substrate that contains the same peptide sequence as KSp35 but possesses a different FRET pair, Tamra-QXL610 (connected through Lys in P5 and Cys in P4', respectively). These changes did not affect the catalytic efficiency of its cleavage or its selectivity towards different rhomboids. Crucially, it absorbs in the red region

of the visible spectrum (550-700 nm). Out of the rhomboids investigated (GlpG, BtioR3, AarA, YqgP and GlpG S201A/H254A as a negative control) KSp35 and KSp76 were not cleaved by the inactive GlpG mutant (as expected) and YqgP. This is not surprising either as YqgP's catalytic activity strongly depends on the presence of lipids (Urban S. and Wolfe M. S., 2005). However, mutation of residues in the recognition motif (P5-P1) of the substrate to amino acids strongly preferred by GlpG (Zoll S. *et al.*, 2014), altered the specificity. Therefore the substrate with the RVRHA sequence in the P5-P1 site is cleaved by GlpG only. This optimisation improved the catalytic efficiency (k_{cat}/K_m) about 23-fold, confirming the previously reported importance of the recognition motif for substrate turnover (Dickey S. W. *et al.*, 2013; Strisovsky K. *et al.*, 2009). To test the suitability of our substrates for HTS we used KSp76 to test rhomboid inhibition by several inhibitors published previously – isocoumarin S037 (Vosyka O. *et al.*, 2013), β -lactam L42 (Pierrat O. A. *et al.*, 2011) and ISKA-cmk (Zoll S. *et al.*, 2014). Our measurements agree with published IC_{50} values (under comparable conditions). In summary, we presented here a series of novel fluorogenic substrates with proposed easy-to-implement modifications to increase specificity towards the targeted rhomboid. The choice of a suitable FRET pair enables their use in the HTS. The newly established fluorescence substrates were subsequently exploited in the development of new rhomboid inhibitors.

We have discovered peptidyl α -ketoamides as new potent and selective rhomboid inhibitors. The crystal structure of GlpG in complex with a peptidyl ketoamide showed a similar binding mode to that previously reported for CMKs (Zoll S. *et al.*, 2014) and peptidyl aldehydes (Cho S. *et al.*, 2016). However, we added to known data by demonstrating that peptidyl ketoamides bind to the S4-S2' GlpG subsites and that the P5 residue (Arg in the peptidyl part) does not dramatically affect the inhibitory potency. In contrast, even though the P5 and P6 residues were not visible in the crystal structure of the GlpG-RKVRMA-CHO complex, *in vitro* analysis shows that their presence improves the inhibitory potency 5-fold (Cho S. *et al.*, 2016). The discrepancy might be explained by looking at the structures and potencies of the aldehyde and ketoamide inhibitors. The long version of the aldehyde inhibitor (RKVRMA-CHO) has a K_i of about 20 μ M whereas ketoamide compound 9 (for which we tested the effect of shortening the peptidyl part) has a K_i of about 0.2 μ M, and is therefore a 100-fold more potent inhibitor. We have demonstrated that the prime side substituent has a large impact on inhibitor potency and this

effect might outweigh the contribution of the peptidyl part. This could explain the apparently conflicting data.

Rhomboids are expected to share a similar mechanism (Lemberg M. K. and Freeman M., 2007a; Lemberg M. K. and Freeman M., 2007b), hence our approach towards inhibitor design could therefore serve as a general strategy for delivery of potent rhomboid inhibitors optimised for a protease of interest. In case of GlpG we started from the analysis based on mutations in the P5-P1 positions of a rhomboid substrate TatA (Zoll S. *et al.*, 2014). There are several other approaches that can address the specificity of a protease, but they often comprise of complex labelling strategies which impedes their convenient use (Mahrus S. *et al.*, 2008). In contrast, multiplex substrate profiling utilises a library of synthetic peptides coupled with mass spectrometry. It is an easy and rapid method to map sequence preferences on both the prime and non-prime sides (O'Donoghue A. J. *et al.*, 2012) and could be used to identify further rhomboid sequence requirements as suggested by a recent work (Lapek J. D., Jr. *et al.*, 2019).

The second part of the inhibitor responsible for its potency and selectivity is the hydrophobic tail substituent. By varying the chemical residue attached to the ketoamide nitrogen we were able to improve the inhibitor potency by over three orders of magnitude. The crystal structure showed that the inhibitor binds from the active site up to the S2' subsite of GlpG. Unfortunately, we were not able to crystallise the complex of GlpG with our best compound (cpd11) but comparison of GlpG complexes with other ketoamides (cpd9 and 10) and previously reported β -lactam L29 (Vinothkumar K. R. *et al.*, 2013) and isocoumarin S016 (Vosyka O. *et al.*, 2013) shows that there is still space for improvement of the hydrophobic substituent (**Figure 25**). The C-terminal tail binds to the S2' cavity which is absent in the native rhomboid, but forms upon ligand binding (Vinothkumar K. R. *et al.*, 2013). Additionally, the introduction of a longer substituent, potentially of a branched and hydrophobic nature, to the prime side might increase the potency of the inhibitor because, as we show in Publication 1, the membrane part of the substrate is also important for substrate turnover. Other potential approaches to further improve ketoamide inhibitors is based on deeper understanding of the rhomboid catalytic mechanism. Specifically, the S1 cavity in GlpG was shown to encompass a water retention site which explains the strong preference for small amino acids in P1 (especially Ala) (Akiyama Y. and Maegawa S., 2007; Strisovsky K. *et al.*, 2009; Zoll S. *et al.*, 2014). Thus, bigger residues might possibly fit into the S1

cavity but interfere with water transport to the active site and result in a more effective and rhomboid-selective inhibitor.

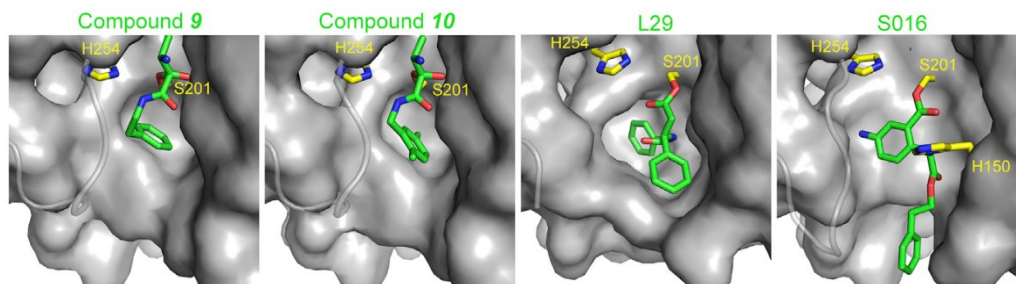


Figure 25: Comparison of the binding modes of different rhomboid inhibitors

All of the illustrated inhibitors bind to the S2' cavity, but each of them occupies a slightly different space. This suggests a direction for the future design of N-terminal hydrophobic substituents of peptidyl ketoamides. This is a comparison of crystal structures PDB: 5MT6, 5MTF, 3ZMI and 3ZEB. The rhomboid surface is depicted in grey, the catalytic dyad in yellow and inhibitors in green.

The ketoamide warhead is already in use as a pharmacophore in treatment of hepatitis C infection (Njoroge F. G. *et al.*, 2008). The family of drugs does not cause any unwanted side effects even though they comprise of an electrophilic warhead that covalently modifies the target protein. As these licensed drugs are also peptidyl ketoamides, we assume that our compounds behave similarly, and as such are suitable for cell biological assays.

Apart from peptidyl ketoamide inhibitors we also tested other classes of molecules potentially active against rhomboids, namely next-generation saccharin inhibitors and benzoxazin-4-one derived inhibitors. Both scaffolds offer a broad range of structural variations. The effect of varying the LG of saccharin inhibitors was initially determined *in silico* with docking studies. Binding of all 50 proposed derivatives to the GlpG active site was simulated based on the structure of GlpG in complex with the phosphonofluoridate inhibitor CAPF. The docking studies revealed that saccharins with the aryl carboxylic acid LG were the most efficient ligands (also confirmed *in vitro*) and that the binding mode of these inhibitors is similar to that of CAPF (Xue Y. *et al.*, 2012). The sulfoxide group interacts with His150 and Gly199 and the LG interacts with the prime side of the enzyme. The preference for aryl carboxylic acid LG could be explained by the fact that the aryl ring participates in π - π interactions with Phe245. This interaction might also contribute to inhibitor selectivity as the aryl ring fits nicely into a hydrophobic region around the GlpG active site. More detailed studies of the electrostatic interactions between inhibitor and enzyme suggested that small electron-withdrawing substituents on the aryl ring at the *ortho* or *para* positions could increase the inhibitor potency as it would become more favourable

for nucleophilic attack by the catalytic Ser. Surprisingly, neither further docking studies nor *in vitro* inhibition assays of GlpG yielded further inhibitors with improved affinity. Briefly, the IC_{50} value of the most potent compound was 0.2 μM which is comparable to the potency of the most efficient β -lactams (Pierrat O. A. *et al.*, 2011) or isocoumarin (Vosyka O. *et al.*, 2013). Unlike β -lactams, which inhibit endogenous GlpG with an IC_{50} of 5-10 μM (Pierrat O. A. *et al.*, 2011), the saccharin inhibitors have a similar potency *in vivo* as they do in *in vitro* studies. Their advantage over isocoumarins is their slowly reversible mechanism.

Saccharin-based inhibitors are expected to form a doubly-linked covalent complex with the target enzyme (Powers J. C. *et al.*, 2002), although it is not clear which second residue, in addition to Ser201, of GlpG they form a covalent bond with. Therefore, GlpG variations with candidate residues mutated to Ala were prepared and incubated with the saccharin inhibitors. Based on the proposed mechanism, the LG should be released and the complexes formed should result in an identical mass shift of 195 Da. The complex was not formed upon His254 deletion implicating His254 as the second binding partner. On the other hand, the catalytic histidine is needed for proper activation of the catalytic serine (Ha Y. *et al.*, 2013) which may be another explanation for why the complex was not formed. Coupled with the fact that the expected change in the protein mobility using SDS-PAGE (Vosyka O. *et al.*, 2013; Zoll S. *et al.*, 2014) was not observed, this suggests the possibility that the doubly-linked product is either not formed or not stable, or that another nucleophile (apart from the His254) in closer proximity to the catalytic serine might be involved in the crosslinking. As the saccharins are slowly reversible inhibitors, the hypothesis including another binding partner is favoured.

A similar approach was utilised in the design of benzoxazin-4-one derived inhibitors. Thirteen candidates underwent molecular docking studies, again based on the structure of GlpG in complex with CAPF (Xue Y. *et al.*, 2012) and 2-styryl substituted compound 3 (**Figure 24**) was shown to be the most promising derivative. Benzoxazinones also share a similar orientation with previously reported inhibitors. In this case, the styryl extension points towards the S2' subsite in the enzyme, as observed with previously characterised inhibitors (Vinothkumar K. R. *et al.*, 2013; Xue Y. *et al.*, 2012). *In vitro* experiments confirmed compound 3 as one of the most potent benzoxazinone derivatives, with an IC_{50} of 4 μM . An electron withdrawing substituent increased the inhibitor activity, similarly to the saccharin inhibitors. The potency of the best benzoxazinone compound is one order of magnitude lower than that of the previously discussed saccharins, as well as of β -lactams

(Pierrat O. A. *et al.*, 2011) and isocoumarins (Vosyka O. *et al.*, 2013). Benzoxazinones are known inhibitors of HLE, chymotrypsin and cathepsin G (Gutschow M. and Neumann U., 1997; Gutschow M. *et al.*, 1999; Krantz A. *et al.*, 1990; Powers J. C. *et al.*, 2002; Teshima T. *et al.*, 1982). Surprisingly, the two most potent derivatives did not show any inhibitory activity against α -chymotrypsin up to 250 μ M concentration. Neither did they inhibit bovine trypsin or human neutrophil elastase at 10 μ M, but several inhibitory variants were able to decrease the activity of these enzymes to 50 % at 50 μ M concentration. One of the inhibitors significantly reduced the activity of cathepsin G at 10 μ M concentration. The specificity profile was not uniform for all variants tested, suggesting that these inhibitors could show some selectivity for GlpG over other serine proteases.

We also collaborated with another group interested in benzoxazin-4-one scaffold-based inhibitors (Steven Verhelst's group, University of Leuven, Leuven, Belgium) and these two papers were published together (Goel P. *et al.*, 2018; Yang J. *et al.*, 2018). The Verhelst group used a different set of substituents attached to the benzoxazin-4-one core, the most effective of which were the 2-alkoxy substituents. Even though they showed slightly higher activity against GlpG, they lost the selectivity over bovine chymotrypsin and trypsin (Yang J. *et al.*, 2018). Therefore, it would be interesting to fully understand the relationship between the inhibitor structure and its potency and selectivity for rhomboids, potentially leading to an inhibitor that combines the optimal features from each type. From the mechanical point of view, benzoxazinones bind covalently to the GlpG which was confirmed by a mass spectrometry by formation of a new peak of a mass corresponding to the expected rhomboid-inhibitor adduct. Their reversibility was confirmed by the dilution experiment.

As inhibitor potency is not the only important feature of the inhibitor that determines its suitability in cell biological assays (not to mention its potential therapeutic use), we tested the peptidyl ketoamides and saccharin inhibitors for their selectivity across about 100 human hydrolases, using the EnPlex method (Bachovchin D. A. *et al.*, 2014) and their selectivity across a panel of different rhomboid proteases (Wolf E. V. *et al.*, 2015a). Both ketoamides and saccharins displayed preference for bacterial and archaeal rhomboid proteases. We also compared their selectivity in the EnPlex assay with β -lactam L41 (Pierrat O. A. *et al.*, 2011) and isocoumarins S006 and S016 (Vosyka O. *et al.*, 2013). Both ketoamides and saccharins have a better selectivity profile than both isocoumarins. What is more, ketoamides inhibited only two out of all the tested hydrolases (dipeptidyl-peptidase 2 and prolylcarboxypeptidase), L41 inhibited just one of them (probable serine carboxypeptidase).

Together with the high potency of peptidyl ketoamides and their safe toxicological profile this makes them the most effective rhomboid inhibitors available to date. Undoubtedly, this scaffold can be further developed to target the rhomboid proteases of medical interest, such as the mitochondrial rhomboid PARL or the *Plasmodium* rhomboid ROM4.

6 SUMMARY

In my thesis I presented several novel tools to study rhomboid proteases, namely new fluorogenic substrates, and potent and selective rhomboid inhibitors. In general, their main advantage is that they should be readily modifiable to target different rhomboids.

With respect to the fluorogenic substrate, our primary goal was to deliver tools that could be used in continuous rhomboid activity assays, would be well tolerated by many rhomboid proteases and would be suitable for HTS. We have met all of these requirements in the substrate presented in chapter 4.1, and subsequently demonstrated their practical use in identifying new rhomboid inhibitors. The substrates were used in *in vitro* assays to characterise the effects of different substituents on inhibitor activity as well as for the kinetic characterisation of the interaction between the most promising compounds and rhomboids.

Regarding the inhibitor design, I present in this thesis three new classes of rhomboid inhibitors. The peptidyl ketoamides are, to date, the best inhibitors available. Their high potency, selectivity and the fact that the same class of compounds is already used as therapeutics makes them ideal candidates for further drug development. Additionally, we have contributed to the improvement of saccharin and benzoxazin-4-one scaffold based inhibitors. Their major advantage is an easy and cheap synthesis route which offers many possibilities for future structure-activity relationship studies that could lead to the development of a potent and selective inhibitor that could be readily synthesised in high yields.

The rhomboid inhibitors, particularly the peptidyl ketoamides, could also be used as a starting point in the search for potential therapeutic molecules against many diseases. Inhibition of parasite invasion of the host cells through inhibition of *Plasmodium* rhomboids might be a feasible malaria treatment (Baker R. P. *et al.*, 2006; Lin J. W. *et al.*, 2013), mitochondrial rhomboid protein PARL probably cooperates on regulation of mitophagy, therefore inhibition of PARL might stimulate mitophagy and potentially decrease the pathological changes associated with Parkinson's disease (Chan E. Y. L. and McQuibban G. A., 2013; Meissner C. *et al.*, 2015), and regulation of EGFR signalling through RBHD4 could be beneficial in the treatment of colorectal cancer (Song W. *et al.*, 2015). Furthermore it is tempting to speculate that inhibitors of active rhomboids could also be tools for the study of the rhomboid pseudoproteases, whose mechanisms remain largely unclear but like contain a structural equivalent of a rhomboid protease active site cavity (in most cases lacking the catalytic residues though).

7 REFERENCES

- Adrain, C. and Freeman, M. (2012). "New lives for old: evolution of pseudoenzyme function illustrated by iRhoms". *Nature Reviews Molecular Cell Biology* **13** (8): 489-498.
- Adrain, C., Zettl, M., Christova, Y., Taylor, N. and Freeman, M. (2012). "Tumor Necrosis Factor Signaling Requires iRhom2 to Promote Trafficking and Activation of TACE". *Science* **335** (6065): 225-228.
- Adrain, C., Strisovsky, K., Zettl, M., Hu, L., Lemberg, M. K. and Freeman, M. (2011). "Mammalian EGF receptor activation by the rhomboid protease RHBDL2". *EMBO Reports* **12** (5): 421-427.
- Akiyama, Y. and Maegawa, S. (2007). "Sequence features of substrates required for cleavage by GlpG, an Escherichia coli rhomboid protease". *Molecular Microbiology* **64** (4): 1028-1037.
- Arutyunova, E., Strisovsky, K. and Lemieux, M. J. (2017). "Activity Assays for Rhomboid Proteases". *Methods in Enzymology* **584**: 395-437.
- Arutyunova, E., Panwar, P., Skiba, P. M., Gale, N., Mak, M. W. and Lemieux, M. J. (2014). "Allosteric regulation of rhomboid intramembrane proteolysis". *EMBO Journal* **33** (17): 1869-1881.
- Bachovchin, D. A., Koblan, L. W., Wu, W., Liu, Y., Li, Y., Zhao, P., Woznica, I., Shu, Y., Lai, J. H., Poplawski, S. E., Kiritsy, C. P., Healey, S. E., DiMare, M., Sanford, D. G., Munford, R. S., Bachovchin, W. W. and Golub, T. R. (2014). "A high-throughput, multiplexed assay for superfamily-wide profiling of enzyme activity". *Nature Chemical Biology* **10** (8): 656-663.
- Baker, R. P. and Urban, S. (2012). "Architectural and thermodynamic principles underlying intramembrane protease function". *Nature Chemical Biology* **8** (9): 759-768.
- Baker, R. P., Wijetilaka, R. and Urban, S. (2006). "Two Plasmodium rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria". *PLoS Pathogens* **2** (10): 922-932.
- Baker, R. P., Young, K., Feng, L., Shi, Y. and Urban, S. (2007). "Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5 as the lateral substrate gate". *Proceedings of the National Academy of Sciences of the United States of America* **104** (20): 8257-8262.
- Ben-Shem, A., Fass, D. and Bibi, E. (2007). "Structural basis for intramembrane proteolysis by rhomboid serine proteases". *Proceedings of the National Academy of Sciences of the United States of America* **104** (2): 462-466.
- Bier, E., Jan, L. Y. and Jan, Y. N. (1990). "Rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous-system development in *Drosophila melanogaster*". *Genes & Development* **4** (2): 190-203.
- Black, R. A., Rauch, C. T., Kozlosky, C. J., Peschon, J. J., Slack, J. L., Wolfson, M. F., Castner, B. J., Stocking, K. L., Reddy, P., Srinivasan, S., Nelson, N., Boiani, N., Schooley, K. A., Gerhart, M., Davis, R., Fitzner, J. N., Johnson, R. S., Paxton, R. J., March, C. J. and Cerretti, D. P. (1997). "A metalloproteinase disintegrin that releases tumour-necrosis factor- α from cells". *Nature* **385** (6618): 729-733.

- Blaydon, D. C., Etheridge, S. L., Risk, J. M., Hennies, H.-C., Gay, L. J., Carroll, R., Plagnol, V., McDonald, F. E., Stevens, H. P., Spurr, N. K., Bishop, D. T., Ellis, A., Jankowski, J., Field, J. K., Leigh, I. M., South, A. P. and Kelsell, D. P. (2012). "RHBDF2 Mutations Are Associated with Tylosis, a Familial Esophageal Cancer Syndrome". *American Journal of Human Genetics* **90** (2): 340-346.
- Bondar, A. N., del Val, C. and White, S. H. (2009). "Rhomboid protease dynamics and lipid interactions". *Structure* **17** (3): 395-405.
- Brossier, F., Jewett, T. J., Lovett, J. L. and Sibley, L. D. (2003). "C-terminal processing of the toxoplasma protein MIC2 is essential for invasion into host cells". *Journal of Biological Chemistry* **278** (8): 6229-6234.
- Brossier, F., Jewett, T. J., Sibley, L. D. and Urban, S. (2005). "A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by Toxoplasma". *Proceedings of the National Academy of Sciences of the United States of America* **102** (11): 4146-4151.
- Brown, M. S., Ye, J., Rawson, R. B. and Goldstein, J. L. (2000). "Regulated Intramembrane Proteolysis". *Cell* **100** (4): 391-398.
- Carruthers, V. B. and Boothroyd, J. C. (2007). "Pulling together: an integrated model of Toxoplasma cell invasion". *Current Opinion in Microbiology* **10** (1): 82-89.
- Civitarese, A. E., MacLean, P. S., Carling, S., Kerr-Bayles, L., McMillan, R. P., Pierce, A., Becker, T. C., Moro, C., Finlayson, J., Lefort, N., Newgard, C. B., Mandarino, L., Cefalu, W., Walder, K., Collier, G. R., Hulver, M. W., Smith, S. R. and Ravussin, E. (2010). "Regulation of Skeletal Muscle Oxidative Capacity and Insulin Signaling by the Mitochondria! Rhomboid Protease PARL". *Cell Metabolism* **11** (5): 412-426.
- Clemmer, K. M., Sturgill, G. M., Veenstra, A. and Rather, P. N. (2006). "Functional characterization of Escherichia coli GlpG and additional rhomboid proteins using an aarA mutant of Providencia stuartii". *Journal of Bacteriology* **188** (9): 3415-3419.
- Daley, D. O., Rapp, M., Granseth, E., Melen, K., Drew, D. and von Heijne, G. (2005). "Global topology analysis of the Escherichia coli inner membrane proteome". *Science* **308** (5726): 1321-1323.
- De Strooper, B., Saftig, P., Craessaerts, K., Vanderstichele, H., Guhde, G., Annaert, W., Von Figura, K. and Van Leuven, F. (1998). "Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein". *Nature* **391** (6665): 387-390.
- Deas, E., Wood, N. W. and Plun-Favreau, H. (2011a). "Mitophagy and Parkinson's disease: the PINK1-parkin link". *Biochimica et Biophysica Acta - Molecular Cell Research* **1813** (4): 623-633.
- Deas, E., Plun-Favreau, H., Gandhi, S., Desmond, H., Kjaer, S., Loh, S. H., Renton, A. E., Harvey, R. J., Whitworth, A. J., Martins, L. M., Abramov, A. Y. and Wood, N. W. (2011b). "PINK1 cleavage at position A103 by the mitochondrial protease PARL". *Human Molecular Genetics* **20** (5): 867-879.
- Dickey, S. W., Baker, R. P., Cho, S. and Urban, S. (2013). "Proteolysis inside the Membrane Is a Rate-Governed Reaction Not Driven by Substrate Affinity". *Cell* **155** (6): 1270-1281.
- Drag, M. and Salvesen, G. S. (2010). "Emerging principles in protease-based drug discovery". *Nature Reviews Drug Discovery* **9** (9): 690-701.

- Dusterhoft, S., Kunzel, U. and Freeman, M. (2017). "Rhomboid proteases in human disease: Mechanisms and future prospects". *Biochimica et Biophysica Acta - Molecular Cell Research* **1864** (11): 2200-2209.
- Ejigiri, I., Ragheb, D. R. T., Pino, P., Coppi, A., Bennett, B. L., Soldati-Favre, D. and Sinnis, P. (2012). "Shedding of TRAP by a Rhomboid Protease from the Malaria Sporozoite Surface Is Essential for Gliding Motility and Sporozoite Infectivity". *PLoS Pathogens* **8** (7):
- Ekici, O. D., Paetzel, M. and Dalbey, R. E. (2008). "Unconventional serine proteases: Variations on the catalytic Ser/His/Asp triad configuration". *Protein Science* **17** (12): 2023-2037.
- Esler, W. P., Kimberly, W. T., Ostaszewski, B. L., Diehl, T. S., Moore, C. L., Tsai, J. Y., Rahmati, T., Xia, W. M., Selkoe, D. J. and Wolfe, M. S. (2000). "Transition-state analogue inhibitors of gamma-secretase bind directly to presenilin-1". *Nature Cell Biology* **2** (7): 428-434.
- Fleig, L., Bergbold, N., Sahasrabudhe, P., Geiger, B., Kaltak, L. and Lemberg, M. K. (2012). "Ubiquitin-Dependent Intramembrane Rhomboid Protease Promotes ERAD of Membrane Proteins". *Molecular Cell* **47** (4): 558-569.
- Gallio, M., Sturgill, G., Rather, P. and Kylsten, P. (2002). "A conserved mechanism for extracellular signaling in eukaryotes and prokaryotes". *Proceedings of the National Academy of Sciences of the United States of America* **99** (19): 12208-12213.
- Ghasriani, H., Kwok, J. K., Sherratt, A. R., Foo, A. C., Qureshi, T. and Goto, N. K. (2014). "Micelle-catalyzed domain swapping in the GlpG rhomboid protease cytoplasmic domain". *Biochemistry* **53** (37): 5907-5915.
- Gilmore, T. D. (2006). "Introduction to NF-kappaB: players, pathways, perspectives". *Oncogene* **25** (51): 6680-6684.
- Goel, P., Jumpertz, T., Ticha, A., Ogorek, I., Mikles, D. C., Hubalek, M., Pietrzik, C. U., Strisovsky, K., Schmidt, B. and Weggen, S. (2018). "Discovery and validation of 2-styryl substituted benzoxazin-4-ones as a novel scaffold for rhomboid protease inhibitors". *Bioorganic & Medicinal Chemistry Letters* **28** (8): 1417-1422.
- Greenblatt, E. J., Olzmann, J. A. and Kopito, R. R. (2011). "Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant alpha-1 antitrypsin from the endoplasmic reticulum". *Nature Structural & Molecular Biology* **18** (10): 1147-1152.
- Greene, A. W., Grenier, K., Aguilera, M. A., Muise, S., Farazifard, R., Haque, M. E., McBride, H. M., Park, D. S. and Fon, E. A. (2012). "Mitochondrial processing peptidase regulates PINK1 processing, import and Parkin recruitment". *EMBO Reports* **13** (4): 378-385.
- Groutas, W. C., Epp, J. B., Venkataraman, R., Kuang, R., Truong, T. M., McClenahan, J. J. and Prakash, O. (1996). "Design, synthesis, and in vitro inhibitory activity toward human leukocyte elastase, cathepsin G, and proteinase 3 of saccharin-derived sulfones and congeners". *Bioorganic & Medicinal Chemistry* **4** (9): 1393-1400.
- Guichard, A., Biehs, B., Sturtevant, M. A., Wickline, L., Chacko, J., Howard, K. and Bier, E. (1999). "Rhomboid and Star interact synergistically to promote EGFR/MAPK signaling during Drosophila wing vein development". *Development* **126** (12): 2663-2676.

- Gutschow, M. and Neumann, U. (1997). "Inhibition of cathepsin G by 4H-3,1-benzoxazin-4-ones". *Bioorganic & Medicinal Chemistry* **5** (10): 1935-1942.
- Gutschow, M., Kuerschner, L., Neumann, U., Pietsch, M., Loser, R., Koglin, N. and Eger, K. (1999). "2-(diethylamino)thieno1,3oxazin-4-ones as stable inhibitors of human leukocyte elastase". *Journal of Medicinal Chemistry* **42** (26): 5437-5447.
- Ha, Y., Akiyama, Y. and Xue, Y. (2013). "Structure and Mechanism of Rhomboid Protease". *Journal of Biological Chemistry* **288** (22): 15430-15436.
- Han, J. Y., Bai, J. C., Yang, Y., Yin, H., Gao, W., Lu, A. G., Liu, F., Ge, H. Y., Liu, Z. M., Wang, J. Y. and Zhong, L. (2015). "Lentivirus-mediated knockdown of rhomboid domain containing 1 inhibits colorectal cancer cell growth". *Molecular Medicine Reports* **12** (1): 377-381.
- Harris, P. K., Yeoh, S., Dluzewski, A. R., O'Donnell, R. A., Withers-Martinez, C., Hackett, F., Bannister, L. H., Mitchell, G. H. and Blackman, M. J. (2005). "Molecular identification of a malaria merozoite surface sheddase". *PLoS Pathogens* **1** (3): 241-251.
- Hedstrom (2002). "Serine Protease Mechanism and Specificity".
- Heinitz, S., Klein, C. and Djarmati, A. (2011). "The p.S77N presenilin-associated rhomboid-like protein mutation is not a frequent cause of early-onset Parkinson's disease". *Movement Disorders* **26** (13): 2441-2442.
- Hempel, F., Bullmann, L., Lau, J., Zauner, S. and Maier, U. G. (2009). "ERAD-derived preprotein transport across the second outermost plastid membrane of diatoms". *Molecular Biology and Evolution* **26** (8): 1781-1790.
- Herlan, M., Vogel, F., Bornhovd, C., Neupert, W. and Reichert, A. S. (2003). "Processing of Mgm1 by the rhomboid-type protease Pcp1 is required for maintenance of mitochondrial morphology and of mitochondrial DNA". *Journal of Biological Chemistry* **278** (30): 27781-27788.
- Higy, M., Junne, T. and Spiess, M. (2004). "Topogenesis of membrane proteins at the endoplasmic reticulum". *Biochemistry* **43** (40): 12716-12722.
- Hlasta, D. J., Ackerman, J. H., Court, J. J., Farrell, R. P., Johnson, J. A., Kofron, J. L., Robinson, D. T., Talomie, T. G., Dunlap, R. P. and Franke, C. A. (1995). "A novel class of cyclic beta-dicarbonyl leaving groups and their use in the design of benzisothiazolone human leukocyte elastase inhibitors". *Journal of Medicinal Chemistry* **38** (23): 4687-4692.
- Huang, Z. J. (1991). "Kinetic assay of fluorescein mono-beta-D-galactoside hydrolysis by beta-galactosidase: a front-face measurement for strongly absorbing fluorogenic substrates". *Biochemistry* **30** (35): 8530-8534.
- Hubbard, S. J. (1998). "The structural aspects of limited proteolysis of native proteins". *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology* **1382** (2): 191-206.
- Chan, E. Y. L. and McQuibban, G. A. (2013). "The mitochondrial rhomboid protease: Its rise from obscurity to the pinnacle of disease-relevant genes". *Biochimica et Biophysica Acta - Biomembranes* **1828** (12): 2916-2925.

- Chatterjee, S., Dunn, D., Tao, M., Wells, G., Gu, Z. Q., Bihovsky, R., Ator, M. A., Siman, R. and Mallamo, J. P. (1999). "P2-achiral, P'-extended alpha-ketoamide inhibitors of calpain I". *Bioorganic & Medicinal Chemistry Letters* **9** (16): 2371-2374.
- Cheng, T. L., Wu, Y. T., Lin, H. Y., Hsu, F. C., Liu, S. K., Chang, B. I., Chen, W. S., Lai, C. H., Shi, G. Y. and Wu, H. L. (2011). "Functions of Rhomboid Family Protease RHBDL2 and Thrombomodulin in Wound Healing". *Journal of Investigative Dermatology* **131** (12): 2486-2494.
- Cho, S., Dickey, S. W. and Urban, S. (2016). "Crystal Structures and Inhibition Kinetics Reveal a Two-Stage Catalytic Mechanism with Drug Design Implications for Rhomboid Proteolysis". *Molecular Cell* **61** (3): 329-340.
- Christova, Y., Adrain, C., Bambrough, P., Ibrahim, A. and Freeman, M. (2013). "Mammalian iRhoms have distinct physiological functions including an essential role in TACE regulation". *EMBO Reports* **14** (10): 884-890.
- Issuree, P. D. A., Maretzky, T., McIlwain, D. R., Monette, S., Qing, X., Lang, P. A., Swendeman, S. L., Park-Min, K.-H., Binder, N., Kallioli, G. D., Yamilina, A., Horiuchi, K., Ivashkiv, L. B., Mak, T. W., Salmon, J. E. and Blobel, C. P. (2013). "iRHOM2 is a critical pathogenic mediator of inflammatory arthritis". *Journal of Clinical Investigation* **123** (2): 928-932.
- Jackson, L. F., Qiu, T. H., Sunnarborg, S. W., Chang, A., Zhang, C., Patterson, C. and Lee, D. C. (2003). "Defective valvulogenesis in HB-EGF and TACE-null mice is associated with aberrant BMP signaling". *EMBO Journal* **22** (11): 2704-2716.
- James, M. N. G. (1993). "Convergence of Active-Center Geometries among the Proteolytic-Enzymes". *Proteolysis and Protein Turnover* 1-8.
- Jeyaraju, D. V., Xu, L., Letellier, M.-C., Bandaru, S., Zunino, R., Berg, E. A., McBride, H. M. and Pellegrini, L. (2006). "Phosphorylation and cleavage of presenilin-associated rhomboid-like protein (PARL) promotes changes in mitochondrial morphology". *Proceedings of the National Academy of Sciences of the United States of America* **103** (49): 18562-18567.
- Jin, S. M., Lazarou, M., Wang, C., Kane, L. A., Narendra, D. P. and Youle, R. J. (2010). "Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL". *Journal of Cell Biology* **191** (5): 933-942.
- Johnson, N., Brezinova, J., Stephens, E., Burbridge, E., Freeman, M., Adrain, C. and Strisovsky, K. (2017). "Quantitative proteomics screen identifies a substrate repertoire of rhomboid protease RHBDL2 in human cells and implicates it in epithelial homeostasis". *Scientific Reports* **7**
- Jones, J. L., Kruszon-Moran, D., Wilson, M., McQuillan, G., Navin, T. and McAuley, J. B. (2001). "Toxoplasma gondii infection in the United States: Seroprevalence and risk factors". *American Journal of Epidemiology* **154** (4): 357-365.
- Jurgens, G., Wieschaus, E., Nussleinvolhard, C. and Kluding, H. (1984). "Mutations Affecting the Pattern of the Larval Cuticle in Drosophila-Melanogaster .2. Zygotic Loci on the 3rd Chromosome". *Wilhelm Roux's Archives of Developmental Biology* **193** (5): 283-295.

- Kapust, R. B. and Waugh, D. S. (1999). "Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused". *Protein Science* **8** (8): 1668-1674.
- Kateete, D. P., Katabazi, F. A., Okeng, A., Okee, M., Musinguzi, C., Asiimwe, B. B., Kyobe, S., Asiimwe, J., Boom, W. H. and Joloba, M. L. (2012). "Rhomboids of Mycobacteria: characterization using an aarA mutant of *Providencia stuartii* and gene deletion in *Mycobacterium smegmatis*". *PLoS One* **7** (9): e45741.
- Keller, S., Sanderson, M. P., Stoeck, A. and Altevogt, P. (2006). "Exosomes: from biogenesis and secretion to biological function". *Immunology Letters* **107** (2): 102-108.
- Kelley, D. E., He, J., Menshikova, E. V. and Ritov, V. B. (2002). "Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes". *Diabetes* **51** (10): 2944-2950.
- Killian, J. A. and von Heijne, G. (2000). "How proteins adapt to a membrane-water interface". *Trends in Biochemical Sciences* **25** (9): 429-434.
- Kim, H., Melen, K., Osterberg, M. and von Heijne, G. (2006). "A global topology map of the *Saccharomyces cerevisiae* membrane proteome". *Proceedings of the National Academy of Sciences of the United States of America* **103** (30): 11142-11147.
- Kimura, A. and Kishimoto, T. (2010). "IL-6: regulator of Treg/Th17 balance". *European Journal of Immunology* **40** (7): 1830-1835.
- Kinch, L. N. and Grishin, N. V. (2013). "Bioinformatics perspective on rhomboid intramembrane protease evolution and function". *Biochimica et Biophysica Acta - Biomembranes* **1828** (12): 2937-2943.
- Kirst, M. E., Meyer, D. J., Gibbon, B. C., Jung, R. and Boston, R. S. (2005). "Identification and characterization of endoplasmic reticulum-associated degradation proteins differentially affected by endoplasmic reticulum stress". *Plant Physiology* **138** (1): 218-231.
- Knop, M., Finger, A., Braun, T., Hellmuth, K. and Wolf, D. H. (1996). "Der1, a novel protein specifically required for endoplasmic reticulum degradation in yeast". *EMBO Journal* **15** (4): 753-763.
- Koonin, E. V., Makarova, K. S., Rogozin, I. B., Davidovic, L., Letellier, M. C. and Pellegrini, L. (2003). "The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers". *Genome Biology* **4** (3): R19.
- Krantz, A., Spencer, R. W., Tam, T. F., Liak, T. J., Copp, L. J., Thomas, E. M. and Rafferty, S. P. (1990). "Design and synthesis of 4H-3,1-benzoxazin-4-ones as potent alternate substrate inhibitors of human leukocyte elastase". *Journal of Medicinal Chemistry* **33** (2): 464-479.
- Kreutzberger, A. J. B. and Urban, S. (2018). "Single-Molecule Analyses Reveal Rhomboid Proteins Are Strict and Functional Monomers in the Membrane". *Biophysical Journal* **115** (9): 1755-1761.
- Kreutzberger, A. J. B., Ji, M., Aaron, J., Mihaljevic, L. and Urban, S. (2019). "Rhomboid distorts lipids to break the viscosity-imposed speed limit of membrane diffusion". *Science* **363** (6426):

- Kunzel, U., Grieve, A. G., Meng, Y., Sieber, B., Cowley, S. A. and Freeman, M. (2018). "FRMD8 promotes inflammatory and growth factor signalling by stabilising the iRhom/ADAM17 sheddase complex". *eLife* **7**
- Lapek, J. D., Jr., Jiang, Z., Wozniak, J. M., Arutyunova, E., Wang, S. C., Lemieux, M. J., Gonzalez, D. J. and O'Donoghue, A. J. (2019). "Quantitative Multiplex Substrate Profiling of Peptidases by Mass Spectrometry". *Molecular & Cellular Proteomics*
- Lazareno-Saez, C., Arutyunova, E., Coquelle, N. and Lemieux, M. J. (2013). "Domain swapping in the cytoplasmic domain of the Escherichia coli rhomboid protease". *Journal of Molecular Biology* **425** (7): 1127-1142.
- Lee, A. G. (2004). "How lipids affect the activities of integral membrane proteins". *Biochimica et Biophysica Acta - Biomembranes* **1666** (1-2): 62-87.
- Lee, J. R., Urban, S., Garvey, C. F. and Freeman, M. (2001). "Regulated intracellular ligand transport and proteolysis control EGF signal activation in Drosophila". *Cell* **107** (2): 161-171.
- Lei, X., Ahn, K., Zhu, L., Ubarretxena-Belandia, I. and Li, Y. M. (2008). "Soluble oligomers of the intramembrane serine protease YqgP are catalytically active in the absence of detergents". *Biochemistry* **47** (46): 11920-11929.
- Lemberg, M. K. (2013). "Sampling the membrane: function of rhomboid-family proteins". *Trends in Cell Biology* **23** (5): 210-217.
- Lemberg, M. K. and Freeman, M. (2007a). "Cutting proteins within lipid bilayers: rhomboid structure and mechanism". *Molecular Cell* **28** (6): 930-940.
- Lemberg, M. K. and Freeman, M. (2007b). "Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases". *Genome Research* **17** (11): 1634-1646.
- Lemberg, M. K. and Adrain, C. (2016). "Inactive rhomboid proteins: New mechanisms with implications in health and disease". *Seminars in Cell and Developmental Biology* **60** 29-37.
- Lemberg, M. K., Menendez, J., Misik, A., Garcia, M., Koth, C. M. and Freeman, M. (2005). "Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases". *EMBO Journal* **24** (3): 464-472.
- Lemieux, M. J., Fischer, S. J., Cherney, M. M., Bateman, K. S. and James, M. N. (2007). "The crystal structure of the rhomboid peptidase from Haemophilus influenzae provides insight into intramembrane proteolysis". *Proceedings of the National Academy of Sciences of the United States of America* **104** (3): 750-754.
- Li, Q. and Sudhof, T. C. (2004). "Cleavage of amyloid-beta precursor protein and amyloid-beta precursor-like protein by BACE 1". *Journal of Biological Chemistry* **279** (11): 10542-10550.
- Liao, H.-J. and Carpenter, G. (2012). "Regulated Intramembrane Cleavage of the EGF Receptor". *Traffic* **13** (8): 1106-1112.
- Lilley, B. N. and Ploegh, H. L. (2004). "A membrane protein required for dislocation of misfolded proteins from the ER". *Nature* **429** (6994): 834-840.
- Lin, J. W., Meireles, P., Prudencio, M., Engelmann, S., Annoura, T., Sajid, M., Chevalley-Maurel, S., Ramesar, J., Nahar, C., Avramut, C. M., Koster, A. J., Matuschewski, K., Waters,

- A. P., Janse, C. J., Mair, G. R. and Khan, S. M. (2013). "Loss-of-function analyses defines vital and redundant functions of the Plasmodium rhomboid protease family". *Molecular Microbiology* **88** (2): 318-338.
- Liu, J., Han, C., Xie, B., Wu, Y., Liu, S., Chen, K., Xia, M., Zhang, Y., Song, L., Li, Z., Zhang, T., Ma, F., Wang, Q., Wang, J., Deng, K., Zhuang, Y., Wu, X., Yu, Y., Xu, T. and Cao, X. (2014). "Rhbdd3 controls autoimmunity by suppressing the production of IL-6 by dendritic cells via K27-linked ubiquitination of the regulator NEMO". *Nature Immunology* **15** (7): 612-622.
- Liu, L.-P. and Deber, C. M. (1998). "Uncoupling Hydrophobicity and Helicity in Transmembrane Segments". *Journal of Biological Chemistry* **273** (11): 23645–23648.
- Liu, X. N., Tang, Z. H., Zhang, Y., Pan, Q. C., Chen, X. H., Yu, Y. S. and Zang, G. Q. (2013). "Lentivirus-mediated Silencing of Rhomboid Domain Containing 1 Suppresses Tumor Growth and Induces Apoptosis in Hepatoma HepG2 Cells". *Asian Pacific Journal of Cancer Prevention* **14** (1): 5-9.
- Liu, Y., Gerstein, M. and Engelman, D. M. (2004a). "Transmembrane protein domains rarely use covalent domain recombination as an evolutionary mechanism". *Proceedings of the National Academy of Sciences of the United States of America* **101** (10): 3495-3497.
- Liu, Y., Stoll, V. S., Richardson, P. L., Saldivar, A., Klaus, J. L., Molla, A., Kohlbrenner, W. and Kati, W. M. (2004b). "Hepatitis C NS3 protease inhibition by peptidyl-alpha-ketoamide inhibitors: kinetic mechanism and structure". *Archives of Biochemistry and Biophysics* **421** (2): 207-216.
- Lohi, O., Urban, S. and Freeman, M. (2004). "Diverse substrate recognition mechanisms for rhomboids: Thrombomodulin is cleaved by mammalian rhomboids". *Current Biology* **14** (3): 236-241.
- Lopez-Serra, P., Marcilla, M., Villanueva, A., Ramos-Fernandez, A., Palau, A., Leal, L., Wahi, J. E., Setien-Baranda, F., Szczesna, K., Moutinho, C., Martinez-Cardus, A., Heyn, H., Sandoval, J., Puertas, S., Vidal, A., Sanjuan, X., Martinez-Balibrea, E., Vinals, F., Perales, J. C., Bramsem, J. B., Orntoft, T. F., Andersen, C. L., Tabernero, J., McDermott, U., Boxer, M. B., Vander Heiden, M. G., Albar, J. P. and Esteller, M. (2014). "A DERL3-associated defect in the degradation of SLC2A1 mediates the Warburg effect". *Nature Communications* **5** 3608.
- Lu, W., Karuppagounder, S. S., Springer, D. A., Allen, M. D., Zheng, L., Chao, B., Zhang, Y., Dawson, V. L., Dawson, T. M. and Lenardo, M. (2014). "Genetic deficiency of the mitochondrial protein PGAM5 causes a Parkinson's-like movement disorder". *Nature Communications* **5** 4930.
- Maegawa, S., Ito, K. and Akiyama, Y. (2005). "Proteolytic action of GlpG, a rhomboid protease in the Escherichia coli cytoplasmic membrane". *Biochemistry* **44** (41): 13543-13552.
- Mahrus, S., Trinidad, J. C., Barkan, D. T., Sali, A., Burlingame, A. L. and Wells, J. A. (2008). "Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini". *Cell* **134** (5): 866-876.
- Maney, S. K., McIlwain, D. R., Polz, R., Pandyra, A. A., Sundaram, B., Wolff, D., Ohishi, K., Maretzky, T., Brooke, M. A., Evers, A., Vasudevan, A. A. J., Aghaeepour, N., Scheller, J., Münk, C., Häussinger, D., Mak, T. W., Nolan, G. P., Kelsell, D. P., Blobel, C. P., Lang,

- K. S. and Lang, P. A. (2015). "Deletions in the cytoplasmic domain of iRhom1 and iRhom2 promote shedding of the TNF receptor by the protease ADAM17". *Science Signaling* **8** (401):
- Manolaridis, I., Kulkarni, K., Dodd, R. B., Ogasawara, S., Zhang, Z. G., Bineva, G., O'Reilly, N., Hanrahan, S. J., Thompson, A. J., Cronin, N., Iwata, S. and Barford, D. (2013). "Mechanism of farnesylated CAAX protein processing by the intramembrane protease Rce1". *Nature* **504** (7479): 301-+.
- McIlwain, D. R., Lang, P. A., Maretzky, T., Hamada, K., Ohishi, K., Maney, S. K., Berger, T., Murthy, A., Duncan, G., Xu, H. C., Lang, K. S., Haeussinger, D., Wakeham, A., Itie-Youten, A., Khokha, R., Ohashi, P. S., Blobel, C. P. and Mak, T. W. (2012). "iRhom2 Regulation of TACE Controls TNF-Mediated Protection Against Listeria and Responses to LPS". *Science* **335** (6065): 229-232.
- McQuibban, G. A., Saurya, S. and Freeman, M. (2003). "Mitochondrial membrane remodelling regulated by a conserved rhomboid protease". *Nature* **423** (6939): 537-541.
- Mehnert, M., Sommer, T. and Jarosch, E. (2014). "Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membrane". *Nature Cell Biology* **16** (1): 77-86.
- Meissner, C., Lorenz, H., Hehn, B. and Lemberg, M. K. (2015). "Intramembrane protease PARL defines a negative regulator of PINK1- and PARK2/Parkin-dependent mitophagy". *Autophagy* **11** (9): 1484-1498.
- Meissner, C., Lorenz, H., Weihofen, A., Selkoe, D. J. and Lemberg, M. K. (2011). "The mitochondrial intramembrane protease PARL cleaves human Pink1 to regulate Pink1 trafficking". *Journal of Neurochemistry* **117** (5): 856-867.
- Moin, S. M. and Urban, S. (2012). "Membrane immersion allows rhomboid proteases to achieve specificity by reading transmembrane segment dynamics". *eLife* **1** e00173.
- Montoya, J. G. and Liesenfeld, O. (2004). "Toxoplasmosis". *Lancet* **363** (9425): 1965-1976.
- Morino, K., Petersen, K. F., Dufour, S., Befroy, D., Frattini, J., Shatzkes, N., Neschen, S., White, M. F., Bilz, S., Sono, S., Pypaert, M. and Shulman, G. I. (2005). "Reduced mitochondrial density and increased IRS-1 serine phosphorylation in muscle of insulin-resistant offspring of type 2 diabetic parents". *Journal of Clinical Investigation* **115** (12): 3587-3593.
- Neal, S., Jaeger, P. A., Duttke, S. H., Benner, C., C, K. G., Ideker, T. and Hampton, R. Y. (2018). "The Dfm1 Derlin Is Required for ERAD Retrotranslocation of Integral Membrane Proteins". *Molecular Cell* **69** (2): 306-320 e304.
- Nguyen, T. N., Padman, B. S. and Lazarou, M. (2016). "Deciphering the Molecular Signals of PINK1/Parkin Mitophagy". *Trends in Cell Biology* **26** (10): 733-744.
- Nickel, W. and Rabouille, C. (2009). "Mechanisms of regulated unconventional protein secretion". *Nature Reviews Molecular Cell Biology* **10** (2): 148-155.
- Njoroge, F. G., Chen, K. X., Shih, N. Y. and Piwinski, J. J. (2008). "Challenges in modern drug discovery: A case study of boceprevir, an HCV protease inhibitor for the treatment of hepatitis C virus infection". *Accounts of Chemical Research* **41** (1): 50-59.
- Noy, P. J., Swain, R. K., Khan, K., Lodhia, P. and Bicknell, R. (2016). "Sprouting angiogenesis is regulated by shedding of the C-type lectin family 14, member A (CLEC14A)

ectodomain, catalyzed by rhomboid-like 2 protein (RHBDL2)". *Faseb Journal* **30** (6): 2311-2323.

