Proteome Analysis of Soybean Root Apoplast Combined with AlphaFold Prediction Reveal *Macrophomina phaseolina* Infection Strategies and Potential Targets for Engineering Resistance

Chetan Veeraganti Naveen Prakash[1#], Muthusaravanan Sivaramakrishnan[1#], Sakshi Goel[1], Vasudha Porwal[1], Pawan Kumar Amrate[2], Manoj Kumar Shrivastava[2], Balakumaran Chandrasekar[1]*

1. Department of Biological Sciences, Birla Institute of Technology and Sciences, Pilani Campus - 333031, Rajasthan, India

2. Department of Genetics and Plant Breeding, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur – 482 004, Madhya Pradesh, India

# equal contribution

* Corresponding author

Abstract

Macrophomina phaseolina (Tassi) Goid. is a hemibiotrophic pathogen that causes charcoal rot (CR) disease in various legumes, including soybean. To date, no reliable resistance gene sources have been identified in soybean or other legumes to combat *M. phaseolina*. Therefore, the identification of mechanistic targets is crucial for improving resistance against the pathogen. The apoplast is a critical region where intense molecular cross-talks occur between plants and pathogens, and the outcome of their interactions is determined in this compartment. Here, we employed label-free quantitative (LFQ) proteomics to investigate the dynamics of soybean root apoplast during *M. phaseolina* infection. We have detected several secreted proteins of *M. phaseolina* and differential regulation of soybean-secreted proteins in root apoplast during infections. Glycome analysis and callose deposition assays have revealed significant changes in soybean root cell wall compositions and potential polysaccharide targets of *M. phaseolina*. AlphaFold 2 (AF2) analysis has revealed several interesting sequence-unrelated structurally similar (SUSS) effectors and effectors with novel structural folds secreted by *M. phaseolina*. Furthermore, AlphaFold Multimer (AFM) analysis has revealed protease-inhibitor interactions displayed by candidate-secreted proteins from soybean and *M. phaseolina*. We have validated these interactions using molecular dynamics (MD) and competitive activity-based protein profiling (ABPP) approaches. Therefore, our work provides insights into Soybean-*M. phaseolina* interactions in the root apoplast and unveil potential candidates for engineering resistance.

Introduction

Macrophomina phaseolina (Tassi) Goid. is a devastating soil-borne fungal pathogen that causes charcoal rot (CR) disease or dry root rot (DRR) disease in several crops under high temperatures and drought stress conditions. *M. phaseolina* has been predicted to pose a serious threat to several legumes under changing climatic conditions (Pandey and Basandrai 2021). Among their various legume hosts, soybean (*Glycine max* L. Merril) is one of the most economically important crops that are vulnerable to CR disease with yield losses ranging from 70-100% (Amrate et al. 2023; Ortiz et al. 2023). During root colonization, *M. phaseolina* exhibits an initial mycelial phase followed by the microsclerotia stage, leading to characteristic charcoal rot disease symptoms (Chowdhury et al. 2017). The microsclerotia phase of the

pathogen plays a critical role in its pathogenesis. Therefore, it is crucial to identify suitable candidates in crop plants for engineering resistance against the pathogen.

Apoplast is an important compartment where the initial dialogue between the plant and pathogen occurs and the dynamics of apoplastic components determine the outcome of plant-pathogen interaction. All the components outside the cell membrane, including the cell wall, constitute the apoplast. The cell wall is a dynamic component composed of various cellulosic and hemicellulosic polysaccharides that serve as a protective barrier to resist pathogen invasion. Furthermore, at the pathogen infection sites, the cell wall is reinforced with appositions such as callose deposition that act as barriers to pathogen invasion (Hosseini-Zahani and Taheri 2023). Fungal pathogens employ several strategies for the successful colonization of their host plants (Buscaill and van der Hoorn 2021). They secrete carbohydrates-active enzymes (CAZymes) that are involved in plant cell wall degradation to gain entry into the host (Sivaramakrishnan et al. 2024). They also manipulate several host-derived factors that act as susceptible factors to favor pathogen colonization (Hong et al. 2024). Furthermore, fungi secrete effectors in apoplast to suppress host immune responses. Some effectors play a role in hijacking host defense machinery by inducing cell death following the inverse gene for gene interactions (Di et al. 2024). Compared to the biotrophic fungal pathogens, only a few necrotrophic effectors have been identified, and their biology is largely understudied (Shao et al. 2021). A large number of necrotrophic effectors are yet to be identified, and machine learning tools such as AlphaFold (AF) would assist in the discovery of novel effectors, which would aid in formulating efficient strategies to control the fungal pathogen (Seong and Krasileva 2023).

Protease-inhibitor interaction is another important aspect that occurs in the apoplast during fungal colonization and determines the outcome of the plant-pathogen interaction (Misas-Villamil et al. 2016). Plant-secreted proteases such as cysteine and serine proteases in the apoplast represent central hubs for plant immune responses (Godson and van der Hoorn 2021). These enzymes mediate the generation of immunogenic elicitors or directly cleave the proteins essential for pathogen virulence and thereby impart resistance in plants (Paulus et al. 2020; S. Wang et al. 2021). The activities of these immune proteases are regulated by the inhibitors that are produced by the pathogen or the plant itself (Wang et al. 2020). AlphaFold multimer (AFM) analysis is emerging as an important tool in plant biology for structural-level understanding and prediction of protein-inhibitor interactions at the plant-pathogen interface (Homma et al. 2023; Mooney and van der Hoorn 2024). AFM analysis has been instrumental in designing strategies for modulating apoplastic protease-inhibitor interactions for enhancing resistance against the oomycete pathogen *Phytophthora infestans* (Schuster et al. 2024). So far, reliable resistance gene sources have neither been identified in soybean nor in other legumes to combat *M. phaseolina*. Therefore, predicting apoplastic protease-inhibitor interactions would assist in formulating strategies for engineering resistance in these crops against the pathogen.

Proteomics is an important approach to investigate the dynamics of apoplast during pathogen infections. To date, the apoplastic proteome studies are largely focused on foliar pathogens, and the studies on root-invading fungal pathogens are limited. This is due to the fact that a large number of plants are needed to collect sufficient apoplastic fluids from roots for proteome analysis. Notably, the apoplastic proteome analysis of soybean roots and their dynamics during *M. phaseolina* has not been investigated. Here, in our study we have performed label-free quantitative proteomics of the apoplastic fluids collected from soybean

roots during *M. phaseolina* infection. In our proteome analysis, we detected several secreted proteins of *M. phaseolina* and differential regulation of several soybeans secreted proteins in root apoplast during infections. Glycome analysis and callose deposition assays have revealed significant changes in soybean root cell wall compositions and potential polysaccharide targets of *M. phaseolina* during colonization. AlphaFold 2 (AF2) analysis has revealed several interesting sequence-unrelated structurally similar (SUSS) effectors and effectors with novel structural folds secreted by *M. phaseolina*. Furthermore, AlphaFold Multimer (AFM) analysis has revealed protease-inhibitor interactions displayed by candidate-secreted proteins from soybean and *M. phaseolina*. We have validated these interactions using molecular dynamics (MD) and competitive activity-based protein profiling (ABPP) approaches. Therefore, our work provides insights into the soybean-*M. phaseolina* interaction in the root apoplast and unveil potential candidates for engineering resistance against *M. phaseolina.*

Material and methods

Isolation and identification of *Macrophomina phaseolina*

Infected root samples were collected from Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV), Jabalpur, Madhya Pradesh, India (23° 12' 47.808" N, 79° 57' 49.752" E, 301.5 m above mean sea level). Root bits were surface sterilized using 0.1% mercuric chloride and rinsed thrice with sterile Milli Q water. The surface sterilized bits were placed on Potato Dextrose Agar and incubated at 28℃. Hyphal tip method was followed to obtain pure cultures of the pathogen (Brown 1924). Molecular identification of the fungus was performed by Sanger sequencing of Internal Transcribed Spacer 1/4 (ITS 1/4) (Primers: ITS1: TCCGTAGGTGAACCTGCGG, ITS4: TCCTCCGCTTATTGATATGC). The obtained consensus sequences of forward and reverse sequencing were analysed in NCBI-BLAST. The phylogenetic tree was constructed with the neighbor-joining method using MEGA 11 software. The Maximum Composite Likelihood model was selected and branch support was calculated by 1,000 Bootstrap repetitions. The ITS1/4 sequence was submitted to NCBI Genbank with accession no. PQ807640.

Plant material

The soybean seeds of varieties JS 20-94, JS 20-34 and JS 20-98 were procured from Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV), Jabalpur, Madhya Pradesh, India and SL-955 was procured from Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, Punjab, India. These varieties were used for the pathogenicity test of the isolated *M. phaseolina* strain.

Root dip method of infection with *M. phaseolina* culture

Infection assay was performed as described by Nelson et al. 2021 with slight modifications. Briefly, the healthy soybean seeds were sterilised by chlorination for 16 hours (Sivaramakrishnan et al. 2023). Sterilized seeds were sown in double autoclaved soil mixture (Soilrite) and were grown in a plant growth chamber (28℃, 75% humidity, 16:8 hour day: night schedule, 6000 lux luminescence) for 2 weeks. These plants were washed in tap water and dipped in 25% inoculum (w/v) of actively growing mycelium of *M. phaseolina* (*M. phaseolina_Jbstrain1)* cultured in Potato Dextrose Broth. Milli Q water was used as control. Later, the plants were repotted and placed in a plant growth chamber (35℃, 65% humidity,

133 16:8-hour day: night schedule, 6000 lux luminescence). Plants were watered at 8-hour
134 intervals and the disease progression was recorded for 3 days.

135 **Microscopic analysis of soybean roots upon *M. phaseolina* infection**

136 Control and infected soybean roots were prepared for microscopic analysis as described by
137 Barrow (2004) with slight modifications. The roots collected at different stages of *M.*
138 *phaseolina* infection were subjected to overnight clearing of roots with 70% ethanol. The
139 cleared roots were boiled in 10% KOH at 100℃ for 1 hour. The roots were washed with tap
140 water, bleached with 10% alkaline H_20_2 for 45 minutes and acidified with 1% HCl for 2 minutes.
141 These cleared roots were mounted on a glass slide using acidified glycerol. Differential
142 Interference Contrast (DIC) imaging was performed using Fully Spectral Confocal Laser
143 Scanning Microscope (FSCLSM, Carl Zeiss LSM 880).

144 **Root apoplastic fluid extraction**

145 The root apoplastic fluids were collected employing vacuum infiltration and centrifugation
146 method (Chandrasekar et al. 2014, Chandrasekar et al. 2022). In brief, the washed and cut
147 pieces of soybean roots were submerged in ice-cold Milli Q water, and vacuum was applied.
148 The water infiltrated into the intercellular spaces of roots as the vacuum was slowly released.
149 The roots were then dried briefly and centrifuged at 1000 g and 4°C for 30 min using a swing
150 bucket rotor. The apoplastic fluid collected was used for proteome analysis and ABPP
151 experiments.

152 **Label-free quantitative proteomics of root apoplast using LC-MS/MS**

153 Apoplastic fluid was reduced with 5 mM TCEP, alkylated with 50 mM iodoacetamide, and then
154 digested with Trypsin (1:50, Trypsin/lysate ratio) for 16 h at 37 °C. Digests were cleaned using
155 a C18 silica cartridge to remove the salt and dried using a speed vac. The dried pellet was
156 resuspended in buffer A (2% acetonitrile, 0.1% formic acid). Experiments were performed on
157 an Easy-nlc-1000 system coupled with an Orbitrap Exploris mass spectrometer. 1ug of peptide
158 sample were loaded on C18 column of 15 cm length, 3.0µm Acclaim PepMap (Thermo Fisher
159 Scientific) and separated with a 0–40% gradient of buffer B (80% acetonitrile, 0.1% formic
160 acid) at a flow rate of 500 nl/min and injected for MS analysis. LC gradients were run for 110
161 minutes. MS1 spectra were acquired in the Orbitrap (Max IT= 60ms, AGQ target= 300%; RF
162 Lens= 70%; R= 60K, mass range= 375−1500 m/z; Profile data). Dynamic exclusion was
163 employed for 30s excluding all charge states for a given precursor. MS2 spectra were collected
164 for top 20 peptides (Max IT= 60ms, R= 15K, AGC target 100%). The mass spectrometer was
165 set in the positive ion mode. The mass spectrometry proteomics data have been deposited to
166 the ProteomeXchange Consortium via the PRIDE [1] partner repository with the dataset
167 identifier PXD059244.

168 **Proteome data analysis**

169 The mass spectrometer was operated using Proteome Discoverer (v2.5) software. For dual
170 Sequest and amanda search, the precursor and fragment mass tolerances were set at 10 ppm
171 and 0.02 Da, respectively. Both peptide spectrum match and protein false discovery rate were
172 set to 0.01 FDR. The obtained proteome data was analysed to predict the putative secretory
173 proteins using Signal P 6.0. The signal P domains containing proteins were then identified and
174 classified as CAZymes using dbCAN3 server (last accessed: June, 2024) using HMMER
175 dbCAN (E-value < 10−15, coverage > 0.35), DIAMOND (E-value < 10−102), and HMMER

dbCAN-sub (E-value < 10−15, coverage > 0.35) tools (Zhang et al., 2018). Only the proteins successfully annotated by at least two of the tools were considered CAZymes. The rest of the proteins were searched for their domain architecture under Pfam which is now hosted by InterPro (Mistry et al. 2021) (last accessed: July, 2024). The secreted proteins were subjected to EffectorP Fungi 3.0 algorithm for the prediction of effector candidates (Sperschneider and Dodds 2022). Volcano plots were generated using EnhancedVolcano package, the PCA was performed using factoextra package and Pearson's correlation was visualised using corrplot package in R software. Z-score values of the differentially regulated plant proteins were calculated using the formula $z= (X-\mu)/\sigma$ where X= individual abundance value (LFQ), μ= mean of the replicate abundance values (LFQ), σ= Standard deviation of the replicate abundance values (LFQ) and was illustrated as a heat map using TBtools-II v1.112. The sequences of functionally characterized AA9s were collected from UniProt database (The UniProt Consortium 2024) and a phylogenetic tree was constructed along with the detected members of AA9 with the neighbor-joining method in MEGA 11 software.

GO annotation of the proteome dataset

Gene Ontology (GO) analysis of the differentially regulated soybean apoplastic proteins based on molecular functions and biological processes was performed using g:GOSt module of g:Profiler with Fisher one-tailed test and Benjamini-Hochberg with an FDR threshold of ≤0.05 to correct for multiple testing. GO analysis fungal-secreted proteins in terms of molecular functions and biological processes was performed using FungiFun v2.2 with hypergeometric distribution test and Benjamini-Hochberg FDR threshold of ≤ 0.05.

FTIR spectroscopy of root AIR samples

FTIR analysis was performed using Fourier Transform Infrared Spectroscopy (Model: Bruker Alpha) over a range of 4,000 to 400 cm^{-1}. For each spectrum, 32 scans were performed at a resolution of 4 cm^{-1}. The spectra were baseline-corrected and smoothened using OPUS spectral processing software with default parameters. The functional groups were annotated based on corresponding wavenumbers.

Monosaccharide composition and glycosyl linkage analysis of root AIR samples

To determine the monosaccharide composition of the control and infected root samples, alditol acetate derivatization was performed (Foster et al. 2010; Pettolino et al. 2012). Initially, alcohol insoluble residue (AIR) samples of the roots were obtained by washing sequentially with 70% ethanol, chloroform:methanol (1:1), and acetone. Then, the AIR samples were hydrolyzed with 2.5 M trifluoroacetic acid (TFA), reduced with sodium borohydride ($NaBH_4$) in NH_4OH followed by acetylation. The alditol acetates were extracted using ethyl acetate and water by phase-separation centrifugation. The organic phase was collected, dried under N_2 and resuspended in acetone. The derivatised sugars were analysed by GC-MS on a 30 m × 0.25 mm SP-2380 Fused Silica Capillary Column (Supelco). Injection was at 80°C with a 3-min hold, ramped to 170°C at 30°C min^{-1}, and then to 240°C at 4°C min^{-1}, with a 20-min hold at the upper temperature. Helium flow was 1.5 mL min^{-1} with split-less injection. The electron impact mass spectrometry was performed with model no. TQ8040 CI/NCI (Shimadzu) at 150 eV and a source temperature of 250°C. The obtained chromatograms were annotated based on retention time and the mass spectrum of standard sugars.

Glycosidic linkage analysis was done as described in (Chandrasekar et al. 2022; Kim et al. 2020; Liu et al. 2015). Briefly, dry DMSO was added to AIR samples and kept for stirring

220 overnight at room temperature. NaOH/DMSO was added, and methylation was carried out by
221 adding methyl iodide under a Nitrogen environment for 4 hours with stirring. The reaction was
222 quenched with Milli Q water, bubbled with nitrogen, and dichloromethane (DCM) was added.
223 The DCM phase was collected, dried, and subjected to hydrolysis and alditol acetate
224 derivatization as followed in monosaccharide composition analysis to obtain partially
225 methylated alditol acetates. These derivatised sugars were analysed using similar parameters
226 as that of monosaccharide composition analysis with GC-MS. The glycosidic linkages were
227 assigned based on retention time and mass spectrum fragmentation patterns compared to the
228 CCRC spectral database (https://www.ccrc.uga.edu/specdb/ms/pmaa/pframe.html).

229 **Crystalline cellulose estimation**

230 Crystalline cellulose estimation of AIR samples of control and infected roots was performed
231 as described by (Pettolino et al. 2012). The pellet after the TFA hydrolysis was treated with
232 Updegraff reagent followed by Saeman hydrolysis. The samples were subjected to anthrone
233 assay to estimate the released glucose, and the readings were taken using a microtiter plate
234 reader at 625nm. Absolute quantification of the crystalline cellulose was calculated against a
235 standard curve plotted using various concentrations of glucose standards from 0-10 µg/ml.

236 **Root callose staining and its quantification**

237 Callose staining of the soybean roots was performed using 0.01% Aniline Blue as described
238 by Schenk and Schikora 2015 with few modifications. In brief, the soybean roots collected at
239 different stages of *M. phaseolina* infection were washed with Milli Q water to remove soil
240 particles and kept in acetic acid: ethanol (1:3) overnight. The roots were transferred to 10%
241 KOH solution and were heated in a water bath at 95℃ for 2 minutes. Subsequently, they were
242 washed with Milli Q water and 1X PBS for 30 minutes each thrice. Then, the roots were
243 washed with 150mM K_2HPO_4 for 1 hour, followed by staining with 0.01% Aniline blue in 150mM
244 K_2HPO_4 for 1 hour and again washed with 150mM K_2HPO_4 for 1 hour. The stained roots were
245 visualized under Zeiss ApoTome 2.0 microscope (Carl Zeiss) using a DAPI filter with an
246 excitation wavelength of 370 nm and an emission wavelength of 509 nm (magnification= 10X).
247 The callose quantification was performed as previously described (Mason et al. 2020). The
248 images were scaled and pre-processed using the "Clear Measure Resize" macro to set the
249 background colour to black, deleted pixels outside the selected area, measured the properties
250 of the selected area, and resized the image to specific dimensions using bilinear interpolation.
251 The resulting images were segmented to detect callose deposits using the Trainable Weka
252 Segmentation plugin of ImageJ with default parameters. The segmented image was converted
253 to an 8-bit format with the background set to black, transformed into a mask, and analysed for
254 particles of a size (0.000005–0.005), circularity parameters (0.25–1.0), and counted (Arganda-
255 Carreras et al. 2017; Schindelin et al. 2012). Callose deposits were expressed as the number
256 of deposits per area.

