1. INTRODUCTION

Functional redundancy describes the situation when two or more genes are responsible for the same biochemical function, resulting in no or minor biological effects if the activity of one gene is lost. In most cases redundancy is a result of gene duplications and it is a widespread phenomenon in higher organisms (Nowak *et al.*, 1997). While it improves the robustness of essential biological reactions it is a major drawback during genetic analysis of knock-out (KO) mutants, as single gene KOs will show no phenotypic alterations. In order to assess this, the current project focussed on a biochemical approach to control functional redundancy of proteolytic enzymes on the enzyme level. To establish such an approach it is necessary to understand how proteolytic activity is regulated.

1.1 Regulation of proteolytic activity

Proteolytic enzymes deal with a plethora of tasks, ranging from general protein degradation to very specific regulatory processes including the activation of zymogens, cleavage of signal peptides or the activation of peptide hormones. These processes are involved in the regulation of various developmental processes, as well as cell death and responses to wounding or pathogenic threats. Peptidases are found in all organisms, constituting 2-4% of all encoded gene products (Farady & Craik, 2010), and they are believed to have evolved over time from general protein degrading enzymes to regulators of increasing specificity (Neurath, 1984). Peptidases cleaving at internal and terminal cleavage sites are called endopeptidases and exopeptidases, respectively. The latter are subdivided in carboxy- and aminopeptidases. Peptidases are further separated based on their catalytic mechanisms into the six different classes of serine, cysteine, glutamic, threonine, aspartate and metallo-proteases. The major difference between these classes is the nature of their catalytic residues, which are involved in the nucleophilic attack on the substrate peptide bond. While serine, cysteine and threonine proteases use Ser, Cys and Thr as nucleophile (Polgár, 2013a,b; Rawlings & Barrett, 2013), metallo-, aspartic and glutamic proteases use water as a nucleophile (Auld, 2013; Wlodawer *et al.*, 2013), and are activated by a metal ion or Asp and Glu, respectively. While the human genome codes for 612 proteases, in Arabidopsis an even higher number of 826 proteases is predicted, including serine proteases as their largest class (Van Der Hoorn, 2008; Farady & Craik, 2010). Considering this high variety of proteases and the fact that the process of proteolysis is essentially irreversible, not only their

importance for various processes is apparent but also a tight regulation of proteolytic activity is essential. Therefore, proteases are tightly controlled at all stages of their lifetime. Besides transcriptional regulation, most proteases are expressed as inactive zymogens, which need to pass through an activation process. This helps to control the respective proteases to be only active at their designated location (Demidyuk *et al.*, 2010). Several proteases are also dependent on cofactors like calcium or zinc (Veltman *et al.*, 1998; Eijsink *et al.*, 2011; Zimmermann *et al.*, 2016). Lack of these cofactors can lead to reduced stability and activity. Last but not least a broad variety of peptidase inhibitors (PIs) are involved in all organisms in the regulation of proteolysis. Even though the inhibition of a protease by a protein may appear as a paradox, proteinaceous PIs comprise the largest group of naturally occurring PIs (Otlewski *et al.*, 2005).

The MEROPS database of proteases and inhibitors currently divides PIs in 39 clans based on tertiary structure, which are further subdivided in 79 families based on protein sequence (Rawlings *et al.*, 2016). Comparing plain gene numbers, proteases are in approximately five fold excess to PIs (Farady & Craik, 2010). In agreement with that, many PIs are rather promiscuous in their choice of proteases, based on the homology between protease active sites (Farady & Craik, 2010).

Pioneering results concerning the identification of the first PI appeared as early as 1936 (Kunitz & Northrop, 1936). The Kunitz's soybean Trypsin Inhibitor (KTI), which is highly expressed in soybean seeds, was the first plant PI isolated, characterized and crystallized. (Kunitz, 1945, 1947; Kunitz & McDonald, 1946). The knowledge of its crystal structure per se and in complex with porcine trypsin (Sweet *et al.*, 1974), helped to build the model of the "standard mechanism" of proteinase/inhibitor interaction of serine proteases, which is the most common mechanism of competitive PIs (Laskowski & Kato, 1980).

In the following, binding mechanisms of protease inhibitors will be described using the Schechter and Berger nomenclature (Schechter & Berger, 1967). Substrate/inhibitor residues down- and upstream of the cleavage site are called prime (P') and non-prime (P) residues, and bind in the respective subsites (S'/S) of the protease active site (Fig. 1.1).

