Structural and Biophysical Insights into Autoproteolytic and Inhibitory Mechanisms Regulating Tryptic Activity in Pancreatitis

I n a u g u r a l d i s s e r t a t i o n

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von

Felix Andreas Nagel

Greifswald, Juli 2022

Dekan: Prof. Dr. Gerald Kerth

1. Gutachterin: Prof. Dr. Mihaela Delcea

2. Gutachter: Prof. Dr. Adnane Achour

3. Gutachter: Prof. Dr. José Toca-Herrera

Tag der Promotion: 26.10.2022

Table of Contents

List of Abbreviations

ACh	Acetylcholine
AP	Acute pancreatitis
ATF6	Activating trancription factor 6
ATP	Adenosinetriphosphate
CCK	Cholecystokinin
CD	Circular dichroism
CFTR	Cystic fibrosis transmembrane conductance regulator
COPI	Coat protein complex I
COPII	Coat protein complex II
CP	Chronic pancreatitis
CTRC	Chymotrypsin C
E. coli	*Escherichia coli*
ER	Endoplasmic reticulum
ERES	Endoplasmic reticulum exit sites
ERGIC	Endoplasmic reticulum - Golgi intermediate compartment
IP$_3$	Inositol 1,4,5-trisphosphate
IRE1	Inositol-requiring enzyme
ITC	Isothermal titration calorimetry
k_a	Association rate
k_{cat}	Conversion rate
K_d	Dissociation constant
k_d	Dissociation rate
K_i	Inhibition constant
K_m	Michaelis-Menten constant
L-BAPA	N$_\alpha$-Benzoyl-L-arginine 4-nitroanilide hydrochloride
M6P	Mannose-6-phosphate
MPR	Mannose-6-phosphate receptor
pdb	Protein data bank
PDI	Protein disulfide bond isomerase
PP	Pancreatic polypeptide
PRSS1	Human cationic trypsin gene
PRSS2	Human anionic trypsin gene
PRSS3	Human mesotrypsin gene
RAP	Recurrent acute pancreatitis
SAPE	Sentinel acute pancreatitis event
SDS PAGE	Sodium dodecyl polyacrylamide gel electrophoresis
SNAP	N-ethylmaleimide-sensitive fusion attachment protein
SNARE	Soluble N-ethylmaleimide sensitive receptor
SPINK1	Serine protease inhibitor Kazal-type 1
SPR	Surface plasmon resonance
TRY1	Human cationic trypsin

TRY2	Human anionic trypsin
TRY3	Human mesotrypsin
UPR	Unfolded protein response
VAMP	Vesicle associated membrane protein
V-ATPase	Vacuolar-type H^+-adenosine 5'-triphoshatase
ZG	Zymogen granule

Common one- and three-letter codes for amino acids and the International System of Units (SI-units), as well as SI-derived units, are used throughout this thesis. The Schechter and Berger nomenclature for inhibitor residues and protease binding pockets is used [1].

Scope and Outline

Pancreatitis is an inflammatory disorder of the pancreas with a mortality rate of 5% and severe negative effects on the quality of life. Of all non-malignant gastrointestinal diseases, it is the most common reason for hospitalization. Pancreatitis is a disease of multiple etiologies with different underlying pathomechanisms. Due to the diversity of mechanisms by which homeostasis within the exocrine pancreas can be disrupted, finding appropriate therapeutic approaches is challenging. Current treatment options are inadequate and are mostly limited to supportive treatment like fluid administration, bowel rest, antibiotics and pain control. Although significant advancements have been achieved in recent decades, the mortality rate for pancreatitis has not decreased. Furthermore, progress is slow due to limited patient sample availability and lack of an appropriate cell model. Taking samples from a human pancreas is typically avoided, because damaging the pancreatic tissue can itself induce pancreatitis. Additionally, while it is possible to keep individual acini in culture, it is not possible to grow pancreatic acinar cells. Thus, less appropriate cell models, often derived from pancreatic cancer samples, have to be used. The most common animal model for pancreatitis is mice, with caerulein administration being the most common method of inducing pancreatitis. However, the use of animal models has significant drawbacks, as they are time-consuming, costly, and pose ethical questions. Furthermore, exposing the pancreas to appropriate stimuli in animal models is difficult. For example, alcohol is the leading cause of pancreatitis in humans, but is typically avoided by animals. Thus, alcohol feeding methods had to be developed to overcome the natural aversion of rodents to alcohol. Results obtained from animal models are also often not transferable into clinical trials and outcomes in humans remain largely unpredictable. Due to the lack of experimental models, our understanding of this highly complex disease is still limited and significant progress is required for the development of effective therapy options.

In this dissertation recombinantly expressed trypsin isoforms and variants of the serine protease inhibitor Kazal-type 1 (SPINK1) inhibitor are used to investigate mechanisms, by which tryptic activity is regulated in pancreatic acinar cells. With premature tryptic activity in the exocrine pancreas being the common focal point of most etiologies connected to pancreatitis, trypsin represents by far the most promising target for treating pancreatitis. Understanding the mechanisms by which the pancreas protects itself and rationalizing mutations that can undermine these protective mechanisms, are important steps towards developing effective therapies.

In this work, a combination of biochemical, biophysical, computational and structural approaches was utilized to unravel molecular mechanisms governing the regulation of tryptic activity.

Article I   **Structural Basis of the Pancreatitis-Associated Autoproteolytic Failsafe Mechanism in Human Anionic Trypsin**

Nagel, F.; Susemihl, A.; Geist, N.; Möhlis, K.; Palm, G.J.; Lammers, M.; Delcea, M. *J. Inflamm. Res.* 2022, *15*, 3633–3642

Article I presents insights into the molecular structure of human anionic trypsin (TRY2), a mostly neglected isoform in pancreatitis research. It is the first crystal structure of human anionic trypsin and the first human trypsin to crystallize as an uninhibited, wild type protein. The structure was solved in an autoproteolytic, dimeric state representative of the autoproteolytic failsafe mechanism. Based on the crystal structure, a model for cationic trypsin (TRY1) was built and the molecular basis for the pathogenicity of the R122H mutation in trypsin was established.

Article II   **A Hypothesized Mechanism for Chronic Pancreatitis Caused by the N34S Mutation of Serine Protease Inhibitor Kazal-Type 1 Based on Conformational Studies**

Kulke, M.; Nagel, F.; Schulig, L.; Geist, N.; Gabor, M.; Mayerle, J.; Lerch, M.M.; Link, A.; Delcea, M. *J. Inflamm. Res. 2021, 14, 2111–2119.*

In addition to self-regulation of the two major trypsin isoforms by autoproteolysis they are inhibited by a trypsin inhibitor called serine protease inhibitor Kazal-type 1 (SPINK1). **Article II** uses molecular dynamics and replica exchange simulations to analyze the conformational space of SPINK1-TRY1 complexes. Focus is a well-known, pancreatitis-associated N34S mutation in SPINK1. Minor differences in the complex structures and conformational flexibility of both SPINK1 variants were detected and hypothesized to influence the proteolytic stability of the inhibitors.

Article III   **Structural and Biophysical Insights into SPINK1 Bound to Human Cationic Trypsin**

Nagel, F.; Palm, G.J.; Geist, N.; McDonnell, T.C.R.; Susemihl, A.; Girbardt, B.; Mayerle, J.; Lerch, M.M.; Lammers, M.; Delcea, M. *Int. J. Mol. Sci.* 2022, *23*.

Article III elucidates the inhibitory mechanism of SPINK1 and provides a comprehensive kinetic, thermodynamic and structural understanding of the interaction with cationic trypsin. Moreover, the effect of the N34S amino acid substitution is analyzed. Crystal structures of SPINK1 WT and N34S in complex with cationic trypsin are almost indistinguishable and refute previously hypothesized pathomechanisms. An unexpected rotamer conformation of His63 in the catalytic triad was observed in SPINK1 inhibited cationic trypsin, which is speculated to contribute to the proteolytic stability of the inhibitor.

Article IV   **Identification of Kazal-inhibitor scaffolds with identical canonical binding loops and their effects on binding properties**

Nagel, F.; Susemihl, A.; Eulberg, T.E.; Delcea, M. *J. Biol. Chem.*, submitted.

Finally, **Article IV** describes a more potent SPINK1 variant and the influence of the Kazal scaffold supporting the canonical binding loop on binding properties is investigated. A T40P mutation in human SPINK1 selectively improved binding to human trypsin isoforms, while showing negligible effects on binding properties with bovine or porcine trypsin.

Introduction

The pancreas

The pancreas is part of the gastrointestinal system and simultaneously functions as an endo- and exocrine gland. It weighs about 100 g and is 14-23 cm long. Its location is in the upper abdomen behind the stomach with the broader pancreatic head lying within the loop of the duodenum and the tail of the pancreas being near the hilum of the spleen [2]. The main pancreatic duct (Wirsung's duct) branches through the pancreas and is joined by the common bile duct near the duodenum (Figure 1). Cells of the pancreas can be categorized according to their involvement in either the exo- or endocrine system. The exocrine part of the pancreas produces digestive enzymes that are secreted into the duodenum via the pancreatic duct. It includes acinar cells, as well as duct cells and the associated connective tissues, nerves, and other vessels. All components of the exocrine pancreas make up 95-98% of the total pancreas mass [3]. The remaining 2% form the endocrine pancreas and are mostly organized in pancreatic islets (islets of Langerhans). On average, the size of pancreatic islets is in the range of 50-250 μm with each islet consisting of about 1,500 individual cells [4]. The endocrine system of the pancreas is critical for maintaining glucose homeostasis, which is mediated by secretion of pancreatic hormones. Although seemingly independent, the exo- and endocrine systems are functionally interconnected, as reflected in the close proximity of acinar and pancreatic islets. Exposing acini to pancreatic hormones like insulin leads to potentiated amylase secretion and subsequent increased polysaccharide digestion [5,6]. This functional interplay is referred to as the "islet-acinar" axis and is still an active field of research with the effects of individual pancreatic hormones on the exocrine system highly debated [7].
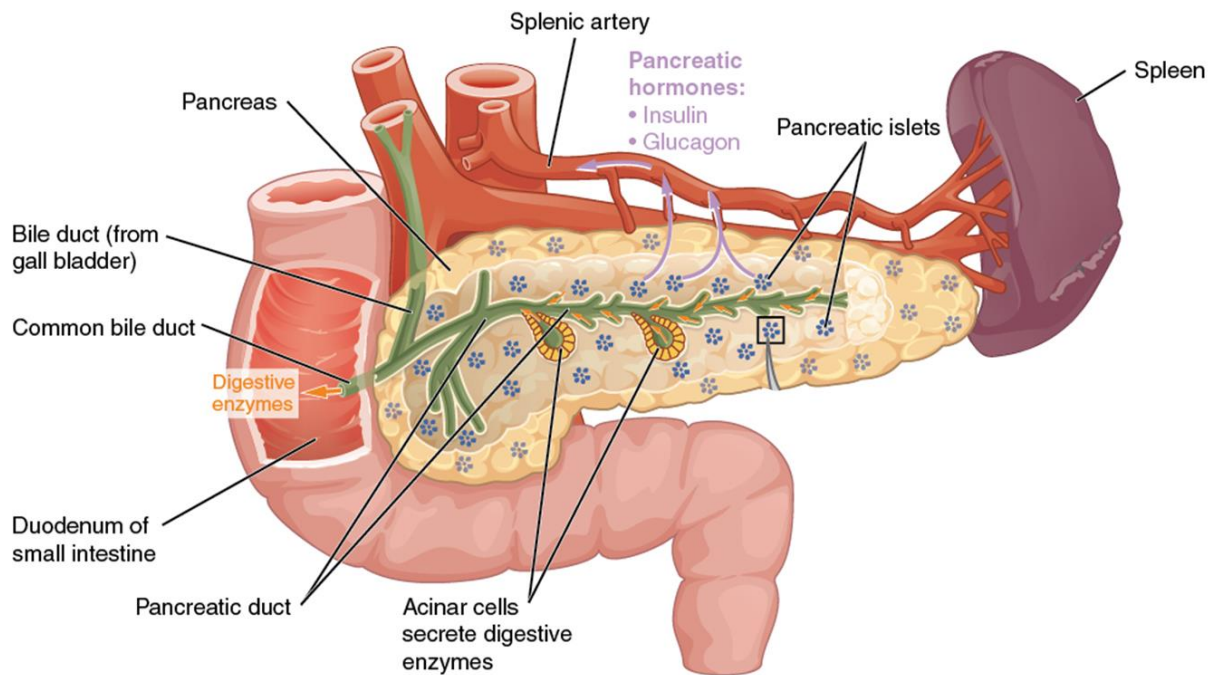
Figure 1: Schematic representation of the pancreas and its relative positioning between spleen and duodenum. The pancreatic duct system is shown in green and acini are shown in yellow. Pancreatic islets are depicted in blue. The figure was adapted from *Betts et al.* [8].

The endocrine pancreas

Pancreatic islets make up about 2% of the pancreatic mass [3]. The existence of these islets was first described in 1869 by Paul Langerhans and their involvement in glucose homeostasis was demonstrated 21 years later by Joseph von Mehring and colleagues [9]. Pancreatic islets are comprised of α, β, δ, ε, and PP (pancreatic polypeptide) cells. These hormone-secreting cells store pancreatic hormones in numerous secretory granules, which can be secreted upon receiving the appropriate stimulus. The endocrine cells are labeled in order of their discovery. Glucagon-producing α cells were discovered first and make up approximately 35% of the islet mass [10,11]. The primary function of glucagon is stimulating glycogenolysis and gluconeogenesis, thereby preventing hypoglycemia [12]. The most common endocrine cell type is insulin-producing β cells. They comprise 50-70% of the islet mass and, in addition to insulin, also produce amylin [13]. Insulin release after food uptake enables replenishment of energy stores by glycogenesis and lipogenesis, promotes the uptake of glucose by fat and skeletal muscle and induces liver glucose uptake by glycogen deposition and prevention of glycogenolysis and gluconeogenesis [14–19]. Insulin is arguably best known for its role in diabetes, caused by insulin insufficiency due to loss of β cells by an autoimmune response (type 1 diabetes) or by insulin resistance (type 2 diabetes), which is commonly caused by

obesity and insufficient exercise [20,21]. A less well-known type of diabetes is type 3c. It is characterized by insufficient endocrine function caused by exocrine damage, frequently due to acute or chronic pancreatitis [22].

Less than 10% of the islet mass is made up of somatostatin-producing δ cells [23]. Somatostatin acts as an inhibitor of both endocrine and gastrointestinal hormones. PP cells produce the pancreatic polypeptide, acting on the endocrine as well as the exocrine pancreas [24,25]. Ghrelin is produced predominantly in the stomach, but also in pancreatic ε cells, where it suppresses insulin release [26].

Through the islet-acinar axis, acinar cells of the exocrine pancreas are exposed to endogenous pancreatic hormones, highlighting their functional relationship [5]. Insulin can bind to insulin receptors on acinar cells, which induces the secretion of amylase [6]. In addition, somatostatin and PP are considered to inhibit pancreatic exocrine secretion [25,27,28]. The effect of glucagon on exocrine functions of the pancreas is still debated and both inhibitory and stimulatory mechanisms have been proposed [7].

The exocrine pancreas

Cells of the exocrine pancreas make up 95-98% of the entire pancreatic mass [3]. They consist of acinar cells, which arrange in so-called "acini" around the pancreatic duct. The duct system is comprised of ductal cells forming the main pancreatic duct and interlobular ducts that drain into the main duct throughout the pancreas (Figure 2). Digestive enzymes released from acinar cells are secreted through acinar tubules into a bicarbonate-rich solution within interlobular ducts. From here they are transported to the main pancreatic duct and finally to the duodenum. Most of the digestive enzymes produced in acinar cells are capable of damaging the surrounding pancreatic tissue. Hence, the integrity of the ductal system is crucial for preventing these enzymes from entering the interstitial space. Leakage from the duct system is mostly prevented by tight junctions between duct cells and acinar cells [29]. It is estimated that up to 90% of protein expression in the exocrine pancreas is devoted to synthesizing about 20 proteins [30]. Due to their destructive nature, every step from synthesis to secretion of digestive enzymes is tightly regulated within acinar cells. Imbalances to this highly intricate system can lead to severe damage of the pancreas. Thus, one of the first modifications to secretory proteins is the cleavage of the signal peptide by signal peptidase [31,32]. This endoplasmic reticulum (ER) specific protease cleaves the signal peptide and traps the protein in the secretory pathway.

Endoplasmic reticulum

Unlike cytosolic proteins, proteins entering the secretory pathway contain an abundance of disulfide bonds. The redox environment of the ER is more reducing than the surrounding cytosol or other cellular compartments, making it suitable for the correct formation of these disulfide bonds. Furthermore, it is known that environmental factors like ethanol feeding in mice leads to increased oxidation in the ER. This suggests an increase in incorrectly formed disulfide bonds, resulting in misfolding, ER-stress and eventually triggering of the unfolded protein response (UPR) system when the ER-associated degradation pathway is overloaded [33–38]. The UPR system is pivotal in sustaining the high expression levels encountered in acinar cells and is regulated by the ER transmembrane proteins inositol-requiring enzyme (IRE1), protein kinase R-like ER kinase and activating transcription factor 6 (ATF6). These proteins act as sensors for the unfolded protein load within the ER lumen and can initiate response mechanisms to reduce ER stress if appropriate [39,40]. Additionally, a protein disulfide bond isomerase (PDI) specifically expressed in acinar cells termed PDIp, further reflects the importance of correctly formed disulfide bonds. PDIp interacts with digestive enzymes and prevents their aggregation [41–43].

Golgi complex

After correct folding, secretory proteins at specialized ER exit sites (ERES) are directed into the ER-Golgi intermediate compartment (ERGIC) within coat protein complex II (COPII)-coated vesicles [44,45]. From here, they are concentrated and traverse through the *cis-, medial-,* and *trans*-Golgi, where they might undergo further post-translational modifications. Concentration of secretory proteins in the ERGIC was demonstrated for chymotrypsin and amylase, which are non-selectively transported from the ER but excluded from COPI vesicles that transport proteins back to the ER [46,47]. At the *cis*-Golgi site, mannose-6-phosphate (M6P) residues are added to proteins intended for the endolysosomal pathway [48]. M6P residues interact with M6P receptors (MPRs) that are primarily located on endo- and lysosomes [49]. At the *trans*-Golgi site, condensing vacuoles are formed and filled with secretory proteins that do not carry M6P residues. Mechanisms by which these secretory proteins are sorted into condensing vacuoles are still an active field of research, but two hypotheses have been formulated [50,51]. The sorting-for-entry hypothesis speculates that, similar to the M6P sorting mechanism, other post-translational modifications, such as sulfation or O-glycosylation, are specifically recognized by respective sorting receptors. However, the existence of these receptors remains to be demonstrated. Both sulfation and O-glycosylation are modifications that occur in the Golgi apparatus and are found in almost all digestive enzymes [52].

The condensing vacuoles are acidic (pH about 5.5), which is due to the acidic nature of the Golgi complex and is further amplified by vacuolar-type H^+-ATPase (V-ATPase) in the

membrane of these vacuoles that acts to acidify the vesicular lumen [53]. During maturation and transport to the apical pole of acinar cells the condensing vacuole becomes smaller. Clathrin-coated vesicles mediate the removal of membrane and small amounts of digestive enzymes. Additionally, water is transported out of these granules resulting in a mature zymogen granule (ZG) with extremely dense protein content [54]. Storage of inactive precursor digestive enzymes called zymogens is aided by concentration of the protein content within ZGs. High protein concentrations cause partial aggregation with some zymogens even existing in a semi-crystalline state that aids with inhibition of these zymogens until they are released [55–57]. Evidence for the limited aggregation within ZGs also gave rise to the sorting-by-retention hypothesis. Soluble, non-zymogen proteins are successively removed during the maturation of the ZG resulting in a granule with very high zymogen content [50]. While condensing vacuoles are initially acidic (pH ~5.5), they lose most of their acidity during maturation [58]. Loss of acidity is speculated to be caused by inactivation of V-ATPase or removal of H^+-pump subunits, both mechanisms remain to be shown. The behavior is anomalous, especially considering that other secretory cells typically show increased acidity during maturation of the secretory vesicles [59,60]. The physiological meaning behind the loss of acidity remains an enigma and has not yet been properly addressed.

Secretion

The vast majority of ZGs secrete their cargo at the apical membrane by exocytosis. The secretion mechanism of pancreatic acinar cells is well studied, historically due to George Palade's work (Nobel Prize for Medicine or Physiology, 1974) who used and established the pancreatic acinar cell as a basic model for secretion [61]. Exocytosis is achieved through soluble N-ethylmaleimide sensitive receptor (SNARE) proteins. In pancreatic acinar cells, ZGs enriched with either the SNARE protein vesicle associated membrane protein 2 (VAMP2) or VAMP8 can be found [62]. VAMP2 interacts with syntaxin 2 and the N-ethylmaleimide-sensitive fusion attachment protein 23 (SNAP23) present in the apical membrane of acinar cells. ZGs enriched with VAMP8 interact with syntaxin 4 and SNAP23 present in the basolateral membrane, which is hence considered to be a pathological release of digestive enzymes [63–65]. During secretion ZG-ZG compound exocytosis is often observed and is mediated by VAMP8, syntaxin 3 and SNAP23.

Fusion of ZGs with the apical plasma membrane is triggered by an increase of intracellular calcium near the apical pole of the cell, often induced by secretagogues [66]. Secretagogues include cholecystokinin (CCK), acetylcholine (ACh) and bombesin. These secretagogues initiate inositol 1,4,5-trisphosphate (IP_3) signaling through interaction with their respective G-protein coupled receptors [67]. Subsequently, IP_3 triggers Ca^{2+}-release ultimately leading to the activation of the ZG secretion machinery. During exocytosis, an F-actin coat forms around

the fusion of the ZG membrane and apical plasma membrane, keeping the formed pore in an energetically unfavorable "open" state that allows stored zymogens to resolubilize and diffuse out of the ZG [68,69]. Exocytosis in pancreatic acinar cells differs greatly from, e.g. neuronal synapses, in that the formed fusion pores are open for up to 20 min and compound exocytosis is rapid with >70% of ZGs being fused within 1 min at the apical pole [70,71]. The underlying reason for the different secretion mechanisms is likely caused by the cell geometry of the pancreatic acinar cell, with the apical plasma membrane constituting only 10% of the total cell surface [70]. Misdirection of the secretion machinery to the basolateral membrane is a major mechanism for the development of pancreatitis and can be stimulated by e.g. alcohol [72,73].

Meal response

Pancreatic secretion is tightly regulated and occurs as a response to food ingestion. The pancreatic meal response can be divided into four phases: *cephalic, gastric, intestinal* and *absorbed nutrient*. The *cephalic phase* is under control of the vagus nerve and describes the physiological response to sensory inputs like smell and taste and as a result, the anticipation of food [74,75]. Interaction of, e.g. fatty acids, with oral receptors also induce and amplify the *cephalic phase* [76]. During this phase, approximately 20-25% of secretion occurs [77–79]. The exocrine secretion response was evaluated in sham fed dogs and humans, where the response in humans lasted for one hour, while in dogs the *cephalic phase* lasted more than four hours [80,81]. The *gastric phase* accounts for about 10% of exocrine secretion and is a result of complex neural reflexes and hormone release upon sensing food in the stomach [78,82,83]. The majority of enzyme secretion is regulated during the *intestinal phase* and accounts for 50-80% of enzyme release. It is triggered particularly by amino acids tryptophan and phenylalanine, but also monoglycerides and peptides [84–88]. Unlike the *cephalic* and *gastric phase*, secretion in the *intestinal phase* is accompanied by the release of high volumes of bicarbonate to neutralize the gastric chyme in the duodenum. Release of bicarbonate indicates concerted secretion in both acinar and duct cells. The response is primarily initiated by the hormones secretin, which is released from S cells after entry of the low pH gastric chyme into the duodenum and CCK [89–92]. The *absorbed nutrient phase* is characterized by a release of exocrine enzymes through nutrients absorbed in blood. However, little is known about the precise mechanisms and the extent to which absorbed nutrients contribute to the release of digestive enzymes [93,94].

Because digestive enzyme secretion is highly regulated by food ingestion and is required to be fast, it is clear that this process cannot be regulated on a transcriptional or translational level. Hence, just like endocrine hormones of the pancreas, digestive enzymes need to be stored for several hours until they are released in response to the next meal. However, digestive and especially proteolytic enzymes of the pancreas are highly promiscuous and

capable of autodigestion and are thus, inherently unstable. This makes storing these enzymes in ZGs challenging and several mechanisms have evolved to inhibit enzymatic activity. The protein composition of ZGs was the subject of many proteomic studies and key digestive enzymes were identified in rat ZGs and their human homologues are summarized in Table 1 [95–97].

Table 1: Representative digestive enzymes found in ZGs. Theoretical molecular weights and isoelectric point (pI) values are calculated without signal peptide [98].

Enzyme	Short name	Gene name	MW (kDa)	pI
Alpha-amylase 1A	AMY1A	*AMY1A*	55.9	6.34
Anionic trypsinogen	TRY2	*PRSS2*	24.9	4.78
Bile salt-activated lipase	BAL	*CEL*	77.2	5.08
Cationic trypsinogen	TRY1	*PRSS1*	25.0	6.1
Chymotrypsinogen B1-2	CTRB1	*CTRB1*	25.8	6.95
Chymotrypsinogen C	CTRC	*CTRC*	27.9	7.68
Pancreatic lipase	PL	*PNLIP*	49.5	6.29
Procarboxypeptidase A1	CPA1	*CPA1*	45.5	5.38
Procarboxypeptidase A2	CPA2	*CPA2*	44.9	5.52
Procolipase	CLPS	*CLPS*	10.1	7.68
Proelastase 1	CELA1	*CELA1*	26.9	8.60

Zymogen storage and regulation

In order to prevent autodigestion, most of the enzymes shown in Table 1 are expressed as inactive precursor proteins called zymogens. These zymogens are activated through cleavage of an N-terminal propeptide that inhibits the enzyme, and at the same time aids with folding of the enzyme in the ER [99]. Trypsin, which is the active form of trypsinogen, represents the most important enzyme in maintaining the inactive state of these enzymes, because it is capable of activating all other zymogens (Figure 2). Trypsinogen constitutes about 20% of the total protein content in the pancreatic juice and is a key player during digestion [100]. Due to the high amounts of trypsinogen in the pancreas and its potential to activate all digestive enzymes within, it is pivotal to tightly regulate trypsin activity in the pancreas. Hence, under physiological conditions, trypsinogen is only activated once it enters the duodenum by an enzyme called enterokinase. Enterokinase is spatially separated from trypsinogen, which prevents unwanted activation of trypsin within pancreatic acinar cells [101]. Despite its name entero*kinase*, the enzyme is actually a transmembrane protease and is more accurately called entero*peptidase* in more recent literature.

Although trypsinogen is spatially separated from its activator enterokinase, trypsinogen possesses weak autoactivation activity [102]. The propeptide has evolved to feature four negatively charged aspartate residues in front of the positively charged lysine residue, after which the propeptide can be cleaved by either enterokinase or, albeit to much lesser extent, trypsin [103]. In addition, autoactivation of trypsin within ZGs of acinar cells is further reduced by the previously mentioned partial aggregation of zymogens, due to the very high protein concentrations within the mature ZGs [55–57].

Despite all these mechanisms that regulate tryptic activity within acinar cells, small amounts of trypsinogen are still able to autoactivate, which is sufficient to start a seemingly unstoppable auto-activation cascade. Thus, a trypsin inhibitor called serine protease inhibitor Kazal-type 1 (SPINK1) expressed in acinar cells is sorted into ZGs together with trypsinogen. Here, SPINK1 acts as a first line of defense against prematurely activated trypsin and essentially fulfills the role of the now cleaved propeptide by inhibiting activated trypsin [104]. Trypsin inhibition by SPINK1 is reversible and follows the standard mechanism, also known as Laskowski mechanism [105]. SPINK1 binds in a substrate-like manner and engages in very tight binding interactions with trypsin. Although trypsin inhibition is beneficial within ZGs, once released into the duodenum, trypsin inhibition by SPINK1 is unwanted. This apparent dilemma between protecting the pancreas through inhibition and maintaining enough tryptic activity for digestion is solved by different expression ratios between trypsin and SPINK1. On average, the molar ratio between trypsin and SPINK1 in acinar cells is five to one. Thus, SPINK1 is theoretically capable of inhibiting a maximum of 20% of the tryptic activity [106,107].
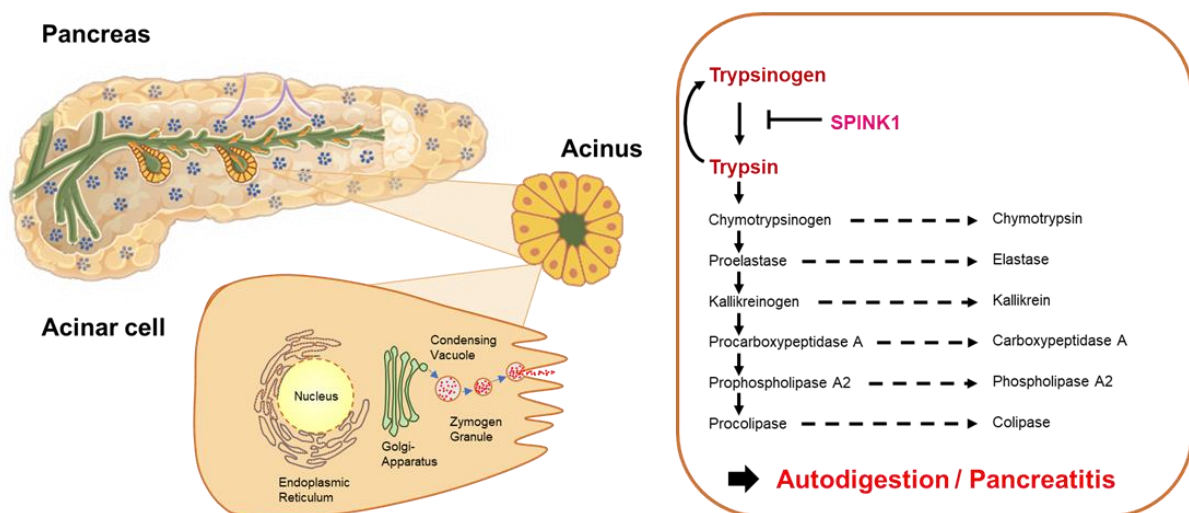

Figure 2: Schematic representation of the exocrine pancreas. Acini surrounding the pancreatic duct and a schematic representation of a pancreatic acinar cell (*left panel*). Protein composition of zymogen granules and their functional relationship in the auto-activation cascade (*right panel*). The image depicting the pancreas was adapted from *Betts et al.* [8].

In case premature trypsin activity within acinar cells exceeds the inhibitory potential of SPINK1, several additional failsafe mechanisms have evolved. One of these mechanisms is the chymotrypsin C (CTRC) dependent degradation of trypsin. Progression through the autoactivation cascade eventually results in the activation of chymotrypsinogen C [108,109]. On the one hand, CTRC can enhance trypsinogen autoactivation rates by modifications of the propeptide. On the other hand, active CTRC results in proteolytic degradation of trypsin, which is considered to be more physiologically relevant than enhancing the autoactivation rate. Thus, activation of CTRC within ZGs results in degradation and inactivation of active trypsin species and aids in protecting the pancreas from autodigestion. However, while this effect was demonstrated for both cationic trypsin (TRY1) and anionic trypsin (TRY2), anionic trypsin was found to be under much tighter control of CTRC, which is explained through an additional cleavage site as well as the lack of a stabilizing Cys139-Cys206 disulfide bond [110].

Additionally, active trypsin can undergo autoproteolysis, effectively self-regulating itself in the absence of other substrates. Two major autolytic sites were identified in TRY1 and TRY2. Cleavage of the Arg122-Val123 bond is considered the first site to be cleaved by autoproteolysis and is essential to the autoproteolytic failsafe mechanism. However, while cleavage of this bond is rapid, it does not inactivate the enzyme. Only in combination with the second autolytic cleavage site Lys193-Asp194 is trypsin inactivated and ultimately degraded [110].

In conclusion, pancreatic acinar cells are responsible for safely synthesizing, storing and secreting highly promiscuous and potentially destructive digestive enzymes. It is obvious how even small imbalances in this tightly regulated process can lead to severe damage of the cells and the surrounding tissue, which is often the cause of many diseases of the pancreas, such as pancreatic exocrine insufficiency, pancreatic cancer, cystic fibrosis or pancreatitis.

Acute, recurrent acute, and chronic pancreatitis

Pancreatitis is an inflammatory disease of the pancreas with significant mortality [111]. The disease can range from mild symptoms that affect only the pancreas, to very severe forms that are accompanied by multisystemic organ failure and ultimately death. The death rate in acute pancreatitis (AP) is ~5% and the most common causes of AP are alcohol abuse and gallstones, which constitute 70-80% of all cases [112,113]. In the past decades, the incidence of acute pancreatitis increased and ranges from 13 to 45/100,000 persons per year [114]. Increases in

metabolic syndrome and gallstone disease due to poor diet have been speculated to be associated with the increase in pancreatitis [115].

Pancreatitis can be categorized into three main forms, *acute pancreatitis (*AP), *recurrent acute pancreatitis* (RAP) and *chronic pancreatitis* (CP), each of which can be subcategorized based on the underlying cause of the disease, e.g. alcohol-induced, autoimmune, hereditary, etc. Although once considered separate diseases, AP and CP are now understood as continuum. Some patients with AP can already show mild morphological changes that are typically associated with CP that can progress to more obvious changes such as calcifications and loss of endo- and exocrine function [116–119]. Multiple theories exist on how AP could eventually progress to CP, ranging from oxidative stress theories to toxic-metabolites, ductal obstruction, and necrosis-fibrosis theories. In an attempt to unify the different potential mechanisms that may all lead to CP, the *sentinel acute pancreatitis event* (SAPE) hypothesis was introduced in 1999 [120,121]. This generalized hypothesis assumes an initiating event (sentinel event) to be necessary, which causes AP and injury to pancreatic acinar cells. In response to the first AP episode, anti-inflammatory mechanisms and other pro-fibrotic events subsequently enable the progression from an initial acute event to chronic pancreatitis.

While the SAPE hypothesis can explain the transition from AP to CP, it makes no assumptions about the origin of the sentinel event itself. Given the complexity of the processes within the exocrine pancreas and the obvious challenges of zymogen synthesis, storage and secretion, it is no surprise that risk factors and pathomechanisms are as manifold as they are diverse. Gallstones are among the most common etiologies of pancreatitis and are found in up to 20% of the population [122,123]. In 8% of patients carrying gallstones, the outflow of pancreatic juice from the pancreatic duct into the duodenum is obstructed [124–126]. The result is increased pressure in the pancreas and damage to acinar cells triggering the disease onset [124].

Alcohol abuse is another well-known risk factor for developing pancreatitis [114,127]. However, despite the relationship of alcohol abuse and pancreatitis being well established, most alcoholics do not develop pancreatitis, indicating that alcohol is merely increasing the susceptibility to other risk factors, rather than being the sole cause of pancreatitis in most cases [128–130]. Ethanol's toxicity is mostly attributed to its metabolites acetaldehyde and fatty acid ethyl esters [131–136]. Fatty acid ethyl esters have been associated with increased intracellular Ca^{2+} concentrations, which was accompanied by premature zymogen activation, ATP depletion and cell necrosis [135,137–140]. Additionally, acetaldehyde was found to redirect VAMP8-containing ZGs from the apical pole of acinar cells to the basolateral membrane resulting in the release of digestive enzymes into the interstitial space through the VAMP8, syntaxin4, SNAP23 fusion complex [136,141]. Other risk factors include smoking,

abdominal trauma, endoscopic procedures, drugs, autoimmune diseases and infections [142,143].

While the etiologies for pancreatitis are manifold, they all ultimately result in cell and/ or tissue damage through premature trypsin activation within acinar cells or the interstitial space. One major mechanism through which trypsin activation is thought to occur is the co-localization theory. During ER-stress, impairment of autophagy and other complex changes to cellular homeostasis, co-localization of lysosomal hydrolases and digestive zymogens is proposed (Figure 3). Fusion of lysosomes with ZGs results in the enrichment of cathepsin B within ZGs [144,145]. Cathepsin B is a known activator for trypsinogen and mouse models have shown significantly decreased trypsinogen activation in cathepsin B deficient mice [146,147]. Often, co-localization occurs due to defects in autolysosome formation during autophagy and results in activation of trypsinogen, which is now considered a major mechanism for the genesis of pancreatitis [148–150].
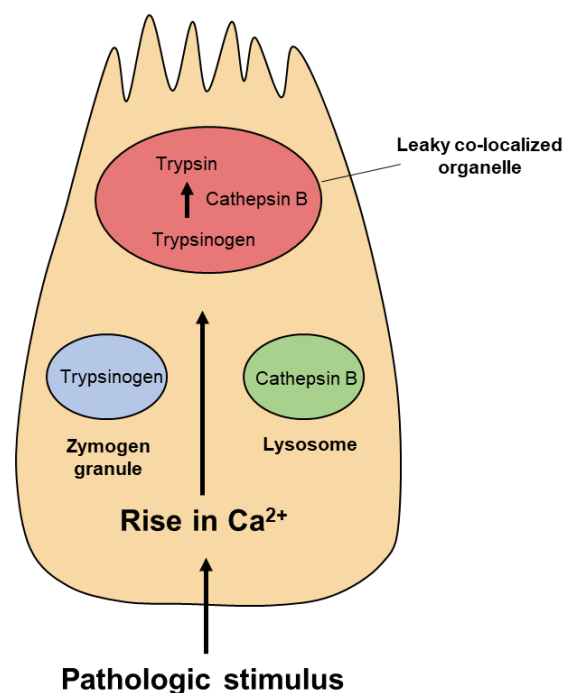

Figure 3: Schematic representation of the co-localization theory. Fusion of lysosomes and zymogen granules cause the activation of trypsinogen by cathepsin B.

Finally, hereditary causes are an important risk factor for pancreatitis, with mutations in the major genes associated with pancreatitis found in over 50% of patients [151,152]. Most commonly, mutations are found in the *PRSS1, CTRC, SPINK1* and *cystic fibrosis transmembrane conductance regulator* (*CFTR*) genes. The gene products cationic trypsin,

SPINK1 and chymotrypsin C are all expressed in acinar cells, while CFTR is the only member expressed in duct cells. No disease-causing mutations are known in the *PRSS2* gene. However, a G191R mutation was found to be enriched in the healthy population, suggesting a protective effect [153]. The G191R mutation makes anionic trypsin more degradation-sensitive by introducing an additional autolytic site, amplifying the self-protecting effect through autoproteolysis. Apart from dominant *PRSS1* mutations, most variants are not the sole cause for pancreatitis and rather predispose for pancreatitis by lowering the threshold requirements for developing the disease. The two genes most commonly found to carry mutations in pancreatitis patients are *PRSS1* and the trypsin inhibitor *SPINK1*. There is a plethora of mutations known in both genes, with most of them representing rare variants (Figure 4) [154–156].

Contrary to the seemingly logical premise, impaired trypsin inhibition by SPINK1 mutations is not a common mechanism for developing pancreatitis [157]. Nevertheless, two decades ago, an N34S variant in SPINK1 was found to be a major risk factor for chronic pancreatitis [104,158,159]. The N34S mutation is very common and can be found in about 0.6% of the general population. In pancreatitis patients, the mutation is overrepresented with about 5.8% carrying the N34S variant, sometimes in combination with other mutations. Initially, it was postulated, that the N34S amino acid substitution would cause impaired trypsin inhibition, leading to increased trypsin autoactivation in pancreatic acinar cells. However, later studies showed that this was not the case, causing other hypotheses to emerge and many research groups calling for more detailed analyses [160,161]. Despite great efforts in the past two decades aiming at elucidating the pathomechanisms underlying the N34S mutation, the mechanism of action has remained enigmatic [162].
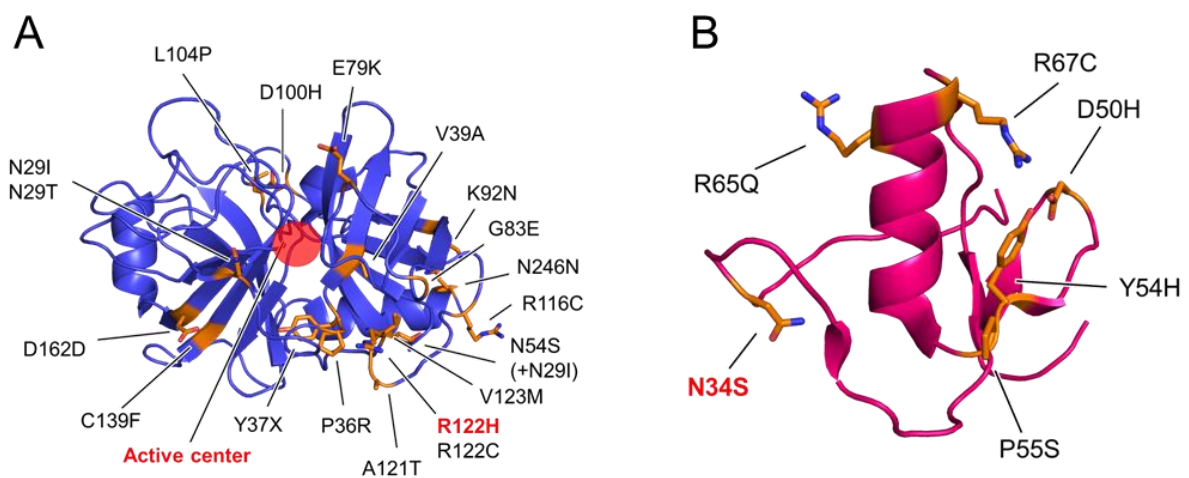

Figure 4: Mutations in cationic trypsin and SPINK1. (A) Crystal structure of cationic trypsin labeled with pancreatitis-associated mutations (pdb: 7QE8). The active site is highlighted in red. (B) Crystal structure of SPINK1 labeled with pancreatitis-associated mutations (pdb: 7QE8). The most common mutations are highlighted in red.

Perhaps the most prominent mutation is R122H in human cationic trypsin. The R122H mutation was identified in 1996 and was the first variant to be associated with pancreatitis [163]. It was quickly discovered that R122H represents a gain-of-function mutation that increases trypsin autoactivation by removing one highly exposed autolytic site in TRY1. Unlike most of the mutations associated with pancreatitis, R122H has a very high penetrance of over 80% and is encountered in over 50% of patients with hereditary pancreatitis [156]. Following its discovery, the R122H variant was subjected to many biochemical studies, with some of them attributing an inhibitory function to the Arg122-loop in addition to its role in autoproteolysis [164]. Weak inhibitory potential of this loop was demonstrated, however the inhibitory mechanism and loop geometry upon binding are still under debate, with some reports claiming similar mechanisms to canonical inhibitors like SPINK1.

Overall, with the emergence of next generation sequencing methods and the accompanying identification of novel genetic risk factors, insights into the cellular regulation of tryptic activity were obtained that may present a starting point for drug targets and design of novel therapeutics.

In summary, pancreatitis is a disease of multiple etiologies and is often caused by many independent events. Thus, it is considered a threshold disease with genetic, dietary and other lifestyle factors as well as traumata contributing to the overall risk. Despite the variety of etiologies and different underlying pathomechanisms, they all share the same outcome of pancreas autodigestion and inflammation, with prematurely activated trypsin in the pancreas being the common focal point.

Results

Crystal structure of human anionic trypsin reveals the molecular basis of the autoproteolytic failsafe mechanism (Article I)

In 1995, the first crystal structure of a human trypsin was reported [165]. Active human cationic trypsin (TRY1) crystallized in the presence of a diisopropyl phosphofluoridate (DFP) inhibitor. Following the discovery of human mesotrypsin (TRY3), also known as brain trypsin, a crystal structure of the benzamidine-inhibited enzyme was reported in 2001 [166]. As of the time of writing this thesis, there are a total of six human cationic trypsin and 15 mesotrypsin structures reported in the protein data bank. Despite being the predominant isoform during pancreatitis, no crystal structure of human anionic trypsin (TRY2) was available. Thus, for **Articles I, III and IV** a refolding protocol for human trypsin isoforms was established. To the best of our knowledge, this is the first purification method described that yields very homogeneous trypsin with high yield from inclusion bodies using a simple, conventional His-tag purification without special or custom purification columns (Figure 5A and 5B). Using a small chromogenic substrate, all trypsin isoforms display similar conversion rates (k_{cat}) and Michaelis-Menten constants (K_m) (Figure 5C). In **Article I** the purified anionic trypsin autoactivated after adding enterokinase at a molar ratio of 1:20,000 to start the autoactivation cascade, which was followed by size exclusion chromatography to remove enterokinase and cleaved propeptides. Following activation, the enzyme was concentrated to 15 mg mL^{-1} and immediately used for crystallization in 22% polyethylene glycol (PEG) 4000, 0.1 M Hepes pH 7.5 and 0.1 M sodium acetate using the hanging drop vapor diffusion method. Anionic trypsin crystallized in space group $P2_12_12_1$ within two to three days and the structure was refined to a resolution of 1.7 Å. Phases were solved by molecular replacement using human cationic trypsin (pdb: 1TRN) as search model (Figure 5D). With SDS-PAGE four additional bands were observed for the wild type enzyme but not for the catalytically inactivated S200A mutant (Figure 5A). The additional bands were attributed to auto-cleavage products of the most well-known autolytic sites after Arg122 and Lys193. Additional bands observed for TRY2 S200A are impurities introduced by the higher enterokinase concentration used due to the absence of autoactivation in the catalytically inactive variant. The electron density of the crystal structure also showed a cleaved Arg122-Val123 bond, while the Lys193-Asp194 bond appeared to be intact with well-defined electron density. Two TRY2 molecules formed a dimer with Arg122 of one enzyme deeply burrowed in the S1 binding pocket of the neighboring TRY2 (Figure 5E). Arg122 and Val123 are separated by more than 8 Å, showing that the loop undergoes induced fit movements to accommodate the created binding interface of the homodimer. Creating the binding interface requires translation and a 180° rotation about the translation axis, which crystallographically

22

results in an endless chain. However, in solution no multimers or even fibers could be observed, most likely due to the weak interaction of the TRY2 dimer.