257 **SUSS effector identification using AlphaFold2**

258 The protein structures were predicted using AF2 as a monomer (Jumper et al. 2021; Mirdita
259 et al. 2022). To refine the complex structure, relaxation was performed using the amber force
260 field. The predicted AF models were compared to the experimentally resolved PDB100
261 database using Foldseek search in TM-align mode with taxonomic restriction to Ascomycota
262 (van Kempen et al. 2024). All TM-scores were normalized relative to the length of the
263 experimentally resolved proteins and TM-score >0.5 were considered as structural analogs.

The superposition of predicted AF models onto the identified PDB hits was performed using the Matchmaker module of UCSF chimera (Meng et al. 2006; Pettersen et al. 2004). The 2D topology maps for the modeled SUSS effectors were predicted using the PDBsum webserver (Laskowski et al. 2017).

AlphaFold Multimer (AFM) analysis

A high-throughput *in silico* pulldown screening with AlphaFold v2.3.2 implemented in the LocalColabFold v1.5.5 (https://github.com/YoshitakaMo/localcolabfold), a local installation of ColabFold was performed. All complexes are predicted as heterodimers using LocalColabFold v1.5.5 with the AF multimer model using templates from the MMseqs2 (combined paired and unpaired MSA) search. The top 20 hits with an E-value of < 0.1 were selected as candidate templates for multiple sequence alignment (MSA) and subsequent input parameters for AFM. We ran the AFM model on the GPU with the following parameter settings: template usage enabled, 20 recycles, early recycle stop tolerance set to 0.5 to control the convergence criterion, no dropout, and energy minimized using amber force field. All runs were performed on a Linux system with an Intel CORE i9 11th generation processor, 32 GB RAM and 6 GB NVIDIA RTX 3060 GPU. We evaluated the predicted protein pairs based on AlphaFold2 performance metrics such as predicted template modelling (pTM) score, predicted local distance difference test (pLDDT), and interface pTM score (ipTM). The above performance metrics were parsed using the in-house Python script (https://github.com/chemicalglycobiologylab) from LocalColabfold .json output files. We also calculated spatial information about interface residues as well as confidence and accuracy metrics at the residue and contact levels using the Colabfold Batch AlphaFold-2 multimer structure analysis pipeline to determine whether a pair of residues is a valid contact, particularly at the heterodimer interface and the catalytic triad (Schmid 2023). We found experimentally resolved protein structures from the PDB with structural folds similar to predicted heterodimers using Foldseek multimer (van Kempen et al. 2024). All TM scores were adjusted based on the query length and experimentally solved template structures. The structural similarities of modelled proteins were assessed using the DALI webserver (Holm 2022). The interface of the predicted heterodimer complexes was analyzed and visualized using the open-source version of the PyMOL Molecular Viewer (Yuan, Chan, and Hu 2017), and contact maps were created using the PDB module of Biopython with a contact cut-off of 8 Å (Cock et al. 2009).

Molecular dynamics simulation studies

The protein-protein complexes are simulated for 100ns using the OPLS-AA force field (Lindahl et al. 2010) implemented in GROMACS 2022.3 (Lemkul 2019). All the structures were equilibrated using Constant temperature, constant volume (NVT) and Constant temperature, constant pressure (NPT) ensembles for a constant temperature of 303 K and 1 atm pressure. Necessary amounts of sodium (Na) and chloride (Cl) were added to the system for neutralization. Periodic boundary conditions were enforced in all directions. All the bond lengths within the system were constrained using the LINear Constraint Solver (LINCS) algorithm. Long-range electrostatics in the system was computed by using the Particle Mesh Ewald method (PME). The V-rescale weak coupling method was used to implement temperature regulation at 310 K. The NVT and NPT ensembles were both equilibrated using the Parrinello-Rahman coupling method. Once the system was equilibrated with desired temperature and pressure, the last stage production run was performed for 100ns. The frames

were saved at 10-ps intervals throughout the MD run, with a time step of 2 fs. The built-in modules of GROningen MAchine for Chemical Simulations (GROMACS), such as gmx_rms, gmx_rmsf, gmx_dist and gmx_hbond, were used to evaluate root mean square deviation (RMSD), root mean square fluctuation (RMSF), pairwise distance between chains, and the number of hydrogen bonds (Hbond) along the entire MD trajectory. The structures were visualized using the PyMOL and Chimera software (Meng et al. 2006; Yuan et al. 2017). The Free energy landscape (FEL) analysis was performed using gmx_anaeig module and free energy landscape v1.0.3 module available at https://github.com/sulfierry/free_energy_landscape. The binding free energy was determined by utilizing the 100 ns trajectory obtained from molecular dynamics simulations, and calculations were performed on every 100 frames extracted from the last 30 ns of the trajectory using the gmx_mmpbsa and hawkdock (Valdés-Tresanco et al. 2021; Weng et al. 2019).

Competitive activity-based protein profiling (ABPP) of apoplastic fluids and fungal culture filtrate with soybean Kunitz

Competitive ABPP was performed as described by (Chandrasekar et al. 2017; Hong and Hoorn 2014; Kaschani et al. 2012). In brief, apoplastic fluids collected from control and infected soybean roots at 3 days post infection (dpi) were preincubated with commercially available soybean Kunitz (Sigma Cat no. 93620-250MG) at concentrations 1.25, 2.5, 5, and 7.5 µg/ mL for 30 minutes. Milli Q was used as control. The apoplastic fluids were treated with the FP-alkyne probes or DMSO (no-probe control) along with 500mM Sodium acetate of pH 6 to maintain the pH of the reaction. The mixtures were kept on a thermoshaker (37℃, 200 rpm, 2 hours). The labeling reactions were quenched by ice-cold acetone precipitation. The protein pellets were resuspended in 1X PBS-SDS buffer, heated at 90°C for 10 min, and used for click chemistry. For click-chemistry, 1mM N_3Rh (Final conc. 50µM), 3.4mM TBTA (Final conc. 100µM), 100mM TCEP (Final conc. 2mM), and 50mM $CuSO_4$ (Final conc. 1mM) was added to the sample and was kept on a thermoshaker under dark conditions (30℃, 800 rpm, 1 hour). The reactions were quenched by acetone precipitation and the protein pellets were dissolved in 1X PBS- SDS buffer and 4X gel loading buffer (Final conc. 1X). The labelled samples were separated on 10% SDS-PAGE gels at 200V for 45 minutes and fluorescently labelled proteins were detected by in-gel fluorescence using a ChemiDoc MP scanner 532/28 excitation/emission filter of Pro-Q Emerald 488 in the Blue Epi filter. Densitometric analysis of the fluorescent bands was performed using Image Lab 6.1.0 software. Competitive ABPP was performed on the culture filtrate that was collected from the 3-day-old liquid cultures inoculated with *M. phaseolina*. The culture filtrate was filtered using Mira cloth and was preincubated with various concentrations of soybean Kunitz. The preincubated culture filtrate was subjected to the above-mentioned ABPP protocol.

Results

Macrophomina phaseolina secretes a large repertoire of CAZymes in the soybean root apoplast

Here, first, we isolated the native *Macrophomina phaseolina* strain from the soybean fields of Jabalpur, India, a hot spot region of charcoal rot (CR) disease. The pure cultures of *M. phaseolina* were isolated from the infected soybean roots using the hyphal tip method (Figure S1A). To confirm the identity of the isolated strain, the Internal Transcribed Spacer (ITS) was amplified and sequenced (Figure S1B). Basic Local Alignment Search Tool (BLAST) analysis against the National Center for Biotechnology Information (NCBI) database has

354 revealed a 100% match with the *M. phaseolina* strains that have been previously isolated
355 (Figure S1C). To test the pathogenicity of the isolated *M. phaseolina* strain (*M.*
356 *phaseolina_Jbstrain1*), the roots of two-week-old seedlings of elite Indian soybean varieties
357 such as JS 20-94, SL 955, JS 20-34 and JS 20-98 were infected using the root dip inoculated
358 method. The characteristic above-ground CR disease symptoms such as leaf chlorosis,
359 marginal necrosis, and wilting were prominently observed during the 3 days post-infection
360 (dpi) (Figure S2). Of the tested soybean varieties, JS 20-94 displayed severe above-ground
361 symptoms (Figure S2). These results indicate that the isolated native *M. phaseolina* strain is
362 pathogenic on various soybean varieties, and the soybean variety JS-2094 is highly
363 susceptible to *M. phaseolina*. Therefore, we used the JS-2094 soybean variety and the
364 isolated virulent *M. phaseolina* strain for our further studies.

365 To monitor the colonization process of isolated *M. phaseolina* strain, the roots of two-
366 week-old seedlings of the JS 20-94 soybean variety were infected with the mycelia of *M.*
367 *phaseolina* using the root-dip inoculation method. The root samples were collected at three
368 different time points: 1 dpi, 2 dpi, and 3 dpi. The differential Interface Contrast (DIC)
369 microscopic analysis of the collected samples has revealed the formation of microsclerotia
370 structures by the pathogen during the process of colonization (Figure 1A). At 3 dpi, the number
371 of microsclerotia was higher compared to earlier time points (Figure 1A). To investigate the
372 proteins that are secreted by *M. phaseolina* and monitor the dynamics of secreted soybean
373 proteins in the apoplast, the roots of two-week-old seedlings of the JS 20-94 soybean variety
374 were infected with the mycelia of *M. phaseolina*. The apoplastic fluids were collected from the
375 infected and control roots at 3 dpi and subjected to shotgun proteome analysis (Figure 1B). In
376 total, we have detected 1543 proteins in our apoplastic proteome data set (Table S1). The
377 Principal Component Analysis (PCA) of these detected proteins has indicated a clear
378 distinction between the control and infected proteome samples and a positive correlation
379 within the respective biological replicates (Figure S3). Analysis of these proteins using the
380 SignalP server has revealed that 483 proteins (31%) carried the secretory signal peptide (SP)
381 at their N-terminus (Figure 1B and Table S2). The remaining proteins detected in the root
382 apoplast might be derived from noncanonically secreted proteins or cytoplasmic
383 contaminations during extraction. Only those proteins that carried the SP were considered for
384 further analysis. Of these 483 proteins, 228 secreted proteins were from *M. phaseolina* and
385 255 secreted proteins from soybean (Figure 1B). First, we analyzed the proteins secreted by
386 *M. phaseolina* in the soybean root apoplast (Table S3). Gene ontology analysis has indicated
387 that the *M. phaseolina* secreted proteins are associated with biological processes such as
388 including carbohydrate metabolism, proteolysis, and cell adhesion (Figure 1C). Furthermore,
389 this analysis has revealed that these proteins are associated with several molecular functions
390 targeting various polysaccharides and protein substrates (Figure 1C). Pfam analysis has
391 revealed that the majority of *M. phaseolina* secreted proteins are CAZymes (126 proteins,
392 55%), followed by 25 proteases, 61 other diverse proteins, and, notably, 16 proteins whose
393 domain information is not available (Figure 1D). The detected 126 CAZymes belong to various
394 classes, such as glycosyl hydrolases (GHs, 77 proteins), auxiliary activity family enzymes
395 (AAs, 22 proteins), carbohydrate esterases (CEs, 15 proteins), polysaccharide lyases (PLs, 8
396 proteins), and other unclassified CAZyme classes (4 proteins). The 77 detected GHs
397 belonged to 36 different GH families, of which members of the GH43 family are detected in
398 large numbers (14 proteins) (Figure 1D). A majority of members of the GH43 family have been
399 shown to display α-L-arabinofuranosidase and/or β-D-xylosidases activities, suggesting the
400 polysaccharides carrying linkages might be targeted by *M. phaseolina* during infections. The

401 auxiliary activity family enzymes (AAs) represent the CAZymes that are detected in large
402 numbers following GHs. The 22 AAs belong to seven different families (AA1, AA2, AA3, AA5,
403 AA7, AA8, and AA9) (Table S3). The members of AA9 are Lytic Polysaccharide
404 Monooxygenases (LPMOs) that predominantly target recalcitrant crystalline polysaccharides
405 of the plant cell wall. The phylogenetic tree analysis of the detected AA9 members with the
406 characterized LPMOs from various fungi suggests that these enzymes might target various
407 plant cell wall polysaccharides such as cellulose, xylan, xyloglucan, gluco-mannans and mixed
408 glucans (Figure S4). The 15 CEs belong to eight different CE families (CE1, CE2, CE4, CE5,
409 CE8, CE12, CE15, and CE16), the eight PLs belong to four different PL families (PL1, PL3,
410 PL4, and PL9). Taken together, these results indicate that *M. phaseolina* secretes a diverse
411 repertoire of various CAZymes belonging to various families in the soybean apoplast, and
412 carbohydrate metabolic processes in the soybean root apoplast might play an important role
413 during infections.

414 **Cell wall composition and callose depositions are altered in soybean roots during *M.***
415 ***phaseolina* colonization**

416 Since a large number of CAZymes were detected in our dataset, we investigated the
417 status of soybean root cell walls during *M. phaseolina* infection. The roots of two-week-old
418 soybean plants were infected with the mycelia of *M. phaseolina*, and the cell walls were
419 extracted from the control and infected samples as alcohol-insoluble-residue (AIR). The cell
420 wall samples were further subjected to various biochemical analyses to monitor the changes
421 in the cell walls during infections. First, we characterized the cell walls isolated from the control
422 and infected soybean roots using the Fourier Transform Infrared Spectroscopy (FTIR). This
423 analysis revealed no significant differences in the spectrum profile of control and infected cell
424 wall samples, suggesting no extensive cell wall degradation at the investigated time point
425 (Figure S5). Next, we performed the monosaccharide composition analysis. Here, the neutral
426 sugars from the cell walls were derivatized as alditol acetates (AA) and these residues were
427 analyzed using gas chromatography-mass spectrometry (GC-MS). This analysis has revealed
428 the presence of different neutral sugars, such as arabinose, xylose, mannose, galactose,
429 rhamnose, and glucose, in the cell walls of the control root sample. Of these neutral sugars,
430 the relative abundance of arabinose and xylose was significantly higher than other detected
431 sugars in the control root sample (Figure 2A), indicating that the polysaccharides derived from
432 xylose and arabinose are abundant in the cell walls of soybean roots. Interestingly, the relative
433 abundance of arabinose and xylose was significantly reduced in the infected cell walls
434 compared to the control samples, suggesting the degradation of polysaccharides derived from
435 these neutral sugars (Figure 2B). To gain further insights into structural changes of soybean
436 root cell walls during *M. phaseolina* infections, we performed glycosyl linkages analysis. The
437 neutral sugars from cell walls were derivatized as partially methylated alditol-acetates (PMAA)
438 and were analyzed using GC-MS. Interestingly, this analysis revealed significant differences
439 in the key glycosidic sugar residues in the control and infected samples. The relative
440 abundance of glycosidic sugar residues such as 2-Rha*p*, 5-Ara*f*, 2-Ara*f*, 4-Gal*p*, and 3,6-Gal*p*
441 was significantly reduced in the cell walls of infected root samples compared to control
442 samples (Figure 2C and Table S4). Meanwhile, the relative abundance of 3-Glc*p* was
443 significantly higher in the cell walls of infected samples than in the control. The relative
444 abundance of other detected glycosidic neutral sugar residues was not significantly different
445 between the control and infection samples (Table S4). The glycosidic sugar residues that are
446 detected in lower relative abundance in the infected samples are characteristic linkages
447 derived from cell wall polysaccharides of pectins and hemicellulose. 2-Rha*p* is the main

448 glycosidic sugar present in the main chain of rhamnogalacturonan-1 (RG1) pectins. 5-Ara*f*, 2-
449 Ara*f* and 4-Gal*p* are associated with the arabinogalactans (AGs) linked with RG1 pectins. 3,6-
450 Gal*p* is the characteristic glycosidic linkage present in type II arabinogalactans. Analysis of
451 crystalline cellulose content in the cell wall samples has indicated that the levels of these
452 polymers are reduced in the infected samples compared to the control (Figure S6). These
453 results suggest that the cell wall polysaccharides RG-1, AGs, and cellulose are potentially
454 targeted by CAZymes of *M. phaseolina* during infection.

455 3-Glc*p* is the characteristic glycosidic sugar residue of β-1,3 glucan polysaccharides.
456 Hence, the higher relative abundance of 3-Glc*p* detected in the infected root samples might
457 be derived from the cell walls of *M. phaseolina* or soybean root callose depositions. Callose
458 deposition is an important immune response displayed by plants to prevent the colonization
459 of fungal pathogens. To investigate if callose depositions are increased in the soybean roots
460 during *M. phaseolina* infections, the control and infected root samples were stained with aniline
461 blue, and the callose was visualized using epifluorescent microscopy. Surprisingly, the number
462 of callose depositions was significantly lower in the infected roots compared to the control at
463 3 dpi (Figure S7). The detection of a high number of callose depositions in the control roots
464 might be due to the conditions we used for the infection assays. Next, we monitored the
465 dynamics of callose deposition by staining the root samples collected at early time points of
466 infection. Remarkably, the number of callose depositions was gradually reduced during *M.*
467 *phaseolina* infection, whereas the callose depositions in the control roots remained unchanged
468 at these time points (Figure 2D). These data indicate that callose depositions are degraded
469 during *M. phaseolina* infection. Furthermore, the increase in the relative abundance of 3-Glc*p*
470 in the infected samples is not due to the callose deposition but rather derived from the cell
471 walls of *M. phaseolina*.

472 **Sequence-unrelated structurally similar (SUSS) effectors and effectors with novel**
473 **structural folds are secreted by *M. phaseolina* in soybean root apoplast.**

474 Plant pathogens secrete protein effectors in apoplast to manipulate the host to favor
475 their colonization process. To investigate the repertoire of candidate effectors secreted by *M.*
476 *phaseolina*, we analyzed 228 *M. phaseolina* proteins detected in the soybean root apoplast
477 using the EffectorP 3.0 machine learning tool (Sperschneider and Dodds 2022). This analysis
478 has revealed that nearly 23% of proteins (49 proteins) are predicted as candidate apoplastic
479 effector proteins, and one protein was predicted as a cytoplasmic effector (Figure 3A). Pfam
480 analysis has revealed that these 49 candidate apoplastic effectors were diverse and belong
481 to various enzyme classes such as glycosyl hydrolases (eight proteins), polysaccharide lyases
482 (seven proteins), serine hydrolases (three proteins), proteases (three proteins) and other
483 proteins (15 proteins). Interestingly, nine candidate apoplastic effectors could not be classified
484 due to a lack of significant sequence-level similarities, and the conserved domain information
485 for these proteins was not available (Figure 3A). To investigate further if these nine effectors
486 have structural homology with other fungal effector proteins, we performed the AlphaFold (AF)
487 analysis. The candidate proteins were modeled as monomers using the AF2 program, and
488 template-free AF2 modeling was implemented in the local colabfold pipeline with 5X models
489 for each unique pair (Figure 3B). The resulting AF models were relaxed using an amber force
490 field for energy minimization, and the relaxed structural models were further analyzed (Figure
491 3B). The top-ranked structural model for each candidate effector was compared with the
492 experimentally resolved protein crystal structures in the PDB100 database using the Foldseek
493 search program. Since *M. phaseolina* belongs to the division Ascomycota of the fungal

494 kingdom, we restricted our Foldseek search to this division. Interestingly, this analysis has
495 revealed that seven proteins shared structural homology with the proteins from the members
496 of the Ascomycota division with a significant TM score > 0.5 (Table S5) and two proteins
497 (UniProt IDs: K2R751 and K2RET3) that did not share similarities at structural levels to
498 proteins from Ascomycota division (Figure S8 A,B). Interestingly, when searching the
499 structural model of these proteins against the AlphaFold protein structure database (AlphaFold
500 DB) revealed that these proteins have structural similarities with proteins from the Ascomycota
501 fungal division, suggesting that these are potentially uncharacterized family of SUSS effectors
502 and their structural folds are evolutionarily expanded across the members of this division
503 (Table S6). Notably, these proteins did not share any structural similarities with any of the
504 structurally resolved proteins of other fungal divisions, indicating that the structural folds in
505 these proteins are possibly novel. Of the seven proteins with predicted structural homologs,
506 five proteins (Uniprot IDs: K2RZ98, K2RDS7 K2QS10, K2RZH1, and K2R1W5) shared
507 similarities with effectors that are secreted by plant-associated fungi and two proteins (Uniprot
508 IDs: K2RGG5 and K2RF41) showed significant structural similarities with proteins from
509 *Neurospora crassa* and *Saccharomyces cerevisiae*, respectively (Figure S9). The protein
510 K2RZ98 shared structural similarity to the Tsp1 effector of *Trichoderma virens* Gv29-8, a
511 beneficial plant fungus with a significant TM-score of 0.801 and an RMSD of 1.23 Å (Figure
512 3C). K2RZ98 displayed a distinct β-barrel fold unique to the Tsp1 effector (Gupta et al. 2021)
513 (Figure 3C). The protein K2RDS7 shared structural similarity to Zt-KP6-1 effector from
514 *Zymoseptoria tritici*, a plant fungal pathogen with a TM-score and RMSD score of 0.503 and
515 2.89 Å, respectively. The K2RDS7 has a characteristic KP6-like fold consisting of two
516 signature α-helices, a variable number of β-sheets (3-5), and unique 2-4 disulfide bonds
517 (Rocafort et al. 2022). Remarkably, three candidate effector proteins, K2QS10, K2RZH1, and
518 K2R1W5, shared structural similarity to the MoErs1 protein with a TM score > 0.5 and a RMSD
519 score of 3-4 Å. MoErs1 is an effector secreted by *Magnaporthe oryzae*, a fungal pathogen
520 that causes the rice blast disease. These three proteins were predicted to significantly share
521 the characteristic of β-trefoil fold displayed by the MoErs1 effector (Y. Liu et al. 2024). Taken
522 together, our AF2 analysis has revealed that SUSS effectors and effectors with novel structural
523 folds are secreted by *M. phaseolina* in soybean root apoplast during infections.

