Flinders UNIVERSITY
ADELAIDE · SOUTH AUSTRALIA

Plant NLR Signalling Domains Self-Associate Preceding Signal Transduction

by

Adam Richard Bentham

Thesis
Submitted to Flinders University
for the degree of

Flinders UNIVERSITY
ADELAIDE · SOUTH AUSTRALIA

Doctor of Philosophy

Science and Engineering

8th of June 2018

CONTENTS

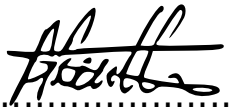
SUMMARY

Plant nucleotide-binding [NB], leucine-rich repeat [LRR] receptors (NLRs) are critical components of the plant innate immune system. Plants have evolved NLRs to detect virulence proteins secreted by microbial pathogens, deemed effectors. Upon effector detection, NLRs signal a form of programmed cell death, called the hypersensitive response (HR), which isolates the pathogen, starving them of nutrients and providing the plant resistance to further colonisation and infection. The N-terminal domains of plant NLRs are the signal transduction domain of these proteins. Transient expression assays *in planta* have demonstrated these domains can signal autonomously of the rest of the protein, and in the absence of an effector. There are two major classes of NLR proteins, segregated by their N-terminal domains. The domains that occupy the N-terminus of a plant NLR can consist of either a Toll-interleukin-1 receptor (TIR) domain or a coiled-coil (CC) domain. While some progress has been made towards understanding the structure and function of TIR and CC domains, there are still many uncertainties concerning how cell death signalling is initiated. In the case of TIR domains, two separate interfaces have been identified to be important for self-association and signal transduction. These interfaces from the TIR domains of the flax NLR L6, and Arabidopsis NLR RPS4, share a conserved overall fold, however signal through different self-association interfaces. Why this is the case is unclear and highlights unknowns in TIR domain signalling. In regard to the CC domains from plant NLRs, only two structures have been determined. These CC domains belong to the barley NLR MLA10, and the potato NLR Rx. Despite these proteins belonging to the same subclass of CC domains, both have a significantly different fold, with MLA10 CC domain forming an obligate helix-loop-helix homodimer, and Rx

forming a monomeric four-helix bundle. These differences raise questions about the true nature of CC domain structure and function. In this thesis, structural, biochemical, and biophysical techniques are used to examine the TIR domain of the Arabidopsis NLR, RPP1, and the CC domain of the wheat NLR, Sr33, in an attempt to contribute more to our understanding of N-terminal domain structure and signalling. The structure of the RPP1 TIR domain was resolved showing that both the previously identified L6 and RPS4 TIR domain interfaces are necessary for RPP1 TIR domain function, and are required simultaneously for effective signalling. The work presented here on the CC domain of Sr33 showed that this domain shares the same fold in solution to Rx, maintaining a monomeric four-helix bundle. Furthermore, a longer Sr33 CC construct than previously reported was shown to be required for the self-association of the Sr33 and MLA10 CC domains, which correlated with the ability of these CC domains to trigger cell death signalling *in planta*. This work on N-terminal domains has helped rectify some inconsistences in the literature surrounding both TIR domains and CC domains, and broaden our understanding of how NLRs signal in a more general sense.

DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed.....  .....

Date.....23/04/2018.....

ACKNOWLEDGMENTS

My family, friends, and supervisors have enabled me to reach a greater potential than I ever previously imagined. The undertaking of this work was a challenge I believed I was up to, but the constant guidance and support of the people around me greatly helped me to stay passionate, motivated, and continue in my academic pursuits even through the most difficult times. Thank you to my supportive parents, Richard and Yvonne, and my great sister and brother-in-law, Jordan and Jonny. I am exceptionally thankful to my three excellent supervisors, Simon Williams, Bostjan Kobe, and Peter Anderson, who have transformed me into the scientist I am today. I would like to thank Simon for constantly pushing me to improve, think critically, and for all the fantastic life advice. You have taught a great deal and been an excellent mentor; I hope my career is something you can be proud of. Bostjan, I would like to thank for taking time to out of many days for talks and discussions on projects, teaching me how to write like a scientist, and running a fantastic lab that perfectly balances hard work with a relaxed environment. Your lab is the perfect environment for any Ph.D. student to flourish. Peter, you inspired me as an undergraduate student with your engaging, enthusiastic teaching and love of science, and have constantly been a source of excellent advice. Thank you for constantly pushing me forward, and supporting me in all the decisions I made throughout my Ph.D.

To all the past and current Kobe and Anderson lab members, and co-authors I have had the pleasure of working with, thank you. In particular Megan Outram and Hayden Burdett, who have both been excellent sources of discussions over coffees and beers on struggles, papers, careers, relationships and everything in between. I look forward to all the catch-ups at all the various

conferences in the future.

I would also like to acknowledge the Grains Research and Development Council (GRDC) and the Thyne-Reid Foundation would have seen fit to award me additional scholarships to help me financially through my candidature. The additional support has assisted me greatly and allowed me to focus on my research without any financial stress.

1. GENERAL INTRODUCTION

1.1 An Overview of Plant–Pathogen Interactions

Although the plant and animal immune systems both use targeted cell death as a weapon against microbial pathogens, there are significant differences in how this process is regulated in the two kingdoms. The main difference lies in the metazoan adaptive circulating immune system; a feature plants do not have. To defend the entire plant from infection, each individual cell must be capable of recognising potential pathogens, and autonomously mounting a defence response (Dodds and Rathjen, 2010). Plants and animals share some elements of pathogen recognition, but the nature of the plant immune system demands a more expansive system of pathogen detection (Nyman et al., 2008).

There are several major classes of microbial pathogens which parasitise plants: bacteria, viruses, fungi, and oomycetes. Based on the nature of infection, these pathogens can be further characterised into another three classes: necrotrophs, hemibiotrophs and biotrophs. Necrotrophic pathogens invade and destroy plant tissue to extract nutrients from the dying tissue, whilst hemibiotrophs colonise the plant before inducing cell death to gain nutrients from the plant (Horbach et al., 2011). Finally, biotrophic pathogens depend on a living host, sequestering nutrients from the host through the manipulation of cell metabolic processes (Glazebrook, 2005).

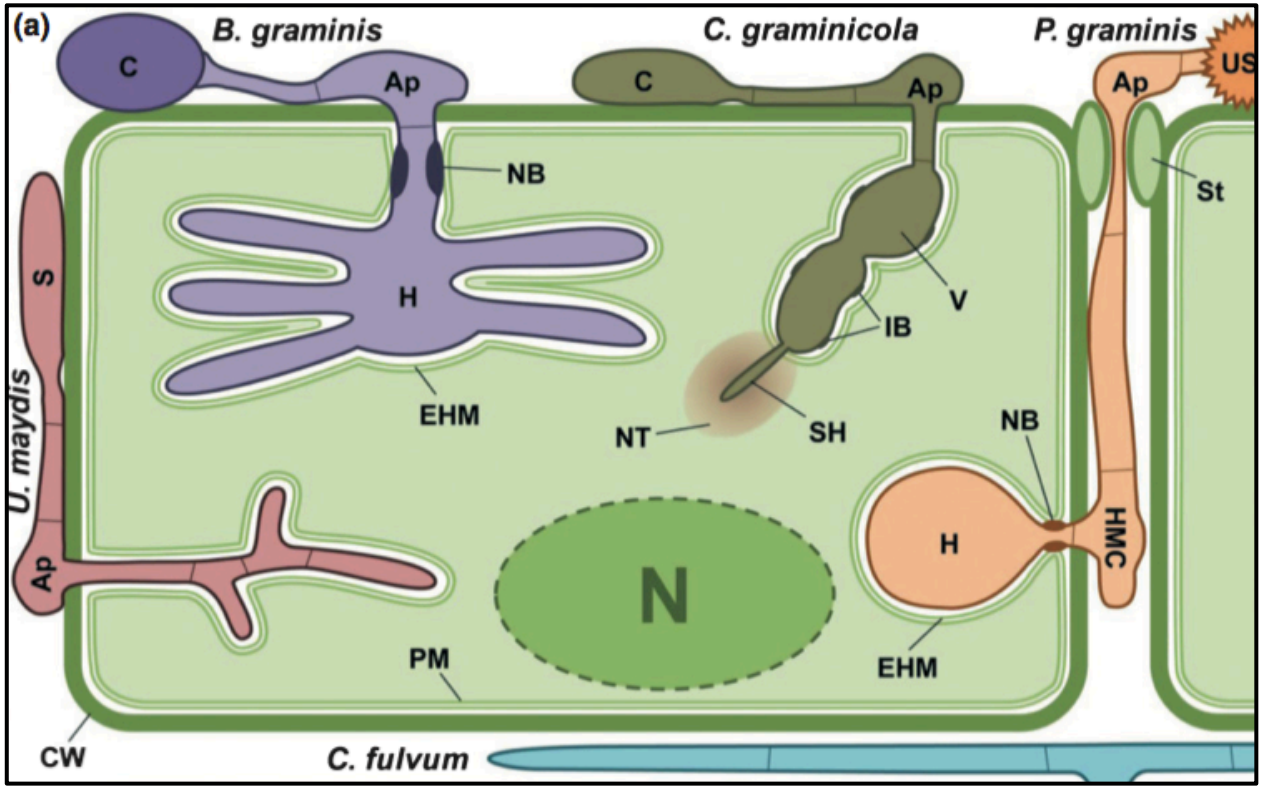
1.1.1 Infection Strategies

Pathogens must navigate a series of plant-derived obstacles to infect and draw nutrients from, or kill, its host (Cantu et al., 2008). The first is access to the interior of the plant. Bacterial pathogens, like *Pseudomonas syringae*, can access the plant through openings such as the stomata or wounds. Once in the apoplast, pathogens

must navigate a complex variety of extracellular receptors on the plasma membrane of the cells to avoid triggering the plants basal immune response (Chisholm et al., 2006).

Biotrophic fungi may gain access to the plants extracellular environment via stomata, wounds or by puncturing the epidermal layer with infection hyphae. There are a variety of species-dependent strategies for extracting nutrients from the host (Doehlemann and Hemetsberger, 2013). *Cladosporium fulvum* thrives in the extracellular spaces between cells; pathogens like these colonise the apoplast, and from there, access the host nutrients. Some fungi, like corn smut (*Ustilago maydis*), invaginate themselves in the cell membrane, forming a biotrophic interface that allows nutrient exchange from host to pathogen (Joosten and de Wit, 1999). Another example of this specialised feeding structure, known as the haustorium, is found in *Melampsora lini* and *Puccinia graminis*, which cause flax and wheat stem rust.

Haustoria create an extensive interaction zone with the host cell. This interaction zone is the extrahaustorial matrix, and is typically separated from the apoplast by a neckband structure. The neckband creates a physical barrier between the extrahaustorial and extracellular matrices, creating a closed interaction system (Catanzariti et al., 2007). These fungal feeding strategies have been summarised in Figure 1.1.



AP	Appressorium
C	Conidium
CW	Cell Well
EHM	Extrahaustorial matrix
H	Haustorium
HMC	Haustorial mother cell
IB	Interfacial bodies
N	Nucleus
NB	Neckband
NT	Necrotic tissue
PM	Plasma membrane
S	Sporidia
SH	Secondary Hypha
St	Stoma
US	Uredospore
V	Vesicle

Figure 1.1 Diagrammatic representation of the infection strategies used by different biotrophic fungi. Seen here are the fungi *Blumeria graminis* (barley powdery mildew), *Colletotrichum graminicola* (cereal and maize anthracnose), *Puccinia graminis* (stem rust), *Ustilago maydis* (corn smut), *Cladosporium fulvum* (tomato leaf mould). Image adapted from Doehlemann and Hemetsberger, 2013.

1.1.2 The Plant's Layers of Defence

The plasma membrane of the plant cell maintains a comprehensive repertoire of transmembrane receptors (Cohn et al., 2001). These receptors serve to detect conserved pathogen-associated molecule patterns (PAMPs) (Felix et al., 1993, Boller, 1995). Detection of a PAMP triggers a strong basal defence response, which denies the pathogen access to nutrients (Gomez-Gomez and Boller, 2000, Chisholm et al., 2006). This PAMP-triggered immunity (PTI) provides a broad defence against a wide variety of microbes, and often succeeds in preventing infection (Cohn et al., 2001, Chisholm et al., 2006).

More specialised bacterial and fungal pathogens have evolved mechanisms to disrupt this basal response through the use of virulence proteins known as "effectors" (Jones and Dangl, 2006). The collective role of these effectors is to block or dampen the basal defence response and facilitate access to host nutrients (Dangl and Jones, 2001). However, these pathogens must contend with a secondary layer of defence. In a response to pathogen effectors, plants have evolved another layer of immunity known as effector-triggered immunity (ETI). This system is orchestrated by an array of effector-specific nucleotide-binding [NB], leucine rich repeat [LRR] receptors (NLRs) (Jones and Dangl, 2006).

Effector-triggered immunity relies on the detection of pathogen effectors by cognate NLR proteins. NLR activation is governed by interactions with an effector, which results in signalling of a defence response known as the hypersensitive response (HR) (Dangl and Jones, 2001). The HR is typified by a burst of reactive oxygen species (ROS), which is often followed by uncontrolled cell death. This response provides a level of disease resistance far greater than that of the PTI response (Tsuda and Katagiri, 2010).

The oxidative burst during HR is coincident with the death of the cell, and creates a toxic, inhospitable environment for invading pathogens (Baker and Orlandi, 1995, Lamb and Dixon, 1997). The release of ROS, in particular H_2O_2, can act as a diffusible signal that activates defence genes in neighbouring cells, but also causes other cells challenged by the pathogen to undergo their own HR (Levine et al., 1994, Alvarez et al., 1998, Apel and Hirt, 2004, Torres et al., 2006).

Once signalled, the HR leads to the up-regulation of defence-related genes and a cascade of programmed cell death in all cells challenged by the pathogen, and in some cases a zone of surrounding cells. This creates a host-mediated quarantine area at the site of infection, and thereby providing the plant with resistance to the pathogen (Heidrich et al., 2012).

The PTI and ETI responses are the two major weapons in a plant's defence arsenal against pathogens. The following sections of the introduction will provide an in-depth discussion of the molecular mechanisms required for effective PTI and ETI functions, and how these mechanisms allow the plant immune system to provide defence to pathogens.

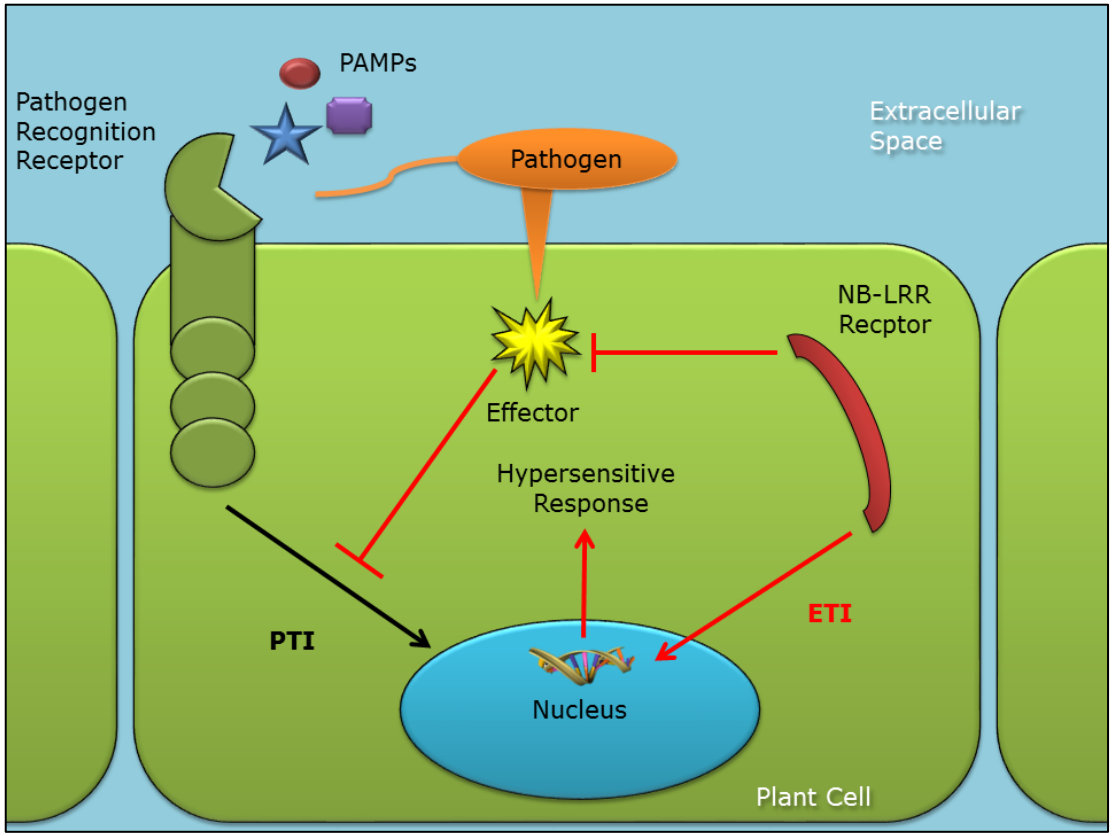


Figure 1.2 A cartoon depiction of the plant innate immune system. Extracellular domains of PRRs on the surface of the plasma membrane detect conserved PAMPs, and begin signalling the PTI response. This response denies potential pathogen from accessing nutrients with in the cell, ultimately ending the pathogens attempts at colonising the plant. To circumvent the PTI response, effectors are secreted into the plant cell to interfere with PTI signalling. In a classically termed "gene-for-gene" reaction to effectors (Flor, 1971), plants have evolved a secondary tier of immunity that relies on the use intracellular NLRs. The NLRs detect virulence proteins and lead to the activation of effector-triggered immunity (ETI), which triggers the hypersensitive response (HR), halting pathogen infection, and providing resistance to the plant.

1.2 PAMP-triggered Immunity

1.2.1 PTI: PAMPs and PAMP Receptors

PAMPs are conserved across a broad range of microbial species; this conservation is often due to being indispensable to the life of the organism (Postel and Kemmerling, 2009). Well-characterised examples of PAMPs include flagellin, the large protein subunit of bacterial flagellum (Chinchilla et al., 2007); the bacterial protein synthesis elongation factor EF-Tu (Kunze et al., 2004), and peptidoglycan (PGN), the major constituent of the Gram-negative bacterial cell wall (Gust et al., 2007). Fungal PAMPs include chitin and ergosterol, both essential components of the fungal cell wall (Boller, 1995, Chisholm et al., 2006).

The detection of PAMPs is performed by protein receptors on the surface of the plasma membrane (Felix et al., 1993, Boller, 1995). These proteins are largely receptor-like proteins (RLPs) and/or receptor-like kinases (RLKs) (Gomez-Gomez et al., 1999, Gomez-Gomez and Boller, 2000). Collectively, these proteins are known as pattern recognition receptors (PRRs), and provide a broad defence against a variety of microbial pathogens (Zipfel, 2009, Dodds and Rathjen, 2010, Monaghan and Zipfel, 2012).

1.2.2 PRR Signalling and Defence Responses

Many PRRs trigger a defence-specific mitogen-activated protein kinase (MAPK)-dependent signalling cascade upon detection of a PAMP (Meng and Zhang, 2013). Mitogen-activated protein kinase signalling by the PTI system results in activation of a wide variety of basal defence mechanisms (Jones and Dangl, 2006, Dodds and Rathjen, 2010, Meng and Zhang, 2013). These mechanisms include, but are not restricted to, the formation of callose-based papillae and deposition of lignin at the cell wall, the release of phytoalexins and other anti-microbial compounds, and the

production of ROS (Mehdy, 1994, Lamb and Dixon, 1997, Hammerschmidt, 1999, Apel and Hirt, 2004, Ferreira et al., 2006).

ROS play a particularly important role in the PTI basal defence response. They act as oxidants in lignin synthesis, they mediate papillae formation, and they exhibit toxicity by damaging peptides and nucleic acids through oxidation (Freudenberg, 1957, Mehdy, 1994, Davin et al., 1995, Barcelo, 1997, Nyska and Kohen, 2002, Apel and Hirt, 2004, Bhuiyan et al., 2009). Furthermore, ROS have their own role in defence response induction through interaction with the salicylic acid (SA) pathway (Draper, 1997, Shirasu et al., 1997). A signal activation loop potentiated by ROS, whereby hydrogen peroxide (H_2O_2) induces SA accumulation, causes the down-regulation of ROS-scavenging processes, leading to a massive increase in cellular ROS levels (Shirasu et al., 1997, Klessig et al., 2000). This high accumulation of ROS results in series of oxidative bursts that further activates systemic defences and is one of the drivers of the HR during ETI (Lamb and Dixon, 1997, Nyska and Kohen, 2002, Apel and Hirt, 2004).

1.2.3 PTI: Effective Defence, but Not Absolute

The PTI system is highly effective at preventing infection and the proliferation of many pathogens. The combined basal responses, activated by MAPK signalling, create an inhospitable environment for the pathogen, and primes the plant defence system to prevent further infection. However, this defence is not insurmountable. As mentioned previously, some pathogens have demonstrated the capability to disrupt the PTI response through the use of effectors (Jones and Dangl, 2006).

1.3 Effector-Triggered Susceptibility

1.3.1 Effectors Secreted by Pathogens Mediate Susceptibility

Many pathogens release a vast cocktail of effectors to mitigate the effects of the basal PTI response (Jones and Dangl, 2006). These secreted effectors can modulate host cell physiology to help support parasitism (Hogenhout et al., 2009, Dodds and Rathjen, 2010). Effector-triggered susceptibility (ETS) is a critical interaction between pathogen and plant that needs to be clearly understood; therefore, ETS is the basis of much ongoing research (Wirthmueller et al., 2013).

The delivery of effectors varies between pathogens. Bacterial pathogens use a type-III secretion system, which involves the injection of effector proteins into the plant cell (Grant et al., 2006). Once inside the host, effectors can be targeted to a variety of subcellular compartments such as the plasma membrane, cytoplasm and nucleus (Nimchuk et al., 2000, Szurek et al., 2001, Caillaud et al., 2012). In contrast, the mechanisms used by fungal or oomycete effectors remain highly conversional and largely unknown (Ellis and Dodds, 2011, Tyler et al., 2013, Wawra et al., 2013). In the case of oomycetes, a general mechanism for effector translocation into the cell remains elusive. However, an N-terminal RxLR motif that appears to play a role in granting the effector access to the host cell has been associated with some effectors (Win et al., 2007). In fungi, a similar, conserved recognisable motif does not appear to be present and the movement of these effectors into host cells is still the topic of much debate.

1.3.2 Effectors are Variable in Structure and Function

Studies on ETS caused by phytopathogenic bacterial effectors have uncovered a plethora of functions and targets with in the host. It is clear that many effectors have the ability to suppress PTI (Jones

and Dangl, 2006, Dodds and Rathjen, 2010). Effectors can act as competitive inhibitors for PRR signalling complexes and suppress downstream MAPK signalling (Gust et al., 2007, Meng and Zhang, 2013).

The best example of interference with PTI signalling is found in the interaction between the *Pseudomonas syringae* effector AvrPtoB and its targets (Abramovitch et al., 2006, Xiao et al., 2007, Gimenez-Ibanez et al., 2009, Cheng et al., 2011). AvrPtoB directly binds to the intercellular kinase domain of some PRRs, as well as kinase-domain containing PRR-interacting proteins. For example, it has been demonstrated that AvrPtoB can form a complex with the BRI1-associated receptor (BAK1), the chitin elicitor receptor kinase 1 (CERK1), and elongation factor receptor (EFR) (Shan et al., 2008, Gimenez-Ibanez et al., 2009). The crystal structures of AvrPtoB and BAK1 show that AvrPtoB acts as a competitive inhibitor to the BAK1 kinase domain, which prevents activation and stops the propagation of defence signalling downstream of CERK1, EFR and FLS2 (Flagellin-Sensing 2) receptors (Shan et al., 2008, Cheng et al., 2011).

Effectors are not limited to competitive inhibition. The HopU1 effector of *P. syringae* has demonstrated a gene silencing-like function. This is achieved by binding to host proteins necessary for binding to PRR mRNA transcripts, leading to a significant reduction in PRR translation (Zhang et al., 2010). In a similar way, the *Xanthomonas campestris pv. vesicatoria* XopD effector, and the transcription activator-like (TAL) effectors, confer virulence through binding to host DNA, altering transcription of specific genes (Szurek et al., 2001, Kim et al., 2008, Kim et al., 2013). XopD contains a DNA-binding domain and EAR-type transcriptional repressor motif, in conjunction with a C-terminal small ubiquitin-like modifier (SUMO) peptidase domain. With this arsenal, this

effector is able to disrupt salicylic acid (SA) signalling and ethylene driven immunity by targeting the ethylene response factor 4 (ERF4) transcription factor for degradation by the proteasome through SUMO1 cleavage (Kim et al., 2013, Kim et al., 2008). TAL effectors operate by binding individual host promoter regions in a sequence specific manner (Robatzek and Wirthmueller, 2013). Despite their varied targets, the ultimate goal of these effectors (HopU1, XopD and TALs) is to alter host transcript levels, producing a pathogen-friendly host proteome for the duration of infection (Wirthmueller et al., 2013).

Another strategy used by phytopathogenic bacteria is to degrade PRRs. It has been reported that AvrPtoB is able to target the CERK1 and other immune kinases for degradation through interaction with host E2 enzymes with its ubiquitin E3 ligase domain (Abramovitch et al., 2006). Some effectors, like HopAR1 (also known as AvrPphB) have proteolytic functions, and are known to interfere with signalling through cleavage of activation loops in several immune kinases downstream of the activated CERK1, FLS2, and EFR PRRs (Shao et al., 2003, Zhu et al., 2004, Kim et al., 2005). HopAR1 is has been demonstrated to do this by using cysteine protease activity, as it shares a fold similar to that of the papain-like cysteine proteases (Zhu et al., 2004).

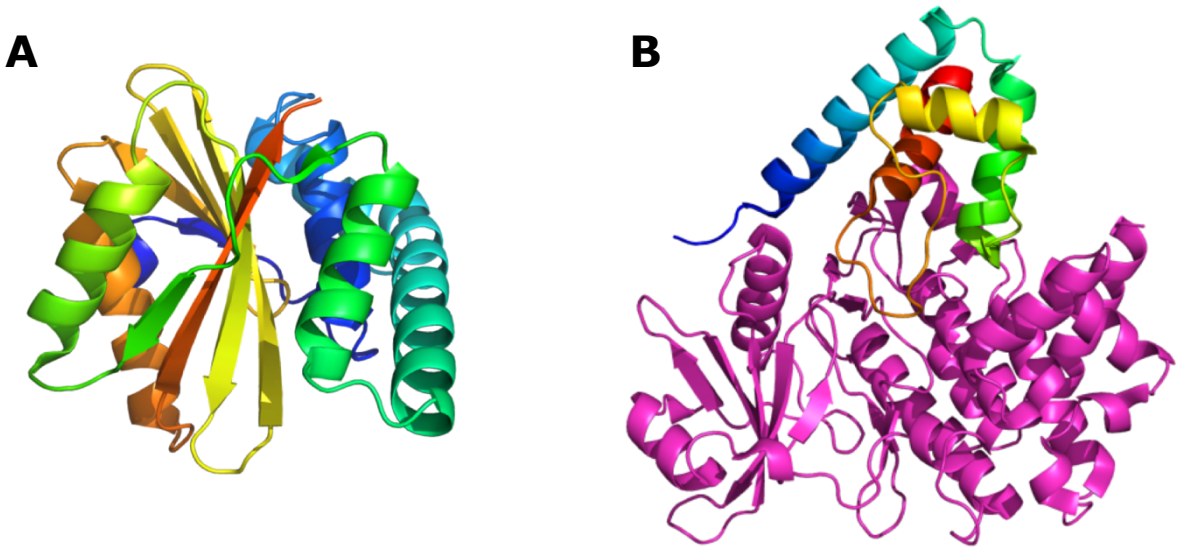


Figure 1.3 Crystal structures of bacterial effectors (**A**) Crystal structure of HopAR1 shows a similar structure to the papain-like cysteine proteases. The HopAR1 is coloured red through to blue indicating the N-terminal region through to the C-terminal region of the protein (PDB ID: 1UKF) (Zhu et al., 2004). (**B**) The crystal structure of AvrPtoB in complex with the BAK1 kinase domain. AvrPtoB is coloured in the same manner as HopAR1, with BAK1 coloured in magenta (PDB ID: 3TL8) (Cheng et al., 2011).

1.3.3 The Current Understanding of the Effectors from Filamentous Pathogens

In contrast to the effectors released by phytopathogenic bacteria, little is known about the function of many fungal and oomycete effectors (Doehlemann and Hemetsberger, 2013, Wirthmueller et al., 2013). However, it is clear that fungi and oomycetes also utilise a cocktail of effectors during infection, and secrete these effectors to promote virulence. How effectors from filamentous pathogens gain access to the interior of the host is uncertain, with much debate surrounding a conserved RxLR motif, which appears to facilitate this action (Bos et al., 2006, Win et al., 2007, Hardham and Cahill, 2010, Ellis and Dodds, 2011, Caillaud et al., 2012, Tyler et al., 2013, Wawra et al., 2013). It has been proposed this motif promotes the binding to phosphatidylinositol phosphates (PIPs) on the surface of the plasma membrane, and this induces endocytosis of the effector (Kale et al., 2010, Sun et al., 2013, Tyler et al., 2013). However, the specificity of this binding, and whether PIP binding is in fact mediated by the RxRL motif, is hotly contested (Ellis and Dodds, 2011). Regardless of opposing data, it is agreed the RxLR motifs in oomycete effectors play a role in translocation from the pathogen into the host (Win et al., 2007, Ellis and Dodds, 2011, Wirthmueller et al., 2013, Wawra et al., 2017).

In regards to fungal and oomycete effector function, very little is known; however it is postulated that the C-terminal regions of filamentous effector proteins dictate biochemical functions (Doehlemann and Hemetsberger, 2013, Wirthmueller et al., 2013). This is most probably due to the N-termini being devoted to translocation from pathogen to host cell. Some potential functions of filamentous effectors can be gleaned from structures solved to date. AvrL567-A and AvrL567-D (*Melampsora lini*) both display and β-sandwich fold, and share some structural similarity to the ToxA

effector from *Pyrenophora tritici-repentis* (Wang et al., 2007). The structures of these effectors show two regions of positive surface charge; these regions have been speculated to represent possible DNA binding sites. *In vitro* DNA binding experiments support this claim, however the biological significance of the similarity to ToxA, and DNA binding, is yet to be demonstrated (Wang et al., 2007). There are other published rust effector structures, but no clear functions. From this one could conclude that effectors have diverse structures and thus likely function in conferring ETS.

1.3.4 Detecting Effectors: A Gene-for-Gene Response

Pathogen virulence relies on a wide and diverse assortment of effectors. The ability to disrupt PTI signalling renders a plant susceptible to infection, and is an important tool for any potential pathogen (Doehlemann and Hemetsberger, 2013). However, in this evolutionary battle between plant and pathogen, not all plants are vulnerable to these effector-wielding pathogens. As mentioned, plants have developed a secondary immune response, adapted to recognise the threat of specific pathogens via race-specific effector detector proteins, NLRs (Cohn et al., 2001).

1.4 Effector-Triggered Immunity

1.4.1 NLRs and Effector Detection

Detection of effectors by NLRs can occur in a direct or indirect manner. Direct detection involves ligand interaction with an effector, whereas indirect detection is through identification of changes in the host that are mediated by an effector (Dodds and Rathjen, 2010).

There is evidence that some NLRs interact directly with effectors via a sensory domain. The best evidence for this comes from the study of the interactions between the flax rust effector AvrL567,

and the sensory leucine-rich repeat (LRR) domains of the NLRs from the flax L locus (Dodds et al., 2006).

Indirect effector detection, also referred to as the guard system, is more complex. A clear example of this is the interactions between the effector AvrPtoB, its cognate NLR Prf, and Prf's signalling partner Pto. Prf is an NLR protein found in *Arabidopsis thaliana* that detects the *P. syringae*-derived AvrPtoB effector (Oh and Martin, 2011). Prf's capacity to detect the presence of AvrPtoB lies in its ability to sense changes in a ligand partner, Pto. Interaction of AvrPtoB with Pto triggers biochemical changes within Pto, which are then detected by Prf. This results in activation of Prf, which then begins defence signalling (Ntoukakis et al., 2013).

Recently, another model of effector detection that shares elements similar to the guard model has been presented. The integrated decoy (ID) model proposes some proteins have in-built decoy domains, not generally conserved in NLRs. These ID domains share homology to potential effector targets within the host, and are thus thought to act as traps for effector proteins (Cesari et al., 2014a). It is suggested that these decoy domains bind effectors, causing the NLR to relay a signal to another partner NLR, which is subsequently activated and initiates defence signalling.

There is now experimental data from multiple heteromeric NLR protein signalling partners, from a variety of plant species, to provide evidence for this model (Cesari et al., 2014b, Cesari et al., 2014a, Williams et al., 2014, Zhai et al., 2014). In particular, it was shown the rice NLR, Pikp-1, utilises a heavy-metal associated (HMA) domain to bind the AvrPikD effector from the rice blast pathogen, *Magnaporthe oryzae* (Maqbool et al., 2015). In this study, the crystal structure of the HMA domain revealed a bound AvrPikD effector molecule, and it was shown, using isothermal

titration calorimetry (ITC), that the effector bound to the HMA domain with nanomolar affinity (Maqbool et al., 2015).

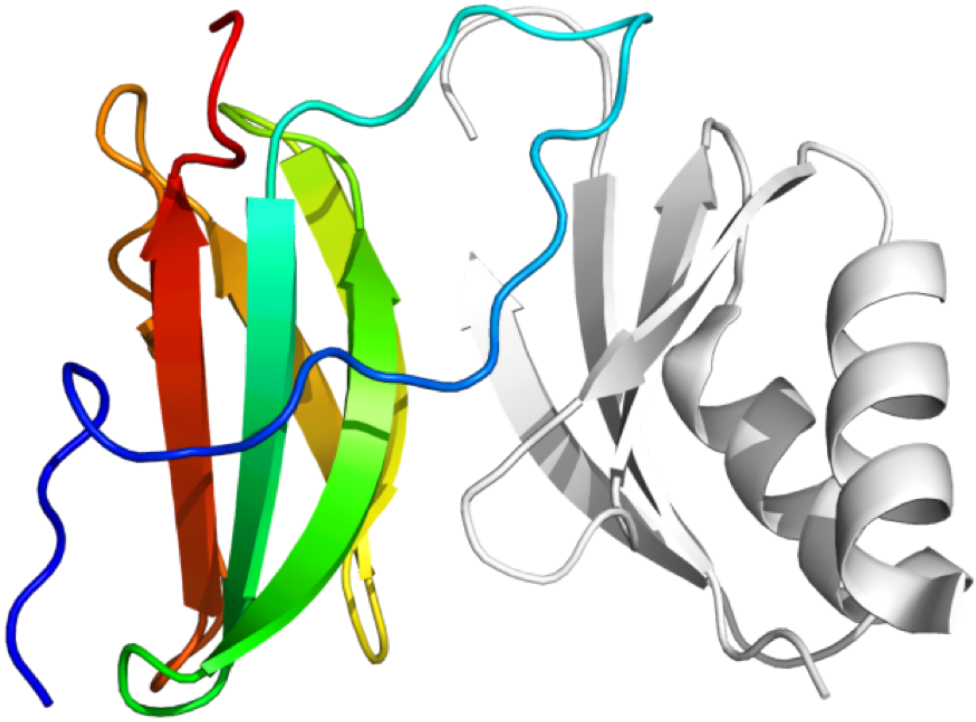


Figure 1.4 The crystal structure of AvrPikD bound to the Pikp-1 HMA domain. AvrPikD is coloured blue through to red from the N- to the C-terminus. The HMA domain of Pikp-1 is coloured in white. PDB ID: 5A6W (Maqbool et al., 2015).

1.4.2 NLR Signalling can be Complex

Some plant NLRs are capable of functioning in pairs, with either homo- or heterodimerisation being critical for function (Eitas and Dangl, 2010). A study of two *Arabidopsis* NLRs, RPS4 (resistance to *Pseudomonas syringae* 4) and RRS1 (resistance to *Ralstonia solanacearum* 1) has exemplified one such partnership. Initially it appeared RRS1 mediated a defence response to the PopP2 effector from *Ralstonia solanacearum* and RPS4 provided resistance to *Pseudomonas syringae* through detection of the AvrRPS4 effector. However, it has since been shown the collaboration of both receptors is required for both the effector recognition and subsequent defence signalling (Huh et al., 2017). Furthermore, the cooperation between these two proteins dictates resistance to a different pathogen, *Colletotrichum higginsianum*, through identification of a factor yet to be characterised, to which neither protein can provide resistance alone (Narusaka et al., 2009). It has been speculated that this NLR partnering may be a frequent feature in plant immunity, as several NLR genes in rice have shown such cooperation in providing resistance to the fungal pathogen *Magnaporthe oryzae* (Ashikawa et al., 2008, Okuyama et al., 2011, Yuan et al., 2011, Cesari et al., 2014b).

While the genetics, and some of the structural basis, of effector recognition is understood, very little is understood about how this translates into plant NLR protein activation and downstream defense signalling. From here the current knowledge of NLR protein structure and function will be discussed, and in doing so, bring to attention gaps in our understanding that need to be addressed.

1.5 Nucleotide-Binding Leucine Rich Repeat Receptors

1.5.1 NLRs exist in plants and animals

The majority of NLR proteins fall under the STAND protein family classification (Lukasik and Takken, 2009). STAND proteins are signal transduction ATPases with numerous domains, and typically act as recognition and signal transduction hubs. These proteins are found in multiple regulatory processes within the cell, contributing to immune function and apoptosis (Danot et al., 2009). Plant NLR proteins have considerable similarity to their animal NLR counterparts, with these proteins sharing comparable domain arrangements (Tameling and Takken, 2008). These two NLR groups exhibit a tripartite domain architecture, consisting of an N-terminal signalling domain, central nucleotide-binding domain (NBD), and C-terminal leucine-rich repeat (LRR).

1.5.2 The Nucleotide-Binding Domain: A Molecular Switch

The nucleotide-binding domain (NBD), known as the NB-ARC domain in plant NLRs (nucleotide binding adaptor shared with APAF-1, some *R* gene products and CED-4), and the NACHT domain in animal NLRs (named after the NAIP, CIITA, HET-E and TP-1 proteins), is the central activation domain. In plant NLRs, it consists of 3 subdomains: the nucleotide-binding (NB) domain, and the ARC1 and ARC2 domains, which come together to form a closed nucleotide-binding region in its resting state (Takken et al., 2006). Each of these domains has highly conserved motifs that play a critical role in the function and activation (Takken et al., 2006, Takken and Goverse, 2012).

The function of the NBD in NLRs is commonly likened to a molecular switch. Studies by Reubold et al. (2011 & 2014), describing the crystal structure of the NB-ARC domain from the human NLR-like protein, apoptotic protease-activating factor 1

(APAF-1), were the first to show the molecular processes involved during activation of the protein were through the NBD. In the proteins "off" state, the NBD is in a closed conformation with an ADP molecule bound. The NLR turns "on" when ADP is exchanged for ATP, resulting in a more open protein conformation (Yan et al., 2005, Reubold et al., 2011, Reubold et al., 2014) (Figure 1.5). Hydrolysis of the ATP to ADP resets the switch which results in a closed protein conformation (Riedl et al., 2005, Yan et al., 2005, Reubold et al., 2011).

Multiple studies have provided strong evidence that nucleotide binding is also critical in controlling the activation states of plant NLR proteins. To date, nucleotide binding through the NB-ARC has been demonstrated to be necessary for function in several plant NLRs: I-2 (tomato), MLA27 (barley), and M and L6 (flax), (Van Ooijen et al., 2008, Maekawa et al., 2011a, Williams et al., 2011, Maekawa et al., 2011b, Bernoux et al., 2016). Studies on the NB-ARC domain of the M protein from flax demonstrated that point mutations in the conserved P-loop and MHD motifs disrupt nucleotide binding, and caused anomalies in activation relative to wild-type M (Williams et al., 2011). Lysine (K) to leucine (L) changes in the P-loop motif of the NB domain abolished nucleotide binding, causing inactivation; and aspartate (D) to valine (V) substitutions in MHD motif of the ARC2 domain prevented ATP hydrolysis, resulting in a constitutively active protein (Williams et al., 2011).

The studies on I-2 (tomato), M (flax) and MLA27 (barley) have given some insight into the role of the NB-ARC and its motifs in plant NLR activation (Van Ooijen et al., 2008, Williams et al., 2011, Maekawa et al., 2011a). The only exception is the tobacco NLR, N, which has been observed to bind ATP when inactive, and binds ADP when activated (Ueda et al., 2006). However, while the ADP/ATP

binding is reversed, the mechanism of activation remains fundamentally unchanged. It is clear that nucleotide binding, exchange, and hydrolysis are necessary for competent NLR function.

The NB-ARC is critical for switching the plant NLR between active and inactive states, but it is not the domain which is responsible for signal transduction. That role belongs to the domains at the N-terminus of the NLR.

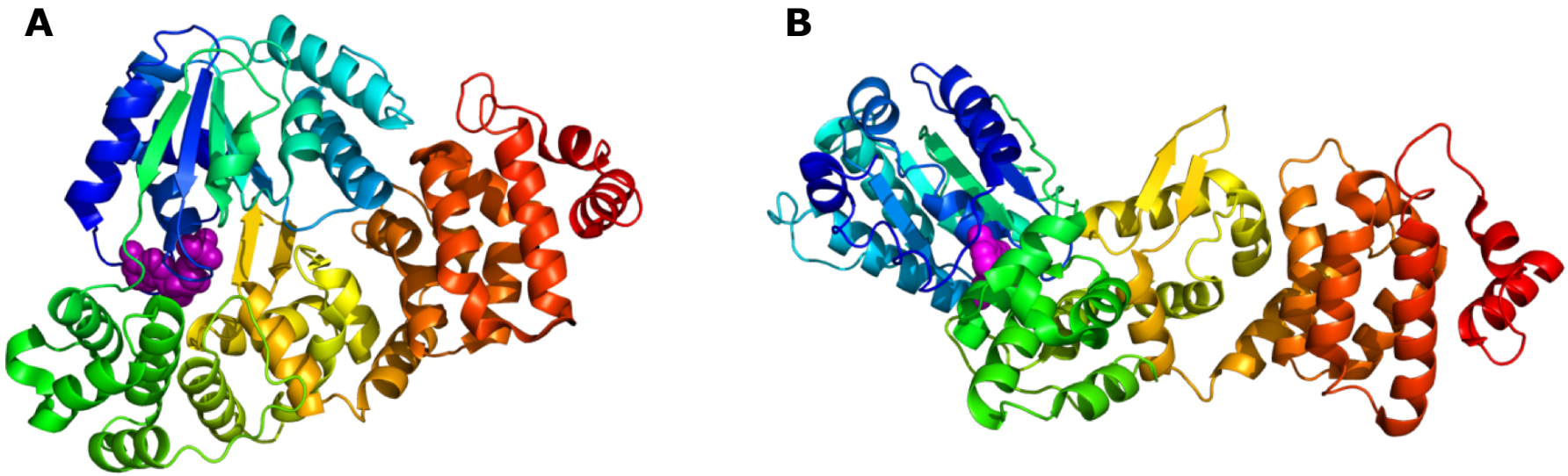


Figure 1.5 The open and closed conformations of the APAF-1 NB-ARC domain represent different states of activity. (**A**) APAF-1 bound to ADP (magenta) in a closed, inactive conformation with the WD1 and WD2 domains in close proximity to the NBD. (PDB ID: 3SFZ) (**B**) APAF-1 bound to ATP (magenta) in an open, active conformation with the WD1 and WD2 domains released from the NBD (PDB ID: 3SHF) (Reubold et al., 2011).

1.5.3 Signalling Through the N-terminal Domain: Coiled-coil and Toll/Interleukin-1 Receptor Domains

The N-terminal domains of plant NLRs separate the two largest classes: coiled-coil (CC) domain-containing (CC-NB-LRRs/CNLs) and Toll/interleukin-1 receptor (TIR) domain-containing (TIR-NB-LRRs/ TNLs) (Lukasik and Takken, 2009). While clearly structurally different, the CC and TIR domains serve the same role within their respective proteins, and are recognised as the signalling domains of plant NLRs (Bernoux et al., 2011, Maekawa et al., 2011). Herein, the current structural and biochemical knowledge of TIR and CC domains will be described, highlighting their importance in NLR signalling.

1.5.4 Current Structural and Biochemical Understanding of Plant TIR Domain Signalling

The TIR domain was characterised as the cytoplasmic domain of the Toll-like and interleukin-1 receptors, which are important for signalling immune responses in mammals (Gay and Keith, 1991). TIR domains can be found in many forms of life, including animals, plants, and bacteria (O'Neill, 2000, Van der Biezen and Jones, 1998, Ve et al., 2015). These domains are typically found associated with multi-domains proteins, as in the case of NLRs; however, TIR-only proteins have also been identified in plants (Meyers et al., 2002, Nishimura et al., 2017). Studies of both plant and mammalian TIR domains have shown their function is linked to self-association and interaction with other TIR domains (Ve et al., 2015). In mammals, multiple interaction interfaces have been identified, and are sufficient to facilitate the formation of large protein complexes (Ve et al., 2017).

There are multiple crystal structures of TIR domains from plant NLRs (Ve et al., 2015). The plant TIR domain exhibits a flavodoxin-

like fold with five α-helices (labelled αA to αE) surrounding a five β-strand β-sheet (labelled βA to βE) (Bernoux et al., 2011b). To date, we have seen the TIR structures from NLRs belonging to many different plant species including L6 (flax); RPS4 and RRS1 (*Arabidopsis*), and RPV1 (Grape) (Bernoux et al., 2011b, Williams et al., 2014, Williams et al., 2016). All the known structures of TIR domains have the same conserved fold, with the exception of RRS1, which has an αD deletion. Furthermore, some studies have shown TIR domains can induce cell death in the absence of an effector, when transiently expressed *in planta* (Krasileva et al., 2010, Bernoux et al., 2011b, Williams et al., 2014).

1.5.4.1 The L6 TIR domain DE interface

All plant TIR domain structures reported in the literature have a key self-association interface that has been validated by mutagenesis (Bernoux et al., 2011b, Williams et al., 2014, Williams et al., 2016). However, that key dimer interface is not consistent throughout all structures. The first plant NLR TIR domain crystal structure was from L6. This domain is capable of signalling autonomously when transient expressed in flax or tobacco, and exhibits a dimerisation interface in its crystal structure that is important for signalling (Figure 1.6). This dimerisation interface consists of interactions between residues of the αD and αE helices, and the interconnecting loops and strands (Bernoux et al., 2011b), and therefore will be referred to as the DE interface. In-solution studies of the L6 TIR domain using multi-angle laser light scattering (MALS) confirmed the ability of the TIR domain for self-association, and it was demonstrated that mutations in residues critical for DE interface formation also prevent cell death signalling when transiently expressed in flax and tobacco. These data provide strong evidence for the necessity of dimer formation for L6 TIR domain signalling.

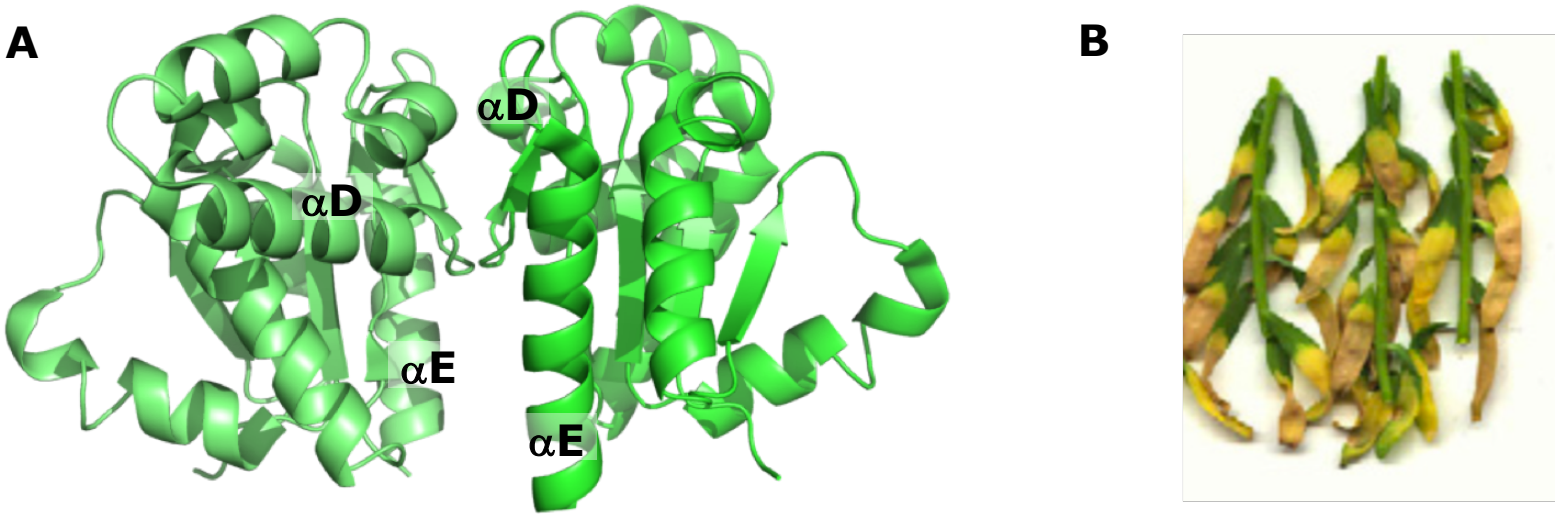


Figure 1.6 The L6 TIR domain forms a homodimer, and is autoactive when expressed in flax. (**A**) The crystal structure of the L6 TIR domain shows dimer formation mediated by interactions between the αD and αE helices (PDB ID: 3OZI). (**B**) Residues 1 – 248 of the L6 TIR domain fused to YFP cause an autoactive phenotype when transiently expressed in flax (Image adapted from Bernoux et. al. (2011)).

1.5.4.2 The RPS4 TIR domain AE interface

A study, subsequent to the determination of the L6 TIR domain structure, characterized a TIR domain heterodimer formed between the RPS4 TIR domain and the RRS1 TIR domain, which revealed a different self-association interface (Williams et al., 2014). This interface consisted of interactions between the αA and αE helices and is deemed the AE interface. This interface is important for the negative regulation of RPS4 TIR domain signalling by the RRS1 TIR domain, but it was also shown to be the interface by which a homodimer of the RPS4 TIR domain was able to signal (Williams et al., 2014). It was demonstrated by mutations to conserved residues in the AE interface, that this surface was important for RPS4 TIR dimer formation in solution. These same mutations prevent cell death signalling by the RPS4 TIR domain when transiently expressed in tobacco (Figure 1.7 A). Furthermore, in the context of the heterodimer, mutations to the AE interface of the RRS1 TIR domain perturb the interaction with the RPS4 TIR domain, and therefore prevent negative regulation of RPS4 TIR domain cell death signalling (Figure 1.7 B). The AE interface has been observed in other plant TIR domains, such as the grape RPV1 TIR domain and *Arabidopsis* TIR-only protein, AtTIR (Williams et al., 2016, Zhang et al., 2017).

Despite the progress made in our understanding of plant TIR domain signalling, there are still some gaps in our knowledge of TIR domain function. The AE interface of the RPS4 TIR domain was not observed in the L6 TIR domain structure, and likewise, the DE interface of the L6 TIR domain was not present in the RPS4 TIR domain structure. With both TIR domains exhibiting the same fold, it begs the question are both interfaces present and required for TIR domain signalling? This question was one of the drivers of the research presented in this thesis.

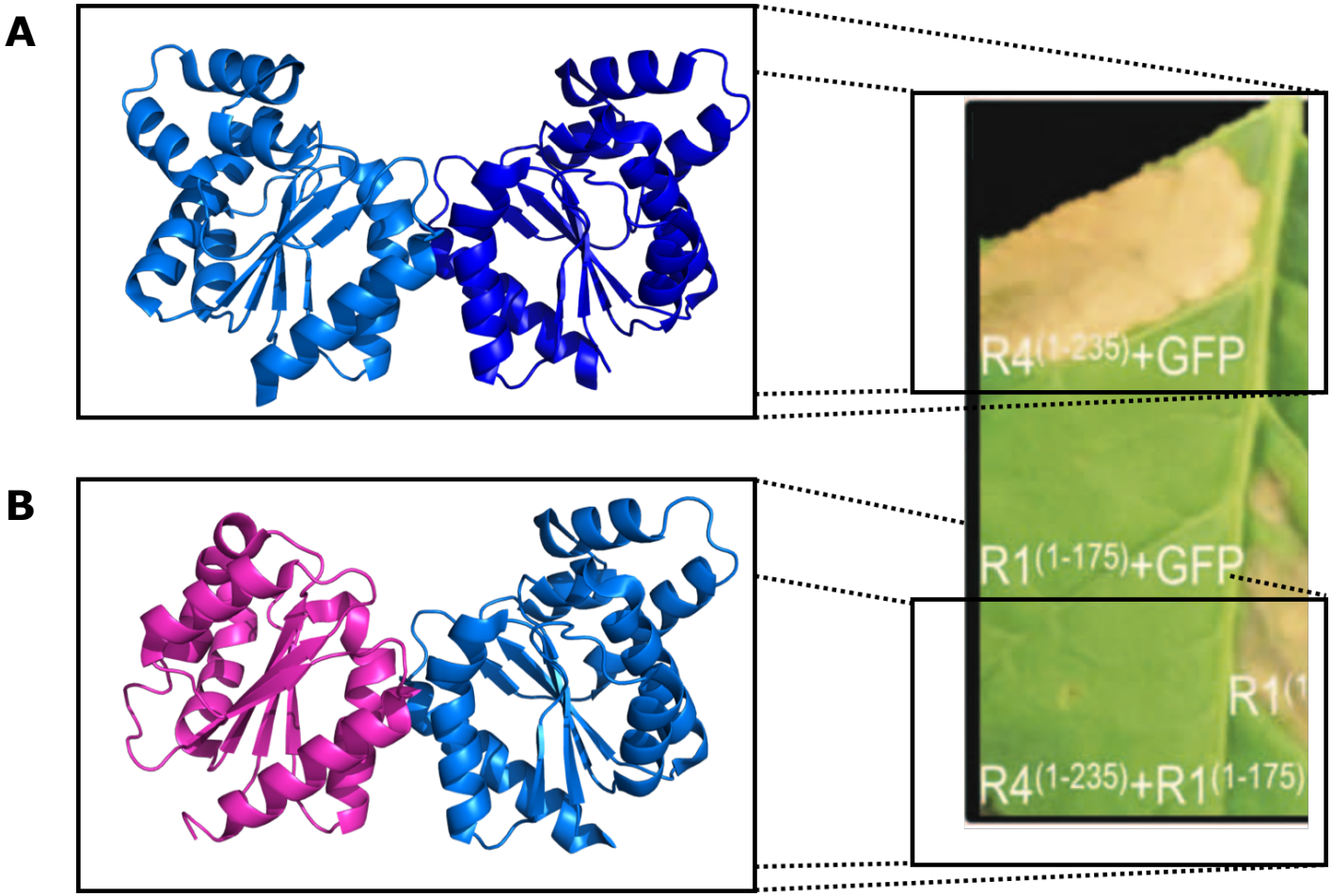


Figure 1.7 The RPS4 TIR domain is autoactive when expressed *in planta*, but negatively regulated by the RRS1 TIR domain. (**A**) The crystal structure of the RPS4 TIR domain homodimer (PDB ID: 4C6R). The first 235 residues of the RPS4 TIR domain (R4), when transiently expressed in tobacco, induce an autoactive phenotype. (**B**) The RRS1 (R1) (purple) and RPS4 (blue) TIR domain heterodimer (PDB ID: 4C6T). Co-expression of the RRS1 TIR domain in conjunction with the RPS4 TIR domain prevents the autoactive signalling of the RPS4 TIR domain. Image adapted from Williams et al. (2014).

1.5.5 Current structural and biochemical understanding of coiled-coil domain signalling

The coiled-coil domain is the other N-terminal domain frequently observed in plant NLRs; however, compared to the TIR domain, little is known about this domain. Early studies of CC-containing NLRs defined the coiled-coil domain to roughly span the first 120 – 200 amino acids (Lupas et al., 1991, Meyers et al., 1999). However, conservation between NLRs in the CC domain region is notoriously low, and initially, CC domains could not be defined by Pfam analysis, resulting in many CNLs being identified through phylogenetic analyses of CNL-specific motifs in the NB-ARC domain (Collins et al., 1998, Meyers et al., 2003).

One of the only conserved motifs found in canonical CC domains is the negatively charged EDVID motif (Collier et al., 2011, Rairdan et al., 2008). The presence of this motif has been documented in CNLs from a variety of different plant species, and has been implicated in intramolecular regulatory interactions with the NB-ARC domain (Bai et al., 2012, Rairdan et al., 2008). As such, CC domains with this motif are typically defined as belonging to the CC_{EDVID} class (Rairdan et al., 2008). However phylogenetic analyses have shown many CNL proteins do not segregate into the same clade as those containing CC_{EDVID} domains, and consequently, are distributed into other classes of CC domains. These other classes include CC domains that share similarity to the cytoplasmic domain of the transmembrane protein RPW8 (CC_R), CC domains that lack an EDVID motif which do not co-segregate with the CC_{EDVID} clade (solely referred to as CC), and CC domains that contain other domain fusions such as the Solanaceae domains or BED DNA-binding domains (CC-SD) (Meyers et al., 2003, Rairdan et al., 2008, Collier et al., 2011).

In many CNLs, CC domain function appears to parallel TIR domain function as an indispensable unit for cell death signalling within the NLR. Transient expression of the first 150 - 200 residues of CC domains of the barley NLR MLA10, and the rice NLR RGA4, has demonstrated CC-dependent cell death signalling; however, this cell death signalling function does not seem to be ubiquitous across all CC domains (Bai et al., 2012, Cesari et al., 2014).

Structural studies of CC domains are far more limited than those of the TIR domains. At the start of this work, only two studies provided crystal structures of plant CC domains, these being the barley NLR MLA10, and the potato NLR Rx (Maekawa et al., 2011a, Hao et al., 2013). Both these proteins belong to the CC_{EDVID} class of CNLs, with an EDVID motif that facilitates interactions with the NB-ARC domain (Rairdan et al., 2008, Bai et al., 2012). However, their sequence identity is exceptionally low (~14%) and the two structures obtained from these studies are significantly different. Furthermore, the first 160 residues of the MLA10 CC domain have been shown to be autoactive when transiently *in planta*, whereas the Rx CC domain does not display an autoactive phenotype (Rairdan et al., 2008, Maekawa et al., 2011a, Bai et al., 2012).

1.5.5.1 The MLA10 CC domain forms an obligate homodimer

The studies of the MLA10 CC domain resulted in the first structure of a plant CC domain. Containing the first 120 residues of the MLA10 protein, the MLA10 CC domain structure forms an obligate homodimer consisting of two monomers with a helix-loop-helix fold (Maekawa et al., 2011) (Figure 1.8 A). With the observation of the MLA10 CC homodimer in the crystal structure, the authors proposed self-association was important for MLA10 CC domain function. The ability of the MLA10 CC domain to self-associate was investigated with a number of biophysical assays, including analytical size-exclusion chromatography (SEC), a

bis(sulfosuccinimidyl)suberate (BS3) chemical crosslinking assay, yeast 2-hybrid analysis, and co-immunoprecipitation, however the latter two experiments were never performed with the CC domain alone (Maekawa et al., 2011).

There has been no *in planta* studies of the MLA10 CC domain consisting of residues 1 – 120, only of the region corresponding to residues 1- 160, so it is not known if the first 120 residues are sufficient for signalling (Maekawa et al., 2011a, Bai et al., 2012). However, mutations that were proposed to disrupt the $MLA10^{1-120}$ CC domain homodimer were shown to disrupt cell death signalling when introduced into the $MLA10^{1-160}$ CC construct (Maekawa et al., 2011a). While this provides some evidence to the proposed dimerisation event as important for MLA10 function, a minimal functional unit that can signal *in planta* and self-associate remains unknown. This represents another aim of the work presented here.

1.5.5.2 The CC domain of Rx forms a heteromeric assembly with its cofactor RanGAP2

The CC domain of the Rx protein does not homodimerise in solution nor does it trigger cell death when transiently expressed *in planta* (Moffett et al., 2002, Rairdan et al., 2008). In the case of Rx, it was observed that transient expression of only the NB-ARC domain in tobacco could initiate cell death (Moffett et al., 2002). The CC domain of Rx (residues 1 – 122; the equivalent length of the MLA10 CC structure) was crystallised in complex with the WPP domain of its cofactor, RanGAP2 (Hao et al., 2013) (Figure 1.8 B). Size-exclusion chromatography of the Rx CC showed the protein is monomeric in solution; however, crystallisation of the Rx CC could not occur without the presence RanGAP2 WPP domain (Hao et al., 2013). This indicates there may be a need for heterotypic interactions for stability of this protein. The role of the CC domain in Rx cell death signalling is unclear, with some speculation it may rather play a role in effector recognition (Moffett et al., 2002).

Other studies of the Rx-RanGAP2 interaction suggest the association of the Rx CC domain with RanGAP2 facilitates translocation to the nucleus where Rx initiates cell death signalling (Sacco et al., 2007).

The differences observed between the functions of MLA10 and Rx CC domains clearly highlight some gaps in our understanding of the function of the CC domains of plant CNLs. Due to the different observations reported by the authors of these two studies, it is difficult to make generalisations about the function of the CC domain. With the biochemical data between Rx and MLA10 suggesting two separate functions, it may be that the structural diversity between the MLA10 and Rx CC domains may explain their functional differences, regardless of belonging to the same subclass of CNL. It is possible both structures represent functionally active molecules, but these data taken together provides no insight into conserved structural elements of the CC domain. More structural studies on CC domains from different plant NLRs are required to provide more insight into plant CC domain structure and function.

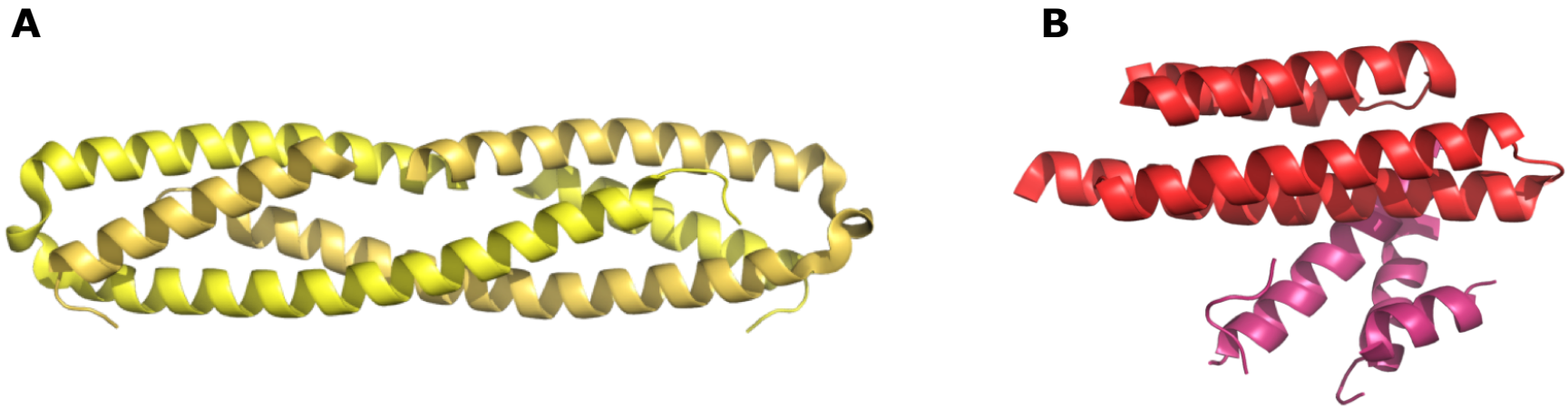


Figure 1.8 The crystal structures of the MLA10 and Rx CC domains. (**A**) The crystal structure of the MLA10 CC domain helix-loop-helix antiparallel homodimer. Monomers are coloured as yellow and gold (PDB ID: 3QFL). (**B**) The crystal structure of the Rx CC domain four-helix bundle (red) in complex with the WPP domain of its cofactor, RanGAP2 (pink) (PDB ID: 4M70).

1.8 Aims: Achieving a better understanding of N-terminal domain signalling in plant NLRs

The research presented in this thesis, in conjunction with that of collaborating researchers, aims to investigate the structure and function of the N-terminal signalling domains of plant NLRs. To this extent, the research presented here focuses on the TIR domain of the Arabidopsis NLR RPP1, and the CC domain of the wheat NLR Sr33. The following sections are devoted to further discussion of these two proteins.

Resistance to *Peronospora Parasitica* 1 (RPP1) is an *Arabidopsis* NLR that provides resistance to the fungal pathogen *Peronospora parasitica* through recognition of the effector, *Arabidopsis thaliana*-related 1 (ATR-1). RPP1 provides an interesting target for study, as different *RPP1* alleles have been identified in different *Arabidopsis* ecotypes. Each of these alleles appear to have different recognition specificities and signalling capacities. In particular, two ecotypes of *Arabidopsis*, Niederzenz and Wassilewskija, express different *RPP1* alleles which differ by only 17 polymorphisms in the TIR domain, however have significantly different signalling phenotypes. Research presented here aims to structurally identify the causes for these differences in phenotype through biophysical analyses and the determination of the crystal structure of the RPP1 TIR domain.

Sr33 is a CNL from wheat that provides resistance to the devastating fungal pathogen, wheat stem rust (*Puccini graminis sp. tritici*). In particular, wheat lines containing the *Sr33* gene show intermediate resistance to one of the most destructive strains of wheat stem rust, UG99, which has been shown to be virulent on 90% of commercial wheat cultivars world-wide. Interestingly, the Sr33 protein is an ortholog of the barley CNL, MLA10. These two

proteins both belong to the CC_{EDVID} subclass of CNL, and share approximately 80% sequence identity in the CC domain.

The valuable resistance Sr33 provides to wheat steam rust makes this protein a research target of significant fundamental and agricultural interest. As there is yet to be any structural characterisation of wheat NLRs, studies of Sr33 will also provide a unique opportunity to begin to understand the structural basis of resistance in wheat. Furthermore, with the lack of knowledge surrounding CC domain function, research into the CC domain of Sr33 provides an opportunity to assess the structure and function of another CC domain to expand our understanding of how these domains contribute to disease resistance.

2. MULTIPLE INTERFACES ARE SIMULTANEOUSLY REQUIRED FOR TIR DOMAIN SELF-ASSOCIATION AND SIGNALLING

The work presented in this chapter has been published as two publications in the journals PLoS Pathogens, and Proceedings of the National Academy of Sciences. The contributions I made to these publications were recognised as second author, and joint first author, respectively. These papers have been referenced extensively within this chapter and have been included as Appendix 1 and Appendix 2. It may benefit the reader to read these papers prior to reading this chapter. All the research presented in this chapter represents my own work, unless otherwise stated.

2.1 Introduction to plant TIR domains

2.1.1 TIR domains self-associate to function

As previously discussed, Toll/interleukin-1 receptor (TIR) signalling domains are involved in innate immune signalling in both plants and animals. The first structure of a TIR domain from a plant NLR was from the flax protein L6 (Bernoux et al., 2011b). In this study, the authors demonstrated that the N-terminal region of the L6 protein (residues 1 – 233), was able to facilitate defence signalling, independent of effector interaction, when transiently expressed in both flax and tobacco. The L6 TIR domain structure was determined through the recombinant expression, and subsequent crystallisation, of residues 23 – 229 from the N-terminal region of the L6 protein. The TIR domain of L6 is contained in residues 59 – 228 (herein referred to as L6 TIR), with residues 29 – 58 at the N-terminus unresolved in the electron density due to disorder or flexibility in the crystal. The L6 TIR maintains a globular, flavodoxin-like fold, with a central five-stranded parallel β sheet surrounded by five α helices. The helices were labelled αA –αE and

the β strands were labelled βA - βE. These names are recognised as standard nomenclature in the literature for describing secondary structural elements within TIR domains, and will be used throughout this thesis.

Structural analyses of the L6 TIR domain revealed a dimerisation interface, termed the DE interface, which consisted of interactions between the αD and αE helices, the βE strand, and the interconnecting loops (Figure 1.6). The DE interface consists of a hydrophobic core surrounded by electrostatic interactions between charged residues. It was shown that disruption of this dimerisation interface through mutagenesis not only perturbed the signalling of the TIR domain, but also prevented signalling in the context of the full length protein in the presence of the flax rust effector AvrL567, when these proteins were transiently expressed in flax (Bernoux et al., 2011b).

Another study investigating the *Arabidopsis* NLRs, RPS4 and RRS1, determined multiple crystal structures of heteromeric and homomeric TIR domain assemblies. The crystal structures of the RPS4 TIR domain homodimer and RPS4 TIR: RRS1 TIR heterodimer revealed a dimerisation interface formed through interactions between residues of the αA and αE helices (termed the AE interface) (Williams et al., 2014). This study categorised an alternative dimerisation interface to the L6 TIR domain DE interface, which was necessary for both RPS4 TIR domain self-association and *in-planta* signalling, and negative regulation of RPS4 TIR domain signalling through heteromeric interactions with the RRS1 TIR domain. This interface is formed by a distinct stacking interaction between two conserved histidine residues, surrounded by a hydrogen bonding network. Like the L6 TIR domain, transient expression of the RPS4 TIR domain in tobacco, in the absence of RRS1 and an effector, also facilitates cell death

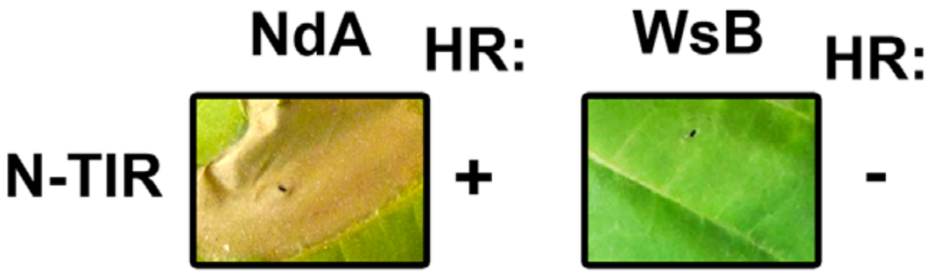
signalling, which can be attenuated by mutations to residues in the RPS4 AE interface. Furthermore, disruption of residues in the AE interface of either RRS1 or RPS4 disrupted heteromeric assembly and negative regulation of RPS4 (Williams et al., 2014).

It is worth noting that many TIR domains that can elicit an autoactive phenotype *in planta* require extra residues that lie outside the structurally defined boundaries of the plant TIR domain. Again, the best examples of this come from the studies of the L6 and RPS4 TIR domains (Bernoux et al., 2011b, Williams et al., 2014). L6 contains ~60 residues at the N-terminus of the TIR domain that are required for autoactive signalling (Unpublished; Simon Williams, personal communication), and RPS4 TIR domain requires additional residues ~60 residues at the C-terminus to facilitate cell death signalling phenotypes in tobacco (Williams et al., 2014). In some cases, the roles of these N-/C-terminal extensions are unclear, however, for L6 the N-terminus is known to be involved in localisation to the Golgi apparatus (Takemoto et al., 2012). It is speculated that this localisation facilitates a type of proximity-induced association, however, this is yet to be published (personal communication, Maud Bernoux). While the studies of Bernoux et. al. (2011) and Williams et. al. (2014) provide very similar structures with similar mechanisms of signalling, they also proposed two distinct interfaces through which signalling is relayed. This presents a number of questions: do different TIRs signal through different protein-protein interaction interfaces? Are there common TIR-domain interaction interfaces? Are the interfaces involved dependent on the mode of recognition, i.e. single NLR protein function (as is the case for L6 NLR) vs dual NLR signalling (as is the case for NLRs, RPS4 and RRS1). Clearly, further investigation of other plant NLRs is required.

2.1.2 Different RPP1 ecotypes present different resistance specificities

The *Arabidopsis* TIR-containing NLR protein, Resistance to *Peronospora parasitica* 1 (RPP1), has evolved to detect the *P. parasitica* effector, *Arabidopsis thaliana* Recognised-1 (ATR-1) (Rehmany et al., 2005, Krasileva et al., 2010). It has been shown allelic variants of *RPP1* from a variety *Arabidopsis* ecotypes have different recognition specificities to different *ATR-1* alleles. *Arabidopsis* ecotypes Wassilewskija (WsB) and Niederzenz (NdA) are two such examples of *Arabidopsis,* which both contain alleles of *RPP1* that differ in *ATR-1* allele recognition, and also differ slightly in gene sequence (Steinbrenner et al., 2015). All *RPP1* alleles encode an N-terminal region with high sequence variability preceding the canonical TIR domain consisting of approximately 80 - 90 amino acids. The protein sequences of RPP1 NdA TIR (1 – 254) and WsB TIR (1 – 248) vary considerably in the N-terminal region (residues 1 – 89 in NdA, and residues 1 – 83 in WsB). However, they differ by only 17 polymorphisms in their respective canonical TIR domains (residues 90 – 254 in NdA and residues 84 – 248 in WsB; Figure 2.1 B). When challenged with a cognate effector, both RPP1 NdA and RPP1 WsB NLRs, subsequently referred to as NdA and WsB, are capable of signalling cell death. However, the transient expression of only the TIR domains of NdA and WsB in tobacco yields significantly different results. As seen with transient expression of L6 and RPS4 TIR domains, the transient expression of the NdA TIR domain with N-terminal extension (1 – 254) in tobacco results in an effector-independent cell death phenotype, whereas infiltration into the same leaf with the equivalent WsB TIR domain construct (1 – 248) elicits no cell death response (Figure 2.1 A, Appendix 1 Figure 3 A).

A



B

```
NdA MGSAMSLSCSKRKATSQDLDSESCKRRKTCSTNDAENCRFIPDESSWSLCANRVISVALT 60
WsB MGSAMSLGCSKRKATNQDVDSESRKRRKICSTNDAENCRFIQDESSWKH------PWSLC 54
    *******.*******.**:**** **** ************ *****.         :*

NdA KFRFQQDNQESNSSSLSLPSPATSVSRNWKHDVFPSFHGADVRRTFLSHILESFRRKGID 120
WsB ANSVVNDTKDTKSSALSLPSPPTSVSRIWKHQVFPSFHGADVRKTILSHILESFRRKGID 114
       . :*.::::**:****** ***** ***:***********:*:**************

NdA TFIDNNIERSKSIGPELKEAIKGSKIAIVLLSRKYASSSWCLDELAEIMICREVLGQIVM 180
WsB PFIDNNIERSKSIGHELKEAIKGSKIAIVLLSKNYASSSWCLDELAEIMKCRELLGQIVM 174
     ************* *****************::*************** ***:******

NdA TIFYEVDPTDIKKQTGEFGKAFTKTCRGKPKEQVERWRKALEDVATIAGYHSHKWCDEAE 240
WsB TIFYEVDPTDIKKQTGEFGKAFTKTCKGKTKEYVERWRKALEDVATIAGYHSHKWRNEAD 234
    ***************************:** ** ********************** :**:

NdA MIEKISTDVSNMLD 254
WsB MIEKIATDVSNMLN 248
    *****:*******:
```

Figure 2.1 Polymorphisms between TIR domains of RPP1 NdA and WsB result in different signalling phenotypes. (**A**) Transient expression of the NdA (1 – 254) and WsB (1 – 248) TIR domains in *N. tabacum* performed by Dr. Karl Schreiber, UC Berkeley. NdA TIR domain presents an autoactive phenotype when transiently expressed in *N. tabacum*, whereas the WsB TIR domain does not. (**B**) Sequence alignment of the RPP1 NdA and WsB TIR domains. Residues that affect protein function and self-association in solution when mutated are coloured in red. Black brackets represent the Pfam (NCBI) TIR domain. Leaf images adapted from Schreiber et al. (2016)

2.2 Aims and Hypotheses

This chapter aims to understand TIR domain signalling in plant NLRs by investigating the role of self-association in the Arabidopsis NLR protein RPP1, utilising structural and biophysical techniques. The hypothesis of this work is that TIR domain signalling is directly related to self-association, and therefore the TIR domain of RPP1 NdA is capable of self-association, while the RPP1 WsB TIR domain is not. Furthermore, the N-terminal extension facilitates TIR domain oligomerisation, and mutants that disrupt signalling will disrupt self-association.

To address these hypotheses, this chapter has the specific aims:

1. Identify the oligomeric state of the RPP1 TIR domains from *Arabidopsis* ecotypes NdA and WsB.
2. Investigate the role of the N-terminal extension in RPP1 TIR domain self-association.
3. Understand the effects of loss-of-function and gain-of-function TIR domain mutants in terms of protein oligomerisation.
4. Solve the structure of the RPP1 TIR domain and characterise residues that are involved in mediating crystal contacts to understand if they have a functional role in mediating RPP1 TIR domain self-association.

2.3 Results

2.3.1 The self-association of the RPP1 TIR domain is transient in-solution, but critical for signalling *in planta*.

2.3.1.1 The N-terminal region of RPP1 facilitates self-association

The hypothesis for this research is that the N-terminal region of the RPP1 TIR domain assists the self-association of the canonical TIR domain. To test the ability of the TIR domain to self-associate, NdA TIR and WsB TIR domains with the N-terminal extensions (NdA^{1-254} and WsB^{1-248}), and NdA TIR and WsB TIR domains without N-terminal extension (NdA^{90-254} and WsB^{84-248}) were expressed and purified, according to methods previously defined by Dr. Karl Schreiber, UC Berkley (published in Schreiber et al. (2016)). After successfully producing these proteins (Figure 2.2; Appendix 1, Materials and Methods), the oligomeric state of the proteins was estimated via measuring the molecular mass of the proteins by size-exclusion chromatography coupled with multi-angle laser light scattering (SEC-MALS).

The SEC-MALS data demonstrate that all the RPP1 proteins tested are capable of self-association, albeit none were measured at the predicted dimeric masses for these proteins. Instead, a range of molecular masses between monomer and dimer were observed, indicating a polydisperse population of molecules. This observation suggests that these proteins are undergoing rapid, transient interactions in solution. Using this method, NdA^{1-254} demonstrated a higher propensity to self-associate than NdA^{90-254}, with average measured mass of 45 kDa (56% larger than the predicted monomeric mass of 28.8 kDa) compared to 25.5 kDa (33.5% larger than the predicted monomeric mass of 19.1 kDa). Similarly, WsB^{1-248} has an enhanced ability to self-associate compared to WsB^{84-248} (20.8% vs 11.6% larger than predicted monomeric mass) (Figure 2.3). Taken together, these data indicate the N-

terminal extension of the RPP1 TIR plays a role in the self-association of the TIR domain. Interestingly, *in planta* data performed by Dr. Karl Schreiber of UC Berkley, demonstrate the N-terminal region of NdA TIR domain is required for cell death activation (Figure 2.1; Appendix 1 Figure 1B, C & S1). Using agrobacterium-mediated transformation, Dr. Schreiber transiently expressed a series of N-terminal RPP1 NdA TIR domain truncations in *N. tabacum* to assess the effect on the autoactivity of the RPP1 NdA TIR domain. Transient expression of NdA^{30-254} and NdA^{90-254} did not present an autoactive phenotype, unlike the wild-type NdA^{1-254}, indicating the N-terminal region of RPP1 NdA TIR domain is required for sufficient signalling. Taken together, the SEC-MALS and *in planta* data suggest that the N-terminal extension of the RPP1 NdA TIR domain is facilitating self-association, which is required for activation of cell-death pathways when transiently expressed in *N. tabacum*.

The SEC-MALS data also demonstrated that NdA TIR domain proteins maintained higher levels of self-association compared to their WsB counterparts (Figure 2.3; Table 2.1). Both NdA^{1-254} and NdA^{90-254} presented measured molecular masses ~22% higher than WsB^{1-248} and WsB^{84-248}. These data indicate that the canonical TIR domain of RPP1 also has a role self-association, and that differences in self-association between NdA and WsB cannot be solely attributed to the variable N-terminal extension.

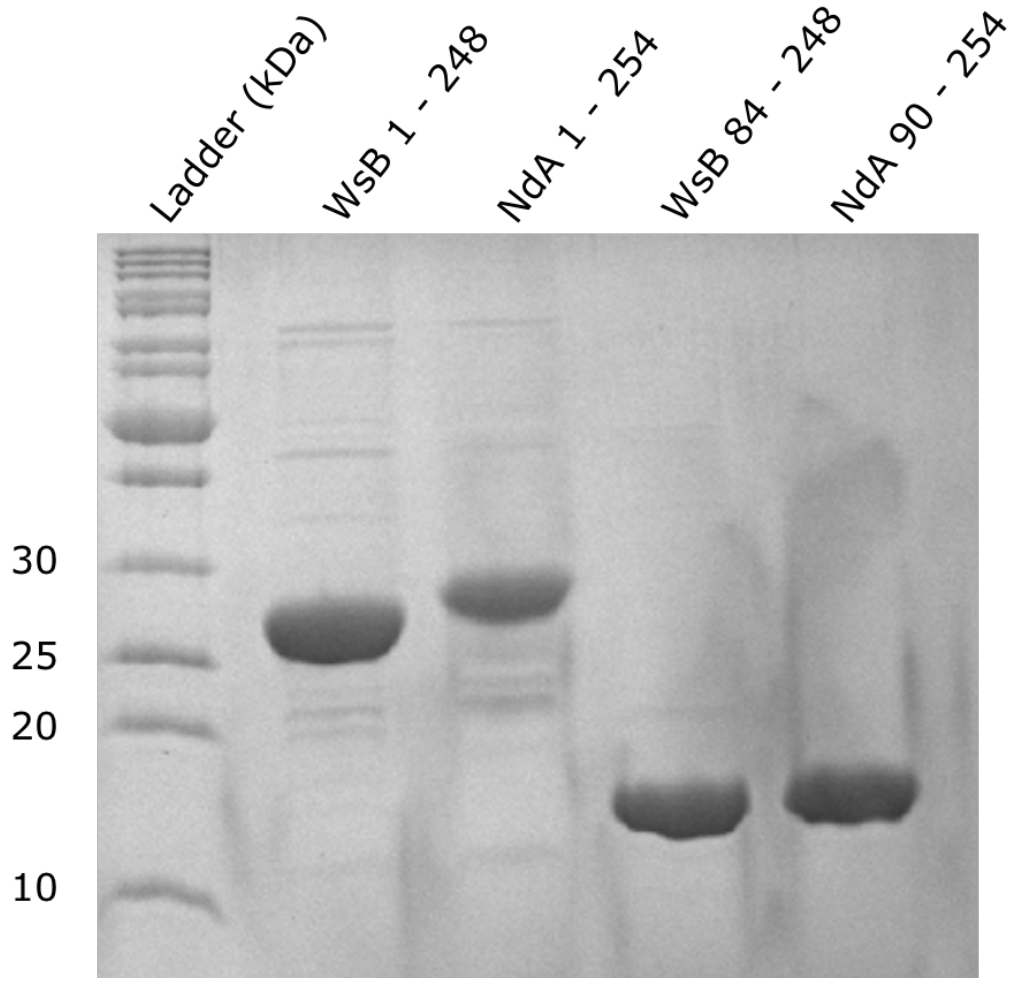


Figure 2.2 Coomassie blue stained SDS-PAGE of purified RPP1 proteins. From the left, RPP1 TIR domains with N-terminal extensions WsB^{1-248} and NdA^{1-254}, and RPP1 TIR domain only proteins WsB^{84-284} and NdA^{90-254}.

2.3.1.2 Mutations in the AE and DE interfaces of RPP1 NdA and WsB TIR domains affect self-association.

Dr. Schreiber, UC Berkeley, designed point mutations in the RPP1 NdA and WsB TIR domains based on polymorphisms between the two TIR domains, and homology modelling using the crystal structure of the L6 TIR domain. NdA^{1-254} alanine mutants, $NdA^{R104A+F106A}$ and $NdA^{G229A+Y230A}$, that corresponded to the AE interface (first identified in RPS4/RRS1 TIR domains) and DE interface (first identified in the L6 TIR domain), respectively, abolished *in planta* cell death responses when transiently expressed *in N. tabacum* (Figure 2.4 B; Appendix 1). Conversely, the WsB mutants that mimicked NdA polymorphic residues in the AE ($WsB^{K98R+I100F}$) and DE (WsB^{R230C}) interfaces induced a cell death signalling phenotype, in contrast to the previously inactive wild-type WsB TIR domain (Figure 2.4 D; Appendix 1).

Work presented here for examination was to test the ability of these RPP1 mutants to self-associate in solution. To do this, $NdA^{R104A+F106A}$, $NdA^{G229A+Y230A}$, $WsB^{K98R+I100F}$ and WsB^{R230C} proteins were expressed and purified, and analysed using SEC-MALS in the same conditions as described in section 2.3.1.1 (Figure 2.4 A & C).

As seen with the initial SEC-MALS analysis of the wild-type RPP1 TIR domains, none of the mutants had a molecular mass equivalent to that of the theoretical monomer or dimer masses, but rather eluted from the column as a polydisperse population of molecules. However, NdA^{1-254} TIR domain loss-of-function mutants $NdA^{R104A+F106A}$ and $NdA^{G229A+Y230A}$ did not demonstrate the same propensity to self-associate as the wild-type NdA^{1-254} TIR domain (Figure 2.4 A). Measured molecular masses for these mutants were both ~25% smaller than wild-type NdA^{1-254} (Table 2.1). The gain-of-function WsB^{1-284} TIR domain mutants $WsB^{K98R+I100F}$ and WsB^{R230C} both presented measured molecular masses higher than wild-type WsB TIR domain (Figure 2.4 C). $WsB^{K98R+I100F}$ and

WsB^{R230C} increased the ability of the WsB^{1-248} TIR domain to self-associate, with measured molecular masses of 48.7 and 41.9 kDa respectively (~72% and ~48% larger than predicted monomeric molecular mass; Table 2.1).