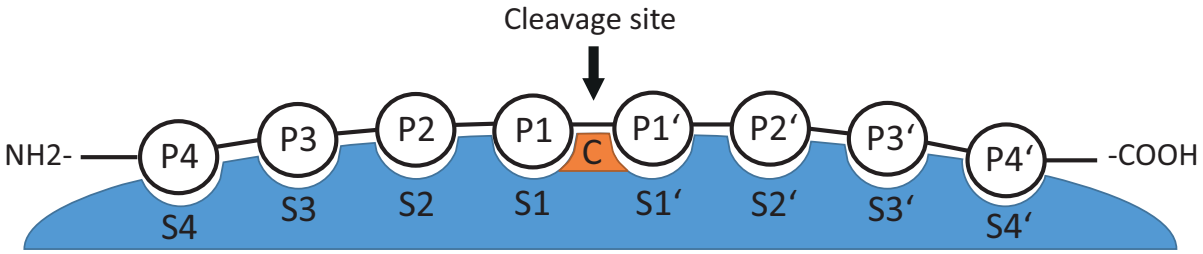

Figure 1.1. Schechter and Berger nomenclature of a protease/substrate complex.

The active site consists of subsites upstream/non-prime (P4-P1) and subsites downstream/prime (P1'-P4') of the cleavage site. C: catalytic center; enzyme in blue; substrate in white. Cartoon modified from Song *et al.*, (2011).

1.2 Competitive inhibitors

Most PIs bind their respective enzymes in a competitive manner in the active center. For serine proteases this has been studied most intensively and was named "standard", or the "Laskowksi" mechanism of protease inhibition (Laskowski & Kato, 1980). It involves a reactive loop also called canonical loop, which is surrounded in most cases by disulfide bridges. The reactive loop specifically binds to the active site of the protease mimicking the substrate, resulting in a substrate-like, yet much slower cleavage (Birk, 2003). Due to the disulfide bridges surrounding the reactive loop, the inhibitor does not change its secondary or tertiary structure upon cleavage and stays attached to the active center of the protease (Fig. 1.2). Several PI families follow the Laskowksi mechanism, including Kazal, Kunitz and Bowman-Birk PIs (Farady & Craik, 2010). PIs of other protease families follow different approaches to inhibit proteases competitively. For example, cysteine PIs of the cystatin family and metallo PIs of the human Tissue Inhibitors of Metalloproteases (TIMPs) family do not follow the "standard mechanism" of protease inhibition. While they still interact with the active site of the protease they avoid a cleavage and therefore do not bind in a substrate-like fashion (Fig. 1.2). Cystatins dodge the cleavage by two separate binding events. While the N terminus of cystatins binds to the non-prime S3-S1 pockets in a substrate-like manner, the prime-side residues of the active site are occupied by two hairpin loops. Therefore, both sides of the active side are blocked without the cystatin getting into contact with the active side residues of the enzyme (Bode & Huber, 2000; Farady & Craik, 2010).

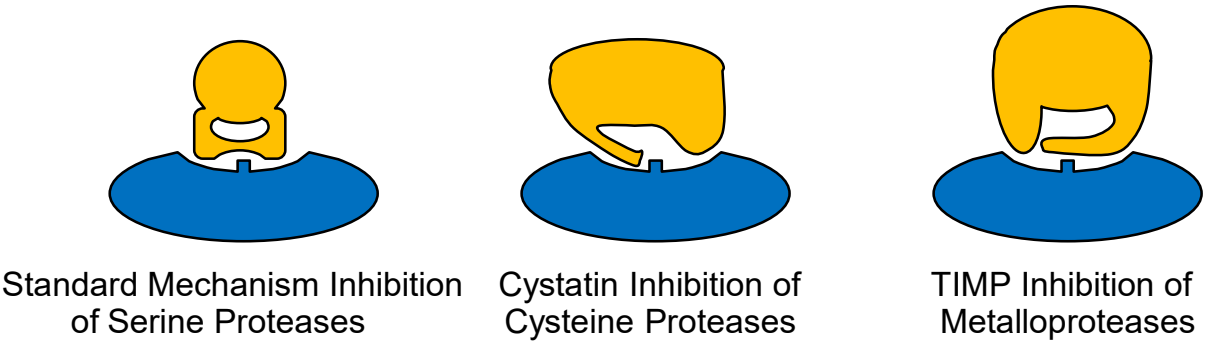

Figure 1.2. Competitive active site inhibition of proteases.

