Characterising Effector Protein, AP1260, of the Myrtle Rust Causing *Austropuccinia psidii.*

Jovarn Victys Sullivan

10th November 2022

Project report submitted in partial fulfilment of the requirements for the degree B.Sc. (Hons.) in the School of Biological Sciences, University of Canterbury

This work was funded by the Ministry of Business, Innovation and Employment (Ngā Rākau Taketake – Myrtle Rust and Kauri Dieback Research, C09X1817)

Abstract

Significance. Myrtle rust is caused by the invasive fungus *Austropuccinia psidii* and is incredibly infectious and physically devastating to Myrtaceae plants. The pandemic biotype of the disease has spread across five continents and was first detected in New Zealand in 2017. The presence of the disease in Australia has already caused major declines in Myrtaceae populations (*e.g.*, eucalyptus) and now threatens New Zealand natives, including taonga species such as pōhutukawa, mānuka and rātā.

Rationale. Transcriptomics by collaborators at Plant and Food Research identified several proteins that are expressed in the early stages of *A. psidii* infection of mānuka—a signature that they are important for the infection of plant cells. In other fungi, these 'effector proteins' are known to manipulate the host plants cellular processes. My research focuses on one of these effector proteins, AP1260. The aim is to define the characteristics of AP1260 through bioinformatic and biophysical analysis.

Results. The results obtained in this study represent the first studies of an *A. psidii* effector protein. Characteristics such as mass, dynamics, stability, and shape of AP1260 were determined. This led to the discovery of a potentially intrinsically disordered region within AP1260.

Conclusions. Characterisation of AP1260 will vastly improve the knowledge of the mechanisms of *A. psidii* infection and may uncover methods to interfere with its function, making its characterisation critical to the biosecurity of nations with Myrtaceae populations. The outcome of this research will be critical for developing methods for controlling and potentially preventing myrtle rust disease.

Introduction

Myrtle rust disease

Austropuccinia psidii (formerly *Puccinia psidii*, Beenken, 2017) is a biotrophic rust fungus belonging to the family Sphaerophragmiaceae, within the order Pucciniales. *A. psidii* is the causal fungus of a plant rust known as guava rust, eucalyptus rust, or more commonly, myrtle rust (Morin et al., 2014; Narouei-Khandan et al., 2020; Soewarto et al., 2021). Myrtle rust gains its name due to its specificity in infecting members of the Myrtaceae family. Myrtaceae is one of the largest angiosperm families in the world containing over 5000 species across 143 genera (Thornhill et al., 2015). Myrtle rust as a disease is rapidly spreading and devastating to the plants it infects, affecting not only the plants but indirectly other species that rely on them (da S. Machado et al., 2015). *A. psidii* infects the new leaves and stems of plants causing the characteristic physical symptom of bright yellow/orange 'rust' like lesions that form on the infected tissue (Beresford et al., 2020; Fensham et al., 2020; Glen et al., 2007). Depending on the susceptibility of the infected host, infected tissue can become malformed or necrotic, with repeat infection being the main cause of death in infected plants (Fensham et al., 2020; Ferrarezi et al., 2022; Glen et al., 2007). Due to *A. psidii's* capacity to infect such a wide range of Myrtaceae species with devastating impact, myrtle rust has become an issue of critical importance for nations where Myrtaceae plants are of cultural, ecological, and economic significance.

A. psidii was first discovered in South America by Winter (1884) and since then it has been observed to infect over 480 species in over 30 countries distributed across five continents (Carnegie & Pegg, 2018; Coutinho et al., 1998; McTaggart et al., 2018, 2020). There are several biotypes of *A. psidii* that established in varying regions, however, out of these, the pandemic biotype is the most widely spread. It has been found across Asia, the Pacific, and Oceania within the last decade with the majority of infections worldwide being attributed to this biotype (Chock, 2020; Glen et al., 2007; Tobias et al., 2020). In Australia, the presence of the *A. psidii* pandemic biotype over the last decade has caused major declines in native Myrtaceae, including several Eucalyptus species as well as many other species as it has become well established due to a lack of adequate management plans (Carnegie & Pegg, 2018; Fensham et al., 2020). The incursion of pandemic biotype in Aotearoa-New Zealand began in 2017, threatening taonga native Myrtaceae species such as pōhutukawa, mānuka,

rātā, kānuka, and ramarama, and economically important exotic species including Eucalyptus species and feijoa (Beresford et al., 2020; Smith et al., 2020). The spread of the pandemic biotype of *A. psidii* threatens all Myrtaceae populations of the countries the fungus reaches.

A. psidii life and disease cycle

Like most rust fungi, *A. psidii* is an obligate biotroph, a trait that is common among plant pathogens. As such, *A. psidii* is only able to live on the hosts that it infects. *A. psidii* derives its nutrients from its hosts, completing the majority of its life cycle, including growth and reproduction, on or within the hosts it infects (Chock, 2020; Lorrain et al., 2019; Voegele & Mendgen, 2011).

The life cycle and disease cycle of *A. psidii* are tightly linked due to its obligate biotroph lifestyle. Infection usually begins with the adhesion of spores to a plant surface. Should the chemical and physical cues of the plant be right, the spores will germinate (Allen et al., 1991; Chock, 2020). Germination leads to the formation of the penetration structure, the appressorium. The appressorium utilizes physical and enzymatic methods to penetrate the plant surface (Talbot, 2019). After penetration is successful, the fungus initiates the process of nutrient acquisition and host immune suppression through a specialized structure called the haustorium. The haustorium is a specialized structure that secretes virulence factors known as effector molecules that suppress host immune systems and acquire nutrients (Goyet et al., 2019). If the haustorium is successful in nutrient acquisition, hyphal growth will continue until sori form at the surface of the cell (Yong et al., 2019). The sori are the part of the fungus that gives the rust look characteristic of their diseases. Sporulation will then occur for the process to repeat (Chock, 2020; Yong et al., 2019).

Myrtle rust effector proteins

The success of *A. psidii* infection is dependent on the formation of haustoria, which forms within the mesophyll cells of the infected plant (Figure 1). These structures form in both susceptible and resistant plants (Chock, 2020; Leite et al., 2013). To inhibit the host's defensive mechanisms, haustoria secrete effector molecules that attempt to inhibit the host's defensive mechanisms (Chock, 2020; Koeck et al., 2011; Yong et al., 2019). Effector proteins are used by many plant pathogens and vary greatly in their functions and structures to overcome basal plant immune responses (Langin et al., 2020) and to increase their fitness

against more specific hosts (Jones & Dangl, 2006). The basal immunity plants have against pathogens is the ability to detect common and broadly conserved non-specific molecular patterns exhibited by the pathogens (Bent & Mackey, 2007). This is known as 'pattern-triggered immunity', which is caused by the recognition of 'pathogen-associated molecular patterns' (Asai & Shirasu, 2015; Nguyen et al., 2021), including bacterial flagellin and fungal chitin (Bent & Mackey, 2007). Plants have developed a method to detect these commonly deployed 'pathogen-associated molecular patterns' using plasma membrane proteins called 'pattern recognition receptors' (Kazan & Lyons, 2014; Thordal-Christensen, 2020). Pattern recognition receptors in plants are heritable, allowing for the innate immunity driven by evolution to be passed on to the offspring of the plant (Bent & Mackey, 2007).

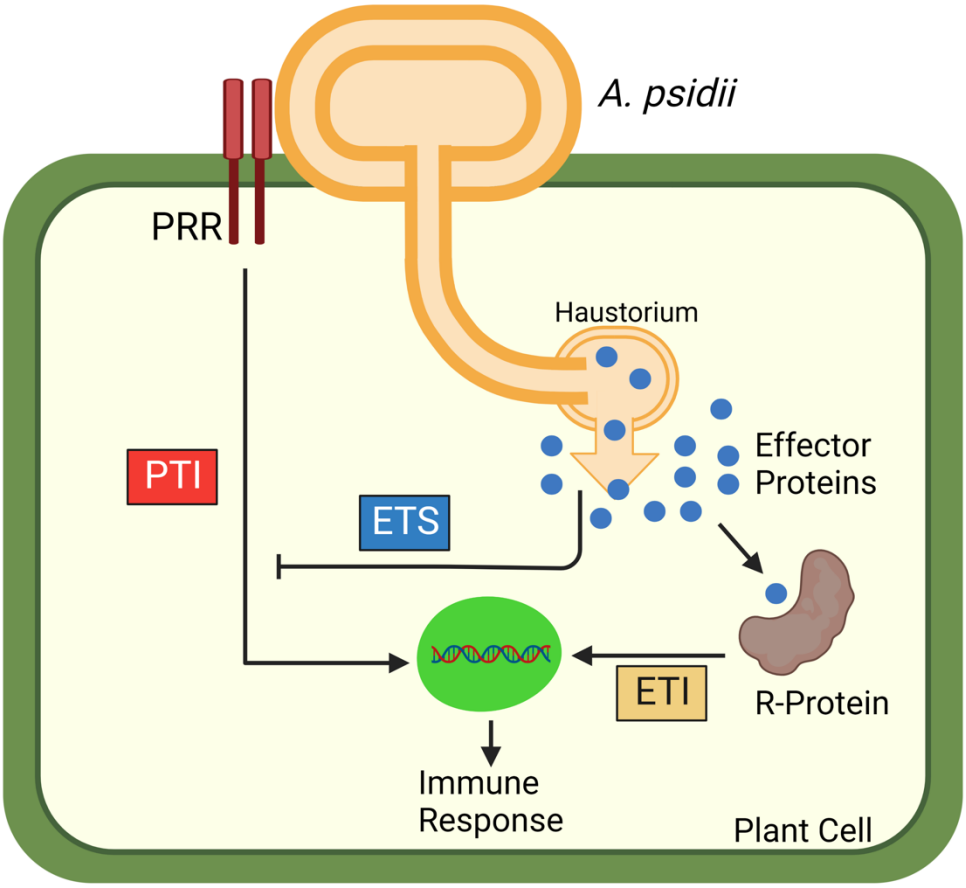


Figure 1 | Overview of plant-*A. psidii* interactions. Common pathogen-associated molecular patterns (PAMPs) are recognized by pattern recognition receptors (PRRs) on the cell surface causing pattern-triggered immunity (PTI). *A. psidii* attempts to interfere with pattern-triggered immunity using effector proteins secreted from the haustorium. If effector proteins successfully inhibit the host immune system, effector-triggered susceptibility (ETS) occurs. If effector proteins are recognized by resistance (R) proteins, effector-triggered immunity (ETI) occurs. Both pattern-triggered immunity and effector-triggered immunity induce immune responses to *A. psidii*. Adapted from Kazan & Lyons, 2014.

To overcome 'pathogen-associated molecular patterns', plant pathogens like *A. psidii* have developed specialized virulence factors called effector molecules. Effectors seek to increase microbial fitness by interfering with 'pattern-triggered immunity' and other basal plant immune responses. When a plant has no resistance to this, a process known as 'effector-triggered susceptibility' is induced where infection by the pathogen is successful (Cui et al., 2015; Jones & Dangl, 2006). In response to the rise of effector proteins in pathogens, some plants have evolved a method to detect them. Known as resistance genes or *R* genes, genes encoding intercellular 'resistance proteins' are the plant's means of either directly or indirectly identifying pathogenic effectors (Asai & Shirasu, 2015; Jones & Dangl, 2006; Kazan

& Lyons, 2014). Detection of effectors by 'resistance proteins' alert the plant to the presence of pathogens, eliciting a strong immune response in a phenomenon known as 'effector-triggered immunity' (Asai & Shirasu, 2015; Cui et al., 2015; Jones & Dangl, 2006; Kazan & Lyons, 2014). This 'evolutionary arms race' has taken place over millions of years as plants and microbial plant pathogens have co-evolved, attempting to gain the upper hand over one another, causing plant defences and microbe mechanisms of infection to drive the evolution of one another (Kazan & Lyons, 2014). Effector proteins are an important tool used by plant pathogens during the infection process. Further study into *A. psidii*'s effector proteins would increase the knowledge of the mechanisms of infection that make it so widespread and devastating to Myrtaceae (Lorrain et al. 2019).

AP1260

Transcriptomics by collaborators at Plant and Food Research (Prof. Grant Smith) identified several proteins that are differentially expressed in the early stages of *A. psidii* infection of Mānuka. This is a signature that they are important for the successful infection of plant cells. The experiments demonstrated that specific genes are overexpressed within the first 24 to 48 hours of infection (currently unpublished). This expression is also seen in other rust fungi during infection (Lorrain et al., 2019). Presently, the functions and structures of these *A. psidii* effector proteins are unknown.

To address the knowledge gap, this project aims to characterize an *A. psidii* effector protein through structural studies. Specifically, this project seeks to define the structural properties of the *A*. *psidii* effector, AP1260.

The size, shape, dynamics, and stability are determined to gain a deeper understanding of the protein and the characteristics that may lend to its potential function. This will be critical for developing methods for controlling and potentially preventing disease. Based on the sequences, AP1260 is predicted to be structurally novel. At present, there are no sequence or structural homologues in the Protein Databank (PDB). The structural determination of AP1260 and other effectors will vastly improve our knowledge of the mechanisms of *A. psidii* infection and may uncover methods to interfere with their function, making their characterisation critical to not only Aotearoa-New Zealand's biosecurity, but the biosecurity of all nations that

A. psidii has spread to. This could then lead to the development of methods for disease treatment and control.

Results

Bioinformatic Analysis

SignalP and EffectorP

The programme SignalP was used to predict if the AP1260 sequence contains a signal peptide. The first 22 residues of AP1260 are predicted to be a signal peptide (Figure 2) with a likelihood of 0.94. For expression and analysis, unless otherwise stated, the signal peptide of AP1260 will be removed from the sequence. EffectorP was used to predict if AP1260 is an effector protein and whether it is cytoplasmic or apoplastic. EffectorP predicted that AP1260 is a cytoplasmic effector (SI Figure 1) with a probability of 0.757.

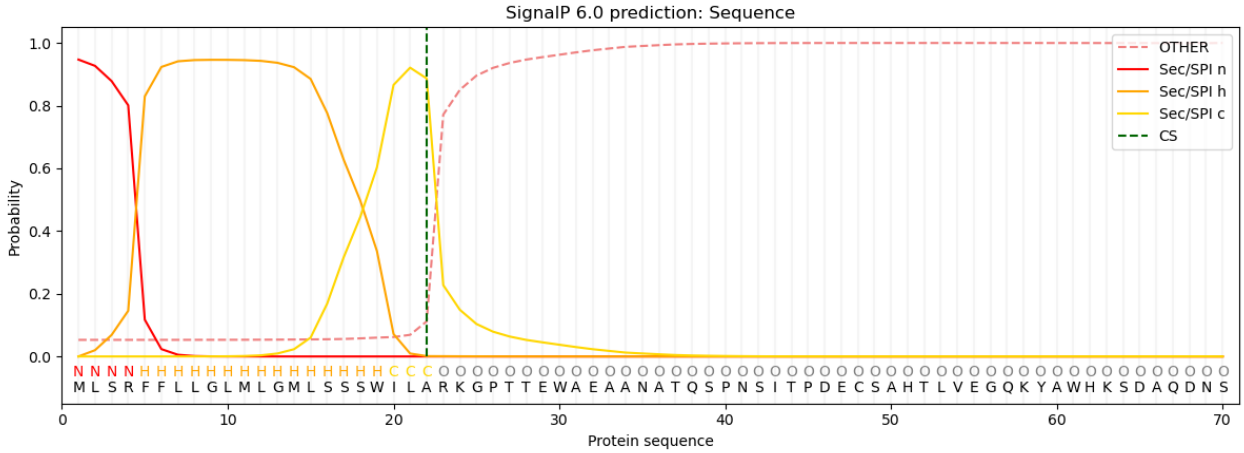


Figure 2 | SignalP analysis of AP1260. The first 22 residues of AP1260 are predicted to be a signal peptide, with a probability of 0.94. CS shows the cleavage site, where the signal peptide is cleaved. Sec refers to the protein sequence targeting the secretory pathway. Sec/SPI shows the subregions of the signal peptide sequence targeted for cleavage by signal peptidase I (SPI). ‘n,’ ‘h,’ and ‘c’ refer to the N-terminus, hydrophobic region, and C-terminus of the signal peptide respectively.