O'Donnell, R. A., Hackett, F., Howell, S. A., Treeck, M., Struck, N., Krnajski, Z., Withers-Martinez, C., Gilberger, T. W. and Blackman, M. J. (2006). "Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite". *Journal of Cell Biology* **174** (7): 1023-1033.

O'Donoghue, A. J., Eroy-Reveles, A. A., Knudsen, G. M., Ingram, J., Zhou, M., Statnekov, J. B., Greninger, A. L., Hostetter, D. R., Qu, G., Maltby, D. A., Anderson, M. O., Derisi, J. L., McKerrow, J. H., Burlingame, A. L. and Craik, C. S. (2012). "Global identification of peptidase specificity by multiplex substrate profiling". *Nature Methods* **9** (11): 1095-1100.

Oda, Y., Okada, T., Yoshida, H., Kaufman, R. J., Nagata, K. and Mori, K. (2006). "Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation". *Journal of Cell Biology* **172** (3): 383-393.

Oikonomidi, I., Burbridge, E., Cavadas, M., Sullivan, G., Collis, B., Naegele, H., Clancy, D., Brezinova, J., Hu, T., Bileck, A., Gerner, C., Bolado, A., von Kriegsheim, A., Martin, S. J., Steinberg, F., Strisovsky, K. and Adrain, C. (2018). "iTAP, a novel iRhom interactor, controls TNF secretion by policing the stability of iRhom/TACE". *eLife* **7**

Okatsu, K., Oka, T., Iguchi, M., Imamura, K., Kosako, H., Tani, N., Kimura, M., Go, E., Koyano, F., Funayama, M., Shiba-Fukushima, K., Sato, S., Shimizu, H., Fukunaga, Y., Taniguchi, H., Komatsu, M., Hattori, N., Mihara, K., Tanaka, K. and Matsuda, N. (2012). "PINK1 autophosphorylation upon membrane potential dissipation is essential for Parkin recruitment to damaged mitochondria". *Nature Communications* **3** 1016.

Olivieri, A., Collins, C. R., Hackett, F., Withers-Martinez, C., Marshall, J., Flynn, H. R., Skehel, J. M. and Blackman, M. J. (2011). "Juxtamembrane Shedding of Plasmodium falciparum AMA1 Is Sequence Independent and Essential, and Helps Evade Invasion-Inhibitory Antibodies". *PLoS Pathogens* **7** (12):

Opitz, C., Di Cristina, M., Reiss, M., Ruppert, T., Crisanti, A. and Soldati, D. (2002). "Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite *Toxoplasma gondii*". *EMBO Journal* **21** (7): 1577-1585.

Parussini, F., Tang, Q., Moin, S. M., Mital, J., Urban, S. and Ward, G. E. (2012). "Intramembrane proteolysis of *Toxoplasma* apical membrane antigen 1 facilitates host-cell invasion but is dispensable for replication". *Proceedings of the National Academy of Sciences of the United States of America* **109** (19): 7463-7468.

Pascall, J. C. and Brown, K. D. (2004). "Intramembrane cleavage of ephrinB3 by the human rhomboid family protease, RHBDL2". *Biochemical and Biophysical Research Communications* **317** (1): 244-252.

Paschkowsky, S., Hamze, M., Oestereich, F. and Munter, L. M. (2016). "Alternative Processing of the Amyloid Precursor Protein Family by Rhomboid Protease RHBDL4". *Journal of Biological Chemistry* **291** (42): 21903-21912.

Paschkowsky, S., Recinto, S. J., Young, J. C., Bondar, A. N. and Munter, L. M. (2018). "Membrane cholesterol as regulator of human rhomboid protease RHBDL4". *Journal of Biological Chemistry* **293** (40): 15556-15568.

Pellegrini, L., Passer, B. J., Canelles, M., Lefterov, I., Ganjei, J. K., Fowlkes, B. J., Koonin, E. V. and D'Adamio, L. (2001). "PAMP and PARL, two novel putative metalloproteases interacting with the COOH-terminus of Presenilin-1 and -2". *Journal of Alzheimer's Disease* **3** (2): 181-190.

Perona, J. J. and Craik, C. S. (1997). "Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold". *Journal of Biological Chemistry* **272** (48): 29987-29990.

Peschon, J. J., Slack, J. L., Reddy, P., Stocking, K. L., Sunnarborg, S. W., Lee, D. C., Russell, W. E., Castner, B. J., Johnson, R. S., Fitzner, J. N., Boyce, R. W., Nelson, N., Kozlosky, C. J., Wolfson, M. F., Rauch, C. T., Cerretti, D. P., Paxton, R. J., March, C. J. and Black, R. A. (1998). "An essential role for ectodomain shedding in mammalian development". *Science* **282** (5392): 1281-1284.

Petersen, K. F., Befroy, D., Dufour, S., Dziura, J., Ariyan, C., Rothman, D. L., DiPietro, L., Cline, G. W. and Shulman, G. I. (2003). "Mitochondrial dysfunction in the elderly: possible role in insulin resistance". *Science* **300** (5622): 1140-1142.

Pierrat, O. A., Strisovsky, K., Christova, Y., Large, J., Ansell, K., Bouloc, N., Smiljanic, E. and Freeman, M. (2011). "Monocyclic beta-Lactams Are Selective, Mechanism-Based Inhibitors of Rhomboid Intramembrane Proteases". *ACS Chemical Biology* **6** (4): 325-335.

Powers, J. C., Asgian, J. L., Ekici, O. D. and James, K. E. (2002). "Irreversible inhibitors of serine, cysteine, and threonine proteases". *Chemical Reviews* **102** (12): 4639-4750.

Rawson, R. B., Zelenski, N. G., Nijhawan, D., Ye, J., Sakai, J., Hasan, M. T., Chang, T. Y., Brown, M. S. and Goldstein, J. L. (1997). "Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs". *Molecular Cell* **1** (1): 47-57.

Reddy, T. and Rainey, J. K. (2012). "Multifaceted Substrate Capture Scheme of a Rhomboid Protease". *Journal of Physical Chemistry B* **116** (30): 8942-8954.

Ren, X., Song, W., Liu, W., Guan, X., Miao, F., Miao, S. and Wang, L. (2013). "Rhomboid domain containing 1 inhibits cell apoptosis by upregulating AP-1 activity and its downstream target Bcl-3". *FEBS Letters* **587** (12): 1793-1798.

Rogaeva, E., Johnson, J., Lang, A. E., Gulick, C., Gwinn-Hardy, K., Kawarai, T., Sato, C., Morgan, A., Werner, J., Nussbaum, R., Petit, A., Okun, M. S., McInerney, A., Mandel, R., Groen, J. L., Fernandez, H. H., Postuma, R., Foote, K. D., Salehi-Rad, S., Liang, Y., Reimsnider, S., Tandon, A., Hardy, J., St George-Hyslop, P. and Singleton, A. B. (2004). "Analysis of the PINK1 gene in a large cohort of cases with Parkinson disease". *Archives of Neurology* **61** (12): 1898-1904.

Ruiz, N., Falcone, B., Kahne, D. and Silhavy, T. J. (2005). "Chemical conditionality: a genetic strategy to probe organelle assembly". *Cell* **121** (2): 307-317.

Sahin, U., Weskamp, G., Kelly, K., Zhou, H. M., Higashiyama, S., Peschon, J., Hartmann, D., Saftig, P. and Blobel, C. P. (2004). "Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR ligands". *Journal of Cell Biology* **164** (5): 769-779.

Sargent, F., Bogsch, E. G., Stanley, N. R., Wexler, M., Robinson, C., Berks, B. C. and Palmer, T. (1998). "Overlapping functions of components of a bacterial Sec-independent protein export pathway". *EMBO Journal* **17** (13): 3640-3650.

- Sekine, S., Kanamaru, Y., Koike, M., Nishihara, A., Okada, M., Kinoshita, H., Kamiyama, M., Maruyama, J., Uchiyama, Y., Ishihara, N., Takeda, K. and Ichijo, H. (2012). "Rhomboid Protease PARL Mediates the Mitochondrial Membrane Potential Loss-induced Cleavage of PGAM5". *Journal of Biological Chemistry* **287** (41): 34635-34645.
- Sherratt, A. R., Blais, D. R., Ghasriani, H., Pezacki, J. P. and Goto, N. K. (2012). "Activity-based protein profiling of the Escherichia coli GlpG rhomboid protein delineates the catalytic core". *Biochemistry* **51** (39): 7794-7803.
- Shi, G., Lee, J., Ye, D. and McQuibban, A. (2011). "Characterizing the function of PARL in PINK1 regulation". *Faseb Journal* **25**
- Shi, G., Lee, J. R., Grimes, D. A., Racacho, L., Ye, D., Yang, H., Ross, O. A., Farrer, M., McQuibban, G. A. and Bulman, D. E. (2011). "Functional alteration of PARL contributes to mitochondrial dysregulation in Parkinson's disease". *Human Molecular Genetics* **20** (10): 1966-1974.
- Shokhen, M. and Albeck, A. (2017). "How does the exosite of rhomboid protease affect substrate processing and inhibition?". *Protein Science* **26** (12): 2355-2366.
- Schapira, A. H. (2008). "Mitochondria in the aetiology and pathogenesis of Parkinson's disease". *Lancet Neurology* **7** (1): 97-109.
- Schechter, I. and Berger, A. (1967). "On the size of the active site in proteases. I. Papain". *Biochemical and Biophysical Research Communications* **27** (2): 157-162.
- Schlondorff, J., Becherer, J. D. and Blobel, C. P. (2000). "Intracellular maturation and localization of the tumour necrosis factor alpha convertase (TACE)". *Biochemical Journal* **347 Pt 1** 131-138.
- Sibley, L. D. (2004). "Intracellular parasite invasion strategies". *Science* **304** (5668): 248-253.
- Simeonov, A., Jadhav, A., Thomas, C. J., Wang, Y., Huang, R., Southall, N. T., Shinn, P., Smith, J., Austin, C. P., Auld, D. S. and Inglese, J. (2008). "Fluorescence spectroscopic profiling of compound libraries". *Journal of Medicinal Chemistry* **51** (8): 2363-2371.
- Singh, N., Joshi, R. and Komurov, K. (2015). "HER2-mTOR signaling-driven breast cancer cells require ER-associated degradation to survive". *Science Signaling* **8** (378): ra52.
- Soldati, D., Foth, B. J. and Cowman, A. F. (2004). "Molecular and functional aspects of parasite invasion". *Trends in Parasitology* **20** (12): 567-574.
- Song, W., Liu, W. J., Zhao, H., Li, S. Z., Guan, X., Ying, J. M., Zhang, Y. F., Miao, F., Zhang, M. M., Ren, X. X., Li, X. L., Wu, F., Zhao, Y. C., Tian, Y. Y., Wu, W. M., Fu, J., Liang, J. B., Wu, W., Liu, C. Z., Yu, J., Zong, S. D., Miao, S. Y., Zhang, X. D. and Wang, L. F. (2015). "Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway". *Nature Communications* **6**
- Spinazzi, M. and De Strooper, B. (2016). "PARL: The mitochondrial rhomboid protease". *Seminars in Cell & Developmental Biology* **60** 19-28.
- Stevenson, L. G., Strisovsky, K., Clemmer, K. M., Bhatt, S., Freeman, M. and Rather, P. N. (2007). "Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase". *Proceedings of the National Academy of Sciences of the United States of America* **104** (3): 1003-1008.

- Strisovsky, K. (2013). "Structural and mechanistic principles of intramembrane proteolysis lessons from rhomboids". *FEBS Journal* **280** (7): 1579-1603.
- Strisovsky, K. (2016). "Rhomboid protease inhibitors: Emerging tools and future therapeutics". *Seminars in Cell & Developmental Biology* **60** 52-62.
- Strisovsky, K., Sharpe, H. J. and Freeman, M. (2009). "Sequence-Specific Intramembrane Proteolysis: Identification of a Recognition Motif in Rhomboid Substrates". *Molecular Cell* **36** (6): 1048-1059.
- Tatsuta, T., Augustin, S., Nolden, M., Friedrichs, B. and Langer, T. (2007). "m-AAA protease-driven membrane dislocation allows intramembrane cleavage by rhomboid in mitochondria". *EMBO Journal* **26** (2): 325-335.
- Teshima, T., Griffin, J. C. and Powers, J. C. (1982). "A new class of heterocyclic serine protease inhibitors. Inhibition of human leukocyte elastase, porcine pancreatic elastase, cathepsin G, and bovine chymotrypsin A alpha with substituted benzoxazinones, quinazolines, and anthranilates". *Journal of Biological Chemistry* **257** (9): 5085-5091.
- Ticha, A., Collis, B. and Strisovsky, K. (2018). "The Rhomboid Superfamily: Structural Mechanisms and Chemical Biology Opportunities". *Trends in Biochemical Sciences* **43** (9): 726-739.
- Ticha, A., Stanchev, S., Skerle, J., Began, J., Ingr, M., Svehlova, K., Polovinkin, L., Ruzicka, M., Bednarova, L., Hadravova, R., Polachova, E., Rampirova, P., Brezinova, J., Kasicka, V., Majer, P. and Strisovsky, K. (2017). "Sensitive Versatile Fluorogenic Transmembrane Peptide Substrates for Rhomboid Intramembrane Proteases". *Journal of Biological Chemistry* **292** (7): 2703-2713.
- Tsruya, R., Wojtalla, A., Carmon, S., Yogeve, S., Reich, A., Bibi, E., Merdes, G., Schejter, E. and Shilo, B. Z. (2007). "Rhomboid cleaves Star to regulate the levels of secreted Spitz". *EMBO Journal* **26** (5): 1211-1220.
- Tyndall, J. D. A., Nall, T. and Fairlie, D. P. (2005). "Proteases universally recognize beta strands in their active sites". *Chemical Reviews* **105** (3): 973-999.
- Ukropcova, B., Sereda, O., de Jonge, L., Bogacka, I., Nguyen, T., Xie, H., Bray, G. A. and Smith, S. R. (2007). "Family history of diabetes links impaired substrate switching and reduced mitochondrial content in skeletal muscle". *Diabetes* **56** (3): 720-727.
- Urban, S. (2009). "Making the cut: central roles of intramembrane proteolysis in pathogenic microorganisms". *Nature Reviews Microbiology* **7** (6): 411-423.
- Urban, S. and Freeman, M. (2002). "Intramembrane proteolysis controls diverse signalling pathways throughout evolution". *Current Opinion in Genetics & Development* **12** (5): 512-518.
- Urban, S. and Freeman, M. (2003). "Substrate Specificity of Rhomboid Intramembrane Proteases Is Governed by Helix-Breaking Residues in the Substrate Transmembrane Domain". *Molecular Cell* **11** (6): 1425-1434.
- Urban, S. and Wolfe, M. S. (2005). "Reconstitution of intramembrane proteolysis in vitro reveals that pure rhomboid is sufficient for catalysis and specificity". *Proceedings of the National Academy of Sciences of the United States of America* **102** (6): 1883-1888.

- Urban, S. and Baker, R. P. (2008). "In vivo analysis reveals substrate-gating mutants of a rhomboid intramembrane protease display increased activity in living cells". *Biological Chemistry* **389** (8): 1107-1115.
- Urban, S. and Moin, S. M. (2014). "A subset of membrane-altering agents and gamma-secretase modulators provoke nonsubstrate cleavage by rhomboid proteases". *Cell Reports* **8** (5): 1241-1247.
- Urban, S., Lee, J. R. and Freeman, M. (2001). "Drosophila Rhomboid-1 defines a family of putative intramembrane serine proteases". *Cell* **107** (2): 173-182.
- Urban, S., Schlieper, D. and Freeman, M. (2002). "Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids". *Current Biology* **12** (17): 1507-1512.
- Uritsky, N., Shokhen, M. and Albeck, A. (2012). "The Catalytic Machinery of Rhomboid Proteases: Combined MD and QM Simulations". *Journal of Chemical Theory and Computation* **8** (11): 4663-4671.
- Uritsky, N., Shokhen, M. and Albeck, A. (2016). "Stepwise Versus Concerted Mechanisms in General-Base Catalysis by Serine Proteases". *Angewandte Chemie-International Edition* **55** (5): 1680-1684.
- Vinothkumar, K. R. (2011). "Structure of Rhomboid Protease in a Lipid Environment". *Journal of Molecular Biology* **407** (2): 232-247.
- Vinothkumar, K. R., Pierrat, O. A., Large, J. M. and Freeman, M. (2013). "Structure of Rhomboid Protease in Complex with beta-Lactam Inhibitors Defines the S2 ' Cavity". *Structure* **21** (6): 1051-1058.
- Vinothkumar, K. R., Strisovsky, K., Andreeva, A., Christova, Y., Verhelst, S. and Freeman, M. (2010). "The structural basis for catalysis and substrate specificity of a rhomboid protease". *EMBO Journal* **29** (22): 3797-3809.
- Viola, K. L. and Klein, W. L. (2015). "Amyloid beta oligomers in Alzheimer's disease pathogenesis, treatment, and diagnosis". *Acta Neuropathologica* **129** (2): 183-206.
- von Heijne, G. (2006). "Membrane-protein topology". *Nature Reviews Molecular Cell Biology* **7** (12): 909-918.
- Vosyka, O., Vinothkumar, K. R., Wolf, E. V., Brouwer, A. J., Liskamp, R. M. and Verhelst, S. H. (2013). "Activity-based probes for rhomboid proteases discovered in a mass spectrometry-based assay". *Proceedings of the National Academy of Sciences of the United States of America* **110** (7): 2472-2477.
- Walder, K., Kerr-Bayles, L., Civitarese, A., Jowett, J., Curran, J., Elliott, K., Trevaskis, J., Bishara, N., Zimmet, P., Mandarino, L., Ravussin, E., Blangero, J., Kissebah, A. and Collier, G. R. (2005). "The mitochondrial rhomboid protease PSARL is a new candidate gene for type 2 diabetes". *Diabetologia* **48** (3): 459-468.
- Walker, B. and Lynas, J. F. (2001). "Strategies for the inhibition of serine proteases". *Cellular and Molecular Life Sciences* **58** (4): 596-624.
- Wang, Y. and Ha, Y. (2007). "Open-cap conformation of intramembrane protease GlpG". *Proceedings of the National Academy of Sciences of the United States of America* **104** (7): 2098-2102.

- Wang, Y., Zhang, Y. and Ha, Y. (2006). "Crystal structure of a rhomboid family intramembrane protease". *Nature* **444** (7116): 179-180.
- Wang, Y., Guan, X., Fok, K. L., Li, S., Zhang, X., Miao, S., Zong, S., Koide, S. S., Chan, H. C. and Wang, L. (2008). "A novel member of the Rhomboid family, RHBDD1, regulates BIK-mediated apoptosis". *Cellular and Molecular Life Sciences* **65** (23): 3822-3829.
- Wang, Y. C., Maegawa, S., Akiyama, Y. and Ha, Y. (2007). "The role of L1 loop in the mechanism of rhomboid intramembrane protease GlpG". *Journal of Molecular Biology* **374** (4): 1104-1113.
- White, S. H. and Wimley, W. C. (1999). "Membrane protein folding and stability: physical principles". *Annual Review of Biophysics and Biomolecular Structure* **28** 319-365.
- Wolf, E. V. and Verhelst, S. H. (2016). "Inhibitors of rhomboid proteases". *Biochimie* **122** 38-47.
- Wolf, E. V., Zeissler, A. and Verhelst, S. H. (2015a). "Inhibitor Fingerprinting of Rhomboid Proteases by Activity-Based Protein Profiling Reveals Inhibitor Selectivity and Rhomboid Autoprocessing". *ACS Chemical Biology* **10** (10): 2325-2333.
- Wolf, E. V., Seybold, M., Hadravova, R., Strisovsky, K. and Verhelst, S. H. (2015b). "Activity-Based Protein Profiling of Rhomboid Proteases in Liposomes". *Chembiochem* **16** (11): 1616-1621.
- Wolf, E. V., Zeissler, A., Vosyka, O., Zeiler, E., Sieber, S. and Verhelst, S. H. (2013). "A new class of rhomboid protease inhibitors discovered by activity-based fluorescence polarization". *PLoS One* **8** (8): e72307.
- Wolfe, M. S., Xia, W. M., Ostaszewski, B. L., Diehl, T. S., Kimberly, W. T. and Selkoe, D. J. (1999). "Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity". *Nature* **398** (6727): 513-517.
- Wu, Z., Yan, N., Feng, L., Oberstein, A., Yan, H., Baker, R. P., Gu, L., Jeffrey, P. D., Urban, S. and Shi, Y. (2006). "Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry". *Nature Structural & Molecular Biology* **13** (12): 1084-1091.
- Wunderle, L., Knopf, J. D., Kuhnle, N., Morle, A., Hehn, B., Adrain, C., Strisovsky, K., Freeman, M. and Lemberg, M. K. (2016). "Rhomboid intramembrane protease RHBDL4 triggers ER-export and non-canonical secretion of membrane-anchored TGF alpha". *Scientific Reports* **6**
- Wust, R., Maurer, B., Hauser, K., Woitalla, D., Sharma, M. and Kruger, R. (2016). "Mutation analyses and association studies to assess the role of the presenilin-associated rhomboid-like gene in Parkinson's disease". *Neurobiology of Aging* **39** 217 e213-215.
- Xue, Y. and Ha, Y. (2012). "Catalytic Mechanism of Rhomboid Protease GlpG Probed by 3,4-Dichloroisocoumarin and Diisopropyl Fluorophosphate". *Journal of Biological Chemistry* **287** (5): 3099-3107.
- Xue, Y. and Ha, Y. (2013). "Large Lateral Movement of Transmembrane Helix S5 Is Not Required for Substrate Access to the Active Site of Rhomboid Intramembrane Protease". *Journal of Biological Chemistry* **288** (23): 16645-16654.

- Xue, Y., Chowdhury, S., Liu, X., Akiyama, Y., Ellman, J. and Ha, Y. (2012). "Conformational Change in Rhomboid Protease GlpG Induced by Inhibitor Binding to Its S' Subsites". *Biochemistry* **51** (18): 3723-3731.
- Yamano, K. and Youle, R. J. (2013). "PINK1 is degraded through the N-end rule pathway". *Autophagy* **9** (11): 1758-1769.
- Yang, J., Barniol-Xicota, M., Nguyen, M. T. N., Ticha, A., Strisovsky, K. and Verhelst, S. H. L. (2018). "Benzoxazin-4-ones as novel, easily accessible inhibitors for rhomboid proteases". *Bioorganic & Medicinal Chemistry Letters* **28** (8): 1423-1427.
- Ye, Y. H., Shibata, Y., Yun, C., Ron, D. and Rapoport, T. A. (2004). "A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol". *Nature* **429** (6994): 841-847.
- Youle, R. J. and Narendra, D. P. (2011). "Mechanisms of mitophagy". *Nature Reviews Molecular Cell Biology* **12** (1): 9-14.
- Zettl, M., Adrain, C., Strisovsky, K., Lastun, V. and Freeman, M. (2011). "Rhomboid Family Pseudoproteases Use the ER Quality Control Machinery to Regulate Intercellular Signaling". *Cell* **145** (1): 79-91.
- Zhou, Y., Moin, S. M., Urban, S. and Zhang, Y. (2012). "An internal water-retention site in the rhomboid intramembrane protease GlpG ensures catalytic efficiency". *Structure* **20** (7): 1255-1263.
- Zimmerman, M., Morman, H., Mulvey, D., Jones, H., Frankshun, R. and Ashe, B. M. (1980). "Inhibition of elastase and other serine proteases by heterocyclic acylating agents". *Journal of Biological Chemistry* **255** (20): 9848-9851.
- Zoll, S., Stanchev, S., Began, J., Skerle, J., Lepsik, M., Peclinovska, L., Majer, P. and Strisovsky, K. (2014). "Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures". *EMBO Journal* **33** (20): 2408-2421.

8 LIST OF PUBLICATIONS

Publications related to dissertation project

a) Research articles

1. Yang, J., Barniol-Xicota, M., Nguyen, M. T. N., **Ticha, A.**, Strisovsky, K., Verhelst, S. H. L. (2018): Benzoxazin-4-ones as novel, easily accessible inhibitors for rhomboid proteases. *Bioorganic and Medicinal Chemistry Letters* **28** (8), 1423-1427. **IF(2017) = 2.442**
2. Goel, P., Jumpertz, T., **Ticha, A.**, Ogorek, I., Mikles, D. C., Hubalek, M., Pietrzik, C. U., Strisovsky, K., Schmidt, B., Weggen, S. (2018): Discovery and validation of 2-styryl substituted benzoxazin-4-ones as a novel scaffold for rhomboid protease inhibitors. *Bioorganic and Medicinal Chemistry Letters* **28** (8), 1417-1422. **IF(2017) = 2.442**
3. **Ticha, A.**, Stanchev, S., Vinothkumar, K. R., Mikles, D. C., Pachel, P., Began, J., Skerle, J., Svehlova, K., Nguyen, M. T. N., Verhelst, S. H. L., Johnson, D. C., Bachovchin, D. A., Lepsik, M., Majer, P., Strisovsky, K. (2017): General and Modular Strategy for Designing Potent, Selective, and Pharmacologically Compliant Inhibitors of Rhomboid Proteases. *Cell Chemical Biology* **24** (12), 1523-1536. **IF(2017) = 5.592**
4. **Ticha, A.**, Stanchev, S., Skerle, J., Began, J., Ingr, M., Svehlova, K., Polovinkin, L., Ruzicka, M., Bednarova, L., Hadravova, R., Polachova, E., Rampirova, P., Brezinova, J., Kasicka, V., Majer, P., Strisovsky, K. (2017): Sensitive Versatile Fluorogenic Transmembrane Peptide Substrates for Rhomboid Intramembrane Proteases. *Journal of Biological Chemistry* **292** (7), 2703-2713. **IF(2017) = 4.010**
5. Goel, P., Jumpertz, T., Mikles, D. C., **Ticha, A.**, Nguyen, M. T. N., Verhelst, S., Hubalek, M., Johnson, D. C., Bachovchin, D. A., Ogorek, I., Pietrzik, C. U., Strisovsky, K., Schmidt, B., Weggen, S. (2017): Discovery and Biological Evaluation of Potent and Selective N-Methylene Saccharin-Derived Inhibitors for Rhomboid Intramembrane Proteases. *Biochemistry* **56** (51), 6713-6725. **IF(2017) = 2.997**

b) Review articles

- Ticha, A.**, Collis, B., Strisovsky, K. (2018): The Rhomboid Superfamily: Structural Mechanisms and Chemical Biology Opportunities. *Trends in Biochemical Sciences* **43** (9), 726-739. **IF(2017) = 15.678**

Other publications

1. Prazienkova, V., **Ticha, A.**, Blechova, M., Spolcova, A., Zelezna, B., Maletinska, L. (2016): Pharmacological characterization of lipidized analogs of prolactin-releasing peptide with a modified C-terminal aromatic ring. *Journal of Physiology and Pharmacology* **67** (1), 121-128.
2. Maletinska, L., Nagelova, V., **Ticha, A.**, Zemenova, J., Pirnik, Z., Holubova, M., Spolcova, A., Mikulaskova, B., Blechova, M., Sykora, D., Lacinova, Z., Haluzik, M., Zelezna, B., Kunes, J. (2015): Novel lipidized analogs of prolactin-releasing peptide have prolonged half-lives and exert anti-obesity effects after peripheral administration. *International Journal of Obesity* **39** (6), 986-993, doi:10.1038/ijo.2015.28.
3. Maletinska, L., **Ticha, A.**, Nagelova, V., Spolcova, A., Blechova, M., Elbert, T., Zelezna, B. (2013): Neuropeptide FF analog RF9 is not an antagonist of NPFF receptor and decreases food intake in mice after its central and peripheral administration. *Brain Research* **1498**, 33-40, doi:10.1016/j.brainres.2012.12.037.

9 APPENDICES

APPENDIX 1.....85

Ticha, A., Stanchev, S., Skerle, J., Began, J., Ingr, M., Svehlova, K., Polovinkin, L., Ruzicka, M., Bednarova, L., Hadravova, R., Polachova, E., Rampirova, P., Brezinova, J., Kasicka, V., Majer, P., Strisovsky, K. (2017): Sensitive Versatile Fluorogenic Transmembrane Peptide Substrates for Rhomboid Intramembrane Proteases. *Journal of Biological Chemistry* **292** (7), 2703-2713. **IF(2017) = 4.010**

APPENDIX 2.....96

Ticha, A., Stanchev, S., Vinothkumar, K. R., Mikles, D. C., Pahl, P., Began, J., Skerle, J., Svehlova, K., Nguyen, M. T. N., Verhelst, S. H. L., Johnson, D. C., Bachovchin, D. A., Lepsik, M., Majer, P., Strisovsky, K. (2017): General and Modular Strategy for Designing Potent, Selective, and Pharmacologically Compliant Inhibitors of Rhomboid Proteases. *Cell Chemical Biology* **24** (12), 1523-1536. **IF(2017) = 5.592**

APPENDIX 3.....114

Goel, P., Jumpertz, T., Mikles, D. C., **Ticha, A.**, Nguyen, M. T. N., Verhelst, S., Hubalek, M., Johnson, D. C., Bachovchin, D. A., Ogorek, I., Pietrzik, C. U., Strisovsky, K., Schmidt, B., Weggen, S. (2017): Discovery and Biological Evaluation of Potent and Selective N-Methylene Saccharin-Derived Inhibitors for Rhomboid Intramembrane Proteases. *Biochemistry* **56** (51), 6713-6725. **IF(2017) = 2.997**

APPENDIX 4.....127

Goel, P., Jumpertz, T., **Ticha, A.**, Ogorek, I., Mikles, D. C., Hubalek, M., Pietrzik, C. U., Strisovsky, K., Schmidt, B., Weggen, S. (2018): Discovery and validation of 2-styryl substituted benzoxazin-4-ones as a novel scaffold for rhomboid protease inhibitors. *Bioorganic and Medicinal Chemistry Letters* **28** (8), 1417-1422. **IF(2017) = 2.442**

Appendix 1



THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 292, NO. 7, PP. 2703–2713, FEBRUARY 17, 2017
© 2017 BY THE AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY, INC. PUBLISHED IN THE U.S.A.

Sensitive Versatile Fluorogenic Transmembrane Peptide Substrates for Rhomboid Intramembrane Proteases*[§]

Received for publication, October 14, 2016, and in revised form, January 5, 2017. Published, JBC Papers in Press, January 9, 2017, DOI 10.1074/jbc.M116.762849

Anežka Tichá^{‡§1}, Stancho Stanchev^{‡1}, Jan Škerle^{‡4,7}, Jakub Began^{‡1,3}, Marek Ingr^{‡**4}, Kateřina Švehlová[‡], Lucie Polovínkin^{‡†5}, Martin Růžička^{‡†6}, Lucie Bednářová[‡], Romana Hadravová[‡], Edita Poláčková[‡], Petra Rampírová[‡], Jana Březinová[‡], Václav Kašička^{‡6}, Pavel Majer[‡], and Kvido Strisovský^{‡7}

From the [‡]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Science, Flemingovo n. 2, Prague 166 10, the [†]Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, Prague 128 43, the [‡]Department of Genetics and Microbiology, Faculty of Science, Charles University, Viničná 5, Prague 128 44, the [§]First Faculty of Medicine, Charles University, Kateřinská 32, Prague 121 08, and the ^{**}Department of Physics and Materials Engineering, Tomas Bata University in Zlín, Faculty of Technology, nám. T.G. Masaryka 5555, 76001, Zlín, Czech Republic

Edited by George N. DeMartino

Rhomboid proteases are increasingly being explored as potential drug targets, but their potent and specific inhibitors are not available, and strategies for inhibitor development are hampered by the lack of widely usable and easily modifiable *in vitro* activity assays. Here we address this bottleneck and report on the development of new fluorogenic transmembrane peptide substrates, which are cleaved by several unrelated rhomboid proteases, can be used both in detergent micelles and in liposomes, and contain red-shifted fluorophores that are suitable for high-throughput screening of compound libraries. We show that nearly the entire transmembrane domain of the substrate is important for efficient cleavage, implying that it extensively interacts with the enzyme. Importantly, we demonstrate that in the detergent micelle system, commonly used for the enzymatic analyses of intramembrane proteolysis, the cleavage rate strongly depends on detergent concentration, because the reaction proceeds only in the micelles. Furthermore, we show that the catalytic efficiency and selectivity toward a rhomboid substrate can be dramatically improved by targeted modification of the sequence of its P5 to P1 region. The fluorogenic substrates that we describe and their sequence variants should find wide use in the detection of activity and development of inhibitors of rhomboid proteases.

Rhomboid intramembrane proteases are evolutionarily widespread and regulate important biological processes including growth factor secretion (1, 2), mitochondrial dynamics (3), invasion of the malaria parasite (4), and membrane protein quality control (5). Rhomboid proteases are increasingly being explored as potential drug targets (6–9), but their selective and potent inhibitors are lacking (reviewed in Ref. 10). Rhomboid inhibitor discovery and development are complicated by the lack of widely usable and easily modifiable *in vitro* activity assays.

Rhomboid activity assays have traditionally relied on recombinant transmembrane protein substrates and gel-based readouts, but such assays are unsuitable for high-throughput screening. A fluorogenic substrate for the *Providencia stuartii* rhomboid protease AarA lacking most of the transmembrane domain of the parent substrate Gurken is cleaved very poorly by other rhomboids including the main model rhomboid protease GlpG of *Escherichia coli* (11). Other published variants of fluorogenic substrates can be used only in liposomes (12) or involve large fluorescent protein moieties making them dependent on expression in a biological system and photochemically less variable (13), which may be important for high-throughput screening of compound libraries where bright red-shifted fluorophores are preferred (14). Moreover, each of the described rhomboid substrates has been used only with one or two related rhomboid proteases, and a strategy to design widely usable or specific substrates has been lacking. Other types of activity assays employing MALDI mass spectrometry (15) and fluorescence polarization (16) have been reported, but MALDI is a low-throughput method that requires sophisticated instrumentation, and fluorescence polarization assays are based on competition of small molecular activity probes with inhibitors and are prone to detergent artifacts (16), making both of these methods unfit for routine kinetics measurements or high-throughput screening.

In view of these limitations, we have sought to develop a robust fluorogenic transmembrane peptide substrate platform for continuous activity assays that would capture all the native enzyme-substrate interactions, be applicable to both the detergent micelle system and liposomes, and would be easily adapt-

* This work was supported by EMBO Installation Grant 2329, Ministry of Education, Youth and Sports of the Czech Republic Projects LK11206 and LO1302, Marie Curie Career Integration Grant Project 304154 (to K. S.), and National Subvention for Development of Research Organizations RVO: 61388963 to the Institute of Organic Chemistry and Biochemistry. The authors declare that they have no conflicts of interest with the contents of this article.

[§] This article contains supplemental information.

¹ Both authors contributed equally to the results of this work.

² Supported by Ph.D. project number 232313 from the Grant Agency of Charles University (GA UK) in Prague.

³ Supported by Ph.D. project number 170214 from the Grant Agency of Charles University (GA UK) in Prague.

⁴ Supported by the Czech Science Foundation Grant P208-12-G016 (Center of Excellence).

⁵ Present address: Institut de Biologie Structurale, 71 avenue des Martyrs, Grenoble, 38044, France.

⁶ Supported by the Czech Science Foundation Grant 15-01948S.

⁷ Recipient of the Purkyne Fellowship of the Academy of Sciences of the Czech Republic. To whom correspondence should be addressed: Institute of Organic Chemistry and Biochemistry, Flemingovo n. 2, Prague, 166 10, Czech Republic. Tel.: 420-220-183-468; E-mail: kvido.strisovsky@uochb.cas.cz.

Fluorogenic Substrates for Rhomboid Proteases

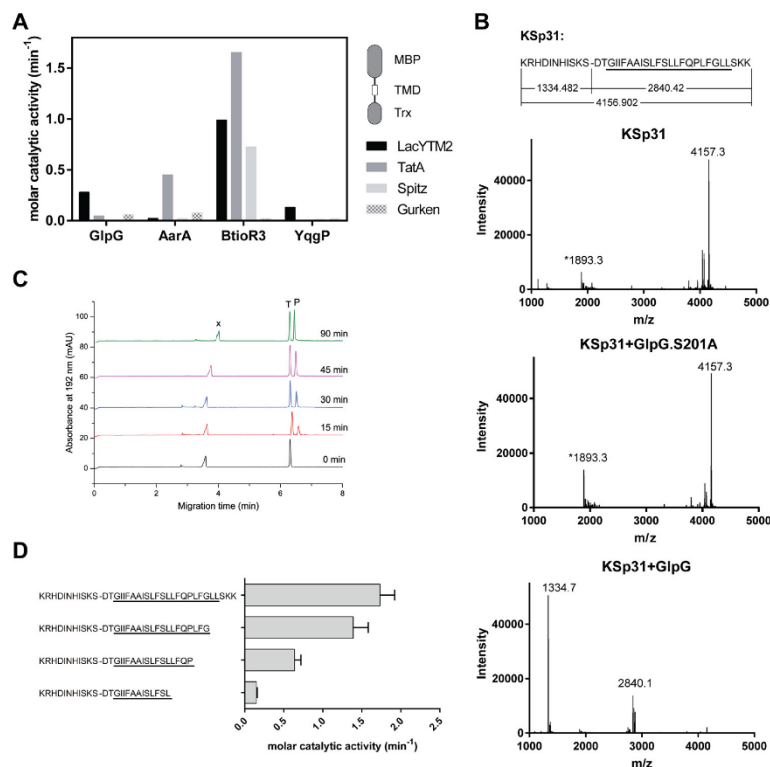


FIGURE 1. Identification of a widely accepted transmembrane substrate for rhomboid proteases. *A*, comparison of cleavage efficiency of model substrates LacYTM2, Gurken, TatA, and Spitz by bacterial rhomboid proteases GlpG (*E. coli*), AarA (*P. stuartii*), YqgP (*B. subtilis*), and BtioR3 (*B. thetaiotaomicron*) *in vitro*. Equal concentrations of purified recombinant substrates were exposed to purified recombinant rhomboid proteases. Cleavage products were separated by SDS-PAGE, stained, and quantified densitometrically to determine initial reaction rates, which were converted to molar catalytic activities to allow comparisons. Displayed values are representative of two independent experiments. *B*, cleavage of synthetic LacYTM2 transmembrane peptide KSp31 by GlpG. Purified synthetic peptide KSp31 was incubated with purified recombinant GlpG or its inactive mutant S201T in the presence of 0.05% (w/v) DDM, and the reaction mixtures were analyzed by MALDI mass spectrometry. The theoretical molecular masses of the expected cleavage products at the native cleavage site are denoted below the peptide sequence, and unambiguously match those experimentally determined and displayed in the mass spectra. The star-marked peak with molecular mass of 1893.3 is an unidentified minor contaminant in the preparation of KSp31. *C*, monitoring of cleavage of peptide substrate KSp31 by rhomboid protease GlpG using CE. The N-terminal cleavage product (P) of KSp31 was separated by free-flow CE in the background electrolyte composed of 100 mM H₂PO₄ and 69 mM Tris, pH 2.5, in bare fused silica capillary at separation voltage \pm 25 kV. Samples for CE were prepared by mixing 20 μ l of reaction mixture at selected reaction times (0–90 min) with 2 μ l of 2.2 mM tyramine (7) as an internal standard. Samples were injected into the capillary by 20 mbar pressure for 10 s. Quantitative analysis was based on the ratio of corrected (migration time normalized) peak areas of peptides of interest and the internal standard. Analyses were performed in triplicate. *P*, cleaved N-terminal peptide; *X*, system peak. *D*, the importance of the transmembrane domain of the substrate for its recognition and cleavage by rhomboid. A series of synthetic peptides covering LacYTM2 with progressive truncations of its transmembrane domain from the C terminus was exposed to GlpG and initial rates of cleavage were quantified by capillary electrophoresis as denoted in panel *C*.

able to diverse rhomboid proteases. Because solid phase synthesis of transmembrane peptides and their purification are non-trivial, and their solution behavior often unpredictable, we place emphasis on choosing a robust system and characterizing it thoroughly, and present a generalizable framework for rhomboid substrate design.

