EXPLORING THE SOURCE OF PLANT NB-LRR DISEASE RESISTANCE PROTEIN SIGNALING: ROLES OF CC AND NB DOMAINS IN PATHOGEN PERCEPTION AND DEFENSE INDUCTION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sarah May Collier

May 2011

EXPLORING THE SOURCE OF PLANT NB-LRR DISEASE RESISTANCE PROTEIN SIGNALING: ROLES OF CC AND NB DOMAINS IN PATHOGEN PERCEPTION AND DEFENSE INDUCTION

Sarah May Collier, Ph.D.

Cornell University 2011

Nucleotide binding, leucine-rich repeat (NB-LRR) proteins constitute a major component of plant resistance to infection by pathogens. While the mechanisms through which NB-LRR proteins are able to perceive pathogen invasion are becoming increasingly well understood, the means by which these proteins are able to translate pathogen detection into the induction of a conserved resistance response in the plant remains uncertain. N-terminal domains have long been assumed to function in NB-LRR protein signal initiation, though only limited evidence exists in support of this hypothesis. Here, we report that the NB domain, rather than the N-terminal CC (coiled-coil) domain of the potato (*Solanum tuberosum*) NB-LRR protein Rx is sufficient for the induction of defense responses, suggesting that the point of Rx signal initiation resides within the NB domain. We have identified regions of the Rx CC domain central to mediating its interaction with the hypothesized Rx recognition co-factor RanGAP2, and regions of the Rx NB domain which are critical for signal initiation. We further report the capability of the Rx and other NB domains to oligomerize, both homotypically and heterotypically. Among NB-LRR proteins able to participate in NB – NB heterotypic oligomerization are members of the CC_R-NB-LRR clade of NB-LRR proteins. We herein present the characterization of this unique and highly conserved class of NB-LRR protein, distinguished by having CC domains most closely

resembling the *Arabidopsis* RPW8 protein. We have found that the CC_R domain is independently capable of defense response induction, setting CC_R-NB-LRR proteins apart from canonical CC-NB-LRR proteins functionally as well as phylogenetically. We additionally report the striking correlation between the occurrence of NRG1-like CC_R-NB-LRR-encoding genes and TIR- (Toll and Interleukin-1 receptor homology) NB-LRR-encoding genes across the genomes of both monocotyledonous and dicotyledonous plant species. Taken together, the findings presented here suggest a functional correlation between CC_R-NB-LRR proteins and canonical NB-LRR proteins, with NB-mediated oligomerization presenting one possible means of realizing this relationship.

BIOGRAPHICAL SKETCH

Sarah M. Collier was born May 18, 1982, to mother Katherine Larson and father Tracy Collier, and along with younger brother Christopher grew up on a small farm on Bainbridge Island, Washington. It is to this upbringing that she owes her passion for plants and for agriculture. After graduating from Bainbridge High School, she attended the University of Washington in Seattle, where she majored in Botany and minored in Norwegian, graduating summa cum laude in 2004 with a Bachelor of Science. While at UW she worked as a research assistant in the plant phylogenetics laboratory of Richard Olmstead, receiving significant training and guidance. During her senior undergraduate year, Sarah studied at the Agricultural University of Norway, in Ås. This afforded an invaluable introduction to agricultural plant science, in addition to an irreplaceable cultural experience.

In 2005 Sarah came to Cornell University in Ithaca, New York, to join the Graduate Field of Plant Breeding. She was initially affiliated with the research group of Molly Jahn, but after Professor Jahn moved from Cornell Sarah joined the laboratory of Peter Moffett at the Boyce Thompson Institute for Plant Research, studying the function of plant disease resistance proteins. After the additional departure of Dr. Moffett from Ithaca, Sarah opted to remain in the lab at BTI, under the joint guidance of Dr. Moffett and Daniel Klessig. Upon completing her Ph.D., she expects to either continue a research career in plant – microbe interactions and crop improvement, or to follow her passion for the intersection of science and policy as it relates to agricultural and environmental sustainability.

for Gene B. Larson and Eleanor A. Collier

I didn't know you well, but I know how much you gave me

and for my parents, who are amazing

ACKNOWLEDGEMENTS

This work would not have been possible without the financial support of the National Science Foundation, both in the form of research grants to the Moffett Lab, and also as a Graduate Research Fellowship to support my work. I am also thankful for funding provided by a Cornell Fellowship, which allowed me to rotate long enough to find my home in the Moffett Lab.

I am grateful to the faculty and staff of the Department of Plant Breeding & Genetics and the Boyce Thompson Institute, and most especially to Molly Jahn, for providing such excellent training and helping to set me on my course. Above all I am grateful to Peter Moffett for taking me into his research group, for providing invaluable training, guidance, and support, and for imbuing every aspect of our work with a genuine passion for science. I am additionally indebted to Gregory Martin, Daniel Klessig, and Walter De Jong, who stepped in and performed above and beyond the expectations of minor committee members when things got really unscripted. It would have been a hard road without their support, guidance, and patience. My thanks also to many past and present members of the Jahn, Moffett, Klessig, and Martin labs for teaching me so much.

The path that brought me to science and to Cornell was laid by many people, chief among them my father, two excellent K-12 science teachers Terri Peterson and Bim Prince, and undergraduate advisors Richard Olmstead and Åsmund Bjørnstad. For their influence I am grateful. Last but decidedly not least I am so thankful for the constant support of my family and friends. Mom, Dad, Chris, and my almost-sister Heather, you are the best companions a person could ask for in life!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ARC	Apaf-1, R proteins, CED-4 homology
CC	coiled-coil
CP	coat protein
CC_R	CC RPW8-like
EGFP	enhanced green florescent protein
FL	full-length
GFP	green florescent protein
HR	hypersensitive response
IB	immunoblot
IP	immunoprecipitation
LRR	leucine-rich repeat
NB	nucleotide binding
PVX	*Potato Virus X*
RNBS	resistance NBS motif
TIR	Toll and Interleukin-1 receptor homology
TMV	*Tobacco Mosaic Virus*
TRV	*Tobacco Rattle Virus*
VIGS	virus-induced gene silencing
WT	wild-type

LIST OF SYMBOLS

α	anti- (of antibodies)
Δ	deletion (of preceding protein fragment)
Θ	mock treatment

CHAPTER 1

NB-LRRs work a "bait and switch" on pathogens*

ABSTRACT

Plant genomes encode large numbers of highly variable nucleotide binding leucine-rich repeat (NB-LRR) disease resistance proteins. These proteins have been studied extensively to understand their evolution and the molecular basis of their function. Multiple studies indicate that the C-terminal LRR domain plays a pivotal role in defining pathogen recognition specificity. However, a growing body of evidence suggests that the N-termini of NB-LRR proteins also function in pathogen recognition. To formulate a framework that can explain the underlying principles governing NB-LRR function while accommodating findings from different experimental systems, we present a "bait and switch" model. This model proposes a two-step recognition process involving interactions with both cellular co-factors (bait) and the LRR domain, which in turn activates the molecular switch leading to disease resistance.

---

* This chapter, with the exception of its final paragraph, has been reformatted and reprinted from: *Trends in Plant Science*, Vol. 14, Collier, S.M. and Moffet, P., NB-LRRs work a "bait and switch" on pathogens, pp. 521-529, Copyright (2009), with permission from Elsevier. All figures presented in this chapter were generated by S.M.C.

NBLing away at pathogens

Plants employ a network of intertwined mechanisms to counter infection by pathogens and parasites. One such line of defense is based on dominant disease resistance (R) genes which mediate resistance to pathogens possessing corresponding avirulence (Avr) genes (Jones and Dangl 2006). While the protein products of Avr genes have little in common structurally, the majority of characterized R genes encode NB-LRR proteins. NB-LRR-encoding genes make up one of the largest and most variable gene families found in plants, with most plant genomes containing several hundred family members (Sacco and Moffett 2009).

Two broad questions have dominated research in this area: *(i)* how do NB-LRR proteins function, and *(ii)* what biological phenomena have led to their evolution and proliferation? In addressing these questions, it is important to note that there are no particular characteristics that distinguish between NB-LRR proteins mediating resistance to different classes of pathogens, including viruses, bacteria, fungi, oomycetes, nematodes, and insects. Very similar R proteins are capable of recognizing completely different pathogens, and the resistance response initiated is not specific to the class of pathogen detected (McDowell et al. 1998; Rossi et al. 1998; Tobias et al. 1999; Bendahmane et al. 2000; Takahashi et al. 2002; Nombela et al. 2003; Rentel et al. 2008). Thus, it stands to reason that the function of NB-LRR proteins is governed by an underlying set of mechanisms, variations of which might manifest differently depending on experimental context. Molecular analysis of NB-LRR function has in recent years been an area of intense inquiry – some of which has yielded seemingly contradictory data regarding the nature of Avr perception. This article reviews recent findings in the field with the objective of formulating a mechanistic framework to accommodate certain common and contrasting themes reported in various NB-LRR

experimental systems. This model in turn highlights a number of unanswered questions regarding NB-LRR protein function and the biological forces that have shaped the evolution of this pathogen surveillance system.

NB-LRR architecture

NB-LRR proteins are so named for the presence of a conserved central nucleotide binding (NB) domain and a more variable C-terminal leucine-rich repeat (LRR) domain. Between these two domains is a region of homology known as the ARC (Apaf-1, R proteins, CED-4) domain (van der Biezen and Jones 1998). The ARC domain can be further subdivided into two structurally and functionally distinct units; ARC1 and ARC2 (Albrecht and Takken 2006; Rairdan and Moffett 2006) (Figure 1.1). Together, the NB and ARC domains comprise a functional nucleotide binding pocket (Tameling et al. 2002). With respect to their N-termini, two major groups can be distinguished within the NB-LRR proteins: those that possess a TIR (Toll and Interleukin-1 Receptor homology) domain, and those that do not. The two groups are also easily and most precisely distinguished by characteristic motifs within their NB and ARC domains (Meyers et al. 1999). NB-LRR proteins lacking TIR domains may possess one of a number of variable N-terminal domains, including some predicted to form coiled-coils (CC), and this class is traditionally referred to as CC-NB-LRR proteins. However, the CC domain is sometimes joined and/or replaced by a Solanaceae domain (Mucyn et al. 2006) (SD) or a predicted BED DNA binding domain (Bai et al. 2002; Tuskan et al. 2006) (B). Alternatively, a number of predicted NB-LRR proteins of both the CC- and TIR-class (as defined by the configuration of their NB-ARC domains) appear to possess no sequence N-terminal to the NB domain (Absent) ((Porter et al. 2009) and references therein) (Figure 1.1). The N-termini of CC-NB-LRR

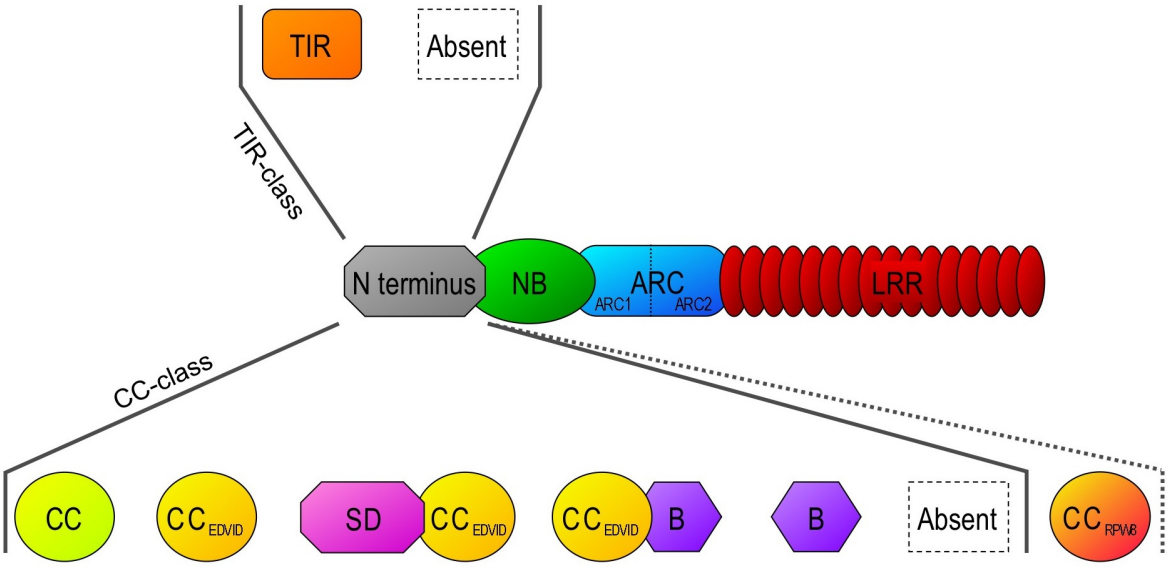


Figure 1.1 *Typical linear architecture of plant NB-LRR proteins, with common N-terminal variations.*

proteins not immediately related by descent show little sequence or structural similarity. Even within the loosely defined CC domains the only apparently widely conserved feature is a small "EDVID" motif, which defines the largest subclass of CC domain (Rairdan et al. 2008). Other non-EDVID CC variants also exist, however, including that of NRG1, a tobacco NB-LRR protein that is not itself an R protein but is required for the function of the N TIR-NB-LRR (Peart et al. 2005). Interestingly, the NRG1 CC domain bears closest similarity to that of RPW8 – a non-NB-LRR R protein of Arabidopsis that provides broad-spectrum resistance against powdery mildew (Xiao et al. 2001) – and unlike other "CC" domains, is able to induce cell death upon over-expression (Peart et al. 2005).

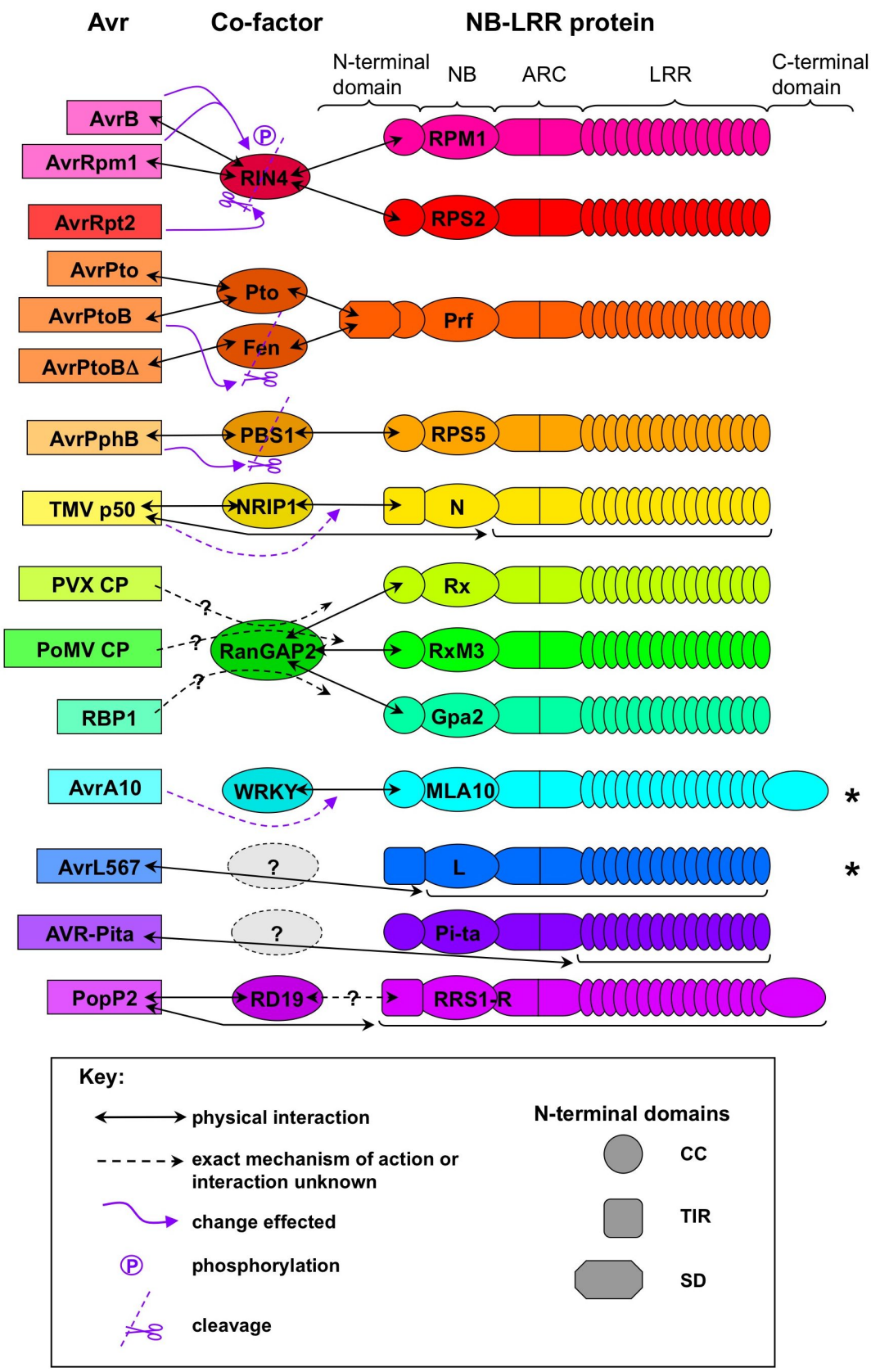
A brief history of NB-LRR models

Molecular genetic studies of closely related NB-LRR proteins with differing recognition specificities indicate that the LRR domain is the primary determinant of NB-LRR protein recognition specificity (Shen et al. 2003; Qu et al. 2006; Rairdan and Moffett 2006; Ellis et al. 2007b). Furthermore, the LRR domains of NB-LRR proteins appear to be under diversifying selection (McDowell and Simon 2006). Taken together with the known role of LRR domains in protein-protein interactions, this originally led to the expectation that Avr and NB-LRR proteins might undergo receptor-ligand interactions. However, early attempts failed to demonstrate such direct interaction and supported the formulation of the guard hypothesis, which suggests that host proteins targeted during pathogenesis are monitored, or guarded, by NB-LRR proteins (Van der Biezen and Jones 1998; Dangl and Jones 2001). In this model the plant

perceives the action of a pathogen effector protein upon host virulence target molecules, and hence these targets, or guardees, become co-factors in recognition. Indeed, subsequent studies have identified a number of cases where recognition appears to be mediated through a cellular co-factor that interacts with both Avr and cognate R protein(s) (Figure 1.2). One major prediction of the model is that all guardees are virulence targets, however this prediction has yet to be unambiguously demonstrated. Additionally, the guard hypothesis is not easily applied to less outwardly hostile Avr proteins from viruses, such as coat proteins and replicases, which have no known function in actively interfering with cellular processes (Moffett and Klessig 2008). A lack of evidence for virulence targeting of recognition co-factors has prompted a broadening of the interpretation of their role; from guardees to molecular decoys (van der Hoorn and Kamoun 2008). The decoy model is based on the supposition that selection pressures have favored recognition co-factors that continue to interact with Avr proteins but are not under pressure to avoid targeting by pathogen-derived proteins. That is, a recognition co-factor does not have to be a true virulence target - it just has to look like one.

