Decoy Substrates for Pathogen Proteases Mediate Innate Immunity *via* a Shared Receptor in *Arabidopsis*

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Louis-Philippe Maier

aus Hamburg

Tübingen

2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	04.05.2023
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Georg Felix
2. Berichterstatter:	Prof. Dr. Thorsten Nürnberger

Index

1 Abstract

1.1 English version

Pathogens and their host organisms are in an ever ongoing evolutionary arms race, wherein the survival of either is detrimental to the other. Driven by forces of selection, pathogens have evolved ingenious strategies to invade the host and compromise its defenses. The host, in turn, has co-evolved an intricate immune system to detect attacking pathogens and defend the organism in a timely manner. This war of attrition is primarily fought out in the extracellular space and takes place on the level of proteins and cells. Extracellular proteases, secreted by pathogens during attack, are known virulence factors. We found that various extracellular proteases from human and plant pathogens are highly immunogenic in the model plant *Arabidopsis thaliana*. Their perception shares a common mechanism that depends on their catalytic activity. It involves the secretion of decoy substrates by the plant cell, whose specific processing by pathogen proteases releases danger-signaling peptides of a shared active motif into the extracellular space. Precursors of these danger-signaling peptides form a superfamily of pro-peptides. Their active motif is perceived by a single cell surface-localized immune receptor. Using the example of the plant pathogenic bacterium and agricultural pest *Xanthomonas arboricola* pv. pruni, we demonstrate that specific perception of a single extracellular pathogen proteases is sufficient to prevent establishment of infection. Our results furthermore suggest not only the possibility to transfer the perception system to different host species, but also to design synthetic decoy substrates for yet unperceived proteases from pathogens of interest.

1.2 German version (Zusammenfassung)

Pathogene und ihre Wirte befinden sich in einem all-währenden evolutionären Wettrüsten, in welchem das Überleben des Einen den Schaden des Anderen bedingt. Pathogene haben raffinierte Strategien entwickelt, die es Ihnen erlauben, ihre Wirte erfolgreich zu infizieren. Der Wirt hat im Gegenzug ein komplexes Immunsystem entwickelt, welches dazu dient, angreifende Pathogene rechtzeitig zu erkennen und sich gegen diese zur Wehr zu setzen. Dieser Zermürbungskrieg wird primär im extrazellulären Raum auf Ebene der Proteine und Zellen ausgefochten. Proteasen, welche von Pathogenen im Laufe der Infektion in den extrazellulären Raum sekretiert werden, sind bekannte Virulenzfaktoren. Wir haben herausgefunden, dass verschiedene Proteasen von Menschen- und Pflanzen-Pathogenen Immunantworten in der Modell-Pflanze *Arabidopsis thaliana* auslösen. Ihre Erkennung beruht auf einem gemeinsamen Mechanismus, der von der katalytischen Aktivität der Proteasen abhängt. Die Pflanze sekretiert eine Mischung von Köder-Substraten, die spezifisch durch Pathogen-Proteasen geschnitten werden, und dadurch Gefahr-signalisierende Peptide in den extrazellulären Raum freisetzen. Vorstufen dieser Peptide gehören zu einer großen Familie von Pro-Peptiden. Ihr immunogenes Peptid-Motif wird durch einen Plasmamembran-lokalisierten Immunrezeptor erkannt. Anhand des Beispiels des pflanzenpathogenen Bakteriums *Xanthomonas arboricola* pv. pruni demonstrieren wir, dass die spezifische Erkennung einer einzelnen Pathogen-Protease ausreicht, um die Infektion der Pflanze zu verhindern. Darüber hinaus suggerieren unsere Ergebnisse sowohl die Möglichkeit, dieses Erkennungs-System auf andere Spezies zu übertragen, als auch artifizielle Köder-Proteine nach Bedarf für anderweitig nicht erkannte Pathogen-Proteasen zu generieren.

2 Introduction

2.1 Conceptual background

Organisms are constantly challenged by a variety of pathogens in their natural environment. The perception of such threats is prerequisite to any kind of active defense and vital to the organism's survival. Animals and plants have evolved complex immune systems for the detection of this danger, independently of each other[1–4]. Plants lack both, specialized motile immune cells and an adaptive immune system. Instead they have evolved an intricate innate immune system enabling them to detect a plethora of diverse disease agents[5].

Cell-surface-localized *pattern recognition receptors* (PRRs)[6] are one important group of sensors employed by the plant immune system. Most prominent among them are the superfamily of *receptor-like kinases* (RLKs), which has experienced a massive expansion in plants[7], and the well-studied family of LRR-RLKs included therein. LRR-RLKs are characterized by an extracellular ecto-domain composed of multiple eponymous leucine-rich repeats, a single-pass transmembrane domain, and a cytosolic kinase domain. The ecto-domain of LRR-RLKs is responsible for high-affinity binding of the specific cognate ligand[8].

With PRR-type LRR-RLKs, the cognate ligands are (macro-) molecules whose presence in the extracellular space, the so-called plant apoplast, indicates imminent pathogen attack to the plant. Thus, these danger signals[9,10] are perceived as a proxy for the attackers themselves. They can be derived from either the attacker, or the plant, and in addition - perhaps - as a product of an interaction of both. Depending on their pathogen or host origin, these signals have been conceptualized as either *pathogen / microbe-associated molecular patterns* (PAMPs / MAMPs[11]; *e.g.* bacteria-derived peptides elf18[12] and flg22[13], which are perceived by the LRR-RLKs EFR[14] and FLS2[15]), or *damage-associated molecular patterns* (DAMPs[16]), respectively.

A curious case are plant-own danger-signaling peptides, which do not quite fit the descriptive title used for canonical DAMPs. While their active levels do change during pathogen attack, they are not simply released passively upon structural *damage* to the plant cell. Instead, it is the plant that actively controls their expression, secretion or post-translational modification in response to perceived danger. These peptides are believed to modulate plant immune responses and assume a function not unlike the cytokine system in animals. They have hence been described as phyto-cytokines[10,17,18]. Examples of ligand / receptor pairs in the model plant *Arabidopsis thaliana* that have been suggested to fit this description include CTNIPs / HSL3[19] (synonymous: SCREWs / NUT[20]), PEPs / PEPRs[21–27], PIPs / RLK7[28–30], PSKs / PSKRs[31–33], PSYs / PSY1R[31,32,34], and RALFs / FER[35]. Recently, the SCOOPs / MIK2[36–38] pair, which is limited to the cabbage plant family (*Brassicaceae*), has been proposed as yet another example of a phyto-cytokine system in this species. Markedly, the above-mentioned examples share a common characteristic, which is that the receptors are genetically encoded in very few copies (usually 1 to 2), while the precursor genes coding for their respective peptide ligands have diversified to varying degree. The evolutionary forces driving this diversification of cytokine-like peptide ligands are still poorly understood and up to scientific debate.

Regardless of their respective ligand, LRR-RLKs appear to share a common activation mechanism. This involves the binding of a suitable ligand to the surface of the receptor ecto-

domain, and is followed by the subsequent recruitment of a SERK-family co-receptor[39]. SERKs are a subfamily of LRR-RLKs with short ecto-domains, the most prominent member of which is SERK3 (synonymous: BAK1). Biochemical and structural studies gave rise to the notion that the ligand acts as a molecular glue[40,41] between the ecto-domains of the LRR-RLK receptor and co-receptor. This complex[42,43] between receptor and co-receptor ecto-domains also brings their transmembrane and kinase domains[43,44] in close proximity to each other. And this proximity is believed to allow for auto- and trans-phosphorylation[45,46] of the kinase domains and subsequent initiation of downstream signaling events.

Early signaling events include rapid changes in ion fluxes across the plant plasma membrane. These are dependent on secondary active transport (*i.e.* an electro-chemical proton gradient) and can be measured as an alkalization of the plant apoplast, which is easily done in plant cell-suspension cultures[13]. Another early event is the extracellular production of reactive oxygen species (ROS) *via* plasma membrane localized NADPH oxidases. And later events include the production of the stress hormone ethylene. Signaling cascades culminate in defense responses that are aimed towards active defense against the perceived pathogen threat. A process that is conceptually summarized in its entirety by the term *pattern-triggered immunity* (PTI)[11]. The relatively low proportion of critically diseased plant individuals found in nature may be interpreted as an indicator of the efficacy of the plant immune system; basal immunity provides plant resistance against an ubiquitous plethora of opportunistic pathogens.

In order to break resistance and establish infection, specialized pathogens, in turn, have co-evolved[47] mechanisms to aid their virulence. These include strategies to invade and colonize the host, to evade perception and defense, or to suppress host immune responses altogether[48,49]. Pathogen genes and their products involved in these mechanisms are referred to as virulence factors. Many virulence factors exert their function by manipulation of host target (macro-) molecules to the pathogen's benefit, and are therefore also referred to as effectors. Secreted pathogen enzymes are particularly well suited to be recruited as efficient virulence factors. This is because they do not interact with virulence targets in a strict bimolecular manner, but instead manipulate an excess of their respective target substrate through their catalytic activity. Among them are secreted pathogen proteases, which have been associated with pathogenesis in humans, livestock, and plants[50,51]. Historically, the plant field has primarily focused on the study of such proteases that are injected from pathogenic bacteria directly into the cytoplasm of the host cell[52]. While apoplastic proteases of pathogenic bacteria have been largely neglected. Considering the ubiquity and importance of proteolytic virulence factors for pathogen fitness, it is self evident that plants may have evolved mechanisms to harness their catalytic activity for highly efficient perception purposes.

2.2 Research question

Gram-negative bacteria of the genus *Xanthomonas* are relevant[53] disease agents with devastating impact on global agricultural production of a variety of crops. *Xanthomonas* strains are highly adapted to their respective host. They cause disease in a manner that depends on both, the bacterial pathovar and the host plant cultivar[54]. The model plant *A. thaliana* is resistant against most *Xanthomonas* strains, including *X. arboricola* pv. pruni[55,56], whose natural hosts are prune trees (*e.g.* almond, cherry, peach and plum).

We were interested in molecular mechanisms that alert the plant to the apoplastic presence of

Xanthomonas and may contribute to the non-host resistance of *A. thaliana* against *X. arboricola* pv. pruni (Xap).

Xanthomonas bacteria contain an enigmatic proteinaceous immunogen, which is released upon cell lysis, and whose perception depends on the PRR-type receptor RLP1 (synonymous: ReMAX)[57,58]. The perceived protein was never identified and thus termed *enigmatic MAMP in Xanthomonas* (eMAX). A large scale purification from biofilm of Xap was performed with the aim of the protein's identification. Surprisingly, the most-active fractions displayed chromatographic and biochemical properties that were distinct from eMAX. A small quantity of this novel immunogenic protein, just sufficient for mass spectrometric analysis, provided a candidate (unpublished work by Vahid Fallahzadeh-Mamaghani under supervision of Prof. Dr. Georg Felix).

Preliminary results obtained with deletion mutants of the candidate-encoding locus in Xap suggested that the immunogenic activity in *Xanthomonas* biofilm depended on it. Its perception was apparently independent of RLP1. And biochemical evidence suggested that the candidate protein, a secreted protease, might be perceived through its enzymatic activity (unpublished work by Louis-Philippe Maier, during his *Master of Science* thesis under supervision of Prof. Dr. Georg Felix).

The aims of this work were (1) the confirmation of preliminary results, (2) the unambiguous proof that the immunogenic activity depended on the candidate's protease activity and not on its molecular structure, (3) identification of the protease's substrate and (4) the cognate receptor involved in its perception.

It is demonstrated that *A. thaliana* perceives the extracellular *Xanthomonas* protease. Not through its molecular structure, but due to its specific proteolytic activity on a plant-secreted precursor of a danger-signaling peptide. The exceptionally sensitive perception of *Xanthomonas via* this protease is sufficient to suppress growth of the pathogen in the non-host. The danger-signaling peptide belongs to a family of secreted pro-peptides, which are all but inactive at physiologically relevant concentrations as precursors, yet highly immunogenic in their mature form. Evidence is provided that unrelated apoplastic proteases of various pathogens process and thereby activate distinct members of the pro-peptide family. All of which signal through a shared LRR-RLK receptor that perceives a common motif among the mature peptides. Based on these observations we propose a novel perception model, in which precursors of host-secreted danger signals act as decoy substrates for pathogen-associated proteases, whose activation-by-processing leads to pathogen detection and resistance.

3 Results

3.1 An extracellular *Xanthomonas* protease is perceived *via* proteolytic activation of an *Arabidopsis* danger signal

Plant pathogenic bacteria of the genus *Xanthomonas* contain an enigmatic PAMP, eMAX, whose perception is dependent on the cognate receptor RLP1 (synonymous: ReMAX). In further attempts to identify this PAMP, we detected a heat-labile immunogenic activity in *Xanthomonas arboricola* pv. pruni (Xap) that rapidly and strongly induced extracellular alkalization in *A. thaliana* cell suspension cultures (*Fig.1A*). However, in contrast to the characteristics of eMAX described previously[57], this immunogenic activity is secreted by the bacteria; it was associated with the intact surface of washed cells and the extracellular fluid of the biofilm. Ultrafiltration experiments suggested the secreted immunogen to be larger than 30 kDa in molecular size. Importantly, when tested in *A. thaliana* mutants lacking *RLP1*, this extracellular immunogen was still active as an inducer of ethylene biosynthesis (*Fig.1B*). While ethylene induction by eMAX, contained in the lysate of the bacterial cells, depended on *RLP1*, as expected. This strongly suggested that the secreted immunogen was distinct from eMAX.

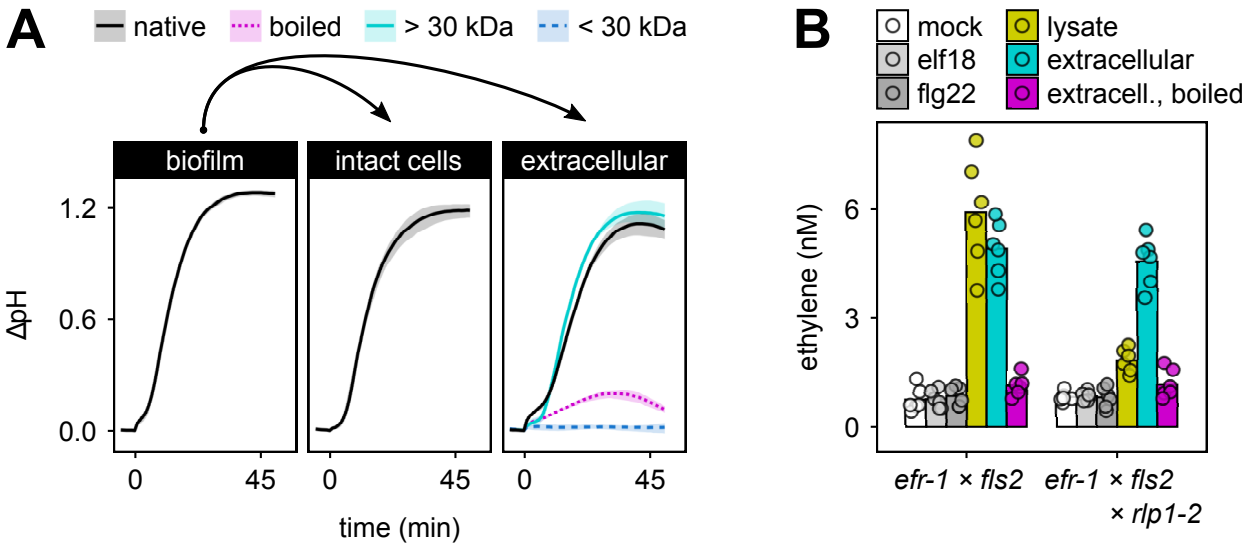


Fig.1: Immunogenic activity in *Xanthomonas* is secreted, heat-labile, and independent of RLP1. (**A**) Extracellular alkalization of *A. thaliana* cell suspension cultures in response to treatment with biofilm fractions [0.06% (w/v) of biofilm mass per cell culture volume] of *X. arboricola* pv. pruni. Biofilm was either used as is, or separated into extracellular fluid and bacterial cells by centrifugation. Bacterial cells were further washed at lower *g*-force to remove debris, and extracellular fluid was subjected to ultrafiltration to assess molecular weight cut-offs of the immunogen. Lines and ribbons indicate the arithmetic mean and total data spread of 3 replicates, respectively. (**B**) Induction of ethylene biosynthesis in leaf discs of indicated *A. thaliana* mutants. Treatment with extracellular fraction, or soluble cell lysate of *Xap* at concentrations equivalent to < 0.05% (w/v) of biofilm. 1 µM of the immunogenic peptides elf18 and flg22 were used as controls of the *efr-1 × fls2* null-mutant background. Bars indicate arithmetic means of 6 replicates. Each experiment of (A) and (B) was repeated at least two times independently, with procedural variations, yet similar results.

For further characterization, large quantities of the novel immunogen were purified from the cell-free biofilm fluid of Xap in a series of chromatographic steps (*for an example chromatogram of the last step, see: Fig.2A*). An aliquot of the final fraction was further separated by semi-native SDS-PAGE, a procedure that retained most of the immunogenic activity. The activity correlated with a protein band migrating at a size of about 130 kDa (*Fig.2B*), which shifted to 60 kDa when samples were denatured by boiling (not shown). Mass spectrometric analysis of the tryptic in-gel digest of this band, identified the protein as a *subtilisin-like protease autotransporter* (SLPA) of *Xanthomonas arboricola* (*NCBI RefSeq* locus ID: WP_026064720; misleading *EMBL* locus name - *pspB* - is avoided, in order to prevent confusion with an unrelated *E. coli* gene of the same acronym).

SLPAs belong to the superfamily of classical autotransporter proteins, which are exclusively found in Gram-negative bacteria. The export of these autotransporters proceeds *via* the type-*Va* secretion system (T5aSS) a complex secretion process that has been comprehensively reviewed[59]. In brief, an N-terminal signal peptide and C-terminal *β*-barrel (*i.e.* the eponymous autotransporter domain) allow for SEC- and BAM-dependent translocation of the interjacent passenger domain (*i.e.* their cargo) through the inner and outer membrane of the bacterial envelope, respectively. In case of SLPAs, the exported passenger consists of a subtilisin-like protease domain (*MEROPS* S8 clan of serine proteases), and a *β*-helical stem C-terminal thereof (*Fig.2C*). The latter has been suggested to function in folding of the cargo during export[60] and its subsequent surface presentatio[n.] SLPAs lack the the self-inhibitory pro-domain that is usually found in subtilisins. Several SLPA family members are released into the extracellular space as soluble proteins, either through autoproteolysis, or independent of their own protease function[61–67].

SLPAs are highly conserved in *Xanthomonas* spp. (two clusters of sequences with > 88% pairwise aa identity within each, and > 56% across clades). Orthologs are found also in infamous pathogens of plants (*Brenneria* spp., *Pectobacterium* spp*., Pseudomonas syringae*, *Xylella fastidiosa*), fungi (*Pseudomonas tolaasii*), animals and humans (*Actinobacillus pleuropneumoniae*[64], *Bordetella pertussis*[62], *Burkholderia pseudomallei*[68], *Campylobacter* spp., *Cronobacter sakazakii*, *Fusobacterium nucleatum*[67], *Glaesserella parasuis*, *Haemophilus* spp., *Mannheimia haemolytica*[69], *Neisseria meningitidis*[70], *Pseudomonas aeruginosa*[65], *Proteus mirabilis*[71], *Serratia marcescens*[72], *Stenotrophomonas maltophilia*) (*Fig.2D*). They are usually encoded as single copy genes. SLPAs are structurally related to the well-studied family of *immunoglobulin A1 protease* autotransporters (IgA1 proteases)[73] and *protease autotransporters of Enterobacteriaceae* (SPATEs)[74], whose enzymatic cargos are chymotrypsin-like proteases (*MEROPS* S6 clan of serine proteases). IgA1 proteases[73] / SPATEs[75] are virulence factors, and the same is assumed for SLPAs[66,68,70,71,76].

Immunogens can be characterized and compared to each other by quantification of their specific activity. Concentrations of active compound that are sufficient to induce half-maximum immune response are expressed as EC_{50} values. Purified $SLPA^{Xap}$ displayed an extraordinary immunogenic activity with an EC_{50} of just 4.5 pM in *A. thaliana* cell suspension cultures (*Fig.2E*). To put this scale of activity into perspective, it shall be considered that a milligram of this protease would suffice to trigger immune responses in metric tons of plant cells. Total recovery from the lossy purification procedure suggested that Xap secreted > 55 pmol of this protease per mg of biofilm.

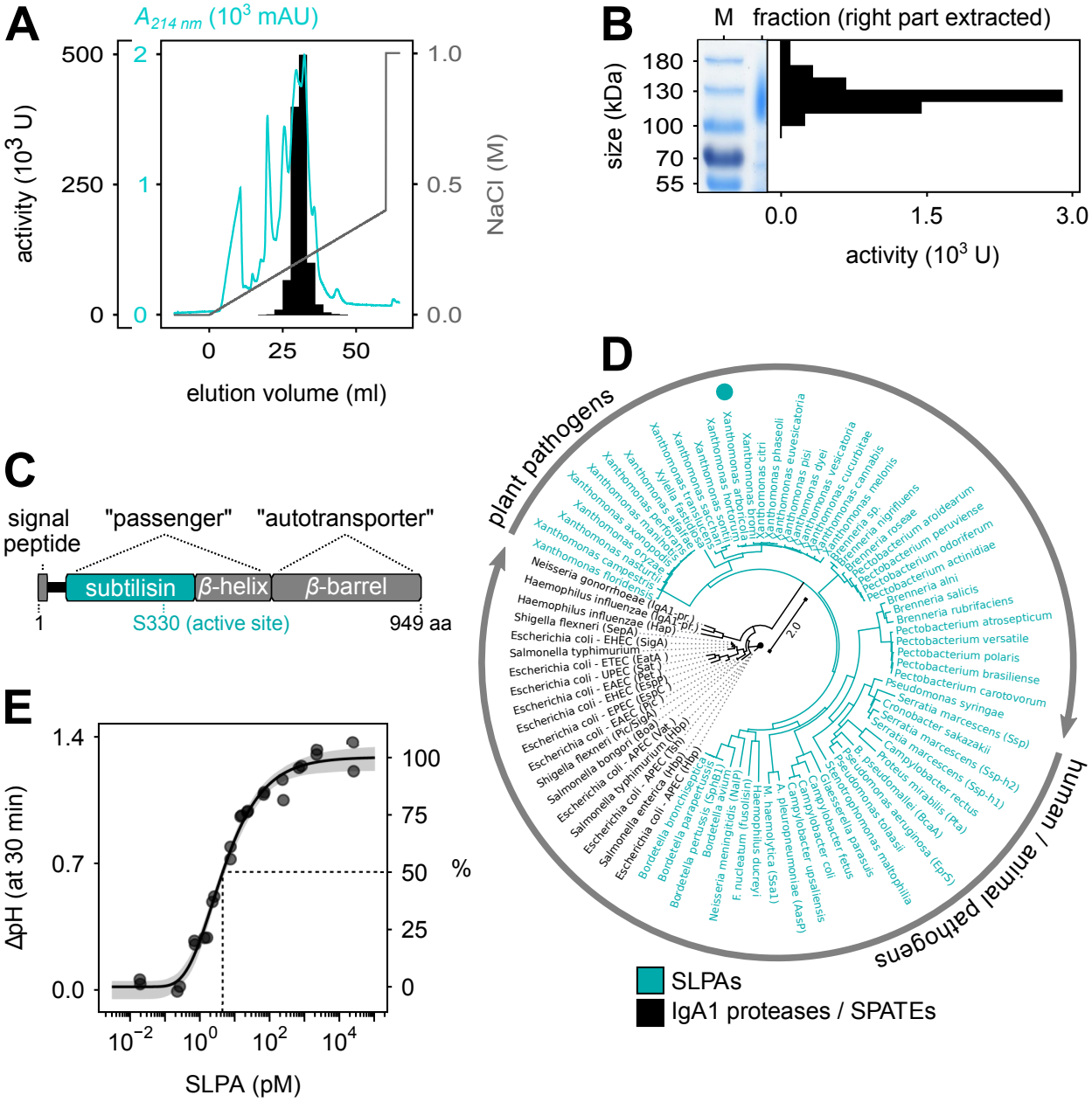


Fig.2: A subtilisin-like protease autotransporter (SLPA) is the immunogen candidate. (**A**) Anion-exchange FPLC (pH 6) was the last step of the immunogen's activity-guided purification; a representative chromatogram of two replicate polishing runs is shown. (**B**) Semi-native SDS-PAGE (*i.e.* sample not boiled; < 0.1% (w/v) of SDS; no reducing agent) of the purified immunogen. The active band was subjected to mass spectrometric analysis and identified the immunogen candidate, that is SLPA. (**C**) Domain structure of SLPA$^{Xap.}$; the subtilisin-like protease domain and the position of the active site serine are highlighted in cyan color. (**D**) Maximum-likelihood phylogenetic tree of SLPA orthologs based on their amino acid sequence (*RAxML*; tree of maximum parsimony among 100 randomly seeded replicates; *BLOSUM62* aa substitution model). Structurally related chymotrypsin-like proteases of the IgA1 protease / SPATE family were used as an outgroup for tree construction. One representative per species was chosen when sequences of multiple pathovars were available. The scale bar indicates radial patristic distances as the average number of substitutions per site. (**E**) Dose-dependent response of *A. thaliana* cell suspension culture to purified SLPA^{Xap} ($EC_{50} = 4.5 \pm 0.6$ pM, $P < 6 \times 10^{-7}$, $R^2_{McF} = 0.99$); plotted line and ribbon indicate best-fit model function and 95% *CI* (1 000 bootstraps), respectively.

The initial purification of SLPA, leading to its mass spectrometry-assisted identification was done by Vahid Fallahzadeh-Mamaghani. Shown figures (A) and (B) are illustrative reproductions obtained during the later preparative scale purification of SLPA as part of this thesis.

Isogenic deletion mutants of *slpa* were generated in Xap. Biofilm fluids of these *Δslpa* mutants did not induce the responses observed with the wild-type strain (*Fig.3A*), corroborating the role of $SLPA^{Xap}$ as the main immunogen in the biofilm fluid of Xap. An important question remained: Is it $SLPA^{Xap}$'s heat-labile tertiary structure, or the thereof dependent catalytic activity of the protease that is perceived by the plant? In order to address this question, complementation of the *Δslpa* strain of Xap was attempted with either the wild type *slpa* allele, or a mutant allele in which the active site serine at position 330 was replaced by an alanine, *slpa(S330A)*. Only the wild type *slpa* allele restored the immunogenic activity in the *Δslpa* background (*Fig.3B*), indicating that the immunogenic activity of $SLPA^{Xap}$ is dependent on its catalytic function.

This implied that the immunogenic effect of $SLPA^{Xap}$ occurred indirectly, *via* the protease-dependent release of peptides that were perceived as danger signals by *A. thaliana*. SLPA's proteinaceous substrate might have been either of microbial, or plant origin. Regardless, the putative peptide product was expected to be heat-stable, in contrast to the heat-labile protease. The bacterial biofilm contained only negligible amounts heat-stable activity, which was independent of *slpa*. Hence, we turned to the plant apoplast as a potential source of the protease's substrate.

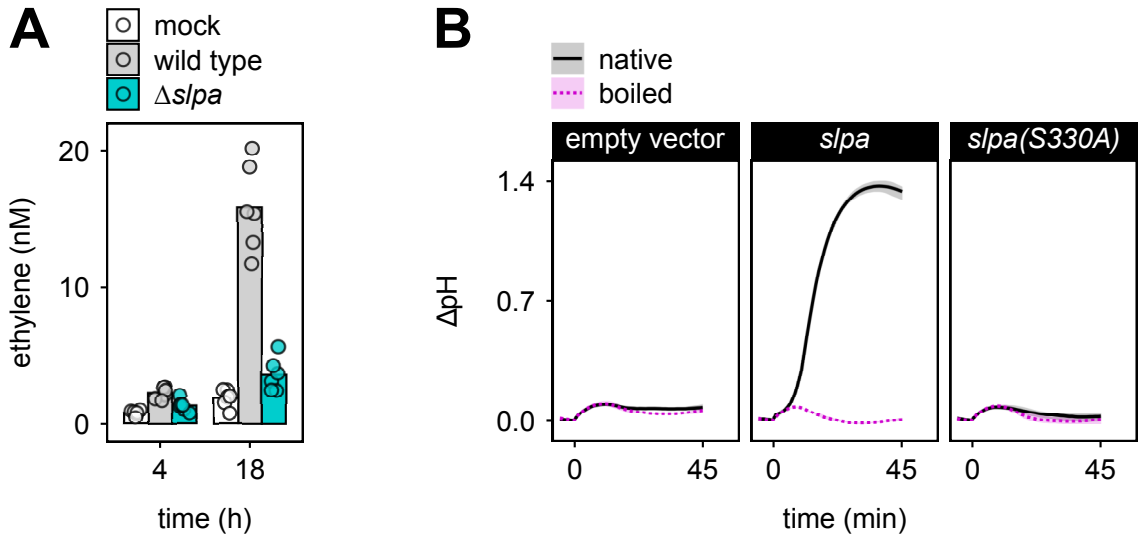


Fig.3: Immunogenic activity of SLPA depends on its protease function. (**A**) Ethylene biosynthesis in leaves of *efr-1 × fls2 × rlp1-2* null-mutants of *A. thaliana* in response to treatment with cell-free biofilm fluid of indicated Xap strains [0.2% (w/v) of biofilm mass]. The response was time-dependent, implying substrate limitation of an enzyme-catalyzed reaction. Bars indicate arithmetic means of 6 replicates; the experiment was independently repeated. (**B**) Extracellular alkalization induced by treatment with biofilm fluids [0.02% (w/v) of biofilm mass] of *Δslpa* complementation strains. Complementation of the null-mutant was attempted with either the wild type allele, an active site mutant, or an empty vector control. Lines and ribbons indicate arithmetic means and total data spread of 3 replicates. The experiment was repeated more than 3 times, independently, with similar results.