Results and Discussion

LacYTM2 Is a Widely Accepted Rhomboid Substrate—To identify a substrate widely accepted by diverse rhomboid proteases, we have measured the efficiency of cleavage of four common model rhomboid substrate transmembrane domains (*P. stuartii* TatA, *Drosophila melanogaster* Gurken and Spitz, and *E. coli* LacYTM2) embedded in a chimeric construct by

four unrelated rhomboid proteases (*E. coli* GlpG, *Bacillus subtilis* YqgP, *P. stuartii* AarA, and *Bacteroides thetaiotaomicron* rhomboid 3 (BtioR3)) (Fig. 1A). Comparison of the efficiencies of cleavage (molar catalytic activities) revealed that the substrate containing the second transmembrane (TM)⁸ helix of *E. coli* LacY protein (LacYTM2) (17) was the most “promiscuous” substrate.

Although it is well accepted that the region around the scissile bond, mainly P4 to P2', is key for the turnover efficiency of

⁸ The abbreviations used are: TM, transmembrane; DDM, *n*-dodecyl- β -D-maltopyranoside; DM, *n*-decyl- β -D-maltopyranoside; CMC, critical micellar concentration; LUV, large unilamellar vesicles; MBP, maltose-binding protein; TAMRA, tetramethylrhodamine; CE, capillary electrophoresis; BGE, background electrolyte.

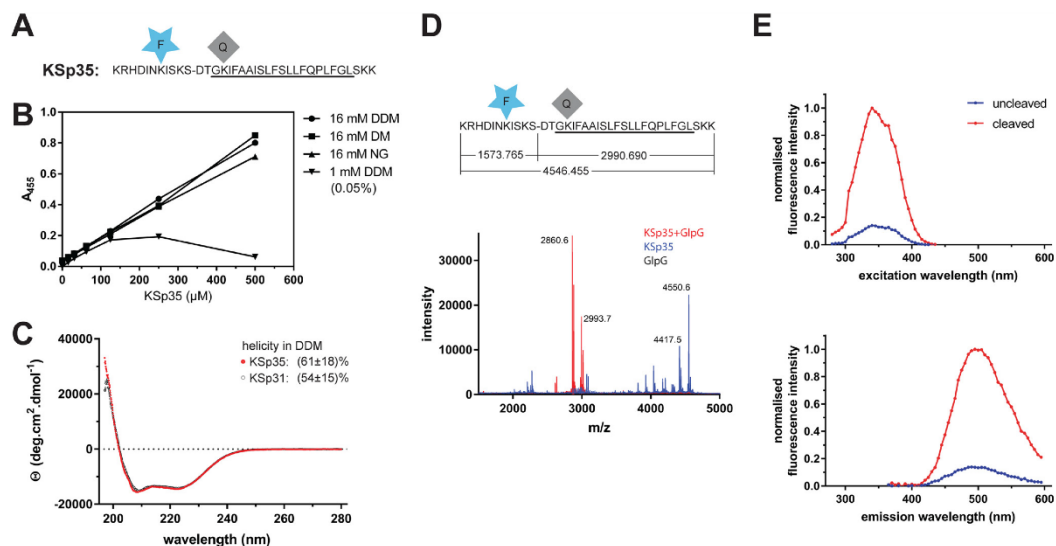


FIGURE 2. Fluorogenic transmembrane peptide substrate based on LacYTM2. *A*, fluorogenic variant of the LacYTM2 transmembrane helix-derived peptide (KSp31) with the P5 and P4' positions replaced by Glu-EDANS and Lys-DABCYL, respectively, yielding fluorogenic substrate KSp35. *B*, solubility of KSp35 in 16 mM detergents DDM, DM, and nonyl glucoside (NG) and at 1 mM DDM. Note that the concentration of DDM micelles is about 100 μM at 16 mM DDM and about 10 μM at 1 mM DDM. The peptide was dissolved to the indicated concentration by dilution from a 10 mM stock solution in DMSO, and after a 2-h incubation at 37 °C the solution was centrifuged at $21,130 \times g$ for 20 min. The absorbance of the supernatant at 455 nm indicated the concentration of the chromophore in solution. *C*, circular dichroism spectra of LacYTM2-derived transmembrane peptide KSp31 and its fluorogenic variant KSp35 in detergent micelles. Peptides were reconstituted into 0.5% (w/v) DDM to 135 μM (KSp31) and 82 μM (KSp35) concentrations. The spectra show similarly significant helical content for both peptides. *D*, identification of the cleavage site in KSp35 by GlpG. Purified 95 μM KSp35 was incubated with 26 μM GlpG for 20 h and analyzed by MALDI. The red peak of the mass of 2993.7 corresponds well to the expected size of the C-terminal cleavage product of 2990.690. The second peak lower by 130 Da is visible in both the blue and red traces is probably a deletion product of chemical synthesis lacking a C-terminal lysine. This variant has proven difficult to purify away, but it is cleaved by GlpG and probably does not influence the kinetics properties of the substrate significantly (see Fig. 1*D*). *E*, excitation and emission spectra of KSp35 and their change upon cleavage by rhomboid GlpG measured in detergent micelles. The spectra of 10 μM KSp35 substrate in reaction buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% (w/v) DDM, 10% (v/v) DMSO) were measured at 37 °C. Excitation wavelengths ranged from 235 to 435 nm with a 10-nm increment and the emission was measured at 493 nm. The emission wavelengths ranged from 365 to 595 nm with a 10-nm increment and excitation at 335 nm.

rhomboid substrates (12, 18), the role of the TM domain of the substrate for recognition and catalysis by rhomboid is less well understood. We have thus next evaluated the importance of the transmembrane region of LacYTM2 for the recognition by *E. coli* GlpG, the main model rhomboid protease, by synthesizing a peptide covering the whole transmembrane region and adjacent juxtamembrane segments of LacYTM2, and a series of its C terminally truncated variants. The full-length LacYTM2 transmembrane peptide KSp31 was cleaved by GlpG efficiently and highly specifically at the expected Ser-Asp cleavage site (Fig. 1*B*). The kinetics of cleavage were monitored by capillary electrophoresis (Fig. 1*C*). The cleavage rate decreased significantly upon truncating the TM helix of LacYTM2 peptide by more than 5 amino acids from the C terminus (Fig. 1*D*), suggesting that most of the TM domain of the substrate is important for the interaction with and recognition by rhomboid. Thus, to develop a widely accepted fluorogenic substrate that would faithfully mimic all the relevant enzyme-substrate interactions including the intramembrane ones, we have used the full-length LacYTM2 transmembrane domain peptide KSp31 as a starting point.

Fluorogenic Transmembrane Peptide Substrate Based on LacYTM2, Basic Properties—To generate a fluorogenic variant of the LacYTM2 peptide, we have replaced the P5 and P4' positions in KSp31 by Glu-EDANS and Lys-DABCYL to yield

KSp35 (Fig. 2*A*). Previously published mutagenic analyses show that these positions are not critical for recognition by rhomboid (18, 19), and they are sufficiently close for Förster resonance energy transfer (FRET) to occur. The KSp35 peptide was soluble up to 500 μM (Fig. 2*B*) in frequently used detergents at 16 mM decyl maltoside (DM), nonyl glucoside (NG), and dodecyl maltoside (DDM). At a total DDM concentration of 16 mM (0.82% (w/v)), the concentration of micelles is about 110 μM , suggesting a partitioning ratio of more than 1 molecule of the substrate per micelle. When DDM was kept at only 1 mM (0.05% (w/v)) total concentration, which yields about 6–10 μM micelles, the solubility of KSp35 became limited to about 100 μM (Fig. 2*B*), indicating that the upper limit of the partitioning ratio is about 10–20 molecules of KSp35 per DDM micelle. The solubility of KSp35 in the absence of detergent was negligible (not shown). Circular dichroism of KSp35 in 0.5% (w/v) DDM (Fig. 2*C*) showed a significant content of α -helical structure ($61 \pm 18\%$), which is consistent with the transmembrane character of the peptide and comparable with the helical content of the parent peptide KSp31 ($54 \pm 15\%$). Cleavage of KSp35 by GlpG occurred at the expected cleavage site (Fig. 2*D*), and was accompanied by an increase in fluorescence at 495 nm (Fig. 2*E*), demonstrating that FRET between the donor and acceptor is occurring in the uncleaved peptide. Collectively, the above results show that KSp35 is a realistic model reflecting all the

Fluorogenic Substrates for Rhomboid Proteases

important interactions between a rhomboid protease and its transmembrane substrate.

Kinetic Characterization of the LacYTM2-based Substrate KSp35 in Detergent Micelle System—In the detergent-solubilized state, most commonly used to study the biochemistry of intramembrane proteolysis, the reaction catalyzed by rhomboid protease occurs in detergent micelles due to the hydrophobicity of both enzyme and substrate. The system is thus microheterogeneous, the effective concentrations of the reactants depend on the volume of the micellar milieu and on the partitioning of reaction components between free solution and the micelles. To characterize the kinetic behavior of the new fluorogenic transmembrane substrates in light of these features of the micellar system, steady-state kinetics was measured with 10 μM substrate, 0.4 μM enzyme, and 0.05% (w/v) DDM, always keeping the concentrations of two components constant and varying the third one around the stated values. At 0.05% (w/v) DDM, the concentration of detergent monomers is 980 μM and micelle concentration about 6–10 μM , calculated assuming critical micellar concentration (CMC) of 0.17 mM (20) and aggregation number between 78 and 149 (20). The molar ratio of enzyme:substrate:micelles is thus 4:100:60–100. In these conditions, assuming that all the reaction partners are evenly distributed among micelles, the average number of substrate molecules per micelle is about 1.5, and only up to 4% of micelles carry an enzyme molecule (micelles containing more than one enzyme molecule are strongly improbable).

The cleavage reactions were started by either mixing two preheated solutions containing substrate or enzyme preincubated with detergent, or adding the DMSO-dissolved substrate into the rest of the preheated reaction mixture. In either case, progress curves are linear from the beginning, which implies that the redistribution of the adsorbed molecules among the micelles is significantly faster than substrate cleavage itself. In accordance with this, the reaction rate is proportional to enzyme concentration within the 0–0.6 μM range (Fig. 3A). Within this concentration range, few enzyme molecules are randomly distributed among many more micelles, providing in principle equal conditions for each enzyme molecule. A similar principle can also explain the observation that the dependence of the reaction rate on substrate concentration is linear in the 0–4 μM range (Fig. 3B). At the upper limit of 4 μM substrate, all micelles can be populated by one (or less likely more) substrate molecule, the linear dependence, furthermore, suggests that this substrate concentration is still below the apparent Michaelis constant of this process.

An important phenomenon is observed when the dependence of the initial rate on detergent concentration is measured. At concentrations above the CMC, the reaction rate rapidly decreases as DDM concentration grows (Fig. 3C), without an obvious impact on the secondary structure content of GlpG (Fig. 3D), suggesting that the effect is caused primarily by the increase in the volume of the micellar phase and consequent decrease of the effective concentrations of both substrate and enzyme. Indeed, mathematical consideration suggests that when substrate and enzyme concentrations are significantly lower than the concentration of micelles (*i.e.* at high DDM concentrations), the probability of location of a substrate molecule

on the same micelle as the enzyme molecule is inversely proportional to the concentration of DDM. Under these conditions, the fraction of substrate-occupied micelles, f_{SM} , is equal to the ratio of the numbers of substrate molecules, $n(S)$, and micelles $n(M)$.

$$f_{SM} = n(S)/n(M) \quad (\text{Eq. 1})$$

The mean number of micelles occupied by both the enzyme and substrate molecules, $n(ESM)$, is then given by this fraction multiplied by the number of enzyme molecules $n(E)$.

$$n(ESM) = f_{SM} \times n(E) = n(S) \times n(E)/n(M) \quad (\text{Eq. 2})$$

Hence, when the DDM concentration is increased at constant $n(S)$ and $n(E)$, then $n(ESM)$ reflecting the reaction rate decreases in accord with the growing value of $n(M)$. This causes the proportional decrease of the reaction rate (in other words, the reaction rate is proportional to $[\text{DDM}]^{-1}$). To inspect whether this model is correct, one can conveniently determine the power of the measured rate dependence on DDM concentration by taking a logarithm of the data from Fig. 3C ($\log a^n = n \times \log a$). The logarithmic plot (Fig. 3C, *open circles, right and upper axes*) can be satisfyingly ($R^2 = 0.9974$) fitted by a second-order polynomial, yielding equation: $y = -0.1436x^2 - 0.3906x + 2.8852$, whose derivative $y' = -0.2872x - 0.3906$ indicates the power of DDM concentration on which the reaction rate depends. This analysis shows that for high DDM concentrations the derivative indeed tends to -1 (for $x = 2$, $y' = -0.965$; thus rate $\sim [\text{DDM}]^{-1}$), which is in accordance with the above assumption, whereas for the lower end of DDM concentrations the absolute value of the power decreases (for $x = 0$, $y' = -0.3906$; thus rate $\sim [\text{DDM}]^{-0.4}$). This is consistent with a model that upon decreasing the detergent concentration (while still being above the CMC), the density of the adsorbed molecules in the micellar phase increases, whereas total concentration of micelles decreases, which leads to less frequent collisions between them and thus less effective redistribution of the adsorbed molecules among the micelles. Possibly, the redistribution efficiency might also be insufficient because of the higher reaction rate caused by the higher reactant concentrations.

Although the reaction kinetics of intramembrane proteases in liposomes has been described in terms of interfacial kinetics (12, 21), that is, expressing the kinetic constants in relationship to the volume or molar fraction of the lipidic phase, (22, 23), the kinetic effects related to the reaction occurring in detergent micelles have surprisingly not yet been considered in enzyme kinetics studies on rhomboid proteases (12, 13) nor other intramembrane proteases, yet they are evidently important for the interpretation of kinetics measurements. Our data show that for reliable and meaningful measurement of apparent Michaelis-Menten kinetics parameters, the micelle concentration must not be limiting the solubility of the substrate, and the detergent concentration must be kept constant. The latter point also means that having a stock solution of the substrate dissolved in detergent (at a higher concentration than intended in the reaction mixture, which frequently can occur during purification and concentration) may lead to underestimation of reaction rates at high substrate concentrations due to a possibly

Fluorogenic Substrates for Rhomboid Proteases

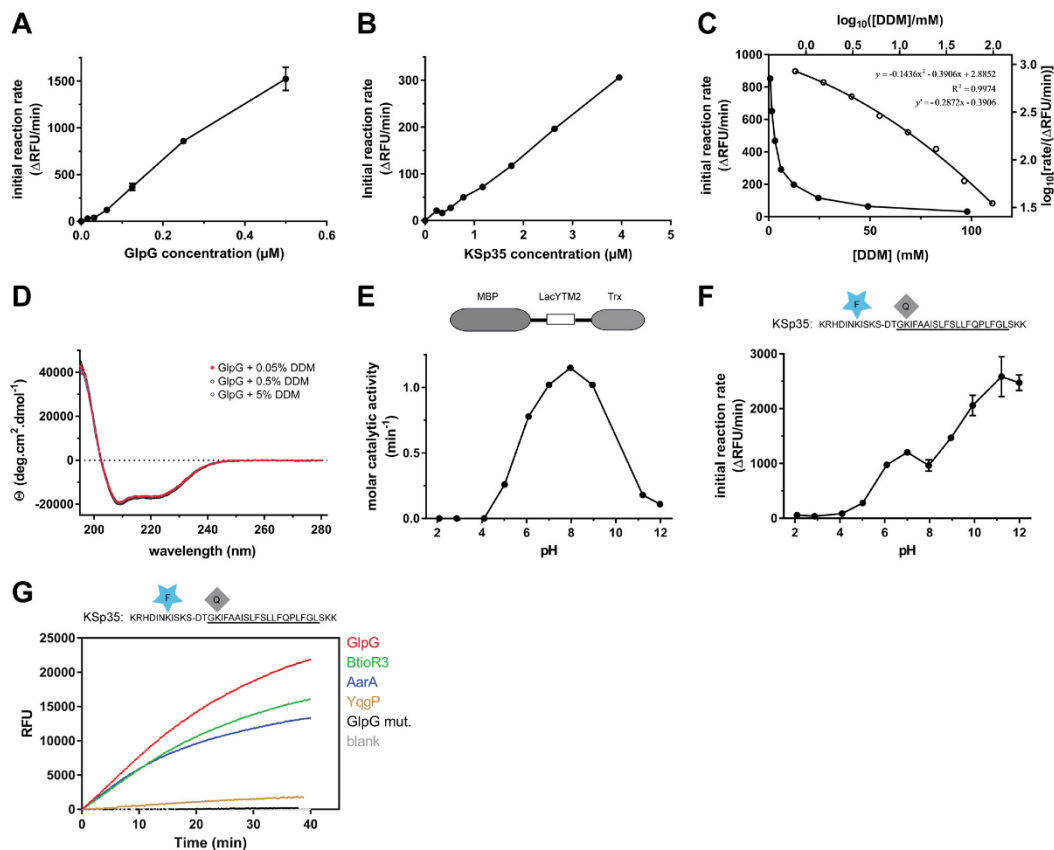


FIGURE 3. Kinetic characterization of fluorogenic transmembrane peptide substrate KSp35 in the detergent micelle system. A, dependence of the initial reaction rate on enzyme concentration. The fluorogenic substrate KSp35 (10 μ M) was incubated with varying concentrations of GlpG in a reaction buffer composed of 20 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% (w/v) DDM, and 10% (v/v) DMSO, and initial reaction rates were measured by following fluorescence at 493 nm. The displayed values are means from duplicate measurements with $2 \times$ S.D. B, dependence of the initial reaction rate on substrate concentration. The rhomboid protease GlpG (0.4 μ M) was incubated with varying concentrations of the fluorogenic substrate KSp35 in a reaction buffer composed of 20 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% (w/v) DDM, 10% (v/v) DMSO, and the initial reaction rates were measured by following fluorescence at 493 nm. Representative values from one of three independent experiments are shown. C, dependence of the initial reaction rate on detergent concentration (solid circles, left and lower axes). The fluorogenic substrate KSp35 (10 μ M) was incubated with 0.4 μ M GlpG at varying concentrations of DDM in a reaction buffer composed of 20 mM HEPES, pH 7.4, 150 mM NaCl, 10% (v/v) DMSO, and initial reaction rates were measured by following fluorescence at 493 nm. Representative values from one of three independent experiments are shown. The open circles (right and upper axes) represent the same plot at the logarithmic scale. When this plot is fitted by second-order polynomial, the equation $y = -0.1436x^2 - 0.3906x + 2.8852$ is obtained, the derivative of which, $y' = -0.2872x - 0.3906$, is equal to the power of DDM concentration with which the reaction rate decreases. For high DDM concentrations the derivative tends to -1 (for $x = 2$, $y' = -0.965$), whereas for lower DDM concentrations the absolute value of the power decreases (for $x = 0$, $y' = -0.3906$). D, overall secondary structure of GlpG is not affected by high concentrations of DDM. CD spectra of GlpG at 0.05, 0.5, and 5% (w/v) (98 mM) DDM were recorded and show no variation in the secondary structure content of GlpG depending on DDM concentration. E, the pH dependence of GlpG activity on the LacYTM2-derived chimeric substrate MBP-LacYTM2-Trx. The substrate (2 μ M) was incubated with 0.1 μ M GlpG in a broad pH range buffer (38) composed of 40 mM H₃PO₄, 40 mM CH₃COOH, and 40 mM H₂BO₃ adjusted to pH values between 2 and 12, and initial reaction rates were measured by SDS-PAGE and densitometry as described under "Experimental Procedures." F, the pH dependence of cleavage of the fluorogenic LacYTM2-derived substrate KSp35 by GlpG. The substrate (10 μ M) was incubated with 0.4 μ M GlpG in a broad pH range buffer (38) composed of 40 mM H₃PO₄, 40 mM CH₃COOH, and 40 mM H₂BO₃ adjusted to pH values between 2 and 12, and initial reaction rates were measured by recording fluorescence at 493 nm. G, selectivity of the fluorogenic substrate KSp35 for diverse bacterial rhomboid proteases. The purified recombinant rhomboid proteases GlpG, AarA, YqgP (all at 0.4 μ M), and BtioR3 (at 0.04 μ M) were incubated with 10 μ M KSp35 in a reaction buffer composed of 20 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% (w/v) DDM, and 10% (v/v) DMSO, and progress curves were measured by recording the increase in fluorescence at 493 nm.

significant increase of detergent concentration in the final reaction mixture, as shown in Fig. 3C. This could result in pseudo-Michaelis kinetics and yield falsely low K_m values. Practical implications are that 1) exact detergent concentrations must be known in any kinetics measurements, and 2) it is advantageous to have the substrate stock solution dissolved in a detergent-free medium or at a detergent concentration lower or equal to

that used in the final assay buffer. The transmembrane substrates presented in this article, generated by chemical synthesis, are in principle avoiding this problem, because their stock solutions are detergent-free dissolved in anhydrous dimethyl sulfoxide. Alternatively, they can be reconstituted into a detergent of choice via disaggregation in hexafluoroisopropanol, as described by Deber *et al.* (24).

Fluorogenic Substrates for Rhomboid Proteases

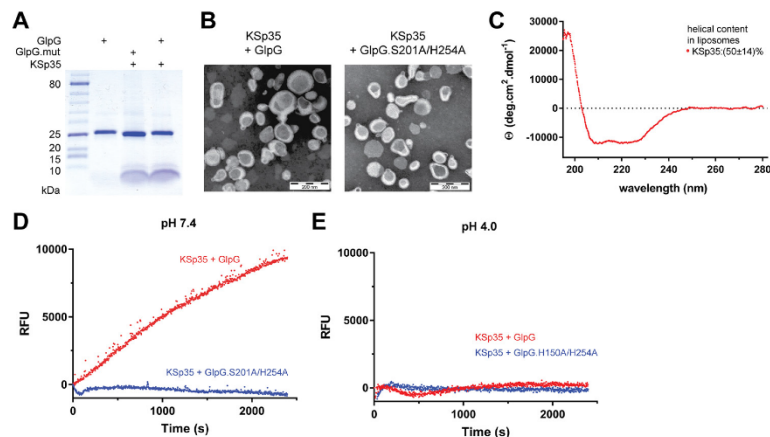


FIGURE 4. The use of the transmembrane peptide substrate in liposomes. *A*, KSp35 was reconstituted into liposomes (LUVs) formed from *E. coli* polar lipid extract in the presence of GlpG or its inactive mutant S201A at pH 4.0. The resulting large unilamellar vesicles were analyzed by SDS-PAGE. *B*, the shape, lamellarity, and approximate size distribution of the KSp35 + GlpG containing proteoliposomes formed at pH 4.0 were characterized by transmission electron microscopy. *C*, the integration of KSp35 into liposomes and its secondary structure content were analyzed by electronic CD. The substrate KSp35 (3 μM) was reconstituted with 2 mg/ml of *E. coli* polar lipid extract yielding an approximate peptide:lipid weight ratio of 1:500. *D*, activity of GlpG in liposomes detected by the KSp35 fluorogenic substrate. The substrate was co-reconstituted with wild type GlpG or its S201A/H254A mutant in a 30:1 molar ratio into LUVs made of *E. coli* polar lipid extract at pH 4.0, proteoliposomes were collected by ultracentrifugation and resuspended in 10 mM HEPES, 150 mM NaCl, pH 7.4, to start the cleavage reaction, which was then followed by measuring fluorescence at 493 nm. *E*, wild type GlpG or its H150A/H254A mutant were co-reconstituted with the substrate KSp35 in a 30:1 molar ratio into LUVs made of *E. coli* polar lipid extract at pH 4.0, proteoliposomes were collected by ultracentrifugation, resuspended in 50 mM sodium acetate, 150 mM NaCl, pH 4.0, and fluorescence was followed at 493 nm.

The pH dependence of cleavage rate of the unmodified LacYTM2 transmembrane segment in the context of an MBP-thioredoxin fusion protein shows a relatively broad maximum around pH 9, with substantial activity of GlpG between pH 6 and 11 and negligible activity below pH 4 and at pH 12 (Fig. 3E), which is largely in agreement with previous studies (12, 13). The dependence of the cleavage rate of KSp35 on pH also shows that GlpG is completely inactive at pH values below and up to 4, but the initial reaction rate of KSp35 cleavage then appears to grow up to pH 12 (Fig. 3F). This effect cannot be ascribed to the pH-dependent change of EDANS fluorescence (data not shown), and could possibly be due to effects of pH on the conformational dynamics of KSp35. However, this is not a concern because in most cases measurements are performed at a physiologically relevant pH near neutral. The apparent catalytic efficiency k_{cat}/K_m of GlpG against KSp35 measured at pH 7.4 and 0.05% (w/v) DDM is $(2.0 \pm 0.5) \times 10^{-3} \text{ min}^{-1} \mu\text{M}^{-1}$, which is comparable with the values reported for the TatA substrate by Dickey *et al.* (12) and Arutyunova *et al.* (13) obtained in similar conditions. Importantly, the LacYTM2-derived fluorogenic peptide substrate KSp35 is cleaved efficiently by unrelated recombinantly purified bacterial rhomboids GlpG, AarA, and BtiOR3, and modestly by YqgP (Fig. 3G), which demonstrates its wide usability, surpassing any other currently available rhomboid substrates.

Use of the Transmembrane Peptide Substrate in Liposomes—Because the natural environment of rhomboid proteases is the lipid membrane, we next tested whether the fluorogenic peptide substrate KSp35 can also be used in liposomes. We co-reconstituted KSp35 with GlpG or its inactive mutant S201A/H254A at pH 4 into large unilamellar vesicles (LUVs) formed from *E. coli* polar lipid extract, and confirmed the composition of the resulting proteoliposomes by SDS-PAGE (Fig. 4A). Neg-

ative stain transmission electron microscopy showed that both empty LUVs and proteoliposomes containing KSp35 in the presence or absence of GlpG or its inactive mutant S201A/H254A had similar morphology and size distribution both at pH 7 and 4 (Fig. 4B). The CD spectrum of LUV-reconstituted KSp35 showed helicity of $50 \pm 14\%$ (Fig. 4C), which is consistent with its transmembrane helix prediction. GlpG is inactive at pH 4 (Fig. 3, E and F), and, consistently, fluorescence of proteoliposomes containing KSp35 and GlpG at pH 4 was at a constant background level (Fig. 4E). Upon neutralization to pH 7.4, time-dependent increase of fluorescence at 495 nm was observed in the presence of wild type GlpG but not in the presence of its active-site mutant S201A/H254A (Fig. 4D). These results collectively demonstrate that the LacYTM2-based fluorogenic transmembrane substrate KSp35 is widely usable both in detergent micelles or liposomes and with diverse rhomboid proteases.

A Red-shifted Variant of the Fluorogenic Transmembrane Substrate for Rhomboids—Large compound libraries for high-throughput screening can often contain compounds that absorb in the UV region (14), and fluorogenic substrates operating at red-shifted wavelengths are less affected by such compound interference. Because EDANS is excited in the UV region, and is thus prone to interference in library screening, we have modified the LacYTM2 peptide backbone by instead attaching the red-shifted TAMRA fluorophore to a Lys introduced into the P5 position and a compatible dark quencher QXL610 to a Cys introduced into the P4' position (Fig. 5A) to yield KSp76. This red-shifted fluorogenic substrate is cleaved by several bacterial rhomboid proteases with efficiencies similar to its UV variant KSp35. The apparent catalytic efficiency k_{cat}/K_m of GlpG cleaving KSp76 is $(1.6 \pm 0.5) \times 10^{-3} \text{ min}^{-1}$

Fluorogenic Substrates for Rhomboid Proteases

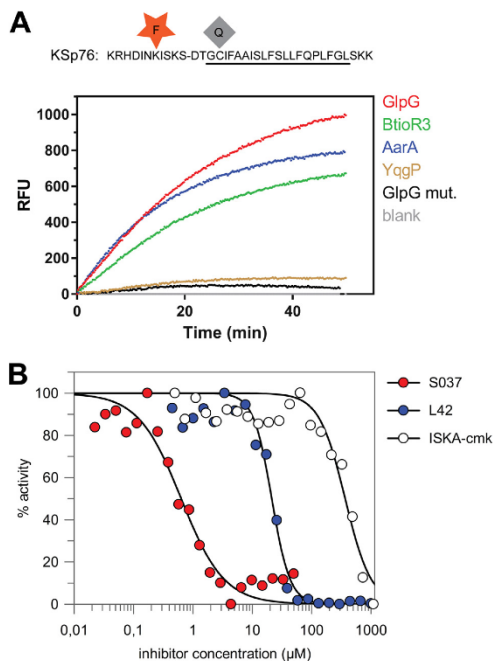


FIGURE 5. Red-shifted variant of the LacYTM2-based fluorogenic substrate. A, modification of Lys in the P5 position of KSp31 by the red-shifted TAMRA fluorophore and P4' Cys by a dark quencher QXL610 yields highly fluorogenic substrate KSp76 that is efficiently cleaved by rhomboid proteases GlpG, AarA, YqgP, and BtioR3 at identical concentrations to those used in Fig. 3G. Excitation wavelength was 553 nm, and emission was followed at 583 nm. B, the red-shifted fluorogenic substrate KSp76 allows measurement of inhibition by compounds that absorb in the UV region, such as isocoumarin, and is thus suitable for high-throughput screening. The dose-response curves of the chloromethylketone ISKA-cmk, β -lactam L42, and isocoumarin S037 were measured after a 60-min preincubation of enzyme with inhibitor. The curves were fitted in GraFit 7 to yield apparent IC_{50} values.

μM^{-1} , which is similar to the EDANS variant KSp35 ($(2.0 \pm 0.5) \times 10^{-3} \text{ min}^{-1} \mu\text{M}^{-1}$) under identical reaction conditions within experimental error (Fig. 6C). The utility of this red-shifted variant of the LacYTM2 substrate is demonstrated by measuring the inhibition curves of chloromethylketone ISKA-cmk (19), β -lactam L42 (11), and isocoumarin S037 (25, 26). Using a 60-min enzyme + inhibitor preincubation time, the measurements yielded apparent IC_{50} values of 370 ± 38 , 12.4 ± 1.6 , and $0.64 \pm 0.08 \mu\text{M}$, respectively (Fig. 5B), which are largely in agreement with published values measured in other assay systems and otherwise comparable conditions (11, 15, 19).

Efficiency and Selectivity of the Substrates Can Be Tuned by Varying Their Non-prime Side Amino Acid Sequence—One of the problems with current rhomboid protease assays is that there has been little rationale about how to modify the substrates to improve their kinetic properties and adapt them for different rhomboid proteases. Recent enzymatic analyses (12, 18) have shown that the region between the P4 and P2' residues determines the k_{cat} of the cleavage reaction, suggesting that selective substrates for rhomboids could be designed by modi-

fying the P4 to P2' region appropriately. A recent mutagenic study of the TatA substrate and structural analysis of a derived rhomboid-substrate-peptide complex revealed amino acids at the P5 to P1 positions of TatA that are preferred by GlpG (19). We tested the impact of these substitutions in the context of the LacYTM2 substrate.

Although single mutations of the P5 amino acid to the preferred Arg, P4 amino acid to Val, and P2 amino acid to His did not improve the cleavage of the purified recombinant MBP-LacYTM2-Trx substrate *in vitro*, mutation of the P1 amino acid to Ala improved the cleavage of mutant 7-fold, and mutation of the P3 residue to Arg improved the cleavage of mutant 16-fold (Fig. 6A). Combining all five mutations yielded a mutant substrate (RVRHA) that was cleaved 64-fold better than the wild type substrate (Fig. 6A), which shows that the effects of the preferred substitutions are additive. When analyzed for cleavage *in vivo*, it turns out that already the wild type MBP-LacYTM2-Trx substrate is such a good substrate of GlpG that it is turned over from 94% (Fig. 6B). The effects of the preferred P5 to P1 mutations thus cannot be assessed in this context as they all exhibit similarly high steady-state turnover (Fig. 6B).

To test this effect in our fluorogenic substrates, we have modified the TAMRA-based LacYTM2-derived fluorogenic substrate by changing the P5 to P1 segment from HISKS to RVRHA to yield KSp64, and compared the kinetic properties of both substrates. The analysis revealed that catalytic efficiency k_{cat}/K_m of GlpG cleaving KSp64 is $(3.7 \pm 0.4) \times 10^{-2} \text{ min}^{-1} \mu\text{M}^{-1}$, which is 23-fold higher than that of the original red-shifted LacYTM2 substrate KSp76 ($(1.6 \pm 0.5) \times 10^{-3} \text{ min}^{-1} \mu\text{M}^{-1}$) (Fig. 6C). The impact of the modifications of the P5 to P1 region on selectivity against other bacterial rhomboid proteases is particularly striking (Fig. 6D), with the initial reaction rate of KSp64 cleavage by GlpG being about 50-fold higher than that of AarA (measured from data displayed in Fig. 6D) and even higher for the other tested rhomboid proteases, revealing a straightforward strategy for designing selective rhomboid substrates.

In summary, we report novel sensitive versatile fluorogenic transmembrane peptide substrates for rhomboid intramembrane proteases that are usable both in detergent micelles and liposomes, are cleaved by diverse rhomboid proteases, and contain a red-shifted fluorophore suitable for high-throughput screening assays. Furthermore, we provide a strategy how to adapt these substrates to individual rhomboid proteases by modifying their P5 to P1 residues, and we demonstrate that controlling the detergent concentration is important for obtaining accurate kinetic data. We expect that the substrates we describe and sequence variants thereof will enable facile detection of activity and development of inhibitors of rhomboid proteases.

Experimental Procedures

General Biochemicals—Lipids were from Avanti Polar Lipids, detergents from Anatrace, buffers and other biochemicals were from Sigma or other suppliers as specified below.

DNA Constructs and Cloning—The expression constructs for rhomboid proteases GlpG, YqgP, and AarA and chimeric MBP-TMD-Trx substrate constructs where TMD = LacYTM2, Gur-

Fluorogenic Substrates for Rhomboid Proteases

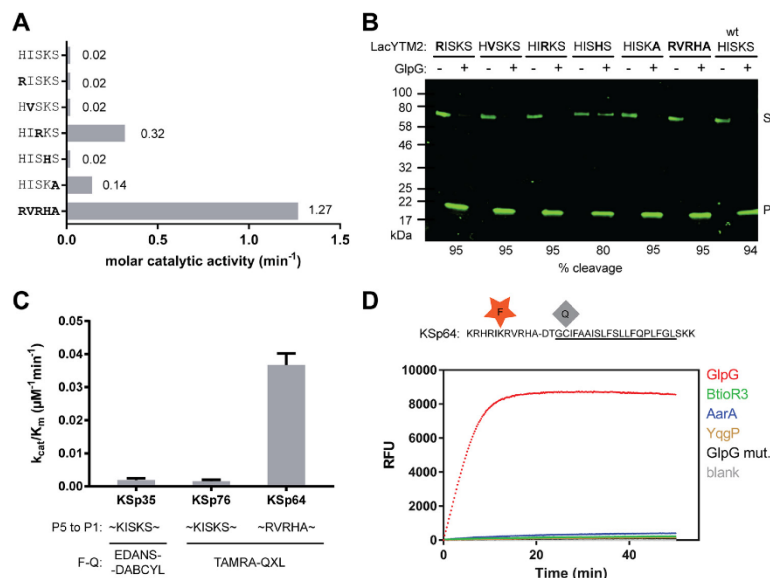


FIGURE 6. The effect of non-prime side substitutions on the catalytic parameters and selectivity of rhomboid substrates. *A*, preferred amino acids in the P5 to P1 positions of the LacYTM2 transmembrane substrate improve its cleavage by GlpG. The LacYTM2 embedded in the MBP-thioredoxin chimera (18) was point-mutated in the P5 to P1 positions according to the sequence preferences of *E. coli* GlpG (19). The recombinant substrates were expressed in *E. coli* Δ glpG, purified, and molar catalytic activity of GlpG in cleaving each of the substrates was determined using gel-based assay (see “Experimental Procedures” for details). The concentration of substrate was always 1.47 μM , concentration of DDM was 0.5% (w/v), the concentration of GlpG was 0.8 μM for wild type substrate (HISKS), and for the RISKS, HVSKS, and HISHS mutants the concentration was 0.08 μM for the HISKA mutant and 0.016 μM for the HIRKS and RVRHA variants (to ensure reliable measurement of the initial reaction rate). Representative values from one of three independent experiments are shown. *B*, the effects of the preferred amino acids in the P5 to P1 region of LacYTM2 on the steady-state level of cleavage by GlpG in biological membranes *in vivo*. Plasmids encoding individual mutant versions of the chimeric mutant LacYTM2 substrates described above were transformed into *E. coli* MC4100 expressing endogenous GlpG, and 2 h after induction of expression of the substrates, the cell lysates were analyzed by immunoblotting using antibody against His tag, located at the C terminus of the constructs. Detection by near-infrared laser scanning, exhibiting linearity over 6 orders of magnitude, enabled reliable quantitation. Integration of product and substrate band intensities yielded steady-state substrate conversion values that are listed below the image. A representative experiment is displayed. *C*, apparent kinetic parameters of fluorogenic rhomboid substrates derived from LacYTM2. Initial reaction rates at very low substrate concentrations were used to calculate catalytic efficiency values (k_{cat}/K_m) of substrates KSp35, KSp64, and Ksp76 cleaved by GlpG at 0.5% (w/v) DDM. The reaction buffer was 20 mM HEPES, pH 7.4, 150 mM NaCl, 10% (v/v) DMSO, enzyme concentration was 0.4 μM , and substrate concentration ranged from 0.5 to 20 μM . Note that a mere optimization of the P5 to P1 region of the substrate increases the catalytic efficiency (k_{cat}/K_m) of its cleavage by GlpG by 23-fold. *D*, influence of the optimization of the P5 to P1 region on the selectivity of a transmembrane substrate for rhomboids. KSp76 underwent cleavage by rhomboid proteases GlpG, AarA, YggP, and BtioR3 at the same concentrations as described in the legends to Figs. 3G and 5A. Note that optimization of the P5 to P1 region of the substrate increases the selectivity for GlpG dramatically.

ken, TatA, or Spitz as described previously (18). The expression construct for rhomboid protease BtioR3 was generated by PCR amplification of the entire ORF encoding the Q8A3X2 (Uniprot ID) protein from *B. thetaiotaomicron* genomic DNA (purchased from ATCC), and its cloning as a C terminally His-tagged construct into pET25b+M as described previously (27). Mutations of the TatA and LacYTM2 recognition motif in the MBP-TMD-Trx construct were generated by overlap assembly PCR (28) and isothermal assembly (29). All constructs were verified by DNA sequencing.

Chemical Synthesis—All reagents were acquired from commercial sources and used without purification. Protected amino acids and amino acid derivatives were purchased from Iris Biotech (Marktredwitz, Germany). Trimellitic anhydride and 3-dimethylaminophenol were from Sigma, QXL610 vinyl-sulfone was from AnaSpec (Fremont, CA), and *N*-(9-fluorenyl) methoxycarbonyl (Fmoc)-Glu(EDANS)-OH from Merck KGaA (Darmstadt, Germany). The detailed synthetic procedures, analytical methods, and compound characterization data are included in the [supporting information](#).

Protein Expression and Purification—Bacterial rhomboid proteases AarA, GlpG, BtioR3, and YggP and the active site mutant GlpG.S201A were overexpressed in *E. coli* C41(DE3) (30) as full-length, C terminally His-tagged proteins from a modified pET25b+ vector (27). The cultures were grown at 37 °C in LB medium to A_{600} of 0.4 and induced by 1 mM isopropyl 1-thio- β -D-galactopyranoside. The expression was continued overnight at 20 °C. Cells were harvested, resuspended in buffer A (25 mM HEPES, pH 7.4, 100 mM NaCl, 10% (v/v) glycerol, 1 mM PMSEF), and lysed by 2 to 3 passes through Avestin EmulsiFlex-C3. Cell debris was removed by a low-speed centrifugation. Cellular membranes were isolated by a 2-h centrifugation at 100,000 $\times g$ and were solubilized in 1.5% (w/v) DDM (solubilization grade, Anatrace) in Buffer B (25 mM HEPES, pH 7.4, 300 mM NaCl, 10% (v/v) glycerol, 10 mM imidazole, EDTA-free Complete Protease Inhibitor mixture (Roche Applied Science)) at room temperature for 1 h. Solubilized proteins were isolated by centrifugation at 100,000 $\times g$ for 30 min and loaded onto nickel-nitrilotriacetic acid HiTrap IMAC HP 1-ml columns (GE Healthcare). Nonspecifically bound proteins were

Fluorogenic Substrates for Rhomboid Proteases

washed off with Buffer C (25 mM HEPES, pH 7.4, 300 mM NaCl, 10% (v/v) glycerol, 0.05% (w/v) DDM) containing 10, 50, and 125 mM imidazole. The protein was eluted with Buffer C containing 250 to 500 mM imidazole. The peak fractions were buffer exchanged into 25 mM HEPES, pH 7.4, 150 mM NaCl, 10% (v/v) glycerol, and 0.05% (w/v) DDM on a HiPrep 26/10 desalting column (GE Healthcare). If needed, proteins were concentrated using Vivaspin ultrafiltration spin cells with 30-kDa MWCO. Protein concentration was determined from absorbance at 280 nm, and the final concentration of DDM was determined as described (31).

Capillary Electrophoresis (CE)—Analyses of standard peptides and enzymatically cleaved peptide substrates were performed on an Agilent CE 7100 instrument (Agilent, Waldbronn, Germany) equipped with photodiode array UV-visible detector operating in the 190–600 nm range. Electropherograms were acquired at 192, 205, and 214 nm and absorbance data at 192 nm were selected for quantitative evaluation due to the highest signal to noise ratio. CE analyses were carried out in a bare fused silica capillary with polyimide outer coating (internal diameter 50 μm , outer diameter 375 μm , effective length to the detector 40 cm, total length 48.5 cm, supplied by Polymicro Technologies, Phoenix, AZ). Peptides were analyzed as cations in acidic background electrolyte (BGE) composed of 100 mM H_3PO_4 , 69 mM Tris, pH 2.5. For highly hydrophobic peptides, this BGE was modified by the addition of 0.05% (w/v) DDM. The temperature of the air-cooled capillary was set to 20 °C and the sample carousel was kept at the same temperature using a circulating water bath. Prior to each CE run, the capillary was successively washed with 100 mM sodium dodecyl sulfate, ethanol, 1 M NaOH, water, 1 M HCl, and the BGE, to remove any possible carryover of hydrophobic peptides and detergents from the previous run. All washes were done at 8 bar pressure for 30 s. Peptide standards used for identification of cleavage products were solubilized in DMSO at 1 mM concentration and mixed with 50 mM HEPES buffer containing 0.05% (w/v) DDM, resulting in 50 μM peptide concentration.

The enzymatic cleavage reactions were carried out in 20 mM HEPES, pH 7.4, with 0.05% (w/v) DDM and 10% (v/v) DMSO, with 250 μM peptide substrate and 2.6 μM full-length GlpG at 37 °C. To measure the initial reaction rates, fractions were collected every 15 min for up to 2 h and the reaction was terminated by the addition of 10 mM HCl. Samples for CE were prepared by mixing 20 μl of peptide solutions with 2 μl of 2.2 mM tyramine (internal standard for quantitative analysis). Sample solutions were injected into the capillary by 20 mbar pressure for 10 s. Separations were performed at +25 kV (anode at the capillary injection end). The electrode vessels were replenished with fresh BGE after each run. All analyses were performed in triplicate. Quantitative analysis was based on the ratio of corrected (migration time normalized) peak areas of peptides of interest and the internal standard (tyramine) (32).