524 **AFM analysis combined with molecular dynamics (MD) reveals the interaction of**
525 **MoErs1-like effector (K2QS10) with secreted cysteine protease (I1JTM0).**

526 MoErs1 is reported to be a species-specific effector secreted by *M. oryzae* to inhibit
527 the activities of OsRD21, a rice cysteine protease (Y. Liu et al. 2024). However, our AF2
528 analysis revealed that the three effectors from *M. phaseolina* shared structural homology with
529 MoErs1, indicating that the MoErs1-like effectors are present in *M. phaseolina*. This prompted
530 us to investigate these proteins further. To investigate if the MoErs1-like secreted effectors
531 from *M. phaseolina* can target soybean cysteine proteases, we performed *in silico* pulldown
532 with the AlphaFold Multimer (AFM) platform (Figure 4A). We used three MoErs1-like secreted
533 effectors (Uniprot IDs: K2RZH1, K2R1W5, and K2QS10) as ligands against five soybean
534 cysteine proteases (Uniprot IDs: I1MXL3, I1MHG0, I1MER7, I1JTM0 and I1M2Y6) as baits.
535 These cysteine proteases were detected in the soybean root apoplast both in control and
536 infected samples (Table S8). In parallel, we performed the AFM analysis of MoErs1 (ligand)
537 against the OsRD21 cysteine protease (bait). The pro-inhibitory domain present in the N-
538 terminus of cysteine proteases was removed before analysis as recommended for AFM
539 analysis of cysteine proteases (Homma et al. 2024). The pipeline for our analysis involves
540 modelling ligand and bait proteins as heterodimers using the template from MMseqs2 search

541 (Mirdita et al. 2019). The AFM-modelled structures were relaxed using an amber force field,
542 and the top five models were considered (Mirdita et al. 2022). The resultant structural models
543 predicted by AFM were screened using ipTM+pTM score, pLDDTscore, and model confidence
544 to select the top interacting partners (Figure 4A). Interestingly, our AFM analysis has revealed
545 several interaction models for MoErs1-like effectors with soybean cysteine proteases
546 with significant ipTM+pTM score > 1.1 and pLDDTscore >0.85 (Figure 4B). Next, we employed
547 model confidence score selection with a > 75% cut-off to select the top interaction models
548 (Figure 4C). Remarkably, of the three MoErs1-like effectors, one MoErs1-like effector
549 (K2QS10) displayed significant interaction with the secreted soybean cysteine protease
550 (I1JTM0) (Figure 4B, C). The interaction model predicted by AFM has a significant ipTM+pTM
551 score of 1.60, a high pLDDT score of 84.36, and a model confidence score of 81% (Figure
552 4D). These cut-off values are comparable with the AFM models for MoErs1 and OsRD21 with
553 an ipTM+pTM score of 1.65, supported by a high pLDDT score of 89.8 and a confidence score
554 of 85.8% (Figure 4D). Notably, the loop regions of the MoErs1-like effector (K2QS10)
555 displayed significant interaction with the amino acids present in the active site of soybean
556 cysteine protease (I1JTM0) and these interactions are comparable to AFM models of MoErs1-
557 OsRD21 complex (Figure 4E). In general, the N-terminus of the plant cysteine proteases
558 carries a self-inhibitory regulatory pro-domain that blocks the active site of these enzymes
559 (Hou et al. 2018). Remarkably, performing the AFM analysis with the pro-inhibitory domain
560 included in the cysteine protease resulted in non-significant models with low ipTM+pTM and
561 pLDDTscore and low model confidence scores (Figure S10). Therefore, this result suggests
562 that the MoErs1-like effector, K2QS10, secreted by *M. phaseolina*, might serve as an active
563 site-directed cysteine protease inhibitor.

564 To further validate the predicted AFM interactions, we performed MD simulations of the
565 MoErs1-like effector and soybean cysteine protease complex. In parallel, we performed these
566 analyses with MoErs1 and OsRD21 cysteine protease complex. All-atom MD simulations were
567 calculated using the OPLS-AA force field with a 100 nanoseconds (ns) time scale. Root mean
568 square deviation (RMSD) analysis was performed to evaluate the reliability and stability of the
569 protein complexes. Additionally, the trajectories obtained from the MD simulations were
570 subjected to MM-PB/GBSA calculations to evaluate the binding affinities of the complexes.
571 The RMSD analysis of the MD trajectories indicated that the MoErs1-like effector (UniprotID.
572 K2QS10) and the cysteine protease (UniprotID. I1JTM0) complex exhibited a constant RMSD
573 after 40 ns, indicating that the protein-inhibitor complex reached equilibrium during the MD run
574 (Figure S11A). The distance between these protein-inhibitor complexes remained constant
575 after 80th ns, indicating that there is no conformational shift at the contact interface (Figure
576 S11B). This is supported by the number of hydrogen bonds calculated at the interface and 2D
577 contact maps showing contact frequency during 80-100th ns (Figures S11C,D). The MM-
578 PB/GBSA calculations for binding affinity have indicated that the protein complex is
579 energetically favourable with negative ΔG values (Table S7). Next, we performed a per-residue
580 decomposition analysis using the HawkDock tool to identify the significant interacting amino
581 acid residue at the contact interface. For this analysis, first, we performed FEL analysis on
582 the MD trajectories to obtain low-energy conformation of the MoErs1-like effector and the
583 cysteine protease (UniprotID. I1JTM0) complex (Figures S11E). Next, we integrated the top
584 interaction model predicted by AFM, the low energy conformer obtained through FEL analysis,
585 and the conformer from 100th ns in the MD trajectories for per-residue decomposition analysis.
586 This analysis has revealed significant interactions of amino acid residues present in the loop
587 regions of the MoErs1-like effector (UniprotID. K2QS10) with the soybean cysteine protease

588 (UniprotID. I1JTM0) (Figure 4E). The amino acid residues present in the loop region of the
589 MoERS1-like effector displayed similar interaction energy values with soybean cysteine
590 protease in all three structural models used for the analysis. This result suggests that the loop
591 regions in the MoErs1-like effector might have a crucial role in inhibiting cysteine protease.
592 This is consistent with the crystal structures of cysteine protease inhibitors such as chagasin
593 and cystatin, where the loop region interacts with the active site to inhibit the activities of
594 cysteine proteases (Bode et al. 1988; Redzynia et al. 2009). Notably, the MD simulations and
595 binding calculation analyses performed for the MoErs1-like effector and soybean cysteine
596 protease complex were comparable with the MoErs-1-OsRD21 complex (Figure 4E and
597 Figure S11A-E). Taken together, the AFM analysis together with the MD simulation indicates
598 that the secreted MoErs1-like effector, K2QS10, might have a role as an apoplastic cysteine
599 protease inhibitor.

600 **Quantitative proteome analysis reveals the dynamics of soybean apoplastic proteins**
601 **during *M. phaseolina* infection.**

602 In our apoplastic proteomic dataset, we detected 255 secreted proteins from soybean
603 that carried a signal peptide. Pfam analysis of 255 secreted soybean proteins has revealed
604 that the majority of these proteins were CAZymes (98 proteins, 38%) followed by proteases
605 (27 proteins, 12%) and protease inhibitors (16 proteins, 7%) (Figure 5A and Table S8). The
606 other 114 proteins were diverse and included CAP proteins (Cysteine-rich secretory protein
607 [CRISP], Antigen 5 [Ag5], pathogenesis-related 1 protein, three proteins), fasciclin-like
608 arabinogalactans (FLAs) (nine proteins), dirigent proteins (seven proteins), thioredoxin (four
609 proteins) and other proteins (Table S8). Of the 255 secreted proteins, 241 proteins were
610 detected in both control and infected samples (Table S9). Eight proteins were detected as
611 uniquely in control samples, and six proteins were detected uniquely in infected samples
612 (Figure S12). To monitor the dynamics of soybean apoplastic proteins during *M. phaseolina*
613 infection, we performed a volcano plot and z-score analysis. In the volcano plot analysis, -
614 $\log_{10}$ p-values of the 241 proteins detected in both control and infected samples were plotted
615 against their respective $\log_2$ fold change of differential abundance values (Figure 5B). Z-scores
616 of the 241 proteins were calculated by normalizing the LFQ values with their standard
617 deviation, and the z-score values were represented as a heat map to visualize the differentially
618 regulated proteins (Figure S13). These analyses have revealed that 95 soybean apoplastic
619 proteins were differentially regulated during *M. phaseolina* infection. 76 proteins were
620 upregulated, 19 proteins were downregulated, and 146 proteins remained unregulated (Figure
621 5B). Gene ontology analysis of the 95 differentially regulated proteins has revealed that the
622 proteins that are associated with several molecular functions and biological processes are
623 detected in our dataset. The proteins associated with lactoperoxidase activity are highly
624 enriched in the upregulated proteins compared to the downregulated proteins (Figure 5C).
625 Furthermore, the proteins associated with several molecular functions such as peptidase
626 inhibitor, heme binding, chitinase activity, alpha-mannosidase activity, chitin-binding, serine
627 peptidase activity, and Flavin Adenine Dinucleotide (FAD) binding are uniquely detected in the
628 upregulated proteins (Figure 5C). Interestingly, the proteins that are associated with the fatty
629 acid binding are uniquely detected only in the downregulated proteins. In addition, the proteins
630 associated with biological processes such as hydrogen peroxide catabolic process and
631 cellular oxidant detoxification are highly enriched in the upregulated proteins compared to the
632 downregulated proteins (Figure 5C). The proteins associated with several biological
633 processes, such as the mannose metabolic process, amino sugar catabolic process,
634 phenylpropanoid biosynthetic process, and cell wall catabolic process, are uniquely detected

635 in the upregulated proteins (Figure 5C). Pfam analysis of 76 upregulated secreted proteins of
636 soybean has revealed these proteins are classified into CAZymes (27, 36%), proteases
637 (seven, 9%), protease inhibitors (nine, 11%), and other diverse proteins (33, 43%) (Figure 5D).
638 The detected CAZymes belonged to various enzyme classes, such as glycosyl hydrolases
639 (GH), carbohydrate esterases (CE), auxiliary activity (AA) and carbohydrate-binding module
640 (CBM). Among the seven upregulated proteases, six proteins were serine proteases from
641 families S8, S10, and S28, and one cysteine protease from family C26. Interestingly, three
642 only leucine-rich repeat-containing proteins (LRR) and two expansin proteins were
643 upregulated during *M. phaseolina* infection. These proteins are previously reported to act as
644 susceptibility factors in other plant species during fungal infections (Abuqamar et al. 2013;
645 Chen et al. 2021). Notably, all the upregulated protease inhibitors (9 proteins) belonged to
646 Kunitz-type serine protease inhibitors. Pfam analysis of 19 downregulated proteins of soybean
647 has revealed these proteins are classified into nine CAZymes from different families, four
648 proteases, and six diverse proteins that include trichome birefringence-like (TBL) protein (one
649 protein), fasciclin-like arabinogalactan-protein (FLA) (two proteins), lipid transfer proteins (two
650 proteins), and lipid recognition protein (one protein) (Figure 5D). The categorization of the
651 differentially regulated proteins based on the pathogenesis-related proteins (PR) has revealed
652 PR proteins belonging to families PR-2, PR-3, PR-5, and PR-9 are upregulated, and the PR
653 proteins belonging to families PR2, PR9, and PR14 are downregulated during *M. phaseolina*
654 infections (Figure 5D). Taken together, this quantitative proteome analysis of the soybean-
655 secreted proteins revealed differential regulation of several apoplastic proteins during *M.*
656 *phaseolina* infection.

657 **AFM analysis combined with activity-based protein profiling (ABPP) reveals**
658 **interactions of soybean Kunitz with serine proteases.**

659 Plant protease inhibitors are involved in the regulation of immune protease activities.
660 In our proteome dataset, we have detected 16 protease inhibitors in the soybean root apoplast.
661 Of these 16 protease inhibitors, 15 belong to Kunitz-type protease inhibitors, and one belong
662 to cysteine protease inhibitor. Interestingly, among these detected protease inhibitors, only the
663 Kunitz-type protease inhibitors were upregulated in abundance during *M. phaseolina*
664 infections (Figure 6A). Nine Kunitz-type protease inhibitors (Uniprot IDs: I1LVB0, C6T2D3,
665 I1MI59, C6T586, I1KYW5, I1KYX0, C6SY46, I1KW54, and I1LVB1) were significantly
666 abundant in the infected samples compared to the control and one Kunitz protein (Uniprot ID:
667 I1KYX1) was uniquely detected only in the infected samples (Figure 6A and Figure S12). To
668 investigate the structural relationship between the upregulated Kunitz proteins, we modeled
669 the ten Kunitz proteins with AF2 as monomers and assessed their structural similarity using
670 the DALI web server (Holm 2022). This analysis has revealed that the detected soybean
671 Kunitz can be classified into two groups: Group 1 consists of six Kunitz proteins (Uniprot IDs:
672 I1MI59, I1KW54, I1LVB1, I1LVB0, C6SY46, C6T2D3) and group 2 consists of 4 Kunitz proteins
673 (Uniprot IDs: I1KYW5, C6T586, I1KYX0, I1KYX1) (Figure S14). This result indicates that
674 Kunitz proteins detected in soybean root apoplast are structurally diverse. Plant Kunitz-type
675 protease inhibitors have been predominately investigated for their role in regulating the
676 activities of serine proteases derived from insects and humans (Mehmood et al. 2024;
677 Okedigba et al. 2023). To investigate the potential targets of the Kunitz proteins that are
678 upregulated in our dataset, we employed AFM analysis to predict interactions of these Kunitz
679 proteins with randomly selected secreted serine proteases detected in soybean root apoplast.
680 We followed the pipeline and the stringent cut-off scores that we used for the AFM analysis of
681 soybean cysteine proteases and MoErs1 (Figure 6B). In the AFM analysis, the ten soybean

682 Kunitz proteins (Nine upregulated and one uniquely expressed) were used as ligands against
683 ten soybean serine proteases and seven *M. phaseolina* serine proteases that were randomly
684 selected from our dataset (Tables S10 and S11). The soybean serine proteases belonged to
685 families S8, S10, and S28, whereas the *M. phaseolina* serine proteases belonged to families
686 S8, S9, and S10 (Tables S10 and S11). First, we validated the AFM pipeline of the serine
687 protease and serine protease inhibitor interactions with P69B, a serine protease from tomato,
688 with its experimentally validated serine protease inhibitor Epi1A from *P. infestans* (Homma et
689 al. 2023; Mooney and van der Hoorn 2024). This analysis has revealed that the AFM analysis
690 predicts the interaction of P69B and Epi1A with significant ipTM+pTM and pLDDT scores of
691 1.5 and >0.85, respectively (Figure S15). Furthermore, replacing Epi1A with Epic2B, a
692 cysteine protease inhibitor from *P. infestans* in the analysis revealed no significant interactions
693 (lower ipTM+pTM values in the range of 0.7–0.9) between Epic2B and P69B (Figure S15).
694 These results indicate that our AFM pipeline can be used to reliably predict the interaction
695 between Kunitz and serine proteases. Interestingly, our AFM analysis of the upregulated
696 Kunitz proteins with the secreted serine proteases has revealed the interaction of Kunitz
697 proteins with both soybean serine proteases and *M. phaseolina* serine proteases with
698 significant ipTM+pTM, pLDDT, and high model confidence scores (Figure 6 C,D). Notably, the
699 AFM analysis of soybean Kunitz proteins with randomly selected secreted chitinases detected
700 in our dataset revealed no significant interaction models for the complexes (Figure S16). This
701 result indicates the reliability of the prediction of soybean Kunitz with the serine proteases.
702 The Kunitz proteins from group 1 and group 2, such as I1KW54, I1MI59, I1LVB1, L1LVB0,
703 C6T2D3, and C6T586, are predicted to interact with serine proteases of both soybean and *M.*
704 *phaseolina* (Tables S10 and S11). The group 2 Kunitz protein, I1KYW5, is predicted to interact
705 only with a serine protease of *M. phaseolina* but not soybean serine proteases. Meanwhile,
706 the other Kunitz protein from group 2, I1KYX1, interacts only with a serine protease of soybean
707 but not *M. phaseolina* serine proteases. The group 1 Kunitz protein, C6SY46, and group 2
708 Kunitz protein, I1KYX0, did not interact with any of the analyzed serine proteases from
709 soybean or *M. phaseolina* (Tables S10 and S11). Taken together, our AFM analysis indicates
710 distinct interactions of soybean Kunitz proteins with serine proteases from soybean and *M.*
711 *phaseolina.*

712 Next, we analyzed the AFM-predicted complexes where the active sites and the
713 catalytic residues for serine proteases are well-annotated. Of the several serine proteases that
714 were used in the AFM analysis, the active site catalytic residues for a soybean serine protease
715 (I1JT08) and two fungal serine proteases (K2RSK5 and K2RLG4) are well-annotated. Analysis
716 of these complexes has revealed that these soybean Kunitz proteins interact with the catalytic
717 active site residues of these serine proteases. The Kunitz proteins are reported to inhibit serine
718 proteases through the stearic blockade by binding to the targets with characteristic loop
719 regions (Guerra et al. 2023). A similar loop region could be visualized for the soybean Kunitz
720 proteins in close proximity near the active site of these serine proteases, suggesting the Kunitz
721 proteins might act as an active site-directed inhibitor of serine protease from soybean and *M.*
722 *phaseolina* (Figure 6E, F). These predicted AFM models of the analyzed complexes are similar
723 to the recent crystal structures of soybean Kunitz and insect serine protease complexes
724 (Mehmood et al. 2024). To further validate these AFM predictions, we performed competitive
725 activity-based protein profiling (ABPP) with the apoplastic fluids. The apoplastic fluids were
726 collected from the control and *M. phaseolina*-infected soybean roots at 3 dpi. The apoplastic
727 fluids were preincubated with different concentrations of the commercially available Kunitz
728 inhibitor of soybean origin. The preincubated proteomes were treated with the

729 fluorophosphate-alkyne probe (FP-alkyne). FP-alkyne is an active site-directed probe that
730 targets several serine hydrolases, including serine proteases (Kaschani et al. 2009). The FP-
731 alkyne-labeled proteomes were subsequently subjected to click chemistry to tag the
732 rhodamine fluorescent moiety. The fluorescently labelled proteins were separated by SDS-
733 PAGE and detected by in-gel fluorescent scanning. This analysis revealed several active
734 serine hydrolases at various molecular weights were detected in the infected and control
735 apoplastic fluid samples compared to the no probe control (Figure 7A,B). Interestingly, the
736 intensities of three bands (band 2, band 3, and band 4) at approx. 75 kDa, 70 kDa, and 55
737 kDa, respectively, were reduced in the infected samples preincubated with the soybean Kunitz
738 (Figure 7A). Furthermore, the intensities of three bands (band 2, band 4, and band 5) at
739 approx. 75 kDa, 55 kDa, and 50 kDa, respectively, were reduced in the control samples
740 preincubated with the soybean Kunitz (Figure 7B). These results suggest the soybean Kunitz
741 protein can interact with the active site of serine proteases in the control and infected
742 apoplastic fluids. To investigate if the soybean Kunitz can inhibit the serine proteases from *M.*
743 *phaseolina*, we performed competitive ABPP using the soybean Kunitz and the culture filtrate
744 of *M. phaseolina.* This analysis has revealed that intensities of two bands (band 1 and band
745 2) at approx. 75 kDa were reduced in *M. phaseolina* culture filtrate samples preincubated with
746 soybean Kunitz. This result suggests that soybean Kunitz protein can also interact with the
747 active site of serine proteases from *M. phaseolina*. Interestingly, we also detected increased
748 labeling intensities of bands in the apoplastic fluid samples and *M. phaseolina* culture filtrates
749 preincubated with soybean Kunitz (Figure 7A-C). The increased intensities might be due to
750 the putative roles of Kunitz in activating serine hydrolases, or indirect activation of some serine
751 hydrolases due to the inhibition of serine protease by Kunitz in the labeling reaction. The exact
752 reason for increased labeling upon Kunitz's pre-treatment requires further investigation. Taken
753 together, our AFM analysis combined with ABPP reveals the interaction of soybean Kunitz with
754 serine proteases of soybean and *M. phaseolina*.