A case for the decoy model can be made with the tomato (*Solanum lycopersicum*) Ser/Thr kinase Pto, which interacts with the NB-LRR protein Prf and mediates recognition of the AvrPto and AvrPtoB proteins of *Pseudomonas syringae*. The actual virulence targets of AvrPto and AvrPtoB, however, appear to include the kinase domains of the receptor-like kinases CERK1, BAK1, EFR1, and FLS2, rather than the kinase domain of Pto (Göhre et al. 2008; Shan et al. 2008; Xiang et al. 2008; Gimenez-Ibanez et al. 2009). This could be interpreted as an indication that Pto is being used as a kind of bait by Prf to interact with effector proteins that normally target other kinases. In this context, we propose that the use of such recognition co-factors as bait

Figure 1.2 *Recognition co-factors in NB-LRR/Avr interactions.* Schematic of characterized physical and functional interactions, as well as hypothesized and partially characterized interactions in otherwise well-characterized systems. Colors denote components of the same or related systems. Known interactions are depicted by solid arrows whereas functional interactions of unknown mechanism are depicted as dashed lines. The association of NRIP1 with N, and WRKY with MLA10, is observed only in the presence of their corresponding Avr proteins. Dashed-line ovals represent hypothetical co-factors yet to be identified *in planta*. Asterisks denote systems containing multiple known alleles with differing recognition specificities such as the *L5*, *L6* and *L7* alleles of *L*, which confer recognition of different version of AvrL567. References, by R protein: RPM1 (Mackey et al. 2002); RPS2 (Axtell and Staskawicz 2003; Day et al. 2005); Prf (Mucyn et al. 2006; Rosebrock et al. 2007; Mucyn et al. 2009); RPS5 (Shao et al. 2003; Ade et al. 2007); N (Ueda et al. 2006; Caplan et al. 2008); Rx (Sacco et al. 2007; Tameling and Baulcombe 2007); RxM3 (Farnham and Baulcombe 2006); Gpa2 (Sacco et al. 2007; Sacco et al. 2009); MLA10 (Shen et al. 2007); L (Dodds et al. 2006; Ellis et al. 2007a); Pi-ta (Jia et al. 2000); RRS1-R (Deslandes et al. 2003; Bernoux et al. 2008).


Avr
Co-factor
NB-LRR protein
N-terminal domain
NB
ARC
LRR
C-terminal domain
AvrB
AvrRpm1
AvrRpt2
RIN4
RPM1
RPS2
AvrPto
AvrPtoB
AvrPtoBΔ
Pto
Fen
Prf
AvrPphB
PBS1
RPS5
TMV p50
NRIP1
N
PVX CP
PoMV CP
RBP1
RanGAP2
Rx
RxM3
Gpa2
AvrA10
WRKY
MLA10
AvrL567
L
AVR-Pita
Pi-ta
PopP2
RD19
RRS1-R
Key:
physical interaction
exact mechanism of action or interaction unknown
change effected
phosphorylation
cleavage
N-terminal domains
CC
TIR
SD

is a common mechanism employed by NB-LRR proteins to sense Avr proteins and activate a molecular switch that results in the induction of resistance responses. As such, the bait and switch model provides a mechanistic explanation of how NB-LRR proteins translate pathogen recognition into defense activation, and is compatible with current conceptualizations of how this recognition system evolved.

The N-terminus: setting the bait

The central importance of the LRR domain in determining recognition specificity would lead to the expectation that either Avr proteins or recognition co-factors bind to this domain. At the same time, it has been generally assumed that the N-terminal domains of NB-LRR proteins function in the initiation of signaling due to similarities in domain structure between plant NB-LRR and animal NACHT-LRR proteins (Fritz et al. 2006; Rairdan and Moffett 2007). However, all plant proteins with either confirmed or proposed roles as co-factors in facilitating Avr recognition have been shown to interact with the N-termini of their corresponding NB-LRR proteins. In many cases, activation of the NB-LRR protein results when the cognate Avr protein interacts with, and in some cases physically alters, the co-factor (Figure 1.2). By extension, we have suggested that the Rx coiled-coil domain (CC)-interacting protein RanGAP2 may also function in Avr recognition, independent of its roles in mitosis and nucleocytoplasmic trafficking (Sacco et al. 2007; Xu et al. 2008). One proposed exception is the case of the barley (*Hordeum vulgare*) MLA10 protein whose CC domain interacts with HvWRKY1 and HvWRKY2 in yeast (*Saccharomyces cerevisiae*), and *in planta* in the presence of Avr_{A10} (Shen et al. 2007). Although it has been proposed that MLA10 induces defense signaling through its interaction with these WRKY

transcription factors, it remains equally plausible that HvWRKY proteins mediate recognition of Avr proteins by MLA10 (Figure 1.2).

In Rx, the sole well-conserved motif present in the CC domain (EDVID) mediates an intramolecular interaction, whereas binding to RanGAP2 is mediated by poorly conserved sequences flanking the EDVID motif as well as at the extreme N-terminus (Rairdan et al. 2008). This suggests that whereas *intra*molecular interactions within NB-LRR proteins would be expected to be conserved, the *inter*molecular interactions of N-terminal domains may be much more diverse. Indeed, the N-terminal-interacting proteins shown in Figure 1.2 include a protein of unknown molecular function (RIN4), representatives from two different subfamilies of Ser/Thr protein kinases (Pto and Fen vs. PBS1), a chloroplast localized sulfurtransferase (NRIP1), a Ran GTPase activating protein (RanGAP2), a Cys protease (RD19), and a transcription factor (HvWRKY). Thus, the diversity of recognition co-factors may expand the variety of pathogen proteins with the potential to be recognized by NB-LRR proteins.

Despite the recent appreciation of the role of N-termini as important points of Avr interaction, in several systems it remains equally clear that the LRR also participates in physical interactions with Avr proteins consistent with its genetically established role in determining recognition specificity (reviewed in (Rairdan and Moffett 2007)). Notably, the L, Pi-ta, and RRS1-R NB-LRR proteins have all been reported to bind directly to their cognate Avr proteins in yeast two-hybrid assays (Figure 1.2) (Jia et al. 2000; Deslandes et al. 2003; Dodds et al. 2006; Ellis et al. 2007a), and the Pi-ta LRR domain was shown to be sufficient for interaction with AvrPi-ta (Jia et al. 2000). In the case of L, binding requires the NB-ARC-LRR portion of the protein and interactions between L and AvrL variants correspond to differences in the LRR

domains encoded by L alleles with different recognition specificities (Ellis et al. 2007a; Ellis et al. 2007b). These examples underscore the need for any widely applicable model of NB-LRR protein function to account for the role of the LRR domain in recognition.

Accumulating bodies of evidence point to *(i)* the N-terminus as an important point of Avr interaction, *(ii)* the LRR domain as the primary determinant of Avr recognition specificity, and *(iii)* direct interactions between LRR domains and Avr proteins. Although these observations arise from different experimental systems, they do not necessarily imply that different NB-LRR proteins employ distinct mechanisms, but may instead represent variations on a theme. The key to reconciling the seemingly contrasting roles for the N-terminal and LRR domains in recognition may lie in studies of the tobacco (*Nicotiana tabacum*) TIR-NB-LRR protein N. N was first observed to bind directly to its cognate Avr, the p50 subunit of the *Tobacco Mosaic Virus* replicase, in yeast two-hybrid and *in vitro* studies (Ueda et al. 2006), while subsequent *in planta* data revealed a requirement for the NRIP1 co-factor, which binds both p50 and the N TIR domain (Caplan et al. 2008). These observations might be integrated through the application of a two-step mechanism involving an initial interaction between NRIP1 and p50 which would facilitate a subsequent interaction with, and recognition by, the N LRR domain. The former interaction may, however, be dispensable if the proteins of interest are forced together under artificial conditions, such as in the yeast two-hybrid system. Consistent with studies of L variants (Dodds et al. 2006; Ellis et al. 2007a), the N LRR domain appears to be more discriminating than its bait. Whereas the NRIP1 protein interacts with both recognized and non-recognized p50 variants, the N LRR domain interacts only with the recognized version of p50 (Ueda et al. 2006; Caplan et al. 2008). These are important observations, considering that very similar

NB-LRR proteins may recognize different Avr proteins, yet at the same time may have identical N-termini that could interact with the same recognition co-factors (Takahashi et al. 2002; Shen et al. 2003; Rairdan and Moffett 2006; Ellis et al. 2007b; Sacco et al. 2007), and by extension the same set of Avr proteins. Thus, we propose in the bait and switch model that although the bait protein may be the initial point of contact, possibly with a range of pathogen-derived proteins, the LRR domain ultimately determines which of these interactions will flip the molecular switch within the NB-LRR protein, leading to initiation of defense responses.

How is the molecular switch wired?

The NB-ARC region is thought to play a critical role in regulating NB-LRR function. Furthermore, the Rx NB domain alone is sufficient to induce defense responses (Rairdan et al. 2008), suggesting that the region of the protein which ultimately interacts with downstream signaling molecules resides within this domain. The NB-ARC nucleotide binding pocket contains a number of motifs highly conserved throughout members of the STAND (signal transduction ATPases with numerous domains) class of proteins (Meyers et al. 1999; Leipe et al. 2004). Most notably, the P-loop of the NB domain, the MHDV motif of the ARC2, and possibly the GLPL motif of the ARC1 domain are predicted to coordinate nucleotide binding, while the kinase 2 motif of the NB domain is thought to be a key mediator of ATP hydrolysis (McHale et al. 2006; Tameling et al. 2006). Additionally, plant NB-ARC domains contain a number of motifs of unknown function (Meyers et al. 1999; Rairdan and Moffett 2007), which present appealing candidates as sites of NB-mediated signal initiation.

The predicted position of the ARC2 domain suggests that it is ideally situated to transmit the message of Avr perception by the LRR to the nucleotide binding

pocket. Notably, the MHDV motif is thought to be analogous to the sensor II motif of AAA+ ATPases (Takken et al. 2006) – which is believed to be involved in nucleotide-dependent conformational changes (Riedl et al. 2005). Indeed, different mutations within the MHDV motif can result in either loss of function or autoactivity, as do mutations elsewhere in the ARC2 domain (Takken et al. 2006), indicating that the ARC2 has a role both in activation and in autoinhibition. Given these observations, it appears that the ARC2 is a critical component of a molecular switch, relaying information regarding a recognition event into the initiation of a signaling event (Rairdan and Moffett 2006; Takken et al. 2006). The histidine residue of the MHDV motif is predicted to contact the bound nucleotide (Takken et al. 2006); thus any conformational change that affects the ARC2 domain could have the potential to affect the protein's nucleotide binding status. In turn, changes in nucleotide binding status are predicted to alter the state of the NB domain and consequently the protein-protein interaction properties of the NB-LRR protein (reviewed in (van Ooijen et al. 2007)).

Although the isolated NB-ARC domains of the Mi-1 and I-2 NB-LRR proteins can bind and hydrolyze ATP (Tameling et al. 2002), the nature of change in nucleotide binding status (i.e. hydrolysis of ATP to ADP, release of ADP, or nucleotide exchange of ADP for ATP) that takes place upon activation of the full-length molecule remains undetermined. Certain mutations of the kinase 2 motif of I-2 can result in NB-LRR autoactivity, and also reduce the rate of ATP hydrolysis *in vitro*, which may indicate that the ATP-bound state is the active conformation of the protein (Tameling et al. 2006; Takken and Tameling 2009). Alternatively, the ATP-bound state may represent a post-activated conformation following upon a transitory signal initiation event. Additionally, the presence of p50 increases ATP hydrolysis by N (Ueda et al. 2006), suggesting multiple rounds of ATP hydrolysis and nucleotide exchange, which would

presumably induce multiple signal initiation events, as previously suggested (Rairdan and Moffett 2006). Whatever the exact mechanism, however, it is clear that some change in nucleotide binding is critical to flipping the NB-LRR switch.

From recognition to signaling: flipping the switch

If recognition involves cooperation between the LRR and N-terminal domains, what then is the nature of the mechanism, or molecular switch, that translates recognition into signal initiation? Structure-function analyses of the potato (*S. tuberosum*) CC-NB-LRR protein Rx indicate that this switch requires some assembly. One necessity for proper Rx function is a physical interaction between the LRR and ARC1 domains. This interaction, as well as a functional nucleotide binding pocket, is itself a prerequisite for a separate intramolecular interaction between the Rx CC domain and the rest of the protein (Moffett et al. 2002; Rairdan et al. 2008). Both of these interactions are required for the NB-LRR protein to fold into a signaling-competent state, and once established appear to retain the protein in an auto-inhibited state until some stimulus induces it to overcome auto-inhibition and adopt its active form (Rairdan and Moffett 2006). In other words, the trap must be set before it can be sprung. In addition NB-LRR protein folding and intramolecular interactions also appear to require co-chaperones such as SGT1, RAR1, and HSP90 (Holt et al. 2005; Leister et al. 2005) which may provide additional safeguards preventing NB-LRR proteins from functioning before being properly assembled.

The interplay between the LRR and ARC2 domains appears to be critical to controlling the transition between auto-inhibited and activated forms of NB-LRR proteins. Numerous mutations throughout both the ARC2 and LRR domains are known to disrupt auto-inhibition, and experimental domain swapping between

related NB-LRR proteins reveals that improper pairing between the ARC2 and LRR domains can also lead to autoactivity (reviewed in (van Ooijen et al. 2007)). It thus appears that upon translation, NB-LRR proteins are folded into a conformation wherein a perfect fit between ARC2 and LRR retains the protein in a delicately balanced state of auto-inhibition, and alteration of this interface allows the NB-LRR protein to proceed to its active state. The presence of hypervariable regions within the ARC2 domains of several groups of NB-LRR homologues suggests that the exact nature of this balance may be unique to each R protein (Sun et al. 2001; Rairdan and Moffett 2006; Seah et al. 2007). By extension each ARC2/LRR interface may have a unique potential for perturbation by different stimuli.

Several examples suggest that bait proteins may play roles in NB-LRR function additional to recognition. A lack of RIN4 results in very poor accumulation of RPM1 (Mackey et al. 2002; Belkhadir et al. 2004) and Pto contributes to Prf accumulation, suggesting that these associations are ultimately required for the proper maturation of the NB-LRR proteins (Mucyn et al. 2006). Indeed, in the absence of RIN4, both RPM1 and RPS2 are constitutively active, suggesting a loss of auto-inhibition (Belkhadir et al. 2004). Pto appears to act as an integral regulatory unit of Prf in that Pto interaction is necessary for Prf to be competent for signaling, but Pto subsequently appears to retain Prf in an auto-inhibited state until the interaction is altered, although the absence of Pto does not lead to autoactivation of Prf (Mucyn et al. 2006; Balmuth and Rathjen 2007; Mucyn et al. 2009). Such a combination of positive and negative regulatory functions of bait proteins is reminiscent of the interaction between the LRR and ARC domains. Furthermore, NB-LRR activation via the disruption of an auto-inhibiting interaction between bait and NB-LRR would appear to be supported by the cases of RPS2 and RPS5 which are activated by the cleavage of their respective recognition co-

factors (Figure 1.2). The observation that the alteration of bait proteins, such as mutation of Pto or ablation of RIN4 (Mackey et al. 2002; Belkhadir et al. 2004; Mucyn et al. 2006), can result in autoactivity of their cognate NB-LRR binding partner raises the question of whether the NB-LRR protein actually recognizes a specific alteration, or absence, of its bait. Alternatively, this autoactivation may be analogous to that induced by mutations in, for example, the ARC2 domain which appear to cause NB-LRR autoactivity by interfering with autoinhibition. That is, if bait proteins play a role in retaining NB-LRR proteins inactive, then their artificial perturbation may mimic the activation process just as mutations within the NB-LRR proteins themselves cause autoactivation. Given that autoactivating mutations are found throughout the different domains of NB-LRR proteins (Bendahmane et al. 2002; Takken et al. 2006) this raises the question of whether there is only one way to activate a given NB-LRR protein. Alternately, flexibility in how NB-LRR proteins can be activated might increase the number of Avr proteins by which they are activated.

We propose that an underlying principle governing NB-LRR function is that bait proteins play an integral role in priming NB-LRR proteins to be functional while retaining the molecular switch in an inactive conformation until a structural change triggers release of auto-inhibition. This structural change may take multiple forms such as modification, removal, or conformational change of the bait protein or simply complex formation with the bait protein. At the same time, at least some bait proteins, such as NRIP1 and WRKY proteins, appear to interact with their cognate NB-LRR proteins only in the presence of the Avr protein (Figure 1.2). In these cases the priming of NB-LRR proteins might occur independently of bait proteins, or priming via interaction with a bait protein may occur concomitantly with alterations that cause activation. Structure-function analysis of the Rx protein indicates that its putative

recognition co-factor, RanGAP2, interacts with regions of the Rx CC domain surrounding the EDVID motif, which itself mediates interaction with the rest of the protein. This EDVID-based intramolecular interaction is dependent on motifs within the ARC2 and LRR domains that are also thought to be involved in nucleotide binding and the molecular switch mechanism (Rairdan et al. 2008). As such, it is conceivable that alterations of bait proteins could be perceived and relayed to the molecular switch via the N-terminal domain. Likewise, through the proximity induced by inter-domain interactions, some recognition co-factors may bring multiple Avr proteins directly into contact with the auto-inhibitory surfaces of the LRR/ARC2 interface. This might reinforce relatively weak, but specific, interactions between Avr and LRR domain which would in turn disrupt the auto-inhibitory effect of specific LRR/ARC2 combinations, as suggested above. Indeed, artificial tethering of RanGAP2 to RBP-1, the Avr protein recognized by Gpa2, results in enhanced activation of Gpa2 (Sacco et al. 2009), suggesting that RanGAP2 might function by facilitating interactions between the R and Avr proteins.

Variations on this theme may explain why some NB-LRR systems (such the Rx/Gpa2, L, and MLA proteins) possess a high degree of recognition specificity between paralogs and allelic variants (Shen et al. 2003; Ellis et al. 2007b), presumably mediated by a common co-factor interacting with multiple similar NB-LRR proteins. At the same time, other cases exist (such as Prf and RPM1) in which a single NB-LRR is activated by multiple Avr proteins which interact with a common bait protein (Figure 1.2). Activation by perturbation of the LRR/ARC2 interface is likely translated into alterations in nucleotide binding status via the MHDV motif of the ARC2. As signal initiation is thought to originate from the NB domain of CC-NB-LRR proteins, it

is conceivable that recognition-induced alterations within the nucleotide binding pocket might release NB signaling potential (Figure 1.3).

While changes within the nucleotide binding pocket ultimately resulting in NB-LRR activity is a well supported concept, it remains to be determined whether signal initiation originates from the NB domain in TIR- as well as CC-NB-LRR proteins. The two classes appear to signal through different pathways (Aarts et al. 1998), and possess characteristic differences within the NB domain (Meyers et al. 1999). While no CC domain has been reported to possess signaling functions on its own, the TIR domains encoded by the flax (*Linum usitatissimum*) *L10* allele as well as the *Arabidopsis* (*A. thaliana*) *RPS4* and *RPPIA* genes have been shown to be sufficient to induce defense responses (Frost et al. 2004; Swiderski et al. 2009). Interestingly however, the flax L6 and L7 proteins, which differ only in their TIR domains, recognize the same versions of AvrL567, but are differentially suppressed by the flax rust inhibitor gene, suggesting that the pathogen can influence recognition specificity through the TIR domain. Together with the documented role of the N TIR domain in recognition co-factor binding, this may indicate a dual role for the TIR domain in signaling and recognition. Indeed, we propose that the ultimate function of the switch mechanism is to alter the protein-protein interaction properties of NB-LRR proteins. Thus, it is possible that a conserved switch mechanism which passes from the LRR/ARC2 interface to the nucleotide binding pocket could alternately modify intramolecular interactions such that signaling could occur via the TIR domain as well.