Mass Spectrometry—The analysis of enzymatic cleavage products of transmembrane peptides was carried out using MALDI-TOF mass spectrometry on an UltrafleXtreme™ MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Germany) with α -cyano-4-hydroxycinnamic acid matrix using a thin-layer method (33). For routine quality control during pep-

ptide synthesis, mass spectra were acquired on a Waters Micro-mass ZQ ESCi multimode ionization mass spectrometer, and LTQ Orbitrap XL (Thermo Fisher Scientific) for HR-MS experiments, in both cases using ESI(+) ionization.

Gel-based Assay for Rhomboid Activity—For gel-based assays used in Fig. 1, the purified recombinant full-length maltose-binding protein thioredoxin fusion proteins harboring the transmembrane domains of TatA, LacYTM2, Gurken, and Spitz (18) were used as substrates. The reaction was carried out in 50 mM Tris, pH 7.4, 100 mM NaCl, 10% (v/v) glycerol, 0.05% (w/v) DDM, and 5 μM substrate. Enzyme concentrations varied to ensure adequate conditions for measurement of initial reaction rates for each enzyme-substrate combination. Time courses were measured by withdrawing 10- μl aliquots from the reaction mixture after 10, 20, 30, 40, 50, 60, and 120 min from the start of the reaction, and stopping the reaction by the addition of SDS-PAGE sample buffer. The reaction mixtures were analyzed by SDS-PAGE, Coomassie staining (Instant-Blue, Expedeon, UK), and densitometry as described (19), and initial reaction rates were converted to molar catalytic activities defined as the number of substrate molecules converted by a molecule of the enzyme per unit of time (consistent with the definition by IUPAC (34, 35)). Variations in conditions used for measurements in Fig. 6 are denoted in the figure legend.

The *in vivo* assay of rhomboid activity was carried out essentially as described (19). Cleavage products were detected by SDS-PAGE and Western blotting using primary anti-penta-His mouse monoclonal antibody (Thermo) and IRDye 800CW goat anti-mouse fluorescent secondary antibody (LiCor). Densitometry was done in ImageStudio software (LiCor) and substrate conversion (α) was calculated from band intensities as $\alpha_\tau = [P]/([S] + [P])$, where $[P]$ and $[S]$ are product and substrate concentrations at time τ , which are proportional to the fluorescence intensity of the product and substrate bands at time τ , because the monoclonal antibody binds to the substrate or product in a constant molar ratio irrespective of their molecular weights.

Fluorescence Assay for Rhomboid Activity—The fluorescence assay of rhomboid activity was performed at 37 °C in 96-well black HTS plates (Greiner Bio-One). The reaction conditions were typically as follows: 20 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% (w/v) DDM, 12% (v/v) DMSO, and 10 μM fluorogenic peptide substrate in a final volume of 50 μl , unless noted otherwise. Concentrations of stock solutions of peptide substrates and inhibitors (if applicable) were determined by quantitative amino acid analysis. Fluorescence was read continuously in a plate reader (Tecan Infinite M1000). Excitation and emission wavelengths were 335 and 493 nm, respectively, for the EDANS-DABCYL substrate, and 553 and 583 nm for the TAMRA-QXL610 substrates. Data were evaluated in i-Control (Tecan), Excel (Microsoft), GraphPad Prism 7 (GraphPad Software, Inc.), and GraFit 7 (Erithacus Software, Ltd.) software.

Inhibition Assays—The inhibition assay was carried out in 20 mM HEPES, pH 7.4, 150 mM NaCl, 12% (v/v) DMSO, 0.05% (w/v) DDM at 37 °C in 96-well black HTS plates (Greiner Bio-one). Purified recombinant full-length GlpG (0.4 μM) was pre-

Fluorogenic Substrates for Rhomboid Proteases

incubated with each inhibitor at different concentrations for 1 h at 37 °C. The cleavage reaction was started by adding 10 μM KSp76 and fluorescence was read continuously to measure initial reaction rates as described above.

Reconstitution into Liposomes—*E. coli* polar lipids (20 mg), with optionally 0.1 mg of Lissamine Rhodamine B-labeled phosphatidylethanolamine (16:0) (Avanti Polar Lipids) added for visibility, were dried in a glass test tube by manual rotation under a nitrogen stream. Residual traces of solvent were removed by overnight incubation in a vacuum chamber (Binder). The resulting lipid film was hydrated in 5 ml of 50 mM acetate, 150 mM NaCl, pH 4.0, by 2 min vortexing followed by a 1-h incubation in a horizontal shaker at 200 rpm and 37 °C, and 3 cycles of freezing in liquid nitrogen and thawing in a 37 °C water bath. The lipid suspension was then extruded through a 200-nm pore membrane by 19 strokes in an Avanti Mini Extruder (Avanti Polar Lipids).

For reconstitution of proteins and peptides into liposomes, these unilamellar LUVs were solubilized in DM to a final ratio of 1.5:1 detergent:lipid, and incubated for 1 h at room temperature under gentle rotation. This mixture was diluted to a final lipid concentration of 2 mg/ml in 50 mM acetate, 150 mM NaCl, pH 4.0, and protein (GlpG or its inactive mutant) dissolved in detergent was added to a final concentration of 8 $\mu\text{g}/\text{ml}$; alternatively, the stock solution of substrate peptide KSp35 in DMSO was diluted to 10 μM . The resulting mixture was incubated at room temperature for 1 h under gentle mixing by inversion. Detergent was removed by overnight dialysis against 500-fold excess of 50 mM acetate, 150 mM NaCl, pH 4, followed by 5 h dialysis against 500-fold excess the same buffer, using 10-kDa MWCO dialysis membranes, which allowed reconstitution of proteoliposomes. These were extruded through 200-nm pore filters 9 times to ensure reproducible size distribution and lamellarity. These final proteoliposomes were harvested by ultracentrifugation (250,000 $\times g$ for 1 h at 4 °C), and resuspended in 10 mM HEPES, pH 7.4, 150 mM NaCl to a concentration of about 33 mg/ml of lipids. The morphology and size distribution of proteoliposomes was analyzed by electron microscopy.

Transmission Electron Microscopy—Liposome samples were negatively stained with 2% phosphotungstic acid on carbon-coated electron microscopy grids and analyzed with a JEOL JEM-1011 device at 80 kV beam acceleration voltage.

CD Spectroscopy—Protein and peptide samples were dissolved in 50 mM phosphate buffer at the indicated concentrations and in the presence of detergent as indicated, or reconstituted in LUVs made of *E. coli* polar lipids and extruded by 100-nm filters to minimize light scattering. Electronic circular dichroism spectra were collected by a Jasco 815 spectrometer (Tokyo, Japan) in the spectral 195–280 nm range using a cylindrical 0.02-cm quartz cell with 0.1-nm step resolution, 5 nm/min scanning speed, 16 s response time, and 1 nm spectral band. After baseline correction, the spectra were expressed as molar ellipticity per residue θ (deg $\text{cm}^2 \text{dmol}^{-1}$). Numerical analysis of the secondary structure and secondary structure assignment were performed using a CDPPro software package and CONTIN program (36, 37).

Author Contributions—K. S. conceived and coordinated the study, designed experiments, and wrote the paper with the input of A. T., M. I., S. S., J. B., J. S., M. R., and V. K. S. S. and P. M. designed and S. S. performed all chemical syntheses. M. R. and V. K. designed and performed all capillary electrophoresis analyses. M. I. analyzed kinetics data, R. H. performed electron microscopy, and L. B. designed and performed all circular dichroism measurements. J. B. performed and evaluated experiments shown in Fig. 6, A and B. JŠ performed and evaluated experiments shown in Fig. 4, A, D, and E. K. Š designed, performed, and analyzed data shown in Fig. 1A. P. R. and E. P. contributed to experiments shown in Fig. 6, A, B, and D. L. P. established the fluorogenic assay and performed and evaluated experiments shown in Fig. 1D. J. Březinová contributed to all mass spectrometry experiments, and A. T. designed, performed, and evaluated all other kinetics and inhibition measurements that are the basis of this manuscript.

Acknowledgments—We thank Steven Verhelst (University of Leuven, Belgium) for his kind gift of isocoumarin S037, Matthew Freeman (Oxford University, United Kingdom) for his kind gift of the inhibitor L42, Zdeněk Voburka and Radko Souček for amino acid analyses, Mirka Blechová for peptide synthesis and purification, and Blanka Collis for critical reading of the manuscript.

References

- Urban, S., Lee, J. R., and Freeman, M. (2002) A family of rhomboid intramembrane proteases activates all *Drosophila* membrane-tethered EGF ligands. *EMBO J.* **21**, 4277–4286
- Lee, J. R., Urban, S., Garvey, C. F., and Freeman, M. (2001) Regulated intracellular ligand transport and proteolysis control EGF signal activation in *Drosophila*. *Cell* **107**, 161–171
- McQuibban, G. A., Saurya, S., and Freeman, M. (2003) Mitochondrial membrane remodelling regulated by a conserved rhomboid protease. *Nature* **423**, 537–541
- O'Donnell, R. A., Hackett, F., Howell, S. A., Treeck, M., Struck, N., Krnjajski, Z., Withers-Martinez, C., Gilberger, T. W., and Blackman, M. J. (2006) Intramembrane proteolysis mediates shedding of a key adhesion during erythrocyte invasion by the malaria parasite. *J. Cell Biol.* **174**, 1023–1033
- Fleig, L., Bergbold, N., Sahasrabudhe, P., Geiger, B., Kaltak, L., and Lemberg, M. K. (2012) Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins. *Mol. Cell* **47**, 558–569
- Riestra, A. M., Gandhi, S., Sweredoski, M. J., Moradian, A., Hess, S., Urban, S., and Johnson, P. J. (2015) A *Trichomonas vaginalis* rhomboid protease and its substrate modulate parasite attachment and cytolysis of host cells. *PLoS Pathog.* **11**, e1005294
- Etheridge, S. L., Brooke, M. A., Kelsell, D. P., and Blyden, D. C. (2013) Rhomboid proteins: a role in keratinocyte proliferation and cancer. *Cell Tissue Res.* **351**, 301–307
- Chan, E. Y., and McQuibban, G. A. (2013) The mitochondrial rhomboid protease: its rise from obscurity to the pinnacle of disease-relevant genes. *Biochim. Biophys. Acta* **1828**, 2916–2925
- Song, W., Liu, W., Zhao, H., Li, S., Guan, X., Ying, J., Zhang, Y., Miao, F., Zhang, M., Ren, X., Li, X., Wu, F., Zhao, Y., Tian, Y., Wu, W., et al. (2015) Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway. *Nat. Commun.* **6**, 8022
- Strisovsky, K. (2016) Why cells need intramembrane proteases: a mechanistic perspective. *FEBS J.* **283**, 1837–1845
- Pierrat, O. A., Strisovsky, K., Christova, Y., Large, J., Ansell, K., Boulou, N., Smiljanic, E., and Freeman, M. (2011) Monocyclic β -lactams are selective, mechanism-based inhibitors of rhomboid intramembrane proteases. *ACS Chem. Biol.* **6**, 325–335
- Dickey, S. W., Baker, R. P., Cho, S., and Urban, S. (2013) Proteolysis inside the membrane is a rate-governed reaction not driven by substrate affinity. *Cell* **155**, 1270–1281

Fluorogenic Substrates for Rhomboid Proteases

13. Arutyunova, E., Panwar, P., Skiba, P. M., Gale, N., Mak, M. W., and Lemieux, M. J. (2014) Allosteric regulation of rhomboid intramembrane proteolysis. *EMBO J.* **33**, 1869–1881
14. Simeonov, A., Jadhav, A., Thomas, C. J., Wang, Y., Huang, R., Southall, N. T., Shinn, P., Smith, J., Austin, C. P., Auld, D. S., and Ingles, J. (2008) Fluorescence spectroscopic profiling of compound libraries. *J. Med. Chem.* **51**, 2363–2371
15. Vosyka, O., Vinothkumar, K. R., Wolf, E. V., Brouwer, A. J., Liskamp, R. M., and Verhelst, S. H. (2013) Activity-based probes for rhomboid proteases discovered in a mass spectrometry-based assay. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2472–2477
16. Wolf, E. V., Zeißler, A., Vosyka, O., Zeiler, E., Sieber, S., and Verhelst, S. H. (2013) A new class of rhomboid protease inhibitors discovered by activity-based fluorescence polarization. *PLoS ONE* **8**, e72307
17. Maegawa, S., Ito, K., and Akiyama, Y. (2005) Proteolytic action of GlpG, a rhomboid protease in the *Escherichia coli* cytoplasmic membrane. *Biochemistry* **44**, 13543–13552
18. Strisovsky, K., Sharpe, H. J., and Freeman, M. (2009) Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. *Mol. Cell* **36**, 1048–1059
19. Zoll, S., Stanchev, S., Began, J., Skerle, J., Lepšik, M., Peclínová, L., Majer, P., and Strisovsky, K. (2014) Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures. *EMBO J.* **33**, 2408–2421
20. VanAken, T., Foxall-VanAken, S., Castleman, S., and Ferguson-Miller, S. (1986) Alkyl glycoside detergents: synthesis and applications to the study of membrane proteins. *Methods Enzymol.* **125**, 27–35
21. Kamp, F., Winkler, E., Trambauer, J., Ebke, A., Fluhrer, R., and Steiner, H. (2015) Intramembrane proteolysis of β -amyloid precursor protein by γ -secretase is an unusually slow process. *Biophys. J.* **108**, 1229–1237
22. Scheel, G., Acevedo, E., Conzelmann, E., Nehr Korn, H., and Sandhoff, K. (1982) Model for the interaction of membrane-bound substrates and enzymes: hydrolysis of ganglioside GD1a by sialidase of neuronal membranes isolated from calf brain. *Eur. J. Biochem.* **127**, 245–253
23. Parry, G., Palmer, D. N., and Williams, D. J. (1976) Ligand partitioning into membranes: its significance in determining K_m and K_s values for cytochrome P-450 and other membrane bound receptors and enzymes. *FEBS Lett.* **67**, 123–129
24. Rath, A., and Deber, C. M. (2013) Design of transmembrane peptides: coping with sticky situations. *Methods Mol. Biol.* **1063**, 197–210
25. Wolf, E. V., Zeissler, A., and Verhelst, S. H. (2015) Inhibitor fingerprinting of rhomboid proteases by activity-based protein profiling reveals inhibitor selectivity and rhomboid autoprocessing. *ACS Chem. Biol.* **10**, 2325–2333
26. Haedke, U., Küttler, E. V., Vosyka, O., Yang, Y., and Verhelst, S. H. (2013) Tuning probe selectivity for chemical proteomics applications. *Curr. Opin. Chem. Biol.* **17**, 102–109
27. Lemberg, M. K., Menendez, J., Misik, A., Garcia, M., Koth, C. M., and Freeman, M. (2005) Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. *EMBO J.* **24**, 464–472
28. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59
29. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345
30. Miroux, B., and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298
31. Urbani, A., and Warne, T. (2005) A colorimetric determination for glycosidic and bile salt-based detergents: applications in membrane protein research. *Anal. Biochem.* **336**, 117–124
32. Solínová, V., Kasicka, V., Koval, D., Barth, T., Ciencialová, A., and Záková, L. (2004) Analysis of synthetic derivatives of peptide hormones by capillary zone electrophoresis and micellar electrokinetic chromatography with ultraviolet-absorption and laser-induced fluorescence detection. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **808**, 75–82
33. Fenyo, D., Wang, Q., DeGrasse, J. A., Padovan, J. C., Cadene, M., and Chait, B. T. (2007) MALDI sample preparation: the ultra thin layer method. *J. Vis. Exp.* **192**, 10.3791/192
34. Nomenclature Committee of the International Union of Biochemistry (1979) Units of enzyme-activity, recommendations 1978. *Eur. J. Biochem.* **97**, 319–320
35. Nomenclature Committee of the International Union of Biochemistry (1983) Symbolism and terminology in enzyme-kinetics, recommendations 1981. *Biochem. J.* **213**, 561–571
36. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260
37. Provencher, S. W., and Glöckner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33–37
38. Mongay, C., and Cerda, V. (1974) Britton-Robinson buffer of known ionic-strength. *Anal. Chim.* **64**, 409–412

General and Modular Strategy for Designing Potent, Selective, and Pharmacologically Compliant Inhibitors of Rhomboid Proteases

Anežka Tichá,^{1,4,9} Stancho Stanchev,^{1,9} Kutti R. Vinothkumar,⁵ David C. Mikles,¹ Petr Pachtl,¹ Jakub Began,^{1,3} Jan Škerle,^{1,2} Kateřina Svehlová,¹ Minh T.N. Nguyen,⁶ Steven H.L. Verhelst,^{6,7} Darren C. Johnson,⁸ Daniel A. Bachovchin,⁸ Martin Lepšík,¹ Pavel Majer,¹ and Kvido Strisovsky^{1,10,*}

¹Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, Prague 166 10, Czech Republic

²Department of Biochemistry, Faculty of Science, Charles University, Hlavova 2030/8, Prague 128 43, Czech Republic

³Department of Genetics and Microbiology, Faculty of Science, Charles University, Viničná 5, Prague 128 44, Czech Republic

⁴First Faculty of Medicine, Charles University, Kateřinská 32, Prague 121 08, Czech Republic

⁵Medical Research Council Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH, UK

⁶Leibniz Institute for Analytical Sciences ISAS, Otto-Hahn-Strasse 6b, 44227 Dortmund, Germany

⁷KU Leuven - University of Leuven, Herestraat 49, Box 802, 3000 Leuven, Belgium

⁸Chemical Biology Program, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, Box 428, New York, NY 10065, USA

⁹These authors contributed equally

¹⁰Lead Contact

*Correspondence: kvido.strisovsky@uochb.cas.cz

<https://doi.org/10.1016/j.chembiol.2017.09.007>

SUMMARY

Rhomboid-family intramembrane proteases regulate important biological processes and have been associated with malaria, cancer, and Parkinson's disease. However, due to the lack of potent, selective, and pharmacologically compliant inhibitors, the wide therapeutic potential of rhomboids is currently untapped. Here, we bridge this gap by discovering that peptidyl α -ketoamides substituted at the ketoamide nitrogen by hydrophobic groups are potent rhomboid inhibitors active in the nanomolar range, surpassing the currently used rhomboid inhibitors by up to three orders of magnitude. Such peptidyl ketoamides show selectivity for rhomboids, leaving most human serine hydrolases unaffected. Crystal structures show that these compounds bind the active site of rhomboid covalently and in a substrate-like manner, and kinetic analysis reveals their reversible, slow-binding, non-competitive mechanism. Since ketoamides are clinically used pharmacophores, our findings uncover a straightforward modular way for the design of specific inhibitors of rhomboid proteases, which can be widely applicable in cell biology and drug discovery.

INTRODUCTION

Rhomboid intramembrane proteases are evolutionarily conserved proteins with numerous important biological functions, including growth factor secretion, regulation of mitochondrial dy-

namics, and membrane protein quality control (Fleig et al., 2012). As such, they are being increasingly explored as potential drug targets, for example, for malaria (Baker et al., 2006; Lin et al., 2013; O'Donnell et al., 2006), cancer (Song et al., 2015), Parkinson's disease (Meissner et al., 2015), and diabetes (reviewed in Chan and McQuibban, 2013). These efforts are, however, hindered by the lack of selective and potent rhomboid inhibitors that could be used for cell biological studies, validation of therapeutic potential of rhomboids, and as templates for drug development (Strisovsky, 2016a). As explained elsewhere in more detail (Strisovsky, 2016a), the currently used inhibitors of rhomboid proteases suffer from drawbacks, making them unsuitable for these purposes. Isocoumarins are highly reactive and lack selectivity (Harper et al., 1985; Powers et al., 2002; Powers et al., 1989), β -lactams have limited potency *in vivo* (half maximal inhibitory concentration [IC₅₀] ~5–10 μ M) (Pierrat et al., 2011), and β -lactones are not very potent (apparent IC₅₀ of ~40 μ M) (Wolf et al., 2013). Furthermore, no rational strategy for modulation of their selectivity exists for any of these inhibitor classes. Here, we address both of these bottlenecks.

The principles of the mechanism and specificity of a protease determine to a large extent the strategies for inhibitor development (Drag and Salvesen, 2010). Rhomboids are serine proteases with a Ser-His catalytic dyad (Wang et al., 2006), and they recognize their transmembrane substrates in a two-tier process. It is assumed that first a portion of the transmembrane domain of the substrate docks into an intramembrane interaction site of rhomboid within the plane of the lipid bilayer, upon which a linear segment of the substrate (possibly generated by local unfolding of the top of the substrate's transmembrane helix) interacts with the water-exposed active site (reviewed in Strisovsky, 2016a; Strisovsky, 2013). This "recognition motif" encompasses the P4 to P2' (Schechter and Berger, 1967) residues of the substrate (Strisovsky et al., 2009), it largely determines the k_{cat} of the



reaction (Dickey et al., 2013), and thus modulates selectivity toward a given rhomboid protease (Ticha et al., 2017). Recent reports have shown that peptidyl chloromethylketones (Zoll et al., 2014) and peptidyl aldehydes (Cho et al., 2016) are weakly inhibiting rhomboid proteases at medium to high micromolar concentrations, but they lack selectivity and their potency is insufficient for use as research tools.

Inspired by the current knowledge of the rhomboid protease mechanism (reviewed in Strisovsky, 2016a), we set out to explore the chemical space of oligopeptides equipped with electrophilic warheads in search of new rhomboid inhibitors of greater potency. Our systematic analysis resulted in the discovery of a modular scaffold based on peptidyl-ketoamide substituted with hydrophobic groups that represents a novel class of potent and selective rhomboid inhibitors. The *in vivo* activity of these compounds is in the low nanomolar range, which is up to three orders of magnitude more potent than any other currently known rhomboid inhibitors. Furthermore, we gained insight into the mode of binding of peptidyl ketoamides by solving their co-crystal structures with rhomboid protease, and we present strategies to modify their selectivity and potency on a systematic basis. We expect this compound class to find a widespread use in cell biology in rhomboid protease related contexts and to provide templates for the development of drugs targeting rhomboid proteases.

RESULTS

The Potency of Substrate-Derived Peptidyl Chloromethylketone Inhibitors Can be Markedly Enhanced by Optimizing the Amino Acid Sequence of the P5 to P1 Region

Rhomboid proteases exhibit discernible sequence preferences in the P5 to P2' region of their substrates (Strisovsky et al., 2009; Zoll et al., 2014). To gain insight into these preferences and their possible interactions, we have generated tetra- and pentapeptidyl chloromethylketones (CMK or cmk henceforth) harboring amino acids preferred in positions P5 to P1 by the *Escherichia coli* rhomboid GlpG (GlpG henceforth), using the sequence background of the *Providencia* TatA (Stevenson et al., 2007), represented by the parent compound Ac-IATA-cmk. We measured the inhibitory properties of this series of compounds using a newly developed *in vitro* assay employing a fluorogenic transmembrane peptide substrate (Ticha et al., 2017) that faithfully represents a native rhomboid substrate. The effects of the mutations were additive, and the inhibitor containing the most favored amino acids in positions P5 to P1 (Ac-RVRHA-cmk) is approximately 26-fold more potent than the parent compound Ac-IATA-cmk (Figure 1A).

To provide mechanistic explanation for the observed increase in inhibitory potency, we determined the structures of GlpG in complex with Ac-RVRHA-cmk and Ac-VRHA-cmk (Figure 1B). The side chain of Arg in the P5 position of Ac-RVRHA-cmk could not be modeled due to poor electron density, and the two structures are otherwise virtually identical; superposition of all corresponding C α atoms yields a root-mean-square deviation of 0.19 Å per atom (using the SSM method as implemented in CCP4MG v2.10.4; Krissinel and Henrick, 2004; Mitchell et al., 1990). Both inhibitors interact with the L3 and L5 loops via

main-chain hydrogen bonds, and via hydrogen bonds involving the side chains of the strongly preferred Arg and His in the P3 and P2 position, respectively (Figure 1C). These interactions are not observed in the structure of the parent compound Ac-IATA-cmk (Zoll et al., 2014), suggesting that they contribute to the higher potency of Ac-RVRHA-cmk and Ac-VRHA-cmk over Ac-IATA-cmk. The interactions of the residues in the P4, P3, and P2 positions with the enzyme are structurally independent, explaining why the effects of substitutions in these positions are additive (Figure 1A). The overall binding mode of both compounds into the rhomboid active site is similar to the binding mode of peptide aldehyde Ac-VRMA-cho (Cho et al., 2016) (Figure 1D). Collectively, these data show that rhomboid subsite preferences are additive in the context of an active site targeted inhibitor and that sequence optimization in this region can significantly increase the inhibitory potency of the compounds.

A Screen of Covalent Reversible Warheads for Inhibition of Rhomboid

Since the sequence-optimized chloromethylketones are poor inhibitors with low micromolar IC₅₀, we searched for alternative, more suitable electrophilic warheads that might improve the inhibitory potency. Furthermore, we reasoned that extending the inhibitor to the prime side of the active site might offer additional binding energy. We therefore synthesized a series of compounds based on the Ac-RVRHA sequence equipped with a selection of electrophilic, reversibly binding warheads commonly used for serine proteases in pharmacological settings (reviewed in Hedstrom, 2002; Walker and Lynas, 2001), including trifluoromethylketones, boronates, acylsulfonamides, thiazolylketones, and ketoamides (Figure 2), the last three of which can be extended into the prime side. We measured the apparent IC₅₀ values of these compounds, and while trifluoromethylketones, acylsulfonamides, and thiazolylketones showed none or very weak inhibition in the millimolar range, the apparent IC₅₀ of the boronate was 8 μM and of the ketoamide 203 μM under identical reaction conditions (Figure 2). Although the peptidyl boronate was the best of the series, it was still a relatively weak inhibitor comparable with the parent chloromethylketone, and it was not clear how to further improve its potency. The ketoamide was about 25-fold less potent, but since it could be extended to the prime side by a modification at the ketoamide nitrogen (Chatterjee et al., 1999; Liu et al., 2004), we next focused our attention on this class of compounds.

Extensions at the Prime Side of Peptidyl Ketoamides Greatly Enhance Their Inhibitory Potency

We hypothesized that extending the peptidyl ketoamides to the prime side of the active site might increase their potency, since the P2' residue (hydrophobic in case of GlpG) was shown to be important for substrate recognition by rhomboids (Dickey et al., 2013; Strisovsky et al., 2009), and interactions of the substrate transmembrane domain beyond P2' potentiate substrate cleavage in a detergent micelle assay (Ticha et al., 2017). We synthesized a series of peptidyl ketoamides based on the Ac-RVRHA sequence, bearing a mostly hydrophobic "tail" of increasing size at the ketoamide nitrogen (Figure 3A) that could reach far into the prime side of the rhomboid active site. The tail substituent indeed had a dramatic effect on the potency of

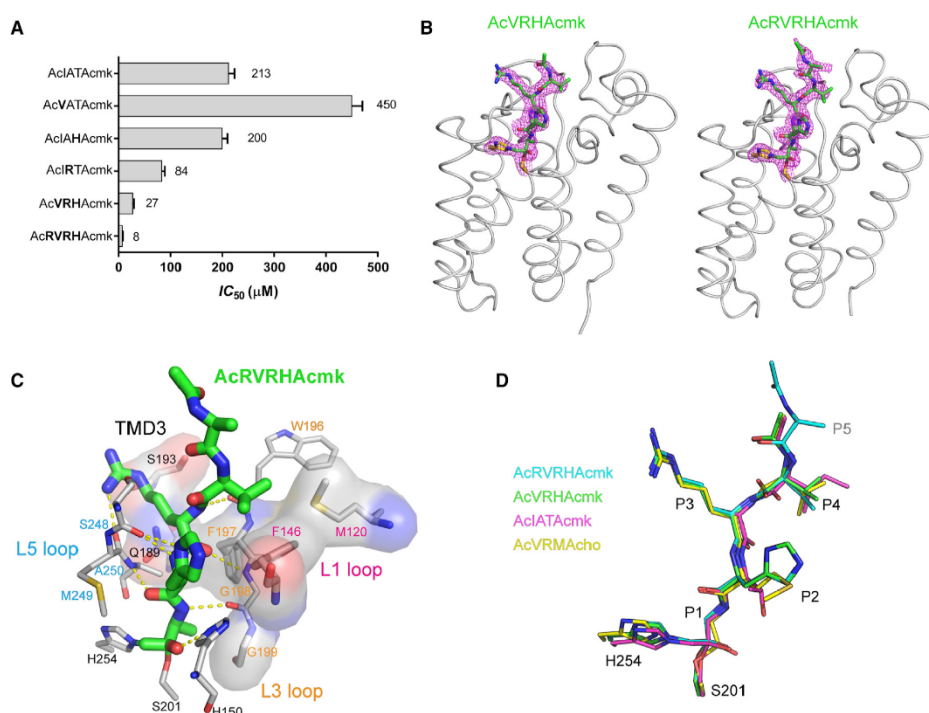


Figure 1. The Potency of Substrate-Derived Inhibitors Can be Improved by Modifying the Amino acid Sequence of the P5 to P1 Region
 (A) The parent inhibitor Ac-IATA-cmk was modified by introducing strongly preferred amino acids (Zoll et al., 2014) into the P4, P3, P2, and P5 positions to yield the listed compounds. Their apparent IC₅₀ values were measured with 1 hr preincubation using 10 μM fluorogenic substrate KSp35 and 0.05% (w/v) DDM. The reported values are best-fit means with SD representative of 2–3 measurements.
 (B) The sequence-optimized peptidyl chloromethylketones were soaked into the native crystals of GlpG and structures of the complexes were solved by X-ray diffraction (for statistics, see Table S1). In the displayed structures, the catalytic dyad is shown as yellow sticks and the inhibitors are shown as green sticks surrounded by the 2mF_o – DF_c electron density map contoured at 1σ and shown 1.6 Å around the stick model. Note that in the Ac-RVRHA-cmk structure (right), the side chain of the Arg residue in the P5 position of the inhibitor has not been modeled due to poor or missing electron density peaks.
 (C) Interactions of RVRHA-cmk with GlpG were analyzed by Ligplot+ (Laskowski and Swindells, 2011). Ligands are shown as thick sticks with carbons in green, proteins as thin sticks with carbons in gray, hydrogen bonds as yellow dashed lines, and amino acids involved in van der Waals contacts are transparent surfaces. The inhibitor forms covalent bonds with S201 and H254 via the chloromethylketone warhead, and it hydrogen bonds with the backbone of residues 196–198 from the L3 loop and residues 248–250 from the L5 loop. van der Waals contacts with the inhibitor are formed by the residues from the L3 loop of GlpG, by S193 and Q189 from TMD3, and by F146 and M120 from the L1 loop that pack against the Val side chain of the P4 position of the inhibitor, as observed previously (Zoll et al., 2014).
 (D) The conformations of peptide inhibitors bound to the active site of GlpG were compared by performing structural alignment of the complexes of VRHAcmk (PDB: 5MT7), RVRHAcmk (PDB: 5MT8), IATAcmk (PDB: 4QO2), and VRMAcho (PDB: 5F5B) in PyMOL (Schrodinger, 2012). Note that the structure of RVRHAcmk suggests where the P5 amino acid points, but the density for this side chain is not visible beyond its β-carbon. The backbone of the ligands in all complexes has virtually identical conformation with the exception of the distortion of the oxyanion by the chloromethylketone, and the biggest differences are found in the conformation of the P2 side chain, which is not surprising, because almost any side chain can be accommodated in this position (Zoll et al., 2014).

the inhibitors *in vitro* (Figure 3A). The most effective compound of the series, bearing a 4-phenyl-butyl tail (compound **11**), already displayed about 1,000-fold lower IC₅₀ than the parent compound **9**. The IC₅₀ of **11** reaches half of the enzyme concentration used in the assay, suggesting that **11** is a potent inhibitor of GlpG.

Next, we examined the relative importance of the peptidyl part for the inhibitory potency. We generated a series of progressively N-terminally truncated variants of **9** and measured their inhibitory potency against GlpG (Figure 3B). Removing the P5 Arg from **9** to yield **12** had virtually no effect on IC₅₀ (0.44 versus 0.55 μM),

while removing the P5 and P4 residues in **13** led to a ~20-fold decrease in potency in comparison with the parent compound **9** (IC₅₀ changes from 0.44 to 9 μM). Removing three residues (from P5 to P3) in **14** led to a dramatic ~150-fold loss of potency, yielding a weak inhibitor with about 65 μM IC₅₀, and the absence of the P5 to P2 residues in **15** resulted in a total ~2,250-fold reduction in potency compared with **9** and IC₅₀ higher than 1 mM. This experiment demonstrates that the non-prime (P4 to P1) and prime sides of the inhibitor contribute to its potency almost equally. The P5 residue can be omitted with only a

Ac-Arg-Val-Arg-His-Ala-X

warhead	Ala-X	IC ₅₀
chloromethylketone		8 μM
trifluoromethylketone		> 1 mM
boronate		8 μM
acylsulfonamide		> 1 mM
thiazolyketone		> 1 mM
α-ketoamide		203 μM

Figure 2. A Screen of Electrophilic Warheads for the Inhibition of Rhomboid Proteases

The optimized parent sequence Ac-RVRHA was linked to electrophilic warheads commonly used for targeting serine proteases (reviewed in Hedstrom, 2002; Walker and Lynas, 2001). The apparent IC₅₀ values of the compounds were measured in 0.05% DDM using 10 μM substrate KSp35 (Ticha et al., 2017) with 1 hr preincubation. Given are the mean values of 2–3 measurements.

marginal effect on inhibitory potency, which can be probably compensated by a suitable prime side tail substituent.

Ketoamides are known to be covalent reversible inhibitors of soluble serine proteases with a classical catalytic triad (Liu et al., 2004). Since rhomboids are unusual serine proteases using only a Ser-His dyad for catalysis (Wang et al., 2006), we investigated the mechanism of rhomboid inhibition by these compounds more closely. Progress curves measured at varying inhibitor concentrations (Figure 4A) had biphasic character; especially at the highest inhibitor concentrations tested, the reaction rate decreased over approximately the first hour and became more or less constant over the next hour (Figures 4A and 4B). This indicates that inhibition was time dependent, which is typical for slow-binding inhibitors (Copeland, 2013b). In addition, upon rapid dilution of inhibitor-saturated enzyme to a subinhibitory concentration, the reaction rate was partially recovered (Figure 4C), together indicating that peptidyl ketoamides exhibit slow-binding reversible behavior (Copeland, 2013a; Singh et al., 2011).

The slow-binding reversible inhibition mechanism can be formally divided into two steps. First, an initial encounter complex (EI) forms, and then a slow step leads to the much more sta-

ble EI* complex (E + I ↔ EI ↔ EI*), usually involving a significant conformational change of the enzyme (Copeland, 2013a). To analyze the contribution of each of these two steps to the mechanism of inhibition of rhomboids by peptidyl ketoamides, we investigated the concentration and time dependence of inhibition kinetics by **10**. The “bending” of biphasic progress curves (Figures 4A and 4B) reflects the rate of “onset of inhibition” described by the rate constant k_{obs} , which can be obtained from progress curve data using non-linear fitting to Equation 1:

$$[P] = v_s t + \frac{(v_i - v_s)}{k_{obs}} [1 - \exp(-k_{obs} t)], \quad (\text{Equation 1})$$

where $[P]$ is the concentration of the reaction products, v_i is the initial reaction rate in the first phase of the biphasic progress curve, and v_s is the steady-state reaction rate (Figure 4B). Analysis of progress curves from Figure 4A showed that v_i was independent of inhibitor concentration, and the plot of k_{obs} against inhibitor concentration fitted well to a linear dependence (Figure 4D). Both phenomena are typical for simple (single-step) slow-binding inhibition (E + I ↔ EI) (Figure 4D); in other words, peptidyl ketoamides behave as “regular” reversible inhibitors but with very low rate constants for association and dissociation (Copeland, 2013a; Morrison, 1982), leading to the slow-binding kinetics. The application of this model yields the apparent inhibitory constant K_i^{app} (i.e., not taking into account the inhibition modality and the influence of the substrate) for **10** of (123 ± 47) nM (Figure 4D).

The true inhibition constant K_i , which is an important, substrate-independent property of an inhibitor, can be calculated from the apparent inhibitory constant K_i^{app} , depending on the inhibitory modality and kinetic parameters of the substrate used. Global non-linear regression fitting of Michaelis curves measured in the presence of increasing concentrations of **10** (plotting v_s against $[S]$) shows that the experimental data are best described by a non-competitive inhibition model (Figure 4E). This inhibition mode means that the inhibitor can bind both to the free enzyme and to the enzyme-substrate complex; in this case specifically, the affinities of the inhibitor to both forms of the enzyme are equal ($\alpha = 1$) (Copeland, 2013a). Although non-competitive modality is non-typical for slow-binding inhibitors, it is conceivable why it is plausible in the case of peptidyl ketoamides and rhomboids. Several studies have suggested that substrate recognition by rhomboid proteases proceeds in two steps, via a docking/interrogation complex, where only a part of substrate’s transmembrane domain interacts with rhomboid, followed by the interaction of the recognition motif with the active site forming the scission-competent complex (Cho et al., 2016; Strisovsky, 2016a, 2016b; Strisovsky et al., 2009) (Figure 4E). Since the active site is unoccupied in the docking complex, binding of an active site-directed inhibitor is possible (Figure 4E), resulting in non-competitive behavior. Under this mechanism of inhibition, the true K_i is identical to K_i^{app} (Copeland, 2013a; Purich, 2010). Similar progress curve analyses of **9** and **11** yield their k_{on} and k_{off} rate constants, their K_i^{app} ($K_i^{app} = k_{off}/k_{on}$), and the true K_i values of (219 ± 76) nM and (45 ± 8) nM, respectively (Figure 4F and Table 1). In summary, this kinetic analysis shows that the peptidyl ketoamides described here are high-affinity inhibitors of rhomboid proteases unprecedented in the literature.

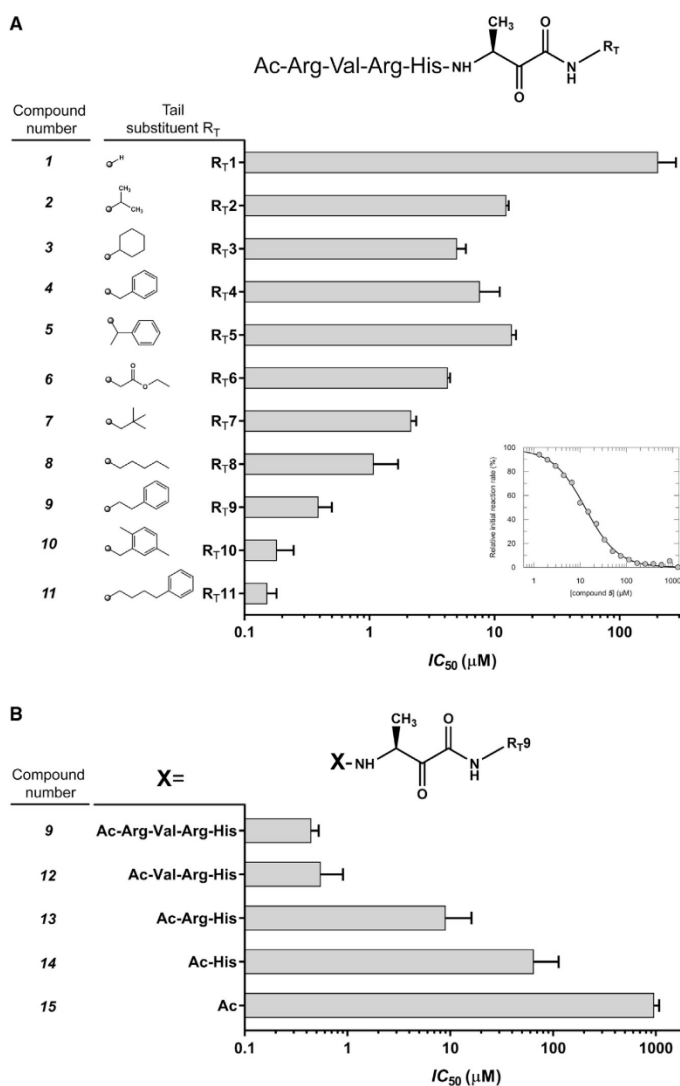


Figure 3. Modification of the Prime-Side Substituent at the Amide Group of Peptidyl Ketoamides Enhances Their Potency by Orders of Magnitude

(A) A screen of the effect of the tail substituent RT on the inhibitory properties of ketoamide inhibitors of GlpG based on the parent compound Ac-RVRHA-CONH₂. The apparent IC₅₀ values of all compounds were measured in 0.05% (w/v) DDM and 10 μM KSp35 (Ticha et al., 2017) with 1 hr preincubation. The IC₅₀ values of the most effective compounds 9, 10, and 11 are three orders of magnitude lower than that of the parent compound 1. The reported values are best-fit means with SD representative of 2–3 measurements. The inset shows a typical inhibition curve.

(B) The significance of the peptidyl part in compound 9. The peptidyl part of 9 was progressively truncated from the N terminus, and the apparent IC₅₀ values of all compounds were measured in 0.05% (w/v) DDM and 10 μM KSp35 (Ticha et al., 2017) with 1 hr preincubation. The reported values are best-fit means with SD representative of 2–3 measurements.

based probe (ABP) competition assays that enable a more general and substrate-independent measurement of inhibitory potency, because they rely solely on the competition between a fluorescently labeled activity-based probe and the tested inhibitor (Nguyen et al., 2015; Serim et al., 2012). The assays we employed used fluorophosphonate ABPs that target the catalytic serine of a wide-range of serine hydrolases (Bachovchin et al., 2014), including rhomboids (Xue et al., 2012), and are thus very practical general detection reagents even for serine hydrolases for which sensitive substrates might not be available.

First, we tested 9, 10, and 11 against a panel of bacterial and eukaryotic rhomboid proteases (Wolf et al., 2015), and found that all three compounds potently competed with ABP labeling of rhomboids from bacterium *Providencia stuartii* (AarA), archaeobacterium *Methanocaldococcus jannaschii* (MjROM), and three closely related rhomboids from bacteria *E. coli* (EcGlpG), *Haemophilus influenzae* (HiGlpG), and *Vibrio cholerae* (VcROM). Compounds 9, 10, and 11 outcompeted the ABP even at a concentration of 500 nM, suggesting that they were potent inhibitors of these rhomboid proteases. In contrast, none of these compounds were able to compete with the ABP labeling of rhomboid protease from the bacterium *Aquifex aeolicus* (AaROM), rhomboids from *Drosophila* (DmRho1) and mouse (MmRHBDL3), and they only partially inhibited labeling of rhomboid protease from bacterium *Thermotoga maritima* (TmROM) at 50 μM (Figure 5A). These data demonstrate that already these first-generation

Selectivity of Peptidyl Ketoamides

Any enzyme inhibitors to be used as specific tools for cell biology or as starting points for drug development must show sufficient level of selectivity toward their intended target. This is particularly important for compounds that react with the catalytic nucleophile common to many serine hydrolases. Only limited tests of selectivity have been conducted for the currently used rhomboid inhibitors isocoumarins, β-lactams and β-lactones, at best interrogating them against trypsin or chymotrypsin (Pierrat et al., 2011; Vosyka et al., 2013). To map the selectivity of peptidyl ketoamides more objectively and widely, we employed activity-

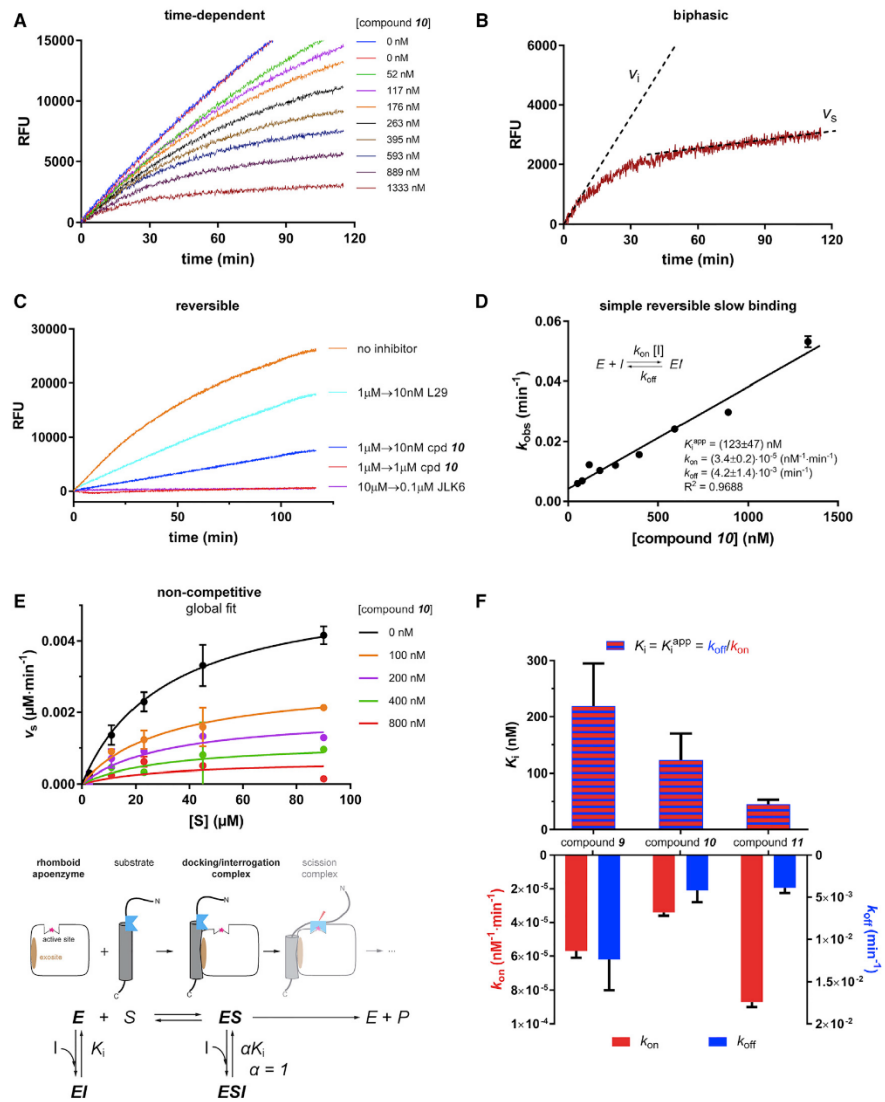


Figure 4. Mechanism of Inhibition of Rhomboid Protease GlpG by Peptidyl Ketoamides Analyzed Using Fluorogenic Transmembrane Peptide Substrates

(A) Progress curves in the presence of increasing concentrations of compound 10 show biphasic character, which is typical for slow-binding inhibitors (Copeland, 2013a; Morrison, 1982). GlpG (0.5 nM) was incubated with 25 μ M substrate KSp93 in the presence of 0.05% (w/v) DDM and 0–1,333 nM 10. Fluorescence at 493 nm was followed to monitor substrate cleavage.