These data show a direct correlation between cell death phenotypes of NdA and WsB TIR domain mutants, and the in-solution self-association of the TIR domains. This suggests that residues from both the AE interface (described in the RPS4 TIR) and the DE interface (described in the L6 TIR) are required for the self-association and signalling of the RPP1 TIR domain. This provides the first evidence that two functional self-association interfaces are required for TIR function in a single protein. However, the molecular mechanisms underlying self-association and the assembly of these AE and DE interfaces in the RPP1 TIR domain still remained unclear.

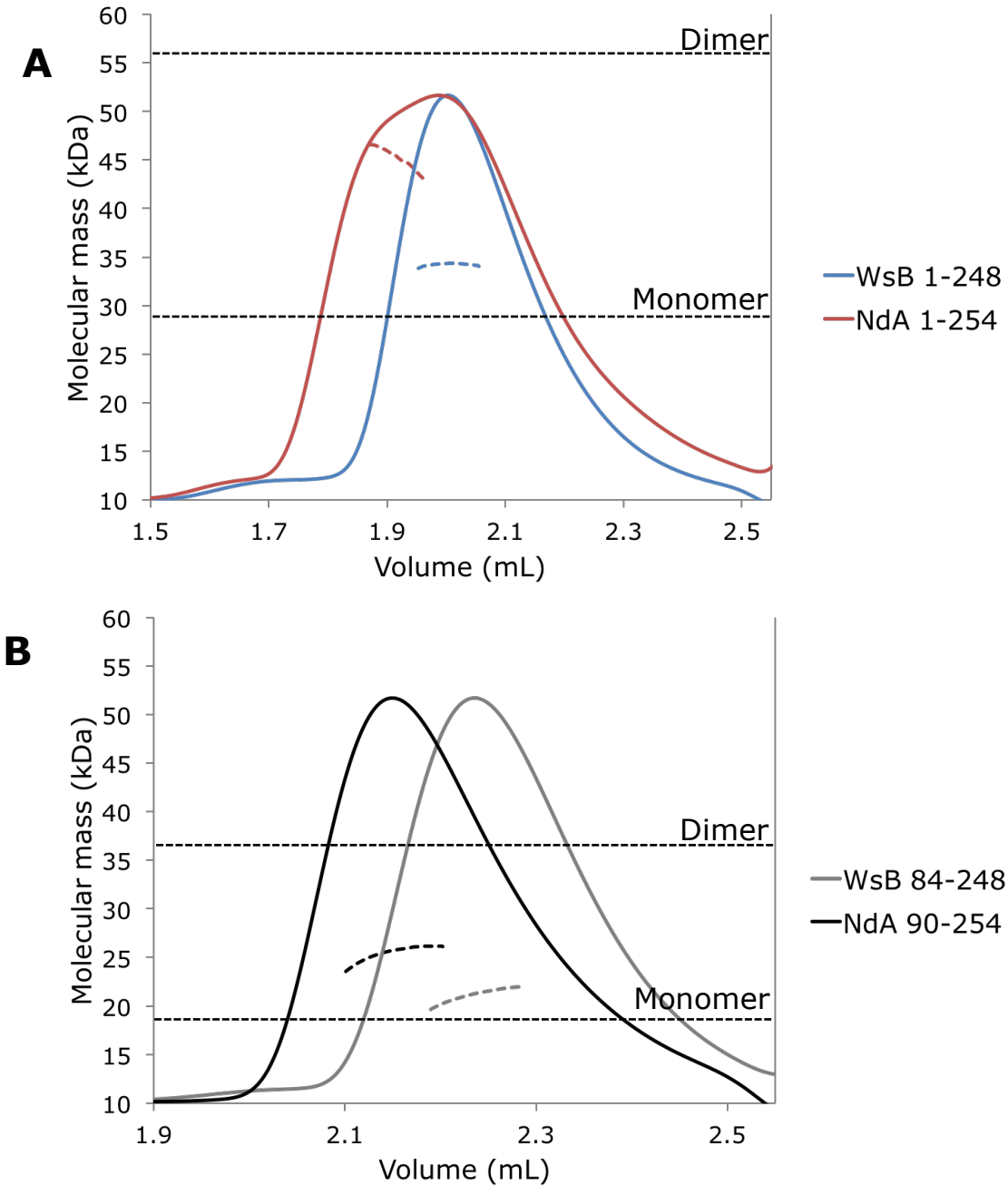


Figure 2.3 Size-exclusion coupled with multi-angle light scattering (SEC-MALS) analysis of RPP1 TIR domain constructs. The elution of the protein from the SEC column is measured as direct refractive index (dRI) and represented by a solid line. Dashed lines with colour corresponding to dRI indicate the molecular mass measured for the protein across the elution peak. Predicted monomeric and dimeric masses are indicated by black dashes lines. (**A**) Comparison of N-terminal extension containing TIR domains NdA$^{1-254}$ and WsB$^{1-248}$ (**B**) Same as (**A**) with the canonical TIR-only NdA$^{90-254}$ and WsB$^{84-248}$.

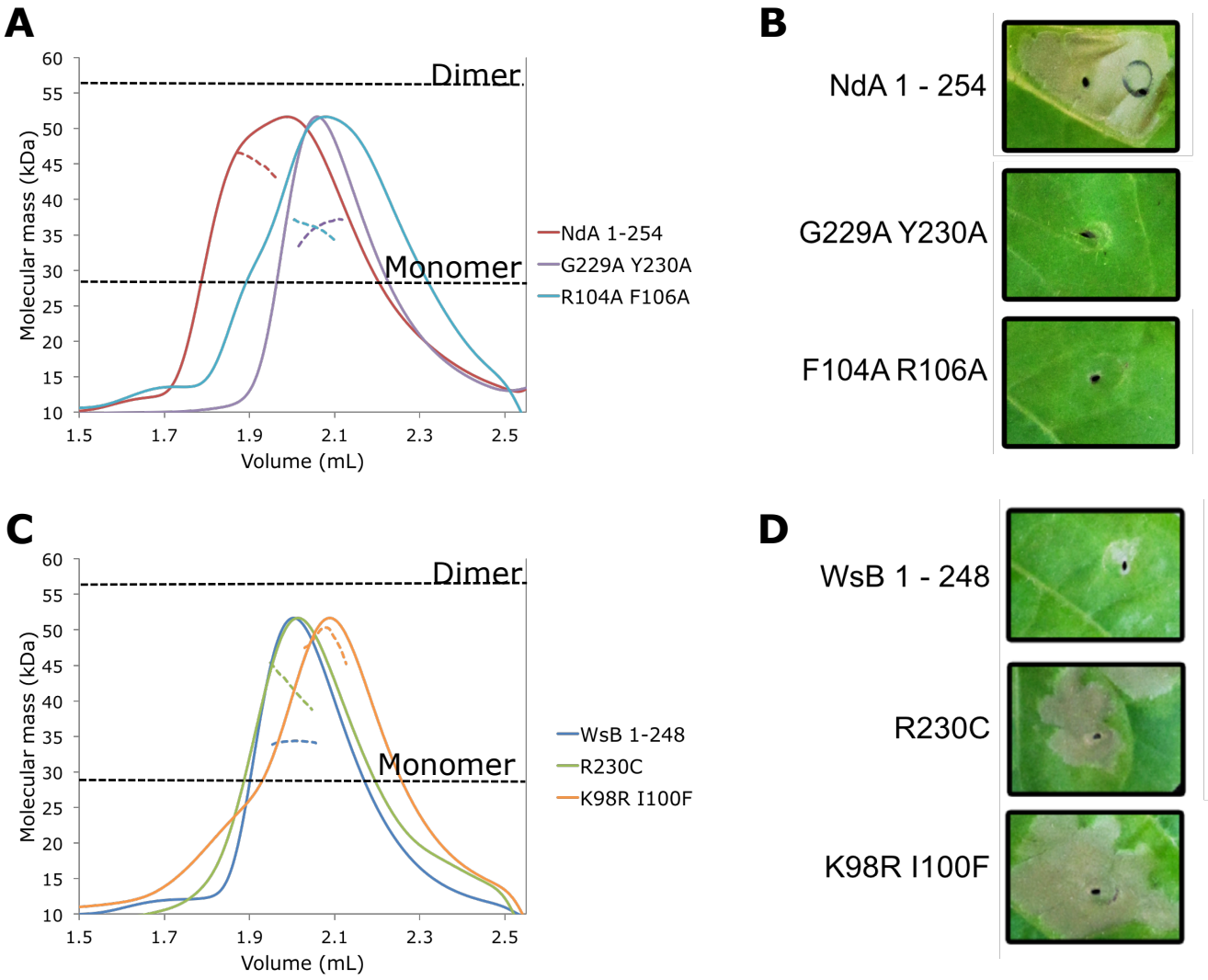


Figure 2.4 *In planta* phenotypes of RPP1 NdA and WsB mutants correlate with changes in ability to self-associate in solution. (**A, C**) SEC-MALS of RPP1 NdA and WsB TIR domain mutants. The elution of the protein from the SEC column is measured as direct refractive index (dRI) and represented by a solid line. Dashed lines with colour corresponding to dRI indicate the molecular mass measured for the protein across the elution peak. Predicted monomeric and dimeric masses are indicated by black dashed lines. (**B, D**) Transient expression of NdA and WsB mutants in *N. tabacum* performed by Dr. Karl Schreiber, adapted from Schreiber et. al. (2016) (**A-B**) RPP1 NdA loss-of-function mutants $NdA^{F104A+R106A}$ and $NdA^{G229A+Y230A}$ have a lower average molecular mass in solution than wild-type RPP1 NdA^{1-254}. (**C-D**) RPP1 WsB gain-of-function mutants WsB^{R230C} and $WsB^{G229+Y230A}$ have a higher average molecular mass in solution than wild-type RPP1 WsB^{1-248}.

Table 2.1 Molecular masses of RPP1 constructs in solution measured via SEC-MALS.

Construct	Theoretical monomer MW (kDa)	Measured MW (kDa)	ΔMW (kDa)	Percentage Difference (%)	Autoactive *in planta* (+/-)
NdA^{1-254}	28.8	45.0	16.2	56.3	+
$NdA^{G229A+Y230A}$	28.8	36.0	7.2	25.0	-
$NdA^{R104A\ F106A}$	28.7	35.9	7.1	24.7	-
WsB^{1-284}	28.3	34.2	5.9	20.8	-
$WsB^{K98R+I100F}$	28.3	48.7	20.4	72.1	+
WsB^{R230C}	28.1	41.9	13.6	48.1	+
NdA^{90-254}	19.1	25.5	6.4	33.5	-
WsB^{84-248}	18.9	21.1	2.2	11.6	-

2.3.2 The crystal structure of RPP1 TIR domain

The SEC-MALS analyses and transient expression assays of RPP1 mutants *in planta* provide evidence for the existence of functional AE and DE interfaces in the RPP1 TIR domain, which have not been witnessed in a single plant TIR domain structure prior to this study. To further investigate these interfaces, structural studies of the purified RPP1 NdA and WsB TIR domain proteins were carried out. The goal of this work was to better understand the molecular mechanisms governing RPP1 TIR domain self-association and subsequent signalling.

2.3.2.1 Crystallisation and data collection of the RPP1 TIR domain

The RPP1 TIR domains were expressed and purified from *E. coli* and subjected to crystallisation trials (Figure 2.5). The yields of NdA^{1-254} and WsB^{1-248} were ~0.25 and ~0.3 mg / L and not amenable to crystallisation trials due to insufficient quantities. However, both NdA^{90-254} and WsB^{84-248} proteins were able to be expressed in large quantities (~5 mg / L) and were subsequently used in crystallisation trials. The NdA^{90-254} protein formed crystals in several conditions from sparse matrix screens (Figure 2.5 A - D), whereas the WsB^{84-248} protein crystals formed in 2.25 M ammonium sulfate, 0.1 M BICINE pH 9. Unfortunately, and despite considerable effort, the conditions could not be optimised to produce crystals suitable for diffraction studies (Figure 2.5 E). For these reasons, I pursued the NdA^{90-254} protein in all subsequent structural studies.

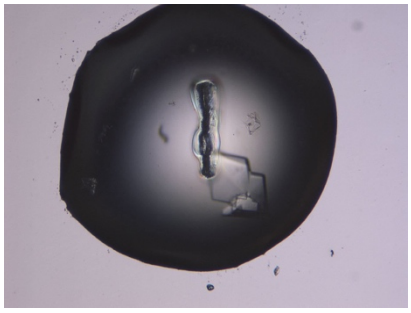
Optimisation of NdA^{90-254} crystals in 0.2 M ammonium citrate tribasic and 20% w/v polyethylene glycol (PEG) 3350 gave rise to robust, cube-shaped crystals. However, these crystals diffracted poorly (<3.5 Å resolution) and the data acquired was highly anisotropic. Further optimisation of these crystals failed to produce

anything that was capable of diffracting to a higher resolution, therefore other crystallisation conditions were investigated. Crystallisation of NdA^{90-254} in 0.1 M sodium citrate pH 5, 20% w/v PEG 6000 produced thick crystals; however multiple crystals grew from a single nucleation point, making them unsuitable for X-ray diffraction (Figure 2.5 B). Optimisation of this condition to 0.2 M citrate pH 5.5, 15% PEG 6000 gave rise to single plate-like rhombus crystals. These crystals were subjected to X-ray diffraction on the MX1 beamline at the Australian Synchrotron, and a dataset to 2.8 Å resolution was obtained.

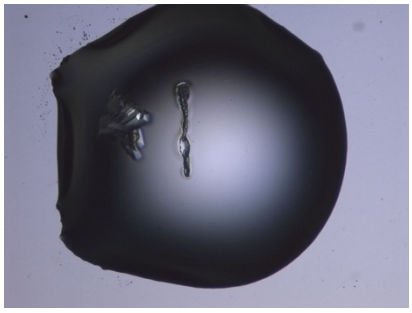
This data allowed for the crystal structure of the NdA^{90-254} TIR (herein referred to as RPP1 TIR) to be solved via molecular replacement using BALBES from the CCP4 program suite (Long et al., 2008). The model was refined with phenix.refine and iMDFF (Adams et al., 2010, Croll and Andersen, 2016). All data analyses and refinement statistics can be found in Table 2.2.

NdA$^{90-254}$

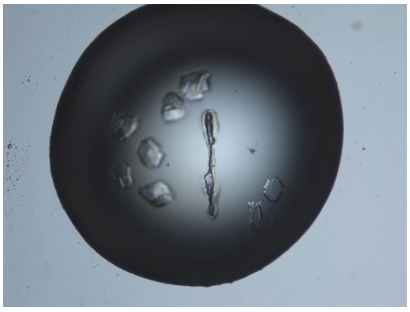
A. 0.2 M ammonium citrate (tribasic), 20% w/v PEG 3350



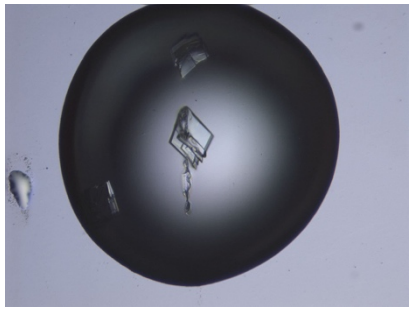
B. 0.1 M citrate pH 5, 20% w/v PEG 6000



C. 0.2 M ammonium sulfate, 20% w/v PEG 3350



D. 8% v/v Tacsimate, 20% w/v PEG 3350



WsB$^{84-248}$

E. 2.25 M ammonium sulfate, 0.1 M BICINE pH 9

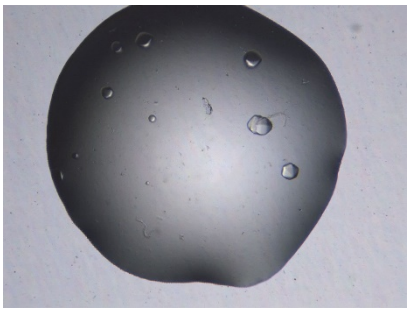


Figure 2.5 Crystal hits for NdA$^{90-254}$ and WsB$^{84-248}$ from sparse matrix screens. All screens were set up with protein concentrations of 10 mg/mL in 10 mM HEPES pH 8, 150 mM NaCl. (A-D) NdA$^{90-254}$ crystal hits. (E) WsB$^{84-284}$ crystal hit. Crystal images are approximately 1000 μM x 1000 μM.

Table 2.2 Crystallographic data for the RPP1^{TIR}

Detector	ADSC Quantum 315r CCD
Wavelength (Å)	0.9537
Temperature (K)	100
Crystal-to-detector distance (mm)	270
Rotation range per image (°)	0.5
Exposure time per image (s)	1
Space group	P 1 $2_1$1
a, *b*, *c* (Å)	81.89, 84.33, 122.75
α, *β*, *γ* (°)	90, 90.1, 90
Average mosaicity (°) [b]	0.31
Resolution range (Å)	42.32-2.58 (2.66-2.58) [a]
Total no. of reflections	394094 (29751)
No. of unique reflections	51965 (4314)
Completeness (%)	99.6 (95.6)
Multiplicity	7.6 (6.9)
Mean $I/\sigma(I)$	14.5 (1.3)
Rmeas (%) [c]	10.7 (184.7)
Rpim (%) [d]	5.4 (96.5)
$CC_{1/2}$ [b]	0.99 (0.59)
Matthews coefficient (Å^3 Da^{-1}) [e]	2.87
Refinement	
Resolution range (Å)	40.94-2.80
R_{work} (%) [f]	21.7
R_{free} (%) [g]	26.8
No. of non-H atoms	
Total	10533
Non-solvent	10533
Water	0
Average isotropic *B* value (Å^2)	79.0
R.m.s.d. from ideal geometry	
Bond lengths (Å)	0.0059
Bond angles (°)	0.989
Ramachadran plot, residues in (%)[h]	
Favoured regions	96.1
Additionally allowed regions	3.8
Outlier regions	0.1

[a] The values in parentheses are for the highest-resolution shell.
[b] Calculated with AIMLESS (Evans and Murshudov, 2013).
[c] Rmeas = $\Sigma hkl\{N(hkl)/[N(hkl)-1]\}1/2\ \Sigma i|Ii(hkl)-\langle I(hkl)\rangle|/\ \Sigma hkl\Sigma iIi(hkl)$, where $Ii(hkl)$ is the intensity of the *i*th measurement of an equivalent reflection with indices *hkl*.
[d] Rpim = $\Sigma hkl\{1/[N(hkl)-1]\}1/2\ \Sigma i|Ii(hkl)-\langle I(hkl)\rangle|/\ \Sigma hkl\Sigma iIi(hkl)$.
[e] Calculated with MATTHEWS_COEF within the CCP4 suite (Winn et al., 2011).
[f] $R_{work} = \Sigma_{hkl}\ ||F_{obs}|-|F_{calc}||/\Sigma_{hkl}\ |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.
[g] R_{free} is equivalent to R_{work} but calculated with reflections (5%) omitted from the refinement process.
[h] Calculated with MolProbity (Davis et al., 2004)

2.3.2.2 The crystal structure of the RPP1 TIR domain exhibits dual interfaces

The RPP1 TIR domain presents the canonical TIR domain flavodoxin-like fold, with a central five-strand β-sheet (designated βA to βE), surrounded by five α-helices (designated αA to αE), with asymmetric unit of the crystal consisting of eight molecules (designated chain A to chain H) (Figure 2.6 A). As previously stated, two functional interfaces have been discovered in plant TIR domains. These interfaces are the AE interface from the RPS4/RRS1 heterodimer, which consists of electrostatic interactions between the αA and αE helices of the two TIR domains, and the DE interface of the L6 TIR homodimer, which consists of interactions between the opposing αD and αE helices, the βE strand, and connecting loops, within the dimer. Analysis of the molecular contacts within the asymmetric unit of the RPP1 TIR domain revealed two interfaces similar to the AE and DE interfaces of the RPS4 and L6 TIR domains, respectively. The AE interface in the RPP1 TIR domain is mediated by interactions between the αA and αE helices of chain A and chain B, and DE interface is formed through interactions between the αD helix, the αE helix, the βE strand, and connecting loops, of chain A and chain C (Figure 2.6 B). This pattern of interaction is repeated with in the asymmetric unit between chain D and chain E, which also form an AE interface, and chain E and chain H, which form a DE interface (Figure 2.6 C).

2.3.2.3 Molecular interactions in the RPP1 AE interface are consistent with previous TIR domain structures

The RPP1 TIR domain AE interface consists of a parallel interface between the αA and αE helices of chains A and B resulting in a buried surface area of ~1232 $Å^2$ (Figure 2.7 A). As seen in the TIR domain structures of RPS4, RRS1, SNC1, RPV1, and the *Arabidopsis* TIR-only protein AtTIR, the interface has the same conserved π-π stacking interaction between the aromatic imidazole rings of two histidine residues ($H109_A$, $H109_B$) in the centre of the

αA helices (Figure 2.7 C). The AE interface is further coordinated by a hydrogen bonding network formed by the residues neighbouring the stacking interaction, $S108_A$ and $S108_B$, interacting with nearby A239 and E238 residues of the αE helices. Additionally, the peripheral acidic R104 residues of the αA helices on both chains interact with basic E243 and D237 residues of the αE helices of the opposing chain, to further stabilise the interaction through the formation of salt bridges (Figure 2.7 B). Sequence analysis of the all current TIR domain structures (L6, RPS4, AtTIR, SNC1, RPP1, and RPV1) shows a clear conservation in the αA helix and the N-terminal end of the αE helix (Figure 2.7 D). This conservation is clearly carried over into the structure of the AE interface, with a structural alignment between all TIR domains with an observed AE interface (RPS4, RRS1, AtTIR, SNC1, and RPP1) demonstrating an exceptionally low deviation in the orientation of the helices and interacting residues (Figure 2.7 C). The structure and sequence conservation of the AE interface combined with *in planta* mutagenesis data highlight the necessity of this interface for TIR domain signalling.

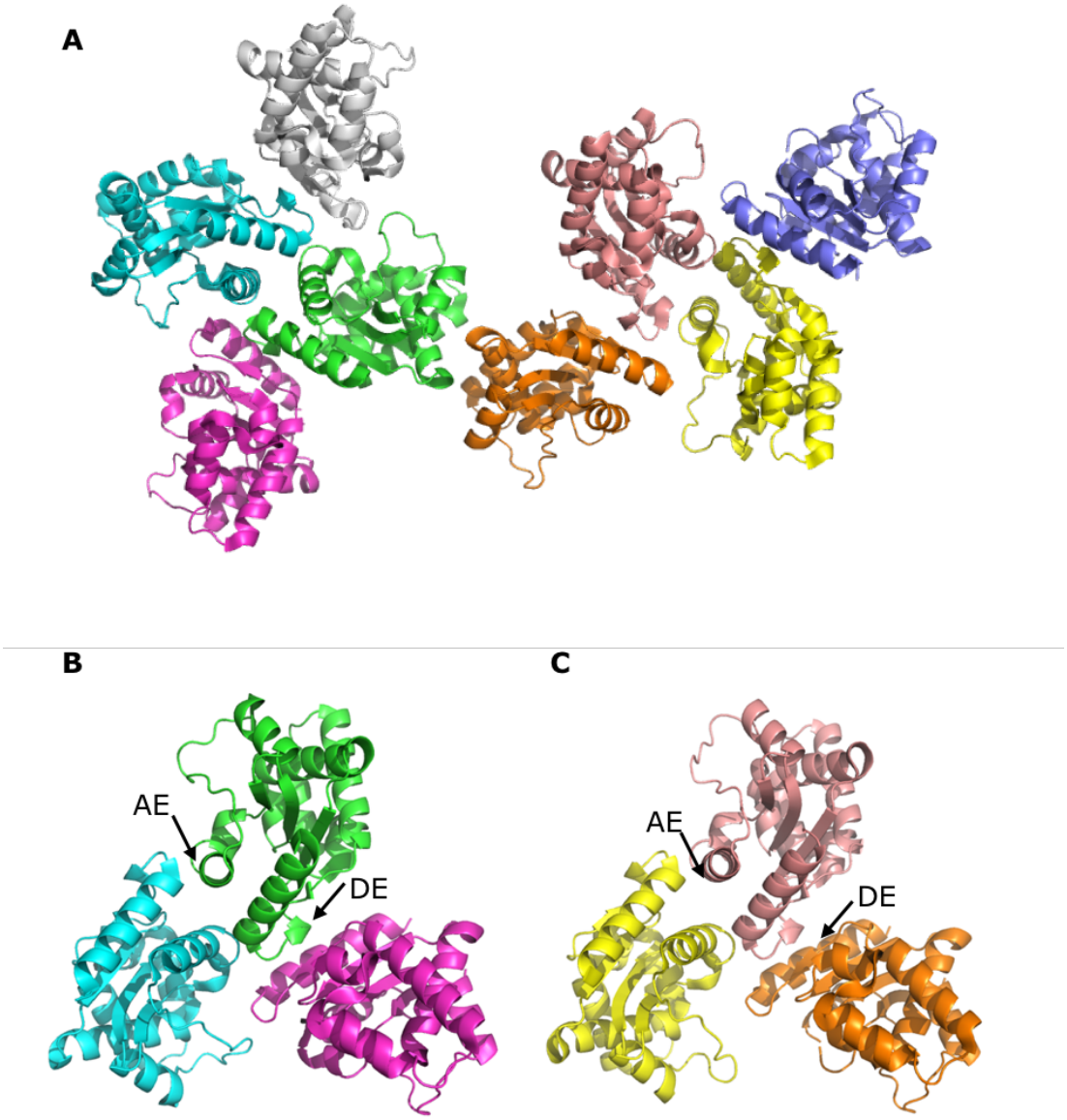


Figure 2.6 The crystal structure of the RPP1 TIR domain at 2.8 Å resolution. (**A**) The asymmetric unit of the crystal contains eight molecules (PDB ID 5TEB). (**B**) Chain A (green) forms an AE interface with chain B (cyan), and a DE interface with chain C (purple). (**C**) As seen in (**B**), but with the AE and DE interfaces formed through interactions between chain D (yellow) and E (pink), and E and H (orange) respectively.

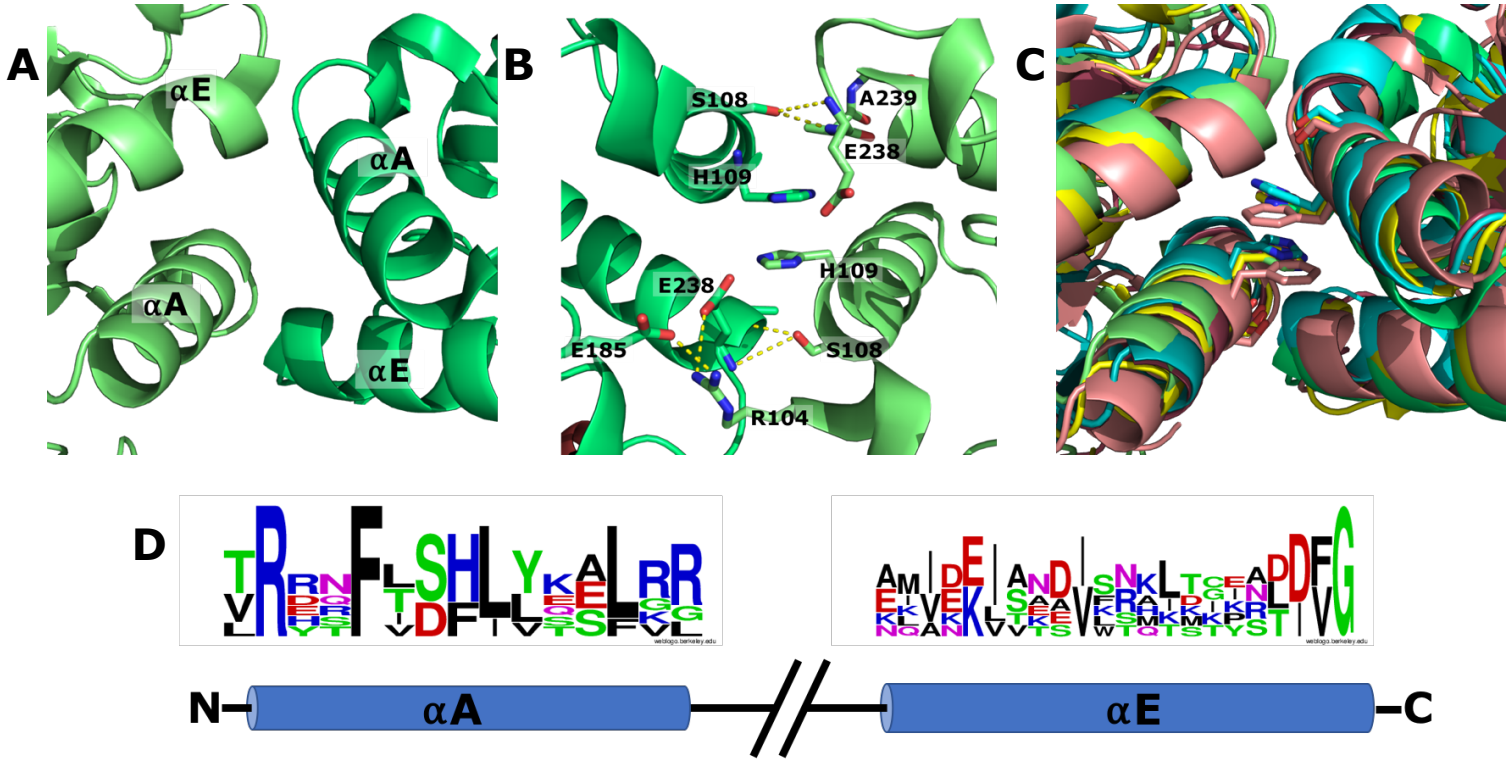


Figure 2.7 The molecular interactions within the AE interface of the RPP1 TIR domain and global conservation of the AE interface in TIR domain structures. (**A**) The AE interface of RPP1 is formed by the interactions between the αA and αE helices with an anti-parallel arrangement. (**B**) The molecular interactions between residues in the AE interface RPP1 TIR domain. The SH motif, first detailed in the RPS4/RRS1 heterodimer AE interface, sits in the centre of the interface. (**C**) Superimposition of all the known crystal structures that present an AE interface shows a high level of structural conservation in the AE interface. The superimposition includes RPP1 TIR (green; PDB ID: 5TEB), AtTIR (pink; PDB ID: 3JRN), RPS4/RRS1 TIR heterodimer (RPS4 in teal, RRS1 in light blue; PDB ID: 4C6T), and the SNC1 TIR domain (yellow; PDB ID: 5TEC). (**D**) Weblogo (Crooks et al., 2004) of the multiple sequence alignment of plant TIR domains with known structures (RPP1, SNC1, RPS4, AtTIR, L6, RPV1; RRS1 omitted due to αD deletion) shows the conservation of residues in the AE interface. Prevalence of a residues in a single position is demonstrated by its size.

2.3.2.2 Molecular interactions in the DE interface of the RPP1 TIR domain.

The DE interface of the RPP1 TIR domain is smaller than the AE interface with a buried surface area of 854 Å^2, mediated by the βE strand, the αD and αE helices, and interconnecting loops, of chains A and C (Figure 2.8 A). The DE interface has a hydrophobic core consisting of multiple non-polar, aliphatic and aromatic residues of the βE strand, forming van der Waals and π-π interactions. The indole side chain of W235$_A$ forms a stacking interaction with the imidazole ring of H231$_C$. Simultaneously, H231$_C$ forms multiple hydrogen bonds with K234$_A$ and C236$_A$. Further down the βE strand, H233$_C$ forms hydrogen bonds with C236 and E240, with the charged amine of H233$_C$ stabilizing the interface through salt bridge formation with the carboxyl group of E240 (Figure 2.8 B). The sequence conservation of the DE interface (as determined by sequence analysis of the L6, SNC1 and RPP1 TIR domain structures which present a DE interface) appears lower than that of the AE interface, with fewer conserved residues mediating interactions in the dimer (Figure 2.8 C).

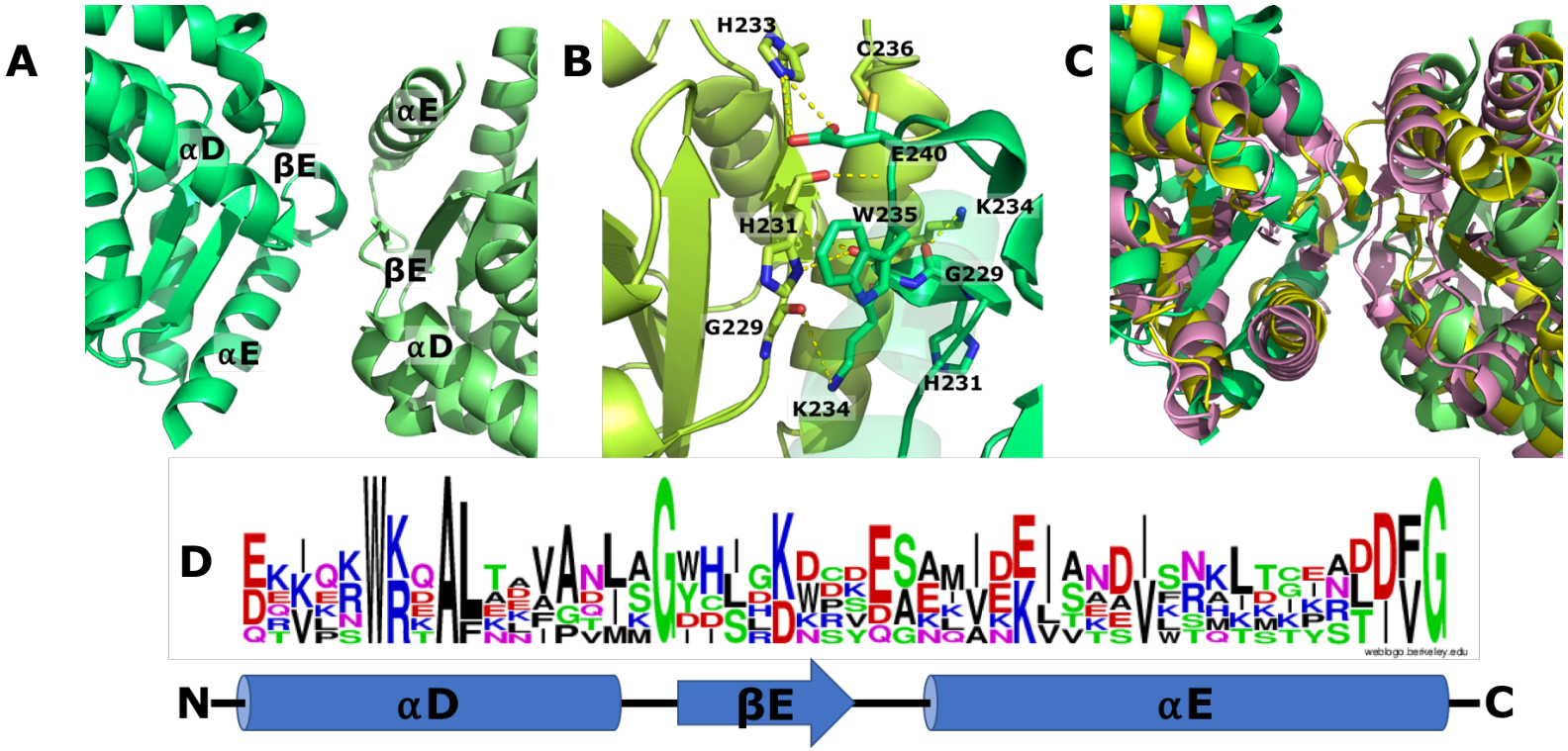


Figure 2.8 The molecular interactions within the DE interface of the RPP1 TIR domain and global conservation of the DE interface in TIR domain structures. (**A**) The DE interface of RPP1 is formed by the interactions between the αD helix, the αE helix, the βE strand and the interconnecting loops. (**B**) The molecular interactions between residues in the DE interface RPP1 TIR domain. The RPP1 DE interface consists of a hydrophobic core surrounded by an electrostatic network of charged residues, similar to that first observed in the L6 TIR domain structure. (**C**) Superimposition of all the known crystal structures that present a DE interface shows a conservation in overall fold; however there are changes in the orientation of the DE interface between RPP1, SNC1, and L6 TIR domains. The alignment includes the RPP1 TIR (green; PDB ID: 5TEB), L6 TIR (pink; PDB ID: 3OZI), and the SNC1 TIR domain (yellow; 5TEC). (**D**) Weblogo (Crooks et al., 2004, Maekawa et al., 2011a) of the multiple sequence alignment of plant TIR domains with a known structure (RPP1, SNC1, RPS4, AtTIR, L6, RPV1; RRS1 omitted due to αD deletion) shows the conservation of residues in the DE interface. Compared to the AE interface, the conservation of specific residues in the DE interface is lower between TIR domains. Prevalence of a residues in a single position is demonstrated by its size.

2.3.3.3 Residues important for cell death signalling are located in the AE and DE interfaces

Joint analysis of the mutational studies (Schreiber et. al. (2016) presented previously in this chapter) and the RPP1 TIR domain structure shows the mutations that correlate with self-association and cell death signalling localise to both the AE and DE interfaces. Notably, R104 and F106 were demonstrated to be essential for cell death signalling and self-association of the TIR *in planta*. In particular, R104 appears to be critical for hydrogen bonding and salt bridge formation in the AE interface; however the imidazole side chain of F106 is not solvent exposed and does not contribute to the formation of the AE interface, but its high conservation across many TIR domains suggests it may play a role in the stabilisation of the structure responsible for that interface (Figure 2.7 B, D). Other residues that are important in self-association and cell death, such as G229, Y230, K234 and C236 (K234 and C236 tested *in planta*, but not tested by MALS (Appendix 1 Figure S4), are all integral to the formation of the DE interface. G229 and Y230 are major components of the core of the DE interface, with K234 and C236 stabilising the interface through hydrogen bonding (Figure 2.8 B).

The DE interface is less well characterised than the AE interface, as previous to this work it was only observed in the L6 crystal structure, unlike the AE interface that can be found in multiple TIR domain structures. Now that we have multiple DE interface-containing structures, we can better characterise conservation in the DE interface. The DE interface as observed in L6, RPP1 and SNC1, is centralised around the βE strand, with a hydrophobic core supported by a conserved glycine (G229 in RPP1; G149 in SNC1; and G201 in L6) and aromatic residue (Y230 in RPP1; Y150 in

SNC1; and W202 in L6), which are supplemented by other neighbouring aromatic and aliphatic residues (Figure 2.8 D). My collaborators, Dr. Xiaoxiao Zhang and Dr. Maud Bernoux, have shown that mutations to these proteins result in interface disruption (Appendix 1 Figure 4 B; Appendix 2 Figure 2 D, Figure 3 B & C, Figure 4 C). Peripheral to the hydrophobic core of the DE interface is a hydrogen-bonding network created by the surrounding charged residues. These residues play an equally important role in stabilising the interface, with charge swap and alanine mutations also resulting in reduced self-association, compromising cell death signalling (Figure 2.4 B & D; Figure 2.8 B; Appendix 1 Figure 4 B; Appendix 2 Figure 2 D, Figure 3 B & C, Figure 4 C). The AE and DE interfaces of TIR domains share common characteristics: a hydrophobic core consisting of interactions between aromatic residues, surrounded by a dense hydrogen-bonding network.

2.3.3.4 The DE interface in the RPP1 TIR domain structure is rotated relative to other TIR domains

One of the most striking features of the DE interface is the low sequence similarity between TIR domain structures when compared to the sequence conservation of the AE interface (Figure 2.9 D). Regardless of this, all the interactions between secondary structural elements required for the formation of the DE interface (i.e. interactions between residues of the αD helix, αE helix, and βE strand) are conserved across all TIR domains with an observed DE interface in their crystal structure (RPP1, L6 and SNC1 TIR domains). However, superimposition of TIR domain dimers presenting a DE interface reveals something intriguing. As seen in Figure 2.8 C, the secondary structural elements of the RPP1, SNC1 and L6 TIR domains forming a DE interface do not align with each other between the three structures. The misalignment of in DE interfaces of L6, SNC1 and RPP1 TIR domains is not due to differences in the fold of the TIR domains, demonstrated by an

alignment of monomers in Figure 2.9 A, rather it appears the orientations of the TIR domains forming the DE interface are different. Using the αE helix of the L6 TIR domain as a reference point, it would appear the DE interface of the RPP1 TIR domain is rotated a significant 97° in comparison to that of the L6 TIR domain (Figure 2.9 B, D). This results in the C-terminal region of the RPP1 TIR domain αE helix turning out further from the DE interface, creating a smaller buried surface area in the RPP1 TIR domain, compared to other TIR domain structures. However, this rotation does not appear to influence the interactions between conserved residues, which maintain the integrity of the DE interface. This rotation of the DE interface is not mirrored in the AE interface, as seen the Figure 2.9 C, as the AE interface of RPP1 TIR domain can be perfectly aligned with that of the RPS4 TIR domain, and other TIR domains as previously shown in Figure 2.7 C. The significance of the differences in orientation of the DE interfaces between TIR domain structures is unclear, but it suggests a level of flexibility in this interface that is not observed in the AE interface.

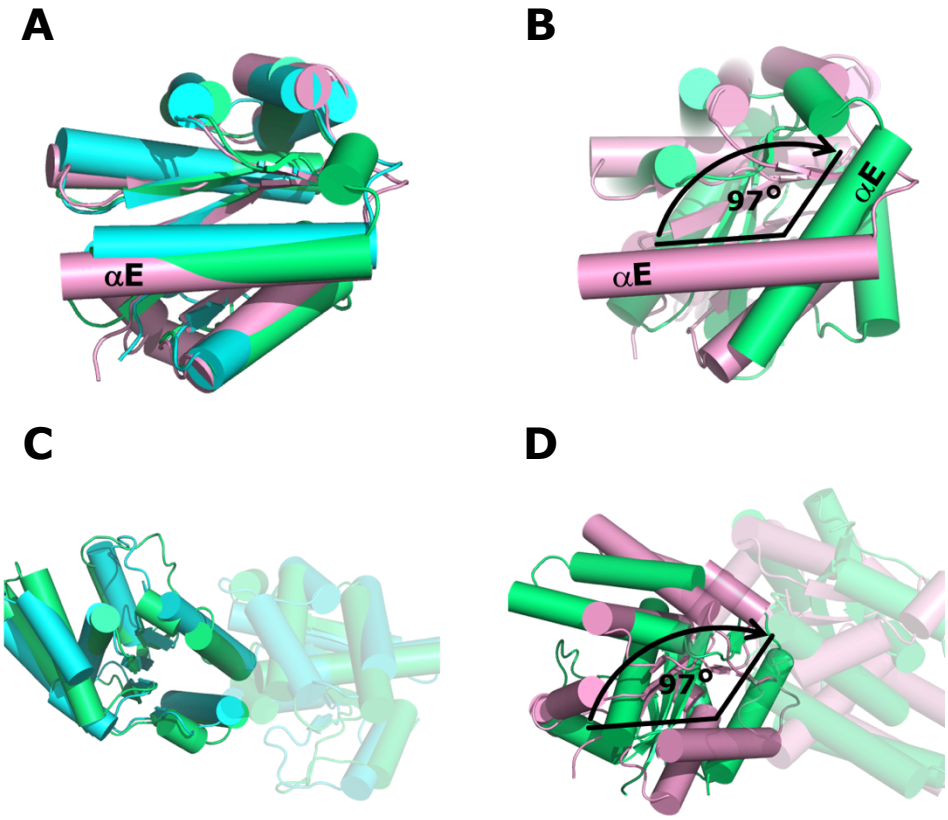


Figure 2.9 The DE interface of the RPP1 TIR domain is rotated relative to the DE interface in the L6 TIR domain structure. (**A**) Superimposition of the RPP1 (green), L6 (pink) and RPS4 (blue) TIR domains shows no difference in overall fold between the three TIR domains (PBD IDs: 5TEB, 3OZI, and 4C6R respectivitely). (**B**) Superimposition of one of the molecules from each of the RPP1 TIR and L6 TIR domain DE interface dimers. Using the αE helix of the L6 TIR domain as a reference point, it appears the RPP1 TIR domain is rotated ~97° relative to the DE interface of L6. (**C**) Superimposition of RPS4 and RPP1 TIR domain AE interface dimers. Very little deviation in the relative orientation of the molecules forming the AE interface is observed between the two proteins. (**D**) Superimposition of RPP1 and L6 TIR domain DE interface dimers. As shown in (**B**), the protein molecules are misaligned after superimposing one molecule of the dimers.

2.4 Discussion

2.4.1 Increased ability of the RPP1 TIR domain to self-associate correlates with cell death signalling

The positive correlation between autoactive RPP1 TIR domain containing proteins and their increased self-association in-solution highlights the importance of self-association in cell death signalling. This data is consistent with the previous studies of the L6 and RPS4 TIR domains (Bernoux et al., 2011a, Williams et al., 2014). Interestingly, the N-terminal extension of the RPP1 TIR domain greatly increased the ability of both NdA and WsB TIR domains to self-associate. Furthermore, when the N-terminal domain was removed, the previously autoactive RPP1 NdA TIR domain was unable to induce cell death (Schreiber et al., 2016). While this N-terminal domain contributes to self-association and cell death, it is still apparent that the canonical TIR domain also plays an important role in RPP1 TIR domain self-association and signalling. This is exemplified by the impact of mutations within the TIR domain, and their effect on self-association and signalling of RPP1 NdA and WsB TIR proteins (Figure 2.4).

It is clear that both the N-terminal extension of the RPP1 TIR domain, and the canonical region of the TIR domain contribute to the self-association of the protein. If either the N-terminal extension, or the canonical TIR domain, are compromised through truncation or mutagenesis, this will result in a reduced ability to self-associate. While at no point was the TIR domain observed to be capable of forming a stable dimer, there is certainly a "self-association threshold" which must be met to facilitate cell death signalling. This was clearly demonstrated by the ability to restore function to the WsB TIR domain through mutations that promoted self-association, and by deactivating the NdA TIR domain through mutations that reduced the capacity of the protein to self-associate.

The transient nature of the interactions of the TIR domains in solution suggests an additional driving force, such as oligomerisation of domains outside of the TIR domain, is required to promote self-association and allow cell-death signalling. This result is consistent with observations of Krasileva et al. (2010), who demonstrated *in planta* autoactivity of the RPP1 WsB TIR domain was dependent on the use of an YFP-tag, a protein known to dimerise. Interestingly, the authors were able to show that the WsB TIR domain was inactive when fused to a mutant of YFP that did not self-associate, as tested by coimmunoprecipitation. At the time of this study, Krasileva et al. stated YFP was facilitating WsB TIR domain dimerisation and therefore promoting signalling, however this was not tested biophysically.

Previous studies of TIR domains have also demonstrated these proteins do not form stable dimers in solution (Bernoux et al., 2016, Williams et al., 2014). Interestingly, in the case of the TIR domain of RPV1, the TIR-alone was capable of causing cell death, however, with several approaches used, self-association was not observed (Williams et al., 2016). In the context of auto-immune signalling, a TIR domain that readily forms a stable dimer presents a danger to the plant cell due to the deadly nature of the signalling response. It has been previously proposed that the transient nature of TIR domain interaction is a safety mechanism to prevent activity in the absence of a pathogen (Bernoux et al., 2011a). In support of this, a recent study demonstrated TIR-derived incompatibility of some NLRs, by which expression of both incompatible NLRs can lead to stunted or failed growth due to hybrid necrosis (Tran et al., 2017).

Due to the transient nature of TIR domain interactions, the presence of "inactive" TIR domains from active NLRs, and the dangers of auto-immune signalling, it is likely TIR domain self-

association is driven by the association of the C-terminal domains of the NLR. This idea has previously been introduced in the study of L6 TIR (Bernoux et al., 2011b). Coined "proximity-induced dimerisation", the activation of a plant NLR by effector sensing would result in receptor oligomerisation, which in turn promotes the TIR domain to self-associate via a conformational change that brings two TIR domains together to signal (Bernoux et al., 2011a, Bentham et al., 2017).

2.4.2 Dual interfaces are required simultaneously for TIR domain signalling

Until this work, there was no evidence of multiple synergistic self-association interfaces in plant TIR domains. The previously characterised AE and DE interfaces were only observed individually, and never seen within the same crystal structure. The RPP1 TIR domain crystal structure, and MALS data of interface mutants, provides strong evidence for the use of dual interfaces in this protein. All mutants that were demonstrated by MALS to affect the ability of the TIR domain to self-associate, located to the AE or DE interfaces of the RPP1 TIR domain described in Figures 2.7 and 2.8. Furthermore, the data supporting dual interfaces in TIR domains are not limited to RPP1. Concurrently, a colleague in the Kobe Laboratory, Dr. Xiaoxiao Zhang solved the structure of the TIR domain from SNC1, and also observed the presence of both the AE and DE interface in this crystal structure (Appendix 2, Figure 1). The identification of dual interfaces in RPP1 and SNC1 TIR domains led to the hypothesis that these interfaces may exist in all plant TIR domains, but were not observed in the crystal structures due to the lattice arrangement during crystal packing. To test this hypothesis, Dr. Maud Bernoux, CSIRO Agriculture and Food Canberra, performed an *in planta* study testing the activity of AE and DE interface mutants in the L6 and RPS4 TIR domains, respectively (Appendix 2, Figure 3). Our suspicions were proven correct, as we observed AE and DE interface mutants attenuated

cell death signalling by the L6 and RPS4 TIR domains. These results remained consistent in the context of the full-length receptors, supporting the idea that the integrity of both interfaces is required for cell death signalling. MALS analyses of recombinant L6 TIR domain with interface mutations, performed by Dr. Zhang, demonstrated the same trend of reduced self-association correlating with perturbed signalling activity *in planta*, as observed with the RPP1 TIR domain (Zhang et al., 2017). DE interface mutants for RPS4 and SNC1 showed little difference to the wild-type protein, however yeast-2 hybrid screens of the mutations show disruption of interaction for both proteins.

Comparison of the crystal structures of the L6, SNC1 and RPP1 TIR domains uncovered intriguing differences. In all three crystal structures, it is apparent that the molecules forming the DE interface exist in different relative orientations. Using the position of the αE helix of the L6 TIR domain as a starting position, a superimposition of the DE interface dimers of SNC1 and RPP1 TIR domains with the DE interface dimer of the L6 TIR domain, reveals the DE interfaces of the RPP1 and SNC1 TIR domains are rotated clockwise 21° and 97°, respectively (Figure 2.9). While the rotation of the molecules creates some variation between the DE interfaces of each TIR domain, the residues essential for the formation and stability of DE interface remain essentially the same. As stated in the results, these differences in the orientation of the DE interfaces of L6, SNC1, and RPP1 TIR domains are not a result of any difference in the overall fold of these proteins. The cause of this difference in the relative orientation of the subunits in the dimer crystal structures is not known. It is possible the cause is different crystal packing of each TIR domain, or inherent flexibility of the DE interface, or a combination of both. What is clear is that the flexibility required by the DE interface to accommodate different orientations is not shared by the AE interface, as the conservation

of its intercalating and structural arrangement is more robust (Figure 2.9).

While there are some differences between TIR domains from different NLRs, collectively, this work suggests that self-association simultaneously through the AE and DE interfaces is required for plant TIR domain signalling. These interfaces have now been demonstrated to be extended to even more plant NLRs (Williams et al., 2016), and also a TIR-only resistance protein, RBA1 (Nishimura et al., 2017).

2.4.3 Dual Interfaces and NLR oligomerisation

The discovery that plant TIR domains signal through multiple conserved interfaces significantly changes our understanding of how TIR domains function. Previous models of NLR activation proposed dimerisation of the full-length receptor, facilitated by the TIR domain, would activate signalling. This would then lead to the formation of a protein scaffold to facilitate binding of a downstream signalling partner (Bernoux et al., 2011a). However, it is apparent that this model does not completely explain the observations of the work presented here, as it does not account for multiple TIR domain interactions. Without any identified downstream signalling partners or independently elucidated TIR domain function, and an lack of structural data outside the TIR domain, we are left to speculate why multiple interfaces are required in TIR domain signalling.

One possibility is that the TIR domain signal could involve high-order oligomerisation. By modelling of the AE and DE interface dimers of L6 and SNC1 TIR domains, we were able to create a "TIR domain superhelix" (Figure 2.10). This superhelix is formed by alternate stacking of AE and DE interfaces creating a large left-handed helix. Interestingly, the formation of helix is highly dependent on the orientation of the DE interface. Comparison

between the L6 and SNC1 TIR domain superhelices shows that the superhelix created by the SNC1 TIR domain structure results in a longer, more extended model, requiring more TIR monomers to complete a single rotation compared to that of the L6 TIR superhelix (Figure 2.10 A & B). In contrast to these two models, the RPP1 TIR domain is unable to be modelled into an opened-ended helix. The greater magnitude of rotation at the DE interface prevents a continuous, alternating stacking of AE/DE interfaces, as the monomers, in this model, would clash (Figure 2.10 C). What is apparent from the superhelix models of L6, SNC1, and RPP1 TIR domains, is that it would require a specific orientation of both AE and DE interfaces to facilitate adequate stacking interactions required for superhelix formation (Zhang et al., 2017). However, the observed flexibility of the DE interface could assist in the formation of this structure, as the ability of the DE interface to assume multiple orientations reduces the specificity of the interaction required to make the superhelix. However, the transient interactions between TIR domains mean this complex would be difficult to assemble autonomously, and would likely require additional domain interactions in the context of a full-length protein.

To this end, oligomerisation of the other domains in the NLR has been proposed numerous times in the literature to be the driving force behind proximity-induced association of CC and TIR domains (Bernoux et al., 2011a, Takken and Goverse, 2012, Schreiber et al., 2016, Bentham et al., 2017). As mentioned, there is a lack of evidence for higher-order assembly in plant NLRs. The only current data to demonstrate the association NB-ARC and LRR domains were acquired via co-immunoprecipitation, a technique that cannot elucidate the stoichiometry of these potential complexes, nor exclude the involvement of other proteins (Mestre and Baulcombe, 2006, Schreiber et al., 2016). The only real way to confirm or

disprove this speculation will be experimental evidence involving either structure of the NB-ARC domain of a plant NLR, or preferably, a structure of a full length NLR.

2.4.4 Mammalian NLR Activation Could Give Insight into Plant NLR function

While higher-order assembly remains speculation in the case of plant NLRs, it has been shown to play a role in mammalian NLR activation. Researchers initially speculated that higher-order assembly was involved in caspase and caspase recruitment domain (CARD) associations in mammals, prior to any structural evidence for higher-order assembly in the mammalian NLR systems (Salvesen and Dixit, 1999, Hu et al., 2015, Zhang et al., 2015). However, the recent cryo-electron microscopy structure of the NAIP2/NLRC4 inflammasome clearly demonstrates the role of the NB domain in the association of the NLR into a large complex (Hu et al., 2015, Zhang et al., 2015). Inflammasome formation is required for signal activation, and facilitates the assembly of the C-terminal CARDs, the domain performing analogous functions to the CC/TIR domains of plant NLRs (Vajjhala et al., 2017). Disruption of the oligomerisation of the inflammasome complex, or compromising the CARD via knockout or mutation results in an inactive receptor (Bouchier-Hayes and Martin, 2002).

A similar NB/LRR-driven oligomeric system may adequately reconcile plant NLR function with current structural observations. In the context of this work, it is apparent the TIR domain has, and requires, some innate capability to self-associate, as seen in MALS for RPP1 NdA and WsB TIR domains, and corroborated by the MALS analysis of the L6, RPS4 and SNC1 TIR domains (Figure 2.3; Appendix 2 Figure 1 A and Figure S4). Disruption of this ability to self-associate results in perturbation of signalling, in both the context of the full-length receptor and the TIR domain only. Furthermore, activity of a TIR domain can be enhanced by

strengthening its ability to self-associate. This is evident by mutational analyses in the RPP1 WsB TIR domain that provide activity and simultaneously enhance self-association (Figure 2.4 D), and previously documented using WsB TIR-YFP-fusions to enforce dimerization (Krasileva et al., 2010). These data collectively point toward the TIR domain needing a driving force to initialise self-association. The discovery of multiple plant TIR domain interfaces that are simultaneously required for self-association and cell death signalling suggests the influence of multiple TIR domain interactions, not only dimerisation, must be necessary for signalling activation. Figure 2.10 demonstrates a potential mechanism for TIR domain oligomerisation, however without solid structural evidence for this association it is impossible to know how these TIR domains would arrange to facilitate signalling. Furthermore, it is impossible to predict how the rest of the NLR will orientate around this structure. It is certainly plausible that a TIR domain super-structure would assume a completely different conformation due to the forces applied to the TIR domains by the oligomerisation of the NB and/or LRR domains. All that can be concluded with certainty is that the AE and DE interfaces are required for the formation of an oligomeric structure, but to confirm the existence of such a structure, further analysis of plant NLRs is required.

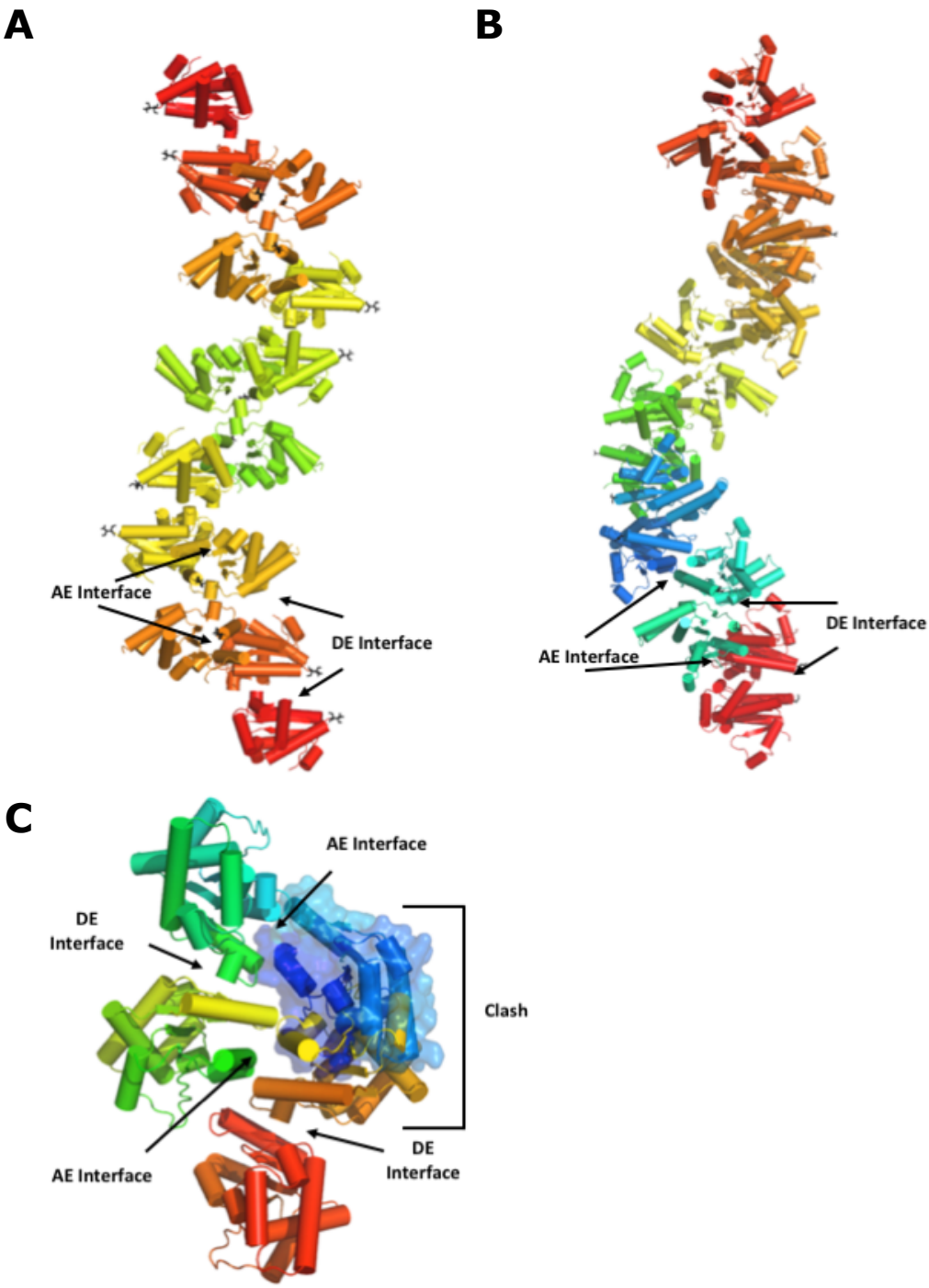


Figure 2.10 Hypothetical TIR domain superhelix formed via alternating stacking of AE and DE interfaces. The propensity of TIR domains to form this helix relies on a suitable rotation of the DE interface. (A) Superhelix formed through stacking of the L6 TIR crystal structure (PDB ID: 3OZI). (B) Same as (A) with SNC1 TIR (PDB ID: 5TEC). (C) Same as (A) with RPP1 TIR (PDB ID: 5TEB).

2.5 Summary

It has been known for some time that homotypic interactions between TIR domains are a prerequisite for NLR signalling. However, this research has played an important role in further understanding the complexities of TIR domain signalling, with the most significant finding of multiple functional TIR self-association interfaces. Biophysical, structural, *in planta* and mutational analyses of several plant TIR domains give irrefutable evidence of the need for simultaneous interaction interfaces for TIR signalling. It is not yet clear exactly how these interfaces allow the TIR domains to perform their concerted function, and there are still significant gaps in our understanding of how these interactions relay signalling of the hypersensitive response. This research challenges the NLR dimerisation hypothesis, rather I propose here a proximity-induced oligomerisation hypothesis for TIR domain signalling, whereby post-activation, multiple NLRs form large oligomers that facilitate the multiple TIR domain interactions required for signal transduction.

3. THE SOLUTION STRUCTURE OF THE SR33 COILED-COIL DOMAIN CHANGES PARADIGMS FOR COILED-COIL SELF-ASSOCIATION

The work presented in this chapter has been published in the journal Proceedings of the National Academy of Sciences. The contributions I made to the paper allow me to be recognised as a joint first author in the publication. This paper has been referenced extensively within this chapter and has been included as Appendix 3. It may benefit the reader to read this paper prior to reading this chapter. All the research presented in this chapter represents my own work, unless otherwise stated.

3.1 Introduction

3.1.1 Current structure/function studies of CC domains provide conflicting results

The majority of structure-function studies of CC domains have been centred around CNLs belonging to the CC_{EDVID} class. Prior to the research performed in this chapter, the only published structures of CC domains are from the MLA10 and Rx CC_{EDVID} domains (herein only referred to as CC domain). However, studies surrounding these proteins not only provide conflicting results about CC domain function, but the structures of the two proteins are vastly different, despite belonging to the same class of CC domain.

The CC domain of the barley NLR MLA10 was the first CC domain structure to be solved. The structure comprises residues 5 – 120 of the MLA10 CC domain ($MLA10^{5-120}$), which form an elongated helical structure, containing two helices linked by a short loop (Maekawa et al., 2011a). These helices were named α1 and α2, however helix α2 is split into two smaller helices, α2a and α2b, due

to a short disordered region in the centre of the helix (Figure 3.1). The interior of the all helices consists largely of hydrophobic residues that appear only to mediate a small number of interactions between the C-terminal end of α1, and α2a. These interactions seem insufficient to stabilise the structure of the monomer. However, it was determined that the MLA10$^{5-120}$ CC domain formed an obligate homodimer with the helices of the two monomers arranged in an anti-parallel, intercalating manner. The helical arrangement in this dimer allows for the hydrophobic residues of each of the helices in each monomer to interact to form a large, buried hydrophobic core of 7950 Å^2 (Figure 3.1 B). The authors justified the propensity of the MLA10$^{5-120}$ CC domain to exist as an obligate dimer with the following series of experiments. It was demonstrated, via size exclusion chromatography, that the MLA10$^{5-120}$ CC domain co-eluted with size markers correlating to its predicted dimer size and by chemical crosslinking with bis(sulfosuccinimidyl)suberate (BS3) the MLA10$^{5-120}$ CC domain was shown to form stable dimers. In yeast two-hybrid assays, a larger CC domain construct (MLA10$^{1-225}$) was capable of self-interactions (Maekawa et al., 2011a). Furthermore, a construct encoding residues 1 – 160, MLA10$^{1-160}$, caused cell death when transiently expressed in *N. benthamiana* as measured by electrolyte-leakage. Mutations to residues that disrupted self-association in Y2H, and perturbed the cell death response induced by the full length MLA10 protein. In the same study, the authors reported that the full length MLA27 (a highly similar NLR from the *MLA* locus) was purified as a monomer. In light of these data, the authors proposed that the first 120 residues of the MLA10 CC domain are the minimal functional unit required for homodimerisation, and the CC domain facilitates homodimerisation of the NLR post-activation.

In contrast to the MLA10$^{5-120}$ CC domain, the first 122 residues of the CC domain of the potato NLR, Rx, (Rx$^{1-122}$) presents a monomeric four-helix bundle fold, co-crystallised with the WPP (Trp-Pro-Pro) domain of its interacting partner, RanGAP2 (Hao et al., 2013) (Figure 3.2). In keeping with the notation used to describe the MLA10$^{5-120}$ CC domain, the helices of Rx are labelled α1 through to α4. While Rx has a typical CC-NB-LRR domain arrangement and contains the EDVID motif like MLA10, the CC domain of Rx does not trigger a cell death response when overexpressed *in planta*; rather cell death can be triggered by overexpressing the Rx NB-ARC domain *in planta* (Bendahmane et al., 1999, Moffett et al., 2002). Intriguingly, the interaction of the Rx$^{1-122}$ CC domain with the WPP domain of RanGAP2 is required for Rx protein function. Mutations that disrupt the Rx$^{1-122}$ CC : WPP domain interaction prevent cell death responses in the context of the full length protein (Rairdan et al., 2008, Hao et al., 2013).

In contrast to the TIR domain, the structural and functional differences observed for of CC domains creates many questions, and makes it hard resolve a single, or multiple functions, for this domain. Even though originally characterised as coming from the same class of CC domain due to the presence of the EDVID motif, there is apparent functional and structural promiscuity between the Rx and MLA10 CC domains. Neither of the crystallised regions of the CC domains have been demonstrated to be autoactive, bringing into question whether either of the CC structures represent a functional fold. To better understand the molecular mechanisms behind CC domain function, more studies are required across a broader range of CNLs.

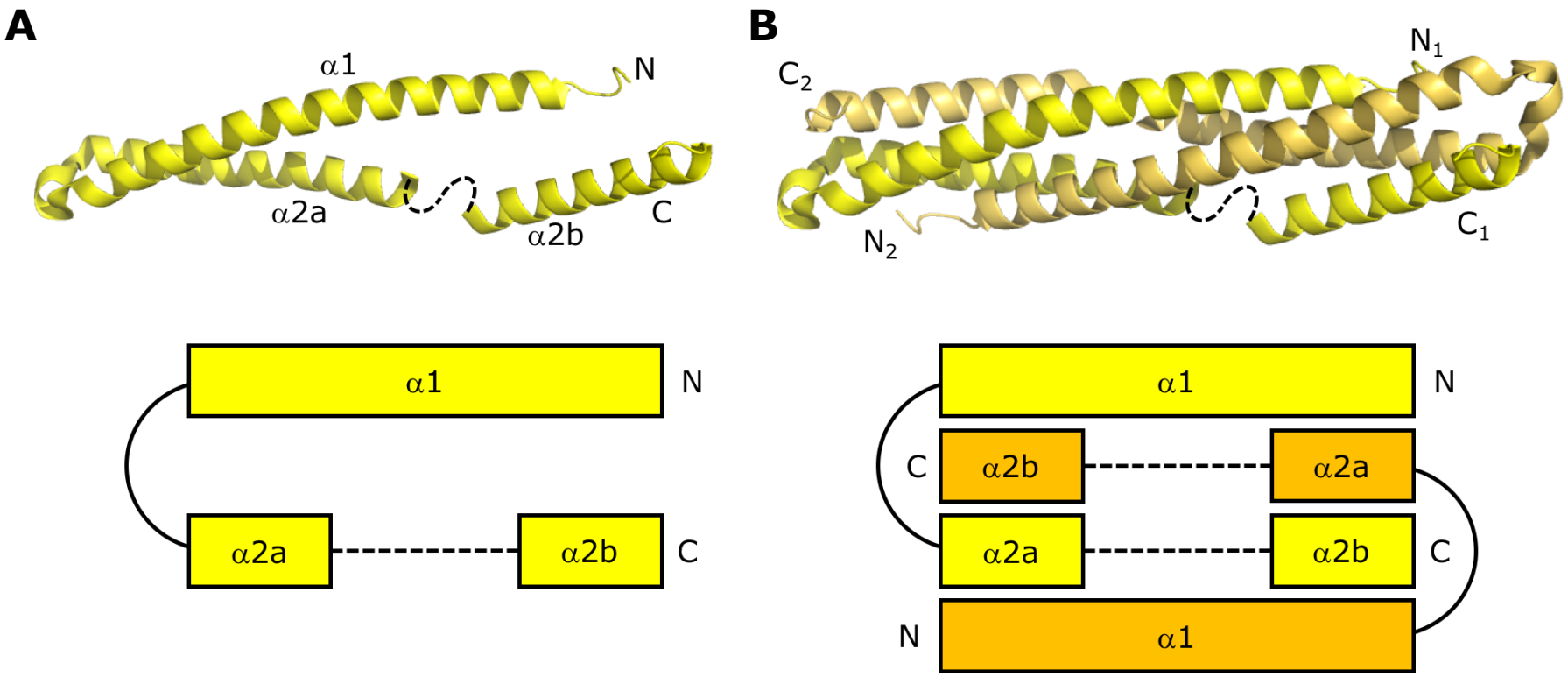


Figure 3.1 The crystal structure of the MLA10$^{5\text{-}120}$ CC domain, with diagrammatic representation of the helical arrangement. The amino and carboxyl terminus are labelled with N and C, respectively. Disordered residues in the structures are represented by a black dashed line. (**A**) The MLA10$^{5\text{-}120}$ monomer consists of two major α-helices annotated α1 and α2. Helix α2 has a short disordered region in the middle of the helix, and is subsequently divided into α2a and α2b. (**B**) The MLA10$^{5\text{-}120}$ homodimer. Interactions between the monomers are mediated by the hydrophobic residues of each of the helices, creating a large buried hydrophobic core. PDB ID: 3QFL (Maekawa et al., 2011b).

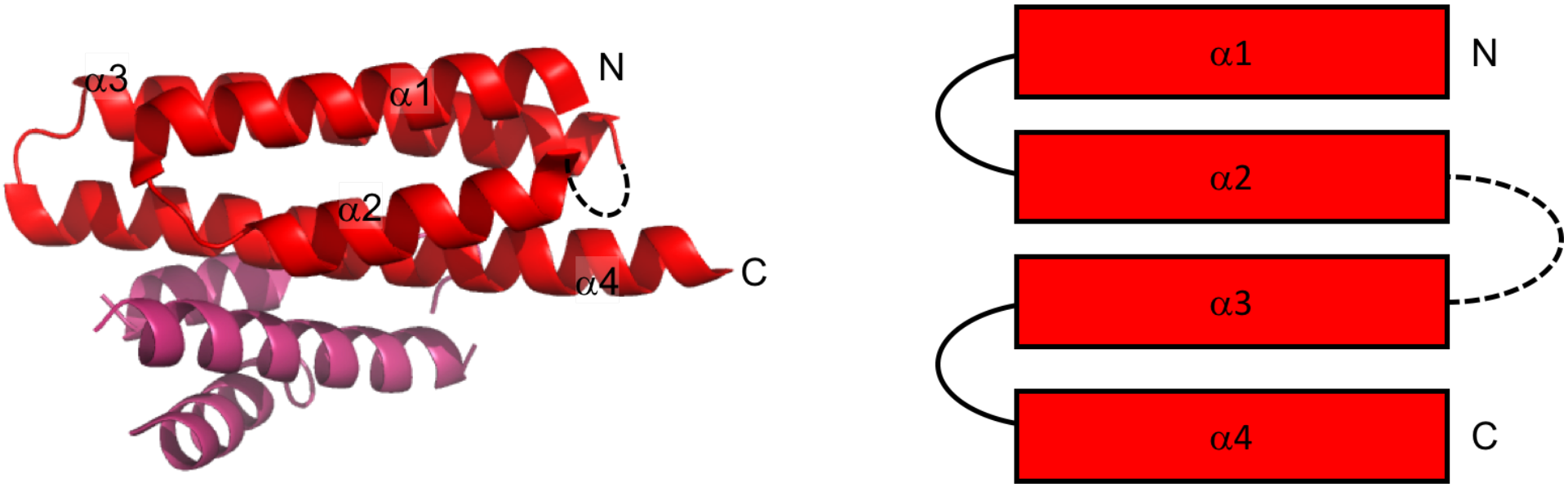


Figure 3.2 The crystal structure of the Rx$^{1-122}$ CC domain in complex with the WPP domain of RanGAP2 with a diagrammatic representation of the helical arrangement. The Rx CC$^{1-122}$ domain has a monomeric four-helix bundle fold consisting of four antiparallel helices, annotated helix α1 through to α4. Interactions with the RanGAP2 WPP domain are mediated by charged residues of α2 and α4. The amino terminus and carboxyl terminus of the protein are indicated by a N and C respectively. Disordered residues are shown as a black dashed line. PDB ID: 4M70 (Hao et al., 2013).

3.1.2 Sr33 is a CNL orthologous to MLA10

Stem rust resistance 33 (*Sr33*) is a gene from the wild wheat relative, *Aegilops tauschii,* that confers resistance to the highly virulent wheat stem rust (*Puccinia graminis sp. tritici*, *Pgt*) strain, UG99. UG99 has demonstrated to be virulent on 90% of wheat cultivars grown globally (Periyannan et al., 2013). This pathogenicity even extends to cultivars carrying the long-standing *Sr31* and *Sr2* resistance genes, which have been protecting wheat crops for over 30 years (Singh et al., 2011, Ayliffe et al., 2013).

The *Sr33* gene provides an intermediate resistance to many *Pgt* strains, and when used in conjunction with *Sr35,* provides a blanket resistance to all characterised strains of wheat stem rust (Periyannan et al., 2013, Saintenac et al., 2013). Genetically mapped in 1991, *Sr33* has been introgressed into common wheat (*Triticum aestivum*) from the wild relative *Aegilops tauschii*, and was able to be isolated and cloned by Periyannan et. al. (2013).

The *Sr33* gene encodes a CNL, and contains predicted motifs conserved in barley and *T. monoccum* mildew resistance proteins, and other CNL proteins (Seeholzer et al., 2010, Jordan et al., 2011). It shares high similarity to the barley powdery mildew pathogen (*Blumeria graminis f. sp. Hordei* or *Bgt*) resistance proteins from the *Mla* locus, and the MLA-like protein from *T. monoccum,* TmMLA1 (Seeholzer et al., 2010, Jordan et al., 2011, Periyannan et al., 2013). Sr33 shares the highest similarity with MLA34 (80% sequence identity), but it is also very similar to MLA10 (81% sequence similarity) (Seeholzer et al., 2010, Periyannan et al., 2013). Like MLA10 and Rx, the Sr33 CC domain is a member of the CC_{EDVID} subclass. Interestingly, Sr33 has low similarity to the "sister" CNLs Lr1, 10 and 21, which provide resistance to leaf rust, despite the similarity in protein class and

pathogens to which they confer resistance (Periyannan et al., 2013).

Data published by Periyannan et. al. (2013) suggest the Sr33 protein does not function through similar signalling pathways as other CNL proteins. Most proteins encoded by the *Mla* and *Lr* loci require interaction with the chaperones SGT1, RAR1 and HSP90 for function (Scofield et al., 2005, Seeholzer et al., 2010). Unlike other cereal CNLs, viral-induced gene silencing (VIGS) of HSP90, SGT1 and RAR1 had no effect on Sr33 dependent immunity. Furthermore, yeast-2 hybrid (Y2H) analysis of interactions between Sr33 and HSP90, SGT1 and RAR1 showed no indication of a direct interaction between these proteins. Finally, Y2H assays of the Sr33 CC domain did not detect any evidence of self-association, unlike that of MLA10 (Maekawa et al., 2011a, Bai et al., 2012, Periyannan et al., 2013).

3.2 Aims and Hypotheses

This chapter aims to investigate the structure and biophysical properties of the Sr33 CC domain, and compare the Sr33 CC domain to the previously characterized MLA10 and Rx CC domains. The hypothesis of this work is that the Sr33 CC domain adopts a similar fold to its ortholog MLA10, and show differences to the Rx CC domain. This is anticipated because of the high similarity of Sr33 to MLA10 (~81% identity in the CC domain) compared to its low similarity to Rx (~18% identity in the CC domain). The specific aims of this work include:

1. Expression and purification of the Sr33 CC domain
2. Biophysical analysis to compare the Sr33, MLA10 and Rx CC domains.
3. Crystallisation and structure determination of the Sr33 CC domain.

3.3 Results

3.3.1 Comparative biophysical analyses of the Sr33$^{6-120}$, MLA10$^{5-120}$ and Rx$^{1-122}$ CC domains reveal similarities

3.3.1.1 Expression and purification of the Sr33$^{6-120}$, MLA10$^{5-120}$ and Rx$^{1-122}$ CC domains

To replicate a Sr33 CC domain construct similar to those reported for MLA10 and Rx (MLA10$^{5-120}$ and Rx$^{1-120}$), a sequence encoding residues 6 – 120 of the Sr33 CC domain (Sr33$^{6-120}$) was cloned into an expression vector suitable for *E. coli* expression.

The Sr33$^{6-120}$, MLA10$^{5-120}$, and Rx$^{1-122}$ CC domains were expressed in *E. coli* BL21 and purified via immobilised metal affinity chromatography (IMAC) and subsequently by size-exclusion chromatography (SEC), as described by in the materials and methods of Casey et al. (2016) (Appendix 3). SDS-PAGE analysis of the CC domains demonstrates each of the proteins to have a molecular mass of between 10 - 15 kDa, which correlate to the expected monomeric mass for each for the proteins (Sr33$^{6-120}$ 13.1 kDa; MLA10$^{5-120}$ 13.4 kDa; and Rx$^{1-122}$ 14.3 kDa) (Figure 3.3 A).

To assess the oligomeric state of the CC domains in solution, each of the proteins was subjected to analytical SEC, with the use of molecular mass standards to determine protein size, as was performed in Maekawa et. al. (2011). The Sr33$^{6-120}$, MLA10$^{5-120}$, and Rx$^{1-122}$ CC domains were all found to elute slightly earlier than the α-chymotrypsin 25 kDa size marker, suggesting based on SEC alone that the proteins are forming dimers (Figure 3.3 B).

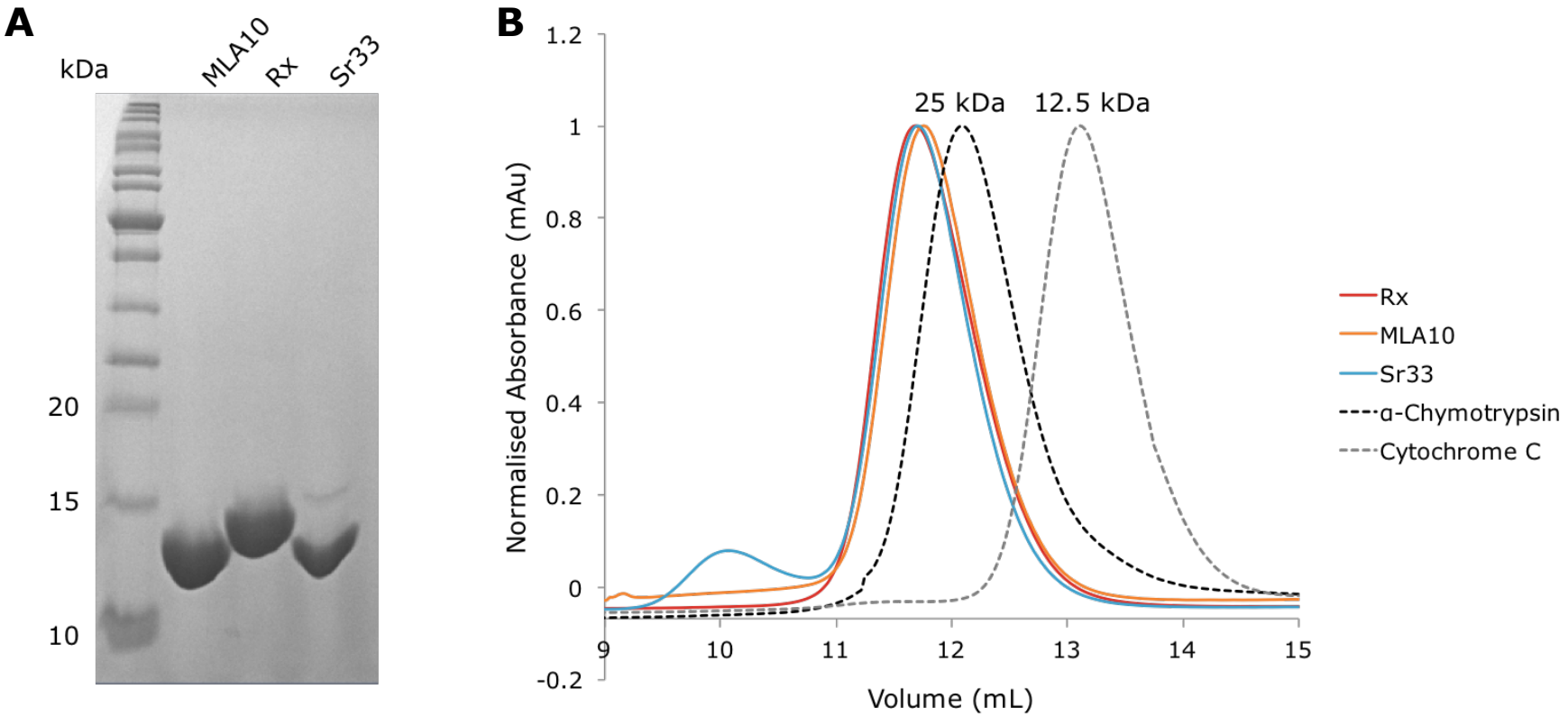


Figure 3.3 Expression and purification of the Sr33, MLA10, and Rx CC domains and comparison by analytical size-exclusion chromatography. (**A**) Sr33, MLA10, and Rx CC domains were run on a 12% SDS-PAGE gel after purification via IMAC and SEC. (**B**) Analytical SEC analysis was performed by separating equal concentrations of the Sr33, Rx and MLA10 CC domains with a Superdex S75 HR 10/300 GL column (GE Healthcare). The CC domains were compared to protein size makers, α-chymotrypsin (25 kDa) and cytochrome C (12.5 kDa) run under the same conditions.

3.3.1.2 Multi-angle light scattering analysis of the Sr33, MLA10, and Rx CC domains reveals they are monomeric in solution

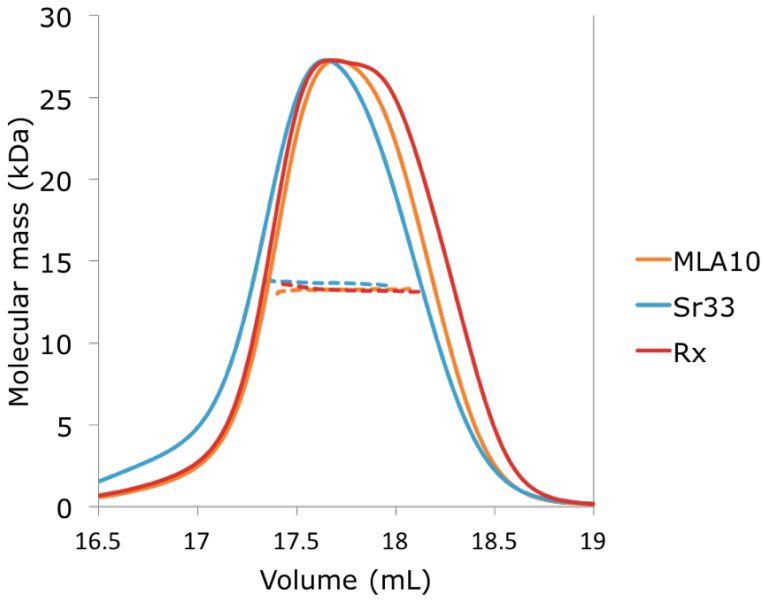
Analytical SEC of $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122} CC domains revealed all proteins migrated through the size-exclusion column at approximately the same rate. Despite this, $MLA10^{5-120}$ and Rx^{1-122} have previously been described as have different oligomeric states, dimer and monomer, respectively. Due to these unexpected results, it was decided the molecular mass should be experimentally determined using the technique of multi-angle light scattering (MALS). SEC separates proteins on the basis of size and shape; however, the inclusion of multi-angle light scattering can be used to derive an absolute molecular mass, not dependent on standards or protein shape. Using SEC-MALS, the molecular mass of $Sr33^{6-120}$ was measured to be approximately 13.7 kDa, with $MLA10^{5-120}$ and Rx^{1-122} both exhibiting molecular masses of 13.3 kDa. For all CC domains, the measured masses correlated to theoretical monomeric mass ($Sr33^{6-120}$ 13.1 kDa, $MLA10^{5-120}$ 13.4 kDa and Rx^{1-122} 14.3 kDa) (Figure 3.4).

3.3.1.3 SAXS analyses of CC domains reveal similarities in shape

The revelation that the $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122} CC domains were all monomeric in solution made determining the fold of these proteins in-solution imperative. Small-angle X-ray scattering (SAXS) is an excellent technique for comparison of protein size, shape, and oligomeric state in solution. The $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} CC domains were subjected to SEC-SAXS analysis with the assistance of Dr. Lachlan Casey, University of Queensland (Figure 3.5; Appendix 3 Figure 3).