Inhibitors bind proteases in a substrate-like manner ("Standard mechanism") or not (Cystatin / TIMP). Inhibitors in yellow, proteases in blue, with the active site indicated as cavity on top. Modified from Farady & Craik, (2010)

TIMPs follow a similar approach. Like cystatins they bind the protease by two separate binding events. But while cystatins avoid the catalytic residues, the N terminus of TIMPs binds to the S1-S3' pockets spanning the active site, chelating the catalytic Zn^{2+} ion by excluding water molecules, necessary as nucleophiles, from the active site. In parallel to that a second loop binds to the non-prime pockets S3-S2 and the N terminus of the metalloprotease (Bode & Huber, 2000; Brew *et al.*, 2000; Farady & Craik, 2010). Therefore, while "standard mechanism" inhibitors bind their protease in a substrate-like fashion, cystatins avoid the active site residues and TIMPs chelate the active site metal ion, blocking the cleavage process. The described competitive inhibition mechanisms appear in some cases not only in single, but also in multi-domain inhibitors. They range from dual inhibitors like in Bowman-Birk PIs, to six, eight or even 15 domains as in circular potato peptidase inhibitor 2, multicystatins or human Kazal inhibitor SPINK5, respectively. Further, these PI domains do not necessarily need to inhibit the same class of proteases (Rawlings, 2010; Grosse-Holz & van der Hoorn, 2016).

1.3 Competitive inhibitors with exosite binding

Besides binding proteases strictly competitively, several inhibitors interact with proteases also outside of the active site. Exosite binding can influence both affinity and specificity. An example for improved affinity due to exosite binding is Rhodniin, a Kazal inhibitor of the blood-sucking assassin bug *Rhodnius prolixus*, which inhibits thrombin with sub-picomolar Ki and, therefore, efficiently avoids blood clotting (van de Locht *et al.*, 1995). In contrast, the exosite interaction of Ecotin leads not only to a higher affinity but also to a broader specificity. Ecotin

is a dimeric protein that forms a heterotetrameric complex with chymotrypsin-fold proteases (Fig. 1.3). This involves three protein interaction interfaces: the dimerization interface, the "standard mechanism" primary interaction in the active site, and a secondary interaction between the partner ecotin subunit and an exosite of the protease (Yang *et al.*, 1998). While the primary interaction site is optimized for chymotrypsin, the secondary interaction site acts compensatory, leading also to a very high affinity against trypsin, which is practically not inhibited by a monomeric Ecotin mutant (Eggers *et al.*, 2001). Another variation of exosite-supported protease inhibition is the well-characterized interaction of inhibitors of the I9 family with subtilisin-like serine proteases (subtilases, SBTs). With the exception of fungal I9 inhibitors (Maier *et al.*, 1979; Dohmae *et al.*, 1995) most PIs of this family can be found as precursor propeptides (PPs) of SBTs (Kantyka *et al.*, 2010; Santamaría *et al.*, 2014). Prior to their function as autoinhibitors, PPs act as intramolecular chaperones, (Ikemura *et al.*, 1987; Shinde & Inouye, 1994; Demidyuk *et al.*, 2010; Dillon *et al.*, 2012; Meyer *et al.*, 2016) to facilitate the correct folding of the pro-subtilase, followed by an autocatalytic cleavage of the PP which then switches from its chaperoning function to inhibition (Li *et al.*, 1995; Huang *et al.*, 1997; Fugere *et al.*, 2002; Nakagawa *et al.*, 2010). The crystal structure of bacterial subtilisin A and plant cucumisin PP/protease complexes show an exosite binding of the β-sheets of the PP with two parallel surface helices of the respective protease, while the C terminus of the PP stays attached to the nonprime subsites of the catalytic center in a product-like manner (Fig. 1.3; Jain *et al.*, 1998; Sotokawauchi *et al.*, 2017). This auto-inhibitory function implies a specific separation of the PP/protease complex. In case of the mammalian prohormone convertase furin as well as in tomato SBT3 this separation happens in a pH-dependent manner in the trans golgi network (Anderson *et al.*, 2002; Meyer *et al.*, 2016).