(B) Biphasic progress curves characterized by an initial reaction rate (v_i) and steady state reaction rate (v_s). The progress curve at 1,333 nM compound 10 from the experiment in (A) is shown in detail, and both reaction rates obtained from non-linear regression into Equation (1) are shown as dotted lines.

(C) Reversibility of inhibition by ketoamides was assessed by the rapid dilution method (Harper et al., 1985; Harper and Powers, 1985). Compound 10 (1 μ M) was pre-incubated with 0.4 μ M GlpG, 0.05% (w/v) DDM at 37 $^\circ$ C for 1 hr, leading to complete inhibition. This solution was then rapidly diluted 100-fold either into the reaction buffer containing 10 μ M substrate KSp64 (Ticha et al., 2017) (yielding final 10 nM inhibitor) or into the reaction buffer with 10 μ M substrate KSp64 and 1,000 nM 10. For comparison, we used β -lactam L29 (Pierrat et al., 2011) at 1 μ M and isocoumarin JLK6 (Vinothkumar et al., 2010) at 10 μ M as known reversible

(legend continued on next page)

Table 1. Summary of the Inhibition Properties of Compounds 9–11

Compound	GlpG			YqgP	
	K_i (nM)	k_{on} (10^{-6} nM $^{-1}$ · min $^{-1}$)	k_{off} (10^{-3} min $^{-1}$)	IC $_{50}$ <i>In Vivo</i> (nM)	IC $_{50}$ <i>In Vivo</i> (nM)
9	220 ± 80	5.7 ± 0.4	12.0 ± 0.4	8.8 ± 0.4	ND
10	120 ± 50	3.4 ± 0.2	4.2 ± 1.4	6.0 ± 0.1	ND
11	45 ± 8	8.7 ± 0.3	3.9 ± 0.6	2.7 ± 0.1	~5–10

Values for GlpG are reported as means ± SD.

peptidyl ketoamides can discriminate between diverse rhomboid proteases.

We next examined peptidyl ketoamides for their possible off-target effects on other serine proteases. To get a representative picture of the selectivity of peptidyl ketoamides, we employed a recently developed EnPlex technology, which allows multiplex analysis of ABP competition with about 100 human serine hydrolases, mostly proteases (Bachovchin et al., 2014). Profiling of **9**, **10**, and **11** showed that in the concentration range where they inhibit rhomboid proteases, they fail to inhibit most of the tested human serine hydrolases with the exception of prolylcarboxypeptidase (PRCP) and the sequence related dipeptidyl-peptidase 7 (DPP7) (Figure 5B). To put this into the context of the current generation of rhomboid inhibitors, isocoumarins S006 and S016 (Vosyka et al., 2013) hit about a dozen serine hydrolases in the same concentration range. The β -lactam L41 (Pierrat et al., 2011) inhibited appreciably only one enzyme (predicted serine carboxypeptidase CPVL), but it is much less potent on rhomboids than **9**, **10**, and **11**, and it does not inhibit GlpG completely *in vivo* (Pierrat et al., 2011). The selectivity profile of ketoamide inhibitors of rhomboids is similar to the profile of clinically used ketoamide inhibitors of the hepatitis C protease (Bachovchin et al., 2014), indicating that the rhomboid-targeting N-modified peptidyl ketoamides are sufficiently selective with minimal risk of cross-reactivity against other serine proteases.

Peptidyl Ketoamides Potently Inhibit Rhomboids in Living Cells

Having established the mechanism of rhomboid inhibition by peptidyl ketoamides in detergent micelles, and having shown that **9**, **10**, and **11** are able to inhibit potently rhomboid prote-

ases from several Gram-negative bacteria (Figure 5A), we next tested whether the inhibitors will be able to target rhomboid proteases embedded in their native lipid bilayer in live cells. First, we expressed the model substrate derived from LacYTM2 in *E. coli* expressing endogenous levels of GlpG, incubated the bacterial cultures in the presence of increasing concentrations of **9**, **10**, and **11**, and detected the steady-state levels of substrate processing by quantitative near-infrared western blotting (Figure 6A). The calculated substrate conversion values relative to the uninhibited reaction were plotted against the inhibitor concentration yielding the *in vivo* IC $_{50}$ values. Strikingly, the most effective compound **11** had an *in vivo* IC $_{50}$ value of 2.7 nM, which is three orders of magnitude lower than any other currently known rhomboid inhibitors (Cho et al., 2016; Pierrat et al., 2011).

We then extended the range of organisms to *Bacillus subtilis*, a representative of Gram-positive bacteria, which have a thick cell wall and include major pathogens such as *Staphylococcus*, *Listeria*, *Streptococcus*, and others. Since the endogenous substrate of the *B. subtilis* rhomboid protease YqgP is unknown, and no robust and rescuable phenotypes have been reported for YqgP, we focused on inhibition of cleavage of a model substrate. Of the common model rhomboid substrates, YqgP cleaves LacYTM2 reasonably well (Ticha et al., 2017). We have thus expressed MBP-LacYTM2-Trx (Strisovsky et al., 2009) from the ectopic *xkdE* locus (Gerwig et al., 2014) in the wild-type *B. subtilis* 168 (BS87) and its *yqgP* deletion mutant (BS88) on an otherwise rhomboid-free background. Although the substrate was to some extent truncated by unknown processes in the $\Delta yqgP$ strain, a specific, closely co-migrating rhomboid-generated N-terminal cleavage product (Figure 6B) was produced in the YqgP wild-type strain BS87 but not in the $\Delta yqgP$

and irreversible inhibitors of rhomboid proteases, respectively. Activity recovery was followed by measuring fluorescence over the course of 120 min with excitation at 553 nm and emission at 583 nm.

(D) Progress curves of KSp93 cleavage at increasing concentrations of **10** measured under (A) were analyzed by non-linear regression as described for slow-binding inhibition (Copeland, 2013a; Morrison, 1982) using GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla, California, USA) to yield the rate constant for the onset of inhibition, k_{obs} . The linear character of the dependence of k_{obs} on inhibitor concentration is typical for a simple slow-binding mechanism (inset), and its linear regression allows determination of the underlying apparent inhibitory constant K_i^{app} and its constituent rate constants k_{on} and k_{off} (inset). The k_{obs} values are reported as best-fit mean ± SD.

(E) The influence of inhibitor concentration on the apparent K_M and k_{cat} suggests the mode of inhibition by compound **10**. Michaelis curves at the indicated inhibitor concentrations were measured by plotting v_s (measured as in Figure 4B) against substrate concentration using 1 nM GlpG, 0.15% (w/v) DDM and highly sensitive substrate KSp96. The data were globally fitted to the models of competitive, non-competitive (figure top), uncompetitive, and mixed inhibition as implemented in GraphPad Prism 7.02, and their statistical analysis yielded the non-competitive mechanism (figure bottom) as the best fit. The middle of the figure shows a schematic mapping of this mechanism onto the consensual model of substrate recognition by rhomboid proteases. The data points in the Michaelis plots (figure top) represent means ± SD of duplicate measurements.

(F) Summary of inhibition kinetics parameters of compounds **9**, **10**, and **11**. The apparent inhibitory constants K_i^{app} (blue-red striped columns) and the constituent rate constants k_{on} (red columns) and k_{off} (blue columns) were determined from progress curve analysis as shown in (A and C) (note that $K_i^{app} = k_{off}/k_{on}$). For non-competitive inhibitors, the true inhibitory constant K_i equals K_i^{app} . Note that **11** is a highly potent inhibitor with K_i of (45 ± 8) nM. Graphs show best-fit means with SDs.

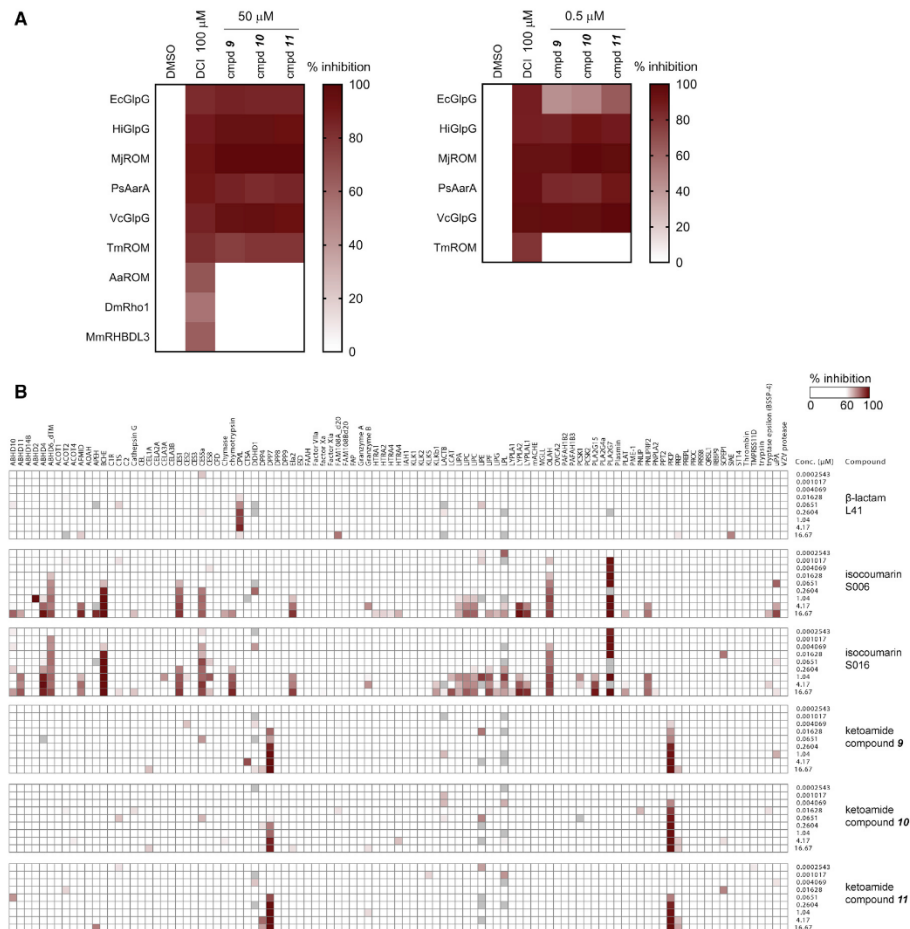


Figure 5. Selectivity of Peptidyl Ketoamides

(A) Selectivity of compounds **9**, **10**, and **11** for nine rhomboid proteases was profiled using activity-based probe competition assay at 50 μ M and 0.5 μ M concentration. The upper limit of enzyme concentration was 0.4 μ M.
 (B) Selectivity of compounds **9**, **10**, and **11** against human serine hydrolases was analyzed using EnPlex as described (Bachovchin et al., 2014).

strain BS88 (Figure 6B). In the absence of any inhibitors, MBP-LacYTM2-Trx was cleaved to about 75% conversion by the endogenous YqgP, and addition of **11** into the growth media completely inhibited substrate cleavage at 50 nM (Figure 6B), indicating that the compound can penetrate the Gram-positive cell wall easily. Moreover, since compound **11** also inhibits several homologs of GlpG (Figure 5A), it is safe to assume that YqgP orthologs in other *Bacilli*, *Lactobacilli*, *Staphylococci*, and *Listeria* might be equally susceptible to inhibition by the described inhibitors, and compound **11** and its analogs can be directly used for chemical proteomics and cell biological studies of rhomboid proteases in Gram-positive bacteria.

N-Modified Peptidyl Ketoamides Bind the Rhomboid Active Site in a Substrate-like Manner Occupying the S4 to S2' Subsites

To understand why peptidyl ketoamides are such efficient rhomboid inhibitors and to establish the basis for structure-guided design of their improved variants, we determined the co-crystal structures of GlpG with **9** and **10** (Figure 7A). The complexes were formed by soaking the inhibitors into apoenzyme crystals, and the structures were solved using diffraction data to 2.16 and 1.78 Å resolution, respectively, allowing detailed comparison of their binding modes. In both cases, the pentapeptide RVRHA binds the active site cavity as an extended β strand,

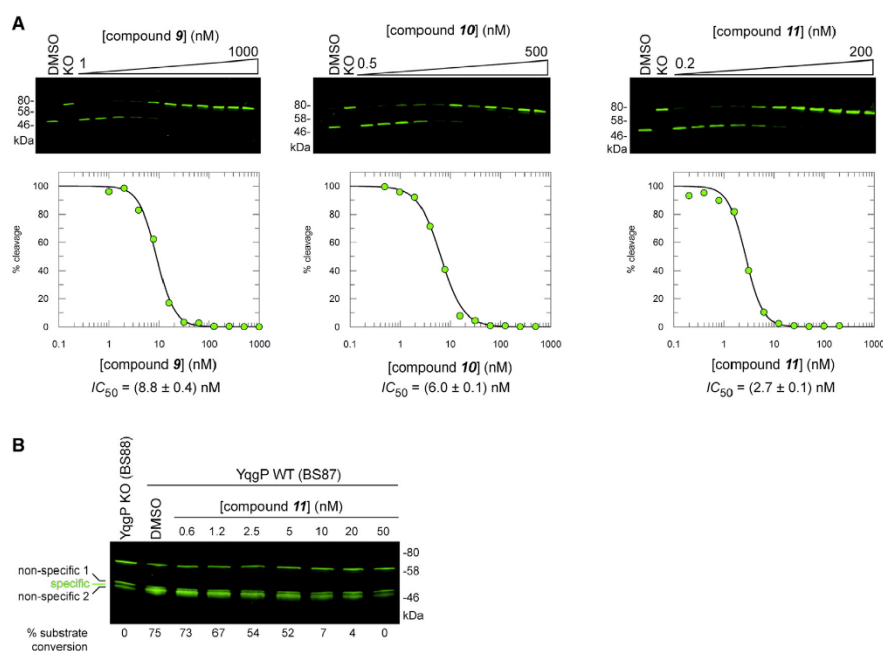


Figure 6. Peptidyl Ketoamides Potently Inhibit Rhomboid Activity in the Membranes of Living Cells

(A) Inhibition of endogenous GlpG by compounds **9**, **10**, and **11** in the membranes of live *E. coli*. The substrate MBP-FLAG-LacYTM2-Trx (Strisovsky et al., 2009) was expressed in wild-type *E. coli* NR698 with genetically permeabilized outer membrane (Ruiz et al., 2005) in the presence of increasing concentrations of inhibitors as described in STAR Methods. Substrate cleavage was measured in cell lysates by immunoblotting for FLAG and quantified using near-infrared fluorescence. The reported *in vivo* IC_{50} values are best-fit means with SD representative of 2–3 measurements. DMSO, dimethylsulfoxide vehicle control; KO, *E. coli* *glpG::tet*.

(B) Inhibition of endogenous YqgP by compound **11** in the membranes of live *B. subtilis*. The substrate AmyE_{SP}-MBP-FLAG-LacYTM2-Trx-HA was expressed in *Bacillus subtilis* 168 (*ycdA::neo*, *xdkE::AmyE_{SP}-MBP-LacYTM2-Trx(erm, lin)*) (BS87) in the presence of increasing concentrations of inhibitors as described in STAR Methods. Substrate cleavage was detected in cell lysates by immunoblotting for FLAG and detection by near-infrared fluorescence. Unspecific cleavage of the substrate was corrected for by subtracting the intensity of the unspecific bands formed in the YqgP knockout control cells (BS88) from the product band and the closely co-migrating unspecific bands observed in the YqgP positive cells (BS87). This treatment was necessary because the specific cleavage product could not be resolved sufficiently well from the non-specific bands to be integrated separately. DMSO, dimethylsulfoxide vehicle control; YqgP KO, *Bacillus subtilis* 168 (*ycdA::neo*, *yqgP::tet*, *xdkE::AmyE_{SP}-MBP-LacYTM2-Trx(erm, lin)*) (BS88).

virtually identically to the binding mode of Ac-RVRHA-cmk (Figure 1C). We do observe electron density for the side chain of arginine in the P5 position in both structures, but its conformation differs between **9** and **10** (Figure 7A), and it is influenced by crystal contacts with the same residue from a neighboring molecule in the crystal (data not shown).

In both inhibitors, the ketoamide warhead is covalently bonded via its proximal carbon to the side-chain oxygen of the catalytic S201, and it engages in a network of six hydrogen bonds in the active site (Figure 7B). The oxyanion formed by the proximal carbonyl oxygen accepts hydrogen bonds from His150 and the main-chain amide nitrogen of the catalytic serine, and the distal ketoamide carbonyl oxygen accepts hydrogen bonds from both H150 and N154, thus amply saturating the hydrogen-bonding groups engaged in the stabilization of the oxyanion (Cho et al., 2016). Furthermore, the ketoamide nitrogen donates a hydrogen bond to H254 and to the S201 side-chain oxygen covalently bound to the warhead. The resulting network

of six hydrogen bonds (Figure 7B) probably helps position the ketoamide warhead in the proximity of the hydroxyl of the catalytic S201 to enhance its chemical reactivity in a conformation-dependent manner.

The tail substituents of **9** and **10** (R_T9 and R_T10) interact with the prime side of GlpG, buried in a cavity delimited by the side chains of amino acids F245, M247, and M249 from the L5 loop, W236 from TMD5, F153 and W157 from TMD2, and residues V204, M208, Y205, H254, H150, and N154 (Figure 7C). The different sets of residues making van der Waals contacts with each ketoamide tail are shown in magenta. The NH group of the side chain of W236 seems to form a weak H-π bond with the phenyl ring of the tail of both compounds, and F245 engages in π...π stacking against the dimethylbenzyl in R_T10 (Figure 7C). Structural alignment of both ketoamide complexes to the complex of the β-lactam L29 (Vinothkumar et al., 2013) and isocoumarin S016 (Vosyka et al., 2013) (Figure 7D) shows that the tails of **9** and **10** bind in a similar area (the S2' subsite)

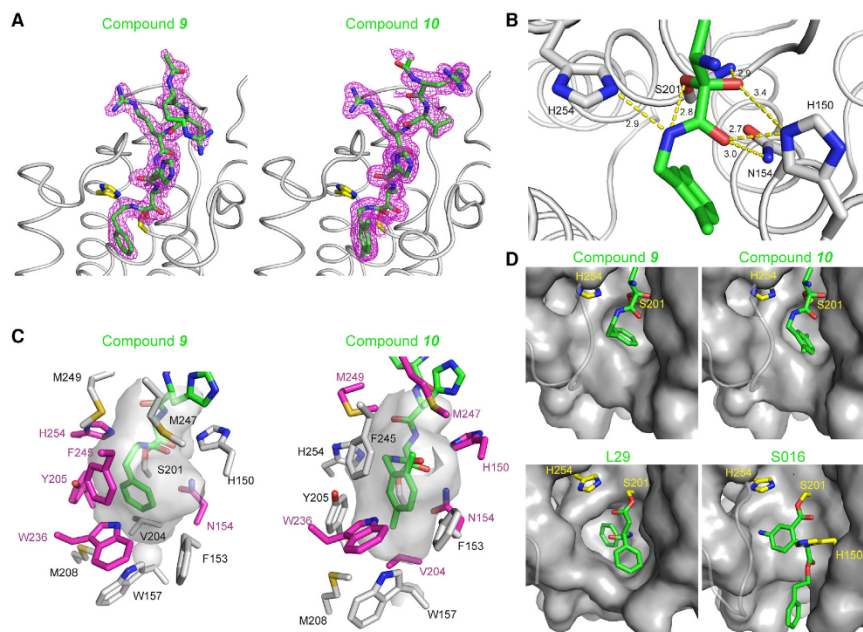


Figure 7. N-Substituted Peptidyl Ketoamides Bind GlpG in a Substrate-like Manner and Occupy the S4 to S2' Subsites of the Rhomboid Active Site

(A) Electron density map and ligand stick model of **9** and **10** in complex with GlpG. Compounds **9** and **10** were soaked into the native crystals of GlpG, and the structures of the complexes were solved by X-ray diffraction (for statistics, see Table S1). The catalytic dyad is shown as yellow sticks and the inhibitors as green sticks surrounded by the $2mF_o - DF_c$ electron density map contoured at 1σ and shown 1.6 Å around the inhibitor model. In the complex of **9**, the electron density for the Arg residue in the P5 position was weaker, and the side chain has been modeled in a different conformation than in the complex of **10**, which was solved to a higher resolution and where the side chain of the Arg in the P5 position is defined clearly.

(B) Hydrogen bond engagement by the warhead of compound **10** in the active site of GlpG was analyzed using the HBplus program (McDonald and Thornton, 1994) implemented in Ligplot+ (Laskowski and Swindells, 2011) with default criteria (donor ... acceptor [D ... A] distance cutoff of 3.9 Å; donor ... acceptor-acceptor antecedent [D ... A-AA] angle of 90°).

(C) Interaction pattern of inhibitor tails in the S2' site of GlpG. The cavity surrounding the tails of **9** and **10** is shown as an inverse surface, and the side chains lining the cavity are shown as sticks. The residues engaged in van der Waals interactions (identified by Ligplot+) with the tails of the inhibitors are shown in magenta.

(D) Comparison of binding modes of the S2' binding moieties in compounds **9**, **10**, L29 (Pierrat et al., 2011), and S016 (Vosyka et al., 2013) in the respective complex structures PDB: 5MT6, 5MTF, 3ZMI, and 3ZEB. Protein is shown as a gray surface, catalytic dyad carbons in yellow, and ligand carbons in green. The L5 loop residues 245–250 are shown as semitransparent loops for clarity. All structures are oriented in the same way.

as the significantly larger groups of inhibitors L29 and S016. This alignment shows that the prime side of the GlpG active site is rather malleable, and larger or branched tails could be accommodated at the amide nitrogen of peptidyl ketoamides. This is likely to provide additional selectivity or binding energy and delineates one possible direction of further development of ketoamides as rhomboid inhibitors. The results presented here open the door to systematic development of rhomboid protease inhibitors in medically relevant contexts such as malaria (Baker et al., 2006; O'Donnell et al., 2006), Parkinson's disease (Chu, 2010; Meissner et al., 2015), and cancer (Song et al., 2015).

DISCUSSION

Here, we discover that peptidyl ketoamides bearing a substantial hydrocarbon modification at the ketoamide nitrogen are efficient

inhibitors of rhomboid intramembrane proteases, superior to any known rhomboid inhibitors in selectivity and by up to three orders of magnitude in potency. We also show that both of these properties are tunable by optimization of the peptide sequence and the character of the ketoamide “tail” substituent, defining a platform for the development of specific and potent rhomboid inhibitors. Since ketoamides are clinically used pharmacophores (Njoroge et al., 2008), our discovery of this pharmacologically compliant chemotype for rhomboid proteases enables the design of rhomboid inhibitors for cell biological and pharmacological use.

Structural analysis of peptidyl ketoamides complexed to GlpG reveals that they bind in a substrate-like manner, occupying the P4 to P2' subsites (Figure 7A). The presence of residues in the P5 and P6 positions has been reported to improve the inhibition potency of peptidyl aldehydes significantly, but these residues

could not be observed in any co-crystal structures (Cho et al., 2016). We do observe weak electron density for the side chain of Arg in the P5 position, but its conformation in the final crystallographic models of the complexes of **9** and **10** differs, indicating some degree of flexibility, and it is probably influenced by crystal contacts. In addition, the P5 residue does not contribute significantly to the inhibition potency of **9** (Figure 3B) and is thus dispensable.

The binding mode of peptidyl ketoamides suggests that they can access the rhomboid active site from bulk solvent, and probably do not need prior partitioning into the membrane. They are covalent (Figure 7) and reversible (Figure 4C), and their kinetics of binding to rhomboid is adequately described by a one-step slow-binding mechanism (Figure 4D). Their inhibition modality is non-competitive (Copeland, 2013a) (Figure 4E), implying that they can bind to the free enzyme as well as the docking/interrogation complex during rhomboid catalysis (Strisovsky, 2016a). This is consistent with the proposed mechanism of inhibition of rhomboid protease GlpG by peptidyl aldehydes (Cho et al., 2016).

For the development of peptidyl ketoamides as rhomboid inhibitors, subsite preferences of the given rhomboid protease must be mapped efficiently. This could be achieved using classical positional scanning peptide libraries starting from a known substrate sequence. Given that the effects of the amino acids in the P5–P1 positions are additive (Figure 1A), the optimal substrate could be formed by combining the single subsite preferences identified in the positional scan. An alternative method for mapping subsite preferences at both the prime and non-prime sides could be multiplex substrate profiling using designed peptide libraries and mass spectrometry (O'Donoghue et al., 2012), although its application to rhomboids has not been tested yet.

The second module determining the potency and selectivity is the tail substituent at the ketoamide nitrogen. Here, the effects of flexibility versus rigidity, branching, and polarity of the substituents need to be investigated to explore the available chemical and conformational space. A more speculative direction of further improvement of the inhibitors may involve cyclization via the tail substituent and the P2 residue, which seems sterically possible and unobstructive in the enzyme-inhibitor complex (Figure 7A). Such cyclization could improve the potency of the inhibitor by conformationally restricting it near the bound conformation.

Finally, peptidyl ketoamides have been used clinically to treat hepatitis C infection (boceprevir, telaprevir) (Njoroge et al., 2008), suggesting that both the intracellular availability and metabolic stability of rhomboid-targeting peptidyl ketoamides can most likely be modified for compliance with pharmacological needs. The potential of rhomboid inhibitors in pharmacologically relevant settings has yet to be proven, but it currently seems that inhibitors of *Plasmodium* rhomboids might be therapeutic for malaria (Baker et al., 2006; Lin et al., 2013), inhibitors of the human mitochondrial rhomboid protease PARL might stimulate mitophagy (Meissner et al., 2015) and thus be disease-modifying in the context of Parkinson's disease (Chan and McQuibban, 2013), and inhibitors of human RHBDL4 could be targeting EGF receptor signaling by transforming growth factor α in colorectal cancer (Song et al., 2015). Specific rhomboid protease inhibitors such

as those that we describe here will serve as key tools for the validation and exploitation of these and other upcoming therapeutic opportunities involving rhomboid proteases.

SIGNIFICANCE

Intramembrane proteases of the rhomboid family are widely conserved and have been implicated in malaria, colon cancer, and Parkinson's disease. They represent potentially attractive drug targets, but until now, no specific, potent, and pharmacologically compatible inhibitors have been available. Here, we discover that peptidyl ketoamides are the first such potent and specific inhibitors of rhomboid proteases, and we delineate a general modular way for their design against diverse rhomboid enzymes. This discovery can have a broad impact on the cell biology of rhomboid proteases and on drug discovery targeting this family of enzymes in the context of infectious diseases, cancer, and neurodegeneration.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENTS AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Constructs and Cloning
 - Protein Expression and Purification
 - Chemical Synthesis
 - Protein Crystallography
 - Rhomboid Activity and Inhibition Assays
 - Inhibitor Selectivity Profiling
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes one table and supplemental text and can be found with this article online at <https://doi.org/10.1016/j.chembiol.2017.09.007>.

AUTHOR CONTRIBUTIONS

K.S. conceived and coordinated the study, designed and evaluated experiments, and wrote the paper with the input of the co-authors. P.M. and S.S. designed and S.S. performed all chemical syntheses, and A.T. designed, performed, and evaluated all kinetics and inhibition measurements with the help of K.S. and J.S. in the initial stages. K.R.V., D.C.M., and P.P. performed all crystallographic experiments and evaluated the data with the input of M.L. M.T.N.N. and S.H.L.V. performed selectivity profiling against rhomboid proteases, J.B. performed all experiments on *B. subtilis*, and D.C.J. and D.A.B. performed and evaluated the EnPlex experiments.

ACKNOWLEDGMENTS

We thank Mirka Blechová and Lenka Monincová for peptide synthesis and HPLC/MS analyses, Radko Souček for amino acid analysis, Jiří Brynda for help with crystallographic data analysis, Pavel Srb and Marek Ingr for help and advice on data fitting and kinetics, Petra Rampirová for DNA cloning

and laboratory assistance, Libor Krásný and Tom Silhavy for reagents and advice, and Colin Adrain and Cyril Bařinka for critical reading of the manuscript. K.S. was a recipient of a Purkyne Fellowship of the Academy of Sciences of the Czech Republic and acknowledges support also from EMBO (installation grant no. 2329), Ministry of Education, Youth and Sports of the Czech Republic (project nos. LK11206 and LO1302), Marie Curie Career Integration Grant (project no. 304154), Gilead Sciences & IOCB Research Centre, and the National Subvention for Development of Research Organisations (RVO: 61388963) to the Institute of Organic Chemistry and Biochemistry, P.P. acknowledges support from the Ministry of Education of the Czech Republic (program "NPU I"), project no. LO1304, and J.S. and J.B. acknowledge support from the Grant Agency of Charles University (GA UK) in Prague (PhD grant project nos. 232313 and 170214). K.R.V. was supported by an MRC grant (MC_U105184322) as part of R. Henderson's group. D.A.B. was supported by Josie Robertson Foundation and the MSKCC (core grant P30 CA008748), and D.C.J. by the NIH (T32 GM115327-Tan). M.L. was supported by the Czech Science Foundation (grant number P208/12/G016). M.T.N.N. and S.H.L.V. are supported by the Deutsche Forschungsgemeinschaft and the Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen.

Received: April 13, 2017
 Revised: August 19, 2017
 Accepted: September 18, 2017
 Published: October 26, 2017

SUPPORTING CITATIONS

The following references appear in the Supplemental Information: Bastiaans et al., 1997; Cao et al., 2010; Coste et al., 1994; D'Andrea and Scola, 2008; Dondoni and Perrone, 1993; Semple et al., 2000; Souček and Urban, 1995; Tulla-Puche et al., 2008; Venkatraman et al., 2006; Yin et al., 2007.

REFERENCES

- Ahrlich, R., Bär, M., Häser, M., Horn, H., and Kölmel, C. (1989). Electronic structure calculations on workstation computers: the program system turbo-mole. *Chem. Phys. Lett.* **162**, 165–169.
- Bachovchin, D.A., Koblan, L.W., Wu, W., Liu, Y., Li, Y., Zhao, P., Woznica, I., Shu, Y., Lal, J.H., Poplawski, S.E., et al. (2014). A high-throughput, multiplexed assay for superfamily-wide profiling of enzyme activity. *Nat. Chem. Biol.* **10**, 656–663.
- Baker, R.P., Wijetilaka, R., and Urban, S. (2006). Two *Plasmodium* rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. *PLoS Pathog.* **2**, e113.
- Bastiaans, H.M.M., vanderBaan, J.L., and Ottenheijm, H.C.J. (1997). Flexible and convergent total synthesis of cyclotheonamide B. *J. Org. Chem.* **62**, 3880–3889.
- Cao, H., Liu, H., and Domling, A. (2010). Efficient multicomponent reaction synthesis of the schistosomiasis drug praziquantel. *Chemistry* **16**, 12296–12298.
- Chan, E.Y., and McQuibban, G.A. (2013). The mitochondrial rhomboid protease: its rise from obscurity to the pinnacle of disease-relevant genes. *Biochim. Biophys. Acta* **1828**, 2916–2925.
- Chatterjee, S., Dunn, D., Tao, M., Wells, G., Gu, Z.Q., Bihovsky, R., Ator, M.A., Siman, R., and Mallamo, J.P. (1999). P2-achiral, P'-extended alpha-ketoamide inhibitors of calpain I. *Bioorg. Med. Chem. Lett.* **9**, 2371–2374.
- Cho, S., Dickey, S.W., and Urban, S. (2016). Crystal structures and inhibition kinetics reveal a two-stage catalytic mechanism with drug design implications for rhomboid proteolysis. *Mol. Cell* **61**, 329–340.
- Chu, C.T. (2010). A pivotal role for PINK1 and autophagy in mitochondrial quality control: implications for Parkinson disease. *Hum. Mol. Genet.* **19**, R28–R37.
- Copeland, R.A. (2013a). Reversible modes of inhibitor interactions with enzymes. In *Evaluation of Enzyme Inhibitors in Drug Discovery* (John Wiley), pp. 57–121.

Copeland, R.A. (2013b). Slow binding inhibitors. In *Evaluation of Enzyme Inhibitors in Drug Discovery* (John Wiley), pp. 203–244.

Coste, J., Frerot, E., and Jouin, P. (1994). Coupling N-methylated amino-acids using pybrop and pyclop halogenophosphonium salts - mechanism and fields of application. *J. Org. Chem.* **59**, 2437–2446.

D'Andrea, S., and Scola, P.M. (2008). Inhibitors of hepatitis C virus. Bristol-Myers Squibb Company. US patent US2008107623 (A1), filed October 25, 2007, and published May 8, 2008.

Dickey, S.W., Baker, R.P., Cho, S., and Urban, S. (2013). Proteolysis inside the membrane is a rate-governed reaction not driven by substrate affinity. *Cell* **155**, 1270–1281.

Dondoni, A., and Perrone, D. (1993). 2-Thiazolyl alpha-amino ketones - a new class of reactive intermediates for the stereocontrolled synthesis of unusual amino-acids. *Synthesis* (Stuttgart), 1162–1176.

Drag, M., and Salvesen, G.S. (2010). Emerging principles in protease-based drug discovery. *Nat. Rev. Drug Discov.* **9**, 690–701.

Eggert, U.S., Ruiz, N., Falcone, B.V., Branstrom, A.A., Goldman, R.C., Silhavy, T.J., and Kahne, D. (2001). Genetic basis for activity differences between vancomycin and glycolipid derivatives of vancomycin. *Science* **294**, 361–364.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132.

Evans, P.R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 282–292.

Fanfrik, J., Holub, J., Ruzickova, Z., Rezac, J., Lane, P.D., Wann, D.A., Hnyk, D., Ruzicka, A., and Hobza, P. (2016). Competition between halogen, hydrogen and dihydrogen bonding in brominated carboranes. *ChemPhysChem* **17**, 3373–3376.

Fleig, L., Bergbold, N., Sahasrabudhe, P., Geiger, B., Kaltak, L., and Lemberg, M.K. (2012). Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins. *Mol. Cell* **47**, 558–569.

Gerwig, J., Kiley, T.B., Gunka, K., Stanley-Wall, N., and Stulke, J. (2014). The protein tyrosine kinases EpsB and PtkA differentially affect biofilm formation in *Bacillus subtilis*. *Microbiology* **160**, 682–691.

Gibson, D.G. (2011). Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol.* **498**, 349–361.

Grimme, S. (2006). Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799.

Harper, J.W., and Powers, J.C. (1985). Reaction of serine proteases with substituted 3-alkoxy-4-chloroisocoumarins and 3-alkoxy-7-amino-4-chloroisocoumarins: new reactive mechanism-based inhibitors. *Biochemistry* **24**, 7200–7213.

Harper, J.W., Hemmi, K., and Powers, J.C. (1985). Reaction of serine proteases with substituted isocoumarins: discovery of 3,4-dichloroisocoumarin, a new general mechanism based serine protease inhibitor. *Biochemistry* **24**, 1831–1841.

Hedstrom, L. (2002). Serine protease mechanism and specificity. *Chem. Rev.* **102**, 4501–4524.

Jensen, F. (2006). *Introduction to Computational Chemistry* (Wiley).

Kabsch, W. (2010). Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132.

Klamt, A., and Schüürmann, G. (1993). COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2*, 799–805.

Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2256–2268.

Laskowski, R.A., and Swindells, M.B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model* **51**, 2778–2786.

Lebedev, A.A., Young, P., Isupov, M.N., Moroz, O.V., Vagin, A.A., and Murshudov, G.N. (2012). JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 431–440.

- Lee, C., Kang, H.J., Hjeltn, A., Qureshi, A.A., Nji, E., Choudhury, H., Beis, K., de Gier, J.W., and Drew, D. (2014). MemStar: a one-shot *Escherichia coli*-based approach for high-level bacterial membrane protein production. *FEBS Lett.* **588**, 3761–3769.
- Lemberg, M.K., Menendez, J., Misik, A., Garcia, M., Koth, C.M., and Freeman, M. (2005). Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. *EMBO J.* **24**, 464–472.
- Lin, J.W., Meireles, P., Prudencio, M., Engelmann, S., Annoura, T., Sajid, M., Chevalley-Maurel, S., Ramesar, J., Nahar, C., Avramut, C.M., et al. (2013). Loss-of-function analyses defines vital and redundant functions of the *Plasmodium* rhomboid protease family. *Mol. Microbiol.* **88**, 318–338.
- Liu, Y., Stoll, V.S., Richardson, P.L., Saldívar, A., Klaus, J.L., Molla, A., Kohlbrenner, W., and Kati, W.M. (2004). Hepatitis C NS3 protease inhibition by peptidyl-alpha-ketoamide inhibitors: kinetic mechanism and structure. *Arch. Biochem. Biophys.* **421**, 207–216.
- McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 32–41.
- McDonald, I.K., and Thornton, J.M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
- Meissner, C., Lorenz, H., Hehn, B., and Lemberg, M.K. (2015). Intramembrane protease PARL defines a negative regulator of PINK1- and PARK2/Parkin-dependent mitophagy. *Autophagy* **11**, 1484–1498.
- Miroux, B., and Walker, J.E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298.
- Mitchell, E.M., Artymiuk, P.J., Rice, D.W., and Willett, P. (1990). Use of techniques derived from graph-theory to compare secondary structure motifs in proteins. *J. Mol. Biol.* **212**, 151–166.
- Morrison, J.F. (1982). The slow-binding and slow, tight-binding inhibition of enzyme-catalysed reactions. *Trends Biochem. Sci.* **7**, 102–105.
- Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367.
- Nguyen, M.T., Van Kersavond, T., and Verhelst, S.H. (2015). Chemical tools for the study of intramembrane proteases. *ACS Chem. Biol.* **10**, 2423–2434.
- Njorge, F.G., Chen, K.X., Shih, N.Y., and Pwiński, J.J. (2008). Challenges in modern drug discovery: a case study of boceprevir, an HCV protease inhibitor for the treatment of hepatitis C virus infection. *Acc. Chem. Res.* **41**, 50–59.
- O'Donnell, R.A., Hackett, F., Howell, S.A., Treeck, M., Struck, N., Krnajska, Z., Withers-Martinez, C., Gilberger, T.W., and Blackman, M.J. (2006). Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. *J. Cell Biol.* **174**, 1023–1033.
- O'Donoghue, A.J., Eroy-Reveles, A.A., Knudsen, G.M., Ingram, J., Zhou, M., Statnekov, J.B., Greninger, A.L., Hostetter, D.R., Qu, G., Maltby, D.A., et al. (2012). Global identification of peptidase specificity by multiplex substrate profiling. *Nat. Methods* **9**, 1095–1100.
- Pierrat, O.A., Strisovsky, K., Christova, Y., Large, J., Ansell, K., Boulou, N., Smiljanic, E., and Freeman, M. (2011). Monocyclic beta-lactams are selective, mechanism-based inhibitors of rhomboid intramembrane proteases. *ACS Chem. Biol.* **6**, 325–335.
- Powers, J.C., Kam, C.M., Narasimhan, L., Oleksyszyn, J., Hernandez, M.A., and Ueda, T. (1989). Mechanism-based isocoumarin inhibitors for serine proteases: use of active site structure and substrate specificity in inhibitor design. *J. Cell. Biochem.* **39**, 33–46.
- Powers, J.C., Asgian, J.L., Ekici, O.D., and James, K.E. (2002). Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* **102**, 4639–4750.
- Purich, D.L. (2010). Kinetic behavior of enzyme inhibitors. In *Enzyme Kinetics: Catalysis & Control* (Elsevier), pp. 485–574.
- Ruiz, N., Falcone, B., Kahne, D., and Silhavy, T.J. (2005). Chemical conditionality. *Cell* **121**, 307–317.
- Schechter, I., and Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162.
- Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675.
- Schrödinger. (2012). The PyMOL Molecular Graphics System, Version 1.5.0.4 (Schrödinger LLC).
- Semple, J.E., Owens, T.D., Nguyen, K., and Levy, O.E. (2000). New synthetic technology for efficient construction of alpha-hydroxy-beta-amino amides via the Passerini reaction. *Org. Lett.* **2**, 2769–2772.
- Serim, S., Haedke, U., and Verhelst, S.H. (2012). Activity-based probes for the study of proteases: recent advances and developments. *ChemMedChem* **7**, 1146–1159.
- Singh, J., Pette, R.C., Baillie, T.A., and Whitty, A. (2011). The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **10**, 307–317.
- Song, W., Liu, W., Zhao, H., Li, S., Guan, X., Ying, J., Zhang, Y., Miao, F., Zhang, M., Ren, X., et al. (2015). Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway. *Nat. Commun.* **6**, 8022.
- Souček, M., and Urban, J. (1995). An efficient method for preparation of optically active N-protected α -amino aldehydes from N-protected α -amino alcohols. *Collect. Czechoslov. Chem. Commun.* **60**, 693–696.
- Stevenson, L.G., Strisovsky, K., Clemmer, K.M., Bhatt, S., Freeman, M., and Rather, P.N. (2007). Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc. Natl. Acad. Sci. USA* **104**, 1003–1008.
- Strisovsky, K. (2013). Structural and mechanistic principles of intramembrane proteolysis - lessons from rhomboids. *FEBS J.* **280**, 1579–1603.
- Strisovsky, K. (2016a). Rhomboid protease inhibitors: emerging tools and future therapeutics. *Semin. Cell Dev. Biol.* **60**, 52–62.
- Strisovsky, K. (2016b). Why cells need intramembrane proteases - a mechanistic perspective. *FEBS J.* **283**, 1837–1845.
- Strisovsky, K., Sharpe, H.J., and Freeman, M. (2009). Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. *Mol. Cell* **36**, 1048–1059.
- Ticha, A., Stanchev, S., Skerle, J., Began, J., Ingr, M., Svehlova, K., Polovinkin, L., Ruzicka, M., Bednarova, L., Hadravova, R., et al. (2017). Sensitive versatile fluorogenic transmembrane peptide substrates for rhomboid intramembrane proteases. *J. Biol. Chem.* **292**, 2703–2713.
- Tulla-Puche, J., Torres, A., Calvo, P., Royo, M., and Albericio, F. (2008). N,N,N',N'-Tetramethylchloroformamidinium hexafluorophosphate (TCFH), a powerful coupling reagent for bioconjugation. *Bioconj. Chem.* **19**, 1968–1971.
- Venkatraman, S., Bogen, S.L., Arasappan, A., Bennett, F., Chen, K., Jao, E., Liu, Y.T., Lovey, R., Hendrata, S., Huang, Y., et al. (2006). Discovery of (1R,5S)-N-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(S)-[[[1,1-dimethylethyl]amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(S)-carboxamide (SCH 503034), a selective, potent, orally bioavailable hepatitis C virus NS3 protease inhibitor: a potential therapeutic agent for the treatment of hepatitis C infection. *J. Med. Chem.* **49**, 6074–6086.
- Vinothkumar, K.R., Strisovsky, K., Andreeva, A., Christova, Y., Verhelst, S., and Freeman, M. (2010). The structural basis for catalysis and substrate specificity of a rhomboid protease. *EMBO J.* **29**, 3797–3809.
- Vinothkumar, K.R., Pierrat, O.A., Large, J.M., and Freeman, M. (2013). Structure of rhomboid protease in complex with beta-lactam inhibitors defines the S2' cavity. *Structure* **21**, 1051–1058.
- Vosyka, O., Vinothkumar, K.R., Wolf, E.V., Brouwer, A.J., Liskamp, R.M., and Verhelst, S.H. (2013). Activity-based probes for rhomboid proteases discovered in a mass spectrometry-based assay. *Proc. Natl. Acad. Sci. USA* **110**, 2472–2477.
- Walker, B., and Lynas, J.F. (2001). Strategies for the inhibition of serine proteases. *Cell. Mol. Life Sci.* **58**, 596–624.
- Wang, Y., and Ha, Y. (2007). Open-cap conformation of intramembrane protease GlpG. *Proc. Natl. Acad. Sci. USA* **104**, 2098–2102.