The Arg122 loop is the first autolytic site to be cleaved by autodigestion in TRY1 and TRY2 and is thus, integral to the autoproteolytic failsafe mechanism. The importance of this loop is also reflected in the fact that the R122H mutation in TRY1 is by far the most important in hereditary pancreatitis, with a penetrance of about 80% and more than 50% of patients carrying this mutation [156]. Hence, the crystal structure of the TRY2 homodimer was used to generate a model of TRY1, which was equilibrated by molecular dynamics simulations with NAMD using the CHARMM36 forcefield (Figure 5F). The model can be used to comprehend the first step of autoproteolysis in human trypsin and highlights why removal of this autolytic site through the R122H mutation results in pancreatitis. In addition, weak inhibitory properties were found for the Arg122 loop [164]. A similar geometry to canonical protease inhibitor binding loops was proposed for the Arg122 loop, but the crystal structure displays a very flexible loop with large structural rearrangements after cleavage, refuting these previous propositions. The inhibitory potential can be rationalized through the binding interface of the homodimer, with not only the Arg122 loop, but also the entire backside of the enzyme contributing to binding.
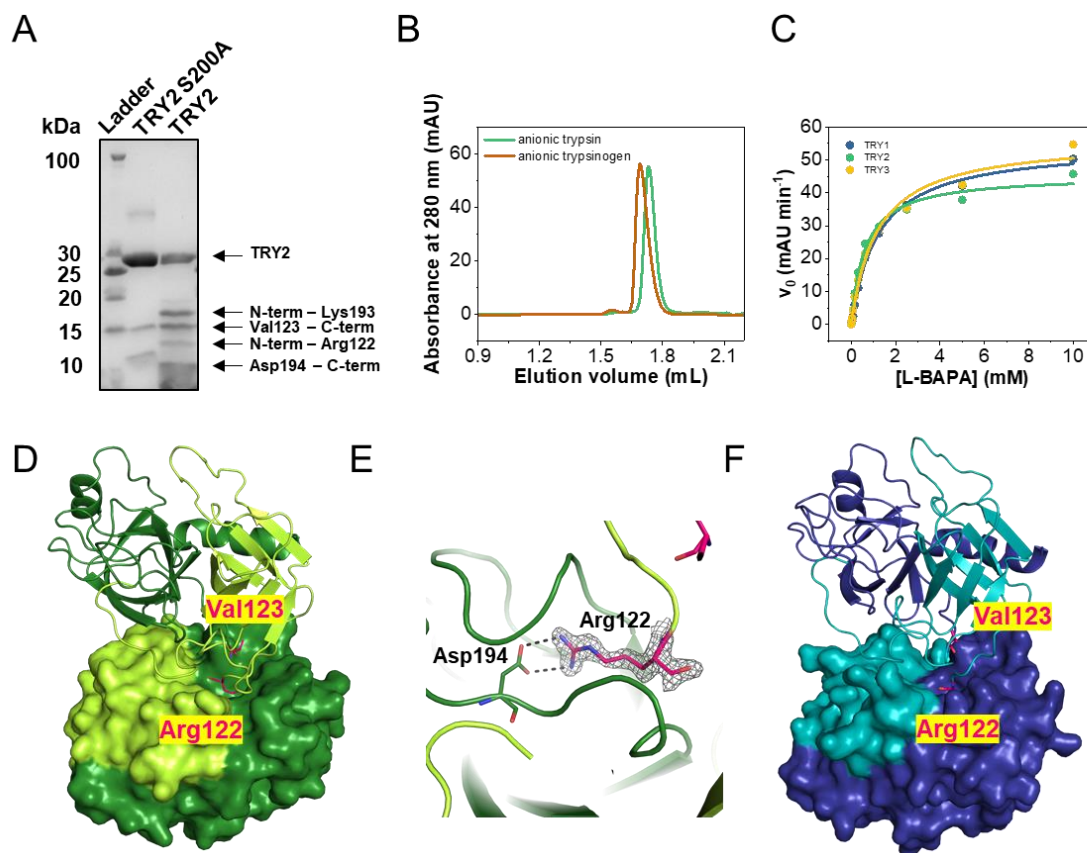

Figure 5: Structure and function of purified human anionic trypsin. (A) Structure of the TRY2-TRY2 dimer. (B) SDS-PAGE of TRY2 S200A and TRY2 WT. (C) Analytical size exclusion chromatography of anionic trypsinogen and trypsin. (D) Interaction of Arg122 with Asp194. The 2Fo-Fc density map is shown 1.6 Å around Arg122 and is contoured at 1.0 σ. (E) Structural model of the TRY1-TRY1 dimer. (F) Michaelis-Menten kinetics of trypsin isoforms with L-BAPA as substrate. Figures are adapted from *Nagel et al.* [167].

Biochemical, biophysical and structural analysis of SPINK1 – TRY1 complexes (Article II and Article III)

The SPINK1 N34S mutation was first discovered and linked to chronic pancreatitis in 2000 [158]. Since then, many hypotheses were formulated attempting to explain the pathogenicity of this common variant. After it became clear that reduced inhibitory potential of the N34S variant is not the mechanism of action, structural differences and rearrangements of hydrogen bond networks at the SPINK1 – TRY1 interface were proposed [104,155].

Article II and Article III aim at elucidating the molecular structure of SPINK1 – TRY1 complexes and focus on differences between SPINK1 WT and the N34S variant. Molecular dynamics simulations were used in **Article II** in order to compare the structural ensembles of both complexes. Two distinct states were investigated with state I representing the "normal" catalytic triad and state II featuring a deprotonated Ser200 and a protonated His63 resembling the transition state during catalysis. Major structural movements in the entire SPINK1 backbone were found in state II for SPINK1 WT and the N34S mutant (Figure 6A and 6B). Structure ensembles clustered in state A/B and C/D. However, pinpointing differences between SPINK1 WT and N34S was difficult regardless of the respective state. Using transition path sampling, a slightly lower energy barrier was found for the SPINK1 N34S transition from state II A/B to state II C/D, which would in principle indicate an easier path towards the transition state and could lead to faster cleavage of the K41-I42 scissile bond in the SPINK1 N34S mutant compared to wild type (Figure 6C). Based on this observation, reduced proteolytic stability of the SPINK1 N34S variant was hypothesized.
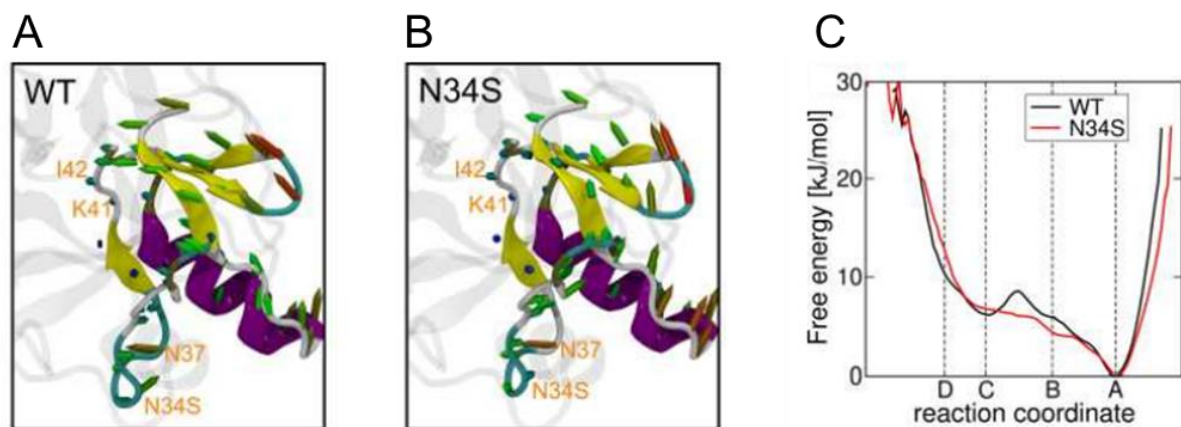

Figure 6: SPINK1 transition from state II A/B to state II C/D. (A) Porcupine plot of the SPINK1 WT transition from state II A/B to C/D. (B) Same as (A) but for SPINK1 N34S. (C). Energy landscape obtained from transition path sampling of the A/B to C/D transition. Figures are adapted from *Kulke et al.* [168].

Article III experimentally ties crystal structures of SPINK1 WT and the N34S variant in complex with TRY1 to their functional properties. TRY1 was expressed, refolded and purified as described in **Article I**. SPINK1 was expressed in SHuffle T7 Express cells at 16°C and purified by nickel affinity chromatography (Figure 7A).

The interaction of both SPINK1 variants was analyzed by colorimetric equilibrium competition binding assays, using N_α-Benzoyl-L-arginine 4-nitroanilide hydrochloride (L-BAPA) as substrate (Figure 7B). Due to the tight binding of SPINK1 to TRY1 a lag phase of 5 h was used to ensure full equilibration of the system before following the substrate conversion at 405 nm. Both SPINK1 WT and SPINK1 N34S display inhibition constants (K_i) below 50 pM, confirming previous observations of the N34S variant not showing reduced inhibitory potential.

Despite showing similar inhibitory properties in equilibrium, point mutations can often introduce changes in binding kinetics that are not observed in equilibrium. Hence, surface plasmon resonance (SPR) was used to determine association (k_a) and dissociation rates (k_d) of the SPINK1 – TRY1 complexes. In order to ensure surface stability, the catalytically inactive TRY1 S200A variant was immobilized to a CM5 sensor chip followed by single cycle kinetics experiments (Figure 7C and 7D). Both SPINK1 variants exhibited similar kinetic properties and very high complex stability with k_d-values of 2×10^{-4} s^{-1}. Dissociation constants (K_d) are similar to the K_i-values obtained using the chromogenic, solution-based assay, although their values differ by one order of magnitude. Both methods show affinities in the sub-nanomolar range.

Isothermal titration calorimetry (ITC) was used to determine binding enthalpies (ΔH) of the SPINK1 – TRY1 complexes. TRY1 S200A was placed in the sample cell and SPINK1 in the syringe. The high affinity of the interactions prohibits direct measurements of K_d using ITC and due to the lack of an appropriate competitor, K_d-values were not determined and taken from previous SPR measurements. Sufficient spacing between injections was ensured before peak integration and binding enthalpies were determined (Figure 7E and 7F). Binding of both SPINK1 variants is enthalpically driven with a ΔH of about -10 kcal mol^{-1}.
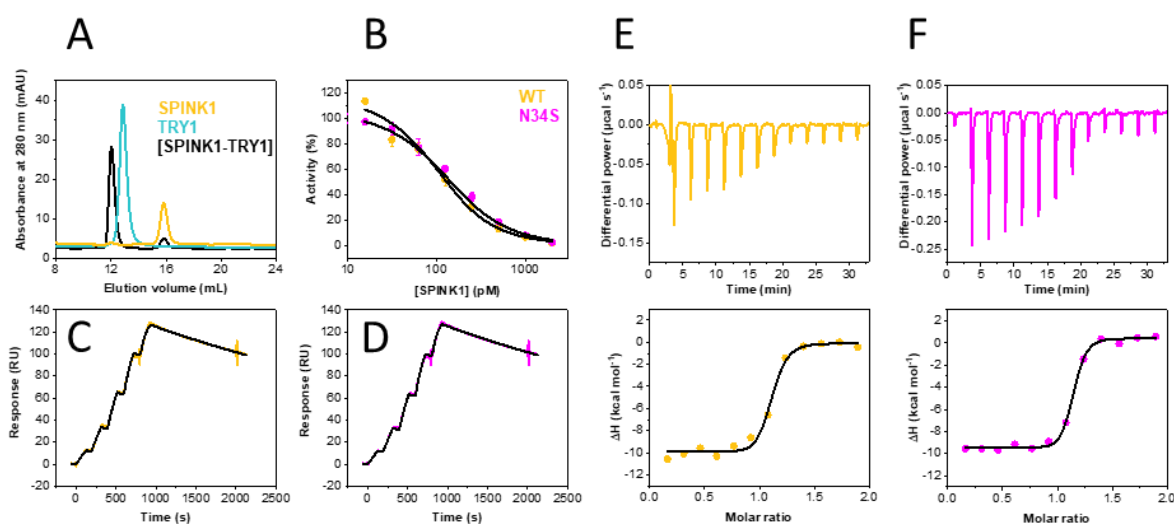
Figure 7: Biophysical characterization of SPINK1 – TRY1 interactions. (A) Analytical size exclusion of free SPINK1 (yellow), free TRY1 (blue) and SPINK1 – TRY1 complexes (black). (B) Trypsin activity assay with SPINK1 WT (yellow) or SPINK1 N34S (pink). (C) SPR single cycle experiments of the SPINK1 WT – TRY1 interaction. (D) Same as (C) but for SPINK1 N34S. (E) Isothermal titration calorimetry of the SPINK1 WT – TRY1 interaction. (F) Same as (E) but for SPINK1 N34S. Figures are adapted from *Nagel et al.* [169]*.*

For crystallization, the catalytically inactive TRY1 S200A variant was used and SPINK1 – TRY1 complexes were purified by size exclusion chromatography. Enzyme – inhibitor complexes crystallized in space group P3$_1$21 using the hanging drop vapor diffusion method with 15% PEG 4000 and 0.3 M (NH$_4$)$_2$SO$_4$ as reservoir solution. Crystals were briefly soaked in reservoir solution containing additional 7% PEG 400 as cryoprotectant before cryocooling in liquid nitrogen. The SPINK1 WT – TRY1 S200A complex was refined up to 2.9 Å and the SPINK1 N34S – TRY1 S200A complex up to 2.1 Å. Attempts of crystallizing SPINK1 complexes with active TRY1 were unsuccessful, most likely due to digestion of the enzyme, even in the inhibited state. Phases were solved by molecular replacement using human cationic trypsin (pdb: 1TRN) and a SPINK1 mutant (pdb: 1CGI) as search models.

In accordance with the standard mechanism of canonical inhibitors, both SPINK1 variants interact with TRY1 in a substrate like manner (Figure 8A). The specificity determining Lys41 residue in SPINK1 penetrates into the S1 binding pocket and interacts with Asp194 in TRY1 (Figure 8B). The canonical conformation of the binding loop is stabilized by Asn56 in SPINK1, which interacts with the Lys41 adjacent P2 and P1' residues, likely supporting the re-synthesis of the K41-I42 peptide bond after cleavage. The carbonyl-carbon of Lys41 faces the oxyanion hole formed by Gly198 and Ala200 and Tyr33 in SPINK1 forms a cation-pi bond with Arg101 in TRY1.

The electron density was generally well defined with the major exception being the loop in SPINK1 carrying the N34S mutation. Probably due to the lack of interactions with TRY1 the loop is very flexible, which is also reflected in the high *B*-factors in that region (Figure 8C). Flexibility of the mutated loop in both structures and the absence of interactions with TRY1,

except for a cation-pi bond between Tyr33 in SPINK1 and Arg101 in TRY1, explains the lack of functional consequences due to the N34S amino acid substitution.
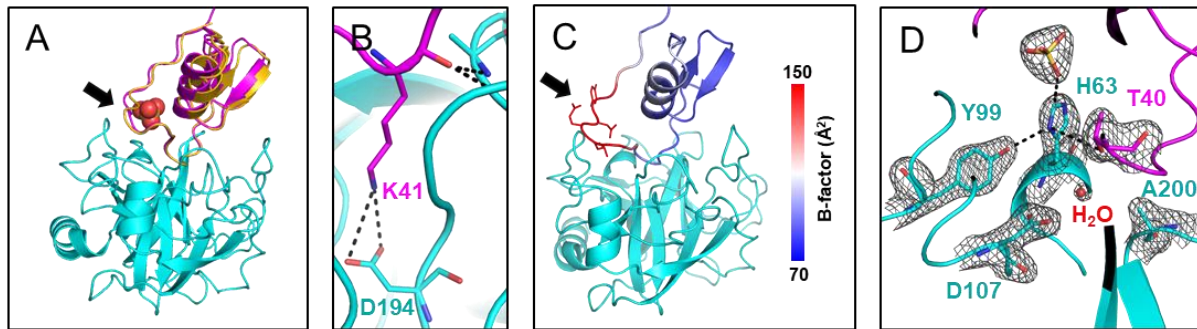

Figure 8: Structure and superposition of SPINK1 – TRY1 complexes. (A) Structure and superposition of TRY1 (cyan) in complex with SPINK1 WT (yellow) and SPINK1 N34S (pink). The black arrow indicates the position of the N34S mutation. (B) Lys41 interacting with Asp194 in the S1 binding pocket. (C) Structure of the SPINK1 N34S – TRY1 structure with SPINK1 colored according to *B*-factors. (D) Catalytic triad of the SPINK1 N34S – TRY1 complex. The 2Fo-Fc density map is shown 1.6 Å around residues of the catalytic triad and Tyr99 and Thr40 in TRY1 and SPINK1, respectively. Electron density around the sulfate ion and a water molecule is shown as well and contoured at 1.5 σ. The figure was adapted from *Nagel et al.* [169].

Unexpectedly, His63 of the catalytic triad was found to be in an atypical *trans* conformation with very well-defined density, indicating little structural flexibility (Figure 8D). Normally in serine proteases, His63 is found sandwiched between Asp107 and Ser200 forming a catalytically productive arrangement of the catalytic triad. In both SPINK1 – TRY1 structures however, His63 is found facing away from the catalytic triad and towards a sulfate ion, which might contribute to stabilizing the *trans* conformation. The unusual rotamer conformation is further stabilized by Tyr99 in TRY1 and Thr40 in SPINK1. It is unclear whether the observed rotamer conformation is an artifact due to the S200A mutation or crystallization conditions, or whether it is part of the inhibitory mechanism of SPINK1. However, no crystal contacts are in close proximity to His63 and most structures reported in the protein data bank featuring the S200A mutation still show His63 in the usual *gauche*+ conformation. It is thus tempting to speculate that the *trans* conformation of His63 might be part of the inhibition mechanism, effectively removing the proton acceptor for Ser200 and thereby preventing the proteolytic cleavage of the SPINK1 inhibitor. Furthermore, the influence of SPINK1 binding on the His63 conformation was demonstrated by molecular dynamics simulations with changes in the rotamer conformation of His63 always being preceded by a conformational change in Thr40 of SPINK1. Thus, the change in rotamer conformation of His63 may have biological relevance and could enhance the proteolytic stability of the SPINK1 inhibitor.

SPINK1 backbone influence on inhibition (Article IV)

Canonical serine protease inhibitors all possess a similar, canonical binding loop and inhibit their target proteases through the standard mechanism. The canonical binding loop features a similar geometry in all 18 families, but it is stabilized by very different supporting scaffolds, indicating that the standard mechanism of inhibition evolved independently several times [170,171]. Based on the similarity of binding loops among the different families, intra- and interscaffolding additivity principles were extrapolated. However, exceptions have been found for both additivity principles. Loop exchange variants between two unrelated canonical inhibitors do not show similar affinity and evolve different consensus sequences in directed evolution experiments [172]. In **Article IV** the influence of the Kazal backbone on binding properties of SPINK1 was analyzed. Through phylogenetic analysis of trypsin-targeting SPINK1 homologues in other mammalian species, two distinct P2-P1 sub-clusters were identified with the amino acid combinations Thr-Lys or Pro-Arg appearing to have co-evolved (Figure 9A). Hence, SPINK1 mutations were introduced to generate the four possible P2-P1 combinations. In addition, mouse and naked mole SPINK1 were expressed and purified, as they feature identical binding loops to the SPINK1 T40P K41R and the T40P variant, respectively. AlphaFold models of mouse and naked mole SPINK1 were generated and compared to the SPINK1 WT crystal structure. No differences in the overall structures were predicted but circular dichroism (CD) experiments indicate slightly different secondary structure compositions. Affinity and binding kinetics towards TRY1 were characterized by colorimetric enzymatic assays and surface plasmon resonance (SPR) (Figure 9B and 9C). The SPINK1 T40P variant displayed at least 10-fold increased affinity, which is almost exclusively attributed to a 10-fold increase in complex half-life (Figure 9D). Van't Hoff and Eyring analyses show that proline as P2 residue results in a more favorable entropy term compared to threonine. Despite featuring the identical T40P binding loop, naked mole SPINK1 displayed 20-fold lower affinity towards TRY1 and was the lowest affinity binder out of all variants tested. TRY1 does not possess a distinct S2 binding pocket, although L104 is able to provide some hydrophobic stabilization. Despite the lack of a designated binding pocket, proline at the P2 position highly influences the binding affinity of the SPINK1 – TRY1 interaction using the human SPINK1 backbone. The increased affinity can be explained through loop rigidification by proline, locking the loop in the canonical conformation and therefore, an overall more favorable geometry.

It is unclear whether the increased affinity observed for the SPINK1 T40P variant is exclusive for the interaction with TRY1. Hence, all SPINK1 variants were tested for affinity and binding kinetics towards human TRY2, bovine trypsin and porcine trypsin. For human TRY2, SPINK1 T40P again showed a drastic increase in affinity and complex half-life and binding of all

SPINK1 variants was generally stronger than for human TRY1. Again, naked mole SPINK1 was among the weakest binders, despite the identical T40P binding loop. Bovine and porcine trypsin were barely affected by any of the variants but appeared to prefer the human Kazal backbone compared to mouse and naked mole (Figure 9F). Thus, substituting the P2 residue in human SPINK1 significantly improves the affinity and complex stability towards the two major human trypsin isoforms, but has no significant influence on the interaction with trypsin from the other tested species.
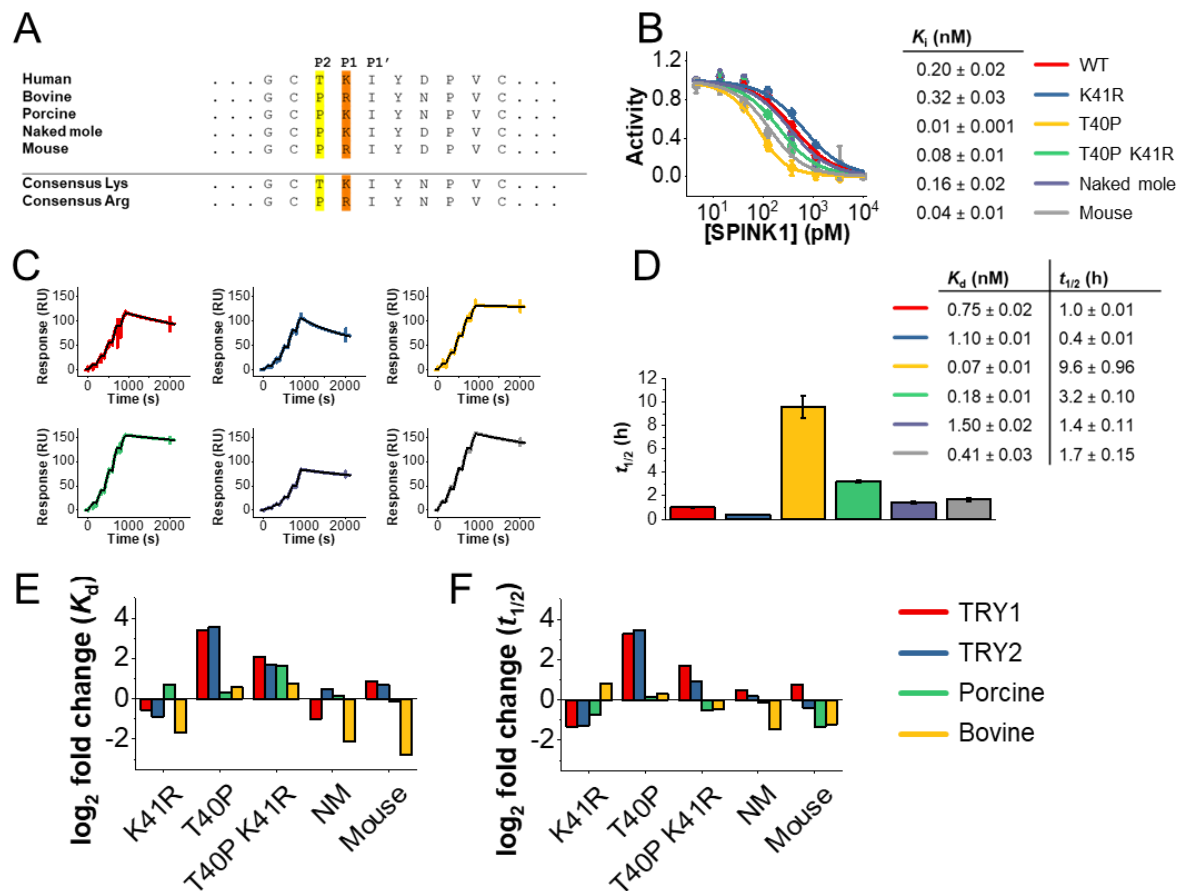
Figure 9: Functional characterization and comparison of SPINK1 variants. (A) Multiple sequence alignment of selected SPINK1 sequences and consensus sequences of SPINK1 homologues containing Lys or Arg as P1 residue. (B) Activity assay of TRY1 with different SPINK1 variants. (C) SPR single cycle kinetic experiments of SPINK1 – TRY1 interactions. (D) Comparison of complex half-life ($t_{1/2}$) of SPINK1 – TRY1 interactions. (E) Changes in affinity of different SPINK1 variants compared to human wild type expressed as $\log_2$(fold change). (F) Changes in $t_{1/2}$ of different SPINK1 variants compared to human wild type expressed as $\log_2$(fold change). The figure was adapted from *Nagel et al.* (**Article IV**).

In addition, the influence of pH on the SPINK1 variant interactions with TRY1 was characterized kinetically. Lowering the pH from 7.4 to 5 resulted in a 16-fold increase in K_d-values for all variants, but the kinetics of mouse and naked mole SPINK1 were much more influenced compared to the human variants indicating a higher susceptibility towards pH changes of the these SPINK1 homologues.

In brief, a SPINK1 T40P variant displaying 10-fold increased affinity and complex stability was generated. The importance of the P2 residue on loop geometry and binding strength as well as differences between trypsin isoforms from different species was highlighted and the influence of the Kazal backbone on enzyme – inhibitor interactions was shown. Because Kazal inhibitors often serve as starting points for generating therapeutic molecules, our data suggest optimizing the canonical binding loop supporting scaffold in addition to the canonical binding loop itself in order to generate higher affinity, more stable and pH resistant inhibitors.

Summary

This thesis provides insights into the regulation of tryptic activity in pancreatic acinar cells. It lays the structural and functional foundation of trypsin self-regulation and trypsin inhibition by its inhibitor SPINK1. The molecular basis of two of the most frequently encountered mutations in pancreatitis patients was established and insights into the inhibition mechanism of SPINK1 were obtained.

The first crystal structure of human anionic trypsin (TRY2) was reported (**Article I**). Perhaps because it is the first structure of any human trypsin isoform that crystallized uninhibited, the structure shows a TRY2 – TRY2 dimer in a state representative of the first step of autoproteolysis. Induced fit movements of the autolytic loop and the resulting interface of the homodimer explain previous observations of the Arg122-containing loop possessing weak inhibitory potential. A corresponding TRY1 – TRY1 dimer was modelled and equilibrated using molecular dynamics simulations. The interface is similar to the TRY2 homodimer and can be used to explain the functional relevance of pancreatitis-associated mutations. Additionally, the R122H mutation in TRY1 is the most frequently encountered mutation in patients with hereditary pancreatitis. In the TRY1 homodimer, Arg122 penetrates deeply into the S1 binding pocket and contributes significantly to the binding affinity. By substituting this highly exposed arginine, an autolytic site in trypsin is removed, cleavage of which is generally considered the first step of an autoproteolytic failsafe mechanism. Furthermore, autoinhibition at high concentrations is also diminished. Knowledge of the trypsin homodimer interface will accelerate the assignment of functions to newly discovered mutations in the future.

In addition to autoproteolysis and autoinhibition, tryptic activity in zymogen granules is inhibited by SPINK1. An N34S mutation, which is common in the general population, has been associated with chronic pancreatitis over two decades ago. Our structure function analysis of SPINK1 – TRY1 complexes indicate small structural differences between the wild type and N34S complexes (**Article II and Article III**). However, these differences are minor and unlikely

to translate into clinical effects. Shortly after the discovery of the N34S mutation it became known that reduced inhibitory activity is not the disease-causing mechanism of this variant. This thesis expands on that knowledge by showing equal kinetic and thermodynamic behavior of both SPINK1 variants and explains the lack of functional consequences of the N34S mutation through high loop flexibility and the absence of productive interactions with TRY1 in the area surrounding the N34S mutation.

Therefore, it is unlikely that the pathomechanism of the N34S mutation occurs at the protein level. Linkage disequilibrium of the N34S variant with mutations in non-sequenced regions upstream of the *SPINK1* promoter region offer one possible explanation and a potential candidate was recently identified [173].

Mechanistically, binding of SPINK1 appears to affect the conformation of the residues forming the catalytic triad, although crystallization conditions and introduced mutations may also have had an influence. Specifically, His63 transitions from the normally observed *gauche*$^+$ to the *trans* conformation, disrupting the productive catalytic triad arrangement. The rotamer change removes the proton acceptor and prevents the formation of the acyl-intermediate. The *trans* conformation is then stabilized by Thr40 in SPINK1. Physiologically, this mechanism likely increases the proteolytic stability of SPINK1 by preventing cleavage of the K41-I42 scissile bond.

Furthermore, a SPINK1 T40P variant displaying 10-fold increased inhibitory potency was generated (**Article IV**). Introduction of the mutation selectively improved binding to the two major human trypsin isoforms, but did not influence binding to porcine or bovine trypsin. It will be interesting to investigate, whether the absence of Thr40 in SPINK1 destabilizes the observed *trans* conformation of His63 in **Article III**. The evolutionary rationale for a threonine as P2 residue in human SPINK1, despite possessing lower affinity compared to proline, might be a tradeoff between binding affinity and proteolytic stability.

References

[1]   I. Schechter, Reprint of "On the Size of the Active Site in Proteases. I. Papain," Biochem. Biophys. Res. Commun. 425 (2012) 497–502. https://doi.org/10.1016/j.bbrc.2012.08.015.

[2]   R.H. Hruban, D.S. Klimstra, M.B. Pitman, AFIP Atlas of Tumor Pathology, Series IV - Tumors of the Pancreas, Washington, DC Am. Regist. Pathol. (2006) 75–110.

[3]   N. Jouvet, J.L. Estall, The pancreas: Bandmaster of glucose homeostasis, Exp. Cell Res. 360 (2017) 19–23. https://doi.org/10.1016/j.yexcr.2017.03.050.

[4]   A. Pisania, G.C. Weir, J.J. O'Neil, A. Omer, V. Tchipashvili, J. Lei, C.K. Colton, S. Bonner-Weir, Quantitative analysis of cell composition and purity of human pancreatic islet preparations, Lab. Investig. 90 (2010) 1661–1675. https://doi.org/10.1038/labinvest.2010.124.

[5]   T. Murakami, T. Fujita, T. Taguchi, Y. Nonaka, K. Orita, The Blood Vascular Bed of the Human Pancreas, with Special Reference to the Insulo-Acinar Portal System.Scanning Electron Microscopy of Corrosion Casts, Arch. Histol. Cytol. 55 (1992) 381–395. https://doi.org/10.1679/aohc.55.381.

[6]   J.A. Williams, H. Sankaran, M. Korc, I.D. Goldfine, Receptors for cholecystokinin and insulin in isolated pancreatic acini: Hormonal control of secretion and metabolism, Fed. Proc. 40 (1981) 2497–2502.

[7]   S.G. Barreto, C.J. Carati, J. Toouli, G.T.P. Saccone, The islet-acinar axis of the pancreas: More than just insulin, Am. J. Physiol. - Gastrointest. Liver Physiol. 299 (2010) G10–G22. https://doi.org/10.1152/ajpgi.00077.2010.

[8]   P.D. J. Gordon Betts, Kelly A. Young, James A. Wise, Eddie Johnson, Brandon Poe, Dean H. Kruse, Oksana Korol, Jody E. Johnson, Mark Womble, Anatomy and Physiology, OpenStax, Houston, Texas, 2013. https://openstax.org/books/anatomy-and-physiology/pages/1-introduction.

[9]   H.M. Dittrich, History of the discovery of pancreatic diabetes by von Mering and Minkowski 1889. A historical overview on the occasion of the 100th anniversary, Z. Gesamte Inn. Med. 44 (1989) 335–340.

[10]   M.A. Lane, The cytological characters of the areas of langerhans, Am. J. Anat. 7 (1907) 409–422. https://doi.org/10.1002/aja.1000070304.

[11]   M. Brissova, M.J. Fowler, W.E. Nicholson, A. Chu, B. Hirshberg, D.M. Harlan, A.C. Powers, Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy, J. Histochem. Cytochem. 53 (2005) 1087–1097. https://doi.org/10.1369/jhc.5C6684.2005.

[12]   N. Bozadjieva, J.A. Williams, E. Bernal-Mizrachi, Glucagon, Pancreapedia Exocrine Pancreas Knowl. Base. (2013).

[13]   S.E. Kahn, D.A. D'Alessio, M.W. Schwartz, W.Y. Fujimoto, J.W. Ensinck, G.J. Taborsky, D. Porte, Evidence of cosecretion of islet amyloid polypeptide and insulin by β-cells, Diabetes. 39 (1990) 634–638. https://doi.org/10.2337/diabetes.39.5.634.

[14]   R.A. Defronzo, From the triumvirate to the ominous octet: A new paradigm for the treatment of type 2 diabetes mellitus, Diabetes. 58 (2009) 773–795. https://doi.org/10.2337/db09-9028.

[15]   D.S. Edgerton, M. Lautz, M. Scott, C.A. Everett, K.M. Stettler, D.W. Neal, C.A. Chu, A.D. Cherrington, Insulin's direct effects on the liver dominate the control of hepatic glucose production, J. Clin. Invest. 116 (2006) 521–527. https://doi.org/10.1172/JCI27073.

[16]   D.K. Sindelar, C.A. Chu, P. Venson, E.P. Donahue, D.W. Neal, A.D. Cherrington, Basal hepatic glucose production is regulated by the portal vein insulin concentration, Diabetes. 47 (1998) 523–529. https://doi.org/10.2337/diabetes.47.4.523.

[17]   M. Hatting, C.D.J. Tavares, K. Sharabi, A.K. Rines, P. Puigserver, Insulin regulation of gluconeogenesis, Ann. N. Y. Acad. Sci. 1411 (2018) 21–35. https://doi.org/10.1111/nyas.13435.

[18]   R.A. DeFronzo, The triumvirate: β-cell, muscle, liver. A collusion responsible for NIDDM, Diabetes. 37 (1988) 667–687. https://doi.org/10.2337/diab.37.6.667.

[19]   G.J. Morton, M.W. Schwartz, Leptin and the central nervous system control of glucose metabolism, Physiol. Rev. 91 (2011) 389–411. https://doi.org/10.1152/physrev.00007.2010.

[20]   N. Sarwar, P. Gao, S.R. Kondapally Seshasai, R. Gobin, S. Kaptoge, E. Di Angelantonio, E. Ingelsson, D.A. Lawlor, E. Selvin, M. Stampfer, C.D.A. Stehouwer, S. Lewington, L. Pennells, A. Thompson, N. Sattar, I.R. White, K.K. Ray, J. Danesh, R.W. Tipping, C.E. Ford, S.L. Pressel, A.R. Folsom, L.E. Chambless, L.E. Wagenknecht, D.B. Panagiotakos, C. Pitsavos, C. Chrysohoou, C. Stefanadis, M. Knuiman, P.H. Whincup, S.G. Wannamethee, R.W. Morris, S. Kiechl, J. Willeit, F. Oberhollenzer, A. Mayr, N.

Wald, S. Ebrahim, J.W. Yarnell, J. Gallacher, E. Casiglia, V. Tikhonoff, P.J. Nietert, S.E. Sutherland, D.L. Bachman, J.E. Keil, I.H. de Boer, J.R. Kizer, K.J. Mukamal, A. Tybjærg-Hansen, B.G. Nordestgaard, M. Benn, R. Frikke-Schmidt, S. Giampaoli, L. Palmieri, S. Panico, D. Vanuzzo, L. Pilotto, A. Gómez de la Cámara, M.A. Rubio, L. Simons, J. McCallum, Y. Friedlander, F.G.R. Fowkes, A.J. Lee, J. Taylor, J.M. Guralnik, C.L. Phillips, R. Wallace, D.G. Blazer, K.T. Khaw, H. Brenner, E. Raum, H. Müller, D. Rothenbacher, J.H. Jansson, P. Wennberg, A. Nissinen, C. Donfrancesco, V. Salomaa, K. Harald, P. Jousilahti, E. Vartiainen, M. Woodward, R.B. D'Agostino, R.S. Vasan, C.S. Fox, M.J. Pencina, E. Bladbjerg, T. Jørgensen, L. Møller, J. Jespersen, R. Dankner, A. Chetrit, F. Lubin, L. Wilhelmsen, H. Eriksson, K. Svärdsudd, L. Welin, A. Rosengren, G. Lappas, C. Björkelund, L. Lissner, C. Bengtsson, P. Cremer, D. Nagel, T.E. Strandberg, R.S. Tilvis, T.A. Miettinen, Y. Kiyohara, H. Arima, Y. Doi, T. Ninomiya, B. Rodriguez, J.M. Dekker, G. Nijpels, E. Rimm, J.K. Pai, S. Sato, H. Iso, A. Kitamura, H. Noda, U. Goldbourt, K. Nyyssönen, T.P. Tuomainen, J.T. Salonen, D. Deeg, J.L. Poppelaars, T.W. Meade, B. Hedblad, G. Berglund, G. Engström, W.M.M. Verschuren, A. Blokstra, M. Cushman, B.M. Psaty, S. Shea, A. Döring, W. Koenig, C. Meisinger, W. Mraz, H. Bas Bueno-De-Mesquita, A. Fletcher, L.H. Kuller, G. Grandits, R. Selmer, A. Tverdal, W. Nystad, R. Gillum, M. Mussolino, S. Hankinson, J.E. Manson, J.A. Cooper, K.A. Bauer, K.W. Davidson, S. Kirkland, J. Shaffer, M.R. Korin, I. Holme, P. Ducimetiere, X. Jouven, S.J.L. Bakker, R.T. Gansevoort, H.L. Hillege, C.J. Crespo, M.R. Garcia Palmieri, P. Amouyel, D. Arveiler, A. Evans, J. Ferrières, H. Schulte, G. Assmann, R.G. Westendorp, B.M. Buckley, C.J. Packard, B. Cantin, B. Lamarche, J.P. Després, G.R. Dagenais, E. Barrett-Connor, D.L. Wingard, R. Bettencourt, V. Gudnason, T. Aspelund, G. Sigurdsson, B. Thorsson, M. Trevisan, J. Witteman, I. Kardys, M. Breteler, A. Hofman, H. Tunstall-Pedoe, R. Tavendale, G.D.O. Lowe, B. V. Howard, Y. Zhang, L. Best, J. Umans, Y. Ben-Shlomo, G. Davey-Smith, A. Onat, G. Hergenç, G. Can, I. Njølstad, E.B. Mathiesen, M.L. Løchen, T. Wilsgaard, B. Zethelius, U. Risérus, C. Berne, J.M. Gaziano, P. Ridker, H. Ulmer, G. Diem, H. Concin, A. Tosetto, F. Rodeghiero, L. Tinker, S. Liu, M. Marmot, R. Clarke, R. Collins, E. Brunner, M. Shipley, J. Buring, J. Shepherd, S.M. Cobbe, I. Ford, M. Robertson, A. Marin Ibañez, E.J.M. Feskens, D. Kromhout, M. Walker, M. Alexander, S. Erqou, P. Haycock, P.L. Perry, S.G. Thompson, S. Watson, A.M. Wood, D. Wormser, Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies, Lancet. 375 (2010) 2215–2222. https://doi.org/10.1016/S0140-6736(10)60484-9.

[21] A.W. Norman, H.L. Henry, Pancreatic Hormones, in: A.W. Norman, H.L.B.T.-H. (Third E. Henry (Eds.), Hormones, Academic Press, San Diego, 2015: pp. 109–139.

https://doi.org/10.1016/b978-0-08-091906-5.00006-9.

[22]   L. Gudipaty, M.R. Rickels, Pancreatogenic (type 3c) diabetes, Pancreapedia Exocrine Pancreas Knowl. Base. (2015).

[23]   D. Baetens, F. Malaisse-Lagae, A. Perrelet, L. Orci, Endocrine pancreas: Three-dimensional reconstruction shows two types of islets of Langerhans, Science. 206 (1979) 1323–1325. https://doi.org/10.1126/science.390711.

[24]   J. Philippe, Somatostatin inhibits insulin-gene expression through a posttranscriptional mechanism in a hamster islet cell line, Diabetes. 42 (1993) 244–249. https://doi.org/10.2337/diab.42.2.244.

[25]   T.E. Adrian, Pancreatic polypeptide, J. Clin. Pathol. (1978) 43–50. https://doi.org/10.1136/jcp.s1-8.1.43.

[26]   J. Tong, R.L. Prigeon, H.W. Davis, M. Bidlingmaier, S.E. Kahn, D.E. Cummings, M.H. Tschöp, D. D'Alessio, Ghrelin suppresses glucose-stimulated insulin secretion and deteriorates glucose tolerance in healthy humans, Diabetes. 59 (2010) 2145–2151. https://doi.org/10.2337/db10-0504.

[27]   J. Morisset, Somatostatin, Pancreapedia Exocrine Pancreas Knowl. Base. (2015).

[28]   K. Shiratori, K.Y. Lee, T.M. Chang, Y.H. Jo, D.H. Coy, W.Y. Chey, Role of pancreatic polypeptide in the regulation of pancreatic exocrine secretion in dogs, Am. J. Physiol. - Gastrointest. Liver Physiol. 255 (1988) G535–G541. https://doi.org/10.1152/ajpgi.1988.255.5.g535.

[29]   H.F. Kern, Fine Structure of the Human Exocrine Pancreas, Pancreas Biol. Pathobiol. Dis. 2 (1993) 9–19.

[30]   L.J. Miller, Pancreatic acinar cell physiology and function, Curr. Opin. Gastroenterol. 8 (1992) 801–807. https://doi.org/10.1097/00001574-199210000-00012.

[31]   M. Voss, B. Schröder, R. Fluhrer, Mechanism, specificity, and physiology of signal peptide peptidase (SPP) and SPP-like proteases, Biochim. Biophys. Acta - Biomembr. 1828 (2013) 2828–2839. https://doi.org/10.1016/j.bbamem.2013.03.033.

[32]   Y. Nyathi, B.M. Wilkinson, M.R. Pool, Co-translational targeting and translocation of proteins to the endoplasmic reticulum, Biochim. Biophys. Acta - Mol. Cell Res. 1833 (2013) 2392–2402. https://doi.org/10.1016/j.bbamcr.2013.02.021.

[33]   R.T. Waldron, H.Y. Su, H. Piplani, J. Capri, W. Cohn, J.P. Whitelegge, K.F. Faull, S. Sakkiah, R. Abrol, W. Yang, B. Zhou, M.R. Freeman, S.J. Pandol, A. Lugea, Ethanol

Induced Disordering of Pancreatic Acinar Cell Endoplasmic Reticulum: An ER Stress/Defective Unfolded Protein Response Model, Cmgh. 5 (2018) 479–497. https://doi.org/10.1016/j.jcmgh.2018.01.001.

[34] M.J. Grey, Proteomic Study Defines How Alcohol Alters ER Structure and Redox Proteome to Trigger ER Stress and Acinar Cell Pathology in Pancreatitis, Cmgh. 5 (2018) 640–641. https://doi.org/10.1016/j.jcmgh.2018.01.018.

[35] H.P. Harding, M. Calfon, F. Urano, I. Novoa, D. Ron, Transcriptional and translational control in the mammalian unfolded protein response, Annu. Rev. Cell Dev. Biol. 18 (2002) 575–599. https://doi.org/10.1146/annurev.cellbio.18.011402.160624.

[36] D. Ron, H.P. Harding, Protein-folding homeostasis in the endoplasmic reticulum and nutritional regulation, Cold Spring Harb. Perspect. Biol. 4 (2012) a013177. https://doi.org/10.1101/cshperspect.a013177.

[37] P. Walter, D. Ron, The unfolded protein response: From stress pathway to homeostatic regulation, Science. 334 (2011) 1081–1086. https://doi.org/10.1126/science.1209038.

[38] R.C. Wek, D.R. Cavener, Translational control and the unfolded protein response, Antioxidants Redox Signal. 9 (2007) 2357–2371. https://doi.org/10.1089/ars.2007.1764.

[39] C.H. Kubisch, C.D. Logsdon, Endoplasmic reticulum stress and the pancreatic acinar cell, Expert Rev. Gastroenterol. Hepatol. 2 (2008) 249–260. https://doi.org/10.1586/17474124.2.2.249.

[40] R.T. Waldron, S. Pandol, A. Lugea, G. Groblewski, Endoplasmic Reticulum Stress and the Unfolded Protein Response in Exocrine Pancreas Physiology and Pancreatitis, Pancreapedia Exocrine Pancreas Knowl. Base. 10 (2015).

[41] M.G. Desilva, Molecular characterization of a pancreas-specific protein disulfide isomerase, PDIp, DNA Cell Biol. 16 (1997) 269–274. https://doi.org/10.1089/dna.1997.16.269.

[42] X.M. Fu, X. Dai, J. Ding, B.T. Zhu, Pancreas-specific protein disulfide isomerase has a cell type-specific expression in various mouse tissues and is absent in human pancreatic adenocarcinoma cells: Implications for its functions, J. Mol. Histol. 40 (2009) 189–199. https://doi.org/10.1007/s10735-009-9230-5.