BLASTP

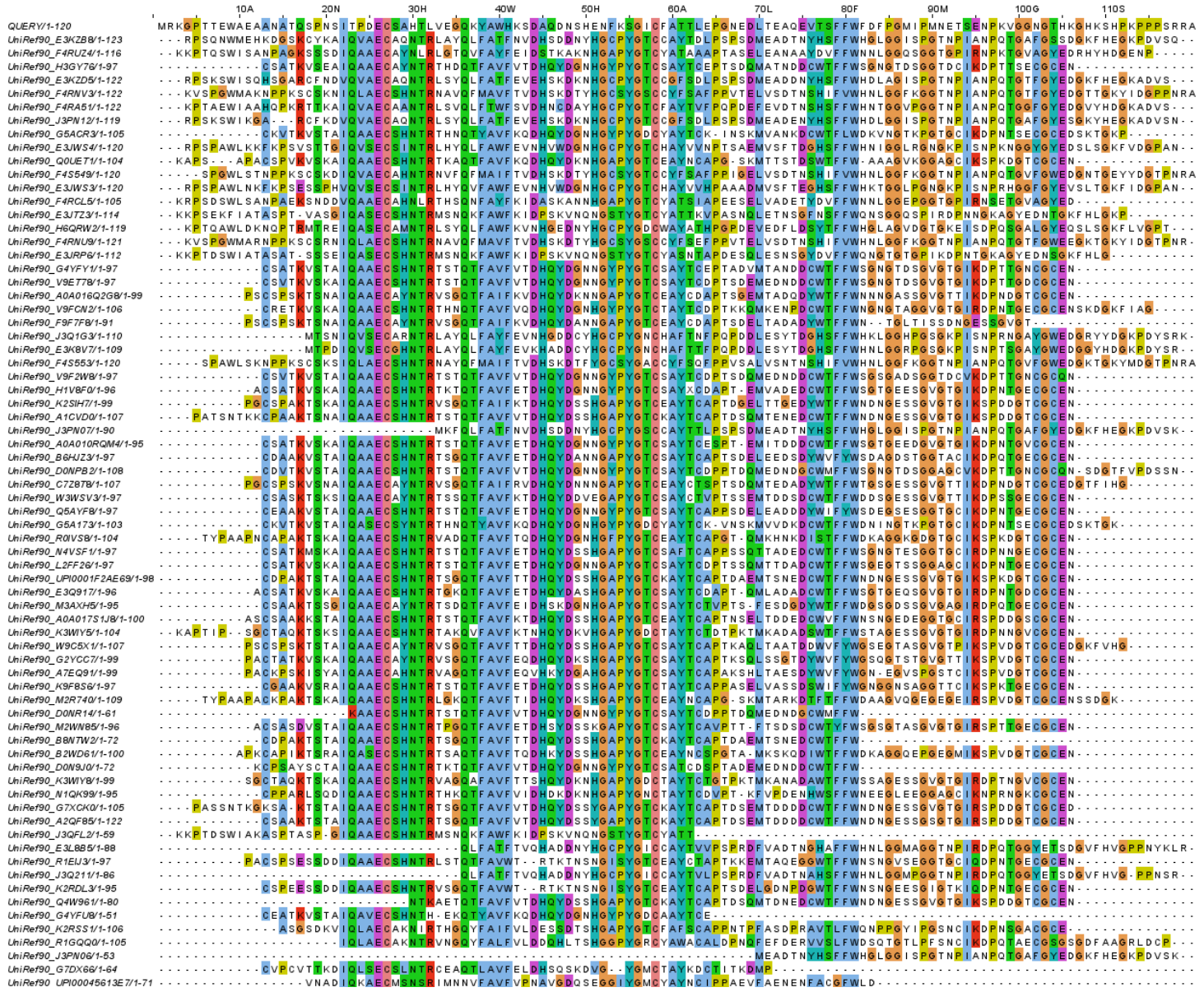
BLASTP was used to search for protein sequences similar to AP1260. BLASTP identity search results were derived from the non-redundant protein sequences database (SI Figure 2). The non-redundant protein sequence database used all non-redundant protein translations from multiple databases and compared them to the AP1260 protein sequence. The identity matches of this search produced several matches however none were higher in percentage identity than 45%. All matches produced were proteins from species in the order Pucciniales, of which *A. psidii* is a member.

JPred4

JPred4 was used to predict the secondary structure of AP1260. JPred4 also uses the PSI-BLAST algorithm to generate a position-specific scoring matrix to further search the protein database and detect distant relationships between proteins (Bhagwat & Aravind, 2007). Secondary structure prediction by JPred4 (SI Figure 3) predicts two α helices and three β strands. Helices are predicted to be from Thr7 to Glu11 and Pro24 to Leu32. Strands are predicted between Lys37 to Ser43, Cys58 to Leu63, and Val76 to Phe83.

JPred4 alignment of AP1260 using the UniRef90 database (Figure 3) provided stronger results than those gained from the previous BLASTP search. There are several conserved residues across related proteins, however, very few residues involved in the predicted secondary structure are conserved between the aligned proteins. Cysteine residues at positions 27 and 58 are highly conserved across all aligned proteins. Two highly conserved cysteines are a strong indicator for a potential disulfide bond within the structure of AP1260.

Figure 3 | JPred4 sequence alignment. Alignment of identity match results from the UniRef90 database. Sequence alignment shows several residues conserved across species. Notably, cysteine residues at positions 27 and 58 are conserved across nearly all matches – a strong indicator of a disulfide bond.



Disorder prediction

DISOPRED3 and ODiNPred were used to predict disordered regions in AP1260. DISOPRED3 predicts the C-terminus of AP1260 to be disordered with a probability of ~0.9 from residue 90 (Figure 4A). ODiNPred predicts the C-terminus of AP1260 to be disordered with a probability of 1.0 from residue 90 (Figure 4B). Both predictions predict a disordered region that begins around residue 90 and ends at the C-terminus.

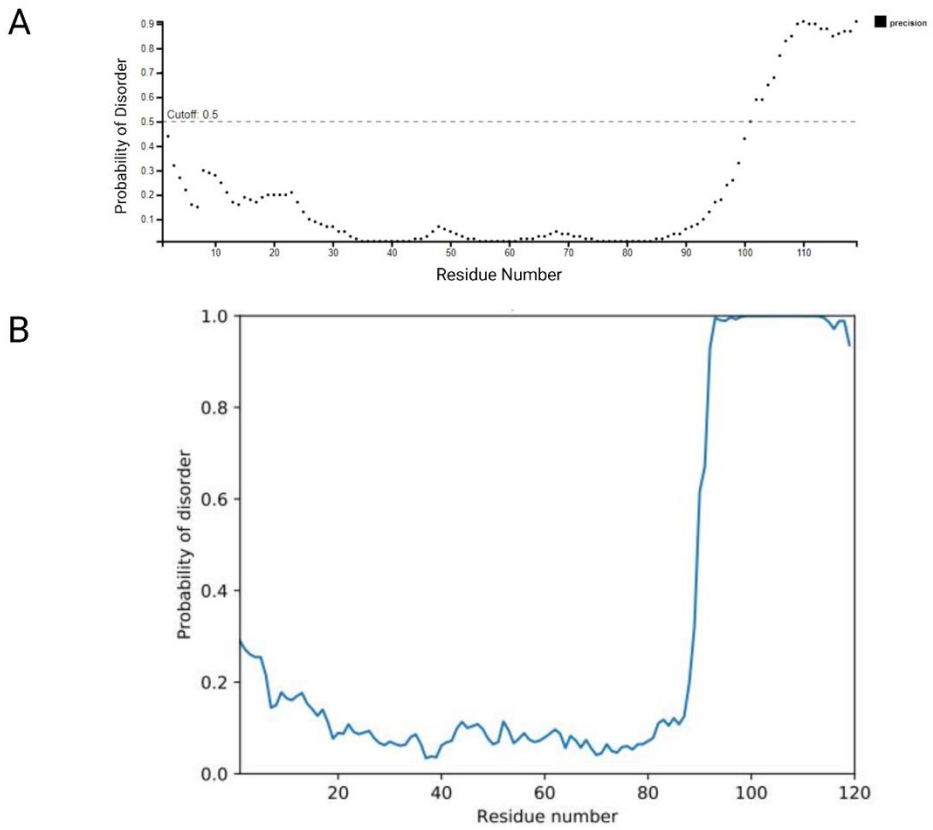


Figure 4 | Disorder prediction. A) DISOPRED3 disorder prediction of AP1260 predicts disorder to begin around residue 90 with a 0.9 probability. B) ODiNPred disorder prediction of AP1260 predicts disorder to begin around residue 90 with a probability of 1.0.

WoLF PSORT

WoLF PSORT predicts the subcellular localization of AP1260 with the signal peptide in fungi to be the golgi, the endoplasmic reticulum, or for the protein to be secreted (SI Figure 4). Without the signal peptide, WoLF PSORT predicts the subcellular localization of AP1260 in fungi to be in the cytoplasm (SI Figure 5). When searching for subcellular localization in plants, with the signal peptide, WoLF PSORT predicts AP1260 to localize in the vacuoles or the

nucleus (SI Figure 6). Without the signal peptide, the plant subcellular localization of AP1260 was predicted to be at the nucleus (SI Figure 7). The results from WoLF PSORT were low in percentage identity matches, however, with the maximum identity percentage being very low at 15.5%.

ProtParam

ProtParam analysis of the AP1260 sequence provided results on the calculated molecular weight, extinction coefficient, amino acid composition, isoelectric point, and the instability index. ProtParam predicts the molecular weight of AP1260 to be 14,035 Da, excluding the T7 tag and His tag, and the isoelectric point to be pH 6.15. AP1260's extinction coefficient is calculated to be 18115 M^{-1} cm^{-1} and its instability index is computed to be 36.38, classifying AP1260 as stable.

AlphaFold

The AlphaFold predicted structural model of AP1260 (Figure 5) shows a short α-helix on the N-terminus followed by a long coil and another α-helix. An anti-parallel β-sheet runs through the centre of the folded region. This β-sheet appears to curve around the second α-helix with a predicted disulfide bond between the α-helix and one of the β-strands by Cys27 and Cys58. The globular region of the protein contains many long coils or loops linking the predicted secondary structure elements. The C-terminus has no visible predicted structure with this 'tail' beginning at the end of the β-sheet at residue 82. Notably, the C-terminal tail has high proline content with seven prolines in the tail region. A charge overview of AP1260 shows the C-terminal tail to have a relatively strong positive character and the globular section of the protein to have both positive and negative character on the surface. Of note is a cleft with a strong negative character on the upper surface of AP1260.

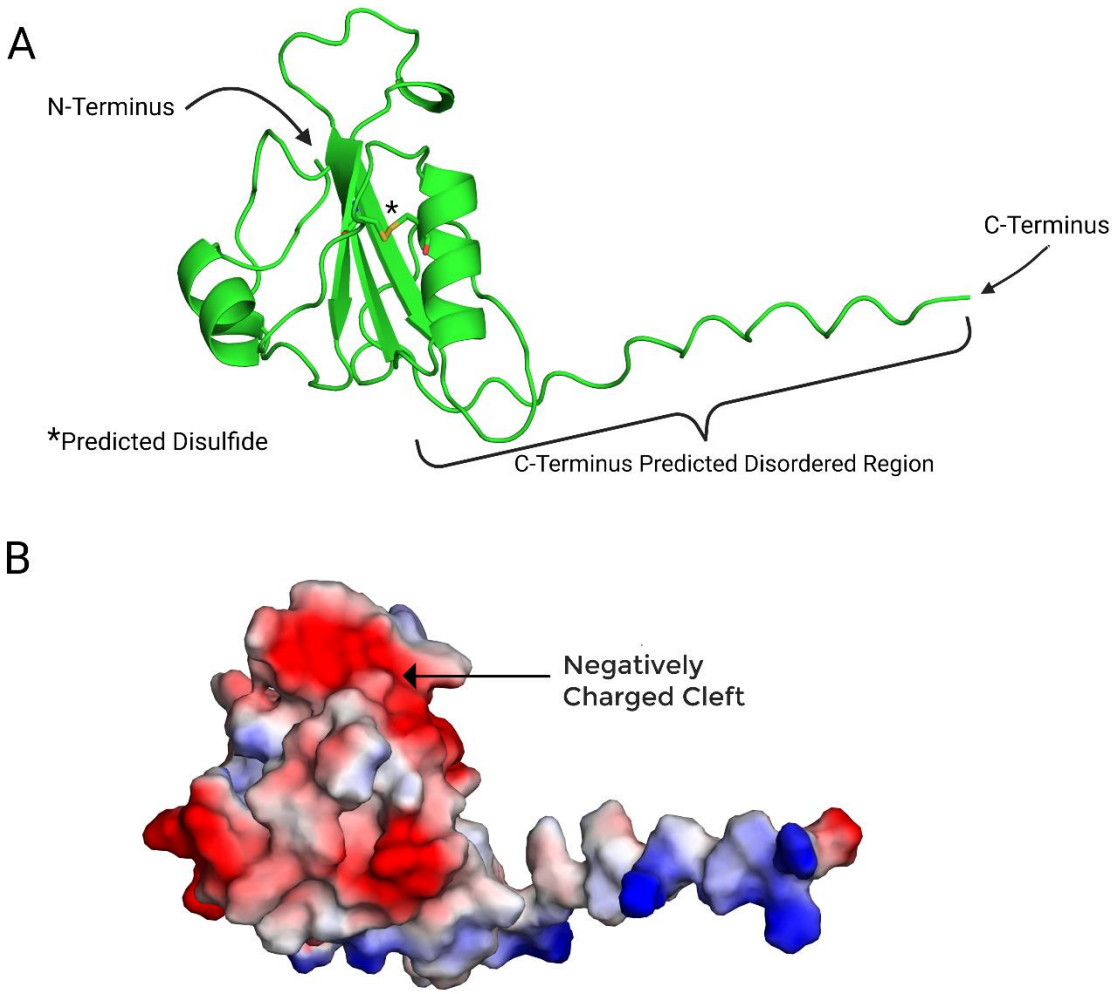


Figure 5 | AlphaFold predicted structure. A) Predicted structure of AP1260 shows a predicted disulfide between Cys27 and Cys58. Helices are predicted at the N-terminus of AP1260 and an antiparallel sheet in the centre of the folded region. The C-terminal end of AP1260 shows no structure. B) Charge overview of AP1260 showing positive character in the tail end and negative character in the folded region.

Biophysical Analysis

Transformation, expression, and purification

To experimentally characterize the biophysical properties of AP1260, it must first be obtained in milligram quantities and purified from other undesired proteins. To obtain AP1260, AP1260 N-terminus, and the AP1260 truncated variants, *E. coli* cells were transformed with plasmids containing kanamycin and ampicillin resistance for selection and a gene encoding an AP1260 variant for overexpression. The transformation of AP1260, AP1260 N-terminus, and the truncated variants was successful, as determined by the presence of 20-30 individual colonies

on LB + kanamycin or ampicillin plates. No colonies had grown on LB + kanamycin or ampicillin plates incubated in the absence of cells, indicating no contamination had occurred.