Wang, Y., Zhang, Y., and Ha, Y. (2006). Crystal structure of a rhomboid family intramembrane protease. *Nature* *444*, 179–180.

Wolf, E.V., Zeissler, A., Vosyka, O., Zeiler, E., Sieber, S., and Verhelst, S.H. (2013). A new class of rhomboid protease inhibitors discovered by activity-based fluorescence polarization. *PLoS One* *8*, e72307.

Wolf, E.V., Zeissler, A., and Verhelst, S.H. (2015). Inhibitor fingerprinting of rhomboid proteases by activity-based protein profiling reveals inhibitor selectivity and rhomboid autoprocessing. *ACS Chem. Biol.* *10*, 2325–2333.

Xue, Y., Chowdhury, S., Liu, X., Akiyama, Y., Ellman, J., and Ha, Y. (2012). Conformational change in rhomboid protease GlpG induced by inhibitor binding to its S' subsites. *Biochemistry* *51*, 3723–3731.

Yin, J., Gallis, C.E., and Chisholm, J.D. (2007). Tandem oxidation/halogenation of aryl allylic alcohols under Moffatt-Swern conditions. *J. Org. Chem.* *72*, 7054–7057.

Zoll, S., Stanchev, S., Began, J., Skerle, J., Lepsik, M., Peclinovska, L., Majer, P., and Strisovsky, K. (2014). Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures. *EMBO J.* *33*, 2408–2421.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-DYKDDDDK	Cell Signaling Technology	Cat#2368
Monoclonal ANTI-FLAG® M2 antibody produced in mouse	Sigma	Cat#F1804
Donkey anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, DyLight 800	Invitrogen	Cat#SA5-10044
Bacterial and Virus Strains		
<i>E. coli</i> NR698	Laboratory of Tom Silhavy (Princeton)	(Ruiz et al., 2005)
<i>E. coli</i> NR698Δ <i>glpG::tet</i>	Laboratory of Matthew Freeman (Oxford)	(Pierrat et al., 2011)
<i>Bacillus subtilis</i> 168	Bacillus Genetic Stock Center	
<i>Bacillus subtilis</i> 168 <i>ycdA::neo</i>	This work	BS2
<i>Bacillus subtilis</i> 168 <i>ycdA::neo, yqgP::tet</i>	This work	BS4
<i>Bacillus subtilis</i> 168 <i>ycdA::neo, xdkE::MBP-LacYTM2-Trx(erm, lin)</i>	This work	BS87
<i>Bacillus subtilis</i> 168 <i>ycdA::neo, yqgP::tet, xdkE::MBP-LacYTM2-Trx(erm, lin)</i>	This work	BS88
<i>E. coli</i> C41(DE3)	Lucigen	Cat#60452-1
Chemicals, Peptides, and Recombinant Proteins		
h-KRHDIN(E-edans)ISKSDTG(K-dabcyI)	(Ticha et al., 2017)	KSp35
IFAAISLFSLLFQPLFGLSKK-nh ₂	This work	KSp93
h-KRHRRI(E-edans)RVRHADTG(K-dabcyI)	(Ticha et al., 2017)	KSp64
IFAAISLFSLLFQPLFGLSKKR-nh ₂	This work	KSp96
h-KRHRIRVR(E-edans)ADTG(K-dabcyI)	This work	
IFAAISLFSLLFQPLFGLSKKR-nh ₂		
TAMRA-XP	Thermo Fisher Scientific	Cat#88318
Revert Total Protein Stain Kit	LI-COR, Inc.	Cat#926-11010
3,4-dichloroisocoumarin (DCI)	Sigma	D7910
AcRVRHAcmk	This paper	
AcVRHAcmk	This paper	
AcIATAcmk	Zoli et al., 2014	
AcVATAcmk	This paper	
AcIAHAcmk	This paper	
AcIRTAcmk	This paper	
AcRVRHA-trifluoromethylketone	This paper	
AcRVRHA-boronate	This paper	
AcRVRHA-acylsulfonamide	This paper	
AcRVRHA-thiazolylketone	This paper	
AcRVRHA-CONH ₂	This paper	
Ac-RVRHA-CONH-isopropyl	This paper	
Ac-RVRHA-CONH-cyclohexyl	This paper	
Ac-RVRHA-CONH-benzyl	This paper	
Ac-RVRHA-CONH-methylbenzyl	This paper	
Ac-RVRHA-CONH-ethoxyacetyl	This paper	

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ac-RVRHA-CONH-neopentyl	This paper	
Ac-RVRHA-CONH-pentyl	This paper	
Ac-RVRHA-CONH-phenylethyl	This paper	
Ac-RVRHA-CONH-(2,4-dimethyl)benzyl	This paper	
Ac-RVRHA-CONH-phenylbutyl	This paper	
Ac-RHA-CONH-phenylethyl	This paper	
Ac-RHA-CONH-phenylethyl	This paper	
Ac-A-CONH-phenylethyl	This paper	
Deposited Data		
Crystal structure of GlpG bound to AcRVRHAcmk	This paper	PDB: 5MT8
Crystal structure of GlpG bound to AcVRHAcmk	This paper	PDB: 5MT7
Crystal structure of GlpG bound to Ac-RVRHA-CONH-phenylethyl	This paper	PDB: 5MT6
Crystal structure of GlpG bound to Ac-RVRHA-CONH-(2,4-dimethyl)benzyl	This paper	PDB: 5MTF
Recombinant DNA		
pMALp2E_MBP-LacYTM2-Trx	(Zoll et al., 2014)	pKS506
pD881-SR	DNA2.0 Inc.	
pD881-SR_MBP-LacYTM2-Trx	This work	pPR61
pET25b+M_GlpG	(Lemberg et al., 2005)	
pGP886_pxyI-AmyE _{SP} -MBP-FLAG-LacYTM2-Trx-HA	This work	pPR200
Software and Algorithms		
XDS	(Kabsch, 2010)	http://xds.mpimf-heidelberg.mpg.de/html_doc/downloading.html
AIMLESS	(Evans, 2011)	http://www.ccp4.ac.uk/html/aimless.html
COOT 0.8.6	(Emsley and Cowtan, 2004)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Jligand	(Lebedev et al., 2012)	http://www.ysbl.york.ac.uk/mxstat/JLigand/
Refmac 5.8.0158	(Murshudov et al., 2011)	
Ligplot+	(Laskowski and Swindells, 2011)	http://www.ebi.ac.uk/thornton-srv/software/LigPlus/download.html
Phaser	(McCoy, 2007)	http://www.ccp4.ac.uk/html/phaser.html
Pymol 1.8.4.0	Schrodinger, LLC	https://www.pymol.org/
Turbomole 7	(Ahlich et al., 1989)	http://www.turbomole.com/
ImageJ 1.49	(Schneider et al., 2012)	https://imagej.net/Welcome
Image Studio Lite Version 5.2	LI-COR, Inc.	https://www.licor.com/bio/products/software/image_studio_lite/download.html
GraphPad Prism 7.02	GraphPad Software, Inc.	https://www.graphpad.com/scientific-software/prism/
GraFit 7	Erithacus Software Ltd.	http://www.erithacus.com/grafit/

CONTACT FOR REAGENTS AND RESOURCE SHARING

Further information and requests for reagents may be directed to, and will be fulfilled by the corresponding author Kvido Strisovsky (kvido.strisovsky@uochb.cas.cz).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Escherichia coli K12 strain NR698 (Ruiz et al., 2005), which has the MC4100 background with the *imp4213* allele carried from BE100 (Eggert et al., 2001) (an in-frame deletion of amino acids 330-352 of LptD) is a gift of Dr. Tom Silhavy (Princeton University). A GlpG-free variant was created by deleting *glpG* using a tetracyclin marker (Pierrat et al., 2011).

e2 Cell Chemical Biology 24, 1523–1536.e1–e4, December 21, 2017

To generate a rhomboid activity free *Bacillus subtilis*, the *ydcA::neo* mutant (BS2, this work) of the wild type *B. subtilis* 168 strain (Bacillus Genetic Stock Center, USA) was modified by deleting the entire the *yggP* gene and replacing it with a tetracyclin resistance gene using homologous recombination, yielding strain BS4 (*ydcA::neo*, *yggP::tet*). Both modifications were verified by genomic PCR of the disrupted locus and Sanger sequencing of the amplified region.

METHOD DETAILS

Constructs and Cloning

To generate a model rhomboid substrate for *in vivo* activity assays in *E. coli*, the MBP-LacYTM2-Trx-Stag-Histag construct was PCR-amplified from pKS506 (Ticha et al., 2017) and cloned into the SapI linearized plasmid pD881-SR (DNA2.0 Inc., Newark, USA) using isothermal assembly (Gibson, 2011), yielding construct pPR61. For expression in *B. subtilis*, the substrate was modified by replacing the MBP signal peptide by the signal peptide from *B. subtilis* AmyE, and the AmyE_{sp}-MBP_{mat}-FLAG-LacYTM2-Trx-HA construct was cloned into the XbaI, Sall digested plasmid pGP886 (Gerwig et al., 2014) (gift of Dr. Libor Krasny, Prague, CR) using isothermal assembly (Gibson, 2011) to yield construct pPR200. This construct was linearized by Scal and integrated into the *xkdE* locus of BS2 and BS4 using erythromycin-lincomycin selection yielding strains BS87 (BS2 *xkdE::Pxyl-LacYTM2(erm)*) and BS88 (BS4 *xkdE::Pxyl-LacYTM2(erm)*).

Protein Expression and Purification

The *E. coli* GlpG for crystallisation was expressed in *E. coli* C41(DE3) (Miroux and Walker, 1996) in PASM 5052 medium as described (Lee et al., 2014). Membrane isolation, purification by metal affinity chromatography, cleavage by chymotrypsin to produce GlpG transmembrane core domain and gel filtration chromatography were carried out as described previously (Vinothkumar et al., 2010; Zoll et al., 2014). The *E. coli* GlpG for inhibition studies was expressed in *E. coli* C41(DE3) (Miroux and Walker, 1996) in LB medium, and solubilised and purified in 0.05% (w/v) DDM as described (Ticha et al., 2017). Other rhomboid proteases were expressed and purified as reported previously (Wolf et al., 2015).

Chemical Synthesis

All reagents were acquired from commercial sources and used without purification. Protected amino acids and amino acid derivatives were purchased from Iris Biotech (Marktredwitz, Germany), Sigma-Aldrich (St. Louis, MO, U.S.A), Thermo Fischer Scientific (Waltham, Massachusetts, U.S.A) and Fluorochem (Hadfield, Derbyshire, UK). Further details on chemical syntheses as well as compound characterisation data by mass spectrometry and NMR are available as Supplemental Information (Methods S1).

Protein Crystallography

Crystals of truncated wild type GlpG apoenzyme were obtained by mixing a solution of 2 - 3 M ammonium chloride or sodium chloride, 0.1 M Bis-Tris, pH 7.0 with protein (4-6 mg/mL) at ratio of 1:1 in hanging drops at 22°C (Vinothkumar et al., 2010; Wang and Ha, 2007). Inhibitors were diluted from 10 mM stock solutions in anhydrous DMSO into buffer resembling the mother liquor to yield final 1 mM inhibitor and 10% DMSO just before soaking. For the chloromethylketone inhibitors, the crystals were incubated with inhibitors at 0.3-0.5 mM concentrations for 24 h. The ketoamide inhibitors were incubated at final concentrations of 0.3-0.5 mM for 30-120 min. All crystals were cryo-protected by adding 25% (v/v) glycerol to the mother liquor and flash frozen in liquid nitrogen.

Data sets of the CMKs and **9** were collected at the I02 beam line at the Diamond Light Source (Harwell) and the data set of **10** was collected at BESSY (Berlin, Germany). Diffraction data were indexed, integrated and scaled with XDS (Kabsch, 2010) and AIMLESS (Evans, 2011). For the structures with inhibitor bound, the coordinates of GlpG (PDB 2XOV) with residues 245-249 (of Loop 5) omitted were used as an input model for Phaser (McCoy, 2007). Restrained refinement was carried out with Refmac (Murshudov et al., 2011) followed by manual model building in COOT (Emsley and Cowtan, 2004). In the final step, TLS was used using the enzyme and the inhibitor peptide as one group (Murshudov et al., 2011). The model, library and link files of the inhibitors were generated with Jligand (Lebedev et al., 2012). In the structures of Ac-(R)VRHA-cmk, H150 was modelled to hydrogen bond to the chloromethylketone oxygen. An additional density was observed close to M149 and H150 raising the possibility that the H150 residue could be also in an alternative conformation, but modeling the alternative conformation or both conformations of H150 was not conclusive in explaining the density. Other similar datasets of CMKs obtained by soaking show that this density might perhaps represent a bound ion, but due to ambiguity we have left the density unmodelled.

In order to find the best possible fit of the molecules of **9** and **10** to the experimental electron densities, quantum mechanical calculations were performed. The model systems comprised the whole inhibitors in their tetrahedral intermediate form with methoxy group representing the S201 side-chain. These models were made in several variants: i) *cis/trans* isomers of the ketoamide proximal/distal carbonyls, ii) *cis/trans* isomers of the distal carbonyl/NH, and iii) different rotameric forms of the His side chain in the P2 position of the inhibitors. All these variants were optimized in Turbomole ver. 7 program (Ahlich et al., 1989) using DFT-D3 method (Grimme, 2006) at B-LYP/DZVP level (Fanfrlik et al., 2016; Jensen, 2006) and COSMO implicit solvent model (Klamt and Schuurmann, 1993). Their intrinsic stabilities were assessed by comparing the final energies, and the conformer with the lowest energy was built into the electron density and chosen as a model for crystallographic refinement.

Noncovalent interactions between the ligands and protein were detected using Ligplot+ (Laskowski and Swindells, 2011) and hydrogen bonds were defined by canonical geometrical criteria (Laskowski and Swindells, 2011; McDonald and Thornton, 1994).

Structural alignments and all structure figures were made with Pymol (Schrodinger, 2012). The coordinates of the structures presented in this manuscript have been deposited in the PDB under the following IDs: 5MT7 (Ac-VRHA-cmk), 5MT8 (Ac-RVRHA-cmk), 5MT6 (compound **9**) and 5MTF (compound **10**). Data collection and refinement statistics are listed in Table S1.

Rhomboid Activity and Inhibition Assays

The activity of GlpG *in vitro* was determined as reported (Ticha et al., 2017). Concentrations of stock solutions of peptide substrates and inhibitors were determined by quantitative amino acid analysis. The IC_{50} and reversibility measurements were performed in 20 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% (w/v) DDM, 12% (v/v) DMSO, and other kinetic measurements in 50 mM potassium phosphate, pH 7.4, 150 mM NaCl, 0.05% (w/v) DDM, 10% (v/v) DMSO, 0.05% (w/v) PEG8000, and 20% (v/v) glycerol unless noted otherwise. The reaction mixture typically consisted of 10 μ M fluorogenic peptide substrate and the measurements were performed without enzyme-inhibitor pre-incubation unless noted otherwise. Note that the fluorogenic substrates used in Figure 4 had nearly identical amino acid sequences but for the point of attachment of the fluorophore or the identity of the fluorophore and quencher (see Key Resources Table).

For measuring the inhibition of GlpG *in vivo*, the *E. coli* strain NR698 with genetically permeabilised outer membrane (Ruiz et al., 2005) and its *glpG* knock-out derivative KS69 (*glpG::tet*) were used as described (Pierrat et al., 2011) with the following modifications. The chimeric substrate encoding LacY transmembrane domain 2 inserted between maltose binding protein and thioredoxin (Strisovsky et al., 2009) was expressed under control of rhamnose promoter (construct pPR61). To evaluate the *in vivo* inhibition by ketoamides, the NR698 cells were inoculated to the density of $OD_{600} = 0.05$ and grown to $OD_{600} = 0.6$ at 37°C. The cells were then incubated with increasing concentrations of inhibitor for 15 min at room temperature, and expression of the chimeric substrate was induced by adding 1 mM L-rhamnose. Cells were grown for further 4 h at 25°C, after which steady-state level of substrate cleavage was evaluated by western blotting with near-infrared fluorescence detection as described (Ticha et al., 2017).

For measuring the inhibition of YqgP *in vivo*, the *B. subtilis* strains BS87 and its *yqgP* knock-out derivative BS88, generated in this work (see Constructs and Cloning section), were used as follows. The chimeric LacYTM2 substrate AmyE_{sp}-MBP-FLAG-LacYTM2-Trx-HA (this work) was expressed under control of xylose promoter from the *xdkE* genomic locus. Fresh LB medium, supplemented with appropriate antibiotic, was inoculated with a few colonies of the *B. subtilis* strain grown overnight on selective LB agar plate and pre-culture was grown for 2 h at 37°C to $OD_{600} = 1$. Pre-culture was then diluted with fresh LB medium to the density of $OD_{600} = 0.05$. At this point, the expression of LacYTM2 was induced by adding 1% (w/v) D-(-)-xylose (Sigma), rhomboid inhibitors were added at a range of concentrations, and the cultures were further incubated for 2.5 h at 37°C (reaching $OD_{600} \sim 1$). Steady-state conversion of the substrate was evaluated by western blotting with near-infrared fluorescence detection as described (Ticha et al., 2017), subtracting the intensity of non-specific bands, closely co-migrating with the specific rhomboid-formed N-terminal cleavage product of the substrate.

Inhibitor Selectivity Profiling

For inhibitor selectivity profiling against rhomboid proteases (Wolf et al., 2015), 400 ng of a purified protein preparation of *E. coli* GlpG was diluted in 30 μ L of reaction buffer (20 mM HEPES, pH 7.4, with 0.05% (w/v) DDM). For other rhomboids, amounts were taken that gave similar labeling intensity during profiling. Rhomboids were incubated for 30 min at room temperature with the indicated concentration of compound, 100 μ M DCI as positive control, or an equal volume of DMSO as negative control. Next, TAMRA-FP serine hydrolase probe (Thermo Fisher #88318) was added to a final concentration of 1 μ M and incubated for 2 h at 37°C in the dark. The reaction was stopped by addition of 4 \times Laemmli buffer and the reaction mixture was resolved on 15% SDS-PAGE. Gels were scanned on a Typhoon Trio+ and analyzed using ImageJ. The intensity of each rhomboid protease band calculated by ImageJ was normalized against its corresponding DMSO-treated counterpart (100% activity) to indicate the residual activity left after inhibition. The remaining activity was used to calculate the percentage of inhibition depicted in the heatmap. Selectivity profiles against human serine hydrolases were determined by EnPlex as described previously (Bachovchin et al., 2014).

QUANTIFICATION AND STATISTICAL ANALYSIS

Enzyme kinetics and inhibition data were analysed in GraphPad Prism v7.02 using in-built algorithms. Means and standard deviations have been derived from the best fit of the data, or based on three independent measurements, as specified, unless noted otherwise. Quantitative western blots were evaluated using near infrared detection with the IRDye 800CW secondary antibody on a LiCor Odyssey CLx infrared scanner with normalisation to total protein using the Revert total protein stain (LiCor).

DATA AND SOFTWARE AVAILABILITY

All crystallographic coordinates of the protein structures presented in this manuscript have been deposited in and will be freely available from the Protein Data Bank (www.rcsb.org) under the following identifiers: 5MT7, 5MT8, 5MT6 and 5MTF.

Discovery and Biological Evaluation of Potent and Selective *N*-Methylene Saccharin-Derived Inhibitors for Rhomboid Intramembrane Proteases

Parul Goel,^{†,∇} Thorsten Jumpertz,[†] David C. Mikles,[‡] Anežka Tichá,[‡] Minh T. N. Nguyen,[§] Steven Verhelst,^{§,||} Martin Hubalek,[‡] Darren C. Johnson,[⊥] Daniel A. Bachovchin,[⊥] Isabella Ogorek,[†] Claus U. Pietrzik,[#] Kvido Strisovsky,[‡] Boris Schmidt,^{*,∇} and Sascha Weggen^{*,†,||}

[†]Department of Neuropathology, Heinrich-Heine University Duesseldorf, Moorenstrasse 5, 40225 Duesseldorf, Germany

[‡]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 10 Praha 6, Czech Republic

[§]Chemical Proteomics Group, Leibnitz Institute for Analytical Sciences (ISAS) e.V., Otto-Hahn-Strasse 6b, 44227 Dortmund, Germany

^{||}Laboratory of Chemical Biology, Department of Cellular and Molecular Medicine, University of Leuven, Herestraat 49, Box 802, 3000 Leuven, Belgium

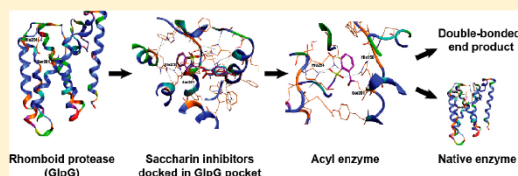
[⊥]Chemical Biology Program, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, Box 428, New York, New York 10065, United States

[#]Institute for Pathobiochemistry, University Medical Center of the Johannes Gutenberg University Mainz, Duesbergweg 6, 55128 Mainz, Germany

[∇]Clemens Schoepf Institute for Organic Chemistry and Biochemistry, Technical University of Darmstadt, Alarich-Weiss-Strasse 4-8, 64287 Darmstadt, Germany

Supporting Information

ABSTRACT: Rhomboids are intramembrane serine proteases and belong to the group of structurally and biochemically most comprehensively characterized membrane proteins. They are highly conserved and ubiquitously distributed in all kingdoms of life and function in a wide range of biological processes, including epidermal growth factor signaling, mitochondrial dynamics, and apoptosis. Importantly, rhomboids have been associated with multiple diseases, including Parkinson's disease, type 2 diabetes, and malaria. However, despite a thorough understanding of many structural and functional aspects of rhomboids, potent and selective inhibitors of these intramembrane proteases are still not available. In this study, we describe the computer-based rational design, chemical synthesis, and biological evaluation of novel *N*-methylene saccharin-based rhomboid protease inhibitors. Saccharin inhibitors displayed inhibitory potency in the submicromolar range, effectiveness against rhomboids both *in vitro* and in live *Escherichia coli* cells, and substantially improved selectivity against human serine hydrolases compared to those of previously known rhomboid inhibitors. Consequently, *N*-methylene saccharins are promising new templates for the development of rhomboid inhibitors, providing novel tools for probing rhomboid functions in physiology and disease.



At least 600 protease-encoding genes have been identified in the human genome. Proteases regulate diverse biological processes, including blood coagulation, antigen presentation, cell motility, apoptosis, and neuronal development.^{1,2} Dysregulation of protease activity can lead to severe diseases like arthritis, cancer, and cardiovascular diseases, which makes proteases attractive targets for the development of new pharmaceuticals.^{3–9} Soluble, globular proteases have been studied intensively for decades and now represent the best characterized and understood enzymes.² In contrast, intramembrane proteases were identified only recently. The founding member of the rhomboid superfamily of intramembrane serine proteases was discovered in *Drosophila* in

2001.¹⁰ Rhomboid intramembrane proteases can be found in all forms of life ranging from archaea and bacteria to fungi, plants, and animals.^{11,12} In contrast to classical serine proteases that use a catalytic triad (Ser/His/Asp),^{13,14} rhomboids use a catalytic dyad (Ser/His) within their active site^{15,16} and their substrates are transmembrane proteins.

Because of their broad phylogenetic distribution, their implication in various diseases, and the availability of high-

Received: October 20, 2017

Revised: November 18, 2017

Published: November 29, 2017



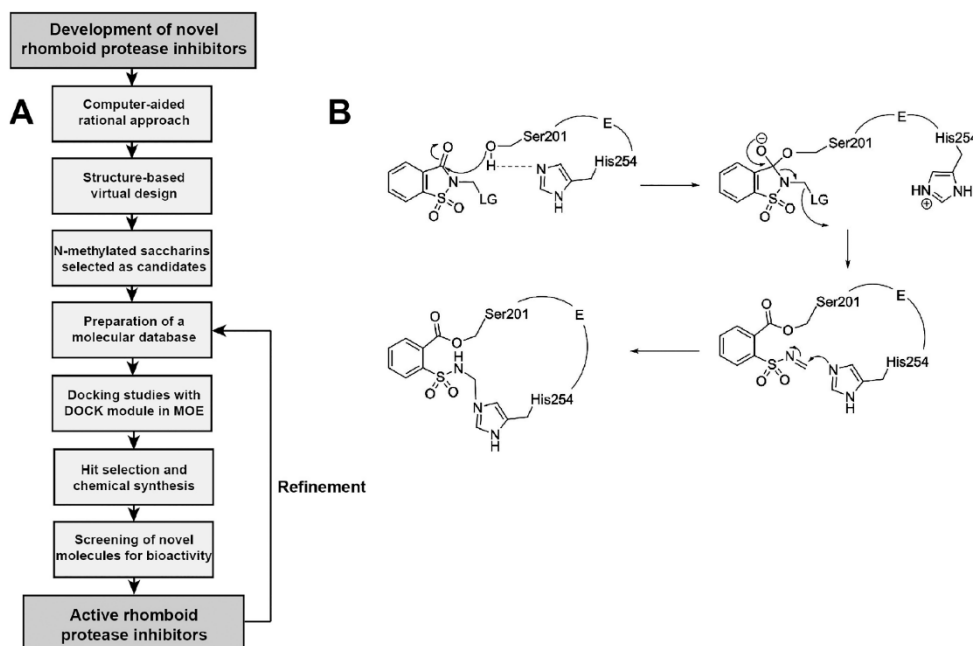


Figure 1. (A) Workflow for the development of novel rhomboid inhibitors. (B) Proposed reaction mechanism of a saccharin-based inhibitor with the catalytic residues in the active site of the *E. coli* rhomboid protease GlpG. The mechanism involves the enzyme-induced formation of a Michael adduct and results in an end product that is covalently cross-linked to the enzyme.

resolution crystal structures, rhomboid proteases are potential targets for drug development, yet potent and selective inhibitors are not available.^{16–20} Isocoumarins were the first inhibitors observed to be effective against rhomboids, including dichloroisocoumarin (DCI),^{10,21} JLK-6 ($IC_{50} \sim 6 \mu M$), and IC16 ($IC_{50} \sim 0.7 \mu M$).^{22,23} However, isocoumarins are in general highly reactive and lack selectivity.^{24,25} In fact, DCI was later established as a pan-rhomboid inhibitor.²⁶ Rational synthesis approaches and the screening of small molecule libraries have identified other classes of rhomboid inhibitors. N-Sulfonylated β -lactams inhibited bacterial rhomboids but showed limited potency *in vivo* ($EC_{50} \sim 5–10 \mu M$) and poor selectivity over soluble serine proteases such as chymotrypsin.²⁷ Fluorophosphonates ($IC_{50} \sim 50 \mu M$)²⁸ and β -lactone-based inhibitors ($IC_{50} \sim 26–44 \mu M$)²⁹ displayed poor potency toward rhomboids. Recently, peptidyl chloromethylketones³⁰ and peptidyl aldehydes³¹ were reported as peptidomimetic inhibitors of rhomboids, but these compounds also lacked sufficient potency and selectivity.

To fill the apparent gap between a comprehensive biochemical understanding of rhomboid enzymes and the absence of selective inhibitors, we applied a computer-aided rational approach with a starting template based on N-methylene-substituted saccharins with a leaving group (LG). This scaffold has several advantages as (i) it is a well-characterized sweetener with a safe toxicological profile,³² (ii) it can act as a suicide inhibitor, resulting in a double-bonded enzyme–inhibitor complex,^{24,33,34} and (iii) it has substantial potential for derivatization. A computational algorithm was used to virtually screen suitable candidate molecules prior to chemical synthesis and compound testing during *in vitro* enzyme activity assays (Figure 1A),³⁵ which yielded a subset of N-methylene-substituted saccharins as potential rhomboid

inhibitors. Mass spectrometry and mutagenesis were used to establish the reaction mechanism of rhomboid inhibition by saccharins, and the *in vivo* potency was determined in *Escherichia coli*. The selectivity of the compounds for a variety of rhomboids and off-target effects against human serine hydrolases were tested using activity-based probe approaches.^{36,37} This effort provided submicromolar, partially reversible, and selective rhomboid protease inhibitors based on a novel and modifiable scaffold.

MATERIALS AND METHODS

Computational Chemistry. Various methods of Molecular Operating Environment (MOE) of the Chemical Computing group³⁸ were used for the docking of saccharins into the rhomboid protease pocket. See the [Supporting Information](#) for details.

Chemical Synthesis. For details of the chemical synthesis of the saccharin inhibitors and the chemical data, see the [Supporting Information](#).

Protein Expression and Purification. Bacterial expression constructs for N-terminally His-tagged *E. coli* GlpG,¹⁶ N-terminally GST-tagged *Aquifex aeolicus* AaROM,²¹ and C-terminally His-tagged Gurken chimeric substrate (MBP-GurkenTMD-Trx-His6)³⁹ were kindly provided by Y. Ha (Yale School of Medicine, New Haven, CT), S. Urban (Johns Hopkins University, Baltimore, MD), and M. Freeman (University of Oxford, Oxford, U.K.). Both rhomboids and the substrate Gurken were expressed in *E. coli* C43(DE3) cells. After induction with isopropyl β -D-1-thiogalactopyranoside at an OD_{600} of 0.8, GlpG and AaROM were expressed overnight at 18 °C and Gurken was expressed for 3 h at 37 °C. The cells were broken by three passages through a nitrogen cavitation

bomb; bacterial membranes were isolated by differential centrifugation, and proteins were solubilized from the membrane preparations with 1.5% dodecyl β -maltoside (DDM) (Glycon Biochemicals). GlpG and Gurken were purified via immobilized metal ion affinity chromatography (IMAC) using a TalonCrude 1 mL HiTrap column and an Akta prime plus chromatography system (GE Healthcare). The GST-tagged AaRom was incubated with glutathione sepharose resin overnight at 4 °C, and the rhomboid was eluted with PreScission (Amersham Pharmacia Biotech) protease cleavage according to the manufacturer's instructions. All recombinant proteins were further concentrated using 10 or 30 kDa molecular weight cutoff Amicon Ultra Centrifugal filters (Millipore). Protein purity was estimated by Coomassie Brilliant Blue (Merck)-stained sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE).

GlpG and AaROM *in Vitro* Activity Assays. Activity assays were performed in a volume of 20 μ L of reaction buffer [50 mM HEPES-NaOH (pH 7.5), 5 mM EDTA, 0.4 M NaCl, 10% (v/v) glycerol, and 0.05% DDM] containing either GlpG (0.35 μ M) or AaROM (0.094 μ M) and Gurken substrate (1.8 μ M). Saccharin inhibitors were preincubated with the rhomboids in reaction buffer for 30 min at 37 °C while being gently shaken. The substrate was added, and the reaction was continued for an additional 90 min at 37 °C. Subsequently, the reaction mixture was separated by SDS–PAGE and stained with Coomassie Brilliant Blue. The N-terminal Gurken cleavage fragment was quantified using ImageJ, normalized to the dimethyl sulfoxide (DMSO) vehicle control condition, and the values were plotted in GraphPad Prism. A nonlinear regression curve fit of the log(inhibitor concentration) versus percent N-terminal Gurken product was used to determine IC₅₀ values.

GlpG Activity in Live *E. coli* Cells. The inhibitory potency of saccharin inhibitors against GlpG in live *E. coli* was determined as described previously.⁴⁰

Cell-Based RHBDL2 Activity Assay. Human HEK293T cells were maintained in Dulbecco's modified Eagle's medium with 10% fetal bovine serum, 1 mM sodium pyruvate, and 100 units/mL penicillin/streptomycin (Thermo Fisher Scientific). Expression constructs for HA-tagged mouse rhomboid RHBDL2 and FLAG-tagged Gurken substrate (kind gifts of M. Freeman²⁷) were transiently co-transfected into HEK293T cells using GeneJuice transfection reagent (Merck). Two hours after the transfection, cells were treated with the saccharin inhibitors or DMSO as a control. After 48 h, cells lysates were prepared with standard 1% NP40 lysis buffer and N-terminal Gurken cleavage fragments were analyzed by Western blotting with the anti-FLAG BioM2 monoclonal antibody (1:1000, Merck).

Mass Spectrometry. The rhomboid GlpG at 0.2 mg/mL in 25 mM HEPES, 150 mM NaCl, and 0.2% (w/v) DDM (pH 8.0) was preincubated with 50 μ M inhibitor or DMSO solvent control at 23 °C overnight. The sample was then diluted with 0.1% formic acid to a concentration of 0.1 mg/mL. Five microliters of the sample was injected onto a desalting column (MassPREP desalting, Waters) and desalted with a fast gradient (4 min) of acetonitrile in water with 0.1% formic acid. The separation was performed by a liquid chromatography system (I-class, Waters), which was coupled online to a mass spectrometer (Synapt G2, Waters) to acquire the intact mass of the protein by electrospray ionization. The raw spectra were subtracted and deconvoluted (MaxEnt 1, Waters) to produce the final spectra shown in panels B and C of Figure 4.

GlpG Enzyme Recovery Assay. Rhomboid activity recovery assays were performed in black 96-well plates. Saccharin inhibitor Bsc5195 (5 μ M), β -lactam L29 (1 μ M), or isocoumarin JLK-6 (10 μ M) was preincubated with 0.4 μ M GlpG in reaction buffer [20 mM HEPES-NaOH (pH 7.4), 150 mM NaCl, and 0.05% (w/v) DDM] for 1 h. Subsequently, the reaction mixture was rapidly diluted 100-fold into reaction buffer containing 10 μ M fluorogenic substrate KSp76.⁴⁰ This quenched fluorescent peptide is cleaved by the rhomboid, leading to activation of a red fluorophore. Substrate cleavage and enzyme activity recovery were monitored continuously by measuring the fluorescence over 100 min with excitation at 553 nm and emission at 583 nm using a microplate reader (Tecan Infinite M1000).

α -Chymotrypsin *in Vitro* Activity Assay. To determine the inhibitory potency of saccharin inhibitors against α -chymotrypsin, the compounds were preincubated in a defined range of concentrations (from 0 μ M to 1 mM) with 4 μ M bovine chymotrypsin (PanReac AppliChem) in reaction buffer [50 mM HEPES-NaOH (pH 7.5), 5 mM EDTA, 0.4 M NaCl, 10% (v/v) glycerol, and 10% DMSO] for 30 min at 25 °C. Then, 2.5 μ L of the preincubated reaction mixture was added to 100 μ L of 0.12 mM *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-pNA, Merck) substrate in reaction buffer. Enzymatic cleavage of the 4-nitroanilide substrate resulted in 4-nitroaniline, which has a yellow color under alkaline conditions.⁴¹ The increase in absorbance at 410 nm over 5 min at 25 °C was measured with a microtiter plate reader (Beckman Coulter). IC₅₀ values were calculated by fitting a nonlinear regression curve to a plot of log(compound concentration) versus percent α -chymotrypsin activity using GraphPad Prism.

Rhomboid Activity Panel. Selectivity profiling of rhomboid inhibitors using activity-based probes was performed as described previously.²⁶

Soluble Serine Hydrolase Panel. Selectivity profiles of the saccharin-based inhibitors against human soluble serine hydrolases were determined as described previously.³⁷

RESULTS AND DISCUSSION

Chemical Rationale. To synthesize a panel of enzyme-activated heterocyclic compounds with inhibitory activity against rhomboid proteases, we considered the saccharin (1,2-benzisothiazol-3-one 1,1-dioxide) template. The basic structure of the saccharin heterocyclic ring system consists of an aromatic ring and a sulfimide ring with lactam and sulfonamide moieties. Saccharin-based inhibitors were first described in 1980 as irreversible, acylating compounds of human leukocyte elastase and chymotrypsin.⁴² The protease's active-site serine initiates attack at the carbonyl group of the saccharin, which leads to formation of an acyl–enzyme complex and enzyme-induced ring opening (Figure 1B). An important advantage of saccharin-based inhibitors is their ability to differentiate between different serine proteases, which allows a rather broad usage. A possible explanation for this observation is the variable stability of the acyl–enzyme complex that is formed when a specific saccharin inhibitor reacts with the active site of a particular serine protease.^{24,34}

We investigated saccharins with a leaving group (LG) attached to the nitrogen atom of the heterocyclic ring system (so-called next-generation *N*-methylene-substituted saccharin inhibitors).^{43–46} Attack of the active-site serine is followed by a prototropic shift, and the release of the leaving group to form a

reactive intermediate that cross-links to the active-site nucleophilic histidine residue. Overall, this leads to a cross-linked structure and an irreversible enzyme–inhibitor complex (Figure 1B).²⁴

Modifications to the saccharin scaffold can be introduced as different N-functional LGs or in the heterocyclic ring of the saccharin core. Because derivatization of the N-functional group can be achieved much easier than at the core, we started to identify the most favorable LGs at this position. A consequence of this strategy is that the reaction between the rhomboid protease and the saccharin-based inhibitors with different LGs could result in the same double-bonded enzyme/inhibitor end product (provided that the reaction mechanism with the rhomboid is the same as with the classical, triad-containing serine proteases). We anticipated that binding of a putative inhibitor to the active site of rhomboid proteases would involve the formation of covalent bonds as well as hydrogen bonds and hydrophobic interactions. The final reaction product should feature two covalent bonds, one toward the catalytic Ser and one toward the catalytic His, and would result in the release of the N-functional LG (Figures 1B and 2C).

Docking Studies. Because our focus was on the effect of different LGs attached to the saccharin core structure, we considered in the docking studies the initial interactions between the inhibitor and the active site of the rhomboid protease rather than the final reaction product. Therefore, a molecular database of 50 saccharin-based serine protease targeting ligands with favorable LGs (Supplementary Table 1) was docked into the active-site cleft of the crystal structure of the *E. coli* rhomboid protease GlpG.

We used a co-crystal structure of GlpG and the inhibitor CAPF (Protein Data Bank entry 3UBB)⁴⁷ that is covalently bound to the active-site Ser201 as a template for the docking studies. The receptor (GlpG, 3UBB) was prepared for docking, and the modeling experiments were performed using the DOCK module with the Amber12:EHT force field in the Molecular Operating Environment (MOE).³⁸ The resulting poses for all 50 saccharin inhibitors were scored by London dG and Affinity dG followed by energy minimization within the enzyme active-site cleft. The output database contained the 10 best poses for each inhibitor along with their scoring function (S).

Subsequently, the molecules were ranked on the basis of the calculated ligand efficiencies (cLEs) followed by their scoring function and ligand interactions (Supplementary Table 2). As a result, saccharins with an aryl carboxylic acid LG were identified as the best-suited ligand and termed “hit molecules” (Supplementary Figure 2). Another observation of the docking studies was that the saccharin derivatives interacted with the rhomboid active site similar to CAPF. The sulfoxide moiety of the saccharin scaffold formed hydrogen bonds to His150 and Gly199, which is depicted in the ligand interaction map of GlpG and the saccharin scaffold with LG38 (Figure 2B; see Supplementary Figure 3 for a ligand interaction map of GlpG with CAPF). Another striking similarity between CAPF and the saccharin derivatives was the occupancy of the S region by the saccharin core ring and the LG extending toward the S' region of GlpG. This observation together with significant π – π interactions between the aryl ring of LG38 and Phe245 of the rhomboid GlpG led us to propose that the LG might play a major role in the recognition process of the inhibitor and enzyme (Figure 2A).

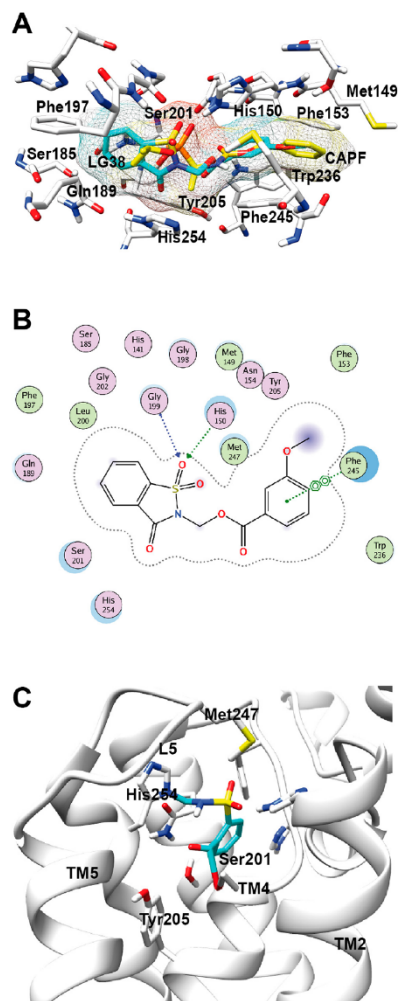
























Figure 2. Docking of saccharin-based inhibitors into the active site of the rhomboid protease GlpG. (A) The model shows the initial interactions of a known rhomboid inhibitor (CAPF, yellow) and a saccharin-based inhibitor (with LG 38, turquoise) with amino acid side chains in the active site of GlpG. (B) Ligand interaction map of GlpG and the saccharin with LG38. The map shows the formation of two H-bonds of the sulfoxide moiety of the saccharin with His150 and Gly199. In addition, a π – π interaction between the aryl ring of the saccharin and Phe245 is visible. (C) Molecular model of N-methylated saccharin within the GlpG active site after the enzymatic reaction between the inhibitor (blue) and the rhomboid protease (gray). The model shows the end product based on the proposed reaction mechanism (see Figure 1B). Two covalent bonds have been formed between the saccharin and Ser201 and His254.

Synthesis and Test of the First Compound Batch.

Calculations from the docking experiments revealed that saccharins with aryl carboxylic acid LGs showed the highest cLE along with saccharins that had aryl sulfides and sulfones as LGs (Supplementary Figure 2 and Supplementary Table 2). Subsequently, we synthesized a preliminary batch of 10 N-methylene saccharin derivatives with different LGs (Supplementary Table 3; for details of the synthesis, see Supplementary

Table 1. Overview of Selected Saccharin-Based Inhibitors with Different Leaving Groups and IC₅₀ Values of <10 μM in the *E. coli* GlpG *in Vitro* Activity Assay⁴⁴

ID	Structure	IC ₅₀ [μM] GlpG	IC ₅₀ [μM] α-chymotrypsin	EC ₅₀ [μM] <i>E. coli</i> <i>in vivo</i>	ID	Structure	IC ₅₀ [μM] GlpG	IC ₅₀ [μM] α-chymotrypsin	EC ₅₀ [μM] <i>E. coli</i> <i>in vivo</i>
DCI		19	3.5	n.d.	BSc5195		0.20	10	0.28
L16		0.9	n.d.	n.d.	BSc5196		0.56	208	1.2
BSc5156		0.61	33	0.99	BSc5197		0.64	68	n.d.
BSc5187		0.54	50	n.d.	BSc5198		0.80	56	0.27
BSc5188		0.36	9	n.d.	BSc5199		>10	39	n.d.
BSc5189		>10	29	n.d.	BSc5200		>10	120	n.d.
BSc5190		>10	27	n.d.	BSc5202		1.26	n.d.	n.d.
BSc5191		>10	14	n.d.	BSc5203		0.33	40	n.d.
BSc5192		0.78	22	n.d.	BSc5204		0.41	12	n.d.
BSc5193		>10	54	n.d.	BSc5205		0.61	18	0.23
BSc5194		0.64	19	n.d.	CAPF		~ 50	n.d.	n.d.