[43] T. Fujimoto, O. Nakamura, M. Saito, A. Tsuru, M. Matsumoto, K. Kohno, K. Inaba, H. Kadokura, Identification of the physiological substrates of PDIp, a pancreas-specific protein-disulfide isomerase family member, J. Biol. Chem. 293 (2018) 18421–18433. https://doi.org/10.1074/jbc.RA118.003694.

[44] K. Kurokawa, A. Nakano, The ER exit sites are specialized ER zones for the transport of cargo proteins from the ER to the Golgi apparatus, J. Biochem. 165 (2019) 109–114. https://doi.org/10.1093/jb/mvy080.

[45] S.I. Bannykh, T. Rowe, W.E. Balch, The organization of endoplasmic reticulum export complexes, J. Cell Biol. 135 (1996) 19–35. https://doi.org/10.1083/jcb.135.1.19.

[46] C. Appenzeller-Herzog, H.P. Hauri, The ER-Golgi intermediate compartment (ERGIC): In search of its identity and function, J. Cell Sci. 119 (2006) 2173–2183. https://doi.org/10.1242/jcs.03019.

[47] A. Oprins, C. Rabouille, G. Posthuma, J. Klumperman, H.J. Geuze, J.W. Slot, The ER to Golgi interface is the major concentration site of secretory proteins in the exocrine pancreatic cell, Traffic. 2 (2001) 831–838. https://doi.org/10.1034/j.1600-0854.2001.21112.x.

[48] J.W. Slot, H.J. Geuze, Immunoelectron microscopic exploration of the Golgi complex, J. Histochem. Cytochem. 31 (1983) 1049–1056. https://doi.org/10.1177/31.8.6863900.

[49] W.J. Brown, M.G. Farquhar, The mannose-6-phosphate receptor for lysosomal enzymes is concentrated in cis Golgi cisternae, Cell. 36 (1984) 295–307. https://doi.org/10.1016/0092-8674(84)90223-X.

[50] P. Arvan, D. Castle, Sorting and storage during secretory granule biogenesis: Looking backward and looking forward, Biochem. J. 332 (1998) 593–610. https://doi.org/10.1042/bj3320593.

[51] Y. Guo, D.W. Sirkis, R. Schekman, Protein sorting at the trans-Golgi network, Annu. Rev. Cell Dev. Biol. 30 (2014) 169–206. https://doi.org/10.1146/annurev-cellbio-100913-013012.

[52] L. Orci, M. Ravazzola, M. Amherdt, A. Perrelet, S.K. Powell, D.L. Quinn, H.P.H. Moore, The trans-most cisternae of the Golgi complex: A compartment for sorting of secretory and plasma membrane proteins, Cell. 51 (1987) 1039–1051. https://doi.org/10.1016/0092-8674(87)90590-3.

[53] M. Forgac, Vacuolar ATPases: Rotary proton pumps in physiology and pathophysiology, Nat. Rev. Mol. Cell Biol. 8 (2007) 917–929. https://doi.org/10.1038/nrm2272.

[54] J. Klumperman, R. Kuliawat, J.M. Griffith, H.J. Geuze, P. Arvan, Mannose 6-phosphate receptors are sorted from immature secretory granules via adaptor protein AP-1, clathrin, and syntaxin 6-positive vesicles, J. Cell Biol. 141 (1998) 359–371. https://doi.org/10.1083/jcb.141.2.359.

[55] S.D. Freedman, G.A. Scheele, Regulated secretory proteins in the exocrine pancreas aggregate under conditions that mimic the trans-Golgi network, Biochem. Biophys. Res. Commun. 197 (1993) 992–999. https://doi.org/10.1006/bbrc.1993.2577.

[56] F.A. Leblond, G. Viau, J. Laine, D. Lebel, Reconstitution in vitro of the pH-dependent aggregation of pancreatic zymogens en route to the secretory granule: Implication of GP-2, Biochem. J. 291 (1993) 289–296. https://doi.org/10.1042/bj2910289.

[57] H. Dartsch, R. Kleene, H.F. Kern, In vitro condensation-sorting of enzyme proteins isolated from rat pancreatic acinar cells, Eur. J. Cell Biol. 75 (1998) 211–222. https://doi.org/10.1016/S0171-9335(98)80115-5.

[58] L. Orci, M. Ravazzola, R.G.W. Anderson, The condensing vacuole of exocrine cells is more acidic than the mature secretory vesicle, Nature. 326 (1987) 77–79. https://doi.org/10.1038/326077a0.

[59] P. Paroutis, N. Touret, S. Grinstein, The pH of the secretory pathway: Measurement, determinants, and regulation, Physiology. 19 (2004) 207–215. https://doi.org/10.1152/physiol.00005.2004.

[60] J.R. Casey, S. Grinstein, J. Orlowski, Sensors and regulators of intracellular pH, Nat. Rev. Mol. Cell Biol. 11 (2010) 50–61. https://doi.org/10.1038/nrm2820.

[61] G. Palade, Intracellular aspects of the process of protein synthesis, Science. 189 (1975) 347–358. https://doi.org/10.1126/science.1096303.

[62] S.W. Messenger, M.A. Falkowski, D.D.H. Thomas, E.K. Jones, W. Hong, H.Y. Giasano, N.M. Boulis, G.E. Groblewski, Vesicle associated membrane protein 8 (VAMP8)-mediated zymogen granule exocytosis is dependent on endosomal trafficking via the constitutive-like secretory pathway, J. Biol. Chem. 289 (2014) 28040–28053. https://doi.org/10.1074/jbc.M114.593913.

[63] H.Y. Gaisano, M. Ghai, P.N. Malkus, L. Sheu, A. Bouquillon, M.K. Bennett, W.S. Trimble, Distinct cellular locations of the syntaxin family of proteins in rat pancreatic acinar cells, Mol. Biol. Cell. 7 (1996) 2019–2027. https://doi.org/10.1091/mbc.7.12.2019.

[64] M.A. Falkowski, D.D.H. Thomas, S.W. Messenger, T.F. Martin, G.E. Groblewski, Expression, localization, and functional role for synaptotagmins in pancreatic acinar cells, Am. J. Physiol. - Gastrointest. Liver Physiol. 301 (2011) G306–G316. https://doi.org/10.1152/ajpgi.00108.2011.

[65] N. Behrendorff, S. Dolai, W. Hong, H.Y. Gaisano, P. Thorn, Vesicle-associated

membrane protein 8 (VAMP8) is a SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) selectively required for sequential granule-to-granule fusion, J. Biol. Chem. 286 (2011) 29627–29634. https://doi.org/10.1074/jbc.M111.265199.

[66]   J. V. Gerasimenko, O. V. Gerasimenko, O.H. Petersen, The role of Ca2+ in the pathophysiology of pancreatitis, J. Physiol. 592 (2014) 269–280. https://doi.org/10.1113/jphysiol.2013.261784.

[67]   J.A. Williams, Receptor-mediated signal transduction pathways and the regulation of pancreatic acinar cell function, Curr. Opin. Gastroenterol. 24 (2008) 573–579. https://doi.org/10.1097/MOG.0b013e32830b110c.

[68]   T. Nemoto, T. Kojima, A. Oshima, H. Bito, H. Kasai, Stabilization of exocytosis by dynamic F-actin coating of zymogen granules in pancreatic acini, J. Biol. Chem. 279 (2004) 37544–37550. https://doi.org/10.1074/jbc.M403976200.

[69]   M.R. Turvey, P. Thorn, Lysine-fixable dye tracing of exocytosis shows F-actin coating is a step that follows granule fusion in pancreatic acinar cells, Pflugers Arch. Eur. J. Physiol. 448 (2004) 552–555. https://doi.org/10.1007/s00424-004-1288-z.

[70]   T. Nemoto, R. Kimura, K. Ito, A. Tachikawa, Y. Miyashita, M. Iino, H. Kasai, Sequential-replenishment mechanism of exocytosis in pancreatic acini, Nat. Cell Biol. 3 (2001) 253–258. https://doi.org/10.1038/35060042.

[71]   P. Thorn, K.E. Fogarty, I. Parker, Zymogen granule exocytosis is characterized by long fusion pore openings and preservation of vesicle lipid identity, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 6774–6779. https://doi.org/10.1073/pnas.0400336101.

[72]   M. Dular, M. Petkovšek, New insights into the Mechanisms of Cavitation Erosion, J. Phys. Conf. Ser. 656 (2015) 2040–2044. https://doi.org/10.1088/1742-6596/656/1/012046.

[73]   T. Takahashi, Y. Miao, F. Kang, S. Dolai, H.Y. Gaisano, Susceptibility Factors and Cellular Mechanisms Underlying Alcoholic Pancreatitis, Alcohol. Clin. Exp. Res. 44 (2020) 777–789. https://doi.org/10.1111/acer.14304.

[74]   N. Furukawa, H. Okada, Effects of stimulation of the hypothalamic area on pancreatic exocrine secretion in dogs, Gastroenterology. 97 (1989) 1534–1543. https://doi.org/10.1016/0016-5085(89)90400-9.

[75]   S. Pandol, Pancreatic Embryology and Development, The Exocrine Pancreas. 12 (2010) 173–221.

[76] T. Hiraoka, T. Fukuwatari, M. Imaizumi, T. Fushiki, Effects of oral stimulation with fats on the cephalic phase of pancreatic enzyme secretion in esophagostomized rats, Physiol. Behav. 79 (2003) 713–717. https://doi.org/10.1016/S0031-9384(03)00201-4.

[77] C. Defilippi, T.E. Solomon, J.E. Valenzuela, Pancreatic secretory response to sham feeding in humans, Digestion. 23 (1982) 217–223. https://doi.org/10.1159/000198753.

[78] A. Anagnostides, V.S. Chadwick, A.C. Selden, P.N. Maton, Sham feeding and pancreatic secretion. Evidence for direct vagal stimulation of enzyme output, Gastroenterology. 87 (1984) 109–114. https://doi.org/10.1016/0016-5085(84)90132-x.

[79] S.J. Konturek, W. Bielanski, T.E. Solomon, Effects of an antral mucosectomy, L-364,718 and atropine on cephalic phase of gastric and pancreatic secretion in dogs, Gastroenterology. 98 (1990) 47–55. https://doi.org/10.1016/0016-5085(90)91289-I.

[80] M. V Singer, Neurohormonal control of pancreatic enzyme secretion in animals, Pancreas Biol. Pathobiol. Dis. 2 (1993) 315–331.

[81] P.N. Maton, Cephalic phase of pancreatic secretion, Gastroenterology. 88 (1985) 604–605. https://doi.org/10.1016/0016-5085(85)90556-6.

[82] J.M. Cargill, K.G. Wormsley, Effect of gastric distension on human pancreatic secretion, Acta Hepatogastroenterol. (Stuttg). 26 (1979) 235–238.

[83] M. Vagne, M.I. Grossman, Gastric and pancreatic secretion in response to gastric distention in dogs., Gastroenterology. 57 (1969) 300–310. https://doi.org/10.1016/s0016-5085(19)33903-4.

[84] V.L. Go, A.F. Hofmann, W.H. Summerskill, Pancreozymin bioassay in man based on pancreatic enzyme secretion: potency of specific amino acids and other digestive products., J. Clin. Invest. 49 (1970) 1558–1564. https://doi.org/10.1172/JCI106373.

[85] R.A. Liddle, I.D. Goldfine, M.S. Rosen, R.A. Taplitz, J.A. Williams, Cholecystokinin bioactivity in human plasma. Molecular forms, responses to feeding, and relationship to gallbladder contraction, J. Clin. Invest. 75 (1985) 1144–1152. https://doi.org/10.1172/JCI111809.

[86] J.H. Meyer, G.A. Kelly, L.J. Spingola, R.S. Jones, Canine gut receptors mediating pancreatic responses to luminal L amino acids, Am. J. Physiol. 231 (1976) 669–677. https://doi.org/10.1152/ajplegacy.1976.231.3.669.

[87] J.H. Meyer, G.A. Kelly, Canine pancreatic responses to intestinally perfused proteins and protein digests, Am. J. Physiol. 231 (1976) 682–691. https://doi.org/10.1152/ajplegacy.1976.231.3.682.

[88] W.E. Dale, C.M. Turkelson, T.E. Solomon, Role of cholecystokinin in intestinal phase and meal-induced pancreatic secretion, Am. J. Physiol. - Gastrointest. Liver Physiol. 257 (1989) G782–G790. https://doi.org/10.1152/ajpgi.1989.257.5.g782.

[89] S.J. Konturek, B. Gabrys, J. Dubiel, Effect of exogenous and endogenous secretin on gastric and pancreatic secretion in cats., Am. J. Physiol. 217 (1969) 1110–1113. https://doi.org/10.1152/ajplegacy.1969.217.4.1110.

[90] S.J. Konturek, J. Tasler, M. Cieszkowski, K. Szewczyk, M. Hladij, Effect of cholecystokinin receptor antagonist on pancreatic responses to exogenous gastrin and cholecystokinin and to meal stimuli, Gastroenterology. 94 (1988) 1014–1023. https://doi.org/10.1016/0016-5085(88)90561-6.

[91] P. Hildebrand, C. Beglinger, K. Gyr, J.B.M.J. Jansen, L.C. Rovati, M. Zuercher, C.B.H.W. Lamers, I. Setnikar, G.A. Stalder, Effects of a cholecystokinin receptor antagonist on intestinal phase of pancreatic and biliary responses in man, J. Clin. Invest. 85 (1990) 640–646. https://doi.org/10.1172/JCI114486.

[92] A. Gabryelewicz, E. Kulesza, S.J. Konturek, Comparison of loxiglumide, a cholecystokinin receptor antagonist, and atropine on hormonal and meal-stimulated pancreatic secretion in man, Scand. J. Gastroenterol. 25 (1990) 731–738. https://doi.org/10.3109/00365529008997600.

[93] L.R. Johnson, K.E. Barret, F.K. Gishan, J.L. Merchant, H.M. Said, J.D. Wood, Physiology of the Gastrointestinal Tract, Elsevier, 2006. https://doi.org/10.1016/B978-0-12-088394-3.X5000-4.

[94] C. Niederau, A. Sonnenberg, J. Erckenbrecht, Effects of intravenous infusion of amino acids, fat, or glucose on unstimulated pancreatic secretion in healthy humans, Dig. Dis. Sci. 30 (1985) 445–455. https://doi.org/10.1007/BF01318177.

[95] X. Chen, A.K. Walker, J.R. Strahler, E.S. Simon, S.L. Tomanicek-Volk, B.B. Nelson, M.C. Hurley, S.A. Ernst, J.A. Williams, P.C. Andrew, Organellar proteomics: Analysis of pancreatic zymogen granule membranes, Mol. Cell. Proteomics. 5 (2006) 306–312. https://doi.org/10.1074/mcp.M500172-MCP200.

[96] M.J. Rindler, C.F. Xu, I. Gumper, N.N. Smith, T.A. Neubert, Proteomic analysis of pancreatic zymogen granules: Identification of new granule proteins, J. Proteome Res. 6 (2007) 2978–2992. https://doi.org/10.1021/pr0607029.

[97] X. Chen, P.J. Ulintz, E.S. Simon, J.A. Williams, P.C. Andrews, Global topology analysis of pancreatic zymogen granule membrane proteins, Mol. Cell. Proteomics. 7 (2008)

2323–2336. https://doi.org/10.1074/mcp.M700575-MCP200.

[98]   E. Gasteiger, C. Hoogland, A. Gattiker, M.R. Wilkins, R.D. Appel, A. Bairoch, others, Protein identification and analysis tools on the ExPASy server, in: Proteomics Protoc. Handb., Springer, 2005: pp. 571–607.

[99]   A.R. Khan, M.N.G. James, Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes, Protein Sci. 7 (1998) 815–836. https://doi.org/10.1002/pro.5560070401.

[100] P.J. Keller, B.J. Allan, The protein composition of human pancreatic juice., J. Biol. Chem. 242 (1967) 281–287. https://doi.org/10.1016/s0021-9258(19)81461-8.

[101] Y. Kitamoto, X. Yuan, Q. Wu, D.W. McCourt, J.E. Sadler, Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of a distinctive assortment of domains, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 7588–7592. https://doi.org/10.1073/pnas.91.16.7588.

[102] J. Kay, B. Kassell, The autoactivation of trypsinogen., J. Biol. Chem. 246 (1971) 6661–6665. https://doi.org/10.1016/s0021-9258(19)34166-3.

[103] J.M. Chen, Z. Kukor, C. Le Maréchal, M. Tóth, L. Tsakiris, O. Raguénès, C. Férec, M. Sahin-Tóth, Evolution of Trypsinogen Activation Peptides, Mol. Biol. Evol. 20 (2003) 1767–1777. https://doi.org/10.1093/molbev/msg183.

[104] R.H. Pfützer, M.M. Barmada, a P. Brunskill, R. Finch, P.S. Hart, J. Neoptolemos, W.F. Furey, D.C. Whitcomb, SPINK1/PSTI polymorphisms act as disease modifiers in familial and idiopathic chronic pancreatitis., Gastroenterology. 119 (2000) 615–623. https://doi.org/10.1053/gast.2000.18017.

[105] M. Laskowski Jr, I. Kato, Protein inhibitors of proteinases, Annu. Rev. Biochem. 49 (1980) 593–626.

[106] H. Rinderknecht, N.H. Stace, I.G. Renner, Effects of chronic alcohol abuse on exocrine pancreatic secretion in man, Dig. Dis. Sci. 30 (1985) 65–71. https://doi.org/10.1007/BF01318373.

[107] V.N. Kalinin, J.T. Kaifi, H. Schwarzenbach, A.S. Sergeyev, B.C. Link, D. Bogoevski, Y. Vashist, J.R. Izbicki, E.F. Yekebas, Association of rare SPINK1 gene mutation with another base substitution in chronic pancreatitis patients, World J. Gastroenterol. 12 (2006) 5352–5356. https://doi.org/10.3748/wjg.v12.i33.5352.

[108] R. Szmola, M. Sahin-Tóth, Chymotrypsin C (caldecrin) promotes degradation of human cationic trypsin: Identity with Rinderknecht's enzyme Y, Proc. Natl. Acad. Sci. U. S. A.

104 (2007) 11227–11232. https://doi.org/10.1073/pnas.0703714104.

[109] Z. Nemoda, M. Sahin-Tóth, Chymotrypsin C (caldecrin) stimulates autoactivation of human cationic trypsinogen, J. Biol. Chem. 281 (2006) 11879–11886. https://doi.org/10.1074/jbc.M600124200.

[110] Z. Jancsó, M. Sahin-Tóth, Tighter control by chymotrypsin C (CTRC) explains lack of association between human anionic trypsinogen and hereditary pancreatitis, J. Biol. Chem. 291 (2016) 12897–12905. https://doi.org/10.1074/jbc.M116.725374.

[111] S.J. Pandol, A.K. Saluja, C.W. Imrie, P.A. Banks, Acute Pancreatitis: Bench to the Bedside, Gastroenterology. 133 (2007) 1056.e1-1056.e25. https://doi.org/10.1053/j.gastro.2007.08.026.

[112] P.A. Banks, M.L. Freeman, R. Fass, D.S. Baroni, E.A. Mutlu, D.E. Bernstein, H.P. Parkman, A.E. Bharucha, C. Prather, W.R. Brugge, D.S. Pratt, L. Chang, A.C. Roach, W. Chey, R.E. Sampliner, M.E. Cohen, S. Sridhar, J.T. Cunningham, N. Vakil, S.A. Edmundowicz, M.A. Valdovinos, J.M. Inadomi, B.C.Y. Wong, T.R. Koch, A.M. Zfass, Practice guidelines in acute pancreatitis, Am. J. Gastroenterol. 101 (2006) 2379–2400. https://doi.org/10.1111/j.1572-0241.2006.00856.x.

[113] N. Hazra, M. Gulliford, Evaluating pancreatitis in primary care: A population-based cohort study, Br. J. Gen. Pract. 64 (2014) e295–e301. https://doi.org/10.3399/bjgp14X679732.

[114] D. Yadav, A.B. Lowenfels, The epidemiology of pancreatitis and pancreatic cancer, Gastroenterology. 144 (2013) 1252–1261. https://doi.org/10.1053/j.gastro.2013.01.068.

[115] J.P. Iannuzzi, J.A. King, J.H. Leong, J. Quan, J.W. Windsor, D. Tanyingoh, S. Coward, N. Forbes, S.J. Heitman, A.A. Shaheen, M. Swain, M. Buie, F.E. Underwood, G.G. Kaplan, Global Incidence of Acute Pancreatitis Is Increasing Over Time: A Systematic Review and Meta-Analysis, Gastroenterology. 162 (2022) 122–134. https://doi.org/10.1053/j.gastro.2021.09.043.

[116] B. Etemad, D.C. Whitcomb, Chronic pancreatitis: Diagnosis, classification, and new genetic developments, Gastroenterology. 120 (2001) 682–707. https://doi.org/10.1053/gast.2001.22586.

[117] C. Nøjgaard, U. Becker, P. Matzen, J.R. Andersen, C. Holst, F. Bendtsen, Progression from acute to chronic pancreatitis: Prognostic factors, mortality, and natural course, Pancreas. 40 (2011) 1195–1200. https://doi.org/10.1097/MPA.0b013e318221f569.

[118] D. Yadav, M. O'Connell, G.I. Papachristou, Natural history following the first attack of

acute pancreatitis, Am. J. Gastroenterol. 107 (2012) 1096–1103. https://doi.org/10.1038/ajg.2012.126.

[119] C. Umapathy, A. Raina, S. Saligram, G. Tang, G.I. Papachristou, M. Rabinovitz, J. Chennat, H. Zeh, A.H. Zureikat, M.E. Hogg, K.K. Lee, M.I. Saul, D.C. Whitcomb, A. Slivka, D. Yadav, Natural History After Acute Necrotizing Pancreatitis: a Large US Tertiary Care Experience, J. Gastrointest. Surg. 20 (2016) 1844–1853. https://doi.org/10.1007/s11605-016-3264-2.

[120] D.C. Whitcomb, Hereditary pancreatitis: New insights into acute and chronic pancreatitis, Gut. 45 (1999) 317–322. https://doi.org/10.1136/gut.45.3.317.

[121] A. Schneider, D.C. Whitcomb, Hereditary pancreatitis: A model for inflammatory diseases of the pancreas, Bailliere's Best Pract. Res. Clin. Gastroenterol. 16 (2002) 347–363. https://doi.org/10.1053/bega.2002.0311.

[122] H. Völzke, S.E. Baumeister, D. Alte, W. Hoffmann, C. Schwahn, P. Simon, U. John, M.M. Lerch, Independent risk factors for gallstone formation in a region with high cholelithiasis prevalence, Digestion. 71 (2005) 97–105. https://doi.org/10.1159/000084525.

[123] S. Buch, C. Schafmayer, H. Vlzke, M. Seeger, J.F. Miquel, S.C. Sookoian, J.H. Egberts, A. Arlt, C.J. Pirola, M.M. Lerch, U. John, A. Franke, O. Von Kampen, M. Brosch, M. Nothnagel, W. Kratzer, B.O. Boehm, D.C. Brring, S. Schreiber, M. Krawczak, J. Hampe, Loci from a genome-wide analysis of bilirubin levels are associated with gallstone risk and composition, Gastroenterology. 139 (2010) 1942–1951. https://doi.org/10.1053/j.gastro.2010.09.003.

[124] M. Rünzi, A. Saluja, M.M. Lerch, R. Dawra, H. Nishino, M.L. Steer, Early ductal decompression prevents the progression of biliary pancreatitis: An experimental study in the opossum, Gastroenterology. 105 (1993) 157–164. https://doi.org/10.1016/0016-5085(93)90021-4.

[125] C.A. Hernández, M.M. Lerch, Sphincter stencels and gallstone migration through the biliary tract, Lancet. 341 (1993) 1371–1373. https://doi.org/10.1016/0140-6736(93)90942-A.

[126] F. Lammert, M. Acalovschi, G. Ercolani, K.J. van Erpecum, K.S. Gurusamy, C.J. van Laarhoven, P. Portincasa, EASL Clinical Practice Guidelines on the prevention, diagnosis and treatment of gallstones, J. Hepatol. 65 (2016) 146–181. https://doi.org/10.1016/j.jhep.2016.03.005.

[127] C.F. Frey, H. Zhou, D.J. Harvey, R.H. White, The incidence and case-fatality rates of acute biliary, alcoholic, and idiopathic pancreatitis in California, 1994-2001, Pancreas. 33 (2006) 336–344. https://doi.org/10.1097/01.mpa.0000236727.16370.99.

[128] P.G. Lankisch, M. Apte, P.A. Banks, Acute pancreatitis, Lancet. 386 (2015) 85–96. https://doi.org/10.1016/S0140-6736(14)60649-8.

[129] A.I. Orabi, A.U. Shah, K. Muili, Y. Luo, S.M. Mahmood, A. Ahmad, A. Reed, S.Z. Husain, Ethanol enhances carbachol-induced protease activation and accelerates Ca2+ waves in isolated rat pancreatic acini, J. Biol. Chem. 286 (2011) 14090–14097. https://doi.org/10.1074/jbc.M110.196832.

[130] A. Lugea, J. Gong, J. Nguyen, J. Nieto, S.W. French, S.J. Pandol, Cholinergic mediation of alcohol-induced experimental pancreatitis, Alcohol. Clin. Exp. Res. 34 (2010) 1768–1781. https://doi.org/10.1111/j.1530-0277.2010.01264.x.

[131] H. Wu, K.K. Bhopale, G.A.S. Ansari, B.S. Kaphalia, Ethanol-induced cytotoxicity in rat pancreatic acinar AR42J cells: Role of fatty acid ethyl esters, Alcohol Alcohol. 43 (2008) 1–8. https://doi.org/10.1093/alcalc/agm044.

[132] C.A. Best, M. Laposata, Fatty acid ethyl esters: Toxic non-oxidative metabolites of ethanol and markers of ethanol intake, Front. Biosci. 8 (2003) 202–217.

[133] J. Werner, M. Laposata, C. Fernandez-Del Castillo, M. Saghir, R. V. Iozzo, K.B. Lewandrowski, A.L. Warshaw, Pancreatic injury in rats induced by fatty acid ethyl ester, a nonoxidative metabolite of alcohol, Gastroenterology. 113 (1997) 286–294. https://doi.org/10.1016/S0016-5085(97)70106-9.

[134] J. Werner, M. Saghir, A.L. Warshaw, K.B. Lewandrowski, M. Laposata, R. V. Iozzo, E.A. Carter, R.J. Schatz, C. Fernández-del Castillo, Alcoholic pancreatitis in rats: Injury from nonoxidative metabolites of ethanol, Am. J. Physiol. - Gastrointest. Liver Physiol. 283 (2002) G65–G73. https://doi.org/10.1152/ajpgi.00419.2001.

[135] D.N. Criddle, J. Murphy, G. Fistetto, S. Barrow, A. V. Tepikin, J.P. Neoptolemos, R. Sutton, O.H. Petersen, Fatty acid ethyl esters cause pancreatic calcium toxicity via inositol trisphosphate receptors and loss of ATP synthesis, Gastroenterology. 130 (2006) 781–793. https://doi.org/10.1053/j.gastro.2005.12.031.

[136] S. Dolai, T. Liang, P.P.L. Lam, N.A. Fernandez, S. Chidambaram, H.Y. Gaisano, Effects of ethanol metabolites on exocytosis of pancreatic acinar cells in rats, Gastroenterology. 143 (2012) 832–843. https://doi.org/10.1053/j.gastro.2012.06.011.

[137] S.G. Voronina, S.L. Barrow, A.W.M. Simpson, O. V. Gerasimenko, G. da Silva Xavier,

G.A. Rutter, O.H. Petersen, A. V. Tepikin, Dynamic Changes in Cytosolic and Mitochondrial ATP Levels in Pancreatic Acinar Cells, Gastroenterology. 138 (2010) 1976–1987. https://doi.org/10.1053/j.gastro.2010.01.037.

[138] P.S. Haber, J.S. Wilson, M. V. Apte, R.C. Pirola, Fatty acid ethyl esters increase rat pancreatic lysosomal fragility, J. Lab. Clin. Med. 121 (1993) 759–764.

[139] J.S. Wilson, M.A. Korsten, M. V. Apte, M.C. Thomas, P.S. Haber, R.C. Pirola, Both ethanol consumption and protein deficiency increase the fragility of pancreatic lysosomes, J. Lab. Clin. Med. 115 (1990) 749–755.

[140] D.N. Criddle, R. Sutton, O.H. Petersen, Role of Ca2+ in pancreatic cell death induced by alcohol metabolites, J. Gastroenterol. Hepatol. 21 (2006) S14–S17. https://doi.org/10.1111/j.1440-1746.2006.04577.x.

[141] L.I. Cosen-Binker, M.G. Binker, C.C. Wang, W. Hong, H.Y. Gaisano, VAMP8 is the v-SNARE that mediates basolateral exocytosis in a mouse model of alcoholic pancreatitis, J. Clin. Invest. 118 (2008) 2535–2551. https://doi.org/10.1172/JC134672.

[142] G. Fluhr, J. Mayerle, E. Weber, A. Aghdassi, P. Simon, T. Gress, T. Seufferlein, J. Mössner, A. Stallmach, T. Rösch, M. Müller, B. Siegmund, P. Büchner-Steudel, I. Zuber-Jerger, M. Kantowski, A. Hoffmeister, J. Rosendahl, T. Linhart, J. Maul, L. Czakó, P. Hegyi, M. Kraft, G. Engel, T. Kohlmann, A. Glitsch, T. Pickartz, C. Budde, C. Nitsche, K. Storck, M.M. Lerch, Pre-Study protocol MagPEP: A multicentre randomized controlled trial of magnesium sulphate in the prevention of post-ERCP pancreatitis, BMC Gastroenterol. 13 (2013) 1–6. https://doi.org/10.1186/1471-230X-13-11.

[143] F.U. Weiss, F. Laemmerhirt, M.M. Lerch, Etiology and risk factors of acute and chronic pancreatitis, Visc. Med. 35 (2019) 73–81. https://doi.org/10.1159/000499138.

[144] O. Watanabe, F.M. Baccino, M.L. Steer, J. Meldolesi, Supramaximal caerulein stimulation and ultrastructure of rat pancreatic acinar cell: Early morphological changes during development of experimental pancreatitis, Am. J. Physiol. - Gastrointest. Liver Physiol. 9 (1984) G457–G467. https://doi.org/10.1152/ajpgi.1984.246.4.g457.

[145] A. Saluja, S. Hashimoto, M. Saluja, R.E. Powers, J. Meldolesi, M.L. Steer, Subcellular redistribution of lysosomal enzymes during caerulein-induced pancreatitis, Am. J. Physiol. - Gastrointest. Liver Physiol. 253 (1987) G508–G516. https://doi.org/10.1152/ajpgi.1987.253.4.g508.

[146] C. Figarella, B. Miszczuk-Jamska, A.J. Barrett, Possible lysosomal activation of pancreatic zymogens. Activation of both human trypsinogens by cathepsin B and

spontaneous acid activation of human trypsinogen 1, Biol. Chem. Hoppe. Seyler. 369 (1988) 293–298.

[147] W. Halangk, M.M. Lerch, B. Brandt-Nedelev, W. Roth, M. Ruthenbuerger, T. Reinheckel, W. Domschke, H. Lippert, C. Peters, J. Deussing, Role of cathepsin B in intracellular trypsinogen activation and the onset of acute pancreatitis, J. Clin. Invest. 106 (2000) 773–781. https://doi.org/10.1172/JCI9411.

[148] O.A. Mareninova, K. Hermann, S.W. French, M.S. O'Konski, S.J. Pandol, P. Webster, A.H. Erickson, N. Katunuma, F.S. Gorelick, I. Gukovsky, A.S. Gukovskaya, Impaired autophagic flux mediates acinar cell vacuole formation and trypsinogen activation in rodent models of acute pancreatitis, J. Clin. Invest. 119 (2009) 3340–3355. https://doi.org/10.1172/JCI38674.

[149] I. Gukovsky, A.S. Gukovskaya, Impaired autophagy underlies key pathological responses of acute pancreatitis, Autophagy. 6 (2010) 428–429. https://doi.org/10.4161/auto.6.3.11530.

[150] A. Habtezion, A.S. Gukovskaya, S.J. Pandol, Acute Pancreatitis: A Multifaceted Set of Organelle and Cellular Interactions, Gastroenterology. 156 (2019) 1941–1950. https://doi.org/10.1053/j.gastro.2018.11.082.

[151] P.M. Vue, K. McFann, M.R. Narkewicz, Genetic Mutations in Pediatric Pancreatitis, Pancreas. 45 (2016) 992–996. https://doi.org/10.1097/MPA.0000000000000589.

[152] J.J. Palermo, T.K. Lin, L. Hornung, C. Alexander Valencia, A. Mathur, K. Jackson, L. Fei, M. Abu-El-Haija, Genophenotypic analysis of pediatric patients with acute recurrent and chronic pancreatitis, Pancreas. 45 (2016) 1347–1352. https://doi.org/10.1097/MPA.0000000000000655.

[153] H. Witt, M. Sahin-Tóth, O. Landt, J.M. Chen, T. Kähne, J.P.H. Drenth, Z. Kukor, E. Szepessy, W. Halangk, S. Dahm, K. Rohde, H.U. Schulz, C. Le Maréchal, N. Akar, R.W. Ammann, K. Truninger, M. Bargetzi, E. Bhatia, C. Castellani, G.M. Cavestro, M. Cerny, G. Destro-Bisol, G. Spedini, H. Eiberg, J.B.M.J. Jansen, M. Koudova, E. Rausova, M. Macek, N. Malats, F.X. Real, H.J. Menzel, P. Moral, R. Galavotti, P.F. Pignatti, O. Rickards, J. Spicak, N.O. Zarnescu, W. Böck, T.M. Gress, H. Friess, J. Ockenga, H. Schmidt, R. Pfützer, M. Löhr, P. Simon, F.U. Weiss, M.M. Lerch, N. Teich, V. Keim, T. Berg, B. Wiedenmann, W. Luck, D.A. Groneberg, M. Becker, T. Keil, A. Kage, J. Bernardova, M. Braun, C. Güldner, J. Halangk, J. Rosendahl, U. Witt, M. Treiber, R. Nickel, C. Férec, A degradation-sensitive anionic trypsinogen (PRSS2) variant protects against chronic pancreatitis, Nat. Genet. 38 (2006) 668–673.

https://doi.org/10.1038/ng1797.

[154] N. Teich, J. Rosendahl, M. Tóth, J. Mössner, M. Sahin-Tóth, Mutations of human cationic trypsinogen (PRSS1) and chronic pancreatitis, Hum. Mutat. 27 (2006) 721–730. https://doi.org/10.1002/humu.20343.

[155] O. Király, T. Wartmann, M. Sahin-Tóth, Missense mutations in pancreatic secretory trypsin inhibitor (SPINK1) cause intracellular retention and degradation, Gut. 56 (2007) 1433–1438. https://doi.org/10.1136/gut.2006.115725.

[156] C. Hu, L. Wen, L. Deng, C. Zhang, A. Lugea, H.Y. Su, R.T. Waldron, S.J. Pandol, Q. Xia, The Differential Role of Human Cationic Trypsinogen (PRSS1) p.R122H Mutation in Hereditary and Nonhereditary Chronic Pancreatitis: A Systematic Review and Meta-Analysis, Gastroenterol. Res. Pract. 2017 (2017). https://doi.org/10.1155/2017/9505460.

[157] A. Szabó, V. Toldi, L.D. Gazda, A. Demcsák, J. Tőzsér, M. Sahin-Tóth, Defective binding of SPINK1 variants is an uncommon mechanism for impaired trypsin inhibition in chronic pancreatitis, J. Biol. Chem. 296 (2021) 100343. https://doi.org/10.1016/J.JBC.2021.100343.

[158] H. Witt, W. Luck, H.C. Hennies, M. Claßen, A. Kage, U. Laß, O. Landt, M. Becker, Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis, Nat. Genet. 25 (2000) 213–216. https://doi.org/10.1038/76088.

[159] N. Teich, N. Bauer, J. Mossner, V. Keim, Mutational screening of patients with nonalcoholic chronic pancreatitis: identification of further trypsinogen variants1, Am. J. Gastroenterol. 97 (2002) 341–346. https://doi.org/10.1111/j.1572-0241.2002.05467.x.

[160] A. Boulling, J.M. Chen, I. Callebaut, C. Férec, Is the SPINK1 p. Asn34Ser missense mutation per se the true culprit within its associated haplotype?, WebmedCentral Genet. 3 (2012) WMC003084.

[161] J.M. Chen, C. Férec, The true culprit within the SPINK1 p.N34S-containing haplotype is still at large, Gut. 58 (2009) 478–480. https://doi.org/10.1136/gut.2008.170191.

[162] É. Kereszturi, O. Király, M. Sahin-Tóth, Minigene analysis of intronic variants in common SPINK1 haplotypes associated with chronic pancreatitis, Gut. 58 (2009) 545–549. https://doi.org/10.1136/gut.2008.164947.

[163] D.C. Whitcomb, M.C. Gorry, R.A. Preston, W. Furey, M.J. Sossenheimer, C.D. Ulrich, S.P. Martin, L.K. Gates, S.T. Amann, P.P. Toskes, R. Liddle, K. McGrath, G. Uomo, J.C.

Post, G.D. Ehrlich, Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene, Nat. Genet. 14 (1996) 141–145. https://doi.org/10.1038/ng1096-141.

[164] Z. Kukor, M. Tóth, G. Pál, M. Sahin-Tóth, Human cationic trypsinogen. Arg117 is the reactive site of an inhibitory surface loop that controls spontaneous zymogen activation, J. Biol. Chem. 277 (2002) 6111–6117. https://doi.org/10.1074/jbc.M110959200.

[165] C. Gaboriaud, L. Serre, O. Guy-Crotte, E. Forest, J.C. Fontecilla-Camps, Crystal structure of human trypsin 1: Unexpected phosphorylation of Tyr151, J. Mol. Biol. 259 (1996) 995–1010. https://doi.org/10.1006/jmbi.1996.0376.

[166] G. Katona, G.I. Berglund, J. Hajdu, L. Gráf, L. Szilágyi, Crystal structure reveals basis for the inhibitor resistance of human brain trypsin, J. Mol. Biol. 315 (2002) 1209–1218. https://doi.org/10.1006/jmbi.2001.5305.

[167] F. Nagel, A. Susemihl, N. Geist, K. Möhlis, G.J. Palm, M. Lammers, M. Delcea, Structural Basis of the Pancreatitis-Associated Autoproteolytic Failsafe Mechanism in Human Anionic Trypsin, J. Inflamm. Res. Volume 15 (2022) 3633–3642. https://doi.org/10.2147/jir.s367699.

[168] M. Kulke, F. Nagel, L. Schulig, N. Geist, M. Gabor, J. Mayerle, M.M. Lerch, A. Link, M. Delcea, A hypothesized mechanism for chronic pancreatitis caused by the n34s mutation of serine protease inhibitor kazal-type 1 based on conformational studies, J. Inflamm. Res. 14 (2021) 2111–2119. https://doi.org/10.2147/JIR.S304787.

[169] F. Nagel, G.J. Palm, N. Geist, T.C.R. McDonnell, A. Susemihl, B. Girbardt, J. Mayerle, M.M. Lerch, M. Lammers, M. Delcea, Structural and Biophysical Insights into SPINK1 Bound to Human Cationic Trypsin, Int. J. Mol. Sci. 23 (2022). https://doi.org/10.3390/ijms23073468.

[170] C.J. Farady, C.S. Craik, Mechanisms of Macromolecular Protease Inhibitors, ChemBioChem. 11 (2010) 2341–2346. https://doi.org/10.1002/cbic.201000442.

[171] D. Krowarsch, T. Cierpicki, F. Jelen, J. Otlewski, Canonical protein inhibitors of serine proteases, Cell. Mol. Life Sci. C. 60 (2003) 2427–2444. https://doi.org/10.1007/s00018-003-3120-x.

[172] E. Boros, F. Sebák, D. Héja, D. Szakács, K. Zboray, G. Schlosser, A. Micsonai, J. Kardos, A. Bodor, G. Pál, Directed Evolution of Canonical Loops and Their Swapping between Unrelated Serine Proteinase Inhibitors Disprove the Interscaffolding Additivity Model, J. Mol. Biol. 431 (2019) 557–575. https://doi.org/10.1016/j.jmb.2018.12.003.

[173] É. Kereszturi, M. Sahin-Tóth, Pancreatic cancer cell lines heterozygous for the SPINK1

p.N34S haplotype exhibit diminished expression of the variant allele, Pancreas. 46 (2017) e54–e55. https://doi.org/10.1097/MPA.0000000000000817.

Author Contributions

Article I   **Structural Basis of the Pancreatitis-Associated Autoproteolytic Failsafe Mechanism in Human Anionic Trypsin**

Nagel, F.; Susemihl, A.; Geist, N.; Möhlis, K.; Palm, G.J.; Lammers, M.; Delcea, M. *J. Inflamm. Res.* 2022, *15*, 3633–3642

Conceptualization, F.N. and M.D.; methodology, F.N., A.S., N.G., K.M., G.J.P.; software, F.N. and N.G.; validation, F.N., A.S., G.J.P. and N.G.; formal analysis, F.N.; investigation, F.N.; resources, M.L. and M.D.; data curation, F.N., G.J.P., N.G. and A.S.; writing—original draft preparation, F.N.; writing—review and editing, A.S., N.G., K.M., G.J.P., M.L. and M.D.; visualization, F.N.; supervision, M.D.; project administration, M.D.; funding acquisition, M.D.

Article II   **A Hypothesized Mechanism for Chronic Pancreatitis Caused by the N34S Mutation of Serine Protease Inhibitor Kazal-Type 1 Based on Conformational Studies**

Kulke, M.; Nagel, F.; Schulig, L.; Geist, N.; Gabor, M.; Mayerle, J.; Lerch, M.M.; Link, A.; Delcea, M. *J. Inflamm. Res. 2021, 14, 2111–2119.*

Conceptualization, M.K., F.N. and M.D.; methodology, M.K., F.N., L.S., N.S. and M.G. software, M.K., L.S. and N.G.; validation, M.K.; formal analysis, M.K.; investigation, M.K.; resources, M.D.; data curation, M.K., F.N., L.S. and N.G.; writing—original draft preparation, M.K. and F.N.; writing—review and editing, L.S., N.G., M.G., J.M., M.M.L., A.L. and M.D.; visualization, M.K.; supervision, M.D.; project administration, M.D.; funding acquisition, M.D.

Article III   **Structural and Biophysical Insights into SPINK1 Bound to Human Cationic Trypsin**

Nagel, F.; Palm, G.J.; Geist, N.; McDonnell, T.C.R.; Susemihl, A.; Girbardt, B.; Mayerle, J.; Lerch, M.M.; Lammers, M.; Delcea, M. *Int. J. Mol. Sci.* 2022, *23*.

Conceptualization, F.N. and M.D.; methodology, F.N., G.J.P., N.G., T.C.R.M., A.S. and B.G.; software, F.N. and N.G.; validation, F.N., G.J.P. and N.G.; formal analysis, F.N.; investigation, F.N.; resources, M.M.L., M.L. and M.D.; data curation, F.N., G.J.P., N.G. and A.S.; writing—original draft preparation, F.N.; writing—review and editing, G.J.P., N.G., T.C.R.M., A.S., J.M., M.M.L., M.L. and M.D.; visualization, F.N.; supervision, M.D.; project administration, M.D.; funding acquisition, M.D.

Article IV   **Identification of Kazal-inhibitor scaffolds with identical canonical binding loops and their effects on binding properties**

Nagel, F.; Susemihl, A.; Eulberg, T.E.; Delcea, M. *J. Biol. Chem.*, submitted.

Conceptualization, F.N., M.D.; methodology, F.N., A.S., T.E.; validation, F.N., A.S., T.E.; formal analysis, F.N.; investigation, F.N., A.S., T.E.; data curation, F.N.; writing-original draft preparation, F.N.; writing-review and editing, F.N., A.S. M.D.; visualization, F.N.; supervision, M.D.; project administration, M.D.; funding acquisition, M.D.

_____   _____

Felix Nagel   Prof. Dr. Mihaela Delcea

Articles

Article I

ORIGINAL RESEARCH

Structural Basis of the Pancreatitis-Associated Autoproteolytic Failsafe Mechanism in Human Anionic Trypsin

Felix Nagel [ID][1], Anne Susemihl [ID][1,2], Norman Geist [ID][1], Kevin Möhlis[1,3], Gottfried J Palm [ID][4], Michael Lammers[4], Mihaela Delcea [ID][1]

[1]Biophysical Chemistry, Institute of Biochemistry, University of Greifswald, Greifswald, Germany; [2]Department of Hematology and Oncology, Internal Medicine C, University of Greifswald, Greifswald, Germany; [3]Helmholtz Institute for Metabolic, Obesity and Vascular Research, Leipzig, Germany; [4]Synthetic and Structural Biochemistry, Institute of Biochemistry, University of Greifswald, Greifswald, Germany

Correspondence: Mihaela Delcea, Biophysical Chemistry, Institute of Biochemistry, University of Greifswald, Greifswald, Germany, Tel +49 3834 420 4423, Fax +49 3834 420 4377, Email delceam@uni-greifswald.de

Objective: The pathophysiological mechanisms underlying chronic pancreatitis (CP) are still poorly understood. Human cationic (TRY1) and anionic (TRY2) trypsins are the two major trypsin isoforms and their activities are tightly regulated within pancreatic acinar cells. Typically, they exist in a molar ratio of 2:1 (cationic:anionic). This ratio is reversed during chronic alcohol abuse, pancreatic cancer, or pancreatitis due to selectively upregulated expression of TRY2, causing anionic trypsin to become the predominant isoform. The involvement of TRY2 in pancreatitis is considered limited due to the absence of disease-causing mutations and its increased prevalence for autoproteolysis. However, exacerbated pancreatitis in TRY2 overexpressing mice was recently demonstrated. Here, we aim to elucidate the molecular structure of human anionic trypsin and obtain insights into the autoproteolytic regulation of tryptic activity.

