Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant development in *Arabidopsis thaliana*

Sachin Rustgi[a,b,1], Edouard Boex-Fontvieille[c,d], Christiane Reinbothe[c], Diter von Wettstein[b,1], and Steffen Reinbothe[c,1]

[a]Department of Plant and Environmental Sciences, Clemson University Pee Dee Research and Education Center, Florence, SC 29506; [b]Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164; [c]Laboratoire de Génétique Moléculaire des Plantes and Biologie Environnementale et Systémique, Université Grenoble Alpes, 38041 Grenoble cedex 9, France; and [d]Laboratoire de Biotechnologies Végétales appliquées aux Plantes Aromatiques et Médicinales, Formations de Recherche en Évolution (FRE) CNRS 3727, Université Jean Monnet, 42100 Saint-Étienne, France

Contributed by Diter von Wettstein, January 4, 2017 (sent for review November 14, 2016; reviewed by Manoj Prasad and Daichang Yang)

Proteolytic enzymes (proteases) participate in a vast range of physiological processes, ranging from nutrient digestion to blood coagulation, thrombosis, and beyond. In plants, proteases are implicated in host recognition and pathogen infection, induced defense (immunity), and the deterrence of insect pests. Because proteases irreversibly cleave peptide bonds of protein substrates, their activity must be tightly controlled in time and space. Here, we report an example of how nature evolved alternative mechanisms to fine-tune the activity of a cysteine protease dubbed RD21 (RESPONSIVE TO DESICCATION-21). One mechanism in the model plant *Arabidopsis thaliana* studied here comprises irreversible inhibition of RD21's activity by Serpin1, whereas the other mechanism is a result of the reversible inhibition of RD21 activity by a Kunitz protease inhibitor named water-soluble chlorophyll-binding protein (WSCP). Activity profiling, complex isolation, and homology modeling data revealed unique interactions of RD21 with Serpin1 and WSCP, respectively. Expression studies identified only partial overlaps in Serpin1 and WSCP accumulation that explain how RD21 contributes to the innate immunity of mature plants and arthropod deterrence of seedlings undergoing skotomorphogenesis and greening.

protease | protease inhibitor | WSCP | Serpin1 | RD21

Proteases selectively catalyze the hydrolysis of peptide bonds and, based on structural and evolutionary criteria, are grouped into families and clans in the MEROPS database (1). To date, 70 families belonging to 12 different clans were identified for cysteine proteases. Family C1 (cysteine protease 1) contains two subfamilies designated C1A and C1B. C1A family members bear a signal peptide for secretion out of the cell and contain disulfide bridges. C1B family members are localized in the cytoplasm and do not contain disulfide bonds (1). In plants, only C1A subfamily members have been identified thus far (2).

Papain-like cysteine proteases (PLCPs) belong to the family C1 of clan CA (3). In *Arabidopsis thaliana*, there are ∼30 genes encoding PLCPs that are grouped into eight subfamilies in phylogenetic comparisons (3). PLCPs exhibit the typical papain-like fold described by Drenth et al. (4), comprising an α-helix-rich domain and a β-barrel-like domain separating the substrate-binding pocket. PLCPs are unusually small proteins of ∼23–30 kDa. They cleave peptide bonds of protein substrates, using a catalytic cysteine residue as nucleophile. Interestingly, PLCPs are produced with an NH_2-terminal auto-inhibitory domain (called prodomain), which covers the substrate binding groove and needs to be proteolytically removed for protease activation (5). The auto-inhibitory prodomain contains a conserved noncontiguous amino acid sequence (ERFNIN) signature (6). Some PLCPs contain additional sequence elements for vacuolar targeting or retention in the endoplasmic reticulum (ER) (6).

A PLCP studied in much detail is RESPONSIVE TO DESICCATION-21 (RD21, encoded by At1g47128). RD21 was originally discovered as a drought-induced gene in *A. thaliana* (7). Later studies suggested a role in plant immunity and resistance against necrotrophic fungal pathogens such as *Botrytis cinerea* (8). RD21 has an interesting structure and biosynthetic pathway. It contains an NH_2-terminal presequence (signal peptide), a 20-kDa auto-inhibitory prodomain, a 33-kDa protease domain, a 2-kDa proline-rich domain, and a 10-kDa granulin-like domain (9, 10). Protease maturation occurs in several steps, comprising signal peptide release resulting in ProRD21, prodomain cleavage to produce an intermediate form of RD21 (iRD21), and final granulin domain removal to produce mature RD21 (mRD21) (11). RD21 accumulates in ER bodies and lytic vacuoles; some studies revealed accumulation in cell walls/apoplastic spaces (12). Evidence suggests the activity of RD21 is regulated by at least two different protease inhibitors (PIs): a serine protease inhibitor (serpin)-like activity (13, 14) and a Kunitz protease inhibitor-type activity (15–17). Serpins have been found in animals and plants and feature a reactive center loop (RCL), which displays a protease target sequence (18). Cleavage of the RCL results in an irreversible, covalent serpin::protease adduct (19–21). Similarly, *Arabidopsis* Serpin1 forms RD21::Serpin1 adduct (henceforth referred to as RD21-Serpin1) in response to fungal pathogens and was suggested to provide some set-point control of RD21 activity during programmed cell death (8, 11, 13, 14).

The second PI thus far known to inhibit RD21 activity is a protein originally identified as water-soluble chlorophyll-binding protein (WSCP) in *Lepidium virginicum* and other *Brassicaceae*, including *Arabidopsis* (16, 17, 22, 23). This protein contains a Kunitz-type PI signature and, in fact, interacts with RD21 in a tissue-specific manner (16, 17, 23). Complexes between RD21 and WSCP were found to accumulate in developing flowers and the apical hook of plants undergoing skotomorphogenesis (16, 17, 23). Underneath the soil or fallen leaves, the germinating sprouts

Significance

Nature has evolved highly specific protease inhibitors comprising serpin and Kunitz proteins. In humans, loss- or gain-of-function mutations in many protease inhibitor (PI) genes lead to compromised innate immune responses, dementia, thrombosis, and other diseases. Here, we report how plants make use of serpin and Kunitz PIs to control papain-like cysteine proteases over plant development and exploit them for defense.