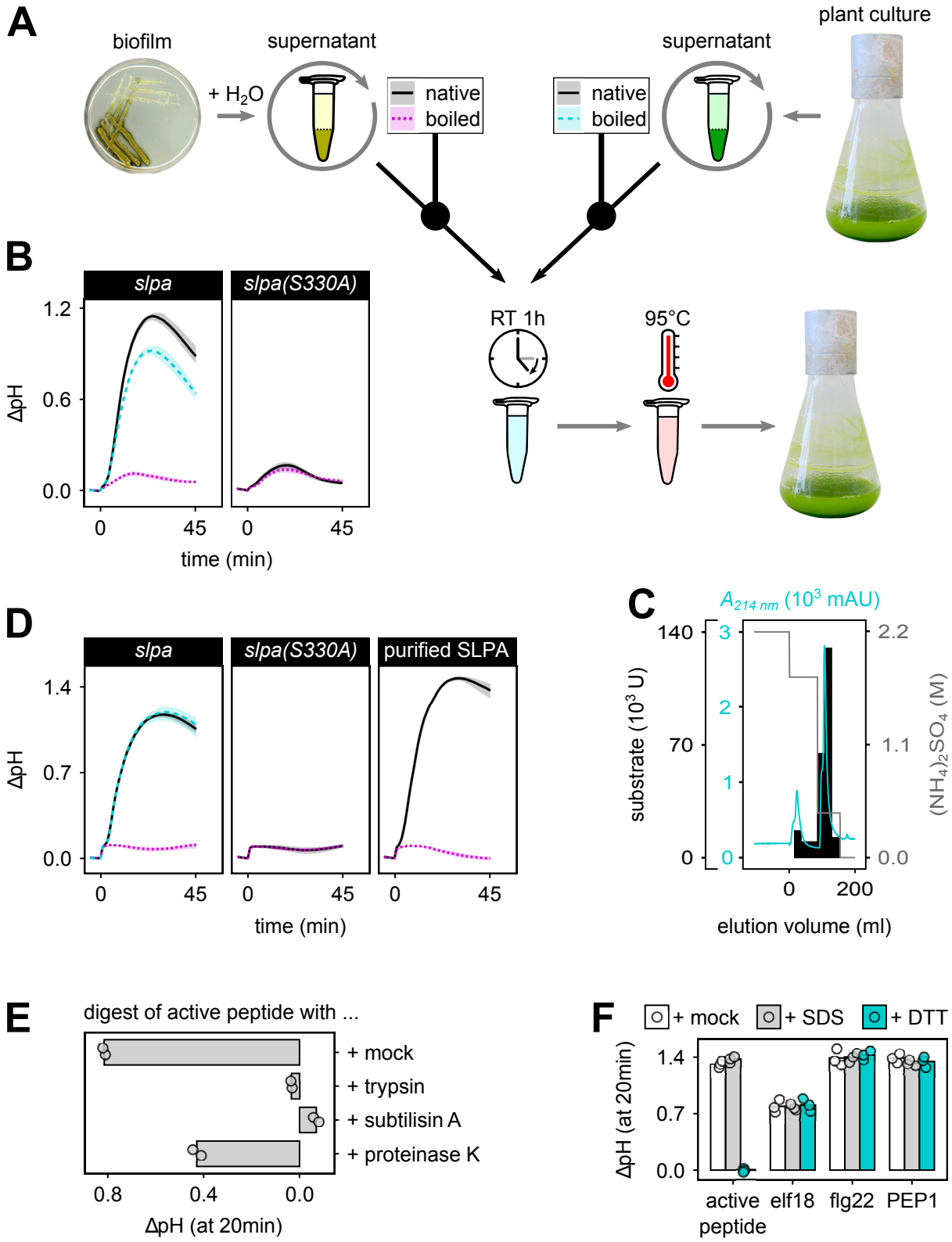


Fig.4: SLPA processes an inactive apoplastic plant pro-peptide into an active peptide signal. (**A**) Graphical overview of the procedure used for *in vitro* digests of plant apoplastic fractions with SLPA^{Xap}. (**B**) The supernatant of *A. thaliana* cell suspension cultures was mixed with biofilm fluids [0.5% (w/v) of biofilm mass] of Xap *Δslpa* complementation strains and incubated for 1 h at RT. The

in vitro digests were subsequently boiled for 20 min at 95°C to inactive the protease SLPA^{Xap}, and used at 8% (v/v) to induce naive *A. thaliana* cell cultures. Bacterial biofilm fluids and plant cell culture supernatant were used either native, or boiled prior to digest. (**C**) The putative apoplastic substrate was chromatographically enriched from 22 liters of *A. thaliana* cell suspension culture supernatant; hydrophobic-interaction chromatography (shown) was the second purification step that followed an initial cation-exchange capture process. Amount of substrate contained in chromatographic fractions was determined by quantification of heat-stable activity that was generated upon digest with SLPA^{Xap}. (**D**) Same *in vitro* digests as in (B), but with enriched substrate fractions [> 3 U/µl] of (C) instead of crude *A. thaliana* cell culture supernatant. The products were used at a final concentration of 0.3% (v/v). Purified SLPA^{Xap} [10 nM] was included in digests as a further control. Plotted lines and ribbons in (B) and (D) indicate arithmetic means and total data spread of 3 replicates. (**E**) The heat-stable immunogenic product of these digests was treated with endopeptidases [1 mg/ml] for 3 h at 37°C. These general proteases were subsequently inactivated by boiling, and products were used at at concentrations of < 3 U/ml to induce cell cultures. (**F**) The active peptide was boiled in either 1% (w/v) SDS, or 30 mM DTT, and used at a final concentration of 10 U/ml. Unrelated immunogenic peptides were treated in the same manner and used as controls at 1 nM concentration. Bars in (E) and (F) indicate arithmetic means of 2 and 3 replicates, respectively. Each experiment was independently repeated more than 3 times, with variations in procedure, but with similar results.

Co-incubation of the cell-free apoplastic fraction of *A. thaliana* culture (*i.e.* cell suspension culture supernatant) together with biofilm fluid of Xap*Δslpax* complementation strains (*Fig.4A*) generated a heat-stable activity in a manner that depended on the wild type *slpa* allele (*Fig.4B*). The substrate was chromatographically enriched from the apoplastic fraction of *A. thaliana* cell suspension cultures (*Fig.4C*) and used in similar *in vitro* digests with purified SLPA^{Xap}, reproducing the phenomenon (*Fig.4D*). Importantly, the generation of the immunogenic product was abolished when protease fractions were heat-denatured prior to digest, while boiling of the substrate was tolerated. Taken together, these results clearly point to a process in which SLPA^{Xap}'s proteolytic action on a plant-secreted pro-peptide generated an active peptide signal that is perceived by the plant. This enzyme-catalyzed reaction may aid in signal amplification through the accumulating peptide product, and may explain the exquisite sensitivity of *A. thaliana* for SLPA^{Xap}.

The proteinaceous character of the putative peptide was confirmed by loss of activity upon further digest with endopeptidases, including the unspecific protease subtilisin A of *Bacillus licheniformis* (*Fig.4E*). The immunogenic peptide was insensitive to denaturation in the presence of SDS, yet sensitive to treatment with reducing agents like DTT (*Fig.4F*). Hence, it appears that it includes at least one disulfide bridge of importance for perception by the plant. Both, the sequence of the peptide and its cognate immune receptor remained to be identified.

3.2 SCOOP pro-peptides act as decoy substrates for extracellular pathogen proteases

In order to identify the plant-derived peptide signal, we established a batch process for the rapid enrichment of the unknown apoplastic precursor from large volumes of *A. thaliana* cell suspension culture supernatant. The enriched fraction was used as a substrate in bulk digests with $SLPA^{Xap}$. The protease was subsequently inactivated by heat-treatment, and the active peptide product was purified to homogeneity through conventional chromatographic means (*Fig.5A*). We obtained a total amount of 17 nmol of the immunogenic peptide product, corresponding to a recovery of about 94 pmol of (pro-) peptide per liter of apolastic fluid.

The purified peptide was highly immunogenic; it displayed an EC_{50} of 23 pM in cell suspension cultures of *A. thaliana*. However, 99.9% of the activity was lost when the peptide fraction was treated with reducing agents (*Fig.5B*).

The mass spectrometric identification of the immunogenic peptide was complicated by several post-translational modifications: The activity was correlated with predominant peptide masses in UPLC-ESI-MS runs (*Fig.5C*). Interrelated pairs of masses were based on 3 520.7 Da and grouped in two clusters, whose difference corresponded to the residue mass of a terminally located lysine (128.10 Da). And differences within each cluster were a multiple of 132.04 Da, indicating the decoration of a single residue with variable-length chains of pentoses. The previous observation that the peptide's immunogenic activity was sensitive to treatment with reducing agents led us to repeat the run with chemically reduced peptide, but otherwise identical conditions. Surprisingly, all of the previously observed peptide masses had disappeared, and instead two new clusters were detected. Masses in the first cluster were based on 1,469.7 Da and differed by the sequential addition of pentose equivalents, while the two masses of the second were based on 2,053.1 Da and differed by the presence of the lysine residue. The sum of pairs of masses from both clusters (minus 2 proton equivalents lost during disulfide bridge formation) equaled the masses observed with the peptide in its oxidized state. In summary, the mass-correlation runs suggested that the immunogenic peptide is a heterodimeric glycopeptide consisting of two disulfide-bridged chains. Chain-1 was glycosylated with up to 3 pentoses, and chain-2 existed in isoforms differing by a terminal lysine. Partial sequences of both chains were obtained by peptide *de novo* sequencing in MS / MS analyses (*Fig.5D*). But no database entries matching these sequences were found among secreted proteins of *A. thaliana*. That is, until it was considered that multiple proline residues of both chains might be hydroxylated. With the changed query a single database hit was found.

The database hit is a secreted prepro-peptide (*UniProt* ID: O04468), which is encoded by the gene At1g65500 (*AGI* locus ID). It contains an N-terminal signal peptide for secretion. The sequence obtained with the purified immunogenic peptide suggests that $SLPA^{Xap}$ specifically cleaves the pro-peptide C-terminal of O/P_{P2}-A/I_{P1}, a motif preceding the N-terminus of chain-1, and forming the C-terminus of both chain-1 and chain-2, respectively.

Corroborating this identification, we found that non-glycosylated synthetic versions of the purified heterodimeric peptide retained excellent immunogenic activity, regardless of whether hydroxy-proline (EC_{50} of 52 pM) or proline (EC_{50} of 91 pM) residues were incorporated into chain-1 (*Fig.6A*). Consequently, neither decoration with pentoses, nor prolyl-4-hydroxylation is crucial for activity.

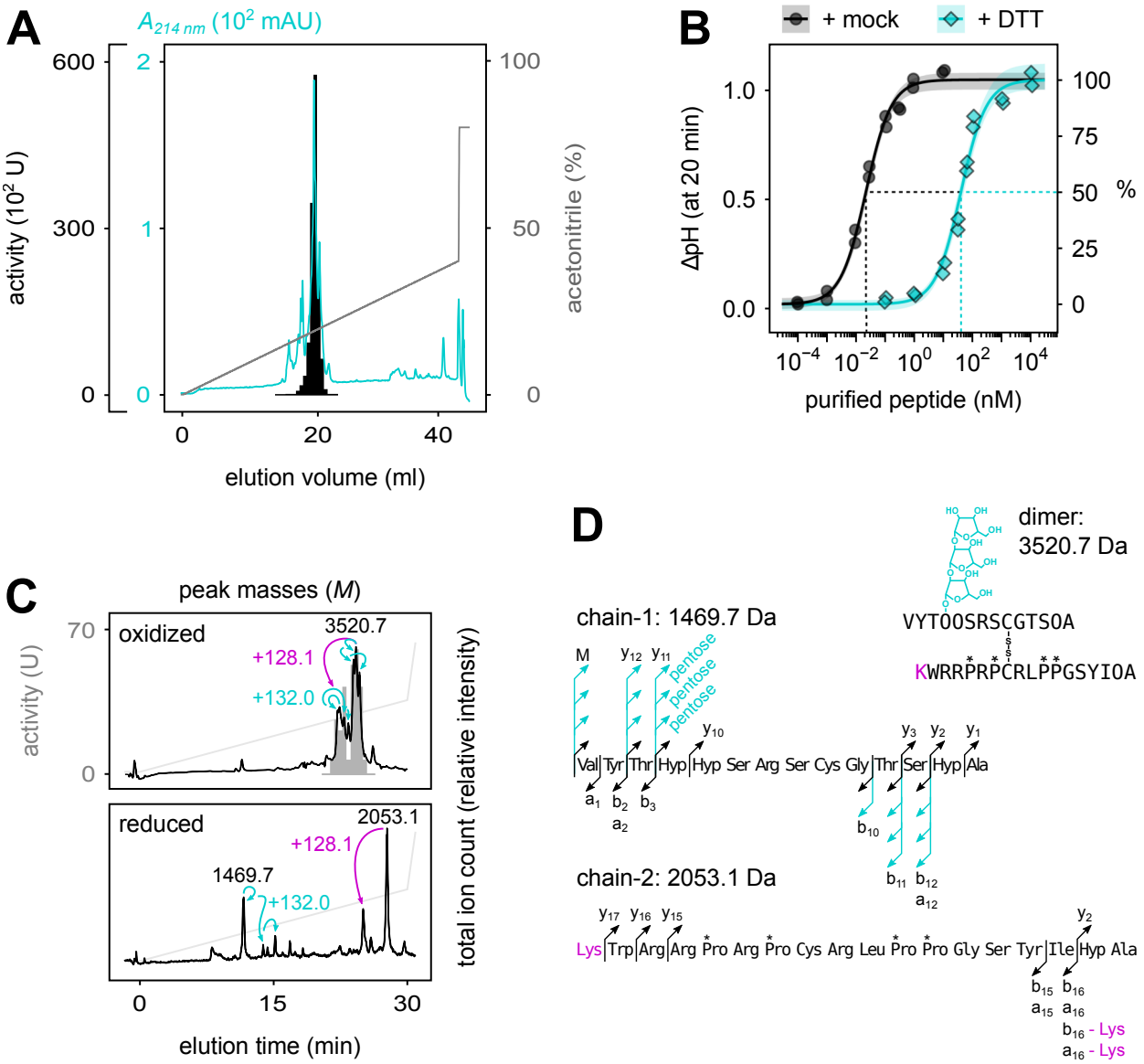


Fig.5: The active peptide signal is a heterodimeric glycopeptide. (**A**) Reversed-phase HPLC was the last step in the peptide's activity-guided purification; a representative chromatogram of two replicate runs is shown. (**B**) The concentrated peptide fraction [10 µM] was either mock-treated, or treated with the reducing agent DTT [30 mM], and boiled at 95 °C for 20 min. Immunogenic activity of the oxidized (EC_{50} = 22.6 ± 2.0 pM, $P < 2 \times 10^{-7}$, R^2_{McF} = 0.99) and reduced (EC_{50} = 40.4 ± 6.1 nM, $P < 2 \times 10^{-5}$, R^2_{McF} = 0.98) samples was assessed in dose-response assays with *A. thaliana* cell suspension culture. Plotted lines and ribbons indicate best-fit model functions and 95% *CI* (1 000 bootstraps), respectively. (**C**) Reversed-phase UPLC-ESI-MS runs with the purified peptide; activity of the oxidized form of the peptide was correlated with interrelated peptide masses. The run was repeated under the same conditions, but with 2-mercaptoethanol - treated (*i.e. reduced*) peptide. (**D**) Masses obtained with the reduced peptide were selected for tandem-MS fragmentation experiments. Due to incomplete fragmentation (energy lost in deglycosylation of precursor masses and predominant dissociation of y_2 ions), a total of 29 fragmentation spectra had to be recorded (not shown), before cumulative evidence allowed for the identification of the heterodimeric peptide. Depicted is the *post-hoc* analysis of both peptide chains' fragmentation patterns. The letter "O" in aa sequences indicates hydroxyproline residues; asterisks (*) label 4 proline residues of chain-2, an undetermined one of which was hydroxylated.
Mass spectrometric analysis was done in collaboration with Dr. Mark Stahl.

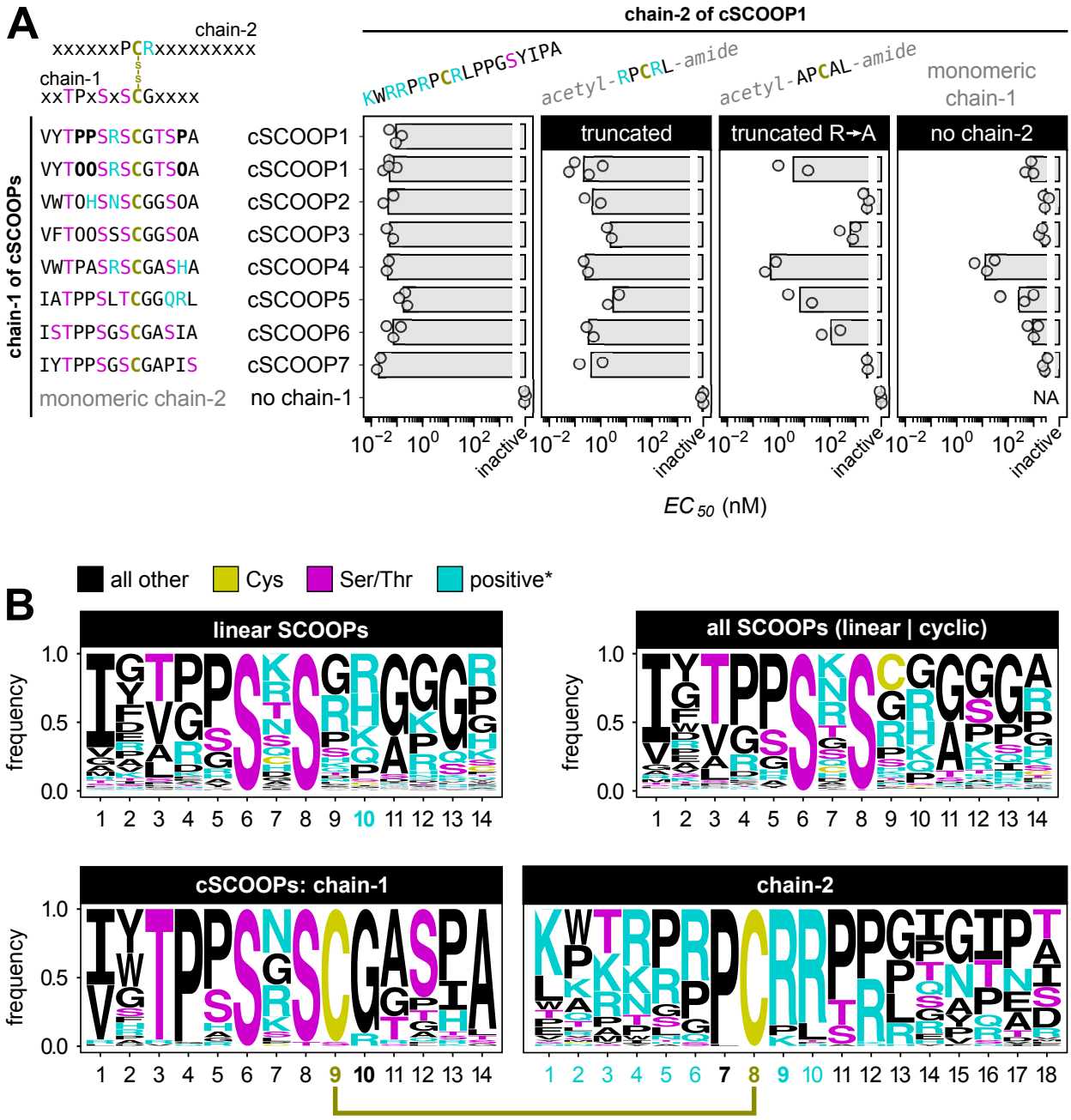


Fig.6: The heterodimeric peptide and its paralogs require a positive charge, provided by chain-2, for full immunogenic activity. They belong to the superfamily of SCOOPs. (**A**) Quantification of the immunogenic activity of synthetic analogs of the heterodimeric peptide (cSCOOP1) and its paralogs. For each family member, chain-1 was either tested in its reduced monomeric form, or as heterodimers with the identified peptide's chain-2. Chain-2 was used in full length, or as truncated versions thereof (wild-type, or with arginine to alanine substitutions). Truncated peptides of chain-2 were N-terminally acetylated and C-terminally amidated to eliminate terminal net charges. Extracellular alkalization was measured at 15 min post peptide treatment. Each data point represents an EC_{50} estimate from an independent dose-response experiment in *A. thaliana* cell suspension cultures; bars indicate geometric means of 2 to 4 repetitions. Conserved sequence positions are indicated and residues of interest are color-coded according to the legend in the lower panel. The letter "O" indicates hydroxyproline residues in peptide aa sequences. (**B**) Sequence logos of linear SCOOPs (active as monomers) and cyclic cSCOOPs (active as heterodimers) generated from

alignments of secreted SCOOP-like prepro-peptides of the *Brassicaceae* plant family. Linear SCOOPs carry a conserved positive charge at position 10 (highlighted in cyan color), which is not found in the homologous chain-1 of cSCOOPs. The positive charge is found on chain-2 of their heterodimers. Our definition of "*positive**" does not refer to the net charge of aa residues, but considers any positively charged moiety; it includes zwitterionic side chains besides basic ones. For tentative evidence of the positive charge's importance to linear SCOOP activity*, see Suppl.Fig.3.*
The tBLASTn search, whose results were used for alignments and creation of sequences logos shown in (B), was done by Prof. Dr. Georg Felix.

However, the immunogenic activity did depend on the heterodimeric structure of the peptide. Monomers of chain-1 and chain-2 alone were all but inactive (EC_{50} > 790 nM) and void of any activity, respectively. This explained the observed loss of activity upon treatment with reducing agents, which broke the disulfide bridge connecting the two chains of the heterodimer.

The prepro-peptide belongs to a family of 7 paralogs in *A. thaliana* (*Suppl.Fig.1*). The family members have erroneously been suggested to be transmembrane proteins that would localize to the plasma membrane[77] or the endoplasmic reticulum[78]; the small prepro-peptides were N- or C-terminally labeled with large GFP tags in these studies (*see Suppl.Fig.2 for alternative localization data*).

They contain a single pair of cysteine residues, and are each expected to form a cyclic structure *via* an intra-molecular disulfide bridge. The sequence of the 14 aa long chain-1 is similar between family members (> 59% pairwise identity) and contains 6 conserved sites, including two serines residues (and occasionally a threonine) at position 6 and 8, a cysteine at position 9, and a glycine at position 10. The 16-18 aa long chain-2 shows little conservation between family members, apart from a high concentration of positively charged residues (mean *pI* of 11.9) around a shared P-C-R motif, of which the central cysteine forms the disulfide bridge to the peptide chain-1.

Considering the sequence variability within chain-2, it was hypothesized that it was not the entire length of this peptide chain, but the positively charged core motif shared between family members that was essential for activity of the heterodimeric peptide. Indeed, a drastically truncated version of the chain-2 peptide was still able to complement the activity of chain-1 (EC_{50} of 230 pM) (*Fig.6A*). Substitution of the positively charged amino acids in this shortened chain-2 for alanine affected the activity considerably (EC_{50} of 3.7 nM; *i.e.* loss of 93% of activity). Similar behavior was also observed with chain-1 peptides of the other family members, whose activity was complemented by dimerization with the identified peptide's chain-2. Thus it appears as if the main function of chain-2 were to provide a positive charge adequately positioned next to the disulfide-bridged chain-1.

We noticed the similarity between chain-1 of the family members and the recently identified superfamily of plant-secreted SCOOP peptides[36–38], with whom they share a conserved S-x-S/T motif. SCOOPs are active as linear single-chain peptides. In order to convey this structural difference of importance to biological activity, we tentatively termed the newly found heterodimeric peptides as cSCOOPs ("*c" for cyclic/cystine-bridged; pronounced "cee-scoop"*) and the purified representative as cSCOOP1. *BLAST* searches in the *NCBI* database were performed, using representative prepro-peptide sequences of linear SCOOPs and cSCOOPs from *A. thaliana* as queries. Significant matches were found exclusively in species of the *Brassicaceae* family, and always with multiple members of both linear SCOOPs and cSCOOPs. A comparison of 172 linear

SCOOPs and 65 chain-1 peptides of cSCOOPs showed overall similarity in the N-terminal 8 amino acids of the peptide motif, including the two conserved serines. Linear SCOOPs and cSCOOPs displayed two characteristic differences in the C-terminal 6 residues (*Fig.6B*): (1) Linear SCOOPs lack the cysteine at position 9 that is indicative of cSCOOPs; the absence of cysteines is a characteristic of linear SCOOPs also proposed by the original study[36] that described the superfamily. And instead of a conserved glycine residue, (2) they carry a positive charge at position 10 (*i.e.* a basic side chain of arginine / histidine / lysine, or the zwitterionic side-chain of glutamine). An exception to this rule are linear SCOOP-like peptides of the EWR1-family[79,80] (synonymous: SCOOP24-28), which display a proline residue at position 10 and - perhaps because of this - low biological activity[80] when compared to other linear SCOOPs, like SCOOP12.

It would be tempting to speculate that cSCOOPs had evolved a heterodimeric structure requiring a linked chain-2, in order to provide chain-1 with the needed positive charge that is already incorporated into the sequence of linear SCOOPs at position 10. Tentative evidence suggested exactly that to be the case (*Suppl.Fig.3*). And precursors of SCOOP15, SCOOP18, and SCOOP19[38], may provide the missing evolutionary link between linear SCOOPs and cSCOOPs: Like linear SCOOPs, they carry a positive charge at position 10. Yet in addition they display a cSCOOP-like structure of two disulfide-bridged peptide chains, with the marked distinction that the cysteine on chain-1 is moved to a different position. It remains to be evaluated in detail whether SCOOP15 / 18 / 19 share structure-based activity requirements with linear SCOOPs or cSCOOPs, and therefore qualify as one, or the other class of SCOOP peptides.

Linear SCOOP peptides have been shown to be perceived by the LRR-RLK MIK2 (*UniProt* ID: Q8VZG8; encoded by *AGI* locus ID: At4g08850). Much like the linear SCOOP12, used as a reference, synthetic cSCOOP1 induced ethylene biosynthesis and extracellular ROS burst in leaves of wild type *A. thaliana*, but not in *mik2* knockout mutants, even when used at peptide concentrations well beyond physiological relevance (*Fig.7A-B*). This indicates that cSCOOPs also are perceived by MIK2.

Nicotiana benthamiana, a species of the *Solanaceae* plant family, encodes neither the MIK2 receptor nor any prepro-peptides of SCOOPs or cSCOOPs. The species is thus insensitive to both cSCOOP1 and the protease $SLPA^{Xap}$.

We used these plants to reconstitute a minimal SCOOP / MIK2 perception system by transient co-expression of *prepro-cSCOOP1* together with *MIK2*. Indeed, when assayed for induction of ethylene biosynthesis, transformed plants gained responsiveness to cSCOOP1 and, importantly, also to treatment with purified protease $SLPA^{Xap}$ (*Fig.7C*). This implies the possibility to transfer MIK2 together with prepro-(c)SCOOPs as a pathogen protease perception system to agriculturally relevant plant species outside of the *Brassicaceae* family.

Wild type *A. thaliana* is not a host of *Xanthomonas arboricola* pv. pruni (Xap); if at all, this apoplastic pathogen grows very poorly in this plant species (*Fig.7D*). However, in the absence of either the bacteria's protease $SLPA^{Xap}$, or the host's immune receptor MIK2, the pathogen was capable to grow and establish infection. On one hand, this shows that $SLPA^{Xap}$, assumed to function as a virulence factor in real hosts, does betray the pathogen and reports its presence to the plant immune system. On the other hand, this results shows that MIK2 is a major resistance factor that keeps *A. thaliana* a non-host for Xap.