Methods: Trypsin isoforms were recombinantly expressed in *E. coli*, purified and refolded. Enzymatic activities of all trypsin isoforms were determined and crystals of TRY2 were grown using the vapor-diffusion method. The structure was solved by molecular replacement and refined to a resolution of 1.7 Å. Equilibration molecular dynamics simulations were used to generate the corresponding TRY1–TRY1 model.

Results: All trypsin isoforms display similar kinetic properties. The crystal structure of TRY2 reveals that the enzyme crystallized in the autoproteolytic state with Arg122 placed in the S1 binding pocket and the corresponding loop cleaved. The TRY2–TRY2 dimer confirms a previously hypothesized autoinhibitory state with an unexpectedly large binding interface.

Conclusion: We provide a structure of TRY2, which is the predominant trypsin isoform in chronic pancreatitis and pancreatic cancer. A proposed autoinhibition mode was confirmed and the structural basis of the autoproteolytic failsafe mechanism elucidated.

Keywords: pancreas, pancreatitis, serine protease, crystal structure, autoproteolysis, R122H

Introduction

Chronic pancreatitis (CP) is a progressive inflammatory disorder of the pancreas that leads to fibrosis and exocrine and endocrine insufficiencies.[1–3] The underlying mechanisms for the development of the disease are diverse and are primarily caused by alcohol abuse or genetic risk factors but can also include smoking, autoimmunity and pancreatic injury.[4] In order to prevent pancreatitis, the tryptic activity within zymogen granules of pancreatic acinar cells has to be tightly regulated. Premature autoactivation is prevented by expression as zymogens with the propeptide inhibiting trypsin. Enterokinase, which activates trypsinogen by cleavage of the propeptide, is spatially restricted to the duodenum. Simultaneously, prematurely activated trypsin is inhibited by the serine protease inhibitor Kazal-type 1 (SPINK1) or can be degraded and cleared by autoproteolysis.[5,6] For the human cationic trypsinogen gene (*PRSS1*) a plethora of disease-causing mutations are known to play a role in hereditary pancreatitis.[7] One of the most prominent mutations is

R122H, which removes one highly exposed autocleavage site and thereby disables one of the failsafe mechanisms against pancreatitis.[8–11] Simultaneously, the Arg122 containing loop was proposed to possess weak inhibitory activity, which is meaningful at high trypsinogen concentrations found inside zymogen granules of pancreatic acinar cells.[12] Conversely, the gene for human anionic trypsin (PRSS2) is better known for its protective mutations.[13] The G191R mutation, in particular, has been shown to be protective against pancreatitis by introducing an additional autocleavage site that makes the enzyme more sensitive to degradation. However, due to tighter control of TRY2 by chymotrypsin C (CTRC), the role of this isoform in hereditary chronic pancreatitis is limited.[14,15] In healthy individuals the molar ratio of cationic trypsin (TRY1) and anionic trypsin (TRY2) is 2:1. During chronic alcohol abuse, pancreatic cancer and CP, this ratio inverts and TRY2 becomes the predominant isoform.[16–19] Until recently, it was unclear whether increased TRY2 secretion predisposes to pancreatitis or alternatively has a protective role against it.[16] However, transgenic expression of TRY2 in mice showed exacerbation of pancreatitis, clearly demonstrating the detrimental role of anionic trypsin.[20] Because TRY2 is selectively up-regulated in many patients with pancreatitis, the authors urged that inhibitors designed for pancreatitis therapy be tested against all isoforms rather than focusing on TRY1.[19,20] To understand the mechanisms by which TRY2 participates in pancreatitis, as well as to design novel inhibitors, a molecular model is indispensable. Probably due to the notoriously unstable nature of the PRSS2 isoform,[16] no crystal structure is available to date. Here, we report successful and easy expression, purification and refolding of all three major trypsin isoforms using a conventional His$_6$ tag. A crystal structure of active human anionic trypsin in a state representing the first step of autoproteolysis was solved. The TRY2 dimer represents an autoinhibited state of the enzyme, adding yet another mechanism by which tryptic activity is regulated inside pancreatic acinar cells.

Materials and Methods
Protein Expression and Purification

cDNA encoding for amino acids 16–247 of all trypsin isoforms was cloned in frame with an N-terminal hexahistidine tag (His$_6$-tag) and an HRV3C cleavage site into a pET47b vector (Novagen, Darmstadt, Germany). N-terminal truncations were introduced to remove the signal peptide sequences for expression in E. coli. Inactivating S200A mutations were introduced in all trypsin isoforms by site directed mutagenesis using QuikChange XL (Agilent Technologies, Santa Clara, CA, USA).

Trypsin isoforms and their inactive S200A variants were overexpressed in BL21 (DE3) E. coli (New England Biolabs, Frankfurt am Main, Germany). Cultures were grown in terrific broth medium with 100 µg mL^{-1} kanamycin at 37 °C. Expression was induced using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an OD$_{600}$ = 2 and carried out overnight at 25 °C. Cells were harvested by centrifugation (4000 g, 4 °C) and resuspended in 0.1 M Tris pH 8 and 5 mM EDTA (Tris-EDTA). Inclusion bodies were prepared essentially as described previously.[6] In brief, cells were lyzed by sonication and inclusion bodies were washed three times with Tris-EDTA before solubilization in 6 M guanidine hydrochloride (Gdn-HCl), 0.1 M Tris pH 8, 2 mM EDTA, 30 mM dithiothreitol (DTT). After incubation at 37 °C for 30 min and centrifugation, the solubilized inclusion bodies were added to refolding solution (0.9 M Gdn-HCl, 0.1 M Tris pH 8, 2 mM EDTA, 1 mM L-cysteine, 1 mM L-cystine, 15 µg mL^{-1} benzamidine) using a syringe pump at 20 µL min^{-1} at 4 °C. After overnight incubation, the refolded trypsin was loaded onto a HisTrap excel column (Cytiva, Freiburg, Germany). The column was extensively washed using 20 mM Hepes pH 8, 150 mM NaCl, 20 mM imidazole, 15 µg mL^{-1} benzamidine and eluted using 20 mM Hepes pH 8, 150 mM NaCl, 500 mM imidazole. Purified trypsinogen isoforms were dialyzed into 2 mM HCl and 2 mM CaCl$_2$ for storage before further use. During this step, residual benzamidine is also removed. All purification steps were carried out at 4 °C. Enterokinase activation was carried out at an enterokinase:trypsin ratio of 1:20,000 for WT and 1:50 for S200A variants in 20 mM Hepes pH 8, 50 mM NaCl and 2 mM CaCl$_2$ and was completed after a few hours at room temperature. Homogeneity was confirmed by analytical size exclusion chromatography using an ÄKTA micro platform (GE Healthcare, Freiburg, Germany) with a Superdex 200 Increase 3.2/300 column (Cytiva, Freiburg, Germany). Calibration curves are reported elsewhere.[21]

Determination of Enzyme Activity

Michaelis-Menten constants (K_m) and turnover numbers (k_{cat}) were determined using L-BAPA (N_α-Benzoyl-L-arginine 4-nitroanilide hydrochloride, Sigma-Aldrich, Taufkirchen, Germany) as substrate. The enzyme concentration was 30 nM and the L-BAPA concentrations ranged from 0.02 mM to 9.5 mM. Assays were carried out in 20 mM Hepes pH 7.4, 150 mM NaCl, 2 mM CaCl$_2$ and 0.05% Tween 20. Using the linear portion of the curves, initial conversion rates were determined at 37 °C. The amount of active enzyme was determined by active site titration using SPINK1. Absorbance of L-BAPA was monitored at 405 nm using a Cytation5 microplate reader (BioTek, VT, USA).

Crystallization and Data Collection

Enterokinase activated TRY2 was concentrated to 15 mg mL^{-1}. Crystals were obtained using the hanging drop vapor diffusion method at 20 °C in a reservoir solution containing 22% polyethylene glycol (PEG) 4000, 0.1 M Hepes pH 7.5 and 0.1 M sodium acetate. Crystals with a size of around 200 μm were cryocooled in liquid nitrogen. The diffraction data were collected at the BESSY 14.2 beamline. Data collection parameters are reported in Table 1.

Structure Determination

Data reduction was performed in XDS[22] and molecular replacement was carried out in Phaser[23] using human cationic trypsin (pdb: 1TRN) as search model. Phenix.refine[24] was used for refinement, and the model was built in Coot.[25] The crystal was isomorphous and contained two molecules per unit cell. NCS restraints were not applied. TLS refinement was used and refinement parameters are given in Table 1.

Molecular Dynamics Simulations

In order to generate a molecular model for the TRY1–TRY1 complex, the obtained crystal structure for TRY2 was mutated to the TRY1 sequence with Coot[25] and subjected to equilibration molecular dynamics simulations with NAMD 2.14.[26] The CHARMM36 forcefield was applied with CHARMM-GUI.[27,28] VMD 1.9.3 was utilized to solvate the protein complex with the TIP3P water model in a cubic cell with side length 8 nm.[29,30] NaCl concentrations were set to 0.15 M and the system was neutralized accordingly. The hydrogen mass repartitioning scheme (HMR) scaled the hydrogen masses in the protein to allow a longer timestep of 4 fs, while the SETTLE algorithm constrained the lengths of all bonds to hydrogen atoms. Pressures and temperatures were adjusted by a Langevin piston barostat at 100 fs period, 200 fs decay time and Langevin thermostat at 1 ps^{-1} damping time, respectively. Long-range electrostatic interactions were described by the smooth particle mesh Ewald method (sPME) and short-range interactions cutoffs were set to 1 nm with a switching function of 0.1 nm. The system was first subjected to standard energy minimization for 50k steps and then slowly heated from 100 K to 300 K in steps of 50 K, for 1 ns each. Afterward, the final production run proceeded at 310 K in an NPT ensemble for 2000 ns and samples were collected every 20 ps. RMSD values converged already after 100 ns, validating full equilibration of the model (Figure S1). RMSD and distance analyses were carried out with VMD.

Results

Human Anionic Trypsin Shows Specific Cleavage at the Arg122-Val123 Peptide Bond

Although the increased tendency for autoproteolysis of TRY2 is clearly established, the overall mechanism and sequence of digestion have remained elusive. To provide a structural framework for rationalizing the autoproteolytic failsafe mechanism, we determined the crystal structure of TRY2 in a state, representing the first step of autoproteolysis. After cleavage of the Arg122-Val123 peptide bond, Arg122 remains bound to TRY2 and is located in the S1 binding pocket (Schechter-Berger nomenclature).[31] The structure was solved at 1.7 Å and two TRY2 molecules formed a homodimer in the asymmetric unit (Figure 1A, Table 1). Specific cleavage of the Arg122-Val123 bond was demonstrated by SDS-PAGE, while the catalytically inactive TRY2 S200A variant remained intact even after enterokinase activation (Figure 1B). Lower bands in the TRY2 S200A lane correspond to impurities introduced by the high amounts of enterokinase needed for activation, due to the absence of autoactivation in the inactive enzyme. Additional bands in the TRY2 WT lane are secondary products due to cleavage at the autolytic Lys193-Asp194 site. Homogeneity of anionic

Table 1 Data Collection and Refinement Statistics of the TRY2 Homodimer

Data Collection	
Beamline	14.2 at BESSY
Wavelength (Å)	0.9184
Unit-cell parameters (Å) a, b, c in space group P2$_1$2$_1$2$_1$	59.06, 80.79, 87.94
Resolution (Å)	50–1.70 (1.80–1.70)
No. of unique reflections (Friedel pairs merged)	47,119 (7342)
Redundancy	12.9 (12.8)
Completeness (%)	99.6 (97.3)
R_{merge}	0.169 (1.252)
cc$_{1/2}$	0.999 (0.82)
$<I/\sigma\ (I)>$	11.9 (2.0)
Wilson B-factor (Å^2)	14.8
Refinement	
Resolution range (Å)	50–1.7 (1.75–1.7)
Completeness (%)	99.4 (95.0)
No. of reflections, working set	45,524 (3902)
No. of reflections, test set	1516 (130)
Final R_{work}	0.187 (0.231)
Final R_{free}	0.218 (0.263)
No. of non-H atoms	
Protein	3368
Water	770
Calcium	2
R.m.s. deviations	
Bond lengths (Å)	0.006
Angles (°)	0.902
Average B factors (Å^2)	
Protein	17.91
Water	29.32
Molprobity analysis	
Ramachandran most favored (%)	97.95
Ramachandran outliers (%)	0
Overall score	1.25
Clash score	4.69
PDB entry	7Z9F
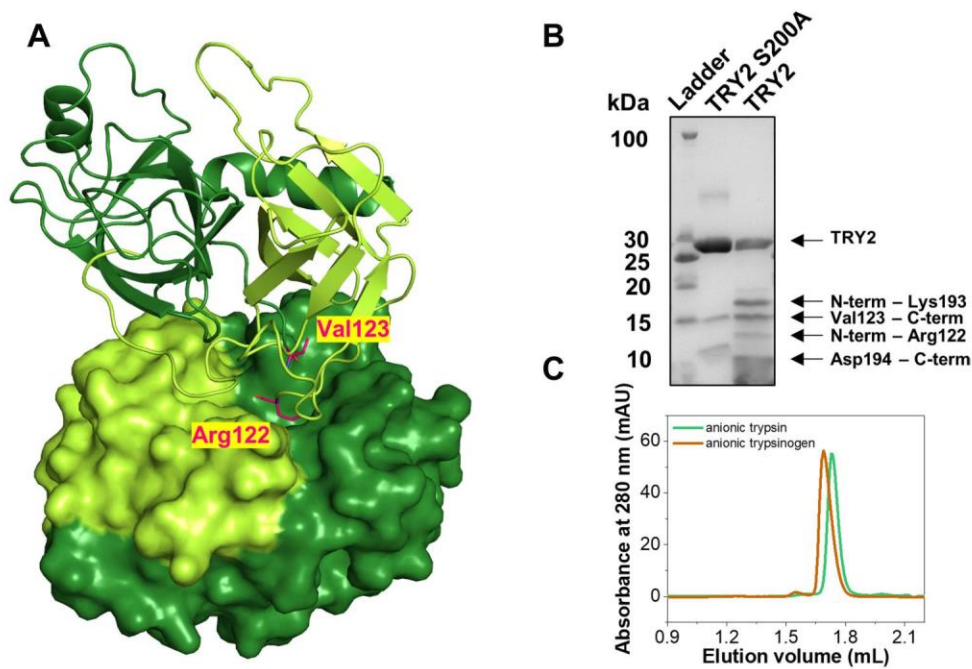
Figure 1 Structure of the TRY2 homodimer. (**A**) TRY2 N-terminus-Arg122 is colored yellow-green and Val123-C-terminus is colored in dark-green. The TRY2 – TRY2 complex shows one monomer in cartoon representation and one in surface representation. One monomer binds on top of TRY2 in a substrate-like manner and is rotated approximately 180° around the y-axis. Arg122 and Val123 are highlighted in pink. (**B**) SDS-PAGE of activated TRY2 WT and the catalytically inactive S200A variant. (**C**) Analytical size exclusion chromatograms of anionic trypsinogen and activated anionic trypsin.

trypsinogen and trypsin was verified by analytical size exclusion chromatography (Figure 1C). Despite the cleavage of TRY2 into two chains, they remain attached by disulfide bonds and the enzyme maintains its activity. While intact TRY2 can be seen in the SDS-PAGE, no electron density for an intact Arg122-Val123 bond was visible in the crystal structure (Figure 2A). At 20 μM all trypsin isoforms elute as monomers with only a small peak visible at 1.55 mL corresponding
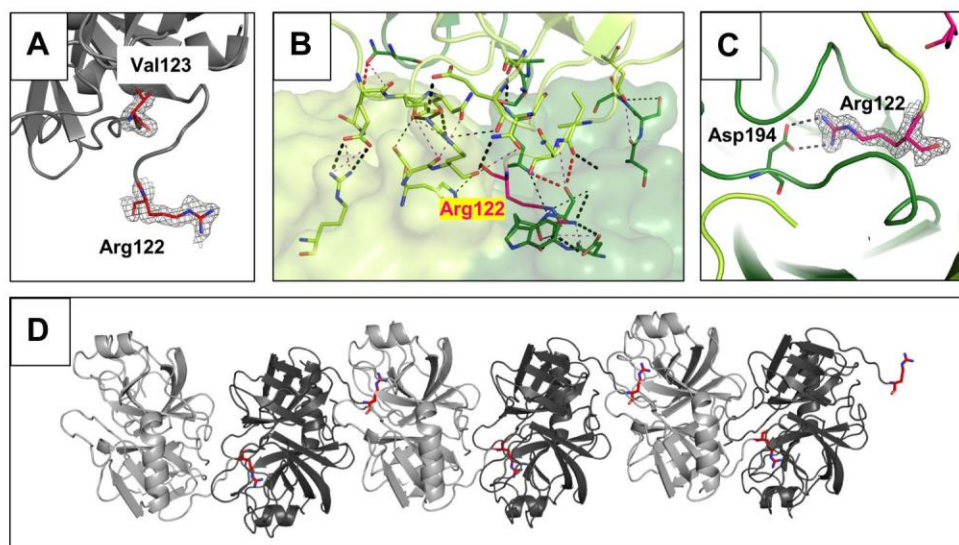
Figure 2 Arg122-Val123 cleavage site and specific interactions of the TRY2 dimer interface. (**A**) Close-up of the cleaved Arg122 loop. The 2Fo-Fc density map is shown 1.6 Å around Arg122 and Val123 and is contoured at 1.0 σ. No additional density for an intact loop is visible. (**B**) Interaction interface with interacting residues represented as sticks. The color scheme is the same as in Figure 1. (**C**) Arg122 specifically interacts with Asp194 in the S1 binding pocket. The 2Fo-Fc density is shown using the same parameters as in (**A**). (**D**) Crystallographic assembly of a TRY2 fiber with repetitive TRY2 dimers. Arg122 is highlighted in red. Two nicked TRY2 monomers (light and dark grey) form the asymmetric unit.

to the size of a trypsin dimer. Based on the TRY2-TRY2 complex crystal structure, we built a model of the TRY1-TRY1 dimer (Figure S2). The equilibrated TRY1 model is highly similar to the TRY2 crystal structure. The distance between Arg122 and Val123 is ~8 Å in the crystal structure of TRY2 and on average ~11 Å for the TRY1 model. Over the course of 2000 ns the shortest distance observed was ~8 Å (Figure S3).

The binding interface of the TRY2 dimer encompasses ~800 Å^2. Arg122 interacts with the specificity determining residue Asp194 (Figure 2B and C). Due to the specific nature of the TRY2-TRY2 interaction, it is likely that the observed dimer not only shows autoproteolysis, but also demonstrates the autoinhibitory state of the enzyme. Inhibitory activity of the Arg122 loop has already been proposed and weak inhibition was demonstrated.[12] Our structure shows that translation and 180° rotation about the translation axis is necessary to create the binding interface for a TRY2 dimer. Crystallographically, this results in an endless chain along the translation axis (Figure 2D). Due to the weakness of the interaction, only dimers, but no oligomers or even fibers were observed in solution. Unlike previous hypotheses, we found that the entire enzyme is involved in the autoinhibition interface, rather than just the Arg122 loop. In addition, the dimerization interface differs significantly from the already known α-chymotrypsin dimer, although the dimerization of α-chymotrypsin also involves an autocleavage site (Figure S4).[32] TRY2 likely undergoes induced fit movements after cleavage of the Arg122-Val123 scissile bond. Similar to known Laskowski inhibitors, TRY2 potentially exists in an equilibrium between cleaved and re-ligated Arg122-Val123 as seen on SDS-PAGE (Figure 1B). However, because of the large distance between both residues after cleavage, the cleavage reaction is probably highly favored.

Trypsin Isoforms Differ in Surface Charge but Not in Kinetic Properties

We compared our TRY2 crystal structure with existing crystal structures of cationic trypsin and mesotrypsin (pdb: 7QE9 and 1H4W, respectively). Despite their high sequence identity (>90%), the three isoforms differ in their isoelectric point.[33] Evolutionary, the existence of different trypsin isoforms is most likely rationalized by each isoform showing varying resistances to inhibitors, commonly encountered in foods. To identify local differences in surface charge, rather than the global charge of the entire enzyme, we used our newly obtained TRY2 structure to compare the overall charge distribution to the other trypsin isoforms (Figure 3). Mesotrypsin appears to be more positively charged near the S1
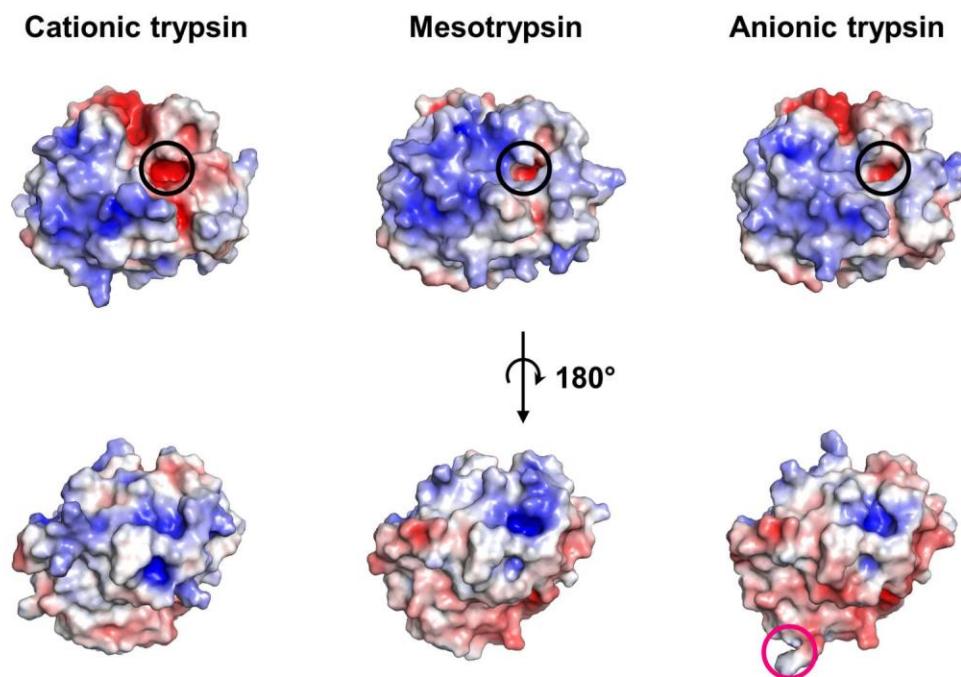

Figure 3 Surface charge comparison of the three major trypsin isoforms. Electrostatics were calculated at pH 7.4 and 150 mM NaCl concentration. The Pdb2pqr server and the adaptive poisson Boltzmann solver plugin in PyMOL were used.[41] The S1 binding pocket is highlighted by a black circle, while Arg122 in TRY2 is highlighted by a pink circle. Color scales are from red (−5 k_bT/e_c) to blue (+5 k_bT/e_c).
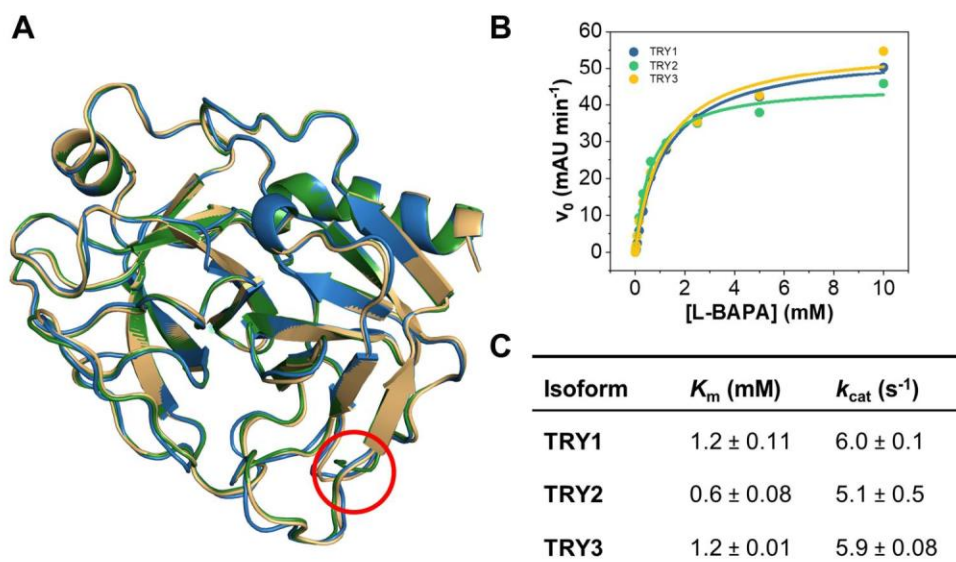
Figure 4 Structural and kinetic comparison of the three major trypsin isoforms. (**A**) Superposition of TRY1-3. TRY1, 2 and 3 are colored blue, green and yellow, respectively. The location of the Arg122 residue is highlighted by a red circle. (**B**) Representative Michaelis-Menten kinetics of all three isoforms using L-BAPA as substrate. (**C**) Kinetic parameters for each isoform. Reported values represent mean ± s.d. of three independent experiments.

binding pocket compared to cationic trypsin, but shows a more negative charge on the backside of the enzyme. Anionic trypsin is the most negatively charged of the three isoforms with differences being most prominent on the backside of the enzyme as well.

While there are differences in surface charge near the S1 binding pocket for all isoforms, the binding pocket itself remains negatively charged, which is reflected in all trypsin variants preferring either lysine or arginine as P1 residue. Despite their differences in charge, all isoforms are structurally very similar and show no variation in the overall scaffold (Figure 4A). In addition, their kinetic profiles are almost indistinguishable when small chromogenic substrates are used (Figure 4B and C). Using L-BAPA as model substrate, TRY1 and mesotrypsin (TRY3) display identical Michaelis-Menten constants (K_m = 1.2 mM) and turnover numbers (k_{cat} = 6 s^{-1}). TRY2 shows slightly lower K_m values (K_m = 0.6 mM) and similar turnover numbers (k_{cat} = 5.1 s^{-1}). Although they are all kinetically similar when using a small chromogenic substrate, larger or proteinaceous substrates might be more affected by the electrostatic environment surrounding the S1 binding pocket of each isoform.

Interestingly, Arg193 in mesotrypsin, which usually prevents inhibitor binding by steric hindrance, is oriented in a way that would not interfere with the formation of a trypsin dimer in the autoinhibited state.[34] Under physiological conditions, Tyr154 is sulfated, which causes a small decrease in affinity for the trypsin inhibitor SPINK1.[35] In the context of trypsin dimerization however, sulfation of this residue might even be beneficial by stabilizing the cleaved Arg122 loop and potentially the newly generated N-terminus (Figure S5). Stabilization of the loop after cleavage might aid in resynthesis of the peptide bond. Our data thereby support a model in which cleavage of the Arg122-Val123 bond is accompanied by self-inhibition of all trypsin isoforms at high concentrations while not affecting the overall activity of the enzyme at low concentrations.

Discussion

We demonstrate that human anionic trypsin selectively cleaves itself at the Arg122-Val123 bond. This cleavage and resulting interactions are of physiological relevance as cleavage of this bond is considered the first step of the autoproteolytic failsafe mechanism. In order to understand the molecular basis of this interaction, we solved the structure of a TRY2–TRY2 dimer at 1.7 Å and used it as a template to build a model of the corresponding TRY1–TRY1 dimer. We observe induced fit movements of Arg122 after cleavage, which is then deeply burrowed within the S1 binding pocket. SDS-PAGE confirms the formation of two fragments of appropriate size and suggests the existence of an equilibrium

between cleavage and re-ligation of the Arg122 loop. Even after cleavage, analytical size exclusion chromatography and enzymatic assays show a functional enzyme.

Our finding that TRY2 cleaves the Arg122-Val123 loop is consistent with biochemical assays involving different concentrations of NaCl and CaCl$_2$, different pH values, and mutagenesis studies involving R122H or R122A mutations.[13,14,16] The importance of this specific loop is also reflected in R122H being the most prominent pancreatitis-causing mutation in the cationic trypsin gene (*PRSS1*). R122H or R122C suppress the autolytic site within TRY1, thereby disrupting autoproteolysis and clearance of prematurely activated trypsin.[8–11] Hence, a molecular model of a TRY1–TRY1 dimer was built on the basis of the TRY2–TRY2 dimer. The observed structures and interactions are essentially the same in both complexes and the findings of the anionic trypsin dimer can be translated to the cationic trypsin dimer.

Previous studies report that the Arg122 loop is not only the main autolytic site, but also functions as a weak inhibitory loop.[12] Given the conformation observed in our crystal structure, it is apparent that binding of Arg122 blocks entry into the S1 binding pocket and thereby inhibits further cleavage. Our SDS-PAGE also shows intact TRY2 even after very long incubation times, which further supports previous hypotheses, where the Arg122-Val123 can be re-ligated by trypsin via a similar mechanism observed in many canonical serine protease inhibitors following the Laskowski mechanism.[36] In addition to providing proof for this proposed mechanism, we also show that not only the Arg122 loop is involved in trypsin autoinhibition, but also the entire backside of the enzyme. Although the entire enzyme participates in inhibition, there are few specific interactions, which probably explain the high K_i values, estimated at around 80 μM.[12] At these K_i values, meaningful autoinhibition can only occur within zymogen granules of pancreatic acinar cells, where trypsin concentrations can exceed 1 mM and enzymes exist in a semi-crystalline state.[18,37]

While the Arg122 loop functions similarly to known Laskowski inhibitors, there are several distinguishing features. Most importantly, in Laskowski inhibitors like SPINK1, the binding loop is very rigid, whereas the Arg122 loop is not stabilized by disulfide bonds or other intramolecular interactions. This causes the Arg122 and Val123 residues to be separated by more than 8 Å after cleavage, which is far from a conformation that would allow efficient re-ligation. Furthermore, during our 2 μs MD simulation of the TRY1 dimer, Arg122 and Val123 never got closer than 8 Å. We therefore assume that, while re-ligation of this bond is possible, the equilibrium is tilted towards cleavage, which is in accordance with the reported weak inhibitory potential of the Arg122 loop.

Due to the tight control of TRY2 by chymotrypsin C, no disease-causing mutations are known within the *PRSS2* gene.[14] Nevertheless, TRY2 contributes to the pancreatitis risk, which is demonstrated by the G191R mutation. By introducing an additional autolytic site, the mutation, which is enriched in the healthy control, causes a small protective effect.[13] In addition, the expression of TRY2 in transgenic mice revealed exacerbated pancreatitis, clearly demonstrating the involvement of TRY2 in pancreatitis.[20] The authors stressed that inhibitors directed at treating pancreatitis should be designed and tested against all trypsin isoforms.

To this day, no pharmacological therapy for pancreatitis exists. Treatment is mostly supportive, including administration of fluids, bowel rest, antibiotics and pain control.[38,39] Due to the complexity and elusiveness of the disease, development of therapeutics is challenging and results obtained from animal models are often not transferable to clinical studies.[40] Our findings add to the understanding of this highly complex disease and provide a structural framework from which novel therapeutics could be developed.

In addition to autoproteolysis, we show that tryptic activity is also regulated by self-inhibition following Arg122 cleavage, adding another mechanism, by which the pancreas elegantly protects itself from prematurely activated trypsin.

Conclusion

We solved the crystal structure of human anionic trypsin in a state representing the first step of autoproteolysis. The structural basis underlying the pathogenicity of the R122H variant in cationic trypsin was elucidated and a previously hypothesized autoinhibitory mechanism confirmed. Our data support a model in which autoproteolysis of the Arg122 loop is accompanied by self-inhibition, which contributes to the tight regulation of tryptic activity in pancreatic acinar cells.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Beyer G, Habtezion A, Werner J, Lerch MM, Mayerle J. Chronic pancreatitis. *Lancet*. 2020;396(10249):499–512. doi:10.1016/S0140-6736(20)31318-0

2. Angelopoulos N, Dervenis C, Goula A, et al. Endocrine pancreatic insufficiency in chronic pancreatitis. *Pancreatology*. 2005;5(2–3):122–131. doi:10.1159/000085264

3. Capurso G, Traini M, Piciucchi M, Signoretti M, Arcidiacono PG. Exocrine pancreatic insufficiency: prevalence, diagnosis, and management. *Clin Exp Gastroenterol*. 2019;12:129–139. doi:10.2147/CEG.S168266

4. Weiss FU, Laemmerhirt F, Lerch MM. Etiology and risk factors of acute and chronic pancreatitis. *Visc Med*. 2019;35(2):73–81. doi:10.1159/000499138

5. Hirota M, Ohmuraya M, Baba H. The role of trypsin, trypsin inhibitor, and trypsin receptor in the onset and aggravation of pancreatitis. *J Gastroenterol*. 2006;41(9):832–836. doi:10.1007/s00535-006-1874-2

6. Nagel F, Palm GJ, Geist N, et al. Structural and biophysical insights into SPINK1 bound to human cationic trypsin. *Int J Mol Sci*. 2022;23(7):3468. doi:10.3390/ijms23073468

7. Hegyi E, Sahin-Tóth M. Genetic risk in chronic pancreatitis: the trypsin-dependent pathway. *Dig Dis Sci*. 2017;62(7):1692–1701. doi:10.1007/s10620-017-4601-3

8. Whitcomb DC, Gorry MC, Preston RA, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat Genet*. 1996;14(2):141–145. doi:10.1038/ng1096-141

9. Teich N, Bauer N, Mössner J, Keim V. Mutational screening of patients with nonalcoholic chronic pancreatitis: identification of further trypsinogen variants. *Am J Gastroenterol*. 2002;97(2):341–346. doi:10.1111/j.1572-0241.2002.05467.x

10. Gorry MC, Gabbaizedeh D, Furey W, et al. Mutations in the cationic trypsinogen gene are associated with recurrent acute and chronic pancreatitis. *Gastroenterology*. 1997;113(4):1063–1068. doi:10.1053/gast.1997.v113.pm9322498

11. Witt H, Luck W, Becker M. A signal peptide cleavage site mutation in the cationic trypsinogen gene is strongly associated with chronic pancreatitis. *Gastroenterology*. 1999;117(1):7–10. doi:10.1016/s0016-5085(99)70543-3

12. Kukor Z, Tóth M, Pál G, Sahin-Tóth M. Human cationic trypsinogen: Arg (117) is the reactive site of an inhibitory surface loop that controls spontaneous zymogen activation. *J Biol Chem*. 2002;277(8):6111–6117. doi:10.1074/jbc.M110959200

13. Witt H, Sahin-Tóth M, Landt O, et al. A degradation-sensitive anionic trypsinogen (PRSS2) variant protects against chronic pancreatitis. *Nat Genet*. 2006;38(6):668–673. doi:10.1038/ng1797

14. Jancsó Z, Sahin-Tóth M. Tighter control by chymotrypsin C (CTRC) explains lack of association between human anionic trypsinogen and hereditary Pancreatitis*♦. *J Biol Chem*. 2016;291(25):12897–12905. doi:10.1074/jbc.M116.725374

15. Weiss FU, Skube ME, Lerch MM. Chronic pancreatitis: an update on genetic risk factors. *Curr Opin Gastroenterol*. 2018;34(5):322–329. doi:10.1097/MOG.0000000000000461

16. Kukor Z, Tóth M, Sahin-Tóth M. Human anionic trypsinogen: properties of autocatalytic activation and degradation and implications in pancreatic diseases. *Eur J Biochem*. 2003;270(9):2047–2058. doi:10.1046/j.1432-1033.2003.03581.x

17. Guy O, Lombardo D, Bartelt DC, Amic J, Figarella C. Two human trypsinogens. purification, molecular properties, and N-terminal sequences. *Biochemistry*. 1978;17(9):1669–1675. doi:10.1021/bi00602a014

18. Rinderknecht H, Renner IG, Carmack C. Trypsinogen variants in pancreatic juice of healthy volunteers, chronic alcoholics, and patients with pancreatitis and cancer of the pancreas. *Gut*. 1979;20(10):886–891. doi:10.1136/gut.20.10.886

19. Rinderknecht H, Stace NH, Renner IG. Effects of chronic alcohol abuse on exocrine pancreatic secretion in man. *Dig Dis Sci*. 1985;30(1):65–71. doi:10.1007/BF01318373

20. Wan J, Haddock A, Edenfield B, Ji B, Bi Y. Transgenic expression of human PRSS2 exacerbates pancreatitis in mice. *Gut*. 2020;69(11):2051–2052. doi:10.1136/gutjnl-2019-320399

21. Susemihl A, Nagel F, Grabarczyk P, Schmidt CA, Delcea M. Easy expression and purification of fluorescent N-terminal BCL11B CCHC zinc finger domain. *Molecules*. 2021;26:24. doi:10.3390/molecules26247576

22. Kabsch W. XDS. *Acta Crystallogr Sect D*. 2010;66(2):125–132. doi:10.1107/S0907444909047337

23. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007;40(4):658–674. doi:10.1107/S0021889807021206

24. Liebschner D, Afonine PV, Baker ML, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr Sect D*. 2019;75(10):861–877. doi:10.1107/S2059798319011471

25. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr*. 2010;66(4):486–501. doi:10.1107/S0907444910007493

26. Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781–1802. doi:10.1002/jcc.20289

27. Huang J, MacKerell AD Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem*. 2013;34(25):2135–2145. doi:10.1002/jcc.23354

28. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem*. 2008;29(11):1859–1865. doi:10.1002/jcc.20945

29. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33–38. doi:10.1016/0263-7855(96)00018-5

30. Hopkins CW, Le Grand S, Walker RC, Roitberg AE. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J Chem Theory Comput*. 2015;11(4):1864–1874. doi:10.1021/ct5010406

31. Schechter I, Berger A. On the size of the active site in proteases. I Papain. *Biochem Biophys Res Commun*. 1967;27(2):157–162. doi:10.1016/S0006-291X(67)80055-X

32. Blevins RA, Tulinsky A. The refinement and the structure of the dimer of alpha-chymotrypsin at 1.67-A resolution. *J Biol Chem*. 1985;260 (7):4264–4275. doi:10.2210/pdb5cha/pdb

33. Scheele G, Bartelt D, Bieger W. Characterization of human exocrine pancreatic proteins by two-dimensional isoelectric focusing/sodium dodecyl sulfate gel electrophoresis. *Gastroenterology*. 1981;80(3):461–473. doi:10.1016/0016-5085(81)90007-X

34. Katona G, Berglund GI, Hajdu J, Gráf L, Szilágyi L. Crystal structure reveals basis for the inhibitor resistance of human brain trypsin11. *J Mol Biol*. 2002;315(5):1209–1218. doi:10.1006/jmbi.2001.5305

35. Szabó A, Toldi V, Gazda LD, Demcsák A, Tőzsér J, Sahin-Tóth M. Defective binding of SPINK1 variants is an uncommon mechanism for impaired trypsin inhibition in chronic pancreatitis. *J Biol Chem*. 2021;296:100343. doi:10.1016/j.jbc.2021.100343

36. Zakharova E, Horvath MP, Goldenberg DP. Structure of a serine protease poised to resynthesize a peptide bond. *Proc Natl Acad Sci USA*. 2009;106 (27):11034–11039. doi:10.1073/pnas.0902463106

37. Goncz KK, Behrsing R, Rothman SS. The protein content and morphogenesis of zymogen granules. *Cell Tissue Res*. 1995;280(3):519–530. doi:10.1007/BF00318356

38. Kambhampati S, Park W, Habtezion A. Pharmacologic therapy for acute pancreatitis. *World J Gastroenterol*. 2014;20(45):16868–16880. doi:10.3748/wjg.v20.i45.16868

39. Leppäniemi A, Tolonen M, Tarasconi A, et al. 2019 WSES guidelines for the management of severe acute pancreatitis. *World J Emerg Surg*. 2019;14(1):27. doi:10.1186/s13017-019-0247-0

40. Pezzilli R. Pharmacotherapy for acute pancreatitis. *Expert Opin Pharmacother*. 2009;10(18):2999–3014. doi:10.1517/14656560903382630

41. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci*. 2001;98(18):10037–10041. doi:10.1073/pnas.181342398
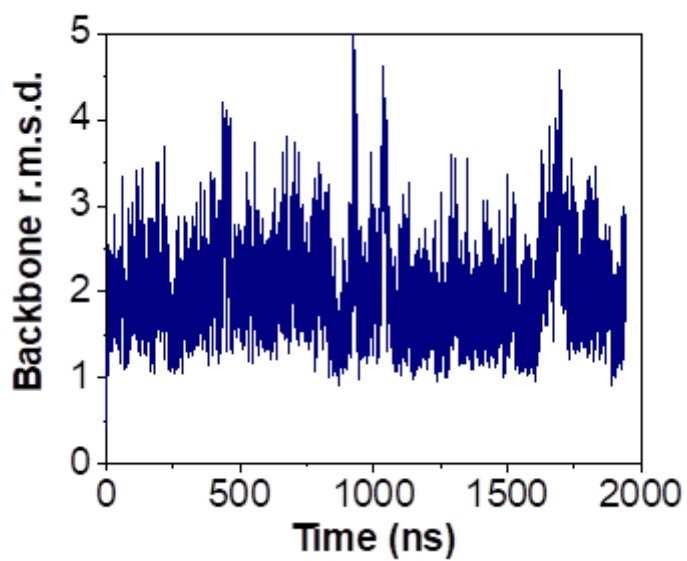
Supporting Information

Figure S10: Backbone r.m.s.d. of the TRY1-TRY1 dimer equilibration.
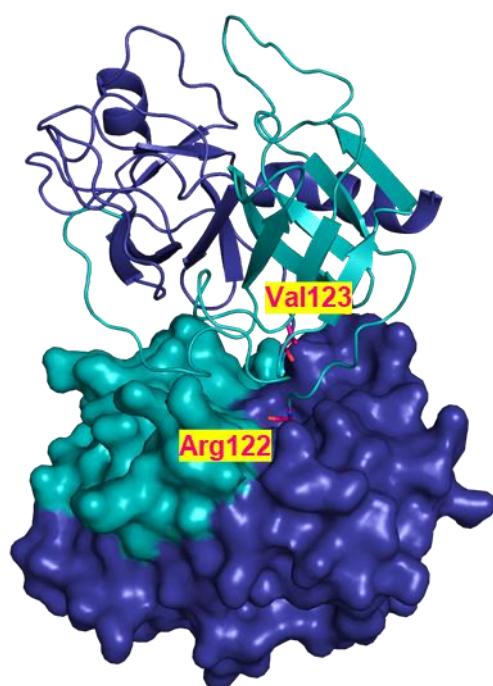

Figure S11: Model of the TRY1 homodimer. TRY1– Arg122 is colored cyan and Val123 is colored in dark-blue. The TRY1 – TRY1 complex shows one monomer in cartoon representation and one in surface representation.
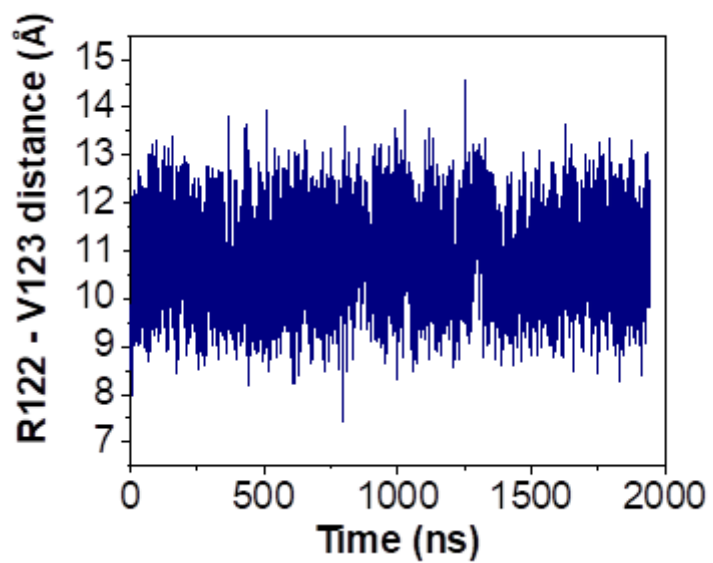
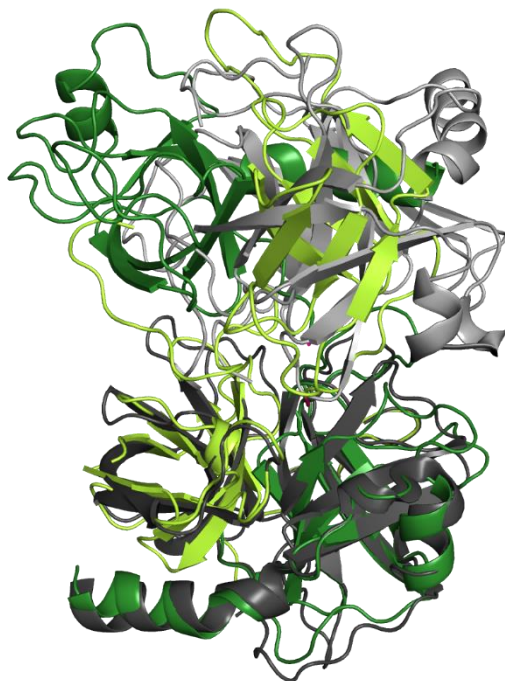**Figure S12: Arg122 – Val123 distance in TRY1 during MDS.**

Figure S13: Superposition of the TRY2 – TRY2 and α-CT dimer. The lower chains were superimposed and TRY2 is shown in green and the α-CT dimer is shown in gray. Data from Blevins et al.[32]
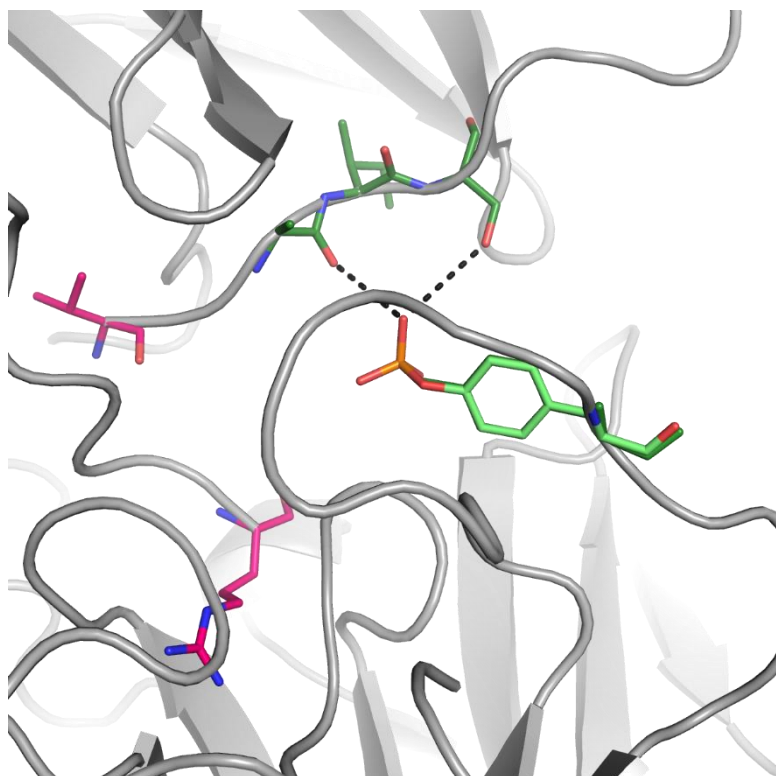
Figure S14: Influence of Tyr154 sulfation. Arg122 and Val123 are highlighted in pink. The sulfate moiety was modeled into the TRY2 structure with the sulfated tyrosine and interacting residues shown in light and dark green, respectively.