Normalised distribution functions, P(r), derived from the scattering data, corresponded with the expected scattering of compact, globular proteins (Figure 3.5 B). Furthermore, the radius of

gyration (R_g) of each of the CC domains was very similar (R_g of $Sr33^{6-120}$ measured at 17.19 Å, and the R_g of $MLA10^{5-120}$ and Rx^{1-122} calculated at 17.62 Å and 17.38 Å, respectively, suggesting all three proteins behave similarly in solution (Figure 3.5 C-D). Finally, Dr. Casey performed molecular mass calculations on the CC domains using a local high-throughput implementation of the volume of correlations method (Rambo and Tainer, 2013) (Figure 3.5 C-D). This analysis resulted in calculated molecular masses of 13.5 kDa, 13.7 kDa, and 13.8 kDa for $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122}, respectively. These data closely resemble the observations obtained from SEC-MALS, and the theoretical monomeric masses of the CC domains, further confirming the $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} CC domains are monomeric in solution.



Construct	Predicted Molecular Mass (kDa)	Measured Molecular Mass (kDa)
Sr33	13.1	13.7
MLA10	13.4	13.3
Rx	14.3	13.3

Figure 3.4 Multi-angle light scattering analysis of Sr33$^{6\text{-}120}$, MLA10$^{5\text{-}120}$, and Rx$^{1\text{-}122}$. Refractive index is plotted as a solid line representing the elution peak of the protein from the size-exclusion column. Molecular mass of the protein is presented as the dashed line under the peak of the corresponding colour. All the CC domains exhibit molecular masses than correlate with their predicted monomeric sizes, which can be viewed in the table under the graph.

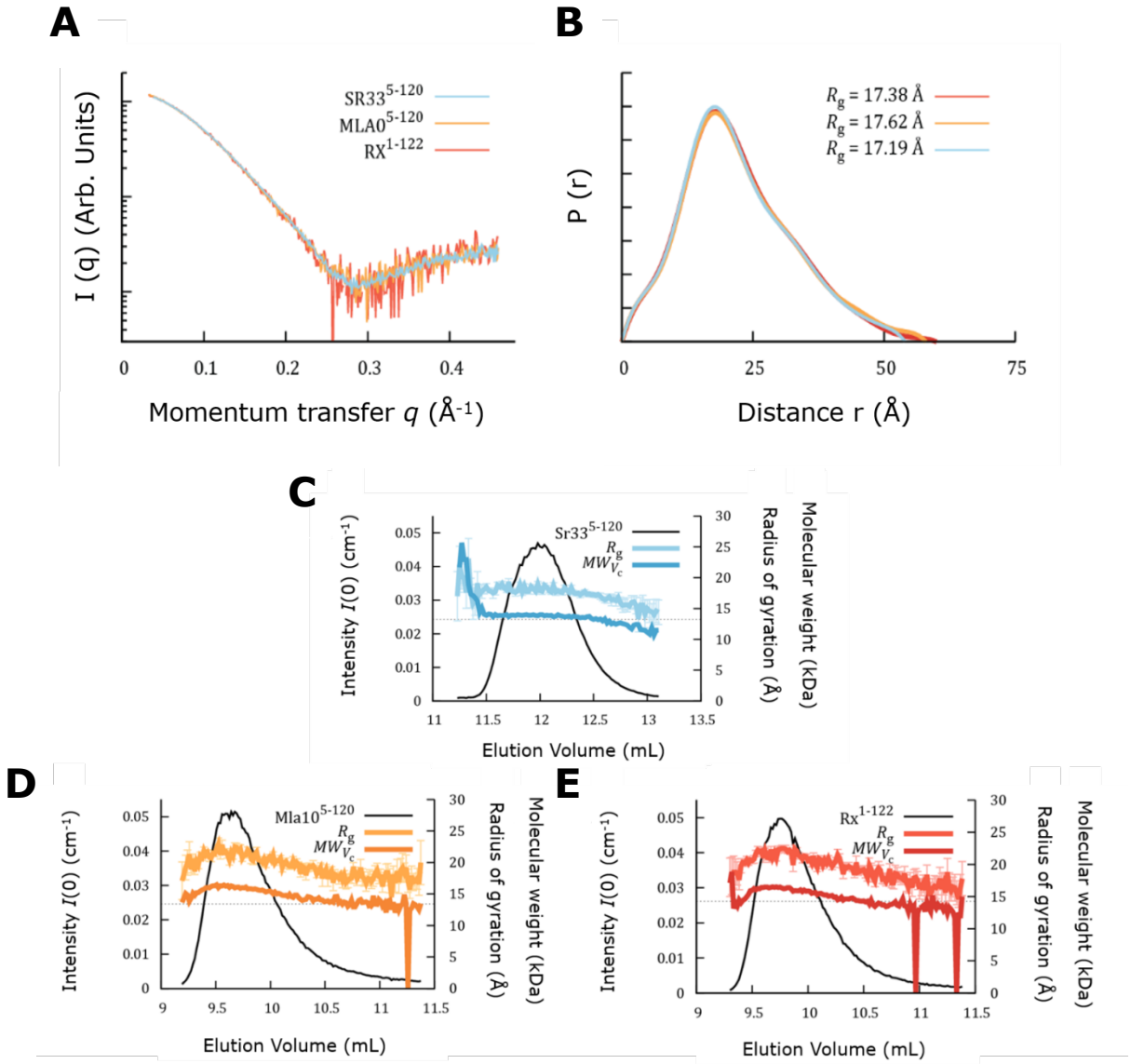


Figure 3.5 Small-angle X-ray scattering of Sr33$^{6\text{-}120}$, MLA10$^{5\text{-}120}$, and Rx$^{1\text{-}122}$ performed by Dr. Lachlan Casey. (**A**) Datasets collected at the Australian Synchrotron SAXS beamline. Rx$^{1\text{-}122}$ and MLA10$^{5\text{-}120}$ scattering data shown is scaled to overlay with Sr33$^{6\text{-}120}$, with intensity in arbitrary units. (**B**) Normalised distribution functions P(r) for Sr33$^{6\text{-}120}$ (blue), MLA10$^{5\text{-}120}$ (yellow), and Rx$^{1\text{-}122}$ (red). (**C-E**) Radius of gyration (R_g) and molecular weight (MW_{VC}) of (C) Sr33$^{6\text{-}120}$, (D) MLA10$^{5\text{-}120}$, and (E) Rx$^{1\text{-}122}$ across the elution peak. Radius of gyration is shown in the lighter shade of colour, and molecular weight in a darker shade, across the elution peak indicated by a solid black line.

3.3.2 Determination of the 3D structures of Sr33$^{6\text{-}120}$ and MLA10$^{5\text{ - }120}$

3.3.2.1 Crystallisation of the MLA10$^{5\text{-}120}$ and Sr33$^{6\text{ -}120}$ CC domains

The Sr33$^{6\text{-}120}$ CC domain was subjected to sparse matrix crystallisation trials in an attempt to yield crystals suitable for X-ray diffraction studies. Parallel to this, an attempt was made to crystallise the MLA10$^{5\text{-}120}$ CC domain in alternative conditions to those used by Maekawa et. al. (2011), to see if a different crystal type could be observed.

Initial screen trials of Sr33$^{6\text{-}120}$ resulted in no crystal hits and despite considerable effort, crystallisation of this protein was unsuccessful (Figure 3.6 A). The MLA10$^{5\text{-}120}$ CC domain crystallised in multiple conditions, including conditions similar to those reported by Maekawa et. al. (2011): 0.1 M sodium acetate trihydrate pH 4.6, 2.0 M sodium formate (Figure 3.6 B). It was decided that 0.1 M MIB buffer pH 7.0 with 25% PEG 1500 resulted in the most optimal crystals for X-ray diffraction studies that were likely different from those analysed by Maekawa et al.

3.3.2.2 The crystal structure of the MLA10$^{5\text{-}120}$ CC domain and the previously reported MLA10 CC domain structure are highly similar

Diffraction data for the MLA10$^{5\text{-}120}$ CC domain were collected to 2.0 Å resolution at the MX2 beamline at the Australian Synchrotron by Dr. Simon Williams. However, the 3D structure could not be solved via molecular replacement with the reported MLA10$^{5\text{-}120}$ CC domain structure. In an effort to determine the crystal structure of MLA10$^{5\text{-}120}$ domain, I generated selenomethionine-labelled MLA10$^{5\text{-}120}$ protein so that single anomalous diffraction (SAD) experiments could be used to solve the phase problem. Fortunately, the protein could be crystallised in the same conditions as the native MLA10$^{5\text{-}120}$ CC domain, from which we were previously able to collect a

dataset (25% PEG 1500, 0.1 M MIB pH 7) (Figure 3.6 C). Further diffraction data were collected at the Australian Synchrotron on the MX2 beamline to 2.1 Å resolution, using a wavelength of 0.9792 Å. The MLA10$^{5\text{-}120}$ crystals had the space group symmetry $P22_12_1$, and the structure was solved using SAD. Model building and refinement using Coot and phenix.refine was attempted, however R_{work} and R_{free} values could not be improved past ~28% and 30%. With the help of Dr. Simon Williams, and collaborators, Dr. Daniel Ericsson and Dr. Dušan Turk, the model and refinement was able to be improved through expanding R_{free} test set from $P22_12_1$ to P1, and reprocessing the data in P1 with final R_{work} and R_{free} values of 25% and 28%, respectively. A full table of crystallographic statistics for the MLA10$^{5\text{-}120}$ CC domain structure can be found in Table 3.1.

From this analysis, eight MLA10$^{5\text{-}120}$ CC domain helix-loop-helix monomers can be found in the asymmetric unit, forming 4 anti-parallel dimers that look very similar to the structure of the MLA10$^{5\text{-}120}$ CC domain published by Maekawa et. al. (2011). As extensive comparisons will follow between the MLA10$^{5\text{-}120}$ CC domain structure first published in Maekawa et. al. (2011), and the structure determined in this research, these proteins will be referred to as the Maekawa structure and the Casey structure, respectively.

The dimers in the Casey structure are formed between chains A and B; C and D; E and F; and G and H. Chains A and B only will be referred to for the rest of the chapter. Similar to the Maekawa structure, the Casey MLA10$^{5\text{-}120}$ CC monomers consist of two helices (deemed α1 and α2) joined by loops. In chain A, α1 maintains a long unbroken helix consisting of residues 5 – 53, which proceeds into a loop (disordered in the structure; residues 54 -57) that joins α1 and α2. Helix α2 spans from residues 58 – 90 before unravelling to form a short loop region and then re-folds to

form the rest of the helix, which continues from residues 98 – 120 (Figure 3.7 A). This small unwound segment of α2 mirrors the disordered residues found in the Maekawa MLA10$^{5\text{-}120}$ structure, and for consistency, the regions of the helix separated by the loop in the Maekawa structure will be referred to as α2a and α2b. Chain B exhibits a similar fold to chain A; however, the loop connecting α1 and α2 is visible in the structure, and the unravelling of helix between α2a and α2b is less severe than in chain A (Figure 3.7 B). The surface area of the homodimer interface is approximately 7,800 Å^2, mediated by interactions between the vast majority of the internalised residues of helices α1, α2a, and α2b (Figure 3.7 C). Superimposition of the Casey MLA10$^{5\text{-}120}$ CC homodimer with the Maekawa MLA10$^{5\text{-}120}$ CC homodimer reveals an overall similar structure. However, the R.M.S.D. of the structure is ~3.6 Å, which is high for two structures of the same protein, and likely reflects the internal organisation of the dimerisation interface, which is in some cases mediated by different interactions between residues of the opposing monomers (Appendix 3, Figure S4).

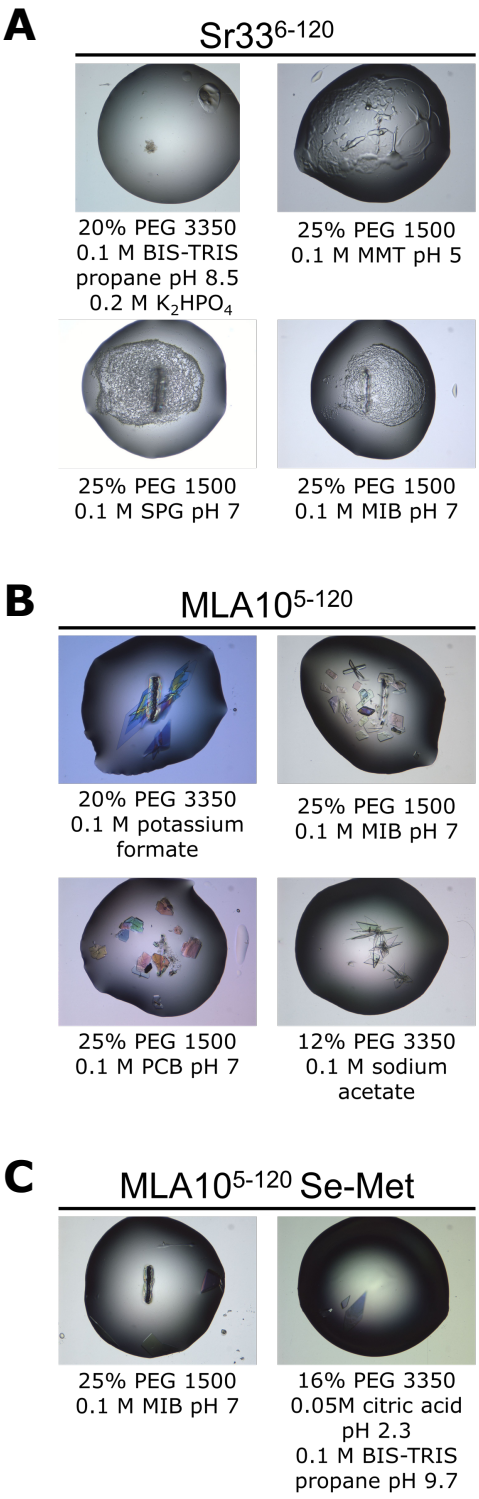


Figure 3.6 Crystallisation of the Sr33$^{6-120}$ and MLA10$^{5-120}$ CC domains. (**A**) Sample of drops from sparse matrix screens of Sr33$^{6-120}$. No crystals were observed. (**B**) Sample of drops from sparse matrix screens of MLA10$^{5-120}$. Crystals readily formed in a variety of conditions. Crystals grown in 25% PEG1500 and 0.1 M MIB pH 7 were chosen for diffraction studies. (**C**) Crystals observed in sparse matrix screens of seleno-methionine (Se-Met) labelled MLA10$^{5-120}$. The Se-Met MLA10$^{5-120}$ is capable of forming crystals in the same condition as the native protein. These crystals were used for SAD at the MX2 beamline at the Australian Synchrotron. Crystal images are approximately 1000 μM x 1000 μM.

Table 3.1 Crystallographic data for the Casey MLA10$^{5\text{-}120}$ CC domain structures

Data processing		
Space group	P 2 2_12_1	P 1
a, b, c (Å)	30.87, 87.56, 92.56	30.72, 87.14, 92.25
α, β, γ (°)	90, 90, 90	89.93, 90.00, 89.98
Resolution (Å)	46.28-2.1 (2.16-2.10) [a]	46.12-2.05 (2.10-2.05)
Rmeas (%) [b]	11.0 (194.0)	6.8 (79.2)
Rpim(%) [c]	3.0 (51.0)	4.8 (56.0)
<I/σ(I)>	15.0 (1.8)	8.2 (1.4)
$CC_{1/2}$ [d]	0.99 (0.89)	0.99 (0.77)
Completeness (%)	100 (100)	96.5 (92.9)
Multiplicity	14.1 (14.4)	1.8 (1.8)
Wilson plot B (Å^2)	44.7	38.9
Observations	216711 (18084)	107102 (7880)
Unique reflections	15392 (1253)	58095 (4314)
Anomalous completeness	100 (100)	-
Anomalous multiplicity	7.7 (7.7)	-
DelAnom correlation between half-sets	0.471 (-0.027)	-
Mid-slope of anomalous normal probability	1.087	-
Estimate of maximum resolution for significant anomalous signal = 3.59 Å, from CCanom > 0.15		
Refinement		
Rwork (%)	27.9 (31.1)	25.2 (37.7)
Rfree (%)	30.1 (35.0)	27.9 (40.7)
Average B-factor (Å^2)	60.35	62.15
R.m.s deviations		
Bond lengths (Å)	0.009	0.001
Bond angles (°)	1.07	0.348
Ramachandran plot (%) [e]		
Favoured	96.41	99.32
Allowed	99.10	100.00
Outliers	0.90	0.00

[a] NB: Values within parentheses indicate the highest resolution bin.
[b] $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th measurement of an equivalent reflection with indices *hkl*.
[c] $R_{pim} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.
[d] Calculated with the program Aimless (25).
[e] As calculated by MolProbity (26).

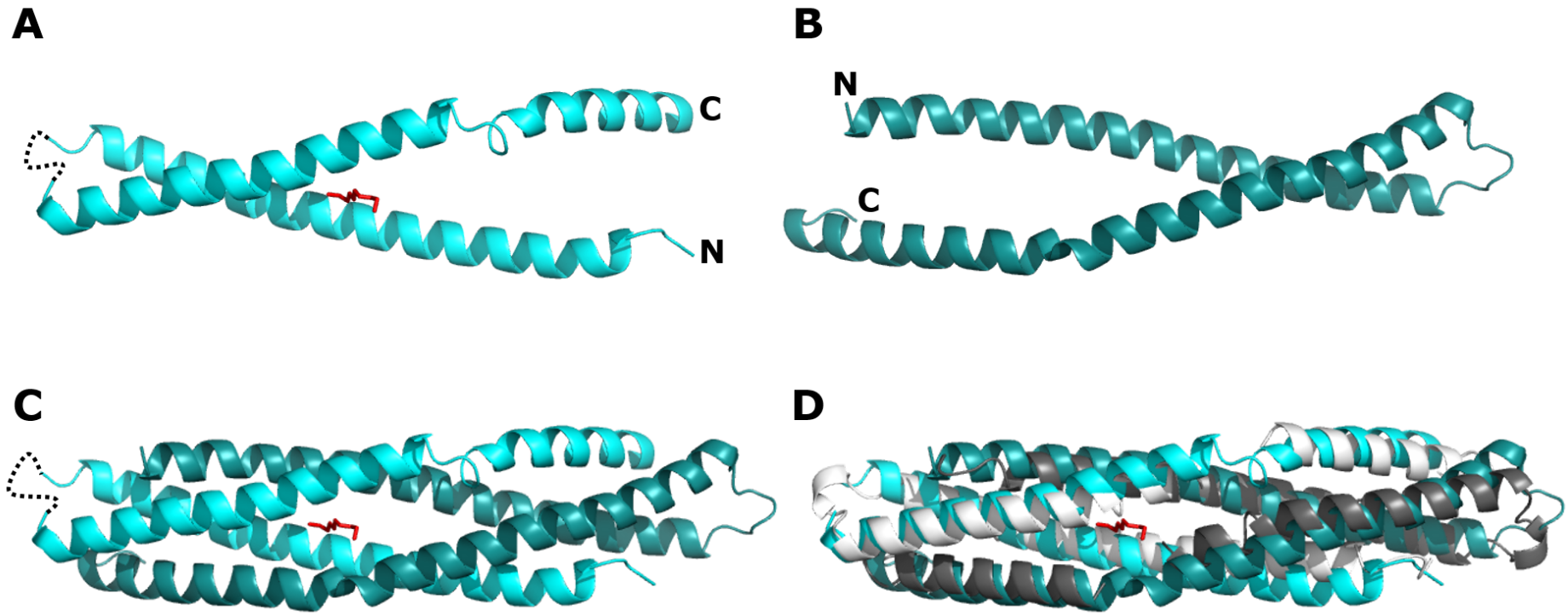


Figure 3.7 Comparison of the Casey and Maekawa MLA10$^{5\text{-}120}$ crystal structures. (**A-B**) The monomeric chains of the Casey structure (PDB ID: 5T1Y) that combine to form the obligate MLA10$^{5\text{-}120}$ dimer seen in (**C**). Residues unresolved in the crystal structure are represented by a black dashed line, and the amino and carboxyl-terminus are indicated by N and C at the appropriate region. (**A**) Chain A of the MLA10$^{5\text{-}120}$ crystal structure, interacting with polyethylene glycol (red). Unravelling of the α2 helix occurs around residues 92 – 97, and residues 54 – 57 are disordered and not observed in the structure. (**B**) Chain B of the MLA10$^{5\text{-}120}$ crystal structure; a kink in the α2 helix, similar to chain A, is observed around residues 92 - 97. (**C**) The MLA10$^{5\text{-}120}$ homodimer formed by interactions between the monomers of chain A and B. (**D**) Superimposition of the Casey MLA10$^{5\text{-}120}$ crystal structure determined in this study with the Maekawa MLA10$^{5\text{-}120}$ crystal structure published in 2011 (monomers represented in white and grey). The two structures align with an R.M.S.D. of 3.6 Å.

3.3.2.3 The NMR structure of Sr33 reveals a monomeric four-helix bundle fold

The MLA10$^{5\text{-}120}$ CC domain structure determined by X-ray crystallography closely resembles the structure reported by Maekawa et. al. (2011), corroborating their initial reports. However, both the Casey and Maekawa MLA10$^{5\text{-}120}$ CC domain structures conflict with the biophysical analyses of CC domains in solution. As no crystals of the Sr33$^{6\text{-}120}$ CC domain were able to be produced, it was decided the structure of the Sr33$^{6\text{-}120}$ should be investigated using nuclear magnetic resonance (NMR) spectroscopy. One major advantage of using NMR to determine the atomic structure of Sr33$^{6\text{-}120}$ compared to X-ray crystallography, is that NMR is an in-solution technique and therefore avoids the creation of structural artefacts that can appear in the crystallisation process.

Dr. Peter Lavrencic (University of Queensland) was able to derive the 3D structure of the Sr33$^{6\text{-}120}$ CC domain using ^{15}N- and ^{13}C-edited NOESY experiments to calculate distance restraints, combined with dihedreal angles obtained from chemical shift analyses (Casey et al., 2016). The structure reveals a monomeric four-helix bundle (Figure 3.8). This Sr33$^{6\text{-}120}$ CC domain structure has a close similarity to the Rx$^{1\text{-}122}$ CC domain structure, which also has a four-helix bundle fold, and shares less similarity with the structure of the CC domain of its ortholog, MLA10, regardless of the high sequence similarity between the two proteins. The helices of the Sr33$^{6\text{-}120}$ NMR structure were labelled α1 to α4, consistent with the notation of the Rx$^{1\text{-}122}$ CC domain.

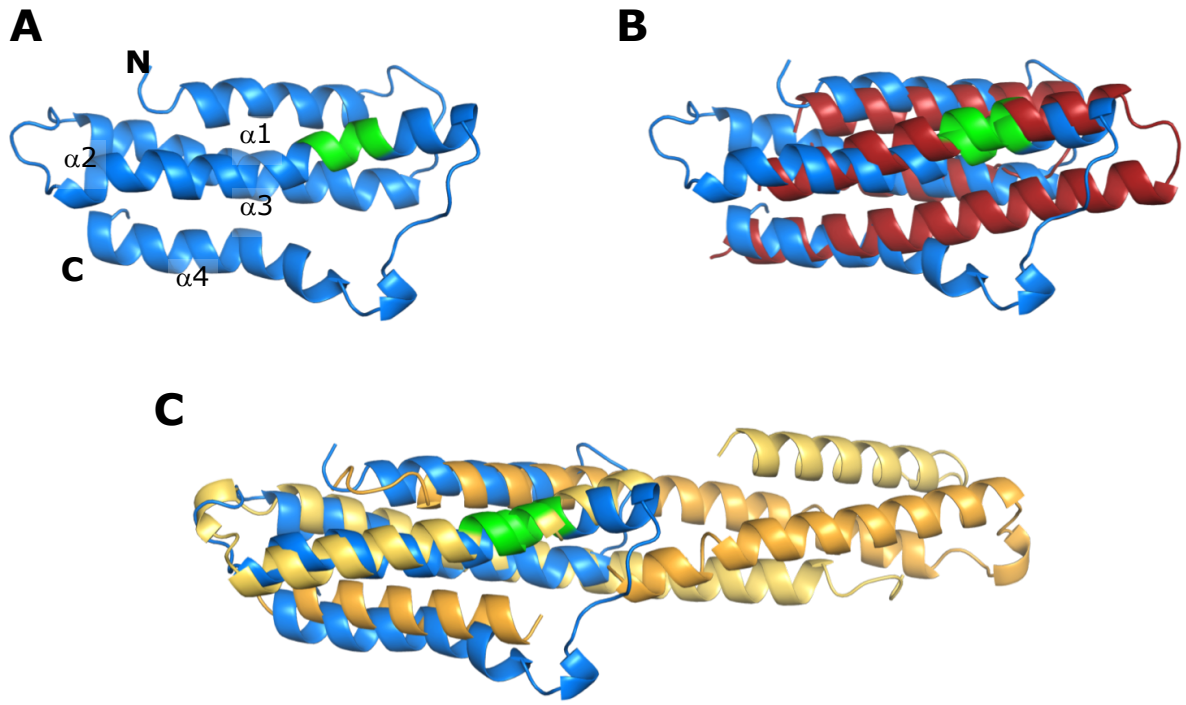


Figure 3.8 The NMR structure of Sr33$^{6\text{-}120}$ CC domain and comparison with the crystal structures of the Rx$^{1\text{-}122}$ and MLA10$^{5\text{-}120}$ CC domains. (**A**) The NMR structure of the Sr33$^{6\text{-}120}$ CC domain reveals a four-helix bundle fold. The EDVID motif typical of this class of CC domain is shown in green. The amino and carboxyl termini are indicated by the labels N and C respectively (PDB ID: 2NCG). (**B**) Structural alignment of the Sr33$^{6\text{-}120}$ NMR structure (blue) with the crystal structure of the Rx$^{1\text{-}122}$ CC domain (red) (PDB ID: 4M70). The Sr33$^{6\text{-}120}$ CC domain shows increased flexibility compared to the Rx$^{1\text{-}122}$ CC domain, however all the helices and the EDVID motif share a similar arrangement. (**C**) Structural alignment of the Sr33$^{6\text{-}120}$ CC domain (Blue) with the MLA10$^{5\text{-}120}$ CC domain homodimer (yellow & gold) (PDB ID: 3QFL). The MLA10 CC domain presents a considerably different fold, however alignment of the Sr33$^{6\text{-}120}$ overlays similar helices between the two proteins, with the EDVID motif occupying a similar position.

3.3.2.4 The Sr33$^{6-120}$ CC domain structure is consistent with biophysical analyses

Further experiments conducted by Dr. Peter Lavrencic demonstrated T_2 relaxation rates of the backbone amides of the Sr33$^{6-120}$ CC domain yield an overall rotational correlation time consistent with a spherical protein ~13 kDa in size, consistent with the molecular masses obtained by SAXS and MALS (Appendix 3, Table S1). With the 3D structures of the Sr33$^{6-120}$, MLA10$^{5-120}$ and Rx$^{1-122}$ CC domains, Dr. Lachlan Casey was able to generate theoretical SAXS scattering curves for each of the structures and compare them to the experimentally acquired scattering for each protein (Figure 3.9 A). It is clear from Figure 3.9 A that the SAXS curves generated from the four-helix bundles of the Sr33$^{6-120}$ and Rx$^{1-122}$ CC domain structures closely fit the previously acquired SAXS datasets, whereas the MLA10$^{5-120}$ CC domain homodimer does not. Furthermore, the structures of Sr33$^{6-120}$ and Rx$^{1-122}$ CC domains can fit into protein envelopes generated from the SAXS data, whereas the MLA10$^{5-120}$ homodimer extends well past the predicted maximum distance (Figure 3.9 B).

The NMR structure of the Sr33$^{6-120}$ CC domain combined with the in-solution comparative biophysical analyses of the Sr33$^{6-120}$, MLA10$^{5-120}$ and Rx$^{1-122}$ proteins provides strong evidence for the structures of each of the CC domains to have a monomeric, four-helix bundle fold in solution. Even though the crystallisation of the MLA10$^{5-120}$ CC domain could be repeated to obtain a similar structure to that reported by Maekawa et al (2011), the crystal structure heavily conflicts with the SAXS and MALS data, indicating this structure does not represent the structure observed in solution.

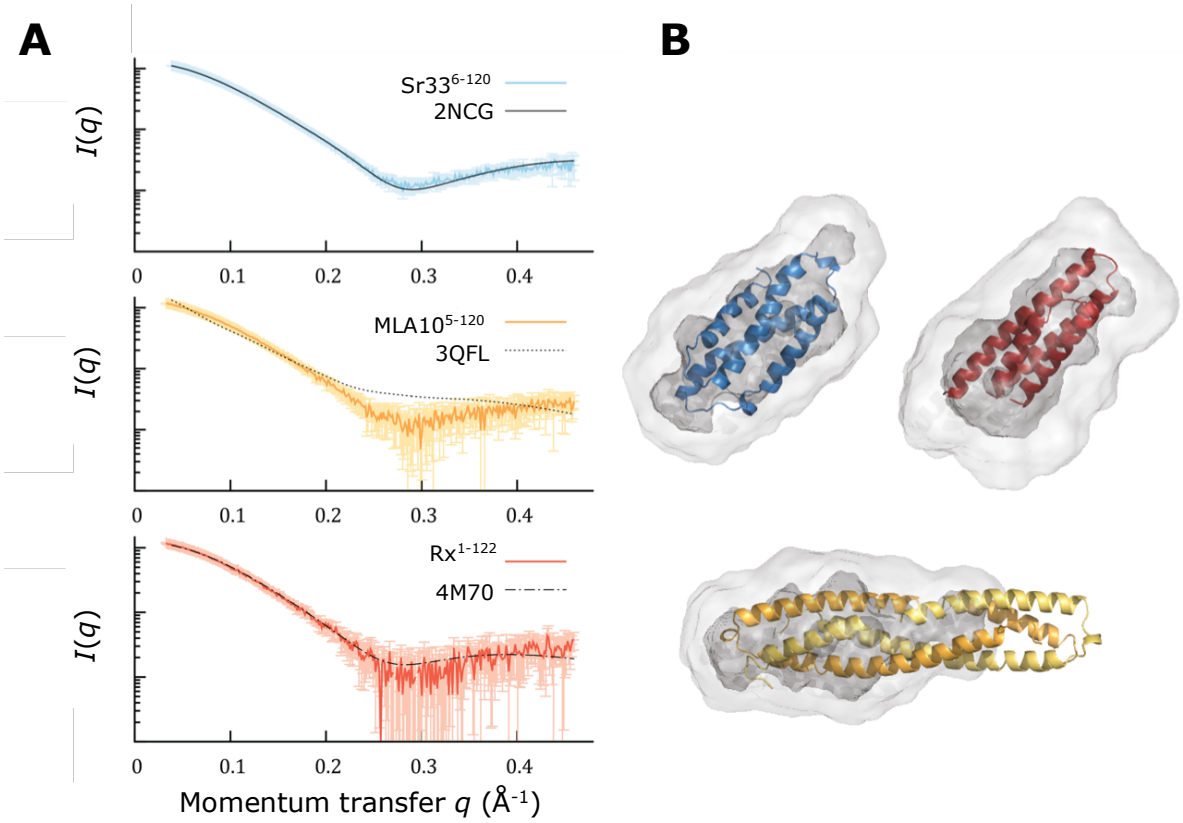


Figure 3.9 Comparison of CC domain structures to experimental SAXS data. (**A**) Comparison of the SAXS datasets of $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} CC domains with theoretical scattering curves generated from each of their respective 3D structures ($Sr33^{6-120}$, PDB ID: 2NCG; $MLA10^{5-120}$, PDB ID: 3FQL; Rx^{1-122}, 4M70). The scattering curves generated from the structures of $Sr33^{6-120}$ and Rx^{1-122} closely match the datasets collected at the SAXS beamline of the Australian Synchrotron. In contrast, the $MLA10^{5-120}$ structure-generated scattering curves do not match the experimental SAXS dataset. (**B**) Docking of $Sr33^{6-120}$ (blue), Rx^{1-122} (red), and $MLA10^{5-120}$ (yellow) 3D structures into protein envelopes generated from the SAXS data collected for each protein.

3.3.3 Longer Sr33 CC domain proteins are capable of self-association in solution, and signalling *in planta*

The structural and biophysical data provide strong evidence of the $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} CC domains are monomeric in solution and have a four-helix bundle fold. Self-association has always been deemed a critical event in the signalling of N-terminal NLR domains, however self-association was not observed for any of the CC domains tested. This revelation prompted a closer look into the sequences of the CC domains of Sr33 and MLA10 to identify a larger CC domain protein capable of self-association.

The sequences of Sr33, MLA10, and the closely related Sr50 (CNL from rye, *Secale cereale*), were analysed with the secondary structure predication program PSIPRED (Jones, 1999, Buchan et al., 2013). PSIPRED revealed the $MLA10^{5-120}$ crystallisation construct cut in the centre of the α4 helix of the CC domain, truncating the α4 helix by approximately 20 residues from the C-terminus. PSIPRED secondary structure predictions of Sr33, MLA10 and Sr50 all revealed the final helix of the CC domain to end at approximately residue 140 (Figure 3.10). Similar secondary structure analyses performed by our collaborator, Dr. Stella Cesari (CSIRO Agriculture and Food, Canberra), using the software COILS determined two separate helical regions spanning residues 20 – 50 and 115 – 144 (Lupas et al., 1991). As there is no *in planta* evidence of autoactivity of the $MLA10^{5-120}$ CC domain, only for $MLA10^{1-160}$, it is possible the $MLA10^{5-120}$ protein only represents a fragment of the CC domain required for signalling. If so, it is possible the CC domain requires the additional residues at the C-terminus of the α4 helix to facilitate self-association.

To test this hypothesis, two sets of longer *Sr33* and *MLA10* CC domain constructs were generated for *E. coli* expression to assess the ability of their protein products to self-associate. The first Sr33

and MLA10 constructs extend to residue to 144, to include the entire α4 helix, as predicted by the COILS software (named $MLA10^{5\text{-}144}$ and $Sr33^{6\text{-}144}$), and the second set of constructs includes residues up to 160 to assess the previously tested "active CC domain" seen in the experiments of Bai et. al. (2012) (named $MLA10^{5\text{-}160}$ and $Sr33^{6\text{-}160}$). Subsequently, $Sr33^{6\text{-}120}$, $MLA10^{5\text{-}120}$ and $Rx^{1\text{-}122}$ will be collectively referred to as C-120 (C-terminus, residue 120); $Sr33^{6\text{-}144}$ and $MLA10^{5\text{-}144}$ will be collectively referred to as C-144 (C-terminus, residue 144); and $Sr33^{6\text{-}160}$ and $MLA10^{5\text{-}160}$ will be collectively referred to as C-160 (C-terminus, residue 160).

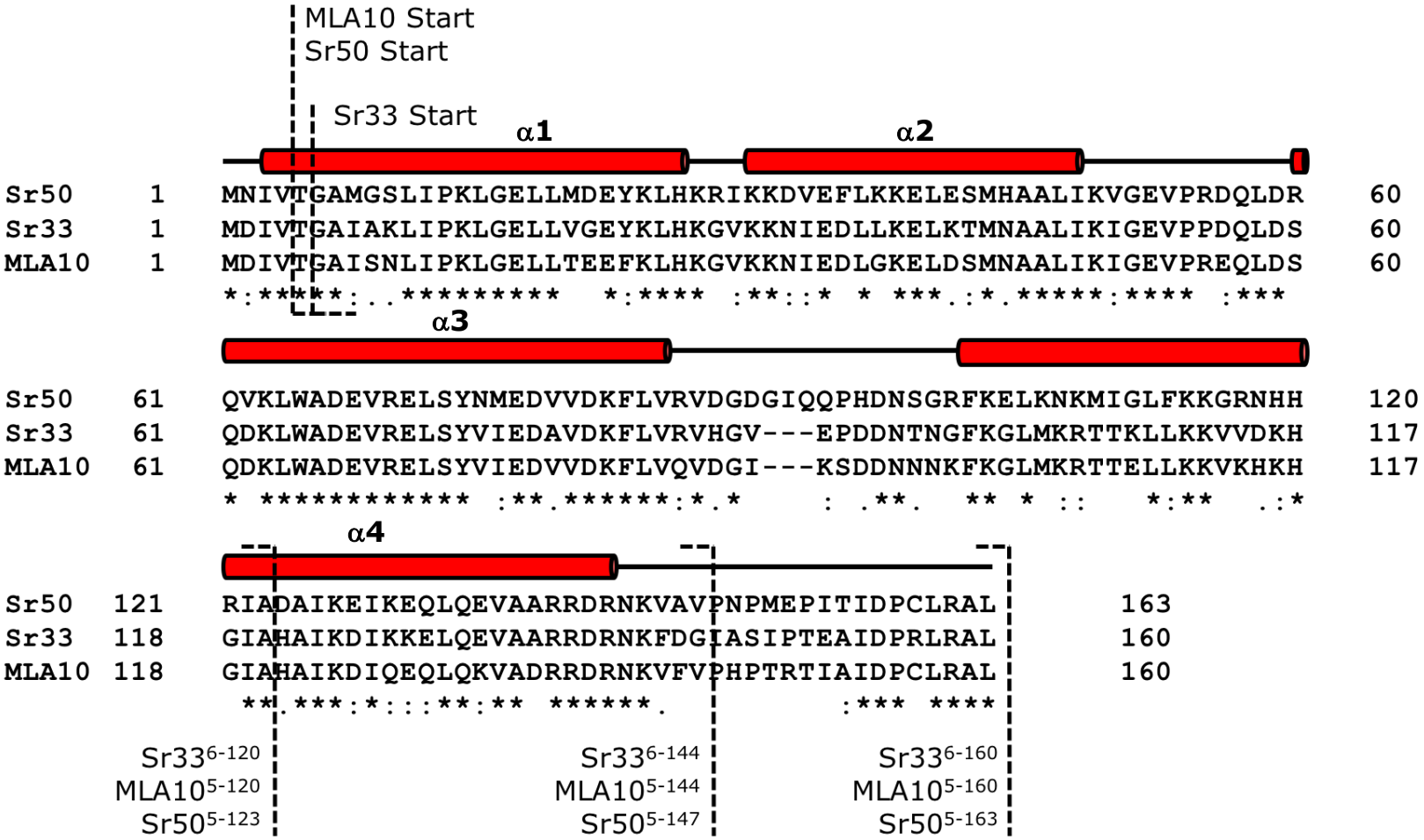


Figure 3.10 Multiple sequence alignment of the CC domains of MLA10, Sr33, and Sr50 overlayed with secondary structure prediction performed with PSIPRED. Dashed lines indicate the construct boundaries for the 120, 144 and 160 constructs. Red cylinders represent predicted helical regions, with a black solid line indicating predicted loops.

3.3.3.1 Additional residues in the α4 helix of the Sr33 and MLA10 CC domains promote self-association in solution

Initial small-scale expression of the Sr33 and MLA10 C-144 and C-160 constructs proved to yield insoluble protein in the same expression conditions as the C-120 constructs. To obtain soluble protein at the quantities required for crystallographic and biophysical studies, a buffer screen was performed that comprised of differing pH and concentrations of sodium chloride (NaCl). For Sr33 and MLA10 C-144, as well as Sr33 C-160, it was found that increasing the pH from 7.5 to 8.5 in combination with increases of NaCl concentration from 200 mM to 300 mM was sufficient to produced large quantities (15 mg / L of culture) of protein (Figure 3.11). However, despite extensive optimisation, the $MLA10^{6-160}$ protein could not be produced in a soluble form.

As performed previously with the C-120 CC domain fragments, the $Sr33^{6-144}$, $Sr33^{6-160}$ and $MLA10^{5-144}$ CC domains were subjected to SEC-MALS to assess their oligomeric state in solution (Figure 3.12). The $MLA10^{5-144}$ CC domain was observed to elute from the column as a single peak with a molecular mass of 22.8 kDa, ~41% larger than the predicted monomeric mass of 16.17 kDa. This is indicative of transient self-association of the $MLA10^{5-144}$ protein, and comparable to the MALS data previously reported for TIR domains, which also demonstrated polydispersity across the elution peak. MALS analyses of $Sr33^{6-144}$ and $Sr33^{6-160}$ CC domains revealed two peaks that were unable to be separated by the column, resulting in a minor shoulder and major peak in the direct refractive index (dRI) chromatogram. The average molecular mass of the protein in the shoulder peak of the $Sr33^{6-144}$ CC domain was measured to be 27.3 kDa, and the major protein peak was measured to correspond 17.5 kDa. Predicted monomer and dimer molecular masses for the $Sr33^{6-144}$ CC domain are 15.94 and 31.88 kDa, respectively. These data demonstrate the shoulder and major peaks of $Sr33^{6-144}$ CC domain correlate to dimer and monomer peaks, respectively,

although the average molecular mass of the shoulder peak is only ~70% larger than the predicted monomeric molecular mass, indicating the protein is still forming transient interactions within this peak. Similar to the $Sr33^{6-144}$ CC domain, the $Sr33^{6-160}$ CC domain presents a shoulder peak that correlates with a dimeric protein species (35 kDa) and a major peak that represents monomeric protein (19.3 kDa), when compared to the expected monomeric and dimeric masses of 17.65 kDa and 35.30 kDa, respectively. In this experiment, the longer constructs C-144 and C-160 of MLA10 and Sr33 CC domains were run in parallel to the C-120 constructs. No shoulder peak is observed in either the $Sr33^{6-120}$ or $MLA10^{5-120}$ samples, and the major peak correlates to monomeric protein as previously reported in this chapter.

Taken together, these data provide clear evidence that the additional residues at the C-terminus of the CC domain are required to facilitate self-association. Interestingly, this self-association is largely transient, as observed with the self-association of TIR domains; however, monomer and dimer protein species can both be observed eluting from the size exclusion column. Previously, *in planta* studies of $MLA10^{1-160}$ demonstrated the protein to be autoactive when transiently expressed in *N. tabacum* (Maekawa et al., 2011a, Bai et al., 2012). This provides a correlation between the ability of the CC domain to self-associate and its ability to signal *in planta*.

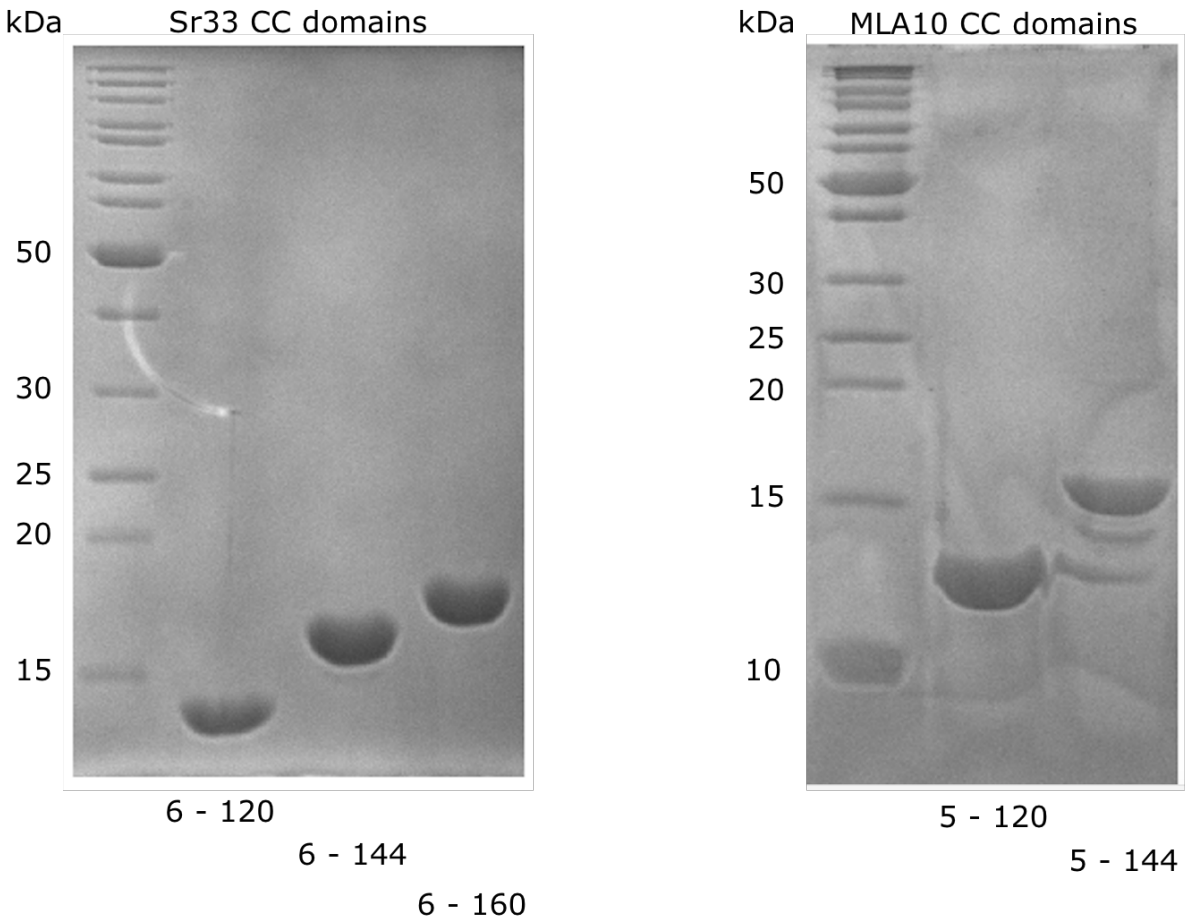
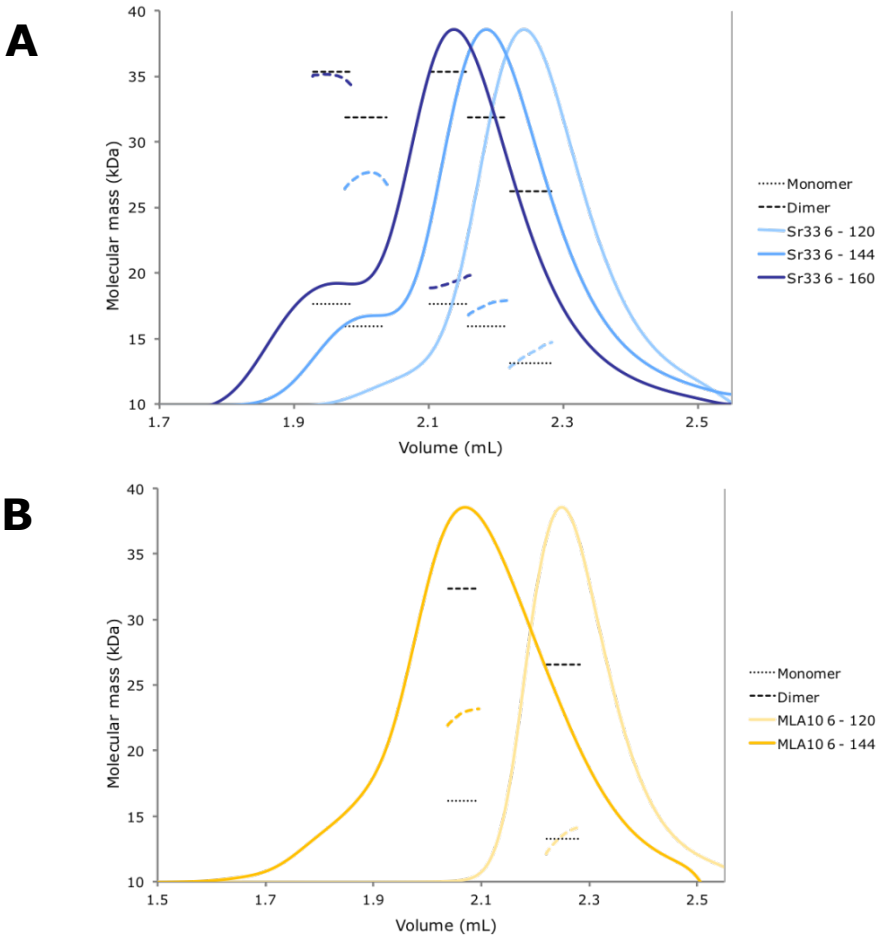


Figure 3.11 Longer CC domain constructs from Sr33 and MLA10 run on SDS-PAGE gel stained with Coomassie blue. Proteins were run on a 15 % SDS-PAGE gel next to an NEB Broad Range protein ladder.



C

	Theoretical Molecular Mass (kDa)		Measured Molecular Mass (kDa)	
Construct	Monomer	Dimer	Shoulder Peak	Major Peak
Sr33 6 -120	13.12	26.24	-	13.9
Sr33 6 - 144	15.94	31.88	27.3	17.5
Sr33 6 - 160	17.65	35.30	35.0	19.3
MLA10 5 -120	13.28	26.56	-	13.4
MLA10 5 - 144	16.17	32.34	-	22.8

Figure 3.12 Multi-angle laser light scattering analyses of the C-120, C-144 and C-160 CC domain constructs of Sr33 and MLA10. Direct refractive index (dRI) on the secondary axis (normalised units) shows the elution of the protein from the column and is represented by a solid coloured line. The molecular mass of the protein is calculated across the elution peak and represented as a dashed line of the colour corresponding to the dRI trace. Theoretical monomer and dimer sizes for each protein are represented by black dashed lines across the peak. (A) MALS analysis of $Sr33^{6\text{-}120}$ (light blue), $Sr33^{6\text{-}144}$ (blue), and $Sr33^{6\text{-}160}$ (dark blue). Shoulder peaks can be observed for each of the C-144 and C-160 CC domain constructs, which have molecular masses close to the expected dimer size of each protein. (B) MALS analysis of $MLA10^{5\text{-}120}$ (yellow) and $MLA10^{5\text{-}144}$ (gold) CC domains. Unlike the extended Sr33 CC domains, the $MLA10^{6\text{-}144}$ CC domain does not separate into two peaks. However, the $MLA10^{5\text{-}144}$ peak exhibits a higher molecular mass than the predicted monomeric size, indicating self-association of the protein within the peak.

3.3.3.2 The CC domains of Sr33 and MLA10 require additional residues for auto-activity in planta

In parallel to the MALS analyses on the longer CC domain constructs of Sr33 and MLA10 performed here, our collaborator Dr. Stella Cesari (CSIRO Agriculture and Food) performed a series of *in planta* transient expression assays using different length CC domain constructs of Sr33, MLA10, and the orthologous rye NLR, Sr50 (Figure 3.13).

As previously reported, the MLA10$^{1-160}$ CC domain presents an autoactive phenotype when transiently expressed in tobacco (Maekawa et al., 2011a, Bai et al., 2012). Dr. Cesari confirmed the auto-activity of this MLA10$^{1-160}$ construct *in planta*, and further demonstrated the Sr33$^{1-160}$ and Sr50$^{1-163}$ CC domains are also capable of cell death signalling when transiently expressed in tobacco. Expression of constructs containing the residues up to the end of the α4 helix, as predicted by COILS (residues 1 – 144), revealed a similar result to the 1 – 160 constructs, with Sr33$^{1-144}$, MLA10$^{1-144}$, and Sr50$^{1-147}$ all presenting autoactive phenotypes when transiently expressed in tobacco. However, Sr33, MLA10 and Sr50 constructs with a truncated α4 helix lost the ability to signal cell death. This result was seen for constructs expressing residues equivalent to 1 – 130, 1 – 135, all the way to 1 – 141, with cell death signalling activity restored in the expression of residues 1 – 142 in all three proteins tested (Figure 3.13).

Coimmunoprecipitation (CoIP) experiments using CFP and HA tagged versions of Sr33, MLA10, and Sr50 were performed for each of the truncations tested in the transient expression assays. The ability to initiate cell death signalling correlated with the proteins ability to self-associate via CoIP, with sensitivity lost for any CC domains with a α4 truncation shorter than 142 residues (Appendix 3).

The work by Dr. Cesari clearly defines a minimal functional unit consisting of residues 1 – 142 (or equivalent) for the CC domains of Sr33, MLA10, Sr50.

To provide biophysical support for the for the *in planta* data acquired by Dr. Cesari, I designed constructs of the Sr33 CC domain consisting of sequential truncations of the α4 helix in an attempt to identify the minimum required residues for the self-association of the CC domain. Sr33 CC domain constructs consisting of residues 6 – 141, 6 – 140, 6-139, 6-138, and 6-128 were expressed and purified in the same conditions as the $Sr33^{6-144}$ CC domain and subjected to SEC-MALS. To increase the separation between Sr33 CC domain monomer and dimer peaks, a higher resolution column than that previously used was acquired to perform SEC-MALS (Superdex S75 Increase GL 5/150).

Each of the Sr33 CC domain truncations tested demonstrated the ability to self-associate to some degree (Figure 3.14; Table 3.2). Sr33 CC domain protein truncations of residues 141 – 139 demonstrated no major difference in the average molecular mass of the dimer peak compared to expected dimer sizes. Each of these proteins demonstrated a strong ability to self-associate, with molecular masses only 2 – 10% below the predicted dimeric mass. However, $Sr33^{6-138}$ and $Sr33^{6-128}$ each demonstrated a large reduction in average molecular masses across the dimer peak, indicating dimer formation is less pronounced than that observed in the longer constructs (~26% and 33% lower than predicted dimeric mass, respectively). Interestingly, the $Sr33^{6-138}$ truncation removes a series of hydrophilic residues from the α4 helix, indicating the possibility of an important role of these residues in the self-association of the CC domain.

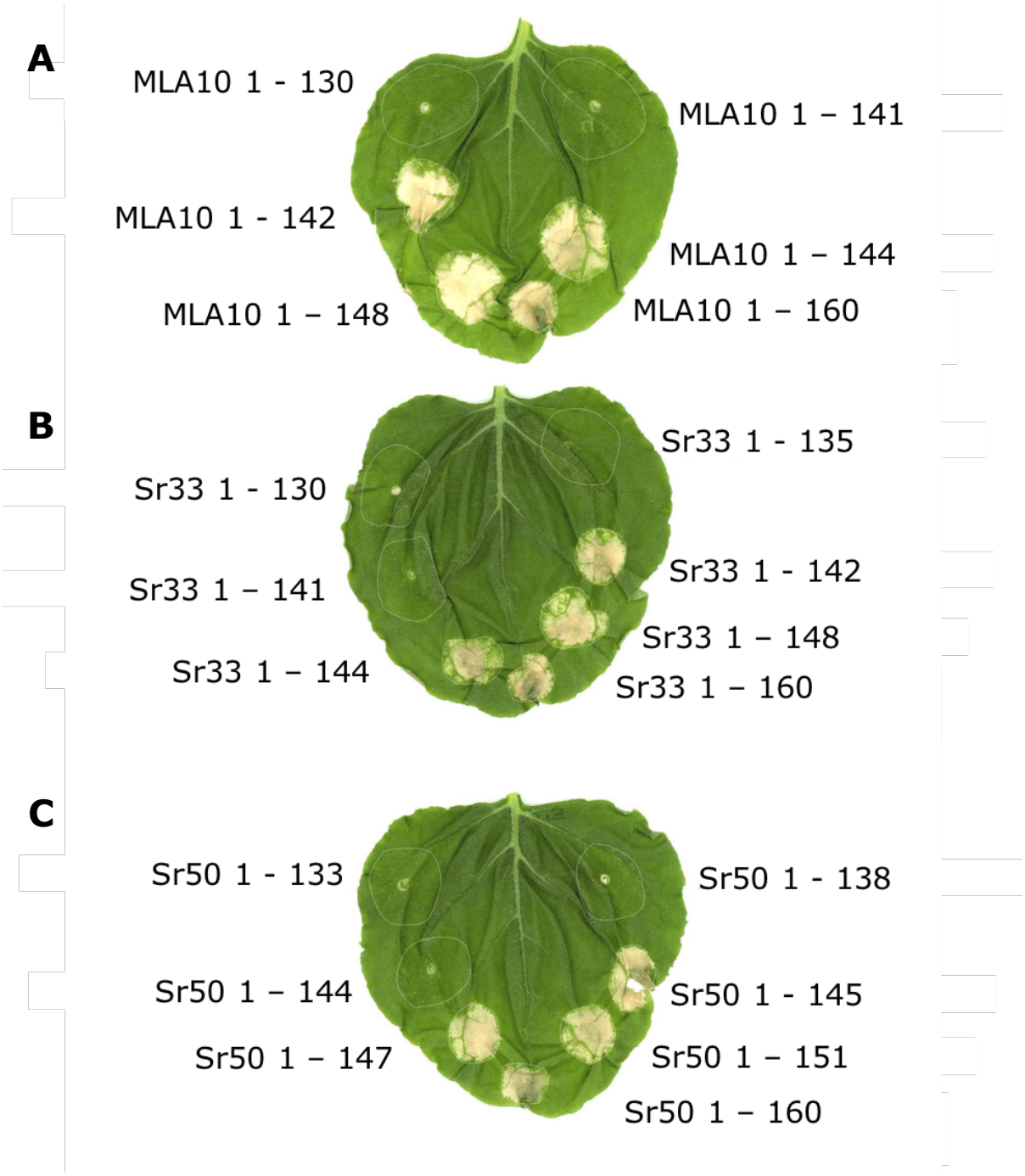


Figure 3.13 *In planta* analysis of autoactive signalling in MLA10, Sr33, and Sr50 CC domains performed by Dr. Stella Cesari, CSIRO Agriculture and Food, Canberra.
(**A**) Transient expression of MLA10 CC domain constructs ranging from 1- 130 to 1 – 160 via agrobacterium-mediated transformation of *N. benthamiana*. (**B**) Same as (**A**) but for Sr33. (**C**) same as (**A** & **B**) but for Sr50.

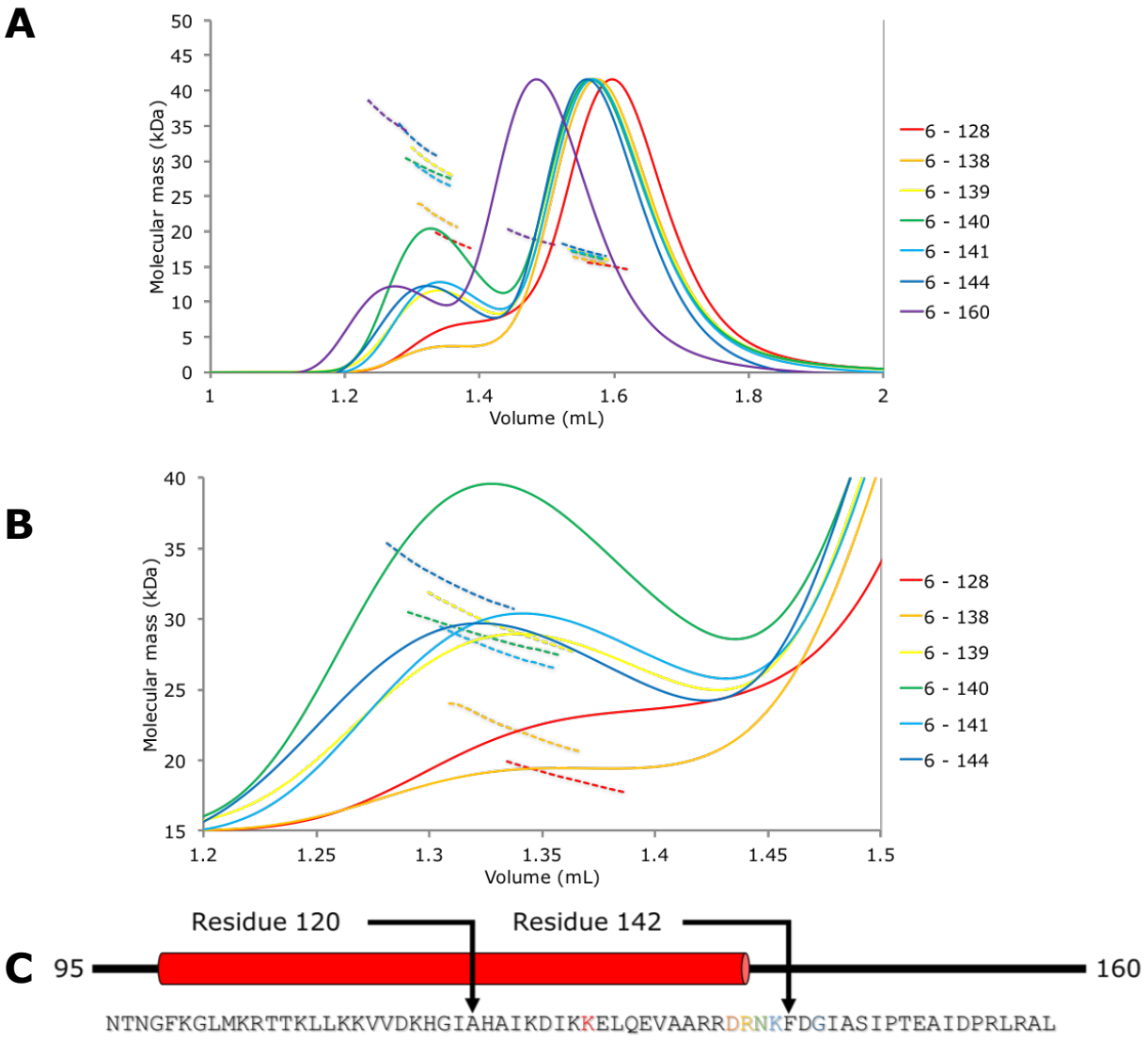


Figure 3.14 Truncations of the α4 helix of the Sr33 CC domain affect self-association. (**A**) MALS analysis of sequential CC domain truncations. Direct refractive index (dRI) on the secondary axis (normalised units) shows the elution of the protein from the column and is represented by a solid coloured line. The molecular mass of the protein is calculated across the elution peak and represented as a dashed line of colour corresponding to the dRI trace. Shorter constructs are in warm colours starting in red ($Sr33^{6\text{-}128}$) and progressing to longer constructs represented by cooler colours ending in purple ($Sr33^{6\text{-}160}$). (**B**) Same as (**A**) however only showing the dimer peaks of each of the truncations. (**C**) Diagrammatic representation of the α4 helix of the Sr33 CC domain, generated from PSIPRED secondary structure prediction, with the corresponding sequence below. Black arrows indicating residues 120 and 142 highlight the terminal residue of the minimal functional units proposed in Maekawa et. al. (2011) and Casey et. al. (2016), respectively. Coloured residues correlate to the terminal residues of the constructs tested via MALS in (**A**) and (**B**).

Table 3.2 Molecular masses of the Sr33 CC domain truncations calculated by SEC-MALS compared with theoretical monomer and dimer masses.

	Theoretical Molecular Mass (kDa)		Measured Molecular Mass (kDa)		Size Difference Compared to Theoretical Molecular Mass (%)	
Construct	Monomer	Dimer	Monomer Peak	Dimer Peak	Monomer	Dimer
6 - 128	14.05	28.10	15.11	18.72	7.54	-33.38
6 - 138	15.22	30.44	15.60	22.23	2.50	-26.97
6 - 139	15.37	30.74	16.64	29.59	8.26	-3.74
6 - 140	15.49	30.92	16.58	28.84	7.04	-6.73
6 - 141	15.62	31.24	16.75	27.83	7.23	-10.92
6 - 144	15.94	31.88	17.30	32.69	8.53	2.54
6 - 160	17.65	35.30	18.99	36.09	7.59	2.24

3.3.3.3 Mutations in the α4 helix of the Sr33 CC domain perturb self-association

The α4 helix in the minimal functional unit of the Sr33 CC domain contains a series of conserved hydrophobic and hydrophilic residues, including a long hydrophilic stretch spanning residues 136 – 141 (RRDRNK) (Figure 3.15 A). To test the importance of these residues in the ability of the Sr33 CC domain to self-associate and signal cell death, Dr. Stella Cesari introduced a combination of charge-swap and alanine mutations in a $Sr33^{1-160}$ background (Cesari et al., 2016).

$Sr33^{1-160}$ CC domain mutants A122E/I123E, I126E, L130E, V133E/A134E, F142E, and a six-alanine mutant of the hydrophilic patch (RRDRNK/6A) all disrupted cell death signalling, with no autoactive phenotypes observed for any of the mutants when transiently expressed in tobacco (Figure 3.15 B). Furthermore, all mutations prevented protein association via CoIP, suggesting these mutations effect the ability of the Sr33 CC domain to self-associate (Cesari et al., 2016). As with the Sr33 CC domain α4 helix truncations, these mutations were tested for their ability to self-associate in solution.

The mutants were generated in the $Sr33^{6-144}$ CC domain as there was no major difference between the abilities of the $Sr33^{6-144}$ and $Sr33^{6-160}$ CC domains to self-associate. Furthermore, higher yields of protein were achieved for the $Sr33^{6-144}$ protein. As such, $Sr33^{6-144}$ will herein be referred to as the wild-type Sr33 CC domain. All the mutations designed by Dr. Cesari, generated protein that could be expressed and purified with the exception of L130E which could not be produced in a soluble form.

All the Sr33 CC domain mutants tested affected the ability of the CC domain to self-associate, however, they did not completely

disrupt dimer formation (Figure 3.16 A & B). As seen with the CC domain truncations, the self-association of the CC domain mutants was not abolished, but impaired in comparison to the wild-type Sr33 CC domain. Each of the mutants demonstrated between ~17% - 25% reduction in average molecular mass across the dimer peak when compared to the wild-type protein, indicating the self-association of the CC domain mutants is compromised (Figure 3.16 C). The exception to this is the RRDRNK/6A mutant from which no accurate molecular mass could be calculated from the dimer peak, due to low light scattering signal, possibly caused by the reduced ability to self-associate.

These data correlate closely with the *in planta* mutant CoIP data published by Cesari et al. (2016). The combination of the transient expression assays with the MALS analysis of the Sr33 CC domain mutants provides strong evidence that the conserved residues in the C-terminal region of the α4 helix of the CC domain are important for the self-association and cell death signalling of the protein.

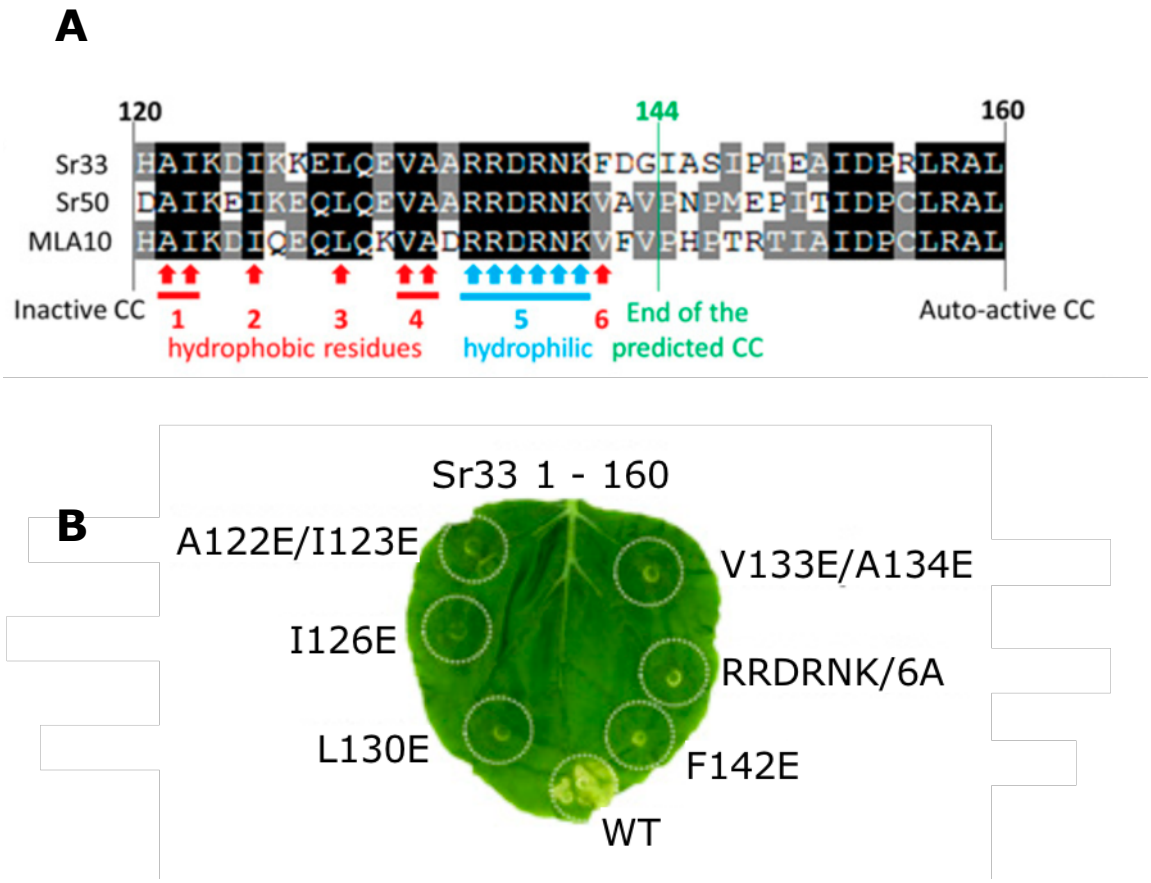
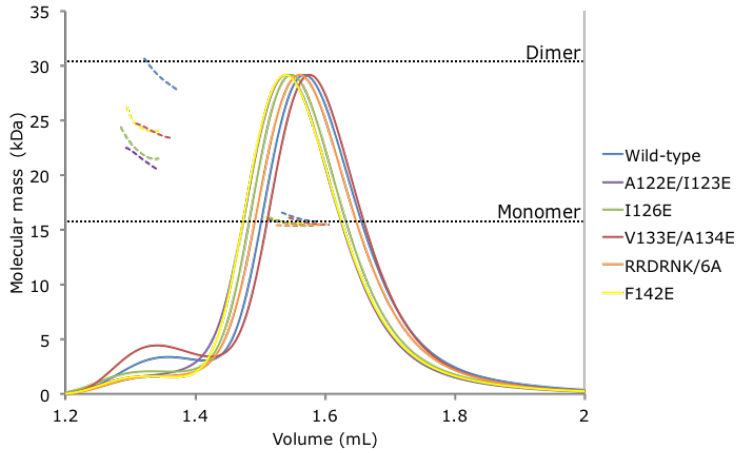
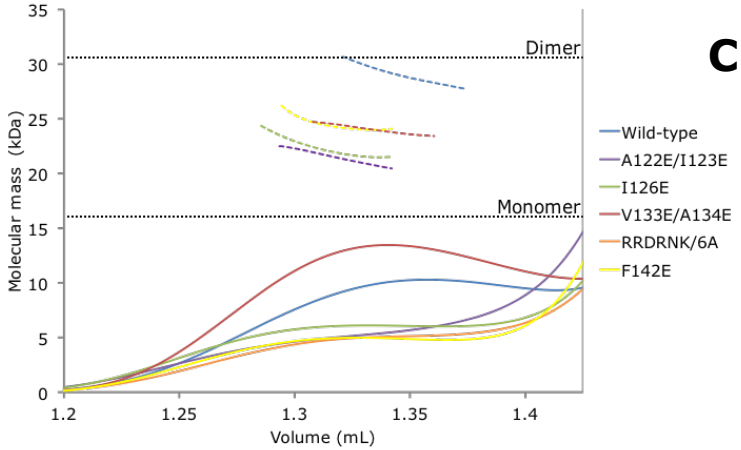


Figure 3.15 Mutations in the α4 helix of the CC domain prevent cell death; image adapted from Cesari et. al. (2016). (**A**) Multiple sequence alignment of the C-terminal end of the last predicted helix of the Sr33, MLA10 and Sr50 CC domains. Conserved hydrophobic and hydrophilic residues are annotated in red and blue, respectively. (**B**) Transient expression of Sr33 CC domain mutants in tobacco reveals mutations to the conserved residues of the final helix disrupt cell death signalling.

A



B



C

Construct	Measured Molecular Mass (kDa)		Size Difference Compared to Wild-type (%)	
	Monomer Peak	Dimer Peak	Monomer	Dimer
Wild-type	16.08	28.99	-	-
A122E/I123E	15.70	21.48	-2.37	-25.89
I126E	15.68	22.39	-2.46	-22.77
V133E/A134E	15.69	24.03	-2.43	-17.11
RRDRNK/6A	15.45	-	-3.88	-
F142E	15.78	24.52	-1.86	-15.41

Figure 3.16 Mutations in the Sr33 CC domain C-terminal helix disrupt self-association in-solution. (**A**) SEC-MALS of Sr33 CC domain α4 helix mutants in comparison to wild-type Sr33 CC domain ($Sr33^{6-144}$). Direct refractive index (dRI) on the secondary axis (normalised units) shows the elution of the protein from the column and is represented by a solid coloured line. The molecular mass of the protein is calculated across the elution peak and represented as a dashed line of colour corresponding to the dRI trace. Expected monomer and dimer sizes for the proteins are shown as black dashed lines across the graph. (**B**) Same as (**A**) however only showing the dimer peak to accentuate the effects of the mutations on dimer formation. (**C**) A table of the measured molecular mass for the wild-type Sr33 CC domain compared to the measured molecular masses of C-terminal helix mutants. All mutants showed a >15 % decrease in molecular mass across the elution peak compared to the wild-type protein, with the exception of RRDRNK/6A, from which no dimer mass could be calculated.

3.3.4 Crystallisation of the active Sr33 CC domain

The biophysical analyses combined with the *in planta* studies of the Sr33 and MLA10 CC domains give compelling evidence that the minimal functional unit of the CC domain extends to residue 142. As there is only structural information of the first 120 residues of the CC domain, the structure of the minimal functional unit of the CC domain is an important structural target. The following research endeavours to address this question with the goal of identifying a key self-association interface.

3.3.4.1 Screening of CC domain constructs for crystallisation

The most desirable outcome of a CC domain crystal structure is the identification of a self-association interface, so we can better understand the molecular mechanism that govern CC self-interactions and the subsequent signalling. To achieve this, CC domain constructs that demonstrated an ability to self-associate via MALS were selected for crystallisation trials. The $Sr33^{6-139}$, $Sr33^{6-140}$, $Sr33^{6-141}$, $Sr33^{6-144}$, $Sr33^{6-160}$ and $MLA10^{5-144}$ proteins were all subjected to a variety of sparse matrix crystallisation screens (Figure 3.17). Unfortunately, no crystals were identified for any of the constructs tested. Interestingly, the majority of conditions tested yielded phase separation, a phenomenon where the protein partitions itself from the crystallisation reagents, forming highly concentrated protein micelles.

Phase separation has been noted to occur close to the point of nucleation required for crystal formation, and crystals can be obtained from phase separation through adjustments to pH, temperature, or protein concentration (Muschol and Rosenberger, 1997, McPherson and Gavira, 2014). In light of this, multiple attempts were made to optimise the crystallisation of these proteins through adjustments to temperature and

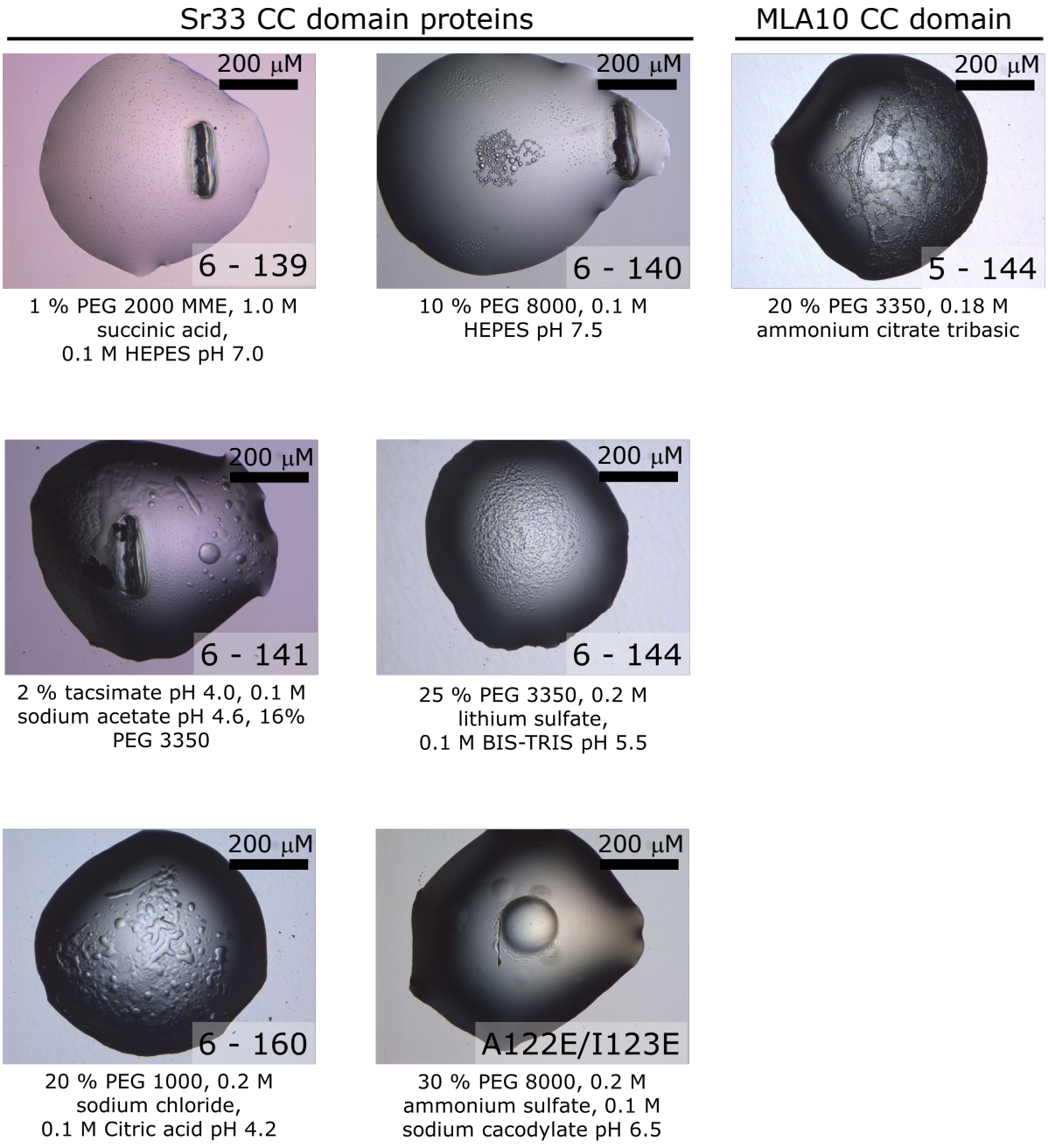


Figure 3.17 Crystallisation trials of different CC domain constructs. All CC domains shown to self-associated via MALS were used in crystallisation trials using a variety of sparse matrix screens supplied by Hampton Research and Molecular Dimensions. Unfortunately, none of the proteins tested were capable of crystal formation. Phase separation was a widely observed phenomenon shared across crystallisation drops of all proteins tested.

protein concentration, however these failed to yield positive results.

In a further attempt to obtain crystals of a full-length CC domain, the A122E/I123E Sr33$^{6-144}$ mutant protein was tested in crystallisation trials. A122E/I123E demonstrated the lowest ability to self-associate via MALS, therefore it was chosen for crystallisation on the hypothesis that a mono-dispersed form of the protein may be more amenable to crystal formation. However, crystallisation screens using the A122E/I123E mutant yielded similar results to the other attempted CC domain crystallisation trials, with the majority of drops presenting phase separation (Figure 3.17). Unfortunately, other Sr33 CC domain mutants were unable to be tested due low protein yields.

3.3.4.2 Alterations to buffer pH change ability to self-associate via MALS and promote crystallisation

It has been noted in the literature that pH plays a significant role in influencing crystallisation of proteins (Newman, 2004). Due to the lack of crystal formation in previous crystallisation screens with multiple constructs of the Sr33 CC domain, the effect of pH on the Sr33 CC domain crystallisation was tested. The purification method of the Sr33$^{6-144}$ protein results in the purified product stored in 10 mM HEPES at pH 7.5 with 200 mM NaCl. For this experiment, the Sr33$^{6-144}$ protein was exchanged into four different buffers, each with a different pH; 0.1 M sodium citrate pH 4.0 and 5.0, BIS-TRIS pH 9, and CHES pH 11, each maintaining the same concentration of 200 mM NaCl. After buffer exchange, the protein was analysed via SEC-MALS to identify any differences in protein oligomerisation. Interestingly, significant peak shifts were observed towards monomeric or dimeric species dependant on the pH of the buffer. Sr33$^{6-144}$ in low pH buffers demonstrated a loss in resolution between monomer/dimer peaks, however the average molecular

mass is approximately 26 kDa, indicating more of the protein is involved into self-association, albeit not forming a stable dimer (Figure 3.18 A). Inversely, $Sr33^{6-144}$ in high pH buffers demonstrated a reduction in the presence of a dimer peak, with a higher proportion of the protein forming a monomeric peak (~17 kDa) (Figure 3.18 A). To check this result was not exclusive to the $Sr33^{6-144}$ protein, $Sr33^{6-141}$ and $Sr33^{6-160}$ were also analysed by SEC-MALS in a buffer with a pH of 4. As seen with the initial experiments with $Sr33^{6-144}$, $Sr33^{6-141}$ and $Sr33^{6-160}$ both lost resolution of monomer and dimer peaks, but presented overall higher average molecular masses across the peak of 23 kDa and 28 kDa, respectively (Figure 3.18 B).

With the intriguing data obtained from MALS, it was decided that crystallisation screens of the $Sr33^{6-144}$ protein would be attempted at a different pH. $Sr33^{6-144}$ in a pH 4.0 sodium citrate buffer was chosen as these conditions appear to have a greater effect on promoting the self-association of the protein in solution, which corresponded with the goal of identifying a dimer interface in the $Sr33^{6-144}$ crystal structure.

Using $Sr33^{6-144}$ in pH 4.0 sodium citrate, some crystalline-like formation was observed in some conditions. These crystals showed growth from 4 days onwards, and developed a dense spherical, walnut-like form (Figure 3.18 C & D). Unfortunately, these crystals are unsuitable for X-ray diffraction studies. Attempts to generate usable crystals through changing protein concentration, precipitant concentration, or alternative crystallisation methods such as streak and micro seeding, failed to improve crystal morphology.

While pH appears to have an effect on the ability of the protein to self-associate, the use of a different pH was not able to produce

crystals satisfactory for X-ray diffraction studies. Due to this, alternative means of crystallisation were pursued.

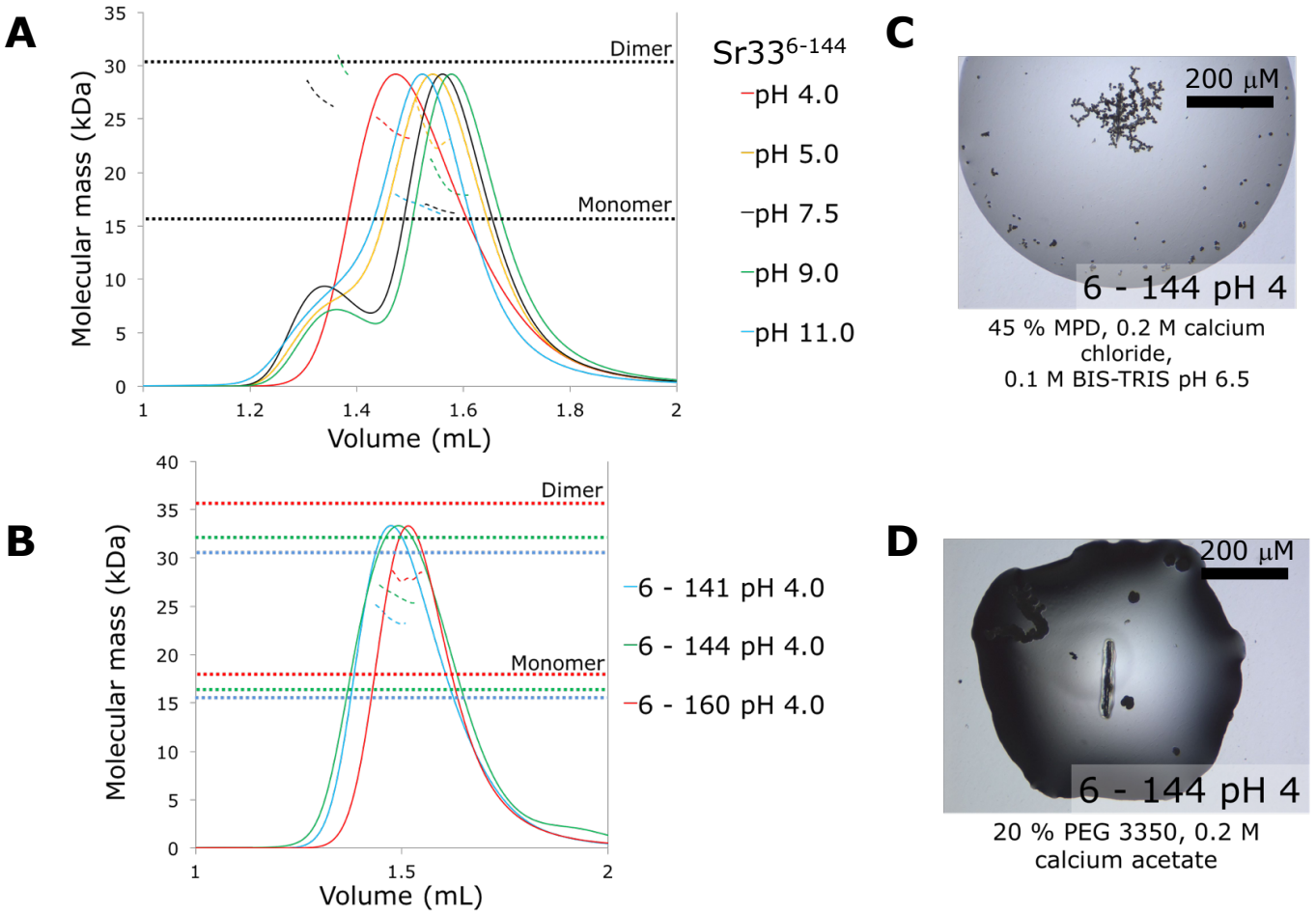


Figure 3.18 The effect of pH on self-association and crystallisation of the Sr33$^{6-144}$ CC domain. (**A**) MALS analysis of Sr33$^{6-144}$ in different buffers with pH of 4, 5, 7.5, 9, and 11. Low pH reduces the separation of the monomer/dimer peaks but increases the overall self-association of the protein. High pH also leads to a loss of peak resolution; however, favours the monomer species of Sr33$^{6-144}$. (**B**) MALS analysis of Sr33$^{6-141}$, Sr33$^{6\ -144}$, and Sr33$^{6-160}$ in a pH 4 buffer. As seen in (**A**), low pH reduces peak resolution to a single peak, however this peak; maintains a molecular mass significantly higher than monomer. (**C & D**) Crystallisation of the Sr33$^{6-144}$ protein in a pH 4 buffer. Walnut-like crystals form after 3 days.