Protein overexpression was induced with IPTG to produce large quantities of AP1260 that was subsequently purified to homogeneity. Diluting the AP1260 sample into low salt after immobilized metal affinity chromatography caused visible aggregation. To resolve this, 30 mM NaCl was added to the solution containing AP1260. Any aggregation present was removable through centrifugation. After the addition of 30 mM NaCl, the AP1260 solution remained stable. Anion exchange chromatography was a new purification step from previous purifications of AP1260. Based on previous purification, a cysteine-linked dimer was occurring, indicating some misfolding as AP1260 has only two cysteines that form a predicted internal disulfide. Use of anion exchange chromatography led to two separated peaks (Figure 6). An SDS-PAGE gel across the two peaks was carried out to determine the separation seen on the trace (Figure 6). The SDS-PAGE gel shows the first peak to contain mainly AP1260 with few other contaminants indicated by the large bands at ~15 kDa. The second peak was seen to contain more of a contaminant at ~10 kDa and the cysteine-linked dimer at ~30 kDa. The addition of anion exchange to the purification process successfully separated the contaminant and cysteine-linked dimer seen in previous purifications from AP1260. Size exclusion chromatography following anion exchange demonstrated a single peak (Figure 6). SDS-PAGE of the size exclusion peak indicates a single species is present (Figure 6). Anion exchange added to the purification process of AP1260 led to a homogenous sample that could be used for biophysical analysis.

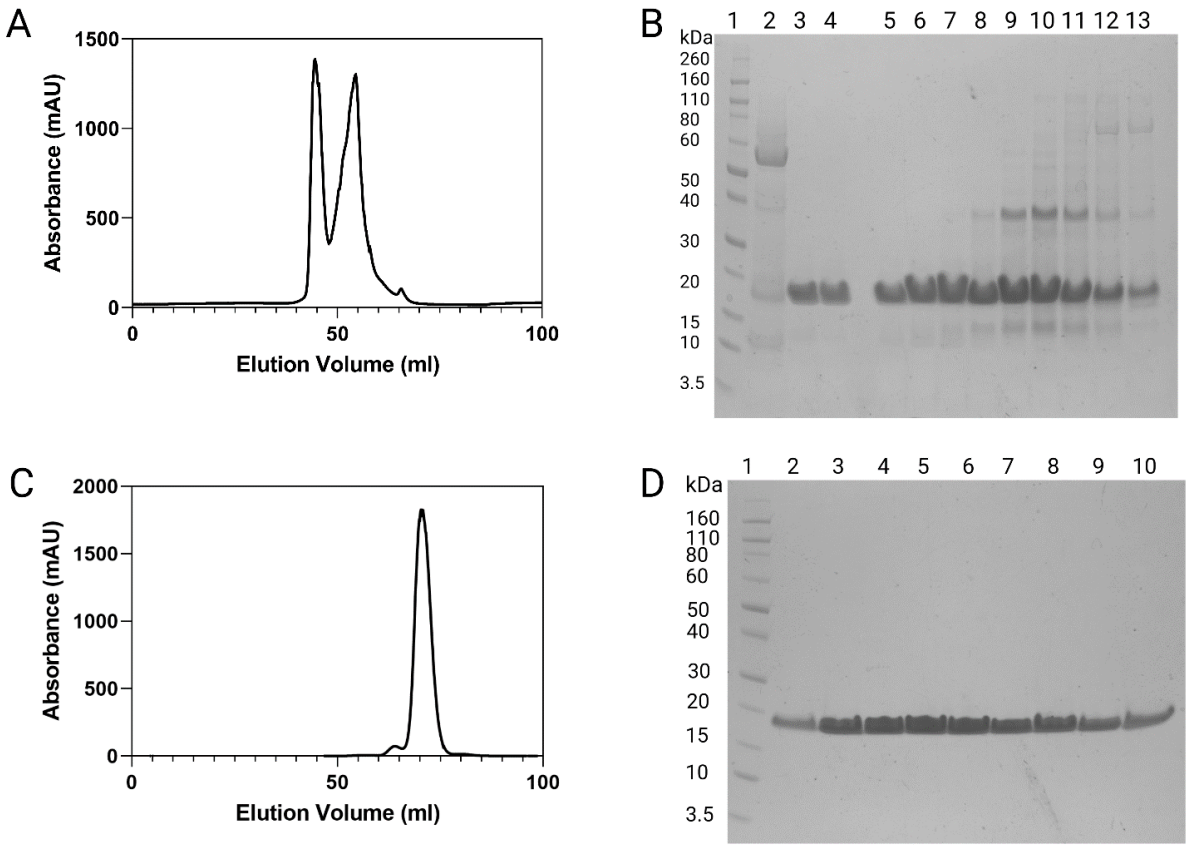


Figure 6 | AP1260 purification results after the addition of anion exchange. A) Anion exchange chromatography trace demonstrating two separated peaks. B) SDS-PAGE gel containing samples across the two anion exchange peaks. Lane 1: Novex® Sharp Pre-stained Protein Standard ladder (molecular weight corresponding to each band is labelled). Lane 2: 7% wash sample to test for contamination. Lanes 3-8: Anion exchange peak one from fraction later pooled for size exclusion chromatography. Lanes 9-13: Anion exchange peak two containing ~10 kDa contaminant and cysteine-linked dimer. C) Size exclusion chromatography trace showing a single species indicating sample homogeneity after purification. D) SDS-PAGE gel showing successful purification after the four-step purification process. Lane 1: Novex® Sharp Pre-stained Protein Standard ladder (molecular weight corresponding to each band is labelled). Lane 2-10: Purified protein samples from size exclusion following anion exchange.

The expression trial of the AP1260 truncated mutants E75, W81, F82, D83, F84, G86, M90, E92, P97, G103, and N-terminus His-tag was conducted for the purposes of crystallography after expression and purification. For cell growth, 20 °C was better suited than 37 °C with cell mass being higher in the mutants grown at 20 °C. Of the truncated mutants, only G103 had suitable expression to be able to purify in milligram quantities. The N-terminal His-tag also was able to be expressed to a sufficient level. SDS-PAGE of cell lysate from cultures that have been induced to overexpress the mutants shows the level of expression of each (Figure X). The N-terminal His-tag construct showed better expression than all the truncated mutants as

indicated by the band at ~15 kDa. Of the mutants, only G103 showed expression as seen by the band at ~12 kDa.

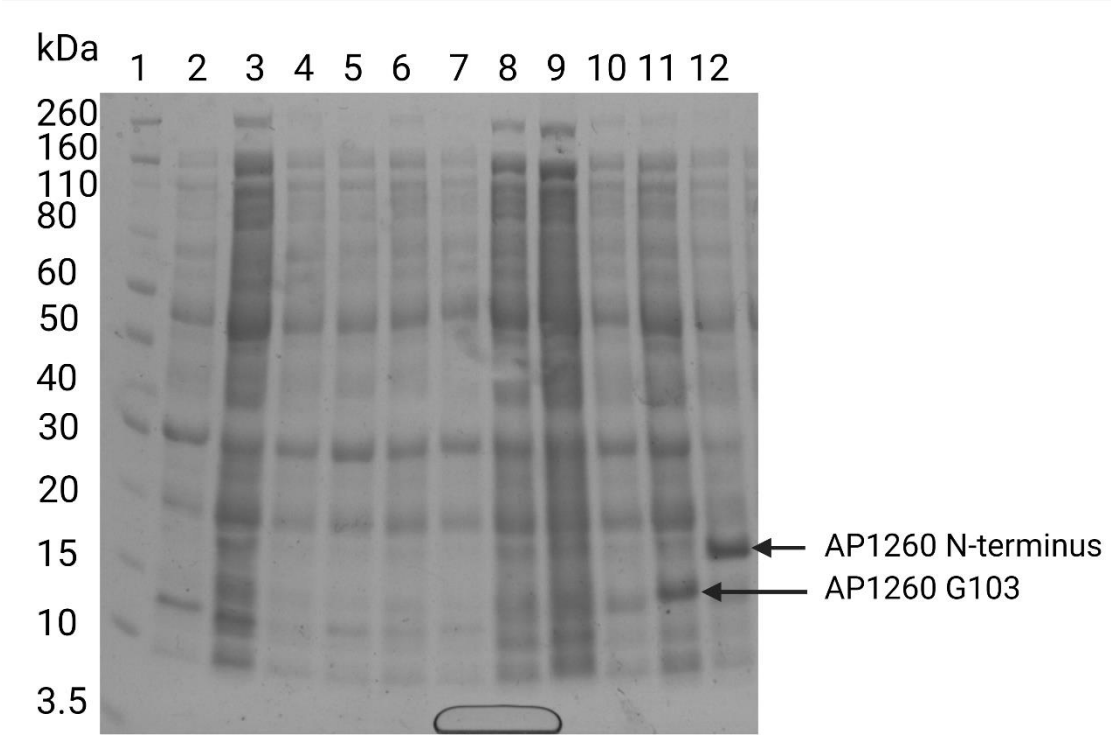


Figure 7 | SDS-PAGE gel of expression trial of AP1260 truncated mutants and N-terminal His-tag at 20 °C. Lane 1: Novex® Sharp Pre-stained Protein Standard ladder (molecular weight corresponding to each band is labelled). Lane 2: Truncated mutant E75. Lane 3: Truncated mutant W81. Lane 4: Truncated mutant F82. Lane 5: Truncated mutant D83. Lane 6: Truncated mutant F84. Lane 7: Truncated mutant G86. Lane 8: Truncated mutant M90. Lane 9: Truncated mutant E92. Lane 10: Truncated mutant P97. Lane 11: Truncated mutant G103. Lane 12: N-terminal His-tag.

Immobilized metal affinity chromatography of G103 and the N-terminal His-tag was done to elucidate their expression. The immobilized metal affinity chromatography trace can be seen in Figure 8. G103 can be seen to show about 20% less signal than N-term. Immobilized metal affinity chromatography demonstrated the successful expression of G103 and the N-terminal His-tag constructs. This step also elucidated that the T7 tag, initially attached to aid in the expression of the original AP1260 construct, is not required to successfully express AP1260.

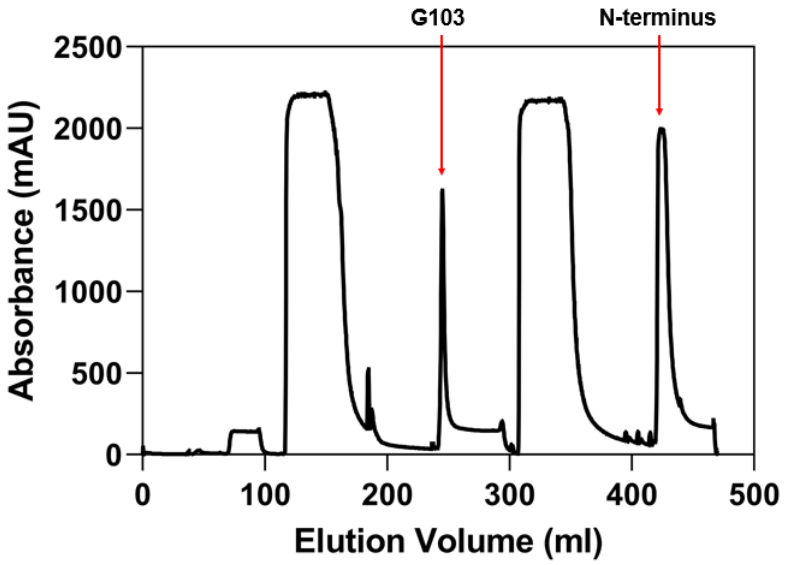


Figure 8 | Immobilized metal affinity chromatography of AP1260 G103 and AP1260 N-terminal His-tag. G103 and N-terminal His-tag were put through an immobilized metal affinity chromatography column back-to-back. Elution beginning at ~250 ml belongs to G103. Elution beginning at ~420 ml belongs to N-terminal His-tag. The signal from the N-terminal construct is ~20% greater than that of G103.

Analytical ultracentrifugation

Analytical ultracentrifugation using a sedimentation velocity experiment was used to determine the molecular weight and oligomerization state of AP1260 in solution. The raw sedimentation data and the residuals for this fit are shown in Figure 9A. The residual plot has no observable patterns indicating the data is well fit to the curve. The results of this experiment demonstrated that AP1260 is monomeric in solution. An S value of 1.67 (Figure 9B) was obtained for AP1260 and a solution mass of 16.1 kDa (Figure 9C). The resulting data and residual plots from the sedimentation velocity experiment, c(*s*) plot and c(*M*) plot are shown in Figure 9. The experiment determined that the AP1260 sample was monodisperse, indicating sample homogeneity with the peak signal being 95% or greater. The frictional ratio (f/f_0) of AP1260 is 1.43 suggesting the shape of AP1260 is slightly asymmetrical.

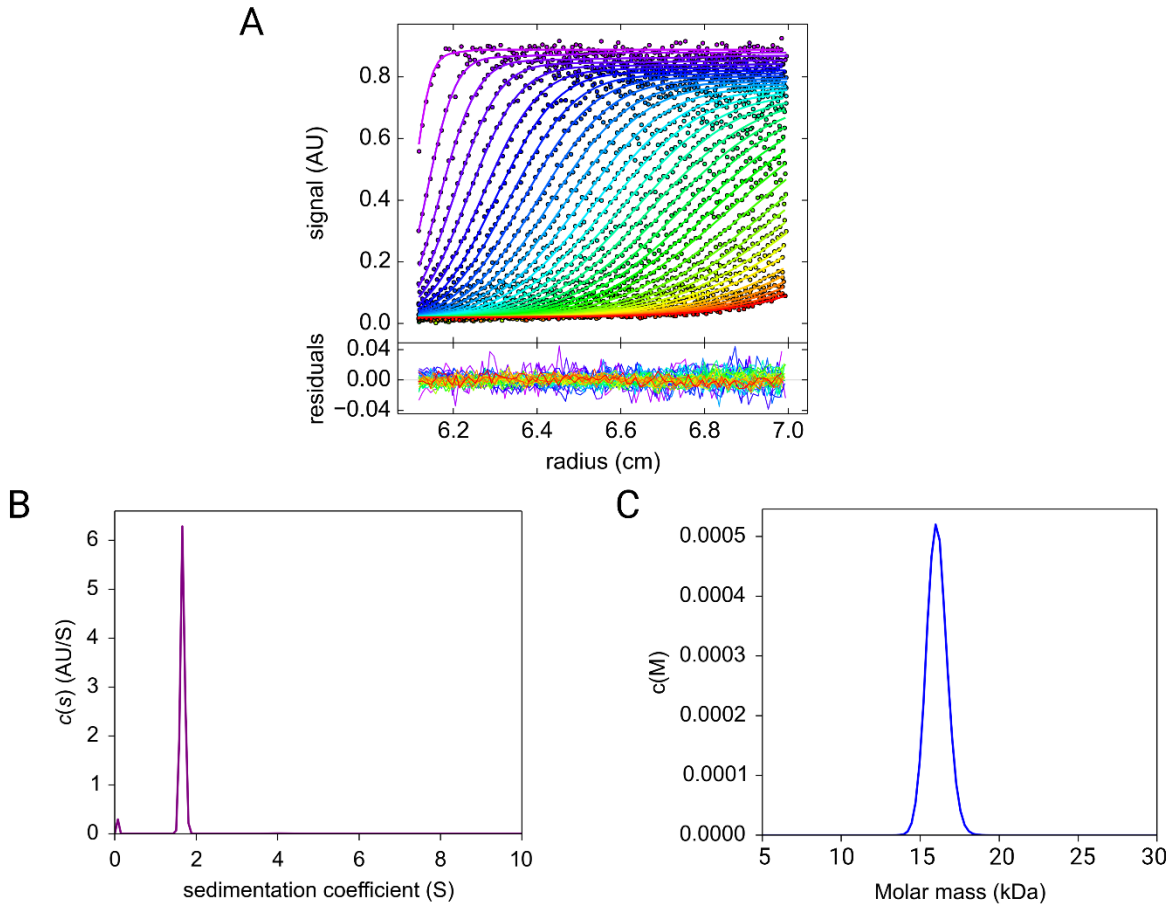


Figure 9 | Analytical ultracentrifugation sedimentation velocity analysis data and residuals, c(*s*) plot and c(*M*) plot. A) Absorbance at 280 nm plotted as a function of radial position (cm). The raw data are represented as open symbols (o) and overlaid with the non-linear least-squares best fit. The residuals for this fit are shown below the plot. B) The c(*s*) model is plotted as a function of the sedimentation coefficient. Data were fitted with an *s* value ranging between 0 and 10 S at a resolution of 300 and a confidence level of 0.95 using SEDFIT. The fit resulted in a frictional ratio (f/f_0) of 1.43. The c(*s*) distribution gave a sedimentation coefficient of 1.67 S. C) The c(*M*) model is plotted as a function of molecular mass (kDa). Data were fitted with a minimum mass of 1 kDa and a maximum mass of 50 kDa at a resolution of 300 and a confidence level of 0.95 using SEDFIT. Data from 5 kDa to 30 kDa are shown. The c(*M*) distribution gave an apparent molecular weight of 16.1 kDa, compared to 14 kDa predicted from the sequence.