Author contributions: S. Rustgi, D.v.W., and S. Reinbothe designed research; S. Rustgi, E.B.-F., C.R., and S. Reinbothe performed research; S. Rustgi, E.B.-F., C.R., and S. Reinbothe analyzed data; and S. Rustgi, D.v.W., and S. Reinbothe wrote the paper.

Reviewers: M.P., National Institute of Plant Genome Research; and D.Y., Wuhan University.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: diter@wsu.edu, srustgi@clemson.edu, or sreinbot@ujf-grenoble.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1621496114/-/DCSupplemental.

invest all the nutrient reserves contained in the seed into an exaggerated hypocotyl growth that aims at bringing the cotyledons above the soil and permitting the switch to photosynthetic growth. Although one might be tempted to expect a role of WSCPs in the greening process, no evidence was obtained for an expression of WSCPs in the cotyledons (16). Instead, promoter-β glucuronidase and protein expression studies revealed an accumulation of WSCP in the apical hook where the protein formed complexes with RD21 and inhibited its activity (16). Unlike RD21's irreversible binding to Serpin1, the interaction with WSCP was reversible and relieved on light exposure (16). Presumably, the chlorophyll precursor chlorophyllide formed during greening bound to WSCP and triggered the dissociation of the RD21::WSCP complex (16). As shown previously, WSCP contains a pigment binding site distantly related to that of the major light-harvesting protein of photosystem II (LHCII) (22) that presumably mediated this effect. Other studies have shown that WSCPs form tetrameric complexes in the presence of chlorophyllide, and thereby shield the pigment against the interaction with molecular oxygen, providing a unique photoprotection mechanism (24).

In the present work, protein biochemistry, 3D structural modeling, and expression profiling were combined to further illuminate the interaction of RD21 with its two PIs, Serpin1 and WSCP. We show that both structural components and fine-tuned expression of Serpin1 and WSCP ensure that RD21's activity is tightly controlled during plant development, thereby avoiding uncontrolled cell death by RD21 operating as prodeath factor.

Results

Protease Activity Profiling. Protease activity profiling was used to trace PLCPs in the apical hook of etiolated *Arabidopsis* plants. Using a biotinylated derivative of the cysteine protease inhibitor E-64, designated DCG-04, which irreversibly reacts with PLCPs (25), four main activity bands (referred to as C1–C4) and two weaker active protease bands (C5 and C6) were identified (Fig. 1 *A* and *B*). All four main bands were present in 2-d-old and 4.5-d-old etiolated seedlings (Fig. 1 *A*, *a*). Their activity pattern, however, changed during greening, with all six labeled bands declining after 12 h of illumination. Remarkably, all but band C1 reappeared at later stages of greening (Fig. 1 *A*, *b*). Protein sequencing and immunoprecipitation studies identified RD21, a PLCP previously identified by DCG-04 activity profiling (25), in protein extracts of 4.5-d-old etiolated seedlings (Fig. 1 *A*, *a*, lane 3 vs. lane 2, product C1). Other bands were not characterized further but may correspond to previously identified proteins (25). Competition experiments showed that DCG-04 labeling in tissue extracts was sensitive to E-64 (Fig. 1 *B*), as reported previously (25).

We next examined the presumed interaction of RD21 with AtWSCP and AtSerpin1. Seedlings expressing Flag-tagged RD21 were grown either in darkness under skotomorphogenetic conditions or exposed to white light for 12 h to induce greening. After purification from seedling extracts, RD21-containing protein complexes were isolated by affinity chromatography and subjected to nondenaturing PAGE. Immunoblotting with antibodies against *Arabidopsis thaliana* (At)WSCP and AtSerpin1 revealed that both etiolated and greening seedlings contained higher molecular mass complexes of RD21 that were composed differently. Etiolated seedlings contained mostly complexes of RD21 with AtWSCP, whereas greening seedlings contained mostly complexes of RD21 with AtSerpin1. Formation of the RD21::AtWSCP complex was sensitive to E-64, regardless of whether the inhibitor was added during or after complex isolation (Fig. 1 *C*). In contrast, formation of the RD21::AtSerpin1 complex in greening seedlings was sensitive to E-64 only as long as the inhibitor was added before complex isolation. When formation of the RD21::AtSerpin1 complex was allowed to proceed in extraction mixtures lacking E-64, the preestablished complexes could no longer be dissociated (Fig. 1 *C*). These results were consistent with the reaction