3.3.4.4 In-situ proteolysis of Sr33 yields crystal formation

A report in Nature Methods has described a technique for the improvement of protein samples for crystallisation deemed *in situ* proteolysis (Dong et al., 2007). This study describes the use of trace amounts of the chymotrypsin and trypsin proteases as additives in crystallisation screens, in an attempt to remove regions of proteins not conducive to crystallisation, such as bulky loops or disordered N-terminal or C-terminal regions, via limited proteolysis. Using the technique, the authors reported nine of 20 proteins that were previously unable to be crystallised formed crystals using *in situ* proteolysis, and 30 of 35 proteins that initially yielded poor crystals were capable of forming crystals suitable for diffraction studies (Dong et al., 2007).

As all previous attempts to crystallise the Sr33 CC domain had failed, utilising this technique seemed a reasonable course of action. Initially the sequence of the $Sr33^{6-160}$ CC domain was analysed for chymotrypsin and trypsin cut sides using ExPasy PeptideCutter (Wilkins et al., 1999). PeptideCutter revealed 29 trypsin cut sites in the $Sr33^{6-160}$ CC domain, nine of which lie outside of secondary structures predicted by PSIPRED (Figure 3.19), which indicated in situ proteolysis was a feasible method to remove disordered regions from the Sr33 CC^{6-160} domain.

Crystallisation screens of $Sr33^{6-160}$ were conducted as previously described, but with the addition of trypsin to a final ratio of 1:1000 w/w (approximate concentration of 10 μg/mL), with a negative control of 10 μg/mL trypsin only, to ensure any crystal formation observed in the $Sr33^{6-160}$ drops could be discounted as the formation of trypsin crystals. Crystal formation in drops containing $Sr33^{6-160}$ with trypsin was observed in two conditions (Figure 3.20). Crystals that grew in 20 % PEG 3350, 8 % Tacsimate pH 4.0 formed as long thin hair-like needles and were unable to be

recreated or optimised outside sparse-matrix screens. However, robust, trapezoid crystals formed in 25 % PEG 3350, 0.2 M magnesium chloride, and 0.1 M BIS-TRIS pH 5.5 were able to be consistently reproduced. These crystals initially diffracted to ~12 Å, however optimisation of PEG3350 concentration from 25 % to 30 % yielded an improvement in resolution to 4.2 Å. A dataset was collected on the MX2 beamline at the Australian Synchrotron, however efforts to solve the structure by molecular replacement have been unsuccessful so far. Unfortunately, due to lack of remaining time in this Ph.D. candidature, critical experiments necessary for solving this structure have not been performed. These experiments include mass spectrometry analysis of the trypsin-treated $Sr33^{6\text{-}160}$ crystals to determine the crystallised sequences, and the generation of selenomethionine labelled $Sr33^{6\text{-}160}$ proteins for anomalous diffractions studies. Production of selenomethionine labelled protein is more technically demanding and costly than using molecular replacement, however would allow the structure of the Sr33 CC domain to be solved with *de novo* phasing using SAD. SAD phasing would be a better alternative to molecular replacement as it may be impossible to solve the structure of the Sr33 CC domain using molecular replacement because no prior CC domain structure encapsulates the entire CC domain.

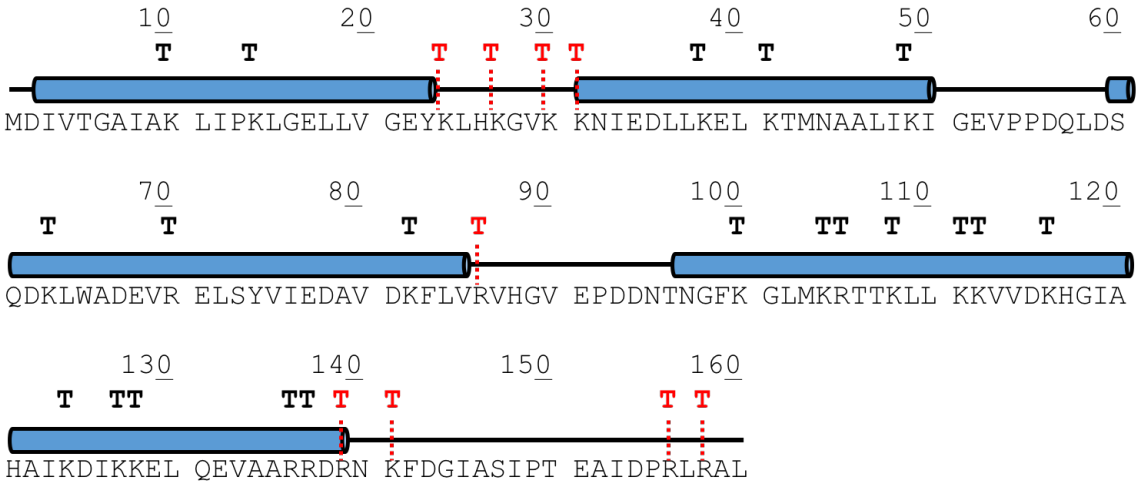
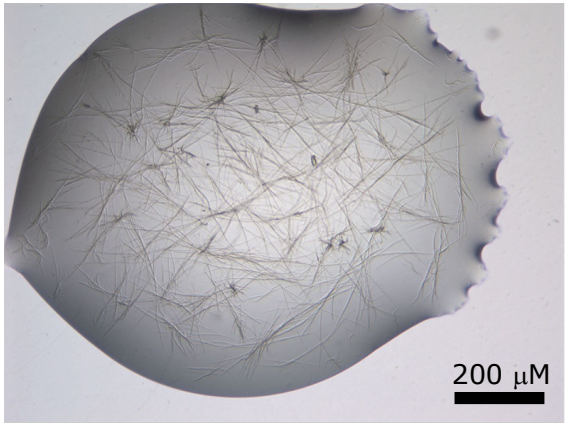
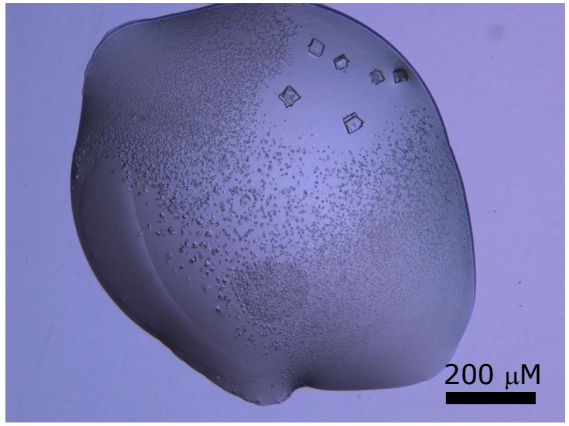


Figure 3.19 Trypsin cleavage sites in the Sr33$^{6-160}$ CC domain. The sequence of Sr33$^{1-160}$ overlayed with the secondary structure prediction obtained by PSIPRED. Predicted trypsin sites are marked with a T above the cleavage site. Nine cut sides lie outside secondary structure boundaries; these cut sites are labelled in red and indicate potential loops which may be removed from Sr33$^{6-160}$ by trypsin.



20 % PEG 3350, 8 % tacsimate pH 4.0



25 % PEG 3350, 0.2 magnesium chloride, 0.1 M BIS-TRIS pH 5.5

Figure 3.20 Crystals of the Sr33$^{6\text{-}160}$ CC domain obtained via *in situ* proteolysis. Crystallisation screens were set up with 10 mg/mL Sr33$^{6\text{-}160}$ with 10 μg/mL trypsin. Crystal formation occurred approximately eight days after the start of the experiment.

3.4 Discussion

Prior to this work it was understood that the MLA10$^{5-120}$ and Rx$^{1-122}$ CC domains have different structures and oligomeric states, and differ functionally (Rairdan et al., 2008, Maekawa et al., 2011a, Bai et al., 2012, Hao et al., 2013). The MLA10$^{5-120}$ CC domain crystal structure revealed an obligate dimer consisting of two helix-loop-helix monomers, and co-migrated with protein standards of expected dimer size in analytical SEC (Maekawa et al., 2011a). In contrast, the Rx$^{1-122}$ CC domain was crystallised as a heteromeric complex, with Rx$^{1-122}$ forming a four-helix bundle interacting with the WPP domain of its cofactor RanGAP2. Rx$^{1-122}$ was purified as a monomer, and showed no evidence of self-association (Hao et al., 2013). The work in this chapter aimed to expand our understand of CC domain structure/function through analysis of the Sr33$^{6-120}$ CC domain, an ortholog of the MLA10$^{5-120}$ CC domain; they share 80% sequence identity.

3.4.1 The structured CC domains of Sr33, MLA10, and Rx CC are monomeric in-solution.

Using both SEC-MALS and -SAXS techniques, it was shown the Sr33$^{6-120}$, MLA10$^{5-120}$ and Rx$^{1-122}$ are monomeric in solution. These data conflict with analytical SEC experiments performed by Maekawa et. al. (2011), which support the formation of a MLA10$^{5-120}$ CC domain dimer. We also observed a very similar migration and elution profile for the CC domains as Maekawa et. al. (2011) utilising analytical SEC. While it could be considered that the analytical SEC and SEC-MALS results conflict with one another, it is important to note migration of a protein through SEC column is influenced by many variables other than size. Size is often not the only determining factor of migration through a SEC column; protein migration is also effected by protein shape, flexibility, and interaction with the SEC column (Wen et al., 1996). Because of these factors, it was decided molecular mass should be measured

using the scattering techniques MALS and SAXS, as they both directly relate signal to average molecular mass of the protein, providing a far more robust and accurate measurement of molecular mass. In the context of this data, it would appear the shape of the CC domains allowed for a faster elution from the SEC column, regardless of their molecular mass, giving the illusion they are larger than they actually are.

While this discovery reconciles the findings of previous reports on the oligomeric states MLA10$^{5-120}$ and Rx$^{1-122}$ CC domains, it also immediately brings into question the differences in structures of MLA10$^{5-120}$ and Rx$^{1-122}$ CC domains. The crystal structures of these proteins have been reported to have different folds, yet our data demonstrates that in solution the folds and oligomeric states are similar. To address this, it was decided the structures of the Sr33$^{6-120}$ and MLA10$^{5-120}$ CC domains should be characterised further to understand these inconsistencies.

3.4.2 A four-helix bundle best represents the structure of CC domains in solution

To further investigate the observed differences with the published literature of MALS data of the C-120 CC domains, small-angle X-ray scattering (SAXS) was performed on the Sr33$^{6-120}$, MLA10$^{5-120}$, and Rx$^{1-122}$ CC domains, as this is an excellent tool for the comparison of protein shape, fold and oligomeric state. The SAXS data correlated with the MALS analyses, demonstrating all the CC domains are monomeric in solution (Figure 3.5). Furthermore, the SAXS data for each of the CC domains are nearly identical, with almost equivalent distribution functions, and radius of gyration calculations for each C-120 CC domains within 0.5 of an angstrom. This data definitively shows that the Sr33$^{6-120}$, MLA10$^{5-120}$, and Rx$^{1-122}$ CC domains have a similar fold in solution.

The Sr33 CC domain could not be crystallised; however, the 3D structure determined by NMR revealed a monomeric four-helix bundle fold, unlike the CC domain of its ortholog, MLA10$^{5-120}$, and much more similar to the monomeric four-helix bundle structure of the Rx$^{1-122}$ CC domain (Figure 3.8).

The crystal structure of the MLA10$^{5-120}$ CC domain conflicts with the in-solution MALS and SAXS data, and the NMR structure of the orthologous Sr33$^{6-120}$ CC domain is significantly different to the MLA10$^{5-120}$ structure; therefore, it was determined to be prudent to perform further experiments to validate each of the structures. To this extent, theoretical SAXS datasets were generated from the crystal structures of MLA10$^{5-120}$ and Rx$^{1-122}$, and the NMR structure of Sr33$^{6-120}$ (Figure 3.9). An overlay of the scattering data generated from the structures of Sr33$^{6-120}$, MLA10$^{5-120}$, and Rx$^{1-122}$ with the experimentally determined SAXS data demonstrated the four-helix bundle structures of the Sr33 and Rx CC domains fit their corresponding experimentally acquired datasets for each protein exceptionally well. In contrast, neither of the SAXS datasets generated from a monomer or dimer based on the crystal structure of MLA10 are consistent with the SAXS dataset obtained for MLA10$^{5-120}$ CC domain. Furthermore, an overlay of the theoretical scattering data generated from the Sr33$^{6-120}$ NMR structure with the MLA10$^{5-120}$ experimental SAXS dataset shows an excellent fit (Casey et al., 2016) (Appendix 3, Figure 3).

Taken together, these data strongly suggest the MLA10$^{5-120}$ CC domain structure published by Maekawa et. al. (2011) does not represent the fold of the MLA10$^{5-120}$ CC domain in solution. Furthermore, it is clear the four-helix bundle structures of Rx$^{1-122}$ and Sr33$^{6-120}$ CC domains very closely denote the overall fold of the proteins in solution. Finally, the SAXS data very clearly shows each of these proteins, Sr33$^{6-120}$, MLA10$^{5-120}$, and Rx$^{1-122}$, share a similar

fold in solution, and therefore it can be deduced the MLA10$^{5\text{-}120}$ CC domain also shares a monomeric, four-helix bundle structure in solution.

As for the structure of the MLA10$^{5\text{-}120}$ CC domain, it is possible this structure formed serendipitously under the conditions of crystallization through a domain swap. The same hydrophobic residues that mediate the dimer interface in the MLA10$^{5\text{-}120}$ dimer also mediate the formation of the hydrophobic core of the Sr33$^{6\text{-}120}$ four-helix bundle (Figure 3.21). Another possibility is that the MLA10$^{5\text{-}120}$ CC domain structure represents an a different "active" conformational fold, promoted by the ingredients of the crystallisation condition. However, without substantial structural or functional evidence to prove either scenario, any conclusions remain speculative.

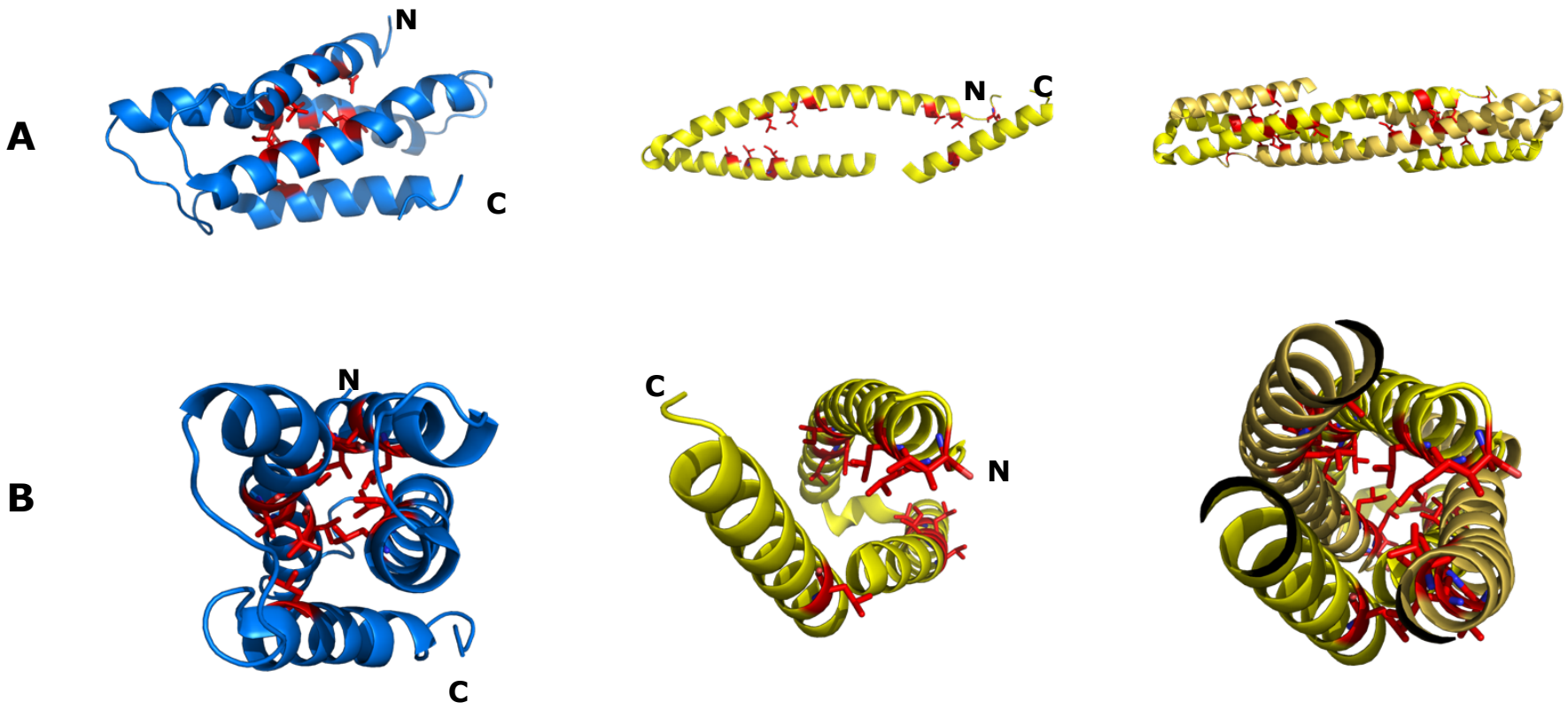


Figure 3.21 Hydrophobic residues that form the Sr33$^{5\text{-}120}$ hydrophobic core also form the dimerisation interface in the MLA10$^{5\text{-}120}$ crystal structure. (**A**) Side view of the Sr33$^{6\text{-}120}$ CC domain NMR structure (blue) (PDB ID: 2NCG), compared with the MLA10$^{5\text{-}120}$ CC domain monomer (yellow) and dimer (yellow and gold) (PDB ID: 3QFL). Hydrophobic residues that are conserved between Sr33 and MLA10 are coloured in red. (**B**) Same as (**A**) except showing the front view looking through the hydrophobic core of the Sr33$^{6\text{-}120}$ CC domain, and the dimer interface of the MLA10$^{5\text{-}120}$ CC domain. The amino and carboxyl termini of the proteins are indicated by N and C resepctively.

3.4.3 Self-association of the CC domain mediates cell death signalling, but requires the inclusion of additional residues at the C-terminus

Self-association of the MLA10 CC domain has been reported to be important for cell death activity (Maekawa et al., 2011a, Bai et al., 2012). However, from the biophysical analyses performed in this chapter, there is no evidence the $Sr33^{6-120}$ CC domain or the $MLA10^{5-120}$ CC domain are capable of self-association. Because of its propensity to be monomeric in solution, it is possible the $MLA10^{5-120}$ CC domain does not represent a minimal functional unit for self-association and cell death signalling as proposed by Maekawa et. al. (2011). This led to the reassessment of the potential boundaries of the CC domain through the use of the secondary structure prediction programs COILS and PSIPRED (Figure 3.10) (Lupas et al., 1991, Jones, 1999). Interestingly, it was found the $MLA10^{5-120}$ and $Sr33^{6-120}$ CC domains constructs truncate the α4 helix of each respective CC domain, which is predicted to extend to residues 140 - 144, dependent on the prediction software used. As removing a significant portion of a helix can have a significant impact on protein structure and function, new constructs encompassing residues 6 – 144 were designed. In conjunction with this, previous reports of MLA10 CC domain activity have only demonstrated constructs containing the first 160 residues to be autoactive *in planta*, with no data for the autoactivity of the $MLA10^{5-120}$ CC domain (Maekawa et al., 2011a, Bai et al., 2012). As self-association has always been associated with cell death in plant NLR signalling, CC domain constructs containing the first 160 residues of Sr33 and MLA10 were generated to assess their ability to self-associate. This experiment aimed to understand if CC domains capable of cell death signalling are able to self-associate.

MALS analyses of $Sr33^{6-144}$, $Sr33^{6-160}$, and $MLA10^{5-144}$ CC domains demonstrated that all these proteins self-associated in solution, with evidence of a dimer peak not seen in in the C-120 constructs (Figure 3.12). This gives a clear indication that this final helix plays a role in the self-association of the CC domain, and the truncation of this helix in the C-120 constructs results in the proteins not being able to self-associate, and thus are unable to signal.

In parallel to this work, the ability to initiate cell death when transiently expressed in tobacco was tested for a series of Sr33, MLA10 and Sr50 CC domain truncations. In all cases, CC domains that did not contain the entire predicted α4 helix were unable to signal cell death (Figure 3.13). Furthermore, it was shown that a minimum of 142 residues in Sr33 CC domain constructs was required to facilitate cell death signalling when transiently expressed in tobacco. This observation was also found for the equivalent length constructs of MLA10 and Sr50, which also demonstrated autoactive phenotypes with the inclusion of these additional α4 helix residues.

Co-immunoprecipitation assays with CFP and HA labelled proteins were used to assess the ability of CC domain constructs to self-associate *in planta*. Dr. Cesari demonstrated CC domain constructs that cannot signal cell death cannot interact with each other in CoIP assays (Casey et al., 2016). To confirm the role of the α4 helix of the $Sr33^{6-144}$ CC domain in self-association, sequential truncations from the C-terminus of the protein were performed and tested for their ability to self-associate via MALS. All the CC domain truncations tested demonstrated some ability to self-associate; however, truncating the CC domain past residue 139, cutting into the the α4 helix, resulted in a significant drop in average molecular mass when measured by MALS (Figure 3.14). This indicates the importance of these C-terminal residues in the self-association of

the CC domain, and likely are involved in forming a self-association interface which is yet to be structurally characterised.

It is clear that the MLA10$^{5-120}$ CC domain does not represent the minimal functional unit required for homodimerisation and cell death signalling as first proposed by (Maekawa et al., 2011a). The CC domain requires a minimum of 142 residues to signal cell death response *in planta*, and requires the entire final helix, which is truncated in the MLA10$^{5-120}$ CC fragment, to facilitate self-association.

3.4.4 Mutations to conserved residues in the final helix of Sr33 hint at an interface important for self-association

The *in planta* data combined with the MALS analyses of the Sr33$^{6-144}$, Sr33$^{6-160}$, and MLA10$^{5-144}$ proteins demonstrates the CC domain is capable of autoactive signalling, and the ability of the protein to self-associate is critical to this function. Multiple sequence alignments of the C-terminus (residues 120 – 142) of the α4 helix show the residues in this region are highly conserved between Sr33, MLA10, and Sr50 (Figure 3.15). It was found that self-association of the CC domain could be perturbed by the mutations to the conserved residues of the α4 helix, and this further correlated with *in planta* data, which demonstrated the same mutations disrupted cell death signalling and interactions via CoIP (Figure 3.15; Figure 3.16).

This work, combined with the Sr33 CC domain truncation data, strongly suggest the conserved residues of the α4 helix are involved in the formation of a self-association interface required for cell death signalling. Interestingly, the CC domain mutants that disrupt self-association and cell death signalling are analogous to the TIR domain interface mutants and their effects on self-association and signalling. In both cases, CC and TIR domains

maintain transient interactions in solution, without stable dimer formation. Mutations that reduce the ability of the CC and TIR domains to self-associate prevent cell death signalling. The only difference between the CC and TIR datasets is that the CC domains elute from the SEC column in two nearly separate monomer and dimer peaks, whereas TIR domains elute as a single polydisperse peak. This difference is likely due to the SEC-MALS of the CC domains being performed with a higher resolution SEC column than that used for the TIR domain SEC-MALS (a Superdex S75 Increase 5/150 versus a Superdex S200 increase 5/150, respectively). It would be interesting to see if monomer and dimer TIR domain peaks could be observed with a higher resolution SEC column. Regardless, it is becoming more apparent that the CC and TIR domains, while maintaining significantly different structural folds, have similar roles and potentially similar mechanisms for signalling.

The identification of a self-association interface is critical to the understanding of the molecular mechanisms behind CC domain signalling. This work provides robust evidence for the existence of an interface formed by the conserved residues of the α4 helix of the CC domain. However, in light of the findings in this chapter, it is clear that without the structure of an active CC domain, this conclusion is only speculative.

3.4.5 Future work required for the determination of an active CC domain structure

The structure of an active Sr33 CC domain remains unsolved. Multiple attempts to crystallise various truncations and mutants of the protein have failed to yield usable crystals (Figure 3.17). The exception to this was the crystallisation of the $Sr33^{6\text{-}160}$ CC domain using *in situ* proteolysis, and the crystallisation of $Sr33^{6\text{-}144}$ in low pH buffers.

In situ proteolysis yielded crystals that diffracted to low resolution (Figure 3.20). However, the possibility exists that with further optimisation of crystallisation a structure could be determined for the proteolysed Sr33$^{6\text{-}160}$ CC domain. First and foremost, the sequence of the protein that has been crystallised must be determined. Due to the nature of the experiment, it is impossible to know where trypsin has digested the protein, and therefore impossible to know the identity of the crystallised protein without the use of mass spectrometry. Unfortunately, due to time constraints, mass spectrometry analyses of Sr33$^{6\text{-}160}$ CC domain crystals have not been performed.

However, potential Sr33$^{6\text{-}160}$ truncations caused by proteolysis that may have been crystallised can be speculated upon through analysis of the predicted trypsin cut sites within the Sr33$^{6\text{-}160}$ CC domain sequence. It has been reported that trypsin has difficulty digesting native proteins, and rarely cuts into secondary structure (Haurowitz et al., 1945, Dong et al., 2007). As reported by Dong et. al. (2007), limited proteolysis by trypsin regularly removes N- or C-terminal loops, with the possibility of removing some loops in between secondary structures of the proteins without deleterious effects to overall protein fold. It is clear from the combination of PSIPRED secondary structure predictions and PeptideCutter trypsin cut site predictions that there are numerous sites in the Sr33$^{6\text{-}160}$ sequences that lie outside the ordered secondary structures of Sr33$^{6\text{-}160}$, including the possibility of removing a loop between the first and second helix of the CC domain (Figure 3.19) (Wilkins et al., 1999, Buchan et al., 2013). The removal of these loops has possibly facilitated crystal growth by allowing for protein units to pack with greater order than previous constructs would allow. But, as previously mentioned, it will require the mass spectrometry of the protein contained within the crystals to know for certain what has been crystallised through the use of this method.

Crystal formation was also observed in experiments that used $Sr33^{6-140}$ protein that had been buffer exchanged into a low pH buffer of 20 mM sodium citrate, 200 mM, sodium chloride, pH 4.0. However, it was noted that crystals formed rapidly, within 12 hours, and only in conditions that contained calcium (Figure 3.18). As these crystals readily formed but could not be optimised into anything usable for diffraction studies, it is possible they did not consist of protein, but rather were formed of calcium citrate. While MALS analyses showed significant shifts in oligomeric state of $Sr33^{6-160}$ determined by the pH of the buffer used, it is hard to rationalise how these changes in buffer pH would have an effect on crystal formation. Many of the crystallisation conditions that the $Sr33^{6-160}$ protein was subjected to in sparse matrix screens contained low pH buffers; it would be expected these buffers would have an equivalent effect on protein oligomerisation in the drop as seen in the MALS analysis. To this extent, it is difficult to know what is occurring during the crystallisation of the protein at low pH, and further analysis of the protein is required to fully understand what effects low pH is having on the protein in solution.

3.5 Summary

Previous studies of the $MLA10^{5-120}$ and Rx^{1-122} CC domains presented two significantly different structures, which also suggested the proteins have different oligomeric states. Through the elucidation of the $Sr33^{6-120}$ NMR structure and comparative biophysical analyses of $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122} using SAXS and MALS, it is clear that each of these proteins are monomeric, and share the same four-helix bundle fold in solution. Furthermore, the $MLA10^{5-120}$ CC domain is not the minimal functional unit for homodimerisation and cell death signalling as proposed in Maekawa et. al. (2011). It has been clearly shown here

that additional C-terminal residues, to a final length of 142 amino acids, are required for cell death signalling when transiently expressed *in planta*, and these additional residues are required for the self-association of the CC domain. Mutations to the residues of the conserved residues of the α4 helix disrupt self-association, and in turn, perturb cell death signalling, linking the ability of the CC domain to self-associate directly to its ability to signal. The crystal structure of an active CC domain is yet to be determined, and is the focus of ongoing research. This work has reconciled some previous inconsistencies in our knowledge of CC domain function, and contributed to the discovery of the minimal functional unit of the plant NLR CC domain, which is significantly broadened our understanding of this domain.

4. GENERAL DISCUSSION

The work presented here has sought to better understand the molecular mechanisms that govern cell death signalling by the N-terminal domains of plant NLRs. For both TIR and CC domains, significant contributions have been made to the current understanding of how these domains self-associate in order to signal, and corrections have been made to previous misconceptions concerning their structure and function. In this chapter, the importance of these findings will be discussed in the context of the current literature surrounding NLR function.

4.1 The function of the N-terminal domains in the context of a full-length plant NLRs

When it is difficult to study an entire protein, due to limited production yield or solubility, structural studies of single domains are useful for understanding the different functions a protein may perform. All contributions to understanding the function of plant NLRs are important; however, it is necessary to interpret the data acquired from the study of single domains in the context of their entire protein. There may be differences in conformation, oligomerisation, and even the structure, of a single domain when compared to the same domain as a part of the full-length protein. Therefore, it is important to consider and discuss the self-association, and potential high-order oligomerisation, of CC and TIR domains in the context of the full-length NLR.

4.1.1 TIR domain self-association in the context of NLR oligomerisation

The discovery that several TIR domains have multiple self-association interfaces that are simultaneously required for cell death signalling suggests these protein function through oligomerisation. As shown in Figure 2.10 in Chapter 2, a model of a TIR domain superhelix can be formed through the alternating

stacking of the AE and DE interfaces. However, this model does not account for the arrangement of other domains in the NLR, which will have significant influence on the orientation of the TIR domain, and ultimately determine the interactions the TIR domain will be capable of forming.

The idea of a helical TIR domain super-structure is not unique to plant NLRs. A recent study has demonstrated the TIR domains of the mammalian Toll-like receptors (TLRs) can associate *in vitro* to form large, helical, open-ended filaments (Ve et al., 2017). Toll-like receptors are a class of transmembrane PAMP-recognition receptors (PRRs) in mammals that signal inflammatory responses through the TIR domain, after detection of non-host molecules by an extracellular LRR domain (Akira et al., 2006, Gay et al., 2014). Ve et. al. (2017) demonstrated the TIR domain of TLR4 nucleates the polymerisation of the TIR domain containing adaptor proteins, MAL (MyD88 adaptor-like protein) and MyD88 (myeloid differentiation primary response gene 88), which self-associate through multiple interfaces to form a large helical assembly (Figure 4.1 A). The mammalian TIR assemblies set a precedent for TIR domain oligomerisation and signalling. However, mammalian TIR domains share significant structural differences to plant TIR domains, especially in the αD helical region, which we known is important for the self-association of plant TIR domains (Ve et al., 2015). As such, the interfaces required for the self-association of these mammalian TIR domains are different to the self-association interfaces observed in plant TIRs, and therefore it is unlikely that plant TIR domains would arrange in an equivalent manner.

To better understand how plant NLRs would assemble to induce interactions between TIR domains, it is more appropriate to draw comparisons to the oligomerisation of animal NLRs. Three dimensional cryo-electron microscopy (cryo-EM) structures are now

available for the entire NLR-like apoptosis activating factor-1 (APAF-1) apoptosome, and the nucleotide oligomerisation domain [NOD]-like receptor [NLR] containing caspase recruitment domain [CARD]-4 (NLRC4) inflammasome (Zhou et al., 2015, Zhang et al., 2015, Hu et al., 2015, Tenthorey et al., 2017).

Similar to plant NLRs, mammalian NLRs have a nucleotide-binding domain (known as the NACHT domain, named after the NAIP, CIITA, HET-E and TP-1 proteins), and a C-terminal LRR domain. These proteins signal; however, through N-terminal CARDs or PYDs (pyrin domains) (Bentham et al., 2017). Also, like plant NLRs, mammalian NLRs exist in an inactive or active conformation, depending on ADP or ATP binding, respectively. The structural studies of the NLRC4 inflammasome by Hu et al. (2015), Zhang et al. (2015) and Tenthorey et al. (2017) revealed NLRC4 is not activated through the direct detection of a ligand, rather is activated by helper NLRs, NLR family inhibitory proteins (NAIPs); specifically NAIP2 and NAIP5 in these studies. In the case of the NAIP2-NLRC4 interaction, NAIP2 becomes active by binding a PAMP, PrgJ, a rod protein of the bacterial T3SS, which then in turn activates NLRC4 via homotypic NACHT domain interactions, resulting in a nucleotide switch from ADP to ATP, causing in a large conformational change in NLRC4. Interestingly, once a molecule of NLRC4 is activated, it is capable of activating other NLRC4 molecules, resulting in auto-assembly of the inflammasome (Zhang et al., 2015, Hu et al., 2015) (Figure 4.1B). This large-scale self-association of NLRC4 through the NACHT domain brings the CARDs of each molecule close enough together to associate, which leads to caspase recruitment and downstream signalling (Zhang et al., 2015, Hu et al., 2015, Tenthorey et al., 2017).

The activation of mammalian NLRs to oligomerise and induce CARD associations could be a very similar mechanism to how plant NLRs

function (Duxbury et al., 2016, Bentham et al., 2017). As discussed in Chapter 1, it is clear that plant TIR domains must be driven to self-associate by external regions, as their ability to self-associate is weak and transient. Studies of mammalian TLR signalling set a precedent for TIR super-structure formation, and therefore it is possible the proximity-induced assembly of a plant TIR-superstructure could be driven by the oligomerisation of the NB-ARC and/or LRR domains. However, it is unknown if plant NLR oligomerisation occurs in a similar manner to that of animal NLRs, or if it would allow for a helical arrangement of TIR domains as seen in TLR signalling, or the model proposed in Figure 2.10. An NLR arrangement similar to animals would could promote either a closed ring-like structure as seen in the studies by Hu et al. (2015) and Zhang et al. (2015), or an open-ended structure as observed by Tenthorney et al. (2017), which would could be conducive to a helical arrangement of TIR domains.

Unfortunately, due to the transient nature of TIR domain self-association, it is unlikely that structural evidence of the formation of a TIR superstructure will be obtained in the absence of the other domains of the NLR. To fully understand whether plant NLRs assemble into large signalling complexes similar to the apoptosome or inflammasome, further studies of entire NLR proteins need to be undertaken.

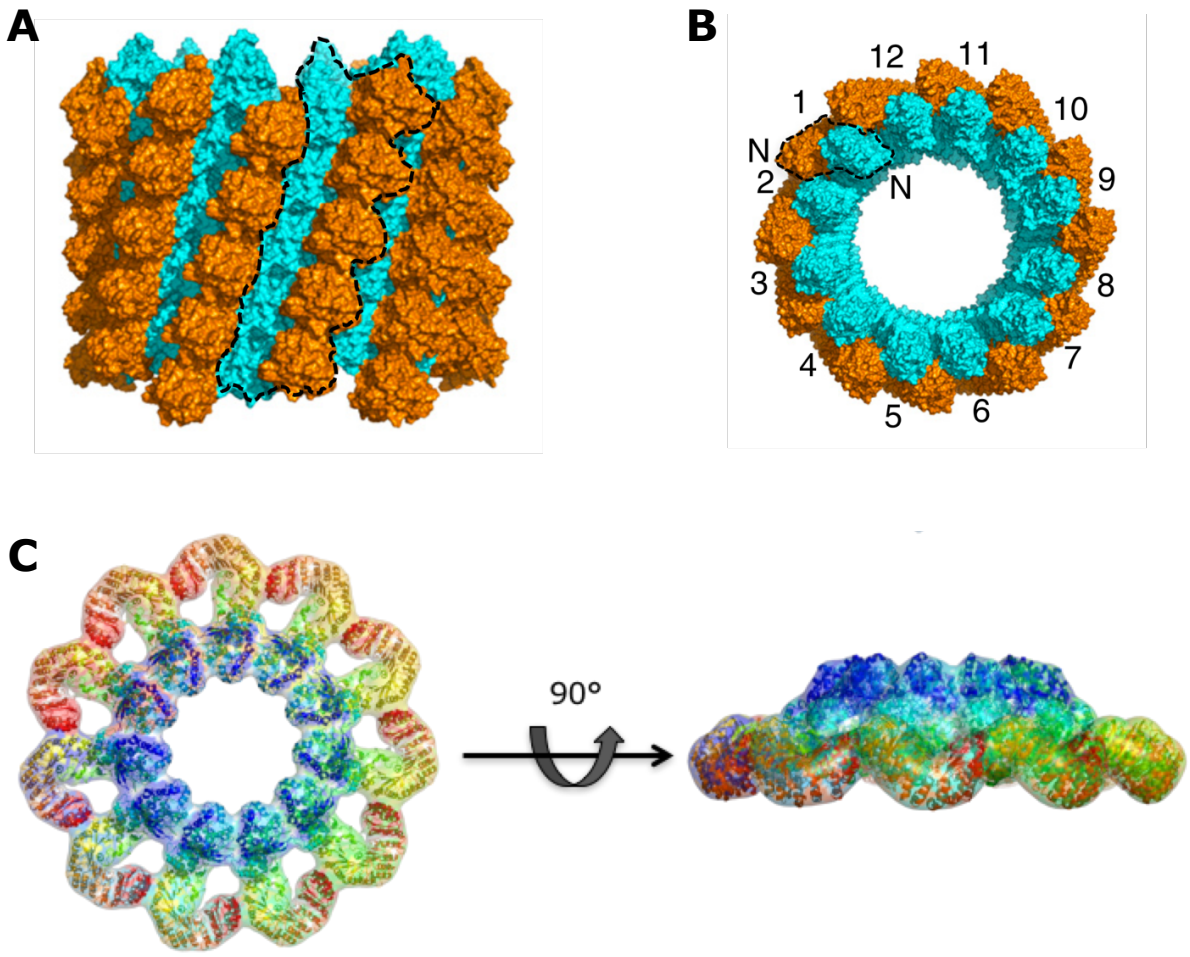


Figure 4.1 Oligomeric structures formed by NLR domains in mammalian immunity. (**A**) Side view of the MAL TIR filament (PDB ID: 5UZB), consisting of 12 two-stranded protofilaments represented by a black dotted line. Inner and outer helical strands are shown in cyan and orange, respectively. (**B**) Top view of the MAL TIR filament. Each protofilament is numbered, with N indicating the position of the N terminus of the inner- and outer-strand subunits. (**C**) The NLRC4 inflammasome (PDB ID: 3JBL). Each inflammasome consists of 10 – 14 monomers which assemble through NACHT domain interactions after activation by NAIP2. Each monomer is coloured from blue to red indicating N-terminus through to C-terminus respectively. Images **A** & **B** are adapted from Ve et. al. (2017); Image **C** is adapted from Bentham et. al. (2017).

4.1.2 Insights into the structure and function of an active CC domain

It has been clearly demonstrated the CC domains of $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122} all share the same fold in solution, and that this fold corresponds to a four-helix bundle, as seen in the $Sr33^{5-120}$ NMR structure, and the Rx^{1-122} crystal structure. Furthermore, these domains do not represent minimal functional units for CC domain signalling, as an additional 22 residues are required for the Sr33 and MLA10 CC domains to cause an autoactive phenotype, when transiently expressed in tobacco. Secondary structure predictions suggest these additional residues are important for the continued formation of the α4 helix, which is necessary for the self-association of the CC domain. Truncations or mutations affecting the conserved residues of the α4 helix prevent cell death and autoactive signalling *in planta.* This region of the protein is likely involved in the formation of a self-association interface; however, the structure of an active CC domain, although attempted in this study, has not been determined, and therefore it is difficult to grasp the precise interactions required for CC domain self-association.

The work in Chapter 3 has identified residues important for the self-association and signalling of the Sr33 CC domain. Mutagenesis of conserved residues in the α4 helix disrupt cell death and perturb self-association. However, MALS of the CC domain mutants shows these proteins still retain the ability to self-associate, although only at a significantly reduced level to that of the wild-type $Sr33^{6-144}$ CC domain. This suggests there are other regions in the protein outside the final helix that are important for self-association. Interestingly, a recent study of the Arabidopsis NLR, RPM1, has demonstrated conserved residues important for the self-association of the RPM1 CC domain that locate to the α2 and α3 helices (El Kasmi et al., 2017). Multiple sequence alignments between Sr33

and RPM1 show significant sequence conservation through this region (El Kasmi et al., 2017). It would be interesting to assess what effects the equivalent mutations in the α2 and α3 helices would have on the self-association of the Sr33 CC domain. These residues could constitute a secondary self-association interface. If this proved to be the case, it would allow for the possibility of CC domain oligomerisation. This would be of significance as there is no functional or structural evidence of CC domain oligomerisation to date, and it would be congruent with the current hypothetical model of plant NLR oligomerisation (Bentham et al., 2017).

As there is no structural knowledge of a CC domain that is capable of inducing cell-death, the structure and function of the CC domain can only be speculated upon. Using the protein structure prediction server (PHYRE2), the α4 helix (residues 99 – 144) can be modelled (Kelley et al., 2015). Interestingly, the modelling process separates the α4 helix into two smaller anti-parallel helices connected by a short linker (herein referred to α4 and α5). When superimposed on the $Sr33^{5\text{-}120}$ NMR structure, the new α4 and α5 helices result in the creation of Sr33 CC domain model that forms a helical bundle of five antiparallel amphipathic helices (Figure 4.2 A). If the CC domain adopts this fold, it would be comparable to the death-domain (DD) fold seen in the effector domains of proteins involved in mammalian immunity and apoptotic pathways. Death-domain proteins include the previously mentioned CARDs and PYDs of mammalian NLRs. These proteins have highly divergent sequences (with as little as 10% sequence identity), but all have the same globular structure consisting of six anti-parallel amphipathic helices that form a helical bundle with Greek key topology (Figure 4.2 C & D). Each member of the DD family is capable of forming a series of homotypic interactions with other DD containing proteins, and regularly assemble into larger oligomeric structures. This structure is necessary for the activation of caspases and kinases, with the

best examples supplied by the structures of the APAF-1 and CED3 apoptosomes, which show extensive CARD interactions (Qi et al., 2010, Zhou et al., 2015).

The Sr33$^{6-144}$ CC domain model presented in Figure 4.2 shares a similar globular fold and helical arrangement to the DD domains of mammalian NLRs. Each helix of the Sr33$^{6-144}$ CC domain five-helix bundle is anti-parallel to the subsequent helix, and shares the same amphipathicity as described for the CARDs and PYDs. Many of the hydrophobic residues are internalised in order to stabilise the helical bundle structure, and charged residues are solvent exposed. It is regularly noted that DD proteins have patches of either acidic or basic residues that are generally located in areas of the protein that govern homotypic interactions. Intriguingly, the residues at the C-terminus of the Sr33$^{6-144}$ CC domain that are important for self-association and cell death signalling are largely basic, making up over 30% of the residues of the final helix. Despite these similarities, a major difference between the modelled Sr33$^{6-144}$ CC domain and DDs is that the CC domain is much larger than the standard size of a DD domain by approximately 50-60 residues. This is reflected in the significantly larger helices of the CC domain.

As there is no structural data for an active Sr33 CC domain, it is impossible to know the final arrangement of the helices of the active CC domain and therefore any claim that the CC domain resembles a DD fold should be taken with caution. However, with what we do know of the functions of plant NLRs in immunity and cell death signalling, combined with the possible structural similarities between the active CC domain and the CARD and PYD effector domains of animal NLRs, it is tempting to suggest the CC domains correspond to plant DDs.

If CC domains do have a DD fold, it would have implications for CNL function, as DD domains are known to oligomerise to form higher-order signalling complexes, in which the DDs form protein scaffolds for the binding of downstream partners, as observed by the homotypic CARD interactions between APAF-1 and caspase-9 (Yuan et al., 2010). Furthermore, we would again expect a scenario where the oligomerisation of the NLR would be mediated through the NB-ARC, which would drive the oligomerisation of the CC domain, as is observed for DD-containing NLRs.

As with TIR domains, the transient nature of CC domain interactions means it is unlikely a CC domain-only oligomer can be structurally characterised. However, it is imperative a structure of the active CC domain is determined to give insight into the mechanisms behind CC domain self-association. The work presented here has provided useful leads in experimental work to see this hope realised.

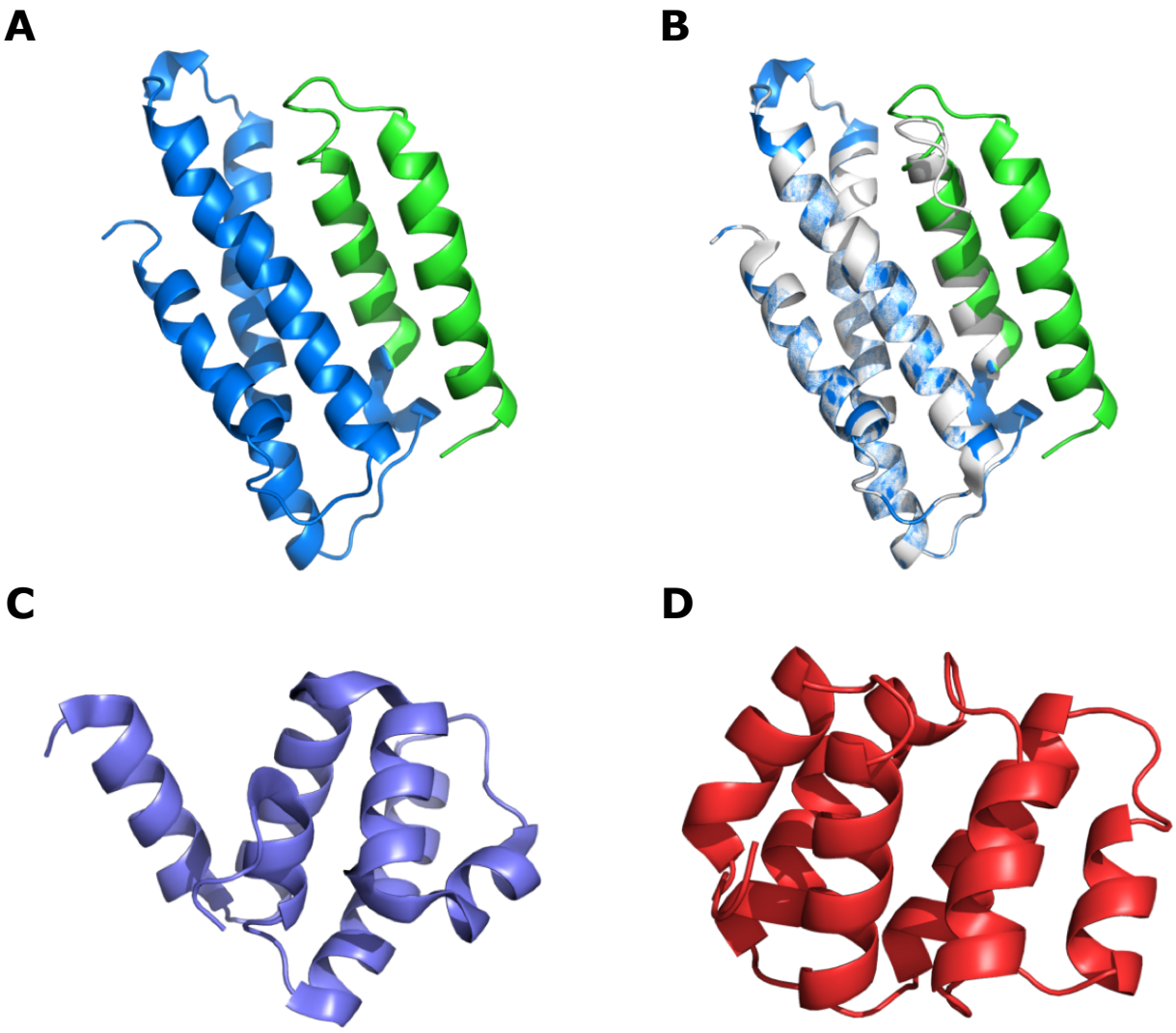


Figure 4.2 A model of the Sr33$^{6-144}$ CC domain in comparison with the structures of the NLRP4 PYD and APAF-1 CARD. (**A**) The model of the Sr33$^{6-144}$ CC domain. The residues encompassing the last predicted helix (residues 99 – 144; shown in green) were modelled with Phyre2 and superimposed on residues 99 – 120 of the Sr33$^{6-120}$ CC domain NMR structure (shown in blue; PDB ID: 2NCG). (**B**) Superimposition of the Sr33$^{6-144}$ model (blue and green) with the Sr33$^{6-120}$ CC domain NMR structure (white). (**C**) The crystal structure of the NLRP4 PYD (PDB ID: 4EWI) has a death-domain fold with six anti-parallel helices in a Greek key topology. (**D**) The crystal structure of the APAF-1 CARD (PDB ID: 1CY5) shares a death-domain fold similar to the NLRP4 PYD.

4.2 A general view of N-terminal domain functions: similarities between TIR and CC containing NLRs

TIR and CC domains are structurally unrelated. However, it is clear there are innate similarities in the function of these N-terminal domains. In both cases, the proteins are required to facilitate cell death signalling by the NLR, and self-association of the domain is a prerequisite for this function. Furthermore, the self-association of both CC and TIR domains is weak and transient. This observation likely reflects the auto-regulatory nature of plant NLRs. As the HR is often typified by cell-death, as such this signalling pathway must be tightly regulated to avoid auto-immune disease. N-terminal domains that formed strong interactions would potential present an auto-immune risk to the plant, as any accidental, global, non-pathogen elicited interaction could be lethal. In support of this, a recent study on incompatible TNLs, DM1 and DM2 (Dangerous Mix 1 & 2) demonstrated auto-immune responses could be caused by spontaneous TIR domain interactions, causing hybrid necrosis resulting in lethal or severely stunted phenotypes (Tran et al., 2017).

Furthermore, there are other similar regulatory mechanisms surrounding TIR and CC domain signalling. It has been demonstrated that TIR domains can be regulated through interactions with the ARC subdomains of the NB-ARC, with signalling differences between the flax TNLs L6 and L7 due to variances in interactions between their respective ARC and TIR domains (Bernoux et al., 2016). The same regulatory process for the CC domain has been observed in MLA10. Studies on the intramolecular interactions between the EDVID motif of the MLA10 CC domain and the ARC subdomains of the MLA10 NB-ARC domain have shown interactions are required for negative regulation of signalling (Bai et al., 2012). The EDVID motif appears necessary

for the association of the CC domain with the NB-ARC to maintain it in an auto-inhibited state. Mutations to the residues of the EDVID motif result in autoactivity of the MLA10 protein. Even though CC domains and TIR domains differ in structurally, auto-regulation through interactions with the NB-ARC appear to be conserved.

4.3 Formation of the plant NLR resistosome

Weak associations between N-terminal domains, multiple self-association interfaces, and auto-inhibition of N-terminal domains through intramolecular interactions with the NB-ARC domain, all point at the need for both TNLs and CNLs to oligomerise in order to signal. Effector detection and conformational changes in the NB-ARC would likely disrupt intramolecular interactions with the N-terminal region and release the domain for signalling. As we know, the self-association of N-terminal domains is transient, and thus it is unlikely their interactions alone could lead to receptor oligomerisation. However, as self-association of the N-terminal domains through multiple interfaces is required for signalling in TIR domains, it is very likely the association of the NB and/or LRR domain drive oligomerisation to initiate signalling.

It has been reported multiple times in the literature that plant NLRs are capable of associating through the NB-ARC and LRR domains; however, the stoichiometry at which they interact is not known, as the evidence comes from CoIP assays (Moffett et al., 2002, Mestre and Baulcombe, 2006, Schreiber et al., 2016). If plant NLRs function in a similar manner to animal NLRs as previously proposed in this discussion, it is possible they form a large oligomeric structure like an apoptosome or inflammasome - in the case a plant NLR, a resistosome. Arguments for plant NLR oligomerisation through the NB-ARC domain have already been made previously in

this discussion; however, initiation of oligomerisation could occur in three possible ways, as shown in Figure 4.3.

Firstly, each NLR that is incorporated into the resistosome complex may have to interact with an effector, as in seen in the APAF-1 apoptosome, which maintains 1:1 stoichiometry with the activating ligand, cytochrome c (Yuan et al., 2010) (Figure 4.3 A). Another alternative is an oligomerisation process similar to that of NLRC4 where a sensor NLR involved in detection of the effector activates the NLR (now deemed helper NLR), stimulating oligomerisation (Figure 4.3 B). Finally, helper NLRs may be negatively regulated by sensor NLRs, but are released and free to signal after interaction between the effector and the helper NLR, as proposed with RPS4 and RRS1 NLRs (Figure 4.3 C) (Williams et al., 2014).

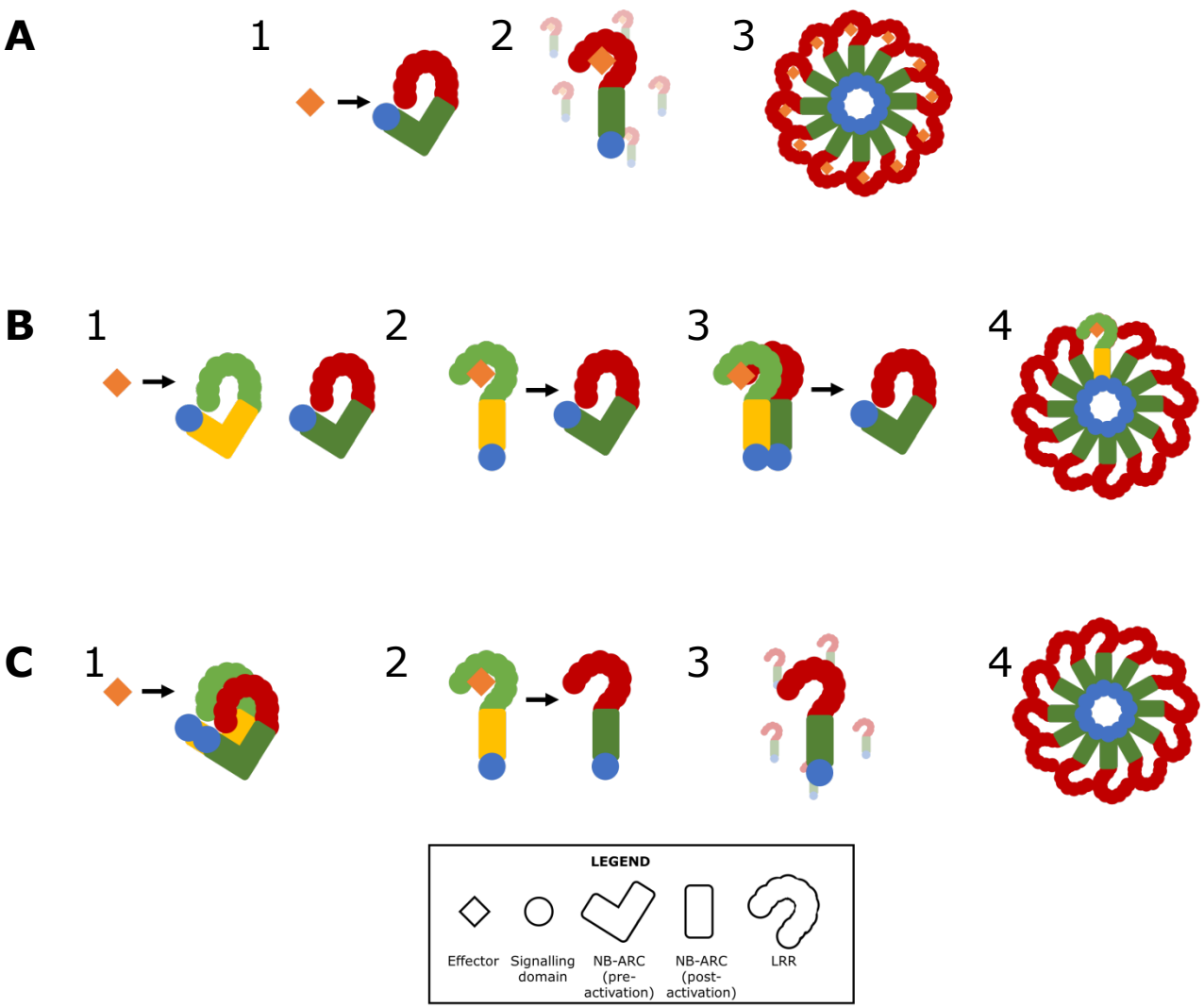


Figure 4.3 Modes of NLR activation post-pathogen perception (**A**) NLRs activate independently through direct detection with an effector and oligomerise, as seen in the APAF-1 apoptosome. (**A1**) An NLR in an inactive state interacts with an effector protein. (**A2**) Effector binding causes a conformational change in the NLR, resulting in an active receptor. (**A3**) NLRs activated by effector binding oligomerise to form a large signalling complex with a 1:1 effector to NLR stoichiometry. (**B**) Effector detection by a sensor NLR stimulates oligomerisation of a helper NLR as seen in the formation of the NLRC4 inflammasome. (**B1**) An effector interacts with a sensor NLR, causing activation. (**B2**) The activated sensor NLR can activate a helper NLR through NBD interactions. (**B3**) Activated sensor and helper NLR pair are able to interact with other inactive helper NLRs through NBD interactions between the activated helper NLR and the inactive helper NLRs. (**B4**) Helper NLRs oligomerise to form a complex with a single sensor NLR bound to a single effector protein. (**C**) Helper NLRs are negatively regulated until effector detection by their cognate sensor NLR partner. (**C1**) Effector is detected by the sensor NLR of the pair. (**C2**) The sensor NLR releases the helper NLR, which can undergo activation. (**C3**) Multiple activated helper NLRs are released from negative regulation. (**C4**) Activated helper NLRs are able to oligomerise.

4.4 Outliers: NLRs that do not fit current models

Many plant NLRs appear to work in similar ways. Autoactive phenotypes can be observed for many TIR domains from a variety of NLRs, and the same can be said for CC domain containing NLRs, from which we also see autoactive phenotypes. However, there are still a few studies that have identified NLRs that do not present similar functions or phenotypes to those discussed in this thesis. The two of the best examples of this are the grape TNL, RPV1, and the potato CNL, Rx.

4.4.1 The RPV1 TIR domain: autoactivity without self-association

RPV1 is a NLR from grape that contains a TIR domain at the N-terminus. This TIR has been demonstrated to signal cell death when transiently expressed *in planta*; however, there is no evidence of self-association of the protein (Williams et al., 2016). The crystal structure of the RPV1 TIR domain has the same flavodoxin-like fold of other TIR domains, and can be observed forming an AE interface dimer (Williams et al., 2016). However, MALS and yeast-2 hybrid analyses do not show any self-association of the protein. This result is different to all other TIR domain studies, including of the results reported in this thesis, as in all other circumstances TIR domains that do not self-associate, due to either a reduced ability, or mutagenesis to prevent self-association, are unable to signal cell death.

It is not certain why the RPV1 TIR domain is able to signal cell death in the absence of measurable self-association. However, as mutations to the residues of the RPV1 TIR domain AE interface also disrupt signalling, suggests this interface is required for function. Furthermore, mutations to residues in the DE interface (predicted through comparison with the L6 TIR domain) also prevent cell death signalling (Williams et al., 2016). This demonstrates both of

the AE and DE interfaces are required for RPV1 TIR domain cell death signalling, as seen in other TIR domains.

It is likely the RPV1 TIR domain does self-associate to signal, as the mutagenesis of conserved interface residues that disrupt RPV1 cell death have prevented other TIR domains self-associating and signalling. If this is true, this implies self-association of the RPV1 TIR domain is potentially driven by interactions with other proteins with *N. tabacum* transient expression assays. Alternatively, it is possible that external factors that could promote self-association, such as compartmentalisation in the cell, which would increase the local concentrations of protein high enough to promote cell death signalling.

Unfortunately, it cannot be definitively said if self-association is important for RPV1 TIR domain signalling as there are no observations of such an interaction, regardless of the likelihood. It is important this observation is not ignored, for it may potentially be a part of an alternative mechanism of TIR domain signalling that does not require self-association, which is yet to be properly characterised.

4.4.1 Rx signals cell death through the NB-ARC, not the CC domain

As previously mentioned, the CC^{1-122} domain of the potato NLR Rx shares significant structural similarity to the CC^{5-120} domains of Sr33 and MLA10. As it is now known the CC^{1-120} domains of CNLs are not the minimal functional units of CC domains that cause cell death signalling and self-association, it is not surprising that the Rx^{1-122} CC domain is not active *in planta*. However, there is no evidence for cell death signalling through a Rx CC domain of any length, rather it has been shown that the expression of the NB subdomain of the Rx NB-ARC is sufficient to trigger cell death signalling in tobacco (Rairdan et al., 2008). Transient expression

assays in tobacco have shown that Rx constructs containing the CC-NB, NB-only, and NB-ARC domains are autoactive, whereas CC-NB-ARC and NB-ARC-LRR constructs are not. By narrowing down the autoactive phenotype of the CC-NB, NB, and NB-ARC constructs, it was determined that the NB domain to be a common denominator in cell death signalling, with the NB domain alone showing the strongest autoactive phenotype (Rairdan et al., 2008).

This phenomenon brings into question the function of the Rx CC domain, and the mode action by which pathogen resistance is conferred by the Rx protein. As the structure of the Rx^{1-122} CC domain shares a similar fold to $MLA10^{5-120}$ and $Sr33^{6-120}$ CC domains, it would be generally assumed these proteins would all share a similar function. However, sequence analysis of the Rx CC domain demonstrates the Rx four-helix bundle is completely contained within the 1 – 122 residue boundary, whereas the fourth helix of the Sr33 and MLA10 CC domains continues on to residues 140 – 144. This indicates there could be significant differences in Rx CC domain function, even though the first 120 domains share the same fold as MLA10 and Sr33.

Unlike MLA10 and Sr33, the Rx CC domain has been shown to interact with a co-factor, RanGAP2. The crystal structure of the Rx CC domain in a complex with the WPP domain of Rx shows the region of the Rx CC domain that interacts with the WPP domain is nearby the EDVID motif (Hao et al., 2013). As previously mentioned, the EDVID motif has been shown to be important for intramolecular interactions between the CC domain and the ARC subdomains of the NB-ARC for both MLA10 and Rx (Rairdan et al., 2008, Bai et al., 2012). It is thought these intramolecular interactions are important for the negative regulation of signalling when the NLR is not challenged by an effector protein, as mutations to the EDVID motif result in autoactive phenotypes for

both MLA10 and Rx (Rairdan et al., 2008, Bai et al., 2012). As the NB subdomain is only capable of triggering cell death, it is thought the function of the Rx CC domain lies in auto-regulation of NB subdomain signalling through interactions with the ARC subdomains, which can be modulated by RanGAP2 binding. This would mean the role of the CC domain would be more towards the indirect recognition of an effector rather than signalling (Rairdan et al., 2008).

Signalling through the NB domain raises many questions about the function of the Rx protein. Interestingly, this activity is not P-loop dependent, meaning nucleotide binding is inconsequential to the activity of the Rx NB subdomain (Rairdan et al., 2008). It has been proposed nucleotide binding is an important part of the conformational change that determines the activity of an NLR (Takken and Goverse, 2012). This suggests the Rx NB domain adopts an active fold when expressed *in planta*. How the NB domain signals cell death is unclear and can only be speculated. It is possible the activity of the Rx NB domain, and observed lack of activity of the Rx CC domain, could hint at the role of the Rx protein as a "sensor NLR". If this is true, it could mean Rx is auxiliary to a helper NLR partner that is activated via interaction with the NB domain of Rx. Sure enough, a recent study has uncovered a family of NLRs in solanaceous plants deemed NRCs (NLR required for cell death), whose presence is required for cell death signalling by multiple NLRs, including Rx (Wu et al., 2017). This type of interaction is also seen between the NB domains of NAIP2/NAIP5 and NLRC4 NLRs from mammals, where NAIP2 activates NLRC4 through homotypic NB domain interactions, which then allows NLRC4 to form an inflammasome complex to relay signalling (Hu et al., 2015, Zhang et al., 2015, Tenthorey et al., 2017). This model would provide support for the hypothesis that the Rx CC domain is involved in negative regulation of the NB-ARC

domain, and once intramolecular interactions between the CC domain and NB-ARC are disturbed, the NB-ARC can form interactions with a signalling partner. However, there is no known NLR-like signalling partner for Rx that has shown the capability to signal cell death, and as transient expression assays are performed in heterologous hosts, it would suggest Rx is interacting with a cross-species signalling partner.

Collectively, the importance of the CC domain of Rx in cell-death activation is unclear; however, there are enough differences between Rx and the other studied CNLs that it is possible this protein signals in a different manner. Until further research of this protein is performed, it is difficult to determine how this protein functions.

4.5 Future Work

Despite this study and the research of many others, our structural knowledge of plant NLRs remains somewhat limited. Currently, only structures of the N-terminal domains are available, and that of an integrated decoy domain. Moreover, the fold an active CC domain structure adopts is still unclear, and additional work to achieve a structure is required if we are to understand CC domain-mediated signalling. Efforts to crystallise the active CC domain have proven difficult; however, there appears to be some success using *in situ proteolysis*, and this should be further explored. If this fails to yield results, a wider variety of CC domains from other CNLs should be tested for their ability to form crystals.

TIR domain oligomerisation is an intriguing concept and should be continued to be investigated. It is clear these domains require the use of multiple interfaces to self-associate and signal, and this phenomenon is ubiquitous for all TIR domains tested thus far. However, capturing a TIR-only complex may prove too difficult, as the weak interactions of TIR domains in solution would not allow for the assembly of complexes in solution without the drive of an exterior force. To truly begin to understand TIR function, and the function of NLRs in general, we need to determine the structure of the full-length receptors, in active and inactive conformations.

Furthermore, the oligomerisation of NLRs to form a resistosome is something that can only ever be speculated upon until we have structural evidence. The studies of mammalian NLRs and NLR-like proteins have demonstrated the usefulness of cryo-EM in the determination of large complexes, with structures of the NLRC4 inflammasome and the CED4 and APAF-1 apoptosomes being elucidated with this technique. However, the first barrier that must

be overcome, if we wish to study full-length receptors, is their expression and purification. This is the limiting factor, as while it is possible to produce these proteins in recombinant expression systems, yields are exceptionally low and they often contain many contaminants. Nevertheless, this problem makes cryo-EM even more desirable, as the amounts of proteins required for this technique are orders of magnitude less than required for crystallography, making it a more feasible choice for structural studies of full-length plant NLRs.

5. CONCLUSIONS

This thesis aimed to better understand the molecular mechanisms by which the N-terminal CC and TIR domains of plant NLRs function. At the onset of this work, it was understood there were two separate receptor-defined interfaces identified for TIR domain signalling, and two CC domain structures with significantly different folds.

From the crystal structure and biophysical analyses of the RPP1 TIR domain and other TIR domains, it is now clear that two interfaces are simultaneously required for TIR domain signalling. This was demonstrated not only by the presence of both interfaces in the RPP1 TIR domain crystal structure, but by MALS analyses. These findings were further corroborated for the TIR domains from other NLRs, including L6, RPS4 and SNC1. Taken together, these data reconcile the two previously published L6 and RPS4 TIR domain structures, and give a clearer insight into TIR domain signalling. The most intriguing outcome of this work is it opens the door to the possibility of TIR domain oligomerisation in plants. However, there is no structural evidence of a TIR domain oligomer in plants thus far, and this requires further research.

The previous structural studies of the $MLA10^{5-120}$ and Rx^{1-122} CC domains gave rise to two contrasting structures, despite both proteins belonging to the same class of CNL proteins. One was an obligate helix-loop-helix dimer, the other a monomeric four-helix bundle. The work on the Sr33 CC domain presented in this thesis has reconciled these differences by demonstrating the $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} CC domains all form the four-helix bundle fold in solution. Furthermore, it was previously proposed the $MLA10^{5-120}$ CC domain presented the minimal functional unit for

homodimerisation and cell death signalling *in planta*. This was proven to be false, as all CC-120 proteins are monomeric in solution, and inactive when transiently expressed *in planta*. It was determined in this work the CC-120 domain of Sr33 and MLA10 truncates the α4 helix of the protein, which extends to residue 142. Expression of Sr33 and MLA10 CC domain proteins with an intact α4 helix *in planta* induces cell death, and this was shown to correlate to self-association in solution. Sequential truncation and mutation of conserved residues in the α4 helix of the CC domain both perturbed self-association and signalling, further demonstrating the importance of this region in the function of the CC domain. From this work, it was determined the minimal functional unit for self-association and cell death signalling is residues 1 – 142 for the MLA10 and Sr33 CC domains. Despite the best possible efforts, no structure of an active CC domain was able to be determined within the time-frame of this candidature. However, significant progress has been made toward the crystallisation of the Sr33 CC domain, and requires only slight optimisation before a structure may be acquired.

Significant progress has been made toward the understanding of N-terminal domain signalling. However, questions of TIR domain, and even possibly CC domain, oligomerisation cannot be answered by studies of the N-terminal domains alone. It is clear to better understand signalling through oligomerisation, studies of full-length receptors need to be undertaken. Structural studies of full-length NLRs are currently limited by yield and purity; however, the advent of cryo-EM makes protein yield less of an issue, and optimisation of protein production to utilise this technique could exceptional results.

There are still inconsistencies in the literature surrounding NLR activation, effector detection, and signalling. To really understand

the molecular mechanisms of NLR function, we need further structural studies of full-length receptors. Only then can we stop speculating on function, and begin to understand the complex interactions between these plant immune proteins and pathogen effectors, which underpin millennia-old plant-pathogen interactions. From that point, we can start to design NLRs that provide durable resistance for crop, and prevent microbial diseases from decimating our much-needed and dwindling food resources.

6.MATERIALS AND METHODS

The research in this thesis has been published in the following articles:

SCHREIBER, K. J., BENTHAM, A., WILLIAMS, S. J., KOBE, B. & STASKAWICZ, B. J. 2016. Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. *PLoS Pathogens*, 12, e1005769;

ZHANG, X., BERNOUX, M., BENTHAM, A. R., NEWMAN, T. E., VE, T., CASEY, L. W., RAAYMAKERS, T. M., HU, J., CROLL, T. I., SCHREIBER, K. J., STASKAWICZ, B. J., ANDERSON, P. A., SOHN, K. H., WILLIAMS, S. J., DODDS, P. N. & KOBE, B. 2017. Multiple functional self-association interfaces in plant TIR domains. *Proceedings of the National Academy of Sciences,* 114**,** E2046-E2052.

CASEY, L. W., LAVRENCIC, P., BENTHAM, A. R., CESARI, S., ERICSSON, D. J., CROLL, T., TURK, D., ANDERSON, P. A., MARK, A. E., DODDS, P. N., MOBLI, M., KOBE, B. & WILLIAMS, S. J. 2016. The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins. *Proceedings of the National Academy of Sciences,* 113**,** 12856-12861.

As such, all the details of the materials and methods for the work performed in this thesis can be found in the Materials and Methods sections of the papers above. These papers can be found attached to this thesis as Appendix 1, Appendix 2, and Appendix 3, respectively. All other techniques not mentioned in these papers have been highlighted and discussed in the results of each section.

7. APPENDIX

7.1 Appendix 1

PLOS | PATHOGENS

RESEARCH ARTICLE

Multiple Domain Associations within the Arabidopsis Immune Receptor RPP1 Regulate the Activation of Programmed Cell Death

Karl J. Schreiber[1], Adam Bentham[2,3], Simon J. Williams[2], Bostjan Kobe[2,4], Brian J. Staskawicz[1]*

1 Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, California, United States of America, **2** School of Chemistry and Molecular Biosciences and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Australia, **3** School of Biological Sciences, Faculty of Science and Engineering, Flinders University, Adelaide, Australia, **4** Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia

* stask@berkeley.edu

OPEN ACCESS

Citation: Schreiber KJ, Bentham A, Williams SJ, Kobe B, Staskawicz BJ (2016) Multiple Domain Associations within the Arabidopsis Immune Receptor RPP1 Regulate the Activation of Programmed Cell Death. PLoS Pathog 12(7): e1005769. doi:10.1371/journal.ppat.1005769

Editor: Jeffery L. Dangl, The University of North Carolina at Chapel Hill, UNITED STATES

Received: March 2, 2016

Accepted: June 24, 2016

Published: July 18, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by grant NSF-IOS-1146793 to BJS from the National Science Foundation, www.nsf.gov/. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Upon recognition of pathogen virulence effectors, plant nucleotide-binding leucine-rich repeat (NLR) proteins induce defense responses including localized host cell death. In an effort to understand the molecular mechanisms leading to this response, we examined the *Arabidopsis thaliana* NLR protein RECOGNITION OF *PERONOSPORA PARASITICA*1 (RPP1), which recognizes the *Hyaloperonospora arabidopsidis* effector *ARABIDOPSIS THALIANA* RECOGNIZED1 (ATR1). Expression of the N-terminus of RPP1, including the Toll/interleukin-1 receptor (TIR) domain ("N-TIR"), elicited an effector-independent cell death response, and we used allelic variation in TIR domain sequences to define the key residues that contribute to this phenotype. Further biochemical characterization indicated that cell death induction was correlated with N-TIR domain self-association. In addition, we demonstrated that the nucleotide-binding (NB)-ARC1 region of RPP1 self-associates and plays a critical role in cell death activation, likely by facilitating TIR:TIR interactions. Structural homology modeling of the NB subdomain allowed us to identify a putative oligomerization interface that was shown to influence NB-ARC1 self-association. Significantly, full-length RPP1 exhibited effector-dependent oligomerization and, although mutations at the NB-ARC1 oligomerization interface eliminated cell death induction, RPP1 self-association was unaffected, suggesting that additional regions contribute to oligomerization. Indeed, the leucine-rich repeat domain of RPP1 also self-associates, indicating that multiple interaction interfaces exist within activated RPP1 oligomers. Finally, we observed numerous intramolecular interactions that likely function to negatively regulate RPP1, and present a model describing the transition to an active NLR protein.

Author Summary

Many plant pathogens inject proteins known as effectors into the cells of their hosts in order to suppress host immune responses and promote pathogen growth. Over time, plants have evolved receptors, described as nucleotide-binding leucine-rich repeat (NLR) proteins, which recognize the activity of pathogen effectors and stimulate defense responses. Plant NLRs contain several domains that exhibit striking functional conservation with NLRs from other eukaryotes. Despite their important contribution to plant immunity, the molecular mechanisms that underlie effector recognition and subsequent immune activation by NLRs remain to be fully elucidated. Here, we focus on RPP1, an NLR from Arabidopsis that recognizes the oomycete effector ATR1. Using transient co-expression of proteins in plants, we demonstrate that recognition of ATR1 stimulates RPP1 oligomerization. This interaction involves multiple domains of RPP1 and is critical for immune activation. In the absence of ATR1, we documented interactions between domains within an individual RPP1 protein, likely occurring to prevent inappropriate immune activation. Finally, we examined differences between RPP1 alleles as well as structural data from animal NLRs to help identify specific amino acids that mediate interactions within and between RPP1 molecules. Collectively, these data allow us to propose a model for the activation of RPP1 following ATR1 recognition.

Introduction

Through interactions with a plethora of phytopathogenic organisms, plants have evolved a sophisticated molecular surveillance system that involves nucleotide-binding leucine-rich repeat (NLR) immune receptors that recognize pathogen-derived virulence effector proteins. Effector recognition elicits a strong immune response that is associated with programmed cell death, known as the hypersensitive response (HR), which ultimately restricts pathogen proliferation [1,2]. Plant NLR proteins are generally modular in structure, with a C-terminal leucine-rich repeat (LRR) domain often being responsible for effector recognition, either directly or indirectly. A central nucleotide-binding domain with homology to Apaf-1, resistance proteins, and CED-4 (NB-ARC) is associated with nucleotide binding and hydrolysis, and acts as a molecular switch for NLR protein activation [3–6]. Upon activation, immune responses are thought to signal through the N-terminal region of the NLR protein, which may comprise an N-terminal coiled-coil (CC) or Toll/interleukin-1 receptor (TIR)-like domain. In the absence of pathogen-associated elicitors, NLR proteins are thought to be maintained in an inactive state by a number of intramolecular interactions [7,6]. Effector recognition by the LRR likely causes a conformational change that increases the accessibility of the NB-ARC domain to allow nucleotide exchange and subsequent NLR protein activation, although the specific structural changes that lead to this active state are unclear.

There is increasing evidence from the study of animal NLR immune receptors that intermolecular associations play a key role in the stimulation of immune responses. Similar to plant NLR proteins, these receptors include a nucleotide-binding domain flanked by a C-terminal ligand-sensing domain and an N-terminal signaling domain that recruits additional immunomodulatory proteins [8]. Following the detection of pathogen- or damage-derived signals, the assembly of higher-order complexes has been documented for proteins such as Apaf-1, NLRC4, CED-4, NALP1, and Dark [9–16]. Analyses of these inflammasome and apoptosome structures indicate that residues within and proximal to the nucleotide-binding domain are the primary regions of contact for oligomerization [13,14,17]. In plants, the only known example of ligand-dependent oligomerization involves the tobacco N protein, whose self-association is

induced by recognition of the tobacco mosaic virus replicase [18]. Currently, it is not known whether this is a unique evolutionary development or representative of a common yet poorly characterized pathway for plant NLR protein activation.

In light of this uncertainty, we examined the *Arabidopsis thaliana* protein RPP1 (RECOGNITION OF *PERONOSPORA PARASITICA*1) [19], a TIR domain-containing NLR protein that recognizes the *Hyaloperonospora arabidopsidis* (formerly *Peronospora parasitica*) effector ATR1 (*ARABIDOPSIS THALIANA* RECOGNIZED1) [20]. This recognition is mediated by the LRR domain of RPP1, with both genetic and *in planta* biochemical evidence suggesting a direct mode of recognition [21,22]. Multiple alleles of RPP1 and ATR1 have been cloned and characterized, revealing significant allelic variation in the NLR protein/effector combinations that lead to cell death. For example, the Niederzenz (NdA) allele of RPP1 recognizes the Emoy2 allele of ATR1, while the Wassilewskija (WsB) allele of RPP1 recognizes the Emoy2, Maks9, and Emco5 ATR1 alleles. Neither RPP1 allele recognizes ATR1_Cala2. These polymorphic phenotypes have enabled an examination of the sequences that contribute to ATR1 recognition [22] and could also be used to dissect the activation of immune responses.

In this study, we sought to functionally characterize the domains of RPP1 with regards to cell death induction following effector recognition. Using transient expression of epitope-tagged proteins, we demonstrate that RPP1 oligomerizes in the presence of a recognized allele of ATR1. Co-immunoprecipitation and mutagenesis experiments were used to investigate the specific domains involved in pre-activation intramolecular interactions and post-activation intermolecular interactions. Phenotypic variation amongst RPP1 alleles and structural data from various animal immune receptors were used to identify specific amino acids that facilitate these protein-protein interactions. Based on these observations, we suggest a model of the molecular events underlying only the second documented example of effector-dependent NLR protein oligomerization in plants.

Results

Allelic variation in TIR domain autoactivity

Our efforts to define the molecular events associated with cell death induction by RPP1 initially focused on the N-terminal region of this protein, which contains a TIR domain (Fig 1A). In other NLRs, the TIR domain executes the cell death response upon effector recognition [23,24]. When the first 254 amino acids of the NdA allele of RPP1 were transiently expressed in *Nicotiana tabacum*, an effector-independent HR developed within 24 hours (Fig 1B). This 254 aa region comprises the minimal autoactive sequence (Fig 1B and 1C) and includes the TIR domain preceded by an N-terminal region with no obvious homology to other functional domains, hereafter collectively referred to as "N-TIR". Interestingly, although the WsB allele of RPP1 can recognize ATR1 and elicit cell death as a full-length protein, the equivalent N-TIR region from this allele did not exhibit autoactivity (Fig 1D). Quantification of electrolyte leakage induced by expression of these constructs confirmed the lack of autoactivity for the WsB allele (S1 Fig). To dissect this phenotypic polymorphism, we constructed chimeras of the two alleles and tested their ability to induce an HR. These analyses indicated that regions both within and N-terminal to the predicted TIR domain of RPP1_NdA conferred autoactivity (Fig 1D). In a WsB background, the inclusion of amino acids 42–92 from RPP1_NdA resulted in a relatively weak HR (Fig 1D). The autoactivity of the reciprocal chimera was not compromised (Fig 1D), indicating that the polymorphisms within this region were dispensable for cell death induction, although the N-terminal region itself was critical for autoactivity (Fig 1C).

To identify specific amino acids within the TIR domain that contribute to autoactivity, we performed site-directed mutagenesis based on amino acid polymorphisms between the NdA,

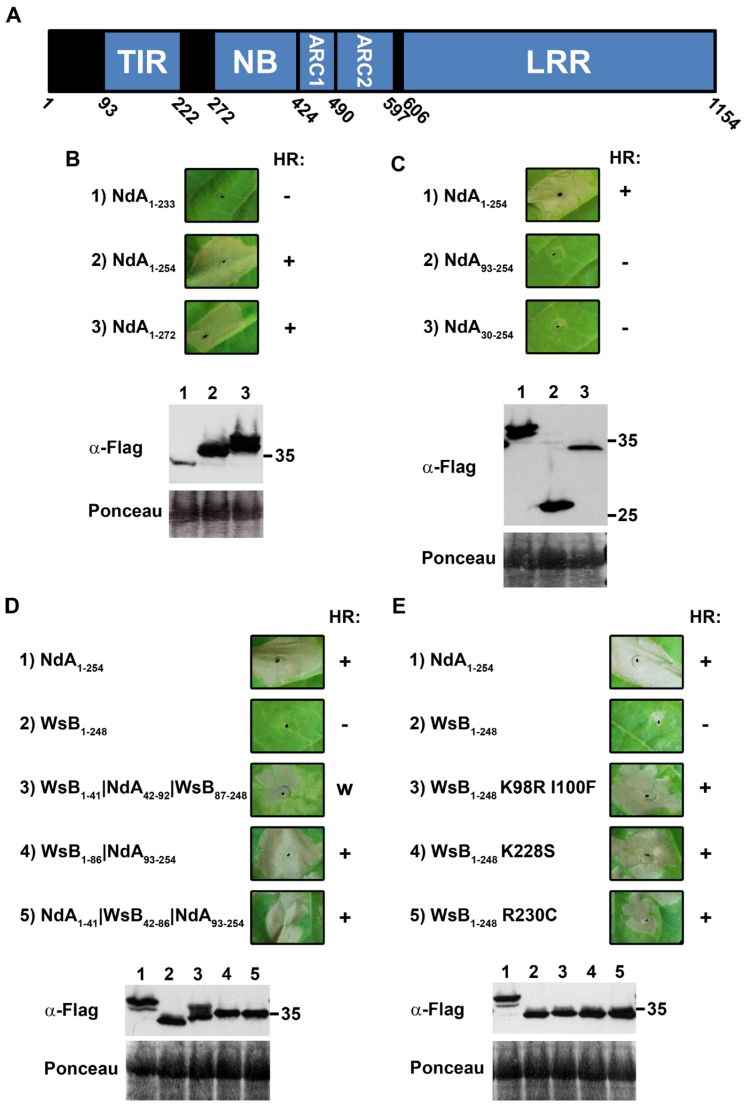
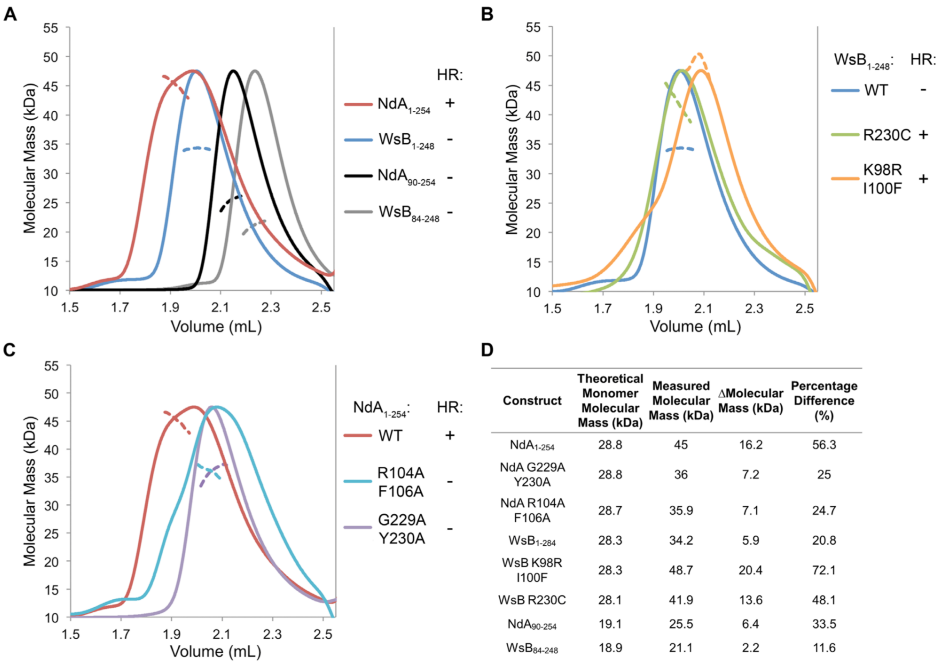

A
TIR
NB
ARC1
ARC2
LRR
1
93
222
272
424
490
597
606
1154
B
HR:
1) NdA1-233
-
2) NdA1-254
+
3) NdA1-272
+
1 2 3
α-Flag
35
Ponceau
C
HR:
1) NdA1-254
+
2) NdA93-254
-
3) NdA30-254
-
1 2 3
35
α-Flag
25
Ponceau
D
HR:
1) NdA1-254
+
2) WsB1-248
-
3) WsB1-41|NdA42-92|WsB67-248
w
4) WsB1-86|NdA93-254
+
5) NdA1-41|WsB42-86|NdA93-254
+
1 2 3 4 5
α-Flag
35
Ponceau
E
HR:
1) NdA1-254
+
2) WsB1-248
-
3) WsB1-248 K98R I100F
+
4) WsB1-248 K228S
+
5) WsB1-248 R230C
+
1 2 3 4 5
α-Flag
35
Ponceau

Fig 1. Identification of sequences required for RPP1 TIR domain autoactivity. (A) Schematic overview of the domain architecture of RPP1. Numbers indicate the amino acid position of predicted domain borders for the Niederzenz (NdA) allele of RPP1. (B,C) Determination of the minimal autoactive TIR domain from the NdA allele of RPP1. Both C-terminal (B) and N-terminal (C) truncations were examined for their ability to elicit an effector-independent hypersensitive response (HR). The specific amino acids comprising each construct are indicated in subscript. (D,E) HR phenotypes associated with chimeras or site-directed mutants of N-TIR domains from the NdA and Wassilewskija (WsB) alleles. Site-directed mutagenesis was guided by the amino acid alignment depicted in S2 Fig. Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient expression and images were captured at 48 hours post-infiltration. HR phenotypes are scored as negative (-), weak (w), or strong (+). An α-Flag antibody was used to evaluate protein expression, while staining of RuBisCO with Ponceau S provided a loading control. The experiment was performed three times with similar results.

doi:10.1371/journal.ppat.1005769.g001

WsB, and Estland (Est-1) alleles of RPP1 (S2 Fig). Est-1 was included because the sequence of its TIR domain closely resembles that of NdA [25], and it is also autoactive (S2 Fig). In a WsB background, the introduction of substitutions K98R I100F, K228S, or R230C (to mimic the NdA or Est-1 sequence) was sufficient to confer autoactivity (Fig 1E, S1 Fig). Constructs with K98R or I100F alone elicited a significantly weaker HR, indicating an additive effect of the two substitutions (S3 Fig). In an NdA background, substitution of the corresponding residues with alanines (R104A F106A) eliminated autoactivity, while replacement with the WsB residues (R104K F106I) had no effect on cell death (S4 Fig). Substitutions K234A and C236A also had no impact on the HR phenotype of the NdA N-TIR domain, although amino acid changes nearby (G229A Y230A) resulted in a loss of autoactivity (S4 Fig).