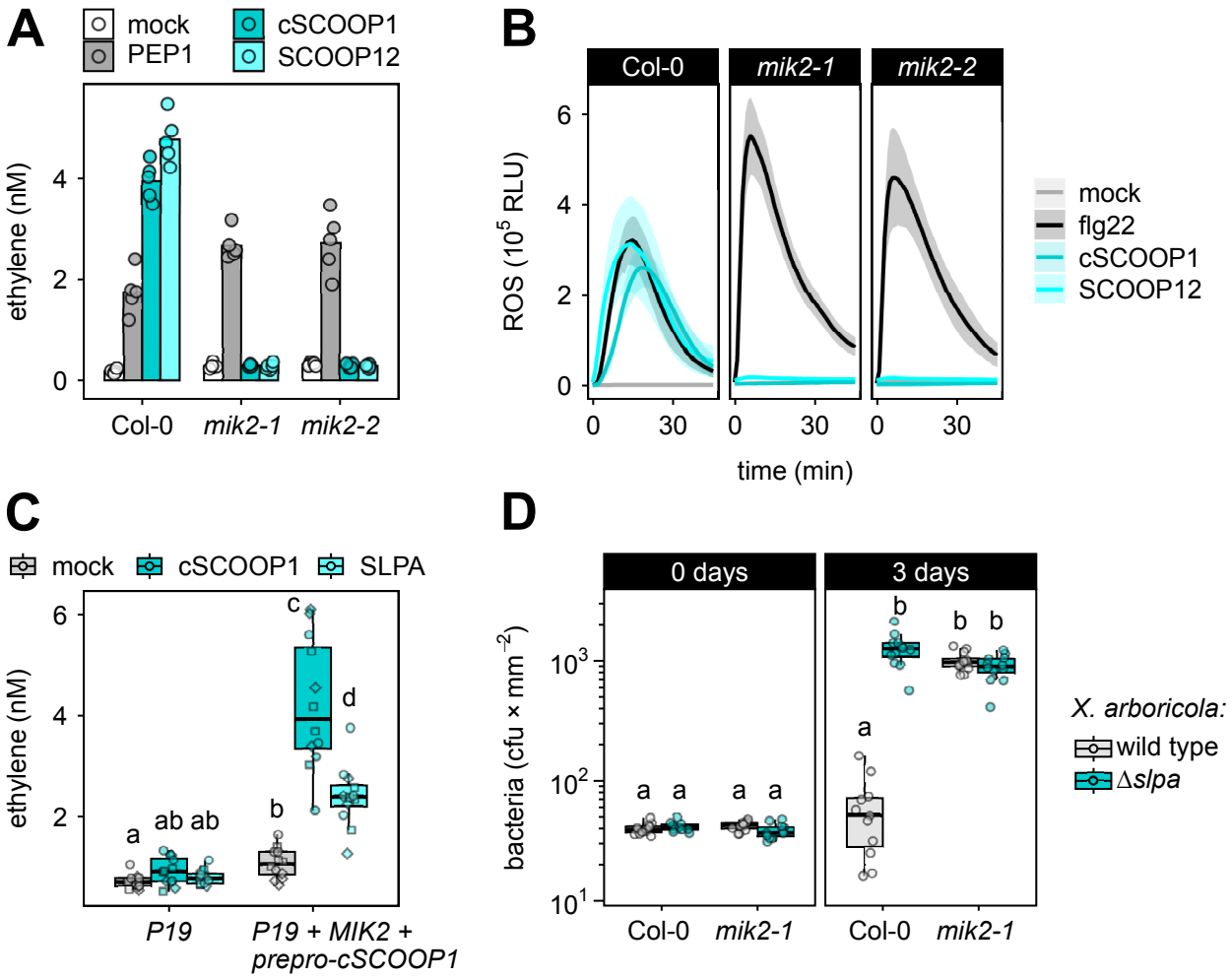


Fig.7: cSCOOP1 and SLPA perception depend on *MIK2*, and are sufficient to suppress growth of *X. arboricola* pv. pruni in *A. thaliana*. (**A**) Induction of ethylene biosynthesis in *A. thaliana* (Col-0) and two independent *mik2* knockout mutant lines in response to treatment with 1 µM of indicated peptides; bars indicate arithmetic means. (**B**) Induction of apoplastic ROS burst by 100 nM of indicated peptides; plotted lines depict arithmetic means and ribbons standard deviation of means, respectively. Each experiment in (A) and (B) was performed at least twice, with 5 replicates each, and with similar results. (**C**) Induction of ethylene biosynthesis in transiently transformed *N. benthamiana* in response to 100 nM cSCOOP1 and 3 nM SLPA^{Xap}; leaves were transformed with either silencing suppressor *P19* alone, or with a combination of *P19* and recombinant *MIK2:GFP* and *prepro-cSCOOP1:HiBiT* under control of the *AtFLS2* and *LjUBQ1* promoter, respectively. Shown are pooled data from 3 independent experiments (indicated by distinct symbols; dots, squares, and diamonds), with 4 replicates each. Experimental date was not a major predictor of ethylene response, while the combination of treatment and plant transformation genotype was (statistical analysis performed on $\log_{10}$-transformed data; *2-way ANOVA*: $F_{5,66} = 84$, $P < 2 \times 10^{-16}$, $R^2_{adj} = 0.85$; post-hoc: *Tukey's HSD* with $P_{adj} < 8 \times 10^{-3}$). (**D**) Growth of wild type and *Δslpa* deletion mutant strains of *X. arboricola* pv. pruni in inoculated leaves of *A. thaliana*. Performance of Xap strains was significantly affected in a time-, *MIK2*- and *slpa*-dependent manner (statistical analysis performed on $\log_{10}$-transformed data; *3-way ANOVA*: $F_{7,88} = 286$, $P < 2 \times 10^{-16}$, $R^2_{adj} = 0.95$; post-hoc: *Tukey's HSD* with $P_{adj} < 4 \times 10^{-10}$). Bacterial growth experiments were performed thrice, each time with 4 to 12 biological replicates, and with similar results. Non-shared lower case letters in (C) and (D) indicate groups with significant differences ($P_{adj} < 0.05$).

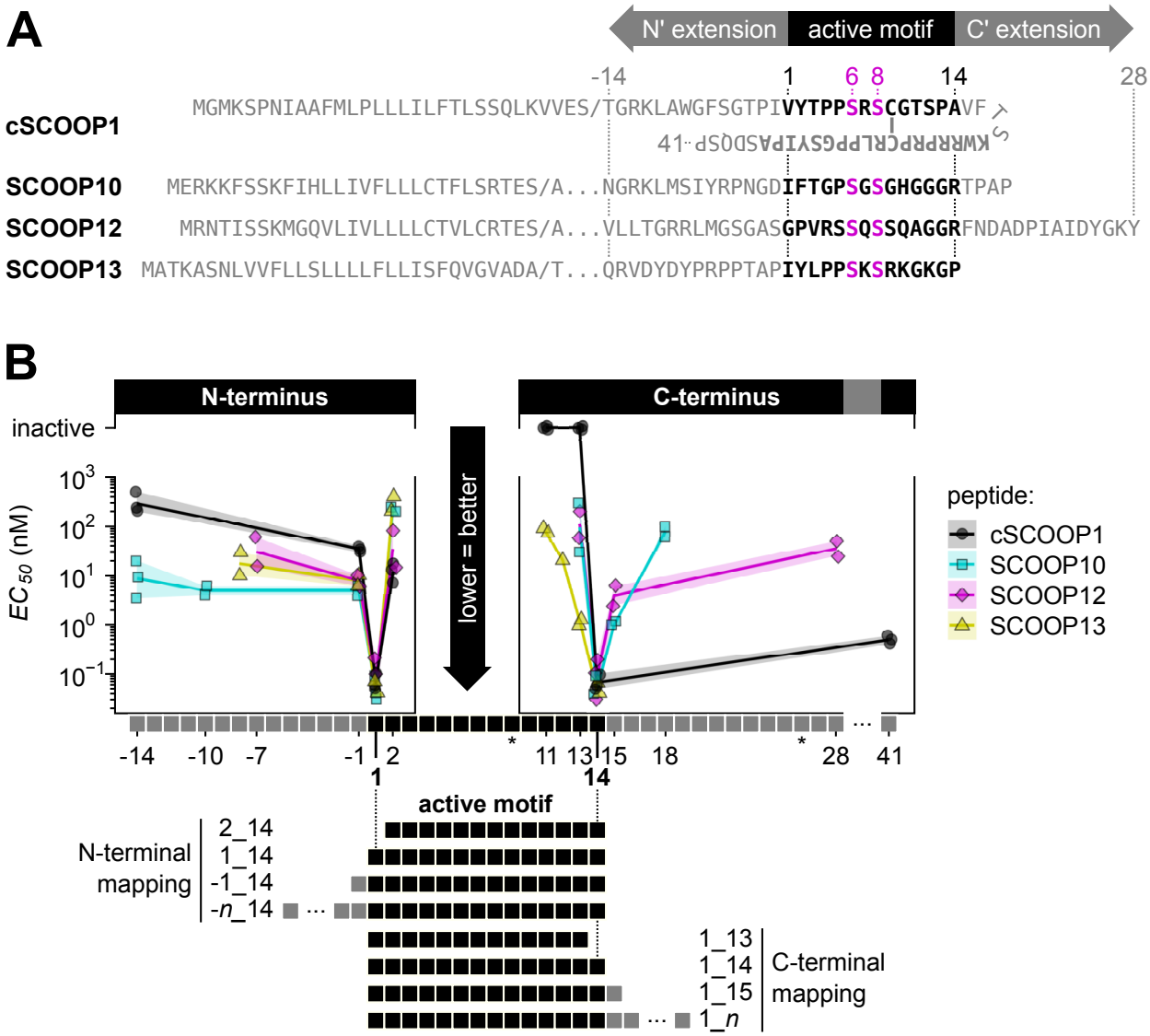


Fig.8: SCOOP peptides share a common active peptide motif of 14 aa length, whose extension / truncation affects their biological activity. (**A**) Alignment of prepro-peptides of cSCOOP1 and 3 representatives of linear SCOOPs according to their active motif (highlighted in black bold letters, the characteristic *S-x-S* motif is depicted in magenta). Slashes (/) separate the N-terminal signal peptide from the remaining pro-peptide. Dots (…) indicate irrelevant sequence stretches within the pro-domains, not shown due to limitations in figure dimensions; these are 36 aa, 7 aa, and 9 aa in length for the precursors of SCOOP10, SCOOP12 and SCOOP13, respectively. N- and C- terminal extensions relative to the active motif are indicated by numbers. For cSCOOP1, the active motif constitutes the highlighted 14 aa of chain-1, disulfide-bridged to its mature chain-2; indicated extension / truncation at either terminus refers to chain-1 only. (**B**) Mapping of the active motif: The specific immunogenic activity of cSCOOP1 and linear SCOOPs with N- and C-terminal extension / truncation of varying lengths was assessed in dose-response experiments with *A. thaliana* cell suspensions cultures, measuring extracellular alkalization at 15 min post peptide treatment. Each data point represents an EC_{50} estimate from an independent experiment, except for data with the active motif (1_14) of each peptide, which are shared between sub-plots of the N-terminal and C-terminal mappings. Lines and ribbons indicate geometric means and total data spread of 2-3 repetitions, respectively. The position of the disulfide-bridge in cSCOOP1 is indicated by asterisks (*).

Seeking confirmation that the 14 aa long sequence of cSCOOP1's chain-1, as purified, represented the mature peptide, we decided to map the active motif of cSCOOP1 and several representative of linear SCOOPs (*Fig.8A*). Synthetic peptide libraries were generated in such a way that either the N-, or C-terminus of the peptides was fixed to the corresponding terminus in cSCOOP1's chain-1. The other terminus was either truncated, or extended into the putative pro-domain of the respective SCOOP pro-peptide. Going forward we will refer to the motif of the different peptide versions according to the position of terminal amino acids relative to cSCOOP1's chain-1: In the notation of "*peptide (N_C)", N < 0* and *C > 14* indicate N- and C-terminal extension, respectively, while *N > 1* and *C < 14* denote the corresponding truncation.

Synthetic cSCOOP1, SCOOP10, SCOOP12 and SCOOP13 peptides were highly active when used in a motif matching chain-1 of the purified cSCOOP1 (1_14). They displayed similar EC_{50} values within a range of 44 - 84 pM (*Fig.8B*). N-terminal extension (-1_14) or truncation (2_14) by just a single amino acid uniformly resulted in an increase of EC_{50} values by 2 to 4 orders of magnitude; this corresponds to a loss of > 98.9% of immunogenic activity. Further extension of the N-terminus had only minor additional effects.

C-terminal truncation by a single amino acid (1_13) completely abolished the activity of cSCOOP1 and lead to a loss of > 99.9% of activity for SCOOP10 and SCOOP12. SCOOP13 was somewhat resistant to C-terminal truncation; loss of one (1_13) and two (1_12) C-terminal amino acids abolished 96.2% and 99.8% of activity, respectively. C-terminal extension of SCOOP10 and SCOOP12, by either a single amino acid (1_15), or all the way to the C-terminus of the respective pro-peptide (1_18 / 1_28), caused loss a of > 95.2% and > 99.7% of activity, respectively. cSCOOP1 differed from these linear SCOOPs in the way that an extension to the C-terminus of the pro-peptide (1_41) was tolerated without major loss of activity (EC_{50} of 493 pM).

In summary, linear SCOOPs and chain-1 of cSCOOPs share a common active motif of 14 aa length. This active motif is positioned in such a way that the characteristic serine residues are the 6th and 8th amino acid of the peptide sequence. As any change in N-terminal length renders these peptides all but inactive, it is apparent that precise N-terminal processing of pro-peptides is required in order to generate the mature immunogenic peptides. Additional C-terminal processing appears to be required in some cases, depending on the structure of the respective pro-peptide: Linear SCOOPs, located *within* the pro-peptide, require precise C-terminal processing for activity. While those, whose active peptide motif is already located at the C-terminus of pro-peptides (*e.g.* SCOOP13), obviously would not require such events. The active motif of cSCOOPs is located on their chain-1 and always positioned *within* their pro-peptide. Digest of enriched pro-cSCOOP1 by SLPA^{Xap} resulted in the purification of a heterodimeric peptide, whose chain-1 sequence coincided exactly with the active motif. This suggests that SLPA^{Xap} processes pro-cSCOOP1, both N- and C-terminally of chain-1, precisely to this active motif. Yet our activity mapping with synthetic peptides implies that C-terminal processing may not be required for activation of cSCOOP1 after all. A mechanistic explanation for the difference in C-terminal processing requirements between linear SCOOPs and cSCOOPs may be reached by considering the latter's cyclic structure. The disulfide-bridge of cSCOOPs should cause their C-terminus to be tucked away, next to chain-1 on the same peptide loop. In this conformation, an extended terminus may not interfere with the active motif's (co-) receptor interactions. But due to their different structure, linear SCOOPs may require precise C-terminal processing, lest the extended C-terminus protrudes and interferes with either receptor binding, or co-receptor recruitment.

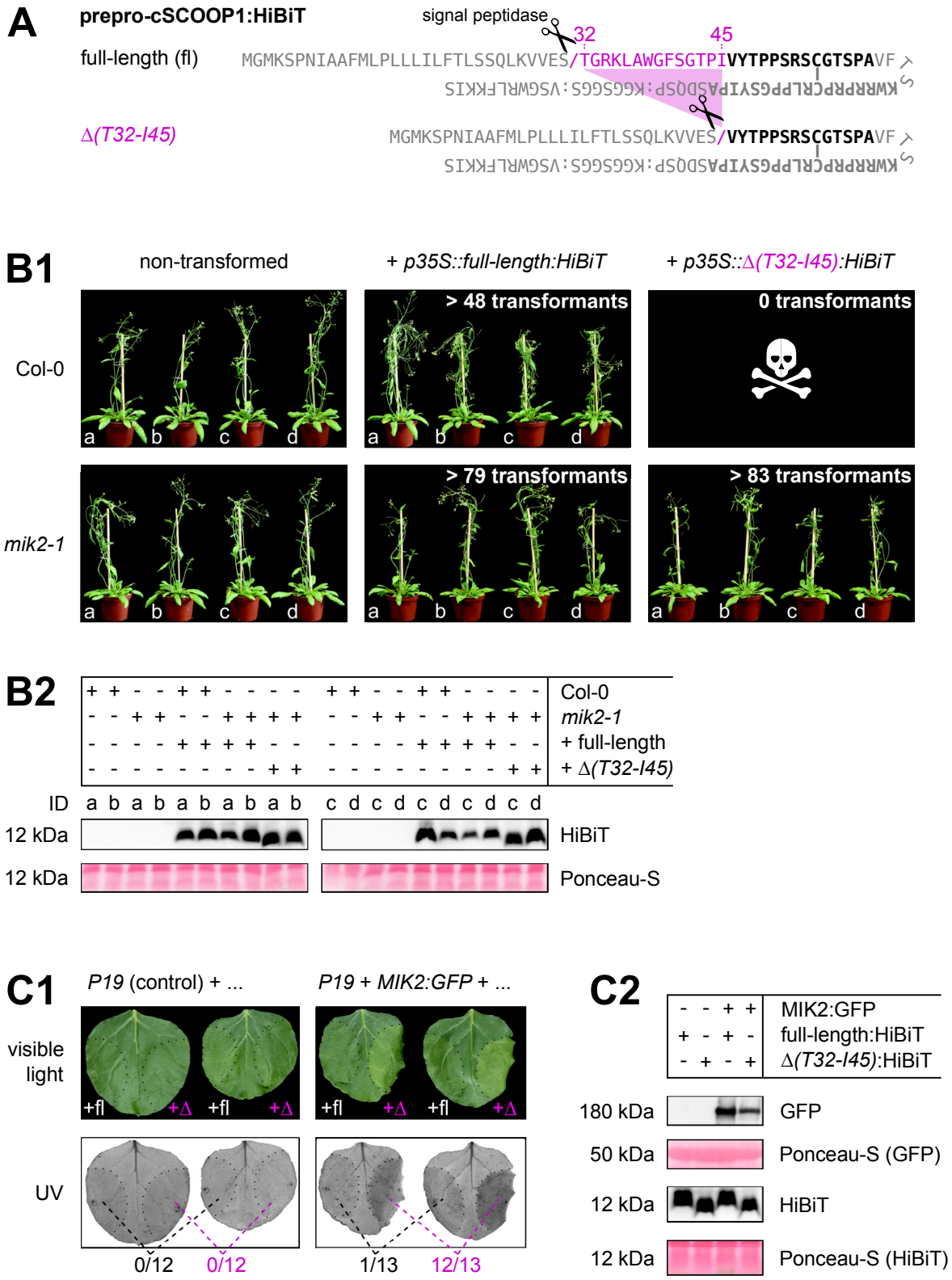


Fig.9: The pro-domain of the cSCOOP1 precursor prevents auto-immunity *in planta*, and its deletion is sufficient to cause weak cell death in a *MIK2*-dependent manner. (**A**) Construct design: Depicted are products of recombinant *prepro-cSCOOP1* expression constructs, C-terminally tagged with a short linker (7 aa) and an HiBiT peptide tag (11 aa). Slashes (/) indicate predicted

processing sites of signal peptidases, C-terminal of which is the signal peptide. Deletion of the N-terminal pro-domain is highlighted with magenta color, while the active motif is highlighted in black bold letters. Col-0 and *mik2-1* plants were transformed with constructs encoding a glufosinate resistance gene and either full-length, or deletion-mutant versions of *prepro-cSCOOP1:HiBiT* under control of the *CaMV35S* promoter. (**B1**) Composite images of 6-weeks old plants from the non-transformed parental lines and glufosinate-resistant transformants (T1 generation). Insets indicate the total number of transformants that were obtained from 3 independent batches of parallel transformations, each with ca. 20 plants of the respective parental lines. No transformants were obtained with the *Δ(T32-I45)* mutant version of the construct in Col-0 background, although threefold the amount of T0 seeds (*i.e.* > 160 000) were sown and subjected to selection. (**B2**) Representative individuals shown in (B1) were selected based on comparable, high level of transgene expression; the respective plants are identified by lower case letters (IDs). (**C1**) Leaves of transiently transformed *N. benthamiana,* co-expressing above-mentioned constructs together with either *P19*, or a combination of *P19* and *MIK2:GFP*, the latter under control of the *CmYLCV* promoter. Chlorosis and cell-death phenotypes were imaged 6 days after transformation under visible light and UV light (Cy5 filter; gray-scale inverted), respectively. Two representative leaves are shown for each construct combination. The experiment was performed 3 times independently, with both *prepro-cSCOOP1:HiBiT* versions under control of two distinct promoters (*CaMV35S* and *LjUBQ1*), and with a total of 12-13 replicates. For each construct combination, fractions indicate the relative frequency with which cell death of a severity comparable to, or stronger than the shown combination of *Δ(T32-I45) + MIK2* was observed. (**C2**) Expression control for construct combinations; pooled replicate samples from transformed leaf areas.

The superfamily of prepro-SCOOPs has originally been identified through data mining of the *A. thaliana* genome in search of genes encoding small secreted proteins related to the stress-induced transcript of the founding family member, *prepro-SCOOP12*[36]. As host-secreted precursors of danger signals, SCOOP peptides have been proposed to act as phyto-cytokines that modulate innate immune responses *in planta*[36–38,80].

Over-expression of recombinant prepro-peptides of several linear SCOOPs and cSCOOPs (including the precursor of cSCOOP1; synonymous: *STMP6*) with C-terminal *β*-glucoronidase (GUS) fusion[77], or with C-terminal GFP tag[78], was reported to result in dwarfed plants. Similarly, over-expression of *prepro-SCOOP12:GFP* allegedly causes dwarfism in a *MIK2*-dependent manner (personal communication with Prof. Dr. Cyril Zipfel and colleagues, whose manuscript is currently in review). Processing of pro-SCOOPs by secreted plant proteases of the subtilase family is believed to contribute to this artificial autoimmune phenotype (unpublished work by Prof. Dr. Cyril Zipfel, Prof. Dr. Andreas Schaller, and colleagues).

In our experiments, we obtained no evidence for the activation of pro-cSCOOP1 by plant proteases. We wondered whether over-expression of *prepro-cSCOOP1* at our hands would result in a similar "*autoimmune*" phenotype. Instead of a large protein (like GFP or GUS), whose size and distinct surface charge may interfere with secretion, localization or apoplastic stability of the small positively charged prepro-peptide, we opted for the small HiBiT tag as a C-terminal label. HiBiT is a peptide tag (11 aa), whose charge (*pI* of 11.6) is almost identical to cSCOOP1's chain-2 ; native and HiBiT-labeled pro-cSCOOP1 are very similar to each other with regard to their biochemical properties.

We used two constructs under control of the strong *CaMV35S* promoter; one encoding full length *prepro-cSCOOP1:HiBiT,* the other encoding *Δ(T32-I45)*, a mutant version in which the N-terminal pro-domain was deleted and the signal peptide directly joined to the N-terminus of the active peptide motif (*Fig.9A*). Proteolytic cleavage of the signal peptide of the *Δ(T32-I45)* version by

plant signal peptidases during co-translational secretion (predicted by *SignalP* vers. 6.0) was expected to result in a processing product that corresponded to a HiBiT-tagged version of the immunogenic cSCOOP1 (1_41) peptide. When transformed into wild type and *mik2-1* null-mutant plants of *A. thaliana*, the full length construct led to robust ectopic expression of the recombinant pro-cSCOOP1 protein (*Fig.9B2*), but without any effect on growth or apparent phenotype of the transformants (*Fig.9B1*). The plants were indistinguishable from the parental lines, whose genetic background did not seem to affect the transformation efficiency either.

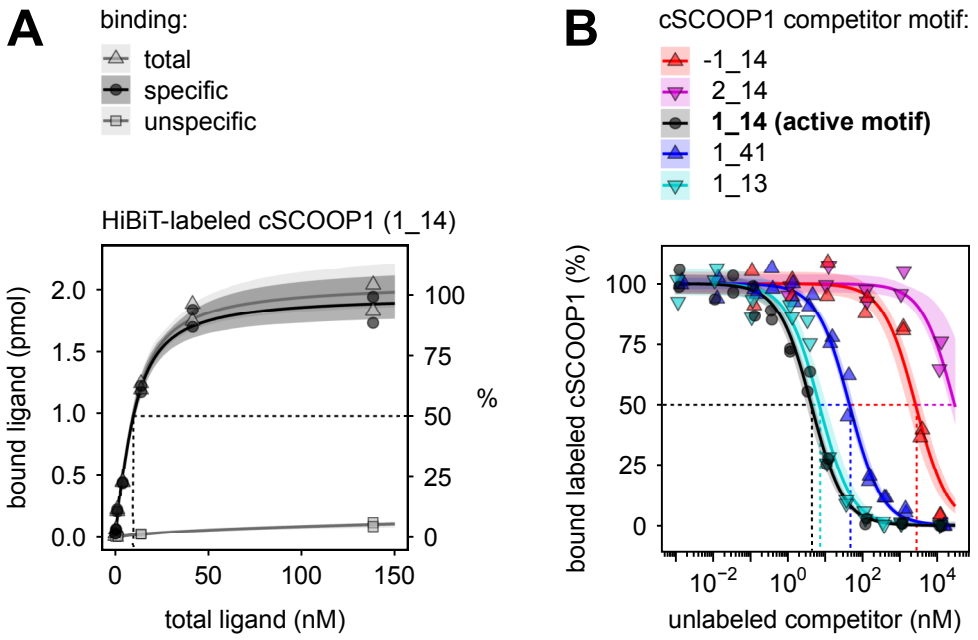
In contrast, despite several attempts, we did not obtain a single transformant with the truncated *Δ(T32-I45)* construct in wild type *A. thaliana* background, while transformation of *mik2-1* plants with the same construct resulted in phenotypically normal, well expressing plants. This suggested *MIK2*-dependent lethality of the construct.

If lethality were caused by the persistent *MIK2*-dependent perception of the auto-active product of the *Δ(T32-I45)* version of *prepro-cSCOOP1:HiBiT*, then it should be possible to observe associated autoimmune phenotypes in transiently transformed *N. benthamiana* leaves expressing both. Indeed, wrinkling of leaves, development of chlorosis, and red light autofluorescence, indicative of early cell death[81], were observed in a manner that depended on both, *MIK2* and the *Δ(T32-I45)* deletion in *prepro-cSCOOP1* (*Fig.9C1*). Co-expression of full-length *prepro-cSCOOP1* together with *MIK2* rarely caused these symptoms, and if so, then they were much less severe when compared to the mutant version. The effect could not be explained by a presumptive difference in protein accumulation, as both pro-peptide versions were expressed to similar levels (*Fig.9C2*). The experiment was repeated with different promoters and also with non-tagged versions of *prepro-cSCOOP1*, yet similar results were obtained (data not shown).

In conclusion, over-expression of wild type *prepro-cSCOOP1* failed to - , and deletion of residue 32 to 45 of the same prepro-peptide sufficed to cause chronic disease symptoms in a *MIK2*-dependent manner. It appears as if the function of the N-terminal pro-domain were to prevent autoimmunity and keep apoplastic pro-cSCOOP1 in an inactive state until activation by the pathogen protease $SLPA^{Xap}$, and perhaps its orthologs, occurred.

Having observed the pronounced effects of changes in terminal length on the peptide's immunogenic activity, it was hypothesized that these were the consequence of an altered binding affinity for MIK2. We addressed the affinity problem with a novel binding assay that utilized the particular detection properties of the HiBiT peptide tag (manuscript accepted for publication). A truncated version of cSCOOP1's chain-2 was N-terminally labeled with the HiBiT peptide and disulfide-bridged to the active motif of chain-1. The affinity of immobilized MIK2 for the HiBiT-labeled cSCOOP1 (1_14) peptide was estimated in saturation binding experiments (*Fig.10A*). The interaction displayed a K_d (dissociation constant) value of about 4.9 nM.

Knowledge of MIK2's affinity for the labeled ligand allowed for the subsequent estimation of its binding affinities for unlabeled cSCOOP1 analogs in competition binding assays (*Fig.10B*). MIK2 bound to cSCOOP1 (1_14) with a K_i (dissociation constant of the unlabeled competitor) of 1.6 nM. The difference in binding affinity between unlabeled and HiBiT-labeled cSCOOP1 (1_14) was in line with their immunogenic activity (EC_{50} of 1.1 nM for the latter; data not shown). N-terminal extension (-1_14) or truncation (2_14) of the active motif by a single amino acid decreased the binding affinity by 3 orders of magnitude, and more than 4 orders of magnitude, respectively.