Circular dichroism

Circular dichroism spectroscopy was used to determine the secondary structure of AP1260 in solution. Data used for analysis were taken from 195 nm to 250 nm, where the HT [V] remained below 600 V to reduce sample noise. The results show that AP1260 is partially unfolded in solution (Figure 10). This resulting data from the circular dichroism experiment was analysed using BeStSel, providing information on the secondary structure of AP1260. The

BeStSel results estimate the secondary structure content of AP1260 to be 13.1% α-helices, 18.1% anti-parallel β-sheet, 2.0% parallel β-sheet 16.9% turns, and 49.8% other or unfolded. The RMSD of the fitted curve from BeStSel was 0.0782, indicating the data is well fitted.

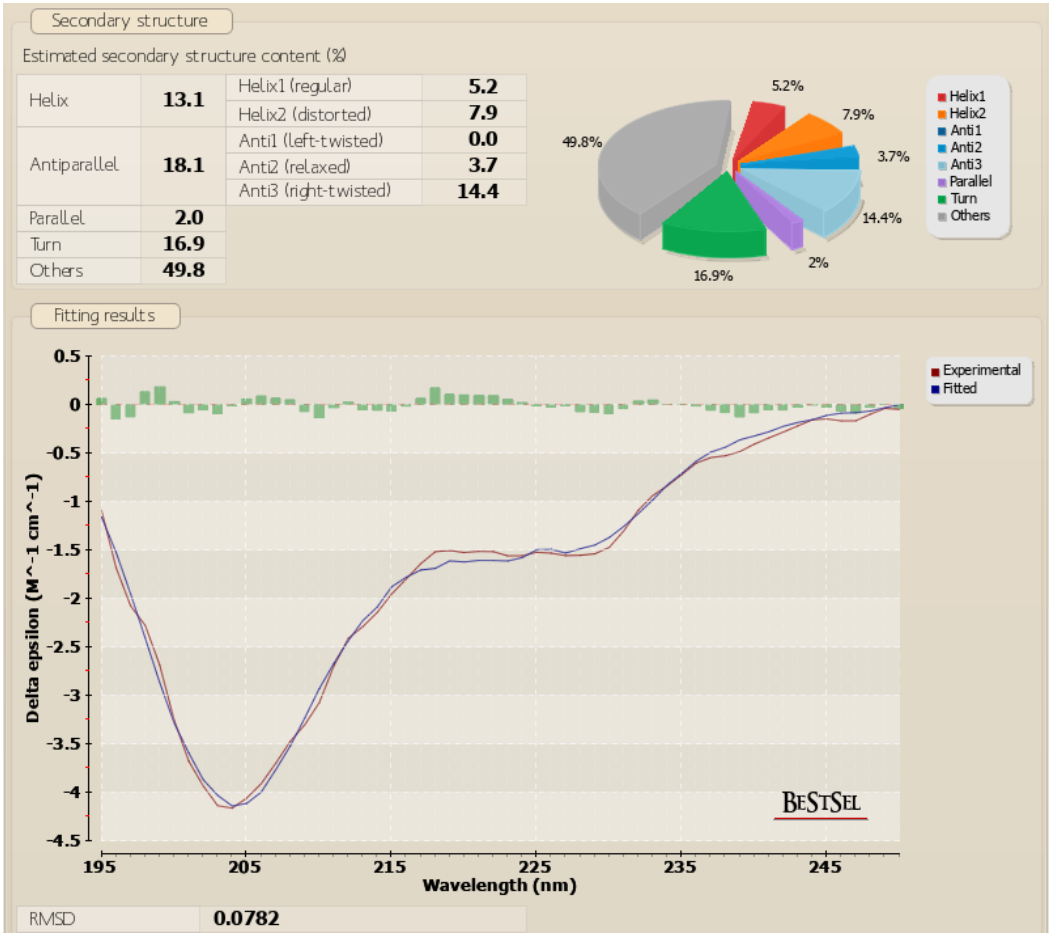


Figure 10 | Circular dichroism spectrum of AP1260. Data were collected at a concentration of 0.1 mg/ml in 20 mM NaH_2PO_4 at pH 7.4. Delta epsilon (M^{-1} cm^{-1}) is plotted as a function of wavelength. Data were taken from 195 nm to 250 nm, where the HT [V] remained below 600 V. BeStSel results of the AP1260 circular dichroism data, estimating secondary structure to be 13.1% α-helices, 18.1% anti-parallel β-sheet, 2.0% parallel β-sheet 16.9% turns, and 49.8% other. The RMSD of the fitted curve from BeStSel was 0.0782.

Differential scanning fluorimetry

Differential scanning fluorimetry was used to determine the thermal stability of AP1260. The first derivative of the melt curve as a function of temperature is presented in Figure 11. AP1260 was measured in triplicate with the results providing a mean melting temperature of 48.1°C. The lower 95% confidence was 45.1°C and the upper 95% confidence was 51.2°C. This significant range in confidence suggests high noise and unusual signal. This may indicate that AP1260 is relatively incompatible with differential scanning fluorimetry. The results gained

may be inaccurate due to this yet still provide some indication of the melting temperature of AP1260.

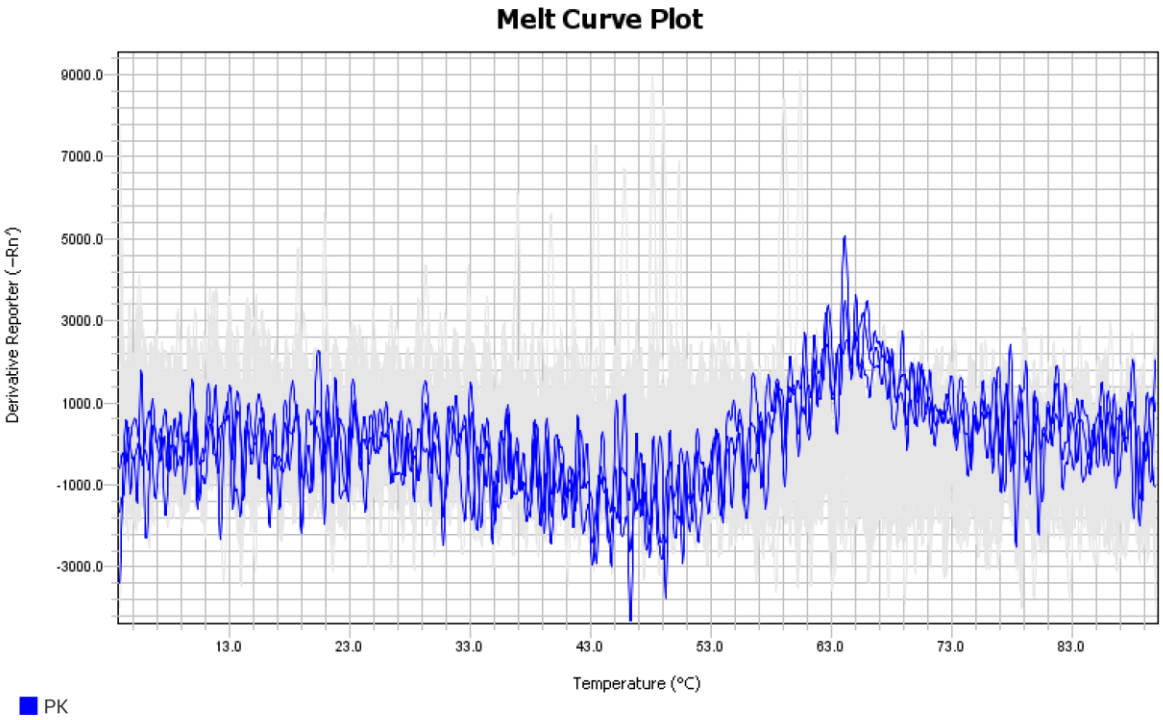


Figure 11 | Differential scanning fluorimetry melt curve. The first derivative of the melt curve is plotted as a function of temperature. Each sample was in 50 mM Tris, 150 mM NaCl, pH 8.0 and was measured in triplicate.

Fluorescence spectroscopy

Fluorescence spectroscopy was used to elucidate the thermal stability of AP1260. Data from emission spectra of AP1260 at various temperatures increasing at 5°C increments from 20 °C to 50°C and immediately back down to 20°C were collected (Figure 12A). These data demonstrate a melting of AP1260 at ~40 °C. This is indicated by increasing temperature causing the emission curves to shift to the right and increase in fluorescence due to buried tryptophan residues becoming open to the solvent. Cooling down to 20 °C initially returns the curve close to the baseline. After several minutes to allow the whole system to return to 20 °C, the emission spectrum near perfectly overlaps with the initial 20 °C spectrum. Based on this overlap of spectra, AP1260 likely refolds after melting, indicated by the reduction of fluorescence back to baseline.

The ratio of 350/332 nm as a function of temperature is shown in Figure 12B. The ratio of 350/332 nm is determined from emission spectra measured from 300 nm to 420 nm at increasing temperatures. Temperatures increasing in 5°C increments from 20 °C to 80 °C demonstrates a melting temperature of AP1260 to be ~40°C.

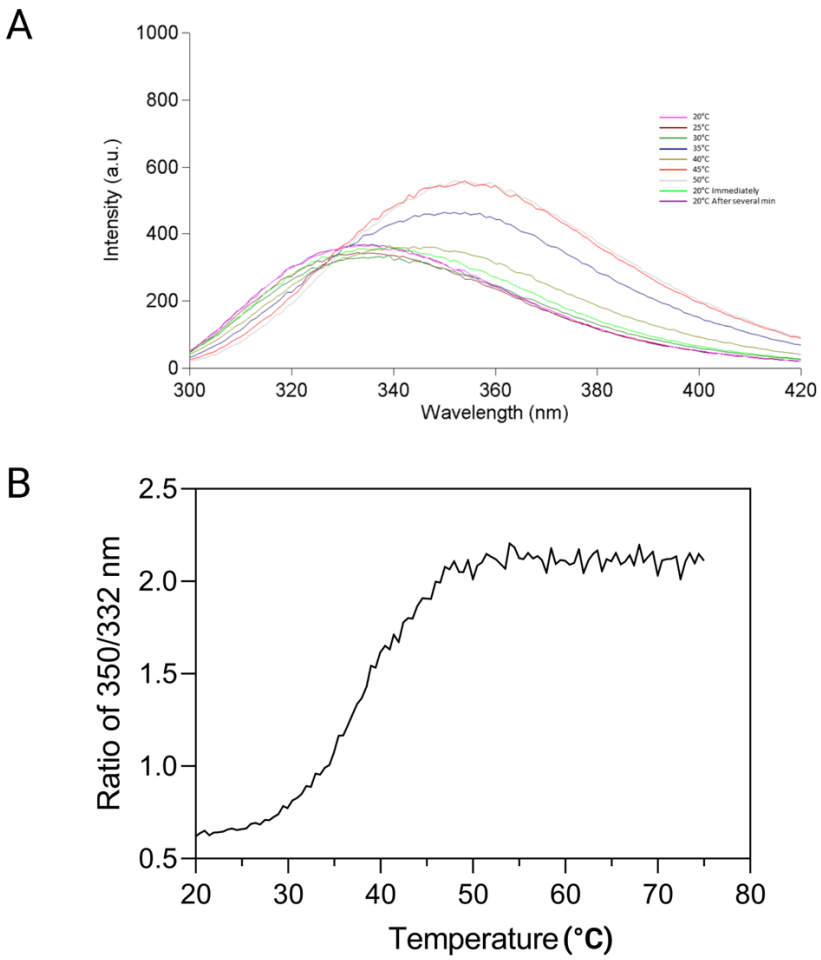


Figure 12 | Fluorescence melt test of AP1260. A) Fluorescence intensity is plotted as a function of wavelength. Temperature was increased from 20 °C to 50 °C and immediately back to 20 °C. Scans indicate melting at ~40 °C as well as refolding upon reduction of temperature. B) The ratio of 350/322 nm data is plotted as a function of temperature. These data indicate the melting temperature of AP1260 to be ~40 °C.

Small angle x-ray scattering

Small angle x-ray scattering was used to obtain low-resolution structural data on AP1260—the first structural data for a myrtle rust effector protein. Of the samples used during small angle x-ray scattering, the samples at a concentration of 10 mg/ml had more signal than those at 5 mg/ml. Of the two data collected using 10 mg/ml samples, the second run had a better

fit at low *q*, therefore analysis was done using these data. Scan 317 of this run was an outlier that could not be easily removed. The scatter was therefore plotted to scan 316, at around 0.27 *q* (Figure 13A). The Guinier plot (Figure 13B) shows the data to be linear at low *q*, indicating that the data is reliable (free of aggregate and/or particle interference). When analysed in a Kratky plot (Figure 13C), a measure of flexibility, the trend indicates that AP1260 is neither fully folded nor fully disordered. The trend directs to AP1260 being partially flexible as the data does not reach zero at high *q*. In the P(r) plot (Figure 13D), the data is randomly distributed and there are no patterns or biases, therefore the data fits exceptionally. Rg across the peak does not change significantly indicating monodispersity with the Rg value from the P(r) and Guinier plots both being 21.1 Å. The P(r) plot also indicates that the particles are elongated or somewhat rod-shaped due to them being skewed to the right. This may also be an indicator of disorder. The Dmax of the particle is also seen to be around 90 Å as determined from the P(r) plot. Small angle x-ray scattering analysis also estimated the molecular weight of AP1260 to be somewhere in the range of 14.5 to 17 kDa.