Article II

HYPOTHESIS

A Hypothesized Mechanism for Chronic Pancreatitis Caused by the N34S Mutation of Serine Protease Inhibitor Kazal-Type 1 Based on Conformational Studies

Martin Kulke [ID][1]
Felix Nagel [ID][1]
Lukas Schulig [ID][2]
Norman Geist [ID][1]
Marcel Gabor[1]
Julia Mayerle[3]
Markus M Lerch[4]
Andreas Link [ID][2]
Mihaela Delcea [ID][1]

[1]Institute of Biochemistry, University of Greifswald, Greifswald, Germany; [2]Institute of Pharmacy, University of Greifswald, Greifswald, Germany; [3]Department of Medicine II, Ludwig-Maximilian University of Munich, Munich, Germany; [4]Department of Medicine a, University Medicine Greifswald, Greifswald, Germany

Correspondence: Mihaela Delcea;
Martin Kulke
Institute of Biochemistry, University of Greifswald, Felix-Hausdorff-Straße 4, Greifswald, 17487, Germany
Tel +49 3834 420 4423
Fax +49 3834 420 4377
Email delceam@uni-greifswald.de; makulke@web.de

Purpose: Although strongly related, the pathophysiological effect of the N34S mutation in the serine protease inhibitor Kazal type 1 (SPINK1) in chronic pancreatitis is still unknown. In this study, we investigate the conformational space of the human cationic trypsin-serine protease inhibitor complex.

Methods: Simulations with molecular dynamics, replica exchange, and transition pathway methods are used.

Results: Two main binding states of the inhibitor to the complex were found, which explicitly relate the influence of the mutation site to conformational changes in the active site of trypsin.

Conclusion: Based on our result, a hypothesis is formulated that explains the development of chronic pancreatitis through accelerated digestion of the mutant by trypsin.

Keywords: trypsin, molecular dynamics simulations, replica exchange, transition path sampling, umbrella sampling

Introduction

Pancreatitis is an inflammatory disorder of the pancreas. While acute and chronic pancreatitis were previously viewed as separated diseases, today they are regarded as a continuum, with nearly 30% of the patients exhibiting overlapping phenotypes that manifest as recurrent pancreatitis.[1] Chronic pancreatitis (CP) develops from recurrent episodes of acute pancreatitis that lead to fibrosis, exocrine pancreatic insufficiency and diabetes. While initial triggers for pancreatitis are diverse, almost all of them result in premature activation of trypsin and subsequently other proteases in pancreatic acinar cells. Therefore, pancreatitis is considered to be an autodigestive disorder caused by trypsin auto-activation, which is also supported by numerous mutations in the *PRSS1* gene coding for human cationic trypsin (TRY1).[2,3] The serine protease inhibitor Kazal-type 1 (SPINK1, also known as PSTI or TATI) represents the first line of defense against the trypsin auto-activation cascade by potently binding and inhibiting active trypsin. The inhibitor has a size of 6.2 kDa and is co-located with trypsinogen and other zymogens in storage organelles called zymogen granules of pancreatic acinar cells. The c.101A>G point mutation is the most common variant of the *SPINK1* gene, which results in a p.

2111

N34S amino acid substitution and represents one of the most clinically relevant risk factors for chronic pancreatitis with almost 10% of the patients carrying the mutation, compared to 1% of the healthy population.[4–6] Although the p.N34S mutation was identified over two decades ago and was subjected to many additional studies, the mechanism of action of this variant remains enigmatic to this day.[4,5,7,8] Hypotheses regarding the disease relevant function of this variant ranged from reduced inhibitory activity, misfolding, reduced expression or secretion to reduced proteolytic stability, structural alterations and secondary trigger mechanisms. While most of the aforementioned hypotheses have now been refuted, very little is known about the SPINK1 structure.[9–16]

Patient studies revealed that the p.N34S mutation alone is not sufficient to develop chronic pancreatitis, but must be combined with additional risk factors.[17–19] It can therefore be assumed that the p.N34S mutation does not lead to a complete destruction of functional SPINK1, but rather reduces either the kinetic properties or the number of correctly folded SPINK1 molecules. In an earlier study, we showed differing secondary structures of SPINK1 wild type (WT) and the N34S mutant by circular dichroism (CD) spectroscopy as well as different secondary structures at lower pH values.[20] However, these minor changes could not be linked to different inhibitory activities of the two variants. The location of differing regions within the proteins could also not be precisely determined, which makes it difficult to assess the relevance of the relatively small changes in the secondary structure and their possible influence on the tertiary structure of the inhibitor. It also remains unclear how

the structures of SPINK1 and its N34S mutant behave in complex with its natural target trypsin (TRY1). In a recent study by Sun et al, the SPINK1-TRY1 complex structure was examined by docking and molecular dynamics simulations and no significant difference in secondary structure content was reported, but the loop region around the mutation site was more flexible compared to the wild type.[21]

In addition, the complex association and dissociation rates of mutant and WT were examined and found to be the same within the margin of error.[11,20] They argued that although these constants are experimentally indistinguishable, their theoretical results support a slight deviation of these rates, which leads to a lower inhibition potential of the p.N34S-mutated SPINK1.[21]

In the present study, we used molecular dynamic simulations (MDS) to elucidate structural and kinetic features of the SPINK1-TRY1 complex and differences due to the p.N34S amino acid substitution in SPINK1. In particular, the secondary structure content in the complex and solution structure of SPINK1 for WT and the p.N34S mutant is examined. For the first time, the inactivation mechanism of SPINK1 through hydrolysis of K41-I42 is investigated and possible differences between WT and mutant are clarified (Figure 1). Based on these results, a hypothesis is formulated that explains a reduction in active p.N34S-mutated SPINK1 compared to WT.

Materials and Methods
Model for the SPINK1-TRY1 Complex
To obtain the SPINK1-TRY1 complex, the crystal structure of human cationic trypsin (PDB code: 1TRN)[22] was
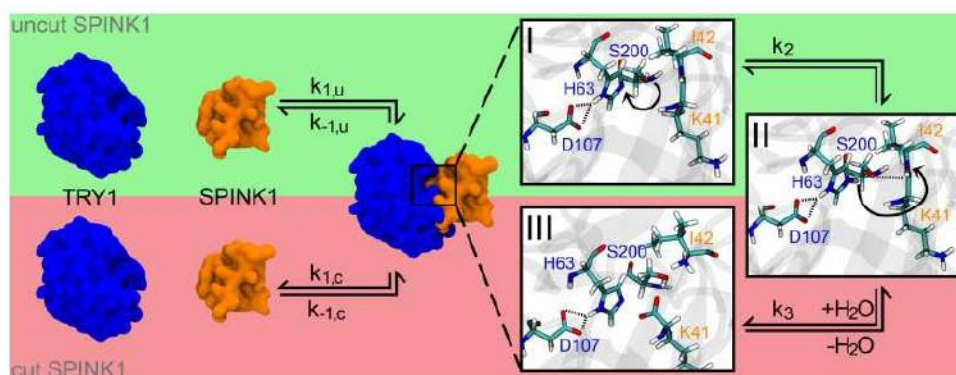

Figure 1 Complex formation for SPINK1 (orange) to TRY1 (blue) including the deactivation pathway of SPINK1 by cutting the K41-I42 peptide bond. The rate constants for the complexation/decomplexation of cut and uncut SPINK1 are indicated with $k_{1,u}/k_{-1,u}$ and $k_{1,c}/k_{-1,c}$, respectively. The zooms on the right partially show the cutting mechanism of SPINK1, which is catalysed by the catalytic triad S200-H63-D107 of TRY1. Starting from state I, the hydroxy hydrogen atom is translated to the far ring nitrogen position of H63 to state II (rate constant k_2). Afterwards, the deprotonated hydroxy oxygen atom of S200 coordinates to the carboxamide carbon atom of K41, while the carboxamide nitrogen atom (I42) attacks the previously translated hydrogen of H63 and the peptide bond is cleaved (rate constant k_3). The system is finally recovered by the addition of water to state III (not shown here).

superposed onto the structure of bovine chymotrypsinogen A in complex with SPINK1 (PDB code: 1CGI).[23] The mutation sites K41Y, I42E and D44R in SPINK1, which were introduced to stabilize the complex structure during crystallization, were mutated back to the WT amino acids. In addition, Y154 was phosphorylated by TRY1 according to 1TRN[22] and a structural calcium ion was added to the coordination site E75-N77-V80-E85. The resulting structure was solvated in explicit water, briefly minimized and equilibrated. The protonation states of the amino acids were then calculated using the H++ server assuming a pH of 7.[24] In addition, a physiological sodium chloride concentration of 0.15 M was added while the total charge in the simulation cell was neutralized. Finally, after adjusting the temperature and pressure in an NVT followed by an NPT ensemble, the complex was equilibrated for 21 ns in an NPT ensemble.

Molecular Dynamics Simulations

The SPINK1-TRY1 system was theoretically investigated with molecular dynamic simulations using the AMBER99SB-ildn force field,[25] which was specially developed for the representation of proteins. The equation of motion was integrated with a Verlet integrator every 2 fs using the GROMACS-5.1 software package.[26] The non-standard amino acid phosphotyrosine was represented with the FFPTM force field,[27] while water is defined by the TIP3P model.[28] The system is simulated with periodic boundary conditions in all spatial directions and the center of mass movement is removed every 200 fs. The temperature is set to 310 K for protein and non-protein separately by a modified V-rescale thermostat[29] with a coupling constant of 100 fs in order to prevent overcooling of the protein. The pressure is controlled by a Berendsen barostat[30] to 1 bar with a coupling constant of 1 ps and isotropically pushing against the system by assuming a compressibility of $4.5e^{-5}$ bar^{-1}. During the simulation, all bonds are constrained to their optimal length using the LINCS algorithm.[31] Non-bonded interactions are considered with 12-6-Lennard-Jones and Coulomb functions up to a distance of 1 nm and the fast smooth Particle-Mesh-Ewald method with a 1 nm cutoff in direct space and 0.12 nm grid spacing in reciprocal space.

Starting from the equilibrated TRY1-SPINK1 system, three independent simulations were carried out for different simulation setups: 1) TRY1 in complex with WT SPINK1, 2) TRY1 in complex with p.N34S mutated SPINK1, 3) WT SPINK1 in solution, 4) p.N34S mutant

SPINK1 in solution, 5) TRY1 in complex with WT SPINK1, but the S200 of TRY1 was deprotonated and the hydrogen was moved to the nitrogen of H63 to simulate an activated catalytic triad (Figure 1, state II) and 6) as 5) but with p.N34S mutated SPINK1. All systems were again equilibrated before collecting data for at least 50 ns. The partial charges for deprotonated serine were determined with the standard Amber protocol using Gaussian 2003.[32]

All simulated protein trajectories are visualized, atom-to-atom distances are analyzed and porcupine diagrams are carried out with VMD 1.9.2.[33] The reaction coordinate was determined by aligning the protein backbone of 14 frames along the trajectory of the TRY1-N34S SPINK1 simulation, which included the flip of N37. After performing a coordinate principal component analysis of these frames using only the backbone of amino acids K102, K178, W216 of TRY1 and N34S, E35, N37, R65 of SPINK1, the first principal component is the reaction coordinate. The secondary structure content was estimated using the CPPTraj program.[34]

Replica Exchange Simulations

Replica exchange simulations based on the TIGER2hs[35] method were carried out for the equilibrated simulation setups 5) and 6). The extended sampling algorithm is implemented in the NAMD 2.13[36] engine and CUDA acceleration. GROMACS topologies and coordinate files were transformed into the Amber format using ParmED,[37] while additionally applying the hydrogen mass partitioning scheme (HMR)[38] to the solute. The time step was increased to 4 fs, whereby the RATTLE[39] algorithm fixed the lengths of all bonds to hydrogen atoms. Short-range interactions based on Lennard-Jones and Coulomb potentials used a 0.9 nm limit and 0.1 nm switching function. Long-distance electrostatics were described by particle mesh Ewald (PME) with a lattice spacing of 0.1 nm. The temperature was adjusted by a Langevin thermostat[40] with a damping coefficient of 1 ps^{-1}, while a Langevin piston barostat[41] applied pressure control to 1 bar at 100 ps intervals with a 200 ps decay time constant. A temperature range of 300–370 K was spanned over 32 replicas in a pseudo-NVT ensemble.[42] A TIGER2hs sampling cycle consisted of 16 ps rapid heating and sampling followed by 4 ps rapid quenching and cooling. Afterwards, the hybrid solvent energy was calculated. The amount of water molecules in the first two water layers was determined by analyzing the radial distribution function along the surface of the protein

after a short MD simulation at 310 K. 1080 explicit water molecules were considered for the hybrid solvent exchange phase.

Implicit solvent energies are evaluated by the Gb_{obcII}[43] model using OpenMM 7.4[44] and periodic boundary conditions. Non-bonded cutoffs were set to half the cell size. Exchange attempts were carried out between the current baseline and another randomly selected replica after the cooling phase according to the TIGER2 scheme.[45] In each cycle, 32 exchange attempts were performed, while immediately swapping the temperatures of replicas after a successful attempt. The conformations from each successful exchange to the baseline were written to the trajectory. Before the next simulation cycle, all non-baseline replicas were assigned new temperatures based on their respective potential energy order.

Transition Path Sampling

Starting from the equilibrated system setups of 5) and 6), umbrella sampling simulations[46] along the reaction coordinate were performed using OpenMM 7.4.[44] Thereby, the reaction coordinate was separated in evenly distributed values ranging from 1 to −1 in 0.05 steps with positive and negative values relating to state A/B and C/D, respectively. To obtain structures on the respective reaction coordinate bins, referred to as umbrellas, a harmonic force was applied in the reaction coordinate space with a force constant of 400 kJ mol^{-1} nm^{-1}, while constraining the structure in place through position restraints on C_α atoms in β-sheets and the geometric center defined by the amino acids K102, K178, W216 of TRY1 and N34S, E35, N37, R65 of SPINK1. Each umbrella was subjected to a 600 ps equilibration in a NPT ensemble with the temperature and pressure set to 310 K by a Langevin thermostat and 1 bar by a Monte Carlo barostat, respectively. Afterwards, each umbrella was sampled for 1 ns, while reducing the force constant for the reaction coordinate restrain in two separate runs to 5 and 10 kJ mol^{-1} nm^{-1}. The result was analyzed with the weighted histogram analysis method (WHAM).[47]

Results

Conformations in the Proximity of the Mutation Site

The secondary structure content of SPINK1 in complex with TRY1 and free in solution was analyzed. In contrast to the circular dichroism data presented in Buchholz et al,[20] the secondary structure content of SPINK1 does not differ significantly between the N34S mutant and WT (Table 1). In the

Table 1 Secondary Structure Content Averaged for All SPINK1 Amino Acids for the TRY1-SPINK1 Complex and Unbound SPINK1 (System Setups 1)–4))

	TRY1-SPINK1		SPINK1 in Solution	
	Wild Type	**N34S**	**Wild Type**	**N34S**
Helix	19%	19%	19%	20%
Sheet	18%	19%	18%	17%
Turn/bend	24%	23%	23%	23%
Coil	39%	39%	40%	40%

complexed state, there are also no differences in the secondary structure content, but we found that the loop region spanned by amino acids 34–38 of SPINK1 has distinct conformations due to the steric hindrance caused by the binding. From these different states, a reaction coordinate is defined that covers the structures A–D (Figure 2A). Typically, the major binding states are A and B, which can also be found in the crystal structure.[22] However, in 1 of 12 simulations of the complex, a change to the states C and D was observed. The structure was stable and did not fold back to A/B (Figure 2C, black N34S curve). As this observation is either a rare event or an artifact of the simulation, replica exchange molecular dynamics simulations (REMD) were performed to better understand the state distribution of the TRY1-SPINK1 complex. Several states were found in C and D in the resulting structural ensemble, supporting the argument that this observation is rare (Figure 2D). Unfortunately, the REMD simulation did not fully converge and the computational resources required to perform a converged simulation were beyond reasonable considerations. Therefore, these data can only be discussed qualitatively and the correct state populations are not accessible via REMD.

Conformational Changes of the Complex Based on the Reaction Coordinate State

To identify the conformational changes associated with the change in the reaction coordinate (Figure 2A), the last 50,000 conformations from REMD simulations of the TRY1-SPINK1 complex were aligned with the backbone of TRY1 and averaged. Simulations were carried out for systems of the N34S mutant and the wild type in state A/B and C/D, respectively. The visualization of the porcupine plots shows that in the transition from state A/B to state C/D, the overall structure of SPINK1 in the binding pocket for both SPINK1 forms rotates counterclockwise with respect to the SPINK1-TRY1 viewing axis (Figure 3A). This leads to interesting aspects of the relevant atom-atom
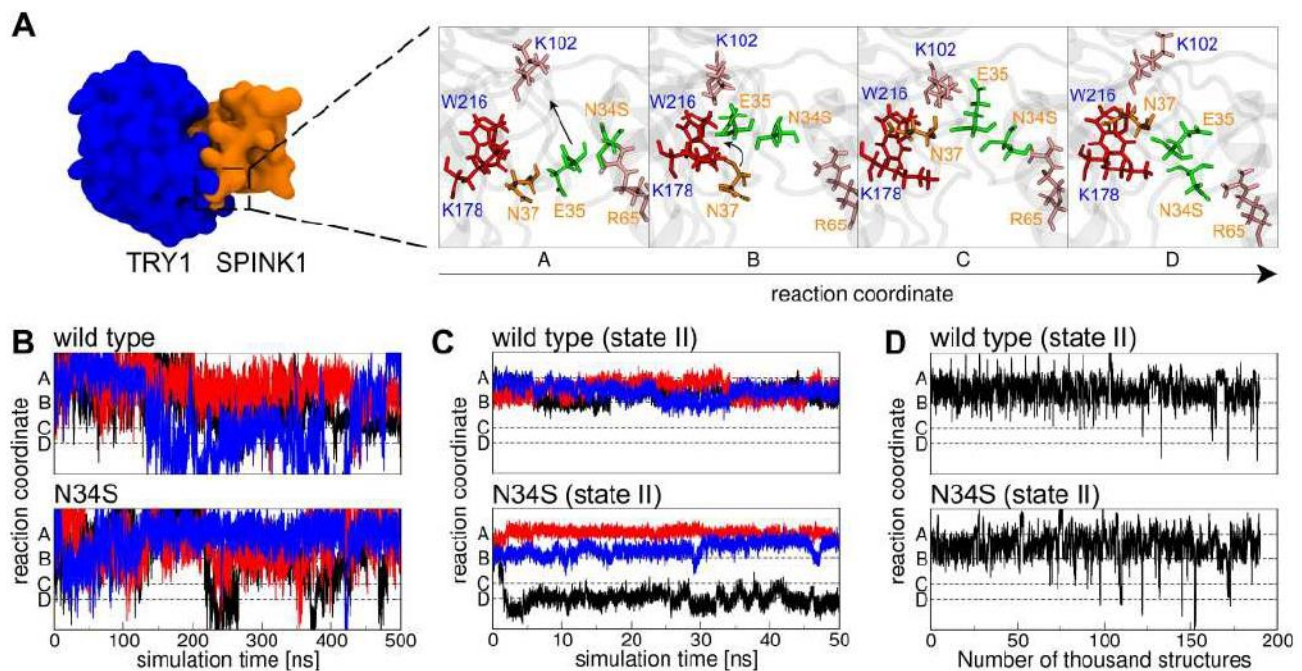
Figure 2 (A) The reaction coordinate as observed in one TRY1-N34S SPINK1 molecular dynamics simulation (cf. (**C**) black curve). Blue and orange labels indicate amino acids from TRY1 and SPINK1, respectively. The amino acids in stick representation are colored by their supposed functions, red binding pocket, orange key, pink steric barrier and green amino acids that have to pass the barrier. (**B**) Conformational state in relation to the reaction coordinate for SPINK1 in solution for WT and N34S mutant. The colors black, blue and red relate to three independent simulations. (**C**) Same as (**B**), but for SPINK1 in complex with TRY1 in state II (cf. Figure 1). (**D**) Conformational state in relation to the reaction coordinate of structures for the TRY1-SPINK1 complex in state II (cf. Figure 1) observed during a replica exchange simulation.

distances for the degradation. The distance between the hydroxy hydrogen atom of S200 and the far ring nitrogen of H63 for state I and II in A/B and C/D is always about 0.2–0.25 nm, which indicates that the reaction from state I to state II is fast and reversible. In state II, however, the distances between the hydrogen of the distant ring of H63 and the carboxamide nitrogen atom of I42 and between the hydroxy oxygen atom of S200 and the carbon atom of the peptide bond of K41 show great differences. They are both significantly reduced by going from state II A/B to C/D for WT and mutant (Figure 3B). In addition, these distances are generally smaller in the WT form.

Discussion

The Reaction Coordinate Includes a Major Energy Barrier

State A is stabilized by the polar interaction between N34S and R65 and is characterized by N37 and E35 being in a downward position with respect to amino acids W216 and K178 of TRY1 (Figure 2A). In state B, the N34S-R65 interaction breaks, causing E35 to move upwards. This allows N37 to reach a binding pocket that is spanned by K178 and W216 in state C. This conformation is stabilized by an interaction formed between E35 and K102. Finally, in state D, E35 moves

back down to lock N37 in the binding pocket. In our opinion, the highest energy barrier for the transition between A/B and C/D is given by the rigidity of the system, and amino acids K102 and K65 act as gatekeepers preventing E35 from moving straight up. However, E35 has to move away for N37 to reach the upper binding pocket as the bulky amino acids K178 and W216 prevent the movement through them.

Sun et al also examined the loop region around the N34S mutation site in molecular dynamic simulations involving the chymotrypsinogen A-SPINK1 complex.[21] They observed greater flexibility in the WT compared to the N34S mutant of this loop and argued that the mutation rigidifies the loop. We believe that their data is explained by the rare event described by the reaction coordinate. Although they performed more extensive molecular dynamics simulations on the complex with a total simulation time of 5 μs, the event may only be observed a few times, resulting in non-converged ensembles for their WT and N34S mutant simulations. If the event occurs more frequently in the WT, the flexibility of the loop is higher because the transitions along the reaction coordinate result in a significant change in the loop dihedral angles (Figure 2A). This suggests that the loop in the mutant is more rigid.
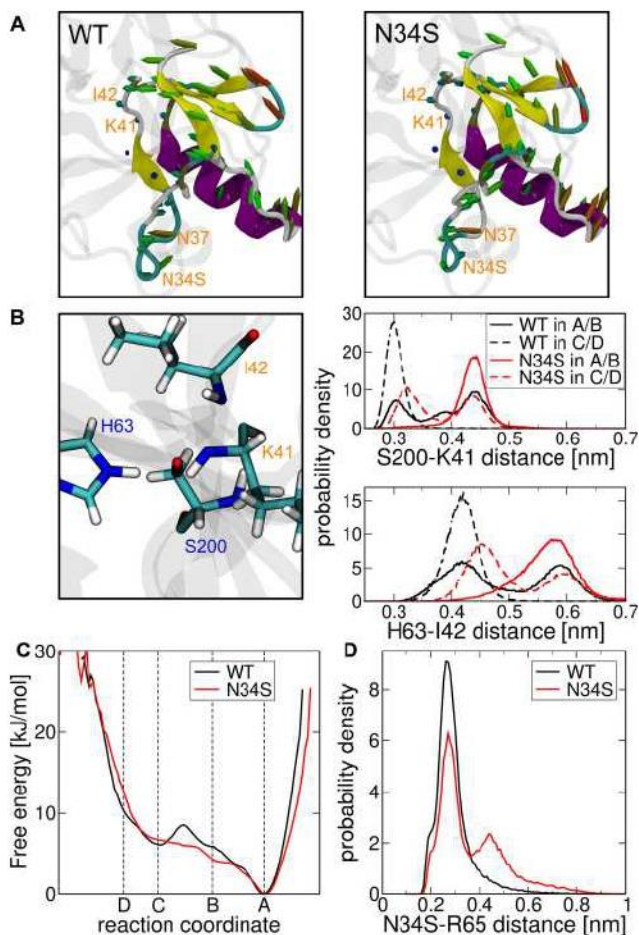
Figure 3 (**A**) Porcupine plots that show the structural difference when state II A/B is transitioned to state II C/D for the mutant right and WT left (cf. Figures 1 and 2A). The colors indicate the secondary structure elements for helix purple, sheet yellow, turns cyan and coil white. The marked arrows indicate the transposition of the C_α atoms with the color gradient blue-green-red for small-medium-large deviations, respectively. (**B**) Distance distributions in the active center between the hydroxy oxygen atom of S200 and the carboxamide carbon atom of K41 (top), and between the far ring hydrogen of H63 and carboxamide nitrogen atom of I42 (bottom). The distributions are measured from respective REMD simulations in state II A/B and state II C/D of the mutant (N34S) and wild type (WT). (**C**) Free energy in kJ/mol along the reaction coordinate determined by transition path sampling of the wild-type solid and mutant dashed. (**D**) Histogram of the minimal distance between the amino acids N34S and R65 for wild-type black and mutant red.

Hypothesis for Differences in the Inhibition Behavior of the N34S Mutant

In formulating a hypothesis as to why the N34S mutant is a risk factor for chronic pancreatitis, the following assumptions are made. First, it is assumed that the binding of mutant and WT SPINK1 to TRY1 occurs mainly in state A or B (Figure 2A), as found in the crystal structure of bovine chymotrypsinogen A in complex with SPINK1 (PDB code: 1CGI). In this context, it should be noted that in solution the mutant loop region of SPINK1 frequently crosses all states A–D (Figure 2B). This does not mean that the binding

occurs in a random state as there is most likely a preference, but our data does not support any statements about the exact probabilities. Secondly, phosphorylated TYR-154 as found in the crystal structure (PDB code: 1TRN)[22] is used in the model instead of sulfated[48] and it is assumed that both post-translational modifications exhibit the same effects in regard to the analyses performed in this study. While the effect of the sulfation is not clear in human cationic trypsin and phosphorylated and sulfated TYR are chemically similar, sulfation instead of phosphorylation can lead to slight differences in the binding conformations.[49] The third assumption is that the difference between mutant and WT in terms of developing chronic pancreatitis is mainly kinetic in nature. This is supported by the fact that no conclusive conformational differences were found in the present or previous studies. The mutation also does not result in a misfolding or a complete absence of SPINK1, as this mutation alone is not enough to develop chronic pancreatitis and the mutation is also present in the healthy population. Contrary, there is experimental evidence that the N34S mutant may be expressed significantly less than the wild type, which would also explain the development of chronic pancreatitis.[50]

In previous studies, kinetic inhibition and surface plasmon resonance spectroscopy experiments revealed that the association and dissociation constants of the SPINK1-TRY1 complex were not significantly different between WT and N34S.[11,20] A recent study to the digestion of SPINK1 mutant and wild type by mesotrypsin showed also no difference.[51] To the best of our knowledge, there is no current study that has investigated the kinetics of peptide hydrolysis for this system. Hence, the rate constants k_2, k_3, $k_{1,c}$ and $k_{-1,c}$ could show differences between mutant and WT (Figure 1). A second consideration is that the rate constant in question should be the slowest of these constants to have the necessary effect. Since digestion appears to be the slowest process and the transition from state I to state II is fast and reversible, we focus on the rate constant k_3. At this point, it can be assumed that digested SPINK1 has a significantly lower inhibitory potential than active SPINK1, since the inhibition strength of SPINK1 in enzyme tests diminishes measurably over time.[52]

Figure 3B shows that the average distances that are important for the digestion of SPINK1 in the active center are significantly smaller in state II C/D compared to state II-A/B. This means that SPINK1 is digested faster in state

II C/D and that the transition from state A/B to C/D is a necessary activation step for digestion. While the populations of state A–D are similar for WT and mutant, the energy barrier between the transition from A/B to C/D is significantly higher for the WT (Figure 3C). We believe that this is related to the N34S-R65 interaction in SPINK1 (Figure 2A). Since the side chain of serine at the mutation site is shorter than that of asparagine, the average minimum distance increases and thus, the strength of the interaction is weakened (Figure 3D). This makes it easier for amino acid E35 to move upwards (Figure 2A–C) and lowers the barriers from A to B and especially from B to C. We propose that this step is the rate-limiting step in the digestive process and therefore the mutant is digested faster, because of the reduced energy barrier between states A/B and C/D.

This statement is of course debatable, as our data does not support conclusive explanations for the rate constant differences between the transition from A/B to C/D and the digestion in state II C/D. In addition, it is unclear whether the digestion rate for the WT in state II A/B is slower than for the mutant in state II C/D (Figure 3B). In this regard, we assume that the energy barriers in Figure 3C are significantly higher for both WT and mutant, supporting the argument that this step determines the rate. This is due to the transition path sampling method used for this data. It depends on the force constant applied and the step size, as well as the atomic position constraint. Based on the restraint methods used, the atomic positions could still perform minor movements, which can lead to smaller measured forces, especially in high-energy states. This leads to a general reduction of the measured energy barriers between states. However, we expect this error to affect WT and mutant in the same way.

Conclusion

In the present study, the conformations of the TRY1-SPINK1 complex were examined in detail using molecular dynamic simulations. Two binding populations A/B and C/D were identified, separated by a major energy barrier, which is significantly lower in the case of the N34S SPINK1 mutant. It has been argued that this leads to a weaker TRY1 inhibition strength of the mutant, as it enables a faster digestion of SPINK1 by TRY1 compared to the WT. The authors want to clarify that the proposed mechanism is hypothetical and needs to be further validated in future experimental and theoretical studies. Nevertheless, it might be an explanation for incomplete penetrance of the disease phenotype in patients with

chronic pancreatitis and a N34S mutation. Since the hypothesis predicts that the mutant will be digested more quickly, a similar enzyme assay as in[51] for N34S and WT SPINK1 with TRY1 can refute or support the proposed mechanism. The second prediction is that amino acids K102 of TRY1 and E35, N37 and R65 of SPINK1 play an important role in the digestion of SPINK1. The N37S and R65Q mutants showed normal binding behavior to TRY1,[51] but it will be interesting to investigate these mutants in the context of digestion. Lastly, in this study, it is assumed that the state of binding of N34S mutant SPINK1 to TRY1 is similar to the state of binding of WT SPINK1 to bovine chymotrypsinogen A. A crystal structure of the N34S SPINK1-TRY1 complex will clarify this assumption.

Finally, a better understanding of the digestive mechanism of SPINK1 caused by TRY1 can help to develop a more robust SPINK1 mutant that is hardly digested and can act as a drug for chronic pancreatitis caused by SPINK1.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Weiss FU, Skube ME, Lerch MM. Chronic pancreatitis. *Curr Opin Gastroenterol.* 2018;34(5):322–329. doi:10.1097/MOG.0000000000 000461
2. Whitcomb DC, Gorry MC, Preston RA, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat Genet.* 1996;14(2):141–145. doi:10.1038/ng1096-141
3. Teich N, Rosendahl J, Tóth M, Mössner J, Sahin-Tóth M. Mutations of human cationic trypsinogen (PRSS1) and chronic pancreatitis. *Hum Mutat.* 2006;27(8):721–730. doi:10.1002/humu.20343
4. Witt H, Luck W, Hennies HC, et al. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. *Nat Genet.* 2000;25(2):213–216. doi:10.1038/76088
5. Pfützer RH, Barmada MM, Brunskill APJ, et al. SPINK1/PSTI polymorphisms act as disease modifiers in familial and idiopathic chronic pancreatitis. *Gastroenterology.* 2000;119(3):615–623. doi:10.1053/gast.2000.18017
6. Aoun E, Chang -C-CH, Greer JB, Papachristou GI, Barmada MM, Whitcomb DC. Pathways to injury in chronic pancreatitis: decoding the role of the high-risk SPINK1 N34S haplotype using meta-analysis. *PLoS One.* 2008;3(4):e2003. doi:10.1371/journal.pone.0002003
7. Boulling A, Chen JM, Callebaut I, Férec C. Is the SPINK1 p. Asn34Ser missense mutation per se the true culprit within its associated haplotype? *Webmed Central Genet.* 2012;3(2): WMC003084. doi:10.9754/journal.wmc.2012.003084

8. Chen JM, Férec C. Chronic pancreatitis: genetics and pathogenesis. *Annu Rev Genomics Hum Genet.* 2009;10(1):63–87. doi:10.1146/annurev-genom-082908-150009

9. Valmu L, Paju A, Lempinen M, Kemppainen E, Stenman U-H. Application of proteomic technology in identifying pancreatic secretory trypsin inhibitor variants in urine of patients with pancreatitis. *Clin Chem.* 2006;52(1):73–81. doi:10.1373/clinchem.2005.056861

10. Kuwata K, Hirota M, Ogawa M. Functional analysis of pancreatic secretory trypsin inhibitor protein with amino acid substitution. *Int Congr Ser.* 2003;1255(C):193–196. doi:10.1016/S0531-5131(03)00201-2

11. Kiraly O, Wartmann T, Sahin-Toth M. Missense mutations in pancreatic secretory trypsin inhibitor (SPINK1) cause intracellular retention and degradation. *Gut.* 2007;56(10):1433–1438. doi:10.1136/gut.2006.115725

12. Boulling A, Le Maréchal C, Trouvé P, Raguénès O, Chen J-M J-M, Férec C. Functional analysis of pancreatitis-associated missense mutations in the pancreatic secretory trypsin inhibitor (SPINK1) gene. *Eur J Hum Genet.* 2007;15(9):936–942. doi:10.1038/sj.ejhg.5201873

13. Kereszturi E, Kiraly O, Sahin-Toth M. Minigene analysis of intronic variants in common SPINK1 haplotypes associated with chronic pancreatitis. *Gut.* 2009;58(4):545–549. doi:10.1136/gut.2008.164947

14. Shimosegawa T, Kume K, Masamune A. SPINK1, ADH2, and ALDH2 gene variants and alcoholic chronic pancreatitis in Japan. *J Gastroenterol Hepatol.* 2008;23(s1):S82–S86. doi:10.1111/j.1440-1746.2007.05291.x

15. Sandhu B, Vitazka P, Ferreira-Gonzalez A, et al. Presence of SPINK-1 variant alters the course of chronic pancreatitis. *J Gastroenterol Hepatol.* 2011;26(6):965–969. doi:10.1111/j.1440-1746.2011.06713.x

16. Boulling A, Masson E, Zou W, et al. Identification of a functional enhancer variant within the chronic pancreatitis-associated SPINK1 c.101A>G (p.Asn34Ser)-containing haplotype. *Hum Mutat.* 2017;38(8):1014–1024. doi:10.1002/humu.23269

17. Threadgold J. The N34S mutation of SPINK1 (PSTI) is associated with a familial pattern of idiopathic chronic pancreatitis but does not cause the disease. *Gut.* 2002;50(5):675–681. doi:10.1136/gut.50.5.675

18. Drenth JPH. Mutations in serine protease inhibitor Kazal type 1 are strongly associated with chronic pancreatitis. *Gut.* 2002;50(5):687–692. doi:10.1136/gut.50.5.687

19. Rai P, Sharma A, Gupta A, Aggarwal R. Frequency of SPINK1 N34S mutation in acute and recurrent acute pancreatitis. *J Hepatobiliary Pancreat Sci.* 2014;21(9):663–668. doi:10.1002/jhbp.111

20. Buchholz I, Nagel F, Klein A, et al. The impact of physiological stress conditions on protein structure and trypsin inhibition of serine protease inhibitor Kazal type 1 (SPINK1) and its N34S variant. *Biochim Biophys Acta.* 2020;1868(1):140281. doi:10.1016/j.bbapap.2019.140281

21. Sun Z, Kolssváry I, Kozakov D, Sahin-Tóth M, Vajda S. The N34S mutation of SPINK1 may impact the kinetics of trypsinogen activation to cause early trypsin release in the pancreas. *bioRxiv.* 2020;1–23. doi:10.1101/2020.08.21.262162

22. Gaboriaud C, Serre L, Guy-Crotte O, Forest E, Fontecilla-Camps J-C. Crystal structure of human trypsin 1: unexpected phosphorylation of Tyr151. *J Mol Biol.* 1996;259(5):995–1010. doi:10.1006/jmbi.1996.0376

23. Hecht HJ, Szardenings M, Collins J, Schomburg D. Three-dimensional structure of the complexes between bovine chymotrypsinogen A and two recombinant variants of human pancreatic secretory trypsin inhibitor (Kazal-type). *J Mol Biol.* 1991;220(3):711–722. doi:10.1016/0022-2836(91)90112-J

24. Anandakrishnan R, Aguilar B, Onufriev AV. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* 2012;40(W1):W537–W541. doi:10.1093/nar/gks375

25. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput.* 2015;11(8):3696–3713. doi:10.1021/acs.jctc.5b00255

26. Abraham MJ, Murtola T, Schulz R, et al. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1–2:19–25. doi:10.1016/j.softx.2015.06.001

27. Khoury GA, Thompson JP, Smadbeck J, Kieslich CA, Floudas CA. Forcefield_PTM: Ab Initio Charge and AMBER forcefield parameters for frequently occurring post-translational modifications. *J Chem Theory Comput.* 2013;9(12):5653–5674. doi:10.1021/ct400556v

28. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79(2):926–935. doi:10.1063/1.445869

29. Basconi JE, Shirts MR. Effects of temperature control algorithms on transport properties and kinetics in molecular dynamics simulations. *J Chem Theory Comput.* 2013;9(7):2887–2899. doi:10.1021/ct400109a

30. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys.* 1984;81(8):3684–3690. doi:10.1063/1.448118

31. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. *J Comput Chem.* 1997;18(12):1463–1472. doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

32. Frisch MJ, Trucks GW, Schlegel HB, et al. *Gaussian 03, Revision C. 02.* Wallingford CT: Gaussian Inc; 2004.

33. Humphrey W, Dalke A, Schulten K. VMD - visual molecular dynamics. *J Mol Graph.* 1996;14(1):33–38. doi:10.1016/0263-7855(96)00018-5

34. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput.* 2013;9(7):3084–3095. doi:10.1021/ct400341p

35. Geist N, Kulke M, Schulig L, Link A, Langel W. Replica-based protein structure sampling methods II: advanced hybrid solvent TIGER2hs. *J Phys Chem B.* 2019;123(28):5995–6006. doi:10.1021/acs.jpcb.9b03134

36. Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *J Comput Chem.* 2005;26(16):1781–1802. doi:10.1002/jcc.20289

37. Case DA, Betz R, Botello-Smith W, et al. *AMBER 2016.* San Francisco: University of California; 2016. doi:10.1021/ct200909j

38. Hopkins CW, Le Grand S, Walker RC, Roitberg AE. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J Chem Theory Comput.* 2015;11(4):1864–1874. doi:10.1021/ct5010406

39. Andersen HC. Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations. *J Comput Phys.* 1983;52(1):24–34. doi:10.1016/0021-9991(83)90014-1

40. Grest GS, Kremer K. Molecular dynamics simulation for polymers in the presence of a heat bath. *Phys Rev A.* 1986;33(5):3628–3631. doi:10.1103/PhysRevA.33.3628

41. Feller SE, Zhang Y, Pastor RW, Brooks BR. Constant pressure molecular dynamics simulation: the Langevin piston method. *J Chem Phys.* 1995;103(11):4613–4621. doi:10.1063/1.470648

42. Kulke M, Geist N, Möller D, Langel W. Replica-based protein structure sampling methods: compromising between explicit and implicit solvents. *J Phys Chem B.* 2018;122(29):7295–7307. doi:10.1021/acs.jpcb.8b05178

43. Nguyen H, Roe DR, Simmerling C. Improved generalized born solvent model parameters for protein simulations. *J Chem Theory Comput.* 2013;9(4):2020–2034. doi:10.1021/ct3010485

44. Eastman P, Swails J, Chodera JD, et al. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLOS Comput Biol.* 2017;13(7):e1005659. doi:10.1371/journal.pcbi.1005659

45. Li X, Latour RA, Stuart SJ. TIGER2: an improved algorithm for temperature intervals with global exchange of replicas. *J Chem Phys.* 2009;130(17):174106. doi:10.1063/1.3129342

46. Kästner J. Umbrella sampling. *Wiley Interdiscip Rev Comput Mol Sci.* 2011;1(6):932–942. doi:10.1002/wcms.66

47. Bauer D. WHM v1.0.0; 2020. Available from: https://github.com/danijoo/WHAM. Accessed April 23, 2021.

48. Sahin-Tóth M, Kukor Z, Nemoda Z. Human cationic trypsinogen is sulfated on Tyr154. *FEBS J.* 2006;273(22):5044–5050. doi:10.1111/j.1742-4658.2006.05501.x

49. Ju T, Niu W, Cerny R, Bollman J, Roy A, Guo J. Molecular recognition of sulfotyrosine and phosphotyrosine by the Src homology 2 domain. *Mol Biosyst.* 2013;9(7):1829. doi:10.1039/c3mb70061e

50. Kereszturi É, Sahin-Tóth M. Pancreatic cancer cell lines heterozygous for the SPINK1 p.N34S haplotype exhibit diminished expression of the variant allele. *Pancreas.* 2017;46(6):e54–e55. doi:10.1097/MPA.0000000000000817

51. Szabó A, Toldi V, Gazda LD, Demcsák A, Tőzsér J, Sahin-Tóth M. Defective binding of SPINK1 variants is an uncommon mechanism for impaired trypsin inhibition in chronic pancreatitis. *J Biol Chem.* 2021;296:100343. doi:10.1016/j.jbc.2021.100343

52. Kuwata K, Hirota M, Shimizu H, et al. Functional analysis of recombinant pancreatic secretory trypsin inhibitor protein with amino-acid substitution. *J Gastroenterol.* 2002;37(11):928–934. doi:10.1007/s005350200156

Article III

Article

Structural and Biophysical Insights into SPINK1 Bound to Human Cationic Trypsin

Felix Nagel [1], Gottfried J. Palm [2], Norman Geist [1], Thomas C. R. McDonnell [3], Anne Susemihl [1,4],
Britta Girbardt [2], Julia Mayerle [5], Markus M. Lerch [6], Michael Lammers [2] and Mihaela Delcea [1,*]

[1] Biophysical Chemistry, Institute of Biochemistry, University of Greifswald, 17489 Greifswald, Germany;
    felix.nagel@uni-greifswald.de (F.N.); norman.geist@uni-greifswald.de (N.G.);
    anne.susemihl@uni-greifswald.de (A.S.)
[2] Synthetic and Structural Biochemistry, Institute of Biochemistry, University of Greifswald,
    17489 Greifswald, Germany; palm@uni-greifswald.de (G.J.P.); bgirbard@uni-greifswald.de (B.G.);
    michael.lammers@uni-greifswald.de (M.L.)
[3] Biochemical Engineering Department, University College London, Bernard Katz, London WC1E 6BT, UK;
    thomas.mcdonnell.11@ucl.ac.uk
[4] Department of Hematology and Oncology, Internal Medicine C, University of Greifswald,
    17489 Greifswald, Germany
[5] Department of Medicine II, University Hospital Munich, Ludwig-Maximillian University Munich,
    81377 Munich, Germany; julia.mayerle@med.uni-muenchen.de
[6] Department of Medicine A, University Medicine Greifswald, 17489 Greifswald, Germany;
    markus.lerch@med.uni-muenchen.de
* Correspondence: delceam@uni-greifswald.de

Abstract: (1) The serine protease inhibitor Kazal type 1 (SPINK1) inhibits trypsin activity in zymogen granules of pancreatic acinar cells. Several mutations in the *SPINK1* gene are associated with acute recurrent pancreatitis (ARP) and chronic pancreatitis (CP). The most common variant is SPINK1 p.N34S. Although this mutation was identified two decades ago, the mechanism of action has remained elusive. (2) SPINK1 and human cationic trypsin (TRY1) were expressed in *E. coli*, and inhibitory activities were determined. Crystals of SPINK1–TRY1 complexes were grown by using the hanging-drop method, and phases were solved by molecular replacement. (3) Both SPINK1 variants show similar inhibitory behavior toward TRY1. The crystal structures are almost identical, with minor differences in the mutated loop. Both complexes show an unexpected rotamer conformation of the His63 residue in TRY1, which is a member of the catalytic triad. (4) The SPINK1 p.N34S mutation does not affect the inhibitory behavior or the overall structure of the protein. Therefore, the pathophysiological mechanism of action of the p.N34S variant cannot be explained mechanistically or structurally at the protein level. The observed histidine conformation is part of a mechanism for SPINK1 that can explain the exceptional proteolytic stability of this inhibitor.