⁴⁴The IC₅₀ of CAPF for GlpG is a value from the literature.⁴⁷

Methods and Supplementary Figure 1) to evaluate their inhibitory potential in an established *in vitro* rhomboid activity assay.^{21,27,39}

The rhomboid protease GlpG from *E. coli* is an extremely well characterized serine intramembrane protease, both biochemically and structurally.^{16,28,48–50} GlpG also served as a standard model for the development of new rhomboid inhibitors and provided new insights into substrate recognition processes and related mechanistic details.^{23,27,30,51} Interestingly, no natural substrates have been described for GlpG, but its sequence preferences for cleavage are well-known.³⁰ We therefore used transmembrane domain 2 of the *Drosophila* protein Gurken as a substrate for the cleavage assays.^{39,52} GlpG and Gurken were expressed and purified as described elsewhere (see the Supporting Information and ref 39). To evaluate the preliminary batch of saccharin inhibitors, GlpG was preincubated with each compound at concentrations ranging from 1 to 250 μM for 30 min at 37 °C prior to the addition of the Gurken substrate. Saccharins with aryl carboxylic acid LGs were able to inhibit GlpG at a concentration of 10 μM, while saccharins with

aryl sulfide and sulfone LGs were inactive at 250 μM. These results were consistent with our findings from the docking experiments, which indicated that carboxylic acid LGs might optimally target the rhomboid protease GlpG. The docking studies showed that the aryl ring contacted Phe245 via π–π interactions and fitted nicely into a hydrophobic region within the GlpG active-site cleft (Figure 2A). Analysis of electrostatic interactions revealed that hydrophobic extensions of a LG in the *para* position might be favorable, whereas a LG substituted at the *ortho* position was engaged in proton donor/acceptor interactions. In addition, strong electron-withdrawing groups could activate the saccharin core structure for an attack by the active-site Ser201. Substitutions with small extensions were possible without steric clashes or external exposure at the *para* and *ortho* positions to optimize the contact between the inhibitor and enzyme as well as activating the inhibitor. Consequently, N-methylated saccharins with an aryl carboxylic acid LG were selected as “lead molecules”.

Synthesis and Test of the Second Compound Batch.

To further examine the most suitable extensions on the aryl

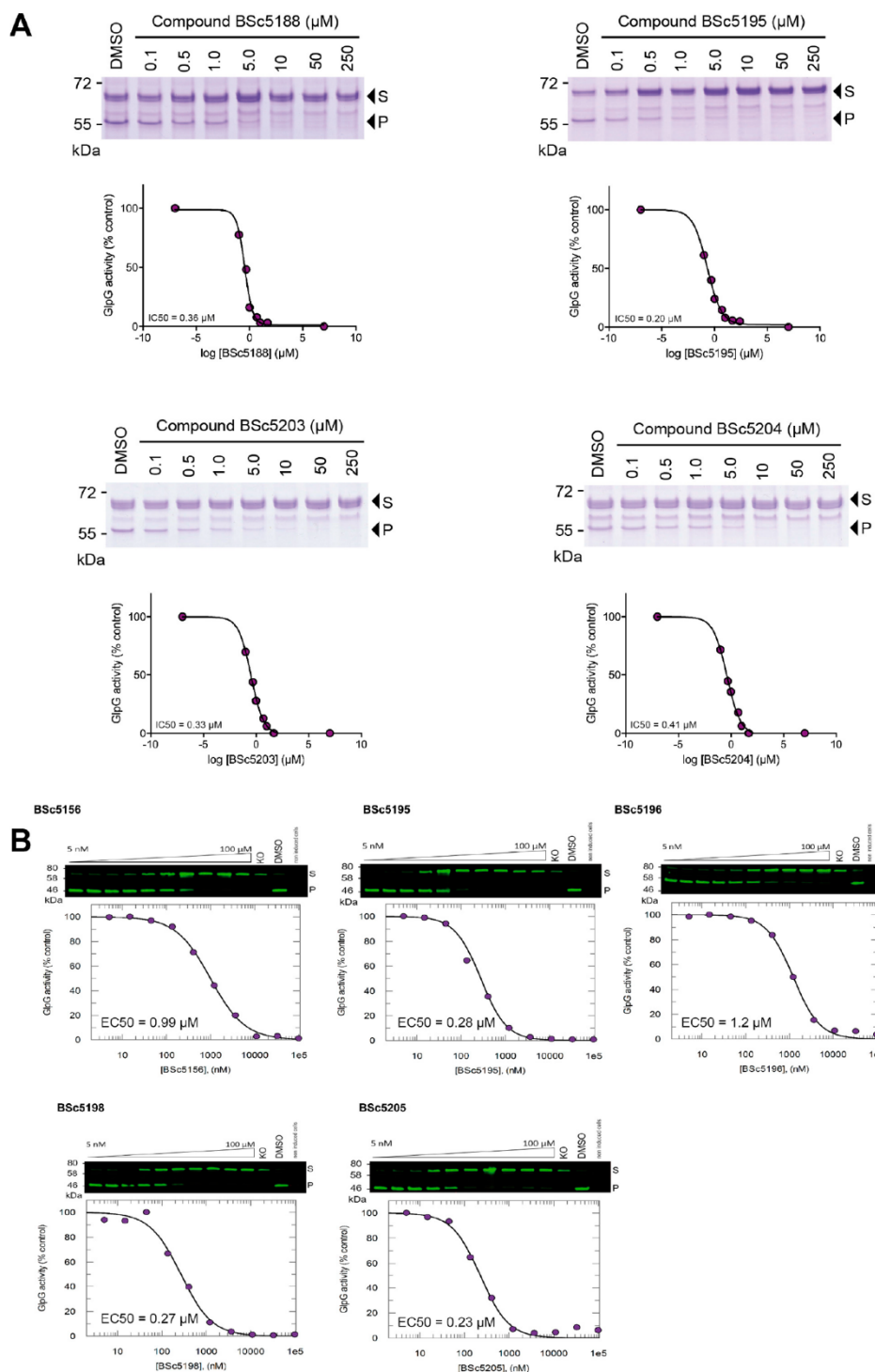


Figure 3. (A) Evaluation of saccharin inhibitors in the GlpG *in vitro* activity assay. The *E. coli* rhomboid GlpG was recombinantly expressed and preincubated with each saccharin inhibitor at concentrations ranging from 0.1 to 250 μM for 30 min at 37 $^{\circ}\text{C}$ prior to the addition of the Gurken substrate. The SDS-PAGE gel shows the reaction products at the end of a 90 min incubation period. The arrows indicate the full-length Gurken substrate (S, ~ 66 kDa) and the N-terminal Gurken cleavage product (P, ~ 55 kDa). Signal intensities of the N-terminal Gurken product were

Figure 3. continued

normalized to the DMSO control condition and plotted against $\log(\text{inhibitor concentration})$ vs percent N-terminal Gurken product was used to determine IC_{50} values. The SDS-PAGE gel was stained with Coomassie Brilliant Blue and quantified with ImageJ, and the values were plotted in GraphPad Prism. Two independent IC_{50} determinations were performed for each compound, and one representative experiment is shown. (B) *In vivo* inhibition of GlpG in live *E. coli* cells. Wild-type strain NR698 with endogenous GlpG activity⁵³ or GlpG-deficient strain NR698 $\Delta glpG$ ⁵⁷ was transformed with the MBP-FLAG-LacYTM2-Trx substrate.^{39,57} Increasing concentrations of the indicated saccharin inhibitors from 5 nM to 100 μM were added to the culture media; substrate expression was induced with 1 mM L-rhamnose, and cells were grown in the presence of the inhibitors for 4 h. Cleavage fragments of the substrate (P, ~46 kDa) were quantified by SDS-PAGE and near-infrared Western blotting of cell lysates, and IC_{50} values were determined as described previously.⁵⁷ Note that in the GlpG-deficient cells the substrate (S) was not processed and no cleavage fragment was produced (lane KO). Two independent IC_{50} determinations were performed for each compound, and one representative experiment is shown.

ring of the LG and the influence on the electron-withdrawing properties of the LG, we performed additional docking experiments with a database of 103 saccharin derivatives. Interestingly, the docking experiments did not reveal any significantly improved interactions between the aryl ring substituents and their neighboring residues within the active site of GlpG. However, to analyze how substitutions of the aryl acid moiety with different electron donor/acceptor groups would affect the inhibitory potency, 20 N-methylated saccharin derivatives (Supplementary Table 4) were synthesized using synthesis scheme 3 (for details, see Supplementary Methods and Supplementary Figure 1), and their biological activity was determined via the *in vitro* GlpG cleavage assay.

All aryl acid saccharin derivatives were able to inhibit GlpG in the *in vitro* activity assay at 250 μM , and compounds that showed activity at 10 μM were further selected for determination of IC_{50} values. First, to benchmark the GlpG *in vitro* activity assay with known rhomboid inhibitors, the IC_{50} values for the β -lactam inhibitor L16 and the isocoumarin DCI were determined and found to be 0.9 and 19 μM , respectively (Table 1), which were comparable to literature data.^{21,27} Subsequent experiments identified 12 saccharin derivatives with IC_{50} values of <1 μM (Figure 3A and Table 1). The aryl acid derivatives substituted with electron-withdrawing groups such as halogens at the *ortho* position (BSc5188, BSc5195, and BSc5197) had IC_{50} values of <1 μM , while an inhibitor substituted with an electron-donating group such as $-\text{NH}_2$ (BSc5193) was unable to completely inhibit GlpG at 10 μM . Therefore, electron-withdrawing groups at the *ortho* position of the aryl carboxylic acid LG appeared to increase the inhibitor potency. However, with this limited range of derivatives, it was not possible to definitively conclude how the electron-withdrawing properties of the LG affected the inhibitor potency. The IC_{50} value of the most potent compound that evolved from our screen was determined to be 200 nM (Figure 3A and Table 1).

Effectiveness of Saccharins against Rhomboid Embedded in the Lipid Membrane. Until now, we have evaluated the activity of the saccharin-based inhibitors only in the GlpG *in vitro* cleavage assay with purified enzyme and substrate solubilized in detergent micelles. In this assay, the enzyme active site is principally more accessible to incoming compounds than in the native environment of rhomboids, which is the lipid membrane. We therefore chose to evaluate four potent compounds in an *in vivo* system of *E. coli*.⁵⁷ The LacYTM2-based chimeric substrate was expressed in *E. coli* in the presence of increasing concentrations of rhomboid inhibitors, and substrate cleavage was assessed by SDS-PAGE and quantitative Western blotting. Compounds BSc5195, BSc5196, BSc5198, and BSc5205 inhibited the

processing of LacYTM2 by GlpG in living cells with IC_{50} values comparable to those obtained in the *in vitro* cleavage assay with purified enzyme and substrate (Figure 3B and Table 1).

Mechanism of Saccharin-Based Inhibitors. Our saccharin-based inhibitors were supposed to react with the nucleophilic side chain of the catalytic serine of a serine protease (Ser201 in the case of the rhomboid protease GlpG).²⁴ The resulting electronic rearrangements should then lead to the release of the acidic leaving group. Compound BSc5193 is an N-methylated saccharin derivative with the fluorescent anthranilic acid (*o*-aminobenzoic acid) as a LG. Indeed, incubation of BSc5193 with the rhomboid protease GlpG caused an increase in fluorescence emission at 400 nm, while there was no fluorescence increase with compound BSc5195, which was used as a negative control (Figure 4A). This supported the idea that the reaction mechanism of N-methylated saccharin-based acid derivatives involved the release of the acidic leaving group.

On the basis of the reaction mechanism of saccharins with classical soluble serine proteases,²⁴ we assumed that after release of the LG the generated Michael acceptor would further react with a nucleophilic residue in its immediate vicinity to form a double-bonded covalent end product (Figure 1B). In the direct vicinity of the catalytic Ser201, two histidine residues could possibly be modified by the saccharin inhibitors: His150 and His254. Alternatively, the Michael acceptor might further react with water, leaving only the catalytic serine modified by the saccharin. To differentiate between these possibilities, we reacted purified wild-type GlpG *in vitro* with five different saccharin inhibitors and studied the difference in mass between free GlpG and the inhibitor-GlpG adduct by electrospray mass spectrometry. All compounds (BSc5156, BSc5195, BSc5196, BSc5198, and BSc5205) generated identical mass shifts of 195 Da, which matched well the theoretical mass difference of 196 Da that would be expected from the reaction of a saccharin with a classical serine protease (Figure 4B). Next, we expressed and purified mutants of GlpG, in which either Ser201 was changed to threonine (S201T) or His150 and His254 were changed to alanine (H150A and H254A, respectively). Again, these GlpG mutants were examined alone or after incubation with saccharin inhibitor BSc5196 by mass spectrometry. BSc5196 was able to bind wild-type GlpG and the H150A mutant, but not the S201T, H254A, or H150A/H254A mutants (Figure 4C).

Overall, the mass spectrometry data were consistent with a reaction mechanism in which the saccharins would modify both the catalytic Ser201 and the catalytic His254, resulting in an end product that forms a covalent cross-link in the enzyme. Nevertheless, we did not consider these data fully conclusive for two reasons. First, in the catalytic Ser201-His254 dyad, the

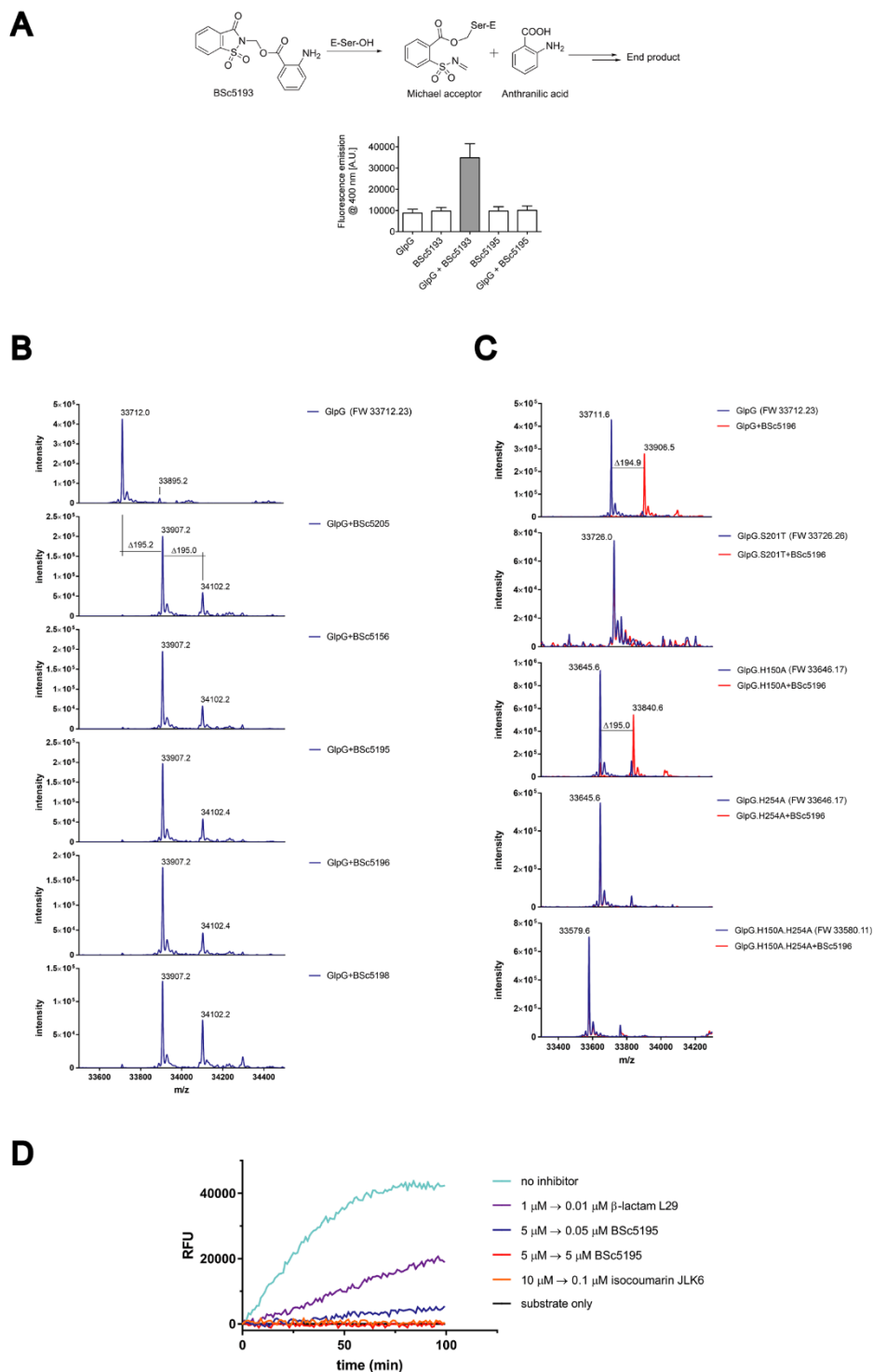


Figure 4. Mechanism of saccharin-based inhibitors. (A) The reaction of N-methylated saccharins with the rhomboid involves the release of the acidic LG. A saccharin derivative with anthranilic acid as the LG (BSc5193) was incubated with recombinantly expressed GlpG. According to the proposed mechanism, nucleophilic attack of Ser201 should lead to the formation of a Michael acceptor and release of the anthranilic acid LG, which becomes fluorescent. Incubation of GlpG with BSc5193 resulted in an increase in fluorescence as compared to those of reaction mixtures containing only the

Figure 4. continued

enzyme or the compound. However, no increase in fluorescence was observed for a different potent saccharin-based inhibitor (BSc5195) with a nonfluorescent LG. Average values of two independent experiments are shown. Error bars represent the standard deviation. (B) Purified wild-type GlpG was reacted *in vitro* with five different saccharins (BSc5205, BSc5156, BSc5195, BSc5196, and BSc5198), and the mass difference between free GlpG and the inhibitor–GlpG adduct was determined by electrospray mass spectrometry. The shift in the *m/z* value was the same after incubation with all five compounds. This indicated that all compounds inhibited GlpG via the same mechanism, which was expected because the covalently bound core structure was the same and only the LG differed between the compounds. Three independent experiments were performed. (C) Upon incubation of wild-type GlpG with the saccharin BSc5196, a shift in the *m/z* value corresponding to a covalent double-bonded end product after reaction with the protease was observed. Mutation of the catalytically active Ser201 to threonine or His254 to alanine abolished the difference in *m/z* observed with the wild-type protein. In contrast, mutation of His150 had no effect and produced a difference in *m/z* as seen with wild-type GlpG. Three independent experiments were performed, and one representative experiment is shown. (D) Reversibility of the saccharin reaction mechanism was investigated by the rapid dilution method. Saccharin BSc5195 (5 μ M), irreversible isocoumarin inhibitor JLK-6 (10 μ M), or reversible β -lactam inhibitor L29 (1 μ M) was preincubated with GlpG for 1 h at 10 times their IC_{50} concentrations. The reaction mixtures were then rapidly diluted 100-fold, and a fluorogenic rhomboid substrate peptide (10 μ M) was added. No recovery of enzyme activity was observed with JLK-6, while substantial recovery was seen with L29 as expected. With the saccharin inhibitor BSc5195, less but noticeable recovery of GlpG activity was observed over the time period of 100 min, demonstrating partial reversibility of the inhibition mechanism. No recovery of activity was apparent when the saccharin inhibitor concentration was maintained at 5 μ M after the dilution step. Three independent experiments were performed, and one representative experiment is shown.

histidine is indispensable for proper activation of the serine.⁵⁴ Consequently, it is highly likely that in the H254A mutant Ser201 is not nucleophilic enough to react with the carbonyl of the saccharin, providing an alternative explanation for why the saccharin failed to bind this mutant. Second, Ser201 and His254 are located on different transmembrane domains, and earlier work had demonstrated that cross-linking the active-site dyad with the isocoumarin inhibitor JLK-6 changed the mobility of the GlpG–inhibitor complex on a SDS–PAGE gel in comparison to that of unmodified GlpG.^{23,30} However, while we could easily reproduce the mobility shift with JLK-6, we did not observe a change in the apparent molecular weight of the rhomboid on SDS–PAGE gels after *in vitro* reaction with any of our saccharin inhibitors (data not shown). One possible explanation could be that the Michael acceptor did not react with His254 but with another nucleophilic residue closer to Ser201 on the same transmembrane domain, facilitating a cross-link that did not result in a conformational change and a corresponding mobility shift. Modeling the spatial proximity of the Michael acceptor to neighboring nucleophilic residues indicated that, aside from His254 (4.9 Å from the Michael acceptor carbon) and His150 (6.4 Å), Tyr205 (3.8 Å) was close in distance, whereas Trp157 and Trp236 were more distant. Other explanations for the lack of a gel shift could be that the saccharin-based inhibitors modified only the catalytically active serine and did not result in an intramolecular cross-link, or that in the double-bonded (cross-linked) end product the bond to the catalytic serine was deacylated, reverting to a single-bonded species (via the Michael acceptor). However, for the latter explanation, mass spectrometry analysis did not provide direct evidence, as it should have resulted in a mass shift of 214 kDa. We explored the possible hydrolysis of the Ser201–saccharin ester bond further by investigating the reversibility of covalent binding over time using the rapid dilution method and a previously described fluorogenic peptidic rhomboid substrate.^{40,55} The inhibitors were preincubated with GlpG for 1 h at 10 times their IC_{50} concentrations. The reaction mixtures were then rapidly diluted 100-fold, and the substrate peptide was added. Processing of this substrate by the rhomboid releases a quencher peptide and activates a red fluorophore, which indicates recovery of enzyme activity. When the assay was performed with the irreversible inhibitor JLK-6, which forms a double-bonded end product with the rhomboid,^{23,30} no recovery of enzyme activity was observed (Figure 4D).

Conversely, with the β -lactam L29, a known reversible inhibitor of rhomboids,²⁷ substantial recovery of activity was evident over a time period of 100 min. With the saccharin inhibitor BSc5195, less but noticeable recovery of GlpG activity was observed, demonstrating partial reversibility of the inhibition mechanism (Figure 4D). Taken together, these data do not easily conform to the reaction mechanism of saccharins with classical soluble serine proteases. Instead, they point to a heterogeneous mechanism in which the saccharin-based inhibitors might form a double-bonded end product by cross-linking Ser201 with another spatially close nucleophilic residue, or might modify only Ser201 because of the low reactivity of the Michael acceptor. This may be followed by slow hydrolysis of the formed ester bond, explaining the partial reversibility of the reaction at low inhibitor concentrations.

Selectivity for Rhomboid Proteases. The saccharin inhibitors were designed to fit into the active-site cleft of the *E. coli* rhomboid protease GlpG. However, we were further interested in knowing whether the compounds would inhibit other rhomboid proteases. To address this issue, we expressed and purified the rhomboid protease from *A. aeolicus* (AaROM) and performed an *in vitro* cleavage assay with AaROM and the Gurken substrate in the presence of the three most potent saccharin derivatives (BSc5188, BSc5195, and BSc5204). All three compounds were able to inhibit the AaROM activity with IC_{50} values between 60 and 90 μ M, indicating an approximately 100-fold lower potency toward AaROM in comparison to that of GlpG (Supplementary Table 5). Next, to investigate whether the saccharin-based compounds might also inhibit the enzymatic activity of a more distant member of the rhomboid protease family, a previously established cell-based assay for murine rhomboid RHBDL2 was employed.^{27,56} Human HEK293T cells were transiently co-transfected with a HA-tagged RHBDL2 construct and a FLAG-tagged Gurken substrate, and three saccharins with IC_{50} values of <1 μ M (BSc5156, BSc5205, and BSc5192) were added to the cell culture media at concentrations of ≤ 250 μ M. Forty-eight hours after the transfection, cell lysates were prepared and Gurken cleavage fragments were detected by Western blotting. The saccharin-based compounds displayed no inhibitory effect on the processing of Gurken by RHBDL2 (Supplementary Figure 4).

To further investigate the selectivity of the saccharin-based inhibitors against different rhomboids, we took advantage of a

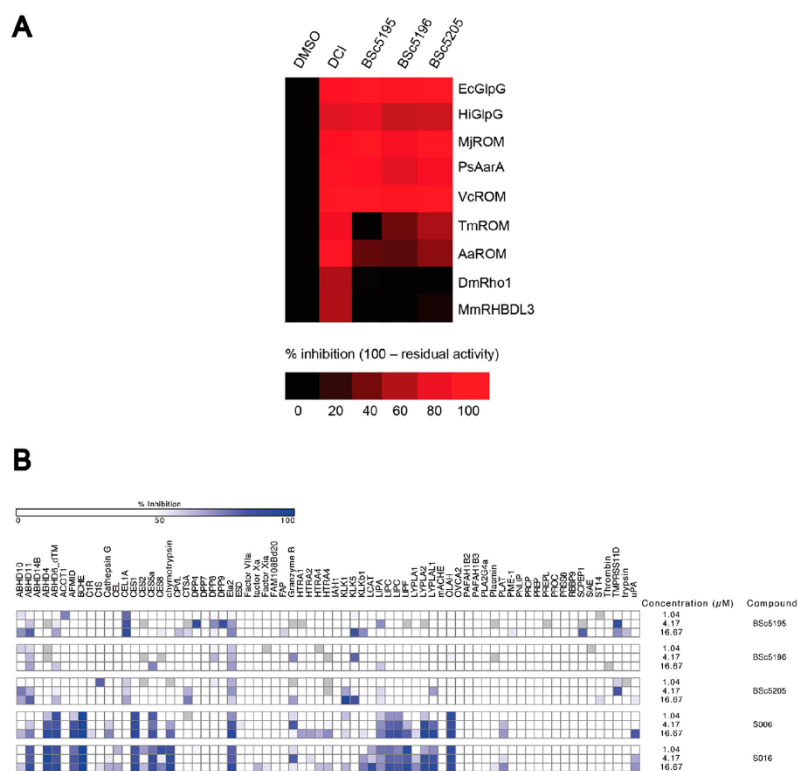


Figure 5. Selectivity of saccharin-based inhibitors. (A) Heat map representation of the activity-based probe (ABP) competition assay with three potent saccharin inhibitors against nine different rhomboid proteases. Recombinantly expressed rhomboids were preincubated with saccharins (50 μM), labeled with the ABP TAMRA-FP, and the labeled enzymes were visualized by SDS–PAGE and fluorescence scanning. The isocoumarin DCI was used as a positive control and inhibited around 80–100% of the activity of all nine rhomboids. Two independent experiments were performed, and average values for each experimental condition are shown. (B) A high-throughput activity-based probe competition assay was used to profile saccharins against human soluble serine hydrolases. Isocoumarins S006 and S016 inhibited multiple human serine hydrolases. In contrast, at concentrations sufficient to inhibit the rhomboid GlpG in the *in vitro* activity assay, saccharin-based compounds BSc5195, BSc5196, and BSc5205 showed far less activity against the panel of human serine hydrolases (no inhibition, white; weak inhibition, gray to light blue; strong inhibition, dark blue). Two independent experiments were performed, and one representative experiment is shown.

recently described activity-based probe (ABP) competition assay.^{23,36} In this assay, small molecule ABPs were used to specifically label the active site of recombinantly expressed and proteolytically active rhomboids. One such ABP is the commercially available TAMRA-FP serine hydrolase probe, which allows the visualization of the labeled enzyme by SDS–PAGE and fluorescence scanning. This fluorophosphate-derived compound binds to the active site of rhomboid proteases and covalently modifies the catalytic serine. However, if after preincubation of the rhomboids with a saccharin-based inhibitor the catalytic site of the enzyme is already modified, the fluorescent ABP would not be able to bind and the fluorescent signal would be reduced. In total, nine different rhomboids from all three phylogenetic trees (archaea, bacteria, and eukaryota) were selected and screened against three potent saccharin-based inhibitors (BSc5195, BSc5196, and BSc5205) and the nonspecific pan-rhomboid inhibitor DCI as a positive control. All compounds were used at a final concentration of 50 μM and were preincubated with the rhomboids before being labeled with the ABP. The results are shown in a heat map to visualize the selectivity of the saccharin-based inhibitors. As expected, the isocoumarin DCI inhibited around 80–100% of the activity of all nine rhomboids (Figure 5A). Interestingly, the

saccharin-based inhibitors were active against only four bacterial rhomboids (*E. coli* GlpG, *Haemophilus influenzae* HiGlpG, *Vibrio cholerae* VcROM, and *Providencia stuartii* PsAarA) and one archaeal rhomboid (*Methanocaldococcus jannaschii* MjROM). Two other bacterial rhomboids (*A. aeolicus* AaROM and *Thermotoga maritima* TmROM) were only partially inhibited, whereas two eukaryotic rhomboids (*Drosophila melanogaster* DmRho1 and *Mus musculus* RHBDL3) were completely spared (Figure 5A). These results are fully consistent with the low potency of the saccharin inhibitors in the AaROM *in vitro* activity assay (Supplementary Table 5) and the lack of an effect in the cell-based assay for eukaryotic murine rhomboid RHBDL2 (Supplementary Figure 4). Taken together, these data indicated that the saccharin-based inhibitors, in contrast to DCI, did not indiscriminately inhibit all rhomboids but were more specific and might have a preference for the evolutionarily older bacterial and archaeal rhomboid proteases.

Selectivity over Classical Serine Proteases. A common problem of inhibitors against broadly distributed enzymes such as proteases is off-target activity. In addition to finding more potent rhomboid inhibitors, we sought to identify more specific inhibitors with a clear preference for rhomboids over soluble

serine proteases. Therefore, we first investigated the potency of 18 aryl acid-substituted saccharins against the well-characterized soluble serine protease α -chymotrypsin in an *in vitro* assay.⁴¹ α -Chymotrypsin was preincubated with saccharins for 30 min at 25 °C prior to the addition of the substrate *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide. This substrate is processed by α -chymotrypsin under alkaline conditions to yield *p*-nitroaniline, which can be detected spectroscopically at 410 nm. The positive control DCI inhibited α -chymotrypsin with an IC₅₀ value of 3.5 μ M, which is close to the value reported in the literature (2.8 μ M).²⁷ The IC₅₀ values of saccharins against α -chymotrypsin were 20–370-fold higher than the IC₅₀ values for inhibition of GlpG in the *in vitro* activity assay (Table 1). This indicated that the saccharin-based inhibitors offered an at least 20-fold window of selectivity for the inhibition of the rhomboid GlpG over the soluble serine protease α -chymotrypsin. However, it is important to note that the IC₅₀ values for α -chymotrypsin and GlpG were determined in different assays and are not directly comparable. In addition, a more exhaustive specificity profiling against 71 serine hydrolases using a high-throughput activity-based probe competition assay³⁷ showed that, while two isocoumarins S006 and S016 inhibited more than a dozen other human serine hydrolases, saccharins BSc5195, BSc5196, and BSc5205 displayed appreciable inhibition of only a few enzymes [e.g., neutrophil elastase, ELA2 (see also Supplementary Figure 5)] in the same concentration range, in which they inhibited the rhomboid GlpG (Figure 5B). Overall, these experiments led us to conclude that *N*-methylene saccharin-derived compounds are potent inhibitors of a subset of rhomboids with selectivity far better than that of isocoumarins.

CONCLUSIONS

In this study, we have described the computer-aided discovery and biological characterization of a novel and improved class of partially reversible inhibitors of rhomboid proteases. In comparison to previously known rhomboid inhibitors, the *N*-methylene saccharins presented here have increased potency and selectivity, and a certain preference for archaeal and bacterial rhomboids. Further structural studies will be needed to determine how saccharins might be able to distinguish between different rhomboids. Importantly, *N*-methylene saccharin-based rhomboid protease inhibitors are easily amenable to derivatization and have a clear potential for further development.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b01066.

Five supplementary figures showing chemical synthesis schemes, docking data, a cell-based activity assay for murine rhomboid RHBDL2, and an *in vitro* activity assay for human neutrophil elastase; five tables displaying chemical structures and summarizing IC₅₀ values of saccharin inhibitors in various biological assays; and detailed synthesis and analytical information about saccharin inhibitors (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: schmidt_boris@t-online.de.

*E-mail: sweggen@hhu.de.

ORCID

Kvido Strisovsky: 0000-0003-3677-0907

Boris Schmidt: 0000-0003-1662-2392

Sascha Weggen: 0000-0002-9800-1145

Author Contributions

P.G. and T.J. contributed equally to this work.

Funding

P.G. was supported by a scholarship of the iGRASPseed graduate school of the Heinrich-Heine-University Duesseldorf. K.S. was a recipient of Purkyne Fellowship of the Academy of Sciences of the Czech Republic and also acknowledges support from EMBO (Installation Grant 2329), Ministry of Education, Youth and Sports of the Czech Republic (Projects LK11206 and LO1302), a Marie Curie Career Integration Grant (Project 304154), and the National Subvention for Development of Research Organisations (RVO, 61388963) to the Institute of Organic Chemistry and Biochemistry. D.A.B. was supported by the Josie Robertson Foundation and MSKCC Core Grant P30 CA008748 and D.C.J. by National Institutes of Health Grant T32 GM115327-Tan. M.T.N.N. and S.V. were supported by the Deutsche Forschungsgemeinschaft, the Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen, the Senatsverwaltung für Wirtschaft, Technologie und Forschung des Landes Berlin, and the Bundesministerium für Bildung und Forschung.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Professors Matthew Freeman (University of Oxford), Sinisa Urban (Johns Hopkins University), and Ya Ha (Yale School of Medicine) for various expression constructs.

REFERENCES

- (1) Puente, X. S., Sanchez, L. M., Overall, C. M., and Lopez-Otin, C. (2003) Human and mouse proteases: a comparative genomic approach. *Nat. Rev. Genet.* 4, 544–558.
- (2) Rawlings, N. D., and Barrett, A. J. (1993) Evolutionary families of peptidases. *Biochem. J.* 290 (1), 205–218.
- (3) Egeblad, M., and Werb, Z. (2002) New functions for the matrix metalloproteinases in cancer progression. *Nat. Rev. Cancer* 2, 161–174.
- (4) Krane, S. M. (2003) Elucidation of the potential roles of matrix metalloproteinases in skeletal biology. *Arthritis Res. Ther.* 5, 2–4.
- (5) Lichtenthaler, S. F., Haass, C., and Steiner, H. (2011) Regulated intramembrane proteolysis—lessons from amyloid precursor protein processing. *J. Neurochem.* 117, 779–796.
- (6) Luttun, A., Dewerchin, M., Collen, D., and Carmeliet, P. (2000) The role of proteinases in angiogenesis, heart development, restenosis, atherosclerosis, myocardial ischemia, and stroke: insights from genetic studies. *Curr. Atheroscler. Rep.* 2, 407–416.
- (7) Mittl, P. R., and Grutter, M. G. (2006) Opportunities for structure-based design of protease-directed drugs. *Curr. Opin. Struct. Biol.* 16, 769–775.
- (8) Siefert, S. A., and Sarkar, R. (2012) Matrix metalloproteinases in vascular physiology and disease. *Vascular* 20, 210–216.
- (9) Turk, B. (2006) Targeting proteases: successes, failures and future prospects. *Nat. Rev. Drug Discovery* 5, 785–799.
- (10) Urban, S., Lee, J. R., and Freeman, M. (2001) Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* 107, 173–182.

- (11) Koonin, E. V., Makarova, K. S., Rogozin, I. B., Davidovic, L., Letellier, M. C., and Pellegrini, L. (2003) The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers. *Genome Biol.* 4, R19.
- (12) Urban, S. (2009) Making the cut: central roles of intramembrane proteolysis in pathogenic microorganisms. *Nat. Rev. Microbiol.* 7, 411–423.
- (13) Storer, A. C., and Menard, R. (1994) Catalytic mechanism in papain family of cysteine peptidases. *Methods Enzymol.* 244, 486–500.
- (14) Hedstrom, L. (2002) Serine protease mechanism and specificity. *Chem. Rev.* 102, 4501–4524.
- (15) Lemberg, M. K., Menendez, J., Misik, A., Garcia, M., Koth, C. M., and Freeman, M. (2005) Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. *EMBO J.* 24, 464–472.
- (16) Wang, Y., Zhang, Y., and Ha, Y. (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature* 444, 179–180.
- (17) Freeman, M. (2009) Rhomboids: 7 years of a new protease family. *Semin. Cell Dev. Biol.* 20, 231–239.
- (18) Freeman, M. (2014) The rhomboid-like superfamily: molecular mechanisms and biological roles. *Annu. Rev. Cell Dev. Biol.* 30, 235–254.
- (19) Strisovsky, K. (2013) Structural and mechanistic principles of intramembrane proteolysis—lessons from rhomboids. *FEBS J.* 280, 1579–1603.
- (20) Strisovsky, K. (2016) Rhomboid protease inhibitors: Emerging tools and future therapeutics. *Semin. Cell Dev. Biol.* 60, 52–62.
- (21) Urban, S., and Wolfe, M. S. (2005) Reconstitution of intramembrane proteolysis in vitro reveals that pure rhomboid is sufficient for catalysis and specificity. *Proc. Natl. Acad. Sci. U. S. A.* 102, 1883–1888.
- (22) Vinothkumar, K. R., Strisovsky, K., Andreeva, A., Christova, Y., Verhelst, S., and Freeman, M. (2010) The structural basis for catalysis and substrate specificity of a rhomboid protease. *EMBO J.* 29, 3797–3809.
- (23) Vosyka, O., Vinothkumar, K. R., Wolf, E. V., Brouwer, A. J., Liskamp, R. M., and Verhelst, S. H. (2013) Activity-based probes for rhomboid proteases discovered in a mass spectrometry-based assay. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2472–2477.
- (24) Powers, J. C., Asgian, J. L., Ekici, O. D., and James, K. E. (2002) Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* 102, 4639–4750.
- (25) Powers, J. C., Kam, C. M., Narasimhan, L., Oleksyszyn, J., Hernandez, M. A., and Ueda, T. (1989) Mechanism-based isocoumarin inhibitors for serine proteases: use of active site structure and substrate specificity in inhibitor design. *J. Cell. Biochem.* 39, 33–46.
- (26) Wolf, E. V., Zeissler, A., and Verhelst, S. H. (2015) Inhibitor Fingerprinting of Rhomboid Proteases by Activity-Based Protein Profiling Reveals Inhibitor Selectivity and Rhomboid Autoprocessing. *ACS Chem. Biol.* 10, 2325–2333.
- (27) Pierrat, O. A., Strisovsky, K., Christova, Y., Large, J., Ansell, K., Bouloc, N., Smiljanic, E., and Freeman, M. (2011) Monocyclic beta-lactams are selective, mechanism-based inhibitors of rhomboid intramembrane proteases. *ACS Chem. Biol.* 6, 325–335.
- (28) Xue, Y., and Ha, Y. (2012) Catalytic mechanism of rhomboid protease GlpG probed by 3,4-dichloroisocoumarin and diisopropyl fluorophosphonate. *J. Biol. Chem.* 287, 3099–3107.
- (29) Wolf, E. V., Zeissler, A., Vosyka, O., Zeiler, E., Sieber, S., and Verhelst, S. H. (2013) A new class of rhomboid protease inhibitors discovered by activity-based fluorescence polarization. *PLoS One* 8, e72307.
- (30) Zoll, S., Stanchev, S., Began, J., Skerle, J., Lepsik, M., Peclinovska, L., Majer, P., and Strisovsky, K. (2014) Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures. *EMBO J.* 33, 2408–2421.
- (31) Cho, S., Dickey, S. W., and Urban, S. (2016) Crystal Structures and Inhibition Kinetics Reveal a Two-Stage Catalytic Mechanism with Drug Design Implications for Rhomboid Proteolysis. *Mol. Cell* 61, 329–340.
- (32) Hlasta, D. J., Ackerman, J. H., Court, J. J., Farrell, R. P., Johnson, J. A., Kofron, J. L., Robinson, D. T., Talomie, T. G., Dunlap, R. P., and Franke, C. A. (1995) A novel class of cyclic beta-dicarbonyl leaving groups and their use in the design of benzisothiazolone human leukocyte elastase inhibitors. *J. Med. Chem.* 38, 4687–4692.
- (33) Walker, B., and Lynas, J. F. (2001) Strategies for the inhibition of serine proteases. *Cell. Mol. Life Sci.* 58, 596–624.
- (34) Zhong, J., and Groutas, W. C. (2004) Recent developments in the design of mechanism-based and alternate substrate inhibitors of serine proteases. *Curr. Top. Med. Chem.* 4, 1203–1216.
- (35) Klebe, G. (2006) Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* 11, 580–594.
- (36) Wolf, E. V., Seybold, M., Hadravova, R., Strisovsky, K., and Verhelst, S. H. (2015) Activity-Based Protein Profiling of Rhomboid Proteases in Liposomes. *ChemBioChem* 16, 1616–1621.
- (37) Bachovchin, D. A., Koblan, L. W., Wu, W., Liu, Y., Li, Y., Zhao, P., Woznica, I., Shu, Y., Lai, J. H., Poplawski, S. E., Kiritsy, C. P., Healey, S. E., DiMare, M., Sanford, D. G., Munford, R. S., Bachovchin, W. W., and Golub, T. R. (2014) A high-throughput, multiplexed assay for superfamily-wide profiling of enzyme activity. *Nat. Chem. Biol.* 10, 656–663.
- (38) *Molecular Operating Environment* (2013) <http://github.com/Yelp/MOE>.
- (39) Strisovsky, K., Sharpe, H. J., and Freeman, M. (2009) Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. *Mol. Cell* 36, 1048–1059.
- (40) Ticha, A., Stanchev, S., Skerle, J., Began, J., Ingr, M., Svehlova, K., Polovinkin, L., Ruzicka, M., Bednarova, L., Hadravova, R., Polachova, E., Rampirova, P., Brezinova, J., Kasicka, V., Majer, P., and Strisovsky, K. (2017) Sensitive Versatile Fluorogenic Transmembrane Peptide Substrates for Rhomboid Intramembrane Proteases. *J. Biol. Chem.* 292, 2703–2713.
- (41) DelMar, E. G., Largman, C., Brodrick, J. W., and Geokas, M. C. (1979) A sensitive new substrate for chymotrypsin. *Anal. Biochem.* 99, 316–320.
- (42) Zimmerman, M., Morman, H., Mulvey, D., Jones, H., Frankshun, R., and Ashe, B. M. (1980) Inhibition of elastase and other serine proteases by heterocyclic acylating agents. *J. Biol. Chem.* 255, 9848–9851.
- (43) Groutas, W. C., Brubaker, M. J., Venkataraman, R., Epp, J. B., Stanga, M. A., and McClenahan, J. J. (1992) Inhibitors of human neutrophil cathepsin G: structural and biochemical studies. *Arch. Biochem. Biophys.* 294, 144–146.
- (44) Groutas, W. C., Chong, L. S., Venkataraman, R., Kuang, R., Epp, J. B., Houser-Archfield, N., Huang, H., and Hoidal, J. R. (1996) Amino acid-derived phthalimide and saccharin derivatives as inhibitors of human leukocyte elastase, cathepsin G, and proteinase 3. *Arch. Biochem. Biophys.* 332, 335–340.
- (45) Groutas, W. C., Epp, J. B., Venkataraman, R., Kuang, R., My Truong, T., McClenahan, J. J., and Prakash, O. (1996) Design, synthesis, and in vitro inhibitory activity toward human leukocyte elastase, cathepsin G, and proteinase 3 of saccharin-derived sulfones and congeners. *Bioorg. Med. Chem.* 4, 1393–1400.
- (46) Groutas, W. C., Kuang, R., Ruan, S., Epp, J. B., Venkataraman, R., and Truong, T. M. (1998) Potent and specific inhibition of human leukocyte elastase, cathepsin G and proteinase 3 by sulfone derivatives employing the 1,2,5-thiadiazolidin-3-one 1,1 dioxido scaffold. *Bioorg. Med. Chem.* 6, 661–671.
- (47) Xue, Y., Chowdhury, S., Liu, X., Akiyama, Y., Ellman, J., and Ha, Y. (2012) Conformational change in rhomboid protease GlpG induced by inhibitor binding to its S' subsites. *Biochemistry* 51, 3723–3731.
- (48) Ben-Shem, A., Fass, D., and Bibi, E. (2007) Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proc. Natl. Acad. Sci. U. S. A.* 104, 462–466.