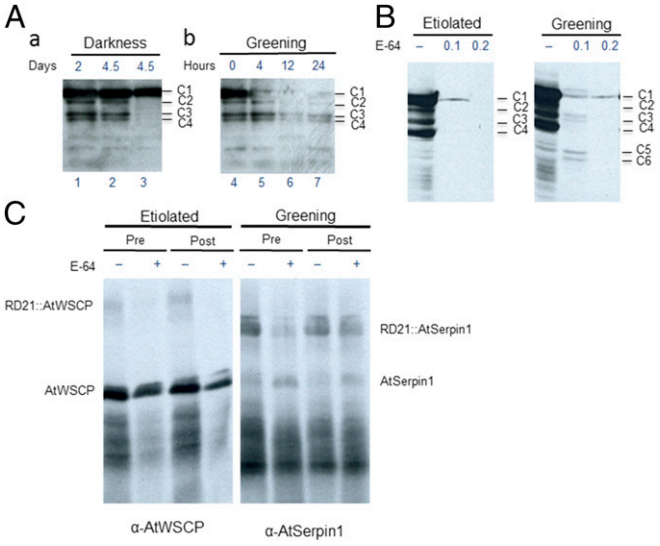

Fig. 1. Protease activity profiling in dark-grown and greening seedlings. (*A*) Time course of labeling of PLCPs by DCG-04 in protein extracts of dark-grown seedlings of different ages (in days after the onset of germination, *a*) and in 4.5-d-old dark-grown seedlings that had been exposed to white light for different periods (*b*). Following a step of precipitation by acetone, biotinylated proteins were purified on streptavidin beads, separated by SDS/PAGE and respective immunoblots probed with streptavidin-HRP and chemiluminescence. The four main bands were designated C1–C4. Lanes 1 and 2, as well as 4–7, show total PLCPs labeled with DCG-04, whereas lane 3 (corresponding to lane 2) shows an immunoprecipitation using RD21-specific antibody. (*B*) Specificity of DCG-04 labeling, as revealed by adding different amounts of E-64 (in millimoles) to the incubation mixtures. In addition to bands C1–C4, two more bands were detected in protein extracts of greening seedlings that are highlighted (C5 and C6). (*C*) Isolation of RD21-containing protein complexes from the protein extracts of dark-grown and greening seedlings. Either before ("Pre") or after ("Post") complex isolation, excess E-64 (2 mM) was added. RD21-containing protein complexes in turn were detected by nondenaturing PAGE and Western blotting, using AtWSCP- and AtSerpin1-specific antibodies, respectively. The positions of free and RD21-complexed AtWSCP (*Left*) and AtSerpin1 (*Right*) are indicated.

mechanism of serpins forming irreversible adducts with their target proteases (19–21).

We next tried to identify what PLCPs are present in etiolated and greening seedlings. DCG-04 was infiltrated into seedlings, and changes in the protein pattern were followed by pulse-labeling with [^{35}S]methionine in combination with SDS/PAGE and autoradiography (Fig. 2), as well as Coomassie staining (*SI Appendix*, Fig. S1). Fig. 2 shows that infiltration of etiolated and greening seedlings with DCG-04 provoked specific changes in the protein synthesis patterns. Several new bands appeared in either case that were supposed to represent either DCG-04-labeled, biotinylated bands or newly synthesized bands. Because the appearance of these bands was sensitive to cycloheximide, a known inhibitor of protein synthesis at cytoplasmic ribosomes, we concluded that these bands were newly synthesized (*SI Appendix*, Fig. S2). All these bands were absent from respective mock incubations in which the seedlings had been infiltrated with Tween-supplemented water (Fig. 2, lane 1; *SI Appendix*, Fig. S3). That the observed changes in the protein synthesis patterns were specific was proven by competition experiments in which we added excesses of known cysteine and serine protease inhibitors, such as E-64, leupeptin, antipain, and PMSF, along with DCG-04 during infiltration (Fig. 2 and *SI Appendix*, Fig. S4). Although E-64, leupeptin, and antipain abrogated the changes triggered by DCG-04, PMSF was ineffective (Fig. 2 and *SI Appendix*, Fig. S4), consistent with the data of van der Hoorn (2). Together, these results suggested highly specific changes in gene

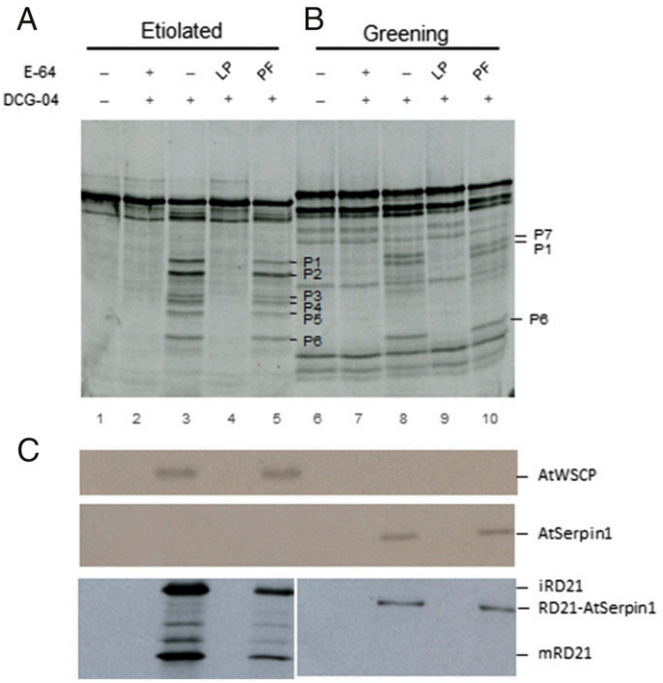

Fig. 2. Induction of PLCPs in planta. (*A–C*) DCG-04 was infiltrated into 4-d-old dark-grown seedlings, whereas controls lacked DCG-04. After infiltration, the seedlings were kept in the dark for another 12 h before being exposed to white light for an additional 12-h period to induced greening. Controls were kept in darkness for the same time, and thus further etiolated. Two hours before harvest, the different dark-grown and greening seedlings were pulse-labeled with [^{35}S] methionine or mock-incubated. Radioactive proteins were used for reducing SDS/PAGE and autoradiography (*A* and *B*), whereas nonradioactive protein extracted from the mock-incubated seedlings was run by nonreducing SDS/PAGE and used for Western blotting (*C*), using antibodies against AtWSCP (*Upper*), AtSerpin1 (*Middle*), and RD21 (*Lower*). Before analysis, the complexes recovered from greening plants were subjected to an immunoprecipitation with antibodies against AtSerpin1. In parallel experiments, E-64, leupeptin (LP), or PMSF (PF) were coinfiltrated along with DCG-04, as indicated. Bands denoted P1–P6 define proteins specifically synthesized in response to DCG-04, whereas bands defined iRD21 and mRD21 are all RD21 specimens formed through synthesis and sequential processing (*cf* Fig. 3). In DCG-04-infiltrated greening seedlings, RD21 was present in a complex with AtSerpin1.

expression that were provoked in the presence of DCG-04 in planta, leading to the induction of new proteins.

Western blot analyses using specific antibodies identified RD21, AtWSCP, and AtSerpin1 among the newly induced proteins. Fig. 2*C*, moreover, showed that AtWSCP was present in protein extracts of etiolated seedlings, but was undetectable in protein extracts from greening seedlings. As shown previously, AtWSCP expression is negatively light-regulated, such that both AtWSCP transcript and protein levels rapidly decline during greening (16). Protein gel blot analyses with antibodies against AtSerpin1, in contrast, revealed just the opposite effects. AtSerpin1 was present in protein extracts of greening seedlings, but not in extracts of etiolated plants (Fig. 2*C*). For RD21, the situation was more complex because of the appearance of multiple bands, iRD21, and mRD21, resulting from the successive removal of the signal peptide and autoinhibitory and granulin domains, respectively (Fig. 3*A*). Interestingly, the abundance of these bands was lower in DCG-04- plus PMSF-pretreated seedlings compared with only DCG-04-infiltrated seedlings (Fig. 2*C*, *Lower*, lane 5 vs lane 3), suggesting that induction and proteolytic processing of ProRD21 involved some serine protease activities. Vacuolar serine proteases, such as Hsr203 and P69B, have been implicated in programmed cell death of tomato seedlings in response to *Cladosporium fulvum* (26).