Evolutionary perspectives

All NB-LRR proteins presumably derive from a common ancestor (Meyers et al. 1999), and this primordial NB-LRR protein may have functioned on the basis of either

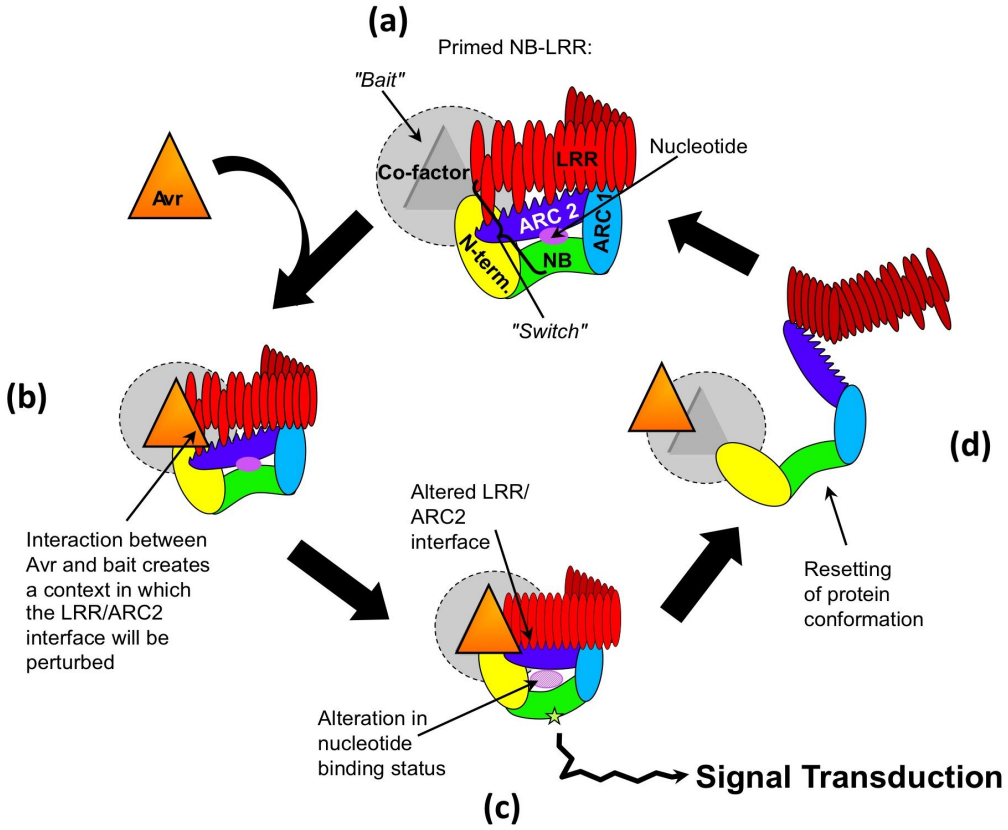


Figure 1.3 *Bait and switch model of NB-LRR protein function.* **(a)** The NB-LRR protein is depicted in its primed (signaling competent) but auto-inhibited (restrained from signaling) state – the "switch" is in the "off" position. This state requires a functional nucleotide binding pocket (presumably bound to ATP or ADP) and multiple intramolecular interactions including a fine-tuned balance between the LRR and ARC2. The recognition co-factor bait may also be a critical component of the primed state. Due to inter-domain interactions, the bait protein might easily be in contact with the recognition and signaling components of the C-terminal half of the protein. **(b)** The pathogen Avr protein is brought into the NB-LRR system via the bait protein, either through direct binding or through alteration to the bait. **(c)** The switch is flipped. Presence of the Avr protein causes perturbation and subsequent conformational changes to the LRR/ARC2 interface. In some cases, this perturbation may be induced by direct and specific interaction between the Avr and LRR domain, facilitated by the bait protein. This change is relayed through the ARC2 into an alteration in nucleotide binding status within the nucleotide binding pocket, which in turn causes further conformational changes, ultimately resulting in signal initiation. One proposed mechanism of signal initiation is that conformational changes within the nucleotide binding pocket allow signaling motif(s) within the NB domain (green star) to associate with downstream signaling components. **(d)** Concurrent with or subsequent to signaling, intramolecular interactions within the NB-LRR protein are likely dissociated and reset, possibly allowing the protein to undergo repeated rounds of recognition, signaling, and resetting given the continued presence of Avr protein (Rairdan and Moffett 2006).

direct or indirect interaction with pathogen-derived molecules. Upon expansion of this protein family, NB-LRR proteins may have co-opted various recognition co-factors which, combined with diversified LRR domains, may have offered a more complex recognition capacity. It has been suggested that a shift in recognition co-factors from guardees to decoys might provide increased evolutionary stability (van der Hoorn and Kamoun 2008), and additional examples are necessary to determine whether guardees or decoys (or both) are the general rule. However, as the bait and switch model is concerned with how bait proteins facilitate recognition and signaling, regardless of whether they function as canonical guardees or decoys, it is compatible with either evolutionary model. At the same time, understanding of the mechanistics of the proteins involved can inform models of how this system has evolved.

The fine balance between the LRR and ARC2 domains may provide NB-LRR systems with additional permutations in recognition specificity. NB-LRR variants whose LRR/ARC2 interface is too unstable or is perturbed by a host protein would be lethal. However, LRR/ARC2 variants perturbed only by pathogen Avr proteins or as a consequence of Avr protein presence or activity would fall under positive selection. At the same time, the number and diversity of NB-LRR proteins within and between species suggests that selective pressure acting on the system as a whole to generate and maintain diversity may be an integral component of the plant innate immune system. While the necessity to bind independently to Avr proteins would put additional and stringent constraints upon NB-LRR protein evolution, binding through recognition co-factors would relieve this constraint – leaving the LRR/ARC2 interface free to evolve solely for perturbation, while physical interaction would be through N-terminally bound co-factors. Such a system allows for greater versatility in R protein

evolution, as in the case of Rx and Gpa2, where multiple Avr proteins presumably interact initially through a common co-factor, but where recognition is achieved only through more specific secondary interactions mediated by the LRR domain (Rairdan and Moffett 2006; Sacco et al. 2007). Indeed, in an important proof of principle, the Rx protein has been randomly mutated within its LRR domain to generate a version of Rx (RxM3) that has an expanded recognition repertoire (Farnham and Baulcombe 2006). However, as its CC domain remains unaltered, RxM3 would still bind the same putative bait protein RanGAP2 (Figure 1.2).

N-terminally mediated interactions with a variety of recognition co-factors accounts for the diversity observed across NB-LRR N-termini, and also expands the quantity of potential Avr proteins available to NB-LRR recognition. By utilizing recognition co-factor bait proteins, each of which may posses the ability to interact with multiple pathogen molecules, a finite number of plant R proteins might be able to detect a far greater quantity of pathogen Avr proteins. Indeed, such appears to be the case with RPM1, RPS2, and Prf, where multiple Avr proteins are detected via their interaction with common co-factors (Figure 1.2) (Kim et al. 2009). This might also explain how the Mi1.2 NB-LRR protein of tomato is able to recognize not only root-knot nematodes (*Meloidogyne* spp.) but also the potato aphid (*Macrosiphum euphorbiae*) and the tobacco whitefly (*Bemisia tabaci*) (Rossi et al. 1998; Nombela et al. 2003). Such a scenario would also allow allelic diversification, resulting in similar NB-LRR proteins with the capacity to recognize series of Avr proteins. At the same time, it is possible that incompatibilities with bait proteins in different genetic backgrounds may compromise NB-LRR protein function or cause autoactivation (Bomblies and Weigel 2007; Alcázar et al. 2009).

The retention of a conserved switching mechanism would also be evolutionarily advantageous, and would suggest relatively conserved signaling pathways. Constant adaptation of bait proteins and the LRR domain allows diverse Avr proteins to be brought into a system where, due to a conserved nucleotide binding pocket-based switch, their detection can be translated into the initiation of a single or limited number of signaling pathways. Thus the bait and switch model helps to further elucidate how NB-LRR proteins are able to detect a myriad of diverse pathogen proteins and translate that detection into a standardized and timely defense response. The bait and switch model also underscores a number of outstanding questions regarding NB-LRR protein function and evolution, including the following:

- Are there constant principles governing what types of interactions will result in NB-LRR activity, or are the conditions of recognition distinct for every system? To answer these questions, more systems comprised of closely related NB-LRR proteins with known co-factors and different recognition specificities for known Avr proteins must be characterized which can provide additional examples of recognition specificity determinants.
- Is there a conserved repertoire of recognition co-factors? If so, what properties make them good baits?
- Given that some NB-LRR/bait interactions are induced, are these co-factors still integral to NB-LRR priming and regulation?
- What nucleotide binding status corresponds to primed vs. active NB-LRR conformations?
- Do variations on the bait and switch theme (disruption of auto-inhibitory effect of bait proteins vs. bait-mediated facilitation of Avr/LRR interactions)

correspond to differences in fast vs. slow evolution of NB-LRR proteins, or to broad vs. highly specific recognition?

- How did the first NB-LRR protein function and how did its properties shape subsequent evolution and radiation of this protein family?

Dissertation overview

Subsequent sections of this dissertation describe structure/function analyses of both the CC and NB domains of Rx, as well as investigations into a potential CC-NB-LRR signaling mechanism and the characterization of a unique and highly conserved class of CC-NB-LRR protein. Chapter 2 describes the characterization of regions of the Rx CC domain involved in RanGAP2 binding, and the physical relationship between RanGAP2 and NB-ARC-LRR interactions mediated by the CC domain. Chapter 3 presents the identification and study of signal initiation by the NB domain of Rx, and also reports the general ability of NB-LRR protein NB domains to oligomerize. Chapter 4 describes the basal CC_R-NB-LRR clade of CC-NB-LRR proteins, which are capable of signal initiation via their unique N-terminal domains, and are also able to undergo NB – NB interaction with canonical NB-LRR proteins. Implications of these findings for our understanding of NB-LRR protein signal initiation are discussed.

CHAPTER 2

Characterization of the Rx CC – RanGAP2 interaction*

ABSTRACT

It is becoming increasingly apparent that the N-terminal domains of nucleotide binding, leucine-rich repeat (NB-LRR) proteins mediate the perception of pathogen avirulence (Avr) proteins, often through their interaction with recognition co-factors. The N-terminal coiled coil (CC) domain of the Rx NB-LRR protein binds to a suspected recognition co-factor, the Ran GTPase-activating protein 2 (RanGAP2) in addition to participating in an intramolecular interaction with C-terminal regions of the protein. We have found that whereas this intramolecular interaction is governed by a highly conserved "EDVID" motif within the CC domain, RanGAP2 binding is influenced by less conserved regions flanking the EDVID motif. This finding is supported by the spatial relationships observed in the crystal structure of the CC – RanGAP2 complex. We further examine the critical function of two residues, one belonging to the Rx CC domain and one belonging to RanGAP2, in mediating this intermolecular interaction.

---

* Portions of this chapter (Figures 2.1 and 2.2) adapted from: Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boettrich, T., and Moffett, P. (2008) The coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signaling. *The Plant Cell* 20: 739-751. www.plantcell.org Copyright American Society of Plant Biologists. Figures 2.1 and 2.3 were generated by S.M.C., data for Figure 2.2 was generated jointly by S.M.C. and T.T.B., and Figure 2.4 was created by S.M.C. utilizing data generated by J. Chai.

INTRODUCTION

Continuing innovation in crop plant disease resistance demands knowledge of the molecular mechanisms contributing to resistance or susceptibility phenotypes. Nucleotide binding leucine-rich repeat (NB-LRR) proteins are now recognized as major constituents of plant pathogen defense, and unraveling the mode of action of these multi-domain proteins is receiving increasing attention. To-date however, the N-terminal domains of NB-LRR proteins remain relatively poorly understood compared to other regions. While the central NB-ARC (Apaf-1, R proteins, and CED4 homology) region is highly conserved and has a clearly established role in nucleotide binding and hydrolysis (Meyers et al. 1999; Leipe et al. 2004; Tameling et al. 2006), and the C-terminal LRR domains are known for their protein-protein interaction capabilities and have been identified as primary determinants of NB-LRR recognition specificity (Shen et al. 2003; Qu et al. 2006; Rairdan and Moffett 2006; Ellis et al. 2007b), functions of the N-terminal CC and TIR domains are less clear. Although the TIR domain is so-named for its homology to Toll and Interleukin-1 receptors, the means by which this homology might further NB-LRR function remains to be determined. The CC domain is named simply for occasional presence of a coiled coil in structural predictions, and contains few other defining or unifying features (Mazourek et al. 2009).

In the absence of any positive indication of function, N-terminal domains were originally though to act in signal initiation, and indeed a number of TIR domains have subsequently been found to be independently capable of inducing resistance responses (Frost et al. 2004; Swiderski et al. 2009; Krasileva et al. 2010). However, the same trend has not been observed in CC-NB-LRR proteins (with the exception of the

atypical NRG1 (Peart et al. 2005), see Chapters 1 and 4), and it remains to be determined where defense signaling originates in this class of proteins. The mechanisms by which both TIR and CC domains perform their functions, signaling or otherwise, are also yet to be determined.

As discussed in Chapter 1, accumulating evidence suggests that the N-terminal domains of NB-LRR proteins frequently play a role in Avr perception through their interaction with recognition co-factors. These recognition co-factors comprise a diverse group of host proteins which interact, either physically or functionally, with pathogen Avr proteins, thus mediating their perception by NB-LRR proteins (Collier and Moffett 2009). In addition to this emerging role in co-factor binding, physical interaction is observed between the N-terminal CC domains of both the potato (*Solanum tuberosum*) Rx and the pepper (*Capsicum annuum*) Bs2 CC-NB-LRR proteins and their NB-ARC-LRR fragments. This interaction between co-expressed fragments is sufficient to reconstitute protein function, and alterations to the interaction are associated with activation of the proteins' signaling capacity (Moffett et al. 2002; Leister et al. 2005; Rairdan and Moffett 2006). Similar interdomain interactions have been demonstrated for the TIR-NB-LRR protein N (Ueda et al. 2006), and may represent a common theme in NB-LRR behavior.

Through a functional dissection of the Rx CC domain, we have previously identified a sole CC-conserved motif (the "EDVID") as the primary region mediating binding to the Rx NB-ARC-LRR (Rairdan et al. 2008). Although additional portions of the CC must be present in order for binding to take place, their precise sequence appears to be of little consequence for this function. In contrast, mutations of single residues within the EDVID motif were seen to cause qualitative and additive reduction both in binding and in reconstitution of Rx function. It is important to note

that while there is little overall sequence conservation between CC domains, we found the EDVID motif to be conserved in the majority of functionally characterized CC-NB-LRR proteins, further supporting the notion that CC – NB-ARC-LRR interaction is a common feature of these proteins.

Here, we present the extension of Rx CC functional dissection to include its ability to bind RanGAP2 (Ran GTPase-activating protein 2) of potato, its hypothesized recognition co-factor. The presence of RanGAP2 is required for proper Rx function, defined as the ability to recognize the coat protein (CP) of potato virus X (PVX) and trigger *in planta* resistance responses resulting in control of viral spread (Sacco et al. 2007; Tameling and Baulcombe 2007). While programmed cell death at the site of infection, termed the hypersensitive response (HR) is a hallmark of NB-LRR protein activation, Rx is able to swiftly suppress viral growth without inducing visible HR. In experimental contexts such as those described here, however, Rx can be seen to be fully capably of HR induction (Bendahmane et al. 1999).

We have assayed a collection of mutated CC variants for their ability both to interact with RanGAP2 as well as to maintain Rx resistance function, and show that regions of the CC critical to RanGAP2 binding are directly adjacent to, but not including, the EDVID motif. We further investigate single residues of both the CC and RanGAP2, identified during crystallization studies, for their role in mediating interaction. Taken together with our previous findings, these results suggest a juxtaposition of RanGAP2 and NB-ARC-LRR binding regions, perhaps on different surfaces of the CC domain. It can also be seen that all loss-of-function mutations within the CC domain correspond to a loss or severe impairment of binding between the CC domain and either the NB-ARC-LRR region or RanGAP2, suggesting that these are the major functions of the Rx CC domain.

RESULTS

Impact of systematic CC mutations on RanGAP2 binding

Because NB-LRR protein CC domains possess little primary sequence conservation, we employed a systematic mutational analysis approach. A series of N- and C-terminal deletion variants of the Rx CC domain, fused to enhanced green fluorescent protein (EGFP) to provide stability to smaller fragments, have previously been assayed for their ability to bind the Rx NB-ARC-LRR and in this manner reconstitute Rx function (Rairdan et al. 2008) (Figure 2.1). We further examined these truncations for the ability to interact with RanGAP2. The CC:EGFP deletion constructs, fused additionally with an HA epitope tag (thus CC:EGFP:HA), were co-expressed in *Nicotiana benthamiana* with FLAG-tagged RanGAP2 (RanGAP2:FLAG) via *Agrobacterium tumefaciens*-mediated transient expression (agroexpression). Plant protein extracts were subsequently subjected to immunoprecipitation and immunoblotting to determine binding ability (Figure 2.2A). The C1 C-terminal deletion retained robust interaction with RanGAP2, while binding of C2-C6 deletion variants was minimal. These results are in accordance with the lack of functional ability exhibited by truncations C3-C6, and may explain the partially compromised function of C2 (Figure 2.1). While N-terminal deletions N4 and N5 were unable to bind RanGAP2, N1-N3 retained some level of interaction, albeit significantly less than wild type (Figure 2.2A). This may account for the lack of function observed for N-terminal deletions, and taken together with C-terminal deletion data suggests that the entire region comprising the first 115 amino acids of the CC domain is required for optimal interaction with RanGAP2.