Given that the autoactivity of some TIR domains is influenced by their ability to self-associate [23,26], we sought to determine if RPP1 exhibits the same behavior. Previous studies of NLR TIR domains, including that from RPP1, noted that self-association could not by detected by co-immunoprecipitation, likely due to weak and/or transient interactions [21,23,24]. As such, we employed a chromatography-based approach to evaluate TIR domain self-association. Size-exclusion chromatography (SEC) is commonly used to characterize protein size and oligomeric state indirectly by the comparison of retention times of proteins of interest with those of protein standards. However, the migration of proteins through SEC columns is influenced by a number of parameters including particle shape, flexibility and composition. For these reasons, we coupled SEC with multi-angle laser light scattering (MALS). The signal from MALS is directly related to the average molecular mass of proteins eluted from a SEC column, making it an ideal technique to investigate and compare the solution properties of purified recombinant TIR domain proteins. When we evaluated the TIR domain alone, the measured average molecular masses of $NdA_{90\text{-}254}$ and $WsB_{84\text{-}248}$ were 33.5% and 11.6% higher, respectively, than would be expected for a monomeric protein (Fig 2A and 2D). This suggests that both proteins self-associate in solution, although to a much greater degree for the NdA allele (Fig 2A). Interestingly, the inclusion of the N-terminal region preceding the TIR domain ($NdA_{1\text{-}254}$ and $WsB_{1\text{-}248}$) further increased the measured molecular mass relative to the expected molecular mass of a monomer. Again, this difference was much more striking for NdA compared to WsB (Fig 2D), suggesting that the N-TIR domain of NdA has a greater tendency to self-associate than that of WsB (Fig 2A). A change in retention time on the SEC column was not observed despite an increase in the measured molecular mass via MALS, which is indicative of a rapid, transient interaction of the protein in solution. In addition, gain-of-autoactivity N-TIR_WsB mutants (R230C and K98R I100F) exhibited significant increases in measured average molecular mass, similar to or even greater than the increase observed for wild-type $NdA_{1\text{-}254}$ (Fig 2B and 2D). Conversely, large reductions in molecular mass were noted for loss-of-autoactivity N-TIR_NdA mutants (R104A F106A and G229A Y230A), such that they resembled the solution properties of wild-type $WsB_{1\text{-}248}$ (Fig 2C and 2D). Altogether, these data indicate that the *in planta* autoactivity of the RPP1 N-TIR domain is correlated with its propensity for self-association in solution.



D

Construct	Theoretical Monomer Molecular Mass (kDa)	Measured Molecular Mass (kDa)	ΔMolecular Mass (kDa)	Percentage Difference (%)
NdA_{1-254}	28.8	45	16.2	56.3
NdA G229A Y230A	28.8	36	7.2	25
NdA R104A F106A	28.7	35.9	7.1	24.7
WsB_{1-284}	28.3	34.2	5.9	20.8
WsB K98R I100F	28.3	48.7	20.4	72.1
WsB R230C	28.1	41.9	13.6	48.1
NdA_{90-254}	19.1	25.5	6.4	33.5
WsB_{84-248}	18.9	21.1	2.2	11.6

Fig 2. The autoactivity of the RPP1 N-TIR domain is correlated with self-association in solution. Purified (N-)TIR domain proteins from the Niederzenz (NdA) and Wassilewskija (WsB) alleles of RPP1 were analyzed by size-exclusion chromatography (SEC) coupled with multi-angle laser light scattering (MALS). For each sample, 175 μg of purified protein was separated on a Superdex Increase 200 5/150 GL SEC column and the molecular mass calculated across the elution peak. The colored solid line represents the normalized refractive index trace (arbitrary units) of the protein eluting from the SEC column. At the elution peak, the averaged molecular mass (kDa) of the proteins was calculated from the protein concentration (derived from the refractive index changes) and light scattering data. The averaged molecular masses across the elution peak are represented by dashed lines of the corresponding color. *In planta* hypersensitive response (HR) phenotypes are indicated for each construct by a "+" (autoactive) or "-"(non-autoactive). These phenotypes are documented in Fig 1, S2 and S4 Figs. (A) Comparison of the solution properties of RPP1 TIR domains with and without native RPP1 N termini. Numbers in the legend refer to the amino acids that comprise each protein sample. (B) Solution properties of the wild-type WsB N-TIR domain and the gain-of-autoactivity mutants, WsB R230C and WsB K98R I100F. (C) Solution properties of the wild-type NdA N-TIR domain and the loss-of-autoactivity mutants, NdA G229A Y230A and NdA R104A F106A. (D) Comparison of theoretical monomer molecular masses and the measured molecular masses for the proteins analyzed in (A-C).

doi:10.1371/journal.ppat.1005769.g002

The NB-ARC domain influences N-TIR domain autoactivity

Despite the inability of the N-TIR domain from RPP1_WsB to elicit an HR, the full-length protein is capable of HR induction following recognition of the effector ATR1 [21,22]. This suggested that additional domains participate in cell death signaling. Indeed, when constructs comprising the N-TIR-NB-ARC1 region (aa 1–490 for NdA, aa 1–484 for WsB) were transiently expressed in *N. tabacum*, sequences from both the WsB and NdA alleles elicited an effector-independent HR (Fig 3A). Intriguingly, N-TIR-NB (NdA: aa 1–424; WsB: aa 1–418) and N-TIR-NB-ARC1-ARC2 (NdA: aa 1–597; WsB: aa 1–598) constructs were not autoactive for either allele. Measurements of electrolyte leakage corroborated these observations, although

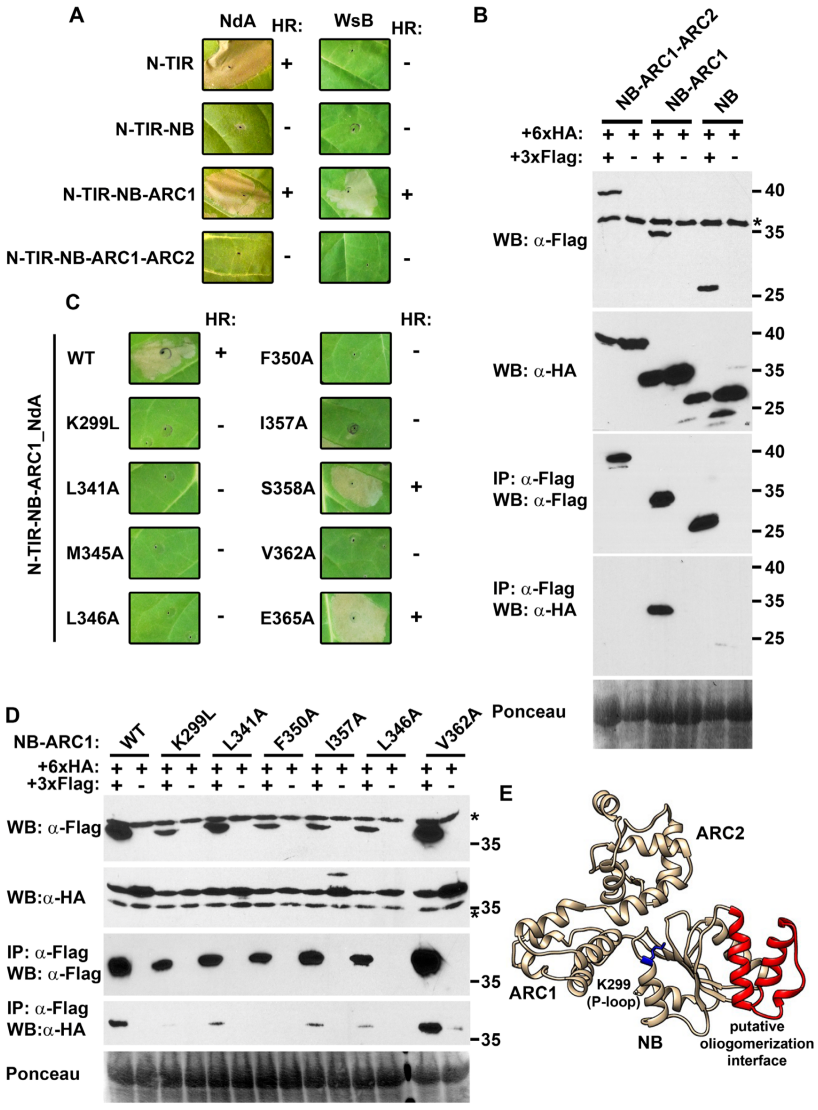

A
NdA
HR:
WsB
HR:
N-TIR
+
-
N-TIR-NB
-
-
N-TIR-NB-ARC1
+
+
N-TIR-NB-ARC1-ARC2
-
-
B
NB-ARC1-ARC2
NB-ARC1
NB
+6xHA:
+3xFlag:
WB: α-Flag
WB: α-HA
IP: α-Flag
WB: α-Flag
IP: α-Flag
WB: α-HA
Ponceau
40
35
25
C
HR:
HR:
N-TIR-NB-ARC1_NdA
WT
+
F350A
-
K299L
-
I357A
-
L341A
-
S358A
+
M345A
-
V362A
-
L346A
-
E365A
+
D
NB-ARC1:
WT
K299L
L341A
F350A
I357A
L346A
V362A
+6xHA:
+3xFlag:
WB: α-Flag
WB:α-HA
IP: α-Flag
WB: α-Flag
IP: α-Flag
WB:α-HA
Ponceau
35
E
ARC2
ARC1
K299
(P-loop)
NB
putative
oligomerization
interface

Fig 3. The NB-ARC domain influences N-TIR domain autoactivity. (A) Hypersensitive response (HR) phenotypes are altered by the successive addition of NB, ARC1, and ARC2 subdomains to the N-TIR domains of the RPP1 alleles Niederzenz (NdA) or Wassilewskija (WsB). Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient expression and images were captured at 48 hours post-infiltration (hpi). The presence or absence of HR is indicated by a "+" or "-", respectively. (B) Detection of self-association between NdA NB-ARC subdomain truncations by co-immunoprecipitation. Differentially epitope-tagged versions of NB, NB-ARC1, or NB-ARC1-ARC2 proteins were transiently expressed in *N. benthamiana* and samples were collected at 48 hpi for co-immunoprecipitation using α-Flag agarose beads. Asterisks indicate non-specific bands. Staining of RuBisCO with Ponceau S provides a loading control. (C) Site-directed mutagenesis of the NB subdomain compromises HR induction in an NdA N-TIR-NB-ARC1 background. Constructs were tested as in (A); protein expression for constructs from (A) and (C) is documented in S9 Fig. (D) Detection of self-association of NB-ARC1 mutants by co-immunoprecipitation, tested as in (B). Experiments were performed at least three times with similar results. (E) Predicted structure of the RPP1 NB-ARC domain derived from homology modeling using the Drosophila Dark protein (PDB:4v4l) as a template. The K299 residue within the P-loop is denoted in blue, while the putative oligomerization interface is highlighted in red. The specific residues at the putative oligomerization interface that were analyzed by mutagenesis are depicted in S10 Fig.

doi:10.1371/journal.ppat.1005769.g003

weak autoactivity was detected for the RPP1_NdA N-TIR-NB construct (S5 Fig). Further dissection of the ARC1 subdomain through C-terminal truncations revealed that a complete ARC1 was required for HR induction (S6 Fig).

Given the importance of N-TIR domain self-association for autoactivity, the contribution of the NB-ARC1 region may also depend upon homotypic interactions. Using co-immunoprecipitation, we demonstrated that the RPP1_NdA NB-ARC1 region (aa 255–490) could self-associate, while very weak interactions were observed for the NB region (aa 255–424) and no self-association was detected with NB-ARC1-ARC2 constructs (aa 255–597) (Fig 3B). Furthermore, self-association was observed with the autoactive N-TIR-NB-ARC1 proteins from alleles NdA and WsB (S7 Fig), providing further evidence that oligomerization significantly influences autoactivity. For the NB-ARC1 subdomain, co-immunoprecipitation yields from self-association experiments were consistently higher for the NdA allele relative to the WsB allele (S8 Fig), so subsequent analyses focused on RPP1_NdA.

The NB-ARC1 interactions that we observed are reminiscent of the *Caenorhabditis elegans* CED-4 protein, whose assembly into an apoptosome is mediated by intermolecular contacts on the surface of the α/β fold [13]. In RPP1, the homologous region is located in two α-helices of the NB subdomain (Fig 3E, highlighted in red, and S10 Fig). Alanine substitutions of surface-exposed residues on these helices generally eliminated the autoactivity of N-TIR-NB-ARC1 constructs, as did a K299L substitution at the predicted nucleotide-binding site, or P-loop (Fig 3C). One exception, however, was noted for the amino acid change E365A, which actually appeared to strengthen the HR (S11 Fig, S12 Fig). Co-immunoprecipitation experiments indicated that NB-ARC1 self-association was reduced in the loss-of-autoactivity mutants (Fig 3D), although *in planta* expression was also lower. This pattern was consistent aside from the N-TIR-NB-ARC1-inactivating substitution V362A, which retained self-association, accompanied by relatively high levels of protein expression. Overall, these data suggested that the NB subdomain, in cooperation with ARC1, is a key mediator of cell death induction.

Effector-dependent oligomerization of full-length RPP1

Following the characterization of effector-independent phenotypes in RPP1 domain truncations, we next sought to examine the function of these domains in the context of full-length RPP1. Transient co-expression of RPP1_NdA and ATR1 in *N. benthamiana* revealed RPP1 self-association in the presence of a recognized allele of ATR1 (ATR1_Emoy2) but not an unrecognized allele (ATR1_Cala2) (Fig 4A). The immunoprecipitated oligomer also included ATR1_Emoy2 (S13 Fig). Based on the oligomerization phenotype, we introduced the same substitutions into the putative oligomerization interface that were examined in the domain truncations, as well as substitutions within the TIR domain and P-loop. In most cases, these

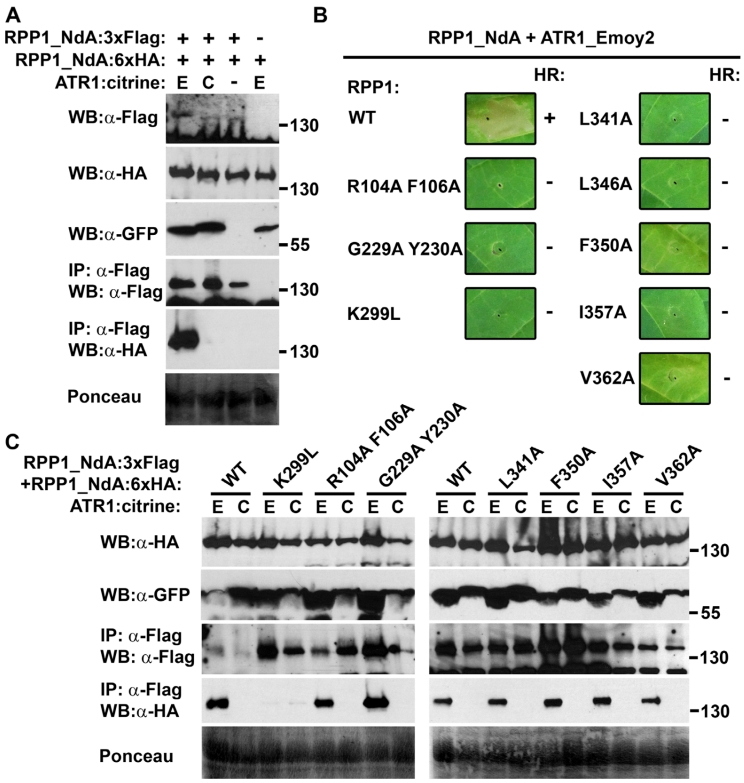


Fig 4. Effector-dependent oligomerization of full-length RPP1 is generally unaffected by mutations in the NB subdomain. (A) RPP1_NdA oligomerization is induced by co-expression with ATR1_Emoy2 (E) but not ATR1_Cala2 (C). Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 36 hours post-infiltration (hpi) for co-immunoprecipitation using α-Flag agarose beads. The expression of ATR1:citrine was detected with an α-GFP antibody. Staining of RuBisCO with Ponceau S provides a loading control. (B) Site-directed mutagenesis of the NB subdomain compromises effector-dependent HR induction by RPP1_NdA. Constructs were transiently co-expressed with ATR1_Emoy2 in *N. tabacum* and images were captured at 48 hpi. The presence or absence of HR is indicated by a "+" or "-", respectively. Protein expression for these constructs is documented in S9 Fig. (C) Detection of self-association of RPP1_NdA NB subdomain mutants by co-immunoprecipitation, tested as in (A). Experiments were performed at least three times with similar results.

doi:10.1371/journal.ppat.1005769.g004

substitutions eliminated the effector-dependent HR (Fig 4B). Interestingly, effector-induced oligomerization was only significantly compromised in the P-loop (K299L) mutant (Fig 4C). Neither R104A F106A and G229A Y230A in the TIR domain, nor mutations on the putative oligomerization surface in the NB domain compromised self-association. Thus, mutations that affected dimerization of individual domains rarely compromised the oligomerization capacity

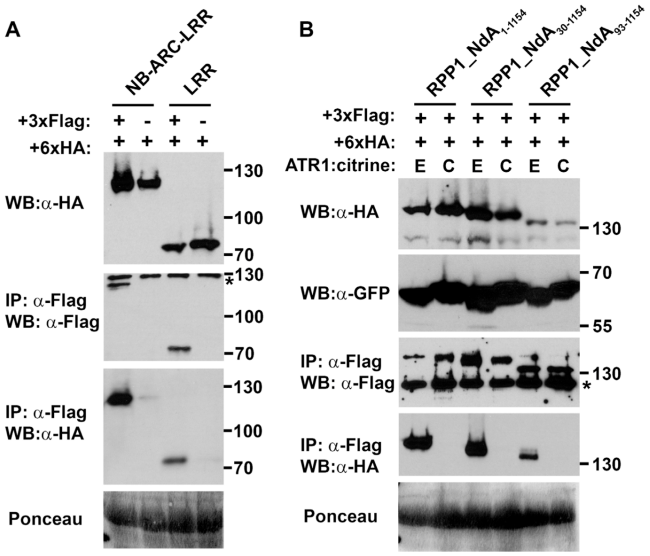


Fig 5. The TIR domain inhibits RPP1 self-association in the absence of the effector ATR1. (A) RPP1_NdA constructs lacking the TIR domain self-associate in an effector-independent manner. Differentially epitope-tagged versions of either NB-ARC-LRR or LRR proteins were transiently expressed in *N. benthamiana* and samples were collected at 48 hours post-infiltration (hpi) for co-immunoprecipitation using α-Flag agarose beads. Asterisks indicate non-specific bands. Staining of RuBisCO with Ponceau S provides a loading control. (B) The TIR domain alone is responsible for inhibiting RPP1 self-association, as N-terminal truncations up to the TIR domain retain effector-dependent self-association. Co-immunoprecipitation experiments were performed using differentially tagged proteins as in (A) except that tissue samples were collected at 36 hpi (E = ATR1_Emoy2, C = ATR1_Cala2). The expression of ATR1:citrine was detected with an α-GFP antibody. The specific amino acids comprising each construct are indicated in subscript. Experiments were performed three times with similar results.

doi:10.1371/journal.ppat.1005769.g005

of the full-length protein. Importantly, this indicates that RPP1 oligomerization is not an indirect byproduct of effector-induced cell death, but rather precedes the HR.

Regulation of RPP1 intermolecular interactions

Considering the differential impact of NB mutations on self-association between NB-ARC1 constructs versus full-length RPP1_NdA, it is possible that the LRR domain participates in the oligomerization event as well. Indeed, we detected LRR (aa 606–1154) and NB-ARC-LRR (aa 255–1154) self-association by co-immunoprecipitation, even in the absence of ATR1 (Fig 5A). The RPP1 LRR domain did not interact with the LRR domain of RPS2 (S14 Fig), suggesting that self-association is not due to non-specific protein-protein interactions. Notably, the addition of the TIR domain abolished this effector-independent self-association (Fig 5). Functionally, we observed that truncation of the first thirty amino acids of RPP1_NdA delayed the HR induced by ATR1_Emoy2 recognition, while removal of the first 92 amino acids (up to the predicted N-terminal border of the TIR domain) completely eliminated the cell death response

(S15 Fig). Both of these truncations retained the capability for effector-dependent oligomerization (Fig 5B). As such, the TIR domain alone is an important negative regulator of RPP1 self-association, while additional N-terminal sequences are required for the function of the TIR domain as a positive regulator of cell death induction.

Intramolecular interactions

The demonstration of effector-induced oligomerization provided a snapshot of RPP1 in a post-activation state, but we also sought to clarify the domain organization of this protein prior to effector recognition. Through co-immunoprecipitation experiments we found that non-autoactive N-TIR domains, including the R104A F106A and G229A Y230A mutants of the NdA allele, interacted with the NB-ARC1 subdomain (Fig 6A). Less NB-ARC1 protein interacted with the wild-type N-TIR domain, although initial levels of N-TIR domain protein were somewhat lower as well. Using the NdA N-TIR R104A F106A mutant, we demonstrated that the NB subdomain comprised the minimal interaction region, although co-immunoprecipitation yields were slightly lower relative to NB-ARC1 (S16 Fig). Mutations in the P-loop (K299L) or at the putative oligomerization interface (L341A, F350A) appeared to reduce the binding of NB-ARC1 to TIR R104A F106A (Fig 6B), although input levels of NB-ARC1 mutant proteins were also lower.

Interactions were also detected between the NB-ARC1 subdomain and the LRR domain. The NB subdomain was again sufficient for this association, while little or no interaction was observed between the LRR and NB-ARC1-ARC2 (Fig 7A, S17 Fig). Given the proximity of ARC2 and the LRR within RPP1, this lack of binding was unexpected. We excluded the possibility that the location of the epitope tags precluded the interaction, because switching the tag to the N-terminus of NB-ARC1-ARC2 did not alter the co-immunoprecipitation results (S17

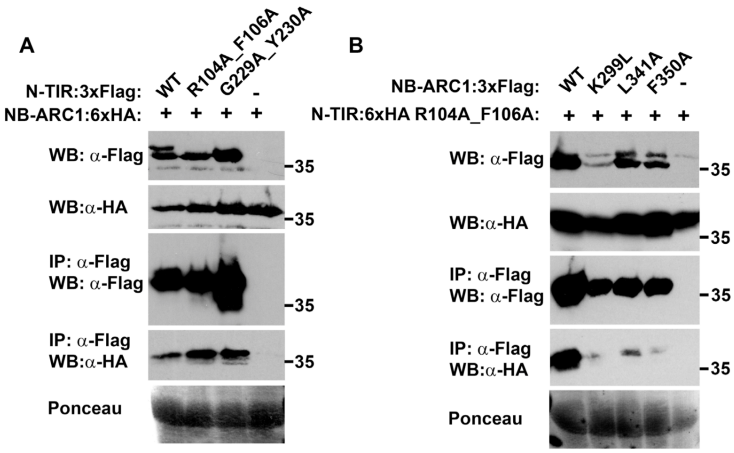


Fig 6. Evaluation of interactions between the N-TIR and NB-ARC1 domains of RPP1_NdA. (A) Non-autoactive N-TIR domain mutants associate with the NB-ARC1 subdomain. (B) Putative oligomerization interface mutants weaken the N-TIR:NB-ARC1 interaction. Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 36 hours post-infiltration (hpi) for (A) and 48 hpi for (B). Co-immunoprecipitations were performed using α-Flag agarose beads. Staining of RuBisCO with Ponceau S provides a loading control. Experiments were performed three times with similar results.

doi:10.1371/journal.ppat.1005769.g006

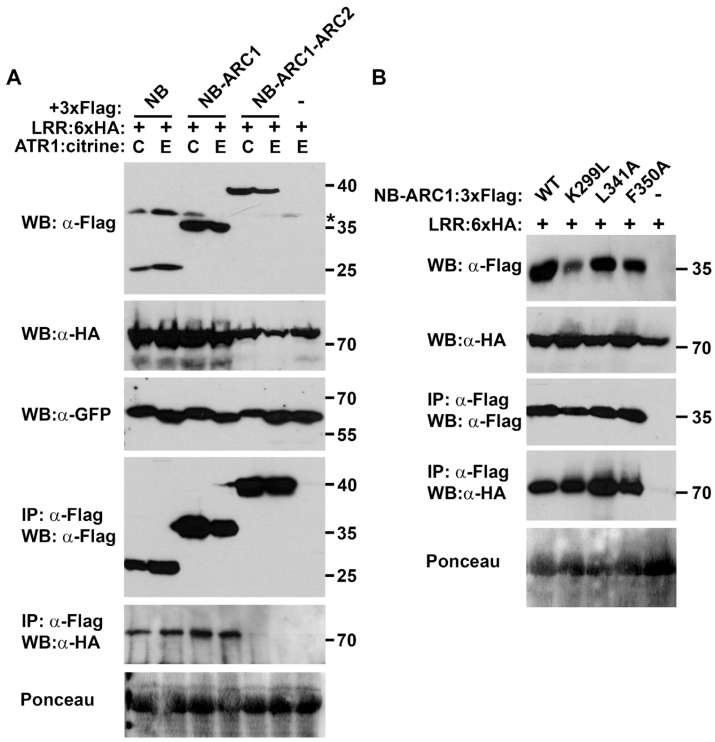


Fig 7. Evaluation of interactions between the LRR and NB(-ARC) domains of RPP1_NdA. The association between the NB(-ARC1) subdomain and the LRR is not disrupted by the presence of ATR1 (A) or by mutations at the putative oligomerization interface (B). Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 48 hours post-infiltration (E = ATR1_Emoy2, C = ATR1_Cala2). Co-immunoprecipitations were performed using α-Flag agarose beads. The expression of ATR1:citrine was detected with an α-GFP antibody. Asterisks indicate non-specific bands. Staining of RuBisCO with Ponceau S provides a loading control. Experiments were performed three times with similar results.

doi:10.1371/journal.ppat.1005769.g007

Fig). Focusing on the interaction between the LRR and NB-ARC1, we observed that the presence of ATR1 did not affect binding (Fig 7A), nor did mutations in the P-loop or at the putative oligomerization interface (Fig 7B). Finally, we detected interactions between the LRR domain and non-autoactive NdA N-TIR domain mutants, with reproducibly higher co-immunoprecipitation yields observed with the R104A F106A mutant versus G229A Y230A (Fig 8). Furthermore, the N-TIR R104A F106A protein associated with LRR and NB-ARC-LRR proteins in an effector-independent fashion (S18 Fig). Taken together, these data suggest that the N-TIR and LRR domains contact the NB domain at different interfaces, and that effector recognition may not induce a complete disruption of these interactions.

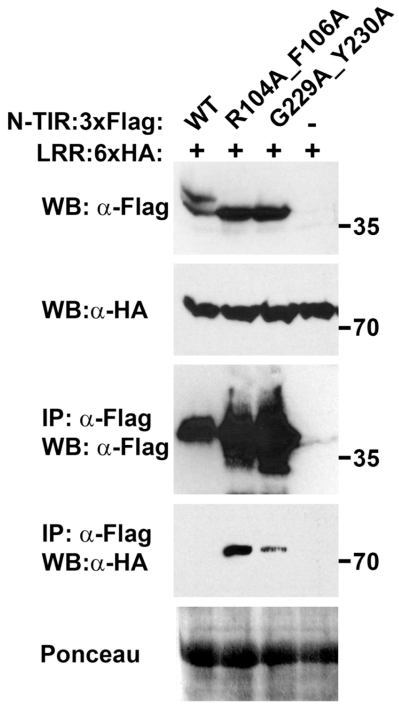


Fig 8. Non-autoactive N-TIR domain mutants associate with the LRR domain of RPP1_NdA. Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 36 hours post-infiltration. Co-immunoprecipitations were performed using α-Flag agarose beads. Staining of RuBisCO with Ponceau S provides a loading control. Experiments were performed three times with similar results.

doi:10.1371/journal.ppat.1005769.g008

Discussion

Effector-independent cell death induction by the N-TIR domain of RPP1

Plant NLR proteins are characterized by a number of conserved domains whose collective activity yields tightly regulated immune responses. When expressed on its own, the N-TIR domain from RPP1_NdA elicited cell death even in the absence of ATR1. Similar autoactivity was noted for the TIR domains of RPS4 from Arabidopsis [24] and L6 from flax [23]. In both cases, the C-terminal border of the minimal autoactive region extended beyond the conserved TIR domain sequence predicted by Pfam [27] to include α-helix E. We also observed this with RPP1. Given the lower protein expression levels associated with further C-terminal truncations, the additional sequence likely affects protein stability.

Interestingly, N-TIR domains from different RPP1 alleles varied substantially in their autoactivity, allowing us to ascribe specific sequences to this phenotypic polymorphism. Using

chimeric constructs, we noted that a region of 50 amino acids directly N-terminal to the predicted TIR domain from RPP1_NdA conferred weak autoactivity to an otherwise inactive background of RPP1_WsB. The functional role of this region is not known and its structure is predicted to be disordered, although our SEC-MALS analysis suggested that it contributes to TIR domain self-association (Fig 2). *In planta*, we noted that this N-terminal sequence is required for TIR domain autoactivity and for effector-dependent cell death induction by full-length RPP1_NdA. It is also worth noting that this region represents the most polymorphic sequence of the N-TIR domain between the NdA and WsB alleles, with 14 aa of insertions/deletions and 67% aa identity (versus 90% identity within the TIR domain).

The analysis of polymorphisms within the predicted TIR domain yielded a small number of residues with significant contributions to autoactivity. In a WsB background, the substitutions K98R I100F, K228S, or R230C enabled the elicitation of cell death at levels near that of the N-TIR domain from RPP1_NdA. Using the crystal structure of the L6 TIR domain [23] as a template for homology modeling, residue K98 is predicted to be exposed on the surface of the protein near α-helix A, while I100 may be located more in the interior of the protein (S2 Fig). Residues K228 and R230 are likely both surface-exposed between β-sheet E and x-helix E. The L6 crystal structure revealed a TIR domain dimerization interface spanning α-helix D to α-helix E, which Bernoux *et al.* [23] functionally interrogated by site-directed mutagenesis. The substitution R73A (near RPP1_WsB K98) abolished autoactivity, although the protein retained the ability to self-associate in a yeast two-hybrid assay. Conversely, an alanine substitution at D208 (corresponding to RPP1_WsB R230) resulted in weak autoactivity despite no detectable self-association. Additional TIR domain structures are available for RPS4 and RRS1, indicating homo- and heterodimerization interfaces at α-helices A and E [26]. Based on the SEC-MALS data, self-association of the RPP1 N-TIR domain may be influenced by both L6- and RPS4-like interfaces (Fig 2). More importantly, we demonstrated that allelic variation in N-TIR domain autoactivity is closely tied to the propensity for N-TIR domain self-association in solution. At least some degree of self-association was observed for all of the proteins that we analyzed, suggesting that some minimum level of association must be achieved for the activation of cell death. This may reflect a requirement for TIR domain dimers of sufficient stability to recruit adaptor proteins [28] for the execution of a cell death program.

The regulatory role of the NB-ARC domain

The NB-ARC domain is generally considered to act as a molecular switch for NLR protein activation [3], but we have demonstrated an important role for this domain as a platform for oligomerization. This function is particularly important for RPP1_WsB, whose N-TIR domain exhibited autoactivity only with the inclusion of the NB and ARC1 subdomains. Krasileva *et al.* [21] previously showed that a C-terminal GFP fusion could facilitate autoactivity of the RPP1_WsB N-TIR domain, and that this effect was dependent upon the ability of GFP to dimerize. A similar dependence on GFP was noted for the CC domains from maize Rp1-D21 and Rp-1D [29]. A major difference, however, is that CC domain constructs containing even a small portion of the NB subdomain significantly weakened cell death induction, and the inclusion of the complete NB-ARC region blocked autoactivity. For the L6 TIR domain, the addition of NB or NB-ARC1 sequences reduced the cell death phenotype to a weak chlorosis which became even weaker, but not absent, with a TIR-NB-ARC1-ARC2 construct [23]. We observed more defined phenotypes with RPP1, in that an N-TIR-NB-ARC1 construct induced macroscopic cell death, while N-TIR-NB was extremely weak and an N-TIR-NB-ARC1-ARC2 construct was essentially inactive. As even a small truncation of the ARC1 subdomain eliminated autoactivity, it appears that the NB-ARC1 region comprises a discrete regulatory module.

Insight into the potential function of this module was provided by co-immunoprecipitation experiments, which indicated that only NB-ARC1 proteins could self-associate. Again, this differentiates RPP1 from L6, where TIR domain self-association gradually diminished with the addition of NB and NB-ARC1 sequences, while TIR-NB-ARC1-ARC2 did not self-associate [23]. In either case, this provides further evidence for the negative regulatory function of the ARC2 subdomain as suggested previously by van Ooijen *et al.* [4].

Following the demonstration of self-association between NB-ARC1 proteins, we investigated the specific residues that influence this interaction. Substitution of a highly conserved lysine within the P-loop motif (K299L) eliminated NB-ARC1 self-association and N-TIR-NB-ARC1 autoactivity (Fig 3C). This residue was shown to be critical for ATP binding by the flax rust resistance protein M [5], and the loss of NLR protein function arising from mutations within the P-loop is well documented [30–34]. Structurally, the P-loop motif is buried within the NB subdomain, so its impact on protein self-association is likely exerted through conformational changes that are thought to accompany nucleotide exchange at this site [3].

The identification of residues at the physical protein interaction interface(s) of the NB-ARC1 region was guided by structural similarity with animal immune receptors. The activated *C. elegans* apoptosome comprises an octamer of CED-4 proteins forming a funnel-shaped structure, within which the α/β domains of each CED-4 member are closely aligned [13]. Structural data indicated that these interactions were stabilized by the stacking of hydrophobic side chains between neighboring α-helices. The importance of these residues for oligomerization was emphasized by mutagenesis within the homologous region of the mammalian protein Apaf-1, which eliminated apoptosome assembly. Likewise, we found that alanine substitutions of surface-exposed residues on the equivalent α-helices of RPP1_NdA reduced the self-association of NB-ARC1 proteins. While we currently lack the structural knowledge to confirm that this region acts as an interaction interface, it is worth noting that most of the surface-exposed residues are relatively hydrophobic, which is typical of many homodimerization interfaces [35–37].

An unusual observation was that the substitution E365A strengthened the intensity of cell death induced by an N-TIR-NB-ARC1 construct or by full-length RPP1_NdA co-inoculated with ATR1_Emoy2. While a glutamate-to-alanine substitution represents an increase in local surface hydrophobicity, similar increases in HR strength resulted from substitutions to the relatively hydrophilic amino acids glutamine and lysine (S11 Fig). As such, the effect may be more influenced by electrostatic forces, because all of these substitutions remove the negative charge at residue 365. In addition, the E365A substitution did not appear to increase the strength of self-association for NB-ARC1 proteins (S12 Fig), suggesting other mechanisms of cell death enhancement such as stabilization of the active state of RPP1. Regardless of the underlying mechanism, the ability to strengthen NLR-mediated immune responses in the absence of autoactivity could be useful for enhancing resistance in agricultural crops, as suggested by "sensitized" NB-ARC domain mutants of other plant NLRs [38–40].

Effector-dependent oligomerization of full-length RPP1

While the assembly of higher-order immunoregulatory structures is well documented in animal systems [8,15], data from plant NLR proteins are relatively sparse and somewhat divergent. Effector-independent self-association was observed in members of the CC domain-containing family of NLR proteins, including RPS5, Rp1-D21, Rp1-D, Rp1-dp2, and MLA [29,41,42]. This behavior has not been observed in TIR domain-containing receptors, although RPS4 and RRS1 form a unique hetero-oligomer in the absence of recognized effectors [26]. Until now, the only example of elicitor-dependent oligomerization involved the tobacco N protein, whose

homotypic interaction occurred only upon recognition of the helicase domain of the tobacco mosaic virus replicase [18]. Similar to our observations, mutations in the P-loop motif of N eliminated oligomerization, while mutations that abolished TIR domain autoactivity did not compromise self-association of full-length N proteins. Mestre and Baulcombe [18] also noted that mutations within the highly conserved "RNBS-A" motif disrupted the function of N while still permitting elicitor-dependent oligomerization. Our mutagenesis experiments in RPP1_NdA included residues L341 and L346, which are part of the RNBS-A motif and whose mutation yielded similar phenotypes. Overall, however, mechanistic interpretation of the results with the N protein was impaired by the inability to transiently express the NB-ARC domain *in planta*. We did not face such an obstacle with RPP1, and were able to demonstrate that residues within the RNBS-A motif and in the adjoining α-helix are involved in self-association of the NB-ARC1 subdomain. It is noteworthy that, while most of the substitutions retained some level of NB-ARC1 self-association, cell death induction was largely eliminated in both N-TIR-NB-ARC1 and full-length contexts. This may indicate that a certain threshold of association is required to activate cell death, and that substitutions of surface-exposed residues result in localized structural perturbations such that this threshold is not exceeded.

At the same time, the retention of effector-dependent oligomerization in mutants of full-length RPP1_NdA suggested that additional domains are involved in this interaction. Indeed, co-immunoprecipitation experiments revealed LRR self-association, even in the absence of ATR1. Similar homotypic interactions were shown for LRRs from RPS5 and Rp1-D21 [29,41], as might be expected given that the full-length proteins also exhibit effector-independent self-association. Full-length RPP1, however, likely exists in monomeric form prior to ATR1 recognition. By examining a series of N-terminal truncations, we found that the TIR domain is required to block effector-independent self-association of RPP1_NdA, likely by occluding the oligomerization interface on the NB domain.

Intramolecular interactions

Resistance proteins control the induction of cell death programs and thus must be tightly regulated. For RPP1, we observed a series of conditional intramolecular interactions with potential negative regulatory functions. Firstly, interactions between the NB-ARC1 and N-TIR domains were most strongly detected with non-autoactive N-TIR domain mutants. Reduced protein accumulation of the wild-type NdA N-TIR domain likely influenced this result, although an appreciable quantity of the wild-type protein was immunoprecipitated. It is relevant to note that, while both CC:CC and CC:NB-ARC interactions were observed for RPS5 and Rp1-D21, the CC domain alone was not autoactive [29,41]. Based on our observations, it is possible that the affinity of N-TIR:NB-ARC1 binding is lower than that of either N-TIR domain self-association or adaptor binding, and that the outcome of these competing interactions is influenced by conformational changes that may accompany RPP1 activation. In this case, the reduced association between the NB-ARC and wild-type N-TIR domain proteins would reflect the active state of RPP1. This principle would also apply to the N-TIR:LRR interaction, which was similarly dependent on N-TIR domain-inactivating mutations. Furthermore, these data suggest that the interaction of the N-TIR domain with other RPP1 domains involves different interfaces than those required for N-TIR domain self-association. As a caveat, results obtained by transient co-expression of separate proteins may not reflect the interactions that occur in *cis*. Additional examples of this interaction from other TIR domain-containing NLR proteins would help to clarify these issues.

In addition to binding the N-TIR domain, we noted that the NB-ARC1 domain also interacts with the LRR domain of RPP1. Mutagenesis experiments revealed residues that

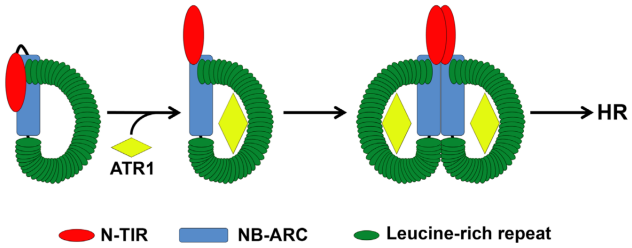


Fig 9. Proposed model of cell death activation by RPP1. In the absence of ATR1, RPP1 is likely maintained in a largely inactive state by a network of N-TIR:NB, NB:LRR, and N-TIR:LRR interactions. Note that the N-TIR:NB interaction occurs at the putative oligomerization interface to prevent effector-independent association. Binding of a recognized allele of ATR1 to the LRR domain may stabilize conformational transitions that reorient the N-TIR domain to expose the oligomerization interface. This allows RPP1 oligomerization via the NB domain, potentially stabilized by LRR:LRR interactions. The reorientation of the N-TIR domain also permits N-TIR domain self-association, which outcompetes N-TIR:NB interactions and ultimately triggers a cell death response (HR). While this model accounts for the effector-independent NB:LRR interaction, the specific interaction interface(s) could differ before and after RPP1 activation.

doi:10.1371/journal.ppat.1005769.g009

differentially affected N-TIR:NB-ARC1 and LRR:NB-ARC1 interactions. Mutations in the NB P-loop motif significantly impaired N-TIR:NB-ARC1 binding but had little impact on the interaction of the LRR with NB-ARC1. A similar discrepancy was observed with the Rx protein from potato, where a P-loop mutation disrupted the CC:NB-ARC-LRR interaction but not the CC-NB-ARC:LRR interaction [43]. The P-loop motif is required for ATP binding [33,44] which is associated with NLR protein activation, perhaps through conformational changes [5,7]. If the CC/TIR and LRR domains interact with different NB-ARC surfaces, localized P-loop-dependent conformational changes in the NB-ARC domain could differentially affect these intramolecular interactions, especially if binding affinities also differ. Evidence for different intramolecular interaction surfaces comes from mutations at the putative oligomerization interface of the NB domain, which also reduced N-TIR:NB-ARC1 interactions to a much greater degree than LRR:NB-ARC1 interactions. Functionally, this supports a role for the TIR domain, and not the LRR, in preventing RPP1 self-association in the absence of ATR1.

For intramolecular interactions involving the LRR domain, we anticipated that effector binding might disrupt these associations, as seen in the NLR proteins Rx1 and Gpa2 [43,45]. We observed, however, that the presence of ATR1 did not affect either LRR:N-TIR or LRR:NB-ARC1 interactions. This may be a consequence of examining domain associations in *trans* rather than in *cis*, although it is difficult to bypass this limitation for *in planta* analyses. For the N-TIR domain, the use of a non-autoactive mutant may have allowed a more stable interaction with the LRR, as the wild-type N-TIR domain did not appear to interact with the LRR. It is also possible that the activation of RPP1 may involve a series of subtle conformational changes that preserve these intramolecular interactions, possibly through shifts in the specific residues bound by each domain.

Overall, our data are compatible with a model (Fig 9) in which RPP1 is maintained in an inactive state by multiple intramolecular interactions. There may be some fluctuation between active and inactive states which is shifted towards the active state by ATR1 binding, as suggested by the recently proposed equilibrium-based switch activation model [6]. Stabilization of the active state by effector binding would be associated with some degree of conformational change sufficient to expose the N-TIR, NB, and LRR domains for self-association. These

changes likely represent a subtle structural rearrangement that may only slightly alter the interfaces of intramolecular interactions, yet facilitate effector-dependent intermolecular associations. The specific details of this conformational transition remain unknown and await the elucidation of the structure of RPP1 as well as its interactions with other downstream signaling components.

Materials and Methods

Cloning

Constructs for transient expression were prepared using the primers listed in S1 Table. In general, sequences were first restriction-cloned into modified pENTR/D-TOPO (Thermo Fisher Scientific, Waltham, MA, USA) vectors so that inserts included 3' sequences encoding either a 3xFlag or 6xhemagglutinin (HA) epitope tag. These entry clones were recombined into the pEarleygate vector pEG100 [46] using LR Clonase (Thermo Fisher Scientific). A construct for transient expression of the RPS2 LRR (aa 490–909, driven by a 35S promoter in vector pGWB14) was kindly provided by Douglas Dahlbeck (University of California, Berkeley). For protein expression in *Escherichia coli*, RPP1 TIR domain sequences were introduced into pLIC171 (T7 promoter, N-terminal His tag) by ligation-independent cloning [47]. All point mutations were introduced using a QuickChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA, USA) and the primers described in S1 Table. Mutagenesis experiments were guided in part by protein homology modeling using the Phyre2 web portal [48] combined with structure visualization using Chimera [49].

Plant materials and growth conditions

Nicotiana tabacum var Turks and *Nicotiana benthamiana* were grown on potting soil in a controlled environment room with a 10-h photoperiod at 24°C. *Agrobacterium tumefaciens* GV3101 (pMP90) was grown at 28°C on Luria-Bertani agar media supplemented with 50 μg/mL gentamycin and 25 μg/mL kanamycin. For transient expression experiments, including co-immunoprecipitations, bacteria from plates were resuspended in 1 mL of induction media (10 mM $MgCl_2$, 10 mM MES, and 150 μM acetosyringone, adjusted to pH 5.6 with KOH), diluted to OD_{600} = 0.45 with induction media, and incubated at room temperature for three hours. For all inoculations, leaves were pierced with a 22 gauge needle prior to infiltration with a 1 mL needleless syringe. Plants were subsequently maintained at room temperature under continuous light until tissue collection at 24 to 48 hours post-infiltration.

Agrobacterium-mediated transient expression and co-immunoprecipitation

To detect transiently expressed proteins, two leaf discs were obtained from inoculated tissues using a #7 cork borer (1.2 cm diameter). The leaf discs were placed in a 1.7 mL SnapLock microfuge tube (Thermo Fisher Scientific) along with a 3 mm glass bead and frozen in liquid nitrogen. Samples were homogenized for 1.5 min in a chilled aluminum block with a Mini-Beadbeater-96 (BioSpec Products, Inc., Bartlesville, OK, USA) and suspended in 130 μL of modified Laemmli buffer (0.24 M Tris-Cl, pH 6.8, 6% SDS, 30% glycerol, 16% β-mercaptoethanol, 0.006% bromophenol blue, and 10 M urea). Samples were vortexed for 30 s, boiled for 5 min, centrifuged at maximum speed for 1 min in a microcentrifuge at room temperature, and the supernatants were transferred to 1.5 mL microfuge tubes. Generally, 12 μL of each sample was separated on 7% or 12.5% discontinuous SDS-PAGE gels and transferred to a nitrocellulose membrane (GVS North America, Sanford, ME, USA). Antibodies for protein detection

included monoclonal ANTI-FLAG M2-Peroxidase (Sigma-Aldrich, St. Louis, MO, USA), anti-HA-Peroxidase (clone 3F10; Sigma-Aldrich), mouse monoclonal anti-GFP (EMD Millipore, Billerica, MA, USA), and monoclonal ANTI-FLAG M2 antibody (Sigma-Aldrich). The latter two antibodies were used in conjunction with a goat-anti-mouse HRP secondary antibody. Antibody binding was visualized with either SuperSignal West Pico or SuperSignal West Femto Maximum Sensitivity Chemiluminescent Substrate (Thermo Fisher Scientific).

Co-immunoprecipitation experiments were conducted essentially as described by Krasileva *et al.* [21] with some modifications. One gram (fresh weight) of leaf tissue was frozen in liquid nitrogen and homogenized with a mortar and pestle, then transferred to a prechilled mortar containing 2 mL of protein extraction buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 0.2% Nonidet P-40, 6 mM β-mercaptoethanol, and cOmplete ULTRA Protease Inhibitor Cocktail [Sigma-Aldrich]) per gram of tissue and homogenized with a prechilled pestle for an additional minute. The homogenate was transferred to 1.5 mL microfuge tubes and centrifuged at 16,000 x g for 20 min at 4°C. Subsequently, 1.4 mL of the supernatant (approximately 70% of the total extract) was transferred to a 1.5 mL microfuge tube containing 15 μL of ANTI-FLAG M2 Affinity Gel (Sigma-Aldrich) and incubated for three hours on a rotator at 4°C. The tubes were centrifuged at 1,500 x g for 10 sec at 4°C, resuspended in 0.5 mL of extraction buffer, and centrifuged in the same manner again, repeating the process for a total of three washes. The beads were resuspended in 50 μL of Laemmli buffer, boiled for 5 min, then centrifuged for 1 min at maximum speed in a microfuge at room temperature. Twelve microliters of each sample was separated on 7% or 12.5% discontinuous SDS-PAGE gels and transferred to a nitrocellulose membrane. Immunodetection of co-immunoprecipitation samples was conducted as described above. For figures in which Flag-tagged inputs are not shown, levels of these proteins were below the threshold of detection with the α-Flag antibody.

Measurement of electrolyte leakage

To quantify the extent of cell death elicited by various constructs, *Agrobacterium* suspensions were prepared as described above and infiltrated into *N. tabacum* leaves such that each technical replicate comprised all constructs infiltrated into sections of the same leaf. When macroscopic HR symptoms became visible (usually within 28 hours post-infiltration), four leaf discs were collected using a #7 cork borer (total area ≈ 4.5 cm^2) and floated on 5 mL of distilled water for 30 min. Leaf discs were then transferred to polypropylene culture tubes containing 6 mL of distilled water and placed under continuous light at room temperature. Twenty-four hours later, electrolyte leakage was measured using a Thermo Orion Model 105A Plus Conductivity Meter (Thermo Fisher Scientific).

Analysis of protein self-association by size-exclusion chromatography and multi-angle laser light scattering

The expression of RPP1 (N-)TIR domain proteins in *E. coli* BL21 (DE3) cells was induced overnight at 20°C via autoinduction [50]. Proteins were purified by immobilised metal-affinity chromatography with the use of a 6x histidine tag. The tag was cleaved from the proteins with TEV protease overnight at 4°C before separation by size-exclusion chromatography (SEC) on a Superdex S75 26/60 HiLoad column (GE Healthcare, Silverwater, NSW, Australia) in a buffer containing 10 mM HEPES pH 7.5, 150 mM NaCl, and 1 mM dithiothreitol (DTT).

For protein molecular mass determination, 175 μg of purified protein was separated on a Superdex Increase 200 5/150 GL SEC column (GE Healthcare) in a buffer containing 10 mM HEPES pH 8.0, 150 mM NaCl, 1 mM DTT, with laser light scattering measured by a Dawn Heleos II 18-angle light-scattering detector coupled with an Optilab TrEX refractive index

detector (Wyatt Technology, Santa Barbara, CA, USA). Astra6.1 software (Wyatt Technology) was used to perform molecular mass calculations. Molecular mass is reported at the protein elution peak (corresponding to the refractive index peak) from the SEC column, where the signal from light-scattering is strong. Estimations of molecular mass were performed across the protein elution peak, with refractive increment (dn/dc values) fixed at 0.186 mL/g, with the assumption that dn/dc is invariable for unmodified proteins [51].

Sequence accession numbers

Sequence data for the genes examined in this study can be retrieved from the National Center for Biotechnology Information using the following accession numbers: RPP1_NdA (HM209027), RPP1_WsB (AAC72978), ATR1_Emoy2 (AAX51198), ATR1_Cala2 (AAX51204), and RPS2 (AAA21874).

Supporting Information

S1 Fig. Quantification of electrolyte leakage induced by allelic chimeras and site-directed mutants of the RPP1 N-TIR domain. Data in (A) and (B) correspond to the constructs tested in (D) and (E) of Fig 1, respectively. The specific amino acids comprising each construct are indicated in subscript. Leaf discs representing approximately 4.5 cm^2 of tissue were collected at approximately 28 hours post-infiltration and electrolyte concentration (conductivity) was measured 24 hours after collection. Error bars indicate standard deviation, and letters above data points indicate statistical significance groups as determined by pairwise Student's t-tests (α = 0.05). Experiments were performed at least three times with similar results.
(TIF)

S2 Fig. Analyses of polymorphic sequences within the RPP1 N-TIR domain. (A) Amino acid alignment of N-TIR domain sequences from the Niederzenz (NdA), Wassilewskija (WsB), and Estland-1 (Est-1) alleles of RPP1. The TIR domain borders predicted by Pfam are delimited by square brackets, and polymorphic amino acids with potential relevance to protein function are highlighted in red. Residues shown to influence autoactivity are indicated with arrowheads. (B) Functionally relevant polymorphic residues are also indicated on a putative structure of the RPP1_WsB TIR domain, derived by homology modeling using structural data from the L6 TIR domain (PDB: 3OZI). Relevant α-helices are also annotated. (C) Autoactivity of N-TIR domains from different RPP1 alleles. Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient expression and images of hypersensitive response (HR) phenotypes were captured at 48 hours post-infiltration. The presence or absence of HR is indicated by a "+" or "-", respectively. The specific amino acids comprising each construct are indicated in subscript. (D) An α-Flag antibody was used to evaluate protein expression, while staining of RuBisCO with Ponceau S provided a loading control. The experiment was performed three times with similar results.
(TIF)

S3 Fig. Additivity of K98R and I100F substitutions in conferring autoactivity to the N-TIR domain from the WsB allele. Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient expression and images of hypersensitive response (HR) phenotypes were captured at 48 hours post-infiltration. HR phenotypes are scored as negative (-), weak (w), or strong (+). The specific amino acids comprising each construct are indicated in subscript. An α-Flag antibody was used to evaluate protein expression, while staining of RuBisCO with Ponceau S provided a loading control. The experiment was performed three times with

similar results.
(TIF)

S4 Fig. Identification of loss-of-function mutations in the N-TIR domain from the NdA allele. Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient expression and images of hypersensitive response (HR) phenotypes were captured at 48 hours post-infiltration. The presence or absence of HR is indicated by a "+" or "-", respectively. The specific amino acids comprising each construct are indicated in subscript. An α-Flag antibody was used to evaluate protein expression, while staining of RuBisCO with Ponceau S provided a loading control. The experiment was performed three times with similar results.
(TIF)

S5 Fig. The NB-ARC domain influences N-TIR domain autoactivity in RPP1_NdA. Leaf discs representing approximately 4.5 cm^2 of tissue were collected at approximately 28 hours post-infiltration and electrolyte concentration (conductivity) was measured 24 hours after collection. Error bars indicate standard deviation, and letters above data points indicate statistical significance groups as determined by pairwise Student's t-tests ($\alpha = 0.05$). The experiment was performed three times with similar results.
(TIF)

S6 Fig. The ARC1 domain comprises a discrete functional module in the regulation of N-TIR domain autoactivity. (A) Hypersensitive response (HR) phenotypes associated with C-terminal truncations of an N-TIR-NB-ARC1 construct from the NdA allele. The amino acids at the C-terminus of each truncation are indicated in subscript. Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient expression and images were captured at 48 hours post-infiltration. HR phenotypes are scored as negative (-), very weak (vw), weak (w), or strong (+). (B) An α-Flag antibody was used to evaluate protein expression, while staining of RuBisCO with Ponceau S provided a loading control. Experiments were performed at least three times with similar results. (C) Predicted structure of the RPP1 NB-ARC domain based on homology modeling using the Drosophila Dark protein (PDB:4v4l) as a template. The three subdomains of this region are highlighted and the locations of each truncation tested in (A) are indicated by arrows.
(TIF)

S7 Fig. Autoactive RPP1 N-TIR-NB-ARC1 proteins are capable of self-association. Differentially epitope-tagged N-TIR-NB-ARC1 proteins from the Niederzenz (N) and Wassilewskija (W) alleles of RPP1 were transiently expressed in *Nicotiana benthamiana* and samples were collected at 48 hours post-infiltration for co-immunoprecipitation using α-Flag agarose beads. Staining of RuBisCO with Ponceau S provides a loading control. The experiment was performed three times with similar results.
(TIF)

S8 Fig. Co-immunoprecipitation yields from self-association experiments using the NB-ARC1 subdomain of RPP1 from alleles Niederzenz (N) and Wassilewskija (W). Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 48 hours post-infiltration for co-immunoprecipitation using α-Flag agarose beads. Staining of RuBisCO with Ponceau S provides a loading control. The experiment was performed three times with similar results.
(TIF)

S9 Fig. *In planta* expression of constructs used for hypersensitive response assays. (A) N-TIR domain addition constructs tested in Fig 3A, including sequences from the RPP1 alleles

Niederzenz (NdA) and Wassilewskija (WsB). (B) N-TIR-NB-ARC1 (NdA) constructs tested in Fig 3C. (C) Full-length RPP1 (NdA) constructs tested in Fig 4. Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 24 hours post-infiltration (hpi) for (A) and (B) or 48 hpi for (C). Staining of RuBisCO with Ponceau S provides a loading control. The experiments were performed three times with similar results.
(TIF)

S10 Fig. Location of residues of interest at the putative oligomerization interface of the RPP1 NB subdomain. The α-helices that comprise the oligomerization interface of the *Caenorhabditis elegans* CED-4 protein (PDB: 2A5Y) are shown on the left, with residues that contribute to oligomerization highlighted in red. The corresponding region of RPP1 is shown at center; residues required for induction of the hypersensitive response are highlighted in red, while those that were dispensable for this response are depicted in green. The predicted structure of the RPP1 NB-ARC domain is shown on the right and is derived from homology modeling using the Drosophila Dark protein (PDB:4v4l) as a template. The putative oligomerization interface region is highlighted in red. Note that the enlarged image of this region is rotated approximately 90° counterclockwise on the y-axis for improved visibility of the residues of interest.
(TIF)

S11 Fig. Amino acid substitutions at glutamate 365 (E365) in RPP1_NdA enhance cell death phenotypes. (A) Hypersensitive response phenotypes elicited by N-TIR-NB-ARC1 constructs or by co-expression of full-length RPP1_NdA and ATR1_Emoy2. Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient expression, using a range of inoculum concentrations (as measured by OD_{600}) to allow a comparison of the strength of cell death induction by wild-type (WT) and E365 substitution mutants. Images were captured at 48 hours post-infiltration (hpi). (B) An α-Flag antibody was used to evaluate protein expression, while staining of RuBisCO with Ponceau S provided a loading control. (C) Quantification of electrolyte leakage induced by various N-TIR-NB-ARC1 constructs. Leaf discs representing approximately 4.5 cm^2 of tissue were collected at approximately 28 hpi and electrolyte concentration (conductivity) was measured 24 hours after collection. The non-autoactive F350A mutant and autoactive N-TIR domain were included as negative and positive controls, respectively. Error bars indicate standard deviation, and letters above data points indicate statistical significance groups as determined by pairwise Student's t-tests ($\alpha = 0.05$). Experiments were performed three times with similar results.
(TIF)

S12 Fig. Mutagenesis of glutamate 365 (E365) does not increase NB-ARC1 self-association. Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 48 hours post-infiltration for co-immunoprecipitation using α-Flag agarose beads. Asterisks indicate non-specific bands. Staining of RuBisCO with Ponceau S provides a loading control. The experiment was performed three times with similar results.
(TIF)

S13 Fig. ATR1_Emoy2 is present in the effector-dependent RPP1_NdA oligomer. Differentially epitope-tagged RPP1_NdA proteins were transiently co-expressed with either ATR1_Emoy2 (E) or ATR1_Cala2 (C) in *Nicotiana benthamiana* and samples were collected at 36 hours post-infiltration for co-immunoprecipitation using α-Flag agarose beads. Due to a non-specific signal at the expected molecular weight of ATR1:citrine when the elution fraction was probed with an α-GFP antibody, HA:ATR1:citrine constructs were used and ATR1 expression detected with an α-HA antibody. Staining of RuBisCO with Ponceau S provides a loading control. The experiment was performed three times with similar results.
(TIF)

S14 Fig. The LRR domain of RPP1_NdA does not interact non-specifically with an unrelated LRR. While self-association of the RPP1_NdA LRR is detected, this protein does not interact with the LRR domain of RPS2. Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 48 hours post-infiltration for co-immunoprecipitation using α-Flag agarose beads. Staining of RuBisCO with Ponceau S provides a loading control. Experiments were performed three times with similar results.
(TIF)

S15 Fig. The first thirty amino acids of RPP1_NdA are not required for cell death induction. Constructs were tested in *Nicotiana tabacum* via *Agrobacterium*-mediated transient co-expression with ATR1_Emoy2, and images of hypersensitive response phenotypes were captured at 48 and 72 hours post-infiltration (hpi). The constructs included RPP1_NdA$_{1\text{-}1154}$ (1), RPP1_NdA$_{30\text{-}1154}$ (2), and RPP1_NdA$_{93\text{-}1154}$ (3), where the specific amino acids comprising each construct are indicated in subscript. An α-HA antibody was used to evaluate protein expression, and the expression of ATR1_Emoy2:citrine was detected with an α-GFP antibody. Staining of RuBisCO with Ponceau S provided a loading control. The experiment was performed three times with similar results.
(TIF)

S16 Fig. The NB subdomain is sufficient for interaction with the N-TIR domain of RPP1_NdA. Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 48 hours post-infiltration for co-immunoprecipitation using α-Flag agarose beads. Staining of RuBisCO with Ponceau S provides a loading control. The experiment was performed three times with similar results.
(TIF)

S17 Fig. The NB subdomain is sufficient for interaction with the LRR domain of RPP1_NdA. An NB-ARC1-ARC2 construct only weakly binds the LRR domain, regardless of whether the epitope tag is located C-terminally (A) or N-terminally (B). Constructs were transiently expressed in *Nicotiana benthamiana* and samples were collected at 48 hours post-infiltration for co-immunoprecipitation using α-Flag agarose beads. Staining of RuBisCO with Ponceau S provides a loading control. Asterisks indicate non-specific bands. The experiments were performed two times with similar results.
(TIF)

S18 Fig. The N-TIR and LRR domains of RPP1_NdA associate in an effector-independent manner. A non-autoactive N-TIR domain mutant (R104A F106A) associates with the LRR, and the interaction is not disrupted by the presence of ATR1. Constructs were transiently expressed in *N. benthamiana* and samples were collected at 48 hours post-infiltration (E = ATR1_Emoy2, C = ATR1_Cala2). Co-immunoprecipitations were performed using α-Flag agarose beads. The expression of ATR1:citrine was detected with an α-GFP antibody. Asterisks indicate non-specific bands. Staining of RuBisCO with Ponceau S provides a loading control. Experiments were performed three times with similar results.
(TIF)

S1 Table. Primers used in this study.
(DOCX)

Acknowledgments

We thank Drs. Jonathan Jones, Jeff Ellis, and Adam Steinbrenner for critical reading of the manuscript as well as the anonymous reviewers for their insightful comments.

Author Contributions

Conceived and designed the experiments: KJS AB SJW BK BJS. Performed the experiments: KJS AB SJW. Analyzed the data: KJS AB SJW BK. Contributed reagents/materials/analysis tools: BK BJS. Wrote the paper: KJS AB BJS.

References

1. Jones JDG, Dangl JL. The plant immune system. Nature. 2006; 444: 323–329. doi: 10.1038/nature05286 PMID: 17108957
2. Coll NS, Epple P, Dangl JL. Programmed cell death in the plant immune system. Cell Death Differ. 2011; 18: 1247–1256. doi: 10.1038/cdd.2011.37 PMID: 21475301
3. Takken FLW, Goverse A. How to build a pathogen detector: Structural basis of NB-LRR function. Curr Opin Plant Biol. 2012; 15: 375–384. doi: 10.1016/j.pbi.2012.05.001 PMID: 22658703
4. Van Ooijen G, Mayr G, Kasiem MM, Albrecht M, Cornelissen BJC, Takken FLW. Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. J Exp Bot. 2008; 59: 1383–1397. doi: 10.1093/jxb/ern045 PMID: 18390848
5. Williams SJ, Sornaraj P, deCourcy-Ireland E, Menz RI, Kobe B, Ellis JG, et al. An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. Mol Plant-Microbe Interact. 2011; 24: 897–906. doi: 10.1094/MPMI-03-11-0052 PMID: 21539434
6. Bernoux M, Burdett H, Williams SJ, Zhang X, Chen C, Newell K, et al. Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. Plant Cell. 2016; 28: 146–159. doi: 10.1105/tpc.15.00303 PMID: 26744216
7. Qi D, Innes RW. Recent advances in plant NLR structure, function, localization, and signaling. Front Immunol. 2013; 4: 1–10. doi: 10.3389/fimmu.2013.00348 PMID: 24155748
8. Chai J, Shi Y. Apoptosome and inflammasome: conserved machineries for caspase activation. Natl Sci Rev. 2014; 1: 101–118. doi: 10.1093/nsr/nwt025
9. Yang X, Chang HY, Baltimore D. Essential role of CED-4 oligomerization in CED-3 activation and apoptosis. Science. 1998; 281: 1355–1357. doi: 10.1126/science.281.5381.1355 PMID: 9721101
10. Acehan D, Jiang X, Morgan DG, Heuser JE, Wang X, Akey CW. Three-dimensional structure of the apoptosome: Implications for assembly, procaspase-9 binding, and activation. Mol Cell. 2002; 9: 423–432. doi: 10.1016/S1097-2765(02)00442-2 PMID: 11864614
11. Yu X, Acehan D, Ménétret J-F, Booth CR, Ludtke SJ, Riedl SJ, et al. A structure of the human apoptosome at 12.8 Å resolution provides insights into this cell death platform. Structure. 2005; 13: 1725–1735. doi: 10.1016/j.str.2005.09.006 PMID: 16271896
12. Yu X, Wang L, Acehan D, Wang X, Akey CW. Three-dimensional structure of a double apoptosome formed by the Drosophila Apaf-1 related killer. J Mol Biol. 2006; 355: 577–89. doi: 10.1016/j.jmb.2005.10.040 PMID: 16310803
13. Qi S, Pang Y, Hu Q, Liu Q, Li H, Zhou Y, et al. Crystal structure of the *Caenorhabditis elegans* apoptosome reveals an octameric assembly of CED-4. Cell. 2010; 141: 446–457. doi: 10.1016/j.cell.2010.03.017 PMID: 20434985
14. Hu Z, Zhou Q, Zhang C, Fan S, Cheng W, Zhao Y, et al. Structural and biochemical basis for induced self-propagation of NLRC4. Science. 2015; 350: 399–404. doi: 10.1126/science.aac5489 PMID: 26449475
15. Yin Q, Fu T, Li J, Wu H. Structural biology of innate immunity. Annu Rev Immunol. 2015; 33: 393–416. doi: 10.1146/annurev-immunol-032414-112258 PMID: 25622194
16. Faustin B, Lartigue L, Bruey J-M, Luciano F, Sergienko E, Bailly-Maitre B, et al. Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. Mol Cell. 2007; 25: 713–724. doi: 10.1016/j.molcel.2007.01.032 PMID: 17349957
17. Hu Z, Yan C, Liu P, Huang Z, Ma R, Zhang C, et al. Crystal structure of NLRC4 reveals its autoinhibition mechanism. Science. 2013; 341: 172–175. doi: 10.1126/science.1236381 PMID: 23765277
18. Mestre P, Baulcombe DC. Elicitor-mediated oligomerization of the tobacco N disease resistance protein. Plant Cell. 2006; 18: 491–501. doi: 10.1105/tpc.105.037234 PMID: 16387833
19. Botella MA, Parker JE, Frost LN, Bittner-Eddy PD, Beynon JL, Daniels MJ, et al. Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. Plant Cell. 1998; 10: 1847–1860. PMID: 9811793
20. Rehmany AP, Gordon A, Rose LE, Allen RL, Armstrong MR, Whisson SC, et al. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. Plant Cell. 2005; 17: 1839–1850. doi: 10.1105/tpc.105.031807.to PMID: 15894715

21. Krasileva KV, Dahlbeck D, Staskawicz BJ. Activation of an Arabidopsis resistance protein is specified by the *in planta* association of its leucine-rich repeat domain with the cognate oomycete effector. Plant Cell. 2010; 22: 2444–2458. doi: 10.1105/tpc.110.075358 PMID: 20601497

22. Steinbrenner AD, Goritschnig S, Staskawicz BJ. Recognition and activation domains contribute to allele-specific responses of an Arabidopsis NLR receptor to an oomycete effector protein. PLoS Pathog. 2015; 11: e1004665. doi: 10.1371/journal.ppat.1004665 PMID: 25671309

23. Bernoux M, Ve T, Williams S, Warren C, Hatters D, Valkov E, et al. Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. Cell Host Microbe. 2011; 9: 200–211. doi: 10.1016/j.chom.2011.02.009 PMID: 21402359

24. Swiderski MR, Birker D, Jones JDG. The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. Mol Plant-Microbe Interact. 2009; 22: 157–165. doi: 10.1094/MPMI-22-2-0157 PMID: 19132868

25. Goritschnig S, Steinbrenner AD, Grunwald DJ, Staskawicz BJ. Structurally distinct *Arabidopsis thaliana* NLR immune receptors recognize tandem WY domains of an oomycete effector. New Phytol. 2016; doi: 10.1111/nph.13823

26. Williams SJ, Sohn KH, Wan L, Bernoux M, Sarris PF, Segonzac C, et al. Structural basis for assembly and function of a heterodimeric plant immune receptor. Science. 2014; 344: 299–303. doi: 10.1126/science.1247357 PMID: 24744375

27. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: The protein families database. Nucleic Acids Res. 2014; 42: 222–230. doi: 10.1093/nar/gkt1223

28. Ve T, Williams SJ, Kobe B. Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. Apoptosis. 2014; 20: 250–261. doi: 10.1007/s10495-014-1064-2

29. Wang G-F, Ji J, El-Kasmi F, Dangl JL, Johal G, Balint-Kurti PJ. Molecular and functional analyses of a maize autoactive NB-LRR protein identify precise structural requirements for activity. PLoS Pathog. 2015; 11: e1004674. doi: 10.1371/journal.ppat.1004674 PMID: 25719542

30. Bendahmane A, Farnham G, Moffett P, Baulcombe DC. Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. Plant J. 2002; 32: 195–204. doi: 10.1046/j.1365-313X.2002.01413.x PMID: 12383085

31. Dinesh-Kumar SP, Tham WH, Baker BJ. Structure-function analysis of the tobacco mosaic virus resistance gene N. Proc Natl Acad Sci U S A. 2000; 97: 14789–14794. doi: 10.1073/pnas.97.26.14789 PMID: 11121079

32. Howles P, Lawrence G, Finnegan J, McFadden H, Ayliffe M, Dodds P, et al. Autoactive alleles of the flax L6 rust resistance gene induce non-race-specific rust resistance associated with the hypersensitive response. Mol Plant-Microbe Interact. 2005; 18: 570–582. doi: 10.1094/MPMI-18-0570 PMID: 15986927

33. Tameling WIL, Elzinga SDJ, Darmin PS, Vossen JH, Takken FLW, Haring MA, et al. The tomato R gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity. Plant Cell. 2002; 14: 2929–2939. doi: 10.1105/tpc.005793.motif PMID: 12417711

34. Tao Y, Yuan FH, Leister RT, Ausubel FM, Katagiri F. Mutational analysis of the Arabidopsis nucleotide binding site- leucine-rich repeat resistance gene RPS2. Plant Cell. 2000; 12: 2541–2554. PMID: 11148296

35. Bahadur RP, Chakrabarti P, Rodier F. Dissecting subunit interfaces in homodimeric proteins. Proteins. 2003; 719: 708–719.

36. Yan C, Wu F, Jernigan RL, Dobbs D. Characterization of protein-protein interfaces. Protein J. 2008; 27: 59–70. doi: 10.1007/s10930-007-9108-x PMID: 17851740

37. Talavera D, Robertson DL, Lovell SC. Characterization of protein-protein interaction interfaces from a single species. PLoS One. 2011; 6. doi: 10.1371/journal.pone.0021053

38. Stirnweis D, Milani SD, Jordan T, Keller B, Brunner S. Substitutions of two amino acids in the nucleotide-binding site domain of a resistance protein enhance the hypersensitive response and enlarge the PM3F resistance spectrum in wheat. Mol Plant-Microbe Interact. 2014; 27: 265–276. doi: 10.1094/MPMI-10-13-0297-FI PMID: 24329172

39. Segretin ME, Pais M, Franceschetti M, Chaparro-Garcia A, Bos JI, Banfield MJ, et al. Single amino acid mutations in the potato immune receptor R3a expand response to *Phytophthora* effectors. Mol Plant-Microbe Interact. 2014; 27: 624–637. doi: 10.1094/MPMI-02-14-0040-R PMID: 24678835

40. Harris CJ, Slootweg EJ, Goverse A, Baulcombe DC. Stepwise artificial evolution of the disease resistance gene. Proc Natl Acad Sci U S A. 2013; 110: 21189–21194. doi: 10.1073/pnas.1311134110 PMID: 24324167

41. Ade J, DeYoung BJ, Golstein C, Innes RW. Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. Proc Natl Acad Sci U S A. 2007; 104: 2531–2536. doi: 10.1073/pnas.0608779104 PMID: 17277084

42. Maekawa T, Cheng W, Spiridon LN, Töller A, Lukasik E, Saijo Y, et al. Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. Cell Host Microbe. 2011; 9: 187–199. doi: 10.1016/j.chom.2011.02.008 PMID: 21402358

43. Moffett P, Farnham G, Peart J, Baulcombe DC. Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. EMBO J. 2002; 21: 4511–4519. doi: 10.1093/emboj/cdf453 PMID: 12198153

44. Hishida T, Iwasaki H, Yagi T, Shinagawa H. Role of Walker Motif A of RuvB protein in promoting branch migration of Holliday junctions. J Biol Chem. 1999; 274: 25335–25342. PMID: 10464259

45. Slootweg EJ, Spiridon LN, Roosien J, Butterbach P, Pomp R, Westerhof L, et al. Structural determinants at the interface of the ARC2 and leucine-rich repeat domains control the activation of the plant immune receptors Rx1 and Gpa2. Plant Physiol. 2013; 162: 1510–28. doi: 10.1104/pp.113.218842 PMID: 23660837

46. Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, et al. Gateway-compatible vectors for plant functional genomics and proteomics. Plant J. 2006; 45: 616–629. doi: 10.1111/j.1365-313X.2005.02617.x PMID: 16441352

47. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 1990; 18: 6069–6074. doi: 10.1093/nar/18.20.6069 PMID: 2235490

48. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015; 10: 845–858. doi: http://dx.doi.org/10.1038/nprot.2015.053 PMID: 25950237

49. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—A visualization system for exploratory research and analysis. J Comput Chem. 2004; 25: 1605–1612. doi: 10.1002/jcc.20084 PMID: 15264254

50. Studier FW. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif. 2005; 41: 207–234. doi: 10.1016/j.pep.2005.01.016 PMID: 15915565

51. Wen J, Arakawa T, Philo JS. Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. Anal Biochem. 1996; 240: 155–166. doi: 10.1006/abio.1996.0345 PMID: 8811899

PNAS PLUS

Multiple functional self-association interfaces in plant TIR domains

Xiaoxiao Zhang[a,b,1], Maud Bernoux[b,1,2], Adam R. Bentham[a,c,1], Toby E. Newman[d,e], Thomas Ve[a,f], Lachlan W. Casey[a], Tom M. Raaymakers[b,g], Jian Hu[b,h], Tristan I. Croll[i], Karl J. Schreiber[j], Brian J. Staskawicz[j,2], Peter A. Anderson[c], Kee Hoon Sohn[d,e], Simon J. Williams[a,k,2], Peter N. Dodds[b,2], and Bostjan Kobe[a,2]

[a]School of Chemistry and Molecular Biosciences, Australian Infectious Diseases Research Centre and Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD 4072, Australia; [b]Agriculture and Food, Commonwealth Scientific and Industrial Research Organisation, Canberra, ACT 2601, Australia; [c]School of Biological Sciences, Faculty of Science and Engineering, Flinders University, Adelaide, SA 5001, Australia; [d]Department of Life Sciences, and School of Interdisciplinary Bioscience and Bioengineering, Pohang University of Science and Technology, Pohang, Gyeongbuk 790-784, Republic of Korea; [e]Bioprotection Research Centre, Institute of Agriculture and Environment, Massey University, Palmerston North 4442, New Zealand; [f]Institute for Glycomics, Griffith University, Southport, QLD 4222, Australia; [g]Department of Biology, Plant–Microbe Interactions, Utrecht University, 3584 CH Utrecht, The Netherlands; [h]College of Biological Sciences, China Agricultural University, Beijing 100094, People's Republic of China; [i]School of Biomedical Sciences, Queensland University of Technology, Brisbane, QLD 4001, Australia; [j]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720; and [k]Research School of Biology, College of Medicine, Biology, and Environment, Australian National University, Canberra, ACT 0200, Australia

Contributed by Brian J. Staskawicz, December 28, 2016 (sent for review November 29, 2016; reviewed by Blake C. Meyers and Ryohei Terauchi)

The self-association of Toll/interleukin-1 receptor/resistance protein (TIR) domains has been implicated in signaling in plant and animal immunity receptors. Structure-based studies identified different TIR-domain dimerization interfaces required for signaling of the plant nucleotide-binding oligomerization domain-like receptors (NLRs) L6 from flax and disease resistance protein RPS4 from *Arabidopsis*. Here we show that the crystal structure of the TIR domain from the *Arabidopsis* NLR suppressor of npr1-1, constitutive 1 (SNC1) contains both an L6-like interface involving helices αD and αE (DE interface) and an RPS4-like interface involving helices αA and αE (AE interface). Mutations in either the AE- or DE-interface region disrupt cell-death signaling activity of SNC1, L6, and RPS4 TIR domains and full-length L6 and RPS4. Self-association of L6 and RPS4 TIR domains is affected by mutations in either region, whereas only AE-interface mutations affect SNC1 TIR-domain self-association. We further show two similar interfaces in the crystal structure of the TIR domain from the *Arabidopsis* NLR recognition of *Peronospora parasitica* 1 (RPP1). These data demonstrate that both the AE and DE self-association interfaces are simultaneously required for self-association and cell-death signaling in diverse plant NLRs.

plant immunity | NLR | TIR domain | plant disease resistance | signaling by cooperative assembly formation

Plants have evolved a sophisticated innate immune system to detect pathogens, in which plant resistance (R) proteins recognize pathogen proteins (effectors) in a highly specific manner. This recognition leads to the effector-triggered immunity (ETI) response that often induces a localized cell death known as the hypersensitive response (1). Most R proteins belong to the nucleotide-binding oligomerization domain (NOD)-like receptor (NLR) family. NLRs are prevalent in the immune systems of plants and animals and provide resistance to a broad range of pathogens, including fungi, oomycetes, bacteria, viruses, and insects (2, 3). NLRs contain a central nucleotide-binding (NB) domain, often referred to as the nucleotide-binding adaptor shared by APAF-1, resistance proteins, and CED-4 (NB-ARC domain) (4) and a C-terminal leucine-rich repeat (LRR) domain. Plant NLRs can be further classified into two main subfamilies, depending on the presence of either a Toll/interleukin-1 receptor domain (TIR–NLR) or a coiled-coil domain (CC–NLR) at their N termini (5).

The CC and TIR domains of many plant NLRs can autonomously signal cell-death responses when expressed ectopically in planta, and mutations in these domains within full-length proteins also compromise signaling, suggesting that these domains are responsible for propagating the resistance signal after activation of the receptor (6–14). Self-association of both TIR (8, 9, 11, 15) and CC (10, 13, 16, 17) domains has been shown to be important for the signaling function. In animal NLRs, the formation of postactivation oligomeric complexes, such as the NLRC4/NAIP inflammasome or the APAF1 apoptosome, is important for bringing together N-terminal domains into a signaling platform (18–20), but there is yet little evidence for such signaling complexes in plants.

Several crystal structures of plant TIR domains have been reported (9, 11, 21–24). These structures reveal a similar overall structure, which consists of a flavodoxin-like fold containing a central parallel β-sheet surrounded by α-helices. This fold is shared with the TIR domains from animal innate immunity proteins, although plant TIR domains generally have an extended αD-helical region that is not found in the animal TIR domains. Whereas the overall structure of plant TIR domains is conserved, the identified self-association interfaces differ. The crystal structure of $L6^{TIR}$ revealed an interface predominantly formed by the αD- and αE-helices (termed here the DE interface) (9). Mutations in this interface disrupt $L6^{TIR}$ self-association and signaling activity (9). In the case of disease resistance protein RPS4 TIR domain ($RPS4^{TIR}$)

Significance

Toll/interleukin-1 receptor/resistance protein (TIR) domains are present in plant and animal innate immunity receptors and appear to play a scaffold function in defense signaling. In both systems, self-association of TIR domains is crucial for their function. In plants, the TIR domain is associated with intracellular immunity receptors, known as nucleotide-binding oligomerization domain-like receptors (NLRs). Previous studies from several plant NLRs have identified two distinct interfaces that are required for TIR:TIR dimerization in different NLRs. We show that the two interfaces previously identified are both important for self-association and defense signaling of multiple TIR–NLR proteins. Collectively, this work suggests that there is a common mechanism of TIR domain self-association in signaling across the TIR–NLR class of receptor proteins.

Author contributions: X.Z., M.B., A.R.B., T.V., L.W.C., B.J.S., K.H.S., S.J.W., P.N.D., and B.K. designed research; X.Z., M.B., A.R.B., T.E.N., L.W.C., T.M.R., J.H., K.H.S., S.J.W., and P.N.D. performed research; M.B., K.H.S., and P.N.D. contributed new reagents/analytic tools; X.Z., M.B., A.R.B., T.V., L.W.C., T.I.C., K.J.S., B.J.S., P.A.A., K.H.S., S.J.W., P.N.D., and B.K. analyzed data; and X.Z., M.B., A.R.B., L.W.C., K.H.S., S.J.W., P.N.D., and B.K. wrote the paper.

Reviewers: B.C.M., Donald Danforth Plant Science Center; and R.T., Iwate Biotechnology Research Center.

The authors declare no conflict of interest.

Data deposition: The protein structures and data used to derive these structures have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 5TEB ($RPP1^{TIR}$) and 5TEC ($SNC1^{TIR}$)].

[1]X.Z., M.B., and A.R.B. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: stask@berkeley.edu, maud.bernoux@csiro.au, simon.williams@anu.edu.au, peter.dodds@csiro.au, or b.kobe@uq.edu.au.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1621248114/-/DCSupplemental.

and disease resistance protein RRS1 TIR domain (RRS1^{TIR}), an interface involving the αA- and αE-helices (the AE interface) was observed in the crystal structures of both individual protein domains and of the RPS4^{TIR}:RRS1^{TIR} heterodimer (11). Dimerization of RPS4^{TIR}:RRS1^{TIR} and self-association of RPS4^{TIR} are dependent on the integrity of the AE interface, and mutations that disrupt this interface prevent both resistance signaling of the RPS4:RRS1 NLR pair and the autoactivity of RPS4^{TIR}.

The different dimerization interfaces in L6^{TIR} and RPS4^{TIR} raise the question of whether either or both of these interfaces have conserved roles in other TIR–NLRs. To address this question, we investigated the structure and function of TIR domains from several plant NLRs. We present the crystal structures of TIR domains from the *Arabidopsis* NLR proteins suppressor of npr1-1, constitutive 1 (SNC1) (25) and recognition of *Peronospora parasitica* 1 (RPP1) (26). The two structures reveal the presence of both DE- and AE-type interaction interfaces. Site-directed mutagenesis of SNC1, L6, and RPS4 reveals that both the AE- and DE-interaction surface regions can be simultaneously involved in self-association (L6 and RPS4) and are required for signaling of these TIR domains. These data imply that self-association through both the AE and DE interfaces plays a general role in TIR-domain signaling in plant immunity.

Results

The Crystal Structure of SNC1^{TIR} Reveals RPS4 and L6-Like Interfaces. We crystallized the TIR domain (residues 8–181) from the *Arabidopsis* NLR protein SNC1 (SNC1^{TIR}) (27) and determined the structure at 2.2-Å resolution (*SI Appendix*, Table S1). The fold features a four-stranded β-sheet (strands βA and βC–βE) surrounded by α-helices (αA–αE) (*SI Appendix*, Figs. S1 and S2*A*). No electron density was observed for the terminal residues (8–9 and 176–181) and residues 46–57 (corresponding to the βB-strand loop helix in L6^{TIR}; *SI Appendix*, Figs. S1 and S2*A*).

In the SNC1^{TIR} crystal, there are two prominent interfaces between molecules (Fig. 1*A*). These interfaces have striking similarities to the AE and DE interface that we previously observed in the structures of RPS4^{TIR} and L6^{TIR}, respectively (Fig. 1 *B* and *C*). Hyun et al. also observed the analogous interfaces in a recent report (24). The AE interface is formed by a symmetrical interaction involving the αA and αE helices (Fig. 1*D*) of the two molecules (hereafter designated molecules "A" and "B"), yielding a total buried surface area of ∼1,000 Å^2. The AE interface in SNC1^{TIR} contains the conserved SH (serine–histidine) motif (*SI Appendix*, Fig. S3 *A* and *B*), which is required for RPS4^{TIR} autoactivity and RPS4^{TIR}:RRS1^{TIR} dimerization and function of the full-length paired-NLR proteins (11). Side chains of the two histidines ($H30_A$ and $H30_B$, located in the center of opposing αA-helices) stack, with the neighboring serines (S29) forming hydrogen-bonding interactions with the backbone of A159 in the opposing αE-helices (Fig. 1*D*). The AE interface is further stabilized by a dense hydrogen-bonding network and electrostatic interactions between charged residues of the αA- and αE-helices that flank the SH motif, including interactions $K33_A$–$E34_B$ and $E164_B$, $H30_A$–$E158_B$, and $E158_A$–$D25_B$.