1. Introduction

The Kazal-type serine protease inhibitor domain (protein family PF00050) presently consists of 9000 members and runs through all domains of life [1,2]. It is mostly indicative of protease inhibitors but has also been found in the extracellular regions of agrins, which are not known to inhibit proteases. The structure of a classical Kazal domain revolves around a central α-helix inserted between two β-strands and a third β-strand toward the C-terminus. In addition, three disulfide bridges are present in the majority of Kazal domains, stabilizing the overall structure and reactive loop [3]. The general inhibition mechanism of this protein

family is competitive and temporary and is accompanied by reversible cleavage of the P1-P1′ scissile bond (Schechter and Berger nomenclature, Figure 1A) [4,5]. A prerequisite for the standard mechanism is for the inhibitor to be proteolytically stable enough to exert its function before it is ultimately degraded by its target protease [4,6–8]. This high proteolytic stability is achieved by low dissociation rates and re-ligation of cleaved scissile bonds, which can be catalyzed due to the absence of water at the interface, although the overall mechanism remains poorly understood [9–13].
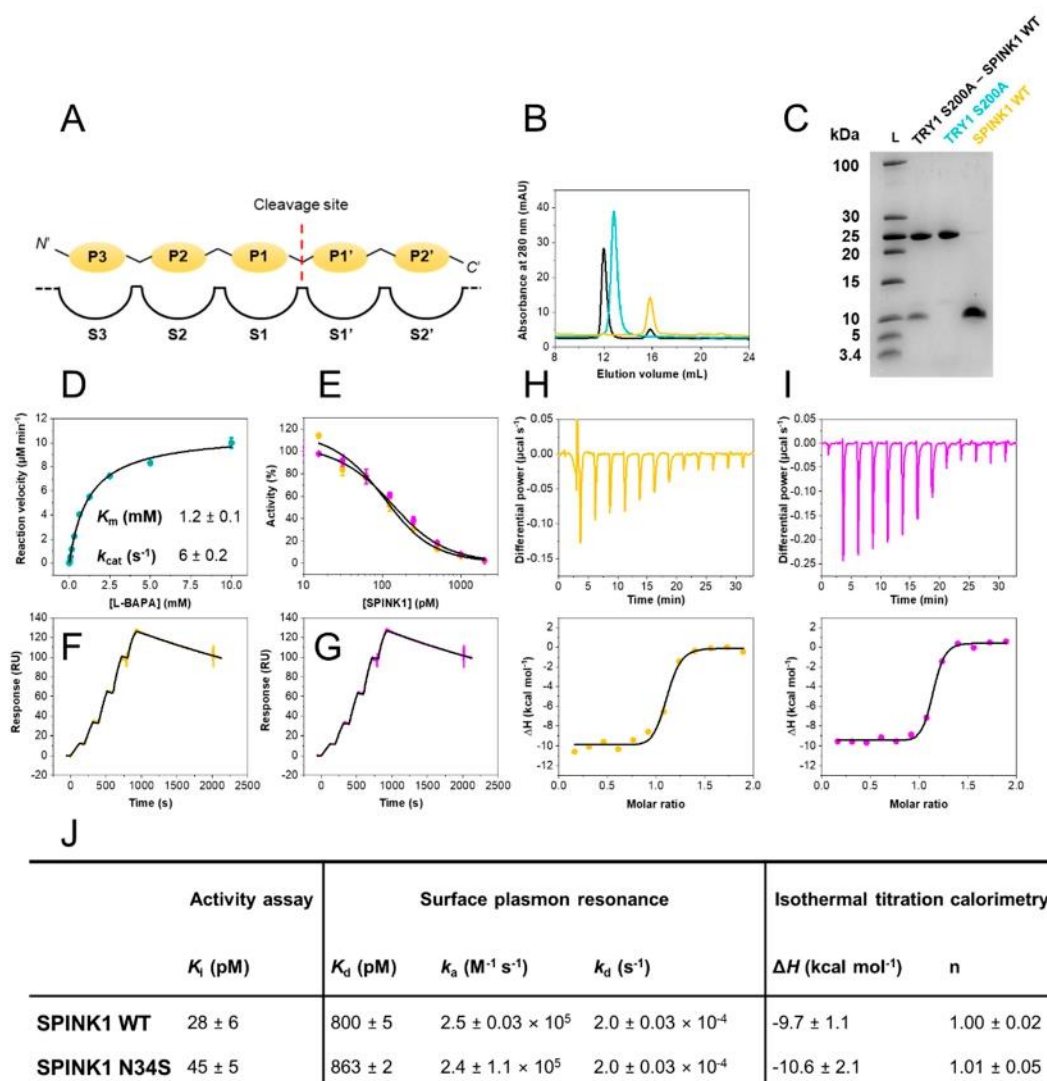

	Activity assay	Surface plasmon resonance			Isothermal titration calorimetry	
	K_i (pM)	K_d (pM)	k_a (M⁻¹ s⁻¹)	k_d (s⁻¹)	ΔH (kcal mol⁻¹)	n
SPINK1 WT	28 ± 6	800 ± 5	$2.5 \pm 0.03 \times 10^5$	$2.0 \pm 0.03 \times 10^{-4}$	-9.7 ± 1.1	1.00 ± 0.02
SPINK1 N34S	45 ± 5	863 ± 2	$2.4 \pm 1.1 \times 10^5$	$2.0 \pm 0.03 \times 10^{-4}$	-10.6 ± 2.1	1.01 ± 0.05

Figure 1. Biophysical characterization of the SPINK1–TRY1 interaction. (**A**) Schechter and Berger nomenclature. (**B**) Size-exclusion chromatograms of purified SPINK1 WT (yellow), TRY1 p.S200A (cyan) and complex (black). (**C**) 16% tricine SDS–PAGE of purified TRY1 p.S200A–SPINK1 WT complex, TRY1 p.S200A and SPINK1 WT. (**D**) Michaelis–Menten kinetics of L-BAPA substrate with TRY1. Error bars represent standard deviations of three independent experiments. K_m and k_{cat} values represent fitted values and their standard error. (**E**) Trypsin activity assay at varying SPINK1 concentrations fitted with Morrison's quadratic equation. Error bars represent standard deviations of three independent experiments. (**F**) Surface plasmon resonance single-cycle kinetic of the TRY1 p.S200A—SPINK1 WT or (**G**) p.N34S interaction fitted with a 1:1 Langmuir interaction model. (**H**) Isothermal titration calorimetry of the TRY1 p.S200A–SPINK1 WT or (**I**) p.N34S interaction fitted with a 1:1 binding site model. (**J**) Summary of equilibrium, kinetic and thermodynamic data. K_i values represent fitted values ± standard errors, while SPR and ITC data are reported as mean ± SD of at least three independent experiments.

In the present study, we investigated human serine protease inhibitor Kazal-type 1 (SPINK1) and its interaction with human cationic trypsin (TRY1). Two decades ago, the inhibitor first gained wider attention when a p.N34S amino acid substitution present in 2% of the population was found to be associated with acute recurrent pancreatitis (ARP) and chronic pancreatitis (CP) [14–17]. Prevalence for developing the disease is increased up to tenfold in carriers of the mutation, and the mechanism of action of this variant remains enigmatic to this day [18–21]. However, recent reviews of available genetic databases indicate a variant in linkage disequilibrium upstream of the *SPINK1* promoter to be causal of the disease [22]. Additionally, the *SPINK1* gene has been identified to be associated with many types of cancer [23–25], which is mostly attributed to potential epidermal growth factor receptor (EGFR) binding [26]. Furthermore, SPINK1 has been reported to inhibit granzyme A, preventing apoptosis in tumor cells [27].

For the past two decades, the structure of the SPINK1 complex with TRY1 has been under debate and was subjected to multiple molecular dynamics simulation studies [28,29], docking simulations [30], homology modeling [15,31] and spectroscopic studies, using circular dichroism [32,33]. Changes within the SPINK1 binding loop and peptide flipping due to the mutation were proposed. Unfortunately, these previous studies were inconclusive regarding differences between SPINK1 WT and the p.N34S mutant and/or did not provide the required resolution for a detailed molecular understanding of the interaction with TRY1.

Here we provide a structure–function analysis for the SPINK1–TRY1 interaction. Kinetic, thermodynamic and affinity binding data show a high affinity interaction with formation of very stable complexes for both SPINK1 WT and p.N34S mutant. We crystallized both SPINK1 variants in a complex with TRY1 and solved both structures at 2.9 and 2.1 Å resolution, respectively. Our structures differ from the aforementioned structural studies because the catalytic His63 adopts an outward facing conformation in both complexes. His63, thereby, faces away from Ser200 and Asp107 of the catalytic triad, rendering the protease inactive. The outward conformation may be further stabilized by a sulfate ion observed in close proximity to His63. We believe that the His63 conformation contributes significantly to the high proteolytic stability of the inhibitor. Hence, our results provide a structural platform that allows for the comparison of SPINK1 WT and the ARP and CP relevant p.N34S mutant bound to their natural target protease TRY1. Furthermore, they reveal a binding mechanism previously not known for this protein family and, thus, may have implications for other Kazal-inhibitors.

2. Results

2.1. Inhibitory Activities of SPINK1 WT and p.N34S Are Similar

Despite it being clearly established that the p.N34S mutation in SPINK1 is associated with a ten-fold increase in prevalence for developing ARP or CP, the mechanism of action of this mutation has remained elusive [18]. To approach this problem, we expressed recombinant SPINK1 variants and human cationic trypsin (TRY1). Protein purity, homogeneity and complex formation were validated by analytical size exclusion chromatography, as well as SDS–PAGE (Figure 1B,C).

Inhibitory constants (K_i) were determined in equilibrium at 37 °C, using Morrison's quadratic equation for tight binding inhibition utilizing previously determined K_m values ($K_m = 1.2$ mM; Figure 1D,E,J). Both SPINK1 variants exhibit K_i values in the sub-nanomolar range ($K_i < 50$ pM). Hence, the p.N34S amino acid substitution does not influence SPINK1 binding to TRY1 in equilibrium.

2.2. SPINK1 WT and p.N34S Show Similar Kinetics for Their Interaction with TRY1

Kinetic rate constants were determined by surface plasmon resonance (SPR). A CM5 sensor chip was prepared with covalently immobilized TRY1 p.S200A. The amino acid substitution of the catalytic serine residue was introduced to ensure surface stability and prevent autodigestion of the enzyme. The interaction with both SPINK1 variants was investi-

gated by single-cycle kinetic experiments at 37 °C (Figure 1F,G,J). Kinetic rate constants were similar with a very high complex stability and association rates k_a of $2.5 \times 10^5 \, M^{-1} \, s^{-1}$. Both SPINK1 variants in complex with TRY1 p.S200A displayed a half-life $t_{1/2} \sim 1$ h. While a single point mutation can often cause similar equilibrium affinities but very different binding kinetics, this is not the case for the SPINK1–TRY1 interaction and the p.N34S amino acid substitution.

2.3. SPINK1 WT and p.N34S Display Similar Thermodynamic Profiles for Their Interaction with TRY1

Isothermal titration calorimetry was carried out with TRY1 p.S200A placed in the sample cell and SPINK1 in the syringe. Dissociation constants (K_d) were not determined, due to the high affinity of the interaction, resulting in c-values > 10,000. Without sufficient curvature, K_d cannot be assessed accurately, but valuable insights into the thermodynamics of the interaction can still be obtained. K_d values from SPR experiments were used for the thermodynamic analysis. Injections were carried out with appropriate spacing to ensure full equilibration before the subsequent injection (Figure 1H,I,J). Determined molar enthalpies for both SPINK1 variants were similar and revealed favorable binding enthalpies at approximately $-10 \, kcal \, mol^{-1}$. The interaction is primarily enthalpically driven with little entropic contribution. Overall, both SPINK1 variants are functionally indistinguishable.

2.4. Structure of SPINK1 Variants Bound to TRY1

We solved the crystal structures of TRY1 p.S200A bound to SPINK1 WT and p.N34S at 2.9 and 2.1 Å resolution, respectively (Table 1 and Figure 2A). The final structure of TRY1 p.S200A contains amino acids 24–247 and amino acids 24–79 for the SPINK1 variants. TRY1 adopts the typical serine protease fold containing two β-barrels sandwiching the active site cleft of the protease. The S1 binding pocket determining the specificity of the protease contains the negatively charged Asp194, which, in turn, leads to the favorable binding of the positively charged Lys41 residue in SPINK1. The inhibitor assumes a typical Kazal domain structure with a reactive loop surrounding a central α-helix. The loop is stabilized by three disulfide bridges, and the domain contains a small portion of β-sheets. SPINK1 binds on top of the TRY1 p.S200A active site in a substrate-like manner, creating an interface of almost 1000 Å² (Figure 2B and Supplementary Figure S1). The structures of both SPINK1 variants are mostly similar, with small differences only present in the loop orientation of the p.N34S mutation site. However, due to high flexibility of this loop, side chain conformations are poorly defined, and changes are unlikely to translate into functional effects (Figure 2C).

2.5. Interactions of SPINK1 Variants with TRY1 p.S200A

Similar to other known Laskowski-inhibitors, both SPINK1 variants bind to TRY1 p.S200A in a substrate-like manner. The specificity determining P1 residue Lys41 is oriented in a way that interacts with Asp194 in TRY1. The oxygen of the carbonyl carbon is stabilized by the oxyanion hole formed by the backbone nitrogen atoms of Ala200 and Gly198 (Figure 3A and Supplementary Figure S2A). In TRY1 WT, Ala200 would be replaced by a serine residue oriented in a way that allows for the cleavage between the P1-P1′ residues in both SPINK1 variants. Apart from the specific Lys41–Asp194 interaction, most other interactions involve the SPINK1 backbone. The loop orientation is also stabilized by intramolecular interactions within the SPINK1 inhibitor. Most prominently, Asn56 interacts with Thr40 and Ile42, which are both direct neighbors to Lys41 (Figure 3B,E and Supplementary Figure S2C). Asn56 is very conserved among Kazal domains and probably exerts similar functions in related Kazal inhibitors. Locking the P1′ and P2 residues into place, Asn56 most likely stabilizes the loop conformation even after cleavage of the P1-P1′ scissile bond, aiding the re-synthesis of the peptide bond. Furthermore, loop flexibility is limited by disulfide bridges, as well as the neighboring β-sheet toward the C-terminal end of the loop. Asn34 or Ser34 does not directly interact with TRY1. However, the neighboring Tyr33 residue engages in a cation–pi

bond with Arg101 in TRY1. This interaction causes Tyr33 to be pulled away from the core interface, making room for a sulfate ion in the SPINK1 loop (Figure 3D and Supplementary Figure S2D). Coincidentally, a p.R101H mutation in rat anionic trypsin was described that creates a metal binding site, causing His63 to assume an energetically more favorable *trans* conformation [34].

Table 1. Data collection and refinement of SPINK1–TRY1 p.S200A complexes.

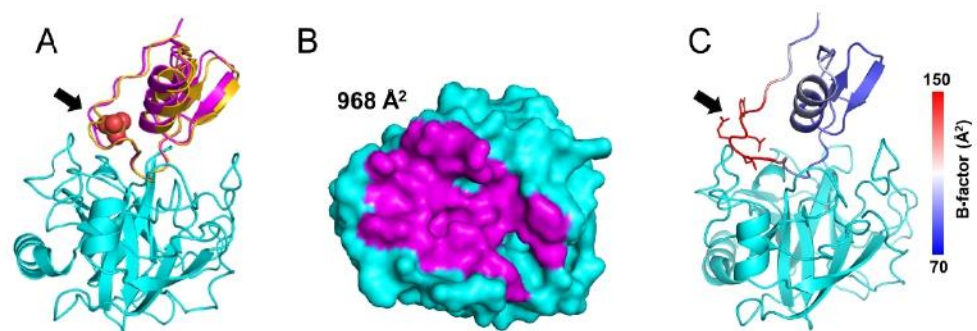|	SPINK1 WT–TRY1 p.S200A	SPINK1 p.N34S–TRY1 p.S200A
Data collection		
Beamline	14.1 at BESSY	14.2 at BESSY
Wavelength (Å)	0.9184	0.9184
Unit-cell parameters (Å) a, c in space group P3$_1$21	77.63, 187.35	76.55, 189.72
Resolution (Å)	50–2.90 (3.08–2.90)	50–2.10 (2.22–2.10)
No. of unique reflections (Friedel pairs merged)	15,112 (2371)	38,593 (6081)
Redundancy	19.1 (19.4)	19.7 (19.4)
Completeness (%)	99.8 (99.0)	99.8 (98.7)
R_{merge}	0.166 (2.651)	0.174 (3.172)
$cc_{1/2}$	0.999 (0.550)	0.999 (0.411)
$<I/\sigma(I)>$	16.3 (1.2)	14.8 (1.0)
Wilson B-factor (Å^2)	85.6	50.1
Refinement		
Resolution range (Å)	50–2.90 (2.99–2.90)	50–2.10 (2.15–2.10)
Completeness (%)	99.7 (98.0)	99.8 (97.1)
No. of reflections, working set	13589 (1191)	36688 (2600)
No. of reflections, test set	1509 (133)	1904 (134)
Final R_{work}	0.222 (0.365)	0.197 (0.334)
Final R_{free}	0.243 (0.374)	0.233 (0.421)
No. of non-H atoms		
Protein	4219	4269
Solvent	16	236
R.m.s. deviations		
Bond lengths (Å)	0.011	0.008
Angles (°)	1.619	1.505
Average B factors (Å^2)		
Protein	100.01	36.73
Solvent	138.68	51.09
Molprobity analysis		
Ramachandran most favored (%)	96.17	96.75
Ramachandran outliers (%)	0.0	0.0
Overall score	1.90	1.80
Clash score	13.48	3.80
PDB entry	7QE8	7QE9

Figure 2. Structure and superposition of TRY1 p.S200A–SPINK1 complexes. The black arrow indicates the position of the p.N34S mutation site. (**A**) Structures and superposition of TRY1 p.S200A (cyan) in complex with SPINK1 WT (yellow) or p.N34S (pink). Sulfate ions are shown as spheres and are colored according to their atom type. (**B**) Binding interface of SPINK1 p.N34S. Due to their similarity, the binding interface of SPINK1 WT was omitted but can be seen in Supplementary Figure S1. (**C**) SPINK1 WT in complex with TRY1 colored by B-factors.
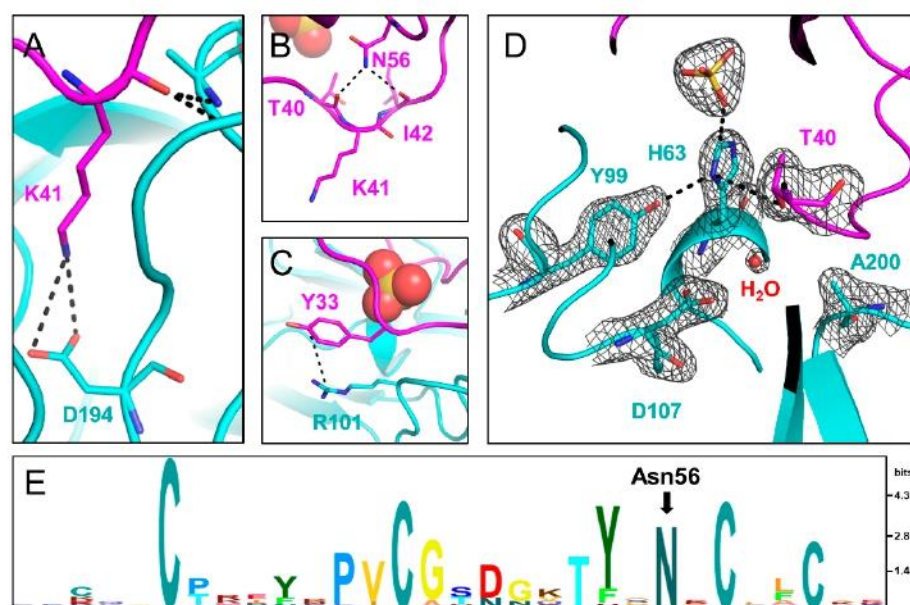
Figure 3. Interactions between TRY1 p.S200A (cyan) and SPINK1 p.N34S (pink). (**A**) Lys41 in SPINK1 interacts with the specificity determining Asp194 in TRY1. (**B**) Asn56 in SPINK1 stabilizes the binding loop by hydrogen bonding with Thr40 and Ile42. (**C**) Tyr33 in SPINK1 forms a cation—pi bond with Arg101 in TRY1 and is hence pulled outward. (**D**) Catalytic triad of the TRY1 p.S200A–SPINK1 complex (cyan). The 2Fo–Fc density map is shown at 1.6 Å around the residues of the catalytic triad and is contoured at 1.5 σ. In the complex structure His63 is rotated toward the sulfate ion and out of the productive catalytic triad arrangement. (**E**) Sequence conservation of the Kazal 1 family displayed by an HMM logo generated in Skylign [35]. Amino acid letter height is calculated based on the information content above background expressed in bits.

2.6. Conformational Change in the TRY1 Catalytic Triad

The structures of the active site regions of the enzyme–inhibitor complex resemble that of a typical Kazal inhibitor–serine protease complex. The scissile peptide group displays a planar geometry with the carbonyl-oxygen pointing toward the oxyanion hole. In TRY1 WT, the carbonyl carbon would be placed ideally for the generation of the acyl-intermediate with Ser200. In the present structure, Ser200 is replaced by an alanine to avoid proteolytic cleavage of the inhibitor or the protease due to residual protease activity observed even in the inhibitor bound complex. Asp107 is positioned in a way that would stabilize the His63 residue in the wild-type protease. Unexpectedly, His63 exhibits an uncommon rotamer conformation, pointing outward and away from the catalytic triad (Figure 3D and Supplementary Figure S2D). Instead, it points toward the sulfate ion bound within the SPINK1 loop. The negative charge of the sulfate ion may compensate for the missing charge of Asp107 in the observed histidine conformation, while a water molecule is located near Asp107. The *trans* conformation is further stabilized by Tyr99 in TRY1 and Thr40 in SPINK1. With this uncommon histidine conformation, TRY1 would not be able to catalyze the cleavage of the P1-P1′ scissile bond, as His63 could not function as a proton acceptor for Ser200, thereby not creating the strongly nucleophilic alkoxide ion. While the p.S200A mutation might destabilize the His63 conformation, many serine protease crystal structures have been solved that contain the same p.S200A mutation, without affecting the histidine conformation (List S1) [9,36,37]. To analyze the impact of SPINK1 binding on the catalytic triad, we employed molecular dynamics simulations by using TRY1 WT and SPINK1 WT complexes, with His63 starting in either *gauche*+ or *trans* (Chi$_1$ = 60° or 180°) conformation. SPINK1 binding sterically blocks the His63 rotamer conformation to flip from one state to the other most of the time (Figure 4A). During the simulation, we observed one *gauche*+ to *trans* and *trans* to *gauche*+ transition, which was preceded by a rotamer change of Thr40 in SPINK1 (Figure 4B,C and Supplementary Video S1). In the *gauche*+ conformation, Ser200

and Asp107 are separated by 7.8 Å, while the distance is reduced to 7.3 Å in the *trans* conformation, potentially blocking easy re-entry of His63 into a productive catalytic triad arrangement (Supplementary Figure S3). Additionally, no relevant crystal contacts are in close proximity to the flipped histidine and the adjacent SPINK1 and TRY1 loops. Therefore, our data provide support for a model in which trypsin inhibition by SPINK1 is in concert with conformational changes within the trypsin active site.
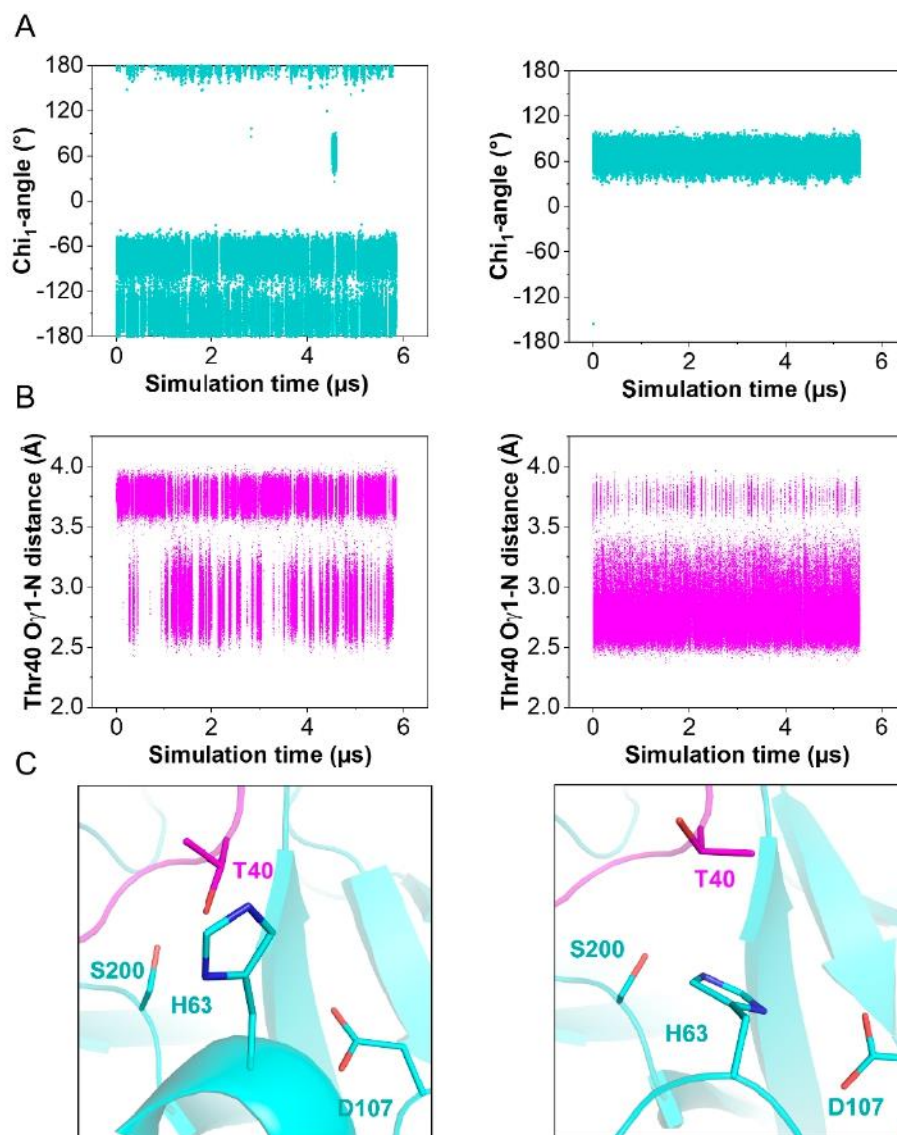
Figure 4. Observed His63 conformations during molecular dynamics simulations. Simulations started with His63 in either *trans* (Chi$_1$ = 180°) or *gauche*$^+$ (Chi$_1$ = −60°) conformation (**left** and **right** panel, respectively) (**A**) Chi$_1$—angles of His63. (**B**) Oγ1—N distance of Thr40 in SPINK1. (**C**) Representative side chain conformations of Thr40 and His63.

3. Discussion

We demonstrated that SPINK1 WT and the p.N34S mutant display similar inhibition constants, binding kinetics and thermodynamic properties upon binding human cationic trypsin (TRY1). This behavior is already well-known within the pancreatic community and has been shown in numerous studies [32,38,39]. Our determined K_d values and association rate constants are higher compared to values reported by Szabó et al. [31] but lower than those reported by Király et al. [38]. Still, all studies show sub-nanomolar affinities and are therefore essentially in agreement with one another. Possibly due to their similar

properties, most carriers of the p.N34S mutation do not develop ARP or CP throughout their entire life [40], despite the prevalence for developing idiopathic chronic pancreatitis (ICP) being increased tenfold. Due to inhibitory properties not being able to explain the pathogenicity of the p.N34S variant, the call for a detailed atomistic model emerged already two decades ago [15]. While several homology models and molecular dynamic simulation studies aimed to meet the demand for such a detailed structure, they are all accompanied by uncertainties, whether a simulated model can accurately reflect reality [41]. Our crystal structures of SPINK1 WT and p.N34S in complex with TRY1 p.S200A provide experimentally determined, high-resolution structures and represent the first structures, where a human Kazal-inhibitor is bound to its natural target protease.

By comparing both structures, we show a slight tilt of the p.N34S variant's α-helix. However, apart from the Arg101–Tyr33 interaction, the loop which contains the p.N34S mutation does not engage in any interactions with TRY1, explaining the high flexibility and limited influence on the interaction of this loop (Figure 2C). While we show minor differences between both structures, they do not translate into different binding constants or other functional differences. Nevertheless, they might contribute to SPINK1 p.N34S being a risk-factor for ARP or CP by other mechanisms. Recently, decreased binding affinity of SPINK1 toward TRY1 was demonstrated after sulfation of Tyr154 in TRY1 [31]. Due to sulfation of the tyrosine, a steric clash with Tyr43 and Pro55 in SPINK1 was proposed, and our crystal structures support this hypothesis (Supplementary Figure S4).

His63 in TRY1 adopts a *trans* conformation rarely found in serine proteases instead of the *gauche*$^+$ conformation [42]. In this conformation, His63 is facing out and away from the catalytic triad. Therefore, the protease's activity is reduced, missing the necessary proton acceptor for Ser200. The remaining members of the catalytic triad remain in their typical conformation. In this study, Ser200 is replaced by an alanine to render the protease inactive for crystallization. The introduced mutation and the crystallization process itself can introduce artefacts that might cause His63 to flip into the *trans* conformation. However, in combination with molecular dynamics simulations, we demonstrate that SPINK1 binding sterically hinders the *trans* to *gauche*$^+$ transition and can affect the His63 conformation. The rotamer change is preceded by a conformational change of Thr40 in SPINK1, and a stabilizing water molecule between Ser200 and Asp107 enters the space usually occupied by His63. The sulfate ion in the crystal structure probably further stabilizes the *trans* conformation and could shift the equilibrium toward the *trans* conformation. Phosphorylation of Ser34 might accomplish the same effect; however, there is currently no indication for post translational modifications in SPINK1. Conversely, a related structure of porcine SPINK1 bound to bovine trypsin with sulfate present displays the *gauche*$^+$ conformation, and other structures containing the p.S200A mutation also show the *gauche*$^+$ conformation [43].

His63 is one of the most important residues in serine protease catalysis [42]. It functions as a proton donor and acceptor throughout the mechanism and allows for the formation of the strongly nucleophilic alkoxide ion. The Ser-His dyad is very conserved among serine proteases and evolved independently numerous times [44,45]. Even though His63 is indispensable for efficient catalysis, substitution by an alanine does not render the protease completely inactive, and, rather, k_{cat} values decrease by three to four orders of magnitude [46,47]. Thus, we assume that the observed conformation of His63 does not inactivate the protease completely, but lowers the rate of SPINK1 proteolysis. Another contributing factor for the proteolytic stability of SPINK1 is Asn56, which bridges the P1$'$ and P2 residues by hydrogen-bonding. Thereby, the binding loop of SPINK1 likely remains rigid even after cleavage of the P1-P1$'$ scissile bond.

4. Conclusions

Our structure–function analysis provides insights into the molecular mechanism, as well as structure, of the SPINK1–TRY1 complex and the p.N34S variant. While we were unable to elucidate the underlying cause of the p.N34S pathogenicity in ARP and CP, we still found minor differences in their complex structures. Previously, structural

differences between both SPINK1 variants were hypothesized to be disease causing or modifying [15,28]. However, our crystal structures prove that both SPINK1 variants are mostly similar, and the minor differences we detected are unlikely to translate into functional or clinical effects.

We are thereby further solidifying the current view of a variant upstream of *SPINK1* promoter being the disease-causing mutation, rather than the p.N34S mutation itself [22]. Moreover, our data may prove useful in the field of cancer research, as the roles of SPINK1 in EGFR binding and granzyme A inhibition are still poorly understood.

In addition, complex formation of both SPINK1 variants is accompanied by flipping of the His63 residue within the catalytic triad. In non-complexed TRY1, His63 adopts a *gauche*$^{+}$ conformation and adopts a *trans* conformation upon complex formation with SPINK1. This mechanism is atypical for Kazal-inhibitors or other proteinaceous inhibitors in general. We speculate that, by this mechanism, SPINK1 elegantly evades proteolysis by TRY1 and thereby displays an increased half-life. Further studies with similar Kazal-inhibitors are planned and are aimed at elucidating whether the observed conformational change in TRY1 upon SPINK1 binding is a mechanism shared among other inhibitors within this family. Such studies will aid in rationalizing the exceptional proteolytic stability of these inhibitors, while also providing a new platform for developing rationally designed proteinaceous protease inhibitors.

5. Materials and Methods

If not otherwise indicated, chemicals were purchased from Sigma (Sigma-Aldrich, Taufkirchen, Germany).

5.1. Protein Expression and Purification

The cDNA coding for amino acids 24-79 of SPINK1 or amino acids 16-247 of TRY1 was cloned into a pET47b expression vector (Novagen, Darmstadt, Germany) in frame with an N-terminal hexahistidine tag (His$_6$-tag) and an HRV3C cleavage site. The N-terminal truncations were introduced to remove the signal peptide sequences for expression in *E. coli*. Mutations were generated by site directed mutagenesis, using QuikChange XL (Agilent Technologies, Santa Clara, CA, USA).

SPINK1 variants were overexpressed in SHuffle T7 Express *E. coli* (New England Biolabs, Frankfurt am Main, Germany) grown in terrific broth medium supplemented with 100 μg mL^{-1} kanamycin at 30 °C. After 16 h of induction at an OD$_{600}$ = 2 with 1 mM isopropyl-β-d-thiogalactoside (IPTG) at 16 °C, cells were harvested by centrifugation and resuspended in lysis buffer containing 20 mM Hepes pH 8, 150 mM NaCl and 20 mM imidazole before lysis by sonication. Clarified lysates were loaded onto a HisTrap HP (Cytiva) affinity column. The column was washed with lysis buffer containing 50 mM imidazole and eluted with lysis buffer containing 500 mM imidazole. The purified SPINK1 was subsequently dialyzed into 20 mM Hepes pH 8, 150 mM NaCl and digested with HRV3C protease (Merck, Darmstadt, Germany). The His$_6$-tag was removed by using the same HisTrap column.

TRY1 variants were overexpressed in NiCo21 (DE3) *E. coli* (New England Biolabs, Frankfurt am Main, Germany) grown in terrific broth medium supplemented with 100 μg mL^{-1} kanamycin at 37 °C. After 16 h of induction at an OD$_{600}$ = 2 with 1 mM IPTG at 25 °C, cells were harvested by centrifugation and resuspended in 0.1 M Tris pH 8 and 5 mM EDTA (Tris-EDTA). Inclusion bodies were prepared as described elsewhere [48]. In brief, inclusion bodies were washed by using Tris-EDTA three times before solubilization in 6 M guanidine hydrochloride (Gdn-HCl), 0.1 M Tris pH 8, 2 mM EDTA and 30 mM dithiothreitol (DTT). After 30 min of incubation at 37 °C and subsequent centrifugation, solubilized inclusion bodies were added to 0.9 M Gdn-HCl, 0.1 M Tris pH 8, 2 mM EDTA, 1 mM L-cystine, 1 mM L-cysteine at a speed of 20 μL min^{-1}, using a syringe pump at 4 °C. After incubation overnight, refolded TRY1 was loaded onto a HisTrap excel (Cytiva, Freiburg, Germany) column. The column was extensively washed with lysis buffer and

eluted by using 500 mM imidazole. Purified TRY1 was dialyzed against 10 mM Hepes pH 8, 75 mM NaCl and 2 mM $CaCl_2$ before digestion by enterokinase (Genescript, Leiden, The Netherlands). His_6-tag and propeptides were removed by size-exclusion chromatography, using a Superdex75 Increase 10/300 GL column (Cytiva, Freiburg, Germany) with 20 mM Hepes pH 8 and 150 mM NaCl.

5.2. Enzyme Inhibition Studies

Inhibition constants (K_i) of SPINK1 WT and N34S mutant were determined by monitoring the conversion of N_α-Benzoyl-L-arginine 4-nitroanilide hydrochloride (L-BAPA, Bachem, Bubendorf, Switzerland) in the presence and absence of varying inhibitor concentrations (15.625–2000 pM) at 405 nm, using a Cytation 5 microplate reader (BioTek, VT, USA). The Michaelis–Menten kinetics of L-BAPA with human cationic trypsin was determined by following the reaction with varying substrate concentrations from 0.02 to 9.5 mM (Equation (1)).

$$v = \frac{v_{max}\,[S]}{K_m + [S]} \tag{1}$$

Initial conversion rates were measured by using 1 mM L-BAPA, 100 pM TRY1 and varying SPINK1 concentrations in 20 mM Hepes pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$ and 0.05% Tween20. All assays were carried out at 37 °C, and initial rates were determined after a 5 h lag phase to ensure sufficient equilibration. Data were fitted by using Morrison's quadratic equation for tight binding inhibitors (Equation (2)).

$$v = v_0 \left(1 - \frac{\left([E] + [I] + K_i\left(1 + \left(\frac{[S]}{K_m}\right)\right)\right) - \sqrt{\left([E] + [I] + K_i\left(1 + \left(\frac{[S]}{K_m}\right)\right)\right)^2 - 4[E]\cdot[I]}}{2[E]} \right) \tag{2}$$

Reported K_i values were determined from fitting the averages of three independent experiments, and standard errors are shown.

5.3. Surface Plasmon Resonance

Surface plasmon resonance (SPR) kinetic analyses were carried out by using a BIAcore T200 instrument (Cytiva, Freiburg, Germany) at 37 °C and a flow rate of 50 µL min^{-1}, and data were recorded at 10 Hz. Standard amine coupling chemistry ((3-dimethylaminopropyl)-3-ethylcarbodiimide(EDC)/N-hydroxysuccinimide(NHS)) was used according to the manufacturer's instructions to covalently immobilize TRY1 p.S200A to a CM5 sensor chip (Cytiva, Freiburg, Germany).

Concentration series of SPINK1 variants were prepared by two-fold dilutions in running buffer (20 mM Hepes pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$, 0.05% Tween20), ranging from 3.125 to 50 nM. Sensorgrams were recorded as single-cycle experiments and were double referenced. After each single-cycle experiment, the surface was regenerated by using 10 mM glycine pH 1.4. For kinetic analyses, sensorgrams were fitted to a 1:1 Langmuir model. Reported kinetic rate constants and dissociation constants represent average and standard deviation of at least three independent experiments.

5.4. Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were carried out by using a MicroCal PEAQ-ITC (Malvern Panalytical, Herrenberg, Germany) instrument at 37 °C and a reference power of 3 µcal s^{-1} in 20 mM Hepes at pH 7.4, 150 mM NaCl and 2 mM $CaCl_2$. Buffers were degassed to ensure signal stability. Delay between injections was set to 240 s. The first injection was rejected from analysis. Subsequently, 18 injections of 3 µL each were carried out. TRY1 p.S200A was placed in the sample cell at 5–10 µM, and SPINK1 was in the syringe at 50–100 µM. Molar binding enthalpies were determined by peak integration, and the heat of dilution was determined from the titration steps at the

end of the experiment and subsequently subtracted. The isotherms were fitted to a one-site binding model. Determined parameters are averages and standard deviations of at least three independent experiments.

5.5. Crystallization and Data Collection

SPINK1–TRY1 p.S200A complexes were concentrated to 20 mg mL^{-1} and crystallized by hanging drop vapor diffusion at 20 °C, using a reservoir solution of 15% PEG 4000 and 0.3 M $(NH_4)_2SO_4$. Crystals between 0.2 and 0.3 mm were soaked briefly in reservoir solution supplemented with 8% PEG 400 as cryoprotectant before cryocooling in liquid nitrogen. Diffraction data were collected at BESSY beamline 14.1 (SPINK1 WT) or 14.2 (SPINK1 N34S). Data collection parameters are given in Table 1.

5.6. Structure Determination

Data reduction was carried out in XDS [49], and phases were solved by molecular replacement in Phaser [50], using human cationic trypsin [51] and a mutated SPINK1 [52] variant as search models (pdb: 1TRN and 1CGI). Refinement was performed by using Refmac5 for the TRY1 p.S200A–SPINK1 N34S complex [53] and in Phenix for the TRY1 p.S200A–SPINK1 WT complex [54]. The higher resolution TRY1 p.S200A–SPINK1 N34S was used as a reference model for weighting restraints during refinement of the SPINK1 WT complex. The model was built in Coot [55]. The crystals are isomorphous and contain two heterodimers per unit cell. NCS restraints were not applied. TLS refinement was used for both structures. Refinement parameters are given in Table 1.

5.7. Molecular Dynamics Simulations

Molecular dynamics simulations of the WT TRY1–SPINK1 complex and His63 starting in either *trans* or *gauche*$^+$ conformations were conducted with NAMD 2.14 [56]. The protein model was built from our WT crystal structure, without the sulfate ion. Force field parameters (CHARMM36) and mutations were added with CHARMM-GUI [57,58]. VMD 1.9.4 was utilized for hydrogen mass repartitioning on the protein and for adding water solvation (TIP3P) and ions to the cubic cell with side length 7 nm [59,60]. Physiological sodium chloride concentration was added at 0.15 mol/L, and the system was neutralized accordingly. The simulation time step was set to 4 fs. Pressures and temperatures were adjusted by a Langevin piston barostat at 100 fs period, 200 fs decay time and Langevin thermostat at 1 ps^{-1} damping time, respectively. The SETTLE algorithm constrained covalent bonds involving hydrogen atoms. Long-range electrostatic interactions were described by the smooth particle mesh Ewald method (sPME), while short-range interactions cutoffs were set to 1 nm with a switching function of 0.1 nm. The system was first subjected to standard energy minimization for 50 k steps and then slowly heated from 100 to 300 K, in steps of 50 K, for 1 ns each. Afterward, the final production runs proceeded at 310 K in an NPT ensemble for 5847 ns and 5048 ns, for *trans* or *gauche*$^+$ conditions, respectively. Samples were collected every 20 ps. Rotational state densities for His63 were extracted with VMD.