RD21 is presumed to interact with AtWSCP in a highly specific manner confined to the apical hook of etiolated seedlings (16). To corroborate this view, RD21-containing protein complexes were isolated from the apical hook of etiolated, DCG-04-infiltrated seedlings and RD21-reactive bands detected by immunoblotting. In control experiments, DCG-04 and E-64 were coinfiltrated into seedlings to compete out the DCG-04-induced changes in protein synthesis. As well, AtSerpin1-containing complexes were identified in the cotyledons of DCG-04-infiltrated, greening seedlings by immunoprecipitation with AtSerpin1 antibodies and the presence of RD21 traced by Western blotting using RD21-specific antibodies. When the obtained protein patterns were compared, four RD21 bands were seen (Fig. 3*B*): three in apical hook extracts and one in cotyledon extracts. The appearance of these four RD21 bands was sensitive to E-64 coinfiltrated with DCG-04 into the seedlings. The descent size of these bands suggested they represented ProRD21, iRD21, and mRD21 (all found in apical hook extracts) vs. RD21::AtSerpin1 (present in cotyledon protein extracts) (Fig. 3*B*). Protein sequencing proved that the identified complex represents the adduct of RD21 with AtSerpin1 (RD21-AtSerpin1).

The results presented thus far suggested that a fraction of RD21 is present in terms of RD21::AtWSCP complexes in etiolated seedlings, whereas another fraction of RD21 bound AtSerpin1 in greening seedlings. When RD21-containing complexes were purified from DCG-04-infiltrated greening seedlings, however, two bands were detected with AtSerpin antibodies, one at ~40 kDa and the lower at ~30 kDa (*SI Appendix*, Fig. S4). Similar results have been described by Lampl et al. (13) and were shown to be a result of cleavage of AtSerpin1 in the RCL. When cysteine protease inhibitors such as E-64, leupeptin, and antipain were coinfiltrated with DCG-04, neither the upper nor the lower AtSerpin1 bands appeared. Thus, induction of AtSerpin1 and subsequent cleavage of the RCL were highly specific and involved some cysteine proteases. In contrast, these findings revealed that despite the

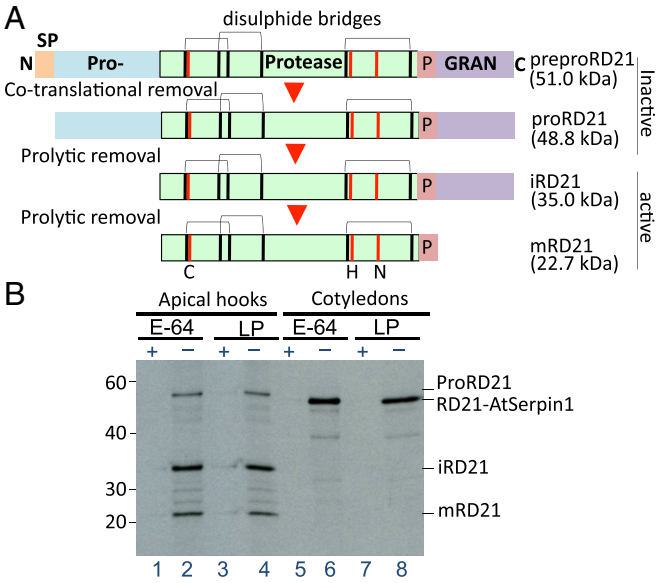

Fig. 3. Structure (*A*) and processing of RD21 (*B*). RD21 was detected in the apical hook of DCG-04-infiltrated dark-grown seedlings and the cotyledons of DCG-04-infiltrated greening seedlings. Seedlings were either infiltrated with DCG-04 alone or coinfiltrated with DCG-04 and either E-64 or LP, as indicated. RD21-containing protein complexes were then isolated as described. Before analysis, the protein extracts prepared from greening seedlings were subjected to immunoprecipitation with an AtSerpin1 antibody. RD21-reactive bands were detected on nonreducing SDS/PAGE gels by Western blotting, using RD21 antibody. Note the lack of ProRD21, iRD21, and mRD21 vs. the RD21-Serpin1 adduct in seedlings that were coinfiltrated with DCG-04 and E-64 or LP.

presence of the RCL-cleaved form, AtSerpin1 did not immediately produce adducts with RD21 (*SI Appendix*, Fig. S4). Otherwise, a third, higher-molecular-weight band should have been seen, which was not the case (*SI Appendix*, Fig. S4). Together, these data suggested that formation of the primary RD21::AtSerpin1 complex in the cotyledons of greening seedlings is initially reversible, whereas establishment of the RD21-AtSerpin1 adduct occurs later and involves the cleavage of AtSerpin1's RCL. In line with this interpretation, infiltration assays in which we tested the role of E-64 applied either simultaneously with DCG-04 or after tissue homogenization provided fundamentally different results. Although E-64 abolished the DCG-04-induced appearance of the RD21-AtSerpin1 adduct, E-64 added during tissue homogenization was ineffective (*SI Appendix*, Fig. S5). In vitro reconstitution experiments carried out at different molar rations of RD21 and AtSerpin1 (*SI Appendix*, Fig. S6*A*) converted the RCL-cleaved AtSerpin1 (AtSerpin1*) into the RD21-AtSerpin1 adduct (*SI Appendix*, Fig. S6*B*). After treatment with β-mercaptoethanol, not present in the experiments described in Figs. 2*C* and 3*B*, a fraction of the RD21-AtSerpin1 adduct was cleaved into RD21 and the RCL-cleaved AtSerpin1 (*SI Appendix*, Fig. S6*C*), proving the thioester linkage between RD21 and AtSerpin1 in the reconstituted complexes.