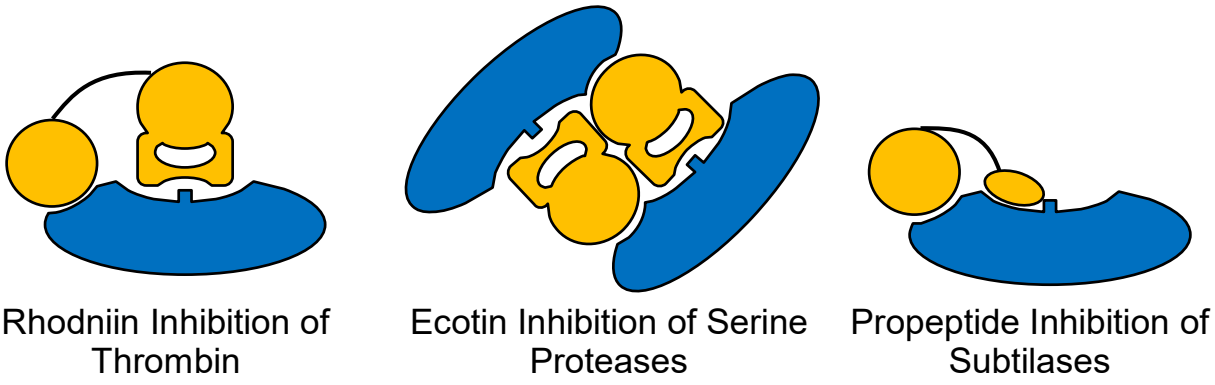

Figure 1.3. Competitive inhibitors with exosite binding.
Inhibitors in yellow, proteases in blue. Modified from Farady & Craik, (2010)

1.4 Suicide inhibitors

Most PIs interact with their target protease in a reversible non-covalent manner. In contrast, protein families I4 (serpins) and I39 (α2-macroglobulin) use irreversible trapping reactions to inhibit proteases. These trapping mechanisms rely on conformational changes upon cleavage of an internal peptide bond in the inhibitor's reactive loop, implying that they work on endopeptidases only (Rawlings, 2010).

Alpha2-macroglobulins are large glycoproteins that have a cage-like structure (Feldman *et al.*, 1985). They act as homodimers or homotetramers where each subunit presents an extended bait loop for cleavage by proteases. Upon cleavage of a bait loop an active thioester is exposed, which is readily hydrolysed by weak nucleophiles like primary amines, reductants or water, and consequently can covalently bind to the protease. Thioester hydrolysis results in a conformation change leading to a closure of the cage-like structure, trapping the active protease inside (Fig. 1.4; Rehman *et al.*, 2013). The trapped protease is not directly inhibited but rather sterically blocked to the access of bigger substrates, while still cleaving smaller peptides entering the cage (Sottrup-Jensen, 1989). The trapped protease/inhibitor complex in animals shows a half-life of several minutes only, as it is rapidly cleared from circulation. (Rehman *et al.*, 2013). In plants α2-macroglobulin genes have been annotated so far only in a few species like cucumber (*Cucumis sativus*), alpine strawberry (*Fragaria vesca*), the alga *Micromonas sp.* RCC299 and black cottonwood (*Populus trichocarpa;* Santamaría *et al.*, 2014), and their physiological function is yet unknown (Grosse-Holz & van der Hoorn, 2016).

In contrast to α2-macroglobulins, members of the irreversible PI family of serpins (I4) are found in all land plants (Santamaría *et al.*, 2014). Like macroglobulins, serpins are relatively large proteinaceous inhibitors ranging from 340 to 440 amino acids (Lampl *et al.*, 2013). While macroglobulins are live "mouse" traps for all kinds of endoproteases (Laskowski & Kato, 1980), serpins are dependent on the formation of the covalent acyl-enzyme intermediate such as in Ser and Cys proteases and kill their "prey". Cleavage of the serpin reactive loop by a protease results in a rapid conformational change. In case of α1-antitrypsin/trypsin interaction the non-prime side of the reactive loop, with the protease bound in the covalent acyl-enzyme intermediate state, moves 70 Å to the opposite pole of serpin, clashing it against the inhibitor (Fig. 1.4). This harsh reaction results in a 37 % reduction of the protease structure, and in an irreversible disruption of its active site by plucking the catalytic serine 6 Å away from the catalytic histidine (Huntington *et al.*, 2000). This trapping mechanism allows for a structural

separation of inhibitory activity from protease specificity (Grosse-Holz & van der Hoorn, 2016).