The DE interface in SNC1^{TIR} involves the αD_1- and αE-helices and the connecting loops and strands (molecules "A" and "C"). There are fewer hydrogen-bonding interactions in it, compared with the AE interface; however, several complementary hydrophobic residues are buried (Fig. 1*E* and *SI Appendix*, Fig. S2*B*). Residues within the βE-strand and αE-helices of the contacting molecules form hydrogen bonds ($K154_C$–$G149_A$ and $Y150_A$; $K154_A$–$R153_C$). $E164_A$ and $E167_A$ form salt bridges with $K154_C$ and $K112_C$, respectively. The DE interface also contains a cation–π interaction between the $W155_A$ aromatic ring and the $R153_C$ side chain (*SI Appendix*, Fig. S2*B*). The SNC1^{TIR} and L6^{TIR} DE interfaces involve similar surface regions (Fig. 1*C*), although after superimposition of one molecule in the pair, the second SNC1^{TIR} and L6^{TIR} molecules are rotated ∼21° relative to each other (Fig. 1*C*). Unlike L6^{TIR}, the αD_3-helix of SNC1^{TIR} does not contribute to the interactions with the neighboring molecule in the crystal lattice, and the interface in SNC1^{TIR} is slightly smaller than

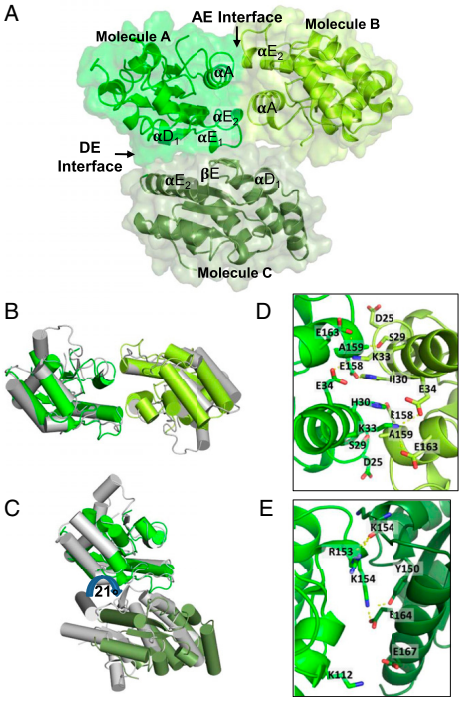


Fig. 1. The crystal structure of SNC1^{TIR} reveals two self-association interfaces. (*A*) SNC1^{TIR} crystal structure contains two major interfaces, involving predominantly αA and αE (AE) and αD and αE (DE) regions of the protein. The green and lime-colored SNC1^{TIR} molecules are observed in the asymmetric unit and interact through the AE interface; the green molecule also interacts with a crystallographic symmetry-related molecule (forest colored) through the DE interface. (*B*) Superposition of the SNC1^{TIR} (green and lime) and RPS4^{TIR} (gray) AE-interface dimers; one chain in the pair was used for superposition. (*C*) As in *B*, but showing the superposition of SNC1^{TIR} and L6^{TIR} (gray) DE-interface dimers; note the ∼21° rotation at the DE interface between the two structures. (*D*) Residues that contribute to the buried surface in the AE-interface interactions in SNC1^{TIR} are highlighted in stick representation. (*E*) As in *D*, showing the DE interface.

in L6^{TIR} (buried surface area 812 Å^2 for SNC1^{TIR} and 890 Å^2 for L6^{TIR}). Sequence analysis of plant TIR domains reveals that most of the coordinating residues involved in the DE interface in SNC1^{TIR} are not conserved (including K112, K154, and E164), with the exception of G149 (*SI Appendix*, Fig. S3 *C* and *D*), contrasting the conservation in the AE interface.

Self-Association of SNC1^{TIR} in Solution Is Disrupted by Mutations in the AE Interface. Reversible self-association in solution is observed for L6^{TIR} and RPS4^{TIR} (9, 11). We used size-exclusion chromatography (SEC) coupled to multiangle light scattering (MALS) to examine the ability of SNC1^{TIR} to self-associate in solution. By SEC-MALS, the average molecular mass of SNC1^{TIR} was higher than the theoretical molecular mass of a monomer (20.1 kDa) and increased with protein loading (Fig. 2*A*). Using the complementary technique small-angle X-ray scattering (SAXS), a similar concentration-dependent increase in average molecular

mass was observed (Fig. 2*B*). These data suggest that SNC1^{TIR} self-associates in solution in a concentration-dependent manner and is in a rapid equilibrium between monomeric and oligomeric (dimeric or higher-order) protein species.

L6^{TIR} and RPS4^{TIR} self-association in solution was shown to be dependent on the DE and AE interfaces, respectively (9, 11). To test whether these protein surfaces play a role in SNC1^{TIR} self-association, key residues involved in forming the two interfaces were mutated (to alanine or amino acid of opposite charge) and the mutant proteins tested using SEC-MALS. These residues include four in the AE interface (S29, H30, K33, and E163), and four in the DE interface (K112, Y150, K154, and E164). Recombinant proteins of all mutants, except Y150A, were successfully produced in *Escherichia coli*. With the exception of E163A, all mutants in the AE interface had average molecular masses close to the expected monomeric mass of SNC1^{TIR} (*SI Appendix*, Table S2 and Fig. S4*A*), and there was no concentration-dependent increase in the molecular mass of the H30A mutant when analyzed by SEC-MALS and SAXS (Fig. 2 *A* and *B*). Therefore, the S29A, H30A, and K33A mutations disrupt SNC1^{TIR} self-association in solution. The E163A mutant had a reduced molecular mass compared with the wild-type protein, suggesting this mutation had a weaker effect on SNC1^{TIR} self-association, probably due to its location at the periphery of the AE interface. By contrast, we could not detect an effect on self-association of mutants in the DE interface (K112A or E, K154A or E, and E164A or K) using SEC-MALS. These observations suggest that the AE interface contributes more than the DE interface to the self-association of the SNC1^{TIR} in solution.

SNC1^{TIR} Autoactivity Is Disrupted by Mutations in Either AE or DE Interfaces. To test the biological relevance of the AE and DE interfaces for SNC1^{TIR} function, we tested the effect of interface mutations on SNC1^{TIR} cell-death signaling. *Agrobacterium*-mediated transient expression of SNC1^{TIR} (residues 1–226) in *Nicotiana benthamiana* induced a visible chlorotic cell-death phenotype 5 d after infiltration. Expression of mutants in the AE interface, including S29A, H30A, and K33A, resulted in a much weaker cell-death response and a significantly reduced level of ion leakage compared with the wild-type protein (*SI Appendix*, Table S2 and Fig. 2 *C*–*F*). The E163A mutant, which showed modestly impaired self-association in solution, did not reduce the level of cell-death phenotype nor ion leakage, compared with the wild-type protein. Overall, these effects correlate well with the effects on self-association, suggesting that the integrity of the AE interface is required for both TIR domain self-association and signaling activity.

Amino acid substitutions of the DE-interface residues also affected SNC1^{TIR} autoactivity (*SI Appendix*, Table S2 and Fig. 2 *C*–*F*). The Y150A mutation, which is at the center of the SNC1^{TIR} DE interface, significantly disrupted autoactivity. Notably, L6^{TIR} has a tryptophan residue (W202) at the equivalent position (*SI Appendix*, Fig. S1) and its substitution with alanine abolished L6^{TIR} self-association in yeast and signaling activity in planta (9). Both K112E and K154E mutations in SNC1^{TIR} led to a reduced cell-death phenotype level, whereas alanine substitution of either residue did not, consistent

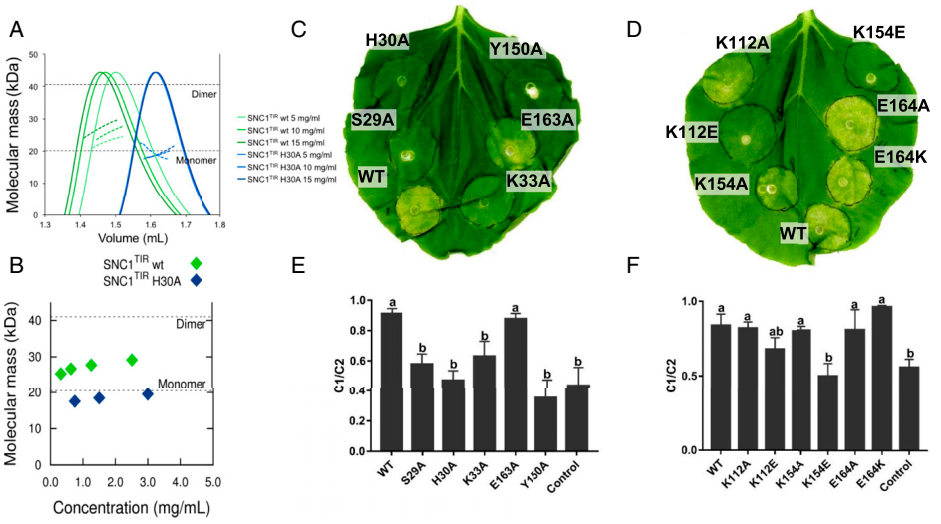


Fig. 2. SNC1^{TIR} self-association and signaling. (*A*) Solution properties of SNC1^{TIR} (WT, wild-type) and SNC1^{TIR} H30A analyzed by SEC-MALS. Green or blue peaks indicate the traces from the refractive index (RI) detector during SEC of SNC1^{TIR} or its H30A mutant, respectively. The lines under the peaks correspond to the average molecular mass distributions across the peak (using equivalent coloring). (*B*) Molecular masses calculated from SAXS data for SNC1^{TIR} (WT, wild-type; green diamonds) and SNC1^{TIR} H30A (blue diamonds), calculated from static samples at discrete concentrations between 3 and 0.25 mg/mL. Dotted lines indicate the theoretical monomeric and dimeric masses. (*C*–*F*) In planta mutational analysis of SNC1^{TIR}. (*C* and *D*) Autoactive phenotype of SNC1^{TIR} (residues 1–226; WT, wild-type) and the corresponding mutants upon *Agrobacterium*-mediated transient expression in *N. benthamiana* leaves. Each construct was coexpressed with the virus-encoded suppressor of gene silencing P19 (33). Photos were taken 5 d after infiltration. (*E* and *F*) Ion-leakage measurement of the infiltrated leaves as shown in *C* and *D*. Each construct was expressed in independent leaves. Leaf disk samples were collected 2 d after infiltration and incubated in Milli-Q water. C1 corresponds to the ions released in solution 24 h after sampling. C2 corresponds to the total ion contents in the sample (see *SI Appendix, Methods* for details). Ion leakage was calculated as C1/C2 ratio. *N. benthamiana* leaves expressing P19 only were used as control. Error bars show SE of means. Statistical differences, calculated by one-way ANOVA and multiple comparison with the control, are indicated by letters.

with electrostatic interactions through the charged side chains. By contrast, neither A nor K substitutions of the E164 residue affected cell-death development. All mutants were detected by immunoblotting and had similar protein expression levels (*SI Appendix*, Fig. S5*A*), indicating the abolition of autoactivity was not due to protein-expression differences.

Residues in the AE Interface Contribute to L6^{TIR} Self-Association and Autoactivity. We previously showed that the DE interface was involved in L6^{TIR} self-association and autoactivity (9), but the AE interface was not observed in the L6^{TIR} crystal structure. To test whether the AE interface is relevant for L6^{TIR} function, we first modeled this potential interface by superimposing the L6^{TIR} molecules onto the RPS4^{TIR} AE-interface dimer (*SI Appendix*, Fig. S6*A*). The L6^{TIR} has a phenylalanine (F79) at the position equivalent to the conserved histidine that forms the core of the AE interfaces in both RPS4^{TIR} and SNC1^{TIR} (Fig. 1*D* and *SI Appendix*, Fig. S6*A*). The equivalent residue is a phenylalanine in AtTIR, where it is involved in an AE interface in a stacking arrangement analogous to the histidine residues in SNC1^{TIR} and RPS4^{TIR} (11, 22). An aspartate residue precedes F79 in L6^{TIR}, occupying the position of the conserved serine in RPS4^{TIR}. The modeling also indicates that residues E74 and Q82 in the αA-helix, and K209 in the αE-helix could form hydrogen bonds in a potential AE-interface interaction in L6^{TIR}.

To test whether the AE interface is involved in L6^{TIR} function, we examined the effect of amino acid substitutions in this interface on its self-association. Mutations of residues F79A and K209E disrupted L6^{TIR} self-association in yeast-two-hybrid (Y2H) assays, whereas the E74K mutation did not (Fig. 3*A*). Protein expression of the BD fusion of the F79A mutant was detected at very low levels (*SI Appendix*, Fig. S5*B*), which may prevent this mutant from triggering yeast growth. However, the K209E and E74K mutant constructs were stably expressed in yeast (*SI Appendix*, Fig. S5*B*).

SEC-MALS and SEC-SAXS analysis of purified recombinant L6^{TIR} (residues 29–233) revealed an average molecular mass of 38.5 kDa, which is between the expected mass for a monomer (23.4 kDa) and dimer (46.8 kDa) (*SI Appendix*, Table S2 and Fig. S4 *B* and *C*), consistent with previous analysis (9). Substitutions of residues F79 or K209 by alanine or negatively charged residues resulted in a decreased (although slightly larger than monomer) average molecular mass (*SI Appendix*, Table S2 and Fig. S4*B*), consistent with the absence of interaction observed in yeast. The E74K mutant could not be produced recombinantly in *E. coli*. Likewise, mutations in the L6^{TIR} DE interface previously shown to affect L6^{TIR} self-association (9) also led to a decreased (although slightly larger than monomer) average molecular mass in solution (*SI Appendix*, Table S2 and Fig. S4*B*). L6^{TIR} with substitutions in both interfaces, including F79A/R164A and F79A/K200E, had an average molecular mass consistent with monomer (*SI Appendix*, Table S2 and Fig. S4*B*), suggesting that self-association in solution was fully abolished in these double interface mutants. Strikingly, the R164E mutation led to a molecular mass close to a trimer (70.2 kDa) and the F79A/R164E double mutant had an average molecular mass between those expected for dimer and trimer (*SI Appendix*, Table S2 and Fig. S4*C*). These observations suggest that the AE and DE interfaces both contribute to L6^{TIR} self-association in solution and in yeast.

We then tested the effect of the AE- and DE-interface mutations on L6^{TIR} autoactivity, using *Agrobacterium*-mediated

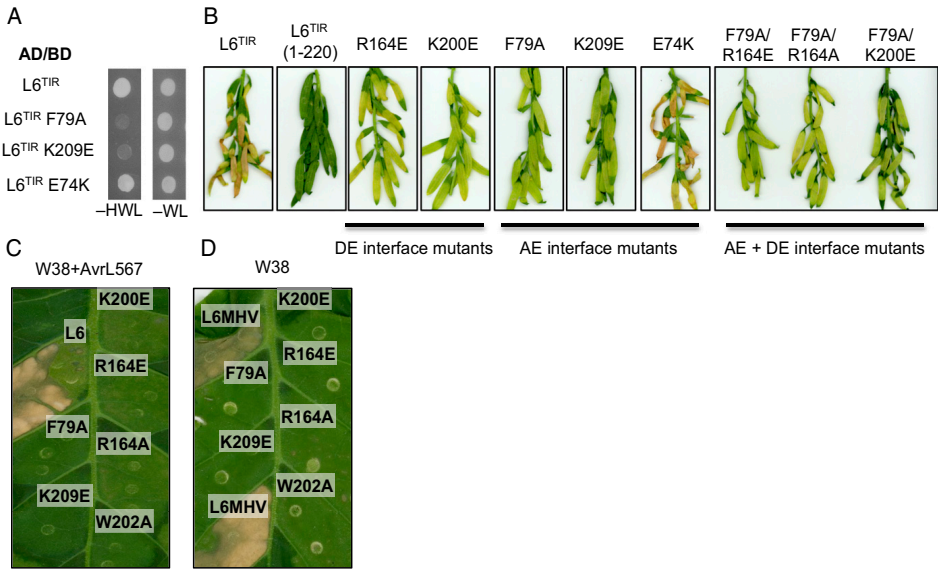


Fig. 3. Mutations in both AE and DE interfaces affect L6^{TIR} self-association and autoactivity and full-length L6 effector-dependent and effector-independent cell-death signaling. (*A*) Mutations in the AE interface disrupt L6^{TIR} self-association in yeast. Growth of yeast cells expressing GAL4-BD and GAL4-AD fusions of L6^{TIR} (residues 29–233) or L6^{TIR} mutants on nonselective media lacking tryptophan and leucine (–WL) or selective media additionally lacking histidine (–HWL). (*B*) Mutations in the AE interface disrupt L6^{TIR} signaling activity in planta. Cell-death signaling activity of L6^{TIR} (residues 1–233) mutants fused to yellow fluorescent protein (YFP), 12 d after agroinfiltration in flax plants. The truncated L6 TIR domain (residues 1–220) was used as a negative control (9). *Agrobacterium* cultures carrying L6^{TIR} mutants were adjusted to OD1. (*C*–*D*) Representative cell-death activity of L6 (*C*) and L6^{MHV} (*D*) mutants, fused to YFP, 3 d after agroinfiltration in wild-type tobacco W38 or transgenic tobacco W38 carrying AvrL567, respectively. *Agrobacterium* cultures carrying L6 and L6^{MHV} mutant were adjusted to OD 0.5.

transient expression in flax. As previously reported (9), mutations in the DE interface such as R164A/E and K200E significantly reduced the L6^{TIR} autoactive phenotype. The F79A and K209E mutations in the AE interface, which affected self-association, also suppressed L6^{TIR} autoactivity, whereas the E74K mutation (which had no effect on self-association in Y2H) did not (Fig. 3*B*). Double mutations in both interfaces, including F79A/R164A, F79A/R164E, and F79A/K200E, resulted in similar phenotypes to the single mutations. All mutants were stably expressed in flax leaves (*SI Appendix*, Fig. S5*C*). These observations suggest that both the AE and DE interfaces contribute to autoactivity of L6^{TIR}; however, neither single nor double mutations completely abolish L6^{TIR} signaling activity.

Both AE and DE Interfaces Are Required for L6 Effector-Dependent and Effector-Independent Signaling Activation. We generated AE- and DE-interface mutants in the full-length L6 protein and tested their effects on effector-dependent and effector-independent cell-death signaling. *Agrobacterium*-mediated transient expression of L6 in transgenic *Nicotiana tabacum* (tobacco) leaves, expressing the corresponding flax-rust (*Melampsora lini*) effector protein AvrL567, induces a strong cell-death response (9) (Fig. 3*C*). Mutations in both the AE (F79A and K209E) and the DE (K200E, R164E, R164A, and W202A) interfaces abolished L6 effector-dependent cell-death signaling (Fig. 3*C*). Immunoblot analysis showed that the K209E construct was not expressed in tobacco, whereas all of the other constructs expressed at a comparable level to the wild-type L6 protein (*SI Appendix*, Fig. S5*D*).

These mutations were also introduced in the autoactive variant of L6, L6^{MHV}, which contains a D-to-V mutation in the MHD motif in the ARC2 subdomain and induces a strong necrotic reaction when transiently expressed in wild-type tobacco W38 without AvrL567 (28). Mutations in both the AE and DE interfaces abolished this autoactive cell-death reaction (Fig. 3*D*), although small cell-death spots were observed with L6^{MHV} R164A mutant. Immunoblotting showed that all mutant proteins were expressed in planta (*SI Appendix*, Fig. S5*E*). Thus, mutations in either AE or DE interfaces suppress L6 effector-dependent and effector-independent cell-death signaling.

Residues in the DE Interface Contribute to RPS4^{TIR} Self-Association and Autoactivity. When overexpressed in planta, RPS4^{TIR} is autoactive and triggers an effector-independent cell-death response (7). RPS4^{TIR} self-associates and can form a heterodimer with RRS1^{TIR} through the AE interface (11). The DE interface is not observed in the crystal structures of RPS4^{TIR}, RRS1^{TIR}, or their heterodimer. To test whether the DE interface could also play a role in RPS4^{TIR} self-association, heterodimerization with RRS1, and autoactivity, we first generated a model of the DE interface in RPS4^{TIR}, by superposition of RPS4^{TIR} onto the L6^{TIR} DE-interface dimer (*SI Appendix*, Fig. S6*B*). The L6 residues R164 and K200 appear to play important roles in stabilizing the DE-interface structure and mutation of either residue suppresses L6^{TIR} self-association and autoactivity (9). Mutation of the equivalent residues in RPS4^{TIR} (R116 or M150), abolished RPS4^{TIR} self-association in Y2H assays (Fig. 4*A*), but did not affect its interaction with RRS1^{TIR} or protein accumulation in yeast (Fig. 4*B* and *SI Appendix*, Fig. S5*F*). The self-association of these RPS4^{TIR} mutants was further examined using SEC-MALS. RPS4^{TIR} (residues 10–178, equivalent to the crystal structure) had an average molecular mass of 21.1 kDa, which is only slightly higher than the expected monomeric mass of 19.6 kDa (*SI Appendix*, Table S2 and Fig. S4*D*). We previously reported a similar in-solution molecular mass of ~23 kDa for RPS4^{TIR} in a slightly different experimental setup (11). The AE-interface mutant S33A, which was previously shown to reduce the RPS4^{TIR} self-association (11), led to an average molecular mass of 19.9 kDa (*SI Appendix*, Table S2). The R116A in the DE interface also resulted in a slight average molecular mass reduction, whereas the M150R mutation was indistinguishable from the wild-type protein (*SI Appendix*, Table S2 and Fig. S4*D*). Although consistent with the DE-interface

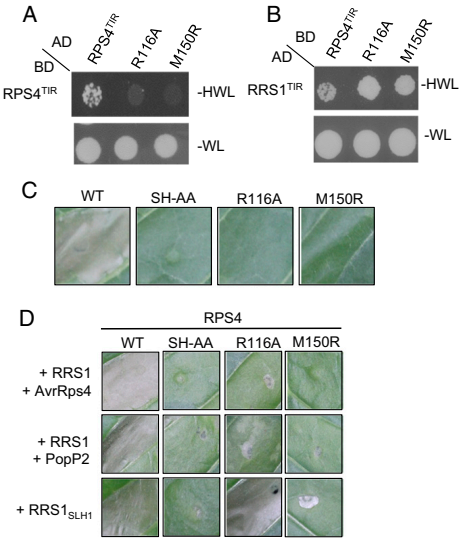


Fig. 4. Mutations in both AE and DE interfaces affect RPS4^{TIR} self-association and autoactivity and full-length RPS4 effector-dependent and effector-independent cell-death signaling. (*A*) Mutations in the DE interface disrupt RPS4^{TIR} self-association in yeast. Growth of yeast cells expressing GAL4-BD fusion and GAL4-AD fusion of RPS4^{TIR} (residues 1–183) or RPS4^{TIR} mutants on nonselective media lacking tryptophan and leucine (–WL) or selective media additionally lacking histidine (–HWL). (*B*) Mutations in the DE interface do not affect RPS4^{TIR} interaction with RRS1^{TIR}. Growth of yeast cells coexpressing GAL4-BD fusion of RPS4^{TIR} or RPS4^{TIR} mutants and GAL4-AD fusion of RRS1^{TIR} (residues 1–185) on –WL or –HWL media. (*C*) Mutations in the DE interface disrupt RPS4^{TIR} signaling activity in planta. Cell-death signaling activity of RPS4^{TIR} (WT, wild-type) and its mutants fused to C-terminal 6xHA tags, 3 d after agroinfiltration in tobacco. (*D*) Representative cell-death activity of full-length RPS4 (WT, wild-type) and its mutants fused to C-terminal 3xHA tags, upon *agro*-mediated transient coexpression with RRS1 and corresponding effectors (AvrRps4 or PopP2), or with RRS1$_{SLH1}$ mutant in W38 tobacco. *Agrobacterium* cultures were adjusted to OD 0.1. Photos were taken 5 d after agroinfiltration.

R116A mutation suppressing RPS4^{TIR} self-association in solution, the low level of self-association of wild-type RPS4^{TIR} detected in this assay and the minor differences observed for the mutants indicate that SEC-MALS may not be sufficiently sensitive to confirm this interaction. Nevertheless, the Y2H data suggest that mutations in both the DE and AE interfaces disrupt RPS4^{TIR} self-association.

We then tested the effect of the DE-interface mutations on RPS4^{TIR} autoactivity. When transiently expressed in tobacco W38, RPS4^{TIR} (residues 1–236) triggered a cell-death response, whereas mutations of the SH motif (SH-AA) in the AE interface as well as either of the R116 and M150 residues in the DE interface abolished this autoactive phenotype (Fig. 4*C*). All mutants were expressed at similar levels to the wild-type RPS4^{TIR} (*SI Appendix*, Fig. S5*G*). These observations suggest that the integrity of both AE and DE interfaces is required for the self-association and autoactivity of RPS4^{TIR}, but the AE interface is the primary interface for RPS4^{TIR} and RRS1^{TIR} heterodimerization.

Both AE and DE Interfaces Are Required for RRS1:RPS4 Effector-Dependent and Effector-Independent Activation. We further examined whether the mutations in the putative DE interface affect effector-dependent activation of the full-length RRS1:RPS4 protein pair. We previously

reported that coexpression of RPS4 and RRS1 with the effectors AvrRps4 or PopP2 in tobacco triggers a strong cell-death response that is abolished by mutations of the SH motif in the AE interface (11). Similarly, mutants in the DE interface also affected RRS1:RPS4 effector-triggered cell death, although the proteins were all expressed (Fig. 4*D* and *SI Appendix*, Fig. S5*H*). The M150R mutant triggered no cell-death response when coexpressed with RRS1 and either effector. Mutation of R116 also disrupted AvrRps4 recognition but induced a weak cell-death response upon PopP2 recognition. To measure the effect of mutations in the DE interface on RPS4 effector-independent signaling, we coexpressed RPS4 mutants with the $RRS1_{SLH1}$ variant, which contains a single amino acid (leucine) insertion in the WRKY domain and activates effector-independent cell death in the presence of RPS4 (11, 29). Mutations of the SH motif in RPS4 abolished cell-death signaling (Fig. 4*D*). The M150R mutant also disrupted RPS4 effector-independent cell death, whereas the R116A mutation did not (Fig. 4*D*). The greater effect of the M150R mutation compared with R116A on full-length RPS4 protein function may be due to its central position in the DE interface, whereas R116 is located at the periphery of the $RPS4^{TIR}$ DE interface. These observations further corroborate that, whereas both interfaces are involved in $RPS4^{TIR}$ self-association and signaling, the AE interface is the primary interface for $RPS4^{TIR}$ and $RRS1^{TIR}$ heterodimerization.

The Crystal Structure of the $RPP1^{TIR}$ Features AE and DE Interfaces. We recently showed that alleles of the *Arabidopsis* NLR protein RPP1 from ecotypes Niederzenz (NdA) and Wassilewskija (WsB) differ in their ability to induce effector-independent cell death via transient expression of the TIR domain in planta (15). RPP1 NdA-1 and WsB alleles differ by 17 substitutions in the TIR domain. Biophysical and functional analyses of proteins where these residues are mutated show a correlation between self-association and the ability for RPP1 TIR domains to induce effector-independent cell death (15). In light of these findings, we undertook structural studies of the RPP1 NdA-1 TIR domain (residues 93–254; $RPP1^{TIR}$). Strikingly, the crystal structure (2.8-Å resolution; *SI Appendix*, Table S1) reveals AE and DE interfaces analogous to $SNC1^{TIR}$ (Fig. 5 *A*–*C* and *SI Appendix*, Fig. S7). Upon superposition of one molecule in the DE-interface dimers, the αE-helix is rotated ~97° in $RPP1^{TIR}$ compared with its position in $L6^{TIR}$ in the other molecule (Fig. 5*C*). Despite this difference to other TIR domain structures, there are common residues within the AE and DE interfaces of the RPP1 crystal structure (*SI Appendix*, Fig. S7).

Discussion

Structural Conservation of TIR Domain Interfaces in Plants. TIR domains feature in innate immunity pathways across phyla (21); however, the molecular mechanisms of signaling by these domains have largely remained elusive. Whereas in mammalian TIR domains no common trends have emerged among the available crystals structures in terms of protein–protein association (21), most plant TIR-domain crystal structures feature structurally analogous AE interfaces (Figs. 1*B* and 5*B*) (23). The exception is $L6^{TIR}$, the crystal structure of which features the DE but not the AE interface. DE interfaces are also observed in the two structures reported here, of $SNC1^{TIR}$ and $RPP1^{TIR}$, and in the structure of AtTIR (*SI Appendix*, Fig. S8), albeit with some deviations in orientation (Figs. 1*C* and 5*C* and *SI Appendix*, Fig. S9). Whereas the presence or absence of an interface in the crystal does not prove or disprove a biological function (30), these observations precipitated a thorough assessment of interfaces in several proteins, as described here. We conclude that self-association through both the AE and DE interfaces plays a general role in TIR–NLR signaling.

TIR-Domain Self-Association and Signaling Through Conserved TIR Domain Interfaces. We have previously shown that both $L6^{TIR}$ and $RPS4^{TIR}$ signaling requires self-association, focusing on the single dimerization interfaces (DE and AE, respectively) observed in these crystal structures. Here we show that mutations in either of the AE or DE interfaces suppress self-association, effector-dependent immunity, and effector-independent autoactivity in both $L6^{TIR}$ and $RPS4^{TIR}$. In Y2H assays, single mutations to residues in distinct (either the AE or DE) interfaces abolish self-association, suggesting that in this assay, interaction through both interfaces is required for detection. Similarly, mutations of residues in either the AE or DE interfaces can independently abolish cell death. Collectively, these data suggest that self-association through AE and DE interfaces is required simultaneously to allow cell-death signaling by L6 and RPS4.

For $SNC1^{TIR}$, both AE and DE interfaces are observed within the crystal structure. Although only mutations in the AE interface were observed to significantly affect $SNC1^{TIR}$ self-association in solution, mutations in either interface suppressed $SNC1^{TIR}$ cell-death signaling, indicating that the intergrity of both interfaces is required for function. In the case of $RPP1^{TIR}$, several previously identified mutations that affect self-association and cell-death signaling (15) map to either the AE or DE interfaces in the $RPP1^{TIR}$ structure. Therefore, the data presented here and previously (9, 11, 15) suggest a correlation between TIR-domain self-association and cell-death signaling. One exception to this correlation is $RPV1^{TIR}$ from *Mirabilis rotundifolia*; no self-association of this protein could be detected in solution or yeast, but nevertheless, mutations within the predicted AE- and DE-interface regions suppressed its cell-death signaling function (23). All of the TIR:TIR domain interactions studied to date are weak and transient, with the exception of the heterodimer association between $RRS1^{TIR}$ and $RPS4^{TIR}$, which appears to play an inhibitory rather than signaling role. Thus, the failure to detect $RPV1^{TIR}$ self-interaction may be due to the weak self-association of this TIR domain, below the detection threshold limit. The weak and transient nature of TIR:TIR domain interactions may be a key regulatory mechanism to reduce the occurrence of cell-death signaling in the absence of an appropriate stimulus. It is also likely that the TIR-domain self-association is stabilized by other domains in the NLR, by other proteins that promote cell-death

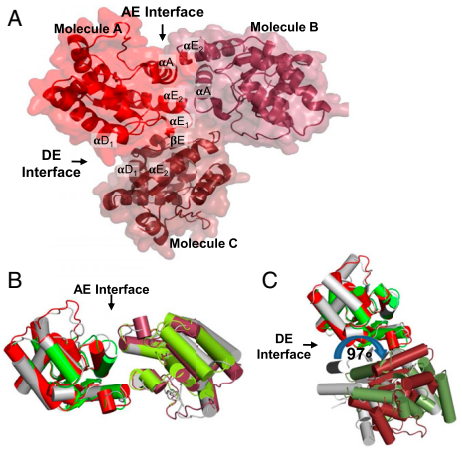


Fig. 5. The AE and DE interface in the crystal structure of $RPP1^{TIR}$. (*A*) Ribbon representation of the RPP1 crystal structure and the AE and DE interfaces, with molecules sharing the AE interface colored red and raspberry and the DE interface, red and ruby. (*B*) Comparison of the AE interface from the $RPS4^{TIR}$ (gray), $SNC1^{TIR}$ (green and lime), and $RPP1^{TIR}$ (red and raspberry) with the chains on the *Left* superimposed, highlighting the strong structural conservation of the interface. (*C*) Comparison of the DE interface from the $L6^{TIR}$ (gray), $SNC1^{TIR}$ (green and forest), and $RPP1^{TIR}$ (red and ruby) structures; only the chains at the *Top* are superimposed, highlighting the rotation observed at the DE interface in these crystal structures.

signaling in planta, or that the TIR domains interact with nonself TIR domains in planta to propagate signaling. Regardless of the exact mechanisms, the required integrity of the AE and DE interfaces suggests that TIR:TIR domain association through both these interfaces is a general requirement for function.

Cooperative Assembly of TIR Domains. Higher-order assembly formation has become an emerging theme in diverse innate immunity pathways. Protein domains from the death-domain family appear to be able to form large, often open-ended helical structures (31). Signaling by cooperative assembly formation (SCAF) explains the ultrasensitive all-or-none response desirable in such pathways (2). One notable feature of the AE and DE interface is that they are not mutually exclusive. In fact, based on the $SNC1^{TIR}$ domain structure, it is possible to build a hypothetical extended TIR domain superhelix propagated through the AE and DE interfaces that are observed in the crystal structure (*SI Appendix*, Fig. S10*A*). An AE interface is also conserved and functional in $L6^{TIR}$; therefore, it is possible for $L6^{TIR}$ to oligomerize through the conserved AE interface and the DE interface observed in the $L6^{TIR}$ crystals. Such an assembly results in a superhelix similar to the one proposed for $SNC1^{TIR}$ (*SI Appendix*, Fig. S10*B*). The same is not possible for $RPP1^{TIR}$, as rotation around the DE interface causes a clash when constructing the hypothetical superhelix (*SI Appendix*, Fig. S10*C*). As such, the varying rotations of different TIR domains suggest there is flexibility in this region, and that a specific geometry is required for TIR domains to form larger oligomeric structures. To date, the only evidence of a TIR-domain self-association complex greater than a dimer was observed in the $L6^{TIR}$ carrying an R164E mutation. The R164E actually suppresses signaling, but the mutant self-associates more strongly in solution compared with wild-type $L6^{TIR}$. Interstingly, the R164A mutation reduces both signaling and self-association, suggesting that the charge substitution mutation at this site may favor an altered specific geometry of association, leading to the formation of an inactive oligomer.

Signaling by Plant TIR Domains in a TIR–NLR Receptor. Ultimately, we need to consider the AE- and DE-interface interactions in the context of a full-length TIR–NLR receptor. Although to date there is no structural data for a full-length receptor, analysis of their mammalian NLR counterparts demonstrates that the nucleotide-binding domain plays a key role in the self-association of the NLRC4 receptor into 10–12 subunit oligomers (18, 19). If plant NLRs follow a nucleated NB-mediated assembly mechanism as observed in animal NLRs (2, 18, 19), this could elegantly induce a proximity-induced assembly of the associated TIR domains through the AE and DE interfaces. Data are starting to emerge on NLR oligomerization upon effector recognition (specifically in tobacco N protein (32) and *Arabidopsis* RPP1 (15)). Many plant TIR domains do not show autoactivity outside the context of the full-length protein (7, 15), suggesting they may not be able to interact adequately on their own without the help from other domains in the NLR.

Materials and Methods

Details of the methods used are provided in *SI Appendix*, *Methods*, including cloning details for vectors and gene constructs, crystallization and structure determination using X-ray crystallography, biophysical experiments including SEC-MALS and SAXS experiments, transient expression in planta, Y2H assays, immunoblot analysis, and ion-leakage measurements.

ACKNOWLEDGMENTS. We thank Stella Césari for discussion and assistance with the in planta immunoprecipitation experiments, Kim Newell for excellent technical support, and Daniel Ericsson for help with X-ray data processing. We acknowledge the use of the University of Queensland Remote Operation Crystallization and X-Ray Diffraction Facility. This research used macromolecular crystallography (MX) and small/wide angle X-ray scattering (SAXS/WAXS) beamlines at the Australian Synchrotron, Victoria, Australia. This research was supported by the Australian Research Council (ARC) Discovery Projects (DP120100685, DP120103558, and DP160102244) and the National Science Foundation (NSF-IOS-1146793 to B.J.S.). B.K. is a National Health and Medical Research Council Research Fellow (1003325 and 1110971). M.B. and S.J.W. are recipients of ARC Discovery Early Career Research Awards (DE130101292 and DE160100893, respectively).

1. Dodds PN, Rathjen JP (2010) Plant immunity: Towards an integrated view of plant-pathogen interactions. *Nat Rev Genet* 11(8):539–548.
2. Bentham A, Burdett H, Anderson PA, Williams SJ, Kobe B (August 24, 2016) Animal NLRs provide structural insights into plant NLR function. *Ann Bot*, 10.1093/aob/mcw171.
3. Duxbury Z, et al. (2016) Pathogen perception by NLRs in plants and animals: Parallel worlds. *BioEssays* 38(8):769–781.
4. van der Biezen EA, Jones JD (1998) The NB-ARC domain: A novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr Biol* 8(7):R226–R227.
5. McHale L, Tan X, Koehl P, Michelmore RW (2006) Plant NBS-LRR proteins: Adaptable guards. *Genome Biol* 7(4):212.
6. Frost D, et al. (2004) Tobacco transgenic for the flax rust resistance gene L expresses allele-specific activation of defense responses. *Mol Plant Microbe Interact* 17(2):224–232.
7. Swiderski MR, Birker D, Jones JD (2009) The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. *Mol Plant Microbe Interact* 22(2):157–165.
8. Krasileva KV, Dahlbeck D, Staskawicz BJ (2010) Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell* 22(7):2444–2458.
9. Bernoux M, et al. (2011) Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host Microbe* 9(3):200–211.
10. Maekawa T, et al. (2011) Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host Microbe* 9(3):187–199.
11. Williams SJ, et al. (2014) Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* 344(6181):299–303.
12. Wang GF, et al. (2015) Molecular and functional analyses of a maize autoactive NB-LRR protein identify precise structural requirements for activity. *PLoS Pathog* 11(2):e1004674.
13. Cesari S, et al. (2016) Cytosolic activation of cell death and stem rust resistance by cereal MLA-family CC-NLR proteins. *Proc Natl Acad Sci USA* 113(36):10204–10209.
14. Kanzaki H, et al. (2012) Arms race co-evolution of Magnaporthe oryzae AVR-Pik and rice Pik genes driven by their physical interactions. *Plant J* 72(6):894–907.
15. Schreiber KJ, Bentham A, Williams SJ, Kobe B, Staskawicz BJ (2016) Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. *PLoS Pathog* 12(7):e1005769.
16. Casey LW, et al. (2016) The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins. *Proc Natl Acad Sci USA* 113(45):12856–12861.
17. Collier SM, Hamel LP, Moffett P (2011) Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Mol Plant Microbe Interact* 24(8):918–931.
18. Hu Z, et al. (2015) Structural and biochemical basis for induced self-propagation of NLRC4. *Science* 350(6259):399–404.
19. Zhang L, et al. (2015) Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization. *Science* 350(6259):404–409.
20. Reubold TF, Wohlgemuth S, Eschenburg S (2011) Crystal structure of full-length Apaf-1: How the death signal is relayed in the mitochondrial pathway of apoptosis. *Structure* 19(8):1074–1083.
21. Ve T, Williams SJ, Kobe B (2015) Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis* 20(2):250–261.
22. Chan SL, Mukasa T, Santelli E, Low LY, Pascual J (2010) The crystal structure of a TIR domain from Arabidopsis thaliana reveals a conserved helical region unique to plants. *Protein Sci* 19(1):155–161.
23. Williams SJ, et al. (2016) Structure and function of the TIR domain from the grape NLR protein RPV1. *Front Plant Sci* 7:1850.
24. Hyun KG, Lee Y, Yoon J, Yi H, Song JJ (2016) Crystal structure of Arabidopsis thaliana SNC1 TIR domain. *Biochem Biophys Res Commun* 481(1–2):146–152.
25. Zhang Y, Goritschnig S, Dong X, Li X (2003) A gain-of-function mutation in a plant disease resistance gene leads to constitutive activation of downstream signal transduction pathways in suppressor of npr1-1, constitutive 1. *Plant Cell* 15(11):2636–2646.
26. Botella MA, et al. (1998) Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants. *Plant Cell* 10(11):1847–1860.
27. Wan L, et al. (2013) Crystallization and preliminary X-ray diffraction analyses of the TIR domains of three TIR-NB-LRR proteins that are involved in disease resistance in Arabidopsis thaliana. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69(Pt 11):1275–1280.
28. Bernoux M, et al. (2016) Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. *Plant Cell* 28(1):146–159.
29. Noutoshi Y, et al. (2005) A single amino acid insertion in the WRKY domain of the Arabidopsis TIR-NBS-LRR-WRKY-type disease resistance protein SLH1 (sensitive to low humidity 1) causes activation of defense responses and hypersensitive cell death. *Plant J* 43(6):873–888.
30. Kobe B, et al. (2008) Crystallography and protein-protein interactions: Biological interfaces and crystal contacts. *Biochem Soc Trans* 36(Pt 6):1438–1441.
31. Wu H (2013) Higher-order assemblies in a new paradigm of signal transduction. *Cell* 153(2):287–292.
32. Mestre P, Baulcombe DC (2006) Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *Plant Cell* 18(2):491–501.
33. Sainsbury F, Thuenemann EC, Lomonossoff GP (2009) pEAQ: Versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol J* 7(7):682–693.

SI Appendix

SI Methods

Vectors and gene constructs. For recombinant protein production in *E. coli*, the cDNAs were cloned into the pMCSG7 vector using ligation-independent cloning (LIC) (1, 2). The resulting constructs contained an N-terminal hexahistidine (His)-tag followed by a TEV (tobacco etch virus) protease cleavage site. The TEV protease cleavage site enables His-tag removal during purification. The plasmids used in yeast-two-hybrid experiments were constructed by Gateway cloning (GWY; Invitrogen) of specific cDNAs into Gateway-compatible yeast-two-hybrid vectors based on pGADT7 and pGBKT7 (Clontech), as previously described (3). For transient expression *in planta*, SNC1 and L6-derived constructs were cloned in the Gateway binary vector pAM-PAT-35 s-GWY-YFPv, as previously described (4). Mutations were introduced by Dpn1-mediated site-directed mutagenesis (Stratagene) following the manufacturer's instructions.

Genomic fragments of full-length RPS4 (4 fragments) and RRS1 (7 fragments) were PCR-amplified from *Arabidopsis* genomic DNA (ecotypes No-0 and Ws-2, respectively). $RPS4^{TIR}$ (residues 1-236) and $RRS1^{TIR}$ (residues 1-175) were PCR-amplified from *Arabidopsis* genomic DNA (ecotypes Col-0 and Ws-2 respectively). PopP2 (residues 149-488) and full-length AvrRps4 were PCR-amplified from previously reported plasmid constructs (5-7). The cloning primers contained a BsaI recognition sequence with specific 4-bp overhangs, and the fragments were cloned into the pICH41021 vector (modified pUC19 in which the BsaI recognition sequence was mutated). 35S binary constructs were generated by Golden Gate assembly of the resulting pICH41021 constructs into pICH86988 (provided by Sylvestre Marillonnet) or EpiGreenB5-GG for AvrRps4 (8). The genes were fused with a C-terminal epitope tag: C-6xHA ($RPS4^{TIR}$); C-3xFlag ($RPS4^{FL}$, $RRS1^{FL}$); C-YFP ($RPS4^{TIR}$, $RRS1^{TIR}$, AvrRps4, PopP2). Accession numbers for the protein sequences used in this study: SNC1 (O23530.1), RPS4 (Q9XGM3.1), L6 (AAA91022), AtTIR (Q9SSN3.1), RPP1 (ADI80539). Primers and construct details are list in Table S4 and S5.

Structure determination of $SNC1^{TIR}$ and $RPP1^{TIR}$. $SNC1^{TIR}$ (residues 8-181) was expressed, purified and crystallized as described previously (9). $RPP1^{TIR}$ (residues 93 – 254) was expressed and purified as described previously (10) and crystallized at a concentration of 10 mg/mL in 15% PEG6000, 0.2 M citrate pH 5.5. X-ray diffraction data was collected on the MX2 beamline at the Australian Synchrotron using the Blu-Ice software (11). The datasets were processed using XDS (12) and scaled using Aimless (13) in the CCP4 suite (14). Molecular replacement (MR) was used to determine the protein structures. For $SNC1^{TIR}$, the $L6^{TIR}$ structure (PDB ID 3OZI) was used as the search model and MR was performed using Phaser (15) and then rebuilt using Autobuild (16), both in the Phenix suite (17). The structure of $RPP1^{TIR}$ was solved using the BALBES pipeline (18) with multiple TIR-domain structure templates (PDB IDs 3OZI, 3JRN, 3H16A, 3SOE, 1FYV). The models were refined using Phenix.Refine (19). $RPP1^{TIR}$ crystal structure contains 8 molecules in the asymmetric unit with pairwise RMSDs of 0.5-1.2 Å. Disulfide-bonded cysteines were observed in the DE-interfaces of some chains. Structure validation was performed using MolProbity (20) and POLYGON within the Phenix suite. Crystallographic parameters are shown in Table S1. Structure analyses and model building were performed using iMDFF

(21), Coot (22), PyMOL (http://www.pymol.org/; DeLano Scientific LLC), Consurf (23), Dali (24) and PISAePDB (25).

Multi-angle light scattering (MALS). The solution properties of proteins were characterized using MALS coupled with size-exclusion chromatography (SEC-MALS). The protein samples were loaded onto a Superdex 200 10/300 size-exclusion column (GE Healthcare), connected with a Dawn Heleos II 18-angle light-scattering detector and an Optilab rEX refractive index detector (Wyatt Technology, Santa Barbara, CA, USA). The experiments were performed at room temperature in the equilibration buffer (10 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM DTT) and a sample volume of 30 µL. The molecular masses of the proteins were calculated using the Astra 5.3 software (26).

Small-angle x-ray scattering (SAXS). SAXS data on wild-type $SNC1^{TIR}$ and the $SNC1^{TIR}$ H30A mutant were collected in static format at the SAXS/WAXS beamline of the Australian Synchrotron. Prior to data collection, purified proteins were thawed and gel-filtered in a buffer containing 10 mM HEPES (pH 7.5), 150 mM NaCl and 1 mM DTT, at 4 °C. Concentrations were obtained by UV absorbance at 280 nm. A four-fold dilution series was prepared using the post-peak gel-filtration buffer, starting at 5 mg/mL for wild-type $SNC1^{TIR}$ and 3 mg/mL for the H30A mutant. At each dilution, 90 µL of sample was injected through a 1.5-mm-diameter quartz capillary at 298 K, at a rate of 1 µL/s, capturing images every 1 s.

The data on wild-type and mutant $L6^{TIR}$ were collected in inline SEC-SAXS format, using a Prominence modular HPLC system (Shimadzu Scientific Instruments) mounted upstream of the capillary, using a WTC-030S5 analytical column and a WTC-030S5G pre-column (Wyatt Technology), in 100 mM NaCl, 10 mM HEPES, pH 7.5, 1 mM DTT at 16 °C. SAXS images were collected constantly during the course of each elution, with 0.05 s intervals between each frame.

A Pilatus 1M detector was used for both the $SNC1^{TIR}$ and $L6^{TIR}$ measurements, yielding a range of momentum transfer $0.011 < q < 0.500$ Å for $SNC1^{TIR}$ and $9.44 \times 10^{-3} < q < 0.477$ for $L6^{TIR}$, where $q = 4\pi.\sin(\theta)/\lambda$. Data reduction and subtraction was performed using scatterBrain (http://www.synchrotron.org.au/index.php/aussyncbeamlines/saxswaxs/software-saxswaxs). Consistent, successive exposures were normalized to transmitted intensity, reduced, scaled to absolute intensity using pure water, averaged and buffer-subtracted. In the static format, buffer measurements were taken before and after the concentration series and compared for evidence of radiation damage. In the SEC-SAXS format, 50 images immediately preceding each peak were designated as buffer. This buffer blank was then subtracted from each individual image to produce a set of subtracted data across the elution. To generate the summed and averaged datasets used in Table S3 in an unbiased manner, sets of contiguous images were taken in which I(0) was at least half the value at the peak maximum.

The ATSAS 2.6 software package was used for subsequent analyses (27). Guinier analysis was performed for $q.Rg < 1.3$ using AUTORG in PRIMUS (28), and data-sets were examined for concentration dependence and linearity. P(r) distributions were obtained for all constructs by indirect transformation in GNOM (29), informed by AUTOGNOM. Molecular masses were estimated from the P(r) distributions using SAXSMoW (30). This process was automated for the SEC-SAXS format due to the large number of individual images, and

molecular mass was instead calculated by the volume of correlation (MM_{Vc}) method (31), which is more amenable to high throughput analysis, for ranges where $q < 0.3$ Å^{-1}.

Transient expression *in planta* and yeast two-hybrid assays. *Agrobacterium tumefaciens* cells were grown overnight at 28 °C in LB media with appropriate antibiotic selections. The cells were pelleted and prepared at a desired optical density in the infiltration mix (10 mM $MgCl_2$, 200 µM acetosyringone), followed by incubation at room temperature for 2 h. Resuspended cells were infiltrated with a 1 mL needleless syringae into leaves of 3-week-old wild-type tobacco (W38), transgenic tobacco expressing the flax rust effector AvrL567 (3) or *N. benthamiana* plants and 4-week-old flax (Hoshangabad). Yeast transformation, using the HF7c yeast strain, and growth assays were performed as described in the Yeast Protocols Handbook (Clontech). Yeast transformants were grown on minimal media lacking tryptophan and leucine (-WL) to select for the presence of both pGADT7 and pGBKT7 plamids, and on minimal media additionally lacking histidine (-HWL) to test for the activation of the reporter gene.

Immunoblot analysis. Yeast protein extraction for immunoblot analysis was performed following a post-alkaline extraction method (32). Plant proteins were extracted by grinding two flax leaves, or two *N. tabacum* or *N. benthamiana* leaf discs (9 mm diameter) in loading buffer collected 72 h, 48 h or 24 h after agroinfiltration, respectively, or as indicated in figure legends.

Proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes (Pall). Membranes were blocked in 5% skim milk and probed with anti-HA-HRP conjugated antibodies (clone F7, Santa Cruz) or anti-GFP mouse monoclonal antibodies (clones 7.1 and 13.1, Roche) followed by goat anti-mouse HRP conjugate (BioRad) for plant protein samples. Yeast protein samples were probed with anti-HA-HRP conjugate rat monoclonal antibodies (clone 3F10, Roche) or with and anti-Myc mouse monoclonal antibodies (clone 9E10, Roche) followed by goat anti-mouse HRP conjugate (BioRad). Protein labelling was detected with the SuperSignal West Femto chemiluminescence kit (Pierce). Membranes were stained with Red Ponceau to visualize protein loading.

Ion-leakage measurements. For each construct tested, three independent leaves were infiltrated and four discs (0.9 mm diameter) were collected per leaf 2 days after infiltration, quickly rinsed with water and placed in 2 mL Eppendorf tubes with 1.8 mL of Milli-Q water, and gently shaken for 24 h before measuring released ion content (C1) using a Horiba B-173 conductivity meter. The samples were then boiled for 10 min and total ion content was measured (C2). Ion leakage was calculated as C1/C2 ratio. *N. benthamiana* leaves expressing P19 only were used as control. Statistic differences were calculated by one-way-ANOVA.

SUPPLEMENTAL REFERENCES

1. Eschenfeldt WH, Lucy S, Millard CS, Joachimiak A, Mark ID (2009) A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol Biol* 498:105-115.
2. Stols L, et al. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* 25:8-15.
3. Bernoux M, et al. (2011) Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host Microbe* 9:200-211.
4. Bernoux M, et al. (2008) RD19, an Arabidopsis cysteine protease required for RRS1-R-mediated resistance, is relocalized to the nucleus by the Ralstonia solanacearum PopP2 effector. *Plant Cell* 20:2252-2264.
5. Sohn KH, Zhang Y, Jones JD (2009) The Pseudomonas syringae effector protein, AvrRPS4, requires in planta processing and the KRVY domain to function. *Plant J* 57:1079-1091.
6. Sohn KH, Hughes RK, Piquerez SJ, Jones JD, Banfield MJ (2012) Distinct regions of the Pseudomonas syringae coiled-coil effector AvrRps4 are required for activation of immunity. *Proc Natl Acad Sci U S A* 109:16371-16376.
7. Sohn KH, et al. (2014) The nuclear immune receptor RPS4 is required for RRS1(SLH1)-dependent constitutive defense activation in Arabidopsis thaliana. *Plos Genet* 10:e1004655.
8. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One* 3:e3647.
9. Wan L, et al. (2013) Crystallization and preliminary X-ray diffraction analyses of the TIR domains of three TIR-NB-LRR proteins that are involved in disease resistance in Arabidopsis thaliana. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69:1275-1280.
10. Schreiber KJ, Bentham A, Williams SJ, Kobe B, Staskawicz BJ (2016) Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. *PLoS Pathog* 12:e1005769.
11. McPhillips TM, et al. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J Synchrotron Radiat* 9:401-406.
12. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125-132.
13. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* 69:1204-1214.
14. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235-242.
15. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.
16. Terwilliger TC, et al. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* 64:61-69.
17. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-221.

18. Long F, Vagin AA, Young P, Murshudov GN (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* 64:125-132.
19. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68:352-367.
20. Davis IW, Murray LW, Richardson JS, Richardson DC (2004) MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* 32:W615-619.
21. Croll TI, Andersen GR (2016) Re-evaluation of low-resolution crystal structures via interactive molecular-dynamics flexible fitting (iMDFF): a case study in complement C4. *Acta Crystallographica Section D* 72:1006-1016.
22. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.
23. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38:W529-533.
24. Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38:W545-549.
25. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774-797.
26. Wen J, Arakawa T, Philo JS (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal Biochem* 240:155-166.
27. Konarev PV, Petoukhov MV, Volkov VV, Svergun DI (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J Appl Cryst* 39:277-286.
28. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS : a Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst* 36:1277-1282.
29. Semenyuk AV, Svergun DI (1991) GNOM – a program package for small-angle scattering data processing. *J Appl Cryst* 24:537-540.
30. Fischer H, de Oliveira Neto M, Napolitano HB, Polikarpov I, Craievich AF (2009) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J Appl Cryst* 43:101-109.
31. Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496:477-481.
32. Kushnirov VV (2000) Rapid and reliable protein extraction from yeast. *Yeast* 16:857-860.
33. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188-1190.
34. Margelevicius M, Venclovas C (2005) PSI-BLAST-ISS: an intermediate sequence search tool for estimation of the position-specific alignment reliability. *BMC Bioinformatics* 6:185.
35. Williams SJ, et al. (2014) Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* 344:299-303.

Tables

Table S1. Crystallographic data.

	SNC1^{TIR}	RPP1^{TIR}
Data collection		
Detector	ADSC Quantum 315r CCD	ADSC Quantum 315r CCD
Wavelength (Å)	0.9537	0.9537
Temperature (K)	100	100
Crystal-to-detector distance (mm)	310	270
Rotation range per image (°)	0.5	0.5
Exposure time per image (s)	1	1
Space group	P 4_32_12	P 1 2_11
a, *b*, *c* (Å)	82.18, 82.18, 124.1	81.89, 84.33, 122.75
α, β, γ (°)	90, 90, 90	90, 90.1, 90
Average mosaicity (°) [b]	0.15	0.31
Resolution range (Å)	19.68-2.20 (2.27-2.20) [a]	42.32-2.58 (2.66-2.58) [a]
Total no. of reflections	372024 (32369)	394094 (29751)
No. of unique reflections	22223 (1888)	51965 (4314)
Completeness (%)	99.8 (99.6)	99.6 (95.6)
Multiplicity	16.7 (17.1)	7.6 (6.9)
Mean $I/\sigma(I)$	17.8 (2.1)	14.5 (1.3)
*R*meas (%) [c]	18.0 (184.9)	10.7 (184.7)
*R*pim (%) [d]	4.3 (44.1)	5.4 (96.5)
$CC_{1/2}$ [b]	0.99 (0.69)	0.99 (0.59)
Matthews coefficient (Å^3 Da^{-1}) [e]	2.63	2.87
Refinement		
Resolution range (Å)	19.68-2.20	40.94-2.80
R_{work} (%) [f]	18.34	21.7
R_{free} (%) [g]	21.66	26.8
No. of non-H atoms		
Total	2642	10533
Non-solvent	2436	10533
Water	206	0
Average isotropic *B* value (Å^2)	45.0	79.0
R.m.s.d. from ideal geometry		
Bond lengths (Å)	0.014	0.0059
Bond angles (°)	1.29	0.989
Ramachadran plot, residues in (%)[h]		
Favoured regions	97.6	96.1
Additionally allowed regions	2.4	3.8
Outlier regions	0	0.1

[a] The values in parentheses are for the highest-resolution shell.
[b] Calculated with AIMLESS (13).
[c] *R*meas = $\Sigma hkl\{N(hkl)/[N(hkl)-1]\}1/2\ \Sigma i|Ii(hkl)-<I(hkl)>|/\ \Sigma hkl\Sigma iIi(hkl)$, where $Ii(hkl)$ is the intensity of the *i*th measurement of an equivalent reflection with indices *hkl*.
[d] *R*pim = $\Sigma hkl\{1/[N(hkl)-1]\}1/2\ \Sigma i|Ii(hkl)-<I(hkl)>|/\ \Sigma hkl\Sigma iIi(hkl)$.
[e] Calculated with MATTHEWS_COEF within the CCP4 suite (14).

[f] $R_{work} = \Sigma_{hkl} ||F_{obs}|-|F_{calc}|| / \Sigma_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.
[g] R_{free} is equivalent to R_{work} but calculated with reflections (5%) omitted from the refinement process.
[h] Calculated with MolProbity (20).

Table S2. Mutational analysis of SNC1^{TIR}, L6^{TIR} and RPS4^{TIR}

	Mutation position	Location in the crystal structure	Interface	TIR domain			Full-length	
				MW (kDa)	Self-association (Y2H)	Autoactivity	Effector-dependent cell-death activity	Effector-independent cell-death activity
SNC1	WT			25.6		++		
	S29A	αA helix	AE	19.6		-		
	H30A	αA helix	AE	19.7		-		
	K33A	αA helix	AE	19.6		-		
	E163A	αE helix	AE	21.3		-/+		
	K112A	αD helix	DE	25.7		+		
	K112E	αD helix	DE	25.1		-		
	Y150A	αE helix	DE	N/A[a]		-		
	K154A	αE helix	DE	24.1		+		
	K154E	αE helix	DE	24		-		
	E164A	αE helix	DE	25.2		++		
	E164K	αE helix	DE	25.2		++		
L6	WT			38.5	+	++	+++	+++
	F79A	αA helix	AE	28.2	-	+/-	-	-
	K209E	αE helix	AE	27.3	-	+/-	-	-
	R164A	αD helix	DE	28.6	-[b]	+/-	+	+
	R164E	αD helix	DE	74.4	-[b]	+/-	+/-	+/-
	K200E	DE loop	DE	27.1	-[b]	+/-	+/-	-
	F79A+R164A		Both	24.6		+/-		
	F79A+R164E		Both	58.9		+/-		
	F79A+K200E		Both	24.4		+/-		
RPS4	WT			21.1	+	+	+	+
	S33A	αA helix	AE	19.9	-[c]			
	S33A+H34A	αA helix	AE	N/A	-[c]	-	-	-
	M150R	DE loop	DE	21.2	-	-	-	-
	R116A	αD helix	DE	20.5	-	-	-[d]	+

List of mutations introduced in the TIR domains. The "MW" column shows the average solution masses measured using SEC-MALS. The theoretical masses for SNC1^{TIR}, L6^{TIR} and RPS4^{TIR} monomers are 20.1 kDa, 23.4 kDa and 19.6 kDa, respectively. The "autoactivity" and "effector-dependent" or "effector-independent cell-death activity" columns show the cell-death symptoms as compared with the wild-type (WT) protein. "+", comparable to WT; "-", weaker than WT.

[a] Not tested by SEC-MALS.

[b] (3)

[c] (35)

[d] +/- when co-expressed with PopP2

Table S3. Properties calculated from averaged $L6^{TIR}$ SAXS datasets.

Protein	$I(0)_{Guin}$ ($x10^{-2}$ cm^{-1})	$I(0)_{P(r)}$ ($x10^{-2}$ cm^{-1})	$R_{g\ Guin}$ (Å)	$R_{g\ P(r)}$ (Å)	MM_{Vc} (kDa)
Wild-type	1.43	1.45	25.12	26.16	29.1
F79A	2.27	2.29	23.92	25.06	26.4
R164A	1.94	1.95	24.23	24.04	28.1
K200E	2.49	2.51	24.78	25.89	27.5
F79A/D208A	2.156	2.17	22.85	23.82	23.8
F79A/R164A	2.11	2.12	22.40	23.13	23.6

Table S4: Constructs used in this study

Use	Construct name	Insert or PCR product	Primers	Template	Plasmid backbone	Cloning method	Reference
Crystal structure and SEC-MALS	SNC1 TIR	SNC1 8-181 cDNA without stop	XZ38/XZ39	synthesized DNA code optimized for E. coli	pMCSG7	Ligation independent cloning	/
	SNC1TIR S29A	SNC1 8-181 cDNA without stop	pXZ1/pXZ2	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR H30A	SNC1 8-181 cDNA without stop	pXZ3/pXZ4	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR K33A	SNC1 8-181 cDNA without stop	pXZ5/pXZ6	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR K112A	SNC1 8-181 cDNA without stop	pXZ7/pXZ8	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR K112E	SNC1 8-181 cDNA without stop	XZ50/XZ51	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR Y150A	SNC1 8-181 cDNA without stop	XZ52/XZ53	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR E163A	SNC1 8-181 cDNA without stop	XZ54/XZ55	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR K154A	SNC1 8-181 cDNA without stop	XZ46/XZ47	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR K154E	SNC1 8-181 cDNA without stop	XZ44/XZ45	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR E164A	SNC1 8-181 cDNA without stop	pXZ9/pXZ10	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	SNC1TIR E164K	SNC1 8-181 cDNA without stop	XZ48/XZ49	pMCSG7 SNC1 TIR	pMCSG7	site directed mutagenesis	/
	RPS4TIR	RPS4 10-178 cDNA without stop	/	/	pMCSG7	Ligation independent cloning	Williams et al., 2014
	RPS4TIR M150R	RPS4 10-178 cDNA without stop	MB296/MB297	pMCSG7 RPS4 TIR	pMCSG7	site directed mutagenesis	/
	RPP1TIR	RPP1 NdA 90-254	/	/	pMCSG7	/	Schreiber et al., 2016
PCR product used to generate entry clone	PCR SNC1TIR	SNC1 1-226 cDNA without stop	MB425/MB432	Arabidopsis Col-0 cDNA	/	PCR amplification	/
Entry clone for cell death assays	pDONR SNC1TIR	SNC1 1-226 cDNA without stop	/	PCR SNC1 TIR	pDONR207	BP Gateway	/
Golden gate intermediate modules	pICH41021 RPS4 TIR	RPS4 1-236 genomic without stop	KSP402/KSP1406	Arabidopsis Col-0 gDNA	pICH41021	Blunt-end ligation	Williams et al., 2014
	pICH41021 RPS4 TIR SH/AA	RPS4 1-236 genomic without stop	PKSP521/PKSP522	pICH41021 RPS4 TIR	pICH41021	site directed mutagenesis	/
	pICH41021 RPS4 TIR R116A	RPS4 1-236 genomic without stop	PKSP523/PKSP524	pICH41021 RPS4 TIR	pICH41021	site directed mutagenesis	/
	pICH41021 RPS4 TIR M150R	RPS4 1-236 genomic without stop	PKSP525/PKSP526	pICH41021 RPS4 TIR	pICH41021	site directed mutagenesis	/
	pCR8 RPS4 module 1	RPS4 genomic module 1	/	Arabidopsis No-0 gDNA	pCR8	/	Sohn et al., 2014
	pCR8 RPS4 module 2	RPS4 genomic module 2	/	Arabidopsis No-0 gDNA	pCR8	/	Sohn et al., 2014
	pCR8 RPS4 module 3	RPS4 genomic module 3	/	Arabidopsis No-0 gDNA	pCR8	/	Sohn et al., 2014
	pCR8 RPS4 module 4 (3xHA fused)	RPS4 genomic module 4 without stop (3xHA fused)	/	Arabidopsis No-0 gDNA	pCR8	/	Sohn et al., 2014
	pCR8 RPS4 module 1 SH/AA	RPS4 genomic module 1	PKSP521/PKSP522	pICH4021 RPS4 module 1	pCR8	site directed mutagenesis	/
	pCR8 RPS4 module 1 R116A	RPS4 genomic module 1	PKSP523/PKSP524	pICH4021 RPS4 module 1	pCR8	site directed mutagenesis	/
	pCR8 RPS4 module 1 M150R	RPS4 genomic module 1	PKSP525/PKSP526	pICH4021 RPS4 module 1	pCR8	site directed mutagenesis	/
	pICSL50009 6xHA	6xHA	/	/	pICSL50009	/	Modified from KSC 715 provided by Dr. Youles, TSL
Cell death assays in flax, N. benthamiana and N.tabacum	SNC1TIR-YFP	SNC1 1-226 cDNA without stop	/	pDONR SNC1TIR	pAM-GWY-YFPv (Bernoux et al. 2008)	LR Gateway	/
	SNC1TIR S29A-YFP	SNC1 1-226 cDNA without stop	MB468/MB469	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR H30A-YFP	SNC1 1-226 cDNA without stop	MB470/MB469	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR K112E-YFP	SNC1 1-226 cDNA without stop	MB472/MB473	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR Y150A-YFP	SNC1 1-226 cDNA without stop	MB474/MB475	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR E164K-YFP	SNC1 1-226 cDNA without stop	MB525/MB526	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR E164A-YFP	SNC1 1-226 cDNA without stop	XZ10/XZ11	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR K112A-YFP	SNC1 1-226 cDNA without stop	XZ36/XZ37	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR K154A-YFP	SNC1 1-226 cDNA without stop	XZ6/XZ7	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIRK154E-YFP	SNC1 1-226 cDNA without stop	XZ8/XZ9	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIRK33A-YFP	SNC1 1-226 cDNA without stop	XZ16/XZ17	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	SNC1TIR E163A-YFP	SNC1 1-226 cDNA without stop	XZ18/XZ19	pAM SNC1TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6TIR-YFP	L6 TIR 1-233 genomic without stop	/	/	pAM-GWY-YFPv (Bernoux et al. 2008)	/	Bernoux et al., 2011
	L6TIR trun-YFP	L6 TIR 1-220 genomic without stop	/	/	pAM-GWY-YFPv (Bernoux et al. 2008)	/	Bernoux et al., 2011
	L6TIR R164E-YFP	L6 TIR 1-233 genomic without stop	/	/	pAM-GWY-YFPv (Bernoux et al. 2008)	/	Bernoux et al., 2011
	L6TIR K200E-YFP	L6 TIR 1-233 genomic without stop	/	/	pAM-GWY-YFPv (Bernoux et al. 2008)	/	Bernoux et al., 2011
	L6TIR F79A-YFP	L6 TIR 1-233 genomic without stop	MB335/MB336	pAM L6 TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6TIR K209E-YFP	L6 TIR 1-233 genomic without stop	MB370/MB134	pAM L6 TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6TIR E74K-YFP	L6 TIR 1-233 genomic without stop	MB337/MB338	pAM L6 TIR-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6TIR F79A/R164A	L6 TIR 1-233 genomic without stop	MB335/MB336	pAM L6 TIR R164A-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6TIR F79A/R164E	L6 TIR 1-233 genomic without stop	MB335/MB336	pAM L6 TIR R164E-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6TIR F79A/K200E	L6 TIR 1-233 genomic without stop	MB335/MB336	pAM L6TIR K200E-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	
	L6-YFP	L6 genomic full length without stop	/	/	pAM-GWY-YFPv (Bernoux et al. 2008)	/	Bernoux et al., 2016
	L6 F79A-YFP	L6 genomic full length without stop	MB335/MB336	pAM L6-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6 K209E-YFP	L6 genomic full length without stop	MB370/MB134	pAM L6-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6 R164E-YFP	L6 genomic full length without stop	MB183/MB132	pAM L6-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6 R164A-YFP	L6 genomic full length without stop	MB131/MB132	pAM L6-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6 K200E-YFP	L6 genomic full length without stop	MB184/MB185	pAM L6-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6 W202A-YFP	L6 genomic full length without stop	MB128/MB127	pAM L6-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6MHV-YFP	L6 genomic full length without stop	/	/	pAM-GWY-YFPv (Bernoux et al. 2008)	/	Bernoux et al., 2016
	L6MHV F79A-YFP	L6 genomic full length without stop	MB335/MB336	pAML6MHV-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6MHV K209E-YFP	L6 genomic full length without stop	MB370/MB134	pAML6MHV-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6MHV R164E-YFP	L6 genomic full length without stop	MB183/MB132	pAML6MHV-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6MHV R164A-YFP	L6 genomic full length without stop	MB131/MB132	pAML6MHV-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6MHV K200E-YFP	L6 genomic full length without stop	MB184/MB185	pAML6MHV-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	L6MHV W202A-YFP	L6 genomic full length without stop	MB128/MB127	pAML6MHV-YFP	pAM-GWY-YFPv (Bernoux et al. 2008)	site directed mutagenesis	/
	RPS4 TIR-6HA	RPS4 1-236 genomic without stop + 6HA	/	/	pICH86988	golden gate	Williams et al., 2014
	RPS4 TIR SH/AA-6HA	RPS4 1-236 genomic without stop + 6xHA	/	/	pICH86988	golden gate	Williams et al., 2014
	RPS4 TIR R116A-6HA	RPS4 1-236 genomic without stop + 6xHA	/	/	pICH86988	golden gate	/
	RPS4 TIR M150R-6HA	RPS4 1-236 genomic without stop + 6xHA	/	/	pICH86988	golden gate	/
	RPS4-3HA	RPS4 genomic full length without stop + 3xHA (x4 modules)	/	/	pICH86988	golden gate	Sohn et al., 2014
	RPS4 SH/AA-3HA	RPS4 genomic full length without stop + 3xHA (x4 modules)	/	/	pICH86988	golden gate	Williams et al., 2014
	RPS4 R116A-3HA	RPS4 genomic full length without stop + 3xHA (x4 modules)	/	/	pICH86988	golden gate	/
	RPS4 M150R-3HA	RPS4 genomic full length without stop + 3xHA (x4 modules)	/	/	pICH86988	golden gate	/
	RRS1-3Flag	Ws-2 RRS1 genomic full length without stop (x7 modules) + 3xFlag	/	/	pICH86988	golden gate	Williams et al., 2014
	RRS1 SLH1-3Flag	Ws-2 RRS1 genomic full length without stop (x7 modules) + 3xFlag	/	/	pICH86988	golden gate	Sohn et al., 2014
	AvrRPS4-YFP	avrRps4(Ppi151) full length without stop + YFP	/	/	pEpiGreenB5-GG	golden gate	Sohn et al., 2012
	PopP2-YFP	PopP2 (Rs GMI1000) 149-488 without stop + YFP	/	/	pICH86988	golden gate	Sohn et al., 2014
Entry clones for Y2H assays	pDONR L6TIR F79A	L6 29-233 cDNA with stop	MB335/MB336	pDONR L6TIR 29-233 (Bernoux et al., 2011)	pDONR207	site directed mutagenesis	/
	pDONR L6TIR K209E	L6 29-233 cDNA with stop	MB339/MB134	pDONR L6TIR 29-233 (Bernoux et al., 2011)	pDONR207	site directed mutagenesis	/
	pDONR L6TIR E74K	L6 29-233 cDNA with stop	MB337/MB338	pDONR L6TIR 29-233 (Bernoux et al., 2011)	pDONR207	site directed mutagenesis	/
	pDONR RPS4TIR R116A	RPS4 1-183 cDNA no stop	MB247/MB248	pDONR RPS4TIR 1-183 (Williams et al., 2014)	pDONR207	site directed mutagenesis	/
	pDONR RPS4TIR M150R	RPS4 1-183 cDNA no stop	MB296/MB297	pDONR RPS4TIR 1-183 (Williams et al., 2014)	pDONR207	site directed mutagenesis	/
Yeast two-hybrid assays	AD L6TIR	L6 29-233 cDNA with stop	/	/	pGADT7-GWY	/	Bernoux et al., 2011
	AD L6TIR F79A	L6 29-233 cDNA with stop	/	pDONR L6TIR F79A	pGADT7-GWY	LR Gateway	/
	AD L6TIR K209E	L6 29-233 cDNA with stop	/	pDONR L6TIR K209E	pGADT7-GWY	LR Gateway	/
	AD L6TIR E74K	L6 29-233 cDNA with stop	/	pDONR L6TIR E74K	pGADT7-GWY	LR Gateway	/
	AD RPS4TIR	RPS4 1-183 cDNA no stop	/	/	pGADT7-GWY	LR Gateway	Williams et al., 2014
	AD RPS4TIR R116A	RPS4 1-183 cDNA no stop	/	pDONR RPS4TIR R116A	pGADT7-GWY	LR Gateway	/
	AD RPS4TIR M150R	RPS4 1-183 cDNA no stop	/	pDONR RPS4TIR M150R	pGADT7-GWY	LR Gateway	/
	AD RRS1TIR	RRS1 1-155 cDNA no stop	/	/	pGADT7-GWY	/	Williams et al., 2014
	BD L6TIR	L6 29-233 cDNA with stop	/	/	pGADT7-GWY	LR Gateway	Bernoux et al., 2011
	BD L6TIR F79A	L6 29-233 cDNA with stop	/	pDONR L6TIR F79A	pGBKT7-GWY	LR Gateway	/
	BD L6TIR K209E	L6 29-233 cDNA with stop	/	pDONR L6TIR K209E	pGBKT7-GWY	LR Gateway	/
	BD L6TIR E74K	L6 29-233 cDNA with stop	/	pDONR L6TIR E74K	pGBKT7-GWY	LR Gateway	/
	BD RPS4TIR	RPS4 1-183 cDNA no stop	/	/	pGBKT7-GWY	LR Gateway	Williams et al., 2014
	BD RPS4TIR R116A	RPS4 1-183 cDNA no stop	/	pDONR RPS4TIR R116A	pGBKT7-GWY	LR Gateway	/
	BD RPS4TIR M150R	RPS4 1-183 cDNA no stop	/	pDONR RPS4TIR M150R	pGBKT7-GWY	LR Gateway	/

Table S5: Primer details

Primer name	Primer sequence 5'-3'
MB127	TTTTAAGTCTCCGACCTTCTTTAGAGCATCT
MB128	AGGTCGGAGACTTAAAAGGATGGCACATCGGAAAGA
MB131	TGGTGGATCCAAGTGACGTAGCACATCAGACTGGA
MB132	TACGTCACTTGGATCCACCATATAAAAAATAGGA
MB134	ATTCTTTCCGATGTGCCATCCTTTTAAGTCT
MB148	AGGTCGGAGACTTAAAAGGAGCGCACATCGGAAAGA
MB183	TGGTGGATCCAAGTGACGTAGAACATCAGACTGGA
MB184	TCTAAAGAAGGTCGGAGACTTAGAAGGATGGCACATCGGA
MB185	TAAGTCTCCGACCTTCTTTAGAGCATCTTTCCA
MB247	AGCTGGAGCCATCCACCGTTGCAGATTTGAAAGGAAAGT
MB248	AACGGTGGATGGCTCCAGCTTGTAGAAGATTGGA
MB296	TTAACTTGATTCCTAACATTAGGGGCATCATCATTGACA
MB297	AATGTTAGGAATCAAGTTAAAAGCTTCCTTCCATTTCT
MB335	ACTCGTGAACAGTTCACCGATGCCCTATATCAGTCTCTCCGT
MB336	ATCGGTGAACTGTTCACGAGTATCTGGACCCCTGAAACT
MB337	TCAGGGGTCCAGATACTCGTAAACAGTTCACCGATTTCCT
MB338	ACGAGTATCTGGACCCCTGAAACTCAAAAACACTT
MB339	TGGCACATCGGAAAGAATGACGAGCAGGGAGCTATAGCA
MB370	TGGCACATCGGAAAGAATGACGAGTATGTAATCCTCATCCT
MB425	CAAAAAAGCAGGCTTAATGGAGATAGCTTCTTCTTCTGGCA
MB432	CAAGAAAGCTGGGTCGTAAAGAGCTCTTCCTATGGTACT
MB468	ATGTCCGTGACTCATTCCTCGCCCATCTTCTCAAGGA
MB469	GAGGAATGAGTCACGGACATCTTCTCCACGAAAGCT
MB470	TGTCCGTGACTCATTCCTCAGCGCTCTTCTCAAGGAGCT
MB471	ATGTCCGTGACTCATTCCTCGCCGCTCTTCTCAAGGAGCT
MB472	ACGTTGATGCTTCGGAAGTTGAAAAACAGACCGGCGAAT
MB473	AACTTCCGAAGCATCAACGTGGAAGAAAATCGGAAT
MB474	TGTTGCAGTTATGGCCGGAGCTGATCTTCGGAAATGGCCT
MB475	TCCGGCCATAACTGCAACAGCTGCTAGAGCTTGCT
MB525	AGCAGCCATGATTGAAAAGCTTGCCGAGGATGTTTTGAGA
MB526	TTCAATCATGGCTGCTTCACTAGGCCATTTCCGAAGA
KSP402	GGTCTCGAATGGAGACATCATCTATTTCCACTGTG
KSP1406	GGTCTCTCGAACCCAACAACTCCAATGATACG
PKSP521	CGGAGATTCGTCGCCGCTCTCGTAACGGCC
PKSP522	GGCCGTTACGAGAGCGGCGACGAATCTCCG
PKSP523	GCCATCCACCGTTGCAGATTTGAAAGGAAA
PKSP524	TTTCCTTTCAAATCTGCAACGGTGGATGGC
PKSP525	CTTGATTCCTAACATTAGGGGCATCATCAT
PKSP526	ATGATGATGCCCCTAATGTTAGGAATCAAG
XZ6	GATATGATCTTCGGGCATGGCCTAGTGA
XZ7	TCACTAGGCCATGCCCGAAGATCATATC
XZ8	GGATATGATCTTCGGGAATGGCCTAGTGAA
XZ9	TTCACTAGGCCATTCCCGAAGATCATATCC
XZ10	AGCCATGATTGAAGCACTTGCCGAGG
XZ11	CCTCGGCAAGTGCTTCAATCATGGCT
XZ14	AGCCATCTTCTCGCAGAGCTCAGG
XZ15	CCTGAGCTCTGCGAGAAGATGGCT
XZ18	GCAGCCATGATTGCAGAGCTTGCC
XZ19	GGCAAGCTCTGCAATCATGGCTGC

XZ36	ATGCTTCGGAAGTTGCAAAACAGACCGG
XZ37	CCGGTCTGTTTTGCAACTTCCGAAGCAT
XZ38	TACTTCCAATCCAATGCGGGTAGCCGTCGTTATGATGTTTTTCCGAG
XZ39	TATTCCACTTCCAATGTTAACCAAAATCATCACTCGGGGTCATGGTTT
XZ44	TTATGATCTGCGTGAATGGCCGAGCG
XZ45	CGCTCGGCCATTCACGCAGATCATAA
XZ46	TATGATCTGCGTGCATGGCCGAGC
XZ47	GCTCGGCCATGCACGCAGATCATA
XZ48	AGCAGCAATGATTGAAAAACTGGCAGAAGATG
XZ49	CATCTTCTGCCAGTTTTTCAATCATTGCTGCT
XZ50	ATGCCAGCGAAGTTGAAAAACAGACCGG
XZ51	CCGGTCTGTTTTTCAACTTCGCTGGCAT
XZ52	CAGTTATGGCAGGTGCAGATCTGCGTAA
XZ53	TTACGCAGATCTGCACCTGCCATAACTG
XZ54	GAAGCAGCAATGATTGCAGAACTGGCAGAA
XZ55	TTCTGCCAGTTCTGCAATCATTGCTGCTTC
pXZ1	GTTCGTGATAGCTTTCTGGCGCATCTGCTGAAAGAA
pXZ2	TTCTTTCAGCAGATGCGCCAGAAAGCTATCACGAAC
pXZ3	CGTGATAGCTTTCTGAGCGCACTGCTGAAAGAACTG
pXZ4	CAGTTCTTTCAGCAGTGCGCTCAGAAAGCTATCACG
pXZ5	GAGCCATCTGCTGGCGGAACTGCGTG
pXZ6	CACGCAGTTCCGCCAGCAGATGGCTC
pXZ7	GATGCCAGCGAAGTTGCGAAACAGACCGGT
pXZ8	ACCGGTCTGTTTCGCAACTTCGCTGGCATC
pXZ9	GAAGCAGCAATGATTGAAGCGCTGGCAGAAGATGTG
pXZ10	CACATCTTCTGCCAGCGCTTCAATCATTGCTGCTTC

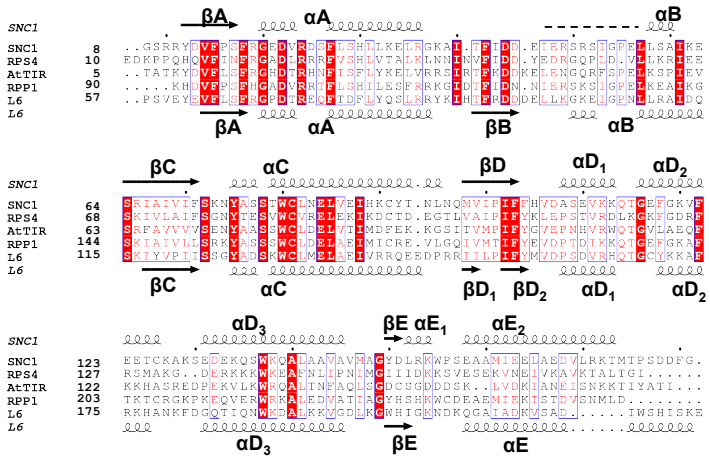


Figure S1. Sequence alignment of representative plant TIR domains.

The elements of secondary structure for SNC1^{TIR} and L6^{TIR} are shown above and below the sequence alignment, respectively. The internal residues missing from the SNC1^{TIR} structure are indicated by dashes (-).

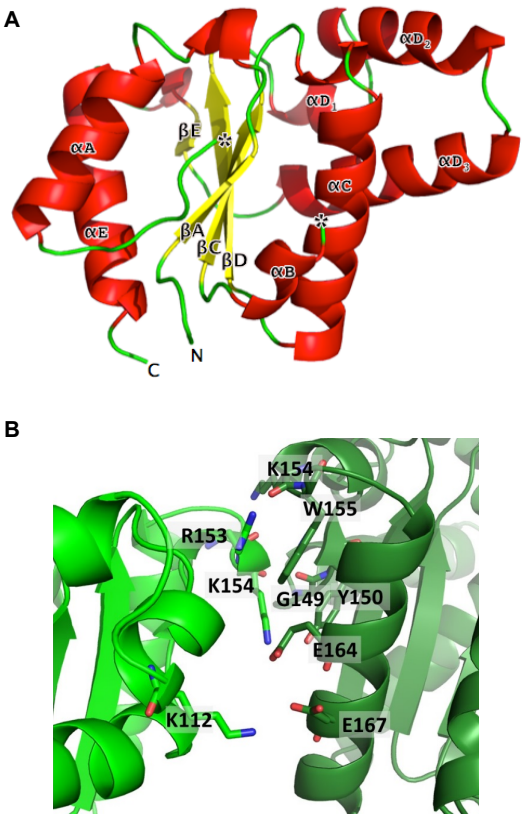


Figure S2. The structure of SNC1^{TIR} and the DE-interface.
(A) The structure of SNC1^{TIR}. The elements of secondary structure follow the nomenclature used in the L6^{TIR} structure (3). The two ends of the flexible internal region are indicted by asterisks (*).
(B) Residues that are involved in DE-interface interactions in SNC1^{TIR} are shown in stick representation.

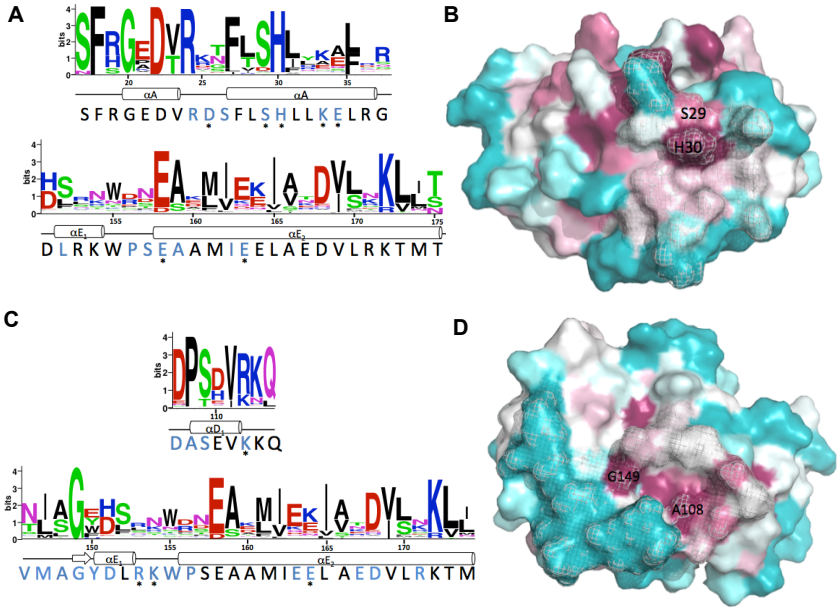


Figure S3. Amino-acid conservation in the interfaces.
(A) and *(C)* show the sequence logo (Weblogo 3.3) (33) based on the sequence alignment of 516 plant TIR domains generated by a PSI-Blast (34) search with SNC1^{TIR} (35~95% identity to SNC1^{TIR}). The sequence and secondary structure elements of SNC1^{TIR} are shown below the logo. Buried surface residues are shown in blue and the key interacting residues are marked by asterisks (*).
(B) and *(D)* show the degree of conservation mapped onto the surface of the SNC1^{TIR} structure (Consurf) (23). Variable and conserved regions are colored in cyan and purple, respectively. The buried surface areas are contoured by mesh. (A) and (B) correspond to the AE-interface, and (C) and (D) correspond to the DE-interface.