755 **Discussion**

756 *Macrophomina phaseolina* is a climate change-relevant fungal pathogen that causes charcoal
757 root rot in various legumes, including soybean. *M. phaseolina* displays a hemibiotrophic
758 lifestyle by employing sequential biotrophic followed by necrotrophic infection strategies to
759 colonize their hosts (Chowdhury et al. 2017). To detect the wide array of effectors secreted by
760 *M. phaseolina* and identify the candidate upregulated plant susceptibility factors, we have
761 chosen the highly susceptible soybean variety, JS 20-94. In our apoplast proteomics study,
762 we detected a large repertoire of CAZymes secreted by *M. phaseolina* in the soybean root
763 apoplast. CAZymes are employed by fungal pathogens to degrade plant cell walls during
764 colonization. The specific members of CAZymes secreted by *M. phaseolina* strongly correlate
765 with the decrease in the characteristic compositions in the plant cell walls during infections.
766 For instance, the decrease in the relative abundance of neutral sugars such as arabinose and
767 xylose and 5-Ara*f* and 2-Ara*f* residues in the cell walls of infected soybean roots might be due
768 to the activities of the members from the GH43 family that have been predominantly shown to
769 exhibit α-arabinofuranosidase and/or β-xylosidase activities (Shrivastava and Goyal 2024).
770 Reduced relative abundance of other glycosyl residues 2-Rha*p*, 4-Gal*p*, and 3,6 Gal*p* in the
771 infected cell walls could be attributed to the activities of the members of the GH78, GH35, and
772 GH53 families that are shown to act on these glycosyl linkages, respectively (Hobson and
773 Deyholos 2013; Nghi et al. 2012; Tanthanuch et al. 2008). Notably, we have detected LPMOs
774 belonging to the AA9 family secreted by *M. phaseolina* in our apoplast proteome dataset.
775 These enzymes were not detected in the recent *M. phaseolina* secretome studies (Arafat et
776 al. 2022; Pineda-Fretez et al. 2023). Fungal LPMOs are emerging as important virulence

factors, and thus, the detection of LPMOs from *M. phaseolina* signifies the importance of our apoplastic proteome analysis. The oxidative activity of these LPMOs by *M. phaseolina* could have reduced crystalline cellulose content in the cell walls of infected soybean root. The glycosidic sugar residues that showed reduced relative abundance in the cell walls of infected samples are characteristic linkages derived from cell wall polysaccharides such as RG-1 and AGs. RG-1 and AGs exist as a complex in the plant cell walls and provide structural support to the plant cell (Tan et al. 2013). Hence, it could be speculated that *M. phaseolina* targets these cell wall components to gain entry into the soybean roots during colonization. Callose, β-1,3-glucan polymer, deposited at pathogen penetration sites helps reinforce weak or compromised cell walls during a pathogen attack (Y. Wang et al. 2021). In our study, we observed a drastic reduction in callose depositions in the infected soybean roots at different stages of *M. phaseolina* infection. The reduction in callose depositions could be due to endo-β-1,3-glucanases activities. The upregulation of plant endo-β-1,3-glucanases has resulted in callose degradation to promote viral infection and bacterial infections (Gaudioso-Pedraza et al. 2018; Shi et al. 2020). In many cases, β-1,3-glucanases are also secreted by pathogenic fungi such as *Colletotrichum graminicola* to degrade callose to promote infection (Gu et al. 2024). In our proteome dataset, we have detected various members of glycosyl hydrolases that are secreted by *M. phaseolina* (GH17, GH55 and GH131) and soybean (GH17 and GH152) that display endo-β-1,3-glucanases activity. It would be interesting to identify the candidate enzyme involved in callose degradation during infection. Taken together, it is evident that specific cell wall components are altered during infections, and hence, engineering soybean with improved resilient characteristics of these cell wall components might be an interesting approach to prevent the invasion of *M. phaseolina*.

In our proteome dataset, we have detected several candidate apoplastic effectors secreted by *M. phaseolina*. Fungal pathogens employ apoplastic effectors to promote host colonization. SUSS effectors are emerging as major players in the apoplast during plant colonization (Seong and Krasileva 2023). In our study, AF2 analysis has revealed interesting SUSS effectors secreted by *M. phaseolina*. First, the Tsp1-like effector (UniProt ID: K2RZ98) with the characteristic β-barrel structural fold. Tsp1, an effector protein, is secreted by *Trichoderma virens,* a beneficial fungus, to module host hormonal pathways, especially inducing salicylic acid in maize plants. Salicylic acid, an antagonist for Jasmonic acid, promoted susceptibility against *Cochliobolus heterostrophus*, a necrotrophic fungal pathogen (Gupta et al. 2021). Second, the effectors with KP6-like fold (UniProt ID: K2RDS7). The effectors with KP6-like fold are secreted by *Z. tritici*, a wheat pathogen, and have been reported to suppress PTI responses induced by flg22, laminarin, and chitin during the early stages of wheat colonization (Thynne et al. 2024). In addition, Zt-KP6-1, an effector with a KP6-like fold, has also been shown to function as an antimicrobial protein against fungal pathogens such as *Botrytis cinerea* and *Z. tritici* (Guillen et al. 2024). Third, MoERS1-like effectors (Uniprot IDs: K2RZH1, K2R1W5, and K2QS10) with the characteristic β-trefoil structural fold. MoERS1 effectors have been reported to be the species-specific effectors secreted by *Magnaporthe oryzae* to inhibit the papain-like cysteine protease *Os*RD21 involved in rice immunity (M. Liu et al. 2024). However, the detection of MoERS1-like effectors in our study indicates that these effectors are not species-specific but seem to be expanded in other fungi. Notably, we have detected two effectors (UniProt IDs: K2R751 and K2RET3) that had no structural similarity with the experimentally resolved structures of fungal proteins in the PDB100 database. However, these effectors have structural similarities to other Ascomycetes fungi based on our search in the AlphaFold protein structure database. Hence, these effectors with novel structural folds might have important functions that require further investigation. Therefore, it is possible that these SUSS effectors secreted by *M. phaseolina* have similar functions as their characterized structural analogs to promote host colonization.

827 Fungi take advantage of certain plant proteins that act as susceptibility factors
828 promoting colonization (Gorshkov and Tsers 2022). In our dataset, we have detected several
829 upregulated proteins whose homologs have been reported to act as susceptibility factors in
830 various plants. First, the upregulation of members of the GH38 family. We have detected three
831 proteins from this family whose abundance is increased during infections. The members of
832 this family have been shown to exhibit α-mannosidase activities and target N-glycans of cell
833 wall glycoproteins for the cell wall loosening process (Ghosh et al. 2011; Méndez-Yáñez et al.
834 2024). Interestingly, increased activities of α-mannosidases have been detected in soybean
835 roots during the colonization of arbuscular mycorrhizal fungi (AMF), and the activities of these
836 enzymes on N-glycans have been speculated for promoting AMF colonization (Dettmann et
837 al. 2005). Second, the upregulation of expansin proteins. In our dataset, we have two expansin
838 proteins upregulated during infections. Expansins are involved in cell enlargement and cell
839 wall loosening at the interaction points between cellulose and hemicellulose (Cosgrove 2005).
840 The absence or downregulation of these proteins has resulted in increased resistance against
841 necrotrophic fungal pathogens such as *Botrytis cinerea* and *Alternaria brassicicola* (Abuqamar
842 et al. 2013; Cantu et al. 2008). Third, the upregulation of members of LRR-only proteins. We
843 have detected three proteins from this family whose abundance is increased during infections.
844 Recently, an apoplast localized LRR-only protein, NTCD4, in Arabidopsis, has been shown to
845 act as a susceptibility factor. NTCD4 interacts with pathogen-derived Necrosis-and ethylene-
846 inducing peptide 1 (Nep1)-like proteins (NLP) to trigger cell death and promote colonization of
847 necrotrophic fungal pathogen *Botrytis cinerea* (Chen et al. 2021). Interestingly, we detected
848 two NLP-like proteins (UniProt IDs: K2RMZ7, K2RNB9) secreted by *M. phaseolina* in soybean
849 root apoplast (Table S3). Hence, it could be speculated that these proteins might exhibit
850 interactions with the upregulated LRR-only proteins to induce cell death and promote *M.*
851 *phaseolina* colonization. Fourth, the upregulation of a GH17 family protein. The members of
852 this family exhibit endo β-1,3-glucanase activities. We have detected one GH17 family protein
853 (Uniprot ID: I1LIL5) whose abundance is increased during infections. Recently, *Hv*BGLUII, a
854 GH17 family protein, has been characterized for its role in releasing susceptible factors from
855 the fungal extracellular polysaccharide (EPS) matrix. The endo β-1,3-glucanase activities of
856 *Hv*BGLUII released a decasaccharide fragment, GD, that scavenging ROS and promotes
857 fungal colonization (Chandrasekar et al. 2022). Taken together, these candidate susceptibility
858 factors can be chosen as potential candidates for engineering CR resistance in soybean using
859 genome editing tools.

860 Apoplastic proteases and their inhibitor interactions are significant determinants of the
861 outcome of plant-fungal interactions. AFM is emerging as an important machining tool for
862 investigating enzyme-inhibitor interactions at the plant-pathogen interface. Our AFM screening
863 of selected secreted proteins has revealed the interaction of cysteine and serine protease with
864 their respective protease inhibitors. The top AFM-predicted models of these protease inhibitor
865 complexes are similar to the reported crystal structures of other protease inhibitor complexes
866 where the loop regions of the protease inhibitor interact strongly with the active site of the
867 proteases (Bode et al. 1988; Mehmood et al. 2024). Papain-like cysteine proteases (PLCPs)
868 and serine proteases represent the central hub for plant immunity, and their activities are
869 modulated by inhibitors during fungal colonization (Coculo et al. 2023; Misas-Villamil et al.
870 2016). In our study, a MoERS1-like effector (UniProt ID: K2QS10) secreted by *M. phaseolina*
871 displayed significant interaction with a ubiquitous PLCP (UniProt ID: I1JTM0) present in the
872 soybean root apoplast. Fungal and oomycetes pathogens secrete proteaceous inhibitors in
873 the apoplast to modulate the activities of immune cysteine proteases during colonization
874 (Misas-Villamil et al. 2016). Therefore, the MoERS1-like effector (UniProt ID: K2QS10)
875 identified in this study might function as a PLCP inhibitor modulating soybean immune
876 response. In addition, our AFM analysis has revealed interactions of soybean Kunitz with the

877 serine protease of soybean and *M. phaseolina*. In our proteome analysis, Kunitz proteins are
878 the only soybean-derived protease inhibitors upregulated during infections. Plant Kunitz
879 proteins are predominantly characterized for their roles in disrupting the digestive functions of
880 insects and animals by inhibiting the activities of serine proteases (Mehmood et al. 2024;
881 Okedigba et al. 2023). However, plant Kunitz proteins also play important roles during plant-
882 pathogen interactions. AtKTI1, a Kunitz protein from Arabidopsis, modulates cell death
883 responses and also serves as a negative regulator for resistance against *Erwinia carotovora*
884 subsp. *carotovora*, necrotrophic bacterial pathogen (Li, Brader, and Palva 2008). Furthermore,
885 Kunitz proteins also module the activities of host serine protease in apoplast during fungal
886 colonization. KPI106, a Kunitz protein from *Medicago truncatula,* is upregulated in root
887 apoplast during colonization of AMF. The KPI106 inhibits an apoplast serine protease, SCP1,
888 which is upregulated during AMF colonization. The *M. truncatula* interaction of tandemly
889 upregulated Kunitz and serine proteases in *M. truncatula* control the colonization of AMF and
890 arbuscule formation (Rech et al. 2013). In our dataset, the abundance of serine proteases and
891 Kunitz are upregulated in soybean root apoplast during infections. Therefore, *M. phaseolina*
892 might follow a strategy similar to AMF by modulating the immune serine protease activities
893 through soybean Kunitz protein to facilitate their colonization. Aside from the negative roles of
894 some members of Kunitz proteins in plant disease resistance, these proteins have also been
895 shown to play positive roles against pathogens. In a recent study, StMLP1, a Kunitz protein
896 from potato, enhanced resistance against *Ralstonia solanacearum*, a bacteria pathogen
897 (Wang et al. 2023). Kunitz proteins have also been shown to act as antimicrobial effects on
898 fungal pathogens under *in vitro* conditions (Clemente et al. 2019; da Silva et al. 2023).
899 However, the exact mechanism by which Kunitz imparts resistance against plant pathogens
900 is not known. Our AFM analysis has revealed significant interactions of soybean Kunitz
901 proteins with the serine proteases secreted by *M. phaseolina*. So far, no reports are available
902 for plant Kunitz proteins targeting fungal serine proteases. Fungal serine proteases are
903 important virulence factors that have been shown to degrade plant immune enzymes, such as
904 chitinases (Han et al. 2019; Jashni et al. 2015). Therefore, inhibiting the virulence-associated
905 fungal serine proteases might be a mechanism of action to impart resistance toward the
906 invading pathogens. Structured-guided engineering of protease-inhibitor complexes is
907 emerging as an important strategy to engineer resistance in plants against pathogens. Here,
908 key amino acid residues are mutated to disrupt the protease-inhibitor interaction, leading to
909 enhanced resistance against pathogens (Schuster et al. 2024). Therefore, the protease-
910 inhibitor complexes that we have predicted using AFM in our study might find their application
911 in structured-guided engineering for improving resistance in soybean against *M. phaseolina*.

912 **Conflicts of Interest**

913 The authors declare no conflicts of interest.

914 **Data Availability Statement**

915 The scripts used for the analysis are available on github
916 (https://github.com/chemicalglycobiologylab) and the data generated during the analysis is
917 available in zonedo (https://zenodo.org/records/13779899). The mass spectrometry
918 proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [1]
919 partner repository with the dataset identifier PXD059244. The data that support the findings of
920 this study are available from the corresponding author upon request.

921 **Acknowledgments**

This research work was supported by SRG grant (SRG/2022/000528) provided by Science and Engineering Research Board (SERB) India and ACRG grant provided by the Birla Institute of Technology and Science, Pilani, (BITS Pilani) Pilani campus. We thank the Central Instrumentation Facility (CIF), BITS Pilani, for providing the GC-MS and confocal microscopy facilities for the research work. We thank the High-Performance Computing (HPC) Facilities, BITS Pilani for MM/PBSA simulation analysis. We also thank the Department of Biological Sciences, BITS Pilani for providing the epifluorescence microscopy and other infrastructural facilities for conducting this research work. We also thank Dr. Siddesh S Kamat and Ojal Saharan, Indian Institute of Science Education and Research (IISER), Pune, for graciously providing the FP-alkyne probe for our ABPP experiments. We thank Dr. B S Gill, Punjab Agricultural University, Ludhiana for providing the soybean seeds. We thank Department of Genetics and Plant Breeding, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur for providing the soybean seeds and the *M. phaseolina*-infected root samples. We thank BITS Pilani, Pilani Campus, for providing Institute Fellowship to Chetan Veeraganti Naveen Prakash and Muthusaravanan Sivaramakrishnan.

Author contributions

BC, CVNP and MS conceived the project; CVNP performed pathogen isolation, infection assays, apoplastic proteome analysis, glycome analysis and competitive ABPP with the help of SG and VP; MS performed AlphaFold and AlphaFold Multimer analysis, MD simulations, callose quantification and deposition assays and Pfam analysis. PKA and MKS provided the seed material, and the *M. phaseolina*-infected root samples; BC supervised the project, designed the experiment, and wrote the manuscript with the contribution from CVNP and MS.

References

Abuqamar, Synan, Suad Ajeb, Arjun Sham, Mohamed Rizq Enan, and Rabah Iratni. 2013. “A Mutation in the Expansin-like 2 Gene Enhances Resistance to Necrotrophic Fungi and Hypersensitivity to Abiotic Stress in *Arabidopsis thaliana*.” *Molecular Plant Pathology* 14(8):813–27. doi: 10.1111/mpp.12049.

Amrate, Pawan K., M. K. Shrivastava, M. S. Bhale, Nisha Agrawal, Giriraj Kumawat, M. Shivakumar, and Vennampally Nataraj. 2023. “Identification and Genetic Diversity Analysis of High-Yielding Charcoal Rot Resistant Soybean Genotypes.” *Scientific Reports* 13(1):8905. doi: 10.1038/s41598-023-35688-2.

Arafat, Md. Yasir, Kanika Narula, Pragya Nalwa, Atreyee Sengupta, Niranjan Chakraborty, and Subhra Chakraborty. 2022. “Proteomic Analysis of Phytopathogenic Fungus *Macrophomina phaseolina* Identify Known and Novel Mycelial Proteins with Roles in Growth and Virulence.” *Journal of Proteins and Proteomics* 13(3):149–57. doi: 10.1007/s42485-022-00095-0.

Arganda-Carreras, Ignacio, Verena Kaynig, Curtis Rueden, Kevin W. Eliceiri, Johannes Schindelin, Albert Cardona, and H. Sebastian Seung. 2017. “Trainable Weka Segmentation: A Machine Learning Tool for Microscopy Pixel Classification.” *Bioinformatics* 33(15):2424–26.