C

cSCOOP1 motif	EC_{50} ± error (nM)	K_d ± error (nM)	P	R^2_{McF}	AIC
1_14, HiBiT-label on chain-2	9.9 ± 1.2	4.9 ± 0.7	3.4×10^{-6}	0.99	182

cSCOOP1 motif	IC_{50} ± error (nM)	K_i ± error (nM)	P	R^2_{McF}	AIC
-1_14	2612 ± 1096	1345 ± 778	3.6×10^{-2}	0.94	14
2_14	>20000	>10000	8.6×10^{-1}	0.82	2
1_14 (active motif)	**4.2 ± 0.6**	**1.6 ± 0.6**	$\mathbf{2.1 \times 10^{-7}}$	**0.99**	**-7**
1_41	44.8 ± 4.9	24.3 ± 7.1	1.0×10^{-8}	0.99	-10
1_13	7.0 ± 1.3	2.8 ± 1.1	2.8×10^{-5}	0.98	9

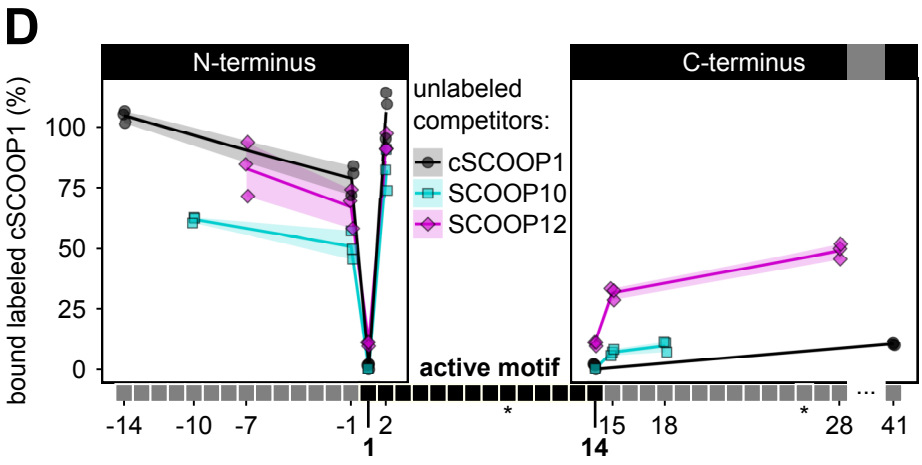


Fig.10: Precise processing of SCOOP peptides to the N-terminus of the shared active motif is required for high-affinity binding to MIK2. (**A**) Saturation kinetics for binding of HiBiT-labeled cSCOOP1 to 10 nM of MIK2. Total and unspecific binding were derived from binding curves obtained in the absence and presence of 10 µM of unlabeled cSCOOP1 competitor peptide, respectively.

Specific binding was modeled as the difference between total and unspecific binding. The experiment was independently repeated with variations in equilibration time and receptor concentration, yet almost identical results. (**B**) Competition binding experiments, in which binding of 150 pM of HiBiT-labeled cSCOOP1 to 3-5 nM of MIK2 was out-competed with increasing concentrations of unlabeled cSCOOP1 analogs; normalization to specific binding at minute competitor concentrations (100 % ≙ difference between upper and lower asymptote of total binding model). Lines and ribbons plotted in (A) and (B) indicate best-fit model functions and 95% *CI* (1 000 bootstraps), respectively. (**C**) Summary of model statistics and affinity constants ($K_d \triangleq K_i$) inferred from experiments shown in (A) and (B); further information on the models used for inference are described in the *Material & Methods* section. (**D**) Competition binding of 100 pM HiBiT-labeled cSCOOP1 and 10 µM of unlabeled SCOOP competitor peptides to MIK2; normalization (%) to total binding in the absence of competitor (*i.e.* mock treatment; not shown). Lines and ribbons indicate geometric means and total data spread of 3 biological replicates, respectively. Asterisks (*) mark the location of the disulfide-bridge in cSCOOP1. Data for the active motif (1_14) of each peptide are shared between sub-plots for the N-terminal and C-terminal mappings in (D). Competition binding experiments illustrated in (B) and (D) were independently repeated with selected peptides and similar results were obtained. Mapping strategy and motif nomenclature are the same as in *Fig. 8*.

And C-terminal extension, all the way to the terminus of the pro-peptide, only led to a mild decrease in binding affinity (K_i of 24.3 nM). The affinities discussed so far correlated well with the peptides' immunogenic activity. An exception was C-terminally truncated cSCOOP1 (1_13), which was completely inactive in *A. thaliana*, but bound to MIK2 with pretty much the same affinity (K_i of 2.8 nM) as the highly active cSCOOP1 (1_14).

Thus, the N-terminus and C-terminus of cSCOOP1 are important for receptor binding and activation, respectively. Similar reports exist for other peptide ligand / LRR-RLK immune receptor pairs[82,83], and the term "*address-message mechanism"*[84] has been coined to describe this phenomenon in general.

The high affinity of MIK2 for the active cSCOOP1 (1_14) peptide was in disagreement with previous reports, which claimed absurdly high K_d values in the micromolar range for the interaction of MIK2 with SCOOP12. It is worth to mention that these authors used all but inactive versions of SCOOP12, that is SCOOP12 (-1_14)[38] and SCOOP12 (2_14)[37]. Semi-quantitative competition binding experiments were performed in order to assess whether observations for cSCOOP1 held true also for other SCOOP peptides. Out-competition of HiBiT-labeled cSCOOP1 was attempted by addition of a large excess (100 000 ×) of unlabeled competitor peptide of interest. Complete out-competition of the labeled ligand was expected for high affinity competitors (*Fig.10D*). Observed effects of N- and C-terminal changes on binding mirrored the ones on immunogenic activity. With the exception of SCOOP10, which - just like cSCOOP1 - appeared to tolerate C-terminal extension without major loss of competition efficacy.

It was intriguing that C-terminally truncated cSCOOP1 (1_13) bound to MIK2 with high affinity (*Fig.10B-C*), yet simultaneously failed to activate the receptor (*Fig.8B*). Going back to the "*address-message"* concept, it was hypothesized that the altered C-terminus of the ligand may interfere with the formation of the receptor / co-receptor complex, and thus prevent the activation of downstream signaling. Indeed, C-terminally truncated cSCOOP1 (1_13) failed to recruit the co-receptor BAK1 into a stable complex with MIK2 in co-immunoprecipitation experiments (*Fig.11A*). Active cSCOOP1 (1_14), in contrast, facilitated this interaction *in planta*. It appears that the N-terminal portion of SCOOP peptides binds to the cognate receptor, while the outermost

C-terminus of the active motif provides a properly sized bolt for SERK-family co-receptors to latch onto.

We tested whether pretreatment with cSCOOP1 (1_13) could be used to block the ligand binding site on the receptor (*Fig.11D*) to inhibit subsequent perception of active SCOOP peptides.

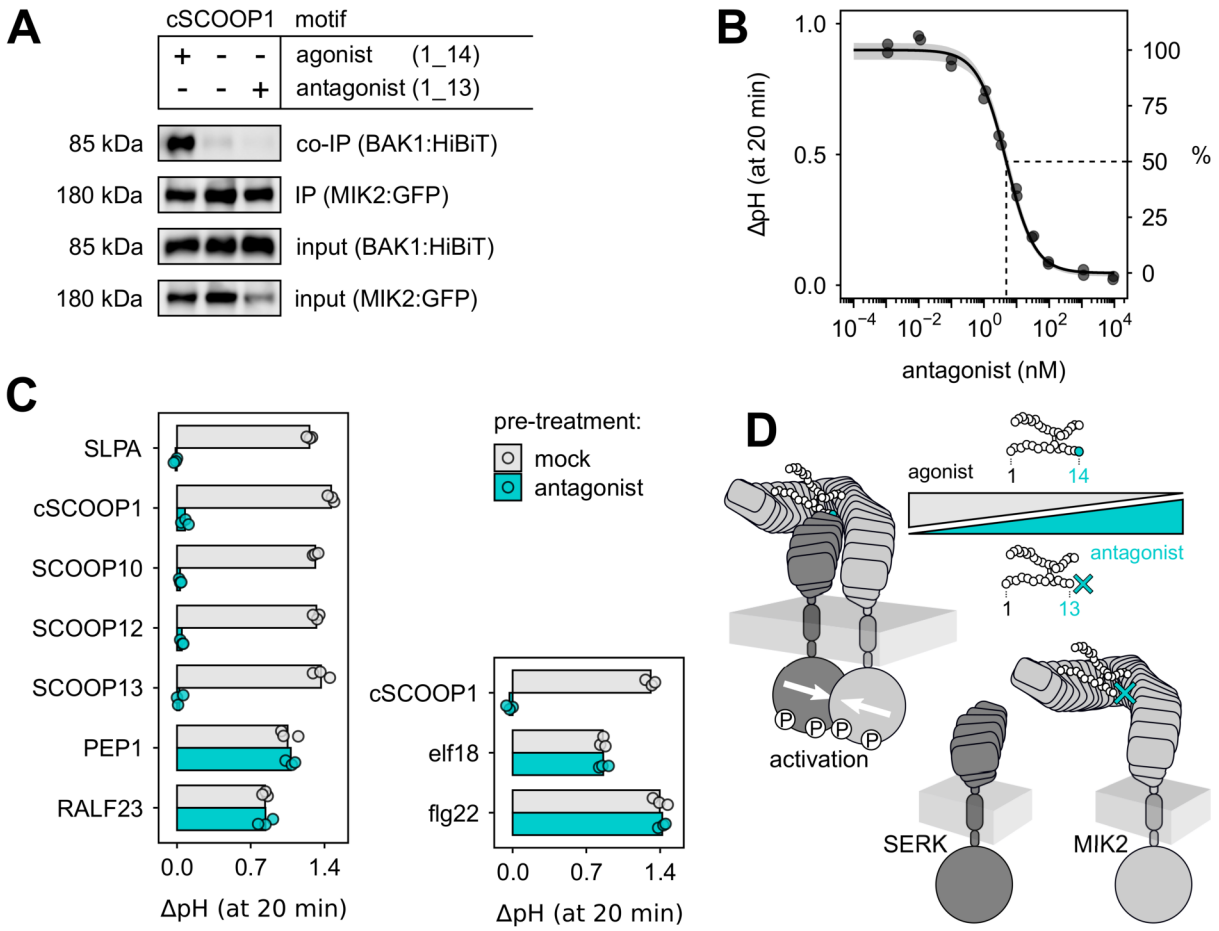


Fig.11: C-terminally truncated cSCOOP1 (1_13) is a potent and specific antagonist of MIK2. (**A**) Co-immunoprecipitation of BAK1 with MIK2 in the presence of 1 µM of cSCOOP1 peptides corresponding to the active motif (1_14), or a C-terminally truncated version (1_13) thereof. The recombinant proteins used for immunoprecipitation were co-expressed in transiently transformed *N. benthamiana*. (**B**) Extracellular alkalization of *A. thaliana* cell suspension cultures in response to 100 pM cSCOOP1 (1_14) was inhibited by pretreatment (5 min) with the antagonistic cSCOOP1 (1_13) in a dose-dependent manner ($IC_{50} = 4.8 \pm 0.4$ nM, $P < 3 \times 10^{-9}$, $R^2_{McF} = 0.99$); plotted line and ribbon indicate best-fit model function and 95% *CI* (1 000 bootstraps), respectively. (**C**) Specificity of the antagonist: *A. thaliana* cell suspension cultures were pretreated with either mock, or 1 µM of antagonistic cSCOOP1 (1_13), and checked for response to $SLPA^{Xap}$ and immunogenic peptides. Concentrations used were 300 pM for $SLPA^{Xap}$ and SCOOPs, 1-10 nM for elf18, flg22 and PEP1, and 10 µM for RALF23. Latter ligands had to be used at higher concentrations due to their lower activity. Bars depict arithmetic means of 3 replicates. Each of the experiments shown in (A) to (C) were independently repeated, with similar results. (**D**) Mechanistic model illustrating the antagonistic effect of cSCOOP1 (1_13) on MIK2-dependent immune responses: Both, agonistic and antagonistic SCOOP peptides bind to MIK2, but the antagonist fails to recruit SERK-family co-receptors into the complex with MIK2 that is required for activation.
Co-immunoprecipitation experiment shown in (A) was done by Dr. Yan Wang, who also provided vector-graphics templates for LRR-RLK structures (modified) shown in (D).

cSCOOP1 (1_13) was found to be a potent antagonist, inhibiting the response to 100 pM of active cSCOOP1 in a dose-dependent manner and with an IC_{50} of 4.8 nM (*Fig.11B*). The antagonistic effect was specific for MIK2. It inhibited responses to cSCOOPs, linear SCOOPs, and *Xanthomonas* SLPA, but did not affect the response to unrelated immunogenic peptides of plant and bacterial origin, whose perception is mediated by other receptors (*Fig.11C*).

Discovery of the antagonist provided us with a powerful biochemical tool to probe the receptor and search for evidence of immunogenic extracellular proteases in other pathogens besides *Xanthomonas*. A brief screen of microbes readily available in our lab demonstrated that extracellular fractions of pathogenic bacteria and fungi contained considerable amounts of heat-labile immunogenic activity of > 30 kDa in size, whose perception depended on non-inhibited MIK2 (*Fig.12A*). Biochemical properties of the immunogenic activity found in extracellular pathogen fractions were reminiscent of $SLPA^{Xap}$ and indicative of secreted proteases. Microbes with such activity included the obligate plant pathogen *Fusarium graminearum* and opportunistic pathogens of both plants and humans, *Pseudomonas aeruginosa* - PA14 and *Sarocladium strictum* (synonymous: *Acremonium strictum*). Crucially, V8 protease (synonymous: SspA; GluC), a secreted virulence factor[85] of the opportunistic human pathogen *Staphylococcus aureus*, did not induce an immune response.

The cell-free apoplastic fluid of *A. thaliana* cell suspension culture was used as a source of substrate(s) in *in vitro* digests with the extracellular pathogen protease fractions (*Fig.12B*). Native proteases of all pathogens, except *S. aureus* V8 protease, generated a heat-stable immunogenic activity, that was absent when proteases were heat-denatured prior to digest. MIK2-dependency was confirmed by antagonist pretreatment of *A. thaliana*, which abolished the activity and identified the immunogenic products as SCOOP peptides. The active product of *S. strictum* protease was sensitive - , and products of both, *F. graminearum* and *P. aeruginosa* proteases were resistant to treatment with reducing agents, characterizing them as a cSCOOP and linear SCOOPs, respectively.

Protease fractions were used in *in vitro* digests with synthetic analogs of pro-SCOOPs, in order to identify their respective substrate. The selection of candidate substrates was guided by educated guesses based on the biochemical properties of the respective peptide products, proteolytic properties of protease fractions, and preliminary screens. The digest of N-terminally extended cSCOOP1 (-14_14) with *X. arboricola* pv. pruni SLPA served as a positive control. $SLPA^{Xap}$ processed the pro-cSCOOP1 analog and released the active cSCOOP1 (1_14) peptide in a manner that depended on its native state (*Fig.12C*). Native protease-dependent generation of the mature active peptide motif was confirmed by mass spectrometric analysis (*Fig.12D*). $SLPA^{Xap}$ was substrate-specific; digests with none of the other pro-peptide analogs resulted in the generation of active products. With the exception of N-terminally extended cSCOOP5 (-14_14), whose digest showed negligible activity, suggesting low processing efficiency.

This pro-cSCOOP5 analog, in turn, was efficiently processed into the active cSCOOP5 (1_14) product by *S. strictum* protease fraction only. And *F. graminearum* and *P. aeruginosa* proteases specifically generated active SCOOP10 (1_14) and SCOOP13 (1_14) peptides from C-terminally extended SCOOP10 (1_18) and N-terminally extended SCOOP13 (-8_14) pro-peptide analogs, respectively.

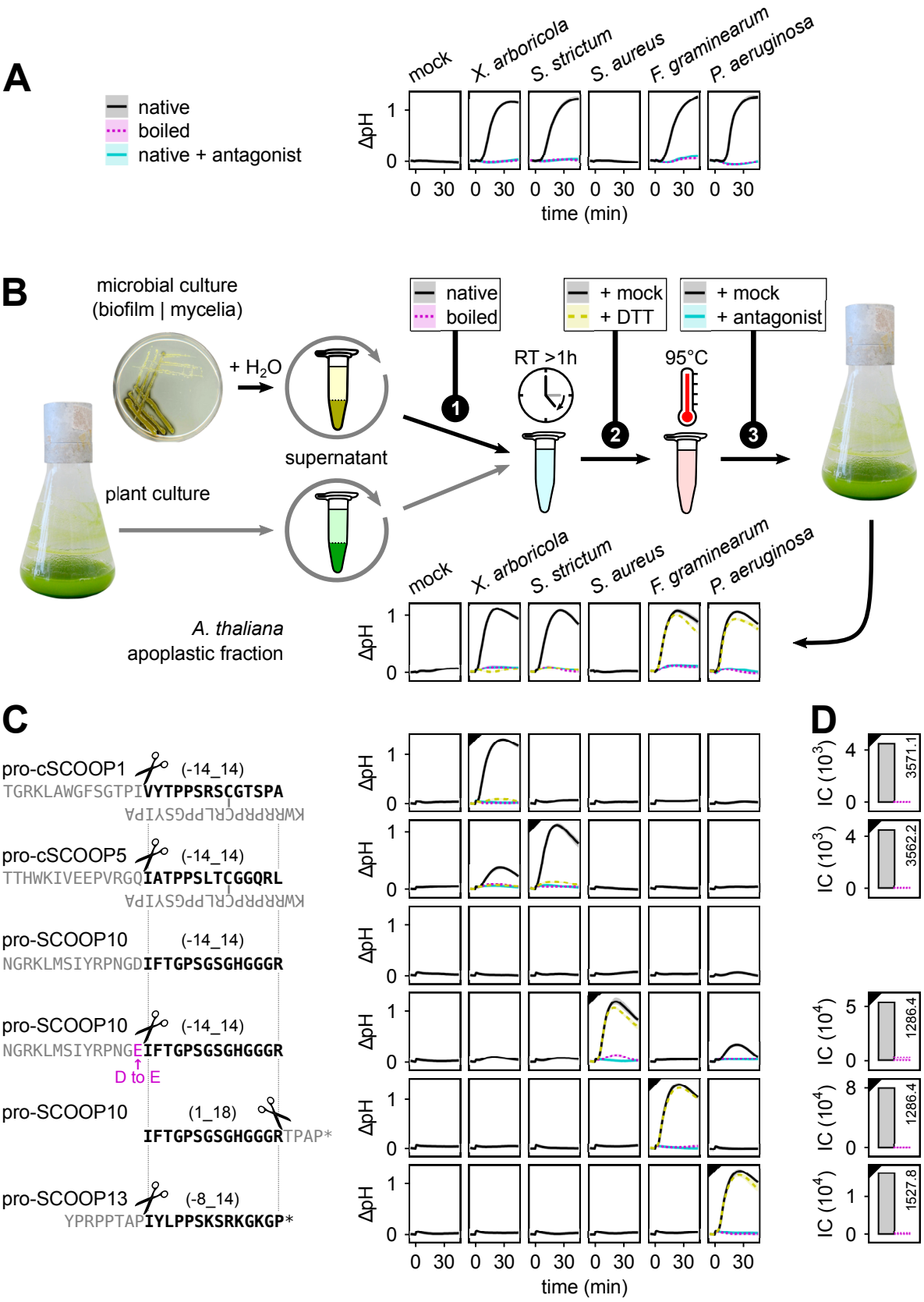


Fig.12: Unrelated pathogen proteases are perceived *via* proteolytic activation of distinct pro-SCOOP precursors. (**A**) Extracellular protease fractions of indicated microbes were used either in

their native state, or after heat-denaturation, to treat *A. thaliana* cell suspension cultures. Protease fractions were prepared as ultra-filtrates of extracellular microbial washes and used at final concentrations of 0.2-2% (w/v) of microbial biomass per cell culture volume, except for *S. aureus* V8 protease [100 nM], which was commercially available. Cell cultures were pretreated with either mock, or 1 µM of antagonistic cSCOOP1 (1_13) to assess MIK2's involvement in perception. (**B**) [1] Native or heat-denatured protease fractions [6% (w/v) of microbial biomass for ultra-filtrates, or 1 µM for V8] were used to digest the cell-free apoplastic fraction of *A. thaliana* cell suspension cultures for 1 h at RT. Digests were subsequently boiled for 20 min to inactive proteases, and used at 10% (v/v) to induce naive or antagonist pretreated cell cultures. [2] Boiling of digests was done in the presence and absence of 10 mM DTT, in order to assess the active peptide product's sensitivity to reducing agents. [3] MIK2 dependency was confirmed by antagonist treatment. (**C**) Digests as in (B), but done with synthetic pro-peptide analogs of selected SCOOPs as substrates in place of plant apoplastic fraction. Products of digests were used in final concentrations equivalent to 0.1-1 nM of substrates. Pro-peptides [100 µM in 50 mM MES buffer (pH 5.6), 100 mM NaCl, 10 mM CaCl2] were digested for 1-3 h with following protease fractions: 10 nM of $SLPA^{Xap}$, 1 µM of *S. aureus* V8, 4-8% (w/v) of microbial biomass for ultra-filtrates of all others. Asterisks in peptide sequences indicate free C-termini of native pro-peptides; the active motif is highlighted with black bold letters. Lines and ribbons plotted in (A) to (C) indicate arithmetic mean and total data spread of 3 replicates, respectively. Each experiment was done thrice, with similar results. (**D**) Mass spectrometric confirmation of native protease-dependent generation of expected active peptide products: The ion count (IC) of masses corresponding to the respective active peptide motif (heterodimers in case of cSCOOPs) was measured for each digest of interest (indicated by black triangles in upper left corner of plots). Digests with native and heat-denatured proteases are depicted with bars of gray and magenta color, respectively; insets indicate average peptide masses queried.
The mass spectrometric quantification of peptide products was done by Dr. Mark Stahl.

Looking at the protease-dependent activation of distinct SCOOP peptide precursors in a strict pro-SCOOP-for-pathogen protease-specific manner, it appears as if the plant secreted distinct decoy substrates for extracellular pathogen proteases for the purpose of the latter's perception. Apoplastic pathogen proteases may inadvertently process these decoy substrates instead of their presumed virulence targets. Precise processing of a decoy by a specific pathogen protease were to release the active danger-signaling SCOOP peptide as a proxy for the presence of the attacking pathogen itself.

Evidence presented strongly suggests that pro-cSCOOP1 is the cognate decoy substrate for *Xanthomonas arboricola* pv. pruni protease SLPA. N-terminal processing of pro-cSCOOP1 was sufficient for activation. And the extension of the concept to other cSCOOPs, like cSCOOP5, seems reasonable. Hence, pro-cSCOOP5 may be the cognate decoy substrate for *S. strictum* protease, which processes the pro-peptide N-terminal of the active peptide motif. An ideal candidate for the cognate decoy substrate for *P. aeruginosa* protease was found with pro-SCOOP13, whose active peptide motif is located at the C-terminus of the pro-peptide. N-terminal processing of this precursor, as done by the pathogen's protease, is sufficient for activation. The active peptide motif of SCOOP10, in contrast, is located within the pro-peptide. Precise processing at both termini is required for the generation of the active peptide. *F. graminearum* protease was able to process the synthetic pro-SCOOP10 analog C-terminal of the active motif only, which makes this pro-peptide a tentative decoy substrate for this pathogen protease. It is unclear, whether the additionally required N-terminal processing were to occur through action of plant-secreted proteases under certain conditions, or whether a different linear pro-SCOOP is the true cognate decoy substrate for *F. graminearum* protease.

Absence of immunogenic activity with *S. aureus* V8 protease suggested that the apoplast of

A. thaliana lacked a cognate decoy substrate for this protease. Which made perfectly sense, considering that this pathogen does not colonize plants; *A. thaliana* wouldn't have experienced forces of selection that were to favor the evolution of a decoy substrate for *S. aureus* protease(s). This made us wonder: Is it possible to design a decoy substrate for *S. aureus* V8 protease and thereby simulate the co-evolution of plant pro-SCOOPs with emerging pathogens *in vitro*? Yes, it was: The substitution of a single amino acid residue, directly N-terminal of the active peptide motif of the pro-SCOOP10 analog, from aspartate to glutamate, resulted in a gain of specific perception of *S. aureus* V8 protease in *A. thaliana* cell cultures (*Fig.12C*).

This not only substantiated our decoy substrate model of extracellular pathogen protease perception and SCOOP precursor activation, but also indicated yet unrealized potential for crop plant protection. The design of novel protease perception capabilities of the plant immune system is possible.

4 Discussion

4.1 Extracellular pathogen proteases and their perception

Orthologs of *Xanthomonas* SLPA protease, encoded by pathogens of animals and humans, have been suggested to be virulence factors. They exert their function either through suppression of host immune responses[66], through induction of necrotic lesions and aid in invasion of tissues[68,71], or through functions as maturases that proteolytically activate other virulence factor proteins[62,64,70,76,86]. The extracellular maturation of host-manipulating effector proteins is thought to be required for transport, and for the prevention of harmful effects to the secreting bacteria themselves. The high degree of conservation of SLPA homologs in all *Xanthomonas* species implies an important function for these plant pathogenic bacteria. Yet, the absence of the protease in the *Δslpa* strain did not affect establishment of infection in *mik2* mutants of *A. thaliana.* However, this experiment, performed under laboratory conditions with injection of Xap strains directly into leaves, does not rule out a function of $SLPA^{Xap}$ in colonization of hosts under natural conditions.

Interestingly, SLPA orthologs are also found in strains of *Xanthomonas* that are adapted to growth in plants of the *Brassicaceae* family, including pathovars of *X. campestris* pv. campestris and *X. campestris* pv. raphani. However, the MIK2-dependent immunogenic activity in biofilm washes of these specialists was much lower (up to 30-fold) when compared to same preparations of non-adapted strains, like *X. arboricola* pv. pruni and *X. axonopodis* pv. citrumelo (data not shown). It remains to be shown whether this loss of immunogenicity is due to altered expression or a changed substrate specificity of SLPAs in *Brassicaceae*-specialized strains, respectively.

A broad body of evidence documents that intracellular plant immune receptors can detect proteases that pathogens inject as effector proteins into the cytoplasm of their plant hosts. The importance of these proteases for pathogen virulence and plant resistance has been well established in the context of so-called effector-triggered immunity (ETI)[52]. In contrast, the immunogenic potential of pathogen proteases released into the apoplastic interphase between the attacker and the host has been reported only in few instances, and their cognate receptors in plants have remained unknown: This includes trypsin-like proteases (*i.e.* Pat-1 / Chp-7 / ChpG) of the Gram-positive bacterium *Clavibacter* spp., which infects plants of the *Solanaceae* family. Their immune perception by resistant host species is dependent on the proteases' catalytic activity[87–90], but neither the immunogenic product nor the corresponding immune receptor has been identified so far. Not the catalytic activity, but the pro-domain of a subtilase secreted by *Rhizoctonia solani* seems involved in the infection of cereals[91] (*Poaceae* plant family). Another example is the subtilisin-like protease AsES of the fungal pathogen *S. strictum*, which is perceived in species of various plant families, including *A. thaliana*[92], *N. benthamiana*[93], and strawberry[94] (*Rosaceae*). Immunogenicity of AsES in *Arabidopsis* is independent of its catalytic activity[92], indicating the perception of a structural epitope of the protein itself. The catalytic function of the secreted trypsin-like serine proteases LysC (synonymous: protease IV; PrpL) of *P. aeruginosa* and ArgC of *X. campestris* has been reported to induce typical immune responses in *A. thaliana*[95]. But their substrates, and the receptor involved in the perception have not yet been identified. LysC and ArgC are site-specific proteases. Their processing, specifically C-terminal of lysine and arginine residues, suggests that the perception of these proteases may involve the activation of

pro-SCOOPs whose active peptide motif is flanked by -, but does not contain the respective protease cleavage sites. Possible candidates for the perception of *P. aeruginosa* LysC are pro-SCOOP23, pro-SCOOP18, and pro-SCOOP19. In turn, *X. campestris* ArgC might activate pro-SCOOP10 or pro-SCOOP11, provided further N-terminal processing by plant proteases were to occur simultaneously.

If direct perception of pathogen proteases through their molecular structure is possible, then why have plants evolved strategies to recognize proteases *via* their proteolytic activity on host substrates? Sensitivity of immune perception may serve as an explanation. Protease are efficient enzymes and the tiny amounts necessary for their function as virulence factors may be insufficient for their structural detection by host immune receptors. Enzyme activity, however, can (inadvertently) generate many products per second, thus providing a potent amplification of potential ligands that the host may utilize for their perception. Such a signal amplification step can explain the exquisite sensitivity of *A. thaliana* cells for $SLPA^{Xap}$ down to 1 pM of the enzyme, while mature SCOOP peptide products stimulate the MIK2 receptor “only” at concentrations in the mid-picomolar range.

The detection of extracellular pathogen proteases through their catalytic action on host substrates is a complex, yet beautifully elegant solution to the problem that is the evolution of efficient immune systems.