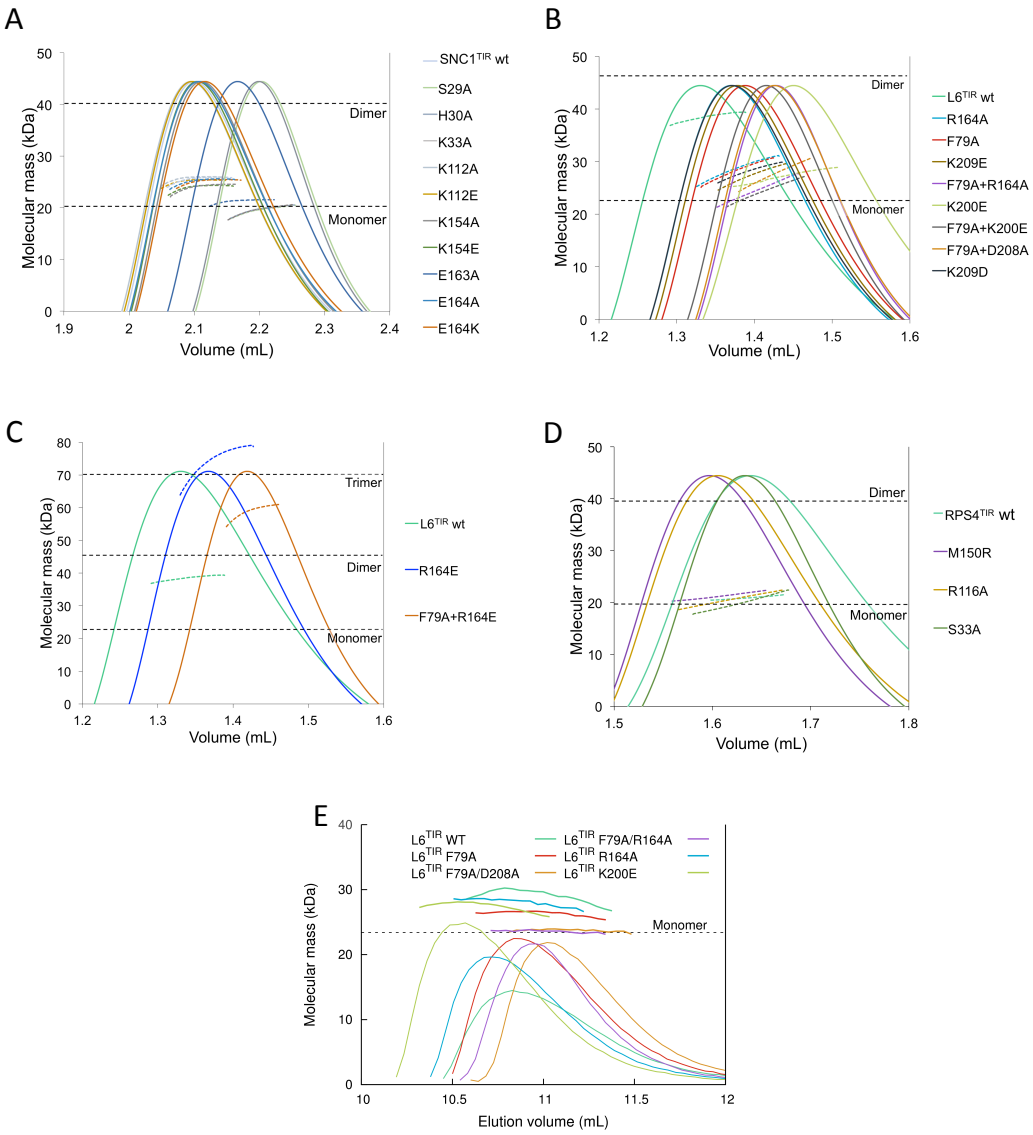

A
Molecular mass (kDa)
Volume (mL)
Dimer
Monomer
SNC1TIR wt
S29A
H30A
K33A
K112A
K112E
K154A
K154E
E163A
E164A
E164K
B
Molecular mass (kDa)
Volume (mL)
Dimer
Monomer
L6TIR wt
R164A
F79A
K209E
F79A+R164A
K200E
F79A+K200E
F79A+D208A
K209D
C
Molecular mass (kDa)
Volume (mL)
Trimer
Dimer
Monomer
L6TIR wt
R164E
F79A+R164E
D
Molecular mass (kDa)
Volume (mL)
Dimer
Monomer
RPS4TIR wt
M150R
R116A
S33A
E
Molecular mass (kDa)
Elution volume (mL)
L6TIR WT
L6TIR F79A
L6TIR F79A/D208A
L6TIR F79A/R164A
L6TIR R164A
L6TIR K200E
Monomer

Figure S4. Solution analyses of AE and DE-interface mutants.
Multi-angle light scattering (MALS) analysis of TIR domain AE and DE-interface mutants (A-B and D-E) and small-angle X-ray scattering (SAXS) analysis of $L6^{TIR}$ AE and DE mutants (C). Refractive index (RI) (normalized arbitrary units) of the proteins as measured by MALS is shown by solid lines and molecular masses (kDa) of the samples, calculated from light scattering measurements under the peak, are shown as dashed lines of the same color as the RI trace. Black dashed lines represent theoretical monomer, dimer and trimer molecular masses.
(A) MALS analysis of $SNC1^{TIR}$ interface mutants.
(B) MALS analysis of $L6^{TIR}$ interface mutants that disrupt self-association.
(C) MALS analysis of $L6^{TIR}$ interface mutants that promote self-association.
(D) MALS analysis of $RPS4^{TIR}$ interface mutants.
(E) Molecular masses from SEC-SAXS of $L6^{TIR}$ mutants. Colored lines show the Guinier *I*(0) trace for each sample across the SEC-SAXS elution, analogous to RI in MALS. Matching coloured lines above each peak show the calculated molecular mass. The theoretical monomeric mass is shown as a dotted line.

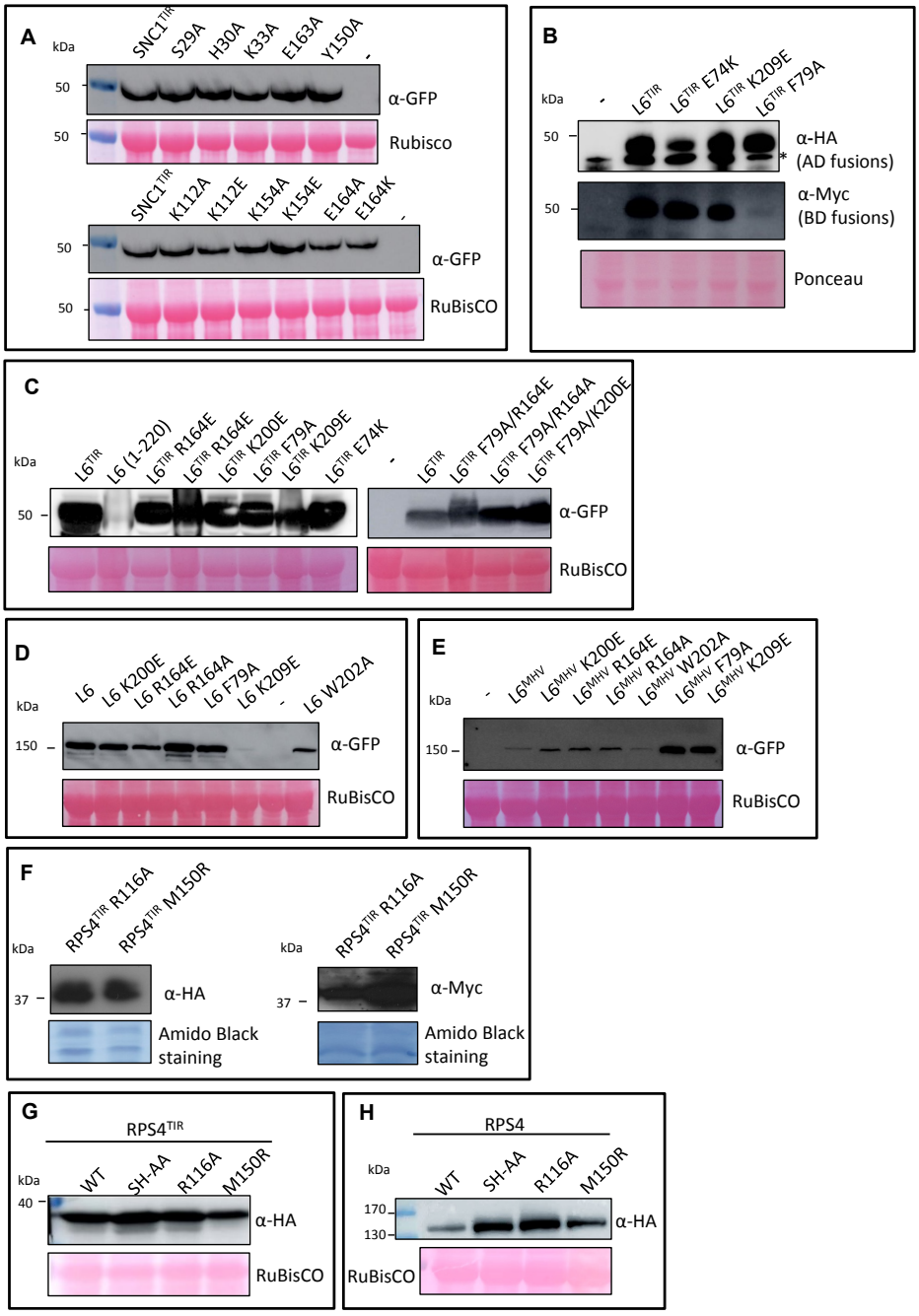

A
kDa
SNC1TIR
S29A
H30A
K33A
E163A
Y150A
-
50
α-GFP
50
Rubisco
SNC1TIR
K112A
K112E
K154A
K154E
E164A
E164K
-
50
α-GFP
50
RuBisCO
B
kDa
-
L6TIR
L6TIR E74K
L6TIR K209E
L6TIR F79A
50
α-HA
(AD fusions)
*
50
α-Myc
(BD fusions)
Ponceau
C
kDa
L6TIR
L6 (1-220)
L6TIR R164E
L6TIR R164E
L6TIR K200E
L6TIR F79A
L6TIR K209E
L6TIR E74K
-
L6TIR
L6TIR F79A/R164E
L6TIR F79A/R164A
L6TIR F79A/K200E
50
α-GFP
RuBisCO
D
kDa
L6
L6 K200E
L6 R164E
L6 R164A
L6 F79A
L6 K209E
-
L6 W202A
150
α-GFP
RuBisCO
E
kDa
-
L6MHV
L6MHV K200E
L6MHV R164E
L6MHV R164A
L6MHV W202A
L6MHV F79A
L6MHV K209E
150
α-GFP
RuBisCO
F
kDa
RPS4TIR R116A
RPS4TIR M150R
37
α-HA
Amido Black
staining
kDa
RPS4TIR R116A
RPS4TIR M150R
37
α-Myc
Amido Black
staining
G
RPS4TIR
kDa
WT
SH-AA
R116A
M150R
40
α-HA
RuBisCO
H
RPS4
kDa
WT
SH-AA
R116A
M150R
170
130
α-HA
RuBisCO

Figure S5. Immunoblot analyses.

(A) Immunoblot detection of SNC1^{TIR}-YFP fusions with anti-GFP antibodies, 24 hours after agroinfiltration in *N. benthamiana* leaves. The negative control is indicated by a dash and corresponds to non-infiltrated *N. benthamiana* leaf tissue.

(B) Immunoblot detection of GAL4-BD and GAL4-AD fusions of L6^{TIR} using anti-Myc and anti-HA antibodies, respectively. * indicates the presence of a non-specific band. The negative control is indicated by a dash and corresponds to untransformed yeast.

(C) Immunoblot detection of L6^{TIR} –YFP mutants using anti-GFP antibodies. Protein samples were taken 3 days after agroinfiltration in Hoshangabad flax leaves. The negative control is indicated by a dash and corresponds to non-infiltrated Hoshangabad flax leaf tissue.

(D-E) Immunoblot detection of full-length L6-YFP (D) and L6^{MHV}–YFP (E) mutants using anti-GFP antibodies. Protein samples were taken 3 days and 48 h after agroinfiltration in W38 tobacco, respectively. The negative control is indicated by a dash and corresponds to non-infiltrated W38 tobacco leaf tissue.

(F) Immunoblot detection of GAL4-AD and GAL4-BD- fusions of RPS4^{TIR} using anti-HA and anti-Myc antibodies, respectively.

(G) Immunoblot detection of RPS4^{TIR}-6HA mutants using anti-HA antibodies. Protein samples were taken 36 h after agroinfiltration in W38 tobacco.

(H) Immunoblot detection of full-length RPS4-6HA mutants using anti-HA antibodies. Protein samples were taken 36 h after agroinfiltration in W38 tobacco.

Protein loading for all plant samples is indicated by red Ponceau staining of the RuBisCO large subunit. Protein loading for yeast samples is indicated by red Ponceau or Amido Black staining.

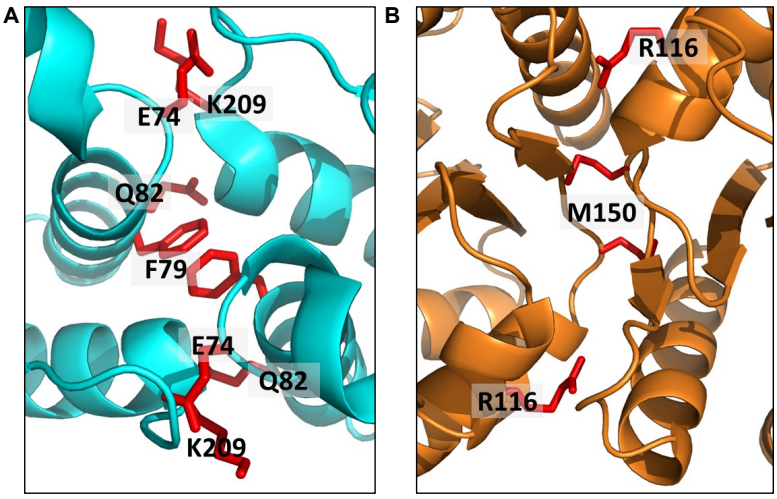


Figure S6. Modelling of the AE-interfaces in L6^{TIR} and DE-interfaces in RPS4^{TIR}.
(A) Structural model of AE-interface dimer of L6^{TIR}; key residues shown as red sticks.
(B) Structural model of DE-interface in RPS4^{TIR}; key residues shown as red sticks.

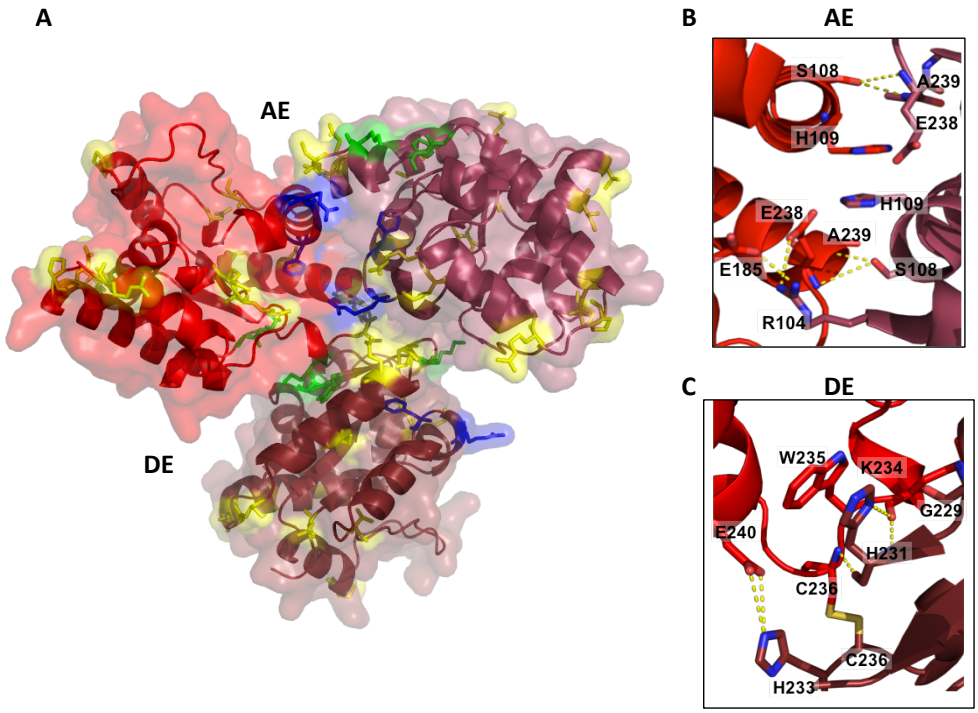


Figure S7. Polymorphic residues in different RPP1 alleles that affect self-association and cell-death signalling localize to the AE and DE-interfaces of RPP1^{TIR}-.
(A) Surface-cartoon representation of RPP1^{TIR} structure, indicating residues polymorphic between alleles from Arabidopsis ecotypes NdA and WsB that affect function (blue), additional polymorphic residues (yellow), and introduced mutations that affect function (green).
(B) The AE-interface of RPP1 (~1290 Å^2 buried; molecules "A" and "B") is defined by histidine stacking ($H109_A$, $H109_B$) in the centre of αA helices and stabilized by a hydrogen-bonding network formed by the neighbouring residues.
(C) The DE-interface of RPP1 is slightly smaller than the AE-interface (buried surface area of ~850 Å^2; molecules "A" and "C"); multiple non-polar aliphatic and aromatic residues from the βE strand form its core through van-der-Waals and π-stacking interactions. DE-interfaces between RPP1^{TIR}, SNC1^{TIR} and L6^{TIR}, share similar residue interactions, with the core containing a conserved glycine (G229/G149/G201, respectively) that is surrounded by aromatic (Y230/Y150/W202, respectively) and aliphatic residues. C236 forms an inter-molecular disulfide bond within the crystals of RPP1 NdA not seen in other structures. Previous studies of RPP1^{TIR} demonstrated that introduction of this non-conserved residue into the WsB allele (WsB R230C) promoted self-association and cell death *in planta* (10). However, this substitution does not generate a stable dimer, and mutation of C236 to alanine in NdA does not disrupt cell-death signalling of the RPP1 NdA TIR domain. Collectively, this suggests that this disulfide formation observed in the crystals structure is unlikely to play a functional role in-planta.

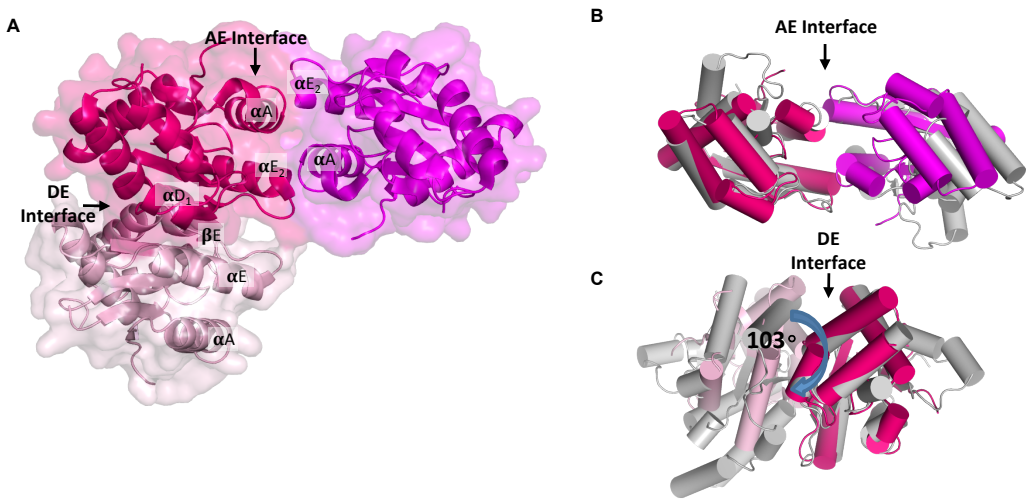


Figure S8. Comparison of AE and DE-interfaces.
(A) Ribbon representation of the AtTIR crystal structure and the AE and DE-interfaces, with molecules sharing the AE-interface colored pink and magenta and the DE-interface pink and light pink.
(B) Comparison of the AE-interface from the $RPS4^{TIR}$ (grey), AtTIR (pink-magenta) with the chains on the left superimposed.
(C) Comparison of the DE-interface from the $L6^{TIR}$ (grey), AtTIR (pink-light pink) structures; only the chains at the top are superimposed.

Figure S10

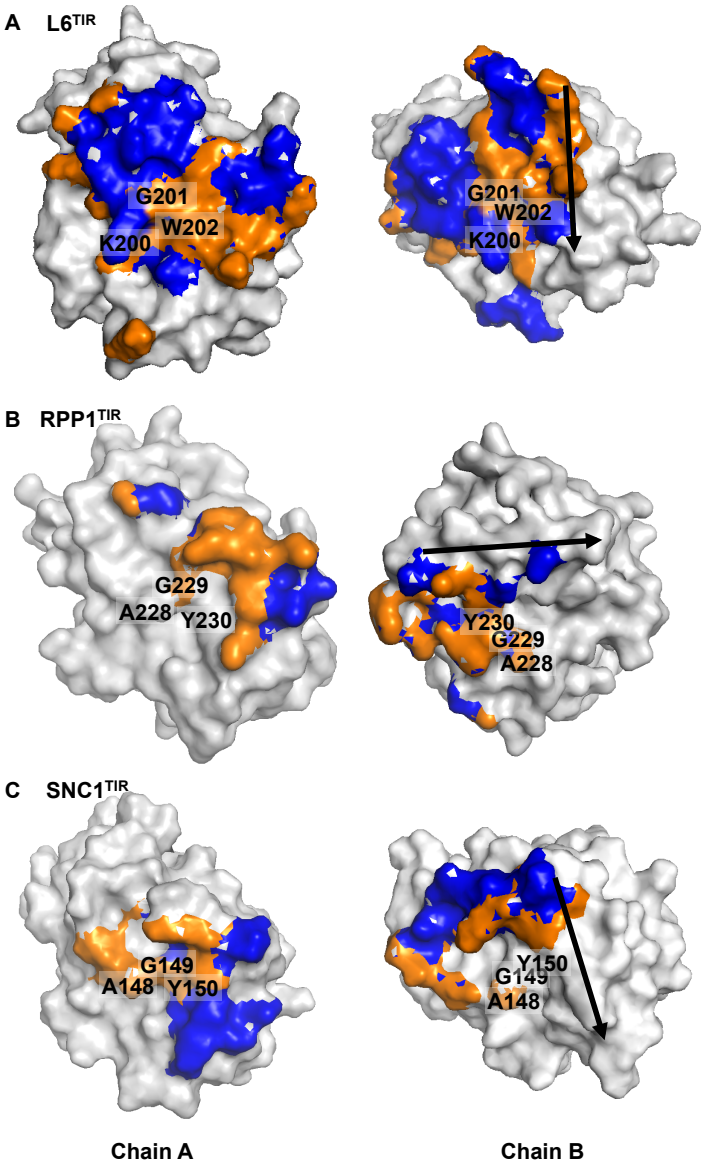


Figure S9. Hydrophobic cores in DE-interfaces.
The TIR-domain structures are shown in surface representation with polar residues buried in the interfaces shown in blue, and hydrophobic residues buried in the interfaces in orange. Chain A of each DE-interface dimer is superimposed onto $L6^{TIR}$ chain A and the relative rotation of chain B is highlighted by an arrow that shows the position and direction of the αE helix. The residues in the hydrophobic cores of the interfaces are marked on top.

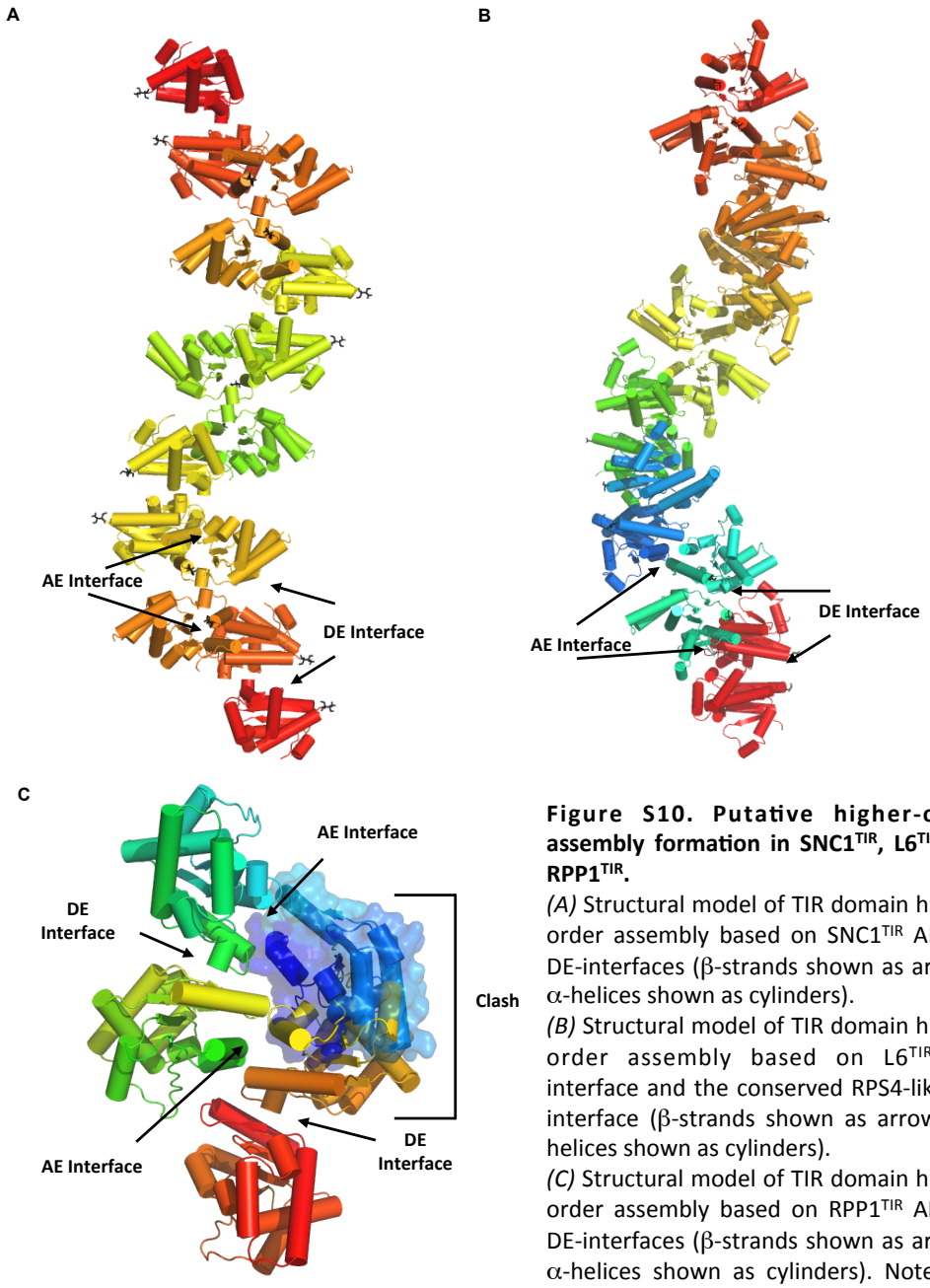


Figure S10. Putative higher-order assembly formation in SNC1^{TIR}, L6^{TIR} and RPP1^{TIR}.

(A) Structural model of TIR domain higher-order assembly based on SNC1^{TIR} AE and DE-interfaces (β-strands shown as arrows, α-helices shown as cylinders).

(B) Structural model of TIR domain higher-order assembly based on L6^{TIR} DE-interface and the conserved RPS4-like AE-interface (β-strands shown as arrows, α-helices shown as cylinders).

(C) Structural model of TIR domain higher-order assembly based on RPP1^{TIR} AE and DE-interfaces (β-strands shown as arrows, α-helices shown as cylinders). Note that such assembly would lead to steric clashes, as indicated in the figure.

PNAS

The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins

Lachlan W. Casey[a,1], Peter Lavrencic[a,b,1], Adam R. Bentham[a,c,1], Stella Cesari[d], Daniel J. Ericsson[e], Tristan Croll[f], Dušan Turk[g], Peter A. Anderson[c], Alan E. Mark[a], Peter N. Dodds[d], Mehdi Mobli[b,2], Bostjan Kobe[a,2], and Simon J. Williams[a,c,h,2]

[a]School of Chemistry and Molecular Biosciences, Institute for Molecular Bioscience and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, QLD 4072, Australia; [b]Centre for Advanced Imaging, University of Queensland, Brisbane, QLD 4072, Australia; [c]School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia; [d]Commonwealth Scientific and Industrial Research Organization Agriculture, Canberra, ACT 2601, Australia; [e]Macromolecular Crystallography (MX) Beamlines, Australian Synchrotron, Melbourne, VIC 3168, Australia; [f]School of Biomedical Sciences, Queensland University of Technology, Brisbane, QLD 4001, Australia; [g]Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, International Postgraduate School Jozef Stefan, Jozef Stefan Institute, 1000 Ljubljana, Slovenia; and [h]Plant Sciences Division, Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia

Edited by Jeffery L. Dangl, Howard Hughes Medical Institute and The University of North Carolina at Chapel Hill, Chapel Hill, NC, and approved September 13, 2016 (received for review June 20, 2016)

Plants use intracellular immunity receptors, known as nucleotide-binding oligomerization domain-like receptors (NLRs), to recognize specific pathogen effector proteins and induce immune responses. These proteins provide resistance to many of the world's most destructive plant pathogens, yet we have a limited understanding of the molecular mechanisms that lead to defense signaling. We examined the wheat NLR protein, Sr33, which is responsible for strain-specific resistance to the wheat stem rust pathogen, *Puccinia graminis f. sp. tritici*. We present the solution structure of a coiled-coil (CC) fragment from Sr33, which adopts a four-helix bundle conformation. Unexpectedly, this structure differs from the published dimeric crystal structure of the equivalent region from the orthologous barley powdery mildew resistance protein, MLA10, but is similar to the structure of the distantly related potato NLR protein, Rx. We demonstrate that these regions are, in fact, largely monomeric and adopt similar folds in solution in all three proteins, suggesting that the CC domains from plant NLRs adopt a conserved fold. However, larger C-terminal fragments of Sr33 and MLA10 can self-associate both in vitro and *in planta*, and this self-association correlates with their cell death signaling activity. The minimal region of the CC domain required for both cell death signaling and self-association extends to amino acid 142, thus including 22 residues absent from previous biochemical and structural protein studies. These data suggest that self-association of the minimal CC domain is necessary for signaling but is likely to involve a different structural basis than previously suggested by the MLA10 crystallographic dimer.

plant innate immunity | resistance protein | NLR proteins | effector-triggered immunity | nuclear magnetic resonance spectroscopy

Plant diseases constitute a major economic and social burden worldwide, and the appearance of new or more virulent pathogens can pose significant challenges. Plants rely on their innate immunity systems to combat pathogens, an important component of which is the recognition of pathogen effector molecules by resistance proteins, a process commonly referred to as effector-triggered immunity. Resistance protein activation triggers a defense mechanism known as the hypersensitive response, which often culminates in localized cell death at the site of infection, leading to general immunity of the whole plant (1).

One such R protein is encoded by the recently discovered wheat gene Sr33 (2). Sr33 confers resistance to the virulent Ug99 strain of wheat stem rust, *Puccinia graminis f. sp. tritici*, a pathogen recognized for its potential threat to global food security. Sr33 is orthologous to the barley powdery mildew resistance protein MLA and rye Sr50 genes (3, 4) and encodes a member of the canonical class of plant resistance proteins, consisting of a central nucleotide-binding (NB) domain, a C-terminal leucine-rich repeat (LRR) domain, and an N-terminal coiled-coil (CC) domain (2). Such proteins have a domain arrangement and function similar to the NB oligomerization domain-like receptors (NLRs) from mammals, and are commonly referred to as plant NLRs (5).

Although the molecular details of plant NLR activation and signaling are not fully understood, targeted studies have helped define the roles of their different domains. The central NB domain appears to control the activation of the protein through NB and nucleotide exchange (6–8). The LRR domain plays a role in effector recognition specificity for a number of plant NLRs, and, in some cases, it is implicated in effector binding (9–11). The LRR domain also appears to have a general autoinhibitory role, and structural and biochemical studies of the human NLR proteins NAIP and NLRC4 support this conclusion (12). In plant NLRs, the

Significance

Plants and animals use intracellular immunity receptors, known as nucleotide-binding oligomerization domain-like receptors (NLRs), to defend themselves against invading microbes. In this study, we report the solution structure of the N-terminal coiled-coil (CC) domain from the wheat stem rust resistance protein Sr33. Remarkably, this structure differs substantially from the published crystal structure of the equivalent region from the orthologous barley powdery mildew resistance protein MLA10. Using a structural, biophysical, and functional approach, we compare the Sr33 CC domain with other structurally defined NLR CC domains. Collectively, this work redefines our current understanding of the structure and function of plant NLR CC domains, which has significant implications for future studies into this important class of defense receptors.

Author contributions: P.A.A., A.E.M., P.N.D., M.M., B.K., and S.J.W. designed research; L.W.C., P.L., A.R.B., S.C., D.J.E., and S.J.W. performed research; P.A.A., A.E.M., P.N.D., M.M., and B.K. contributed new reagents/analytic tools; L.W.C., P.L., A.R.B., S.C., D.J.E., T.C., D.T., P.N.D., M.M., B.K., and S.J.W. analyzed data; and L.W.C., P.L., A.R.B., S.C., P.N.D., M.M., B.K., and S.J.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The protein structures and data used to derives these structures have been deposited with the Protein Data Bank, www.pdb.org [PDB ID codes 2NCG (Sr33(6-120) and 5T1Y (Mla10(5-120))].

See Commentary on page 12619.

[1]L.W.C., P.L., and A.R.B. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: s.williams8@uq.edu.au, m.mobli@uq.edu.au, or b.kobe@uq.edu.au.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1609922113/-/DCSupplemental.

N-terminal domain generally consists of either a CC domain, as in Sr33, or a Toll/interleukin-1 receptor/resistance protein (TIR) domain. Both CC and TIR domains are implicated in downstream signaling, and have been shown to be necessary and sufficient for cell death responses in a number of systems (6, 13–18).

Oligomerization is key to animal NLR activation and signaling, as demonstrated by the structural characterization of the NAIP2/NLRC4 inflammasome NLR pair (19–21). Our understanding of these processes in plant NLRs is more limited. To date, effector-induced self-association of a full-length plant NLR has been demonstrated for only the TIR-NLRs, including the tobacco mosaic virus resistance protein N (22) and the *Arabidopsis* protein RECOGNITION OF PERONOSPORA PARASITICA1 (RPP1) (23). Self-association in the absence of effector proteins has been observed in CC-NLR proteins from *Arabidopsis* RPS5 (24), barley (MLA10) (14), maize (Rp1-D) (18), and wheat (Sr33 and Sr50) (25). However, MLA10, Sr33, and Sr50 also display autoactivity when overexpressed (16, 25), suggesting autoassociation may mimic the activated state.

To date, structure-guided studies of plant NLRs have been restricted to the N-terminal TIR or CC domain (6, 14, 17, 26), with the exception of the noncanonical integrated-sensor heavy metal-associated domain from the rice NLR Pik (27). Although the plant NLR TIR domains have a conserved fold (6, 17), the structures of the two known CC domain fragments from barley MLA10 and potato Rx are strikingly different. The N-terminal amino acids 5–120 of MLA10 (designated MLA10$^{5\text{-}120}$) crystallized as an antiparallel homodimer adopting a helix–loop–helix fold. However, the equivalent region from Rx (Rx$^{1\text{-}122}$) adopted a compact, four-helical bundle when crystallized in a 1:1 heteroassociation with the Trp-Pro-Pro (WPP) domain from its cofactor protein RanGAP2 (26). MLA10$^{5\text{-}120}$ was also reported to dimerize in vitro, whereas Rx$^{1\text{-}122}$ was monomeric in vitro in the absence of cofactors. A larger CC-containing fragment of MLA10 (MLA10$^{1\text{-}160}$) was capable of inducing cell death *in planta* (14), whereas studies of Rx found no cell death induction by CC domain fragments (28). Although the CC domains from Rx and MLA10 share low sequence identity in this region (<20%), both contain the EDVID motif and are classified within the CC$_{\text{EDVID}}$ domain class from plant NLRs (29). These studies suggested that significant structural and mechanistic variation is present among the CC$_{\text{EDVID}}$ domains of plant NLRs.

Recently, Cesari et al. (25) found that the CC-containing fragments of MLA10, Sr33, and Sr50 corresponding to MLA10 residues 1–160 induced cell death and self-associated *in planta*, whereas the smaller 1–120 fragments (equivalent to the MLA10 structure) did not. Thus, to better understand the role of the CC domains in NLR protein signaling, we undertook a structural and functional study of the CC region of the wheat stem rust NLR protein Sr33. Here, we present the solution 3D structure of the Sr33 CC domain (residues 6–120; Sr33$^{6\text{-}120}$), determined by NMR spectroscopy. The structure resembles the CC domain of Rx more closely than the CC domain of MLA10. Prompted by this finding, we carried out a detailed biophysical comparison of the CC domains of Sr33, MLA10, and Rx, which suggests that these CC-NLRs all adopt a common fold and are monomeric in solution. We also find that the minimal functional regions for MLA10, Sr33, and Sr50 CC domain self-association and cell death signaling exceed the boundaries used in in vitro and structural studies to date.

Results

The NMR Structure of Sr33$^{6\text{-}120}$ Reveals a Compact, Four-Helix Bundle. Previous structural studies of the Sr33 ortholog MLA10 encompassed residues 5–120 of the CC domain (MLA10$^{5\text{-}120}$). For our investigation, we initially focused on an equivalent region within the Sr33 CC domain. Soluble Sr33 CC domain protein was produced by recombinant expression of a construct comprising residues 6–120 in *Escherichia coli*, and the atomic structure was determined by NMR spectroscopy.

Sr33$^{6\text{-}120}$ appears to be monomeric under the conditions used for the NMR structural studies. This conclusion is evidenced by the sharp line widths of the resonances. In addition, the average T_2 relaxation rates of the backbone amides yield an overall rotational correlation time of the protein of ∼8.7 ns. This rotational correlation time corresponds to a spherical protein with a molecular mass of ∼13 kDa, compared with the theoretical monomeric molecular mass of 13.1 kDa (30) (*SI Appendix*, Table S1). Analysis of the assigned chemical shifts (31) revealed four distinctive α-helical regions (α1, residues 7–19; α2, residues 28–51; α3, residues 60–87; and α4, residues 99–115). Other than the termini, two small regions (comprising residues 22–23 and residues 88–91) were found to have near-random-coil chemical shifts, indicating that these regions are highly dynamic. The residues that correspond to 88–91 in the MLA10$^{5\text{-}120}$ crystal structure were also poorly defined; however, the residues corresponding to 22–23 appear in an ordered helical region in the MLA10$^{5\text{-}120}$ crystal structure. Dihedral angles derived from the chemical shift analysis were used, together with distance restraints from ^{15}N- and ^{13}C-edited NOESY experiments, to calculate the 3D structure of the protein. The structure shows that the protein is folded into a four-helix bundle (Fig. 1*A* and *SI Appendix*, Fig. S1). The average root-mean-square distance (RMSD) for the amide backbone atoms (N, Cα, and C′) of residues 6–89 and 98–110 in the ensemble of the 20 lowest energy structures is 0.93 Å (*SI Appendix*, Table S1).

The EDVID motif is conserved within this class of CC domains and is important in mediating intramolecular domain interactions in full-length CC-NLR proteins (26, 28). In Sr33$^{6\text{-}120}$, the equivalent motif (residues 77–81) encodes residues EDAVD,

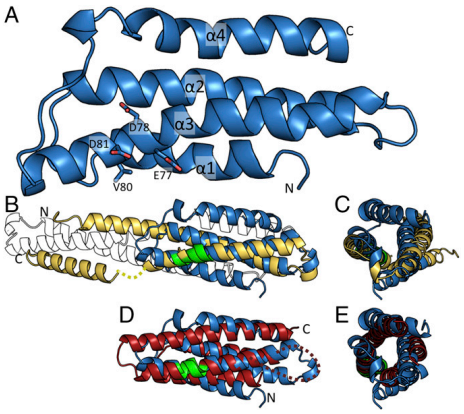


Fig. 1. Solution structure of Sr33 reveals a four-helix bundle fold. (*A*) NMR structure of Sr33$^{6\text{-}120}$ in cartoon representation, with the individual helices and N and C termini labeled. The conserved EDVID motif (EDAVD in Sr33) is shown in stick representation (colored green in *B* and *D*). (*B*) Superposition of the Sr33$^{6\text{-}120}$ structure (blue) and the crystal structure of MLA10$^{5\text{-}120}$ (yellow) in cartoon representation. Missing residues in the MLA10$^{5\text{-}120}$ structure (amino acids 91–95) are shown by a dotted line. The crystallographic dimer observed for MLA10$^{5\text{-}120}$ is shown as a black and white outline. (*C*) Superposition, as shown in *B*, rotated 90° around the *y* axis. (*D*) Superposition of the Sr33$^{6\text{-}120}$ structure (blue) and the crystal structure of Rx$^{1\text{-}122}$ (red) in cartoon representation. Missing residues in the Rx$^{1\text{-}122}$ structure (amino acids 40–50) are shown by a dotted line. (*E*) Superposition, as shown in *D*, rotated 90° around the *y* axis.

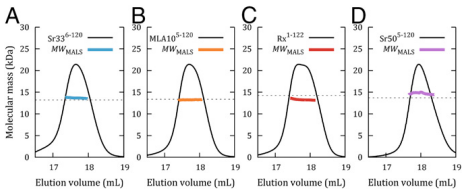


Fig. 2. Molecular mass calculations based on SEC-MALS analysis for $Sr33^{6-120}$ (*A*), $MLA10^{5-120}$ (*B*), Rx^{1-122} (*C*), and $Sr50^{5-123}$ (*D*). For all proteins, solid black lines represent the normalized refractive index trace (arbitrary units, *y* axis) for proteins eluted from an in-line Superdex 200 10/300 column. Colored lines under the peaks correspond to the averaged molecular mass (right-hand *y* axis) distributions across the peak as determined by MALS (MW_{MALS}). Dotted lines indicate the predicted molecular masses of a monomer. The average MW_{MALS} values compared with predicted monomeric molecular mass values are 13.7/13.1 kDa for $Sr33^{6-120}$, 13.3/13.4 kDa for $MLA10^{5-120}$, 13.3/14.3 kDa for Rx^{1-122}, and 14.7/14.1 kDa for $Sr50^{5-123}$.

which reside in the α3 helix, with residues E77, D78, V80, and D81 all surface-exposed (Fig. 1*A*). Despite the high sequence similarity between $Sr33^{6-120}$ and $MLA10^{5-120}$, these structures differ significantly and the solution structure of $Sr33^{6-120}$ resembles the structure of Rx^{1-122} more closely (Fig. 1 *B–E*).

Biophysical Characterization Shows That $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} Are Predominantly Monomeric and Adopt Compact, Globular Conformations. The NMR results suggested that $Sr33^{6-120}$ is monomeric in solution. This result differed from our expectations based on the crystal structure of $MLA10^{5-120}$. We investigated this finding further using in vitro biophysical techniques and also included the $MLA10^{5-120}$ and Rx^{1-122} constructs used in previous structure determination studies (14, 26), as well as $Sr50^{5-120}$, a rye ortholog of Sr33 (4), in our analysis. Size-exclusion chromatography (SEC)-coupled multiangle light scattering (MALS) experiments on these four proteins found an average molecular mass in solution very close to the predicted monomer sizes, indicating that they were all predominantly monomeric in solution, even at loading concentrations of 30 mg/mL (Fig. 2 *A–D*). This finding was corroborated by SEC-coupled small-angle X-ray scattering (SAXS) for $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} (*SI Appendix*, Fig. S2 *A–C* and Table S2), which yielded molecular masses between 12 and 14 kDa for $Sr33^{6-120}$ and between 12 and 16 kDa for $MLA10^{5-120}$ and Rx^{1-122}, varying with concentration across the elution. Averaged datasets from different fractions reflect this finding (*SI Appendix*, Fig. S2 *D–I*), but the data also suggested that the peak center fractions were suffering from some nonspecific aggregation (*SI Appendix*, Fig. S2 *G–I*), which is not present in the tail fractions. SAXS data also contain information about the shape of particles in solution. The scattering from the tail fractions of all three proteins is indistinguishable within experimental error (Fig. 3*A*), and their calculated properties (*SI Appendix*, Table S2) and real-space distributions also agree (Fig. 3*B*). Moreover, the experimental data are consistent with the predicted scattering of the four-helix bundle arrangement observed in the structures of $Sr33^{6-120}$ and Rx^{1-122}. They are not consistent with the structures of either the dimer or individual protomers from the $MLA10^{5-120}$ crystal structure (Fig. 3*C* and *SI Appendix*, Table S3). The same conclusion can be drawn if one attempts to superimpose the ab initio reconstructions from the SAXS data onto the corresponding high-resolution structures (Fig. 3*D*). All three datasets yield compact, globular-shaped envelopes, into which the NMR structure of $Sr33^{6-120}$ and the crystal structure of Rx^{1-122} can be docked within the proposed envelope. In contrast, the envelopes are clearly smaller than the extended conformation seen in the $MLA10^{5-120}$ crystals.

Maekawa et al. (14) found that the $MLA10^{5-120}$ protein coeluted from an analytical SEC column with a 25-kDa protein standard and also that treatment with the amine-to-amine cross-linker bis(sulfosuccinimidyl)suberate (BS3) resulted in covalent dimer formation. We repeated these experiments using our own purified $MLA10^{5-120}$ protein and observed similar results. $MLA10^{5-120}$ indeed elutes at a similar volume to the 25-kDa chymotrypsin marker during SEC (*SI Appendix*, Fig. S3). However, this behavior was also observed for Sr33, and for Rx, which had been previously found to be monomeric (26). We also observed a similar magnitude of cross-linked dimer to that observed by Maekawa et al. (14) after incubation with BS3 (*SI Appendix*, Fig. S3). In both experiments, the majority of the $MLA10^{5-120}$ protein remains monomeric even after 2 h of incubation with BS3. In contrast, the flax-rust AvrM effector protein, which forms a stable dimer in solution (32), was immediately cross-linked as a dimer on the addition of BS3 (*SI Appendix*, Fig. S3). Thus, these data do not support the formation of a stable $MLA10^{5-120}$ dimer in solution. Collectively, biophysical analysis shows that the monomer is the predominant species in solution for $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122}, and that this monomer is compact, globular, and consistent with four-helix bundle structures.

Crystal Structure of $MLA10^{5-120}$. The crystal structure of $MLA10^{5-120}$ reported by Maekawa et al. (14) shows a helix–loop–helix structure that forms a dimer through a large interface. The published $MLA10^{5-120}$ crystal structure was obtained at low pH (4.6) and high salt (2.0 M sodium formate). To investigate whether dimer formation is dependent upon the crystallization conditions, we attempted crystallization of $MLA10^{5-120}$ using alternative precipitants and

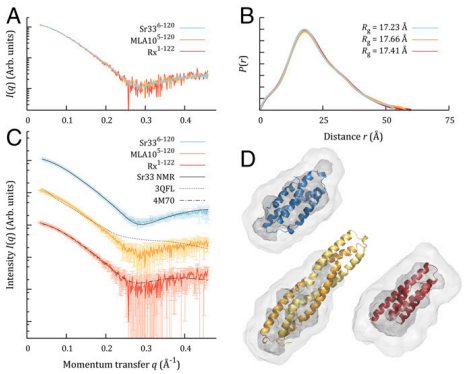


Fig. 3. SAXS data from monomeric fractions of $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} are consistent with compact, globular particles. (*A*) Datasets from SEC-SAXS are shown as colored lines, with the $MLA10^{5-120}$ and Rx^{1-122} data scaled to overlay with the $Sr33^{6-120}$ data. Arb., Arbitrary. (*B*) Normalized distance distribution functions, *P*(*r*), are shown as colored lines matching the scattering curve from which they were calculated. All distributions have been scaled to the maxima of the highest peak. (*C*) SEC-SAXS datasets again plotted as colored lines, now arbitrarily offset in *y* for clarity. Experimental errors are displayed at 1σ in lighter colors. The theoretical scattering predicted from each 3D structural model is shown as a black line against the corresponding dataset. (*D*) The first member of the $Sr33^{6-120}$ NMR ensemble (blue), the Rx^{1-122} crystal structure (red), and the dimeric $MLA10^{5-120}$ crystal structure (yellow) are shown in cartoon representation, docked into ab initio envelopes calculated from their respective scattering datasets. Ab initio models are shown in transparent surface representation, with the average model from 16 independent runs shown in light gray and the filtered model shown in darker gray.

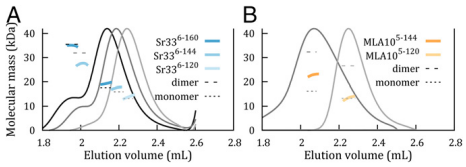


Fig. 4. Solution studies of CC domains with extended sequences of Sr33 and MLA10. Molecular mass calculations from SEC-MALS analysis for $Sr33^{6\text{-}120}$, $Sr33^{6\text{-}144}$, and $Sr33^{6\text{-}160}$ (*A*) and $MLA10^{5\text{-}120}$ and $MLA10^{5\text{-}144}$ (*B*). Solid gray, dark gray, and black lines represent the refractive index for the three proteins, respectively, when eluted from an in-line Superdex 200 5/150 GL column, normalized to the height of the major peak for clarity. Dotted lines indicate the predicted molecular masses of both monomeric and dimeric species, and colored lines show the experimental molecular mass distributions as determined by MALS (values are shown in *SI Appendix*, Table S5).

neutral pH. We found that $MLA10^{5\text{-}120}$ crystallized readily in 25% (wt/vol) polyethylene glycol 1500 at pH 7.0. These crystals diffracted X-rays to ∼2.0-Å resolution, and the structure was solved using single-wavelength anomalous diffraction (*SI Appendix*, Table S4).

The resulting structure resembles the structure described previously for $MLA10^{5\text{-}120}$ (14), having a Cα RMSD of 3.7 Å from 110 residues (*SI Appendix*, Fig. S4), and it forms a similar crystallographic dimer. We observed electron density for all residues (residues 5–120) in the best chain, including the previously undefined residues 91–95. However, there are notable differences between the two structures, particularly with respect to the interactions between the monomers forming the crystallographic dimer. In some regions, the interacting residues in the two protomers are offset by approximately one helical turn relative to the previously published crystallographic dimer (*SI Appendix*, Fig. S4). Comparison of the dimeric interface using PISA (Protein, Interfaces, Structures, and Assemblies) (33) shows that the structure presented here includes a more extensive hydrogen-bonding network between complementary residues, suggesting that the conformational dynamics of $MLA10^{5\text{-}120}$ permit different structural rearrangements during crystallization.

Extended CC Domain Fragments of Sr33 and MLA10 CC Domains Show an Increased Propensity to Self-Associate. Recent *in planta* results indicate that the residues between 120 and 160 are required for signaling and self-association of the Sr33, Sr50, and MLA10 N-terminal domains (25). Secondary-structure predictions (34) of these proteins predict a helix extending from residue 98 to residue 138 (*SI Appendix*, Fig. S5), which would be truncated in the fragments used for structure determination. Using a modified purification buffer system (*SI Appendix*, *SI Methods*), we expressed and purified $Sr33^{6\text{-}144}$, $Sr33^{6\text{-}160}$, and $MLA10^{5\text{-}144}$ to homogeneity (*SI Appendix*, Fig. S5 *B* and *C*). SEC-MALS revealed increased self-association in these longer CC domain fragments (Fig. 4). An earlier peak with a molecular mass near the molecular mass expected for the dimer was apparent for $Sr33^{6\text{-}144}$ and $Sr33^{6\text{-}160}$, while a larger peak at the position expected of the monomer was still present for both. $MLA10^{5\text{-}144}$ eluted in a single peak that was more extended and asymmetrical than the single peak of $MLA10^{5\text{-}120}$, with a molecular mass 40% higher than the molecular mass expected for the monomer, indicating a polydisperse population of molecules (Fig. 4 and *SI Appendix*, Table S5). These experiments show that the additional residues promote self-association in both Sr33 and MLA10.

Defining the Minimal CC-Domain Signaling Unit in Sr33, MLA10, and Sr50. To identify the minimal N-terminal fragment necessary for the signaling function of these proteins, we generated six truncations of the MLA10, Sr33, and Sr50 N-terminal domains at positions surrounding the predicted end of the α4 helix (Fig. 5*A*). These truncations were transiently expressed in *Nicotiana benthamiana* under the control of the 35S promoter and fused to a C-terminal HA tag. Fragments truncated at, or beyond, the equivalent of MLA10 residue 142 induced cell death similar to the autoactive 1–160 fragments (16, 25), whereas shorter fragments were inactive (Fig. 5*A*). Immunoblotting showed that the proteins were stable and accumulated to similar levels (*SI Appendix*, Fig. S6). These results demonstrate that the minimal N-terminal cell death signaling domains in MLA10, Sr33, and Sr50 extend to the amino acid position equivalent to 142 in MLA10 (*SI Appendix*, Fig. S5).

Autoactive Fragments of MLA10, Sr33, and Sr50 Self-Associate *in Planta*. To investigate if cell death induction was correlated with *in planta* self-association, we performed coimmunoprecipitation (co-IP) experiments using fragments of MLA10, Sr33, and Sr50 equivalent to $MLA10^{1\text{-}141}$, $MLA10^{1\text{-}142}$, and $MLA10^{1\text{-}144}$ fused to CFP or HA tags (Fig. 5*B*). The CFP-fused CC fragments were all expressed (*SI Appendix*, Fig. S6*B*) and displayed equivalent cell death activity as the corresponding HA-tagged fragments (*SI Appendix*, Fig. S6*C*). Expression of all proteins in the input was verified by immunoblotting using anti-GFP and anti-HA antibodies (Fig. 5*B*). CFP-fused proteins were enriched after immunoprecipitation with anti-GFP beads, and HA-fused autoactive fragments of MLA10, Sr33, and Sr50 coprecipitated with their respective CFP-fused fragment. This binding was specific, because they did not coprecipitate, or coprecipitated to a much weaker extent, with a divergent CC domain from the rice RGA4 protein (35) used as a negative control (Fig. 5*B*). In the case of the inactive CC domains ($MLA10^{1\text{-}141}$ and equivalents), these proteins showed much lower levels of self-association.

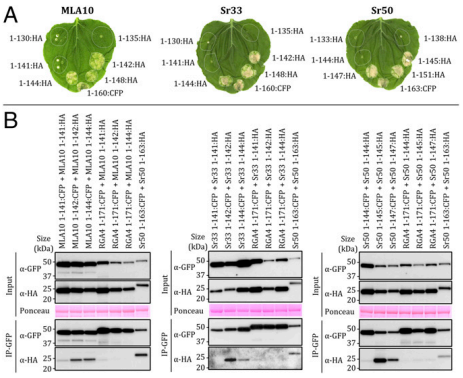


Fig. 5. Minimal autoactive domains of MLA10, Sr33, and Sr50 self-associate *in planta*. (*A*) MLA10, Sr33, and Sr50 protein fragments fused to HA or CFP were transiently expressed in *N. benthamiana*. The autoactive $MLA10^{1\text{-}160}$:CFP, $Sr33^{1\text{-}160}$:CFP, and $Sr50^{1\text{-}163}$:CFP constructs were used as positive controls. Cell death was visualized 5 d after infiltration. Equivalent results were obtained in three independent experiments. (*B*) Indicated proteins, transiently expressed in *N. benthamiana* leaves, were extracted 20 h after infiltration and analyzed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP) (Input). Proteins were immunoprecipitated with anti-GFP beads (IP-GFP) and analyzed by immunoblotting with anti-GFP and anti-HA antibodies. RGA4 (CC domain):CFP fusion was used as a control for specificity. $Sr50^{1\text{-}163}$ was used as a positive control (25). Ponceau staining of the RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase) large subunit shows equal protein loading.

Therefore, these data indicate a clear correlation between self-association and *in planta* signaling activity.

Discussion

Conservation of Plant NLR CC Domain Structure. Before this study, the structures of two plant CC domain fragments had been reported. The structure of the CC domain from the potato NLR protein Rx, a four-helix bundle, is "strikingly different" (26) from the structure of the CC domain from the barley NLR protein MLA10, a helix–turn–helix dimer (14). Given the low sequence identity between MLA10 and Rx within the structured region (~18%) and the fact that Rx^{1-122} was crystallized bound to the RanGAP2 WPP domain, it has been unclear whether these structural differences may represent divergent types of CC domains or different conformational states. Hao et al. (26) noted that the identification of any representative structure for the class would require characterization of further CC_{EDVID} structures. Interestingly, such variation has not been observed between structures of TIR domains, the other class of plant NLR N-terminal domains, which have a conserved fold despite low sequence identities (36).

The NMR structure of the CC region in $Sr33^{6-120}$, solved here, reveals a compact, four-helix bundle similar to the crystal structure of the Rx^{1-122} (26), rather than the dimeric arrangement in the crystal structure of $MLA10^{5-120}$ (14), despite the much lower sequence identity (18% versus 82%). Furthermore, biophysical characterization by SEC-MALS showed that these regions of Sr33, MLA10, Rx, and Sr50 were predominantly monomeric in solution (Fig. 2). Similarly, $Sr33^{6-120}$, $MLA10^{5-120}$, and Rx^{1-122} were indistinguishable by SEC-SAXS shape analysis, and these data were consistent with the four-helix bundle structures of the $Sr33^{6-120}$ NMR structure and Rx^{1-122} crystal structure, but not with the dimeric $MLA10^{5-120}$ crystal structure (Fig. 3). These results indicate that plant NLR CC domains likely have conserved structures, as is the case for TIR domains.

Self-Association and CC Domain Signaling. Our findings that neither $Sr33^{6-120}$ nor $MLA10^{5-120}$ self-associate in solution are consistent with recent observations that constructs of Sr33, Sr50, and MLA10 comprising residues 1–120 (or equivalent) of the CC domain do not self-associate or induce a cell death phenotype when transiently expressed in tobacco (25). In contrast, longer constructs comprising residues 1–160 are capable of both *in planta* self-association (based on co-IP) and cell death activity (25). Building upon these observations, we found that the minimal functional unit for cell death signaling of these proteins extends to a position slightly C-terminal to the predicted end of the last α-helix within the CC domain (Fig. 5). Co-IP experiments demonstrate that CC domain fragments that are capable of causing cell death can also self-associate when expressed *in planta*, whereas inactive fragments displayed strongly impaired self-association. Solution studies using these longer active constructs showed that the inclusion of additional residues at the C terminus ($MLA10^{5-144}$, $Sr33^{6-144}$, and $Sr33^{6-160}$) also facilitates self-association in vitro (Fig. 4), corroborating the link between self-association and signaling. It is, however, important to note that even for the longest constructs, both monomer and dimer forms of these proteins are present. Collectively, these results establish a correlation between self-association and biological activity, and suggest that CC domain self-association is the switch regulating cell death induction.

The CC Dimer. Given these findings, it is important to address the tendency of $MLA10^{5-120}$ to crystallize in a dimeric conformation, as well as the previously presented evidence (14) for such a dimer in solution. To interrogate the behavior of this protein in solution, it is necessary both to confirm previous experiments and to relate these experiments to the biophysical results. In their study, Maekawa et al. (14) suggested that $MLA10^{5-120}$ existed as a dimer in solution based on the crystal structure, the slow migration of the protein by SEC, and the appearance of cross-linked bands following incubation with a chemical cross-linking agent, and the fact that the recombinant protein was no longer soluble after the putative dimer interface was disrupted by mutations (14).

We also observed that $MLA10^{5-120}$, Rx^{1-122} [reported monomer (26)], and $Sr33^{6-120}$ have similar SEC elution times and that these elution times are consistent with the migration rate of a protein approximately double their monomeric molecular masses. Although migration by SEC is often used to estimate molecular masses by comparison with known protein standards, migration rates do not depend solely on molecular mass. The rate of migration is also influenced by particle shape, flexibility, composition, and rates of association and disassociation (37). In contrast, scattering techniques, such as MALS and SAXS, directly relate signal to average mass. These approaches provide a more robust and quantitative means of molecular mass measurement than migration by SEC (37, 38). In the case of $MLA10^{5-120}$, these methods show only a small degree of transient self-association at most. Consistent with these data, $MLA10^{5-120}$ remains predominantly monomeric even after extended incubation with a cross-linking agent. In contrast, the dimeric AvrM protein is rapidly and efficiently cross-linked (*SI Appendix*, Fig. S3). Chemical cross-linking can be promoted by transient or even nonspecific interactions, and we suggest that the observed behavior is more indicative of weak reversible self-association as opposed to obligate dimer formation.

The process of crystallization selects for states that promote a highly ordered arrangement in the crystal, not necessarily the most prevalent conformation in solution. The comparison of the structure of the $MLA10^{5-120}$ dimer observed in the crystals to the four-helix bundle structures suggests that the monomers have undergone a domain swap during crystallization. *SI Appendix*, Fig. S7 shows that it is possible to superimpose two copies of the $Sr33^{6-120}$ four-helix bundle side-by-side onto the MLA10 dimer. Repositioning α1 and α4 in $Sr33^{6-120}$ to form a continuous helix would regenerate the helix–loop–helix observed in the $MLA10^{5-120}$ dimer, while retaining the internal hydrophobic contacts. Importantly, this analysis also explains the insolubility of the interface-disrupting mutants reported by Maekawa et al. (14), because the residues that form the dimeric interface in the crystal also form the hydrophobic core of the four-helix bundle monomer, and mutation of these residues may thus be expected to disrupt the protein fold (*SI Appendix*, Fig. S8).

Our results demonstrate that the region between residues 120 and 142 in both Sr33 and MLA10 is necessary for both self-association and signaling activity, and suggest that the dimeric form is responsible for CC domain signaling. However, the structure and nature of this dimer remain unknown. It remains plausible that the $MLA10^{5-120}$ crystal structures capture some part of the activated signaling dimer. It is noteworthy that domain swapping is common among signaling proteins (39); however, it is not trivial to reconcile this conformation with the importance of the additional C-terminal residues. The C termini are at opposing ends of the rod-shaped dimer, and an extension of the C-terminal helices would project in opposite directions. In the event that these regions folded back onto the body of the domain-swapped CC domain, the C-terminal extensions would be unlikely to interact directly. However, it is possible that these regions may further stabilize a domain-swapped CC domain dimer through interactions with other regions of the protein (*SI Appendix*, Fig. S9). It is equally plausible that the MLA10 conformation observed in the crystal structures is a product of the crystallization process and is not biologically relevant. In this case, the additional C-terminal residues may promote a thus far uncharacterized self-association event between the monomers in the four-helix bundle conformation. Ultimately, a full structural

exploration of the longer, active constructs will be necessary to reveal the structural basis of the self-association and its role in CC-domain signaling.

Mechanism of Signaling by NLR CC Domains. It has been hypothesized that the self-association of the TIR domains postactivation is positively regulated via self-association of other domains from the full-length plant NLR (40). Structural studies demonstrate self-association of NB and LRR domains in animal NLRs (19–21), and this self-association has been shown by co-IP experiments for the *Arabidopsis* CC-NLR RPS5 (24) and TIR-NLR RPP1 (23). We propose a similar model of signaling for the Sr33, Sr50, and MLA10 CC-NLRs, in which the transient self-association of the CC domain is stabilized by the full-length NLR to achieve the activated state. These associations would presumably facilitate the recruitment of downstream signaling molecules, as is the case in animal NLRs (19–21, 41) and Toll-like receptors (42), and resemble the mechanism proposed for TIR-NLRs (5, 8).

We demonstrate that both closely and distantly related CC-NLR proteins have structurally similar CC domains, reconciling previously conflicting data and models of activation of this important domain. We show that self-association of the CC domains correlates strongly with cell death activity in the MLA10 and Sr33 CC-NLR proteins, and we define residues comprising the minimal functional unit for these proteins (both biophysically and *in planta*). Our data redefine the structural understanding of the CC domains from CC-NLR proteins. This work will provide a foundation for further structural studies of the more complete, signaling-competent NLR CC domain.

Materials and Methods

Details of the methods used are provided in *SI Appendix, SI Methods*, including cloning, protein production and crystallization, structure determination (NMR spectroscopy and X-ray crystallography), and SEC-MALS and SEC-SAXS experiments; constructs for in planta analyses; transient protein expression; and co-IP and cell death assays in *N. benthamiana*. Primers and construct details are listed in *SI Appendix*, Tables S6 and S7.

ACKNOWLEDGMENTS. We thank the Australian Synchrotron scientists for their support and Sambasivam Periyannan and Evans Lagudah for access to the Sr33 cDNA. We acknowledge the University of Queensland Remote Operation Crystallization and X-ray Diffraction Facility, the Queensland NMR Network, and the SAXS-wide-angle X-ray scattering and MX Beamlines at the Australian Synchrotron. This research was supported by Australian Research Council (ARC) Grants DP120100685 and DP160102244, National Health and Medical Research Council Research Fellow Grants 1003325 and 1110971 (to B.K.), ARC Future Fellow Grant FTI10100925 (to M.M.), and ARC Discovery Early Career Researcher Award DE160100893 (to S.J.W.). S.C. is the recipient of Commonwealth Scientific and Industrial Research Organization Postdoctoral Fellowship.

1. Dodds PN, Rathjen JP (2010) Plant immunity: Towards an integrated view of plant-pathogen interactions. *Nat Rev Genet* 11(8):539–548.
2. Periyannan S, et al. (2013) The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science* 341(6147):786–788.
3. Seeholzer S, et al. (2010) Diversity at the Mla powdery mildew resistance locus from cultivated barley reveals sites of positive selection. *Mol Plant Microbe Interact* 23(4):497–509.
4. Mago R, et al. (2015) The wheat Sr50 gene reveals rich diversity at a cereal disease resistance locus. *Nat Plants* 1:15186.
5. Bentham A, Burdett H, Anderson PA, Williams SJ, Kobe B (August 25, 2016) Animal NLRs provide structural insights into plant NLR function. *Ann Bot*, 10.1093/aob/mcw171.
6. Bernoux M, et al. (2011) Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host Microbe* 9(3):200–211.
7. Williams SJ, et al. (2011) An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Mol Plant Microbe Interact* 24(8):897–906.
8. Bernoux M, et al. (2016) Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. *Plant Cell* 28(1):146–159.
9. Dodds PN, Lawrence GJ, Ellis JG (2001) Six amino acid changes confined to the leucine-rich repeat β-strand/β-turn motif determine the difference between the P and P2 rust resistance specificities in flax. *Plant Cell* 13(1):163–178.
10. Ravensdale M, et al. (2012) Intramolecular interaction influences binding of the flax L5 and L6 resistance proteins to their AvrL567 ligands. *PLoS Pathog* 8(11):e1003004.
11. Krasileva KV, Dahlbeck D, Staskawicz BJ (2010) Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell* 22(7):2444–2458.
12. Hu Z, et al. (2013) Crystal structure of NLRC4 reveals its autoinhibition mechanism. *Science* 341(6142):172–175.
13. Swiderski MR, Birker D, Jones JD (2009) The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. *Mol Plant Microbe Interact* 22(2):157–165.
14. Maekawa T, et al. (2011) Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host Microbe* 9(3):187–199.
15. Collier SM, Hamel LP, Moffett P (2011) Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Mol Plant Microbe Interact* 24(8):918–931.
16. Bai S, et al. (2012) Structure-function analysis of barley NLR immune receptor MLA10 reveals its cell compartment specific activity in cell death and disease resistance. *PLoS Pathog* 8(6):e1002752.
17. Williams SJ, et al. (2014) Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* 344(6181):299–303.
18. Wang G-F, Balint-Kurti PJ (2015) Cytoplasmic and nuclear localizations are important for the hypersensitive response conferred by maize autoactive Rp1-D21 protein. *Mol Plant Microbe Interact* 28(9):1023–1031.
19. Zhang L, et al. (2015) Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization. *Science* 350(6259):404–409.
20. Hu Z, et al. (2015) Structural and biochemical basis for induced self-propagation of NLRC4. *Science* 350(6259):399–404.
21. Diebolder CA, Halff EF, Koster AJ, Huizinga EG, Koning RI (2015) Cryoelectron tomography of the NAIP5/NLRC4 inflammasome: Implications for NLR activation. *Structure* 23(12):2349–2357.
22. Mestre P, Baulcombe DC (2006) Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *Plant Cell* 18(2):491–501.
23. Schreiber KJ, Bentham A, Williams SJ, Kobe B, Staskawicz BJ (2016) Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. *PLoS Pathog* 12(7):e1005769.
24. Ade J, DeYoung BJ, Golstein C, Innes RW (2007) Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proc Natl Acad Sci USA* 104(7):2531–2536.
25. Cesari S, et al. (2016) Cytosolic activation of cell death and stem rust resistance by cereal MLA-family CC-NLR proteins. *Proc Natl Acad Sci USA* 113(36):10204–10209.
26. Hao W, Collier SM, Moffett P, Chai J (2013) Structural basis for the interaction between the potato virus X resistance protein (Rx) and its cofactor Ran GTPase-activating protein 2 (RanGAP2). *J Biol Chem* 288(50):35868–35876.
27. Maqbool A, et al. (2015) Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. *eLife* 4:e08709.
28. Rairdan GJ, et al. (2008) The coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signaling. *Plant Cell* 20(3):739–751.
29. Collier SM, Moffett P (2009) NB-LRRs work a "bait and switch" on pathogens. *Trends Plant Sci* 14(10):521–529.
30. Anglister J, Grzesiek S, Ren H, Klee CB, Bax A (1993) Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *J Biomol NMR* 3(1):121–126.
31. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
32. Ve T, et al. (2013) Structures of the flax-rust effector AvrM reveal insights into the molecular basis of plant-cell entry and effector-triggered immunity. *Proc Natl Acad Sci USA* 110(43):17594–17599.
33. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.
34. Buchan DWA, Minneci F, Nugent TC, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res* 41(Web Server issue): W349–W357.
35. Césari S, et al. (2014) The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. *EMBO J* 33(17):1941–1959.
36. Ve T, Williams SJ, Kobe B (2015) Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis* 20(2):250–261.
37. Wen J, Arakawa T, Philo JS (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal Biochem* 240(2):155–166.
38. Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496(7446):477–481.
39. Liu Y, Eisenberg D (2002) 3D domain swapping: as domains continue to swap. *Protein Sci* 11(6):1285–1299.
40. Bernoux M, Ellis JG, Dodds PN (2011) New insights in plant immunity signaling activation. *Curr Opin Plant Biol* 14(5):512–518.
41. Lu A, et al. (2014) Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell* 156(6):1193–1206.
42. Lin SC, Lo YC, Wu H (2010) Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature* 465(7300):885–890.

SI-Appendix

The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins

Lachlan W. Casey[a,1], Peter Lavrencic[a,b,1], Adam Bentham[a,c,1], Stella Cesari[d], Daniel J. Ericsson[e], Tristan I. Croll[f], Dusan Turk[g], Peter A. Anderson[c], Alan E. Mark[a], Peter N. Dodds[d], Mehdi Mobli[b,2], Bostjan Kobe[a,2] and Simon J. Williams[a,c,h,2]

[a]School of Chemistry and Molecular Biosciences, Institute for Molecular Bioscience and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Queensland 4072, Australia. [b]Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland 4072, Australia. [c]School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia. [d]CSIRO Agriculture, Canberra ACT 2601, Australia. [e]Macromolecular Crystallography Beamlines, Australian Synchrotron, Melbourne, Victoria 3168, Australia. [f]School of Biomedical Sciences, Queensland University of Technology, Brisbane, Queensland 4001, Australia. [g]Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins; International Postgraduate School Jozef Stefan; Jozef Stefan Institute, 1000 Ljubljana, Slovenia. [h]Plant Sciences Division, Research School of Biology, The Australian National University, Canberra 2601, Australia.

[1]These authors contributed equally to this work

[2]To whom correspondence may be addressed**.** Email: s.williams8@uq.edu.au, b.kobe@uq.edu.au, or m.mobli@uq.edu.au.

SI Methods

Cloning, expression and purification.

The cDNAs coding for the proteins under study were cloned into the pMCSG7 vector by ligation-independent cloning (LIC). Primers designed for LIC consisted of gene-specific sequence, flanked by LIC overhangs to facilitate cloning into expression vectors. Details of primers and constructs used in cloning and expression are given in Tables S6 and S7. For Rx^{1-122}, a gBlock® of the codon optimised (*E. coli* expression) CC fragment with LIC sites was ordered from Integrated DNA Technologies (IDT) and cloned into expression vector pMCSG7.

For biophysical studies, the proteins were expressed in *Escherichia coli* BL21 (DE3) at 20˚C, using the autoinduction expression. Cells were lysed via sonication in the lysis buffer (consisting of 50 mM HEPES pH 8.0, 300 mM NaCl, and 1 mM dithiothreitol [DTT]) for $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122}. A similar lysis buffer was used for the longer CC domain fragments (corresponding to $Sr33^{6-144}$, $Sr33^{6-160}$, $MLA10^{5-144}$ and $MLA10^{5-160}$); however, the pH was adjusted to 7.5 and 500 mM of NaCl was used. The proteins were separated from clarified cell lysate via immobilized metal affinity chromatography (IMAC), facilitated by N-terminal 6 x histidine tags. Proteins were eluted from the IMAC column using elution buffer (consisting of 50 mM HEPES pH 7.5 and 8 [protein-dependent], 250 mM NaCl, and 250 mM imidazole). Post elution, excess imidazole was removed via buffer exchange, and proteins were maintained in a buffer consisting of 50 mM Tris-HCl pH 8.0, 250 mM NaCl and 1 mM DTT. Overnight treatment with TEV (tobacco etch virus) protease at 20˚C was used to remove the histidine tag, leaving a three-residue N-terminal overhang (Ser-Asp-Ala). SDS-PAGE analysis of proteins was used to follow the purification and removal of the histidine tag. The cleaved protein was re-applied to the nickel affinity chromatography column to remove the histidine tagged TEV protease and other contaminants. The proteins were further purified using a Superdex 75 HiLoad 26/60 size-exclusion chromatography (SEC) column (GE Healthcare) equilibrated with 10 mM HEPES pH 8.0, 150 mM NaCl and 1 mM DTT. Amicon® Ultra centrifugal filters (15 mL) (Merck Millipore) were used to concentrate proteins to appropriate concentrations for biophysical analysis, post-SEC.

We experienced considerable difficulties with expression and purification of $Sr50^{6-123}$. After much effort and optimisation we were able to obtain quantities of $Sr50^{6-123}$ that facilitated SEC-MALS analysis. This was achieved when using lysis, wash and elution buffers consisting of 1 M NaCl and 50 mM HEPES pH 8.5. Despite this, we still observed

significant protein loss during chromatography and concentration steps. These issues precluded $Sr50^{6-153}$ from further analysis using SEC-SAXS and structural studies.

Protein expression for NMR spectroscopy

E. coli BL21 cells expressing the $Sr33^{6-120}$ protein (see above) were grown in M9 minimal media containing ^{13}C-labelled glucose, and ^{15}N-labelled ammonium chloride. Protein expression was induced using 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) at 20˚C for overnight protein expression. The $^{13}C/^{15}N$-labelled $Sr33^{6-120}$ protein was purified using nickel affinity and size-exclusion chromatography as described above.

The correlation time of the protein was estimated based on transverse relaxation rates (T_2), measured as described previously (1). The correlation time was converted to a molecular mass using the Stoke-Einstein equations as described in (2), using a modified equation for estimation of protein volumes according to (3) with the addition of 2 Å to account for the hydration shell.

NMR data acquisition

The $^{13}C/^{15}N$-labelled $Sr33^{6-120}$ sample containing 5% D_2O was filtered using a low-protein-binding Ultrafree-MC centrifugal filter (0.22 μm pore size; Millipore, MA, USA), then 300 μL was added to a susceptibility-matched 5 mm outer-diameter microtube (Shigemi Inc., Japan).

NMR data were acquired at 25˚C using a 900 MHz AVANCE spectrometer (Bruker BioSpin, Germany) equipped with a cryogenically cooled probe. Data used for resonance assignment were acquired using non-uniform sampling (NUS); sampling schedules that approximated the rate of signal decay along the various indirect dimensions were generated using sched3D (4). The decay rates used were 1 Hz for all constant-time ^{15}N dimensions, 30 Hz for all ^{13}C dimensions, and 15 Hz for the semi-constant indirect ^{1}H dimension. ^{13}C- and ^{15}N-edited HSQC-NOESY experiments were acquired using linear sampling. Separate experiments were acquired for the aliphatic and aromatic regions of the ^{13}C dimension.

NUS data were processed using the Rowland NMR toolkit (www.rowland.org/rnmrtk/toolkit.html); maximum entropy parameters were selected automatically as described previously (5). NMR spectra were analyzed and assigned using the program CcpNmr (6). ^{1}HN, ^{15}N, ^{13}C backbone resonance assignments were obtained from the analysis of amide-proton strips in 3D HNCACB, CBCA(CO)NH, and HNCO spectra. Sidechain ^{1}H and ^{13}C chemical shifts were obtained primarily from 3D

H(CC)(CO)NH-TOCSY and (H)CC(CO)NH-TOCSY spectra, respectively. The remaining side-chain assignments were derived from 3D H(C)CH-TOCSY and ^{15}N- and ^{13}C-edited NOESY-HSQC spectra

NMR structure determination

Distance restraints for structure calculations were derived from 3D ^{13}C- and ^{15}N-edited NOESY-HSQC spectra acquired with a mixing time of 120 ms. NOESY spectra were manually peak-picked and integrated using the box-sum method in CcpNmr. The peak lists were then assigned and an ensemble of structures calculated automatically using the torsion angle dynamics package CYANA (7). The tolerances used in the structure calculations were 0.03 ppm in the indirect ^{1}H dimension, 0.02 ppm in the direct ^{1}H dimension, 0.2 ppm for the aromatic ^{13}C and ^{15}N dimensions, and 0.4 ppm for the aliphatic ^{13}C data.

Backbone dihedral-angle restraints (112 for both ϕ and ψ) were derived from TALOS+ chemical shift analysis (8); the restraint range was set to twice the estimated standard deviation. All X-Pro peptide bonds were clearly identified as *trans* on the basis of characteristic NOEs and the C_β and C_γ chemical shifts for the Pro residues.

CYANA was used to calculate 200 structures from random starting conformations, then the 20 conformers with the lowest CYANA target function were chosen to represent the structural ensemble. During the automated NOESY assignment/structure calculation process CYANA assigned 94.4% of all NOESY crosspeaks (3186 out of 3372) for Sr33.

Analytical size-exclusion chromatography (SEC) and cross-linking

The purified MLA10$^{5\text{-}120}$, Sr33$^{6\text{-}120}$ and Rx$^{1\text{-}122}$ protein (450 μg) was separated on a Superdex 75 10/300 GL SEC column with a mobile phase consisting of 10 mM HEPES pH 7.5 and 150 mM NaCl. Protein size markers chymotrypsin (25 kDa) and cytochrome c (15 kDa) were separated using the same conditions as for MLA10$^{5\text{-}120}$. Cross-linking experiments were performed as described in (9). In brief, 20 μL of MLA10$^{5\text{-}120}$ (in the SEC buffer) at a concentration of 150 μM was mixed with 5 μL of BS3 (bis(sulfosuccinimidyl)suberate) at a concentration of 20 mM. The reaction was incubated on ice and monitored at various time points from 0-120 minutes. The reaction was quenched with equal volumes of 1 M Tris pH 7.5, before the samples were separated using 13% SDS-PAGE.

Size-exclusion chromatography (SEC)-coupled multi-angle light scattering (MALS)

SEC-MALS was performed using an in-line Superdex 200 100/300 GL or Superdex 200 Increase 5/150 GL SEC column (GE Healthcare) combined with a Dawn Heleos II 18-angle light-scattering detector coupled with an Optilab TrEX refractive index detector (Wyatt Technology, Santa Barbara, CA, USA). Purified proteins were separated at 0.5 mL/min (10/300) or 0.25 mL/min (5/150) in 10 mM HEPES pH 8.0 and 150 mM NaCl. Molecular-mass calculations were performed using the Astra6.1 software (Wyatt Technology). Input of the refractive increment (dn/dc values) was set at 0.186 in the molecular-mass calculations, based on the premise that dn/dc is constant for unmodified proteins (10). The molecular mass was determined across the protein elution peak.

Size-exclusion chromatography (SEC)-coupled small-angle X-ray scattering (SAXS)

SEC-SAXS was performed during two shifts at the SAXS/WAXS beamline of the Australian Synchrotron on a Pilatus 1M detector, using an in-line WTC-030S5 SEC column and a 2 mL WTC-030S5G pre-column (Wyatt Technology), together with a Prominence modular HPLC system (Shimadzu Scientific Instruments). All experiments were conducted at 16°C using 10 mM HEPES (pH 7.5), 150 mM NaCl buffer with 1 mM DTT. Eluate from the column was directed through a 1 mm quartz capillary mounted in the beam. For all samples, the injected volume was 95 µL at 30 mg/mL protein concentration, as determined by UV absorbance at 280 nm. High concentrations were used to maximize signal after dilution during gel-filtration, as the expected particle size is small.

The data for Sr33 was collected in 5 s exposures at 0.05 s intervals with a flow rate of 0.25 mL/min. A Wyatt WTC-030S5G pre-column was used upstream of the WTC-030S5. The sample-to-detector distance was 1.6 m, and a wavelength of 1.12713 Å yielded a range of momentum transfer ($0.009 < q < 0.478$ Å^{-1}, where $q = 4\pi.\sin(\theta)/\lambda$). The data for MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ were collected during a different shift, in 2 s exposures at 0.05 s intervals, with a flow rate of 0.5 mL/min. The WTC-030S5 without pre-column was used for these samples. A sample-to-detector distance of 1.4 m was used to obtain data over the range $0.010 < q < 0.614$ Å^{-1}.

Data reduction, normalisation and subtraction was performed using scatterBrain (http://www.synchrotron.org.au/index.php/aussyncbeamlines/saxswaxs/software-saxswaxs). Unless noted otherwise, subsequent analyses were performed using the tools in version 2.6 of the ATSAS program suite (11).

100 frames immediately preceding each peak were summed and normalized for exposure time to obtain buffer blanks. Initially, these buffers were subtracted from each

individual image to generate a series of subtracted frames across the elution peak, from which $I(0)$ and R_g were individually calculated using the Guinier approximation, as implemented in batch-mode AUTORG, for points such that $q.R_g$ <1.3. Molecular masses were calculated using a local high-throughput implementation of the volume of correlation (V_c) method developed by Rambo and Tainer (12), for points up to q = 0.3.

These metrics were evaluated for variation across the peak. To obtain the final scattering curves for analysis, the original images from elution ranges corresponding to the peak centre and the peak tail, were summed and normalized in scatterBrain, and then subtracted from the corresponding blank.

Guinier analysis and the determination of $I(0)$, R_g and MM_{Vc} were performed on the summed and averaged curves in the same manner as for individual frames. Data-points closer to the beamstop than the first Guinier point were discarded. Data points where q > 0.46 Å^{-1} were also discarded, due to poor signal-to-noise. Distance distributions, $P(r)$, were obtained by indirect transformation in GNOM (https://www.embl-hamburg.de/biosaxs/gnom.html), informed by AUTOGNOM. In addition to MM_{Vc}, molecular masses were also estimated from the Porod volume calculated by GNOM, using the empirical ratio developed by Pethoukhov and coworkers of $MM_{Porod} = V_{Porod}*0.625$ (11).

Theoretical scattering was calculated from atomic models using FoXS (13). Short stretches of residues not visible in the electron density of the published MLA10$^{5\text{-}120}$ crystal structure were added to both chains using the loop-building routines in MODELLER (14) independently from the SAXS data.

Crystallization and crystal structure determination of MLA10$^{5\text{-}120}$

Native and selenomethionine-labelled MLA10$^{5\text{-}120}$ protein at 10 mg/mL and 6 mg/mL, respectively, in 10 mM HEPES (pH 8.0), 100 mM NaCl, and 1 mM DTT were used in crystallization trials. Crystallization experiments were initially performed with native protein using hanging-drop vapour diffusion in 96-well plates. Several commercial screens were used, including Index, PEG/Ion and PEGRx (Hampton Research) and Pact Premier and JCSG+ (Molecular Dimensions). 100 nl protein solution and 100 nl well solution were prepared on hanging-drop seals (TTP4150-5100 sourced from Millennium Science, Australia) using a Mosquito robot (TTP Lab-Tech, UK) and equilibrated against 75 ml reservoir solution. The drops were monitored and imaged using the Rock Imager system (Formulatrix, USA). Numerous promising hits were observed within 24 hours; however, the crystals grown in Pact Premier, condition B4 (MIB buffer pH 7.0, 25% PEG 1500) were pursued for data collection. Crystals grown in

larger 1:1 μL (protein: well solution) drops were cryo-protected using the well-solution containing 20% glycerol prior to flash-cooling in liquid nitrogen. X-ray diffraction data of the native crystals were collected from a single crystal at the Australian Synchrotron MX2 beamline to ~2.0 Å resolution using a wavelength of 0.9537 Å. The crystal-to-detector distance was set to 200 mm and the oscillation range was 0.5°. Data collection was performed using Blu-Ice software, indexed and integrated using XDS (15) and scaled with AIMLESS within the CCP4 suite (16). With the native data-set molecular replacement was attempted using the published $MLA10^{5-120}$ structure (PDB ID 3QFL; (9)) in monomeric, dimeric and various truncated forms, as well as the structure of Rx^{1-122} (PDB ID 4M70; (17)); however, a solution could not be obtained. Subsequently, selenomethionine-labelled protein (confirmed by mass spectrometry) was crystallized as described for the native protein. X-ray diffraction data of selenomethionine-labelled crystals were collected from a single crystal at the Australian Synchrotron MX2 beamline to ~2.1 Å resolution using a wavelength 0.9792 Å. The crystal-to-detector distance was set to 200 mm and the oscillation range was 0.5°. Data collection was performed using Blu-Ice software, indexed and integrated using XDS (15) and scaled with AIMLESS within the CCP4 suite (16).