Bode, W., R. Engh, D. Musil, U. Thiele, R. Huber, A. Karshikov, J. Brzin, J. Kos, and V. Turk. 1988. “The 2.0 A X-ray Crystal Structure of Chicken Egg White Cystatin and Its Possible Mode of Interaction with Cysteine Proteinases.” *The EMBO Journal* 7(8):2593–99. doi: 10.1002/j.1460-2075.1988.tb03109.x.

Brown, W. 1924. “A Method of Isolating Single Strains of Fungi by Cutting Out a Hyphal Tip.” *Annals of Botany* 38(150):402–4.

Buscaill, Pierre, and Renier AL van der Hoorn. 2021. “Defeated by the Nines: Nine Extracellular Strategies to Avoid Microbe-Associated Molecular Patterns Recognition in Plants.” *The Plant Cell* 33(7):2116–30.

Cantu, D., A. R. Vicente, L. C. Greve, F. M. Dewey, A. B. Bennett, J. M. Labavitch, and A. L. T. Powell. 2008. “The Intersection between Cell Wall Disassembly, Ripening, and Fruit Susceptibility to *Botrytis cinerea*.” *Proceedings of the National Academy of Sciences* 105(3):859–64. doi: 10.1073/pnas.0709813105.

Chandrasekar, Balakumaran, Thomas Colby, Asif Emran Khan Emon, Jianbing Jiang, Tram Ngoc Hong, Joji Grace Villamor, Anne Harzen, Herman S. Overkleeft, and Renier A. L. van der Hoorn. 2014. “Broad-Range Glycosidase Activity Profiling *.” *Molecular & Cellular Proteomics* 13(10):2787–2800. doi: 10.1074/mcp.O114.041616.

Chandrasekar, Balakumaran, Tram Ngoc Hong, and Renier A. L. van der Hoorn. 2017. “Inhibitor Discovery by Convolution ABPP.” *Methods in Molecular Biology (Clifton, N.J.)* 1491:47–56. doi: 10.1007/978-1-4939-6439-0_4.

Chandrasekar, Balakumaran, Alan Wanke, Stephan Wawra, Pia Saake, Lisa Mahdi, Nyasha Charura, Miriam Neidert, Gereon Poschmann, Milena Malisic, Meik Thiele, Kai Stühler, Murali Dama, Markus Pauly, and Alga Zuccaro. 2022. “Fungi Hijack a Ubiquitous Plant Apoplastic Endoglucanase to Release a ROS Scavenging β-Glucan Decasaccharide to Subvert Immune Responses.” *The Plant Cell* 34(7):2765–84. doi: 10.1093/plcell/koac114.

Chen, Jun-Bin, Shu-Wen Bao, Ya-Li Fang, Lu-Yang Wei, Wang-Sheng Zhu, You-Liang Peng, and Jun Fan. 2021. “An LRR-Only Protein Promotes NLP-Triggered Cell Death and Disease Susceptibility by Facilitating Oligomerization of NLP in *Arabidopsis*.” *New Phytologist* 232(4):1808–22. doi: 10.1111/nph.17680.

Chowdhury, Supriyo, Arpita Basu, and Surekha Kundu. 2017. “Biotrophy-Necrotrophy Switch in Pathogen Evoke Differential Response in Resistant and Susceptible Sesame Involving Multiple Signaling Pathways at Different Phases.” *Scientific Reports* 7(1):17251. doi: 10.1038/s41598-017-17248-7.

Clemente, Marina, Mariana G. Corigliano, Sebastián A. Pariani, Edwin F. Sánchez-López, Valeria A. Sander, and Víctor A. Ramos-Duarte. 2019. “Plant Serine Protease Inhibitors: Biotechnology Application in Agriculture and Molecular Farming.” *International Journal of Molecular Sciences* 20(6):1345. doi: 10.3390/ijms20061345.

Cock, Peter J. A., Tiago Antao, Jeffrey T. Chang, Brad A. Chapman, Cymon J. Cox, Andrew Dalke, Iddo Friedberg, Thomas Hamelryck, Frank Kauff, Bartek Wilczynski, and Michiel J. L. de Hoon. 2009. “Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics.” *Bioinformatics* 25(11):1422–23. doi: 10.1093/bioinformatics/btp163.

Coculo, Daniele, Daniele Del Corpo, Miguel Ozáez Martínez, Pablo Vera, Gabriella Piro, Monica De Caroli, and Vincenzo Lionetti. 2023. “*Arabidopsis* Subtilases Promote Defense-Related Pectin Methylesterase Activity and Robust Immune Responses to *Botrytis* Infection.” *Plant Physiology and Biochemistry* 201:107865. doi: 10.1016/j.plaphy.2023.107865.

1010 Cosgrove, Daniel J. 2005. “Growth of the Plant Cell Wall.” *Nature Reviews Molecular Cell*
1011 *Biology* 6(11):850–61. doi: 10.1038/nrm1746.

1012 Dettmann, Steffen, Heidemarie Thierfelder, B. N. Prasad-’, Abdelaal Yousef Shamseldinl, and
1013 Dietrich Wemerl. n.d. “Glycosidase and Glycosyltransferase Activity Increase in
1014 Arbuscular Mycorrhiza Infected Legume Roots.”

1015 Di, Ruolin, Lixiang Zhu, Zhen Huang, Minyan Lu, Liuyu Yin, Caixia Wang, Yixue Bao,
1016 Zhenzhen Duan, Charles A. Powell, Qin Hu, Jisen Zhang, Muqing Zhang, and Wei
1017 Yao. 2024. “*Fusarium sacchari* FsNis1 Induces Plant Immunity.” *Gene* 907:148260.
1018 doi: 10.1016/j.gene.2024.148260.

1019 Foster, Cliff E., Tina M. Martin, and Markus Pauly. 2010. “Comprehensive Compositional
1020 Analysis of Plant Cell Walls (Lignocellulosic Biomass) Part II: Carbohydrates.” *Journal*
1021 *of Visualized Experiments : JoVE* (37):1837. doi: 10.3791/1837.

1022 Gaudioso-Pedraza, Rocio, Martina Beck, Lisa Frances, Philip Kirk, Carolina Ripodas, Andreas
1023 Niebel, Giles E. D. Oldroyd, Yoselin Benitez-Alfonso, and Fernanda de Carvalho-
1024 Niebel. 2018. “Callose-Regulated Symplastic Communication Coordinates Symbiotic
1025 Root Nodule Development.” *Current Biology* 28(22):3562-3577.e6. doi:
1026 10.1016/j.cub.2018.09.031.

1027 Ghosh, Sumit, Vijaykumar S. Meli, Anil Kumar, Archana Thakur, Niranjan Chakraborty, Subhra
1028 Chakraborty, and Asis Datta. 2011. “The N-Glycan Processing Enzymes α-
1029 Mannosidase and β-D-N-Acetylhexosaminidase Are Involved in Ripening-Associated
1030 Softening in the Non-Climacteric Fruits of Capsicum.” *Journal of Experimental Botany*
1031 62(2):571–82. doi: 10.1093/jxb/erq289.

1032 Godson, Alice, and Renier A. L. van der Hoorn. 2021. “The Front Line of Defence: A Meta-
1033 Analysis of Apoplastic Proteases in Plant Immunity.” *Journal of Experimental Botany*
1034 72(9):3381–94. doi: 10.1093/jxb/eraa602.

1035 Gorshkov, Vladimir, and Ivan Tsers. 2022. “Plant Susceptible Responses: The Underestimated
1036 Side of Plant–Pathogen Interactions.” *Biological Reviews* 97(1):45–66. doi:
1037 10.1111/brv.12789.

1038 Gu, Xiaoyu, Zhiyan Cao, Zhiqiang Li, Haiyue Yu, and Wende Liu. 2024. “Plant Immunity
1039 Suppression by an β-1,3-Glucanase of the Maize Anthracnose Pathogen
1040 *Colletotrichum graminicola*.” *BMC Plant Biology* 24(1):339. doi: 10.1186/s12870-024-
1041 05053-0.

1042 Guerra, Yasel, Vinicio Armijos-Jaramillo, Tirso Pons, Eduardo Tejera, and Colin Berry. 2023.
1043 “Canonical or Noncanonical? Structural Plasticity of Serine Protease-Binding Loops in
1044 Kunitz-STI Protease Inhibitors.” *Protein Science* 32(2):e4570. doi: 10.1002/pro.4570.

1045 Guillen, Karine de, Léa Mammri, Jérôme Gracy, André Padilla, Philippe Barthe, François Hoh,
1046 Mounia Lahfa, Justine Rouffet, Yohann Petit-Houdenot, Thomas Kroj, and Marc-Henri
1047 Lebrun. 2024. “*Zymoseptoria tritici* Effectors Structurally Related to Killer Proteins
1048 UmV-KP4 and UmV-KP6 Are Toxic to Fungi, and Define Extended Protein Families in
1049 Fungi.” 2024.10.14.618152.

1050 Gupta, Gagan D., Ravindra Bansal, Hiral Mistry, Bharati Pandey, and Prasun K. Mukherjee.
1051 2021. “Structure-Function Analysis Reveals *Trichoderma virens* Tsp1 to Be a Novel
1052 Fungal Effector Protein Modulating Plant Defence.” *International Journal of Biological*
1053 *Macromolecules* 191:267–76. doi: 10.1016/j.ijbiomac.2021.09.085.

Han, Li-Bo, Yuan-Bao Li, Fu-Xin Wang, Wen-Yan Wang, Jun Liu, Jia-He Wu, Nai-Qin Zhong, Shen-Jie Wu, Gai-Li Jiao, Hai-Yun Wang, and Gui-Xian Xia. 2019. "The Cotton Apoplastic Protein CRR1 Stabilizes Chitinase 28 to Facilitate Defense against the Fungal Pathogen *Verticillium dahliae*." *The Plant Cell* 31(2):520–36. doi: 10.1105/tpc.18.00390.

Hobson, Neil, and Michael K. Deyholos. 2013. "Genomic and Expression Analysis of the Flax (*Linum usitatissimum*) Family of Glycosyl Hydrolase 35 Genes." *BMC Genomics* 14(1):344. doi: 10.1186/1471-2164-14-344.

Holm, Liisa. 2022. "Dali Server: Structural Unification of Protein Families." *Nucleic Acids Research* 50(W1):W210–15. doi: 10.1093/nar/gkac387.

Homma, Felix, Jie Huang, and Renier A. L. van der Hoorn. 2023. "AlphaFold-Multimer Predicts Cross-Kingdom Interactions at the Plant-Pathogen Interface." *Nature Communications* 14(1):6040. doi: 10.1038/s41467-023-41721-9.

Homma, Felix, Joy Lyu, and Renier A. L. van der Hoorn. 2024. "Using AlphaFold Multimer to Discover Interkingdom Protein–Protein Interactions." *The Plant Journal* 120(1):19–28. doi: 10.1111/tpj.16969.

Hong, Pingjing, Zibo Zhang, Yunzhi Zhou, Xianqi Lu, Ehsan Sadeghnezhad, Qianqian Pang, Zhen Tao, Yuanxin Cheng, Bo Wang, and Haifeng Jia. 2024. "Polygalacturonase Inhibiting Protein Enhances Cell Wall Strength of Strawberry Fruit for Resistance to *Botrytis cinerea* Infection." *Scientia Horticulturae* 327:112850. doi: 10.1016/j.scienta.2024.112850.

Hong, Tram, and Renier Hoorn. 2014. "DIGE-ABPP by Click Chemistry: Pairwise Comparison of Serine Hydrolase Activities from the Apoplast of Infected Plants." *Methods in Molecular Biology (Clifton, N.J.)* 1127:183–94. doi: 10.1007/978-1-62703-986-4_15.

Hosseini-Zahani, Fatemeh, and Parissa Taheri. 2023. "The Role of Oxidative Burst, Antioxidant Genes and Enzymes in Association with Callose in Tomato Reaction to Various Taxonomic Groups of *Rhizoctonia* Spp." *Physiological and Molecular Plant Pathology* 128:102174. doi: 10.1016/j.pmpp.2023.102174.

Hou, Shuguo, Pierce Jamieson, and Ping He. 2018. "The Cloak, Dagger, and Shield: Proteases in Plant–Pathogen Interactions." *Biochemical Journal* 475(15):2491–2509. doi: 10.1042/BCJ20170781.

Jashni, Mansoor Karimi, Ivo H. M. Dols, Yuichiro Iida, Sjef Boeren, Henriek G. Beenen, Rahim Mehrabi, Jérôme Collemare, and Pierre J. G. M. de Wit. 2015. "Synergistic Action of a Metalloprotease and a Serine Protease from *Fusarium oxysporum* f. sp. *lycopersici* Cleaves Chitin-Binding Tomato Chitinases, Reduces Their Antifungal Activity, and Enhances Fungal Virulence." *Molecular Plant-Microbe Interactions®* 28(9):996–1008. doi: 10.1094/MPMI-04-15-0074-R.

Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and

Demis Hassabis. 2021. “Highly Accurate Protein Structure Prediction with AlphaFold.” *Nature* 596(7873):583–89. doi: 10.1038/s41586-021-03819-2.

Kaschani, Farnusch, Christian Gu, Sherry Niessen, Heather Hoover, Benjamin F. Cravatt, and Renier. A. L. van der Hoorn. 2009. “Diversity of Serine Hydrolase Activities of Unchallenged and *Botrytis*-Infected *Arabidopsis thaliana*.” *Molecular & Cellular Proteomics* 8(5):1082–93. doi: 10.1074/mcp.M800494-MCP200.

Kaschani, Farnusch, Sabrina Nickel, Bikram Pandey, Benjamin F. Cravatt, Markus Kaiser, and Renier A. L. van der Hoorn. 2012. “Selective Inhibition of Plant Serine Hydrolases by Agrochemicals Revealed by Competitive ABPP.” *Bioorganic & Medicinal Chemistry* 20(2):597–600. doi: 10.1016/j.bmc.2011.06.040.

van Kempen, Michel, Stephanie S. Kim, Charlotte Tumescheit, Milot Mirdita, Jeongjae Lee, Cameron L. M. Gilchrist, Johannes Söding, and Martin Steinegger. 2024. “Fast and Accurate Protein Structure Search with Foldseek.” *Nature Biotechnology* 42(2):243–46. doi: 10.1038/s41587-023-01773-0.

Kim, Sang-Jin, Balakumaran Chandrasekar, Anne C. Rea, Linda Danhof, Starla Zemelis-Durfee, Nicholas Thrower, Zachary S. Shepard, Markus Pauly, Federica Brandizzi, and Kenneth Keegstra. 2020. “The Synthesis of Xyloglucan, an Abundant Plant Cell Wall Polysaccharide, Requires CSLC Function.” *Proceedings of the National Academy of Sciences* 117(33):20316–24. doi: 10.1073/pnas.2007245117.

Laskowski, Roman, Jagoda Jabłońska, Lukas Pravda, Radka Svobodová Vařeková, and Janet Thornton. 2017. “PDBsum: Structural Summaries of PDB Entries.” *Protein Science* 27. doi: 10.1002/pro.3289.

Lemkul, Justin. 2019. “From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package [Article v1.0].” *Living Journal of Computational Molecular Science* 1(1):0–53. doi: 10.33011/livecoms.1.1.5068.

Li, Jing, Günter Brader, and E. Tapio Palva. 2008. “Kunitz Trypsin Inhibitor: An Antagonist of Cell Death Triggered by Phytopathogens and Fumonisin B1 in *Arabidopsis*.” *Molecular Plant* 1(3):482–95. doi: 10.1093/mp/ssn013.

Lindahl, Erik, Pär Bjelkmar, Per Larsson, Michel A. Cuendet, and Berk Hess. 2010. “Implementation of the Charmm Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models.” *Journal of Chemical Theory and Computation* 6(2):459–66. doi: 10.1021/ct900549r.

Liu, Lifeng, Jonathan Paulitz, and Markus Pauly. 2015. “The Presence of Fucogalactoxyloglucan and Its Synthesis in Rice Indicates Conserved Functional Importance in Plants.” *Plant Physiology* 168(2):549–60. doi: 10.1104/pp.15.00441.

Liu, Muxing, Fangfang Wang, Bo He, Jiexiong Hu, Ying Dai, Weizhong Chen, Mingxi Yi, Haifeng Zhang, Yonghao Ye, Zhongli Cui, Xiaobo Zheng, Ping Wang, Weiman Xing, and Zhengguang Zhang. 2024. “Targeting *Magnaporthe oryzae* Effector MoErs1 and Host Papain-like Protease OsRD21 Interaction to Combat Rice Blast.” *Nature Plants* 10(4):618–32. doi: 10.1038/s41477-024-01642-x.

Liu, Yanan, Edan Jackson, Xueru Liu, Xingchuan Huang, Renier A. L. van der Hoorn, Yuelin Zhang, and Xin Li. 2024. “Proteolysis in Plant Immunity.” *The Plant Cell* koae142. doi: 10.1093/plcell/koae142.

Mason, Kelly N., Gayani Ekanayake, and Antje Heese. 2020. “Chapter 10 - Staining and Automated Image Quantification of Callose in *Arabidopsis* Cotyledons and Leaves.” Pp. 181–99 in *Methods in Cell Biology*. Vol. 160, *Plant Cell Biology*, edited by C. T. Anderson, E. S. Haswell, and R. Dixit. Academic Press.

Mehmood, Sohaib, Soren Skou Thirup, Sarah Ahmed, Nabila Bashir, Ahsan Saeed, Maria Rafiq, Qamar Saeed, Muhammad Najam-ul-Haq, Binish Khaliq, Muhammad Ibrahim, Wadi B. Alonazi, and Ahmed Akrem. 2024. “Crystal Structure of Kunitz-Type Trypsin Inhibitor: Entomotoxic Effect of Native and Encapsulated Protein Targeting Gut Trypsin of *Tribolium castaneum* Herbst.” *Computational and Structural Biotechnology Journal* 23:3132–42. doi: 10.1016/j.csbj.2024.07.023.

Méndez-Yáñez, Angela, Darwin Sáez, Francisca Rodríguez-Arriaza, Claudio Letelier-Naritelli, Felipe Valenzuela-Riffo, and Luis Morales-Quintana. 2024. “Involvement of the GH38 Family Exoglycosidase α-Mannosidase in Strawberry Fruit Ripening.” *International Journal of Molecular Sciences* 25(12):6581. doi: 10.3390/ijms25126581.

Meng, Elaine C., Eric F. Pettersen, Gregory S. Couch, Conrad C. Huang, and Thomas E. Ferrin. 2006. “Tools for Integrated Sequence-Structure Analysis with UCSF Chimera.” *BMC Bioinformatics* 7(1):339. doi: 10.1186/1471-2105-7-339.

Mirdita, Milot, Konstantin Schütze, Yoshitaka Moriwaki, Lim Heo, Sergey Ovchinnikov, and Martin Steinegger. 2022. “ColabFold: Making Protein Folding Accessible to All.” *Nature Methods*. doi: 10.1038/s41592-022-01488-1.

Mirdita, Milot, Martin Steinegger, and Johannes S"oding. 2019. “MMseqs2 Desktop and Local Web Server App for Fast, Interactive Sequence Searches.” *Bioinformatics* 35(16):2856–58. doi: 10.1093/bioinformatics/bty1057.

Misas-Villamil, Johana C., Renier A. L. van der Hoorn, and Gunther Doehlemann. 2016. “Papain-like Cysteine Proteases as Hubs in Plant Immunity.” *New Phytologist* 212(4):902–7. doi: 10.1111/nph.14117.