Expression Profiling of RD21, AtSerpin, and AtWSCP. RT-PCR-based transcript profiling was used to further explore the differential roles of AtWSCP and AtSerpin1 as RD21 inhibitors over seedling development in the dark and during greening. Concomitant Western blot analyses were conducted to quantify RD21, AtWSCP, and AtSerpin1 protein levels. As shown in Fig. 4*A*, *AtWSCP* transcript levels increased over the course of seedling development in the dark. In contrast, *RD21* transcript levels remained fairly constant, whereas *AtSerpin1* transcript levels slightly declined. eFP Browser searches (27) revealed 1.37- to 8.03-fold increase in *AtWSCP* transcript abundance during the first 4–12 h of dark growth, but only 0.01- to 0.35-fold increases for *RD21* (Fig. 4*B*). For *AtSerpin1*, a slight decline in transcript abundance became apparent in the dark during later stage of development (Fig. 4*B*). For all three gene products studied, protein levels followed similar courses as their respective transcript, indicating that their expression was largely controlled transcriptionally. Confirming previous results (16), AtWSCP expression was negatively light-regulated during greening, whereas that of RD21 and AtSerpin1 did not change significantly (Fig. 4*C*).

At the level of organ specificity, AtWSCP transcript and protein were found to accumulate in the apical hook, but not in the cotyledons, of dark-grown and greening seedlings (Fig. 4*D*). These findings corroborated previous promoter-β glucuronidase studies (16). In contrast, opposing effects were observed for the cotyledons of greening plants, where only AtSerpin1 transcript and protein, but not AtWSCP transcript and protein, accumulated to detectable levels (Fig. 4*D*). Interestingly, two RD21-reactive protein bands of the sizes expected for iRD21 and mRD21, respectively, were detected, of which the upper band, presumably representing iRD21, was present only in the cotyledons of greening seedlings.

In silico localization studies using the Cell eFP Browser (27) provided additional insights into the differential regulation of AtWSCP, AtSerpin1, and RD21 expression and function. AtWSCP and RD21 displayed overlapping localization patterns in the Golgi apparatus, ER, vacuole, and extracellular space. In contrast, AtSerpin1 localized to the cytosol of all cell types analyzed (*SI Appendix*, Fig. S7). At first glance, these localization patterns were not suggestive of RD21::AtSerpin1 interactions. It must be noted, however, that earlier studies localized AtSeprin1 and AtMC9 in the ER and apoplast (28). Although RD21 has thus far not been found in the apoplast, this possibility cannot be denied, as the RD21 homolog in tomato, the immune protease C14, was detected in the apoplast (12). In addition, it has been demonstrated that vacuolar proteases including RD21 are released in response to fungal elicitors causing the hypersensitive response (8), and thus could be ready to interact with cytosolic AtSerpin1. Together, the noticed differences in expression and compartmentalization are suggestive of unique, nonoverlapping roles of AtWSCP and AtSerpin1 during plant development.

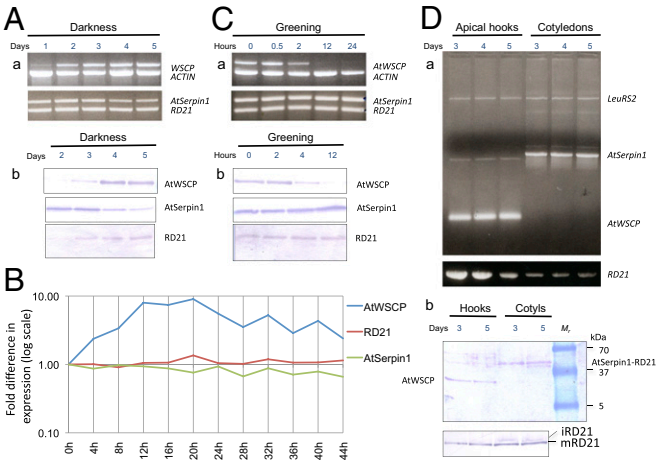

Fig. 4. Expression of *AtWSCP*, *AtSerpin1*, and *RD21* transcripts and proteins during skotomorphogenesis and greening. (*A*) Comparative analysis of AtWSCP, AtSerpin1, and RD21 expression during the course of skotomorphogenesis (in days after the onset of germination; *a*) and during greening of 4.5-d-old etiolated seedlings (in hours; *b*). (*a* and *b*) RT-PCR and Western blot analyses, respectively. RT-PCR mixtures contained specific primers for simultaneous *AtWSCP* and *ACTIN* vs. *RD21* and *AtSerpin1* transcript detection, whereas protein gel blots were probed individually, using the indicated antibodies. (*B*) In silico transcript analysis using the eFP Browser. *AtWSCP*, *RD21*, and *AtSerpin1* transcript abundances are shown for dark-grown seedlings at different points of development. The x-fold differences in the expression levels of different genes were calculated on the basis of absolute expression levels recorder for each gene in the control (at time 0) and plotted on a logarithmic scale to magnify the differences. Note that because the basal expression levels are very distinct for each of the genes studied, only relative changes in transcript abundance were given. (*C*) Comparative transcript and protein accumulation kinetics of AtWSCP, RD21, and AtSerpin1 over the course of greening. Transcript and protein identification and quantification were made as described in *A*. (*D*) Transcript and protein accumulation kinetics of RD21, AtWSCP, and AtSerpin1 in the apical hook of dark-grown seedlings and the cotyledons of greening seedlings. Note that quantification of *AtWSCP* and *AtSerpin1* transcripts (*a*) and proteins (*b*) was made simultaneously (*a* and *b*, *Upper*, respectively), whereas RD21 transcript and protein quantification was made in replicate assays (*a* and *b*, *Lower*, respectively). RT-PCR mixtures also contained primers for quantification of leucyl-tRNA synthetase 2 (LeuRS2), used as constitutive expression marker. Note that proteins for Western blotting in *A* and *C* were run by reducing SDS/PAGE, whereas those in *D* were obtained by either nonreducing SDS/PAGE (*Upper*) or reducing SDS/PAGE (*Lower*). Positions of molecular weight markers are given for the Western blots.