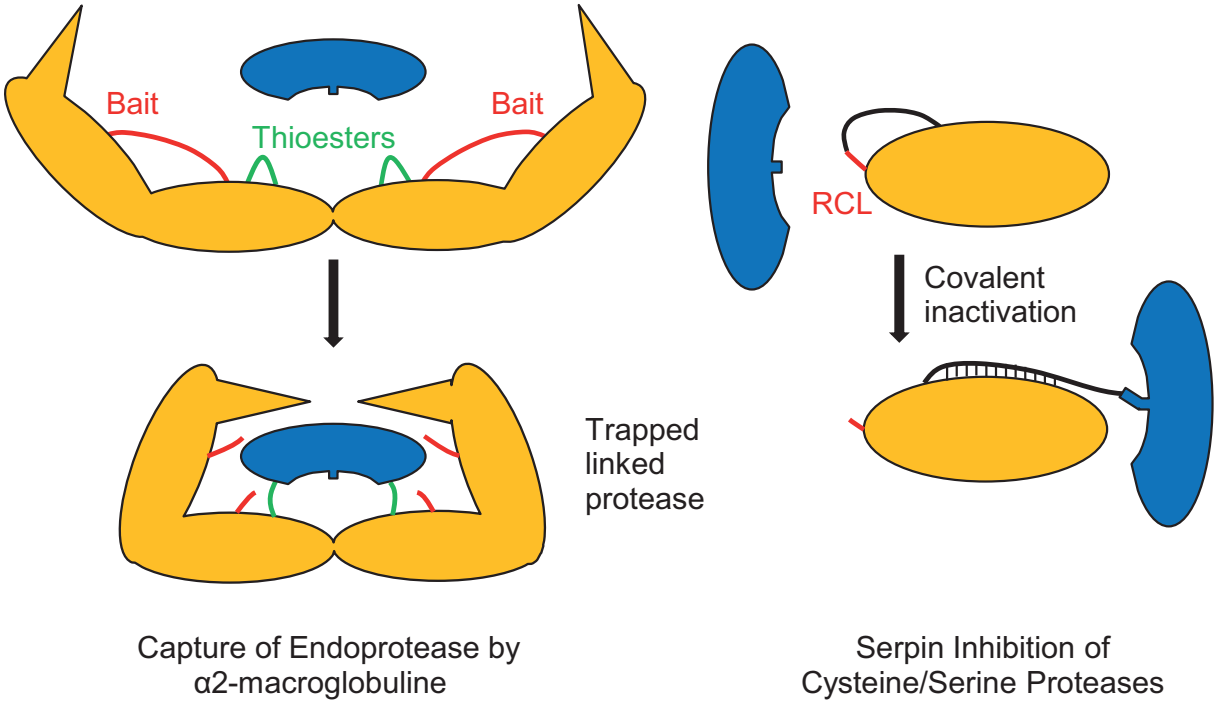

Figure 1.4. Suicide inhibitors of endopeptidases.

Reactive center loops RCL and bait loop in red, Thioesters in green, Inhibitors in yellow, proteases in blue. Modified from Farady & Craik, (2010); Grosse-Holz & van der Hoorn, (2016)

1.5 Physiological roles of proteinase inhibitors in plants

Besides the well-known fact that PIs are involved in the regulation of proteolytic activity, their precise targets and physiological roles in plants remain largely unknown. In general three different roles are discussed, that are not necessarily mutually exclusive. These functions include the regulation of endogenous proteases during development and in response to biotic and abiotic stress, the inhibition of exogenous herbivore digestive and pathogen effector proteases, and due to their stability and high abundance in tubers and seeds, a role as storage proteins (Hartl *et al.*, 2011). Regarding the latter role, the aforementioned Soybean Kunitz Trypsin Inhibitor (KTI) is, with up to 13mg/g fresh weight, highly abundant and responsible for most of the trypsin inhibitory activity in soybean seeds (Freed & Ryan, 1978). Likewise, serine protease inhibitors of the Bowman-Birk and the trypsin/α-amylase family are constitutively expressed in leguminous seeds and cereal grains, respectively (McManus *et al.*, 2000). The PI content of potato tubers can even comprise up to 50% of soluble protein, with

serine protease inhibitors of the Kunitz family and potato PI-1 and PI-2 being the most abundant (Pouvreau *et al.*, 2001, 2003). These serine PIs are thought to play dual roles as passive protectant of storage tissues against pathogens and insects and as cysteine-rich seed-storage proteins (McManus *et al.*, 2000).