Proteolytic activation of pro-SCOOPs through pathogen proteases is highly reminiscent of reports describing a similar mechanism in the activation of a proinflammatory cytokine in humans and mice. A secreted papain-like cysteine protease (SpeB) of the opportunistic human pathogen *Streptococcus pyogenes* was shown to cleave the N-terminal pro-domain of the interleukin-1β precursor (pro-IL-1β), facilitating its activation[96,97]. Mature IL-1β cytokine signals through the cognate interleukin-1 receptors (IL1Rs) of the immunoglobulin-like Toll / interleukin-1-like receptor family. Interestingly, the authors speculated that the vertebrate cytokine system may have evolved from an ancient function of pro-cytokines as putative pathogen protease sensors[97].

A clear example of extracellular proteolytic activation of a host-secreted danger-signal has been documented in *Drosophila melanogaster*. Host cells secrete a precursor of the chymotrypsin-like serine protease Persephone (pro-Psh) into the hemolymph of the arthropod organism. The N-terminal pro-domain of Psh keeps the danger-signaling protease in an inactive state. Processing of pro-Psh within the inhibitory pro-domain, as done by unrelated extracellular proteases of bacterial pathogens, initiated Psh activation[98]. Mature Psh is at the start of a proteolytic activation cascade that culminates in the activation a plasma membrane localized PRR-type immune receptors of the Toll-like receptor family.

Toll-like receptors themselves have been proposed as sensors of extracellular pathogen protease activity. Activation of the avian-specific plasma membrane localized TLR15 was reported to rely on the unspecific degradation of its LRR ecto-domain. The catalytic activity of several extracellular proteases of fungal and bacterial pathogens was able to induce TLR15-dependent immune signaling[99]. Deletion of its entire ecto-domain was likewise sufficient for activation, and implied that the LRRs of TLR15 are of inhibitory function in the absence of protease activity.

The most-studied example of extracellular protease sensors is the family of accordingly named protease-activated receptors (PARs). PARs are relatively small multi-pass transmembrane proteins and G protein-coupled receptors. Mammals encode 4 paralogs with functions in blood

coagulation, proinflammatory response, surveillance of the integrity of the intestinal lining, and additional regulatory processes[100]. The N-terminus of PARs is exposed to the extracellular space. It contains the precursor of their respective peptide ligand, with a short terminal pro-domain preventing its interaction with the sensing part of the receptor. Specific proteolytic processing removes the inhibitory pro-domain and allows for binding of the tethered ligand and activation of the receptor[101]. PARs perceive not only host- , but also microbe-secreted proteases, and have been implicated in allergy, pathogenesis, and innate immunity[102]. One of the microbe-secreted proteases shown to proteolytically activate human PAR receptors is EprS[65], a direct ortholog of $SLPA^{Xap}$ in *P. aeruginosa*. Because the protease-sensing pro-peptides are part of the PARs' structure, the activating proteases have to reach these receptors on the cell surface to generate their danger-signaling ligands in the first place. The plant receptor MIK2 is uncoupled from both, its cognate peptide ligands, and their production by proteases. This may be important for the perception of SLPAs in plants, as their molecular size is likely to prevent their diffusion through intact plant cell walls and thus their direct interaction with plasma membrane resident sensors.

4.2 The active peptide motif of SCOOPs in context of perception specificity

Plant encoded SCOOP and cSCOOP peptides, apart from their conserved S-x-S/T motif, appear to share little sequence overlap, yet all of them function as agonists of MIK2 ([37,38,80], and this study). Based on a sequence comparison of numerous SCOOPs from various *Brassicaceae* and biochemical analysis, we propose three major criteria for high affinity receptor interaction and immunogenic activity. These are (1) a length of exactly 14 aa, with (2) serine / threonine residues at position 6 and 8, and (3) a positive charge either on the same peptide at position 10, or provided by a second disulfide-bridged peptide chain. Other positions in the peptide are quite variable, although with a clear bias for certain amino acids and exclusions of others, for example the ones with acidic side chains.

Autoclaved *Fusarium oxysporum* cultures (*comment: thermal hydrolysis of proteins into random peptide fragments*) were shown to induce immune responses in a MIK2-dependent manner[103] and *mik2* mutants were found to be more susceptible to this pathogen[38,80,104]. This was interpreted as evidence that MIK2 may have evolved as a plant SCOOP peptide receptor from an ancient function as a receptor of SCOOP-like PAMPs[38]. An *in silico* search for SCOOP-like motifs in the proteome of *F. oxysporum*[38] resulted in hits within cytoplasmic proteins (*UniProt* ID: A0A0D2XZ19; hypothetical protein FOXG_09241 and its ortholog FGSG_07177; see *Suppl.Fig4A*). The authors speculated that predicted surface exposure of SCOOP-Like motifs within these proteins were sufficient for immunogenic activity. If this were true, then their synthetic peptides should be immunogenic regardless of terminal extension. This was not the case: While the synthetic peptides show immunogenic activity, they do require the exact same dimensions as the mature plant SCOOPs (*Suppl.Fig4B*). With respect to the loose structural requirements of SCOOP peptides, it is of no surprise that sequences with S-x-S/T motifs also occur in plant pathogens. In fact, the proteomes of most organism, including humans, contains thousands of protein fragments that look like SCOOPs. The function of these SCOOP-like sequences as immunogens, however, depends on proper trimming to the 14 aa long active peptide motif and their exposure to MIK2 in the plant apoplast. We observed that *F. graminearum* secretes considerable amounts of a heat-labile immunogenic protease, whose catalytic activity is perceived *via* MIK2. In contrast, washed

mycelial lysates of *F. graminearum*, were all but inactive. And residual activity was abolished, when preparations were boiled to inactivate traces of the protease, providing no hint of heat-stable SCOOP-like peptide activity, even in the cytosol of the fungus (*Suppl.Fig4C*). Secretion of a similar extracellular protease by the closely related *F. oxysporum* may likewise serve as an alternative explanation to the reported MIK2-dependent resistance against this fungal pathogen.

If the overall sequence composition of SCOOP peptides is of little importance to their activity, then why does the ubiquity of matching motifs in the secretome of the plant not cause autoimmunity under normal physiological conditions? Or asked differently, how does the immune receptor MIK2 achieve perception specificity? Cumulative evidence presented suggests, that precise processing of pro-SCOOPs to the shared active motif is a strict requirement for immunogenic activity at physiologically relevant peptide concentrations. Thus, perception specificity of MIK2 may be achieved through this processing requirement; it appears to be the exact length of mature SCOOP peptides that acts as a danger signal in the plant. Conceptually, the two conserved serine / threonine residues and the positively charged site of the active motif may function in the uniform orientation of distinct SCOOP peptides during their length assessment by MIK2 and SERK-family co-receptors: The length of danger-signaling SCOOPs were to be measured from both peptide termini. Correct N- and C-terminal dimensions were primarily assessed through the peptide's differential binding affinity for the receptor, and efficacy in co-receptor recruitment, respectively.

Precursors of SCOOP peptides are expected to remain in the plant apoplast in an inactive pro-peptide state, until proteolytic activation by extracellular pathogen (or plant) proteases occurred. This activation model is supported by reports of the mass spectrometric detection of forms of pro-SCOOP10[105] and pro-SCOOP12[38], but not their activated peptide form, in healthy *A. thaliana* plants.

4.3 Decoy substrate model of pro-SCOOP function and evolution

The superfamily of *prepro-SCOOPs* contains at least 34 members in *Arabidopsis* ([36,38,80], and our work) and many more SCOOP-like peptides among the predicted secretome of *A. thaliana* (unpublished work by Prof. Dr. Cyril Zipfel and colleagues). All active SCOOP peptides described so far signal through the same receptor, MIK2.

In all of our experiments we obtained no evidence that pro-cSCOOP1 were a plant protease-activated phyto-cytokine. However, the proteolytic activation of pro-SCOOPs through extracellular plant and pathogen proteases are not mutually exclusive concepts; they might act in a concerted manner. Yet the corresponding plant proteases and the conditions under which they may activate SCOOP peptide signals still remain to be reported convincingly. In our study we obtained clear evidence that not only pro-cSCOOP1 but also at least two other members of the family, pro-cSCOOP5 and pro-SCOOP13, are activated through specific extracellular proteases of *X. arboricola* pv. pruni, *S. strictum*, and *P. aeruginosa*, respectively. Most importantly, in the case of $SLPA^{Xap}$, the occurrence of the protease that activates pro-SCOOP1 has real consequences for the MIK2-dependent resistance of *Arabidopsis* against the secreting pathogen (Fig. 7D).

An intriguing thought is the possibility that the protease-sensing function of pro-SCOOPs may extend beyond microbial pathogens. It could include the perception of animal plant pests, like

parasitic cyst nematodes, aphids, or herbivorous insects. A recent report of the involvement of *MIK2* and *prepro-SCOOP12* in the resistance of *A. thaliana* against the chewing larvae of *Spodoptera littoralis*[106] might point in this direction. After all, salivary glands of herbivores are known to express proteases[107,108].

Extracellular pathogen proteases specifically processed and activated distinct SCOOP pro-peptides in the plant apoplast. This pro-SCOOP-for-pathogen protease-specific perception was reminiscent of the classical gene-for-gene model of plant immunity, as proposed by H. H. Flor[109]. Flor's model was the theoretical basis for our modern understanding of effector-triggered immunity (ETI) and innate immune systems in general. Within the ETI research community, an "*effector target [, that is] required for [receptor] function but with no function in host defense or susceptibility in the absence of its cognate [receptor, and whose] alteration [through effector action] does not result in enhanced pathogen fitness in plants that lack the [receptor, but] triggers innate immunity in plants that carry [it]*", is defined as a "*decoy*"[110]. As we couldn't ask for a more fitting description of our observations, we borrow the term and refer to pro-SCOOPs that are specifically activated by extracellular pathogen proteases (*i.e.* the *"effectors"*) as their cognate decoy substrates (*Fig.13*).

The co-evolution of new perception capabilities of plant hosts with their pathogens, which constantly evolve to evade perception[47], is assumed to be the driving force behind duplication and diversification of resistance genes. LRR-RLKs, with at least 216 family members[111] in *A. thaliana* alone, are a prime example of this hypothesis. The evolution of LRR-RLK receptors is limited by structural constraints. Their intracellular kinase and the backbone of their LRR ecto-domain are both quite conserved[7,112]. And surface residues involved in their shared mode of peptide ligand binding and co-receptor recruitment are apparently subject to stabilizing selection[8].

Promiscuous binding of ligands, as is the case for MIK2's interaction with the diverse SCOOPs, may outsource the specific sensor function from the ligand binding site on the ecto-domain of the receptor to the processing requirement of the cognate ligand precursors themselves. Freed from the structural constraints of the receptor, the pro-peptides were allowed to co-evolve rapidly with the pathogens whose extracellular proteases they are supposed to perceive. Thus, the proteolytic activation by distinct extracellular pathogens proteases may very well be the driving force behind the diversification of pro-SCOOPs in *A. thaliana* and other species of the *Brassicaceae* family. According to this hypothesis, their shared receptor could be conserved and optimized for binding affinity and signaling output, while the decoy substrates were to undergo rapid duplication and hypermutation. Non-functionalizing mutations in duplicated decoys would not result in deleterious effects on the host's immune performance, as MIK2's function in the perception of other pathogen protease-activated SCOOPs were unaffected. Such pro-SCOOP alleles would not experience negative selection. Non-functional decoy substrates could simply continue to evolve until their processing through specific proteases of emerging pathogens would result in the release of an immunogenic peptide, facilitating their positive selection in the population. In support of this model, a single point mutation in the pro-domain of the SCOOP10 precursor was found to be sufficient for a gain of perception of an otherwise non-immunogenic extracellular pathogen protease.

The rather loose structural requirements for the immunogenic activity of SCOOP peptides are a necessity for their function as decoy substrates: Simultaneous variations within variable residues of the active peptide motif and in flanking regions on the pro-peptide are required in order to evolve novel perception capabilities for unrelated proteases of different substrate specificity. To function

as a cognate decoy substrate for a pathogen protease, a pro-SCOOP has to contain the enzyme's substrate recognition sites either on its pro-domain, or in terminal regions of the active peptide motif. Such, that subsequent site-specific cleavage occurred exactly at the termini of the active motif. Yet the generated SCOOP peptide must not contain even a single additional protease cleavage site within the active motif itself, else it were degraded into inactive fragments.

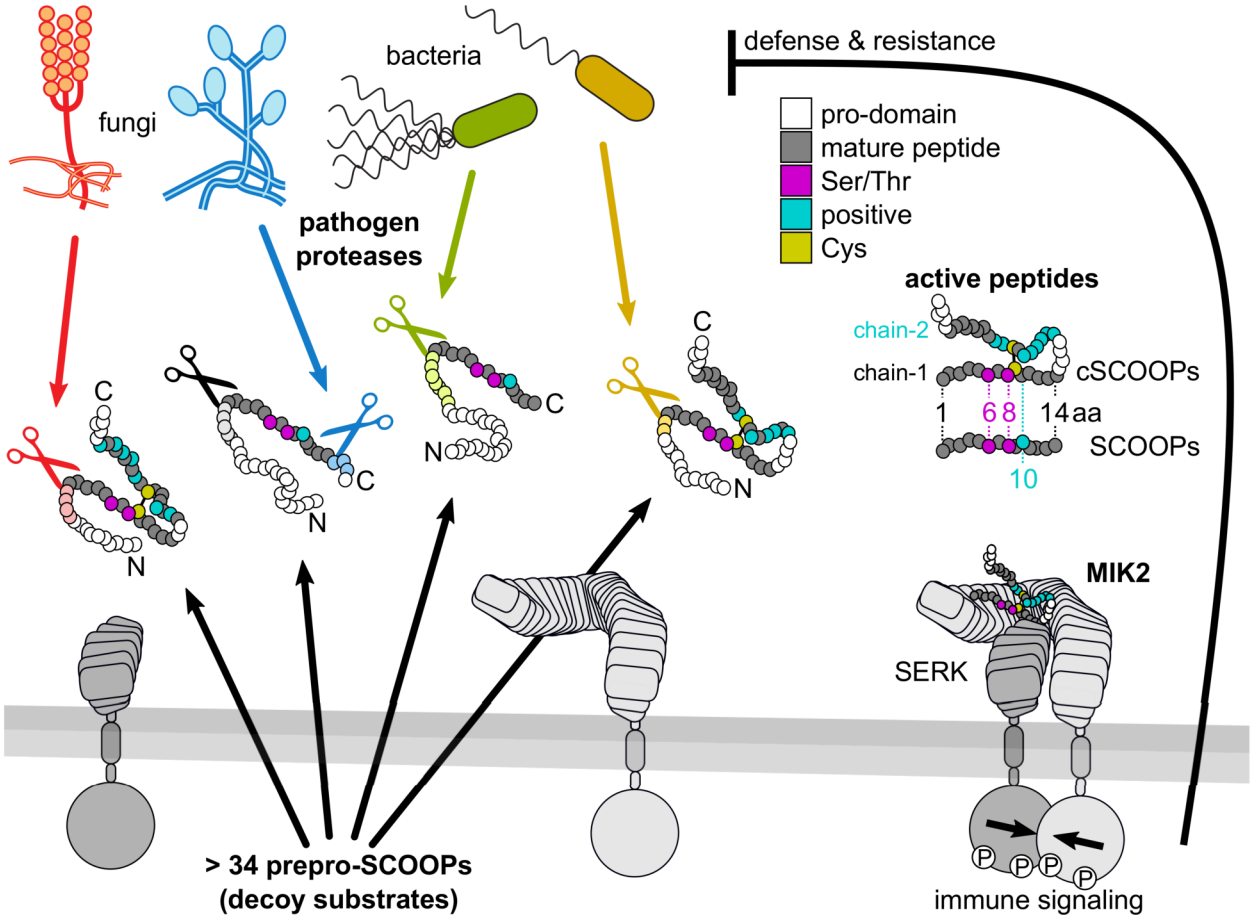


Fig.13: Decoy substrate model of pro-SCOOPs as sensors of extracellular pathogen proteases. The plant cell secretes a bouquet of (all but) inactive pro-SCOOPs into the apoplast. SCOOP peptides on these precursors are flanked by variable regions, or contain them on their termini. These variable regions are processing sites for distinct substrate-specific proteases. Attacking pathogens secrete proteases into the plant apoplast, where they encounter the SCOOP precursors. The precursors await activation in a pro-SCOOP-for-pathogen protease - specific manner. Pathogen proteases inadvertently release active danger-signaling SCOOP peptides from these decoy substrates through precise processing. Structural requirements for immunogenic activity differ between two classes of the mature peptides. Linear SCOOPs and cyclic cSCOOPs share a loosely defined active motif, whose most important criteria are a length of exactly 14 aa, with conserved serine / threonine residues at position 6 and 8, and a positive charge located C-terminally thereof. The positive charge is located at position 10 in linear SCOOP peptides and on a second peptide chain in heterodimeric cSCOOPs, respectively. Linear SCOOPs require precise processing at both termini for immunogenic activity, while N-terminal processing alone might be sufficient for activation of cSCOOPs. Both classes of active SCOOP peptides bind to the same cognate receptor, MIK2. They alert plant cells proximal to the site of protease digest to the immediate pathogen threat. Alarmed plant cells initiate a general immune response directed against the attacker(s). The result is a broad non-host resistance.
Depictions of microbes and LRR-RLKs were adapted from Boller and Felix[11] and Dr. Yan Wang.

We propose to try and apply this MIK2-dependent decoy substrate system in the protection of crops against protease-secreting plant pests. Cognate decoy substrates may be realized for any extracellular protease of pathogens of interest. They can be designed in a straightforward manner, or discovered through the screen of synthetic analogs of naturally occurring pro-SCOOPs. Efficient stacking and polycistronic expression of multiple decoy substrates from a single transgene, that also encodes MIK2, may be employed for a broad gain of immunogenic perception of unrelated pathogens.

5 Material & Methods

5.1 Expression constructs

Constructs for expression *in planta* were assembled using the versatile and scarless *Golden Gate* cloning system of Binder *et alia*[113]. cDNA sequences of *A. thaliana* (Col-0) *MIK2* and *prepro-cSCOOP1* were cloned as individual low-level modules in *pUC57* (level-I; *gentamicin*^{R}). Other level-I modules encoded a variety of promoters, tags (or stop codons), terminators, and selection markers. They were either part of the original cloning system, kindly provided by Dr. Robert Morbitzer and the lab of Prof. Dr. Thomas Lahaye (*Universität Tübingen;* Germany), or generated by us. BpiI-catalyzed cut-ligations into *pUC57 (LI + BpiI)* generated the truncated *prepro-cSCOOP1 Δ(T32-I45)* mutant module (level-I). Level-I modules of choice were combined into higher-level expression constructs in *Golden Gate* cut-ligations with BsaI and BpiI. *MIK2* and labeled versions of *prepro-cSCOOP1* were C-terminally tagged with *sGFP* and the *HiBiT* tag, respectively. *BAK1* was likewise cloned with a C-terminal *HiBiT* tag. The protein encoding sequences were assembled into *pXpre-S* (level-II module; *spectinomycin*^{R}) -, or further stacked into multi-gene expression cassettes in *pXpre-K* (level-III; *kanamycin*^{R}) - based binary expression vectors. Their transcription was driven by promoters indicated in the results section.

HiBiT is a recently developed small peptide tag (11 aa; VSGWRLFKKIS), whose detection does not rely on antibody-conjugates, but the efficient reconstitution of a truncated protein (LgBiT) of the small *Oplophorus gracilirostris* luciferase Nluc[114,115]. LgBiT and the substrate of reconstituted Nluc, furimazine, are commercially available. The *Nano-Glo® HiBiT Blotting System* kit (by *Promega*) was used for detection of recombinant HiBiT-tagged proteins on blots according to the manufacturer's instructions. The kits also enabled the highly sensitive detection of HiBiT labels in solution.

Bacterial expression constructs, used in complementation of *Xanthomonas* mutants, were generated similar to the plant constructs. First, Gibson assembly[116] was used to modify the bacterial broad host range expression vector *pDSK-602*[117] to make it compatible with the *Golden Gate* system. The resulting vector backbone, *pDSK-GG2* (level-II, *spectinomycin*^{R}) encodes *repA/B/C* and *mobA/B/C* genes that are required for bacterial conjugation, which is useful for strains that are not amenable to conventional transformation methods. *pDSK-GG2* contains a fixed *lpp* terminator, but allows the use of promoters and tags (or stop codons) of choice. The open reading frame of the *X. arboricola* pv. pruni *slpa* gene was cloned as a level-I module in *pUC57*, and the active site mutant, *slpa(S330A)*, was generated by BpiI cut-ligation into *pUC57(LI + BpiI)*. Either version was assembled into *pDSK-GG2* and expressed under control of the *araC/pBAD* promoter. This L-arabinose-inducible promoter displayed weak expression in *Xanthomonas* when grown on media void of D-glucose or L-arabinose. Complementation strains carrying this construct displayed SLPA^{Xap} activity levels similar to the ones detected in biofilm fluid of the non-transgenic wild type strain of Xap. And hence, fluids of such non-induced biofilms were used in bio-assays. Induction with > 30 mM L-arabinose resulted in increased expression and was immensely useful for preparative purification of the protease from biofilm of the complementation strain. This is because apparent toxicity of SLPA^{Xap} in *E. coli* - B strains prevented heterologous expression.

Correct assembly of all expression clones was confirmed by sequencing.

5.2 Plants

Axenic cell suspension cultures of *Arabidopsis thaliana* (Ler-0)[118] and (Col-0) - *efr-1 × bak1-4* (generated in our lab) were maintained as described by the authors. That is, cells were grown in liquid medium on an orbital shaker at 24 °C, 130 rpm, and constant light. 1-4 ml of 7 day-old pre-culture was used to inoculate 100 ml of fresh medium in weekly subculturing. The culture medium contained 0.44%(w/v) *Murashige & Skoog* salts and vitamins (M0222; by *Duchefa Biochemie*), 3% (w/v) D-sucrose, 0.5 mg/l of 2-(naphthalen-1-yl)acetic acid (NAA), 50 µg/l of N(6)-furfuryl-amino-purine (kinetin), and was adjusted to pH 5.6 with KOH. 4 to 6 day-old cultures were used in bio-assays and for purification of apoplastic pro-peptides. The Ler-0 culture was used whenever elf18 or flg22 peptides were involved in activity assays, else experiments were usually done with the other culture line.

A. thaliana plants were individually grown in peat-based potting soil and environmentally isolated growth chambers under short day conditions (cycle of 8 h light and 16 h darkness, 22-24 °C, 40-60% relative air humidity). 4 to 6 week-old plants were typically used in bio-assays.

Arabidopsis was transformed according to the established floral dip protocol[119]: 1-2 weeks after plants started bolting, any but the unopened flowers were removed with scissors. Pruned inflorescences were dipped into *Agrobacterium tumefaciens* resuspensions (strain GV3101, carrying binary expression plasmids with glufosinate resistance genes; used at an OD_{600} of 0.5, in 5% [w/v] D-sucrose and 0.02% [v/v] of the detergent *Silwet®* L-77) for 1 min under gentle agitation. Inflorescences were subsequently covered with plastic bags, and incubated over night in darkness. Plastic bags were removed and plants were further grown at long day conditions (cycle of roughly 16 h light and 8 h darkness, 22-26 °C, 40-60% relative air humidity) for seed development. T0 seeds were sown in bulk on nutrient-rich potting soil and allowed to germinate and grow for 1 week. Transformants were selected with a commercial glufosinate [~4 mg/l] - based herbicide solution (*Basta®*; by *BASF*), with a gap of 7 days in between 2 sprays. Resistant seedlings (T1 generation) were transferred to individual pots.

The *A. thaliana* (Col-0) triple mutant line *efr-1 × fls2 × rlp1-2* was previously described[57]. Null alleles of *MIK2*, that is *mik2-1*[120] (*SALK* ID: SALK_061769) and *mik2-2*[104] (*SALK* ID: SALK_046987), were ordered from the *Nottingham Arabidopsis Stock Centre* (*NASC:* https://arabidopsis.info/). Received lines were segregating; offspring homozygous for the T-DNA inserts was selected according to phenotyping with SCOOP12 peptide in bio-assays, and confirmed by PCR-based genotyping of the affected locus.

Nicotiana benthamiana plants were grown in a greenhouse under long day conditions. For transient transformation, middle-aged leaves of 4 to 5 week-old plants were infiltrated from the abaxial side with mixtures of *A. tumefaciens* resuspensions (GV3101, carrying binary expression plasmids; at an OD_{600} of 0.2 per construct, in 10 mM $CaCl_2$ and 150 µM acetosyringone), using a blunt-ended syringe. Constructs encoding the viral silencing suppressor protein P19 under control of the *CaMV35S* promoter were co-transformed. Transformed leaves were usually given 2 to 3 days of time for expression before they were used in experiments.

5.3 Pathogens

Xanthomonas spp. and *Pseudomonas aeruginosa* - PA14 were grown at 28 °C on solid NYG medium (0.5% [w/v] peptone, 0.3% [w/v] yeast extract, 2% [w/v] glycerol, 1% [w/v] agar).

Cultures of *Sarocladium strictum* (mislabeled as *Ustilago maydis* - UM521; identity as *S. strictum* implied by re-sequencing of the internal transcribed spacer between the fungal 18 S and 28 S rRNA genes) and *Fusarium graminearum* - IMB 12337 were obtained from *Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH* (*DMSZ;* https://www.dsmz.de), and maintained at RT on solid MP medium (1% [w/v] malt extract, 0.25% [w/v] peptone, 1.5% [w/v] agar).

The empty vector *pBBR1MCS-5*[121] (*gentamicin*^{R}) was introduced into *X. arboricola* pv. pruni for its resistance gene that was used as a positive selection marker and facilitated the contamination-free work with the bacteria on media containing 20 µg/ml of gentamicin.

Δslpa deletion mutants in *X. arboricola* pv. pruni were generated *via* homologous recombination using suicide vector *pOGG2*[122] (*spectinomycin*^{R}): Sequences flanking the open reading frame of *slpa* were assembled into *pOGG2 via* BsaI cut-ligation in such a manner that the 500 bp long homology arms of the locus were joined to either end of a carbenicillin resistance gene (*bla*) in their original orientation. The recombination construct was transformed into donor-strain *E. coli* - DH5α (*λpir*), and introduced into *X. arboricola* pv. pruni through conjugation in tri-parental mating with helper-strain *E. coli* - HB101 carrying *pRK2013*[123] (required for conjugative transfer and replication of *pOGG2*, *kanamycin*^{R}) in the absence of antibiotics, as previously described[123]: Trans-conjugants were selected on media containing 100 µg/ml of spectinomycin in addition to gentamicin. Colonies were re-streaked and grown for another 3 days to provide sufficient time for homologous recombination between the construct and the *slpa* locus to occur. They were subsequently subjected to bottleneck selection to facilitate the loss of the recombined plasmid; dilutions were subcultured over several days in the absence of spectinomycin. Clones that had lost *pOGG2,* were obtained by conditional negative selection on 5% (w/v) D-sucrose containing media, an effect that is caused by the product of the levansucrase gene (*sacB*) on the backbone of *pOGG2*. The media also contained 100 µg/ml of carbenicillin for parallel selection of clones with recombined *slpa* alleles. Loss of suicide plasmid *pOGG2* and helper plasmid *pRK2013* was suggested by the colonies' sensitivity to spectinomycin and kanamycin, respectively. Seamless substitution of the open reading frame of *slpa* with the *bla* gene was confirmed by PCR-based genotyping and sequencing of the locus beyond the length of the homology regions.

Complementation constructs encoded on *pDSK-GG2* were introduced into the *Δslpa* strain of *X. arboricola* pv. pruni through conjugation as described above. Trans-conjugants were selected for simultaneous resistance to spectinomycin, carbenicillin and gentamicin, and sensitivity to kanamycin. PCR-based genotyping confirmed the introduction of the plasmid-encoded *slpa* and *slpa(S330A)* constructs.

5.4 Bio-assays

Extracellular alkalization was measured with *A. thaliana* cell suspension cultures as previously described[82]: Culture flasks were opened and placed on an orbital shaker running at 130 rpm and allowed to acclimatize for several hours. 1 ml aliquots were distributed to 15 ml flat-bottom glass vials positioned on same shaker. Small combined-glass electrodes (by *Mettler Toledo*) were added to each vial and changes in pH continuously measured using regularly calibrated pH-meters (632 / 691, by *Metrohm*). Cells were treated with immunogens once pH was stable, and subsequent changes in pH were recorded.

ROS bursts in response to treatment with immunogenic peptides were detected in leaf pieces

according to an established protocol[124]. In short, middle-aged leaves were cut into rectangles ($4 \times 4\ mm^2$) and floated over night on water. They were then transferred to white flat-bottom 96-well plates containing 0.2 mM luminol (L-012, by *Wako Chemicals*) and 20 µg/ml of horseradish peroxidase (by *Applichem* and *Sigma Aldrich*). Plates were moved to a luminometric plate reader and allowed to acclimatize for 30 min. Wells were then treated with peptides, whose induction of extracellular ROS was recorded as an increase in luminescence over time.