Discussion

A protease activity profiling approach was undertaken to identify PLCPs in *Arabidopsis* seedlings undergoing skotomorphogenesis and greening. Using DCG-04, a biotinylated analog of the cysteine protease inhibitor E-64 (25), several activity bands were detected in protein extracts of both seedling types; their abundance changed differentially during skotomorphogenesis and greening. As expected, the appearance of these bands was sensitive to E-64 and other cysteine protease inhibitors, indicating that these bands were a result of active cysteine proteases. Among the identified bands was RD21, a previously identified cysteine proteinase in *Arabidopsis*. In an attempt to identify PLCPs including RD21 in planta, proteins were discovered that were specifically induced on DCG-04 treatment. Their appearance was abrogated in the presence of E-64 and other cysteine protease inhibitors in planta, indicating that the effect of DCG-04 was specific. Because DCG-04

is a structural analog of E-64 differing from E-64 only by the composition of the peptide scaffold and attached biotin moiety (25), this result is suggestive of a new mechanism of elicitor-induced change in gene expression triggered by DCG-04. A possible scenario could be that DCG-04 targeted cysteine proteases such as RD21 and irreversibly inactivated them. To compensate for this effect, plants responded by up-regulating the expression of PLCPs, including RD21, but also that of their respective PIs including AtWSCP and AtSerpin1. How the respective changes in gene expression are brought about and whether these involve transcriptional and/or posttranscriptional steps remains to be addressed. Inhibitor studies with cycloheximide unveiled cytoplasmic protein synthesis as a requirement for accumulation of RD21 and other cysteine proteases and their respective PIs. In humans, for example, deleting papain-like cysteine proteases such as cysteine cathepsins revealed the operation of compensatory mechanisms in cancer (29). Proteases, in fact, are key components for regulating diseases and tumorigenic processes, including angiogenesis, tumor growth, and invasion. Elevated protease expression was found to be associated with poor patient prognosis across numerous tumor types (29). Other studies demonstrated important roles of lysosomal cysteine cathepsins in aging and neurodegeneration (30). Last but not least, cysteine cathepsins were reported to trigger caspase-dependent cell death through cleavage of BID (BH3 interacting-domain death agonist) and antiapoptotic Bcl-2 (B-cell lymphoma 2) homologs (31). To account for these very different functions, protease gene expression and activity must be tightly controlled over development and in specific tissues (32, 33).

RD21 is a cysteine endoprotease of ubiquitous occurrence and unique features in plants. RD21 is synthesized with an NH_2-terminal presequence for intracellular targeting through the secretary pathway, adjacent propeptide exhibiting auto-inhibitory activity, and COOH-terminal granulin domain of unknown function (11, 16). The NH_2-terminal propeptide is removed by a yet-unresolved mechanism either requiring an autocatalytic processing at low pH or the catalytic cleavage by a specific processing peptidase (11). Within the seed, the propeptide keeps RD21 in a silent state until germination occurs and drops vacuolar pH below 5, allowing intramolecular conformational changes for propeptide removal and protease activation.

When newborn seedlings develop further, they can either switch to photomorphogenetic growth in sunlight or skotomorphogenesis in darkness. In either case, RD21 is operative. Etiolated seedlings, for example, need to protect their Achilles' heels against crustacean devourers (16, 17). RD21 is sequestered in an inactive form in a complex with WSCP in the apical hook of etiolated seedlings. When the seedlings begin to de-etiolate, RD21 is released from the complex to counteract digestive proteases in the arthropod gut, conferring efficient protection (16). In greening and mature seedlings, RD21 is operative as a prodeath factor that triggers programmed host cell death once released from the vacuole (13, 14). Binding of RD21 to AtSerpin1, however, is a mechanism to control RD21's activity as prodeath factor (13, 14). Changes in vacuole permeability and protease compartmentalization are supposed to account for this set-point control mechanism (8, 14). In support of this model, mutant plants lacking RD21 or overexpressing AtSerpin1 both were found to exhibit significantly less elicitor-stimulated programmed cell death than plants lacking AtSerpin1 (8, 13, 14).

Molecular modeling was used to pinpoint the differential interactions of RD21 with AtSerpin1 and AtWSCP, respectively. According to Lampl et al. (13), AtSerpin1 has a structure very similar to that of animal serpins, with distinguishing, plant-specific features. As found for other serpins, AtSerpin1 displays a RCL, which contains a protease target sequence as bait. RCL cleavage gives rise to an irreversible, covalent serpin-protease adduct. By virtue of RD21's cysteine protease activity, this reaction mechanism is induced, triggering the irreversible inactivation of RD21. Molecular modeling of RD21 revealed a typical papain-like structure of RD21, with two almost equally sized lobes dubbed R (right) and L (left), divided by an active site cleft (16) that participated in the RD21::AtSerpin1 interaction. As we report here, the primary RD21 interaction with AtSerpin1 is reversible, giving rise to RD21::AtSerpin1 complex (Fig. 5 *A* and *B* and *SI Appendix*). Cleavage of the AtSerpin1's RCL then leads to a second type of complex containing RD21; still, formation of this complex was sensitive to E-64, and thus was reversible. Only at a later stage were RD21-AtSerpin1 adducts formed that accumulated in the cotyledons of greening plants but were not found in the apical hook of etiolated seedlings because of highly specific expression patterns.

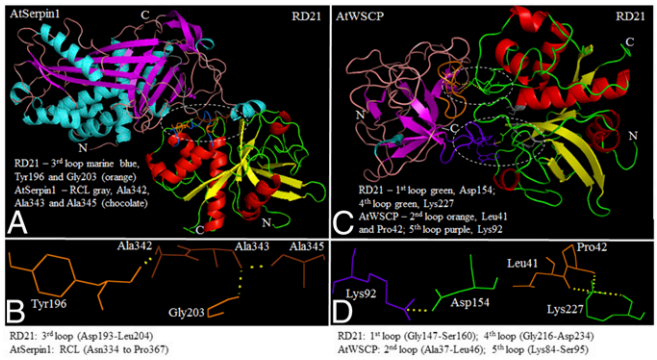

Fig. 5. Interaction models of RD21 with AtWSCP and AtSerpin1. (*A*) Interaction model of RD21 with AtSerpin1. (*C*) As *A*, but showing the interaction model for RD21 and AtWSCP. *A* and *C* depict ribbon diagrams, where β-strands and α-helices of RD21 are shown in yellow and red, respectively (unless otherwise indicated), and those of AtWSCP and AtSerpin1 in magenta and cyan (unless otherwise indicated). (*B* and *D*) Interacting amino acids, respectively, in RD21::AtSerpin1 and RD21::AtWSCP complexes. Amino acid numbering for RD21 and AtSerpin1 is based on their respective locations in the full-length preproteins, and for AtWSCP numbering is based on the mature protein.