References

1. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.E.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [CrossRef] [PubMed]
2. Rawlings, N.D.; Barrett, A.J.; Bateman, A. MEROPS: The peptidase database. *Nucleic Acids Res.* **2010**, *38*, D227–D233. [CrossRef] [PubMed]
3. Papamokos, E.; Weber, E.; Bode, W.; Huber, R.; Empie, M.W.; Kato, I.; Laskowski, M. Crystallographic refinement of Japanese quail ovomucoid, a Kazal-type inhibitor, and model building studies of complexes with serine proteases. *J. Mol. Biol.* **1982**, *158*, 515–537. [CrossRef]
4. Laskowski, J.M.; Kato, I. Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **1980**, *49*, 593–626. [CrossRef]
5. Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162. [CrossRef]
6. Bode, W.; Huber, R. Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **1992**, *204*, 433–451. [CrossRef]
7. Otlewski, J.; Jelen, F.; Zakrzewska, M.; Oleksy, A. The many faces of protease-protein inhibitor interaction. *EMBO J.* **2005**, *24*, 1303–1310. [CrossRef]
8. Rawlings, N.D.; Tolle, D.P.; Barrett, A.J. Evolutionary families of peptidase inhibitors. *Biochem. J.* **2004**, *378*, 705–716. [CrossRef]
9. Zakharova, E.; Horvath, M.P.; Goldenberg, D.P. Structure of a serine protease poised to resynthesize a peptide bond. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11034–11039. [CrossRef]
10. Longstaff, C.; Campbell, A.F.; Fersht, A.R. Recombinant chymotrypsin inhibitor 2: Expression, kinetic analysis of inhibition with alpha-chymotrypsin and wild-type and mutant subtilisin BPN′, and protein engineering to investigate inhibitory specificity and mechanism. *Biochemistry* **1990**, *29*, 7339–7347. [CrossRef]
11. Lu, W.Y.; Starovasnik, M.A.; Dwyer, J.J.; Kossiakoff, A.A.; Kent, S.B.; Lu, W. Deciphering the role of the electrostatic interactions involving Gly70 in eglin C by total chemical protein synthesis. *Biochemistry* **2000**, *39*, 3575–3584. [CrossRef] [PubMed]
12. Peräkylä, M.; Kollman, P.A. Why Does Trypsin Cleave BPTI so Slowly? *J. Am. Chem. Soc.* **2000**, *122*, 3436–3444. [CrossRef]
13. Radisky, E.S.; Koshland, D.E.J. A clogged gutter mechanism for protease inhibitors. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10316–10321. [CrossRef] [PubMed]
14. Witt, H.; Luck, W.; Hennies, H.C.; Classen, M.; Kage, A.; Lass, U.; Landt, O.; Becker, M. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. *Nat. Genet.* **2000**, *25*, 213–216. [CrossRef]
15. Pfützer, R.H.; Barmada, M.M.; Brunskill, A.P.J.; Finch, R.; Hart, P.S.; Neoptolemos, J.; Furey, W.F.; Whitcomb, D.C. SPINK1/PSTI polymorphisms act as disease modifiers in familial and idiopathic chronic pancreatitis. *Gastroenterology* **2000**, *119*, 615–623. [CrossRef] [PubMed]
16. Aoun, E.; Chang, C.C.H.; Greer, J.B.; Papachristou, G.I.; Barmada, M.M.; Whitcomb, D.C. Pathways to injury in chronic pancreatitis: Decoding the role of the high-risk SPINK1 N34S haplotype using meta-analysis. *PLoS ONE* **2008**, *3*, e2003. [CrossRef] [PubMed]
17. Chandak, G.R.; Idris, M.M.; Reddy, D.N.; Mani, K.R.; Bhaskar, S.; Rao, G.V.; Singh, L. Absence of PRSS1 mutations and association of SPINK1 trypsin inhibitor mutations in hereditary and non-hereditary chronic pancreatitis. *Gut* **2004**, *53*, 723–728. [CrossRef]
18. Kereszturi, E.; Király, O.; Sahin-Tóth, M. Minigene analysis of intronic variants in common SPINK1 haplotypes associated with chronic pancreatitis. *Gut* **2009**, *58*, 545–549. [CrossRef]
19. Lukas, J.; Pospech, J.; Oppermann, C.; Hund, C.; Iwanov, K.; Pantoom, S.; Petters, J.; Frech, M.; Seemann, S.; Thiel, F.-G.; et al. Role of endoplasmic reticulum stress and protein misfolding in disorders of the liver and pancreas. *Adv. Med. Sci.* **2019**, *64*, 315–323. [CrossRef]
20. Threadgold, J.; Greenhalf, W.; Ellis, I.; Howes, N.; Lerch, M.M.; Simon, P.; Jansen, J.; Charnley, R.; Laugier, R.; Frulloni, L.; et al. The N34S mutation of SPINK1 (PSTI) is associated with a familial pattern of idiopathic chronic pancreatitis but does not cause the disease. *Gut* **2002**, *50*, 675–681. [CrossRef]
21. Chen, J.-M.; Férec, C. The true culprit within the SPINK1 p.N34S-containing haplotype is still at large. *Gut* **2009**, *58*, 478–480. [CrossRef] [PubMed]

22. Pu, N.; Masson, E.; Cooper, D.N.; Génin, E.; Férec, C.; Chen, J.-M. Chronic Pancreatitis: The True Pathogenic Culprit within the SPINK1 N34S-Containing Haplotype Is No Longer at Large. *Genes* **2021**, *12*, 1683. [CrossRef]

23. Mehner, C.; Radisky, E.S. Bad Tumors Made Worse: SPINK1. *Front. Cell Dev. Biol.* **2019**, *7*, 10. [CrossRef] [PubMed]

24. Räsänen, K.; Itkonen, O.; Koistinen, H.; Stenman, U.-H. Emerging Roles of SPINK1 in Cancer. *Clin. Chem.* **2016**, *62*, 449–457. [CrossRef]

25. Lin, T.-C. Functional Roles of SPINK1 in Cancers. *Int. J. Mol. Sci.* **2021**, *22*, 3814. [CrossRef] [PubMed]

26. Ozaki, N.; Ohmuraya, M.; Hirota, M.; Ida, S.; Wang, J.; Takamori, H.; Higashiyama, S.; Baba, H.; Yamamura, K. Serine Protease Inhibitor Kazal Type 1 Promotes Proliferation of Pancreatic Cancer Cells through the Epidermal Growth Factor Receptor. *Mol. Cancer Res.* **2009**, *7*, 1572–1581. [CrossRef]

27. Lu, F.; Lamontagne, J.; Sun, A.; Pinkerton, M.; Block, T.; Lu, X. Role of the inflammatory protein serine protease inhibitor Kazal in preventing cytolytic granule granzyme A-mediated apoptosis. *Immunology* **2011**, *134*, 398–408. [CrossRef]

28. Kulke, M.; Nagel, F.; Schulig, L.; Geist, N.; Gabor, M.; Mayerle, J.; Lerch, M.M.; Link, A.; Delcea, M. A Hypothesized Mechanism for Chronic Pancreatitis Caused by the N34S Mutation of Serine Protease Inhibitor Kazal-Type 1 Based on Conformational Studies. *J. Inflamm. Res.* **2021**, *14*, 2111. [CrossRef]

29. Sun, Z.; Kolossváry, I.; Kozakov, D.; Sahin-Tóth, M.; Vajda, S. The N34S mutation of SPINK1 may impact the kinetics of trypsinogen activation to cause early trypsin release in the pancreas. *bioRxiv* **2020**. [CrossRef]

30. Ohmuraya, M.; Sugano, A.; Hirota, M.; Takaoka, Y.; Yamamura, K.I. Role of intrapancreatic SPINK1/Spink3 expression in the development of pancreatitis. *Front. Physiol.* **2012**, *3*, 126. [CrossRef]

31. Szabó, A.; Toldi, V.; Gazda, L.D.; Demcsák, A.; Tőzsér, J.; Sahin-Tóth, M. Defective binding of SPINK1 variants is an uncommon mechanism for impaired trypsin inhibition in chronic pancreatitis. *J. Biol. Chem.* **2021**, *296*, 100343. [CrossRef] [PubMed]

32. Buchholz, I.; Nagel, F.; Klein, A.; Wagh, P.R.; Mahajan, U.M.; Greinacher, A.; Lerch, M.M.; Mayerle, J.; Delcea, M. The impact of physiological stress conditions on protein structure and trypsin inhibition of serine protease inhibitor Kazal type 1 (SPINK1) and its N34S variant. *Biochim. Biophys. Acta Proteins Proteom.* **2020**, *1868*, 140281. [CrossRef] [PubMed]

33. Boros, E.; Sebák, F.; Héja, D.; Szakács, D.; Zboray, K.; Schlosser, G.; Micsonai, A.; Kardos, J.; Bodor, A.; Pál, G. Directed Evolution of Canonical Loops and Their Swapping between Unrelated Serine Proteinase Inhibitors Disprove the Interscaffolding Additivity Model. *J. Mol. Biol.* **2019**, *431*, 557–575. [CrossRef] [PubMed]

34. McGrath, M.E.; Haymore, B.L.; Summers, N.L.; Craik, C.S.; Fletterick, R.J. Structure of an engineered, metal-actuated switch in trypsin. *Biochemistry* **1993**, *32*, 1914–1919. [CrossRef] [PubMed]

35. Wheeler, T.J.; Clements, J.; Finn, R.D. Skylign: A tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinform.* **2014**, *15*, 7. [CrossRef] [PubMed]

36. Salameh, M.A.; Soares, A.S.; Navaneetham, D.; Sinha, D.; Walsh, P.N.; Radisky, E.S. Determinants of Affinity and Proteolytic Stability in Interactions of Kunitz Family Protease Inhibitors with Mesotrypsin. *J. Biol. Chem.* **2010**, *285*, 36884–36896. [CrossRef]

37. Salameh, M.A.; Soares, A.S.; Hockla, A.; Radisky, E.S. Structural Basis for Accelerated Cleavage of Bovine Pancreatic Trypsin Inhibitor (BPTI) by Human Mesotrypsin. *J. Biol. Chem.* **2008**, *283*, 4115–4123. [CrossRef]

38. Király, O.; Wartmann, T.; Sahin-Tóth, M. Missense mutations in pancreatic secretory trypsin inhibitor (SPINK1) cause intracellular retention and degradation. *Gut* **2007**, *56*, 1433–1438. [CrossRef]

39. Kuwata, K.; Hirota, M.; Shimizu, H.; Nakae, M.; Nishihara, S.; Takimoto, A.; Mitsushima, K.; Kikuchi, N.; Endo, K.; Inoue, M.; et al. Functional analysis of recombinant pancreatic secretory trypsin inhibitor protein with amino-acid substitution. *J. Gastroenterol.* **2002**, *37*, 928–934. [CrossRef]

40. Yadav, D.; Lowenfels, A.B. The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology* **2013**, *144*, 1252–1261. [CrossRef]

41. Lazim, R.; Suh, D.; Choi, S. Advances in molecular dynamics simulations and enhanced sampling methods for the study of protein systems. *Int. J. Mol. Sci.* **2020**, *21*, 6339. [CrossRef] [PubMed]

42. Jing, H.; Babu, Y.S.; Moore, D.; Kilpatrick, J.M.; Liu, X.-Y.; Volanakis, J.E.; Narayana, S.V.L. Structures of native and complexed complement factor D: Implications of the atypical his57 conformation and self-inhibitory loop in the regulation of specific serine protease activity11. *J. Mol. Biol.* **1998**, *282*, 1061–1081. [CrossRef] [PubMed]

43. Bolognesi, M.; Gatti, G.; Menegatti, E.; Guarneri, M.; Marquart, M.; Papamokos, E.; Huber, R. Three-dimensional structure of the complex between pancreatic secretory trypsin inhibitor (Kazal type) and trypsinogen at 1.8 A resolution: Structure solution, crystallographic refinement and preliminary structural interpretation. *J. Mol. Biol.* **1982**, *162*, 839–868. [CrossRef]

44. Perona, J.J.; Craik, C.S. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **1997**, *272*, 29987–29990. [CrossRef]

45. Perona, J.J.; Craik, C.S. Structural basis of substrate specificity in the serine proteases. *Protein Sci.* **1995**, *4*, 337–360. [CrossRef]

46. Corey, D.R.; Craik, C.S. An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *J. Am. Chem. Soc.* **1992**, *114*, 1784–1790. [CrossRef]

47. Corey, D.R.; Willett, W.S.; Coombs, G.S.; Craik, C.S. Trypsin specificity increased through substrate-assisted catalysis. *Biochemistry* **1995**, *34*, 11521–11527. [CrossRef]

48. Sahin-Tóth, M.; Szabó, A. Expression, purification and activity measurements of pancreatic proteases. *Pancreapedia Exocrine Pancreas Knowl. Base* **2011**. [CrossRef]

49. Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132. [CrossRef]

50. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674. [CrossRef]

51. Gaboriaud, C.; Serre, L.; Guy-Crotte, O.; Forest, E.; Fontecilla-Camps, J.-C. Crystal Structure of Human Trypsin 1: Unexpected Phosphorylation of Tyr151. *J. Mol. Biol.* **1996**, *259*, 995–1010. [CrossRef] [PubMed]

52. Hecht, H.J.; Szardenings, M.; Collins, J.; Schomburg, D. Three-dimensional structure of the complexes between bovine chymotrypsinogen A and two recombinant variants of human pancreatic secretory trypsin inhibitor (Kazal-type). *J. Mol. Biol.* **1991**, *220*, 711–722. [CrossRef]

53. Murshudov, G.N.; Vagin, A.A.; Dodson, E.J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. Sect. D* **1997**, *53*, 240–255. [CrossRef]

54. Liebschner, D.; Afonine, P.V.; Baker, M.L.; Bunkoczi, G.; Chen, V.B.; Croll, T.I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A.J.; et al. Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. Sect. D* **2019**, *75*, 861–877. [CrossRef] [PubMed]

55. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 486–501. [CrossRef]

56. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]

57. Huang, J.; MacKerell, J.A.D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135–2145. [CrossRef]

58. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [CrossRef]

59. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]

60. Hopkins, C.W.; Le Grand, S.; Walker, R.C.; Roitberg, A.E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11*, 1864–1874. [CrossRef]

Supporting Information
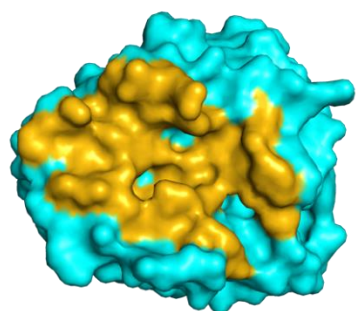
935.5 Å²

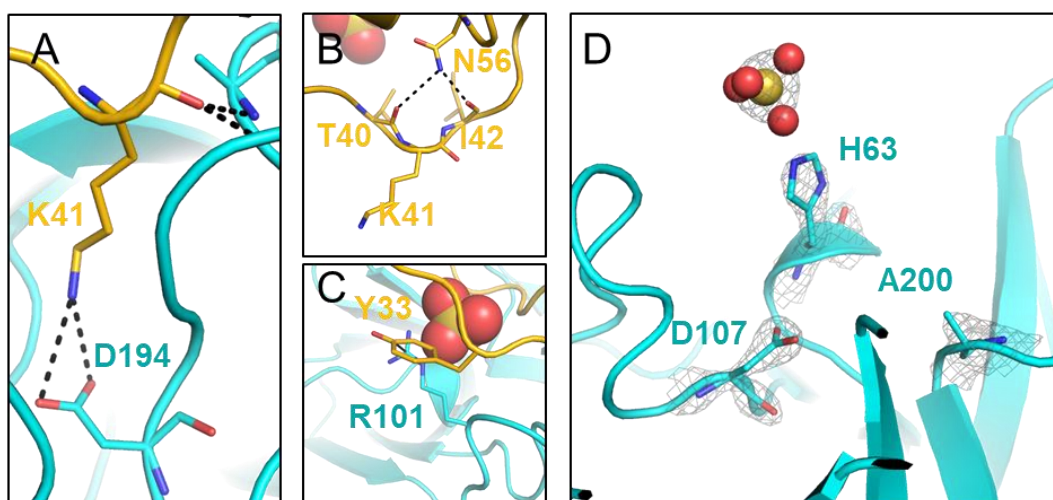**Figure S1.** TRY1 p.S200A – SPINK1 WT interaction interface.

Figure S2. Interactions between TRY1 p.S200A and SPINK1 WT. SPINK1 WT is shown in yellow and TRY1 p.S200A in cyan. (**A**) Lys41 in SPINK1 interacts with the specificity determining Asp194 in TRY1. (**B**) Asn56 in SPINK1 stabilizes the binding loop by hydrogen bonding with Thr40 and Ile42. (**C**) Tyr33 in SPINK1 forms a cation – pi bond with Arg101 in TRY1 and is hence pulled outwards. (**D**) Catalytic triad of the TRY1 p.S200A – SPINK1 complex. In the complex structure His63 faces towards the sulfate ion and away from the catalytic triad. The 2Fo – Fc density map is shown 1.6 Å around the residues of the catalytic triad and is contoured at 1.0 σ.
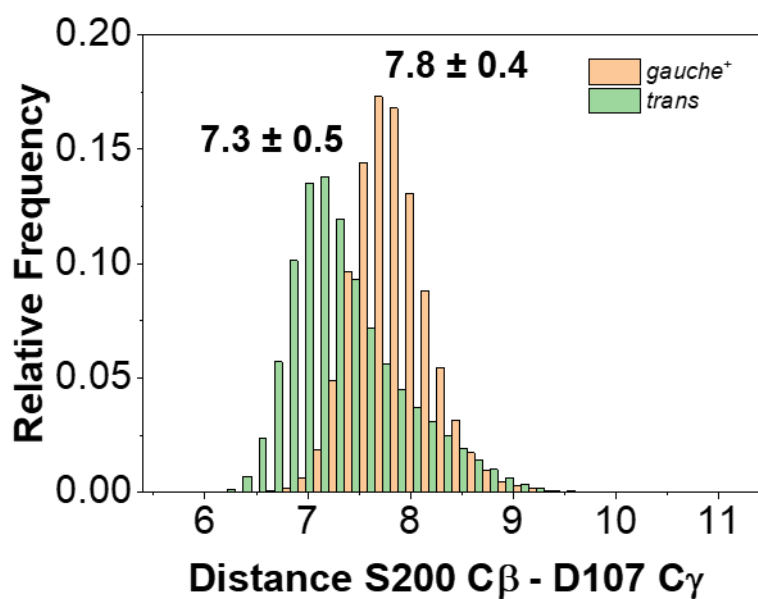
Figure S3. Distribution of distances between Ser200 and Asp107 of the catalytic triad during MDS. Distances were evaluated between complexes with His63 in *trans* (green) or *gauche*[+] (orange) conformation.
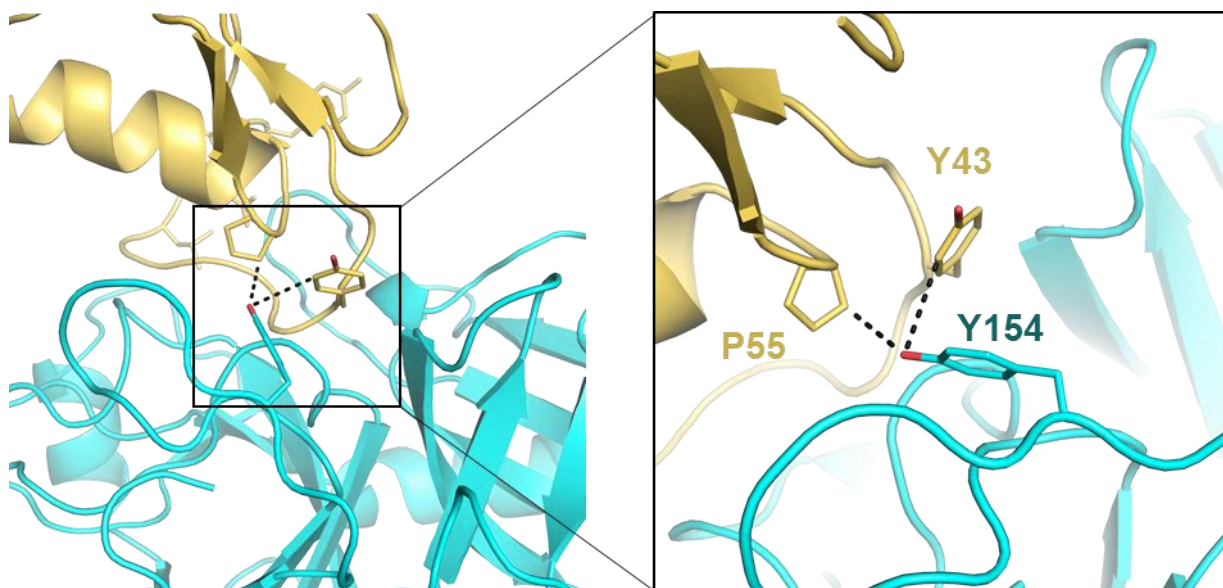

Figure S4. Interactions of Tyr154 with SPINK1 WT. SPINK1 WT is shown in yellow and TRY1 p.S200A in cyan. Tyr154 interacts with Tyr43 and Pro55 in SPINK1. Physiologically Tyr154 is sulfated, which weakens the interaction of SPINK1 with TRY1 due to steric hindrance.

List S1. Serine protease structures with p.S200A mutations. 4/74 structures (5.4 %) show alternative His63 conformations. Structures with alternative His63 conformations are highlighted in red.

4FXG	5JPM	2XW9	2XWB	5ZLZ	5BRR	4XSK	3PB1	2O8W	2O8U	4ZKS	5WXT	2PUX
2PV9	3K65	4HZH	4RN6	3SQE	3LU9	1TB6	4DT7	3SQH	3B9F	2B5T	1DM4	1JOU
1JMO	4H6S	4H6T	3KCG	4IS5	4ISL	5LYO	6BQM	6AOD	6TS4	6I44	1L4D	1BML
1DDJ	1L4Z	6D3Z	1MZA	1MZD	2GD4	5VOE	5VOF	2R9P	3L3T	5C67	3P92	4U32
6HAR	3L33	6GFI	5JBT	4U30	6BX8	3P95	4DG4	4B2B	4B2C	4B2A	4B1T	1OPH
4WXV	2RA3	3TGJ	1YLC	3FP8	1YKT	3FP7	1YLD	1K9O				

Article IV

Identification of Kazal-inhibitor scaffolds with identical canonical binding loops and their effects on binding properties

Authors: Felix Nagel[1], Anne Susemihl[1,2], Tobias Felix Eulberg[1], Mihaela Delcea[1*]

[1] Biophysical Chemistry, Institute of Biochemistry, University of Greifswald, Greifswald, Germany

[2] Department of Hematology and Oncology, Internal Medicine C, University of Greifswald, Greifswald, Germany

*Corresponding author: Mihaela Delcea, delceam@uni-greifswald.de

Running title: Influence of Kazal-inhibitor scaffolds on binding properties

Keywords

pancreas, serine protease, protease inhibitor, circular dichroism (CD), surface plasmon resonance (SPR), SPINK1

Abbreviations

CD, circular dichroism; k_a, association rate; K_d, dissociation constant; k_d, dissociation rate; K_i, inhibition constant; NM, naked mole; SPINK1, serine protease inhibitor Kazal type 1; SPR, surface plasmon resonance; $t_{1/2}$, half-life; TRY1, human cationic trypsin; TRY2, human anionic trypsin; ΔG, change in free energy; ΔH, change in enthalpy; ΔS, change in entropy

Abstract

Kazal inhibitors are widely used as scaffolds for therapeutic molecules, taking advantage of the easily exchangeable canonical binding loop. Different Kazal inhibitor backbones have been suggested to be therapeutically useful, but the impact of different Kazal-like scaffolds on binding properties is still largely unknown. Here, we identified trypsin-targeting human serine protease inhibitor Kazal type 1 (SPINK1) homologues in different mammalian species that cluster in two P2-P1 combinations implying the co-evolution of these residues. We generated loop exchange variants of human SPINK1 for comparison with Kazal inhibitors from related species. Using comprehensive biophysical characterization of the inhibitor-enzyme interactions we found not only affinity but also pH resistance to be highly backbone dependent. Differences are mostly observed in complex stability, which varies by over one order of magnitude. We provide clear evidence for high backbone dependency within the Kazal family. Hence, when designing Kazal inhibitor based therapeutic molecules, testing different backbones after optimizing the canonical binding loop can be beneficial and may result in increased affinity, complex stability, specificity and pH resistance.

Introduction

Canonical serine protease inhibitors are divided into 18 families which all share a common mechanism (1, 2). Serine proteases are inhibited through reversible tight-binding interactions, commonly described as the "standard mechanism" (3). The 18 inhibitor families are contained within 13 different clans, indicating that this mechanism of inhibition evolved independently numerous times. Binding of a canonical inhibitor is mediated by the binding loop present in all inhibitors of this family. The inhibitor interacts with the protease in a substrate-like manner, which is followed by cleavage of the P1-P1' scissile bond (Schechter and Berger nomenclature) (4). Protease and inhibitor exist in an equilibrium between intact and cleaved reactive site peptide bond and after complex dissociation the free inhibitor molecule can exist in both the intact and cleaved state (5, 6). The reactive site of the inhibitor is contained within structurally similar loops whose conformation and geometry upon complex formation are essentially identical in all 18 families (3, 7–9). Despite the varying supporting scaffolds of each protein family, binding was thought to be independent of the scaffold and mostly mediated by the canonical binding loop itself, with each amino acid contributing independently towards binding (10, 11). Hereof, inter- and intra-scaffolding additivity principles were extrapolated, however exceptions have recently been found (12). Loop exchange variants between a Kazal and Pacifastin inhibitor showed that the Kazal inhibitor loop distorts the weak pacifastin scaffold, while the Pacifastin binding loop was distorted by the stable Kazal inhibitor scaffold.

However, the loop exchange variants were modified since the Pacifastin inhibitor possesses a completely conserved P3' cysteine residue that participates in disulfide bonding within the inhibitor and therefore, could not be omitted without destabilizing the overall folding of the inhibitor. Hence, as an alternative approach, we focused on loop exchange variants of a single inhibitor family. Through phylogenetic analysis of the human serine protease inhibitor Kazal type 1 (SPINK1) we identified two different P2-P1 amino acid combinations among different mammal species. It appears that either a Thr-Lys or Pro-Arg pair has evolved, which would contradict the aforementioned additivity principles, as a crosstalk between the P2 and P1 residues is implied. Therefore, we generated the four possible P2-P1 combinations (Thr-Lys, Pro-Lys, Thr-Arg, Pro-Arg) in SPINK1 and identified Kazal inhibitors of other organisms that feature binding loops with sequences identical to the mutated human SPINK1 loop. Due to the lack of productive contacts between the Kazal domain backbone and the target protease outside the canonical binding loop, identical binding behavior was expected (Figure S1). However, here we show, that binding of Kazal inhibitors can be highly backbone dependent with the Kazal scaffold influencing complex stability, as well as pH dependency of the interaction. A similar effect was previously observed for loop exchange variants of different Kunitz type inhibitors, but only minor effects of the scaffold on the overall affinity were found (13). Recently, SPINK2 was proposed as a potential scaffold for designing novel therapeutic molecules. Furthermore, already active patents describe the therapeutic use of SPINK1 and variants thereof as anticancer or anticoagulation agents (14–16). Our findings indicate that optimizing the supporting Kazal domain may be a viable approach for increasing the inhibitor's specificity, pH resistance, complex stability and affinity.

Results

Phylogenetic Analysis of Trypsin-Binding Kazal Inhibitors

We carried out phylogenetic analyses of human SPINK1 related proteins and identified two predominant P2 and P1 residues (Figure 1A). Consistent with the lack of a dedicated S2 pocket in trypsin, the P2 residue in the different SPINK1 homologues was the least conserved residue within the canonical binding loop (Figure 1B). The P1 residues were all positively charged, containing either lysine or arginine. Sub-clustering according to P1 residues revealed distinct P2-P1 combinations with either Thr-Lys or Pro-Arg in the canonical binding loop (Figure 1C). Although not possessing a binding pocket for the P2 residue, trypsin can influence the orientation of the P2 residue by L104 and H63 in the case of a threonine side chain (Figure 1D). Proline however, does not interact with the protease, but assists the overall loop rigidity and aids in maintaining a favorable geometry. We proceeded to generate human SPINK1

(hSPINK1) variants containing the four possible P2-P1 combinations and identified mouse (mSPINK1) and naked mole SPINK1 (nmSPINK1) proteins as homologues containing identical canonical binding loops to the generated variants.
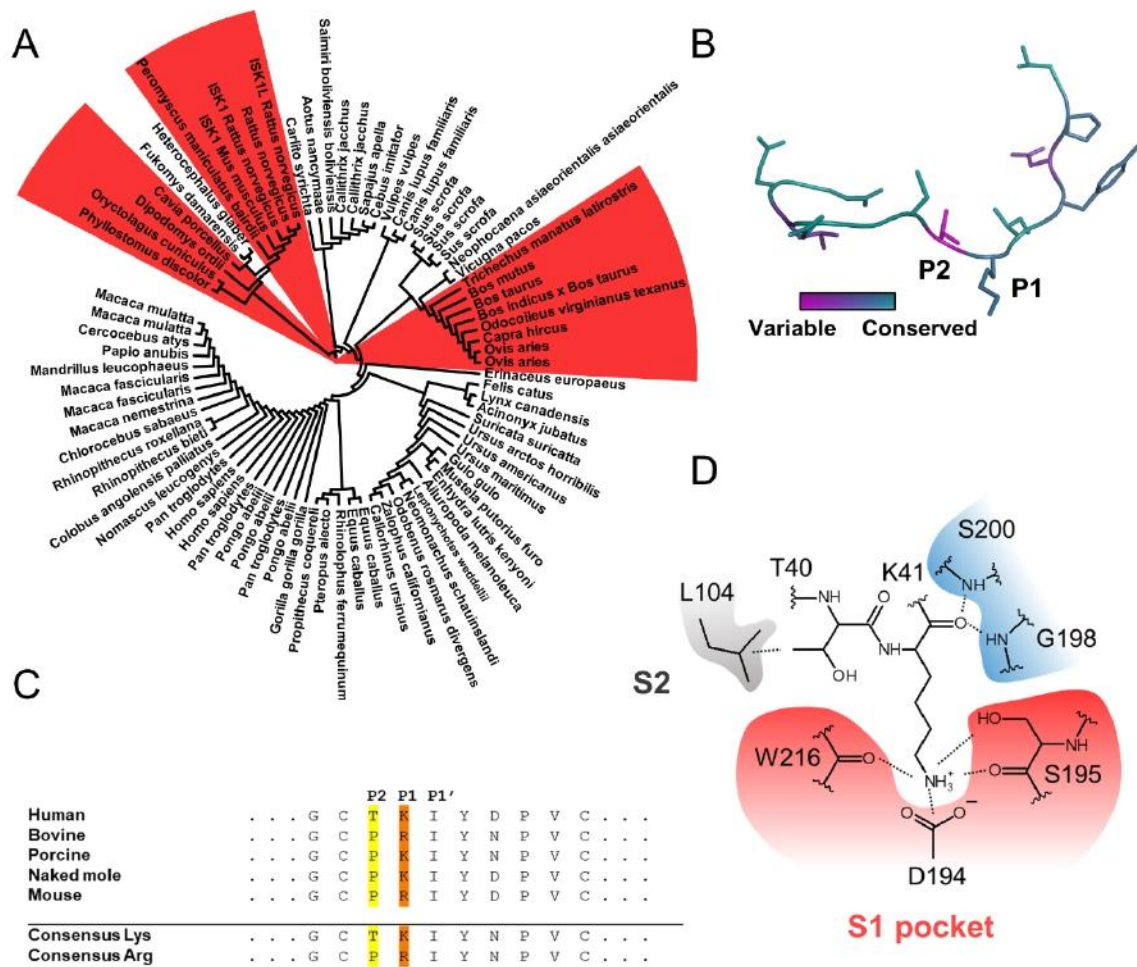

Figure 1. SPINK1 sequence and structure conservation. (A) Phylogenetic analysis of trypsin-targeting SPINK1 homologues. Duplicates indicate different isoforms within the same species. SPINK1 variants featuring arginine as P1 residue are highlighted in red. (B) Structure of the canonical binding loop (PDB ID: **7QE8**) colored according to conservation using the multiple sequence alignment from (A). Conservation scores were determined using the ConSurf server (17). (C) Multiple sequence alignment of selected SPINK1 variants and consensus sequences of (A) sub-clustered according to P1 residues. (D) Structure of binding pockets in trypsin. The S1 pocket is highlighted in red, the oxyanion hole in blue and L104 interacting with the P2 residue is colored gray.

Kinetic and Thermodynamic Profiling of SPINK1 – Trypsin Interactions Reveals High Backbone Dependency

In addition to available crystal structures, missing hSPINK1 structures for the generated mutants, as well as mSPINK1 and nmSPINK1, were modelled using AlphaFold (18). All structures and models show or predict a similar Kazal-like fold (Figure 2A). Circular dichroism spectroscopy shows similar spectra for all human variants, indicating that the substitution of the P2-P1 residues did not destabilize the overall fold. mSPINK1 shows higher alpha-helix

content compared to the human counterparts, while nmSPINK1 shows a lower alpha-helix content (Figure 2B). Purity and homogeneity of the proteins was verified by tricine SDS-PAGE (Figure 2C) and analytical size exclusion chromatography (Figure 2D). The refolded human cationic trypsin (TRY1) was characterized using L-BAPA as model substrate (Figure 2E). The obtained K_m-values were used for determining K_i-values of the SPINK1 variants with TRY1 at 37°C, pH 7.4 (Figure 2F). Highest affinity binders were hSPINK1 T40P and mSPINK1. Surprisingly, nmSPINK1 showed an almost 20-fold decrease in affinity compared to hSPINK1 T40P, despite containing an identical binding loop. Kinetics of the interactions were characterized by surface plasmon resonance (SPR) with TRY1 as ligand and SPINK1 variants as analytes (Figure 2G). Binders displaying the highest affinity were again both human T40P variants as well as mSPINK1. Additionally, the higher affinity is almost exclusively caused by higher complex stability, which results in a half-life of 10 h for hSPINK1 T40P compared to 1 h for hSPINK1 WT (Figure 2H). Even though nmSPINK1 has the same binding loop as hSPINK1 T40P, the complex stability is much lower, indicating a major impact of the Kazal backbone on the overall binding properties of the inhibitor. Finally, thermodynamic analyses between all hSPINK1 variants show that the proline substitution at the P2 site results in a more favorable entropic contribution to binding, while minor enthalpy-entropy compensation is also observed (Figure 2I). T40P was omitted from the analyses due to very slow dissociation rates at lower temperatures. However, the overall impact of the proline substitution can be inferred from comparing the K41R and T40P K41R variants. Through transition state theory, we came to a similar conclusion, where the T40P substitution results in a lower enthalpic energy barrier of the transition state, but a much higher entropic barrier resulting in an overall more favorable binding of proline variants (Figure 2J).
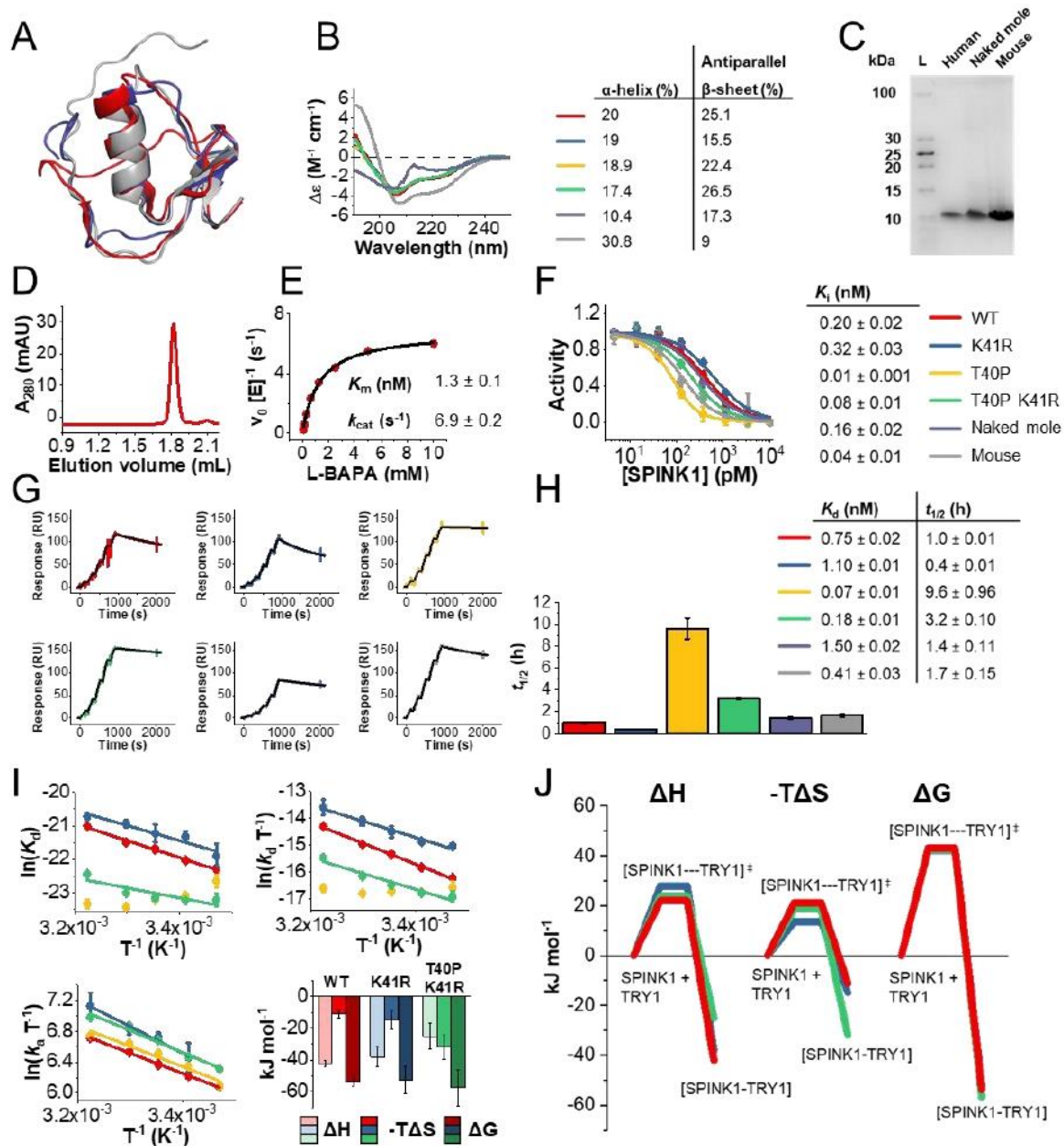
Figure 2. Kinetic and thermodynamic characterization of SPINK1-TRY1 interactions. (A) Superposition of the crystal structure of SPINK1 WT (PDB ID: **7QE8**, red) and AlphaFold models of naked mole (purple) and mouse SPINK1 (gray). (B) Far UV circular dichroism spectra of SPINK1 variants with their secondary structure contents determined by deconvolution in BeStSel (19). (C) Tricine SDS-PAGE of purified human, naked mole and mouse SPINK1. (D) Analytical size exclusion chromatography of purified SPINK1. (E) Michaelis-Menten kinetics of TRY1 and L-BAPA at pH 7.4 and 37°C. (F) TRY1 equilibrium activity assays at 37°C and pH 7.4 with different SPINK1 variants and K_i-values obtained with Equation 5. (G) Single cycle kinetic experiments of SPINK1-TRY1 interactions using surface plasmon resonance at pH 7.4 and 37°C. Experimental data are shown in color (as in F) and 1:1 Langmuir fits are represented in black. (H) Complex half-life ($t_{1/2}$) of different SPINK1-TRY1 interactions obtained from (G) and corresponding K_d-values. (I) Van't Hoff and Eyring analyses of interactions between human SPINK1 variants and TRY1 at different temperatures. (J) Energy landscapes of selected SPINK1-TRY1 interactions.

Kinetic Profiles of SPINK1 Variants Interacting with Different Trypsin Homologues

Due to the unexpected preference of TRY1 for proline as P2 residue despite the wild type inhibitor carrying a threonine in that position, we extended the characterization to other related proteases. Aiming to identify whether this P2 preference is a general feature for trypsin homologues or specific to human trypsin, we characterized the interaction with human anionic trypsin (TRY2), bovine, and porcine trypsin, and all SPINK1 variants (Figure S2, Table S1).

Comparison between affinities of the different SPINK1 variants with different trypsin isoforms reveals the most prominent difference for the hSPINK1 T40P variant with TRY1 and TRY2 (Figure 3A). Conversely, bovine and porcine trypsin are barely affected by substituting the P2 residue with proline. Bovine trypsin is not affected by minor changes in the canonical binding loop within the hSPINK1 scaffold but strongly disfavors the mSPINK1 and nmSPINK1 variants despite them having identical binding loops to the human counter parts. Except for the hSPINK1 T40P K41R – porcine trypsin combination, all affinity changes are caused by alterations in complex stability rather than association rates (Figure 3B). Hence, substituting the P2 residue in hSPINK1 WT with a proline residue selectively improved binding towards human trypsin isoforms, but not towards trypsin from the other species tested.
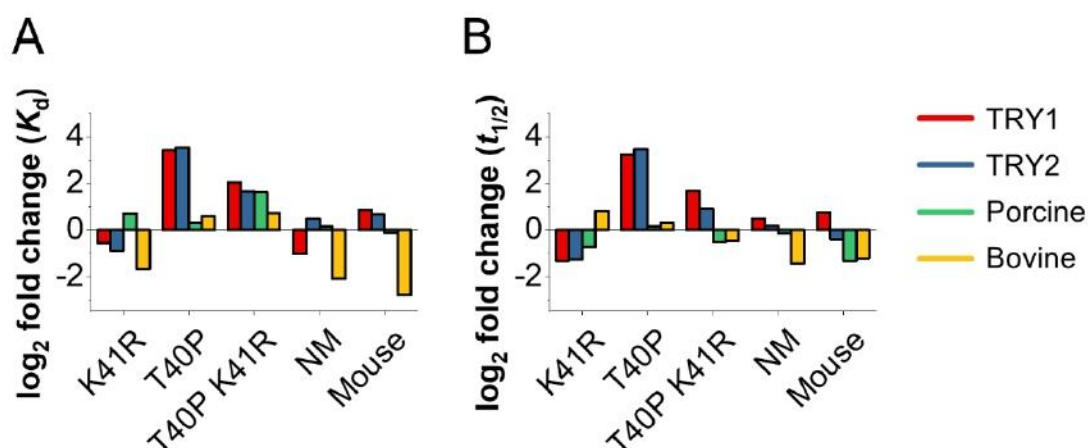

Figure 3. Influence of mutations in the canonical binding loop and changes in inhibitor scaffold on affinity and complex stability. (A) Changes in affinity obtained from SPR experiments of different SPINK1 variants towards different trypsin variants compared to human SPINK1 WT. (B) Changes in complex stability of different SPINK1 variants towards different trypsin variants compared to human SPINK1 WT.

Free energy contributions of Scaffolds, P1, and P2 residues

We showed that substitution of the P2 and P1 residues, as well as the protein scaffold, can influence the binding to trypsin variants (Table 1). Our data enabled us to dissect the individual free energy contributions of each feature. Using the available kinetic data, we can compare

the contribution to the transition free energy of association and dissociation as well as in equilibrium (Equation 1-3). The inhibitor scaffold has a greater influence on overall binding than replacing the P1 residue with lysine or arginine. Proline as P2 residue exerts the greatest influence on binding, most likely by stabilizing the overall loop geometry due to limited degrees of freedom.

$$\Delta\Delta^{\ddagger}G_a = -RT\ln\left(\frac{k_a(\text{variant A})}{k_a(\text{variant B})}\right) \tag{1}$$

$$\Delta\Delta^{\ddagger}G_d = -RT\ln\left(\frac{k_d(\text{variant A})}{k_d(\text{variant B})}\right) \tag{2}$$

$$\Delta\Delta G_d = -RT\ln\left(\frac{K_d(\text{variant A})}{K_d(\text{variant B})}\right) \tag{3}$$

Table 1. Impact of binding loop mutations and inhibitor scaffold on affinity and kinetic rate constants of SPINK1 variants interacting with TRY1 at 37°C.

Variant A	Variant B	k_a difference (log$_2$ fold)	k_d difference (log$_2$ fold)	K_d difference (log$_2$ fold)	$\Delta\Delta^{\ddagger}G_a$ (kJ mol^{-1})	$\Delta\Delta^{\ddagger}G_d$ (kJ mol^{-1})	$\Delta\Delta G_d$ (kJ mol^{-1})
Influence of point mutations within the canonical binding loop							
WT	K41R	-0.8	-1.3	-0.6	1.4	2.4	1.0
WT	T40P	-0.2	3.3	3.4	0.3	-5.8	-6.1
WT	T40P K41R	-0.4	1.7	2.1	0.7	-3.0	-3.7
Influence of scaffold substitution							
T40P	Naked mole	1.6	-2.8	-4.4	-2.9	5.0	7.9
T40P K41R	Mouse	0.3	-0.9	-1.2	-0.5	1.6	2.1

Kazal Inhibitor Backbones Influence pH Dependency

We investigated the SPINK1 – TRY1 interaction on pH sensitivity. Lowering the pH below the catalytic optimum of trypsin of pH ~7 is known to decrease the affinity for the SPINK1 inhibitor. Using SPR, we observed lower affinity for all SPINK1 variants at pH 5 (Figure 4A). For human SPINK1, K_d-values increased 16-fold compared to the same interaction at pH 7.4, regardless of the P2-P1 variant used (Figure 4B). Decreases in affinity were also prominent for the naked mole and mouse variants, which would indicate similar susceptibility to pH changes. Upon analyzing the kinetic rate constants however, we found the largest decrease in complex stability for naked mole and mouse SPINK1. Simultaneously, naked mole and mouse SPINK1 association rates increased up to 16-fold compensating for the lower complex stability

(Figure 4B and 4C). Hence, we show that both association and dissociation rate are highly backbone dependent and that different Kazal motifs can lower the pH sensitivity of the inhibitor. However, caution should be exercised when considering equilibrium binding data. Here, K_d-values would suggest that naked mole and mouse SPINK1 are as susceptible to pH changes as the human counterparts, when in actuality, their interactions are the most influenced by the lower pH out of all variants.
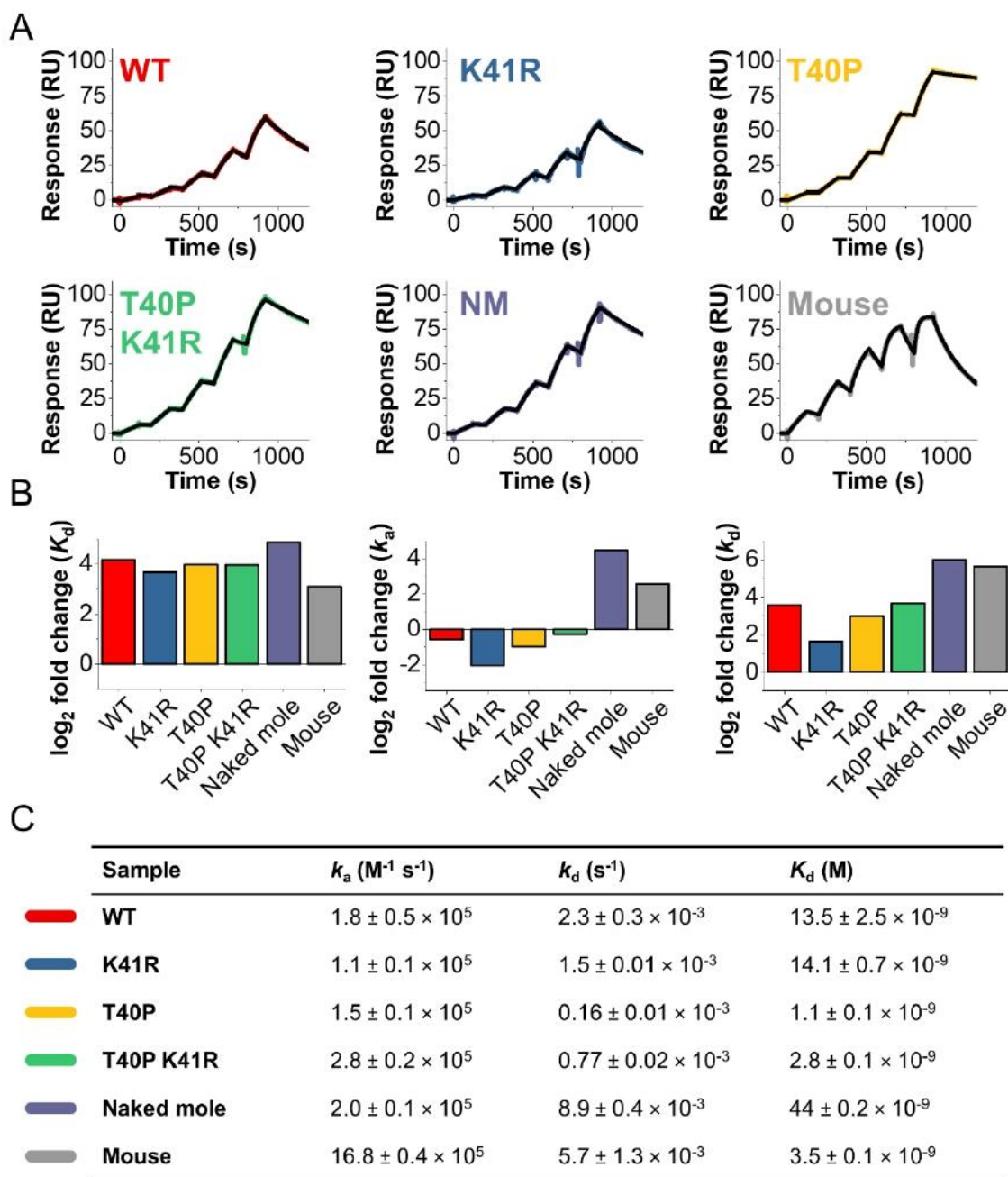

Figure 4. Influence of pH on binding kinetics of SPINK1-TRY1 interactions. (A) Single cycle kinetics of SPINK1-TRY1 interactions using SPR at 37°C and pH 5. Experimental data is shown in colors and 1:1 Langmuir fits are shown in black. (B) Change in affinity and kinetic rate constants compared to pH 7.4. (C) Summary of kinetic rate constants and dissociation constants obtained from (A).