Induction of ethylene biosynthesis in response to immunogenic peptides was measured as previously described[13]. Leaf pieces were prepared in the same manner as for ROS burst measurements. 3 pieces each were transferred to 6 ml glass reaction tubes containing 500 µl of water. Following treatment with peptides, reaction tubes were immediately closed with rubber stoppers and agitated on an orbital shaker at 100 rpm on a slope. Head-space was sampled 4 to 6 h post treatment and ethylene contained therein quantified through gas chromatography (GC-14 A, by *Shimadzu*): The gas sample was separated on an alumina column (1.2 m long, 3 mm inner diameter, self-packed with Al_2O_3 grains) *via* isocratic elution at 140 °C with an inert carrier gas (250 kPa N_2). Signals were detected with an FID (50 kPa H_2, 25 kPa air). Ethylene peaks were identified according to their retention time and were integrated for quantification. The assay was modified for measurement of immunogenic protease activity. Whole leaves, or 3 leaf discs of 10 mm diameter, were (3 ×) vacuum-infiltrated with protease fractions (or mock; in 0.1% [w/v] BSA, 10 mM $CaCl_2$) for 3 min in order to mimic their apoplastic localization. Infiltrated leaf material was transferred to 15 ml flat-bottom glass vials, containing 200 µl of water to prevent desiccation. Vials were sealed with rubber stoppers and incubated at RT for 10-18 h, without agitation. Sufficient incubation time was necessary for the substrate-limited enzyme reaction.

Infection assays of *A. thaliana* with *X. arboricola* pv. pruni strains were done as follows: Fresh bacterial biofilm was harvested from solid media 24 h hours after plating of pre-cultures. Biofilm was gently resuspended in 10 mM $CaCl_2$, adjusted to OD_{600} of 0.4, and further diluted by factor 1 000 ×. Final resuspensions were infiltrated into middle-aged leaves of well-watered *A. thaliana* plants (3 leaves per plant ≙ 1 biological replicate) using a blunt-ended syringe. Leaf surfaces were carefully dried with soft paper towels to remove visible external traces of resuspensions. Infiltrated plants were transferred to mini-greenhouses and covered with plastic lids for high humidity conditions. Plants were moved to short day conditions with decreased temperature (18-22 °C). Lower temperature was used because the inside of the mini-greenhouses heated up slightly above ambient temperature under light exposure. Leaf samples were harvested immediately following infiltration, at day 0, and 3 days later with a cork punch (1 disc of 5 mm diameter per leaf). Leaf discs of each replicate were transferred together to 1.5 ml reaction tubes containing 4 glass beads (3 mm diameter) in 200 µl $MgCl_2$. Bacteria were extracted from leaf samples by homogenization with a mixer mill (*MM400*, by *Retsch*) for 2 × 1 min at 1 600 rpm. Homogenates were serial diluted in 10 mM $MgCl_2$ and spotted on NYG solid media supplemented with 20 µg/ml of gentamicin and 50 µg/ml of cycloheximide for selection. Colonies were counted 3 days later.

5.5 Binding and interaction assays

We utilized the exceptionally sensitive detection properties of the HiBiT peptide label and established a novel luciferase-based *in vitro* peptide-ligand / receptor binding assay (manuscript accepted for publication). In short, cSCOOP1's chain-1 was heterodimerized with a truncated chain-2 that contained an N-terminal HiBiT peptide tag (VSGWRLFKKIS:G:RRPRPCRL; chain-2

is underlined). The purified peptide was allowed to bind to *in planta*-expressed and affinity-immobilized MIK2:GFP. Unbound peptide ligand was removed through rapid washed, and receptor-bound ligand was eluted by heat-denaturation of the complex. The labeled ligand was quantitatively detected through HiBiT-mediated reconstitution of the Nluc luciferase activity in presence of an excess of the LgBiT luciferase fragment and substrate.

In more detail: Recombinant *MIK2:GFP* was transiently expressed in leaves of *N. benthamiana* under control of the *CmYLCV* promoter. Expressing leaves were harvested, shock-frozen in cryogenic nitrogen, and ground into a fine powder. Frozen plant powder (200 mg per sample) was resuspended in ice-cold extraction buffer (50 mM MES at pH 5.7, 100 mM NaCl) at a concentration of at 20% (w/v). Cell debris, containing the membrane fraction and hence the receptor, was pelleted by centrifugation at 4 °C and 5 000 × *g* for 5 min. The supernatant was discarded and the cell debris was washed twice with an excess of same buffer by repeated cycles of resuspension and centrifugation. The washed pellet was resuspended in ice-cold solubilization buffer (20% [w/v] in 25 mM Tris at pH 8.0, 150 mM NaCl, 1% [v/v] Nonidet P-40, 0.5% [w/v] sodium deoxycholate) and incubated at 4 °C for 1 h by over-head shaking at 15 rpm. Insoluble debris was subsequently removed from the solubilizate by centrifugation at 4 °C and > 16 000 × *g* for 25 min. The supernatant was transferred to new 1.5 ml reaction tubes. *GFP-Trap® magnetic agarose beads* (10 µl of slurry per sample, by *Chromotek*) was equilibrated in solubilization buffer and added to the solubilizate. The mix was incubated for one more hour at 4 °C, with agitation. The GFP-tagged receptor was subsequently pulled down with the affinity beads and the help of a magnetic rack. The beads were washed (2 ×) with solubilization buffer, (2 ×) with IP wash buffer (25 mM Tris at pH 8.0, 150 mM NaCl), and (2 ×) with binding buffer (50 mM MES at pH 5.7, 100 mM NaCl, 10 mM $CaCl_2$, 10 mM $MgCl_2$). Each time, they were completely resuspended in 1 ml of ice cold buffer and pulled back down again with the magnet for the removal of the supernatant. 200 µl of binding buffer, containing 100 or 150 pM of HiBiT-labeled cSCOOP1, alone or in combination with increasing concentrations of unlabeled competitor peptides, was added to the beads. The binding mixture was allowed to equilibrate for more than 1 hour with gentle over-head shaking. The ligand / receptor complex was subsequently pulled down with the magnet and rapidly washed (4 ×) with ice-cold binding wash buffer (50 mM MES at pH 5.7, 150 mM NaCl, 50 mM $CaCl_2$, 10 mM $MgCl_2$, 0.1% (v/v) Tween-20) to remove non-bound ligand. Washed beads were resuspended in 200 µl of same buffer and transferred to a new tube (safe-lock). The binding buffer was removed, and replaced with the same volume of elution buffer (100 mM Tris at pH 8, 100 mM NaCl). Bound ligand was released from the receptor through heat-denaturation at 95 °C for 10 min. White flat-bottom 96-well plates were blocked with 100 µl of 1% (w/v) BSA and 0.1 M NaCl per well. The blocking solution was removed 20 min later and replaced by 10 µl of binding assay eluates. 90 µl of freshly prepared detection solution (1 × *Nano-Glo®* blotting buffer, 0.5% [v/v] LgBiT protein solution, 0.2% [v/v] furimazine substrate solution; of the *Nano-Glo® HiBiT Blotting System* kit by *Promega*) was added to each well. The plate was mixed by gentle tapping, incubated for 20 min in the dark, and then transferred to a luminomeric plate reader for detection of the HiBiT label. HiBiT-labeled standard peptides were used in dilution series for calibration of the system, and allowed for conversion of arbitrary luminescence units to (femto-) molar amounts of bound ligand. Leftover volumes of the eluates were used in replicates for quantitative Western blotting with anti-GFP antibody. The intensities of the corresponding protein bands were compared to dilutions series of a GFP standard of known concentration, and used to determine the concentration of MIK2:GFP that was present during the binding equilibration step. Models used in the estimation of binding

affinities are described in the statistics section.

The ligand-induced interaction of MIK2 with SERK-family receptors was assessed *in planta* in co-immunoprecipitation experiments. The procedure was modified from earlier descriptions[42]. *MIK2:GFP* and *BAK1:HiBiT* were transiently expressed in *N. benthamiana* under control of the *CmYLCV* promoter and the *CaMV35S* promoter, respectively. Expressing leaves were infiltrated with 1 µM of indicated peptides in water and harvested 3 min later. They were frozen and ground into a fine powder. 200 mg of powdered leaf material was resuspended in 1 ml of solubilization buffer supplemented with 1% (v/v) of *plant protease inhibitor cocktail* (PPI, by *Sigma Aldrich*). The resuspension was incubated for 1 h at 4 °C with overhead-shaking. Insoluble material was precipitated by centrifugation at 4 °C and > 40 000 × *g* for 30 min. The solubilizate was transferred to new tubes, and aliquots were taken as input samples. 10 µl of equilibrated *GFP-Trap® magnetic agarose beads* slurry was added to the solubilizates. The mixture was incubated for another hour at 4 °C, with agitation. Recombinant MIK2:GFP was immunoprecipitated and pulled down with the help of a magnetic rack. Immobilized beads were rapidly washed, (2 ×) with each, ice-cold solubilization buffer and IP wash buffer. Input samples and washed beads were boiled in SDS sample buffer, separated on replicate SDS-PAGE gels, and subjected to Western blot analysis for the detection of the tagged proteins.

5.6 Chromatography

Activity-guided purification of the immunogenic extracellular *Xanthomonas* protease ($SLPA^{Xap}$): Biofilm resuspensions (OD_{600} of 2 in NYG liquid medium) of *X. arboricola* pv. pruni were plated on NYG solid media containing the appropriate antibiotics and grown for < 2 days at RT. 350 g (5 m^2) of biofilm was harvested and resuspended in ice-cold water at a concentration of < 20% (w/v). All subsequent purification steps were carried out at 4 °C. Bacterial cells were removed by successive centrifugation at 5 000 × *g* and 10 000 × *g*, each for 30 min. Supernatant was treated with buffer (20 mM Tris, pH 8 at 4 °C) and loaded in two batches on a large self-packed AEC column (310 ml of *Q-Sepharose™ Fast Flow*, by *GE Healthcare Life Sciences*). The column was washed with 120 mM NaCl, and the active fraction was rapidly step-eluted at a flow-rate of 25 ml/min with 600 ml of 400 mM NaCl in same buffer. $(NH_4)_2SO_4$ and NaCl were added step-wise to a final concentration of 1 M and 0.3 M, respectively. The sample was loaded on a pre-equilibrated HIC column (*Fractogel TSK-Butyl-650-S*, 15 ml; by *Merck*). The activity eluted within a gradient of 100-0% of $(NH_4)_2SO_4$ [1 M] & NaCl [0.3 M] in 20 mM Tris (pH 8) over 150 ml at 4 ml/min. The most-active fractions were diluted with 10 mM Tris (pH 8) to a conductivity of < 7 mS/cm, loaded on an AEC column (*HitTrap™ Q-Sepharose™ High Performance*, 5 ml, by *GE Healthcare Life Sciences*), and separated by gradient elution with 0-40% of 1 M NaCl over 60 ml in same buffer, at a flow-rate of 5 ml/min. Fractions of the activity peak were dialyzed against water (2 × 3 h, MWCO of 10 kDa). Desalted samples were flash-frozen in N_2(fl.), lyophilized, and re-eluted at a concentration of 5 mg/ml in 1 M NaCl and 20 mM Tris (pH 8). The sample was separated in three batches on an SEC column (*HiLoad™ Superdex™ 16/600 200 prep grade*, by *GE Healthcare Life Sciences*) with isocratic elution at a flow-rate of 3 ml/min. Most-active fractions of replicate runs were pooled and desalted on ultrafilters (*Vivaspin® 20*, MWCO of 10 kDa, by *Sartorius*). The active fraction was polished on an high resolution AEC column (*Mono Q™ 4.6/100 PE*, by *GE Healthcare Life Sciences*), with a shallow gradient elution of 0-40% of 1 M NaCl in 20 mM MES (pH 6) over 60 ml at 2 ml/min. The activity peak was up-concentrated on same ultrafilters, aliquoted, flash-frozen, and stored at -80 °C.

The apoplastic pro-peptide (pro-cSCOOP1), whose activation by $SLPA^{Xap}$ resulted in an active peptide product, was enriched from the apoplastic fraction of *A. thaliana* cell cultures through large scale purification efforts: Huge amounts of apoplastic volume were rapidly harvested and enriched with a batch-capture protocol that was specifically developed for this purpose: 180 liters of plant cell suspension cultures were processed in batches of > 15 liters, each. All steps were carried out at 4 °C. Plant cells were quickly removed from the suspension through filtration with sieves and *Miracloth®* filtration material (by *Merck Millipore*). Filtrate was collected in 5 liter beakers on magnetic stirrers and cooled down to < 10 °C within minutes. This was achieved through the generous use of cool packs (-20 °C) that were immersed into the solution. Buffer (20 mM MES, pH 6) and CEC resin (2% [w/v], *SP-Sepharose™ Fast Flow*, by *GE Healthcare Life Sciences*) was directly added, and the slurry subsequently incubated for 30 min under constant stirring. The slurry was then rapidly filtered through nylon membranes (20-60 µm mesh), and the filter-cake was quickly packed into a self-made column. The resin was washed with 200 mM, and step-eluted with 1 liter of 2 M NaCl in 20 mM MES (pH 6), at a flow rate of > 100 ml/min. $(NH_4)_2SO_4$ was added step-wise to the batch eluate to a final concentration of 2 M. The sample was loaded onto a pre-equilibrated self-packed HIC column (*Butyl-S Sepharose™ 6 Fast Flow*, 25ml, by *GE Healthcare Life Sciences*), washed with 1.8 M, and step-eluted with 100 ml of 0.4 M $(NH_4)_2SO_4$ in 20 mM MES (pH6), at a flow rate of 10 ml/min. Aliquots were taken of the fractions before they were flash-frozen in N_2(fl.) and stored at -80 °C. The aliquots were digested for 1 h at RT with 1 nM of purified protease $SLPA^{Xap}$. The protease was subsequently inactivated by boiling. The activity of the heat-stable peptide product was quantified in *A. thaliana* cell suspension cultures to determine the amount of pro-peptide substrate in the fractions.

Activity-guided purification of the active peptide product (mature cSCOOP1): Fractions with considerable amounts of pro-peptide substrate were collected, pooled, and bulk-digested with 10 nM of purified $SLPA^{Xap}$ for 4 h at RT. The protease was inactivated by boiling for 30 min. Insoluble protein precipitate was removed by centrifugation at 4 °C and > 10 000g for 30 min. The sample was treated with 0.5% (v/v) TFA, and split into 8 aliquots. Peptides contained therein were extracted using C18 RP-SPE columns (C18 ec *Chromabond®*, 1 000 mg, by *Macherey Nagel*). The columns were washed with 0.1% (v/v) TFA, and step-eluted with 30% and 60% (v/v) ethanol in 0.1% (v/v) TFA. The eluate was separated on a CEC-HPLC column (*Bio-SCX NP5 SS*, by *Agilent Technologies*), with a gradient of 10-45% of 2 M NaCl in 20 mM MES (pH 6) over 20 ml at a flow-rate of 0.6 ml/min. The activity peak was loaded on an RP-HPLC column (*ZORBAX Rx-C18 4.6x250*, by *Agilent Technologies*) and separated in two successive runs of shallow gradient elution at different pH: First at pH 4.5, with 5-40% (v/v) acetonitrile in 10 mM ammonium acetate over 25 ml at 0.4 ml/min. And, following the removal of the organic solvent in a centrifugal sample concentrator (under vacuum, at 60 °C), then in two batches at pH2, with 0-40% acetonitrile in 0.1% (v/v) TFA over 40 ml at a flow rate of 1 ml/min. The fractions of the activity peak were dried as before and re-eluted in water.

Preparation of crude extracellular protease fractions from microbial pathogens: Bacterial biofilm resuspensions (OD_{600} of 2 in NYG liquid medium) were plated on solid media and grown over night. Fresh biofilms of *X. arboricola* pv. pruni and *P. aeruginosa*-PA14 were harvested 20-24 h and 12-14 h later, respectively. They were resuspended in ice-cold water at concentrations of 20% (w/v) of biofilm mass. Bacterial cells and cell fragments were removed by successive centrifugation at 5 000 × *g* for 10 min and (2 ×) > 14 000 × *g* for > 20 min, each at 4 °C. Agar blocks with mycelia were excised from fungal pre-cultures and used to inoculate fresh solid media. The

fungi colonized the plates over the course of several days. Mycelia of *F. graminearum* and *S. strictum* were given 5 days and 9 days after transfer, respectively. They were carefully removed from the solid media with tweezers and flat spatulas, transferred to reaction tubes, and covered with water at a ratio equivalent to 20% (w/v) of mycelial biomass. Following 30 min of incubation, mycelia were removed by centrifugation at 1 500 × *g* for 5 min. The supernatant was centrifuged again, (2 ×) at 4 °C and > 14 000 × *g* for > 20 min, to remove debris. Supernatants of bacterial and fungal pathogens were loaded on ultrafiltration columns (*Nanosep® Omega™*, MWCO of 30 kDa, 500 µl, by *PALL*) for enrichment of large extracellular proteins. They were up-concentrated (10 ×), and subsequently washed (2 ×) with ice-cold water by repeated centrifugation at 8 000 × *g* and 4 °C. Aliquots were prepared, flash-frozen in N_2(fl.) and stored at -80 °C. The extracellular fraction of *S. strictum* tended to clog the 30 kDa ultrafilters. Hence, it was first filtered through ultrafiltration columns (same kind) with a MWCO of 100 kDa to remove tiny particulate matter. The flow-through was subsequently used with 30 kDa filters as done with the other microbial preparations.

S. aureus V8 protease was obtained from commercial sources (synonymous: *Endoproteinase GluC*, by *NEB*).

5.7 Mass spectrometry

Tryptic in-gel digests of the immunogenic protein purified from *Xanthomonas* ($SLPA^{Xap}$) were performed according to standard procedures through services of the *Proteome Center Tuebingen* (*University of Tübingen*; Germany). The service also included the generic mass spectrometric analysis of the tryptic peptide fragments and provided a list of candidate proteins.

For analysis of purified and synthetic (pro-) peptides, 0.1-1 nmol of samples were injected onto a C18 RP-UPLC column (*Acquity UPLC HSS T3 2.1x100;* by *Waters*). They were eluted with a shallow gradient of 5-30% (v/v) (or higher) methanol in 0.1% (v/v) formic acid at a flow rate of 200 µl/min over 30 min. Eluting peptides were detected with an ESI (positive ion mode) - coupled Q-TOF mass spectrometer (*Synapt-G2 Mass Spectrometer*; by *Waters*).

The mass spectrometer was operated in scanning mode for MS runs with the purified $SLPA^{Xap}$ - activated apoplastic peptide signal of *A. thaliana* (cSCOOP1). The run was performed in replicates, (2 ×) with the oxidized and the chemically reduced (30 mM 2-mercaptoethanol; 95 °C for 30 min) form of the heterodimeric peptide, under otherwise identical conditions. The ion source was disconnect from the UPLC system in one of the runs with the oxidized form and eluting fractions were collected for the correlation of immunogenic activity with peptide masses. The reduced peptide was subjected to further MS / MS analyses: Predominant ions of all isoforms of the parental peptide chains were selected and fragmented *via* collision-induced dissociation (CID) over a wide range of fragmentation energies. Masses of fragment ions (m/z) were recorded. Runs had to be repeated many times due to incomplete fragmentation, caused by CID-energy depletion through predominant fragmentation into y_2-ions and deglycosylated isoforms of the parental masses. The post-translational modifications of the peptide prevented the identification through conventional comparison to proteomic databases. The structure of the peptide was manually reconstructed instead.

Quantification of peptides in digests of synthetic pro-SCOOPs with pathogen proteases: Digests were performed as described in the results section. Insoluble proteins were precipitated from the heat-denatured digests through centrifugation at > 40 000 × *g* for 30 min. The peptide-containing

supernatant was directly injected into the RP UPLC-ESI-MS system and separated as described above. MS analyses were performed in scanning mode, and total ion counts were detected. Counts of all ion masses (m/z) that matched the active peptide motif (14 aa), or its pro-peptide analog, were integrated.

5.8 Synthetic peptides

More than a hundred SCOOP-related peptides were order from custom synthesis services (mostly *GenScript*); they were ordered at crude purity for economic reasons. Regardless, received preparations were typically of sufficient purity. MIK2-dependency of peptide activity was confirmed for each preparation, excluding putative contamination with other immunogenic peptides (like the frequently ordered flg22), that would affect bio-assays even in trace amounts. Relevant peptides with lower purity, or such ones that displayed irregular activity patterns, were either re-ordered, or further purified at our hands.

Heterodimeric synthetic cSCOOPs were ordered as individual peptide chains. They dimerized spontaneously when exposed to slightly alkaline conditions and atmospheric oxygen: 1 mM of each, chain-1 and chain-2, were co-incubated in 50 mM ammonium acetate (pH ~ 9) under vigorous agitation for > 1h at 60 °C. The dimerization buffer was removed by C18 RP-SPE. The representative heterodimers of cSCOOP1 were further purified to > 90% purity, using the described RP-HPLC procedure.

Dilution series of peptides stocks were prepared in 1% (w/v) BSA and 0.1 M NaCl to prevent putative loss of dilute peptide due to unspecific binding to plastic surfaces. Same solution without peptides was used for mock treatments.

5.9 Bioinformatical analyses

Analysis of sequences, phylogenetic modeling and genetic construct design was done with the help of the proprietary *Geneious Prime* software suite (vers. 2022, by *Dotmatics*; https://www.geneious.com/).

Signal peptides and their processing sites by signal peptidases were predicted using the *SignalP* server (vers. 6.0[125]; https://services.healthtech.dtu.dk/services/SignalP-6.0/). Other protein domains and amino-acid residues of interest were annotated according to homology to *InterProScan* (vers. 5[126]; http://www.ebi.ac.uk/interpro/search/sequence/) and *BLASTp* (https://blast.ncbi.nlm.nih.gov/Blast.cgi) database hits.

Phylogenetic comparison of serine proteases was assisted by the *MEROPS Peptidase Database* (vers. 12.1; https://www.ebi.ac.uk/merops). *UniProt* (https://www.uniprot.org/) was searched for SLPA family members, using name tags of the characteristic domains as queries. Protein hits were filtered for full length, and for canonical SLPA protein structure according to the respective predictions in the *AlphaFold Protein Structure Database* (https://alphafold.ebi.ac.uk/). Progressive pairwise alignments of genuine SLPAs and IgA1 proteases / SPATEs were performed with *MAFFT*[127] on basis of the *BLOSUM62* aa substitution matrix[128]. The joined alignment was used for phylogenetic tree construction with the hill-climbing algorithm of *RAxML* (vers. 8[129]).

Partial sequences obtained during mass spectrometric *de novo* sequencing of the purified active peptide (cSCOOP1) were used as queries to search the *UniProt* protein database with the *UniProt*

Peptide Search tool (https://www.uniprot.org/peptide-search). Results were filtered for signal peptide-containing sequences. The *tBLASTn* (https://blast.ncbi.nlm.nih.gov/) tool was used to mine protein coding nucleotide sequences of the *NCBI* database in search of homologs of prepro-SCOOPs. 19 representatives, containing both, cSCOOP and linear SCOOP members, were used as queries. Results were filtered for signal peptide containing hits. Genuine prepro-peptides were manually aligned according to the biochemically mapping of their active peptide motif.

5.10 Statistical analyses and modeling

Statistical analyses and plots were done in the *R* programming language (*r-base*, vers. 4.2.2[130] https://www.R-project.org) using the *Rstudio IDE* (vers. 2022.12[131], by *Posit Software*; https://posit.co/products/open-source/rstudio/), with following additional package libraries: *broom*, *dplyr*, *egg*, *emmeans*, *ggplot2*[132], *ggseqlogo*[133], *modelr*, *multcomp*[134], *multcompView*, *nlme*, *purrr*, *scales, tibble*, *tidyr.*

Compliance with parametric model assumptions, such as the homogeneity of variance and normality of (transformed) data, were assessed in QQ-plots and with Shapiro-Wilk tests prior to statistical inference. Selection of *ANOVA* models (1-way vs. 2-way vs. 3-way) was guided by measures of model fitness and parsimony; the simpler of two models was chosen, when both were of comparable descriptive power. *Post-hoc* analyses (*Tukey's HSD*) were only performed when *ANOVAs* already indicated highly significant differences between data grouping levels, and *P*-values were adjusted for repeated hypothesis testing. Geometric means were adequately used in place of arithmetic means when sets of data with exponential or logarithmic distribution (*i.e.* % and EC_{50} values) were described and compared to each other.

Quantitative dose-dependent plant immune responses were modeled as general four parameter logistic growth functions[135] of peptide concentration. Non-cooperativity (*i.e. Hill* slope of 1) was assumed in simplification of the models, without loss of model fitness. This reduced their complexity to 3 parameters, that is (1) EC_{50} or IC_{50}, (2) lower asymptote, and (3) upper asymptote:

EC_{50} (agonist)
$$response = min + \frac{(max - min)}{1 + 10^{(\log_{10}(EC_{50}) - \log_{10}[peptide])}}$$

IC_{50} (antagonist)
$$response = min + \frac{(max - min)}{1 + 10^{(\log_{10}[peptide] - \log_{10}(IC_{50}))}}$$

Dose-response data obtained with SLPA^{Xap} were inadequately described by the four parameter model. This is because the dose-response curve displayed an asymmetric shape with an unusually shallow slope at the upper asymptote. The protease was clearly substrate-limited at high concentrations. Asymmetry was addressed with a five parameter model[136,137], that introduced a symmetry factor and enabled the confident estimation of the protease's EC_{50} value:

$$response = min + \frac{(max - min)}{\left\{1 + 10^{\left\{Hill \times \left\{\left\{\log_{10}(EC_{50}) + (\frac{1}{Hill}) \times \log_{10}\left\{2^{(\frac{1}{sym})} - 1\right\}\right\} - \log_{10}[protease]\right\}\right\}}\right\}^{sym}}$$

Same 3 parameter logistic growth functions were used for the estimation of EC_{50} and IC_{50} values in binding models that described observed saturation binding and competition binding data, respectively.

Importantly, the EC_{50} value was not synonymous to the affinity (K_d) of the receptor for the labeled peptide ligand. This was due to the high affinity of the observed interaction and the enrichment of the receptor in the binding assay: Simplifications of the mass action law of ligand / receptor interaction, which are often used in saturation binding experiments to derive the K_d directly from the EC_{50} value, include the intrinsic assumption, that the concentration of free ligand $[L]_{free}$ at the EC_{50} roughly equals its total concentration $[L]_{total}$:

general: $K_d = \frac{[L]_{free} \times [R]_{free}}{[LR]}$ $\frac{[LR]}{[R]_{total}} = \frac{[L]_{free}}{K_d + [L]_{free}}$ $[L]_{total} > [L]_{free}$ $EC_{50} > K_d$

simplified: $K_d = \frac{[L]_{total} \times [R]_{free}}{[LR]}$ $\frac{[LR]}{[R]_{total}} = \frac{[L]_{total}}{K_d + [L]_{total}}$ $[L]_{total} \sim [L]_{free}$ $EC_{50} \sim K_d$

Which implies that the concentration of the ligand / receptor complex [LR] is negligible (*i.e.* $[LR] \ll [L]_{free}$). This assumption is only true, if the total receptor concentration $[R]_{total}$ is much lower than both, the receptor's affinity for the labeled ligand K_d, and the free ligand concentration $[L]_{free}$ at the EC_{50} (excellently reviewed by Jarmoskaite *et alia*[138]). These assumptions are obviously violated for high affinity interactions (*i.e.* low K_d), in particular when the receptor protein were enriched for the binding assay. This violation is indicated by an EC_{50} value within the same order of magnitude as the total receptor concentration $[R]_{total}$. Estimates of K_d values would be artificially inflated if the simplified model were applied regardless.

An alternative binding model for use with these so-called "*intermediate binding regimes*" has been proposed as a solution to the technical problem (*equation 6* discussed by Jarmoskaite *et alia*[127]):

$$\frac{[LR]}{[R]_{total}} = unspecific + [R]_{total} \times \frac{([R]_{total} + [L]_{total} + K_d) - \sqrt{(([R]_{total} + [L]_{total} + K_d)^2 - 4 \times [R]_{total} \times [L]_{total})}}{2 \times [R]_{total}}$$

The trade-off in its application is, that it requires estimates of the total receptor concentration, present during the binding equilibration step. The total concentration of recombinant MIK2:GFP was determined as described, included as a parameter in the model, and used to accurately estimate the K_d of MIK2 for HiBiT-labeled cSCOOP1.