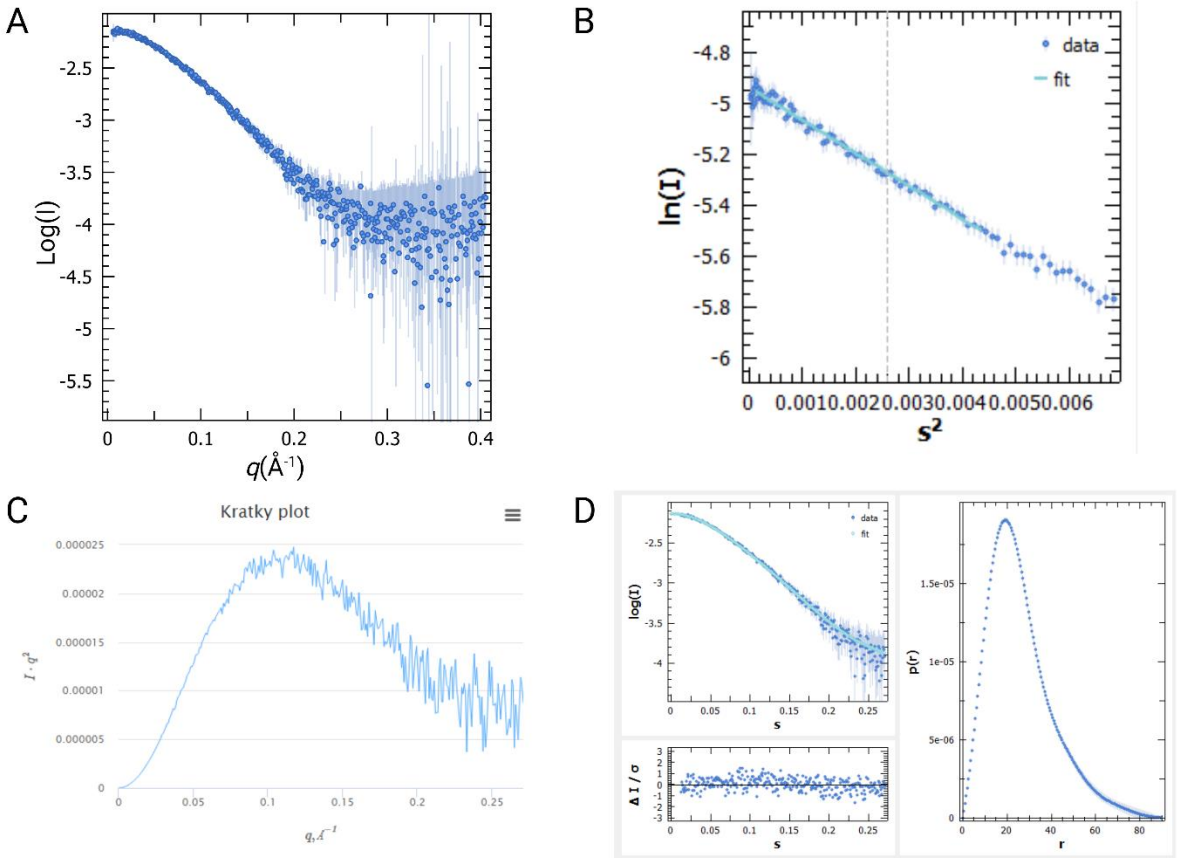


Figure 13 | Small angle x-ray scattering of AP1260. A) Small angle x-ray scatter of AP1260 at 10 mg/ml in 50 mM Tris, 150 mM NaCl at pH 8.0. B) Small angle x-ray scattering Guinier plot

of AP1260 demonstrates data is linear at low *q*. C) Small angle x-ray scattering Kratky plot. Indicates AP1260 is neither fully folded nor fully disordered. The trend indicates AP1260 is partially flexible. D) Small angle x-ray scattering P(r) plot and corresponding fit data. P(r) plot is right-skewed and indicates a Dmax of 90 Å.

X-ray crystallography

X-ray crystallography was attempted on AP1260 to obtain a high-resolution structure. AP1260 crystallography trails using JCSG+, PACT Premier, and SG1 crystallography screens with protein concentrated to 9 mg/ml had no observed crystal growth. Furthermore, AP1260 concentrated to 22 mg/ml using JCSG+ and PACT Premier crystallography screens also has no observed crystal growth. Aggregation was observed in each crystallography screen at both 9 mg/ml and 22 mg/ml. However, not all conditions in the screens contained aggregation. Clear conditions still had no observed crystal growth. N-Terminus AP1260 crystallography trials using protein concentrated to 22 mg/ml and JCSG+, PACT Premier, and SG1 crystallography screens observed no crystal growth. Not all crystallography conditions had observed aggregation yet no crystal growth occurred.

Discussion

A combination of bioinformatic analyses was used to predict the structure, potential disorder, and subcellular localization of AP1260. Biophysical analysis methods including, analytical ultracentrifugation, circular dichroism, differential scanning fluorimetry, fluorescence spectroscopy, and small angle x-ray scattering were then applied to obtain information on the mass, shape, stability, and dynamics of AP1260.

As the structure, function, and binding partners of AP1260 are still under investigation, the results of this research provide a beginning toward the development of methods that target AP1260. Due to AP1260's early expression during myrtle rust infection, these methods will inform a means through which myrtle rust can be controlled.

AP1260 is a novel effector protein

Bioinformatic analysis using the AP1260 sequence was used to predict several of AP1260's characteristics and compare it to related proteins submitted to numerous databases. BLASTP was to search for submitted protein sequences related to AP1260. The results from this search, as presented in supplementary Figure 2, showed several matches, however, the highest match provided an identity score of 45%. All matches were from other rust fungi, none of which were published, preventing any insight into the potential function of AP1260 that could have been gained. The lack of highly related matches show that no previous work has been done on AP1260 or any closely related *A. psidii* effectors. JPred4 alignment using the PSI-BLAST algorithm was used to detect distant relationships to AP1260. Alignment results are presented in Figure 3. From the alignment, many conserved residues can be observed through the sequence. Interestingly, the only two cysteine residues in the protein are highly conserved across nearly all matches. This indicates an internal disulfide within the protein.

EffectorP analysis of the AP1260 sequence provided a prediction that AP1260 is a cytoplasmic effector, presented in supplementary Figure 1. EffectorP predicted AP1260 to be a cytoplasmic effector. Cytoplasmic effectors of plant pathogens act within plant cells where they suppress or manipulate host processes (Asai & Shirasu, 2015; Wang et al., 2017). The localisation of AP1260 within plant cells was predicted using WoLF PSORT, presented in supplementary Figures 4-7. WoLF PSORT predicted the localisation of AP1260 within plant cells to be the nucleus and vacuoles. The predicted localisation site of AP1260 being the

nucleus and vacuoles may provide insight into the function of AP1260. Both the nucleus and vacuole are essential in plant cells so could be potential targets of AP1260. Bacterial effectors that target the nucleus have been observed to alter transcription in the host cells (Canonne & Rivas, 2012), a process through which they seek to increase their fitness. The vacuole is an essential organelle involved in compartmentalising things such as sugars, proteins, and secondary metabolites, making the vacuole play a crucial role in signalling in response to stress (Zhang et al., 2014). Due to the importance that processes in both the nucleus and vacuole contribute to plant defence, the prediction that these may be targets of AP1260 is rational.

AlphaFold was used to predict the structure of AP1260 to gain a basis from which structural information can be compared. The provided structure, presented in Figure 5, has a globular region at the N-terminus containing some helices and an antiparallel sheet. The C-terminus of the structure from the end of the third β-strand is unfolded or disordered. There is also a predicted internal disulfide between the two cysteine residues as might be thought from the JPred4 alignment. Secondary structure prediction by JPred4 (SI Figure 3) predicted secondary structure elements seen in the AlphaFold structure. The helices and strands predicted by JPred4 are seen between the same residues in the AlphaFold structure, the two predictions corroborating on predicted secondary structure elements within the protein. The AlphaFold predicted structure showed moderate confidence in the folded regions and low confidence in the loops and the C-terminal tail. The C-terminal tail of AP1260 being seemingly disordered in the predicted structure is supported by disorder prediction which predict with a probability of ~1.0 for AP1260 to be disordered from residue 90, located at the C-terminus. This can be seen in the disorder predictions presented in Figure 4. Although the AlphaFold structure provides some structural information, the low confidence in the prediction requires further biophysical analysis to get high confidence and high-resolution structural information.

As there were no experimentally determined characteristics for AP1260, bioinformatic analysis of the sequence in this research provides some basic understanding of the protein.

Novel AP1260 biophysical properties have been determined

Biophysical analysis of AP1260 uncovered many of its characteristics. Biophysical methods were used to determine the mass, oligomeric state, thermal stability, shape and dynamics,

and the secondary structure of AP1260. Understanding the biophysical characteristics of AP1260 will deepen the understanding of *A. psidii*'s mechanism of infection.

Analytical Ultracentrifugation was used to determine the mass and oligomeric state of AP1260. Results from the sedimentation velocity experiment are presented in Figure 9. The c(*s*) distribution gave a sedimentation coefficient of 1.67 S. From this it was determined that AP1260 exists as a monomer in solution. The c(*M*) distribution supports the c(*s*) distributions findings. An apparent molecular weight of 16.1 kDa was obtained, as would be expected from the mass obtained from ProtParam which excluded the T7 and histidine tags. With the apparent mass being that of one AP1260 protein, the sedimentation coefficient of 1.67 S relates to monomeric AP1260. The frictional ratio (f/f_0) of AP1260 was determined to be 1.43, indicating AP1260 is slightly asymmetrical in solution as might have been thought from the predicted structure.

Differential scanning fluorimetry was used to obtain results on the thermal stability of AP1260 as presented in Figure 11. A melting temperature of 48.1°C was derived. On account of the significant range in confidence, high noise, and unusual signal, fluorescence spectroscopy was used to elucidate AP1260's thermal stability as presented in Figure 12. The results from intrinsic fluorescence melt tests provided a melting temperature of ~40°C, similar to the melting temperature provided from differential scanning fluorimetry. A melting temperature of 40°C is curious because of *A. psidii* geographic range, inhabiting mainly semi-tropical regions where the ambient temperature can sometimes reach 40°C. Further results from fluorescence spectroscopy revealed that when the temperature is returned from 50°C to 20°C, after a few minutes, the fluorescence nearly perfectly overlaps the baseline indicating AP1260 refolds after being denatured through heat. Denaturing and refolding at temperatures observed in *A. psidii*'s host range, may be implicated in the function of AP1260, a mechanism which requires further research.

Small angle x-ray scattering was used to characterise AP1260's shape and dynamics in solution. The results obtained from small angle x-ray scattering are presented in Figure 13. A mass in the range of 14-17 kDa was obtained from the results which also indicated monodispersity. These data support the mass obtained from analytical centrifugation and ProtParam and indicated that only a single species was present in solution indicating AP1260 to be stable. The Kratky plot obtained from the results indicated that AP1260 is not fully

folded and not fully disordered. The trend in the Kratky plot indicates that AP1260 is partially folded, supporting the predicted structure from AlphaFold with the globular region and disordered C-terminus. Data from the P(r) plot indicated AP1260 is elongated or somewhat rod-shaped, or indicating disorder within the protein. This is supported by the frictional ratio obtained from analytical centrifugation indicating asymmetry, and the predicted structure of AP1260 with a disordered region.

Circular dichroism spectroscopy was used to determine the secondary structure content of AP1260. The circular dichroism results are presented in Figure 10. The results provided the secondary structure content of AP1260, determining that 13.1% of the protein is α-helices, 20.1% β-sheet, 16.9% turns, and 49.8% random coils, loops, or unfolded. These results support the JPred4 secondary structure prediction with both helices and sheets being present within the content of the secondary structure. The JPred4 alignment predicts that the majority of the AP1260 sequence does not have secondary structure which would be denoted as 'other' within the circular dichroism content. Circular dichroism data also supports the AlphaFold predicted structure with both helices and sheets being present in the secondary structure content of AP1260.

AP1260 exhibits an intrinsically disordered region

Through both bioinformatic and biophysical analysis, AP1260 has been determined to be flexible, partially folded, or partially disordered. Bioinformatics using AlphaFold provided a predicted structure with a globular region at the N-terminus and an unfolded tail at the C-terminus. Supported by disorder prediction, the C-terminus of the protein is predicted to be disordered. These predictions of AP1260's structure lend credence to the C-terminal tail of AP1260 being an intrinsically disordered region. Intrinsically disordered regions do not contain enough hydrophobic residues to facilitate cooperative folding. More often they contain a higher content of polar or charged residues. Due to their lack of tertiary structure, intrinsically disordered regions can fluctuate between many conformations in a dynamic equilibrium (Babu, 2016). AP1260's tail region has high proline and polar residue content as well as containing strong positive character as shown in Figure 5B. The bioinformatic analysis of AP1260 supports that AP1260 contains an intrinsically disordered region.

Biophysical analysis of AP1260 demonstrates an intrinsically disordered region may be present. The frictional ratio obtained from analytical ultracentrifugation results indicates AP1260 is slightly asymmetrical, not fully asymmetric, or globular. This is corroborated by the small angle x-ray scattering Kratky and P(r) plots. The Kratky plot indicates that AP1260 is partially folded, or partially unfolded, likely in the C-terminal tail region. The P(r) plot is skewed to the right, indicating an elongated or somewhat rod-shaped particle. This P(r) plot trend is also an indicator of disorder. The P(r) also provides a Dmax of 90 Å, a large distance for AP1260 which is only 120 residues long. An intrinsically disordered tail region would elucidate the Dmax of the particle observed in small angle x-ray scattering. Circular dichroism spectroscopy provided secondary structure information for AP1260. Of the secondary structure content, nearly 50% is denoted as being 'other'. Within this category, any content of the protein that does not form secondary structure is contained. An intrinsically disordered region would fall within this category. Biophysical analysis of AP1260 corroborates with bioinformatic analysis of AP1260 in that it is likely that the C-terminal tail of AP1260, predicted by AlphaFold, is an intrinsically disordered region.

X-ray crystallography was attempted to gain a high-resolution, high-confidence structure of AP1260. However, the original AP1260 construct was unable to form crystals in the conditions attempted for crystal growth. The potential intrinsic disorder of the tail region was theorised to be the reason for this, preventing crystal contacts from forming. Therefore, the other AP1260 variants were attempted. The N-terminal variant was attempted to remove any extra disorder that may have been added due to the presence of the histidine tag on the C-terminus. The truncated variants were then attempted to remove the tail in the hopes of increasing crystal contact formation. As seen from the expression SDS-PAGE gel presented in Figure 7, only the longest truncation, G103, and the N-terminus variant showed adequate expression for potential analysis. The lack of expression of the truncated variants unveils that the C-terminal tail of AP1260 may be important to its successful expression. Further research is required to optimise the crystallography of AP1260 to gain a high-resolution structure.

Further research that would be beneficial in the understanding of AP1260 is using nuclear magnetic resonance to gain more structural information – due to AP1260's potentially intrinsically disordered region. Yeast two hybrid is another technique that could be used in future for obtaining the binding partner(s) of AP1260 in plant cells. Along with yeast two

hybrid, localisation studies would show exactly where AP1260 localises within plant cells and may aid in finding binding partners. Knockout studies, knocking out AP1260 in an infecting *A. psidii* individual would test the importance of AP1260 in the infection process. The work completed in this study in conjunction with future work provide a basis from which myrtle rust treatments targeting AP1260 can be produced.

Conclusion

Austropuccinia psidii's infectious disease, myrtle rust, is a severe threat to Myrtaceae populations in all locations it is present in. Research focusing on *A. psidii*'s effector proteins provides a unique means through which the infection process can be understood and provide unique targets for inhibitory treatments. Analysis of the novel *A. psidii* effector protein, AP1260, determined the mass, thermal stability, secondary structure, and dynamics of the protein. This analysis led to the discovery of a potential intrinsically disordered region within the protein. Further experiments are required to obtain a high-resolution structure and determine the function of AP1260; however, this research lay the ground work required for these experiments to happen. The characterisation of AP1260 and other effectors will vastly improve the knowledge of the mechanisms of *A. psidii* infection and may uncover methods to interfere with their function, making their characterisation critical to the biosecurity of all nations with myrtle rust. In time this could lead to the development of methods for disease treatment.

Materials and Methods

Materials

Components for Luria broth (LB) media and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) were purchased from Invitrogen. Isopropyl-β-D-1-thiogalactopyranoside (IPTG) was purchased from GoldBio. Millipore ultrapure water (Milli-Q) was filtered in the lab and used to make all solutions.

Media and antibiotic preparation

LB medium (Miller) was purchased in powdered form and made by adding 25 g of powder to every 1 L of Milli-Q water. The media was sterilized by autoclaving then stored at room temperature. LB agar was made by adding 15 g of agar powder to every 1 L of LB medium then autoclaving to sterilize. When required, agar was melted then cooled before adding an antibiotic then poured into sterile Petri dishes in a laminar flow cabinet. The agar was left to resolidify before the plates were sealed and stored at 4 °C for short-term use.

Kanamycin and ampicillin were used as the selective antibiotics throughout this project. Stocks of both kanamycin and ampicillin were made at 1000 x concentration with Milli-Q water. The stocks were sterilized with a 0.22 μm syringe filter (Jet BioFil) before storage at -20 °C.