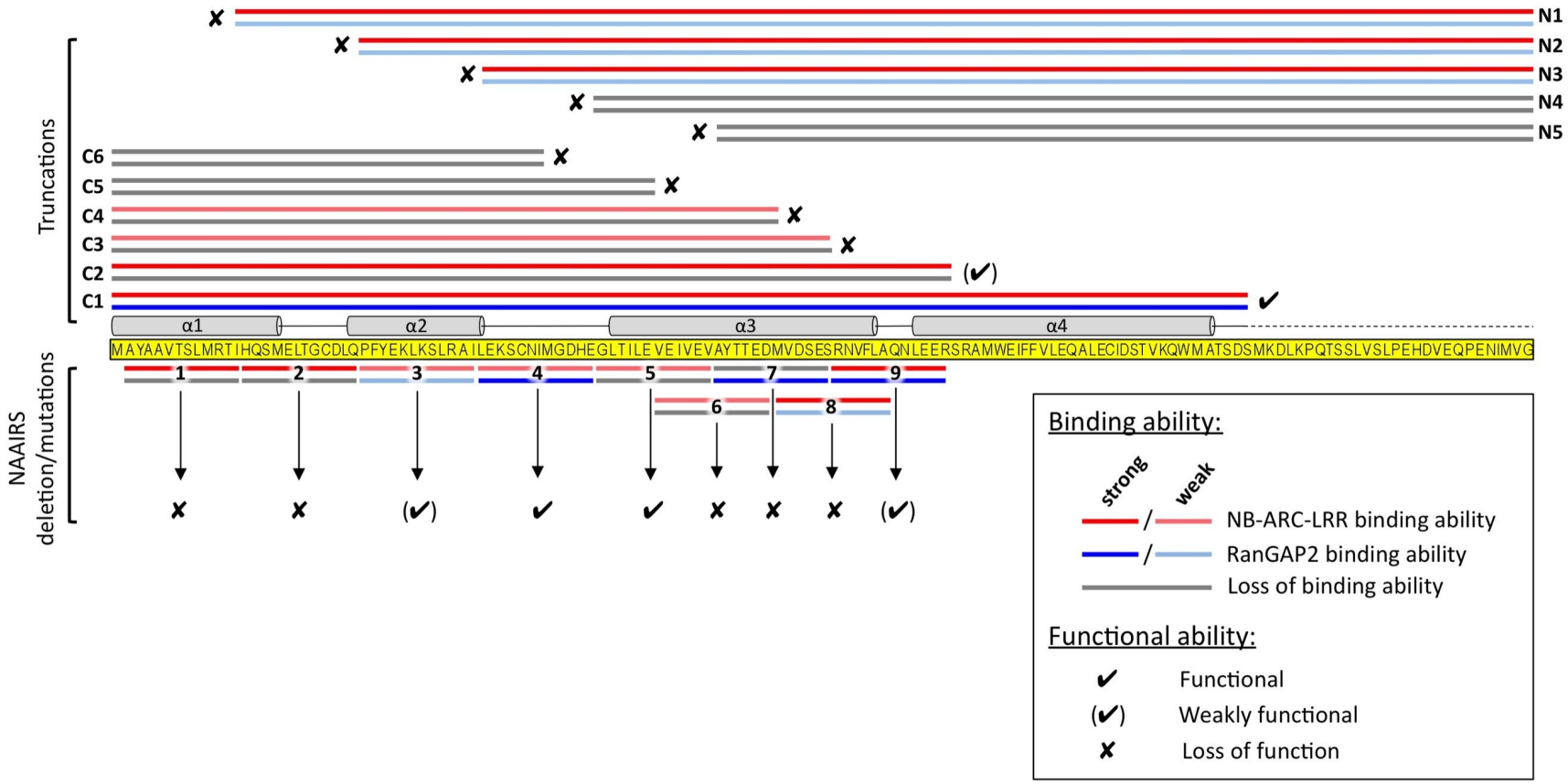


Figure 2.1 *Summary of Rx CC truncations and mutations.* The sequence of the Rx CC domain (1-144) is shown in yellow, with secondary structural features (as presented in Figure 2.4) depicted above with α-helices shown as gray cylinders. Locations of N- and C-terminal CC:EGFP truncations are depicted at top, and locations of NAAIRS CC:HA internal deletion/mutations are depicted at bottom. The ability of each variant to confer CP-dependent HR when co-expressed with NB-ARC-LRR, as reported in (Rairdan et al. 2008), is indicated as functional ability. The ability of each variant to interact with the NB-ARC-LRR, as reported in (Rairdan et al. 2008), is indicated as a red or gray line (top line), while the ability to bind RanGAP2 is indicated as a blue or gray line (bottom line).

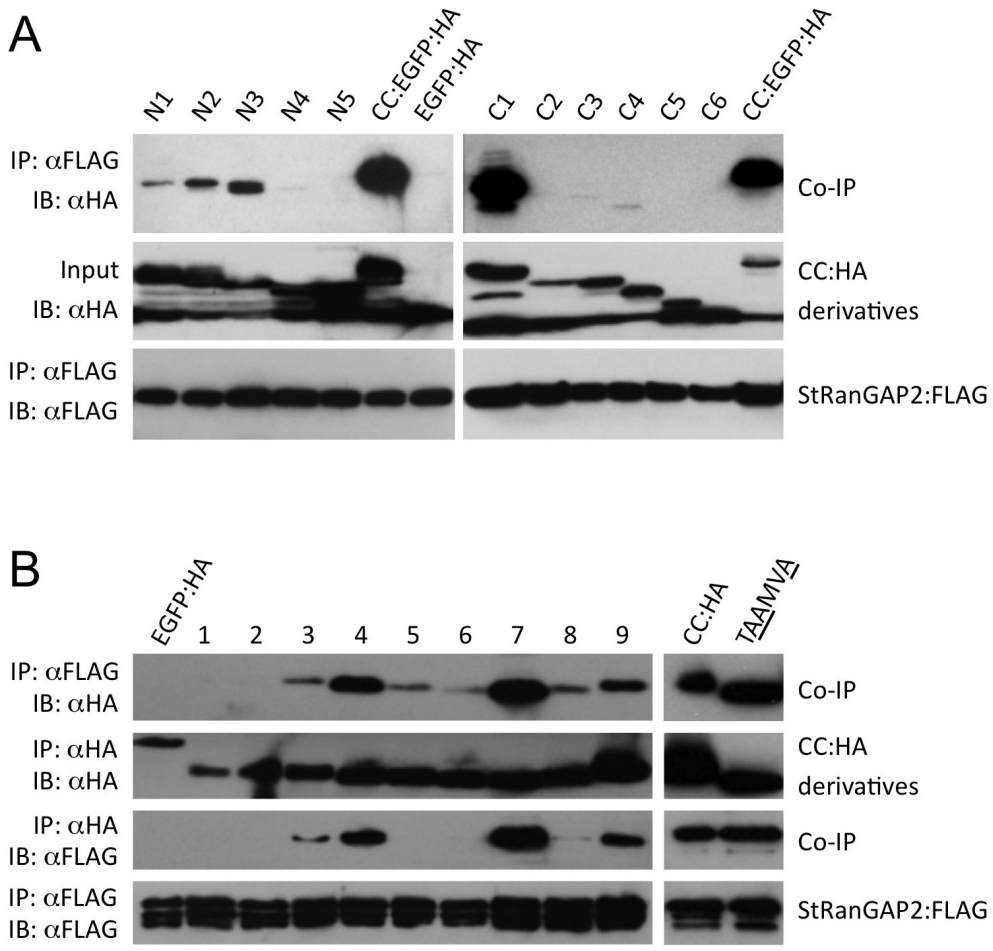


Figure 2.2 *Delineation of RanGAP2 binding regions.* **A,** Interaction of CC truncation fragments with RanGAP2. RanGAP2:FLAG was agroexpressed in *N. benthamiana* leaves along with EGFP:HA, CC:EGFP:HA, or one of the N- and C-terminal deletion derivatives of Rx CC:EGFP:HA as depicted in Figure 2.1. Protein extracts were subjected to anti-(α)FLAG immunoprecipitation (IP) followed by immunoblotting (IB) with the indicated antibodies. **B,** Interaction of CC internal deletion/mutations with RanGAP2. RanGAP2:FLAG was agroexpressed in *N. benthamiana* leaves along with EGFP:HA, CC:HA, internal deletion derivatives of Rx CC:HA as depicted in Figure 2.1, or a mutant derivative (TEDMVD to TAAMVA) of CC:HA. Protein extracts were subjected to anti-FLAG or anti-HA immunoprecipitation followed by immunoblotting with the indicated antibodies.

The region delimited by the N3 and C1 deletions, which appears to be most critical for RanGAP2 binding, encompasses the EDVID motif (EDMVD in Rx) required for CC interaction with the NB-ARC-LRR. We therefore investigated the involvement of the EDVID motif in RanGAP2 interaction using a series of small internal deletions which replace 12 residues with the amino acid sequence NAAIRS in the context of the full CC domain (residues 1 – 144) (Figure 2.1). The NAAIRS motif is found in both α-helices and β-sheets (Wilson et al. 1985), suggesting that it may be less likely to cause gross disruptions in secondary structure, and it has been used successfully to identify critical regions of the *Arabidopsis* SNI1 protein (Mosher et al. 2006). These internal deletion mutations have been previously characterized for their ability to interact both physically and functionally with the Rx NB-ARC-LRR (Rairdan et al. 2008) (Figure 2.1). In coimmunoprecipitation experiments internal deletion mutants 5, 6, and 8, which flank the EDVID motif, showed little or no interaction with RanGAP2 (Figure 2.2B). In contrast, internal deletion mutant 7, which deletes the EDVID motif, showed robust interaction with RanGAP2 (Figure 2.2B). Likewise a TEDMVD-to-T<u>AA</u>MV<u>A</u> mutant of Rx CC:HA, which specifically alters highly conserved acidic residues of the EDVID motif (Rairdan et al. 2008), also showed robust interaction with RanGAP2 (Figure 2.2B). These results suggest that regions of the CC involved in RanGAP2 binding overlap with, but are distinct from, those mediating interaction with the NB-ARC-LRR (namely the EDVID motif). Interestingly, although deletion analysis revealed that the N terminus of the CC domain is not strictly required for RanGAP2 binding (Figure 2.2A), NAAIRS mutations 1 and 2, both located in this region, significantly compromised RanGAP2 binding (Figure 2.2B). Thus while the N terminus of the CC domain likely does not directly mediate interaction with

RanGAP2, it may be in proximity to the interaction surface such that distortions of N-terminal conformation are disruptive to RanGAP2 binding.

Structurally informed CC and RanGAP2 mutational analysis

The Rx CC has recently been co-crystallized with the RanGAP2, allowing the interacting structure to be determined (J. Chai, personal communication). Structurally-directed mutagenesis identified two mutations, W90D of the Rx CC and A89D of RanGAP2, both of which significantly compromise the *in vitro* interaction (J. Chai, personal communication). We incorporated these mutations into our Rx CC:HA and RanGAP2:FLAG agroexpression clones in order to confirm their relevance *in planta*. While HR induction following Rx reconstitution and elicitation (co-expression of CC + NB-ARC-LRR + PVX CP) was comparable between the W90D CC mutant and wild-type (Figure 2.3A), the W90D mutation was seen to compromise the ability of reconstituted Rx to control viral spread when co-expressed with PVX tagged with GFP (PVX:GFP). This effect was visible both as reduced incidence of HR lesions and as increased GFP florescence in W90D treatments as compared to wild-type (Figure 2.3A). These findings indicate that while the W90D mutation may not entirely abolish Rx function, it does result in significant functional impairment. We were unable to evaluate the functional relevance of the RanGAP2 A89D mutation, because native *N. benthamiana* RanGAP is sufficient for Rx function.

Coimmunoprecipitation experiments with the Rx CC and RanGAP2 point mutants demonstrated an impaired capacity for *in planta* interaction (Figure 2.3B). While either mutation drastically reduced interaction, the effects of the mutations also appear to be somewhat additive, as interaction between the two mutant variants was

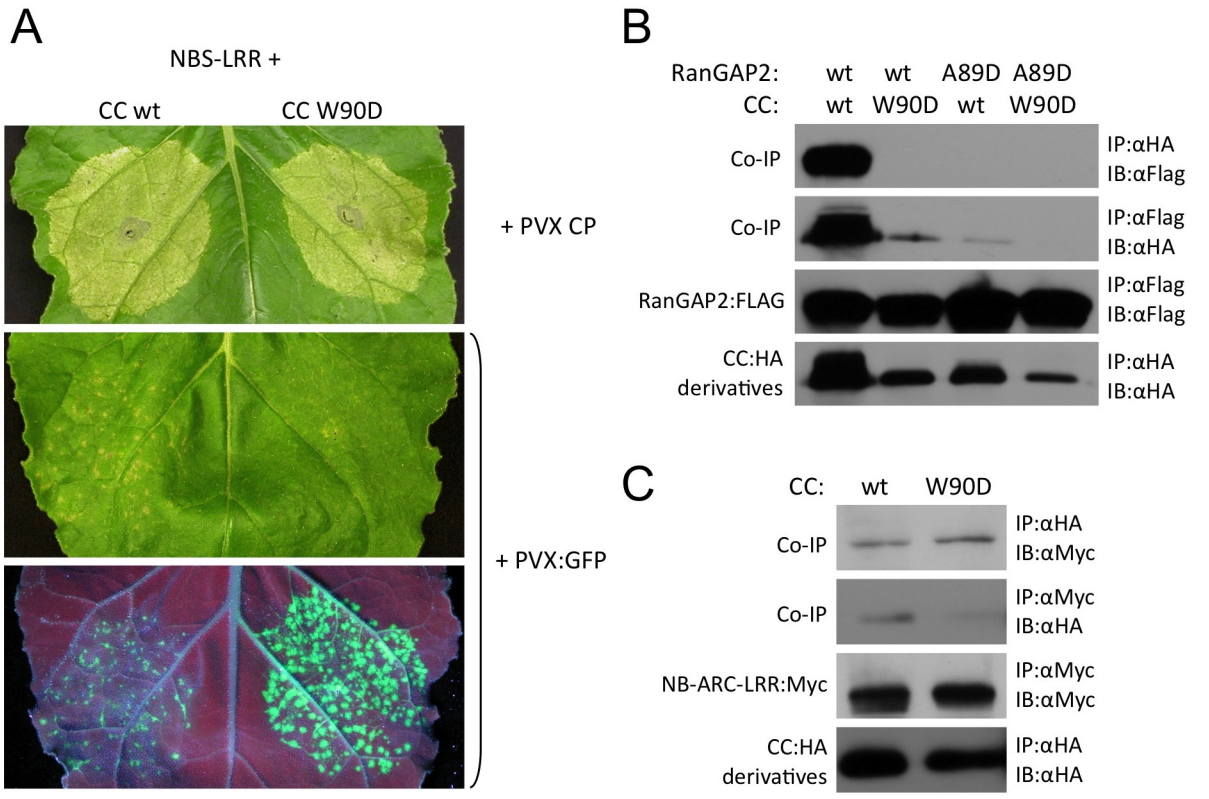


Figure 2.3 *Functional analysis of Rx CC and RanGAP2 point mutants.* **A,** Impact of the W90D mutation on Rx-mediated resistance. The Rx NB-ARC-LRR fragment was coexpressed with the Rx CC domain, either wild type (wt) or the W90D mutant, in the presence of the PVX CP (top panel) or PVX:GFP (bottom panels) in *N. benthamiana* leaves. Leaves coinfiltrated with CP were photographed under white light after three days (top panel), and those coinfiltrated with PVX:GFP were photographed after six days under white light (middle panel) or UV light (bottom panel). **B,** Impact of point mutations on CC – RanGAP2 binding. Rx CC:HA (wt or W90D) was agroexpressed in *N. benthamiana* leaves along with RanGAP2:FLAG (wt or A89D). Protein extracts were subjected to anti-HA and anti-FLAG immunoprecipitation followed by immunoblotting as indicated. **C,** NB-ARC-LRR binding by Rx CC W90D. Rx CC (wt or W90D) was agroexpressed in *N. benthamiana* leaves together with Rx NB-ARC-LRR fused to six tandem c-Myc epitope tags (NB-ARC-LRR:Myc). Protein extracts were subjected to anti-HA and anti-Myc immunoprecipitation followed by immunoblotting as indicated.

even weaker than between either mutant and the wild-type partner. The residual binding ability retained by the Rx CC W90D mutant may explain its ability to induce HR upon co-expression with Rx NB-ARC-LRR + PVX CP. Importantly, while the Rx CC W90D mutation has a drastic effect on binding to RanGAP2, it has little if any impact on the intramolecular interaction between CC and NB-ARC-LRR domains (Figure 2.3C), further supporting the physical distinction between residues mediating these two interactions.

DISCUSSION

We and others have previously demonstrated that the Rx CC domain plays dual roles, mediating both an intramolecular interaction with the NB-ARC-LRR regions as well as an intermolecular interaction with the WPP domain of its hypothesized recognition co-factor, RanGAP2 (Moffett et al. 2002; Sacco et al. 2007; Tameling and Baulcombe 2007; Rairdan et al. 2008). Here, we present a deletion analysis of the Rx CC domain which reveals that the region encompassing amino acids 38 – 116 (deletion fragments N3 – C1) is of particular importance in mediating interaction with RanGAP2 (Figure 2.1, 2.2A). Interestingly, this region also contains the EDVID motif known to be central to the intramolecular interaction with the NB-ARC-LRR region (Rairdan et al. 2008). Analysis of higher-resolution NAAIRS deletion/mutations, however, revealed that the EDVID motif is dispensable for RanGAP2 binding (Figure 2.1, 2.2B). Rather, those regions immediately flanking the EDVID motif, defined by mutations 5, 6, and 8, emerge as being critical to the RanGAP2 interaction (Figure 2.1, 2.2A). Notably therefore, whereas the

intramolecular interaction with the NB-ARC-LRR region is mediated by the most highly conserved region of the CC domain (Rairdan et al. 2008), interaction with RanGAP2 is mediated by regions which display little conservation between NB-LRR proteins ((Rairdan et al. 2008) and Figure 2.2B). As discussed in Chapter 1, this finding supports the emerging understanding that whereas *intra*molecular interactions within NB-LRR proteins are likely well-conserved, *inter*molecular interactions mediated by their N-terminal domains are likely diverse, possibly enabling interactions with various recognition co-factors and/or Avr proteins.

Whether CC domains perform any additional function(s) such as signal initiation is currently unknown. However, the observation that all alterations to the CC domain which compromised Rx-mediated resistance responses corresponded to, and could thus be ascribed to, a loss of either NB-ARC-LRR or RanGAP2 binding (Figure 2.1) suggests that these two interactions comprise the fundamental roles of the Rx CC domain. That is, our mutational analysis of the CC domain revealed no 'unexplained' loss-of-function mutants that would indicate another critical function outside of the two already described interactions, making CC-mediated signal initiation an unlikely scenario.

While neither the interaction with the NB-ARC-LRR region nor with RanGAP2 are likely to have a direct role in signal initiation (Sacco et al. 2007), the proximity of these two interactions is nonetheless intriguing. Given that the Rx CC – NB-ARC-LRR interaction is abolished upon PVX CP perception (Moffett et al. 2002) as well as by various inactivating and autoactivating mutations throughout the NB-ARC-LRR region (Rairdan et al. 2008), it would seem that this interaction is sensitive to the recognition and/or activation state of the protein. Furthermore, the dual interactions of the CC domain make it ideally situated to transform CP-related alterations to

RanGAP2 into conformational changes within the NB-ARC-LRR region. This possibility is supported by the juxtaposition of binding regions revealed by truncation and NAAIRS analysis (Figure 2.1, 2.2B).