Binding affinities (K_i) of MIK2 for unlabeled cSCOOP1 competitor peptides were modeled with following considerations: The kinetics of competition binding are described by the growth term of the logistic regression function that is used to model IC_{50} values. According to the mass action law, this growth term equals the concentration of receptor-bound labeled ligand $[LR]_i$ at the considered total concentration of competitor $[I]_{total}$, divided by the initial concentration of receptor-bound labeled ligand in absence of competitor $[LR]_0$:

growth term: $$\frac{[LR]_i}{[LR]_0} = \frac{1}{1+10^{(\log_{10}([I]_{total}) - \log_{10}(IC50))}}$$

Per definition, this term equates to 0.5 at a total competitor concentration equal to its IC_{50}, and rearrangement of the formula allows to solve for the root of the polynomial function:

growth term at $[I]_{total} = IC_{50}$: (solve for root; K_i)

$$0 = \frac{[LR]_i}{[LR]_0} - 0.5$$

$[LR]_i$ and $[LR]_0$ can both be solved with an exact mass action law-based cubic equation of competition binding kinetics, as previously illustrated[139]:

equation-a:

$$a = K_d + K_i + [L]_{total} + [I]_{total} - [R]_{total}$$

equation-b:

$$b = K_i \times ([L]_{total} - [R]_{total}) + K_d \times ([I]_{total} - [R]_{total}) + K_i \times K_d$$

equation-c:

$$c = -K_d \times K_i \times [R]_{total}$$

equation-t:

$$t = acos\left\{\frac{-2a^3 + 9ab - 27c}{2 \times \sqrt{(a^2 - 3b)^3}}\right\}$$

equation for $[LR]_i$ & $[LR]_0$:

$$[LR] = \frac{[L]_{total} \times (2 \times \sqrt{(a^2 - 3b)} \times \cos(\frac{t}{3}) - a)}{3 \times K_d + (2 \times \sqrt{(a^2 - 3b)} \times \cos(\frac{t}{3}) - a)}$$

Equation-a, -b, -c, and -t are inserted into the final *equation for $[LR]_i$ & $[LR]_0$*. The IC_{50}, derived from the generic competition binding model, is used as a value for the parameter $[I]_{total}$, and an infinitesimally small value (*e.g.* $[I]_{total} \sim 10^{-21}$) may be used to estimate the upper asymptote described by $[LR]_0$. This simplifies the equation drastically, to just 4 parameters: These are (1) the total concentration of labeled ligand $[L]_{total}$ (*is known*), (2) the IC_{50} (*modeled*), (3) the total receptor concentration $[R]_{total}$ (*quantified in anti-GFP Western blots*), and (4) the affinity of the receptor for the labeled ligand K_d (*experimentally determined prior*). Substitutions of these parameters allows to solve the growth term equation for K_i as its root, and hence, the estimation of the binding affinity of MIK2 for the unlabeled cSCOOP1 peptides. The values of K_d and K_i have the same dimensions and can therefore be directly compared to each other.

6 Contributions

Dr. Vahid Fallahzadeh-Mamaghani conducted the original purification of the SLPA^{Xap} protein, leading to its mass spectrometric identification as a candidate. The mass spectrometric analysis was performed by services of the *Proteome Center Tuebingen* (*University of Tübingen;* Germany).

Dr. Mark Stahl was instrumental in the mass spectrometric identification of the purified cSCOOP1 peptide (shown in *Fig.5C-D*). He was also involved in the quantification of mature peptides in *in vitro* digests of synthetic pro-SCOOP analogs with pathogen proteases (shown in *Fig.12D*).

Dr. Yan Wang performed co-IP experiments (shown in *Fig.11A*) with provided plants and peptides. She was so kind to share vector graphics templates depicting LRR-RLK proteins for use in model cartoons (shown in *Fig.11* and *Fig.13*).

Dr. Judith Fliegmann took care of the genotypic confirmation of *mik2* null-mutant alleles.

Prof. Dr. Georg Felix provided *BLAST* results that were used in sequence analysis of SCOOP peptides (shown in *Fig.6B*).

The generation of *Δslpa* deletion strains in Xap, and scouting experiments for some of the phenomena described in chapter 3.1 were done by myself (Louis-Philippe Maier) already as part of my *Master of Science* thesis under Prof. Dr. Georg Felix' supervision. Crucially, these experiments were repeated with the inclusion of critical controls and new constructs. The work described was further corroborated by additional experimental procedures and analyses.

7 Acknowledgments

This work was funded by the *Deutsche Forschungsgemeinschaft* (*DFG*, *German Research Foundation*) - project number 441178209. Furthermore, this work was supported by the *Ministry for Science, Research and Art of Baden-Wuerttemberg* (Az: 7533-30-20/1).

We thank Dr. Sandra Schwarz (*Universitätsklinikum Tübingen;* Germany) for providing colonies of *P. aeruginosa* - PA14, and Dr. Laurent Noël (*Centre National de la Recherche Scientifique;* France) and the lab of Prof. Dr. Thomas Lahaye (*Universität Tübingen;* Germany) for having shared strains of *Xanthomonas* spp. . We are also grateful to Dr. Robert Morbitzer from the Lahaye lab, who kindly provided many *Golden Gate* cloning modules for our use.

Anna Sophia Moormann, Marvin Noss, and Adrian Vojtassak, are noteworthy for their work in our laboratory. Their *Bachelor of Science* theses and internships have contributed to our understanding of the perception system. I apologize for all of the futile experiments and screens I had suggested. Your had persevered nonetheless..!

Dr. Yan Wang, Dr. Maike Osterhus (née Lammers), Dr. Lei Wang, Nga Pham, plus my other peers and cherished former colleagues, are thanked for their support, collaboration, and all the shenanigans and fun activities that we were able to enjoy together at the department.

I am indebted and much obliged to Prof. Dr. Georg Felix and Prof. Dr. Thorsten Nürnberger for their supervision of my *PhD*-studies. Their guidance and advice has always been appreciated. Considering all of my shortcomings, I am particularly grateful for their patience and lenience.

Prof. Dr Andreas Schaller is thanked for his help with the conceptualization of the research project, the collaborative application for funding, and the prolific discussions.

Dr. Judith Fliegmann is thanked for encouragement, support and collaborative work on recent manuscripts. Her organizing ability is dearly missed in other laboratories.

I am grateful to Ilonka Bock, Petra Neumann, and Caterina Brancato who helped with general plant work and maintenance of plant cell cultures. Without their continued efforts, the purification of the cSCOOP1 precursor might not have been possible.

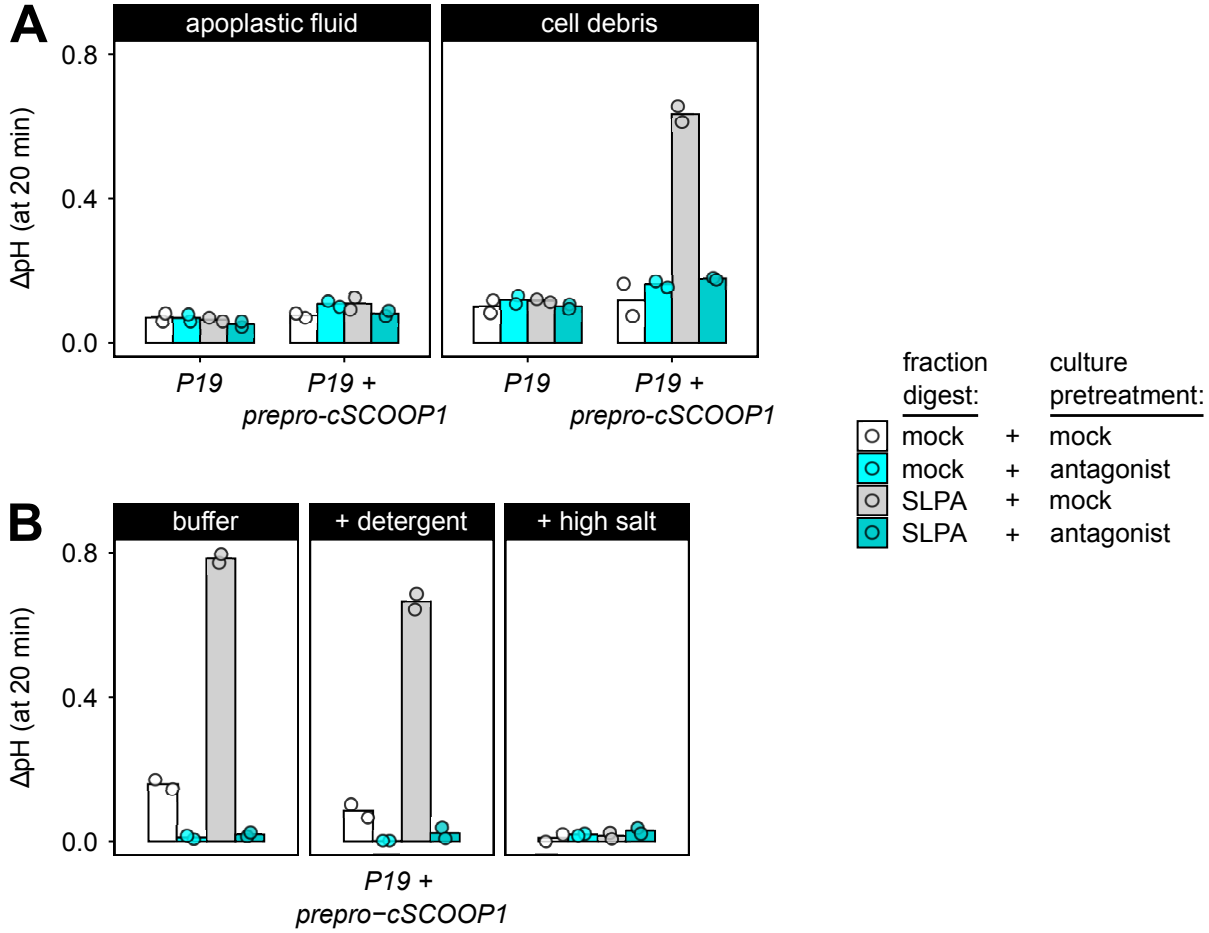
The entire team of the *ZMBP* plant cultivation facilities is thanked for having continuously provided excellent plant growth conditions and abundant amounts of material, even for my excessive experiments. Their work is crucial for everyone's research at the institute. The *ZMBP* workshop and our electricians, in particular Michael Wedel, are thanked for their help with technical issues. They have fixed many of our broken devices in urgent need of repair. I am thankful also to Dieter Steinmetz, and Dr. Joachim Kilian for their help with matters concerning *IT* and maintenance of our gas chromatographic systems, respectively.

Starting with my *Master of Science* studies at the *ZMBP*, I had the pleasure to spend almost a decade of my early career at this excellent research institute. I was able to experience the inner workings of the *Plant Biochemistry Department*, and beyond, in various research groups. The inclusive atmosphere, the flat hierarchy, and the countless scientifically stimulating discussions were much appreciated. I hope that future generations of researchers at the *ZMBP* will remain attached to its current values, such that other (aspiring) scientists to come will be able to benefit from the same experience.

8 Supplementary

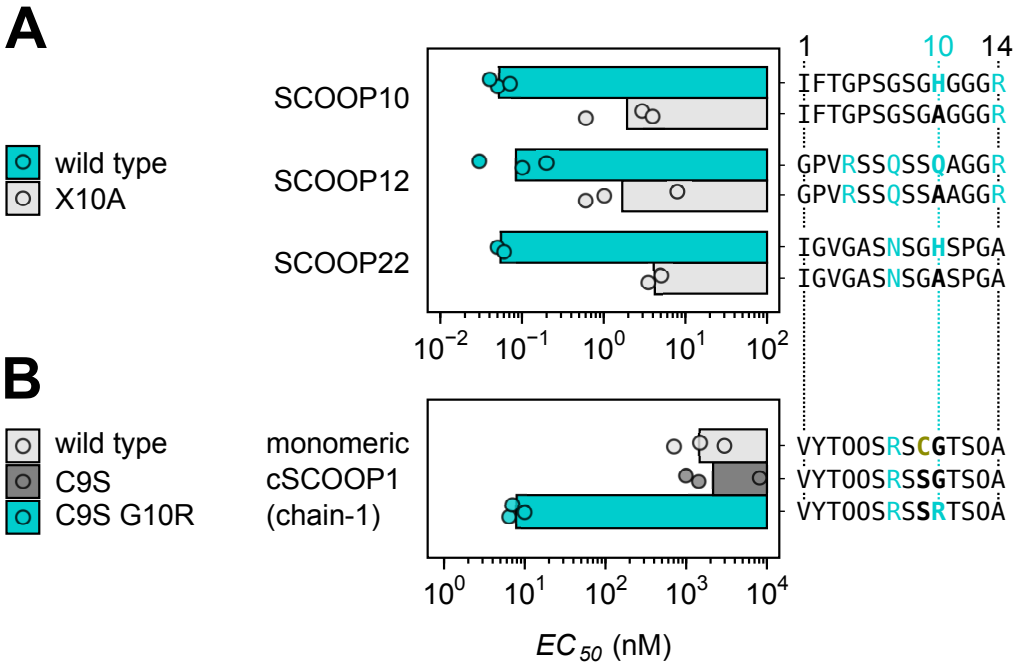
8.1 Supplementary results

Suppl.Fig.1: Alignment of cSCOOP prepro-peptides in *A. thaliana*. Alignment of cSCOOP precursors encoded by the genome of *A. thaliana* (Col-0). The active peptide motif is indicated in black bold letters, and the characteristic S-x-S/T motif is highlighted in magenta color. Slashes (/) indicate predicted processing sites of signal peptidases, N-terminal of which are the signal peptides that are to be removed during co-translational secretion. Dashes (-) are gaps in the alignment. *AGI* locus-IDs (TAIR10) are displayed in brackets behind each prepro-peptide name.

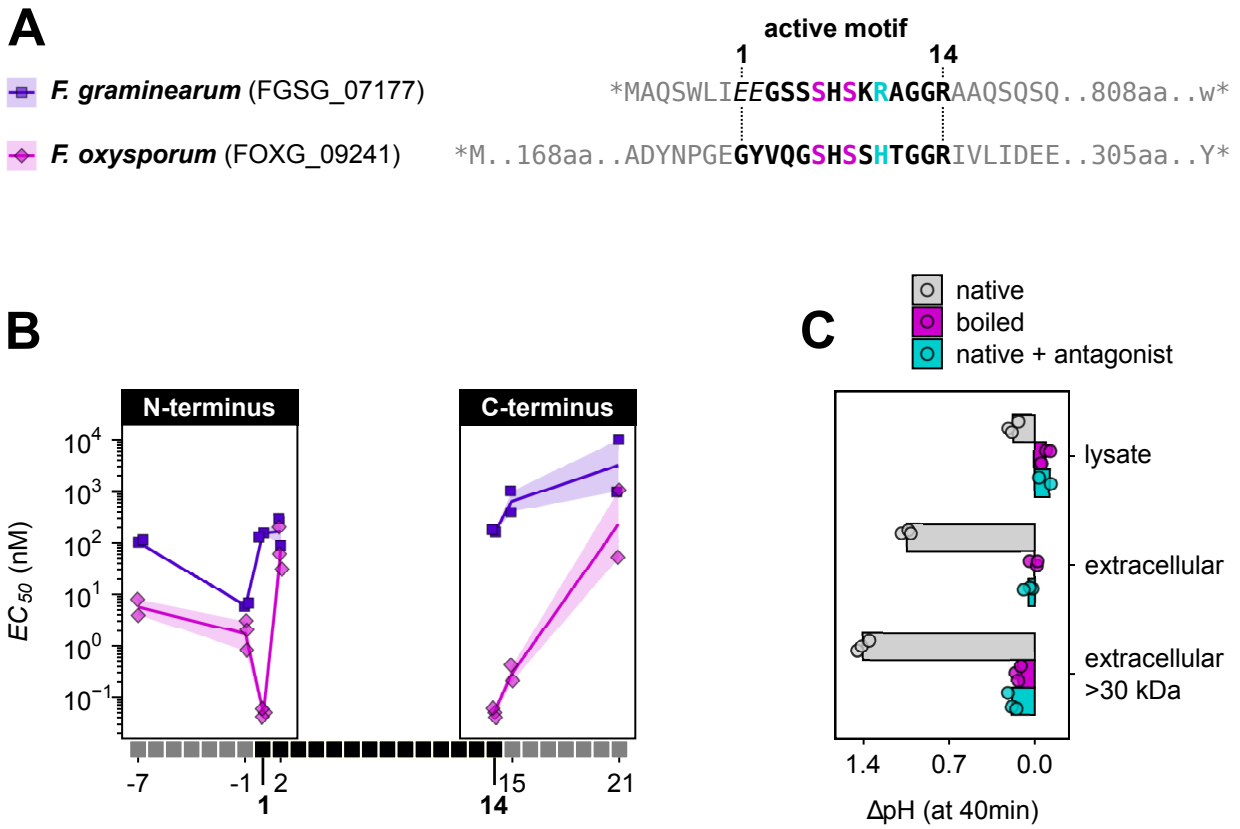


Suppl. Fig.2: Localization of pro-cSCOOP1 according to SLPA^{Xap}-dependent production of the immunogenic cSCOOP1 peptide. The pro-peptide is neither membrane-integral, nor soluble in leaf apoplastic fluid; it appears to bind to cell wall fractions *via* ionic interactions.
Fractionation: Constructs of prepro-cSCOOP1 under control of the *CaMV35S* promoter were expressed for 3 days in transiently transformed *N. benthamiana* leaves. Apoplastic fluid was recovered from leaves by vacuum-infiltration of water (1 × 3 min) and subsequent centrifugation at 500 × *g* and 4 °C for > 15 min. Apoplastic fluid and extracted leaves were frozen in N_2(fl.), the latter were cryogenically ground into a fine powder. The frozen leaf powder was further resuspended in buffer (50 mM MES, pH 5.5, 100 mM NaCl) and centrifuged at 5 000 × *g* and 4 °C for 5 min. The supernatant was discarded and the cell debris pellet was resuspended and washed an additional 2 times with the same buffer. Apoplastic fluid was thawed and spiked with buffer stocks to adjust its composition to the same as the cell-debris resuspension.
Digest: Apoplastic fluid and cell debris resuspensions were treated with 10 mM $CaCl_2$ and either 10 nM SLPA^{Xap}, or mock, and incubated for 15 min on a rotary shaker. The reaction was stopped by boiling of digests at 95 °C for 30 min. Insoluble material (protein precipitates and cell debris) were removed by centrifugation at > 14000 × *g* and 4 °C for 30 min.
Peptide extraction: Supernatants were adjusted to pH ~ 2 with TFA, loaded on C18 RP-SPE columns, washed with 0.1% (v/v) TFA, and eluted with ethanol [30% and 60% steps, pooled]. Eluates were dried in a centrifugal sample concentrator under vacuum, and re-eluted in water.
Activity assay: Peptide fractions were tested at concentrations equivalent to < 50% (w/v) of leaf fresh weight per ml of cell culture. *A. thaliana* cell suspension cultures were used either naive, or pretreated with 1 µM of the antagonistic cSCOOP1 (1_13) peptide, to confirm MIK2-dependency of the measured peptide activity. (*For an introduction to the antagonist, see Fig.11.*)
Results: (**A**) MIK2-dependent Immunogenic activity was generated when cell debris, but not apoplastic fluid, of prepro-cSCOOP1-expressing leaves was digested with SLPA^{Xap}. This suggests that

pro-cSCOOP1 was not soluble in the plant apoplast, but bound to plant material. (**B**) Cell debris was washed either with the same buffer as before, or with detergent - (+ 1% Tween-20) or salt - spiked (+ > 1M of each NaCl and $CaCl_2$) buffers. Washed cell debris was subsequently washed an additional 2 times with a large excess of the original buffer for equilibration, to guarantee similar conditions in the subsequent digest with $SLPA^{Xap}$. Salts, but not detergent, were able to wash pro-cSCOOP1 away from the debris. This implies that prepro-cSCOOPs' original characterization as transmembrane-domain - containing peptides[77] is incorrect; pro-cSCOOP1 did not associate with the membrane fraction that was visibly washed away with detergent. Instead, it appears that the positively charged pro-peptide bound to insoluble plant material *via* ionic interactions. The white granular material left after detergent washes was a cell wall fraction. Hence, pro-cSCOOP1 seems to bind tightly to insoluble components of the plant cell wall. Binding to poly-anionic pectins of the cell wall is likely. Bars indicate arithmetic means of 2 fractionation replicates. Each experiment was repeated at least two times, independently, and with similar results, regardless of whether non-tagged constructs, or recombinant versions with C-terminal HiBiT tags were used for expression.



Suppl.Fig.3: A conserved positive charge at position 10 of the active peptide motif is required for sensitive perception of linear SCOOPs. Corresponding mutations in cSCOOP1's chain-1 turn the inactive monomer into a moderately active linear SCOOP. Quantification of the immunogenic activity of synthetic SCOOP and cSCOOP peptides in *A. thaliana* cell suspension cultures. Each data point represent an EC_{50} estimate from an independent dose-response experiment (extracellular alkalization at 15 min post treatment); bars depict geometric means of 2 to 3 repetitions. (**A**) Positively charged aa residues (highlighted in cyan color) at position 10 were replaced by alanine, resulting in a consistent decrease of activity by 2 orders of magnitude. (**B**) cSCOOP1's chain-1 is all but inactive in its monomeric form. The same is true for a version of chain-1 in which the disulfide-bridging cysteine at position 9 is replaced by serine. Additional substitution of the glycine at position 10 for a positively charged residue (as it the case in linear SCOOPs) increases its immunogenic activity by more than 2 orders of magnitude. *Note that for the wild type SCOOP12 (1_14) peptide, the shown data points are of the same experiments as in Fig.8B.*



Suppl.Fig.4: *Fusarium* sp. does not contain active SCOOP-like peptides, and synthetic peptides have the same length requirements for activity as plant SCOOPs. (**A**) Alignment of two hypothetical proteins of *Fusarium* spp., encoding transcription factors with predicted nuclear localization. The active motif is highlighted in black bold letters; the SCOOP-characteristic S-x-S/T motif at position 6 - 8 and positive charge at position 10 are indicated in magenta and cyan color, respectively. The non-canonical acidic terminus of the putative peptide fragment of *F. graminearum*, expected to be causative of lower activity when compared to the ortholog in *F. oxysporum*, is indicated in italic letters. Protein termini and distance to them are indicated with asterisks (*) and numbers, respectively. (**B**) Mapping of the active peptide motif, with the same strategy, as done for genuine plant SCOOPs in *Fig.8*. Each data point represents an EC_{50} estimate from an independent dose-response experiment in *A. thaliana* cell suspension cultures at 15 min post treatment. Ribbons indicate total data spread of 2-3 experiments. (**C**) Activity of native or heat-treated fractions of *F. graminearum* in *A. thaliana* cell suspension cultures at a concentration equivalent to 0.4% (w/v) of fungal biomass. Plant cultures were used either naive, or pretreated with 1 µM of the antagonistic cSCOOP1 (1_13) peptide to assess MIK2-dependency of observed responses. *Preparation of fractions:* 5-day old mycelia were harvested from solid media, avoiding damage to the fungal cells, washed with water (extracellular fraction), flash-frozen, and cryogenically ground into a fine powder. The fungal powder was resuspended in ice-cold water, extracted for 15 min at 4°C, and subsequently centrifuged for the precipitation of debris (*i.e.* for removal of bulk chitin). The supernatant (lysate) was checked for activity. The extracellular fraction was either used as is, or selected for larger extracellular proteins through ultrafiltration (MWCO of 30 kDa).

8.2 Abbreviations

aa (unit)	amino acids; sequence length
$A_{214\,nm}$ (chromatography)	absorbance at 214 nm wavelength; absorption maximum of peptide bonds
AEC (chromatography)	anion exchange chromatography
AGI (database)	*Arabidopsis genome identifier;* TAIR10 version of annotations; *https://www.arabidopsis.org*
AIC (statistics)	*Akaike's information criterion*; relative estimate of the amount of information lost by application of a statistical model in the description of observed data (the lower or more negative, the better)
ANOVA (statistics)	*analysis of variance*; a test for significant differences among groups
BAK1	*brassinosteroid insensitive 1 - associated receptor kinase 1*; a SERK
BAM (secretion)	*barrel assembly machinery;* protein complex involved in secretion of autotransporter proteins (type-*Va* secretion system) through the outer membrane of Gram-negative bacteria
bp (unit)	nucleotide base pairs; sequence length
BSA (chemical)	*bovine serum albumin* (*Bos taurus*); a blood protein fraction used for blocking of non-specifically binding surfaces in biochemical assays
C-terminus	carboxylate terminus of peptides / proteins
CEC (chromatography)	cation exchange chromatography
CI (statistics)	confidence interval
co-IP	co-immunoprecipitation
cSCOOP	*cyclic / cystine-bridged SCOOPs;* a family of SCOOPs
CTNIPs (syn.: SCREWs)	*CTNIP* (aa motif) – containing peptides (synonymous: *small phyto-cytokines regulating defense and water loss*); plant-secreted family of pre(pro)-peptides
d (unit)	days
DAMP	*damage-associated molecular pattern*
DTT (chemical)	dithiothreitol
EC_{50} (constant)	concentration of agonist to achieve half-maximum amplitude of response, or of labeled ligand to achieve 50% receptor occupancy
EFR (*efr-1*)	*EF-Tu receptor*; LRR-RLK; receptor of elf18 (*null mutant allele*)
elf18	N-terminally acetylated, 18 aa peptide of bacterial *elongation factor-Tu* (EF-Tu) protein
eMAX	*enigmatic MAMP in Xanthomonas*; a proteinaceous elicitor associated with the lysate of *Xanthomonas* bacteria
ENA (database)	*European nucleotide archive; https://www.ebi.ac.uk/ena/browser/home*
ESI (MS)	electro-spray ionization
ETI	effector-triggered immunity
EWR1	*enhancer of vascular wilt resistance 1*; a small family of SCOOP-like prepro-peptides with an atypical peptide motif
FER	*feronia;* malectin-RLK; receptor of RALFs; involved in development and (negative) regulation of PTI

FID (chromatography)	flame-ionization detector (GC)
flg22	22 aa peptide of bacterial flagellin protein
FLS2 (*fls2*)	*flagellin-sensing 2;* LRR-RLK; receptor of flg22 (*null mutant allele*)
FPLC (chromatography)	fast-protein liquid chromatography
$F_{x, y}$ (statistics)	*F-statistic* value of a 2 / 3-way *ANOVA* model, with *x* and *y* degrees of freedom
GC (chromatography)	gas chromatography
GFP	*green fluorescent protein* (*Aequorea victoria*); used as a protein tag
GUS	*β-glucoronidase* (*E. coli*); a reporter gene, rarely also used as a protein tag
h (unit)	hours
HiBiT	small peptide fragment of nano-luciferase (*Oplophorus gracilirostris*); used as a protein/peptide tag
HIC (chromatography)	hydrophobic interaction chromatography
HPLC (chromatography)	high-performance liquid chromatography
HSL3 (syn.: NUT)	*HAESA-like 3* (synonymous: *plant screw unresponsive receptor*); LRR-RLK; receptor of CTNIPs; involved in positive regulation of stomatal aperture and PTI
IC_{50} (constant)	concentration of antagonist to achieve 50% inhibition of a response amplitude, or of non-labeled ligand to displace 50% of receptor-bound labeled ligand
ID	database identifier of a gene locus, or protein
IgA1	*immunoglobulin-A1*; human antibody
IP	immunoprecipitation
K_d (constant)	*dissociation constant* of a labeled ligand; a measure of receptor affinity
K_i (constant)	*dissociation constant* of a non-labeled ligand; a measure of receptor affinity and equivalent to K_d values
LRR	*leucine-rich repeat;* a repetitive protein domain
M (unit; in context of MS)	mass of an analyte (MS)
m/z (unit)	mass-to-charge ratio of an analyte (MS)
mAU (unit)	milli-absorbance units (chromatography)
MES (chemical)	2-(N-morpholino)-ethane-sulfonic acid
MIK2 (*mik2-1; mik2-2*)	*male discoverer 1 - interacting receptor-like kinase 2* (misnomer; not involved in reproduction); LRR-RLK; receptor of SCOOPs (*null mutant alleles*)
min (unit)	minutes
MS	mass spectrometry; detection of non-fragmented precursor ions
MS/MS	tandem mass spectrometry; analysis of ion fragmentation spectra
MWCO	molecular weight cut-off (kDa); assumes globular protein shape
NADPH oxidases	*nicotinamide adenine dinucleotide phosphate* (reduced) - dependent plasma membrane localized enzymes that reduce apoplastic H_2O to ROS
N-terminus	amine terminus of peptides / proteins
OD_{600} (unit)	optical density (*i.e.* absorbance) at a wavelength of 600 nm
P (statistics)	*probability* value
P19	19 kDa transcriptional silencing - suppressor protein (*Tombusvirus*); used in

	transient expression in *N. benthamiana*
P_{adj} (statistics)	probability value, adjusted for repeated testing
PAGE (chromatography)	polyacrylamide-gel electrophoresis
PAMP (synonymous: MAMP)	*pathogen-associated molecular pattern*
PAR	*protease-activated receptor;* (4×) G-protein coupled receptors; found in mammals
PEPRs	*plant elicitor peptide 1 receptors*; (2×) LRR-RLKs; receptors of PEPs; involved in positive regulation of PTI and DAMP signaling
PEPs	*plant elicitor peptide 1 - like peptides*; non-secreted family of plant pro-peptides, that are released upon loss of plasma membrane integrity
peptide	mature, fully-active peptide
pI (constant)	*isoelectric point;* pH value at which a peptide / proteins predominantly exists in a zwitterionic state of a net-charge of 0
PIPs	*PAMP-induced peptides;* plant-secreted family of pre(pro)-peptides
prepro-peptide	signal peptide - containing pro-peptide, N-terminal signal peptide is proteolytically removed during secretion
pro-peptide	pro-domain - containing precursor of a peptide, whose release and activation requires processing
PRR	*pattern-recognition receptor;* a receptor that is involved in the perception of PAMPs or DAMPs
PSKRs	*phytosulfokine receptors;* (2 ×) LRR-RLKs; receptors of PSKs; involved in development and negative regulation of PTI and cell death
PSKs	*phytosulfokines;* plant-secreted family of di-tyrosyl-sulfated prepro-peptides
PSY1R	*plant peptide containing sulfated tyrosine 1 - receptor;* LRR-RLK; receptor of PSYs; involved in development and negative regulation of PTI and cell death
PSYs	*plant peptide containing sulfated tyrosine 1 - like peptides;* plant-secreted family of tyrosyl-sulfated prepro-peptides
PTI	*pattern-triggered immunity*
pv.	pathovar
Q-ToF (MS)	quadrupol - time-of-flight mass analyzer (MS)
R^2 (statistics)	*coefficient of determination*; indicates the total amount of observed variation described by a statistical model (1 ≙ 100%)
R^2_{adj} (statistics)	*coefficient of determination*, adjusted for repeated testing
R^2_{McF} (statistics)	McFadden's *pseudo* R^2*;* alternative estimate of statistical model fitness for non-linear regression models
RALFs	*rapid alkalinization factors*; plant-secreted family of disulfide-bridges - containing prepro-peptides
RLK	*receptor-like kinase* (in plants, usually: serine / threonine / tyrosine kinases)
RLK7	*receptor-like kinase 7;* LRR-RLK; involved in positive regulation of PTI
RLP1 (*rlp1-2*)	*receptor-like protein 1;* LRR-receptor like protein without kinase domain; receptor of eMAX (*null mutant allele*)
RLU (unit)	*relative light unit;* an arbitrary measure of luminescence
ROS	*reactive oxygen species*; primarily the O_2^- radical and H_2O_2
RP (chromatography)	reversed phase chromatography