Mistry, Jaina, Sara Chuguransky, Lowri Williams, Matloob Qureshi, Gustavo A. Salazar, Erik L. L. Sonnhammer, Silvio C. E. Tosatto, Lisanna Paladin, Shriya Raj, Lorna J. Richardson, Robert D. Finn, and Alex Bateman. 2021. “Pfam: The Protein Families Database in 2021.” *Nucleic Acids Research* 49(D1):D412–19. doi: 10.1093/nar/gkaa913.

Mooney, Brian Críostóir, and Renier AL van der Hoorn. 2024. “Novel Structural Insights at the Extracellular Plant-Pathogen Interface.” *Current Opinion in Plant Biology* 82:102629.

Nelson, Jonathan R., Sujeet Verma, Nahla V. Bassil, Chad E. Finn, James F. Hancock, Glenn S. Cole, Steven J. Knapp, and Vance M. Whitaker. 2021. “Discovery of Three Loci Increasing Resistance to Charcoal Rot Caused by *Macrophomina phaseolina* in Octoploid Strawberry.” *G3 Genes|Genomes|Genetics* 11(3):jkab037. doi: 10.1093/g3journal/jkab037.

Nghi, Do Huu, Britta Bittner, Harald Kellner, Nico Jehmlich, René Ullrich, Marek J. Pecyna, Paula Nousiainen, Jussi Sipilä, Le Mai Huong, Martin Hofrichter, and Christiane Liers. 2012. “The Wood Rot Ascomycete *Xylaria polymorpha* Produces a Novel GH78 Glycoside Hydrolase That Exhibits α-l-Rhamnosidase and Feruloyl Esterase Activities and Releases Hydroxycinnamic Acids from Lignocelluloses.” *Applied and Environmental Microbiology* 78(14):4893–4901. doi: 10.1128/AEM.07588-11.

1185 Okedigba, Ayoyinka O., M. Luciana Rosso, Daisy Yu, Chao Shang, Haibo Huang, Bo Zhang,
1186 and Daniel G. S. Capelluto. 2023. "Comparative Binding Affinity Analysis of Soybean
1187 Meal Bowman-Birk and Kunitz Trypsin Inhibitors in Interactions with Animal Serine
1188 Proteases." *ACS Food Science & Technology* 3(8):1344–52. doi:
1189 10.1021/acsfoodscitech.3c00158.

1190 Ortiz, Viviana, Hao-Xun Chang, Hyunkyu Sang, Janette Jacobs, Dean K. Malvick, Richard
1191 Baird, Febina M. Mathew, Consuelo Estévez de Jensen, Kiersten A. Wise, Gloria M.
1192 Mosquera, and Martin I. Chilvers. 2023. "Population Genomic Analysis Reveals
1193 Geographic Structure and Climatic Diversification for *Macrophomina phaseolina*
1194 Isolated from Soybean and Dry Bean across the United States, Puerto Rico, and
1195 Colombia." *Frontiers in Genetics* 14. doi: 10.3389/fgene.2023.1103969.

1196 Pandey, Abhay K., and Ashwani K. Basandrai. 2021. "Will Macrophomina Phaseolina Spread
1197 in Legumes Due to Climate Change? A Critical Review of Current Knowledge." *Journal*
1198 *of Plant Diseases and Protection* 128(1):9–18. doi: 10.1007/s41348-020-00374-2.

1199 Paulus, Judith K., Jiorgos Kourelis, Selva Ramasubramanian, Felix Homma, Alice Godson,
1200 Anja C. Hörger, Tram Ngoc Hong, Daniel Krahn, Laura Ossorio Carballo, Shuaishuai
1201 Wang, Joe Win, Matthew Smoker, Sophien Kamoun, Suomeng Dong, and Renier A.
1202 L. van der Hoorn. 2020. "Extracellular Proteolytic Cascade in Tomato Activates
1203 Immune Protease Rcr3." *Proceedings of the National Academy of Sciences*
1204 117(29):17409–17. doi: 10.1073/pnas.1921101117.

1205 Pettersen, Eric F., Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M.
1206 Greenblatt, Elaine C. Meng, and Thomas E. Ferrin. 2004. "UCSF Chimera—A
1207 Visualization System for Exploratory Research and Analysis." *Journal of*
1208 *Computational Chemistry* 25(13):1605–12. doi: 10.1002/jcc.20084.

1209 Pettolino, Filomena A., Cherie Walsh, Geoffrey B. Fincher, and Antony Bacic. 2012.
1210 "Determining the Polysaccharide Composition of Plant Cell Walls." *Nature Protocols*
1211 7(9):1590–1607. doi: 10.1038/nprot.2012.081.

1212 Pineda-Fretez, Amiliana, Adriana Orrego, Julio César Masaru Iehisa, María Eugenia Flores-
1213 Giubi, Javier E. Barúa, Rosa Sánchez-Lucas, Jesús Jorrín-Novo, and María Cristina
1214 Romero-Rodríguez. 2023. "Secretome Analysis of the Phytopathogen *Macrophomina*
1215 *phaseolina* Cultivated in Liquid Medium Supplemented with and without Soybean Leaf
1216 Infusion." *Fungal Biology* 127(5):1043–52. doi: 10.1016/j.funbio.2023.04.001.

1217 Rech, Stefanie S., Sven Heidt, and Natalia Requena. 2013. "A Tandem Kunitz Protease
1218 Inhibitor (KPI106)–Serine Carboxypeptidase (SCP1) Controls Mycorrhiza
1219 Establishment and Arbuscule Development in *Medicago truncatula*." *The Plant Journal*
1220 75(5):711–25. doi: 10.1111/tpj.12242.

1221 Redzynia, Izabela, Anna Ljunggren, Anna Bujacz, Magnus Abrahamson, Mariusz Jaskolski,
1222 and Grzegorz Bujacz. 2009. "Crystal Structure of the Parasite Inhibitor Chagasin in
1223 Complex with Papain Allows Identification of Structural Requirements for Broad
1224 Reactivity and Specificity Determinants for Target Proteases." *The FEBS Journal*
1225 276(3):793–806. doi: 10.1111/j.1742-4658.2008.06824.x.

1226 Rocafort, Mercedes, Joanna K. Bowen, Berit Hassing, Murray P. Cox, Brogan McGreal, Silvia
1227 de la Rosa, Kim M. Plummer, Rosie E. Bradshaw, and Carl H. Mesarich. 2022. "The
1228 *Venturia inaequalis* Effector Repertoire Is Dominated by Expanded Families with
1229 Predicted Structural Similarity, but Unrelated Sequence, to Avirulence Proteins from

Other Plant-Pathogenic Fungi." *BMC Biology* 20(1):246. doi: 10.1186/s12915-022-01442-9.

Schenk, Sebastian, and Adam Schikora. 2015. "Staining of Callose Depositions in Root and Leaf Tissues." *BIO-PROTOCOL* 5(6). doi: 10.21769/BioProtoc.1429.

Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, and Benjamin Schmid. 2012. "Fiji: An Open-Source Platform for Biological-Image Analysis." *Nature Methods* 9(7):676–82.

Schmid, Ernst. 2023. "Colabfold Batch AlphaFold-2-Multimer Structure Analysis Pipeline."

Schuster, Mariana, Sophie Eisele, Liz Armas-Egas, Till Kessenbrock, Jiorgos Kourelis, Markus Kaiser, and Renier A. L. Van Der Hoorn. 2024. "Enhanced Late Blight Resistance by Engineering an EpiC2B-insensitive Immune Protease." *Plant Biotechnology Journal* 22(2):284–86. doi: 10.1111/pbi.14209.

Seong, Kyungyong, and Ksenia V. Krasileva. 2023. "Prediction of Effector Protein Structures from Fungal Phytopathogens Enables Evolutionary Analyses." *Nature Microbiology* 8(1):174–87. doi: 10.1038/s41564-022-01287-6.

Shao, Dandan, Damon L. Smith, Mehdi Kabbage, and Mitchell G. Roth. 2021. "Effectors of Plant Necrotrophic Fungi." *Frontiers in Plant Science* 12. doi: 10.3389/fpls.2021.687713.

Shi, Feifei, Ying Wang, Fang Zhang, Xingxing Yuan, Huatao Chen, Xuehao Chen, Xin Chen, and Xiaoyan Cui. 2020. "Soybean Endo-1,3-Beta-Glucanase (GmGLU) Interaction With Soybean Mosaic Virus-Encoded P3 Protein May Contribute to the Intercellular Movement." *Frontiers in Genetics* 11. doi: 10.3389/fgene.2020.536771.

Shrivastava, Madhulika, and Arun Goyal. 2024. "Unveiling the Structural and Functional Perspectives of a Bifunctional α-l-Arabinofuranosidase/Endo-β-1,4-Xylanase (BoGH43_35) from *Bacteroides ovatus*." *Archives of Biochemistry and Biophysics* 764:110232. doi: 10.1016/j.abb.2024.110232.

da Silva, Marciele S., Layrana de Azevedo dos Santos, Gabriel Bonan Taveira, Celso Shiniti Nagano, Renata Pinheiro Chaves, de Oliveira Carvalho, Rosana Rodrigues, and Valdirene Moreira Gomes. 2023. "Trypsin/α-Amylase Inhibitors from Capsicum Chinense Seeds: Characterization and Antifungal Activity against Fungi of Agronomic Importance." *Protein and Peptide Letters* 30(3):260–74. doi: 10.2174/0929866530666230221141804.

Sinha, Nilanjan, Sourav Kumar Patra, and Sanjay Ghosh. 2022. "Secretome Analysis of *Macrophomina phaseolina* Identifies an Array of Putative Virulence Factors Responsible for Charcoal Rot Disease in Plants." *Frontiers in Microbiology* 13. doi: 10.3389/fmicb.2022.847832.

Sivaramakrishnan, Muthusaravanan, Sakshi Goel, Nikhil Ratnaparkhi, and Balakumaran Chandrasekar. 2023. "Chemiluminescence-Based Assay to Monitor Early Oxidative Bursts in Soybean (*Glycine Max*) Lateral Roots." *Current Protocols* 3(8):e869. doi: 10.1002/cpz1.869.

Sivaramakrishnan, Muthusaravanan, Chetan Veeraganti Naveen Prakash, and Balakumaran Chandrasekar. 2024. "Multifaceted Roles of Plant Glycosyl Hydrolases during

Pathogen Infections: More to Discover." *Planta* 259(5):113. doi: 10.1007/s00425-024-04391-5.

Sperschneider, Jana, and Peter N. Dodds. 2022. "EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic Effectors in Fungi and Oomycetes." *Molecular Plant-Microbe Interactions: MPMI* 35(2):146–56. doi: 10.1094/MPMI-08-21-0201-R.

Tan, Li, Stefan Eberhard, Sivakumar Pattathil, Clayton Warder, John Glushka, Chunhua Yuan, Zhangying Hao, Xiang Zhu, Utku Avci, Jeffrey S. Miller, David Baldwin, Charles Pham, Ronald Orlando, Alan Darvill, Michael G. Hahn, Marcia J. Kieliszewski, and Debra Mohnen. 2013. "An *Arabidopsis* Cell Wall Proteoglycan Consists of Pectin and Arabinoxylan Covalently Linked to an Arabinogalactan Protein." *The Plant Cell* 25(1):270–87. doi: 10.1105/tpc.112.107334.

Tanthanuch, Waraporn, Mallika Chantarangsee, Janjira Maneesan, and James Ketudat-Cairns. 2008. "Genomic and Expression Analysis of Glycosyl Hydrolase Family 35 Genes from Rice (*Oryza sativa* L.)." *BMC Plant Biology* 8(1):84. doi: 10.1186/1471-2229-8-84.

The UniProt Consortium. 2024. "UniProt: The Universal Protein Knowledgebase in 2025." *Nucleic Acids Research* gkae1010. doi: 10.1093/nar/gkae1010.

Thynne, Elisha, Haider Ali, Kyungyong Seong, Mohammad Abukhalaf, Marco A. Guerreiro, Victor M. Flores-Nunez, Rune Hansen, Maja Salman, Jason J. Rudd, and Kostya Kanyuka. 2024. "An Array of *Zymoseptoria tritici* Effectors Suppress Plant Immune Responses." *bioRxiv* 2024–03.

Valdés-Tresanco, Mario S., Mario E. Valdés-Tresanco, Pedro A. Valiente, and Ernesto Moreno. 2021. "gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS." *Journal of Chemical Theory and Computation* 17(10):6281–91. doi: 10.1021/acs.jctc.1c00645.

Wang, Bingsen, Yuqi Wang, Wenfeng He, Mengshu Huang, Liu Yu, Dong Cheng, Juan Du, Botao Song, and Huilan Chen. 2023. "StMLP1, as a Kunitz Trypsin Inhibitor, Enhances Potato Resistance and Specifically Expresses in Vascular Bundles during *Ralstonia solanacearum* Infection." *The Plant Journal* 116(5):1342–54. doi: 10.1111/tpj.16428.

Wang, Shuaishuai, Rongkang Xing, Yan Wang, Haidong Shu, Shenggui Fu, Jie Huang, Judith K. Paulus, Mariana Schuster, Diane G. O. Saunders, Joe Win, Vivianne Vleeshouwers, Yuanchao Wang, Xiaobo Zheng, Renier A. I van der Hoorn, and Suomeng Dong. 2021. "Cleavage of a Pathogen Apoplastic Protein by Plant Subtilases Activates Host Immunity." *New Phytologist* 229(6):3424–39. doi: 10.1111/nph.17120.

Wang, Yan, Yuanchao Wang, and Yiming Wang. 2020. "Apoplastic Proteases: Powerful Weapons against Pathogen Infection in Plants." *Plant Communications* 1(4). doi: 10.1016/j.xplc.2020.100085.

Wang, Ying, Xifeng Li, Baofang Fan, Cheng Zhu, and Zhixiang Chen. 2021. "Regulation and Function of Defense-Related Callose Deposition in Plants." *International Journal of Molecular Sciences* 22(5):2393. doi: 10.3390/ijms22052393.

Weng, Gaoqi, Ercheng Wang, Zhe Wang, Hui Liu, Feng Zhu, Dan Li, and Tingjun Hou. 2019. "HawkDock: A Web Server to Predict and Analyze the Protein-Protein Complex Based on Computational Docking and MM/GBSA." *Nucleic Acids Research* 47(W1):W322–30. doi: 10.1093/nar/gkz397.

Yuan, Shuguang, H. C. Stephen Chan, and Zhenquan Hu. 2017. “Using PyMOL as a Platform for Computational Drug Design.” *WIREs Computational Molecular Science* 7(2):e1298. doi: 10.1002/wcms.1298.

Main Figures

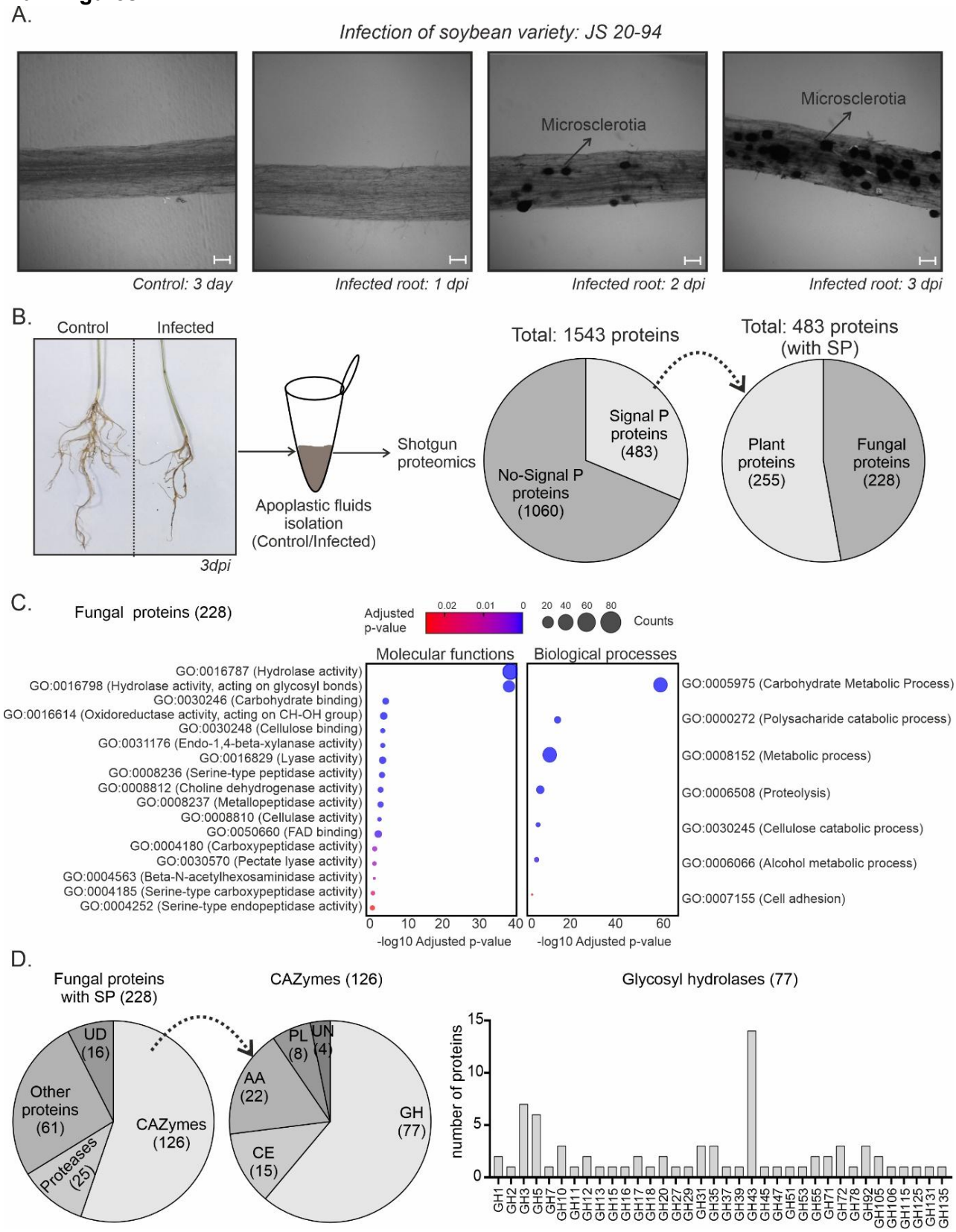

Figure 1: *Macrophomina phaseolina* secretes a large repertoire of carbohydrate-active enzymes (CAZymes) in the soybean root apoplast during infection. (A) Microscopy of *M. phaseolina* soybean root colonization. The two-week-old seedlings of highly susceptible soybean variety, JS 20-94, were treated with 25% (w/v) of homogenised actively growing *M. phaseolina* culture using root-dip inoculation

method. The roots collected from 1 dpi, 2 dpi and 3 dpi were washed and cleared overnight with 70% ethanol. Fungal colonisation in the soybean roots was visualised using Differential Interface Contrast (DIC) microscopic analysis (magnification= 10X, scale = 100µm). The arrow marks indicate the microsclerotia structure formed in the infected roots at 2 dpi and 3 dpi, respectively. (B) Shotgun proteome analysis of apoplastic fluids collected from control and *M. phaseolina*-infected soybean roots. Apoplastic fluids were collected from 2-week-old control and *M. phaseolina*-infected soybean plants at 3 dpi. Trypsin-digested peptides of the control and infected apoplastic fluids was subjected to shot-gun proteome analysis using LC-MS/MS. Signal P 6.0 (http://www.cbs.dtu.dk/services/ SignalP/) tool was used to identify the proteins with signal peptides (C) Gene ontology analysis of proteins secreted by *M. phaseolina*. Gene ontology analysis was performed using FungiFun v2.2 with a hypergeometric distribution test and Benjamini-Hochberg FDR threshold of ≤ 0.05. The colour of the bubbles represents the adjusted p-value denoting its statistical significance, and the size of the bubbles represents the counts, i.e., the number of genes involved in a particular molecular function or biological process. GO IDs represent taxonomy that statistically determines the involvement of those sets of genes in the specific biological processes and molecular functions as indexed in QuickGo (https://www.ebi.ac.uk/QuickGO/) (D) Classification of proteins secreted by *M. phaseolina*. The proteins were classified into CAZymes, proteases, other proteins, and unknown domain (UD) based on their domain architecture, which was predicted by Pfam domain analysis. The CAZymes were classified based on the families such as glycosyl hydrolases, carbohydrate esterases, auxiliary activities, and polysaccharide lyases using dbCAN3 server. dpi, days post-infection, SP, signal peptide, UD, unknown domain, GH, glycosyl hydrolases, PL, polysaccharide lyases, CE, carbohydrate esterases, AA, auxiliary activities, UN, unclassified CAZymes