Transformation

The experimental work for the project began by transforming *E. coli* SHuffle T7 express competent cells, from New England Biolabs. Multiple constructs were used in this project. The full length AP1260 with a C-terminus T7 tag for expression and histidine tag for purification (Referred to as AP1260) was contained in a pET-24a(+) expression vector with kanamycin resistance. Truncated mutants, E75, W81, F82, D83, F84, G86, M90, E92, P97, G103 (full length proteins truncated at, glutamic acid at position 75, tryptophan at position 81, phenylalanine at position 82, aspartic acid at position 83, phenylalanine at position 84, glycine at position 86, methionine at position 90, glutamic acid at position 92, proline at position 97, and glycine at position 103) all with C-terminal His-tags, and a full length construct with an N-terminal His-tag were contained in a pET-21a(+) expression vector with ampicillin resistance. All gene constructs were purchased from GenScript.

The competent cells were slowly defrosted from -80 °C on ice. Approximately 100 ng of DNA was added to an aliquot of the cells. The cells were then incubated on ice for 30 min. After incubation the cells were heat shocked at 42 °C for 30 s. The cells were then cooled on ice for 5 min, before the addition of 950 µL of SOC medium and incubation at 37 °C and 180 revolutions per minute (rpm) for 60 min. Aliquots of these cells were plated onto agar plates with either kanamycin or ampicillin for the selection of the transformants and incubated at 37 °C overnight. A control plate was incubated without cells to confirm that there was no contamination occurring.

Glycerol Stocks

Stocks of the transformed cells were made by inoculating 5 mL of LB with a single transformant colony in the presence of kanamycin or ampicillin. The cultures were incubated at 37 °C and 180 rpm overnight. Following incubation 200 µL of 50% glycerol (v/v) was added to aliquots of 200 µL culture before they were flash frozen using liquid nitrogen. Stocks of transformed cells were then stored at -80 °C until usage.

Expression

Overnight cultures were made by taking a scratching of a frozen glycerol stock and adding it to 50 mL of LB with kanamycin or ampicillin. The overnight cultures were incubated at 37 °C and 180 rpm overnight. Following incubation, these cultures were then used to make large scale cultures using a 1/100 dilution into LB with kanamycin or ampicillin, which were incubated at 37 °C and 180 rpm until an appropriate optical density at 600 nm (OD_{600}) was reached (~0.6). Once achieved, IPTG was used to induce protein overexpression in the cultures by adding to a final concentration of 1 mM and incubating at 20 °C and 180 rpm overnight.

An expression trial of the truncated mutants and N-terminus construct was done in duplicate by taking a scratching of a frozen glycerol stock and adding it to 5 mL of LB with ampicillin. The cultures were incubated at 37 °C and 180 rpm overnight. Following incubation, these cultures were then used to make 50 mL cultures using a 1/100 dilution into LB ampicillin, which were incubated at 37 °C and 180 rpm until an appropriate optical density at 600 nm (OD_{600}) was reached (~0.6). Once achieved, IPTG was used to induce protein overexpression

in the cultures by adding to a final concentration of 1 mM. One batch was incubated at 20 °C and 180 rpm overnight and the other at 37°C and 180 rpm overnight.

Purification

Following protein overexpression, cells were harvested through centrifugation at 6000 x*g* for 6 min at 4 °C using a Sorvall™ Lynx 6000 Superspeed Centrifuge (Thermo Scientific). The cell pellets were stored at -20 °C until required.

All purification steps from here onward were performed at 4 °C or on ice. Cell pellets were thawed and resuspended in a solution of 50 mM 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris), 500 mM NaCl, and 20 mM imidazole at pH 8.0 at approximately 4 mL/g of cell pellet. Cells were lysed using a Hielscher UP200S ultrasonic processor with an S7 sonotrode set to 70% amplitude and 0.5 s on/off bursts for three times for 5 min to allow for cooling in between. Cell debris were removed by centrifugation at 30,000 x*g* for 30 min at 4 °C using a Sorvall™ Lynx 6000 Superspeed Centrifuge (Thermo Scientific). The resulting supernatant was decanted and stored at 4 °C for further purification.

A four-step chromatography purification protocol was used to purify AP1260 and its variants from the *E. coli* host proteins. In order, the steps were immobilized metal affinity chromatography (IMAC), a desalting column, anion exchange chromatography (AEC), and size-exclusion chromatography (SEC). The chromatography steps were performed using ÄKTA protein purification systems (Cytiva Life Sciences) and the buffers presented in Table 1. Before each step, the relevant chromatography column was pre-equilibrated with at least three column volumes of the first buffer in that step.

Table 1. Chromatography buffers and components

Buffer type	Components
IMAC buffer A	50 mM Tris, 500 mM NaCl, 20 mM imidazole, pH 8.0
IMAC buffer B	50 mM Tris, 150 mM NaCl, 500 mM imidazole, pH 8.0
Anion exchange buffer A	50 mM Tris, 30 mM NaCl, pH 8.0
Anion exchange buffer B	50 mM Tris, 1 M NaCl, pH 8.0
Size-exclusion buffer	50 mM Tris, 150 mM NaCl, pH 8.0

Immobilized Metal Affinity Chromatography

IMAC was the first chromatography step used for purifying AP1260 and its variants. The buffers that were used and their components are presented in Table 1. The supernatant obtained following cell lysis and centrifugation was loaded on to a pre-equilibrated 5 mL HisTrap FF column (Cytiva) with Ni^{2+} as the immobilized metal. Once entirely loaded, the column was washed with three column volumes of IMAC buffer A and 6% IMAC b. Proceeding washing, three column volumes of IMAC buffer B were applied, to elute the proteins that had bound to the column. The eluted proteins were fractioned and collected in 96 well plates. The fractions corresponding to absorbance peaks at 280 nm were analysed by SDS-PAGE for the presence of AP1260. The fractions identified to contain AP1260 were pooled to be used in the next purification step.

Desalting

Desalting was the next step used for purifying AP1260. The buffer that was used and its components are presented in Table 1. The pooled proteins from the previous step were loaded on to a pre-equilibrated 53 mL HiPrep 26/10 Desalting column (Cytiva). The column was then washed with three column volumes of anion exchange buffer A. Fractions of the flow-through were collected in a 96 well plate. The fractions that corresponded to absorbance peaks at 280 nm were analysed by SDS-PAGE for the presence of AP1260. Fractions containing AP1260 were pooled for use in the next purification step.

Anion Exchange Chromatography

Anion exchange chromatography was the third chromatography step used for purifying AP1260. The buffers that were used and their components are presented in Table 1. The pooled proteins from the previous step were loaded on to a pre-equilibrated 5 mL HiTrap Q Sepharose High Performance column (Cytiva) that has positively charged resin. Once entirely loaded, the column was washed with three column volumes of anion exchange buffer A. Following washing, a continuous gradient of anion exchange buffer B was applied, reaching 100% over three column volumes, to elute proteins that had bound to the column. The eluted proteins were fractioned and collected in 96 well plates. The fractions corresponding to absorbance peaks at 280 nm were analysed using SDS-PAGE for the presence of AP1260. The fractions identified to contain AP1260 were pooled for use in the next purification step.

Size-Exclusion Chromatography

Size-exclusion chromatography was the fourth and final chromatography step used for purifying AP1260. This step used a single buffer, the components of which are presented in Table 1. The pooled protein from the previous step was concentrated to ~1% of the column volume (~1.2 mL) then loaded on to a pre-equilibrated 120 mL HiLoad 16/600 Superdex 75 pg column (Cytiva). The protein was then eluted with the size-exclusion buffer. Fractions of the eluted proteins were collected in 96 well plates. Fractions of the eluted proteins that corresponded to absorbance peaks at 280 nm were analysed using SDS-PAGE for the presence of AP1260. Fractions containing AP1260 were pooled into a final protein solution for use in the following experiments.

Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Samples for SDS-PAGE were prepared by combining 10 µL of protein sample and 10 µL of loading dye containing 200 mM β-mercaptoethanol and 0.2% (w/v) bromophenol blue in Milli-Q water. Protein-dye samples were heated at 100 °C for 5 min. SDS-PAGE was performed using Bolt™ 4-12% Bis-Tris precast protein gels in an SDS-PAGE gel tank (Invitrogen). Gels were submerged in 2-(N-morpholino)ethanesulfonic acid (MES) buffer and the outer buffer chamber was filled. Samples were electrophoresed alongside a Novex® Sharp Pre-stained Protein Standard. Electrophoresis was run at 200 V for 22 min. Following electrophoresis, gels

were stained using Coomassie G-250. Images of the gels were taken using a CHEMI GENIUS^2 bio-imaging system (Syngene).

Measuring Protein Concentration

The concentration of protein solutions was estimated by measuring the absorbance at 280 nm with a NanoDrop™ One spectrophotometer (Thermo Scientific). The extinction coefficient for AP1260 was estimated to be 18115 M^{-1} cm^{-1} with the ProtParam tool from ExPASy (Gasteiger et al., 2003). The extinction coefficient and absorbance values were used with the Beer-Lambert law to obtain the protein concentration.

Analytical ultracentrifugation

Analytical Ultracentrifugation was performed with a Beckman Coulter ProteomeLab™ XL-I Protein Characterization System. The AP1260 sample (in 50 mM Tris, 150 mM NaCl, pH 8.0) was used at 0.53 mg/mL. Volumes of 400 µL of reference and 380 µL of sample were loaded into 12mm double sector cells with quartz windows and mounted in an An-60 Ti four-place rotor. The samples were centrifuged at a rotor speed of 50,000 rpm at 20 °C. Radial intensity data were collected at a single wavelength (280 nm) without averaging, using a 0.003 cm step size for a total of 650 scans.

SEDNTERP was used to calculate the partial specific volumes of AP1260 (0.71 g/mL), along with solvent density (1.0041 g/mL) and viscosity (0.01016 poise) (Laue et al., 1992). Data obtained from the experiment were fitted to a continuous size distribution [c(*s*)] model using the program SEDFIT (Schuck, 2000). The data were then fitted to a continuous mass distribution [c(*M*)] model in SEDFIT to calculate the apparent molecular mass of the species observed in the [c(*s*)] distribution.

Small angle x-ray scattering

Small angle x-ray scattering was performed using the SAXS/WAXS beamline at the Australian Synchrotron. Diffraction data were collected from AP1260 samples concentrated to 5 mg/mL and 10 mg/mL (60 µL injections). The data were processed in CHROMIXS (Panjkovich & Svergun, 2018), providing the SAXS scatter plot, P(r) plot, Guinier plot, and Kratky plot.

Circular dichroism

Circular dichroism spectroscopy was performed using a JASCO J-815 CD spectrophotometer. The AP1260 sample was analysed at a protein concentration of 0.1 mg/mL in a buffer of 20 mM NaH_2PO_4 at pH 7.4. The measured wavelength range of 195 – 250 nm with 0.1 nm intervals. To reduce the signal-to-noise ratio, any data points exceeding a high-tension voltage of 700 were removed. The quartz cell had a path length of 1 mm and the data were collected at 20 °C with 5 accumulations. To analyse the data gained from CD, BeStSel single spectrum analysis was used (Micsonai et al., 2022).

Differential scanning fluorimetry

Differential scanning fluorimetry was performed using a QuantStudio™ 3 Real-Time PCR System (Applied Biosystems). SYPRO® Orange dye (10x final dilution) was used to determine the thermal stability of the protein samples by measuring the fluorescence at 520 nm. AP1260 samples were analysed at a protein concentration of 1 mg/mL. Samples were measured in triplicate in separate wells of a 96 well plate. The protocol used incremental temperature increases of 0.03 °C per second across a temperature range of 4 °C to 95 °C

Fluorescence spectroscopy

Fluorescence spectroscopy was performed using a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies). Excitation wavelengths of 280 nm and emission wavelengths of 300 to 420 nm were measured at various temperatures to determine the thermal stability of the AP1260 protein samples. Fluorescence scans were taken at incremental temperature increases of 5°C over a temperature range of 20°C to 80°C. AP1260 fluorescence scans were also taken at increments of 5°C with temperature increasing from 20°C to 50°C before being reduced immediately back to 20°C. AP1260 samples were analysed at a protein concentration of 0.1 mg/mL.

X-ray crystallography

Protein samples were concentrated to 9 mg/mL and 22 mg/mL for AP1260 and 22 mg/mL for the N-terminus variant using Microsep™ Advance centrifugal concentrators with a molecular cut-off weight of 10 kDa. Crystallization was performed using the sitting-drop vapour-diffusion method in 96 well crystallization trays. For protein concentrated to 9 mg/mL and the N-terminus variant, the Molecular Dimensions crystal screens JCSG+, PACT Premier, and SG1

were used for to screen for successful crystallography conditions (Newman et al., 2005; Page et al., 2003). For protein concentrated to 22 mg/mL, the Molecular Dimensions screens JCSG+ and PACT Premier were used to find successful crystallization conditions (Newman et al., 2005). The drops were laid using a TTP Labtech mosquito® Crystal unit with protein samples of 200 nl and reservoir volumes of 200 nl in each drop at 20 °C in an attempt to produce crystals.

Bioinformatics search

A bioinformatics search into AP1260 was done using the protein amino acid sequence. The protein sequence was analysed using ProtParam (Gasteiger et al., 2005). The protein sequence was input into the Protein Basic Local Alignment Search Tool (BLASTP) (Altschul et al., 1990) and JPred4 (Drozdetskiy et al., 2015) with the alignments being edited in Jalview Version 2 (Waterhouse et al., 2009). The protein sequence was also input into EffectorP 3.0 (Sperschneider & Dodds, 2022), SignalP 6.0 (Teufel et al., 2022), and WoLF PSORT (Horton et al., 2007). To predict regions of disorder, the sequence of AP1260 was input into DISOPRED3 (Jones & Cozzetto, 2015) and ODiNPred (Dass et al., 2020). To predict structure, the sequence was input into AlphaFold (Jumper et al., 2021).

Software

GUSSI was used to plot data fit residuals and [c(*s*)] and [c(*M*)] distributions (Brautigam, 2015). BioRender was used to generate and label figures. GraphPad Prism 9 was used to graph purification traces using simple linear regression (*GraphPad Prism Version 9.4.1 for Windows*).