The interaction model of RD21 with AtWSCP is fundamentally different from that with AtSerpin1. According to Boex-Fontvieille et al. (16), AtWSCP most closely resembles soybean Kunitz-type trypsin inhibitor and tamarind Kunitz inhibitor, with an α-turn and 10 antiparallel β-strands forming a barrel-like structure (*SI Appendix*, Fig. S8) (34, 35). According to studies on oryzacystatin-I and papain-like proteases, as well as on tamarind Kunitz inhibitor and its interactions with factor Xa and trypsin (34, 35), we proposed an AtWSCP::RD21 interaction model (16) that was validated here. Specifically, the second loop (Ala37-Leu46; orange in Fig. 5*C*), which spans between β-strands 2 and 3 and encompasses the LHCII signature, and the fifth loop (Lys84-Ser95; purple in Fig. 5*C*), which connects β-strands 5 and 6 and forms RCL, are proposed to establish interactions with RD21. In this interaction model, Tyr88 and Pro89 in the RCL of AtWSCP are predicted to intrude into the active site region of RD21 containing Cys161 and His297, and thereby to block its protease activity (16). Moreover, one amino acid residue, Lys92, in the RCL and two amino acid residues, Leu41 and Pro42, in the LHCII signature sequence are predicted to form hydrogen bonds with amino acid residues Asp154 and Lys227, respectively, in RD21 (Fig. 5*D*). Together, these hydrogen bonds are expected to stabilize the observed AtWSCP::RD21 interaction. In contrast, the presence and close physical proximity of the LHCII signature of AtWSCP to the catalytic triad of RD21 could explain the previously observed light-triggered, presumably chlorophyllide-mediated dissociation of the AtWSCP::RD21 complex (16).

Previous genome annotations suggest a single protease inhibitor may target up to five, and sometimes even more, proteases (36). We therefore identified proteases other than RD21 that could bind AtSerpin1 and WSCP, respectively. Using molecular modeling based on the validated RD21::AtSerpin1 and RD21::AtWSCP

interactions (13, 16), new PI–target protease interactions were revealed and included metacaspase-9 (MC9), interacting with AtSerpin1, and the proaleurain maturation protease (PMP), as well as XBCP3 (xylem bark cysteine peptidase 3), interacting with AtWSCP (*SI Appendix*, Fig. S9). The large similarity of interactions noted in either group of PIs to that with RD21 demonstrates that there is a topological conservation in the established structural models for AtSerpin1 and MC9 vs. AtWSCP, XBCP3, and PMP. The latter result confirms and extends findings by Halls et al. (15), who identified AtWSCP as a potent inhibitor of the recombinant PMP and papain. In fact, there is a small family of granulin domain-containing protease in *Arabidopsis* encoded by At3g19390 (PMP), At4g34460 (*XBCP3*), At1g47120 (*RD21*), and At5g43060 (*RD21B*), of which two, RD21 and PMP, were proven to interact with AtWSCP in yeast two-hybrid screens (17). It must be noted, however, that these screens failed to demonstrate AtWSCP::XBCP3 interactions modeled here (*SI Appendix*, Fig. S10). For expression of the *XBCP3* cDNA in yeast, a partial cDNA clone was used that encodes a truncated version of XBCP3 lacking its predicted NH_2-terminal signal peptide and interaction domain with RD21. Thus, truncation of the NH_2-terminal region of XBCP3 could explain the observed lack of interaction in the Y2H screen. In addition, fusion of XBCP3 and RD21 to either a DNA binding domain or a transcription activation domain might have interfered with their interaction in yeast. For AtMC9, the situation is less complicated because Vercammen et al. (28) demonstrated its interaction with AtSerpin1. Together, our biochemical, modeling, and expression profiling studies shed new light onto the complex network of protease–protease inhibitor interactions controlling innate immunity and arthropod deterrence during plant development.

Materials and Methods

Protease Activity Profiling. Protease activity profiling was conducted essentially as described by van der Hoorn et al. (25). For further details, see *SI Appendix*.

Pulse-Labeling of Total Leaf Proteins with ^{35}S-Methionine. Pulse-labeling of protein with [^{35}S]methionine was done as previously described (37). For further details, see *SI Appendix*.

Protein Complex Isolation and Analysis. Conditions for the isolation of HM_r complexes containing Flag-tagged RD21 or hexa-histidine-tagged AtWSCP or AtSerpin1 were as described by Boex-Fontvieille et al. (16) (*cf SI Appendix*).

Expression Profiling and In silico-Localization Studies. In silico-expression and localizations studies were performed with the eFP Browser (27). Conventional RT-PCR was carried out following (16, 17) (*cf SI Appendix*).

ACKNOWLEDGMENTS. This work was supported by a grant of the Life Sciences Discovery Fund (3143956-01) to S. Rustgi and D.v.W. and a grant of the Chaire d'Excellence Program of the French Ministry of Research and Education to C.R.