In some cases it is even difficult to distinguish between seed storage proteins and PIs. For example, the major storage albumin from *Theobroma cacao* seeds comprising 25-30 % of total seed protein is homologous with Kunitz protease inhibitors (Spencer & Hodge, 1991).

1.6 Proteinase inhibitors in development and regulation of programmed cell death

While the PIs that also act as storage proteins are inhibitors of serine proteinases, cysteine proteinases are responsible for the mobilisation of storage proteins (Grudkowska & Zagdańska, 2004; Martínez *et al.*, 2012; Díaz & Martinez, 2013). In non-tuberized potato stolons protease activity is high and can be inhibited almost completely *in vitro* by the eight-headed Potato Multicystatin (PMC). During potato tuber development protease activity declines, which is correlated with an increased expression and a 230-fold higher level of PMC in the soluble protein fraction, and the accumulation of potato seed-storage proteins patatin and potato PI-1 and PI-2. This suggests an important role for PMC in the regulation of seed storage protein accumulation (Weeda *et al.*, 2009). Conversely, during aging of potato tubers and sprouting, PMC expression is downregulated and protein levels are progressively declining, which correlates with an increased cysteine protease activity and mobilisation of protein reserves from patatin and PI-1 and 2 (Kumar *et al.*, 1999; Weeda *et al.*, 2010). While all results concerning PMC are conclusive, direct genetic evidence of PMC being the regulator of storage protein accumulation and mobilisation is still missing.

Similar to PMC, chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*) phytocystatin 1 (BrCYS1) expression is high in dry seeds and declines during the first five days of germination (Hong *et al.*, 2007). Its homolog in Arabidopsis, phytocystatin 6 (AtCYS6) is highly expressed in seeds, seedlings and flowers. Its expression is diminished already after one day, while expression stays high upon treatment with abscisic acid (ABA). Constitutive overexpression of either cystatins leads to retarded germination, while it is promoted in AtCYS6 knockout mutants (Hwang *et al.*, 2009). The ABA dependent regulation of AtCYS6 and germination retarding function of both

cystatins suggests a similar regulating role for BrCYS1 and AtCYS6 in seeds during storage protein mobilisation as PMC in potato tubers.

Serine Protease Inhibitor 1 (AtSerpin1) is involved in the regulation of hypocotyl cell elongation (Ghorbani *et al.*, 2016). It controls the activity of SBT6.1 in processing the propeptide of Golven 1 (GLV1/ RGF/ CLEL6). GLV1 promotes hypocotyl cell elongation and GLV1-overexpressing (OE) plants show a curvy root phenotype. GLV1 gain-of-function is lost in *sbt6.1* and *sbt6.2* KO plants as well as in plants overexpressing AtSerpin1. Further, AtSerpin1 inhibits SBT6.1 activity *in vitro*. These observations support a model in which GLV1 peptide-promoted hypocotyl cell elongation depends on SBT6.1/6.2 processing, which is itself controlled by AtSerpin1. (Ghorbani et al., 2016).

Interestingly, AtSerpin1 is not, as its name implies just a serine protease inhibitor. Besides SBT6.1 (Ghorbani *et al.*, 2016) and trypsin (Vercammen *et al.*, 2006), also metacaspases 4 and 9 (Vercammen *et al.*, 2006) and RD21 (Lampl *et al.*, 2010) are inhibited by AtSerpin1 *in vitro*, which belong to the caspase (C14) and papain (C1) family of cysteine peptidases, respectively,