References

Allen, E. A., Hoch, H. C., Steadman, J. R., & Stavely, R. J. (1991). Influence of Leaf Surface Features on Spore Deposition and the Epiphytic Growth of Phytopathogenic Fungi. In J. H. Andrews & S. S. Hirano (Eds.), *Microbial Ecology of Leaves* (pp. 87–110). Springer. https://doi.org/10.1007/978-1-4612-3168-4_5

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Asai, S., & Shirasu, K. (2015). Plant cells under siege: Plant immune system versus pathogen effectors. *Current Opinion in Plant Biology*, *28*, 1–8. https://doi.org/10.1016/j.pbi.2015.08.008

Babu, M. M. (2016). The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochemical Society Transactions*, *44*(5), 1185–1200. https://doi.org/10.1042/BST20160172

Beenken, L. (2017). Austropuccinia: A new genus name for the myrtle rust Puccinia psidii placed within the redefined family Sphaerophragmiaceae (Pucciniales). *Phytotaxa*, *297*(1), 53. https://doi.org/10.11646/phytotaxa.297.1.5

Bent, A. F., & Mackey, D. (2007). Elicitors, Effectors, and *R* Genes: The New Paradigm and a Lifetime Supply of Questions. *Annual Review of Phytopathology*, *45*(1), 399–436. https://doi.org/10.1146/annurev.phyto.45.062806.094427

Beresford, R. M., Shuey, L. S., & Pegg, G. S. (2020). Symptom development and latent period of Austropuccinia psidii (myrtle rust) in relation to host species, temperature, and ontogenic resistance. *Plant Pathology*, *69*(3), 484–494. https://doi.org/10.1111/ppa.13145

Bhagwat, M., & Aravind, L. (2007). PSI-BLAST Tutorial. In *Comparative Genomics: Volumes 1 and 2*. Humana Press. https://www.ncbi.nlm.nih.gov/books/NBK2590/

Brautigam, C. A. (2015). Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods in Enzymology*, *562*, 109–133. https://doi.org/10.1016/bs.mie.2015.05.001

Canonne, J., & Rivas, S. (2012). Bacterial effectors target the plant cell nucleus to subvert host transcription. *Plant Signaling & Behavior*, *7*(2), 217–221. https://doi.org/10.4161/psb.18885

Carnegie, A. J., & Pegg, G. S. (2018). Lessons from the Incursion of Myrtle Rust in Australia. *Annual Review of Phytopathology*, *56*(1), 457–478. https://doi.org/10.1146/annurev-phyto-080516-035256

Chock, M. K. (2020). The global threat of Myrtle rust (AUSTROPUCCINIA psidii): Future prospects for control and breeding resistance in susceptible hosts. *Crop Protection*, *136*, 105176. https://doi.org/10.1016/j.cropro.2020.105176

Coutinho, T. A., Wingfield, M. J., Alfenas, A. C., & Crous, P. W. (1998). Eucalyptus Rust: A Disease with the Potential for Serious International Implications. *Plant Disease*, *82*(7), 819–825. https://doi.org/10.1094/PDIS.1998.82.7.819

Cui, H., Tsuda, K., & Parker, J. (2015). Effector-Triggered Immunity: From Pathogen Perception to Robust Defense. *Annual Review of Plant Biology*, *66*, 487–511.

da S. Machado, P., Alfenas, A. C., Alfenas, R. F., Mohammed, C. L., & Glen, M. (2015). Microsatellite analysis indicates that Puccinia psidii in Australia is mutating but not recombining. *Australasian Plant Pathology*, *44*(4), 455–462. https://doi.org/10.1007/s13313-015-0364-5

Dass, R., Mulder, F. A. A., & Nielsen, J. T. (2020). ODiNPred: Comprehensive prediction of protein order and disorder. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-71716-1

Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). JPred4: A protein secondary structure prediction server. *Nucleic Acids Research*, *43*(W1), W389–W394. https://doi.org/10.1093/nar/gkv332

Fensham, R. J., Carnegie, A. J., Laffineur, B., Makinson, R. O., Pegg, G. S., & Wills, J. (2020). Imminent Extinction of Australian Myrtaceae by Fungal Disease. *Trends in Ecology & Evolution*, *35*(7), 554–557. https://doi.org/10.1016/j.tree.2020.03.012

Ferrarezi, J. A., McTaggart, A. R., Tobias, P. A., Hayashibara, C. A. A., Degnan, R. M., Shuey, L. S., Franceschini, L. M., Lopes, M. S., & Quecine, M. C. (2022). Austropuccinia psidii uses tetrapolar mating and produces meiotic spores in older infections on Eucalyptus grandis. *Fungal Genetics and Biology*, *160*, 103692. https://doi.org/10.1016/j.fgb.2022.103692

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, *31*(13), 3784–3788.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571–607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Glen, M., Alfenas, A. C., Zauza, E. A. V., Wingfield, M. J., & Mohammed, C. (2007). *Puccinia psidii*: A threat to the Australian environment and economy – a review. *Australasian Plant Pathology*, *36*(1), 1. https://doi.org/10.1071/AP06088

Goyet, V., Wada, S., Cui, S., Wakatake, T., Shirasu, K., Montiel, G., Simier, P., & Yoshida, S. (2019). Haustorium Inducing Factors for Parasitic Orobanchaceae. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/article/10.3389/fpls.2019.01056

GraphPad Prism version 9.4.1 for Windows. (n.d.). GraphPad Software. www.graphpad.com

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). WoLF PSORT: Protein localization predictor. *Nucleic Acids Research*, *35*(Web Server issue), W585–W587. https://doi.org/10.1093/nar/gkm259

Jones, D. T., & Cozzetto, D. (2015). DISOPRED3: Precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*, *31*(6), 857–863. https://doi.org/10.1093/bioinformatics/btu744

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), Article 7117. https://doi.org/10.1038/nature05286

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), Article 7873. https://doi.org/10.1038/s41586-021-03819-2

Kazan, K., & Lyons, R. (2014). Intervention of Phytohormone Pathways by Pathogen Effectors. *The Plant Cell*, *26*(6), 2285–2309. https://doi.org/10.1105/tpc.114.125419

Koeck, M., Hardham, A. R., & Dodds, P. N. (2011). The role of effectors of biotrophic and hemibiotrophic fungi in infection. *Cellular Microbiology*, *13*(12), 1849–1857. https://doi.org/10.1111/j.1462-5822.2011.01665.x

Langin, G., Gouguet, P., & Üstün, S. (2020). Microbial Effector Proteins – A Journey through the Proteolytic Landscape. *Trends in Microbiology*, *28*(7), 523–535. https://doi.org/10.1016/j.tim.2020.02.010

Laue, T., Shah, B., Ridgeway, T., & Pelletier, S. (1992). Computer-aided Interpretation of Sedimentation Data for Proteins. *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. https://scholars.unh.edu/faculty_pubs/801

Leite, T. F., Moon, D. H., Lima, A. C. M., Labate, C. A., & Tanaka, F. A. O. (2013). A simple protocol for whole leaf preparation to investigate the interaction between Puccinia psidii and Eucalyptus grandis. *Australasian Plant Pathology*, *42*(1), 79–84. https://doi.org/10.1007/s13313-012-0179-6

Lorrain, C., Gonçalves dos Santos, K. C., Germain, H., Hecker, A., & Duplessis, S. (2019). Advances in understanding obligate biotrophy in rust fungi. *New Phytologist*, *222*(3), 1190–1206. https://doi.org/10.1111/nph.15641

McTaggart, A. R., du Plessis, E., Roux, J., Barnes, I., Fraser, S., Granados, G. M., Ho, W. W. H., Shuey, L. S., & Drenth, A. (2020). Sexual reproduction in populations of Austropuccinia psidii. *European Journal of Plant Pathology*, *156*(2), 537–545. https://doi.org/10.1007/s10658-019-01903-y

McTaggart, A. R., Shuey, L. S., Granados, G. M., du Plessis, E., Fraser, S., Barnes, I., Naidoo, S., Wingfield, M. J., & Roux, J. (2018). Evidence that Austropuccinia psidii may complete its sexual life cycle on Myrtaceae. *Plant Pathology*, *67*(3), 729–734. https://doi.org/10.1111/ppa.12763

Micsonai, A., Moussong, É., Wien, F., Boros, E., Vadászi, H., Murvai, N., Lee, Y.-H., Molnár, T., Réfrégiers, M., Goto, Y., Tantos, Á., & Kardos, J. (2022). BeStSel: Webserver for secondary structure and fold prediction for protein CD spectroscopy. *Nucleic Acids Research*, *50*(W1), W90–W98. https://doi.org/10.1093/nar/gkac345

Morin, L., Talbot, M. J., & Glen, M. (2014). Quest to elucidate the life cycle of Puccinia psidii sensu lato. *Fungal Biology*, *118*(2), 253–263. https://doi.org/10.1016/j.funbio.2013.12.004

Narouei-Khandan, H. A., Worner, S. P., Viljanen, S. L. H., van Bruggen, A. H. C., & Jones, E. E. (2020). Projecting the suitability of global and local habitats for myrtle rust (Austropuccinia psidii) using model consensus. *Plant Pathology*, *69*(1), 17–27. https://doi.org/10.1111/ppa.13111

Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I., & Perrakis, A. (2005). Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: The PACT/JCSG+ strategy. *Acta Crystallographica. Section D, Biological Crystallography*, *61*(Pt 10), 1426–1431. https://doi.org/10.1107/S0907444905024984

Nguyen, Q.-M., Iswanto, A. B. B., Son, G. H., & Kim, S. H. (2021). Recent Advances in Effector-Triggered Immunity in Plants: New Pieces in the Puzzle Create a Different Paradigm. *International Journal of Molecular Sciences*, *22*(9), Article 9. https://doi.org/10.3390/ijms22094709

Page, R., Grzechnik, S. K., Canaves, J. M., Spraggon, G., Kreusch, A., Kuhn, P., Stevens, R. C., & Lesley, S. A. (2003). Shotgun crystallization strategy for structural genomics: An optimized two-tiered crystallization screen against the Thermotoga maritima proteome. *Acta Crystallographica. Section D, Biological Crystallography*, *59*(Pt 6), 1028–1037. https://doi.org/10.1107/s0907444903007790

Panjkovich, A., & Svergun, D. I. (2018). CHROMIXS: Automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics*, *34*(11), 1944–1946. https://doi.org/10.1093/bioinformatics/btx846

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical Journal*, *78*(3), 1606–1619.

Smith, G. R., Ganley, B. J., Chagné, D., Nadarajan, J., Pathirana, R. N., Ryan, J., Arnst, E. A., Sutherland, R., Soewarto, J., Houliston, G., Marsh, A. T., Koot, E., Carnegie, A. J., Menzies, T., Lee, D. J., Shuey, L. S., & Pegg, G. S. (2020). Resistance of New Zealand Provenance Leptospermum scoparium, Kunzea robusta, Kunzea linearis, and Metrosideros excelsa to Austropuccinia psidii. *Plant Disease*, *104*(6), 1771–1780. https://doi.org/10.1094/PDIS-11-19-2302-RE

Soewarto, J., Somchit, C., du Plessis, E., Barnes, I., Granados, G. M., Wingfield, M. J., Shuey, L., Bartlett, M., Fraser, S., Scott, P., Miller, E., Waipara, N., Sutherland, R., & Ganley, B. (2021). Susceptibility of native New Zealand Myrtaceae to the South African strain of Austropuccinia psidii: A biosecurity threat. *Plant Pathology*, *70*(3), 667–675. https://doi.org/10.1111/ppa.13321

Sperschneider, J., & Dodds, P. N. (2022). EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic Effectors in Fungi and Oomycetes. *Molecular Plant-Microbe Interactions®*, *35*(2), 146–156. https://doi.org/10.1094/MPMI-08-21-0201-R

Talbot, N. J. (2019). Appressoria. *Current Biology*, *29*(5), R144–R146. https://doi.org/10.1016/j.cub.2018.12.050

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, *40*(7), Article 7. https://doi.org/10.1038/s41587-021-01156-3

Thordal-Christensen, H. (2020). A holistic view on plant effector-triggered immunity presented as an iceberg model. *Cellular and Molecular Life Sciences*, *77*(20), 3963–3976. https://doi.org/10.1007/s00018-020-03515-w

Thornhill, A. H., Ho, S. Y. W., Külheim, C., & Crisp, M. D. (2015). Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Molecular Phylogenetics and Evolution*, *93*, 29–43. https://doi.org/10.1016/j.ympev.2015.07.007

Tobias, P. A., Schwessinger, B., Deng, C. H., Wu, C., Dong, C., Sperschneider, J., Jones, A., Lou, Z., Zhang, P., Sandhu, K., Smith, G. R., Tibbits, J., Chagné, D., & Park, R. F. (2020). Austropuccinia psidii, causing myrtle rust, has a gigabase-sized genome shaped by transposable elements. *G3: Genes|Genomes|Genetics*, *11*(3), jkaa015. https://doi.org/10.1093/g3journal/jkaa015

Voegele, R. T., & Mendgen, K. W. (2011). Nutrient uptake in rust fungi: How sweet is parasitic life? *Euphytica*, *179*(1), 41–55. https://doi.org/10.1007/s10681-011-0358-5

Wang, S., Boevink, P. C., Welsh, L., Zhang, R., Whisson, S. C., & Birch, P. R. J. (2017). Delivery of cytoplasmic and apoplastic effectors from Phytophthora infestans haustoria by distinct secretion pathways. *New Phytologist*, *216*(1), 205–215. https://doi.org/10.1111/nph.14696

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Winter, G. (1884). Repertorium. Rabenhorstii fungi europaei et extraeuropaei. Cent XXXI et XXXII. *Hedwigia*, *23*, 164–172.

Yong, W. T. L., Ades, P. K., Tibbits, J. F. G., Bossinger, G., Runa, F. A., Sandhu, K. S., & Taylor, P. W. J. (2019). Disease cycle of Austropuccinia psidii on Eucalyptus globulus and Eucalyptus obliqua leaves of different rust response phenotypes. *Plant Pathology*, *68*(3), 547–556. https://doi.org/10.1111/ppa.12959

Zhang, C., Hicks, G. R., & Raikhel, N. V. (2014). Plant vacuole morphology and vacuolar trafficking. *Frontiers in Plant Science*, *5*. https://www.frontiersin.org/articles/10.3389/fpls.2014.00476

Acknowledgements

Firstly, I would like to thank my supervisor Ren Dobson for his support, advice, and enthusiasm throughout the project. Your interest in the results and experiments throughout the project made obtaining them fulfilling and exciting, and continuously increased my enthusiasm for the project.

Thank you to Michael Currie, without whom none of this would have been possible. Teaching me how to use new equipment, analyse results, and fielding my endless questions made me feel right at home in the lab environment. Your assistance made the whole year less stressful and chatting about the results and experiments deepened my understanding of many things.

Thank you to the entire Dobson lab group. I felt welcomed from day one. Weekly lab meetings and frequent chats produced a friendly and productive environment. Thank you all for your constructive feedback and for answering all my questions. Thanks to this I feel I was able to produce my best work.

I greatly appreciate and thank the Australian Institute of Nuclear Science and Engineering (AINSE) for their financial assistance.

Supplementary information

1 protein was provided as input.

Predicted effectors: 1 (100.0%)
Predicted cytoplasmic effectors: 1 (100.0%)
Predicted apoplastic effectors: 0 (0.0%)

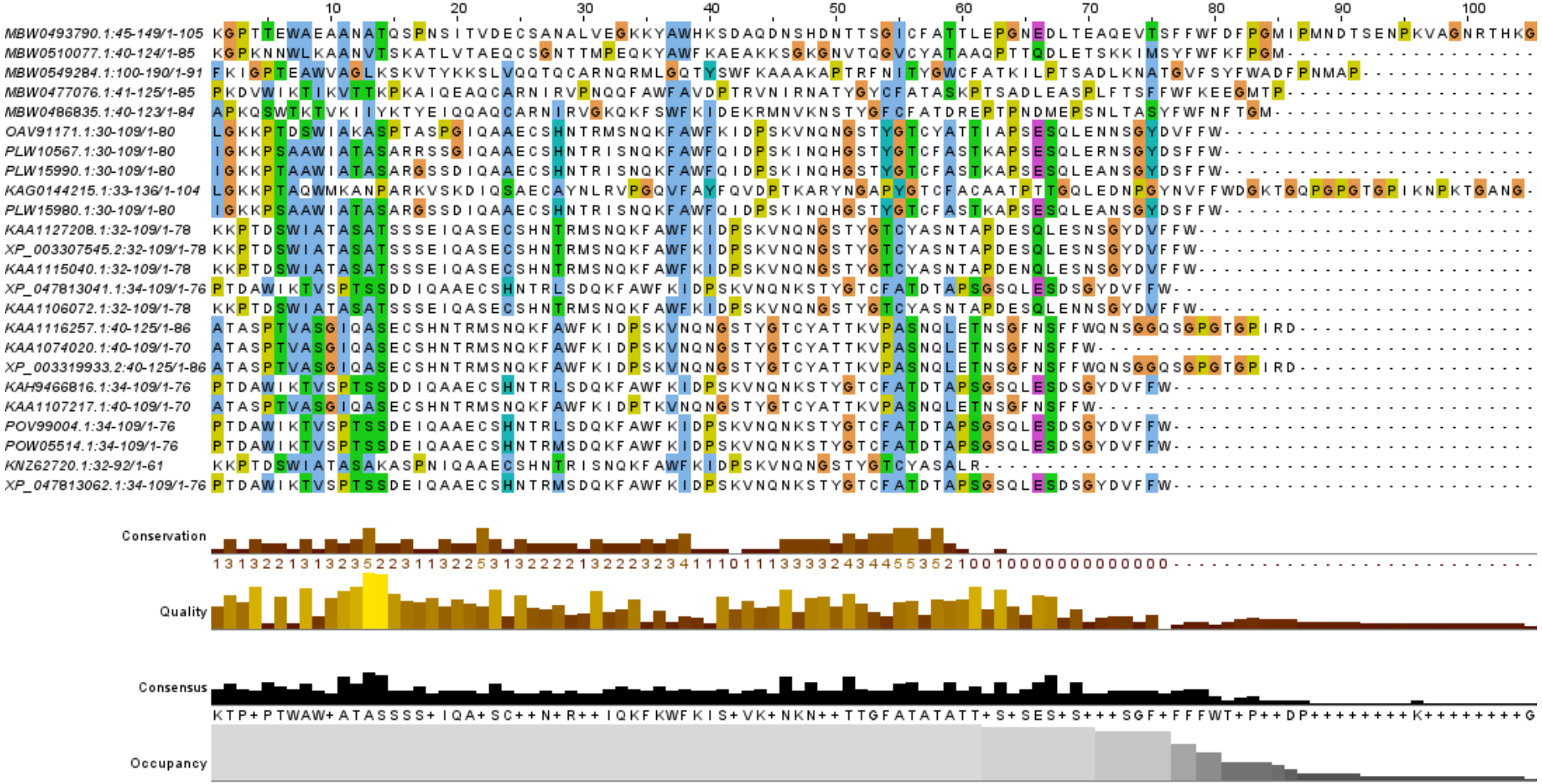
0.0% of the cytoplasmic effectors are predicted dual-localized (cytoplasmic/apoplastic).