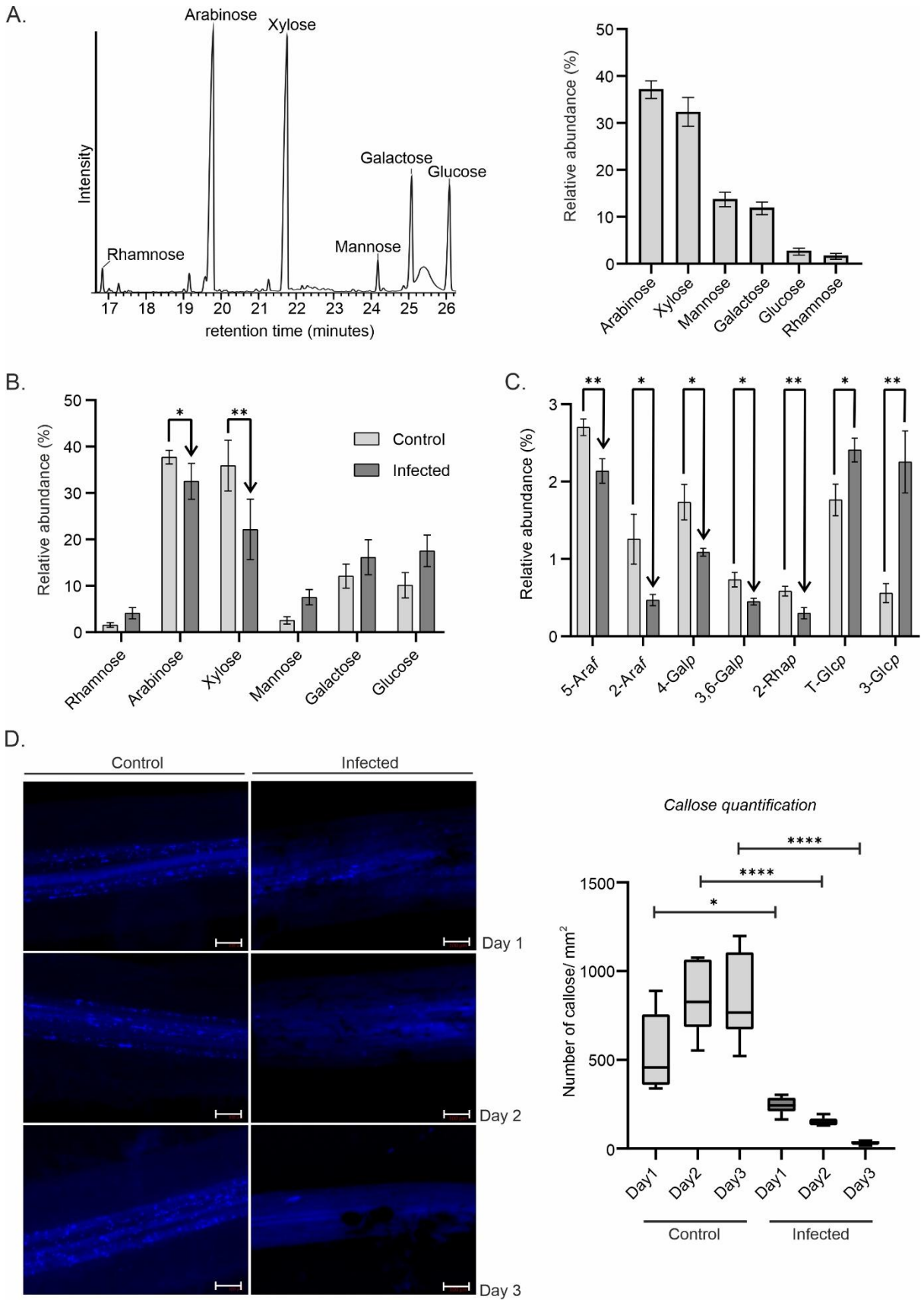

Figure 2: Cell wall composition and callose depositions are altered during *M. phaseolina* infection. (A) Arabinose and xylose are the major neutral sugars detected in the soybean root cell walls. Cell walls were isolated from the soybean roots as alcohol insoluble fraction (AIR) and subjected to monosaccharide composition analysis. The neutral sugars were released from the cell walls by mild

TFA (Trifluoroacetic acid) hydrolysis and derivatised as alditol acetate (AA) and the AA residues were analysed using GC-MS. The detected neutral sugars were annotated based on their retention time and the MS spectra of the respective sugar standards. The experiment was conducted with five independent biological replicates. The neutral sugars are expressed as a percentage of the total peak areas of sugars detected. (B) The relative abundance of arabinose and xylose are reduced in the soybean root cell walls infected with *M. phaseolina*. The relative abundance of the neutral sugars detected in the control and infected root AIR samples is expressed as a percentage of the total peak areas of sugars detected. The statistical significance was determined using two-tailed t-test (*: p-value ≤ 0.05, **: 0.001 < p-value ≤ 0.01). The experiment was conducted with five independent biological replicates. (C) The relative abundance of characteristic glycosyl linkages derived from arabinogalactan and the rhamnogalacturonan 1 (RG1) are reduced in the soybean root cell walls infected with *M. phaseolina*. Glycosyl linkage analysis was performed with the AIR samples isolated from control and *M. phaseolina*-infected soybean roots. The sugar residues were derivatized as partially methylated alditol acetates (PMAA) and the residues were analysed using GC-MS. The glycosyl linkages were annotated based on their retention time and the publicly available CCRC spectra library. The glycosyl linkages with significant differences in relative abundance between the control and infected AIR samples are represented. The statistical significance was determined using two-tailed t-test (*: p-value ≤ 0.05, **: 0.001 < p-value ≤ 0.01). (D) The callose levels in the soybean roots are reduced during *M. phaseolina* infection. Aniline blue staining was performed on soybean roots collected at 1 dpi, 2 dpi, and 3 dpi. The stained lateral roots were observed under an epifluorescence microscope using a DAPI filter with an excitation wavelength of 370 nm and an emission wavelength of 509 nm (magnification= 10X, scale =100µm). Callose depositions were quantified using the trainable weka segmentation v3.3.4 plugin of ImageJ 1.54f. The experiment was conducted with six independent biological replicates, two replicates per plant in a completely randomized experimental design. The statistical significance was determined using two-tailed t-test (*: p-value ≤ 0.05, ****: p-value ≤ 0.0001).

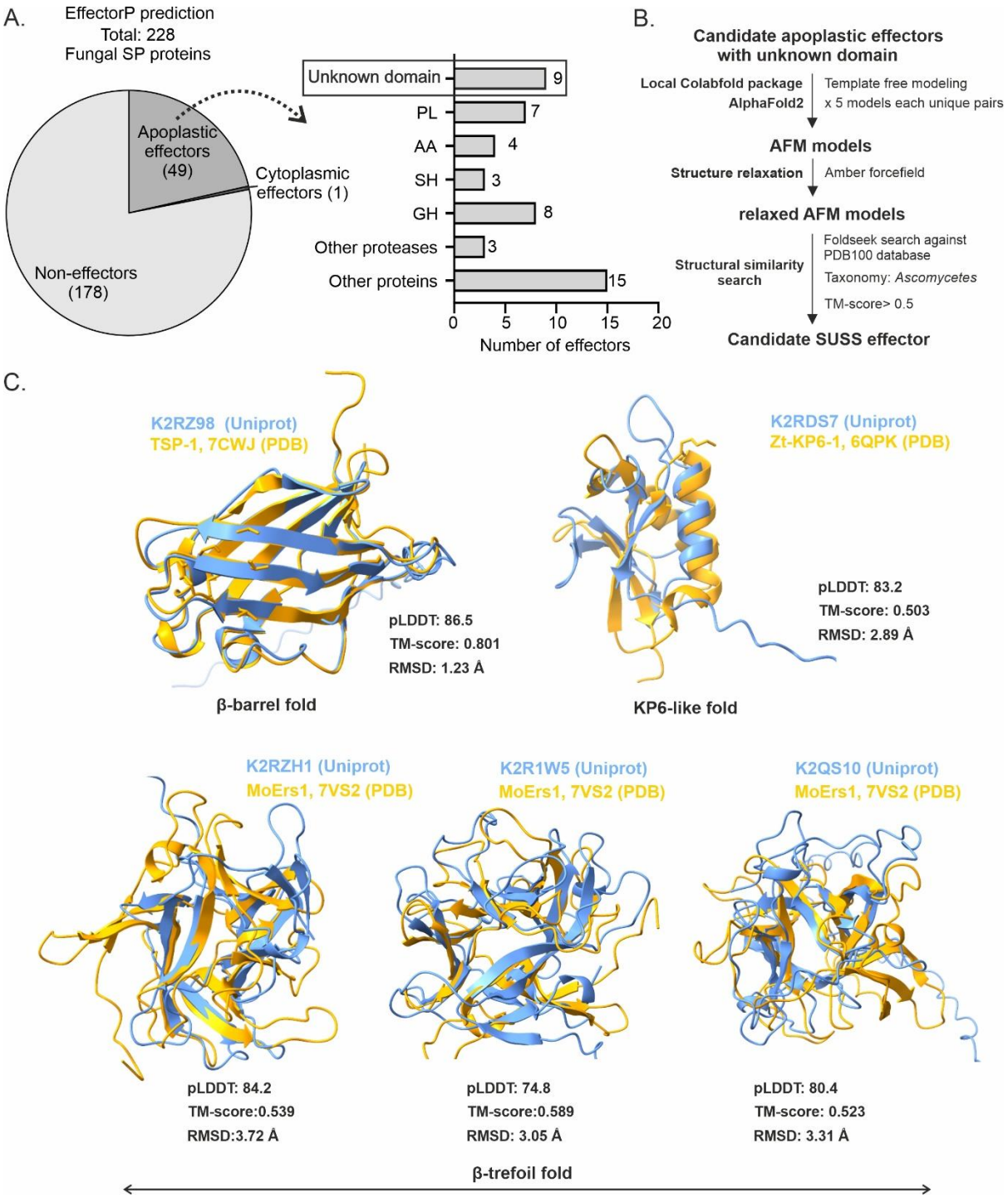

Figure 3: *M. phaseolina* secretes SUSS effectors and effectors with novel structural folds in the soybean root apoplast during infection. (A) Apoplastic effectors are secreted by *M. phaseolina*. *M. phaseolina* proteins carrying signal peptide were analysed using EffectorP Fungi 3.0 (https://effectorp.csiro.au/) for effector prediction. The predicted apoplastic effector candidates were classified based on their domain architecture using Pfam analysis. (B) Pipeline for predicting SUSS effector candidates using the AF2 and Foldseek program. The candidate apoplastic effectors with unknown domain were modelled using AF2 implemented in colabfold pipeline without template. The AF2 modelled effector candidates were screened against PDB100 database using Foldseek search with taxonomy restriction to Ascomycetes. The protein candidates with TM-score >0.5 were considered structurally similar. (C) AF2 predicts SUSS effector candidates secreted by *M. phaseolina*. Superposition (structural alignment) of AF2 predicted putative SUSS effectors from *M. phaseolina* to experimentally resolved 3D tertiary structures available in PDB100 database were performed using

Foldseek search and matchmaker module of chimera molecular viewer. The predicted putative SUSS effectors are represented in blue whereas crystal structures of known effectors are represented in yellow. The protein K2RZ98 shows similarity to Tsp-1 effector from a beneficial fungi *Trichoderma virens* (PDB ID: 7CWJ) as suggested by TM-score (0.801) and RMSD (1.23 Å). K2RDS7 protein shows folding similar to Zt-KP6-1 effector (PDB ID: 6QPK) from *Zymoseptoria tritici* as represented by TM-score (0.503) and RMSD (2.89 Å). The proteins K2RZH1, K2R1W5, and K2QS10 show similarity to MoERS-1 effector from *Pyricularia oryzae* (PDB ID: 7VS2). SP, signal peptide, PL, polysaccharide lyases, AA, auxiliary activities, SH, serine hydrolases, GH, glycosyl hydrolases, AF2, AlphaFold 2, SUSS, sequence unrelated structurally similar, PDB, Protein data bank, pLDTT, predicted local distance difference test, TM-score, template modeling scores, RMSD, root mean square deviation.

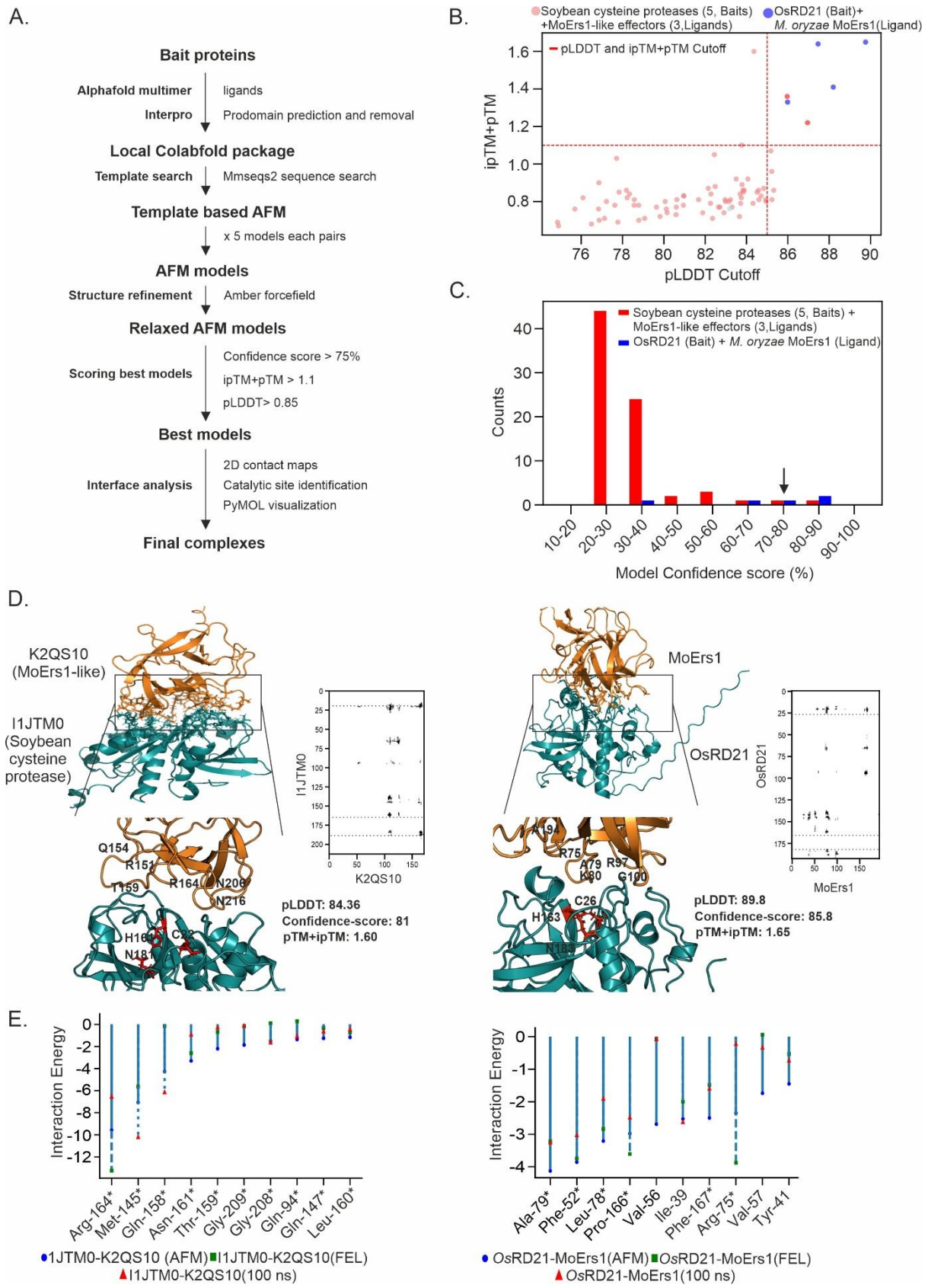

Figure 4: AFM analysis predicts significant interaction of a MoErs1-like effector (UniprotID. K2QS10) with a soybean cysteine protease (UniprotID. IJTM0). (A) The workflow for predicting interactions between proteins (bait and ligands) using AFM. The pipeline involves modeling ligand and bait proteins as heterodimers using the template from MMseqs2 search. The AFM-modeled structures were relaxed

using an amber force field, and the top five models were considered. The resultant structural models predicted by AFM were screened using the stringent ipTM+pTM score (> 1.1), pLDDTscore (>0.85), and model confidence (> 75%) to select the top interacting partners. The protein complexes satisfying this threshold were further analyzed for interfacial contacts using 2D contact maps, and the interactions were visualized using the PyMOL software. (B) AFM analysis reveals interactions of MoErs1-like effectors with soybean cysteine proteases. The cysteine proteases detected in the soybean root apoplast and the rice RD21 protease (OsRD21) were screened against MoErs1-like effectors secreted by *M. phaseolina* and the MoErs1 effector from *M. oryzae* using AFM. This analysis was implemented in the local colabfold pipeline using the ipTM + pTM and pLDDT scores. 120 AFM models are distributed over the full pTM+ ipTM score and pLDDT range, with a subpopulation of high confidence AFM model predictions with pTM+ ipTM score > 1.1 and pLDDT>0.85 (top right corner). The top-ranked AFM models are colored in darker shades of red and blue (Red represents the models of MoErs1-like effector and soybean cysteine proteases, and blue represents the models of MoErs1 and OsRD21). (C) AFM predicts the interaction of MoErs1-like effector (UniprotID. K2QS10) with a soybean cysteine protease (UniprotID. IJTM0) with a high model confidence score. The AFM interaction model with a confidence score > 75% is indicated with an arrow. (D) The MoErs1-like effector (UniprotID. K2QS10) and MoErs1 interact in close proximity to the active sites of respective cysteine proteases. The AFM-predicted interaction models were visualized using the PyMOL software and presented as a cartoon backbone 3D representation. The heterodimer interface of interaction between K2QS10 (MoErs1-like effector) and I1JTM0 (soybean cysteine proteases) and MoErs1 and OsRD21 is zoomed and represented as an insert. The putative active site catalytic triads in cysteine proteases were predicted using InterPro and represented as red sticks. The amino acid residues forming hydrogen bonds near the active site are represented as single-letter codes with residue numbers. The interfacial contacts in the protein complexes are presented as 2D contact maps. (E) The loop region of MoErs1-like effector and MoErs1 displays energetically favourable interactions with the respective cysteine proteases. The per-residue decomposition plots of MoErs1-like effector and MoErs1 with the respective cysteine protease complex were using the HawkDock tool. The decomposition plots represent the interaction energy of amino acid residue in the loop region of MoErs1-like effector and MoErs1. The dark blue dot represents the protein complex from the AFM model, the green square represents the low energy conformer obtained through FEL analysis, and the red triangle represents the conformer from 100th ns in the MD trajectories. The light blue line in the plot indicates the range of interaction energy values for each amino acid, and the blue line is dotted in the plot for easy visualization of other data points. The symbol * represents amino acids at loop regions of MoErs1 and MoErs1-like effectors that interact with the respective cysteine proteases. pLDTT, predicted local distance difference test, pTM, predicted template modelling, ipTM, interface predicted template modelling, AF2, AlphaFold 2, SUSS, Sequence unrelated structurally similar, PDB, Protein data bank, TM-score, template modeling scores, 2D, two-dimensional, FEL, Free energy landscape, AFM, AlphaFold multimer, ns, nano seconds.