The juxtaposition of RanGAP2 and NB-ARC-LRR binding regions becomes even more striking when co-immunoprecipitation results are superimposed upon the structure of interacting Rx CC and RanGAP2 WPP domains (Figure 2.4). The CC domain is composed of four helices: α1 (amino acids 1 – 17), α2 (25 – 38), α3 (51 – 78), and α4 (82 – 112). It can be seen that regions of importance to RanGAP2 binding comprise the two adjacent α3 and α4 helices. The α4 helix, which includes the functionally critical W90 residue, is closely associated with the region of RanGAP2 containing the A89 residue involved in CC binding (Figure 2.3A and 2.3B). The α3 helix, which houses the EDVID motif flanked by regions important to RanGAP2 binding, is located at a greater distance from the RanGAP2 WPP domain, presumably in contact with the Rx NB-ARC-LRR via the EDVID motif. The close proximity observed among all four helices also sheds light on previously perplexing observations. For instance, while no functionally critical residues were identified within helices α1 and α2, it is easy to imagine how their deletion or disruption might also disrupt the physical relationship between helices α3 and α4, and consequently the relationships between the WPP, CC, and NB-ARC-LRR domains. Collectively, our results suggest that the Rx CC domain possesses distinct surfaces in contact with RanGAP2 and with the NB-ARC-LRR, and that these two interacting surfaces are themselves in close proximity. Thus rather than mediating signal initiation, the Rx CC domain occupies a position from which it might relay information regarding

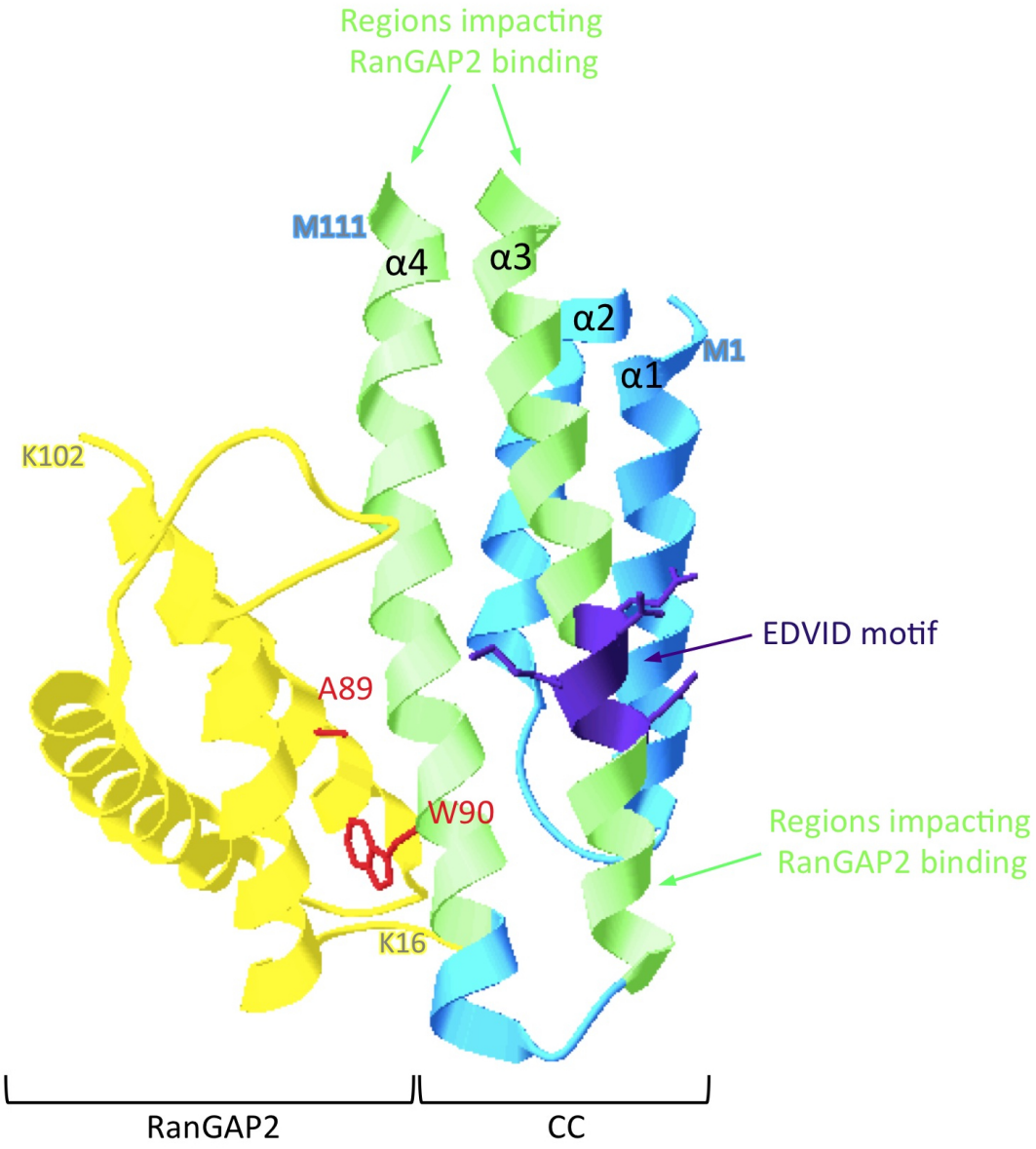


Figure 2.4 *Structure of interacting Rx and RanGAP2.* The RanGAP2 WPP domain is shown in yellow and the Rx CC domain is shown with a base of blue. The protein structures are depicted as ribbon diagrams, with CC residues belonging to the EDVID motif shaded purple, and those found to be impact to RanGAP2 binding (Figure 2.2 deletion/mutations 5, 6, and 8 plus the region delineated by truncations C1-C2) shaded green. Side-chains of EDVID motif residues are displayed in purple, and those corresponding to Rx CC W90 and RanGAP A89 are displayed and labeled in red. The location and position of the first and last amino acid present in the structural model of each protein fragment is indicated in gray. Helices are numbered α1 – α4 based on their order of appearance in the CC domain primary sequence, and correspond to those depicted in Figure 2.1. Structure kindly provided by J. Chai.

alterations in Avr or recognition co-factor conformation to the rest of the protein, or vice versa.

MATERIALS AND METHODS

Plasmid construction

Rx CC:EGFP:HA truncations and Rx CC:HA NAAIRS deletion mutations have been described previously (Rairdan et al. 2008), as have Rx NB-ARC-LRR:Myc (Moffett et al. 2002) and RanGAP2:FLAG (Sacco et al. 2007). W90D and A89D site-directed mutants were generated by extension-overlap PCR (Vallejo et al. 2003) using KOD high-fidelity thermostable polymerase (Novagen) and the Expand high fidelity PCR system (Roche) according to the manufacturers' instructions. Each mutation utilized a set of four primers: a pair of flanking cloning primers including restriction sites and Kozak consensus sequence where necessary, and a pair of internal primers to induce the desired mutation. For Rx CC W90D the flanking primers were: 5'-CGTATCTAGAACCACCATGGCTTATGCTGCTGTTAC-3' and 5'-CGTAGGATCCGCCAACCATTATATTCTCG-3'; and the internal primers were: 5'-GCAGGGCTATGGACGAGATTTTTTTCG-3' and 5'-CGAAAAAAATCTCGTCCATAGCCCTGC-3'. For RanGAP2 A89D the flanking primers were: 5'-CGTATCTAGAACCACCATGGATGCCACAACAGC-3' and 5'-CGTAGTATACCAAAGCTTCTATTGC-3'; and the internal primers were: 5'-GTGCAGCTTTATGACAGGGAATGCAGC-3' and 5'-GCTGCATTCCCTGTCATAAAGCTGCAC-3'. Underlined nucleotides indicate the sources of mutation. The Rx CC W90D PCR product was cloned into the XbaI and

BamHI sites of the pBIN61 vector containing an HA tag (Bendahmane et al. 2002), and the RanGAP2 A89D PCR product was cloned into the XbaI and BstZ171 sites of pBIN61-RanGAP2:FLAG (Sacco et al. 2007). All constructs generated by PCR were sequence-verified.

Transient protein expression and analysis

Binary vectors were transformed into *A. tumafaciens* strain C58C1 carrying the virulence plasmid pCH32, and agroexpression was performed as described previously (Bendahmane et al. 2000) at OD_{600} = 0.2 for each construct. All proteins were co-expressed in *N. benthamiana* leaves under control of the CaMV 35S promoter of the pBIN61 vector (Bendahmane et al. 2002). Protein extracts were collected approximately 48h post-infiltration; extraction, coimmunoprecipitation, and immunoblotting were performed as previously described (Rairdan and Moffett 2006). The CP-dependent HR assay consisted of agroinfiltrating Rx fragments as well as the PVX CP (Bendahmane et al. 2000) at OD_{600} = 0.2. The PVX:GFP resistance assay consisted of agroinfiltrating Rx variants at OD_{600} = 0.2 along with *A. tumafaciens* strain GV3101 carrying the plasmid pGr208, which expresses the PVX:GFP cDNA, as well as the helper plasmid pSoup (Peart et al. 2002b) at OD_{600} = 0.001. GFP fluorescence was monitored 5d later using a handheld UV lamp. Each experiment was repeated at least three times with similar results.

Protein structural analysis

Structure of the interacting Rx CC domain and RanGAP2 WPP domain is the result of crystallography studies lead by Jijie Chai, National Institute of Biological

Sciences, Beijing, China. Visualization and manipulation of 3D structure was performed using SwissPdb Viewer (Guex and Peitsch 1997).

CHAPTER 3

The Rx NB domain in signal initiation and oligomerization*

ABSTRACT

Despite our expanding understanding of the mechanistic underpinnings of plant disease resistance proteins, the means by which plant CC-NB-LRR (coiled coil, nucleotide binding, leucine-rich repeat) proteins trigger plant defense responses upon recognition of pathogens remains poorly understood. Although the N-terminal CC domain has long been assumed to function in signal initiation, no such activity has been reported. Rather, we have found that the NB domain of the CC-NB-LRR protein Rx is sufficient to mediate the induction of defense responses. We have isolated this activity to a specific structural region of the domain, which corresponds to regions mediating oligomerization in metazoan apoptosome-forming proteins Apaf-1 and CED-4. We further report the ability of the Rx NB domain to oligomerize both homotypically, and heterotypically with other NB-LRR protein NB domains.

---

* Portions of this chapter (Figures 3.1 and 3.2) adapted from: Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boettrich, T., and Moffett, P. (2008) The coiled-coil and nucleotide binding domains of the potato Rx disease resistance protein function in pathogen recognition and signaling. *The Plant Cell* 20: 739-751. www.plantcell.org Copyright American Society of Plant Biologists. All data presented in this chapter was generated by S.M.C. except for a portion of Figure 3.1A and Figure 3.2A, which were generated by G.J.R.

INTRODUCTION

Nucleotide binding, leucine-rich repeat (NB-LRR) proteins constitute a major component of plant disease resistance. They serve to recognize a myriad of different pathogen avirulence (Avr) proteins and to translate that recognition into the initiation of a strong resistance response in the plant. This response is often accompanied by programmed cell death, termed the hypersensitive response (HR), at the site of infection, and typically provides effective control of pathogen spread. Despite our ever-increasing understanding of NB-LRR protein function (see Chapter 1), the mechanism by which these proteins are able to transform the recognition of diverse Avr proteins into the initiation of a conserved resistance signal remains largely unknown.

NB-LRR proteins derive their name from the presence of a central NB domain and a C-terminal LRR domain. Between these two domains lies a region of homology referred to as the ARC (Apaf-1, R proteins, CED-4). The NB and ARC domains together comprise a functional nucleotide binding pocket (Tameling et al. 2002). Amino acid sequences within the NB-ARC region are highly conserved, and include a number of functional motifs which indicate homology with members of the STAND (signal transduction ATPases with numerous domains) class of proteins (Meyers et al. 1999; Leipe et al. 2004). In addition to its presumed role in nucleotide binding, the NB-ARC region of the potato (*Solanum tuberosum*) NB-LRR protein Rx has also been found to be involved in intramolecular interactions with both its LRR domain and its N-terminal domain (Moffett et al. 2002; Rairdan et al. 2008), and such intramolecular interactions are likely common to other NB-LRR proteins as well (see Chapter 1).

At their N-termini, NB-LRR proteins typically possess either a TIR (Toll and Interleukin-1 receptor homology) or a CC (coiled-coil) domain, and the two classes are further distinguished by characteristic differences within their NB-ARC domains (Meyers et al. 1999). As discussed in Chapters 1 and 2, it is becoming increasingly evident that NB-LRR N-terminal domains play a role in Avr perception, often through interaction with recognition co-factors. Interestingly, N-terminal domains were initially assumed to function in signal initiation, and a number of TIR domains have subsequently been found to be capable of independently inducing HR (Frost et al. 2004; Swiderski et al. 2009; Krasileva et al. 2010). However, with the exception of the atypical NRG1 (Peart et al. 2005) (see Chapter 4), no similar function has been identified among CC-NB-LRR proteins. The origin of NB-LRR protein signal initiation, particularly for members of the CC class, thus remains to be determined.

Over-expression of full-length NB-LRR proteins in plant tissue is sometimes observed to activate defense responses even in the absence of a corresponding Avr protein (Oldroyd and Staskawicz 1998; Tao et al. 2000; Bendahmane et al. 2002). Thus over-expression of NB-LRR protein fragments, as in the cases of TIR-mediated HR induction mentioned previously, can be a useful tool for delimiting the regions of the protein responsible for triggering resistance responses. Over-expression of full-length Rx induces HR in very young tobacco (*Nicotiana tabacum*) leaves in the absence of its cognate Avr protein, the coat protein (CP) of *Potato Virus X* (PVX). Deletion of the Rx ARC and LRR domains results in a protein fragment (CC-NB) with enhanced HR-inducing activity (Bendahmane et al. 2002), suggesting that resistance signal initiation originates from within the CC-NB region of Rx. Having thoroughly examined the Rx CC domain and found no evidence for a role in signaling (Chapter 2), we next investigated the Rx NB domain as a potential source of Rx signal initiation. Here, we

show that when effectively stabilized the Rx NB domain possesses a clear capacity for HR induction. Through mutational analysis of this domain we have found that signal initiation can be distinguished from the region's role in nucleotide binding. We further present a novel capacity for oligomerization among NB domains as a possible mechanism of signal transduction.

RESULTS

NB-mediated HR induction

Examination of the Rx CC-NB and NB domains alone has previously been impeded by significant and complete protein instability, respectively. However, we have found that fusion of these fragments to enhanced green florescent protein carrying an HA epitope tag (EGFP:HA) greatly enhances their stability (Figure 3.1A). Whereas the weakly accumulating CC-NB:HA does not induce a visible HR under our experimental conditions when transiently expressed via *Agrobacterium* (agroexpression) in tobacco leaves, the more abundant CC-NB-EGFP:HA induces a robust HR visible within 48 hours (Figure 3.1B). The CC:EGFP:HA fusion protein however does not induce an HR despite accumulating to similar levels as CC-NB:EGFP:HA (Figure 3.1), likely indicating that the Rx CC domain does not function in signal initiation. Alternately, this observation could be due to a requirement for interaction with the NB domain. The CC and NB domains do not interact *in trans*, however (Rairdan et al. 2008), and mutation of the highly conserved P-loop (PL) of the kinase-1a motif – a critical component of nucleotide binding, mutation of which also abrogates the ability of the CC and NB-ARC-LRR fragments to interact *in trans* (Moffett et al. 2002) – does not compromise HR induction by Rx CC-NB:EGFP:HA

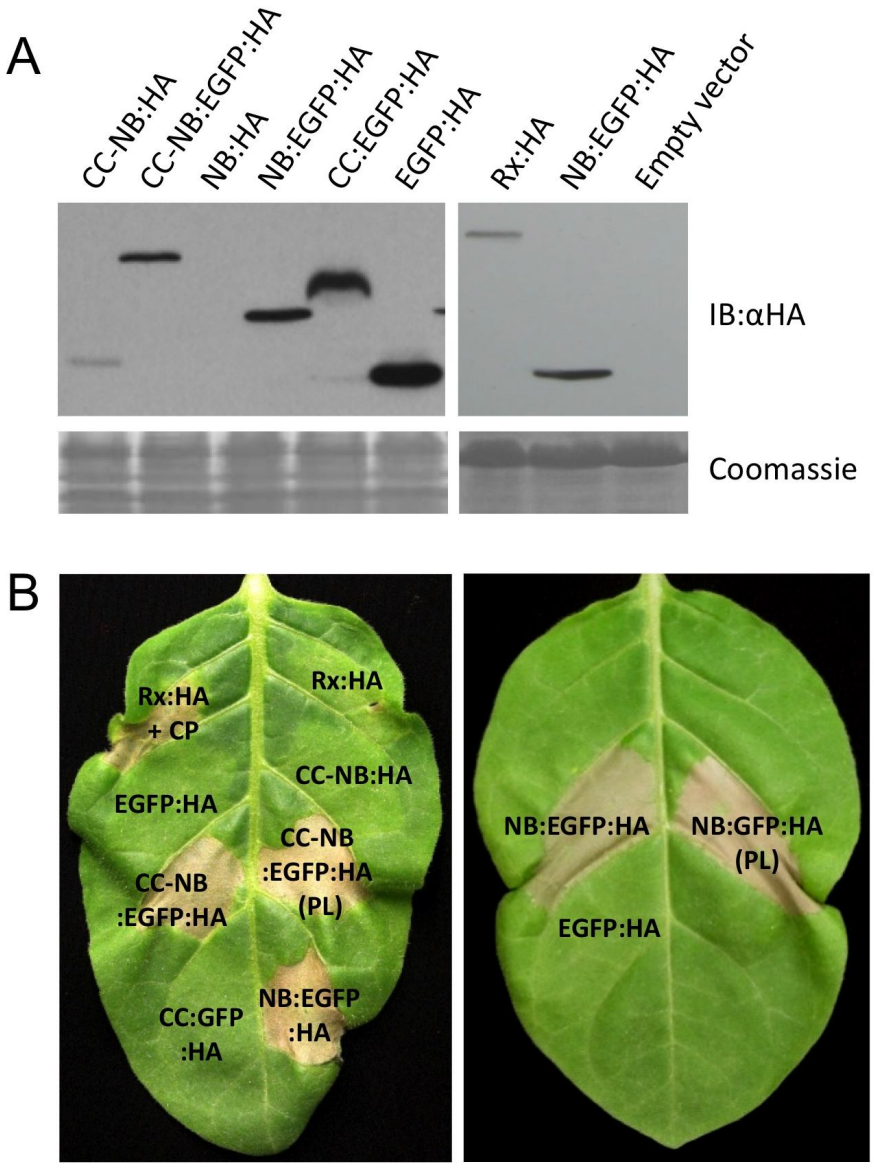


Figure 3.1 *The Rx NB domain is capable of inducing an HR.* **A,** Relative expression levels of Rx derivative fragments. The indicated fragments were agroexpressed in *N. benthamiana* leaves, and proteins were extracted for anti-(α)HA immunoblotting (IB) two days later (top panels). Equal loading was confirmed by coomassie staining (bottom panels). Proteins on the left and right sets of panels were separated on different percentage SDS-PAGE gels for optimal resolution. **B,** Expression of Rx fragments in *N. tabacum* leaves. The indicated fragments were agroexpressed in tobacco, and infiltrated leaves were photographed two days later.

(Figure 3.1B – CC-NB:EGFP:HA(PL)). CC – NB interaction thus seems unlikely to be a prerequisite for signal initiation.

Similar to the CC domain, stability of the Rx NB domain is greatly enhanced by fusion to EGFP, allowing it to accumulate to levels similar to the CC-NB:EGFP:HA fragment (Figure 3.1A). Transient expression of both the Rx NB:EGFP:HA and NB:EGFP:HA(PL) induce an HR of equal intensity as CC-NB:EGFP:HA, suggesting that the NB domain is sufficient for the initiation of downstream defense responses, and that nucleotide binding is not a prerequisite for this function. We also compared the accumulation of the Rx NB:EGFP:HA to that of FL Rx:HA (Figure 3.1A), and found that EGFP-stabilized NB accumulates to higher levels than the full-length protein, similar to the difference in accumulation observed between CC-NB:EGFP:HA and CC-NB:HA (Figure 3.1A). Thus HR-induction appears to be correlated with protein stability, in addition to the previously reported negative regulatory functions of C-terminal portions of Rx (Bendahmane et al. 2002). These results, in combination with findings presented in Chapter 2, suggest that the region of Rx responsible for the initiation of defense signaling lies within the NB domain.