The crystals of $MLA10^{5-120}$ appeared to have the symmetry of the space group $P22_12_1$ and the structure was solved using single-wavelength anomalous diffraction (SAD) through the CRANK2 pipeline (18). Model building and refinement was done through cycles of Coot (19) and refinement in BUSTER-TNT (20). Refinement, however, proved unstable with BUSTER-TNT, unable to converge on a stable anisotropy ratio. Furthermore, R_{work} and R_{free} would stall at ~28% and ~30%, respectively. These factors could be improved by expanding the R_{free} test set from $P22_12_1$ to P1 space groups, and reprocessing the data to the P1 space group. In addition, we combined direct interactive modeling using interactive molecular dynamics flexible fitting (iMDFF) in VMD (21) and Phenix.refine (22) to generate the final model. Statistics for the refined atomic model are presented in Table S4.

Constructs for *in planta* analyses

Details of primers and constructs used in this study are given in Table S6 and S7. For transient expression in *N. benthamiana*, molecular cloning was performed by a combination of Quikchange site-directed mutagenesis (Agilent Technologies) and Gateway recombination (Life Technologies) as detailed in Table S7. The $MLA10_{1-160}$, $Sr33_{1-160}$ and $Sr50_{1-163}$ constructs cloned in pDONR207 (29) were used as templates for site-directed deletion to generate the $MLA10_{1-130}$, $MLA10_{1-135}$, $MLA10_{1-141}$, $MLA10_{1-142}$, $MLA10_{1-144}$, $MLA10_{1-148}$, $Sr33_{1-130}$, $Sr33_{1-135}$, $Sr33_{1-141}$, $Sr33_{1-142}$, $Sr33_{1-144}$, $Sr33_{1-148}$, $Sr50_{1-}$

$_{133}$, $Sr50_{1-138}$, $Sr50_{1-144}$, $Sr50_{1-145}$, $Sr50_{1-147}$ and $Sr50_{1-151}$ ENTRY constructs. These constructs were then recombined by LR reaction in the binary vector pBIN19-35S::GTW:3HA or pBIN19-35S::GTW:CFP by LR coning to obtain expression vectors.

Transient protein expression and cell death assays in *N. benthamiana*

N. benthamiana plants were grown in a growth chamber at 23°C with a 16 hours light period. For *N. benthamiana* leaf transformations, pBIN19-derived vector constructs were transformed into *Agrobacterium tumefaciens* strain GV3101_pMP90. Bacterial strains were grown in Luria-Bertani liquid medium containing 50 mg/ml rifampicin, 15 mg/ml gentamycin and 25 mg/ml kanamycin at 28°C for 24 hours. Bacteria were harvested by centrifugation, resuspended in infiltration medium (10 mM MES pH 5.6, 10 mM $MgCl_2$ and 150 µM acetosyringone) to an OD_{600nm} ranging from 0.5 to 1, and incubated for 2 hours at room temperature before leaf infiltration. Three leaves from two plants were infiltrated for each combination of constructs and the experiment was repeated three times independently. The infiltrated plants were incubated in growth chambers under controlled conditions for all following assays. For documentation of cell death, leaves were scanned five days after infiltration.

Protein extraction western blot and co-immunoprecipitation

Protein extraction, from *N. benthamiana* leaves and co-IP experiments were performed as described (23). For immunoblotting analysis, proteins were separated by SDS-PAGE and transferred to a nitrocellulose membrane. Membranes were blocked in 5% skimmed milk and probed with anti-HA-HRP antibodies (Roche) or anti-GFP antibodies (Roche) followed by goat anti-mouse antibodies conjugated with horseradish peroxidase (Pierce). Labeling was detected using the SuperSignal West Femto chemiluminescence kit (Pierce). Membranes were stained with Ponceau S to confirm equal loading.

Table S1. NMR structure statistics[a]

Experimental restraints[b]		
	Inter-proton distance restraints	
	Intra-residue	580
	Sequential	186
	Medium-range (i–j < 5)	293
	Long-range (i–j > 5)	249
	Dihedral-angle restraints	224
	Total number of restraints per residue	13.32
RMSD from the mean of the atomic coordinates of the ensemble (Å)[c]		
	Backbone atoms (residues 6–89 & 98–110)	0.93 ± 0.21
	All heavy atoms (residues 6–89 & 98–110)	1.31 ± 0.19
Stereochemical quality[d]		
	Residues in most favoured Ramachandran region (%)	93.1
	Ramachandran outliers (%)	0 ± 0
	Unfavourable side-chain rotamers (%)	0 ± 0
	Clashscore, all atoms[b]	0 ± 0

[a]All statistics are given as mean (some ± SD).

[b]Only structurally relevant restraints, as defined by CYANA, are included.

[c]Mean r.m.s. deviation calculated over the entire ensemble of 20 structures.

[d]As reported by CYANA (7, 24).

Table S2. Properties derived from averaged SAXS datasets

Protein	Fraction	Elution range (mL)	$I(0)_{\text{Guin}}$ (cm^{-1})	$I(0)_{P(r)}$ (cm^{-1})	$R_{g\ \text{Guin}}$ (Å)	$R_{g\ P(r)}$ (Å)	MM_{Vc} (kDa)	MM_{Porod} (kDa)
$Sr33^{6-120}$	Centre	11.77 – 12.31	4.12 e^{-2}	4.14 e^{-2}	18.84	19.55*	13.8	14.4
	Tail	12.35 – 12.69	1.28 e^{-2}	1.28 e^{-2}	17.08	17.23	13.5	11.4
$MLA10^{5-120}$	Centre	9.29 – 10.04	3.69 e^{-2}	3.76 e^{-2}	20.67	23.02*	16.0	15.6
	Tail	10.42 – 11.29	0.46 e^{-2}	0.46 e^{-2}	17.72	17.66	13.7	12.8
Rx^{1-122}	Centre	9.48 – 10.07	3.95 e^{-2}	4.01 e^{-2}	20.70	23.29*	16.3	15.8
	Tail	10.65 – 11.32	0.31 e^{-2}	0.31 e^{-2}	17.10	17.41	13.8	12.2

* Values for $R_{g\ P(r)}$ that differ from $R_{g\ \text{Guin}}$ by greater than 5%

The theoretical monomeric molecular masses of $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122} are 13.1 kDa, 13.4 kDa and 14.3 kDa, respectively.

Table S3. Goodness-of-fit (χ) scores for averaged SAXS datasets compared to structures

Sample	Fraction	Atomic structure (PDB ID)					
		$Sr33^{6-120}$	Rx^{1-122}	$MLA10^{5-120}$			
		NMR*	4M70	3QFL monomer	3QFL dimer	MX* monomer	MX* dimer
$Sr33^{5-120}$	Tail	0.67	1.63	12.03	8.41	11.54	8.08
$MLA10^{5-120}$	Tail	0.51	0.87	3.99	2.88	3.85	2.75
Rx^{1-122}	Tail	0.48	0.73	3.44	2.71	3.41	2.58

* Structure presented in this work.

Table S4. Crystallographic table for MLA10$^{5\text{-}120}$

Data processing		
Space group	P 2 2_12_1	P 1
a, b, c (Å)	30.87, 87.56, 92.56	30.72, 87.14, 92.25
α, β, γ (°)	90, 90, 90	89.93, 90.00, 89.98
Resolution (Å)	46.28-2.1 (2.16-2.10) [a]	46.12-2.05 (2.10-2.05)
Rmeas (%) [b]	11.0 (194.0)	6.8 (79.2)
Rpim(%) [c]	3.0 (51.0)	4.8 (56.0)
<I/σ(I)>	15.0 (1.8)	8.2 (1.4)
$CC_{1/2}$ [d]	0.99 (0.89)	0.99 (0.77)
Completeness (%)	100 (100)	96.5 (92.9)
Multiplicity	14.1 (14.4)	1.8 (1.8)
Wilson plot B (Å^2)	44.7	38.9
Observations	216711 (18084)	107102 (7880)
Unique reflections	15392 (1253)	58095 (4314)
Anomalous completeness	100 (100)	-
Anomalous multiplicity	7.7 (7.7)	-
DelAnom correlation between half-sets	0.471 (-0.027)	-
Mid-slope of anomalous normal probability	1.087	-
Estimate of maximum resolution for significant anomalous signal = 3.59 Å, from CCanom > 0.15		
Refinement		
Rwork (%)	27.9 (31.1)	25.2 (37.7)
Rfree (%)	30.1 (35.0)	27.9 (40.7)
Average B-factor (Å^2)	60.35	62.15
R.m.s deviations		
Bond lengths (Å)	0.009	0.001
Bond angles (°)	1.07	0.348
Ramachandran plot (%) [e]		
Favoured	96.41	99.32
Allowed	99.10	100.00
Outliers	0.90	0.00

[a] NB: Values within parentheses indicate the highest resolution bin.
[b] $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)\text{-}1]\}^{1/2} \sum_i|I_i(hkl)\text{- }<I(hkl)>|/ \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of an equivalent reflection with indices hkl.
[c] $R_{pim} = \sum_{hkl}\{1/[N(hkl)\text{-}1]\}^{1/2} \sum_i|I_i(hkl)\text{- }<I(hkl)>|/ \sum_{hkl}\sum_i I_i(hkl)$.
[d] Calculated with the program Aimless (25).
[e] As calculated by MolProbity (26).

Table S5. Summary of predicted and experimental average molecular masses for constructs as determined by MALS.

	Theoretical Molecular Mass (kDa)		Experimental Molecular Mass (kDa)	
Construct	Monomer	Dimer	Shoulder	Main peak
$Sr33^{6\text{-}120}$	13.12	16.24	-	13.9
$Sr33^{6\text{-}144}$	15.94	31.88	27.3	17.5
$Sr33^{6\text{-}160}$	17.65	35.30	35.0	19.3
$MLA10^{5\text{-}}$	13.28	26.56	-	13.4
$MLA10^{5\text{-}}$	16.17	32.34	-	22.8

Table S6. Primers used in this study for *in-planta* and *in vitro* studies

Primer name	Primer sequence 5'-3'
oCS281	GACATCCAAGAGCAACTCGACCCAGCTTTCTTGTAC
oCS282	GTACAAGAAAGCTGGGTCGAGTTGCTCTTGGATGTC
oCS283	GCAACTCCAAAAGGTGGCTGATGACCCAGCTTTCTTGTAC
oCS284	GTACAAGAAAGCTGGGTCATCAGCCACCTTTTGGAGTTGC
oCS285	GATAGGCGTGACAGGAACAAGGACCCAGCTTTCTTGTAC
oCS286	GTACAAGAAAGCTGGGTCCTTGTTCCTGTCACGCCTATC
oCS287	TAGGCGTGACAGGAACAAGGTAGACCCAGCTTTCTTGTAC
oCS288	GTACAAGAAAGCTGGGTCTACCTTGTTCCTGTCACGCCTA
oCS289	CGTGACAGGAACAAGGTATTTGTTGACCCAGCTTTCTTGTAC
oCS290	GTACAAGAAAGCTGGGTCAACAAATACCTTGTTCCTGTCACG
oCS291	GGTATTTGTTCCTCATCCTACGGACCCAGCTTTCTTGTAC
oCS292	GTACAAGAAAGCTGGGTCCGTAGGATGAGGAACAAATACC
oCS293	GACATCAAGAAGGAACTCGACCCAGCTTTCTTGTAC
oCS294	GTACAAGAAAGCTGGGTCGAGTTCCTTCTTGATGTC
oCS295	CTCCAGGAGGTGGCTGCTGACCCAGCTTTCTTGTAC
oCS296	GTACAAGAAAGCTGGGTCAGCAGCCACCTCCTGGAG
oCS297	CTAGGCGTGACAGGAACAAGGACCCAGCTTTCTTGTAC
oCS298	GTACAAGAAAGCTGGGTCCTTGTTCCTGTCACGCCTAG
oCS299	CGTGACAGGAACAAGTTCGACCCAGCTTTCTTGTAC
oCS300	GTACAAGAAAGCTGGGTCGAACTTGTTCCTGTCACG
oCS301	AGGAACAAGTTCGATGGTGACCCAGCTTTCTTGTAC
oCS302	GTACAAGAAAGCTGGGTCACCATCGAACTTGTTCCT
oCS303	GAACAAGTTCGATGGTATTGCTTCTATTGACCCAGCTTTCTTGTAC
oCS304	GTACAAGAAAGCTGGGTCAATAGAAGCAATACCATCGAACTTGTTC
oCS305	GAAATCAAGGAGCAACTCGACCCAGCTTTCTTGTAC
oCS306	GTACAAGAAAGCTGGGTCGAGTTGCTCCTTGATTTC
oCS307	CTCCAGGAGGTGGCTGCTGACCCAGCTTTCTTGTAC
oCS308	GTACAAGAAAGCTGGGTCAGCAGCCACCTCCTGGAG
oCS309	GCTAGGCGTGACAGGAACAAGGACCCAGCTTTCTTGTAC
oCS310	GTACAAGAAAGCTGGGTCCTTGTTCCTGTCACGCCTAGC
oCS311	TAGGCGTGACAGGAACAAGGTAGACCCAGCTTTCTTGTAC
oCS312	GTACAAGAAAGCTGGGTCTACCTTGTTCCTGTCACGCCTA
oCS313	CGTGACAGGAACAAGGTAGCTGTTGACCCAGCTTTCTTGTAC
oCS314	GTACAAGAAAGCTGGGTCAACAGCTACCTTGTTCCTGTCACG
oCS315	GCTGTTCCTAATCCTATGGACCCAGCTTTCTTGTAC
oCS316	GTACAAGAAAGCTGGGTCCATAGGATTAGGAACAGC
MLA10_5_FW	TACTTCCAATCCAATGCGACCGGTGCCATTTCCAACCTGATTCC
MLA10_120_RV	TATTCCACTTCCAATGTTAAGCTATCCCATGCTTATGCTTGACTTTCTTC
MLA10_144_RV	TATTCCACTTCCAATGTTAAACAAATACCTTGTTCCTGTCACGCCTATC
MLA10_160_RV	TATTCCACTTCCAATGTTACAAAGCTC
Sr50_5_FW	TACTTCCAATCCAATGCGACGGGGGCCATGG
Sr50_123_RV	TATTCCACTTCCAATGTTAAGCTATGCGATGGTGATTC
Sr33_6_FW	TACTTCCAATCCAATGCGACGGGTGCCA
Sr33_120_RV	TATTCCACTTCCAATGTTAAGCTATTC
Sr33_144_RV	TATTCCACTTCCAATGTTAACCATCGAACTTGTTCCTGTCACGCC
Sr33_160_RV	TATCCACTTCCAATGTTATAGAGCACGG

Table S7. Constructs used in this study for *in-planta* and *in vitro* analysis

Use	Construct	Plasmid name	Insert or PCR product	Primers	Template	Plasmid backbone	Cloning method	Reference
Entry clones for *N. benthamiana* assays	MLA10$^{1-160}$	pSC260	MLA10$^{1-160}$ (without stop)	/	/	pDONR207	/	29
	Sr33$^{1-160}$	pSC298	Sr33$^{1-160}$ (without stop)	/	/	pDONR207	/	29
	Sr50$^{1-163}$	pSC262	Sr50$^{1-163}$ (without stop)	/	/	pDONR207	/	29
	MLA10$^{1-130}$	pSC392	MLA10$^{1-130}$ (without stop)	oCS281/282	pSC260	pDONR207	Quikchange lightning (Agilent)	/
	MLA10$^{1-135}$	pSC393	MLA10$^{1-135}$ (without stop)	oCS283/284	pSC260	pDONR207	Quikchange lightning (Agilent)	/
	MLA10$^{1-141}$	pSC394	MLA10$^{1-141}$ (without stop)	oCS285/286	pSC260	pDONR207	Quikchange lightning (Agilent)	/
	MLA10$^{1-142}$	pSC395	MLA10$^{1-142}$ (without stop)	oCS287/288	pSC260	pDONR207	Quikchange lightning (Agilent)	/
	MLA10$^{1-144}$	pSC396	MLA10$^{1-144}$ (without stop)	oCS289/290	pSC260	pDONR207	Quikchange lightning (Agilent)	/
	MLA10$^{1-148}$	pSC397	MLA10$^{1-148}$ (without stop)	oCS291/292	pSC260	pDONR207	Quikchange lightning (Agilent)	/
	Sr33$^{1-130}$	pSC398	Sr33$^{1-130}$ (without stop)	oCS293/294	pSC298	pDONR207	Quikchange lightning (Agilent)	/
	Sr33$^{1-135}$	pSC399	Sr33$^{1-135}$ (without stop)	oCS295/296	pSC298	pDONR207	Quikchange lightning (Agilent)	/
	Sr33$^{1-141}$	pSC400	Sr33$^{1-141}$ (without stop)	oCS297/298	pSC298	pDONR207	Quikchange lightning (Agilent)	/
	Sr33$^{1-142}$	pSC401	Sr33$^{1-142}$ (without stop)	oCS299/300	pSC298	pDONR207	Quikchange lightning (Agilent)	/
	Sr33$^{1-144}$	pSC402	Sr33$^{1-144}$ (without stop)	oCS301/302	pSC298	pDONR207	Quikchange lightning (Agilent)	/
	Sr33$^{1-148}$	pSC403	Sr33$^{1-148}$ (without stop)	oCS303/304	pSC298	pDONR207	Quikchange lightning (Agilent)	/
	Sr50$^{1-133}$	pSC404	Sr50$^{1-133}$ (without stop)	oCS305/306	pSC262	pDONR207	Quikchange lightning (Agilent)	/
	Sr50$^{1-138}$	pSC405	Sr50$^{1-138}$ (without stop)	oCS307/308	pSC262	pDONR207	Quikchange lightning (Agilent)	/
	Sr50$^{1-144}$	pSC406	Sr50$^{1-144}$ (without stop)	oCS309/310	pSC262	pDONR207	Quikchange lightning (Agilent)	/
	Sr50$^{1-145}$	pSC407	Sr50$^{1-145}$ (without stop)	oCS311/312	pSC262	pDONR207	Quikchange lightning (Agilent)	/
	Sr50$^{1-147}$	pSC408	Sr50$^{1-147}$ (without stop)	oCS313/314	pSC262	pDONR207	Quikchange lightning (Agilent)	/
	Sr50$^{1-151}$	pSC409	Sr50$^{1-151}$ (without stop)	oCS315/316	pSC262	pDONR207	Quikchange lightning (Agilent)	/
Cell death assays and co-IPs in *N. benthamiana*	RGA4$^{1-171}$:CFP	pSC167	RGA4$^{1-171}$	/	/	pBin19-35s::GTW:CFP	/	17
	MLA10$^{1-160}$:CFP	pSC302	MLA10$^{1-160}$	/	/	pBin19-35s::GTW:CFP	/	29
	Sr33$^{1-160}$:CFP	pSC301	Sr33$^{1-160}$	/	/	pBin19-35s::GTW:CFP	/	29
	Sr50$^{1-163}$:HA	pSC280	Sr50$^{1-163}$	/	/	pBin19-35s::GTW:HA	/	29
	Sr50$^{1-163}$:CFP	pSC303	Sr50$^{1-163}$	/	/	pBin19-35s::GTW:CFP	/	29
	MLA10$^{1-130}$:HA	pSC410	MLA10$^{1-130}$	/	pSC392	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	MLA10$^{1-135}$:HA	pSC411	MLA10$^{1-135}$	/	pSC393	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	MLA10$^{1-141}$:HA	pSC412	MLA10$^{1-141}$	/	pSC394	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	MLA10$^{1-142}$:HA	pSC413	MLA10$^{1-142}$	/	pSC395	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	MLA10$^{1-144}$:HA	pSC414	MLA10$^{1-144}$	/	pSC396	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	MLA10$^{1-148}$:HA	pSC415	MLA10$^{1-148}$	/	pSC397	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/

	$Sr33^{1-130}$:HA	pSC416	$Sr33^{1-130}$	/	pSC398	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr33^{1-135}$:HA	pSC417	$Sr33^{1-135}$	/	pSC399	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr33^{1-141}$:HA	pSC418	$Sr33^{1-141}$	/	pSC400	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr33^{1-142}$:HA	pSC419	$Sr33^{1-142}$	/	pSC401	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr33^{1-144}$:HA	pSC420	$Sr33^{1-144}$	/	pSC402	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr33^{1-148}$:HA	pSC421	$Sr33^{1-148}$	/	pSC403	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr50^{1-133}$:HA	pSC422	$Sr50^{1-133}$	/	pSC404	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr50^{1-138}$:HA	pSC423	$Sr50^{1-138}$	/	pSC405	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr50^{1-144}$:HA	pSC424	$Sr50^{1-144}$	/	pSC406	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr50^{1-145}$:HA	pSC425	$Sr50^{1-145}$	/	pSC407	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr50^{1-147}$:HA	pSC426	$Sr50^{1-147}$	/	pSC408	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$Sr50^{1-151}$:HA	pSC427	$Sr50^{1-151}$	/	pSC409	pBin19-35s::GTW:HA	LR Gateway (Life Technologies)	/
	$MLA10^{1-141}$:CFP	pSC428	$MLA10^{1-141}$	/	pSC394	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$MLA10^{1-142}$:CFP	pSC429	$MLA10^{1-142}$	/	pSC395	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$MLA10^{1-144}$:CFP	pSC430	$MLA10^{1-144}$	/	pSC396	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$Sr33^{1-141}$:CFP	pSC431	$Sr33^{1-141}$	/	pSC400	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$Sr33^{1-142}$:CFP	pSC432	$Sr33^{1-142}$	/	pSC401	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$Sr33^{1-144}$:CFP	pSC433	$Sr33^{1-144}$	/	pSC402	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$Sr50^{1-144}$:CFP	pSC434	$Sr50^{1-144}$	/	pSC406	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$Sr50^{1-145}$:CFP	pSC435	$Sr50^{1-145}$	/	pSC407	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
	$Sr50^{1-147}$:CFP	pSC436	$Sr50^{1-147}$	/	pSC408	pBin19-35s::GTW:CFP	LR Gateway (Life Technologies)	/
Recombinant expression of proteins in *E. coli*.	$MLA10^{5-120}$	pMCSG7	$MLA10^{5-120}$	MLA10_5_FW / MLA10_120_RV	/	pMCSG7	Ligation-independent cloning	1
	$MLA10^{5-144}$	pMCSG7	$MLA10^{5-144}$	MLA10_5_FW / MLA10_144_RV	/	pMCSG7	Ligation-independent cloning	1
	$MLA10^{5-160}$	pMCSG7	$MLA10^{5-160}$	MLA10_5_FW / MLA10_160_RV	/	pMCSG7	Ligation-independent cloning	1
	$Sr33^{6-120}$	pMCSG7	$Sr33^{6-120}$	Sr33_6_FW / Sr33_120_RV	/	pMCSG7	Ligation-independent cloning	1
	$Sr33^{6-144}$	pMCSG7	$Sr33^{6-144}$	Sr33_6_FW / Sr33_144_RV	/	pMCSG7	Ligation-independent cloning	1
	$Sr33^{6-160}$	pMCSG7	$Sr33^{6-160}$	Sr33_6_FW / Sr33_160_RV	/	pMCSG7	Ligation-independent cloning	1
	$Sr50^{5-123}$	pMCSG7	$Sr50^{5-123}$	Sr50_5_FW / Sr33_123_RV	/	pMCSG7	Ligation-independent cloning	1
	Rx^{1-122}	pMCSG7	Rx^{1-122}	/		pMCSG7	Ligation-independent cloning	1

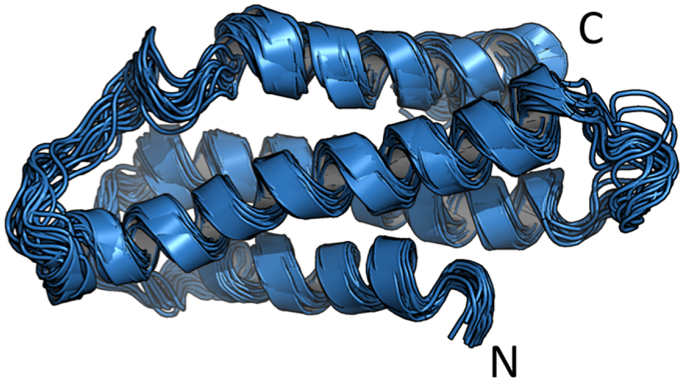


Fig. S1. Twenty superimposed lowest-energy structures of Sr33$^{6\text{-}120}$ (PDB ID 2NCG).

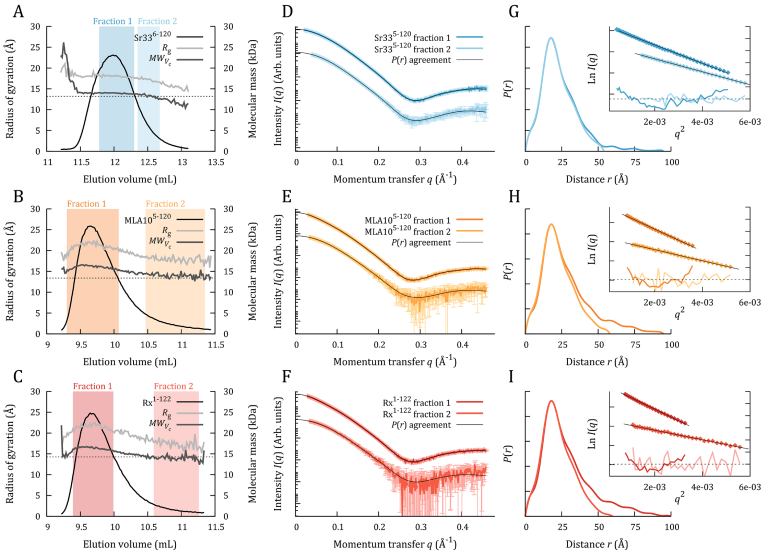


Fig. S2. Analysis of scattering curves averaged over the peak centre and dilute fractions from SEC-SAXS. (**A-C**) Evolution of particle Guinier R_g and molecular mass during in-line SEC-SAXS. For all proteins, the trace of zero-angle intensity, $I(0)$, is plotted as a black line arbitrarily scaled against the y axis, while the properties R_g and MM_{VC} are plotted as light and dark grey lines, respectively. The predicted monomeric molecular mass of each construct is shown as a black dotted line. Fractions averaged for analysis are marked by coloured shading. Note that the use of a 2 mL pre-column for Sr33$^{6-120}$ shifts that peak by the corresponding volume. (**D-F**) Experimental data-sets plotted as coloured lines, with experimental errors displayed at 1σ in lighter colour. Solid black lines indicate the fit of the corresponding distance distribution. The data-sets are arbitrarily offset along the *y*-axis for ease of visualization. (**G-I**) Normalized distance distribution functions, $P(r)$, are shown as coloured lines matching the scattering curve from which they were calculated. $P(r)$s have been normalized to reciprocal-space zero-angle intensity. The Guinier regions of the data-sets are shown in the insets, transformed as q^2 vs ln $I(q)$. Individual data-points are plotted as coloured diamonds, and a linear regression fit to each is shown as a black line. The data-sets are again offset in *y* for visualization. The residuals of each linear fit are also shown as coloured lines, plotted against the right hand axis. Aggregation in the peak fractions is apparent as a "smiling" curvature in the residuals., while the tail fraction residuals are normally distributed.

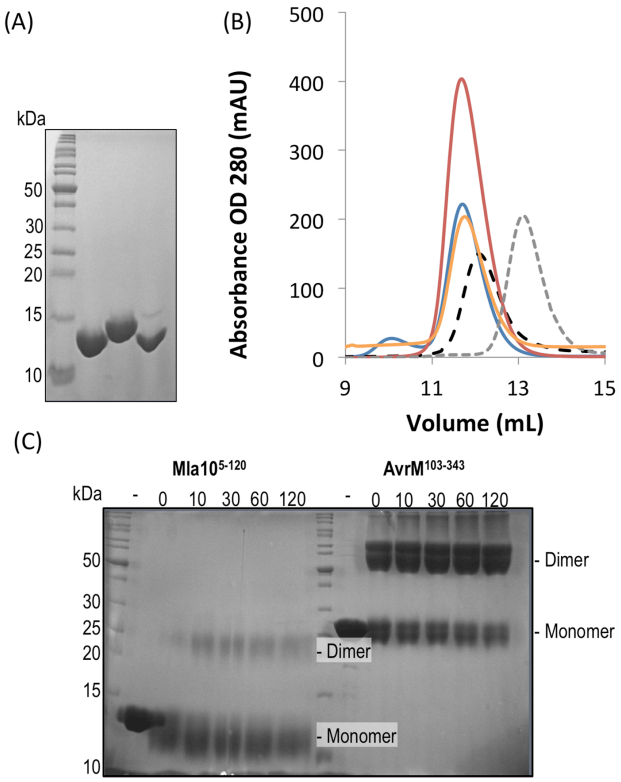


Fig. S3. Purification and SEC analysis **(A)** Coomassie blue-stained SDS-PAGE of purified (left-right) MLA10$^{5\text{-}120}$, Rx$^{1\text{-}122}$ and Sr33$^{6\text{-}120}$ proteins. **(B)** MLA10$^{5\text{-}120}$ (orange), Rx$^{1\text{-}122}$ (red) and Sr33$^{6\text{-}120}$ (blue) were separated on a Superdex S75 10/300 size-exclusion chromatography column and compared with known standards chymotrypsin (25 kDa – black dashed line) and cytochrome c (12 kDa – grey dashed line). **(C)** Chemical crosslinking of MLA10$^{5\text{-}120}$ and AvrM$^{103\text{-}343}$. The protein was incubated with the cross-linker BS3 and sampled at time points 0, 10, 30, 60, 120 min (- represents the protein without BS3 added). The protein samples were separated by SDS-PAGE and Coomassie blue-stained.

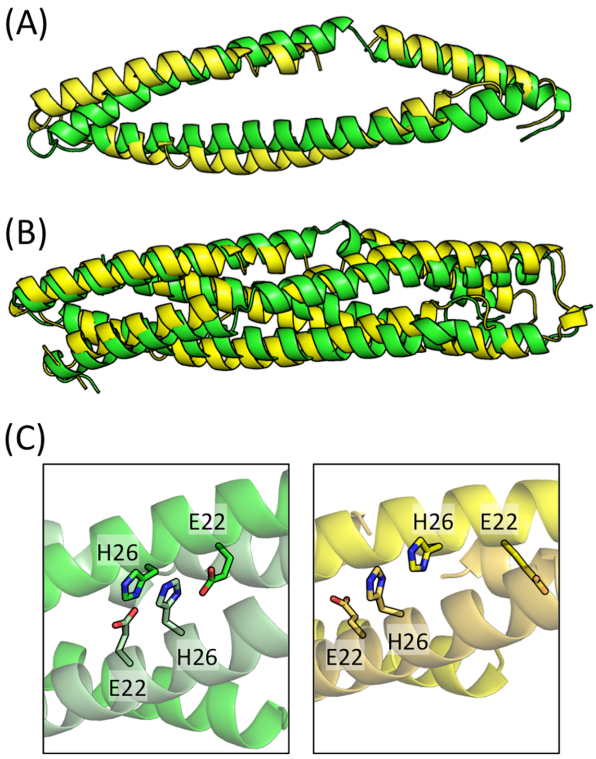


Fig. S4. A comparison of the MLA10$^{5\text{-}120}$ crystal structure solved in this study (PDB 5T1Y, shown in green) to that solved previously (9) (PDB ID 3QFL, shown in yellow). The RMSD for the monomer (**A**) and crystallographic dimer (**B**) is 3.6 Å and 3.7 Å, respectively. While overall the structures look similar there are differences between them with respect to the interactions between residues that coordinate the crystallographic dimer. (**C**) For example, in the structure solved here H26 and E22 from different protomers form a hydrogen bond (green, left), yet they do not interact in the 3QFL crystallographic dimer (yellow, right).

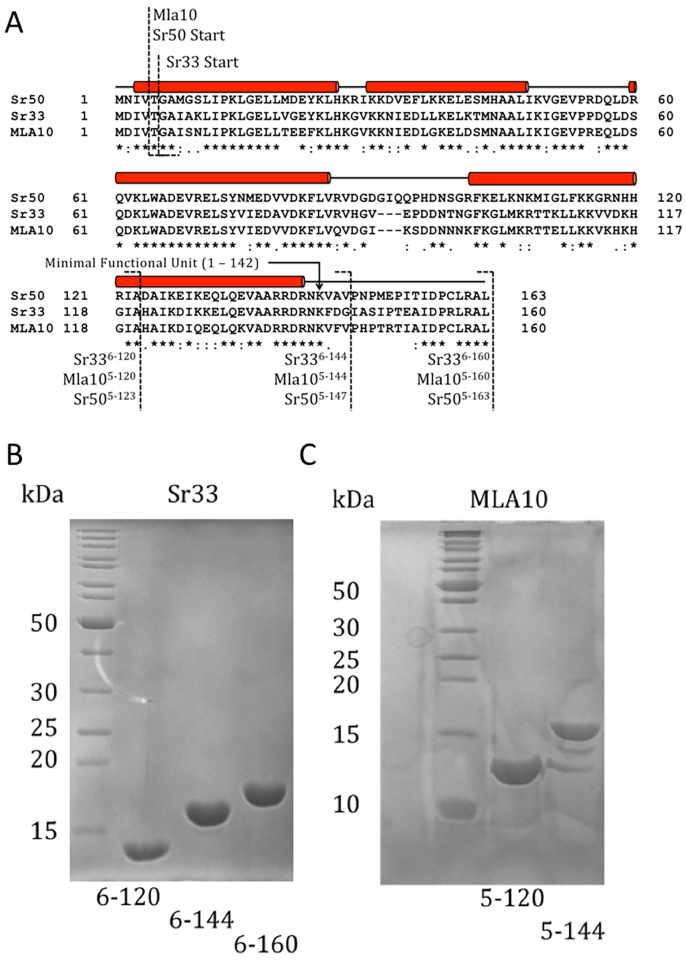


Fig. S5. Secondary structure prediction for MLA10, Sr33 and Sr50 CC domains from protein sequences using PSIPRED (27). Construct boundaries are marked with dashed lines. The minimal functional units for these constructs are indicated with an arrow. (**B, C**) Coomasie blue-stained SDS-PAGE of purified (left to right) $Sr33^{6\text{-}120}$, $Sr33^{6\text{-}144}$ and $Sr33^{6\text{-}160}$ (**B**); and $MLA10^{5\text{-}120}$ and $MLA10^{5\text{-}144}$ (**C**) proteins, used in solution studies (**Fig. 4**).

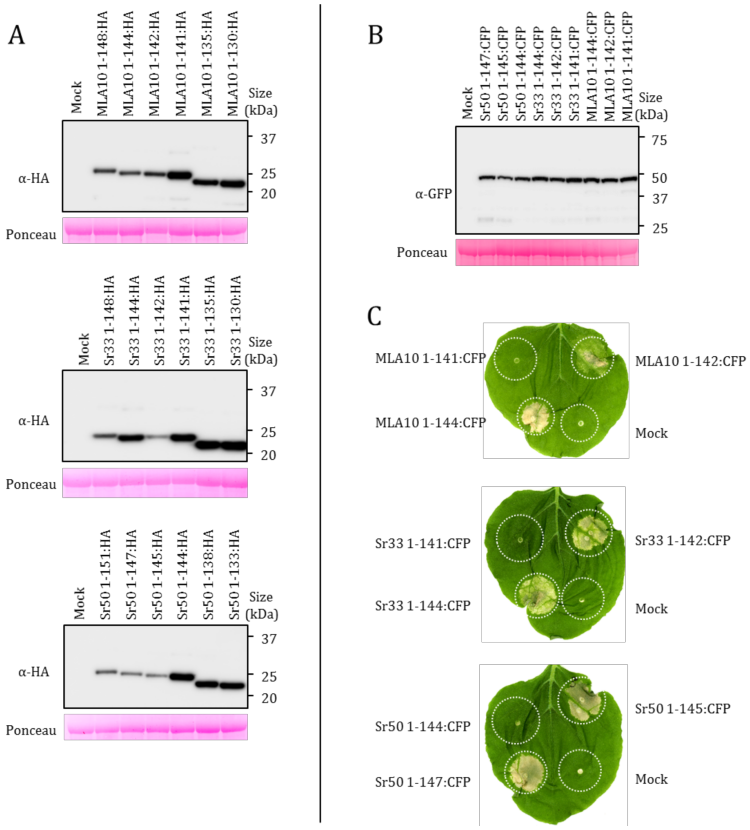


Fig. S6. HA and CFP-tagged CC fragments of MLA10, Sr33 and Sr50 expressed in *N. benthamiana*. **(A, B)** The indicated proteins were extracted from transiently transformed *N. benthamiana* leaves 20 hours after infiltration and were analyzed by immunoblotting with anti-GFP or anti-HA antibodies. Ponceau staining of RuBisCO was used to verify equal protein loading. **(C)** The indicated constructs were transiently expressed in *N. benthamiana*. Cell death was visualized five days after infiltration.

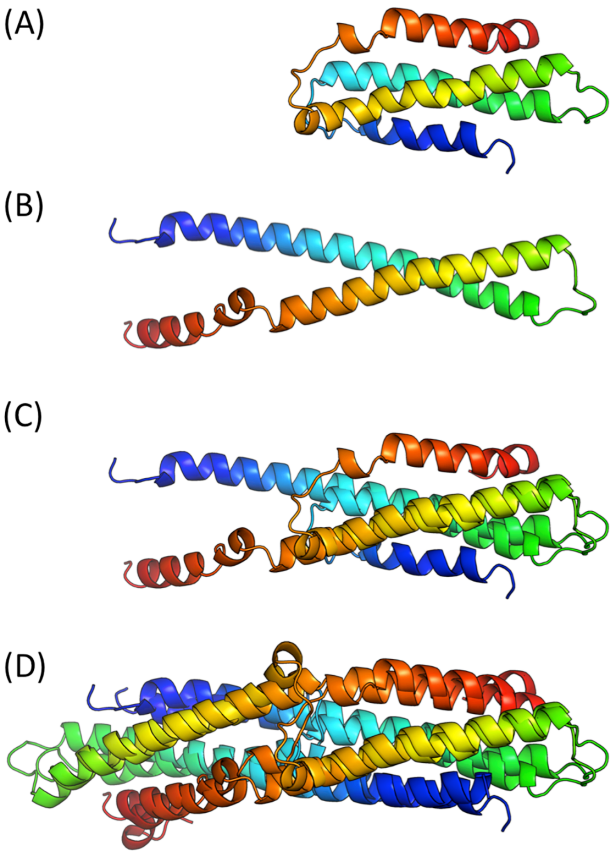


Fig. S7. Comparison of the MLA10$^{5\text{-}120}$ crystal structure and the Sr33$^{6\text{-}120}$ solution structure and the rationales' for the dimer of MLA10$^{5\text{-}120}$ representing a crystallisation induced domain-swapped dimer. (**A**) and (**B**) represent the NMR structure of Sr33$^{6\text{-}120}$ (PDB 2NCG) and the crystal structure of MLA10$^{5\text{-}120}$ (PDB 5T1Y), respectively. These are shown in cartoon and coloured using a rainbow spectrum (blue: N-terminus – red: C-terminus). Superposition of Sr33$^{6\text{-}120}$ onto MLA10$^{5\text{-}120}$ monomer (**C**) and dimer (**D**) (in (D) two Sr33 molecules were superimposed).

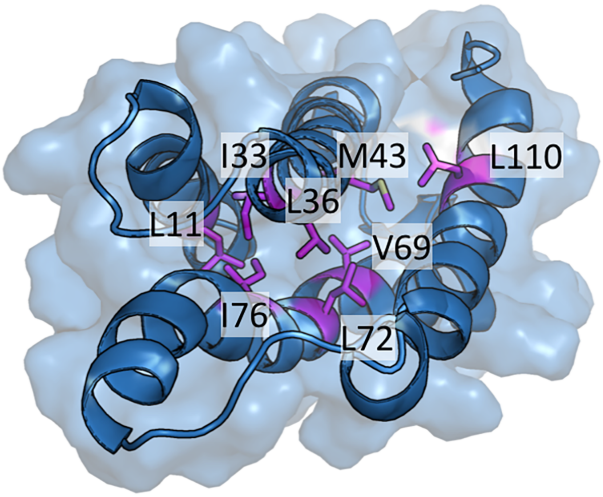


Fig. S8. Maekawa et al (9) reported that mutations in MLA10$^{5\text{-}120}$, including L11E, I33E, L36E, M43E, V69E, L72E, I76E, and L110E, could not be produced in a stable and soluble form when expressed in *E. coli*. Here the equivalent mutations are indicated in the Sr33$^{6\text{-}120}$ structure in stick representation, colored magenta and labelled. These residues form part of the four-helix bundle hydrophobic core in the Sr33$^{6\text{-}120}$ monomer. We suggest that glutamate mutations at these positions would have a destabilizing effect on the CC domain four-helix bundle fold.

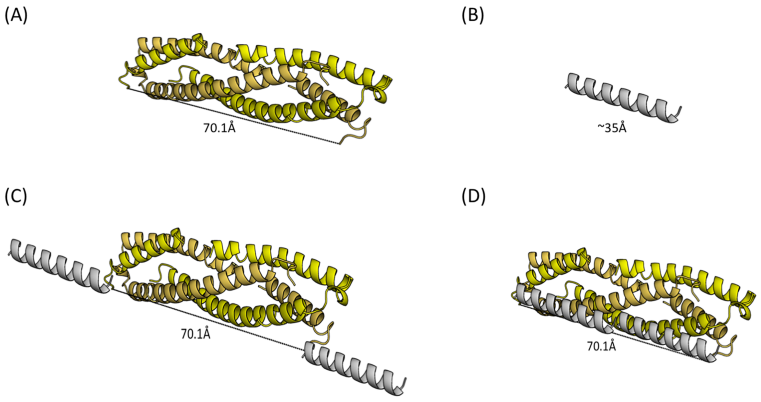


Fig. S9. Self-association and the contribution of the additional C-terminal residues. **(A)** In the MLA10 crystallographic domain-swapped dimer, the C-termini are ~70 Å apart, projecting in opposing directions. The residues 120-144 are predicted to be predominantly helical (see **Fig. S5**). **(B)** When modelled using UCSF Chimera (https://www.cgl.ucsf.edu/chimera/), a helix comprising these residues extends ~35 Å. **(C)** If the helices were to continue without break, they would project away from the body of MLA10 crystallographic domain-swapped dimer. In this situation, they would not support dimer formation in the context of the domain-swap dimer. **(D)** In an event that they folded towards each other as a modelled helix, they would not extend the distance to interact; however, it is plausible that these regions may provide additional contacts that could stabilize further the domain-swap structure.

Reference:

1. Anglister J, *et al.* (1993) Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *J Biomol NMR* 3(1):121-126.
2. Cavanagh J, *et al.* (2007) *Protein NMR Spectroscopy: Principles and Practice* (Elsevier) 2nd Ed pp 912-912.
3. Fischer H, *et al.* (2004) Average protein density is a molecular-weight-dependent function. *Prot Sci* 13:2825-2828.
4. Mobli M, *et al.* (2010) A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment. *J Magn Reson* 204:160-164.
5. Mobli M, *et al.* (2007) An automated tool for maximum entropy reconstruction of biomolecular NMR spectra. *Nat Methods* 4:467-468.
6. Vranken WF, *et al.* (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59:687-696.
7. Guntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278:353-378.
8. Cornilescu G, *et al.* (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 13:289-302.
9. Maekawa T, *et al.* (2011) Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host Microbe* 9(3):187-199.
10. Wen J, *et al.* (1996) Size-Exclusion Chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal Biochem* 240(2):155-166.
11. Petoukhov MV, *et al.* (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Cryst* 45(2):342-350.
12. Rambo RP & Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496(7446):477-481.
13. Schneidman-Duhovny D, *et al.* (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res* 38:W540-544.
14. Fiser A, *et al.* (2000) Modeling of loops in protein structures. *Prot Sci* 9(9):1753-1773.
15. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125-132.
16. Winn MD, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235-242.
17. Hao W, *et al.* (2013) Structural basis for the interaction between the potato virus X resistance protein (Rx) and its cofactor Ran GTPase-activating protein 2 (RanGAP2). *J Biol Chem* 288(50):35868-35876.
18. Pannu NS, *et al.* (2011) Recent advances in the CRANK software suite for experimental phasing. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):331-337.
19. Emsley P, *et al.* (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
20. Bricogne G, *et al.* (2016) *BUSTER version 2.10.2* (Global Phasing Ltd., Cambridge, UK).
21. Croll TI & Andersen GR (2016) Re-evaluation of low-resolution crystal structures via interactive molecular-dynamics flexible fitting (iMDFF): a case study in complement C4. *Acta Crystallogr D Biol Crystallogr* 72(9):1006-1016.
22. Afonine PV, *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352-367.
23. Cesari S, *et al.* (2016) Cytosolic activation of cell death and stem rust resistance by cereal MLA-family CC–NLR proteins. *Proc Natl Acad Sci U S A* 113(36):10204-10209

24. Herrmann T, *et al.* (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* 319:209-227.
25. Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Bio Crystallogr* 69(Pt 7):1204-1214.
26. Chen VB, *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Bio Crystallogr* 66(Pt 1):12-21.
27. Buchan DWA, *et al.* (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res* 41:W349-357.

BIBLIOGRAPHY

ABRAMOVITCH, R. B., JANJUSEVIC, R., STEBBINS, C. E. & MARTIN, G. B. 2006. Type III effector AvrPtoB requires intrinsic E3 ubiquitin ligase activity to suppress plant cell death and immunity. *Proceedings of the National Academy of Sciences,* 103**,** 2851-2856.

ADAMS, P. D., AFONINE, P. V., BUNKÓCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEADD, J. J., HUNG, L. W., KAPRAL, G. J., GROSSE-KUNSTLEVE, R. W., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D*
, 66**,** 213-221.

AKIRA, S., UEMATSU, S. & TAKEUCHI, O. 2006. Pathogen recognition and innate immunity. *Cell,* 124**,** 783-801.

APEL, K. & HIRT, H. 2004. Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology,* 55**,** 373-399.

ASHIKAWA, I., HAYASHI, N., YAMANE, H., KANAMORI, H., WU, J. Z., MATSUMOTO, T., ONO, K. & YANO, M. 2008. Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer Pikm-specific rice blast resistance. *Genetics,* 180**,** 2267-2276.

AYLIFFE, M., PERIYANNAN, S. K., FEECHAN, A., DRY, I., SCHUMANN, U., WANG, M. B., PRYOR, A. & LAGUDAH, E. 2013. A simple method for comparing fungal biomass in infected plant tissues. *Molecular Plant-Microbe Interactions,* 26**,** 658-667.

BAI, S., LIU, J., CHANG, C., ZHANG, L., MAEKAWA, T., WANG, Q., XIAO, W., LIU, Y., CHAI, J., TAKKEN, F. L., SCHULZE-LEFERT, P. & SHEN, Q. H. 2012. Structure-function analysis of barley NLR immune receptor MLA10 reveals its cell compartment specific activity in cell death and disease resistance. *PLoS Pathogens,* 8**,** e1002752.

BARCELO, A. R. 1997. Lignification in plant cell walls. *In:* JEON, K. W. (ed.) *International Review of Cytology.*

BENDAHMANE, A., KANYUKA, K. & BAULCOMBE, D. C. 1999. The Rx gene from potato controls separate virus resistance and cell death responses. *Plant Cell,* 11**,** 781-791.
BENTHAM, A., BURDETT, H., ANDERSON, P. A., WILLIAMS, S. J. & KOBE, B. 2017. Animal NLRs provide structural insights into plant NLR function. *Annals of Botany,* 119**,** 827-702.
BERNOUX, M., BURDETT, H., WILLIAMS, S. J., ZHANG, X., CHEN, C., NEWELL, K., LAWRENCE, G. J., KOBE, B., ELLIS, J. G., ANDERSON, P. A. & DODDS, P. N. 2016. Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. *Plant Cell,* 28**,** 146-159.
BERNOUX, M., ELLIS, J. G. & DODDS, P. N. 2011a. New insights in plant immunity signaling activation. *Current Opinion in Plant Biology,* 14**,** 512-518.
BERNOUX, M., VE, T., WILLIAMS, S., WARREN, C., HATTERS, D., VALKOV, E., ZHANG, X. X., ELLIS, J. G., KOBE, B. & DODDS, P. N. 2011b. Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host & Microbe,* 9**,** 200-211.
BHUIYAN, N. H., SELVARAJ, G., WEI, Y. D. & KING, J. 2009. Gene expression profiling and silencing reveal that monolignol biosynthesis plays a critical role in penetration defence in wheat against powdery mildew invasion. *Journal of Experimental Botany,* 60**,** 509-521.
BOLLER, T. 1995. Chemoperception of microbial signals in plant cells. *Annual Review of Plant Physiology and Plant Molecular Biology,* 46**,** 189-214.
BOS, J. I. B., KANNEGANTI, T.-D., YOUNG, C., CAKIR, C., HUITEMA, E., WIN, J., ARMSTRONG, M. R., BIRCH, P. R. J. & KAMOUN, S. 2006. The C-terminal half of Phytophthora infestans RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in Nicotiana benthamiana. *Plant Journal,* 48**,** 165-176.
BOUCHIER-HAYES, L. & MARTIN, S. J. 2002. CARD games in apoptosis and immunity. *EMBO Rep,* 3**,** 616-21.
BUCHAN, D. W. A., MINNECI, F., NUGENT, T. C. O., BRYSON, K. & JONES, D. T. 2013. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research,* 41**,** W349-357.
CAILLAUD, M.-C., PIQUEREZ, S. J. M., FABRO, G., STEINBRENNER, J., ISHAQUE, N., BEYNON, J. & JONES, J. D. G. 2012. Subcellular localization of the Hpa RxLR effector repertoire identifies a tonoplast-associated protein HaRxL17 that confers enhanced plant susceptibility. *Plant Journal,* 69**,** 252-265.
CANTU, D., VICENTE, A. R., LABAVITCH, J. M., BENNETT, A. B. & POWELL, A. L. T. 2008. Strangers in the matrix: plant cell

walls and pathogen susceptibility. *Trends in Plant Science,* 13**,** 610-617.
CASEY, L. W., LAVRENCIC, P., BENTHAM, A. R., CESARI, S., ERICSSON, D. J., CROLL, T., TURK, D., ANDERSON, P. A., MARK, A. E., DODDS, P. N., MOBLI, M., KOBE, B. & WILLIAMS, S. J. 2016. The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins. *Proceedings of the National Academy of Sciences,* 113**,** 12856-12861.
CATANZARITI, A. M., DODDS, P. N. & ELLIS, J. G. 2007. Avirulence proteins from haustoria-forming pathogens. *Fems Microbiology Letters,* 269**,** 181-188.
CESARI, S., BERNOUX, M., MONCUQUET, P., KROJ, T. & DODDS, P. N. 2014a. A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. *Frontiers in Plant Science,* 5**,** 606.
CESARI, S., KANZAKI, H., FUJIWARA, T., BERNOUX, M., CHALVON, V., KAWANO, Y., SHIMAMOTO, K., DODDS, P., TERAUCHI, R. & KROJ, T. 2014b. The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. *Embo Journal,* 33**,** 1941-1959.
CESARI, S., MOORE, J., CHEN, C., WEBB, D., PERIYANNAN, S., MAGO, R., BERNOUX, M., LAGUDAH, E. S. & DODDS, P. N. 2016. Cytosolic activation of cell death and stem rust resistance by cereal MLA-family CC–NLR proteins. *Proceedings of the National Academy of Sciences,* 113**,** 10204-10209.
CHENG, W., MUNKVOLD, K. R., GAO, H., MATHIEU, J., SCHWIZER, S., WANG, S., YAN, Y.-B., WANG, J., MARTIN, G. B. & CHAI, J. 2011. Structural analysis of Pseudomonas syringae AvrPtoB bound to host BAK1 reveals two Similar kinase-interacting domains in a type III effector. *Cell Host & Microbe,* 10**,** 616-626.
CHINCHILLA, D., ZIPFEL, C., ROBATZEK, S., KEMMERLING, B., NURNBERGER, T., JONES, J. D. G., FELIX, G. & BOLLER, T. 2007. A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature,* 448**,** 497-512.
CHISHOLM, S. T., COAKER, G., DAY, B. & STASKAWICZ, B. J. 2006. Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell,* 124**,** 803-814.
COHN, J., SESSA, G. & MARTIN, G. B. 2001. Innate immunity in plants. *Current Opinion in Immunology,* 13**,** 55-62.
CROLL, T. I. & ANDERSEN, G. R. 2016. Re-evaluation of low-resolution crystal structures via interactive molecular-dynamics flexible fitting (iMDFF): a case study in complement C4. *Acta Crystallographica Section D*
, 72**,** 1006-1016.

CROOKS, G. E., HON, G., CHANDONIA, J. M. & BRENNER, S. E. 2004. WebLogo: a sequence logo generator. *Genome Research,* 14**,** 1188-1190.
DANGL, J. L. & JONES, J. D. G. 2001. Plant pathogens and integrated defence responses to infection. *Nature,* 411**,** 826-833.
DANOT, O., MARQUENET, E., VIDAL-INGIGLIARDI, D. & RICHET, E. 2009. Wheel of life, wheel of death: A mechanistic insight into signaling by STAND proteins. *Structure,* 17**,** 172-182.
DAVIN, L. B., XIA, Z. Q. & LEWIS, N. G. 1995. Regulation of lignin synthesis. *Plant Physiology,* 108**,** 42-42.
DODDS, P. N., LAWRENCE, G. J., CATANZARITI, A. M., TEH, T., WANG, C. I. A., AYLIFFE, M. A., KOBE, B. & ELLIS, J. G. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences,* 103**,** 8888-8893.
DODDS, P. N. & RATHJEN, J. P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics,* 11**,** 539-548.
DOEHLEMANN, G. & HEMETSBERGER, C. 2013. Apoplastic immunity and its suppression by filamentous plant pathogens. *New Phytologist,* 198**,** 1001-1016.
DONG, A., XU, X. & EDWARDS, A. M. 2007. In situ proteolysis for protein crystallization and structure determination. *Nature Methods,* 4**,** 1019-1021.
DRAPER, J. 1997. Salicylate, superoxide synthesis and cell suicide in plant defence. *Trends in Plant Science,* 2**,** 162-165.
DUXBURY, Z., MA, Y., FURZER, O. J., HUH, S. U., CEVIK, V., JONES, J. D. G. & SARRIS, P. F. 2016. Pathogen perception by NLRs in plants and animals: Parallel worlds. *Bioessays,* 38**,** 769-781.
EITAS, T. K. & DANGL, J. L. 2010. NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Current Opinion in Plant Biology,* 13**,** 472-477.
EL KASMI, F., CHUNG, E.-H., ANDERSON, R. G., LI, J., WAN, L., EITAS, T. K., GAO, Z. & DANGL, J. L. 2017. Signaling from the plasma-membrane localized plant immune receptor RPM1 requires self-association of the full-length protein. *Proceedings of the National Academy of Sciences,* 114**,** E7385-E7394.
ELLIS, J. G. & DODDS, P. N. 2011. Showdown at the RXLR motif: Serious differences of opinion in how effector proteins from filamentous eukaryotic pathogens enter plant cells. *Proceedings of the National Academy of Sciences,* 108**,** 14381-14382.
FELIX, G., REGENASS, M. & BOLLER, T. 1993. Specific perception of subnanomolar concentrations of chitin fragments by tomato cells: Induction of extracellular alkalinization,

changes in protein phosphorylation, and establishment of a refractory state. *Plant Journal,* 4**,** 307-316.
FERREIRA, R. B., MONTEIRO, S., FREITAS, R., SANTOS, C. N., CHEN, Z. J., BATISTA, L. M., DUARTE, J., BORGES, A. & TEIXEIRA, A. R. 2006. Fungal pathogens: The battle for plant infection. *Critical Reviews in Plant Sciences,* 25**,** 505-524.
FLOR, H. H. 1971. Current status of gene-for-gene concept. *Annual Review of Phytopathology,* 9**,** 275-296.
FREUDENBERG, K. 1957. Structure and formation of lignin. *Industrial and Engineering Chemistry,* 49**,** 1384-1384.
GAY, N. J. & KEITH, F. J. 1991. Drosophila Toll and IL-1 receptor. *Nature,* 351**,** 355-356.
GAY, N. J., SYMMONS, M. F., GANGLOFF, M. & BRYANT, C. E. 2014. Assembly and localization of Toll-like receptor signalling complexes. *Nature Reviews Immunology,* 14**,** 546-558.
GIMENEZ-IBANEZ, S., HANN, D. R., NTOUKAKLS, V., PETUTSCHNIG, E., LIPKA, V. & RATHJEN, J. P. 2009. AvrPtoB targets the LysM receptor kinase CERK1 to promote bacterial virulence on plants. *Current Biology,* 19**,** 423-429.
GLAZEBROOK, J. 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology.*
GOMEZ-GOMEZ, L. & BOLLER, T. 2000. FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell,* 5**,** 1003-1011.
GOMEZ-GOMEZ, L., FELIX, G. & BOLLER, T. 1999. A single locus determines sensitivity to bacterial flagellin in Arabidopsis thaliana. *Plant Journal,* 18**,** 277-284.
GRANT, S. R., FISHER, E. J., CHANG, J. H., MOLE, B. M. & DANGL, J. L. 2006. Subterfuge and manipulation: Type III effector proteins of phytopathogenic bacteria. *Annual Review of Microbiology.*
GUST, A. A., BISWAS, R., LENZ, H. D., RAUHUT, T., RANF, S., KEMMERLING, B., GOTZ, F., GLAWISCHNIG, E., LEE, J., FELIX, G. & NURNBERGER, T. 2007. Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in Arabidopsis. *Journal of Biological Chemistry,* 282**,** 32338-32348.
HAMMERSCHMIDT, R. 1999. Phytoalexins: What have we learned after 60 years? *Annual Review of Phytopathology,* 37**,** 285-306.
HAO, W., COLLIER, S. M., MOFFETT, P. & CHAI, J. J. 2013. Structural basis for the interaction between the Potato Virus X resistance protein (Rx) and its cofactor Ran GTPase-activating Protein 2 (RanGAP2). *Journal of Biological Chemistry,* 288**,** 35868-35876.
HARDHAM, A. R. & CAHILL, D. M. 2010. The role of oomycete effectors in plant-pathogen interactions. *Functional Plant Biology,* 37**,** 919-925.

HAUROWITZ, F., TUNCA, M., SCHWEMN, P. & GOKSU, V. 1945. The action of trypsin on native and denatured proteins. *Journal of Biological Chemistry,* 157**,** 621-625.
HEIDRICH, K., BLANVILLAIN-BAUFUMÉ, S. & PARKER, J. E. 2012. Molecular and spatial constraints on NB-LRR receptor signaling. *Current Opinion in Plant Biology,* 15**,** 385-391.
HOGENHOUT, S. A., VAN DER HOORN, R. A. L., TERAUCHI, R. & KAMOUN, S. 2009. Emerging concepts in effector biology of plant-associated organisms. *Molecular Plant-Microbe Interactions,* 22**,** 115-122.
HORBACH, R., NAVARRO-QUESADA, A. R., KNOGGE, W. & DEISING, H. B. 2011. When and how to kill a plant cell: Infection strategies of plant pathogenic fungi. *Journal of Plant Physiology,* 168**,** 51-62.
HU, Z., ZHOU, Q., ZHANG, C., FAN, S., CHENG, W., ZHAO, Y., SHAO, F., WANG, H.-W., SUI, S.-F. & CHAI, J. 2015. Structural and biochemical basis for induced self-propagation of NLRC4. *Science,* 350**,** 399-404.
HUH, S. U., CEVIK, V., DING, P., DUXBURY, Z., MA, Y., TOMLINSON, L., SARRIS, P. F. & JONES, J. D. G. 2017. Protein-protein interactions in the RPS4/RRS1 immune receptor complex. *PLOS Pathogens,* 13**,** e1006376.
JONES, D. T. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology,* 292**,** 195-202.
JONES, J. D. G. & DANGL, J. L. 2006. The plant immune system. *Nature,* 444**,** 323-329.
JOOSTEN, M. & DE WIT, P. 1999. The tomato - Cladosporium fulvum interaction: A versatile experimental system to study plant-pathogen interactions. *Annual Review of Phytopathology,* 37**,** 335-367.
JORDAN, T., SEEHOLZER, S., SCHWIZER, S., TOLLER, A., SOMSSICH, I. E. & KELLER, B. 2011. The wheat Mla homologue TmMla1 exhibits an evolutionarily conserved function against powdery mildew in both wheat and barley. *Plant Journal,* 65**,** 610-621.
KALE, S. D., GU, B., CAPELLUTO, D. G. S., DOU, D., FELDMAN, E., RUMORE, A., ARREDONDO, F. D., HANLON, R., FUDAL, I., ROUXEL, T., LAWRENCE, C. B., SHAN, W. & TYLER, B. M. 2010. External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. *Cell,* 142**,** 284-295.
KELLEY, L. A., MEZULIS, S., YATES, C. M., WASS, M. N. & STERNBERG, M. J. E. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols,* 10**,** 845-858.
KIM, H. S., DESVEAUX, D., SINGER, A. U., PATEL, P., SONDEK, J. & DANGL, J. L. 2005. The Pseudomonas syringae effector AvrRpt2 cleaves its C-terminally acylated target, RIN4, from

Arabidopsis membranes to block RPM1 activation. *Proceedings of the National Academy of Sciences,* 102**,** 6496-6501.
KIM, J.-G., STORK, W. & MUDGETT, M. B. 2013. Xanthomonas type III effector XopD desumoylates tomato transcription factor SlERF4 to suppress ethylene responses and promote pathogen growth. *Cell Host & Microbe,* 13**,** 143-154.
KIM, J.-G., TAYLOR, K. W., HOTSON, A., KEEGAN, M., SCHMELZ, E. A. & MUDGETT, M. B. 2008. XopD SUMO protease affects host transcription, promotes pathogen growth, and delays symptom development in Xanthomonas-infected tomato leaves. *Plant Cell,* 20**,** 1915-1929.
KLESSIG, D. F., DURNER, J., NOAD, R., NAVARRE, D. A., WENDEHENNE, D., KUMAR, D., ZHOU, J. M., SHAH, J., ZHANG, S. Q., KACHROO, P., TRIFA, Y., PONTIER, D., LAM, E. & SILVA, H. 2000. Nitric oxide and salicylic acid signaling in plant defense. *Proceedings of the National Academy of Sciences,* 97**,** 8849-8855.
KRASILEVA, K. V., DAHLBECK, D. & STASKAWICZ, B. J. 2010. Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell,* 22**,** 2444-2458.
KUNZE, G., ZIPFEL, C., ROBATZEK, S., NIEHAUS, K., BOLLER, T. & FELIX, G. 2004. The N terminus of bacterial elongation factor Tu elicits innate immunity in Arabidopsis plants. *Plant Cell,* 16**,** 3496-3507.
LAMB, C. & DIXON, R. A. 1997. The oxidative burst in plant disease resistance. *Annual Review of Plant Physiology and Plant Molecular Biology,* 48**,** 251-275.
LONG, F., VAGIN, A. A., YOUNG, P. & MURSHUDOV, G. N. 2008. BALBES: a molecular-replacement pipeline. *Acta Crystallographica Section D*
, 64**,** 125-132.
LUKASIK, E. & TAKKEN, F. L. W. 2009. STANDing strong, resistance proteins instigators of plant defence. *Current Opinion in Plant Biology,* 12**,** 427-436.
LUPAS, A., VAN DYKE, M. & STOCK, J. 1991. Predicting coiled coils from protein sequences. *Science,* 252**,** 1162-1624.
MAEKAWA, T., CHENG, W., SPIRIDON, L. N., TOLLER, A., LUKASIK, E., SAIJO, Y., LIU, P. Y., SHEN, Q. H., MICLUTA, M. A., SOMSSICH, I. E., TAKKEN, F. L. W., PETRESCU, A. J., CHAI, J. J. & SCHULZE-LEFERT, P. 2011a. Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host & Microbe,* 9**,** 187-199.
MAEKAWA, T., KUFER, T. A. & SCHULZE-LEFERT, P. 2011b. NLR functions in plant and animal immune systems: so far and yet so close. *Nat Immunol,* 12**,** 817-826.

MAQBOOL, A., SAITOH, H., FRANCESCHETTI, M., STEVENSON, C. E. M., UEMURA, A., KANZAKI, H., KAMOUN, S., TERAUCHI, R. & BANFIELD, M. J. 2015. Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor. *eLife,* 4**,** e08709.

MCPHERSON, A. & GAVIRA, J. A. 2014. Introduction to protein crystallization. *Acta Crystallographica Section F,* 70**,** 2-20.

MEHDY, M. C. 1994. Active oxygen species in plant defense against pathogens. *Plant Physiology,* 105**,** 467-472.

MENG, X. & ZHANG, S. 2013. MAPK cascades in plant disease resistance signaling. *Annual Review of Phytopathology,* 51**,** 245-266.

MESTRE, P. & BAULCOMBE, D. C. 2006. Elicitor-mediated oligomerization of the tobacco N Disease resistance protein. *Plant Cell,* 18**,** 491-501.

MEYERS, B. C., MORGANTE, M. & MICHELMORE, R. W. 2002. TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in Arabidopsis and other plant genomes. *Plant Journal,* 32**,** 77-92.

MOFFETT, P., FARNHAM, G., PEART, J. & BAULCOMBE, D. C. 2002. Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *Embo Journal,* 21**,** 4511-4519.

MONAGHAN, J. & ZIPFEL, C. 2012. Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology,* 15**,** 349-357.

MUSCHOL, M. & ROSENBERGER, F. 1997. Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *The Journal of Chemical Physics,* 107**,** 1953-1962.

NARUSAKA, M., SHIRASU, K., NOUTOSHI, Y., KUBO, Y., SHIRAISHI, T., IWABUCHI, M. & NARUSAKA, Y. 2009. RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *Plant Journal,* 60**,** 218-226.

NEWMAN, J. 2004. Novel buffer systems for macromolecular crystallization. *Acta Crystallographica Section D* *,* 60**,** 610-612.

NIMCHUK, Z., MAROIS, E., KJEMTRUP, S., LEISTER, R. T., KATAGIRI, F. & DANGL, J. L. 2000. Eukaryotic fatty acylation drives plasma membrane targeting and enhances function of several type III effector proteins from Pseudomonas syringae. *Cell,* 101**,** 353-363.

NISHIMURA, M. T., ANDERSON, R. G., CHERKIS, K. A., LAW, T. F., LIU, Q. L., MACHIUS, M., NIMCHUK, Z. L., YANG, L., CHUNG, E. H., EL KASMI, F., HYUNH, M., OSBORNE NISHIMURA, E., SONDEK, J. E. & DANGL, J. L. 2017. TIR-only protein RBA1 recognizes a pathogen effector to regulate cell death in Arabidopsis. *Proceedings of the National Academy of Sciences,* 114**,** E2053-e2062.

NTOUKAKIS, V., BALMUTH, A. L., MUCYN, T. S., GUTIERREZ, J. R., JONES, A. M. E. & RATHJEN, J. P. 2013. The tomato Prf complex Is a molecular trap for bacterial effectors based on Pto transphosphorylation. *PLoS Pathogens,* 9**,** e1003123.

NYMAN, T., STENMARK, P., FLODIN, S., JOHANSSON, I., HAMMARSTROM, M. & NORDLUND, P. 2008. The crystal structure of the human Toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *Journal of Biological Chemistry,* 283**,** 11861-11865.

NYSKA, A. & KOHEN, R. 2002. Oxidation of biological systems: Oxidative stress phenomena, antioxidants, redox reactions, and methods for their quantification. *Toxicologic Pathology,* 30**,** 620-650.

O'NEILL, L. 2000. The Toll/interleukin-1 receptor domain: a molecular switch for inflammation and host defence. *Biochemical Society Transactions,* 28**,** 557-563.

OH, C. S. & MARTIN, G. B. 2011. Tomato 14-3-3 protein TFT7 interacts with a MAP kinase kinase to regulate immunity-associated programmed cell death mediated by diverse disease resistance proteins. *Journal of Biological Chemistry,* 22**,** 14129-14136.

OKUYAMA, Y., KANZAKI, H., ABE, A., YOSHIDA, K., TAMIRU, M., SAITOH, H., FUJIBE, T., MATSUMURA, H., SHENTON, M., GALAM, D. C., UNDAN, J., ITO, A., SONE, T. & TERAUCHI, R. 2011. A multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS-LRR protein genes. *Plant Journal,* 66**,** 467-479.

PERIYANNAN, S., MOORE, J., AYLIFFE, M., BANSAL, U., WANG, X. J., HUANG, L., DEAL, K., LUO, M. C., KONG, X. Y., BARIANA, H., MAGO, R., MCINTOSH, R., DODDS, P., DVORAK, J. & LAGUDAH, E. 2013. The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science,* 341**,** 786-788.

POSTEL, S. & KEMMERLING, B. 2009. Plant systems for recognition of pathogen-associated molecular patterns. *Seminars in Cell & Developmental Biology,* 20**,** 1025-1031.

QI, S., PANG, Y., HU, Q., LIU, Q., LI, H., ZHOU, Y., HE, T., LIANG, Q., LIU, Y., YUAN, X., LUO, G., LI, H., WANG, J., YAN, N. & SHI, Y. 2010. Crystal structure of the Caenorhabditis elegans apoptosome reveals an octameric assembly of CED-4. *Cell,* 141**,** 446-457.

RAIRDAN, G. J., COLLIER, S. M., SACCO, M. A., BALDWIN, T. T., BOETTRICH, T. & MOFFETT, P. 2008. The coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signaling. *Plant Cell,* 20**,** 739-751.

RAMBO, R. P. & TAINER, J. A. 2013. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature,* 496**,** 477-481.

REHMANY, A. P., GORDON, A., ROSE, L. E., ALLEN, R. L., ARMSTRONG, M. R., WHISSON, S. C., KAMOUN, S., TYLER, B. M., BIRCH, P. R. & BEYNON, J. L. 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. *Plant Cell,* 17**,** 1839-1850.

REUBOLD, T. F., HAHNE, G., WOHLGEMUTH, S. & ESCHENBURG, S. 2014. Crystal structure of the leucine-rich repeat domain of the NOD-like receptor NLRP1: Implications for binding of muramyl dipeptide. *Febs Letters,* 588**,** 3327-3332.

REUBOLD, T. F., WOHLGEMUTH, S. & ESCHENBURG, S. 2011. Crystal structure of full-length Apaf-1: How the death signal Is relayed in the mitochondrial pathway of apoptosis. *Structure,* 19**,** 1074-1083.

RIEDL, S. J., LI, W. Y., CHAO, Y., SCHWARZENBACHER, R. & SHI, Y. G. 2005. Structure of the apoptotic protease-activating factor 1 bound to ADP. *Nature,* 434**,** 926-933.

ROBATZEK, S. & WIRTHMUELLER, L. 2013. Mapping FLS2 function to structure: LRRs, kinase and its working bits. *Protoplasma,* 250**,** 671-681.

SACCO, M. A., MANSOOR, S. & MOFFETT, P. 2007. A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant Journal,* 52**,** 82-93.

SAINTENAC, C., ZHANG, W. J., SALCEDO, A., ROUSE, M. N., TRICK, H. N., AKHUNOV, E. & DUBCOVSKY, J. 2013. Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. *Science,* 341**,** 783-786.

SALVESEN, G. S. & DIXIT, V. M. 1999. Caspase activation: The induced-proximity model. *Proceedings of the National Academy of Sciences,* 96**,** 10964-10967.

SCHREIBER, K. J., BENTHAM, A., WILLIAMS, S. J., KOBE, B. & STASKAWICZ, B. J. 2016. Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. *PLoS Pathogens,* 12**,** e1005769.

SCOFIELD, S. R., HUANG, L., BRANDT, A. S. & GILL, B. S. 2005. Development of a virus-induced gene-silencing system for hexaploid wheat and its use in functional analysis of the Lr21-mediated leaf rust resistance pathway. *Plant Physiology,* 138**,** 2165-2173.

SEEHOLZER, S., TSUCHIMATSU, T., JORDAN, T., BIERI, S., PAJONK, S., YANG, W. X., JAHOOR, A., SHIMIZU, K. K., KELLER, B. & SCHULZE-LEFERT, P. 2010. Diversity at the Mla powdery mildew resistance locus from cultivated barley reveals sites of positive selection. *Molecular Plant-Microbe Interactions,* 23**,** 497-509.

SHAN, L., HE, P., LI, J., HEESE, A., PECK, S. C., NUERNBERGER, T., MARTIN, G. B. & SHEEN, J. 2008. Bacterial effectors

target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. *Cell Host & Microbe,* 4**,** 17-27.
SHAO, F., GOLSTEIN, C., ADE, J., STOUTEMYER, M., DIXON, J. E. & INNES, R. W. 2003. Cleavage of Arabidopsis PBS1 by a bacterial type III effector. *Science,* 301**,** 1230-1233.
SHIRASU, K., NAKAJIMA, H., RAJASEKHAR, V. K., DIXON, R. A. & LAMB, C. 1997. Salicylic acid potentiates an agonist-dependent gain control that amplifies pathogen signals in the activation of defense mechanisms. *Plant Cell,* 9**,** 261-270.
SINGH, R. P., HODSON, D. P., HUERTA-ESPINO, J., JIN, Y., BHAVANI, S., NJAU, P., HERRERA-FOESSEL, S., SINGH, P. K., SINGH, S. & GOVINDAN, V. 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. *Annual Review of Phytopathology,* 49**,** 465-481.
STEINBRENNER, A. D., GORITSCHNIG, S. & STASKAWICZ, B. J. 2015. Recognition and activation domains contribute to allele-specific responses of an Arabidopsis NLR receptor to an oomycete effector protein. *PLoS Pathogens,* 11**,** e1004665.
SUN, F., KALE, S. D., AZURMENDI, H. F., LI, D., TYLER, B. I. & CAPELLUTO, D. G. S. 2013. Structural basis for interactions of the Phytophthora sojae RxLR effector Avh5 with phosphatidylinositol 3-phosphate and for host cell entry. *Molecular Plant-Microbe Interactions,* 26**,** 330-344.
SZUREK, B., MAROIS, E., BONES, U. & VAN DEN ACKERVEKEN, G. 2001. Eukaryotic features of the Xanthomonas type III effector AvrBs3: protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. *Plant Journal,* 26**,** 523-534.
TAKEMOTO, D., RAFIQI, M., HURLEY, U., LAWRENCE, G. J., BERNOUX, M., HARDHAM, A. R., ELLIS, J. G., DODDS, P. N. & JONES, D. A. 2012. N-terminal motifs in some plant disease resistance proteins function in membrane attachment and contribute to disease resistance. *Molecular Plant-Microbe Interactions,* 25**,** 379-392.
TAKKEN, F. L. W., ALBRECHT, M. & TAMELING, W. I. L. 2006. Resistance proteins: molecular switches of plant defence. *Current Opinion in Plant Biology,* 9**,** 383-390.
TAKKEN, F. L. W. & GOVERSE, A. 2012. How to build a pathogen detector: structural basis of NB-LRR function. *Current Opinion in Plant Biology,* 15**,** 375-384.
TAMELING, W. I. L. & TAKKEN, F. L. W. 2008. Resistance proteins: scouts of the plant innate immune system. *European Journal of Plant Pathology,* 121**,** 243-255.
TENTHOREY, J. L., HALOUPEK, N., LÓPEZ-BLANCO, J. R., GROB, P., ADAMSON, E., HARTENIAN, E., LIND, N. A., BOURGEOIS, N. M., CHACÓN, P., NOGALES, E. & VANCE, R. E. 2017. The structural basis of flagellin detection by NAIP5: A strategy to limit pathogen immune evasion. *Science,* 358**,** 888-893.

TRAN, D. T. N., CHUNG, E.-H., HABRING-MÜLLER, A., DEMAR, M., SCHWAB, R., DANGL, J. L., WEIGEL, D. & CHAE, E. 2017. Activation of a plant NLR complex through Heteromeric Association with an autoimmune risk variant of another NLR. *Current Biology,* 27**,** 1148-1160.

TYLER, B. M., KALE, S. D., WANG, Q., TAO, K., CLARK, H. R., DREWS, K., ANTIGNANI, V., RUMORE, A., HAYES, T., PLETT, J. M., FUDAL, I., GU, B., CHEN, Q., AFFELDT, K. J., BERTHIER, E., FISCHER, G. J., DOU, D., SHAN, W., KELLER, N. P., MARTIN, F., ROUXEL, T. & LAWRENCE, C. B. 2013. Microbe-independent entry of oomycete RxLR effectors and fungal RxLR-Like effectors Into plant and animal cells Is specific and reproducible. *Molecular Plant-Microbe Interactions,* 26**,** 611-616.

UEDA, H., YAMAGUCHI, Y. & SANO, H. 2006. Direct interaction between the tobacco mosaic virus helicase domain and the ATP-bound resistance protein, N factor during the hypersensitive response in tobacco plants. *Plant Molecular Biology,* 61**,** 31-45.

VAJJHALA, P. R., VE, T., BENTHAM, A., STACEY, K. J. & KOBE, B. 2017. The molecular mechanisms of signaling by cooperative assembly formation in innate immunity pathways. *Molecular Immunology,* 86**,** 23-37.

VAN DER BIEZEN, E. A. & JONES, J. D. 1998. Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochemical Sciences,* 23**,** 454-456.

VAN OOIJEN, G., MAYR, G., KASIEM, M. M. A., ALBRECHT, M., CORNELISSEN, B. J. C. & TAKKEN, F. L. W. 2008. Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany,* 59**,** 1383-1397.

VE, T., VAJJHALA, P. R., HEDGER, A., CROLL, T., DIMAIO, F., HORSEFIELD, S., YU, X., LAVRENCIC, P., HASSAN, Z., MORGAN, G. P., MANSELL, A., MOBLI, M., O'CARROLL, A., CHAUVIN, B., GAMBIN, Y., SIERECKI, E., LANDSBERG, M. J., STACEY, K. J., EGELMAN, E. H. & KOBE, B. 2017. Structural basis of TIR-domain-assembly formation in MAL- and MyD88-dependent TLR4 signaling. *Nature Structural & Molecular Biology,* 24**,** 743-751.

VE, T., WILLIAMS, S. J. & KOBE, B. 2015. Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis,* 20**,** 250-261.

WANG, C.-I. A., GUNCAR, G., FORWOOD, J. K., TEH, T., CATANZARITI, A.-M., LAWRENCE, G. J., LOUGHLIN, F. E., MACKAY, J. P., SCHIRRA, H. J., ANDERSON, P. A., ELLIS, J. G., DODDS, P. N. & KOBE, B. 2007. Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity. *Plant Cell,* 19**,** 2898-2912.

WAWRA, S., DJAMEI, A., ALBERT, I., NUERNBERGER, T., KAHMANN, R. & VAN WEST, P. 2013. In vitro translocation experiments with RxLR-reporter fusion proteins of Avr1b from Phytophthora sojae and AVR3a from Phytophthora infestans fail to demonstrate specific autonomous uptake in plant and animal cells. *Molecular Plant-Microbe Interactions,* 26**,** 528-536.

WAWRA, S., TRUSCH, F., MATENA, A., APOSTOLAKIS, K., LINNE, U., ZHUKOV, I., STANEK, J., KOŹMIŃSKI, W., DAVIDSON, I., SECOMBES, C. J., BAYER, P. & VAN WEST, P. 2017. The RxLR Motif of the Host Targeting Effector AVR3a of Phytophthora infestans Is Cleaved before Secretion. *Plant Cell,* 29**,** 1184-95.

WEN, J., ARAKAWA, T. & PHILO, J. S. 1996. Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Analytical Biochemistry,* 240**,** 155-166.

WILKINS, M. R., GASTEIGER, E., BAIROCH, A., SANCHEZ, J. C., WILLIAMS, K. L., APPEL, R. D. & HOCHSTRASSER, D. F. 1999. Protein identification and analysis tools in the ExPASy server. *Methods in Molecular Biology,* 112**,** 531-552.

WILLIAMS, S. J., SOHN, K. H., WAN, L., BERNOUX, M., SARRIS, P. F., SEGONZAC, C., VE, T., MA, Y., SAUCET, S. B., ERICSSON, D. J., CASEY, L. W., LONHIENNE, T., WINZOR, D. J., ZHANG, X. X., COERDT, A., PARKER, J. E., DODDS, P. N., KOBE, B. & JONES, J. D. G. 2014. Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science,* 344**,** 299-303.

WILLIAMS, S. J., SORNARAJ, P., DECOURCY-IRELAND, E., MENZ, R. I., KOBE, B., ELLIS, J. G., DODDS, P. N. & ANDERSON, P. A. 2011. An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular Plant-Microbe Interactions,* 24**,** 897-906.

WILLIAMS, S. J., YIN, L., FOLEY, G., CASEY, L. W., OUTRAM, M. A., ERICSSON, D. J., LU, J., BODEN, M., DRY, I. B. & KOBE, B. 2016. Structure and function of the TIR domain from the grape NLR protein RPV1. *Frontiers in Plant Science,* 7**,** 1850.

WIN, J., MORGAN, W., BOS, J., KRASILEVA, K. V., CANO, L. M., CHAPARRO-GARCIA, A., AMMAR, R., STASKAWICZ, B. J. & KAMOUN, S. 2007. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Cell,* 19**,** 2349-2369.

WIRTHMUELLER, L., MAQBOOL, A. & BANFIELD, M. J. 2013. On the front line: structural insights into plant-pathogen interactions. *Nature Reviews Microbiology,* 11**,** 761-776.

WU, C. H., ABD-EL-HALIEM, A., BOZKURT, T. O., BELHAJ, K., TERAUCHI, R., VOSSEN, J. H. & KAMOUN, S. 2017. NLR network mediates immunity to diverse plant pathogens.

Proceedings of the National Academy of Sciences, 114**,** 8113-8118.
XIAO, F., HE, P., ABRAMOVITCH, R. B., DAWSON, J. E., NICHOLSON, L. K., SHEEN, J. & MARTIN, G. B. 2007. The N-terminal region of Pseudomonas type III effector AvrPtoB elicits Pto-dependent immunity and has two distinct virulence determinants. *Plant Journal,* 52**,** 595-614.
YAN, N., CHAI, J. J., LEE, E. S., GU, L. C., LIU, Q., HE, J. Q., WU, J. W., KOKEL, D., LI, H. L., HAO, Q., XUE, D. & SHI, Y. G. 2005. Structure of the CED-4-CED-9 complex provides insights into programmed cell death in Caenorhabditis elegans. *Nature,* 437**,** 831-837.
YUAN, B., ZHAI, C., WANG, W. J., ZENG, X. S., XU, X. K., HU, H. Q., LIN, F., WANG, L. & PAN, Q. H. 2011. The Pik-p resistance to Magnaporthe oryzae in rice is mediated by a pair of closely linked CC-NBS-LRR genes. *Theoretical and Applied Genetics,* 122**,** 1017-1028.
YUAN, S., YU, X., TOPF, M., LUDTKE, S. J., WANG, X. & AKEY, C. W. 2010. Structure of an apoptosome-procaspase-9 CARD complex. *Structure,* 18**,** 571-583.
ZHAI, C., ZHANG, Y., YAO, N., LIN, F., LIU, Z., DONG, Z. Q., WANG, L. & PAN, Q. H. 2014. Function and interaction of the coupled genes responsible for Pik-h encoded rice blast resistance. *PLoS One,* 9**,** e98067.
ZHANG, J., LI, W., XIANG, T., LIU, Z., LALUK, K., DING, X., ZOU, Y., GAO, M., ZHANG, X., CHEN, S., MENGISTE, T., ZHANG, Y. & ZHOU, J.-M. 2010. Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a Pseudomonas syringae effector. *Cell Host & Microbe,* 7**,** 290-301.
ZHANG, L., CHEN, S., RUAN, J., WU, J., TONG, A. B., YIN, Q., LI, Y., DAVID, L., LU, A., WANG, W. L., MARKS, C., OUYANG, Q., ZHANG, X., MAO, Y. & WU, H. 2015. Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization. *Science,* 350**,** 404-409.
ZHANG, X., BERNOUX, M., BENTHAM, A. R., NEWMAN, T. E., VE, T., CASEY, L. W., RAAYMAKERS, T. M., HU, J., CROLL, T. I., SCHREIBER, K. J., STASKAWICZ, B. J., ANDERSON, P. A., SOHN, K. H., WILLIAMS, S. J., DODDS, P. N. & KOBE, B. 2017. Multiple functional self-association interfaces in plant TIR domains. *Proceedings of the National Academy of Sciences,* 114**,** E2046-E2052.
ZHOU, M., LI, Y., HU, Q., BAI, X.-C., HUANG, W., YAN, C., SCHERES, S. H. W. & SHI, Y. 2015. Atomic structure of the apoptosome: mechanism of cytochrome c- and dATP-mediated activation of Apaf-1. *Genes & Development,* 29**,** 2349-2361.
ZHU, M. F., SHAO, F., INNES, R. W., DIXON, J. E. & XU, Z. H. 2004. The crystal structure of Pseudomonas avirulence

protein AvrPphB: A papain-like fold with a distinct substrate-binding site. *Proceedings of the National Academy of Sciences,* 101**,** 302-307.
ZIPFEL, C. 2009. Early molecular events in PAMP-triggered immunity. *Current Opinion in Plant Biology,* 12**,** 414-420.