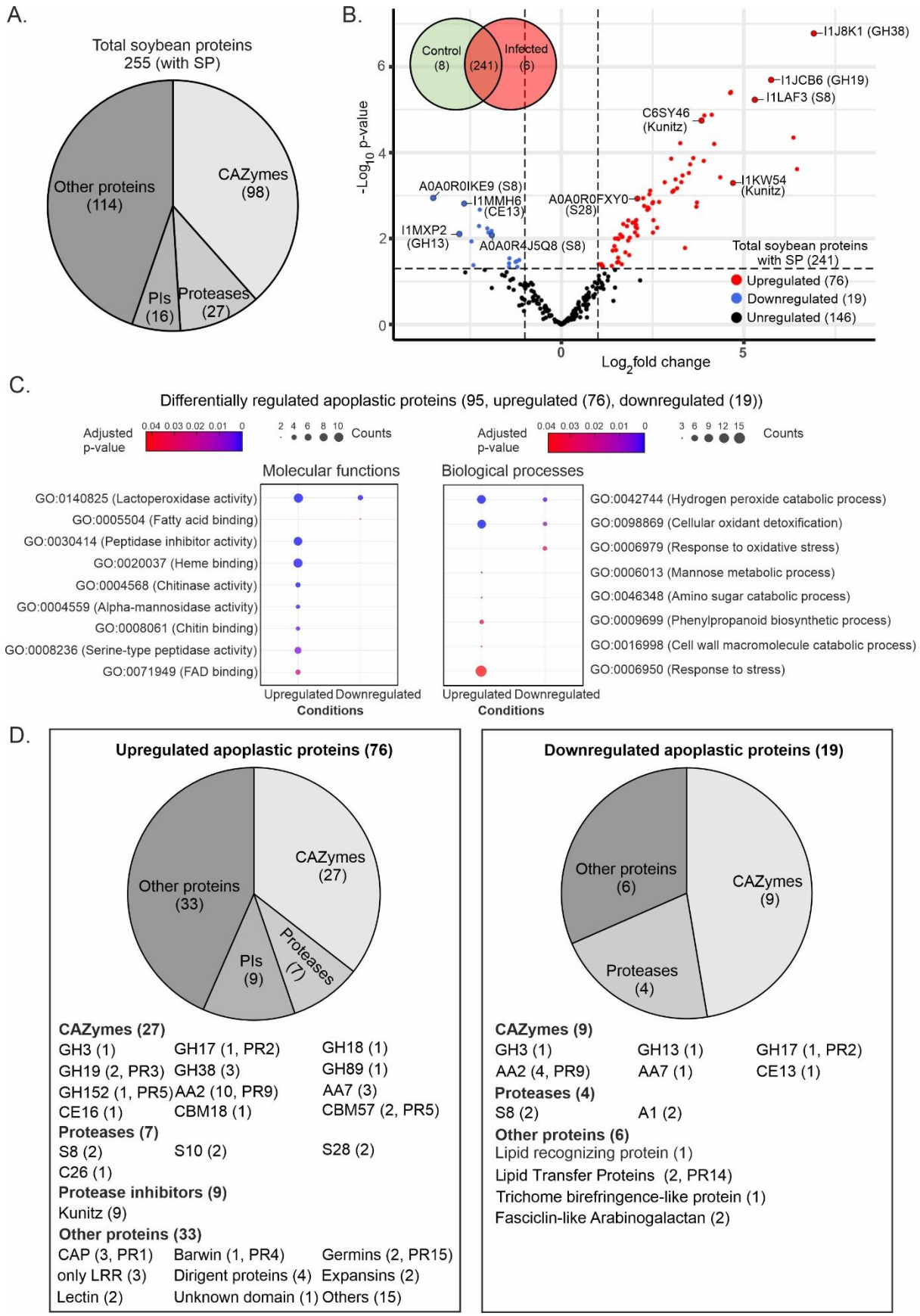

Figure 5: Quantitative proteome analysis reveals the dynamics of soybean-secreted proteins in root apoplast during *M. phaseolina* infection. (A) A diverse class of soybean-secreted proteins are detected in the root apoplast. 255 soybean proteins with signal peptides were classified based on their domain

architecture using dbCAN3 and Pfam analysis. (B) Differential regulation of soybean apoplastic proteins upon *M. phaseolina* infection. Soybean apoplastic proteins detected in control and infected apoplastic fluids are represented as a Venn diagram. Soybean apoplastic proteins are ranked in a volcano plot according to their statistical P value (y-axis) as -log 10 and their relative Label-Free Quantification (LFQ) values (Log_2 fold change, x-axis) between control *and M. phaseolina*-infected soybean root samples (n=3) using the EnhancedVolcano package in the R program. Log p-value signifies its statistical significance, i.e., the lower the p-value, the greater the significance of the value. Log_2 fold change was calculated to understand the differential regulation of the apoplastic proteins in the infected apoplast compared to the control apoplast. The upregulated proteins with log_2 fold change > 1 and p-value < 0.05 are represented in red colour, downregulated proteins with log_2 fold change < -1 and p-value < 0.05 are represented in blue colour and the rest of the proteins are represented as black dots. The proteins with bigger dots represent the top two highly regulated proteins of the major classes. (C) Gene ontology analysis of the differentially regulated soybean apoplastic proteins. GO annotation was performed for upregulated and downregulated soybean apoplastic proteins using g:GOSt module of g:Profiler based on molecular functions and biological processes. Statistical significance was calculated using the Fisher one-tailed test in combination with the Benjamini-Hochberg with an FDR threshold of ≤0.05 to correct for multiple testing. The colour code signifies the adjusted p-value and the bubble size denotes the counts of genes. The adjusted p-value is a statistical measure denoting the significance of enrichment, where a low value indicates the result is less likely to occur by chance. GO IDs represent taxonomy that statistically determines the involvement of those sets of genes in the specific biological processes and molecular functions as indexed in QuickGo (https://www.ebi.ac.uk/QuickGO/). (D) Classification of differentially regulated soybean apoplastic proteins. Significantly upregulated and downregulated soybean proteins were classified based on their domain architecture using Pfam analysis. The proteins of their respective classes are listed. SP, signal peptide, GH, glycosyl hydrolases, PIs, protease inhibitors, CE, carbohydrate esterases, CBM, carbohydrate binding module, AA, auxiliary activities, CAP, Cysteine-rich, Antigen-5, PR-1, LRR, Leucine-rich repeats, S#, serine proteases, C#, cysteine proteases, A#, aspartic proteases, PR proteins, pathogenesis-related proteins. #, number

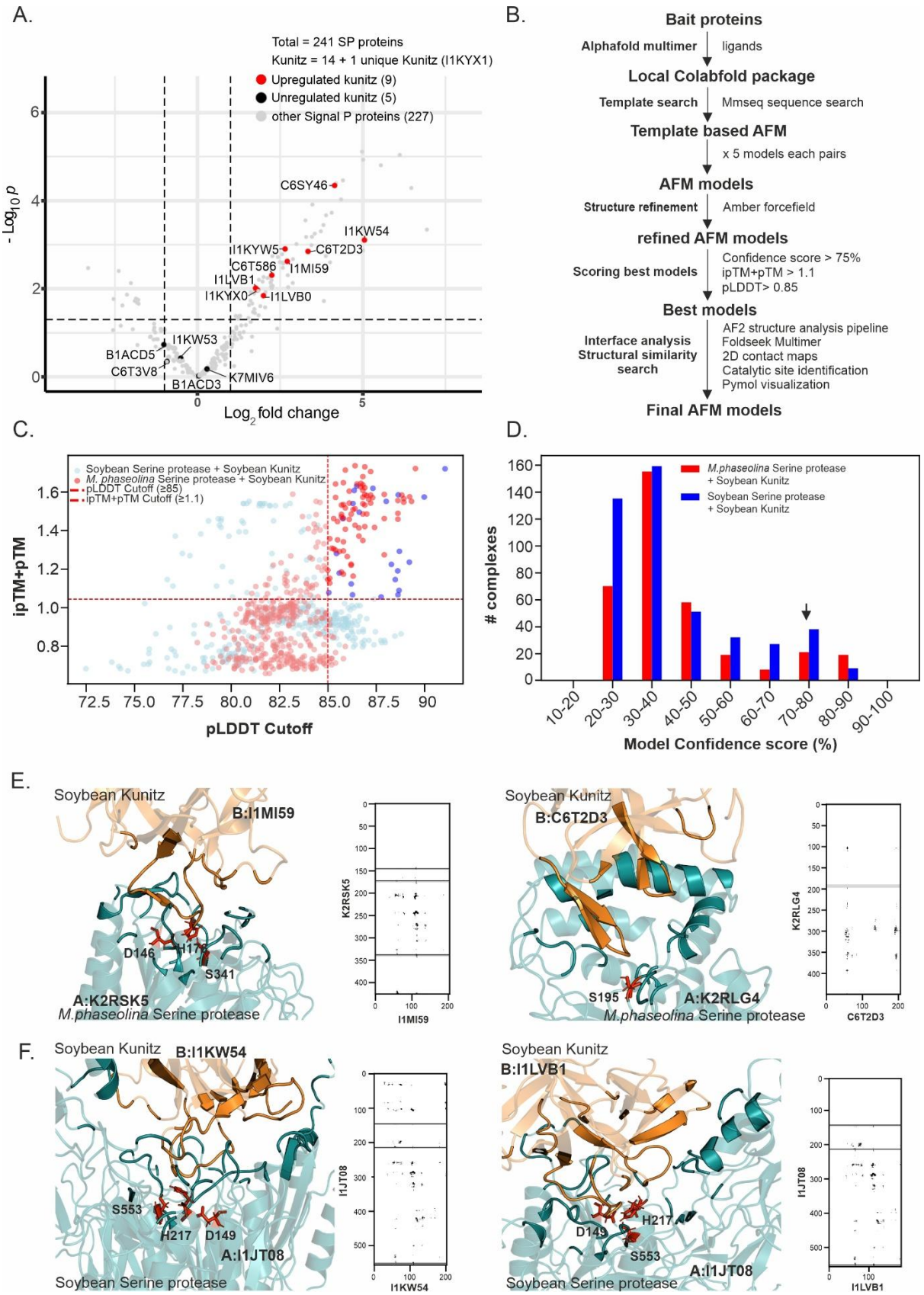

Figure 6: AlphaFold Multimer (AFM) analysis predicts the interaction of soybean Kunitz and serine proteases. (A) Soybean Kunitz proteins are upregulated in soybean root apoplast during *M. phaseolina* infection. Soybean apoplastic proteins are ranked in a volcano plot according to their statistical P value (y-axis) as -log 10 and their relative Label-Free Quantification (LFQ) values (Log_2 fold

change, x-axis) between control *and M. phaseolina*-infected soybean root samples (n=3) using EnhancedVolcano package of R software. Log p-value signifies its statistical significance, i.e., the lower the p-value, the greater the significance of the value. Log_2 fold change was calculated to understand the differential regulation of the apoplastic proteins in the infected apoplast compared to the control apoplast. Red and black dots represent the upregulated and unregulated soybean Kunitz, respectively, and the grey dot represents the other signal peptide containing proteins identified in the soybean apoplast. (B) The workflow for predicting interactions between proteins (bait and ligands) using AFM. The pipeline involves modeling ligand and bait proteins as heterodimers using the template from MMseqs2 search. The AFM-modeled structures were relaxed using an amber force field, and the top five models were considered. The resultant structural models predicted by AFM were screened using the stringent ipTM+pTM score (> 1.1), plDDTscore (>0.85), and model confidence (> 75%) to select the top interacting partners. The protein complexes satisfying this threshold were further analyzed for interfacial contacts using 2D contact maps, and the interactions were visualized using the PyMOL software. (C) AFM analysis reveals interactions of soybean Kunitz with serine proteases of soybean and *M. phaseolina*. The upregulated soybean Kunitz in our dataset was screened against the randomly selected ten secreted serine proteases of soybean and seven secreted serine proteases *M. phaseolina* using AFM. This analysis was implemented in the local colabfold pipeline using the ipTM + pTM and pLDDT scores. 120 AFM models are distributed over the full pTM+ ipTM score and pLDDT range, with a subpopulation of high confidence AFM model predictions with pTM+ ipTM score > 1.1 and pLDDT>0.85 (top right corner). The top-ranked AFM models are colored in darker shades of red and blue (Red represents the models of *M. phaseolina* serine proteases and Soybean Kunitz, and blue represents the models of Soybean serine proteases and Soybean Kunitz). (D) AFM predicts the interaction of soybean Kunitz with serine proteases of soybean and *M. phaseolina* with a high model confidence score. The AFM interaction model with a confidence score > 75% is indicated with an arrow. (E) The Soybean Kunitz proteins (UniprotIDs: I1MI59 and C6T2D3) interact with the active site of the *M. phaseolina* serine proteases (UniprotIDs: K2RSK5 and K2RLG4). (F). The Soybean Kunitz proteins (UniprotIDs: I1KW54 and I1LVB1) interact with the active site of the soybean serine protease (UniprotIDs: I1JT08). The AFM-predicted interaction models in Figure 6E and Figure 6F were visualized using the PyMOL software and presented as a cartoon backbone 3D representation. The heterodimer interface of interaction between soybean Kunitz and serine proteases is zoomed and represented. The annotated active site catalytic triads in serine proteases are represented as red sticks. The interfacial contacts in the protein complexes are presented as 2D contact maps. SP, signal peptide, AFM, AlphaFold Multimer, AF2, AlphaFold2, 2D, two-dimensional, pTM, predicted template modeling scores, ipTM, interface predicted template modeling score, pLDDT, predicted local distance difference test score.

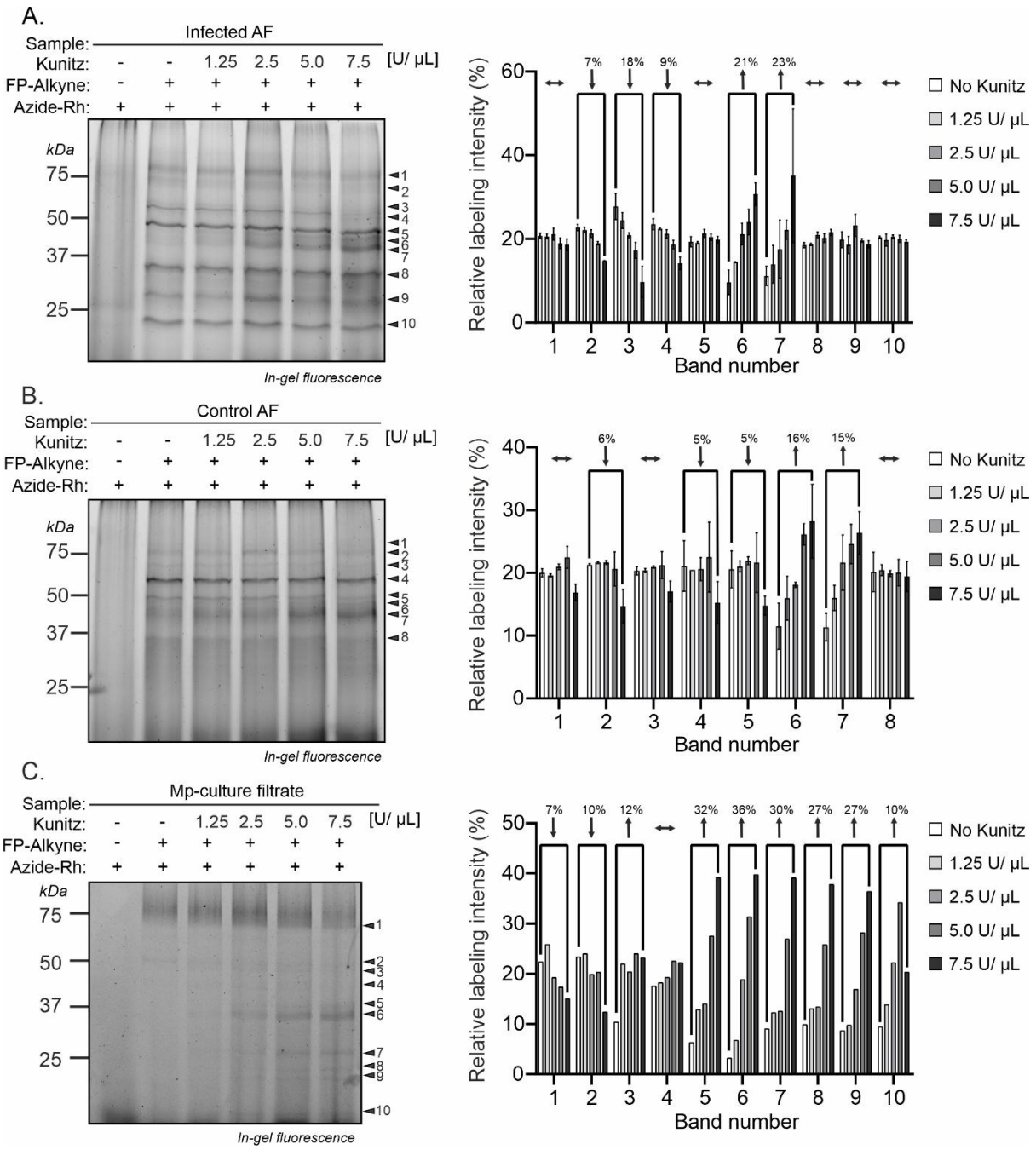

Figure 7: Competitive Activity-based Protein profiling (ABPP) reveals the interaction of soybean Kunitz with serine proteases of soybean and *M. phaseolina* at their active site. (A) Competitive ABPP reveals the suppression of FP-labeled signals in the *M. phaseolina* infected apoplastic fluids preincubated with commercial soybean Kunitz. (B) Competitive ABPP reveals the suppression of FP-labeled signals in control apoplastic fluids preincubated with commercial soybean Kunitz. (C) Competitive ABPP reveals the suppression of FP-labeled signals in the culture filtrates of *M. phaseolina*. The apoplastic fluids in Figure 7A and Figure 7B were isolated from soybean roots at 3 days post-infection (dpi). The culture filtrate in Figure 7C was collected from the 3-day-old liquid cultures inoculated with *M. phaseolina*. The apoplastic fluids or the culture filtrate samples were pre-incubated with various concentrations of commercially available soybean Kunitz (1.25, 2.5, 5, 7.5 U/ µL) for 30 minutes. The preincubated proteome was treated with the FP-alkyne probe for labelling and subsequently subjected to Cu-catalysed click chemistry reaction with Azide Flour 488. The fluorescently-labeled proteins were separated in 10% SDS PAGE gels under 200 v for 45 min and were detected by in-gel fluorescence using ChemiDoc MP (Biorad) scanner 532/28 excitation/emission filter of Pro-Q Emerald 488 in the blue epi filter. The arrow indicates the bands detected and are annotated accordingly. Densitometric

analysis of the annotated bands was performed using Image Lab 6.1. software and the relative labeling intensity of each band was calculated and represented as a percentage. The percentage decrease or increase of labeling intensities respective to the untreated sample is indicated with the arrow mark. AF, apoplastic fluids, Mp, *M. phaseolina*, kDa, Kilo Dalton, FP, fluorophosphonate, Rh, Rhodamine.