SGT1 dependence of NB-mediated HR

Rx-mediated resistance is compromised in plants where the levels of SGT1 (suppressor of the G2 allele of *skp1*) have been reduced by virus-induced gene silencing (VIGS) (Moffett et al. 2002; Peart et al. 2002b). In order to determine whether HR induced by the Rx NB:EGFP:HA relies on the same signaling and/or regulatory components as the full-length protein, we silenced *SGT1* in plants carrying *Rx* as a transgene. We also silenced *EDS1* (enhanced disease susceptibility 1), which is

required for resistance mediated by the N NB-LRR protein against *Tobacco Mosaic Virus*, but not for Rx-mediated resistance against PVX (Peart et al. 2002a). Agroexpression of the autoactive Rx(D460V) mutant, the PVX CP, and the Rx CC-NB:EGFP:HA and NB:EGFP:HA fragments produced an HR in *EDS1*-silenced plants as well as those infected with the empty VIGS vector (TV:00), but produced no response in *SGT1*-silenced plants (Figure 3.2A). Rx NB:EGFP:HA accumulated to similar levels in *SGT1*-silenced plants as in plants infected with a VIGS vector containing a β glucuronidase (GUS) insert (Figure 3.2B). Requirement for SGT1 suggests that the HR induced by Rx NB:EGFP:HA is similar to responses mediated by the full-length protein, and is unlikely the result of general toxicity of the NB fragment or the NB:EGFP:HA fusion protein.

Rx2 and Gpa2 NB-mediated HR

Having determined that the Rx NB domain is sufficient for the induction of HR, we next examined whether the NB domains of other NB-LRR proteins would display similar activity. Transient expression of the NB domains of the Bs2, RPS5, RPS2, Prf, N, RPP1, RPS4, and Bs4 NB-LRR proteins fused to EGFP:HA did not result in HR induction, although in many of these cases EGFP:HA fusion appears to be insufficient to stabilize the protein fragment (Figure 3.3). We also examined the NB domains of two *Rx* homologs, *Rx2* and *Gpa2*, both of which originate from potato and possess high levels of sequence similarity to Rx within their NB domains. Rx2 also recognizes the PVX CP, and Gpa2 confers resistance to the potato cyst nematode *Globodera pallida* (Bendahmane et al. 2000). Both the Rx2 and Gpa2 NB domains were stable when expressed as EGFP:HA fusions, and were found to induce an HR upon over-expression (Figure 3.3).

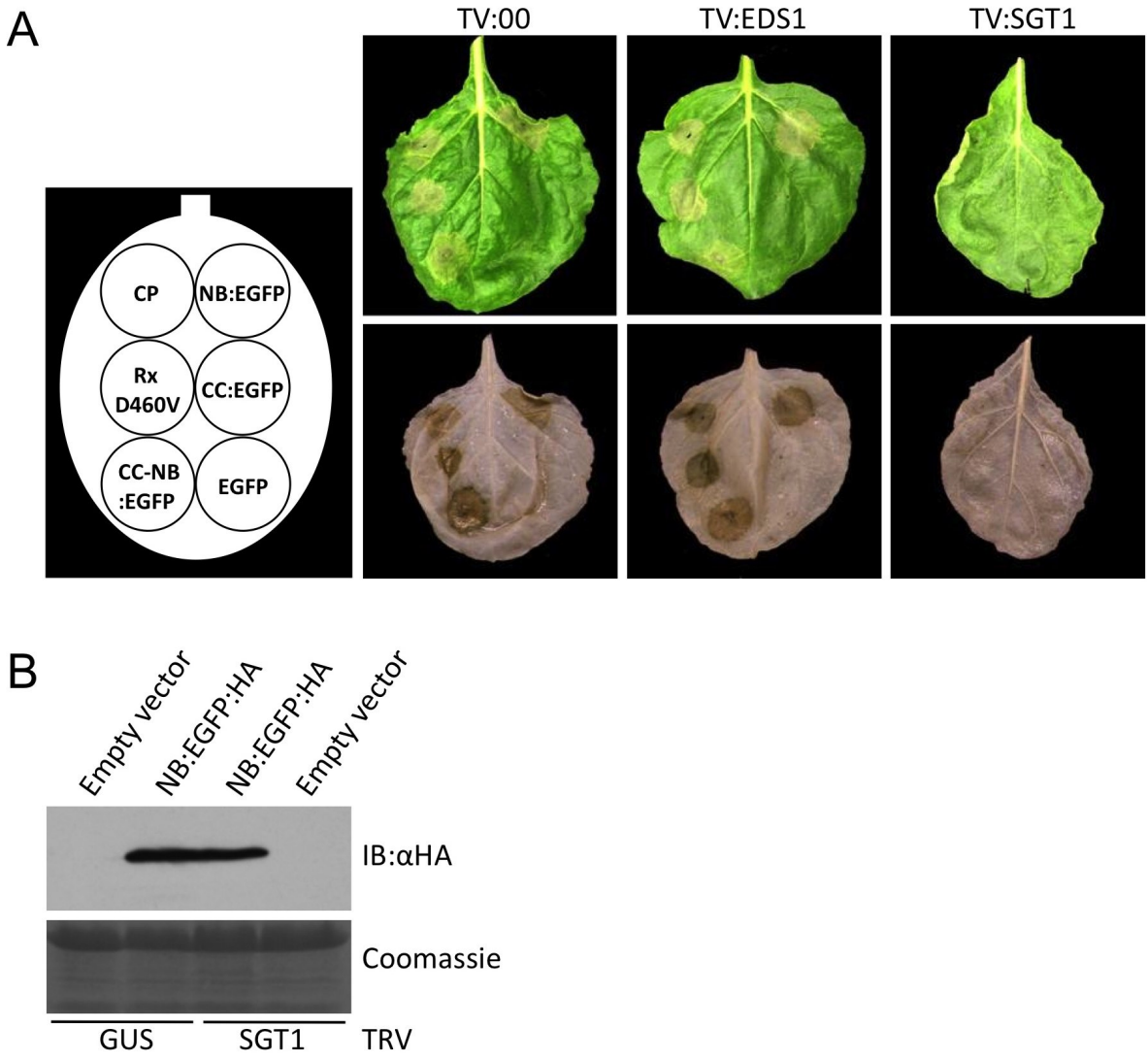


Figure 3.2 *Rx NB-mediated HR is dependent on SGT1*. **A,** Activity of Rx derivatives in *EDS1* and *SGT1* silenced leaves. *N. benthamiana* plants expressing *Rx* as a transgene were treated with *Tobacco Rattle Virus* (TRV) vectors carrying either no insert (TV:00), or sequences derived from *N. benthamiana SGT1* or *EDS1* genes. Three weeks after TRV infection, silenced leaves were agroinfiltrated with the Rx derivatives (or PVX CP) indicated in the panel at left. Leaves were photographed three days later (top panels), cleared with ethanol, and photographed again (bottom panels). **B,** Stability of Rx NB:EGFP:HA in *SGT1*-silenced plants. Rx NB:EGFP:HA and an empty vector control were agroexpressed in *N. benthamiana* plants silenced either for *SGT1* or with a TRV vector carrying a GUS insert. Protein extracts were collected two days later for anti-HA immunoblotting (top panel). Equal loading was verified by coomassie staining (bottom panel).

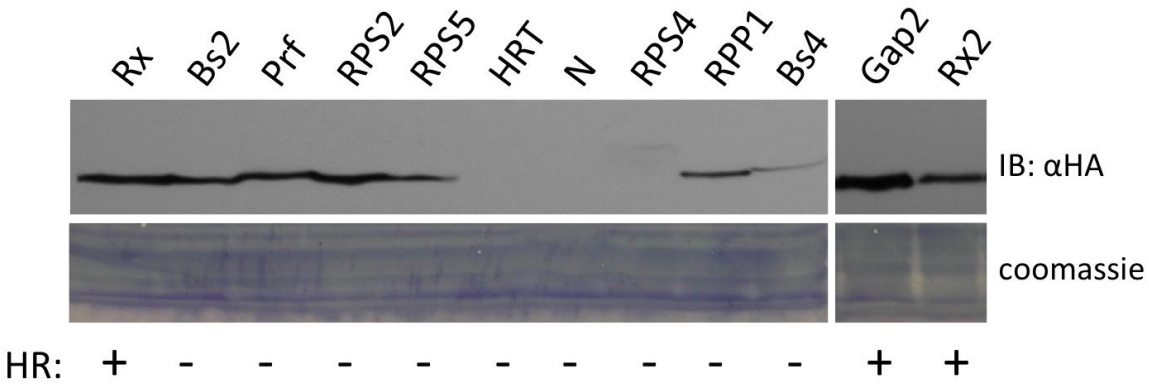


Figure 3.3 *The NB domains of other NB-LRR proteins.* NB domains of the indicated NB-LRR proteins were agroexpressed in *N. benthamiana* leaves. Protein samples were collected after 36 hours and subjected to anti-HA immunoblotting (top panel). Equal loading was verified by coomassie staining (middle panel). Infiltrated patches were monitored for five days for presence (+) or absence (-) of HR development (bottom panel).

Functional analysis of conserved NB domain motifs

The portion of the Rx NB domain initially identified as being capable of Avr-independent HR induction consists of amino acids 139-293. This region contains a number of conserved motifs, some of which are involved in nucleotide binding and hydrolysis, and some of which have no known function but are conserved among NB-LRR proteins (Figure 3.4A). Among the former group – those motifs critical for nucleotide binding in many ATP/GTP-binding proteins – the anchor and kinase-1a motifs are known to coordinate the bound nucleotide, while the kinase-2 motif serves to bind a Mg^{2+} ion required for phosphotransfer reactions. The region also contains a kinase-3a motif thought to interact with the purine base of ATP (Traut 1994; Tameling et al. 2002; Leipe et al. 2004). In addition to the P-loop of the kinase-1a which extends into the nucleotide binding pocket, all three kinase motifs typically contribute to a core structure of several aligned β-strands which are interspersed with α-helices (Traut 1994). Due to this broadly conserved tertiary structure, the region corresponding to the NB domain of NB-LRR proteins is often referred to as the α/β fold in other systems. Motifs found in the NB domain which are unique to NB-LRR proteins consist of RNBS-A, -B, and -C (for resistance NBS) motifs, as well as an "LXXLL" motif (Meyers et al. 1999; Mestre and Baulcombe 2006). Of these, the RNBS-B is synonymous with the kinase-3a of other STAND proteins, despite sharing little sequence similarity (Meyers et al. 1999; Tameling et al. 2002), and the RNBS-C is located principally in the ARC domain with only its N-terminus present in the NB domain. The sequence of the RNBS-A motif differs markedly between CC-NB-LRR

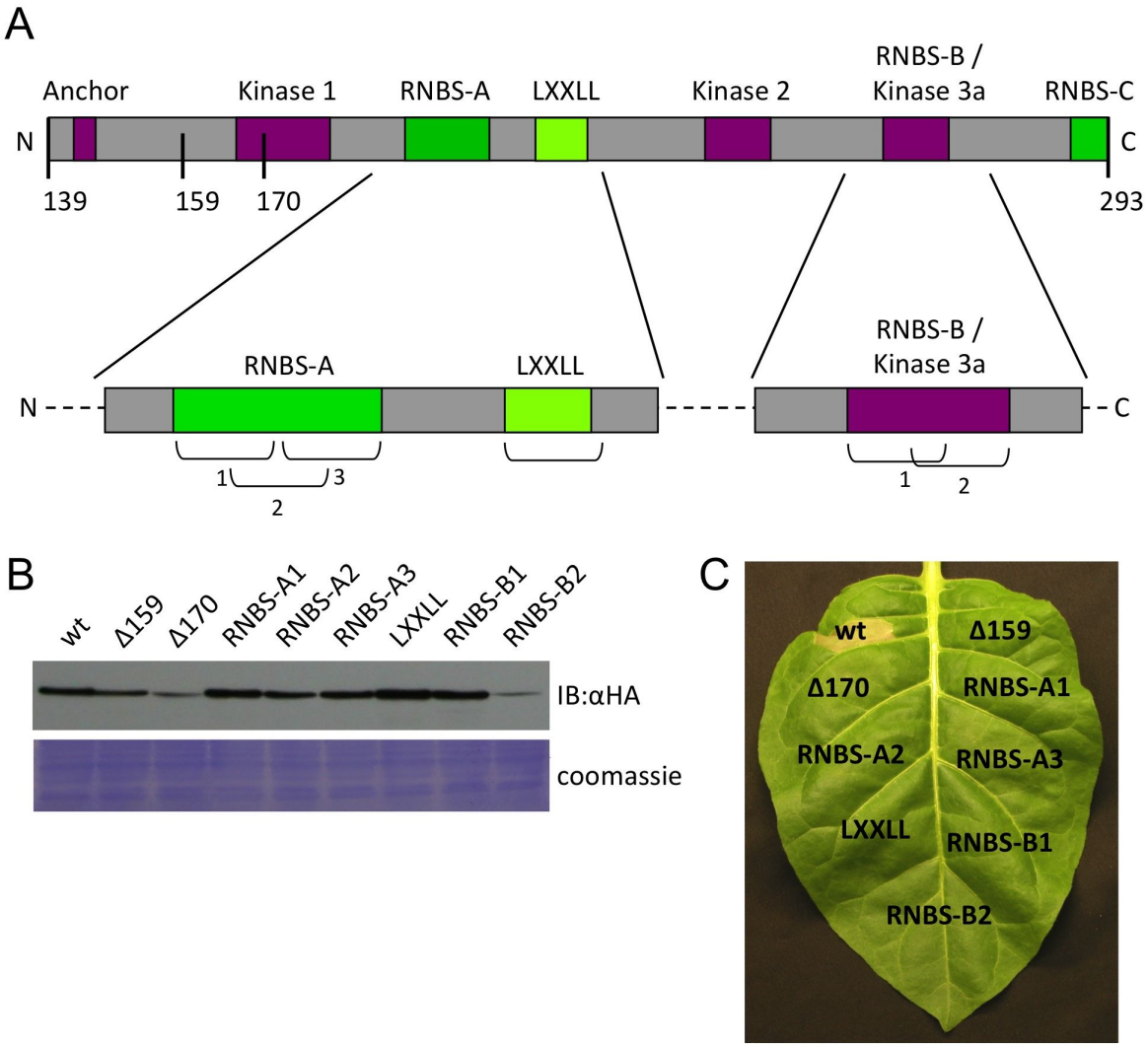


Figure 3.4 *Functional analysis of conserved motifs.* **A,** Schematic representation of the Rx NB domain. Amino acids 139 – 293 of Rx are depicted (top panel) with conserved motifs involved in nucleotide binding highlighted in purple, and those of unknown function highlighted in green. Location of Δ159 and Δ170 truncations are also indicated. Regions of interest are magnified in the bottom panel, with locations of NAAIRS substitution mutations indicated by brackets. **B,** Stability of Rx NB derivatives. The indicated Rx NB:EGFP:HA truncations and NAAIRS replacements were agroexpressed in *N. benthamiana* leaves, and protein samples were collected 36 hours later for anti-HA immunoblotting (top panel). Equal loading was confirmed by coomassie staining (bottom panel). **C,** Activity of Rx derivatives. Rx NB:EGFP:HA truncation and NAAIRS replacement derivatives were agroexpressed in *N. tabacum* and photographed three days later.

and TIR-NB-LRR proteins, and is a reliable means of distinguishing the two NB-LRR classes (Meyers et al. 1999).

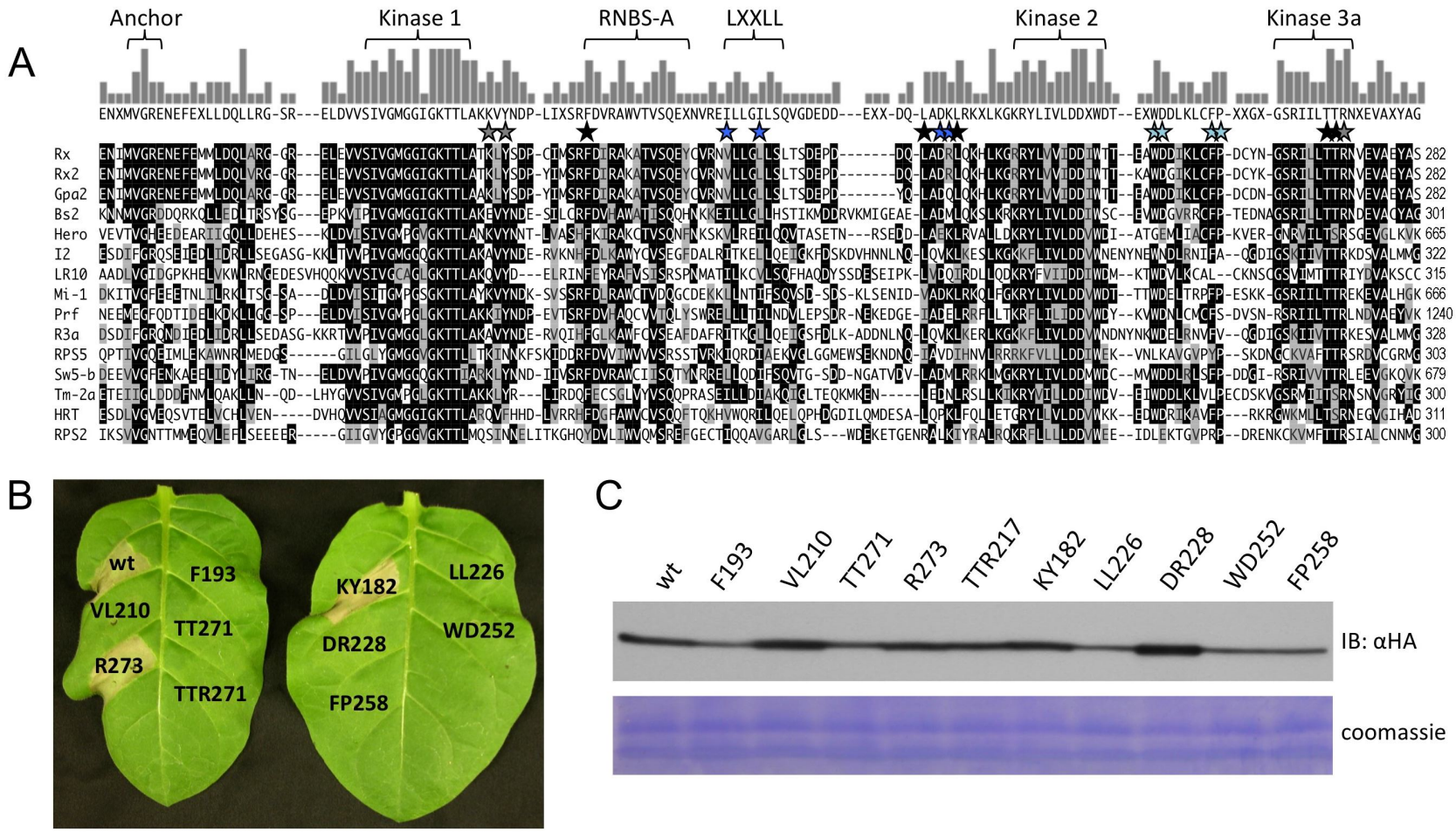
In an effort to narrow the minimal signal initiation region of Rx, we created a series of truncations within the NB domain. An Rx 149-296 fragment which deletes the anchor motif retained its ability to induce an HR (see Appendix C), further indicating that signal initiation is mechanistically distinct from nucleotide binding. Further N-terminal truncations, however, resulted in loss of HR-inducing activity accompanied by a reduction in stability, despite EGFP:HA fusion (Figure 3.4B and 3.4C). Due to the persistent complication of NB fragment instability, we next turned to a site-directed mutagenesis approach. As signal initiation is likely a conserved function among NB-LRR proteins and appears to be distinct from nucleotide binding, we focused on motifs conserved specifically among NB-LRR proteins, systematically replacing six-amino acid stretches of the RNBS-A, LXXLL, and RNBS-B motifs with the amino acid sequence "NAAIRS" in the Rx NB:EGFP:HA context (Figure 3.4A). The NAAIRS sequence was chosen because it is found in both α-helices and β-sheets (Wilson et al. 1985), suggesting that it is unlikely to cause gross disruptions to secondary structure (see Chapter 2). While NAAIRS substitutions RNBS-B2 resulted in a loss of HR-inducing activity accompanied by a significant reduction in protein accumulation, other substitutions accumulated to levels comparable to wild-type (Figure 3.4B). None of these stable variants retained the ability of wild-type to induce HR upon agroexpression (Figure 3.4C), although the RNBS-A3 variant was observed sometimes to induce mild chlorosis, indicating that the C-terminal portion of the RNBS-A motif may not be strictly required for signal initiation. These findings further suggest that the N-terminal portion of the RNBS-A motif and/or the LXXLL motif

may be involved in the induction of downstream defense responses. As the structure of the NB domain is clearly sensitive to perturbations, however, it is also possible that these large substitutions might distort the protein fragment such that it is no longer signaling-competent, without themselves residing within the signal initiation region.