RT	room temperature (22-24 °C)
SALK (resource & database)	*Salk Institute Genomic Analysis Laboratory;* http://signal.salk.edu/
SCOOPs	*serine-rich endogenous peptides*; plant-secreted super-family of prepro-peptides, specific to the *Brassicaceae* plant family
SDS (chemical)	sodium dodecyl sulfate
SEC (secretion)	secretory machinery; protein complex involved in prokaryotic secretion of pre-proteins (through the inner membrane of Gram-negative bacteria); also involved in eukaryotic secretion
SEC (chromatography)	size exclusion chromatography
SERK	*somatic embryogenesis receptor kinase;* a family of LRR-RLKs; co-receptors of receptor LRR-RLKs
SLPA (*Δslpa*)	*subtilisin-like protease autotransporter;* family of bacteria-secreted extracellular proteases (*isogenic deletion mutant*)
sp.	species (singular)
SPATE	*serine-protease autotransporter of the Enterobacteriaceae;* family of bacteria-secreted extracellular proteases
SPE (chromatography)	solid phase extraction; a chromatographic technique
spp.	species (plural)
TFA (chemical)	trifluoro-acetic acid
Tris (chemical)	tris-(hydroxymethyl)-amino-methane
Tukey's HSD (statistics)	Tukey's *honestly significant difference;* a parametric *post-hoc* test to identify significantly different groups
U (in context of activity)	activity unit; 1 unit is equivalent to the amount of active compound that achieves half-maximum extracellular alkalization in 1 ml of *A. thaliana* cell suspension culture (related to EC_{50})
UPLC (chromatography)	ultra-performance liquid chromatography
Xap	*Xanthomonas arboricola* pv. pruni; a prune tree - specialized pathovar of *Xanthomonas*
Δ(X-Y) (allele)	indicates a deletion of encoded aa residues from *X to Y*
ΔpH (response)	total change of pH over a given amount of time (response amplitude)

9 References

1. Gómez-Gómez, L. & Boller, T. Flagellin perception: a paradigm for innate immunity. *Trends in plant science* **7**, 251–256 (2002).
2. Nürnberger, T., Brunner, F., Kemmerling, B. & Piater, L. Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews* **198**, 249–266 (2004).
3. Zipfel, C. & Felix, G. Plants and animals: a different taste for microbes? *Current Opinion in Plant Biology* **8**, 353–360 (2005).
4. Jones, J. D. G., Vance, R. E. & Dangl, J. L. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, aaf6395 (2016).
5. Spoel, S. H. & Dong, X. How do plants achieve immunity? Defence without specialized immune cells. *Nat Rev Immunol* **12**, 89–100 (2012).
6. Albert, I., Hua, C., Nürnberger, T., Pruitt, R. N. & Zhang, L. Surface sensor systems in plant immunity. *Plant Physiology* **182**, 1582–1596 (2020).
7. Afzal, A. J., Wood, A. J. & Lightfoot, D. A. Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *MPMI* **21**, 507–517 (2008).
8. Hohmann, U., Lau, K. & Hothorn, M. The structural basis of ligand perception and signal activation by receptor kinases. *Annu Rev Plant Biol* **68**, 109–137 (2017).
9. Matzinger, P. Tolerance, danger, and the extended family. *Annual Review in Immunology* **12**, 991–1045 (1994).
10. Gust, A. A., Pruitt, R. & Nürnberger, T. Sensing danger: Key to activating plant immunity. *Trends in Plant Science* **22**, 779–791 (2017).
11. Boller, T. & Felix, G. A Renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology* **60**, 379–406 (2009).
12. Kunze, G. *et al.* The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *The Plant Cell Online* **16**, 3496–3507 (2004).
13. Felix, G., Duran, J. D., Volko, S. & Boller, T. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal* **18**, 265–276 (1999).
14. Zipfel, C. *et al.* Perception of the bacterial PAMP EF-Tu by the receptor EFR testricts Agrobacterium-mediated transformation. *Cell* **125**, 749–760 (2006).
15. Gómez-Gómez, L. & Boller, T. FLS2: an LRR receptor–like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell* **5**, 1003–1011 (2000).
16. Tanaka, K. & Heil, M. Damage-associated molecular patterns (DAMPs) in plant innate immunity: applying the danger model and evolutionary perspectives. *Annu. Rev. Phytopathol.* **59**, 53–75 (2021).

17. Yamaguchi, Y. & Huffaker, A. Endogenous peptide elicitors in higher plants. *Current Opinion in Plant Biology* **14**, 351–357 (2011).

18. Hou, S., Liu, D. & He, P. Phytocytokines function as immunological modulators of plant immunity. *Stress Biology* **1**, 8 (2021).

19. Rhodes, J. *et al.* Perception of a conserved family of plant signalling peptides by the receptor kinase HSL3. *eLife* **11**, e74687 (2022).

20. Liu, Z. *et al.* Phytocytokine signalling reopens stomata in plant immunity and water loss. *Nature* **605**, 332–339 (2022).

21. Huffaker, A., Pearce, G. & Ryan, C. A. An endogenous peptide signal in *Arabidopsis* activates components of the innate immune response. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10098–10103 (2006).

22. Yamaguchi, Y., Pearce, G. & Ryan, C. A. The cell surface leucine-rich repeat receptor for *At*Pep1, an endogenous peptide elicitor in *Arabidopsis*, is functional in transgenic tobacco cells. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10104–10109 (2006).

23. Huffaker, A. & Ryan, C. A. Endogenous peptide defense signals in *Arabidopsis* differentially amplify signaling for the innate immune response. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10732–10736 (2007).

24. Krol, E. *et al.* Perception of the *Arabidopsis* danger signal peptide 1 involves the pattern recognition receptor *At*PEPR1 and its close homologue *At*PEPR2. *Journal of Biological Chemistry* **285**, 13471–13479 (2010).

25. Yamaguchi, Y., Huffaker, A., Bryan, A. C., Tax, F. E. & Ryan, C. A. PEPR2 is a second receptor for the Pep1 and Pep2 peptides and contributes to defense responses in *Arabidopsis*. *The Plant Cell* **22**, 508–522 (2010).

26. Bartels, S. *et al.* The family of Peps and their precursors in *Arabidopsis*: differential expression and localization but similar induction of pattern-triggered immune responses. *Journal of Experimental Botany* **64**, 5309–5321 (2013).

27. Hander, T. *et al.* Damage on plants activates Ca^{2+}-dependent metacaspases for release of immunomodulatory peptides. *Science* **363**, eaar7486 (2019).

28. Hou, S. *et al.* The secreted peptide PIP1 amplifies immunity through Receptor-Like Kinase 7. *PLoS Pathogens* **10**, e1004331 (2014).

29. Najafi, J. *et al.* PAMP-INDUCED SECRETED PEPTIDE 3 (PIP3) modulates immunity in *Arabidopsis thaliana*. *Journal of Experimental Botany* erz482 (2019).

30. Shen, J. *et al.* Secreted Peptide PIP1 Induces Stomatal Closure by Activation of Guard Cell Anion Channels in *Arabidopsis*. *Front. Plant Sci.* **11**, 1029 (2020).

31. Mosher, S. *et al.* The tyrosine-sulfated peptide receptors PSKR1 and PSY1R modify the immunity of *Arabidopsis* to biotrophic and necrotrophic pathogens in an antagonistic manner. *The Plant Journal* **73**, 469–482 (2013).

32. Shen, Y. & Diener, A. C. *Arabidopsis thaliana* RESISTANCE TO FUSARIUM OXYSPORUM 2 implicates tyrosine-sulfated peptide signaling in susceptibility and resistance to root infection. *PLoS Genetics* **9**, e1003525 (2013).

33. Igarashi, D., Tsuda, K. & Katagiri, F. The peptide growth factor, phytosulfokine, attenuates pattern-triggered immunity. *The Plant Journal* **71**, 194–204 (2012).

34. Pruitt, R. N. *et al.* The rice immune receptor XA21 recognizes a tyrosine-sulfated protein from a Gram-negative bacterium. *Science Advances* **1**, e1500245 (2015).

35. Stegmann, M. *et al.* The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* **355**, 287–289 (2017).

36. Gully, K. *et al.* The SCOOP12 peptide regulates defense response and root elongation in *Arabidopsis thaliana*. *J Exp Bot* **70**, 1349–1365 (2019).

37. Rhodes, J. *et al.* Perception of a divergent family of phytocytokines by the *Arabidopsis* receptor kinase MIK2. *Nat Commun* **12**, 705 (2021).

38. Hou, S. *et al.* The *Arabidopsis* MIK2 receptor elicits immunity by sensing a conserved signature from phytocytokines and microbes. *Nature Comm.* **12**, 5494 (2021).

39. Ma, X., Xu, G., He, P. & Shan, L. SERKing coreceptors for receptors. *Trends in Plant Science* **21**, 1017–1033 (2016).

40. Sun, Y. *et al.* Structural Basis for flg22-Induced Activation of the *Arabidopsis* FLS2-BAK1 Immune Complex. *Science* **342**, 624–628 (2013).

41. Santiago, J., Henzler, C. & Hothorn, M. Molecular mechanism for plant steroid receptor activation by somatic embryogenesis co-receptor kinases. *Science* **341**, 889–892 (2013).

42. Chinchilla, D. *et al.* A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* **448**, 497–500 (2007).

43. Hohmann, U. *et al.* Mechanistic basis for the activation of plant membrane receptor kinases by SERK-family coreceptors. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3488–3493 (2018).

44. Albert, M. & Felix, G. Chimeric receptors of the *Arabidopsis thaliana* pattern recognition receptors EFR and FLS2. *Plant Signaling & Behavior* **5**, 1430–1432 (2010).

45. Schulze, B. *et al.* Rapid heteromerization and phosphorylation of ligand-activated plant transmembrane receptors and their associated kinase BAK1. *Journal of Biological Chemistry* **285**, 9444–9451 (2010).

46. Schwessinger, B. *et al.* Phosphorylation-dependent differential regulation of plant growth, cell Death, and innate immunity by the regulatory receptor-like kinase BAK1. *PLoS Genetics* **7**, e1002046 (2011).

47. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).

48. Deslandes, L. & Rivas, S. Catch me if you can: bacterial effectors and plant targets. *Trends in Plant Science* **17**, 644–655 (2012).

49. Buscaill, P. & van der Hoorn, R. A. L. Defeated by the nines: nine extracellular strategies to avoid microbe-associated molecular patterns recognition in plants. *The Plant Cell* **33**, 2116–2130 (2021).

50. Jashni, M. K., Mehrabi, R., Collemare, J., Mesarich, C. H. & de Wit, P. J. G. M. The battle in the apoplast: further insights into the roles of proteases and their inhibitors in plant–pathogen interactions. *Front. Plant Sci.* **6**, (2015).

51. Figaj, D., Ambroziak, P., Przepiora, T. & Skorko-Glonek, J. The role of proteases in the virulence of plant pathogenic bacteria. *IJMS* **20**, 672 (2019).

52. Mooney, B. C., Mantz, M., Graciet, E. & Huesgen, P. F. Cutting the line: manipulation of plant immunity by bacterial type III effector proteases. *Journal of Experimental Botany* **72**, 3395–3409 (2021).

53. Mansfield, J. *et al.* Top 10 plant pathogenic bacteria in molecular plant pathology: Top 10 plant pathogenic bacteria. *Molecular Plant Pathology* **13**, 614–629 (2012).

54. An, S.-Q. *et al.* Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. *FEMS Microbiology Reviews* **44**, 1–32 (2020).

55. European and Mediterranean Plant Protection Organization. Xanthomonas arboricola pv. pruni: Diagnostics. *EPPO Bulletin* **36**, 129–133 (2006).

56. Fischer-Le Saux, M., Bonneau, S., Essakhi, S., Manceau, C. & Jacques, M.-A. Aggressive emerging pathovars of Xanthomonas arboricola represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Appl. and Environm. Microbiol.* **81**, 4651–4668 (2015).

57. Jehle, A. K., Fürst, U., Lipschis, M., Albert, M. & Felix, G. Perception of the novel MAMP eMax from different *Xanthomonas* species requires the *Arabidopsis* receptor-like protein ReMAX and the receptor kinase SOBIR. *Plant Signaling & Behavior* **8**, e27408 (2013).

58. Jehle, A. K. *et al.* The receptor-like protein ReMAX of Arabidopsis detects the microbe-associated molecular pattern eMax from Xanthomonas. *The Plant Cell* **25**, 2330–2340 (2013).

59. Albenne, C. & Ieva, R. Job contenders: roles of the β-barrel assembly machinery and the translocation and assembly module in autotransporter secretion. *Molecular Microbiology* **106**, 505–517 (2017).

60. Rojas-Lopez, M. *et al.* Identification of the autochaperone domain in the type Va secretion system (T5aSS): prevalent feature of autotransporters with a β-helical passenger. *Front. Microbiol.* **8**, 2607 (2018).

61. Shikata, S., Shimada, K., Ohnishi, Y., Horinouchi, S. & Beppu, T. Characterization of secretory intermediates of *Serratia marcescens* serine protease produced during its extracellular secretion from Escherichia coli cells. *The Journal of Biochemistry* **114**, 723–731 (1993).

62. Coutte, L., Antoine, R., Drobecq, H., Locht, C. & Jacob-Dubuisson, F. Subtilisin-like autotransporter serves as maturation protease in a bacterial secretion pathway. *European Molecular Biology Journal* **20**, 5040–5048 (2001).

63. Turner, D. P. J., Wooldridge, K. G. & Ala'Aldeen, D. A. A. Autotransported serine protease A of *Neisseria meningitidis*: an immunogenic, surface-exposed outer membrane, and secreted Protein. *Infection and Immunity* **70**, 4447–4461 (2002).

64. Ali, T., Oldfield, N. J., Wooldridge, K. G., Turner, D. P. & Ala'Aldeen, D. A. A. Functional Characterization of AasP, a maturation protease autotransporter protein of *Actinobacillus pleuropneumoniae*. *Infection and Immunity* **76**, 5608–5614 (2008).

65. Kida, Y., Taira, J., Yamamoto, T., Higashimoto, Y. & Kuwano, K. EprS, an autotransporter protein of *Pseudomonas aeruginosa*, possessing serine protease activity induces inflammatory responses through protease-activated receptors. *Cellular Microbiology* **15**, 1168–1181 (2013).

66. Del Tordello, E., Vacca, I., Ram, S., Rappuoli, R. & Serruto, D. *Neisseria meningitidis* NalP cleaves human complement C3, facilitating degradation of C3b and survival in human serum. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 427–432 (2014).

67. Doron, L. *et al.* Identification and characterization of fusolisin, the *Fusobacterium nucleatum* autotransporter serine protease. *PLoS ONE* **9**, e111329 (2014).

68. Campos, C. G., Borst, L. & Cotter, P. A. Characterization of BcaA, a putative classical autotransporter protein in *Burkholderia pseudomallei*. *Infect and Immun* **81**, 1121–1128 (2013).

69. Lo, R. Y., Strathdee, C. A., Shewen, P. E. & Cooney, B. J. Molecular studies of Ssa1, a serotype-specific antigen of *Pasteurella haemolytica* A1. *Infection and immunity* **59**, 3398–3406 (1991).

70. Van Ulsen, P. *et al.* A Neisserial autotransporter NalP modulating the processing of other autotransporters. *Mol. Microbiol.* **50**, 1017–1030 (2003).

71. Alamuri, P. & Mobley, H. L. T. A novel autotransporter of uropathogenic *Proteus mirabilis* is both a cytotoxin and an agglutinin. *Mol Microbiol* **68**, 997–1017 (2008).

72. Miyazaki, H., Yanagida, N., Horinouchi, S. & Beppu, T. Specific excretion into the medium of a serine protease from *Serratia marcescens*. *Agricul and Biol Chem* **54**, 2763–2765 (1990).

73. Vidarsson, G. *et al.* Working mechanism of immunoglobulin A1 (IgA1) protease: Cleavage of IgA1 antibody to *Neisseria meningitidis* PorA requires de novo synthesis of IgA1 protease. *Infect Immun* **73**, 6721–6726 (2005).

74. Dutta, P. R., Cappello, R., Navarro-García, F. & Nataro, J. P. Functional comparison of serine protease autotransporters of *Enterobacteriaceae*. *Infect Immun* **70**, 7105–7113 (2002).

75. Ruiz-Perez, F. & Nataro, J. P. Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cell Mol Life Sci* **71**, 745–770 (2014).

76. Mazar, J. & Cotter, P. A. Topology and maturation of filamentous haemagglutinin suggest a new model for two-partner secretion. *Molecular Microbiology* **62**, 641–654 (2006).

77. Yu, Z. *et al.* The *Brassicaceae*-specific secreted peptides, STMPs, function in plant growth and pathogen defense. *J Integr Plant Biol* **62**, 403–420 (2020).

78. Neukermans, J. *et al.* ARACINs, *Brassicaceae*-specific peptides exhibiting antifungal activities against necrotrophic nathogens in *Arabidopsis*. *Plant Physiol* **167**, 1017–1029 (2015).

79. Yadeta, K. A., Valkenburg, D.-J., Hanemian, M., Marco, Y. & Thomma, B. P. H. J. The *Brassicaceae*-specific EWR1 gene provides resistance to vascular wilt pathogens. *PLoS ONE* **9**, e88230 (2014).

80. Zhang, J. *et al.* EWR1 as a SCOOP peptide activates MIK2-dependent immunity in *Arabidopsis*. *Journal of Plant Interactions* **17**, 562–568 (2022).

81. Landeo Villanueva, S., Malvestiti, M. C., Ieperen, W., Joosten, M. H. A. J. & Kan, J. A. L. Red light imaging for programmed cell death visualization and quantification in plant–pathogen interactions. *Mol. Plant Pathol.* **22**, 361–372 (2021).

82. Meindl, T., Boller, T. & Felix, G. The plant wound hormone systemin binds with the N-terminal part to Its receptor but needs the C-terminal part to activate it. *The Plant Cell* **10**, 1561–1570 (1998).

83. Meindl, T., Boller, T. & Felix, G. The nacterial elicitor flagellin activates its receptor in tomato cells according to the address-message concept. *The Plant Cell* **12**, 1783–1794 (2000).

84. Schwyzer, R. ACTH: A Short introductory review. *Ann NY Acad Sci* **297**, 3–26 (1977).

85. Singh, V. & Phukan, U. J. Interaction of host and *Staphylococcus aureus* protease-system regulates virulence and pathogenicity. *Med Microbiol Immunol* **208**, 585–607 (2019).

86. Jurnecka, D., Man, P., Sebo, P. & Bumba, L. *Bordetella pertussis* and *Bordetella bronchiseptica* filamentous hemagglutinins are processed at different sites. *FEBS Open Bio* **8**, 1256–1266 (2018).

87. Nissinen, R. *et al.* The putative secreted serine protease Chp-7 is required for Full virulence and induction of a nonhost hypersensitive response by *Clavibacter michiganensis* subsp *sepedonicus*. *MPMI* **22**, 809–819 (2009).

88. Verma, R. K. & Teper, D. Immune recognition of the secreted serine protease ChpG restricts the host range of *Clavibacter michiganensis* from eggplant varieties. *Molecular Plant Pathology* **23**, 933–946 (2022).

89. Hwang, I. S. *et al.* An apoplastic effector Pat-1Cm of the Gram-positive bacterium *Clavibacter michiganensis* ects as both a pathogenicity factor and an immunity elicitor in plants. *Front Plant Sci* **13**, 888290 (2022).

90. Lu, Y., Hatsugai, N., Katagiri, F., Ishimaru, C. A. & Glazebrook, J. Putative serine protease effectors of *Clavibacter michiganensis* induce a hypersensitive response in the apoplast of *Nicotiana* species. *MPMI* **28**, 1216–1226 (2015).

91. Zheng, A. *et al.* The evolution and pathogenic mechanisms of the rice sheath blight pathogen. *Nat Commun* **4**, 1424 (2013).

92. Caro, M. del P. *et al.* The fungal subtilase AsES elicits a PTI-like defence response in *Arabidopsis thaliana* plants independently of its enzymatic activity. *Molecular Plant Pathology* **21**, 147–159 (2020).

93. Caro, M. D. P. *et al.* A fungal protease named AsES triggers antiviral immune responses and effectively restricts virus infection in arabidopsis and *Nicotiana benthamiana* plants. *Annals of Botany* **129**, 593–606 (2022).

94. Chalfoun, N. R. *et al.* Purification and characterization of AsES protein. *J Biol Chem* **288**, 14098–14113 (2013).

95. Cheng, Z. *et al.* Pathogen-secreted proteases activate a novel plant immune pathway. *Nature* **521**, 213–216 (2015).

96. Kapur, V., Majesky, M. W., Li, L. L., Black, R. A. & Musser, J. M. Cleavage of interleukin 1 beta (IL-1 beta) precursor to produce active IL-1 beta by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7676–7680 (1993).

97. LaRock, C. N. *et al.* IL-1β is an innate immune sensor of microbial proteolysis. *Sci. Immunol.* **1**, (2016).

98. Issa, N. *et al.* The circulating protease Persephone is an immune sensor for microbial proteolytic activities upstream of the Drosophila Toll pathway. *Molecular Cell* **69**, 539-550.e6 (2018).

99. de Zoete, M. R., Bouwman, L. I., Keestra, A. M. & van Putten, J. P. M. Cleavage and activation of a Toll-like receptor by microbial proteases. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4968–4973 (2011).

100. Ossovskaya, V. S. & Bunnett, N. W. Protease-Activated Receptors: Contribution to Physiology and Disease. *Physiological Reviews* **84**, 579–621 (2004).

101. Hollenberg, M. D. Proteinase-Activated Receptors. *Pharmacological Reviews* **54**, 203–217 (2002).

102. Shpacovitch, V., Feld, M., Bunnett, N. W. & Steinhoff, M. Protease-activated receptors: novel PARtners in innate immunity. *Trends in Immunology* **28**, 541–550 (2007).

103. Coleman, A. D. *et al.* The Arabidopsis leucine-rich repeat receptor-like kinase MIK2 is a crucial component of early immune responses to a fungal-derived elicitor. *New Phytol* **229**, 3453–3466 (2021).

104. Van der Does, D. *et al.* The Arabidopsis leucine-rich repeat receptor kinase MIK2/LRR-KISS connects cell wall integrity sensing, root growth and response to abiotic and biotic stresses. *PLoS Genetics* **13**, e1006832 (2017).

105. Guillou, M.-C. *et al.* The PROSCOOP10 gene encodes two extracellular hydroxylated peptides and impacts flowering time in *Arabidopsis*. *Plants* **11**, 3554 (2022).

106. Stahl, E. *et al.* The MIK2/SCOOP signaling system contributes to *Arabidopsis* resistance against herbivory by modulating jasmonate and indole glucosinolate biosynthesis. *Front Plant Sci* **13**, 852808 (2022).

107. Celorio-Mancera, M. de la P. *et al.* Sialome of a generalist Lepidopteran herbivore: identification of transcripts and proteins from *Helicoverpa armigera* labial salivary glands. *PLoS ONE* **6**, e26676 (2011).

108. Rivera-Vega, L. J., Galbraith, D. A., Grozinger, C. M. & Felton, G. W. Host plant driven transcriptome plasticity in the salivary glands of the cabbage looper (*Trichoplusia ni*). *PLoS ONE* **12**, e0182636 (2017).

109. Flor, H. H. Current status of the gene-for-gene concept. *Ann Rev of Phytopath* **9**, 275–296 (1971).

110. van der Hoorn, R. A. L. & Kamoun, S. From guard to decoy: A new model for perception of plant pathogen effectors. *Plant Cell* **20**, 2009–2017 (2008).

111. Shiu, S.-H. & Bleecker, A. B. Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10763–10768 (2001).

112. Chakraborty, S., Nguyen, B., Wasti, S. D. & Xu, G. Plant Leucine-Rich Repeat Receptor Kinase (LRR-RK): structure, ligand perception, and activation mechanism. *Molecules* **24**, 3081 (2019).

113. Binder, A. *et al.* A modular plasmid assembly kit for multigene expression, gene silencing and silencing rescue in plants. *PLoS ONE* **9**, e88218 (2014).

114. Hall, M. P. *et al.* Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem Biol* **7**, 1848–1857 (2012).

115. Dixon, A. S. *et al.* NanoLuc complementation reporter optimized for accurate measurement of protein interactions in cells. *ACS Chem Biol* **11**, 400–408 (2016).

116. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343–345 (2009).

117. Murillo, J. *et al.* Characterization of pPT23B, the plasmid involved in syringolide production by *Pseudomonas syringae* pv. tomato PT23. *Plasmid* **31**, 275–287 (1994).

118. May, M. J. & Leaver, C. J. Oxidative stimulation of glutathione synthesis in *Arabidopsis thaliana* suspension cultures. *Plant Phys* **103**, 621–627 (1993).

119. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**, 735–743 (1998).

120. Wang, T. *et al.* A receptor heteromer mediates the male perception of female attractants in plants. *Nature* **531**, 241–244 (2016).

121. Kovach, M. E. *et al.* Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**, 175–176 (1995).

122. Schulze, S. *et al.* Analysis of new type III effectors from *Xanthomonas* uncovers XopB and XopS as suppressors of plant immunity. *New Phytologist* **195**, 894–911 (2012).

123. Figurski, D. H. & Helinski, D. R. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1648–1652 (1979).

124. Butenko, M. A. *et al.* Tools and strategies to match peptide-ligand receptor pairs. *The Plant Cell* **26**, 1838–1847 (2014).

125. Teufel, F. *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol* **40**, 1023–1025 (2022).

126. Jones, P. *et al.* InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).

127. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acid Research* **30**, 3059–3066 (2002).

128. Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences* **89**, 10915–10919 (1992).

129. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

130. R Core Team. R: A language and environment for statistical computing. (2022).

131. Posit team. RStudio: Integrated Development Environment for R. (2022).

132. Wickham, H. *ggplot2: elegant graphics for data analysis*. (Springer-Verlag New York, 2016).

133. Wagih, O. ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* **33**, 3645–3647 (2017).

134. Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. *Biom. J.* **50**, 346–363 (2008).

135. Hill, A. V. The possible effects of the aggregation of the molecules of huemoglobin on its dissociation curves. *The Journal of Physiology* **40**, iv–vii (1910).

136. Richards, F. J. A flexible growth function for empirical use. *J Exp Bot* **10**, 290–301 (1959).

137. Giraldo, J., Vivas, N. M., Vila, E. & Badia, A. Assessing the (a)symmetry of concentration-effect curves: empirical versus mechanistic models. (2002).

138. Jarmoskaite, I., AlSadhan, I., Vaidyanathan, P. P. & Herschlag, D. How to measure and evaluate binding affinities. *eLife* **9**, e57264 (2020).

139. Wang, Z.-X. An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Letters* **360**, 111–114 (1995).