Plant metacaspases show distant relation with caspases (Uren *et al.*, 2000) and are upregulated upon cell death (Hoeberichts *et al.*, 2003). In contrast to animal caspases involved in apoptosis, plant metacaspases do not show caspase-like Asp specificity and rather cleave substrates after Lys and Arg residues (Vercammen *et al.*, 2004; Watanabe & Lam, 2005). From the nine metacaspases in Arabidopsis, AtMC1 and AtMC2 antagonistically regulate hypersensitive cell death. AtMC4 mediates programmed cell death (PCD) activation by the fungal toxin fuminosin B1, and abiotic stress and AtMC8 is required for UV-C stress induced PCD (Tsiatsiani *et al.*, 2011). AtMC9 is the only metacaspase, which needs acidic conditions to become proteolytically active and participates, in concert with the Xylem Cysteine Proteases XCP1 and XCP2, in the cellular autolysis of the xylem (Salvesen *et al.*, 2016). Further, AtMC9 activates the peptide hormone Grim Reaper by release of an 11 amino acid peptide, which is sufficient to induce Reactive Oxygen Species (ROS)-dependent cell death by binding to Pollen-Specific Receptor-like Kinase 5 (PRK5; Wrzaczek *et al.*, 2015). Recently, AtMC1 was described to bind to AtSerpin1 in a non-canonical fashion (*i.e.* without cleavage of the reactive center loop) and to be inhibited in autoprocessing, resulting in reduced AtMC1-regulated PCD *in vivo* (Lema Asqui *et al.*, 2017).

Besides SBT6.1, AtMC1, AtMC4 and AtMC9, AtSerpin1 also inhibits the endoplasmic reticulum (ER) and vacuole localized pro-death cysteine protease Responsive-to-

Desiccation-21 (RD21). RD21 is a positive regulator of PCD upon drought and oxidative stress and in response to cell death elicitors of nectrotrophic pathogens (Lampl *et al.*, 2013; Koh *et al.*, 2016). Drought stress leads to the accumulation of ROS, and further to a disintegration of vacuolar membranes. Necrotrophic fungi, like *Sclerotinia sclerotiorum* and *Botrytis cinerea,* secrete oxalic acid to stimulate changes in vacuolar permeability. Vacuole rupture then leads to a release of RD21 and other vacuolar proteases, like Vacuolar Processing Enzymes (VPE) to the cytoplasm, resulting in autolysis of the plant cell. AtSerpin1 OE and *rd21* KO plants show reduced PCD upon drought and oxidative stress, as well as reduced growth of the necrotrophic fungi *S. sclerotiorum* and *B. cinerea*. This suggests cytoplasmic AtSerpin1 acting as a set-point switch by regulating the activity of vacuolar RD21 leaking to the cytoplasm (Lampl *et al.*, 2013). According to the MEROPS database, RD21 prefers Arg in P1 (Rawlings *et al.*, 2016). Therefore, AtSerpin1 inhibits serine and cysteine proteases as long as these show a cleavage preference after basic residues (Fluhr *et al.*, 2012).

AtSerpin1 is not the only PI described to regulate RD21. Kunitz-type proteinase inhibitor AtWSCP contributes to cell death regulation in the female reproductive tract, by binding to and inhibiting RD21. At*wscp* KO mutant flowers exhibit precocious cell death in the transmitting tract and unnatural death of septum epidermis cells, while ectopic expression of AtWSCP reverses both effects (Boex-Fontvieille et al., 2015). Several other examples show proteinase inhibitors as negative regulators of endogenous proteases involved in PCD such as, soybean CystatinN2 (GmCYSN2). Ectopic expression of GmCYSN2 but not of Bowman/Birk or Kunitz inhibitors leads to reduced hypersensitive response (HR) triggered by oxidative stress or avirulent *Pseudomonas syringae* infection in soybean tissue culture (Solomon et al., 1999). Further, AtCysa and Cysb are induced upon salt, cold and oxidative stress and overexpression of either results in enhanced tolerance against various stresses in yeast and Arabidopsis (Zhang *et al.*, 2008).

Arabidopsis Kunitz serine protease inhibitor AtKTI1 is induced by *Erwinia* culture filtrate, salicylic acid (SA), ROS and wounding. RNAi-silenced lines show increased lesion formation after fumonisin B1 treatment and in imcompatible interaction with *Pseudomonas syringae* avrB. Additionally, silenced plants show spontaneous lesions at early developmental stages suggesting a defect in HR regulation (Li *et al.*, 2008). Therefore, PIs are main players in the control of programmed cell death.

Last but not least, an interesting dual role has been proposed for Arabidopsis Serpin 4 and 5 (SRP4/5). Expression of both genes is induced upon ultra-violet (UV) irradiation and