Point mutation of conserved residues

In addition to previously defined NB-LRR-conserved motifs, amino acid sequence comparison of aligned NB domains of various CC-NB-LRR proteins reveals a number of other well-conserved regions as well (Figure 3.5A). From among these, we selected six residues or pairs of nearby residues with likelihood of playing a functional rather than structural role and introduced alanine mutations at these sites (Figure 3.5A, Table 3.1). We also generated a pair of alanine mutations (VL210,214AA) within the LXXLL motif, in an effort to minimize the possibility of gross structural alterations complicating analysis of this region. We additionally generated a series of three alanine mutations (TT271,272AA; R273A; and TTR271,272,273AAA) at the C-terminus of the RNBS-B in order to better assess the functions of this motif. The HR-inducing capacity and the stability of these Rx NB:EGFP:HA point mutant variants was subsequently evaluated (Figure 3.5B and 3.5C, Table 3.1). While the majority of alanine mutations were reasonably stable, there was some fluctuation observed in protein levels relative to wild-type (Table 3.1). This is likely due to the rapid cell death associated with the expression of those variants – including wild-type – which retain their HR-inducing activity. All protein samples were collected approximately 36 hours after agroinfiltration, at a point when transiently expressed proteins had accumulated to levels sufficient for detection, but before the onset of visible necrosis. However, development of NB:EGFP:HA-mediated HR following agroexpression

Figure 3.5 *Mutational analysis of conserved residues.* **A,** Multiple sequence alignment of CC-NB-LRR NB domains. Consensus and consensus strength are shown at top, with conserved motifs designated by brackets. Aligned sequences are shown in the bottom panel, with conserved residues shaded black and residues with similar functional groups shaded gray. Stars indicate Rx residues selected for alanine mutation, with color designating the effect of each mutation. Gray – no effect on NB-mediated HR; black – unstable protein, no conclusion possible; dark blue – compromised NB-mediated HR; light blue – compromised NB-mediated HR and CC binding. **B,** HR-inducing activity of Rx alanine mutants. The Rx NB:EGFP:HA derivatives were agroexpressed as indicated in *N. tabacum* leaves and photographed after two days. **C,** Relative stability of Rx alanine mutants. The indicated Rx NB:EGFP:HA derivatives were agroexpressed in *N. benthamiana*. Protein samples were collected after 36 hours for anti-HA immunoblotting (top panel). Equal loading was confirmed by coomassie staining (bottom panel). A representative immunoblot is presented. Determination of relative stability was based on four replicate analyses (Table 3.1).


A
Anchor
Kinase 1
RNBS-A
LXXLL
Kinase 2
Kinase 3a
ENXMVGRENEFEXLLDQLLRG-SR---ELDVVSIVGMGGIGKTTLAKKVYNDP-LIXSRFDVRAWVTVSQEXNVREILLGILSQVGDEDD---EXX-DQ-LADKLRKXLKGKRYLIVLDDXWDT--EXWDDLKLCFP-XXGX-GSRIILTTRNXEVAXYAG
Rx ENIMVGRENEFEMMLDQLARG-GR---ELEVVSIVGMGGIGKTTLATKLYSDP-CIMSRFDIRAKATVSQEYGVRNVLLGLLSLTSDEPD-------DQ-LADRLQKHLKGRRYLVVIDDIWTT--EAWDDIKLCFP-DCYN-GSRILLTTRNVEVAEYAS 282
Rx2 ENIMVGRENEFEMMLDQLVRG-GR---ELEVVSIVGMGGIGKTTLATKLYSDP-YIMSRFDIRAKATVSQEYGVRNVLLGLLSLTSDEPD-------DQ-LADRLQKHLKGRRYLVVIDDIWTT--KAWDGIKLCFP-DCYK-GSRILLTTRNVEVAEYAS 282
Gpa2 ENIMVGRENEFEMMLDQLARG-GR---ELEVVSIVGMGGIGKTTLAAKLYSDP-YIMSRFDIRAKATVSQEYGVRNVLLGLLSLTSDEPD-------YQ-LADQLQKHLKGRRYLVVIDDIWTT--EAWDDIKLCFP-DCDN-GSRILLTTRNVEVAEYAS 282
Bs2 KNNMVGRDDQRKQLLEDLTRSYSG---EPKVIPIVGMGGIGKTTLAKEVYNDE-SILGRFDVRAWATISQQHNKKEILLGLLHSTIKMDDRVKMIGEAE-LADMLQKSLKRKRYLIVLDDIWSC--EVWDGVRRCFP-TEDNAGSRILLTTRNDEVAQYAG 301
Hero VEVTVGHEEDEARIIGQLLDEHES---KLDVISIVGMPGVGKTTLANKVYNNT-LVASHFKIRAKCTVSQNFNKSKVLREILQQVTASETN---RSEDD-LAEKLRVALLDKRYLIVLDDVWDI--ATGEMLIACFP-KVER-GNRVILTSRSGEVGLKVK 665
I2 ESDIFGRQSEIEDLIDRLLSEGASG-KKLTVVPIVGMGGQGKTTLAKAVYNDE-RVKNHFDLKAWYCVSEGFDALRITKELLQEIGKFDSKDVHNNLNQ-LQVKLKESLKGKKFLIVLDDVWNENYNEWNDLRNIFA--QGDIGSKIIVTTRKDSVALMMG 322
LR10 AADLVGIDGPKHELVKWLRNGEDESVHQQKVVSIVGCAGLGKTTLAKQVYD---ELRINFEYRAFVSISRSPNMATILKCVLSQFHAQDYSSDESEIPK-LVDQIRDLLQDKRYFVIIDDIWDM--KTWDVLKCAL--CKNSQGSVIMTTTRIYDVAKSCC 315
Mi-1 DKITVGFEEETNLIRKLTSG-SA---DLDVISITGMPGSGKTTLAYKVYNDK-SVSSRFDLRAWCTVDQGCDEKKLLNTIFSQVSD-SDS-KLSENID-VADKLRKQLFGKRYLIVLDDVWDT--TTWDELTRPFP-ESKK-GSRIILTTREKEVALHGK 666
Prf NEEMEGFQDTIDELKDKLLGG-SP---ELDVISIVGMPGLGKTTLAKKIYNDP-EVTSRFDVFAQCVVTQLYSWRELLLTILNDVLEPSDR-NEKEDGE-IADELRRFLLTKRFLILIDDVWDY--KVWDNLCMCFS-DVSN-RSRIILTTRLNDVAEYVK 1240
R3a DSDIFGRQNDIEDLIDRLLSEDASG-KKRTVVPIVGMGGLGKTTLAKAVYNDE-RVQIHFGLKAWFCVSEAFDAFRITKGLLQEIGSFDLK-ADDNLNQ-LQVKLKERLKGKKFLIVLDDVWNDNYNKWDELRNVFV--QGDIGSKIIVTTRKESVALMMG 328
RPS5 QPTIVGQEIMLEKAWNRLMEDGS------GILGLYGMGGVGKTTLLTKINNKFSKIDDRFDVVIWVVVSRSSTVRKIQRDIAEKVGLGGMEWSEKNDNQ-IAVDIHNVLRRRKFVLLLDDIWEK--VNLKAVGVPYP--SKDNGCKVAFTTRSRDVCGRMG 303
Sw5-b DEEVVGFENKAEELIDYLIRG-TN---ELDVVPIVGMGGQGKTTIARKLYNND-IIVSRFDVRAWCIISQTYNRRELLQDIFSQVTG-SDD-NGATVDV-LADVLRRKLVGKRYLIVLDDMWDC--MVWDDLRLSFP-DDGI-RSRIVVTTRLEEVGKQVK 679
Tm-2a ETEIIGLDDDFNMLQAKLLN--QD--LHYGVVSIVGMPGLGKTTLAKKLYR---LIRDQFECSGLVYVSQQPRASEILLDIAKQIGLTEQKMKEN-----LEDNLRSLLKIKRYVILLDDIWDV--EIWDDLKLVLPECDSKVGSRMIITSRNSNVGRYIG 300
HRT ESDLVGVEQSVTELVCHLVEN-----DVHQVVSIAGMGGIGKTTLARQVFHHD-LVRRHFDGFAWVCVSQQFTQKHVWQRILQELQPHDGDILQMDESA-LQPKLFQLLETGRYLLVLDDVWKK--EDWDRIKAVFP---RKRGWKMLLTSRNEGVGIHAD 311
RPS2 IKSVVGNTTMMEQVLEFLSEEEER-----GIIGVYGPGGVGKTTLMQSINNELITKGHQYDVLIWVQMSREFGECTIQQAVGARLGLS---WDEKETGENRALKIYRALRQKRFLLLLDDVWEE--IDLEKTGVPRP--DRENKCKVMFTTRSIALCNNMG 300
B
wt
F193
VL210
TT271
R273
TTR271
KY182
LL226
DR228
WD252
FP258
C
wt
F193
VL210
TT271
R273
TTR217
KY182
LL226
DR228
WD252
FP258
IB: αHA
coomassie

varies within a window of several hours. Hence variation in the precise point prior to or during the early stages of programmed cell death at which samples were collected constitutes a likely confounding factor, with the potential to skew comparisons between HR-inducing and non-HR-inducing variants. We therefore rated each variant's stability based on four replicate trials (Table 3.1), and excluded from consideration any mutations which resulted in reduced stability relative to wild-type in 50% or more of trials. Thus NB:EGFP:HA F193A, LL226,230AA, and TT271,272AA were deemed to be relatively unstable and therefore functionally uninformative (Figure 3.5A and 3.7).

Table 3.1 Summary of NB point mutation impacts on stability and function*

Replicate / Mutation	NB:EGFP								NB-ARC-LRR						
	Stability[a]				HR[b]				HR[c]			IP[d]		Co-IP[e]	
	A	B	C	D	A	B	C	D	A	B	C	A	B	A	B
wild-type	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
KY182,184AA	++	+	++	+	+	+	+	+	+	(+)	+	+	++	+	++
F193A	+	(+)	(-)	(+)	-	-	-	-	+	(+)	+	+	+	(+)	(+)
VL210,214AA	+	+	+	+	-	-	-	-	+	+	+	+	+	+	+
LL226,230AA	+	(+)	(-)	(+)	-	-	-	-	+	(+)	+	+	+	(-)	+
DR228,229AA	++	+	++	++	(+)	-	(-)	-	+	(+)	+	+	++	+	++
WD252,253AA	+	+	+	(+)	-	-	-	-	-	-	-	+	(+)	-	-
FP258,259AA	+	+	+	(+)	-	-	-	-	-	-	(-)	+	(+)	-	(-)
TT271,272AA	+	(+)	(-)	(+)	-	-	-	-							
R273A	+	(+)	+	+	+	+	+	+							
TTR271,272,273AAA	+	+	(-)	+	-	-	-	-							

* + = as wild-type, ++ = greater than wild-type, (+) and (-) = less and significantly less than wild-type, respectively
[a] relative NB:EGFP:HA protein accumulation in immunoblot
[b] HR induction by NB:EGFP:HA variants
[c] HR induction following co-expression of NB-ARC-LRR:Myc + CC:HA + CP
[d] relative accumulation of NB-ARC-LRR:Myc protein following immunoprecipitation (IP)
[e] strength of CC:HA – NB-ARC-LRR:Myc interaction following coimmunoprecipitation (Co-IP)

Rx NB:EGFP:HA alanine mutants KY182,184AA (KY182), VL210,214AA (VL210), DR228,229AA (DR229), WD252,253AA (WD252), FP258,259AA, R273A (R273), and TTR271,272,273AAA (TTR271) were all reasonably stable (Table 3.1, Figure 3.5C). Interestingly, all of these variants save KY182 and R273 exhibited compromised HR-induction, with DR228 being partially compromised and the rest displaying a complete loss of HR-inducing activity (Table 3.1, Figure 3.5B). Residues VL210,214, DR228,229, WD252,253, and FP258,259 all therefore represent possible candidates for signal initiation regions.

In order to better understand the functional impact of the alanine mutations described above, we transferred some of these into the longer NB-ARC-LRR context to assess their ability to bind the Rx CC domain *in trans*, as well as their ability to reconstitute Rx function when co-expressed with the CC domain and PVX CP. NB-ARC-LRR variants fused to six c-Myc epitope tags (NB-ARC-LRR:Myc) were co-expressed with Rx CC:HA and subjected to coimmunoprecipitation and immunoblotting (Figure 3.6C). While mutant variants F193, VL210, KY182, LL226, and DR228 all retained the ability to interact with the CC domain, WD252 and FP258 did not. Reconstitution of Rx function mirrored CC binding ability. Thus upon co-expression with Rx CC:HA plus CP, NB-ARC-LRR:Myc variants F193, VL210, KY182, LL226, and DR228 all induced an HR in *N. benthamiana* leaves similar to wild-type, whereas the WD252 variant was not HR-inducing, and the FP258 variant was only weakly so (Figure 3.6A). It is surprising that while the latter two mutations had no significant impact on NB:EGFP:HA stability, they did result in decreased NB-ARC-LRR:Myc accumulation (Table 3.1, Figure 3.6B). Failure of these two variants to bind the CC domain and to reconstitute Rx function might thus be partially due to compromised stability, or may alternately indicate that WD252,253 and FP258,259

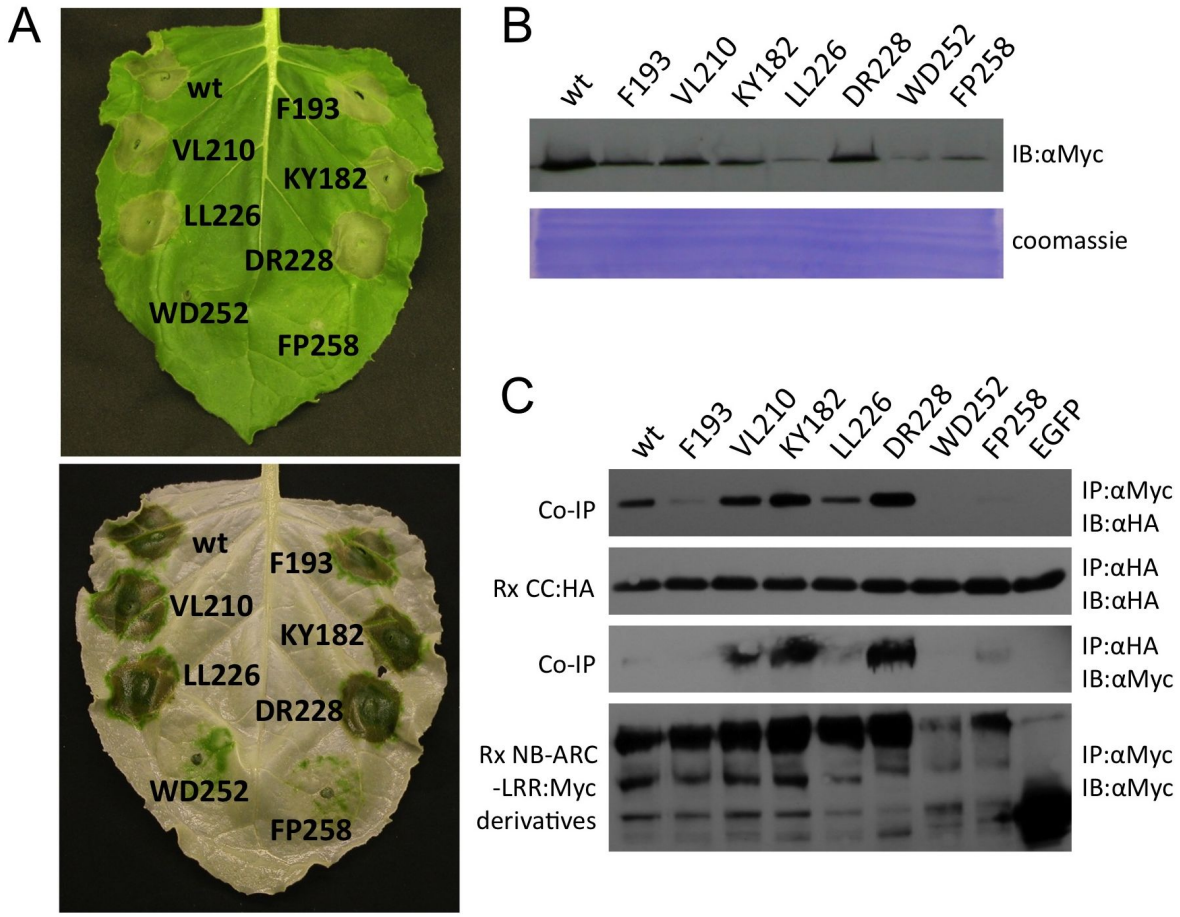


Figure 3.6 *Impact of alanine mutations on Rx NB-ARC-LRR.* **A,** HR induction by reconstituted Rx alanine mutants. The indicated derivatives of Rx NB-ARC-LRR:Myc were co-expressed with Rx CC:HA and PVX CP in *N. benthamiana* leaves. Leaves were photographed three days after agroinfiltration (top panel), cleared with ethanol, and photographed again (bottom panel). **B,** Relative stability of alanine mutation variants. The indicated Rx NB-ARC-LRR:Myc derivatives were agroexpressed in *N. benthamiana* leaves. Protein samples were collected two days later and subjected to anti-HA immunoblotting (top panel). Equal loading was confirmed by coomassie staining. **C,** CC binding capacity of Rx NB-ARC-LRR alanine mutants. The indicated derivatives of Rx NB-ARC-LRR:Myc were co-expressed with Rx CC:HA in *N. benthamiana* leaves. Two days later protein extracts were collected and subjected coimmunoprecipitation (Co-IP). Protein extracts were subjected to either anti-HA or anti-Myc immunoprecipitation (IP), followed by either anti-HA or anti-Myc immunoblotting (IB), as indicated.

residues are involved to some extend in mediating the interaction between the CC domain and the NB-ARC-LRR region. The mediation of such an interaction may have structural significance, perhaps in turn offering an explanation for the instability associated with mutation at these sites in a longer context. The observation that all mutations which abolished NB:EGFP-mediated HR were capable of HR-induction in the context of reconstituted Rx suggests that the threshold for signaling may be higher in NB:EGFP than in NB-ARC-LRR + CC + CP. That is, mutations which are seen to abolish HR-induction in the NB:EGFP context likely take the protein fragment below a threshold for autoactivity, while not actually compromising signal initiation by the reconstituted whole protein.