Discussion

We demonstrate by various biophysical techniques that the interaction between Kazal inhibitors and different trypsin variants can be backbone dependent despite no apparent productive contacts outside the canonical binding loop. Using phylogenetic analysis of trypsin-targeting Kazal inhibitors, homologous to human SPINK1, we identified two major P2-P1 pairs occurring naturally. Thr-Lys or Pro-Arg combinations appear to be preferred in the majority of cases. We generated the four possible P2-P1 combinations in human SPINK1 and recombinantly expressed naked mole and mouse SPINK1, which natively feature identical binding loops to the SPINK1 T40P or SPINK1 T40P K41R variants, respectively.

Unexpectedly, despite the lack of a designated S2 binding pocket in trypsin, the hSPINK1 T40P variant displayed a remarkable complex half-life of 10 hours, which represents a 10-fold increase compared to hSPINK1 WT. Most likely, proline in the P2 position helps rigidifying the binding loop, locking it in the canonical conformation (3, 9, 20). In contrast to Kunitz inhibitors, replacement of the P1 Lys residue with Arg weakened the interaction, demonstrating a clear lysine preference of human trypsins for Kazal inhibitors (13, 21). Although the Pro-Lys combination resulted in by far the most potent inhibitor, this combination is rarely encountered in nature. The evolutionary rationale behind the low occurrence of this variant remains to be elucidated, but unwanted cross-reactions and lower specificity in secondary functions apart from trypsin inhibition may offer one possible explanation.

Despite naked mole and mouse SPINK1 featuring identical binding loops to the human T40P and T40P K41R variants, K_d-values varied up to 10-fold under physiological conditions. At lower pH values the differences amplified to up to 40-fold. In previous studies, loop exchange variants of the Kunitz inhibitors amyloid precursor protein inhibitor domain (APPI) and bovine pancreatic trypsin inhibitor (BPTI) interacting with mesotrypsin were described, but only minor differences in affinity between both scaffolds were found (13). However, large differences in k_{cat}-values were reported indicating an influence of the inhibitor scaffold beyond supporting the canonical binding loop. The authors concluded that the binding affinity depends mostly on the binding loop, whereas proteolysis of the inhibitor by mesotrypsin can be highly scaffold dependent. Our results indicate, that in the case of Kazal inhibitors, the inhibitor's scaffold can contribute significantly to the overall affinity. Against the backdrop of recent proposals such as using SPINK2 as scaffold for therapeutic inhibitors and several filed patents describing SPINK1 and related mutants as therapeutic molecules in coagulation and cancer therapy, our findings suggest that, while very potent inhibitors can be obtained from a single Kazal scaffold, performing a secondary scaffold screen will likely create more potent, specific, and resistant inhibitors (14–16).

Furthermore, we show scaffold and loop selectivity to be protease- and species specific, implying that care must be taken when optimizing inhibitors in different animal models. For pancreatitis and trypsin targeting molecules specifically, transferring results obtained from animal models to clinical trials is often not possible and outcomes are mostly unpredictable (22). Here, we offer insights into the kinetic and thermodynamic behavior of different trypsin-inhibiting Kazal inhibitors, partly explaining the lack of correlation between results from animal models and clinical trials.

In conclusion, we identified a hSPINK1 T40P variant displaying 10-fold increased potency for trypsin inhibition. We also show high scaffold dependency and illuminate the underlying kinetic and thermodynamic contributions to the different binding behaviors of scaffolds containing identical binding loops. Our data show that inhibitor-protease pairs form one highly integrated unit, and a secondary scaffold screen after optimizing the canonical binding loop is advisable. Optimizing the binding loop supporting scaffold can increase affinity, specificity, pH resistance and complex stability resulting in higher potency inhibitors. In the context of pancreatitis, our results offer a biophysical framework for protease-inhibitor interactions and may help transfer results obtained from animal models to clinical trials.

Experimental procedures

Protein Expression and Purification

SPINK1 variants and human trypsin isoforms were expressed and purified as previously described (23). In brief, SPINK1 was expressed in SHuffle T7 Express *E. coli* (New England Biolabs, Frankfurt am Main, Germany). After induction and overnight expression at 16°C, cells were lysed by sonication and purified by immobilized metal ion affinity chromatography (IMAC). His-tags were removed by cleavage with the HRV3C protease and proteins were purified again by IMAC. Human trypsin isoforms were expressed in *E. coli* BL21 (DE3) (New England Biolabs) overnight at 30°C. Inclusion bodies were washed before solubilization and refolded in 0.9 M Gdn-HCl, 0.1 M Tris pH 8, 2 mM EDTA, 1 mM L-cysteine, 1 mM L-cystine at 4°C overnight. Refolded trypsin isoforms were purified by IMAC and activated with enterokinase before use. Bovine and porcine trypsin were commercially available (Sigma Aldrich, Taufkirchen, Germany) and were thus, not expressed recombinantly. Protein purity and homogeneity was verified by tricine SDS PAGE and analytical size exclusion chromatography. Calibration of the analytical size exclusion can be found elsewhere (24).

Circular Dichroism Spectroscopy

Circular dichroism (CD) spectra of SPINK1 variants were recorded between 190 and 250 nm using a Chirascan V100 CD spectrometer (Applied Photophysics, Leatherhead, UK) in 1 mm quartz cuvettes containing 250 µg mL^{-1} of protein in 10 mM Tris pH 7.4 and 10 mM NaF, with adaptive sampling turned on and a step size of 0.5 nm. Three scans were accumulated and averaged before deconvolution using Bestsel (19).

Activity Assays

Inhibition constants (K_i) of SPINK1 variants were determined by monitoring the conversion rate of N$_\alpha$-Benzoyl-L-arginine 4-nitroanilide hydrochloride (L-BAPA, Bachem, Bubendorf, Switzerland) at varying inhibitor concentrations ranging from 4.5 pM to 10 nM. The absorbance was recorded at 405 nm using a Cytation 5 microplate reader (BioTek, VT, USA). Michaelis-Menten kinetics of L-BAPA with different trypsin variants were determined using L-BAPA concentrations from 0.02 to 10 mM and fitted using Equation 4.

$$v = \frac{v_{\max}[S]}{K_{\mathrm{m}} + [S]} \qquad (4)$$

For K_i determination, 100 pM of enzyme and 1 mM L-BAPA were used. Assays were carried out at 37°C in 20 mM Hepes pH 7.4, 150 mM NaCl, 2 mM CaCl$_2$ and 0.05% Tween20 and initial rates were determined after a 24 h lag phase at 37°C. Data were fitted with Morrison's quadratic equation using previously determined K_{m}-values (Equation 5).

$$v = v_{\max}\left(1 - \frac{\left([E] + [I] + K_i(1 + (\frac{[S]}{K_{\mathrm{m}}}))\right) - \sqrt{\left([E] + [I] + K_i(1 + (\frac{[S]}{K_{\mathrm{m}}}))\right)^2 - 4[E] \cdot [I]}}{2[E]}\right) \qquad (5)$$

Reported K_i-values are determined from fitting the averages of three independent experiments.

Surface Plasmon Resonance

Kinetic analyses were carried out using a BIAcore T200 instrument (Cytiva, Freiburg, Germany) at 37°C, a flow rate of 50 µL min^{-1} and a data collection rate of 10 Hz. Trypsin was covalently immobilized to a CM5 sensor chip (Cytiva) using an amine coupling kit according to the manufacturer's instructions (Cytiva). A concentration series of SPINK1 was prepared as two-fold dilutions from 3.125 to 50 nM in HBS-P+ buffer supplemented with 2 mM CaCl$_2$. Single cycle kinetics were recorded and double referenced. The surface was regenerated using 10 mM glycine pH 1.4. Sensorgrams were fitted to a 1:1 Langmuir model and reported rate constants represent average and standard deviation of at least three independent experiments. For van't Hoff, Eyring and pH analyses the temperature or buffer system was varied, respectively.

Phylogenetic Analyses

We used 79 sequences from mammal species homologous to human SPINK1 (Table S1). Out of the 79 sequences, 62 featured lysine as P1 residue and 17 contained an arginine in the corresponding position. Multiple sequence alignments (MSA) were carried out using the MUSCLE algorithm (25). Phylogenies were estimated using PhyML 2.5.0+ within the MPI Bioinformatics Toolkit and visualized in iTOL v5 (26–28).

Structure Prediction

Apart from SPINK1 WT, for which a crystal structure was available (PDB ID: **7QE8**), SPINK1 structures were predicted using AlphaFold within ChimeraX (18, 29). For inhibitor – enzyme complex prediction, AlphaFold-Multimer was used (30). Multimer scores were above 90 in all cases, indicating high confidence complexes, which are consistent with available crystal structures. Amber forcefields for structure relaxation were used in all cases and the best model of each prediction was chosen for visualization and analysis.

References

1. Farady, C. J., and Craik, C. S. (2010) Mechanisms of macromolecular protease inhibitors. *Chembiochem.* **11**, 2341–2346

2. Krowarsch, D., Cierpicki, T., Jelen, F., and Otlewski, J. (2003) Canonical protein inhibitors of serine proteases. *Cell. Mol. Life Sci. C.* **60**, 2427–2444

3. Laskowski Jr, M., and Kato, I. (1980) Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593–626

4. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162

5. Luthy, J., Praissman, M., Finkenstadt, W., and Laskowski, M. J. (1973) Detailed mechanism of interaction of bovine -trypsin with soybean trypsin inhibitor (Kunitz). I. Stopped flow measurements. *J. Biol. Chem.* **248**, 1760–1771

6. Zakharova, E., Horvath, M. P., and Goldenberg, D. P. (2009) Structure of a serine protease poised to resynthesize a peptide bond. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11034–11039

7. Jackson, R. M., and Russell, R. B. (2000) The serine protease inhibitor canonical loop conformation: examples found in extracellular hydrolases, toxins, cytokines and viral proteins11Edited by R. Huber. *J. Mol. Biol.* **296**, 325–334

8. Bode, W., and Huber, R. (1993) Natural protein proteinase inhibitors and their interaction with proteinases. *EJB Rev.*

9. Apostoluk, W., and Otlewski, J. (1998) Variability of the canonical loop conformations in serine proteinases inhibitors and other proteins. *Proteins Struct. Funct. Bioinforma.* **32**, 459–474

10. M., L. S., Wuyuan, L., A., Q. M., Stephen, A., Izydor, A., Wojciech, A., Theresa, B., Wen, C. Y., James, C., G., J. M. N., Ikunoshin, K., Clyde, K., William, K., Tomoko, K., Tiao-Yin, L., Michio, O., Jacek, O., Soon-Jae, P., Sabiha, Q., Michael, R., Misao, T., Nicholas, W., Harry, W., Anna, W., Maciej, W., Tadeusz, W., Richard, W., Wenlei, Z., and Michael, L. (2001) Predicting the reactivity of proteins from their sequence alone: Kazal family of protein inhibitors of serine proteinases. *Proc. Natl. Acad. Sci.* **98**, 1410–1415

11. Qasim, M. A., Ganz, P. J., Saunders, C. W., Bateman, K. S., James, M. N. G., and Laskowski, M. (1997) Interscaffolding additivity. Association of P-1 variants of eglin C and of turkey ovomucoid third domain with serine proteinases. *Biochemistry.* **36**, 1598–

1607

12. Boros, E., Sebák, F., Héja, D., Szakács, D., Zboray, K., Schlosser, G., Micsonai, A., Kardos, J., Bodor, A., and Pál, G. (2019) Directed Evolution of Canonical Loops and Their Swapping between Unrelated Serine Proteinase Inhibitors Disprove the Interscaffolding Additivity Model. *J. Mol. Biol.* **431**, 557–575

13. Salameh, M. A., Soares, A. S., Navaneetham, D., Sinha, D., Walsh, P. N., and Radisky, E. S. (2010) Determinants of Affinity and Proteolytic Stability in Interactions of Kunitz Family Protease Inhibitors with Mesotrypsin. *J. Biol. Chem.* **285**, 36884–36896

14. Schulte, S., Kronthaler, U., Schmidbauer, S., Weimer, T., and Hofmann, K. (2019) Therapeutic application of kazal-type serine protease inhibitors

15. Lu, X., and Block, T. (2012) Identification of modulators of serine protease inhibitor Kazal and their use as anti-cancer and anti-viral agents

16. Nishimiya, D., Kawaguchi, Y., Kodama, S., Nasu, H., Yano, H., Yamaguchi, A., Tamura, M., and Hashimoto, R. (2019) A protein scaffold, engineered SPINK2, for generation of inhibitors with high affinity and specificity against target proteases. *Sci. Rep.* **9**, 11436

17. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350

18. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature.* **596**, 583–589

19. Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y.-H., Goto, Y., Réfrégiers, M., and Kardos, J. (2018) BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res.* **46**, W315–W322

20. Negulescu, H., Guo, Y., Garner, T. P., Goodwin, O. Y., Henderson, G., Laine, R. A., and Macnaughtan, M. A. (2015) A Kazal-Type Serine Protease Inhibitor from the Defense Gland Secretion of the Subterranean Termite Coptotermes formosanus Shiraki. *PLoS One.* **10**, e0125376

21. Grzesiak, A., Krokoszynska, I., Krowarsch, D., Buczek, O., Dadlez, M., and Otlewski, J. (2000) Inhibition of Six Serine Proteinases of the Human Coagulation System by Mutants of Bovine Pancreatic Trypsin Inhibitor*. *J. Biol. Chem.* **275**, 33346–33352

22. Pezzilli, R. (2009) Pharmacotherapy for acute pancreatitis. *Expert Opin. Pharmacother.* **10**, 2999–3014

23. Nagel, F., Palm, G. J., Geist, N., McDonnell, T. C. R., Susemihl, A., Girbardt, B., Mayerle, J., Lerch, M. M., Lammers, M., and Delcea, M. (2022) Structural and Biophysical Insights into SPINK1 Bound to Human Cationic Trypsin. *Int. J. Mol. Sci.* . 10.3390/ijms23073468

24. Susemihl, A., Nagel, F., Grabarczyk, P., Schmidt, C. A., and Delcea, M. (2021) Easy Expression and Purification of Fluorescent N-Terminal BCL11B CCHC Zinc Finger Domain. *Mol.* . 10.3390/molecules26247576

25. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797

26. Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A. N., and Alva, V. (2018) A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J. Mol. Biol.* **430**, 2237–2243

27. Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321

28. Letunic, I., and Bork, P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296

29. Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., and Ferrin, T. E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82

30. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J., and Hassabis, D. (2022) Protein complex prediction with AlphaFold-Multimer. *bioRxiv.* 10.1101/2021.10.04.463034

Data availability

All the relevant data in this study are available upon reasonable request.

Author contributions

Conceptualization, F.N., M.D.; methodology, F.N., A.S., T.E.; validation, F.N., A.S., T.E.; formal analysis, F.N.; investigation, F.N., A.S., T.E.; data curation, F.N.; writing-original draft preparation, F.N.; writing-review and editing, F.N., A.S. M.D.; visualization, F.N.; supervision, M.D.; project administration, M.D.; funding acquisition, M.D.; All authors have read and agreed to the published version of the manuscript.

Funding

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.
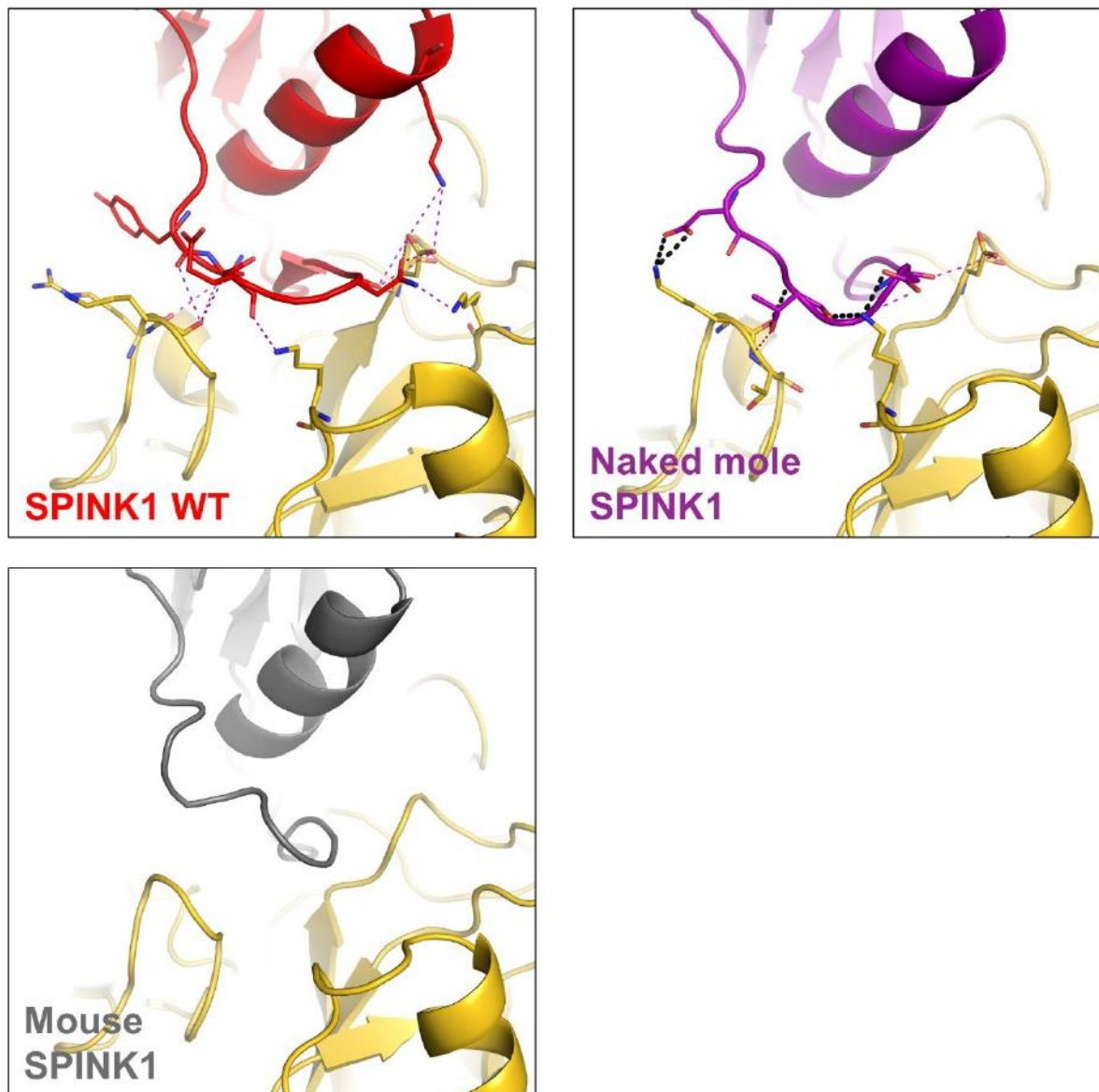
Supporting Information

Figure S1: SPINK1 complexes with TRY1. The SPINK1 WT complex is shown in red (PDB ID: **7QE8**). The naked mole complex (purple) and mouse complex (gray) were predicted by AlphaFold multimer. Apart from naked mole SPINK1, no meaningful polar contacts between SPINK1 and TRY1 were observed or predicted outside the canonical binding loop. Favorable polar contacts are highlighted with black dashes and close (<4.0 Å) but not ideal contacts are shown in purple dashes.
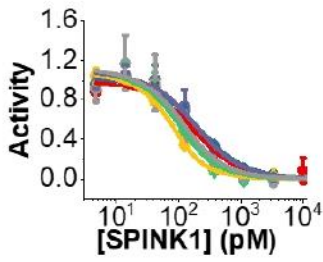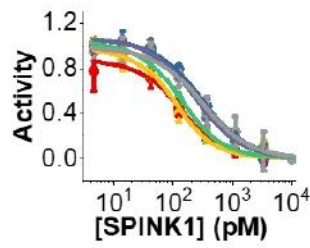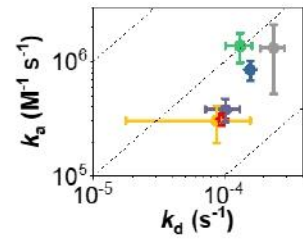
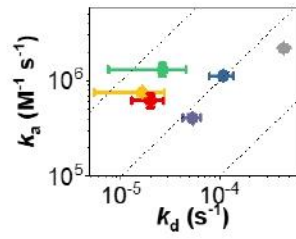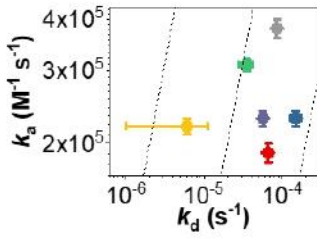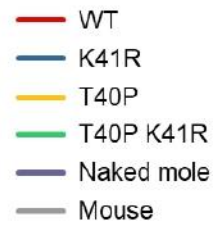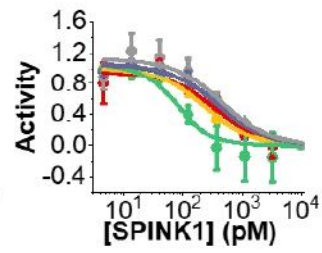**Figure S2: Interactions of SPINK1 variants with different trypsins.** Equilibrium binding data (top) and kinetic SPR analyses as k_a-k_d maps (bottom).

Table S1: Summary of binding parameters for the interactions between SPINK1 variants and trypsin.

	K_i (nM)	K_d (nM)	$t_{1/2}$ (h)
Anionic trypsin			
WT	0.04 ± 0.01	0.35 ± 0.03	2.9 ± 0.2
K41R	0.05 ± 0.008	0.65 ± 0.1	1.2 ± 0.2
T40P	0.01 ± 0.001	0.03 ± 0.02	32 ± 26
T40P K41R	0.018 ± 0.006	0.11 ± 0.02	5.5 ± 1.1
Naked mole	0.049 ± 0.013	0.25 ± 0.06	3.3 ± 0.2
Mouse	0.025 ± 0.011	0.22 ± 0.03	2.2 ± 0.1
Bovine trypsin			
WT	0.053 ± 0.014	0.03 ± 0.01	9.7 ± 3.4
K41R	0.163 ± 0.022	0.096 ± 0.021	17 ± 5
T40P	0.039 ± 0.013	0.02 ± 0.014	12 ± 8
T40P K41R	0.058 ± 0.012	0.018 ± 0.012	7 ± 5
Naked mole	0.125 ± 0.019	0.129 ± 0.02	3.6 ± 0.7
Mouse	0.172 ± 0.021	0.207 ± 0.02	4.2 ± 0.3
Porcine trypsin			
WT	0.182 ± 0.063	0.31 ± 0.06	2.0 ± 0.1
K41R	0.223 ± 0.035	0.19 ± 0.04	1.2 ± 0.07
T40P	0.113 ± 0.015	0.25 ± 0.21	2.2 ± 1.7
T40P K41R	0.017 ± 0.011	0.1 ± 0.01	1.4 ± 0.3
Naked mole	0.205 ± 0.036	0.28 ± 0.1	1.8 ± 0.5
Mouse	0.243 ± 0.091	0.34 ± 0.25	0.8 ± 0.2

Table S2: Sequences and accession numbers of SPINK1 homologues used for phylogenetic analysis

MUSCLE alignment of sequences with P1 = Lys in CLUSTAL format

```
>ISK1_HORSE
GNTYPNECMLCVENQKRQMPVLIQRSGPC--D-SL-GREAKCNNNAGGCTKIYNPVCGT
>ISK1_HUMAN
GNTYPNECVLCFENRKRQTSILIQKSGPCGAD-SL-GREAKCYNELNGCTKIYDPVCGT
>ISK1_PIG
GITYSNECVLCSENKKRQTPVLIQKSGPC--T-SP-QREATCTSEVSGCPKIYNPVCGT
>ISK1_CANLF
GITYANECLLCLENKKRQTSILVEKSGPC--NNML-QRQANCNLKVNGCNKIYNPICGS
>A0A452GEN5_HORSE
GNTYPNECMLCVENQKRQMPVLIQRSGPCRAD-SL-GREAKCNNNAGGCTKIYNPVCGT
>A0A2K6PAM9_RHIRO
GKTYPNECVLCFENQKRQTSILIRKSGPCRAD-SL-GREAKCYNELTGCTKIYDPVCGT
>H2QRQ7_PANTR
GNTYPNECVLCFENRKRQTSILIQKSGPCGAD-SL-GREAKCYNELNGCTKIYDPVCGT
>A0A6D2W914_PANTR
GNTYPNECVLCFENRKRQTSILIQKSGPCGAD-SL-GREAKCYNELNGCTKIYDPVCGT
>A0A2J8NJ67_PANTR
GNTYPNECVLCFENR-------------GAD-SL-GREAKCYNELNGCTKIYDPVCGT
>F7HE31_MACMU
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>G7MVE7_MACMU
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A2K5NJ41_CERAT
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A096NAT0_PAPAN
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A2K5ZTE8_MANLE
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>G7P8L5_MACFA
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A2K5TSP7_MACFA
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A2K6D3G8_MACNE
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A663DAL1_PONAB
GNTYPNECVLCFENRKRQTSILIQKSGPCGAD-SL-GREAKCYNELNGCTKIYDPVCGT
>H2PGZ4_PONAB
GNTYPNECVLCFENRKRQTSILIQKSGPCGAD-SL-GREAKCYNELNGCTKIYDPVCGT
>G3QZB2_GORGO
GNTYPNECVLCFENRKRQTSILIQKSGPCGAD-SL-GREAKCYNELNGCTKIYDPVCGT
>D6RIU5_HUMAN
GNTYPNECVLCFENR-------------GAD-SL-GREAKCYNELNGCTKIYDPVCGT
>A0A2J8X5Q8_PONAB
GNTYPNECVLCFENR-------------GAD-SL-GREAKCYNELNGCTKIYDPVCGT
>A0A2K5JYE4_COLAP
GKTYPNECVLCSENQKRQTSILIRKSGPCGAD-SL-EREAKCYNELTGCTKIYEPVCGT
>G1RG35_NOMLE
```

```
GNTYPNECVLCFENQKRQTSILIQKSGPCGAD-SL-GREAKCYNELNGCTKIYDPVCGT
>A0A2K6L943_RHIBE
GKTYPNECVLCFENQKRQTSILIRKSGPCRAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A0D9RI87_CHLSB
GKTYPNECVLCFENQKRQTSVLIRKSGPCGAD-SL-GREAKCYNELTGCTKIYDPVCGT
>A0A2Y9I1X1_NEOSC
GETYSNECVLCVQNKKRQIPILIQKSGPCRAG-LV-GREAKCNTETTACTKIYNPVCGT
>A0A2U3XMH3_LEPWE
GETYSNECVLCVQNKKRQIPILIQKSGPCRAG-LV-GREAKCNTETTACTKIYNPVCGT
>A0A1S2ZAJ5_ERIEU
GITYSNECMLCQENRKRQSPVLIQKSGPCGAD-SL-KREAKCNNEISGCTKIYNPVCGT
>A0A0P6IXR0_HETGA
GISYGNECTLCAENLKRQVPDLIRKSGPCAAY-S-SGRKANCDYTMDGCPKIYDPVCGT
>A0A6J3FK91_SAPAP
GVTYGNECMLCLENQRRQVPVYIKKSGVCGAE-PE-GRKPNCYNTVNGCTKIFNPVCGT
>A0A2K5PC02_CEBCA
GVTYGNECMLCLENQRRQVPVYIKKSGVCGAE-PE-GRKPNCYNTVNGCTKIFNPVCGT
>A0A2K5CSI3_AOTNA
GITYGNECMLCLENQRRQVPVSIKKSGVCGAE-PE-GRKPNCYNTVNGCTKIYNPVCGT
>F7HEY3_CALJA
GVTYGNECMLCLENQRRQVPVYIKKSGVCGAE-PE-GRKPNCYNTVNGCTKIFNPVCGT
>A0A2K6S432_SAIBB
GVTYGNECMLCLENQRRQVPVYIKKSGVCGAE-PE-GRKPNCYNTVNGCTKIFNPVCGT
>A0A5F4W7I5_CALJA
GVTYGNECMLCLENQ--------------GAE-PE-GRKPNCYNTVNGCTKIFNPVCGT
>A0A2Y9KJ82_ENHLU
GETYSNECVLCVQNKKRQTPVFIQKSGPCRAG-LE-GREAKCNTETNACTKIYDPVCGT
>A0A3P4LTF1_GULGU
GETYSNECVLCVQNK--------------RAG-LE-GREAKCNTETNACTKIYDPVCGT
>M3Y8P9_MUSPF
GETYSNECVLCVQNKKRQTPVFIKKSGPCRAG-LE-GREAKCNTETNACTKIYDPVCGT
>A0A3Q7QQL7_CALUR
GETYSNECVLCVQNKKHQIPVLIQKSGPCRAG-LV-GREAKCNTETTACTKIYNPVCGT
>A0A6J2E0M7_ZALCA
GETYSNECVLCVQNKKHQIPVLIQKSGPCRAG-LV-GREAKCNTETTACTKIYNPVCGT
>A0A384C477_URSMA
GETYSNECVLCVQNKKRQTPVLIQKSGPCRAG-LV-GREAKCNSETTACTKIYNPVCGT
>A0A3Q7WJ38_URSAR
GETYSNECVLCVQNKKRQTPVLIQKSGPCRAG-LV-GREAKCNSETTACTKIYNPVCGT
>A0A452Q7H3_URSAM
GETYSNECVLCVQNKKRQTPVLIQKSGPCRAG-LV-GREAKCNSETTACTKIYNPVCGT
>A0A2U3WYF4_ODORO
GETYSNECVLCVQNKKRQIPVLIQKSGPCRAG-LV-GREAKCNTETSACTKIYNPVCGT
>A0A1U7T0L7_TARSY
GITYSNECMLCLENQKRQIPVLIQKSGPCRAD-SP-GRKANCNNEVNGCNKIYNPVCGT
>A0A287BPH0_PIG
GITYSNECVLCSENKKRQTPVLIQKSGPCRAT-SP-QREATCTSEVSGCPKIYNPVCGT
>A0A4X1T366_PIG
GITYSNECVLCSENKKRQTPVLIQKSGPCRAT-SP-QREATCTSEVSGCPKIYNPVCGT
>A0A4X1T0Q5_PIG
GITYSNECVLCSENK--------------RAT-SP-QREATCTSEVSGCPKIYNPVCGT
>A0A091DNB8_FUKDA
```

```
GISYGNECTLCAENV-------------TAY-S-SGRKANCDYTMDGCPKIYDPVCGT
>A0A6J0A3A1_ACIJB
GKTYSNECLLCVENKKRQNPVLIKKSGPCRAD-LL-ERQAKCDSEIPGCTKIYAPVCGT
>A0A2I2U9I8_FELCA
GKTYSNECLLCVENKKRQNPVLIKKSGPCRAD-LL-ERQAKCDSEITGCTKIYAPVCGT
>A0A667FRQ2_LYNCA
GKTYSNECLLCVENKKHQNPVLIKKSGPCRAD-LL-ERQAKCDSEIPGCTKIYAPVCGT
>A0A5S6CIS3_CANLF
GITYANECLLCLENKKRQTSILVEKSGPCRANNML-QRQANCNLKVNGCNKIYNPICGS
>A0A3Q7R004_VULVU
GITYANECLLCLENKKRQTSILVEKSGPCRANNML-QRQANCNLKVNGCNKIYNPICGS
>A0A671ET95_RHIFE
GNTYPNECELCMENKKHQSPILIKKSGPCRAD-TL-EREAICRNHVNGCTKIYNPVCGT
>G1L0R7_AILME
GETYSNECVLCVQNKKRQIPVLIQKSGPCLC--IF-FCEAKCNSETTACTKIYNPVCGT
>A0A673VI12_SURSU
GKTYSNECVLCVENQKRQILVLIQKSGPCRAA-MP-GREANCNSQVTGCTKIYNPVCGT
>A0A2K6FQW4_PROCO
GHTYPNECTLCHDNKERRSPVLIQKEGACRPN-V-EGRKPVCV-DFVGCHKIYDPVCGS
>L5L1D3_PTEAL
GTTYPNECELCMENKKRQISVLIKTDGPCRAD-PM-GRKAVCTSKLSGCPKIYNPVCGT
>A0A341AU37_NEOAA
GVTYSNECVLCVENMKHKTPVLIQKSGRCRAN-FL-ERKAKCTSEVNGCPKIYNPVCGT
>A0A6I9I5X5_VICPA
GVTYSNECMLCIENTKRQVPVLIQKSGPCNAN-FL-GREAKCSNEVSGCSKIYNPVCGT
```

MUSCLE alignment of sequences with P1 = Arg in CLUSTAL format

```
>ISK1_MOUSE
GITYANECVLCFENRKRIEPVLIRKGGPCFSAKV-TGKEASCHDAVAGCPRIYDPVCGT
>ISK1_RAT
GITYPSECSLCFENRKFGTSIHIQRRGTC-PAEV-NGKTPNCPKQIMGCPRIYDPVCGT
>ISK1_BOVIN
GVTYSNECLLCMENKERQTPVLIQKSGPCGAN-IL-GREAKCTNEVNGCPRIYNPVCGT
>ISK1_SHEEP
GVTYANECLLCMENKERQTPVLIQKSGPC--N-IL-GREAKCTNEVNGCPRIYNPVCGT
>ISK1L_RAT
GKTYANECILCFENRKFGTSIRIQRRGLC-TAKV-IGKKANCPNTLVGCPRDYDPVCGT
>H0VGW8_CAVPO
GISYSNECLLCVENRKRQVPVLIKKSGPCTA--S-VERQANCDHTRLGCPRIYSPVCGT
>A0A6I9LEH1_PERMB
GITYSNECTLCSENRERQVPVLIRKYGPCTAD-FV-GKKADCNYGVTGCPRIYDPVCGM
>A0A2Y9E249_TRIMA
GNTYSNECMLCFENKKRQIQVRVLKSGPCRAN-CL-GREAVCPSDVNGCPRIYNPVCGT
>A0A4W2DEG9_BOBOX
GVTYSNECLLCMENKERQTPVLIQKSGPCGAN-IL-GREAKCTNEVNGCPRIYNPVCGT
>L8HVR5_9CETA
GVTYSNECLLCMENK-------------GAN-IL-GREAKCTNEVNGCPRIYNPVCGT
>A0A452G8R9_CAPHI
GVTYANECLLCMENKERQTPVLIQKSGPCGAN-IL-GREAKCTNEVNGCPRIYNPVCGT
>A0A6J0XCB7_ODOVR
GVTYSNECLLCMENKERQTPVLIQKSGPCGAN-IL-GREAKCTNGVNGCPRIYNPVCGT
```

```
>W5P0U0_SHEEP
GVTYANECLLCMENKERQTPVLIQKSGPCGAN-IL-GREAKCTNEVNGCPRIYNPVCGT
>M0R594_RAT
GKTYANECILCFENR--------------TAKV-IGKKANCPNTLVGCPRDYDPVCGT
>G1SFU9_RABIT
GNTYSNECILCTENQKRQIPVLIRKSGPCRAD-SL-GRMANCNNEVTGCPRNYDPVCGT
>A0A1S3FY80_DIPOR
GTTYSNECMLCDENRKRQVPVLIQKSGPCTAN-SL-GRKANCNEENTGCPRIYQPVCGT
>A0A6J2N2U9_9CHIR
GHTYSNECELCMENKRRKIPVLIKKYGPCGDDSVXSHFXAKCNHEVSGCTRNYNPICGT
```

Eigenständigkeitserklärung

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

_____

Felix Andreas Nagel

List of Publications

1.   **Nagel, F.**, Susemihl, A., Geist, N., Möhlis, K., Palm, G. J., Lammers, M., and Delcea, M. (2022) Structural Basis of the Pancreatitis-Associated Autoproteolytic Failsafe Mechanism in Human Anionic Trypsin. J. Inflamm. Res. **15**, 3633–3642

2.   **Nagel, F.**, Palm, G. J., Geist, N., McDonnell, T. C. R., Susemihl, A., Girbardt, B., Mayerle, J., Lerch, M. M., Lammers, M., and Delcea, M. (2022) Structural and Biophysical Insights into SPINK1 Bound to Human Cationic Trypsin. Int. J. Mol. Sci. . 10.3390/ijms23073468

3.   Susemihl, A., **Nagel, F.**, Grabarczyk, P., Schmidt, C. A., and Delcea, M. (2021) Easy Expression and Purification of Fluorescent N-Terminal BCL11B CCHC Zinc Finger Domain. Mol. . 10.3390/molecules26247576

4.   Kulke, M., **Nagel, F.**, Schulig, L., Geist, N., Gabor, M., Mayerle, J., Lerch, M. M., Link, A., and Delcea, M. (2021) A Hypothesized Mechanism for Chronic Pancreatitis Caused by the N34S Mutation of Serine Protease Inhibitor Kazal-Type 1 Based on Conformational Studies. J. Inflamm. Res. **14**, 2111–2119

5.   Dharmaraj, K., Dattler, D., Kahlert, H., Lendeckel, U., **Nagel, F.**, Delcea, M., and Scholz, F. (2021) The effects of the chemical environment of menaquinones in lipid monolayers on mercury electrodes on the thermodynamics and kinetics of their electrochemistry. Eur. Biophys. J. **50**, 731–743

6.   Buchholz, I., **Nagel, F.**, Klein, A., Wagh, P. R., Mahajan, U. M., Greinacher, A., Lerch, M. M., Mayerle, J., and Delcea, M. (2020) The impact of physiological stress conditions on protein structure and trypsin inhibition of serine protease inhibitor Kazal type 1 (SPINK1) and its N34S variant. Biochim. Biophys. acta. Proteins proteomics. **1868**, 140281

Curriculum vitae

Name	Felix Nagel
Sex	Male
Date of birth	11/8/1994
Marital status	Single
Nationality	German
Phone	0176 82349409
Email	felix.nagel@uni-greifswald.de
Address	Rudolf-Petershagen-Allee 25
	17489 Greifswald
	Germany

Education

2019 – expected 2022	University of Greifswald, Biochemistry, degree sought: PhD
	Thesis topic: Structural and Biophysical Insights Into Autoproteolytic and Inhibitory Mechanisms Regulating Tryptic Activity in Pancreatitis
2016 – 2019	University of Greifswald, Biochemistry, MSc (final grade **1.7**)
	Thesis topic: Kinetic and Thermodynamic Analysis of Serine Protease Interactions with the ICP associated SPINK1 N34S mutant (final grade **1.1**)
2013 – 2016	University of Tübingen, Biochemistry, BSc (final grade **2.4**)
	Thesis topic: Purification, Crystallization, and Structural Analysis of scFvs from JCPyV Neutralizing Antibodies (final grade **1.0**)
2005 – 2013	Gymnasium in der Taus, Backnang, Highschool diploma

Research Skills

Methods

- Construct design, Cloning, Protein Expression in *E. coli* and HEK293 cells
- Inclusion body refolding
- Protein purification (Affinity chromatography, Size exclusion chromatography, Ion exchange chromatography, Ultracentrifugation using density gradients)
- Colorimetric enzymatic assays
- Circular dichroism spectroscopy
- Fluorescence spectroscopy
- Surface plasmon resonance
- Isothermal titration calorimetry
- Crystallography
- X-ray diffraction
- cryoEM (basic knowledge)
- Structure refinement and phasing (Molecular replacement)

Software

- MS Office
- Linux (basics)
- Python (basics)
- Origin
- Snapgene
- PyMOL
- Coot
- Phenix
- Refmac
- CCP4
- XDS

Languages

- German (native speaker)
- English (fluent, FCE level B)

Publications and Prizes

Prizes   Prize for exceptional work during the Master thesis awarded by the "Gesellschaft für Biochemie und Molekularbiologie" (GBM-Masterpreis **2019**)

Publications   **Nagel, F.**; Susemihl, A.; Geist, N.; Möhlis, K.; Palm, G.J.; Lammers, M.; Delcea, M. Molecular Basis of the Pancreatitis-associated Autoproteolytic Failsafe Mechanism in Human Anionic Trypsin. *J. Inflamm. Res.* **2022**,15,3633-3642.

Nagel, F.; Palm, G.J.; Geist, N.; McDonnell, T.C.R.; Susemihl, A.; Girbardt, B.; Mayerle, J.; Lerch, M.M.; Lammers, M.; Delcea, M. Structural and Biophysical Insights into SPINK1 Bound to Human Cationic Trypsin. *Int. J. Mol. Sci.* **2022**, *23*.

Susemihl, A.; **Nagel, F.**; Grabarczyk, P.; Schmidt, C.A.; Delcea, M. Easy Expression and Purification of Fluorescent N-Terminal BCL11B CCHC Zinc Finger Domain. *Mol.* **2021**, *26*.

Kulke, M.; **Nagel, F**.; Schulig, L.; Geist, N.; Gabor, M.; Mayerle, J.; Lerch, M.M.; Link, A.; Delcea, M. A Hypothesized Mechanism for Chronic Pancreatitis Caused by the N34S Mutation of Serine Protease Inhibitor Kazal-Type 1 Based on Conformational Studies. *J. Inflamm. Res.* **2021**, *14*, 2111.

Dharmaraj, K.; Dattler, D.; Kahlert, H.; Lendeckel, U.; **Nagel, F.**; Delcea, M.; Scholz, F. The effects of the chemical environment of menaquinones in lipid monolayers on mercury electrodes on the thermodynamics and kinetics of their electrochemistry. Eur. Biophys. J. **2021**, 50, 731–743,

Buchholz, I.; **Nagel, F.**; Klein, A.; Wagh, P.R.; Mahajan, U.M.; Greinacher, A.; Lerch, M.M.; Mayerle, J.; Delcea, M. The impact of physiological stress conditions on protein structure and trypsin inhibition of serine protease inhibitor Kazal type 1 (SPINK1) and its N34S variant. Biochim. Biophys. acta. Proteins proteomics **2020**, 1868, 140281.

In Preparation	**Nagel, F.**; Susemihl, A.; Eulberg, T.F.; Delcea, M. Identification of Kazal-inhibitor scaffolds with identical canonical binding loops and their effects on binding properties. **Submitted**

Geist, N.; **Nagel, F.**; Delcea, M. Molecular Interplay of ADAMTS13-MDTCS and von Willebrand Factor-A2: Deepened Insights from Extensive Atomistic Simulations

McDonnell, C.R.T.; **Nagel, F.;** Spiteri, V.; Pericleous, C.; Farinha, F.; Munoz, M.C.; Griffin, M.; Nicolaides, A.; Bakshi, J.; Perkins, J.S.; Delcea, M.; Dalby, P.; Giles, I.; Rahman, A. Antibodies Cross-Reacting to Both Domain I and Complement Factor B predict the development of atherosclerosis over a decade in advance in SLE patients.

Harprecht, C.; Ströh, L.J.; Senn, L.; O'Hara, B.A.; **Nagel, F.**; Kuhn, J.; Stierhof, Y.D.; Atwood, W.J.; Combaluzier, B.; Stehle, T.

Other experiences

2018	Research stay at the University College London.
	Focus: Refolding of the beta2 glycoprotein1 domain 1 (2 weeks)
2017	Student assistant at the Institut für Biochemie, Greifswald.
	Focus: Protein purification, enzymatic assay development (1 month)
2012	Practical at the Institut für Biochemie und molekulare Biologie, Freiburg.
	Focus: Structural and functional analysis of medically relevant membrane proteins (2 weeks)

_____
Felix Andreas Nagel

Danksagung

Vielen Dank an…

… Mihaela Delcea für deine gute und stets optimistische Betreuung meiner Promotion, die Unterstützung eigene Ideen zu verfolgen und die Hilfe mich bei dir wissenschaftlich und persönlich weiterzuentwickeln. Egal unter welchen Umständen, du warst immer für mich da, wenn ich dich gebraucht habe.

… Michael Lammers für die gute Zusammenarbeit und die schnelle und unkomplizierte Kooperation in mittlerweile einigen Projekten. Danke auch für die wissenschaftlichen Ratschläge und sehr hilfreichen Diskussionen.

… Anne „Drift" Susemihl für das Lesen dieser Arbeit. Dafür, dass ich mich immer auf dich verlassen konnte, das abwechselnde PI und TA spielen und deine Art, dank der ich immer wieder Spaß an, oder wenigstens während, der Arbeit hatte. Wenn die Arbeit nämlich da ist, muss sie auch gemacht werden und mit wir meine ich dich. Ohne dich wären die letzten Jahre sehr trostlos geworden. Danke für alles!

… Una Janke für das gemeinsame Durchstehen der letzten Jahre. Durch dich wurden auch die schwersten Zeiten erträglich. Danke aber auch für die guten Zeiten im Labor, den Spaß den wir gemeinsam hatten, den morgendlichen Kaffee, das abendliche Bier und alles andere.

… Alexandra Mitlehner für die schöne Zeit zusammen und die vielen Wochenenden während der Corona Pandemie auch wenn es geregnet hat wie schei** und da Stacheldraht drum ist.

… Norman Geist für die MD Simulationen, hilfreichen Diskussionen über Proteinstrukturen und jedes noch so kleine und nervige IT Problem, das du behoben hast.

… Kevin Möhlis und Tobias Eulberg, für die Zusammenarbeit am SPINK1 Projekt und die spaßige Zeit mit euch.

… Anke Balzer für die Hilfe jeglicher Art und für jede noch so anstrengende Auseinandersetzung mit der Beschaffung.

… den restlichen AK Delcea für die tolle Arbeitsatmosphäre und das gute Zusammenarbeiten.

… meine Familie für die Unterstützung in den letzten 27 Jahren und das Daumendrücken aus der Ferne. Danke vor allem auch an Fabi für das Lesen dieser Arbeit.