Predicted effectors (download as FASTA file here):

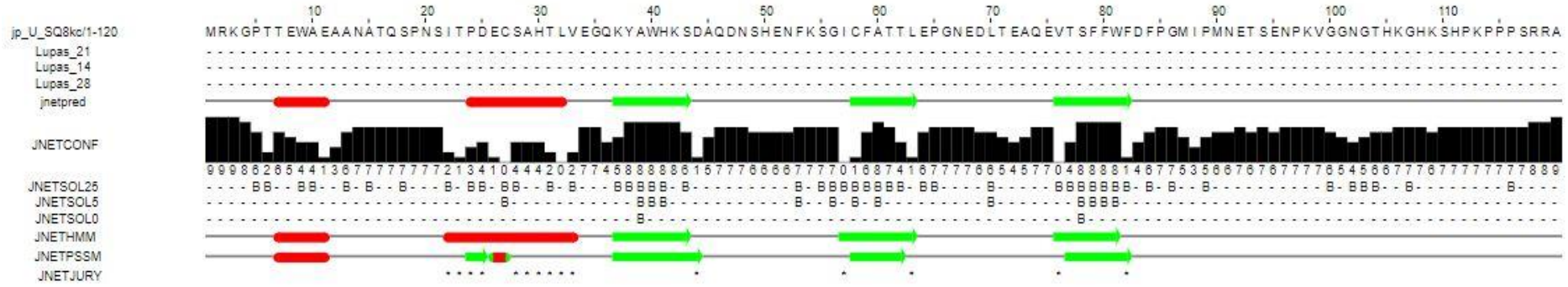
Summary table

#Identifier	Cytoplasmic effector	Apoplastic effector	Non-effector	Prediction
default	Y (0.757)	-	-	Cytoplasmic effector

Supplementary Figure 1 | EffectorP prediction. EffectorP predicts that AP1260 is a cytoplasmic effector with a probability of 0.757.



Supplementary Figure 2 | BLASTP sequence alignment. Alignment of identity match results from the non-redundant protein sequences database. All identity matches were no higher in identity than 45% and were proteins belonging to species in the order Pucciniales.



Supplementary Figure 3 | JPred4 secondary structure prediction. JPred4 analysis of the AP1260 amino acid sequence predicted two α helices and three β strands. Helices are predicted to be from Thr7 to Glu11 and Pro24 to Leu32. Strands are predicted between Lys37 to Ser43, Cys58 to Leu63, and Val76 to Phe83.

id	site	distance	identity	comments
GDA1_SCHPO	golg	337.201	7.3741%	[Uniprot] SWISS-PROT45:Type II membrane protein. Golgi.
KRE2_YEAST	golg	337.908	10.4072%	[Uniprot] SWISS-PROT45:Type II membrane protein. Golgi.
ERG2_YEAST	E.R.	385.289	12.1622%	[Uniprot] SWISS-PROT45:Endoplasmic reticulum.
CHI1_APHAL	extr	395.555	10.1655%	[Uniprot] SWISS-PROT45:Secreted.
LIP1_CANAL	extr	402.22	8.97436%	[Uniprot] SWISS-PROT45:Secreted.
GUB_ORPSP	extr	418.202	11.4286%	[Uniprot] SWISS-PROT45:Secreted.
CEL1_AGABI	extr	439.995	15%	[Uniprot] SWISS-PROT45:Secreted.
SDH3_YEAST	mito	443.904	16.1616%	[Uniprot] SWISS-PROT45:Integral membrane protein. Mitochondrial inner membrane.
MNS1_YEAST	E.R.	446.091	7.10383%	[Uniprot] SWISS-PROT45:Type II membrane protein. Endoplasmic reticulum. GO:0005783; C:endoplasmic reticulum; Evidence:IDA
GUN1_TRIRE	extr	452.269	9.78261%	[Uniprot] SWISS-PROT45:Secreted.
GAA1_YEAST	E.R.	453.986	7.81759%	[Uniprot] SWISS-PROT45:Integral membrane protein. Endoplasmic reticulum. GO:0030176; C:integral to endoplasmic reticulum membrane; Evidence:IDA
MANB_ASPNG	extr	455.21	5.6928%	[Uniprot] SWISS-PROT45:Secreted. GO:0005576; C:extracellular; Evidence:NAS
KTR2_YEAST	golg	462.004	10.3529%	[Uniprot] SWISS-PROT45:Type II membrane protein. Golgi. GO:0005794; C:Golgi apparatus; Evidence:IDA
CTR4_SCHPO	plas	469.006	10.7266%	[Uniprot] SWISS-PROT45:Integral membrane protein.
CARP_YEAST	vacu	473.546	9.1358%	[Uniprot] SWISS-PROT45:Lysosome-like vacuoles.
CSG2_YEAST	plas	475.277	9.02439%	[Uniprot] SWISS-PROT45:Integral membrane protein. GO:0030176; C:integral to endoplasmic reticulum membrane; Evidence:IDA
COQ2_YEAST	mito	482.907	10.2151%	[Uniprot] SWISS-PROT45:Integral membrane protein. Mitochondrial inner membrane.
LIPA_CANAL	extr	484.348	9.89247%	[Uniprot] SWISS-PROT45:Secreted.
GAOA_DACDE	extr	485.696	7.64706%	[Uniprot] SWISS-PROT45:Secreted.
AXHA_ASPTU	extr	487.992	11.747%	[Uniprot] SWISS-PROT45:Secreted.
COX9_YEAST	mito	495.807	7.0922%	[Uniprot] SWISS-PROT45:Mitochondrial inner membrane. GO:0005751; C:respiratory chain complex IV (sensu Eukarya); Evidence:IPI
GHT1_SCHPO	plas	501.252	8.07899%	[Uniprot] SWISS-PROT45:Integral membrane protein.
CCPR_YEAST	mito	502.856	11.0803%	[Uniprot] SWISS-PROT45:Mitochondrial matrix. GO:0005739; C:mitochondrion; Evidence:IDA
GUX2_TRIRE	extr	503.303	8.70488%	[Uniprot] SWISS-PROT45:Secreted.
GHT4_SCHPO	plas	503.593	7.71993%	[Uniprot] SWISS-PROT45:Integral membrane protein.
THIX_YEAST	plas	511.899	7.69231%	[Uniprot] SWISS-PROT45:Integral membrane protein.
CX5A_YEAST	mito	511.907	11.7647%	[Uniprot] SWISS-PROT45:Mitochondrial inner membrane. GO:0005751; C:respiratory chain complex IV (sensu Eukarya); Evidence:IPI

Supplementary Figure 4 | WoLF PSORT fungi search with signal peptide. WoLF PSORT predicts the subcellular localization of AP1260 with the signal peptide in fungi to be the Golgi, the endoplasmic reticulum, or for the protein to be secreted.

id	site	distance	identity	comments
G6PI_ASPOR	cyto	364.161	7.23327%	[Uniprot] SWISS-PROT45:Cytoplasmic (By similarity).
POR1_YEAST	mito	374.88	11.6608%	[Uniprot] SWISS-PROT45:Mitochondrial outer membrane. GO:0005741; C:mitochondrial outer membrane; Evidence:IDA
G6PI_KLULA	cyto	378.185	7.0018%	[Uniprot] SWISS-PROT45:Cytoplasmic.
G6PI_YEAST	cyto	384.592	5.60579%	[Uniprot] SWISS-PROT45:Cytoplasmic.
RT01_YEAST	mito	387.734	10.9034%	[Uniprot] SWISS-PROT45:Mitochondrial. GO:0005763; C:mitochondrial small ribosomal subunit; Evidence:IDA
UCRQ_YEAST	mito	417.661	11.6667%	[Uniprot] SWISS-PROT45:Mitochondrial inner membrane. GO:0005750; C:respiratory chain complex III (sensu Eukarya); Evidence:IDA
IDI1_YEAST	cyto	423.53	11.4983%	[Uniprot] SWISS-PROT45:Cytoplasmic.
PUR4_YEAST	cyto	424.361	3.68189%	[Uniprot] SWISS-PROT45:Cytoplasmic. GO:0005737; C:cytoplasm; Evidence:IDA
UNG_SCHPO	nucl	424.801	11.4907%	[Uniprot] SWISS-PROT45:Nuclear.
ERV1_YEAST	mito	432.219	12.1693%	[Uniprot] SWISS-PROT45:Mitochondrial. GO:0005758; C:mitochondrial intermembrane space; Evidence:IDA
COXE_SCHPO	mito	432.39	12.2137%	[Uniprot] SWISS-PROT45:Mitochondrial inner membrane.
ATPE_SCHPO	mito	435.45	10.8333%	[Uniprot] SWISS-PROT45:Mitochondrial.
UBC3_SCHPO	nucl	437.898	9.55414%	[Uniprot] SWISS-PROT45:Nuclear.
CPH1_CANAL	nucl	439.636	6.40244%	[Uniprot] SWISS-PROT45:Nuclear.
CAO1_CANTR	pero	441.839	5.74887%	[Uniprot] SWISS-PROT45:Peroxisomal.
BCA2_YEAST	cyto	442.233	10.1064%	[Uniprot] SWISS-PROT45:Cytoplasmic. GO:0005634; C:nucleus; Evidence:IDA
CPYI_YEAST	cyto	444.597	14.1553%	[Uniprot] SWISS-PROT45:Cytoplasmic. GO:0005625; C:soluble fraction; Evidence:IDA
PBP1_YEAST	cyto_nucl	452.603	6.23269%	[Uniprot] SWISS-PROT45:Nuclear and cytoplasmic. GO:0005634; C:nucleus; Evidence:IDA
NU57_YEAST	nucl	452.641	6.65434%	[Uniprot] SWISS-PROT45:Cytoplasmic and nucleoplasmic side of nuclear envelope (symmetric distribution); nuclear pore complex. GO:0005643; C:nuclear pore; Evidence:IDA
TREA_YEAST	cyto	453.902	5.7257%	[Uniprot] SWISS-PROT45:Cytoplasmic. GO:0005737; C:cytoplasm; Evidence:IDA
SC14_YEAST	golg	455.705	8.58086%	[Uniprot] SWISS-PROT45:Associated with the Golgi complex as a peripheral membrane protein. GO:0005829; C:cytosol; Evidence:IDA
PYRD_YEAST	cyto	459.21	13.0573%	[Uniprot] SWISS-PROT45:Cytoplasmic. GO:0005737; C:cytoplasm; Evidence:IDA
GUN1_HUMIN	extr	461.38	7.71144%	[Uniprot] SWISS-PROT45:Secreted.
NMT_ASPFU	cyto	463.025	8.33333%	[Uniprot] SWISS-PROT45:Cytoplasmic.
SMI1_YEAST	nucl	465.102	8.11881%	[Uniprot] SWISS-PROT45:Nuclear. GO:0005634; C:nucleus; Evidence:IDA
CYC_EMENI	mito	465.669	15%	[Uniprot] SWISS-PROT45:Mitochondrial matrix.
YB2G_SCHPO	nucl	470.012	5.74374%	[Uniprot] SWISS-PROT45:Nuclear.

Supplementary Figure 5 | WoLF PSORT fungi search without signal peptide. Without the signal peptide, WoLF PSORT predicts the subcellular localization of AP1260 in fungi to be in the cytoplasm.

id	site	distance	identity	comments
SPR1_IPOBA	vacu	234.258	11.8721%	[Uniprot] SWISS-PROT45:Vacuolar.
At1g23420.1	nucl	344.479	14.2857%	[Arath]
At4g35350.1	vacu	346.211	12.1127%	[Arath]
At5g64210.1	mito	350.251	9.06516%	[Arath]
UCRI_PEA	chlo	357.755	13.0435%	[Uniprot] SWISS-PROT45:Chloroplast thylakoid.
SPRA_IPOBA	vacu	358.227	15.5251%	[Uniprot] SWISS-PROT45:Vacuolar.
At2g44920.1	chlo	364.118	13.3333%	[Arath] Subclass:thylakoid
VPEG_ARATH	vacu	369.995	8.36735%	[Uniprot] SWISS-PROT45:Vacuolar.
PROD_ARATH	mito	376.501	7.21443%	[Arath] [Uniprot] SWISS-PROT45:Mitochondrial matrix. Evidence:IDA Pubmed:8776899
At3g12480.1	nucl	385.627	11.6041%	[Arath]
At3g21920.1	extr	403.401	10.0719%	[Arath]
At1g21210.1	plas	407.051	6.36856%	[Arath]
At1g68765.1	extr	408.788	11.3475%	[Arath]
At1g76450.1	chlo	412.981	13.7652%	[Arath] Subclass:thylakoid

Supplementary Figure 6 | WoLF PSORT plant search with a signal peptide. When searching for subcellular localization in plants, with the signal peptide, WoLF PSORT predicts AP1260 to localize in the vacuoles or the nucleus.

14 Nearest Neighbors

id	site	distance	identity	comments
At5g18560.1	nucl	309.374	11.8541%	[Arath]
At1g04370.1	nucl	335.622	12.0301%	[Arath]
At5g17430.1	nucl	339.176	5.85198%	[Arath]
CYC_FRIAG	mito	342.142	8.33333%	[Uniprot] SWISS-PROT45:Mitochondrial matrix.
CYC_WHEAT	mito	345.028	10%	[Uniprot] SWISS-PROT45:Mitochondrial matrix.
At3g20060.1	nucl	352.594	12.7072%	[Arath]
SODC_NICPL	cyto	356.013	11.9205%	[Uniprot] SWISS-PROT45:Cytoplasmic.
At4g24470.1	nucl	357.561	9.70874%	[Arath]
At1g24590.1	nucl	358.383	11.7647%	[Arath]
CYC_PHAAU	mito	359.063	13.3333%	[Uniprot] SWISS-PROT45:Mitochondrial matrix.
At5g61890.1	nucl	364.476	10.4839%	[Arath]
At5g61760.1	nucl	367.266	10%	[Arath]
At1g21910.1	nucl	369.102	12.1739%	[Arath]
At2g41710.1	nucl	370.708	8.74704%	[Arath]

Supplementary Figure 7 | WoLF PSORT plant search without signal peptide. Without the signal peptide, the plant subcellular localization of AP1260 was predicted to be at the nucleus.