(49) Baker, R. P., and Urban, S. (2015) Cytosolic extensions directly regulate a rhomboid protease by modulating substrate gating. *Nature* 523, 101–105.

(50) Dickey, S. W., Baker, R. P., Cho, S., and Urban, S. (2013) Proteolysis inside the membrane is a rate-governed reaction not driven by substrate affinity. *Cell* 155, 1270–1281.

(51) Vinothkumar, K. R., Pierrat, O. A., Large, J. M., and Freeman, M. (2013) Structure of rhomboid protease in complex with beta-lactam inhibitors defines the S2' cavity. *Structure* 21, 1051–1058.

(52) Urban, S., and Freeman, M. (2003) Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain. *Mol. Cell* 11, 1425–1434.

(53) Ruiz, N., Falcone, B., Kahne, D., and Silhavy, T. J. (2005) Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell* 121, 307–317.

(54) Ha, Y., Akiyama, Y., and Xue, Y. (2013) Structure and mechanism of rhomboid protease. *J. Biol. Chem.* 288, 15430–15436.

(55) Harper, J. W., Hemmi, K., and Powers, J. C. (1985) Reaction of serine proteases with substituted isocoumarins: discovery of 3,4-dichloroisocoumarin, a new general mechanism based serine protease inhibitor. *Biochemistry* 24, 1831–1841.

(56) Adrain, C., Strisovsky, K., Zettl, M., Hu, L., Lemberg, M. K., and Freeman, M. (2011) Mammalian EGF receptor activation by the rhomboid protease RHBDL2. *EMBO Rep.* 12, 421–427.

(57) Ticha, A., Stanchev, S., Vinothkumar, K. R., Mikles, D. C., Pachel, P., Began, J., Skerle, J., Svehlova, K., Nguyen, M. T. N., Verhelst, S. H. L., et al. (2017) General and modular strategy for designing potent, selective and pharmacologically compliant inhibitors of rhomboid proteases. *Cell Chem. Biol.* 24, 1–14.



Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

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

Discovery and validation of 2-styryl substituted benzoxazin-4-ones as a novel scaffold for rhomboid protease inhibitors

Parul Goel^{a,d}, Thorsten Jumpertz^a, Anežka Tichá^b, Isabella Ogorek^a, David C. Mikles^b, Martin Hubalek^b, Claus U. Pietrzik^c, Kvido Strisovsky^b, Boris Schmidt^{d,*}, Sascha Weggen^{a,*}^a Department of Neuropathology, Heinrich-Heine University Duesseldorf, Moorenstrasse 5, 40225 Duesseldorf, Germany^b Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 10, Praha 6, Czech Republic^c Institute for Pathobiochemistry, University Medical Center of the Johannes Gutenberg University Mainz, Duesbergweg 6, 55128 Mainz, Germany^d Clemens Schoepf Institute for Organic Chemistry and Biochemistry, Technische Universitaet Darmstadt, Alarich-Weiss-Strasse 4-8, 64287 Darmstadt, Germany

ARTICLE INFO

Article history:

Received 14 November 2017

Revised 5 February 2018

Accepted 8 February 2018

Available online 9 February 2018

Keywords:

Rhomboid proteases
Intramembrane proteases
Benzoxazinones
Molecular docking
Inhibition

ABSTRACT

Rhomboids are intramembrane serine proteases with diverse physiological functions in organisms ranging from archaea to humans. Crystal structure analysis has provided a detailed understanding of the catalytic mechanism, and rhomboids have been implicated in various disease contexts. Unfortunately, the design of specific rhomboid inhibitors has lagged behind, and previously described small molecule inhibitors displayed insufficient potency and/or selectivity. Using a computer-aided approach, we focused on the discovery of novel scaffolds with reduced liabilities and the possibility for broad structural variations. Docking studies with the *E. coli* rhomboid GlpG indicated that 2-styryl substituted benzoxazinones might comprise novel rhomboid inhibitors. Protease *in vitro* assays confirmed activity of 2-styryl substituted benzoxazinones against GlpG but not against the soluble serine protease α -chymotrypsin. Furthermore, mass spectrometry analysis demonstrated covalent modification of the catalytic residue Ser201, corroborating the predicted mechanism of inhibition and the formation of an acyl enzyme intermediate. In conclusion, 2-styryl substituted benzoxazinones are a novel rhomboid inhibitor scaffold with ample opportunity for optimization.

© 2018 Elsevier Ltd. All rights reserved.

Results and discussion

Rhomboids are intramembrane serine proteases present in prokaryotic, archaeal and eukaryotic organisms.¹ In 2001, the first rhomboid was discovered in *Drosophila* and shown to perform a critical proteolysis step in EGF-receptor signaling.^{2,3} Since then, rhomboids have been implicated in a wide range of biological processes including bacterial quorum sensing⁴, mitochondrial dynamics and integrity^{5,6}, and protein quality control.⁷ In addition, rhomboids have been identified as putative drug targets in the context of multiple diseases⁸ such as cancer⁹, diabetes^{10,11}, parasitic diseases^{12,13}, and Parkinson's disease.⁵ The crystal structures of rhomboids from *E. coli* and *H. influenzae* have been solved and revealed that rhomboids are serine-histidine dyad proteases composed of 6 core transmembrane helices, which form a V-shaped cavity and expose the active site to a partially hydrophilic environment.^{14,15} These structures together with numerous biochemical

studies have provided a detailed understanding of the catalytic mechanism of rhomboid proteases¹⁶, but this has not yet translated into the development of potent, selective and drug-like inhibitors.¹⁷ Through different strategies, from the testing of candidate molecules to rational synthesis to the screening of small molecule libraries, isocoumarins^{3,18,19}, fluorophosphonates²⁰, β -lactams²¹, and β -lactones²² were found to be effective against rhomboids, but these inhibitors generally displayed low potency and/or insufficient selectivity.^{18,20,21} Effectively, inherent liabilities as exemplified by the high reactivity of isocoumarins likely preclude or limit further development of these compound classes.^{23,24}

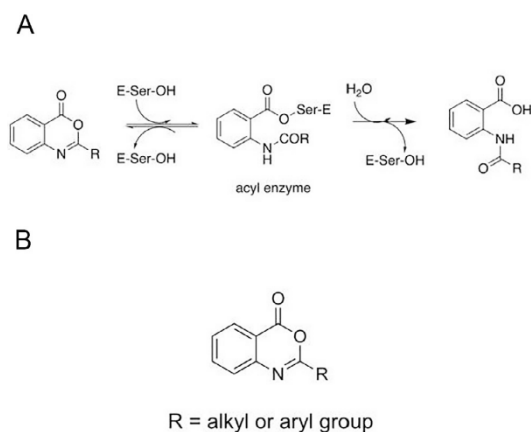
Accordingly, using a computer-aided candidate approach, we focused on the discovery of novel scaffolds with reduced liabilities and the possibility for broad structural variations. One scaffold we selected was 2-substituted derivatives of 4H-3,1-benzoxazin-4-ones, which were previously used as heterocyclic acylating agents against serine proteases such as HLE, α -chymotrypsin, and cathepsin G.^{23,25–28} The mechanism of inhibition involves the formation of an O-acyl enzyme intermediate. The nucleophilic serine reacts with the C-4 carbonyl of the benzoxazinone, which results in opening of the heterocyclic ring and formation of the O-acyl enzyme

* Corresponding authors.

E-mail addresses: schmidt_boris@t-online.de (B. Schmidt), sweggen@hhu.de (S. Weggen).<https://doi.org/10.1016/j.bmcl.2018.02.017>

0960-894X/© 2018 Elsevier Ltd. All rights reserved.

intermediate (Fig. 1A).²³ The enzyme selectivity and potency of acylating agents is promoted by fast acylation and slow deacylation, which is dependent on the substitution of the aromatic ring and the C-2 position in case of benzoxazinones.^{25,29,30} A major advantage of benzoxazin-4-ones is that the core structure consists of two fused aromatic rings, which allows extensive structural modifications and optimization with respect to the target enzyme. For initial docking studies into the rhomboid active site, we assembled a molecular database of thirteen 2-alkyl or 2-aryl substituted benzoxazinones (Fig. 1B).



Compd	-R	Compd	-R
1		8	
2	-CH ₃	9	
3		10	
4		11	
5		12	
6		13	
7			

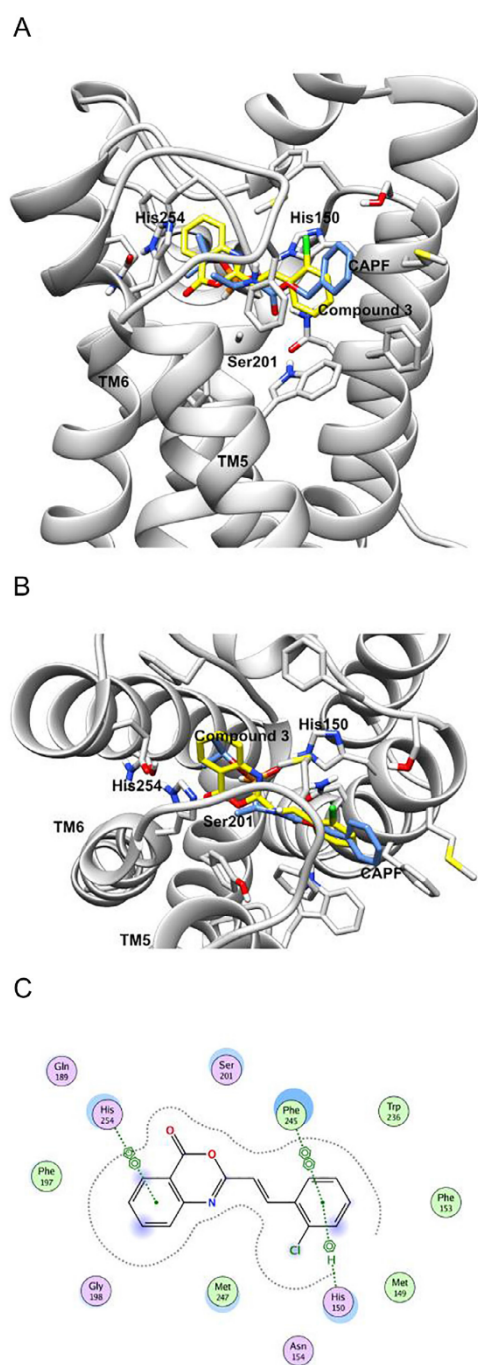
Fig. 1. (A) Mechanism of inhibition of soluble serine proteases by 2-substituted benzoxazin-4-one derivatives.³⁰ (B) Molecular database of 2-alkyl or aryl substituted benzoxazin-4-ones assembled for docking studies into the rhomboid active site.

In the docking studies, we focused on the initial interactions between the benzoxazinones and the active site of the rhomboid protease rather than the final reaction product. For preparation of the docking receptor, we used the co-crystal structure of the *E. coli* rhomboid GlpG and the fluorophosphonate inhibitor CAPF (PDB ID: 3UUB), in which the active site Ser201 is covalently bound to CAPF.³¹ The molecular modelling experiments were performed in the molecular operating environment software (MOE) with the DOCK module and the MMFF94x force field, and scored by London dG and Affinity dG followed by energy minimization within the enzyme active site cleft.^{32,33} The output data was ranked based on the calculated ligand efficiencies (cLE = docking score/number of heavy atoms)³⁴, which revealed that the 2-styryl substituted benzoxazinone **3** was the most favorable of all 2-substituted benzoxazinones (cLE = -0.3164). A comparative analysis of the protein/ligand docking results of compound **3** and CAPF indicated that **3** was adequately fitting into the binding pocket of the enzyme and was not exposed to the external environment (Fig. 2A and B). The core heterocyclic ring of **3** was oriented towards the S1 subsite while the 2-styryl extension pointed towards the S2' subsite of the rhomboid, which had been defined in previous structures of GlpG in complex with different inhibitors.^{31,35} Moreover, close interactions of **3** with the neighbouring residues His254 and Phe245 as shown in the ligand interaction map were observed and suggested to further explore the scaffold (Fig. 2C).

To validate the docking results, all derivatives listed in Fig. 1B were synthesized by methods shown in schemes 1 and 2 (Fig. 3A).³⁶ The benzoxazinone derivatives were then evaluated for their inhibitory potency in an established *in vitro* enzyme activity assay with the *E. coli* rhomboid GlpG and the transmembrane domain 2 of the *Drosophila* protein Gurken as a substrate.^{21,37–39} Each of the benzoxazinones were pre-incubated with GlpG at a single concentration of 250 μ M for 30 min at 37 °C. Subsequently, the Gurken substrate was added, the reaction was continued for another 90 min at 37 °C, and the N-terminal Gurken substrate cleavage fragment was visualized by SDS-PAGE and quantified using ImageJ. Only the 2-styryl substituted benzoxazinones **3**, **5** and **11** showed activity at this concentration. For IC₅₀ determinations, fluorogenic rhomboid substrates were applied as described previously.^{40–42} Consistent with the docking results, compound **3** was a potent rhomboid inhibitor with an IC₅₀ value of 4.4 \pm 1.6 μ M (Fig. 3B). Compound **5** was equally potent (IC₅₀ 3.7 \pm 1.3 μ M) while **11** displayed around 10-fold lower activity (IC₅₀ 48 \pm 14.1 μ M). Among these three compounds with a single substitution at the aromatic ring of the styryl substituent, an electron withdrawing group appeared to increase the potency, which could be further explored in subsequent studies.

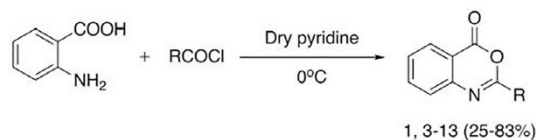
Next, we evaluated the active benzoxazinones in a well-established *in vitro* activity assay for the soluble serine protease α -chymotrypsin.^{43,44} The compounds were pre-incubated with bovine α -chymotrypsin for 30 min at 25 °C. Subsequently, the substrate N-succinyl-Ala-Ala-Pro-Phe-p-nitroanilide was added, which under alkaline conditions is turned over by α -chymotrypsin to p-nitroaniline, a yellow compound that can be detected spectroscopically at 410 nm. 3,4-Dichloroisocoumarin (DCI) was used as a positive control and inhibited α -chymotrypsin with an IC₅₀ value of 3.5 μ M, comparable to previously reported values.²¹ In contrast, the benzoxazinones **3** and **5** did not display any inhibitory activity in the α -chymotrypsin *in vitro* activity assay at the highest concentration of 250 μ M (data not shown). In addition, we examined the active benzoxazinones in similar *in vitro* protease activity assays for bovine trypsin, human neutrophil elastase, and human cathepsin G.⁴⁵ At a concentration of 10 μ M, none of the benzoxazinones inhibited trypsin or neutrophil elastase, while DCI almost completely abolished the activity of both enzymes. At a concentration

of 50 μM , trypsin activity was reduced around 50% by compound **3**, and neutrophil elastase activity was reduced around 50% by compound **11** (Supplemental Figs. 1 and 2). In contrast, the compounds were more effective against cathepsin G. Both **3** and **5** but not **11**

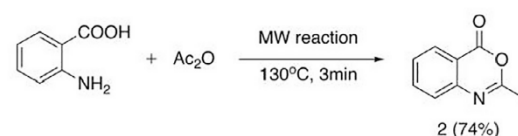


A

Scheme 1



Scheme 2



B

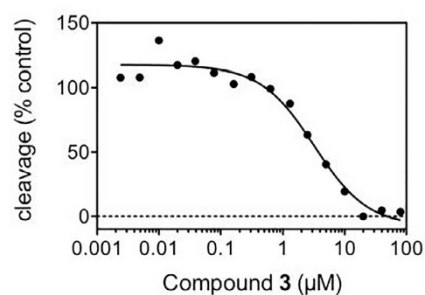


Fig. 3. (A) Synthesis of 2-substituted benzoxazin-4-ones. Scheme 1 shows the synthesis of 2-alkyl/aryl substituted benzoxazinones (compounds **1** and **3-13**) obtained by the reaction of corresponding acyl/benzoyl chlorides with anthranilic acid. Scheme 2 shows the synthesis of 2-methyl benzoxazin-4-one (compound **2**) obtained by a microwave reaction of anthranilic acid with acetic anhydride. (B) Biological evaluation of 2-alkyl and 2-aryl substituted benzoxazinone inhibitors in rhomboid *in vitro* activity assays. All compounds were first screened in a gel-based activity assay for the *E. coli* rhomboid GlpG at a single concentration of 250 μM as described.³⁹ Only the 2-aryl substituted benzoxazinones were found to be active at this concentration and were selected for IC_{50} determinations. In these assays, the rhomboid GlpG was pre-incubated with increasing concentrations of the compounds **3**, **5** and **11** for 1 h at 37 $^{\circ}\text{C}$. Subsequently, the reaction was started by the addition of fluorogenic substrates.⁴⁰ These quenched fluorescent peptides are cleaved by the rhomboid, leading to the activation of a fluorophore. Fluorescence intensities were normalized to the DMSO control condition and plotted against log (inhibitor concentration) in GraphPad Prism software. The figure shows a representative dose-response curve for compound **3**. A nonlinear regression curve fit was used to determine apparent IC_{50} values as described.⁴⁰ Compounds **3** ($4.4 \pm 1.6 \mu\text{M}$) and **5** ($3.7 \pm 1.3 \mu\text{M}$) were equally potent while **11** ($48 \pm 14.1 \mu\text{M}$) displayed around 10-fold lower activity. Two independent IC_{50} determinations were performed for each compound, and IC_{50} values represent averages \pm SD.

Fig. 2. Molecular docking model of the rhomboid protease GlpG in complex with a benzoxazinone-based inhibitor. (A) Front view of GlpG docked with the 2-styryl substituted compound **3** (yellow) and the known rhomboid inhibitor CAPF (blue). The model indicated that both compounds occupied a similar space in the rhomboid pocket and were not exposed to the external surface. (B) Top view of compound **3** and CAPF in the GlpG pocket. (C) Ligand interaction map of compound **3** in the GlpG pocket. Significant interactions were observed with neighbouring residues His254 and Phe245. Green: hydrophobic, pink: polar, red: exposed.

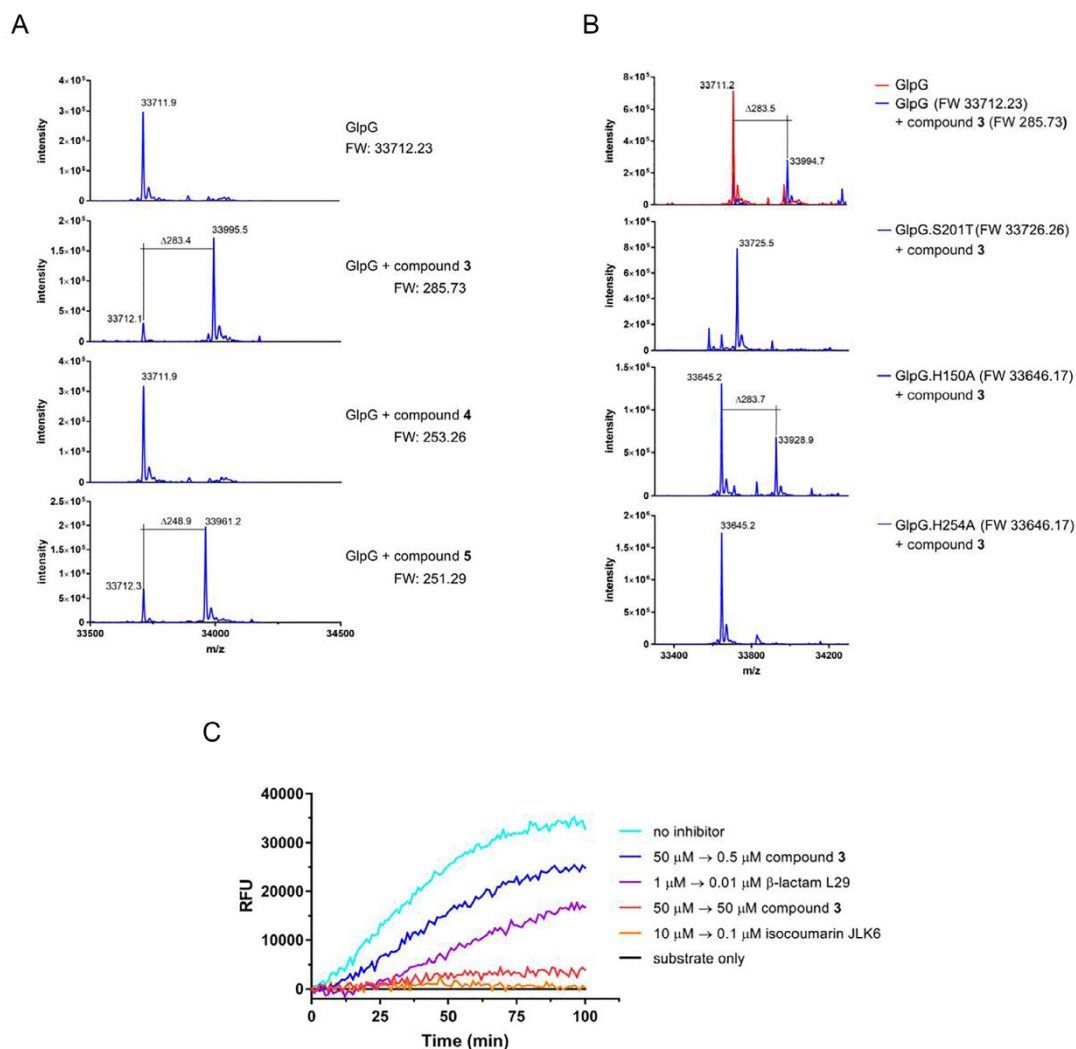


Fig. 4. Reaction mechanism of 2-styryl substituted benzoxazinone inhibitors. (A) Wild type GlpG was recombinantly expressed and purified as described previously,⁴⁰ reacted *in vitro* with compounds **3** or **5**, and binding to the rhomboid was examined by electrospray mass spectrometry. A second major peak was observed in the spectra corresponding to the inhibitor-bound enzyme with the expected mass shift compared to the free enzyme. In contrast, no binding was seen when the compounds were omitted (vehicle control) or when the rhomboid was reacted with the inactive benzoxazinone **4**. The results of one of two independent experiments are shown. (B) The 2-styryl substituted benzoxazinone **3** was incubated with different rhomboid mutants and binding was analyzed as in A. Compound **3** was able to bind to wild type GlpG and the H150A mutant with identical mass shifts between the inhibitor-bound and the free enzyme. However, no binding was observed to either the S201T or the H254A mutant. This indicated that Ser201 and proper activation of this catalytic serine by His254 was required for binding of the benzoxazinone to the rhomboid and formation of the O-acyl enzyme intermediate. The results of one of two independent experiments are shown. (C) Reversibility of the reaction was tested by the rapid dilution method. Compound **3** (50 μ M) was pre-incubated with the rhomboid GlpG for 1 h. Subsequently, the reaction mixture was rapidly diluted 100-fold with reaction buffer containing a fluorogenic rhomboid substrate (10 μ M). GlpG activity largely recovered over a time period of 120 min demonstrating reversibility of the inhibition mechanism. No recovery was observed when the reaction mixture was diluted into buffer containing substrate and 50 μ M compound **3**. The known reversible and irreversible rhomboid inhibitors β -lactam L29 and isocoumarin JLK-6 were used as controls and displayed the expected behaviour. The results of one of two independent experiments are shown.

caused significant enzyme inhibition at 10 μ M (Supplemental Fig. 3). Overall, these results indicated that 2-styryl substituted benzoxazinones might possess some selectivity for the rhomboid GlpG over other soluble serine proteases. Notably, in an accompanying paper, Yang, Verhelst and colleagues presented benzoxazinones with a 2-alkoxy substituent as rhomboid inhibitors.⁴⁶ These derivatives displayed slightly higher potency for GlpG but no apparent selectivity over bovine chymotrypsin and trypsin,

suggesting potential advantages of the docking approach to optimize the scaffold for this particular target.

The inhibitory activity of 2-styryl substituted benzoxazinones in the GlpG *in vitro* activity assays indicated that this scaffold contained active rhomboid inhibitors. However, to ensure that the reduced GlpG activity was not caused by non-specific effects and that the mechanism of inhibition conformed to the known reaction mechanism of benzoxazinones with soluble serine proteases²³, we

used mass spectrometry to study the residues in the rhomboid's active site that were reacting with the benzoxazinones. GlpG mutants were generated, expressed and purified, in which Ser201 was exchanged to threonine or in which His150 or His254 were exchanged to alanine. The recombinant proteins were then reacted *in vitro* with the 2-styryl substituted benzoxazinones **3** or **5**, and binding of the compounds to the rhomboid was examined by electrospray mass spectrometry. When the 2-styryl substituted benzoxazinones were reacted with wild type GlpG, two major peaks were observed in the spectra: the first smaller peak corresponding to the free unbound GlpG and the second major peak corresponding to the inhibitor-bound enzyme (Fig. 4A). Importantly, the shifts in the rhomboid's mass for compounds **3** (283.4 Da) and **5** (248.9 Da) were close to the theoretical mass differences of 285.73 and 251.29 Da that would occur when the benzoxazinones reacted with the enzyme according to the mechanism shown in Fig. 1A. In contrast, no binding and mass shift was seen when the compounds were omitted (vehicle control) or when the rhomboid was reacted with the inactive benzoxazinone compound **4**, supporting the specificity of the mass spectrometry analysis (Fig. 4A). Further analysis showed that compound **3** was not able to bind to either the S201T or the H254A mutant (Fig. 4B). This indicated that the benzoxazinone indeed reacted with the catalytic Ser201 to form an O-acyl enzyme intermediate (Fig. 4B). In the catalytic Ser201-His254 dyad, the histidine is required to properly activate the serine.¹⁶ Hence, in the H254A mutant the serine is likely not nucleophilic enough to react with the C-4 carbonyl of the benzoxazinone, which could explain that **3** failed to bind this mutant. However, mutating His150, which contributes to the oxyanion hole of the active site and is covalently modified by some isocoumarins^{19,47}, did not prevent binding of **3** to the rhomboid as expected (Fig. 4B).

Formation of the O-acyl enzyme intermediate should be reversible over time (Fig. 1A). To test this, we used the rapid dilution method^{48,49} and a fluorogenic peptidic rhomboid substrate.⁴⁰ Compound **3** (50 μ M) was pre-incubated with GlpG for 1 h. Subsequently, the reaction mixture was rapidly diluted 100-fold and the substrate KSp76 was added. Cleavage of this substrate by the rhomboid leads to the release of a quencher peptide and activation of a red fluorophore, which indicates recovery of enzyme activity. This showed that GlpG activity largely recovered over a time period of 100 min demonstrating reversibility of the inhibition mechanism (Fig. 4C). Similar results were obtained with the β -lactam L29, a known reversible inhibitor of rhomboids.²¹ In contrast, pre-incubation with the isocoumarin JLK-6, an irreversible inhibitor that forms a double-bonded end product with the rhomboid¹⁹, did not allow recovery of enzyme activity (Fig. 4C).

In summary, by combining molecular docking studies with enzyme activity assays, we found that 2-styryl substituted benzoxazin-4-ones are a novel template to generate inhibitors for rhomboid proteases. Mechanistic studies indicated that benzoxazinone inhibitors covalently modified the catalytic serine residue in the rhomboid active site, analogous to the reaction mechanism of benzoxazinones with soluble serine proteases. While the identified active 2-styryl substituted benzoxazinones at present have moderate potency, this new scaffold allows extensive structural variations with considerable potential to increase potency and specificity, towards the goal of improved small molecule inhibitors for rhomboid proteases.

Acknowledgements

We thank Professors Matthew Freeman (University of Oxford) and Ya Ha (Yale School of Medicine) for bacterial expression constructs. P.G. was supported by a scholarship of the iGRASPseed graduate

school of the Heinrich-Heine-University Duesseldorf. K.S. was a recipient of the Purkyne Fellowship of the Academy of Sciences of the Czech Republic and also acknowledges support from the Ministry of Education, Youth and Sports of the Czech Republic (projects no. LK11206 and LO1302), Marie Curie Career Integration Grant (project no. 304154), and the National Subvention for Development of Research Organisations (RVO: 61388963) to the Institute of Organic Chemistry and Biochemistry.

Conflicts of interest

None.

A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.bmcl.2018.02.017>.

References

- Freeman M. The rhomboid-like superfamily: molecular mechanisms and biological roles. *Annu Rev Cell Dev Biol.* 2014;30:235–254.
- Lee JR, Urban S, Garvey CF, Freeman M. Regulated intracellular ligand transport and proteolysis control EGF signal activation in *Drosophila*. *Cell.* 2001;107:161–171.
- Urban S, Lee JR, Freeman M. *Drosophila* rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell.* 2001;107:173–182.
- Clemmer KM, Sturgill GM, Veenstra A, Rather PN. Functional characterization of *Escherichia coli* GlpG and additional rhomboid proteins using an aarA mutant of *Providencia stuartii*. *J Bacteriol.* 2006;188:3415–3419.
- Chan EY, McQuibban GA. The mitochondrial rhomboid protease: its rise from obscurity to the pinnacle of disease-relevant genes. *Biochim Biophys Acta.* 2013;1828:2916–2925.
- McQuibban GA, Saurya S, Freeman M. Mitochondrial membrane remodeling regulated by a conserved rhomboid protease. *Nature.* 2003;423:537–541.
- Fleig L, Bergbold N, Sahasrabudhe P, Geiger B, Katak L, Lemberg MK. Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins. *Mol Cell.* 2012;47:558–569.
- Dusterhoft S, Kunzel U, Freeman M. Rhomboid proteases in human disease: Mechanisms and future prospects. *Biochim Biophys Acta.* 2017;1864:2200–2209.
- Song W, Liu W, Zhao H, et al. Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway. *Nat Commun.* 2015;6:8022.
- Civitarese AE, MacLean PS, Carling S, et al. Regulation of skeletal muscle oxidative capacity and insulin signaling by the mitochondrial rhomboid protease PARL. *Cell Metab.* 2010;11:412–426.
- Walder K, Kerr-Bayles L, Civitarese A, et al. The mitochondrial rhomboid protease PSARL is a new candidate gene for type 2 diabetes. *Diabetologia.* 2005;48:459–468.
- Parussini F, Tang Q, Moin SM, Mital J, Urban S, Ward GE. Intramembrane proteolysis of *Toxoplasma* apical membrane antigen 1 facilitates host-cell invasion but is dispensable for replication. *Proc Natl Acad Sci U S A.* 2012;109:7463–7468.
- Santos JM, Ferguson DJ, Blackman MJ, Soldati-Favre D. Intramembrane cleavage of AMA1 triggers *Toxoplasma* to switch from an invasive to a replicative mode. *Science.* 2011;331:473–477.
- Lemieux MJ, Fischer SJ, Cherney MM, Bateman KS, James MN. The crystal structure of the rhomboid peptidase from *Haemophilus influenzae* provides insight into intramembrane proteolysis. *Proc Natl Acad Sci U S A.* 2007;104:750–754.
- Wang Y, Zhang Y, Ha Y. Crystal structure of a rhomboid family intramembrane protease. *Nature.* 2006;444:179–180.
- Ha Y, Akiyama Y, Xue Y. Structure and mechanism of rhomboid protease. *J Biol Chem.* 2013;288:15430–15436.
- Strisovsky K. Rhomboid protease inhibitors: Emerging tools and future therapeutics. *Semin Cell Dev Biol.* 2016;60:52–62.
- Urban S, Wolfe MS. Reconstitution of intramembrane proteolysis *in vitro* reveals that pure rhomboid is sufficient for catalysis and specificity. *Proc Natl Acad Sci U S A.* 2005;102:1883–1888.
- Vosyka O, Vinothkumar KR, Wolf EV, Brouwer AJ, Liskamp RM, Verhelst SH. Activity-based probes for rhomboid proteases discovered in a mass spectrometry-based assay. *Proc Natl Acad Sci U S A.* 2013;110:2472–2477.
- Xue Y, Ha Y. Catalytic mechanism of rhomboid protease GlpG probed by 3,4-dichloroisocoumarin and diisopropyl fluorophosphonate. *J Biol Chem.* 2012;287:3099–3107.
- Pierrat OA, Strisovsky K, Christova Y, et al. Monocyclic β -lactams are selective, mechanism-based inhibitors of rhomboid intramembrane proteases. *ACS Chem Biol.* 2011;6:325–335.

22. Wolf EV, Zeissler A, Vosyka O, Zeiler E, Sieber S, Verhelst SH. A new class of rhomboid protease inhibitors discovered by activity-based fluorescence polarization. *PLoS One*. 2013;8:e72307.
23. Powers JC, Asgian JL, Ekiçi OD, James KE. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev*. 2002;102:4639–4750.
24. Powers JC, Kam CM, Narasimhan L, Oleksyszyn J, Hernandez MA, Ueda T. Mechanism-based isocoumarin inhibitors for serine proteases: use of active site structure and substrate specificity in inhibitor design. *J Cell Biochem*. 1989;39:33–46.
25. Gutschow M, Kuerschner L, Neumann U, et al. 2-(Diethylamino)thieno[3,4-b]oxazin-4-ones as stable inhibitors of human leukocyte elastase. *J Med Chem*. 1999;42:5437–5447.
26. Gutschow M, Neumann U. Inhibition of cathepsin G by 4H-3,1-benzoxazin-4-ones. *Bioorg Med Chem*. 1997;5:1935–1942.
27. Krantz A, Spencer RW, Tam TF, et al. Design and synthesis of 4H-3,1-benzoxazin-4-ones as potent alternate substrate inhibitors of human leukocyte elastase. *J Med Chem*. 1990;33:464–479.
28. Teshima T, Griffin JC, Powers JC. A new class of heterocyclic serine protease inhibitors. Inhibition of human leukocyte elastase, porcine pancreatic elastase, cathepsin G, and bovine chymotrypsin A alpha with substituted benzoxazinones, quinazolines, and anthranilates. *J Biol Chem*. 1982;257:5085–5091.
29. Gutschow M, Schlenk M, Gab J, et al. Benzothiazinones: a novel class of adenosine receptor antagonists structurally unrelated to xanthine and adenine derivatives. *J Med Chem*. 2012;55:3331–3341.
30. Zhong J, Groutas WC. Recent developments in the design of mechanism-based and alternate substrate inhibitors of serine proteases. *Curr Top Med Chem*. 2004;4:1203–1216.
31. Xue Y, Chowdhury S, Liu X, Akiyama Y, Ellman J, Ha Y. Conformational change in rhomboid protease GlpG induced by inhibitor binding to its S' subsites. *Biochemistry*. 2012;51:3723–3731.
32. MOE. <http://github.com/Yelp/MOE>, 2013.
33. For a detailed description of the docking procedures, please refer to the Supplementary Material.
34. Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today*. 2004;9:430–431.
35. Vinothkumar KR, Pierrat OA, Large JM, Freeman M. Structure of rhomboid protease in complex with beta-lactam inhibitors defines the S2' cavity. *Structure*. 2013;21:1051–1058.
36. For detailed synthesis information and the chemical data, please refer to the Supplementary Material.
37. Strisovsky K, Sharpe HJ, Freeman M. Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. *Mol Cell*. 2009;36:1048–1059.
38. Bacterial expression constructs for N-terminally His-tagged GlpG and C-terminally His-tagged Gurken chimeric substrate (MBP-GurkenTMD-Trx-His6) were kind gifts of Ya Ha (Yale School of Medicine) and Matthew Freeman (University of Oxford). Both enzyme and substrate were expressed in *E. coli* C43(DE3) cells. After induction with IPTG, GlpG was expressed overnight at 18 °C, while Gurken was expressed for 3 h at 37 °C. The bacterial cells were broken with a nitrogen cavitation bomb, membranes were isolated by differential centrifugation and proteins were solubilized with 1.5% dodecyl-beta-maltoside (DDM, Glycon Biochemicals). GlpG and Gurken were purified via immobilized metal-ion affinity chromatography using a Talon[®]Crude 1 ml HiTrap column and an Akta prime plus chromatography system (GE Healthcare). The recombinant proteins were further concentrated using 10 or 30 kDa molecular weight cut-off Amicon[®] Ultra Centrifugal filters (Millipore). Protein purity was estimated by Coomassie Brilliant Blue (Merck) stained SDS-PAGE. Enzyme activity assays were performed in 20 µl reaction buffer [50 mM HEPES-NaOH (pH 7.5); 5 mM EDTA; 0.4 M NaCl; 10% Glycerol (v/v); 0.05% DDM] containing GlpG (0.35 µM) and Gurken substrate (1.8 µM). Benzoxazinone inhibitors were pre-incubated with GlpG in reaction buffer for 30 min at 37 °C with gentle shaking. The substrate was added and the reaction was continued for another 90 min at 37 °C. Subsequently, the reaction mixture was separated by SDS-PAGE and stained with Coomassie Brilliant Blue. The N-terminal Gurken cleavage fragment was quantified using ImageJ, normalized to the DMSO vehicle control condition, and the values were plotted in GraphPad Prism software.
39. Goel P, Jumpertz T, Mikles DC, et al. Discovery and biological evaluation of potent and selective n-methylene saccharin-derived inhibitors for rhomboid intramembrane proteases. *Biochemistry*. 2017;56:6713–6725.
40. Ticha A, Stanchev S, Skerle J, et al. Sensitive versatile fluorogenic transmembrane peptide substrates for rhomboid intramembrane proteases. *J Biol Chem*. 2017;292:2703–2713.
41. Ticha A, Stanchev S, Vinothkumar KR, et al. General and modular strategy for designing potent, selective, and pharmacologically compliant inhibitors of rhomboid proteases. *Cell Chem Biol*. 2017;24:e1524.
42. Rhomboid inhibition assays were performed in black 96-well plates. Increasing concentrations of the 2-styryl substituted benzoxazinones **3**, **5**, or **11** were pre-incubated with 0.4 µM GlpG in reaction buffer [20 mM HEPES-NaOH (pH 7.4); 150 mM NaCl; and 0.05% (w/v) DDM] for 1 h. Subsequently, the reaction was started by the addition of 10 µM of the fluorogenic substrates KSp35 (for compounds **5** and **11**) or KSp76 (for compound **3**). Both substrates have comparable kinetic properties and are interchangeable for IC₅₀ determinations. IC₅₀ values were determined as described previously.⁴⁰
43. DelMar EG, Largman C, Brodrick JW, Geokas MC. A sensitive new substrate for chymotrypsin. *Anal Biochem*. 1979;99:316–320.
44. The ability of the benzoxazinone derivatives to inhibit the classical serine protease α-chymotrypsin was determined in an *in vitro* activity assay. The compounds were pre-incubated in concentrations up to 250 µM with 4 µM of bovine α-chymotrypsin (PanReac AppliChem) in reaction buffer [50 mM HEPES-NaOH (pH 7.5); 5 mM EDTA; 0.4 M NaCl; 10% Glycerol (v/v); 10% DMSO] for 30 min at 25 °C. Subsequently, 2.5 µl of this mixture were added to 100 µl reaction buffer containing 0.12 mM of the substrate N-Succinyl-Ala-Ala-Pro-Phe-p-nitroanilide (Suc-AAPF-pNA, Merck). Cleavage of the substrate by α-chymotrypsin resulted in 4-nitroaniline, which is yellow under alkaline conditions. A microtiter plate reader (Beckman Coulter) was used to record the increase in absorbance at 410 nm over 5 min at 25 °C. IC₅₀ values were determined by fitting a nonlinear regression curve to a plot of log (compound concentration) versus % α-chymotrypsin activity using GraphPad Prism software.
45. Enzyme *in vitro* activity assays were used to determine the inhibitory activity of the 2-styryl substituted benzoxazinones against the soluble serine proteases trypsin, neutrophil elastase and cathepsin G. In the trypsin assay, increasing concentrations of the benzoxazinones or DMSO vehicle were pre-incubated with 10 nM trypsin from bovine pancreas (Sigma-Aldrich) for 30 min at 30 °C in a total volume of 90 µl assay buffer (PBS). The reaction was started by the addition of 10 µl substrate solution (1 mM Z-Gly-Gly-Arg-aminomethylcoumarin in PBS, Bachem), and the release of the fluorescent cleavage product aminomethylcoumarin was monitored at 340 nm excitation/440 nm emission with a microplate reader (Tecan Safire). In the cathepsin G assay, increasing concentrations of the compounds or DMSO vehicle were pre-incubated with 200 nM human neutrophil cathepsin G (Enzo Life Sciences) for 30 min at 37 °C in a total volume of 90 µl assay buffer [100 mM HEPES; 500 mM NaCl (pH 7.4)]. The reaction was started by the addition of 10 µl substrate solution (2 mM N-Succinyl-Ala-Ala-Pro-Phe-p-nitroanilide in assay buffer, Bachem). Enzymatic cleavage of the substrate resulted in the release of the yellow colored p-nitroaniline, and absorbance was recorded at 410 nm. To measure the activity of the benzoxazinones against human neutrophil elastase, the Neutrophil Elastase Colorimetric Drug Discovery Kit (Enzo Life Sciences) was used according to the manufacturer's protocol. In brief, increasing concentrations of the benzoxazinones or DMSO vehicle were pre-incubated with 0.22 mU neutrophil elastase for 20 min at 37 °C in a total volume of 95 µl assay buffer [100 mM HEPES (pH 7.25); 500 mM NaCl; 0.05% Tween-20]. The reaction was started by the addition of 5 µl substrate solution (2 mM N-Methoxysuccinyl-Ala-Ala-Pro-Val-p-nitroanilide in assay buffer), and the formation of p-nitroaniline was recorded as described for the cathepsin G assay. In all three assays, changes in absorbance/fluorescence were measured at 1 min time intervals for 5 min, and the initial reaction slopes were used to determine the reaction velocity (V). To calculate the % remaining protease activity, the formula (V inhibitor / V DMSO control) × 100 was used.
46. Yang J, Barniol-Xicota M, Nguyen MTN, Ticha A, Strisovsky K, Verhelst SHL. Benzoxazin-4-ones as novel, easily accessible inhibitors for rhomboid proteases. *Bioorg Med Chem Lett*. 2018;28:1423–1427.
47. Vinothkumar KR, Strisovsky K, Andreeva A, Christova Y, Verhelst S, Freeman M. The structural basis for catalysis and substrate specificity of a rhomboid protease. *EMBO J*. 2010;29:3797–3809.
48. Harper JW, Hemmi K, Powers JC. Reaction of serine proteases with substituted isocoumarins: discovery of 3,4-dichloroisocoumarin, a new general mechanism based serine protease inhibitor. *Biochemistry*. 1985;24:1831–1841.
49. Rhomboid activity recovery assays were performed in black 96-well plates. Compound **3** (100 µM), β-lactam L29 (1 µM) or isocoumarin JLK-6 (10 µM) were pre-incubated for 1 h with 0.4 µM GlpG in reaction buffer [20 mM HEPES-NaOH (pH 7.4); 150 mM NaCl; 0.05% (w/v) DDM]. Afterwards, this solution was rapidly diluted 100-fold into reaction buffer containing 10 µM of the fluorogenic substrate KSp76. This quenched fluorescent peptide is cleaved by the rhomboid leading to activation of a red fluorophore. Substrate cleavage and enzyme activity recovery were monitored continuously by measuring fluorescence over a time period of 120 min with excitation at 553 nm and emission at 583 nm using a microplate reader (Tecan Infinite M1000).

**Univerzita Karlova v Praze, 1. lékařská fakulta
Kateřinská 32, Praha 2**

**Prohlášení zájemce o nahlédnutí
do závěrečné práce absolventa studijního programu
uskutečňovaného na 1. lékařské fakultě Univerzity Karlovy v Praze**

Jsem si vědom/a, že závěrečná práce je autorským dílem a že informace získané nahlédnutím do zpřístupněné závěrečné práce nemohou být použity k výdělečným účelům, ani nemohou být vydávány za studijní, vědeckou nebo jinou tvůrčí činnost jiné osoby než autora.

Byl/a jsem seznámen/a se skutečností, že si mohu pořizovat výpisy, opisy nebo kopie závěrečné práce, jsem však povinen/a s nimi nakládat jako s autorským dílem a zachovávat pravidla uvedená v předchozím odstavci.

Příjmení, jméno (hůlkovým písmem)	Číslo dokladu totožnosti vypůjčitele (např. OP, cestovní pas)	Signatura závěrečné práce	Datum	Podpis