1. Rawlings ND, Waller M, Barrett AJ, Bateman A (2014) MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 42(Database issue):D503–D509.
2. van der Hoorn RA (2008) Plant proteases: From phenotypes to molecular mechanisms. *Annu Rev Plant Biol* 59:191–223.
3. Beers EP, Jones AM, Dickerman AW (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in Arabidopsis. *Phytochemistry* 65(1):43–58.
4. Drenth J, Jansonius JN, Koekoek R, Swen HM, Wolthers BG (1968) Structure of papain. *Nature* 218(5145):929–932.
5. Taylor MA, et al. (1995) Recombinant pro-regions from papain and papaya proteinase IV-are selective high affinity inhibitors of the mature papaya enzymes. *Protein Eng* 8(1):59–62.
6. Grudkowska M, Zagdańska B (2004) Multifunctional role of plant cysteine proteinases. *Acta Biochim Pol* 51(3):609–624.
7. Koizumi M, Yamaguchi-Shinozaki K, Tsuji H, Shinozaki K (1993) Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in *Arabidopsis thaliana*. *Gene* 129(2):175–182.
8. Shindo T, Misas-Villamil JC, Hörger AC, Song J, van der Hoorn RAL (2012) A role in immunity for Arabidopsis cysteine protease RD21, the ortholog of the tomato immune protease C14. *PLoS One* 7(1):e29317.
9. Yamada K, Matsushima R, Nishimura M, Hara-Nishimura I (2001) A slow maturation of a cysteine protease with a granulin domain in the vacuoles of senescing *Arabidopsis* leaves. *Plant Physiol* 127(4):1626–1634.
10. Yamada T, et al. (2001) Isolation of the protease component of maize cysteine protease-cystatin complex: Release of cystatin is not crucial for the activation of the cysteine protease. *Plant Cell Physiol* 42(7):710–716.
11. Gu C, et al. (2012) Post-translational regulation and trafficking of the granulin-containing protease RD21 of *Arabidopsis thaliana*. *PLoS One* 7(3):e32422.
12. Bozkurt TO, et al. (2011) *Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. *Proc Natl Acad Sci USA* 108(51):20832–20837.
13. Lampl N, et al. (2010) *Arabidopsis* AtSerpin1, crystal structure and in vivo interaction with its target protease RESPONSIVE TO DESICCATION-21 (RD21). *J Biol Chem* 285(18): 13550–13560.
14. Lampl N, Alkan N, Davydov O, Fluhr R (2013) Set-point control of RD21 protease activity by AtSerpin1 controls cell death in *Arabidopsis*. *Plant J* 74(3):498–510.
15. Halls CE, et al. (2006) A Kunitz-type cysteine protease inhibitor from cauliflower and *Arabidopsis*. *Plant Sci* 170(6):1102–1110.
16. Boex-Fontvieille E, Rustgi S, von Wettstein D, Reinbothe S, Reinbothe C (2015) Water-soluble chlorophyll protein is involved in herbivore resistance activation during greening of *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 112(23):7303–7308.
17. Boex-Fontvieille E, Rustgi S, Reinbothe S, Reinbothe C (2015) A Kunitz-type protease inhibitor regulates programmed cell death during flower development in *Arabidopsis thaliana*. *J Exp Bot* 66(20):6119–6135.
18. Huntington JA (2006) Shape-shifting serpins–advantages of a mobile mechanism. *Trends Biochem Sci* 31(8):427–435.
19. Schechter I, Berger A (2012) On the size of the active site in proteases. I. Papain. 1967. *Biochem Biophys Res Commun* 425(3):497–502.
20. Huntington JA, Read RJ, Carrell RW (2000) Structure of a serpin-protease complex shows inhibition by deformation. *Nature* 407(6806):923–926.
21. Gettins PGW (2002) Serpin structure, mechanism, and function. *Chem Rev* 102(12): 4751–4804.
22. Satoh H, Uchida A, Nakayama K, Okada M (2001) Water-soluble chlorophyll protein in *Brassicaceae* plants is a stress-induced chlorophyll-binding protein. *Plant Cell Physiol* 42(9):906–911.
23. Bektas I, Fellenberg C, Paulsen H (2012) Water-soluble chlorophyll protein (WSCP) of *Arabidopsis* is expressed in the gynoecium and developing silique. *Planta* 236(1): 251–259.
24. Horigome D, et al. (2007) Structural mechanism and photoprotective function of water-soluble chlorophyll-binding protein. *J Biol Chem* 282(9):6525–6531.
25. van der Hoorn RA, Leeuwenburgh MA, Bogyo M, Joosten MH, Peck SC (2004) Activity profiling of papain-like cysteine proteases in plants. *Plant Physiol* 135(3):1170–1178.
26. Sueldo D, et al. (2014) Dynamic hydrolase activities precede hypersensitive tissue collapse in tomato seedlings. *New Phytol* 203(3):913–925.
27. Winter D, et al. (2007) An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2(8):e718.
28. Vercammen D, et al. (2006) Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. *J Mol Biol* 364(4):625–636.
29. Akkari L, et al. (2016) Combined deletion of cathepsin protease family members reveals compensatory mechanisms in cancer. *Genes Dev* 30(2):220–232.
30. Stoka V, Turk V, Turk B (2016) Lysosomal cathepsins and their regulation in aging and neurodegeneration. *Ageing Res Rev* 32:22–37.
31. Droga-Mazovec G, et al. (2008) Cysteine cathepsins trigger caspase-dependent cell death through cleavage of bid and antiapoptotic Bcl-2 homologues. *J Biol Chem* 283(27):19140–19150.
32. Turk V, Stoka V, Turk D (2008) Cystatins: Biochemical and structural properties, and medical relevance. *Front Biosci* 13:5406–5420.
33. Turk V, et al. (2012) Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochim Biophys Acta* 1824(1):68–88.
34. Song HK, Suh SW (1998) Kunitz-type soybean trypsin inhibitor revisited: Refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from *Erythrina caffra* and tissue-type plasminogen activator. *J Mol Biol* 275(2):347–363.
35. Patil DN, Chaudhary A, Sharma AK, Tomar S, Kumar P (2012) Structural basis for dual inhibitory role of tamarind Kunitz inhibitor (TKI) against factor Xa and trypsin. *FEBS J* 279(24):4547–4564.
36. Farady CJ, Craik CS (2010) Mechanisms of macromolecular protease inhibitors. *ChemBioChem* 11(17):2341–2346.
37. Reinbothe S, Reinbothe C, Parthier B (1993) Methyl jasmonate-regulated translation of nuclear-encoded chloroplast proteins in barley (*Hordeum vulgare* L. cv. Salome). *J Biol Chem* 268(14):10606–10611.