Structural analysis

In order to investigate the structural basis for our functional findings related to signal initiation, we constructed a three-dimensional model of the Rx NB-ARC nucleotide binding pocket based on the α/β fold regions of multiple templates including *Homo sapiens* Apaf-1 and *Caenorhabditis elegans* CED-4, whose structures have previously been solved (Riedl et al. 2005; Qi et al. 2010). While structure of the nucleotide binding pocket interior is highly conserved, it is less certain whether exterior surface structures are equally well-conserved. Therefore, in order to verify as accurate a representation as possible we employed two different automated prediction servers. The model presented in Figure 3.7 was constructed using HHPred (Söding 2005) for template identification and alignment and MODELLER (Sali et al. 1995) for creation of the three-dimensional structure. The predicted Rx NB-ARC structure differs from that of Apaf-1 by an RMSD (root mean square deviation) value of 1.48Å, and from that of CED-4 by 1.49Å. The structure presented in Figure 3.8B was created

by I-TASSER (Roy et al. 2010), and differs from Apaf-1 and CED-4 by 0.90Å and 1.64Å RMSD values, respectively. The two Rx NB-ARC structural predictions differ from one another by an RMSD of 1.39Å. The predicted structures of the NB domain can be seen to be largely similar, with a core of five β-strands surrounded by a pair of α-helices proximal to the nucleotide binding pocket and two pairs of α-helices and/or loops located distally to the nucleotide binding pocket (Figures 3.7 and 3.8B). As the two models are in agreement with each other as well as with previously published models of the NB-ARC domains of Mi-1.2 and I-2 (van Ooijen et al. 2008a; van Ooijen et al. 2008b) regarding these features, we take them to be a fairly accurate representation.

Upon superimposing the location of our alanine mutations onto the Rx NB-ARC three-dimensional structure, it can be seen that while mutations not impacting NB:EGFP-mediated signaling are located in and around the nucleotide binding pocket, those which abolished NB:EGFP-mediated HR induction are located within the distal loops and α-helices (Figure 3.7). Further, while WD252 and FP258 mutations which affected both HR-induction and CC binding are located within one set of paired helices and/or loops, the two mutations which impacted HR-induction exclusively – VL210 and DR228 – are located within what appears to be the more exposed and distal pair of helices, indicating that this region may be important to signal initiation. This observation strengthens the supposition that signal initiation and nucleotide binding are distinct functions. In particular, it is interesting to note that mutation of the R273 residue, which belongs to the kinase-3a motif and is predicted to extend into the nucleotide binding pocket and contact the bound nucleotide (Traut 1994), has no effect on signal initiation.

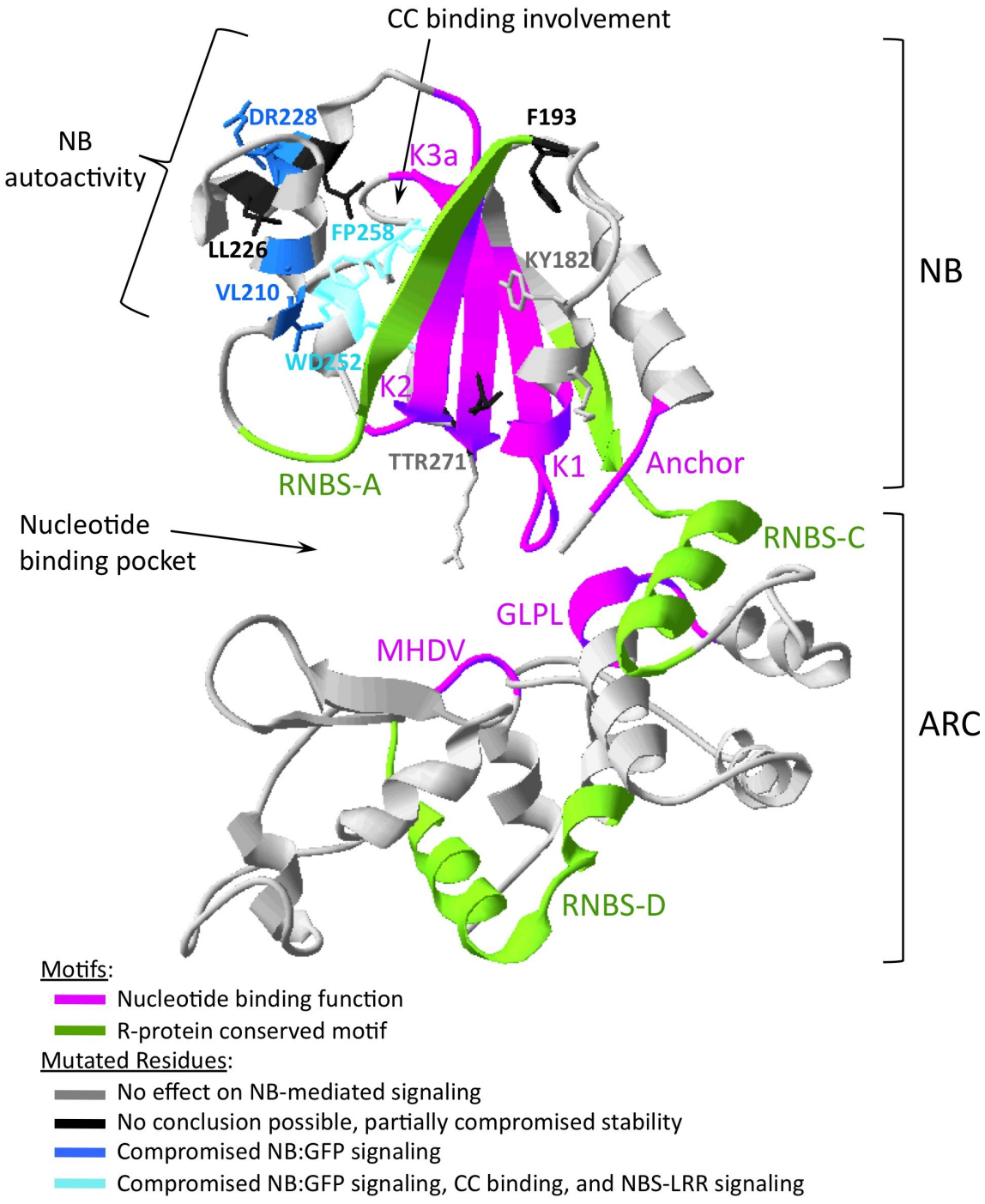
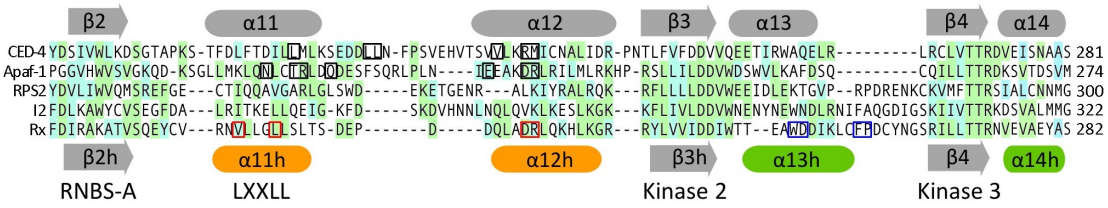


Figure 3.7 *Structural model of the Rx NB-ARC region.* A three-dimensional model of the Rx NB-ARC region was constructed based on the structures of Apaf-1 and CED-4. Locations of previously identified conserved motifs are shaded as indicated. Residues targeted by alanine substitutions in this study are labeled and shaded as indicated. A probable Rx signaling surface responsible for autoactivity is also indicated, as is the location of mutations found to impact CC binding.

As the distal structures containing residues potentially involved in signal initiation are both spatially removed from the nucleotide binding pocket and do not house any motifs known to be involved in nucleotide binding, we were curious what other functions might be attributed to the corresponding region in better-studied systems. Apaf-1 and CED-4 are components of the metazoan apoptotic machinery which, when properly stimulated, assemble into heptameric and octameric rings, respectively (Riedl et al. 2005; Qi et al. 2010). These wheel-like complexes constitute the apoptosome, which is responsible for the activation of downstream caspases leading to cell death. A high-resolution crystal structure of the complete CED-4 apoptosome has recently been described (Qi et al. 2010), and shows that the α/β fold domain is central to mediating oligomerization between CED-4 monomers. Within this domain, helices α11 and α12 were found to be particularly closely involved in oligomerization, containing multiple residues which reside at the oligomerization interface. Mutation of these residues was found to cripple apoptosome formation in both CED-4 and Apaf-1, further supporting the role of this region in mediating oligomerization. We aligned the NB domain of Rx with the α/β fold domains of CED-4 and Apaf-1 in order compare the location of our Rx mutations of interest (Figure 3.8A). The I2 and RPS2 NB-LRR protein NB domains were also included for comparison. The Rx VL210 mutation can be seen to lie within a predicted helix homologous to helix α11 of CED-4 (α11h), while the Rx DR228 mutation resides in predicted helix α12h, corresponding to CED-4 helix α12 (Figure 3.8A and 3.8B). Thus the region we have identified as a potential source of signal initiation for Rx corresponds to the region mediating oligomerization in CED-4 and Apaf-1. Strikingly,

A



B

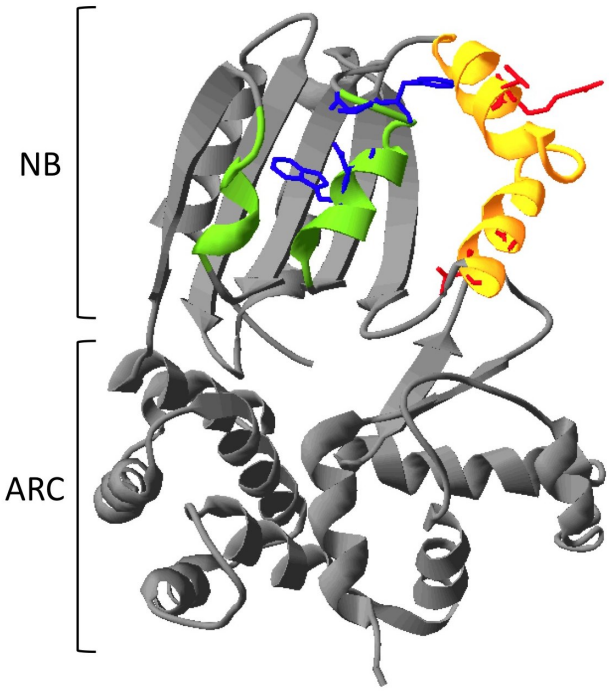


Figure 3.8 *Structure/function comparison of Rx and CED-4.* **A,** Alignment of amino acid sequence from the α/β fold domains of CED-4 and Apaf-1 with portions of the RPS2, I2, and Rx NB domains. Secondary structural features (α-helices and β-strands) of CED-4 are depicted at top, and predicted homologous (h) structural features of Rx are depicted at bottom, labeled additionally with the corresponding previously defined conserved motif, where applicable. CED-4 and Apaf-1 residues which when mutated by Qi et al. (2010) compromised apoptosome assembly are boxed in black. Residues of Rx which compromised NB-mediated HR induction when substituted with alanine are boxed in red. Rx alanine mutations which compromised both HR induction and CC binding are boxed in blue. **B,** Structural model of the Rx NB-ARC region. This model is presented at an approximately 180° rotation on the vertical axis relative the model in Figure 3.7. Helices α11h and α12h are shaded orange, while helices α13h and α14h are shaded green. Sidechains of residues where alanine mutation compromises NB-mediated HR induction are shown in red. Sidechains of residues where mutations compromise both HR induction and CC binding are shown in blue. Red and blue residues correspond to those boxed in (**A**).

the Rx residues targeted by the DR228 mutation correspond precisely to residues found to mediate oligomerization among CED-4 and Apaf-1 monomers (Figure 3.8A) (Qi et al. 2010). Conversely, mutations found to impact both Rx signaling and CC binding can be seen to reside in and around predicted helix α13h (Figure 3.8A and 3.8B). No particular role has yet been ascribed to this structural feature, other than its position flanking the kinase-2 motif, and unlike helices α11 and α12, there appears to be no indication of a specific role for helix α13 in oligomerization.

NB-mediated oligomerization

Given the structural correlation between the sites of Rx signal initiation and CED-4/Apaf-1 oligomerization, we were curious whether the Rx NB domain might also mediate oligomerization. We therefore performed coimmunoprecipitation experiments between Rx NB:EGFP wild-type and mutant variants fused to HA and c-Myc epitope tags. While Rx NB:EGFP does not interact with the full-length protein, the Rx NB domain was found to be capable of homotypic oligomerization in the context of NB:EGFP (Figure 3.9A). Surprisingly, all alanine mutants tested, including those leading to signaling inactivity, also retained the ability to interact with wild-type NB:EGFP (Figure 3.9A). To evaluate whether the presence of wild-type protein masks an impact of mutations on oligomerization, we also tested HA- and c-Myc-tagged mutant variants against themselves (Figure 3.9B). Contrary to expectations, most variants retained the ability to oligomerize to varying extents, with active variants tending to bind with less efficiency than inactive variants. As such wild-type and KY182 variants were seen to bind most weakly, followed by the slightly active DR228 variant. While indicating that the Rx NB domain is indeed capable of oligomerization,

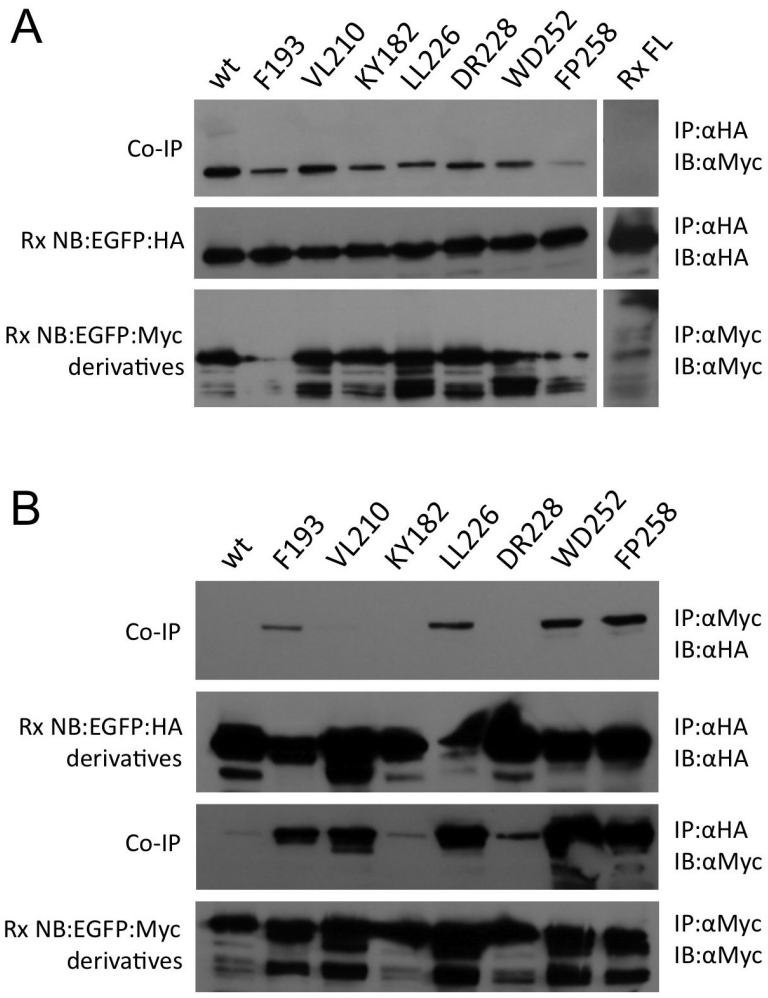


Figure 3.9 *Homotypic oligomerization of the Rx NB domain.* **A,** Coimmunoprecipitation of Rx NB:EGFP alanine substitution variants with wild-type Rx NB:EGFP. The indicated Rx NB:EGFP:Myc variants were coexpressed with wild-type Rx NB:EGFP:HA in *N. benthamiana*. 36 hours after agroinfiltration protein samples were collected and subjected to immunoprecipitation and immunoblotting as indicated (left panels). In a separate experiment, full-length Rx was similarly examined for coimmunoprecipitation with wild-type Rx NB:EGFP:HA (right panels). **B,** Coimmunoprecipitation of Rx NB:EGFP alanine substitution variants. HA- and Myc-tagged versions of the indicated Rx NB:EGFP derivatives were co-expressed in *N. benthamiana* leaves. Protein samples were collected 36 hours after agroinfiltration and subjected to immunoprecipitation and immunoblotting as indicated.

these findings do not present a clear functional mechanism, as HR-inducing activity seems if anything to be correlated with loss of oligomerization. It should be noted, however, that onset of HR continues to be a confounding factor in assays involving active Rx NB variants. This is evident upon contrasting the strong wt – wt binding from Figure 3.9A with the nearly absent wt – wt binding present in Figure 3.9B. As these experiments were performed under similar conditions with similar sampling time points, variation observed in protein behavior is most likely related to variation in the precise stage prior to visible cell death at which samples were collected.

As we have thus-far only identified Rx and closely related proteins as being capable of NB-mediated HR induction, we sought to investigate whether the phenomenon of NB-mediated oligomerization would also be restricted to Rx. We therefore co-expressed the EGFP-stabilized NB domains of Bs2, Prf, and RPS5 with that of Rx, followed by coimmunoprecipitation. In this case we used the VL210 mutant variant of Rx NB:EGFP, as it retains the ability to bind wild-type Rx NB:EGFP (Figure 3.9A), yet its inability to induce HR allowed for a longer period of protein accumulation *in planta* prior to extraction. Intriguingly, we found that the EGFP-stabilized NB domains of Bs2 and RPS5 interacted robustly with that of Rx, with Prf interacting to a lesser extent (Figure 3.10A). Additionally, we found that an Rx NB:EGFP truncation which deletes the kinase-1a motif (Rx Δ190) and abolishes HR induction (data now shown) retains its ability to interact with wild-type Rx NB:EGFP (Figure 3.10A). These observations suggest that, even if the two activities are functionally related, the threshold for NB autoactivity is higher than the threshold for oligomerization.

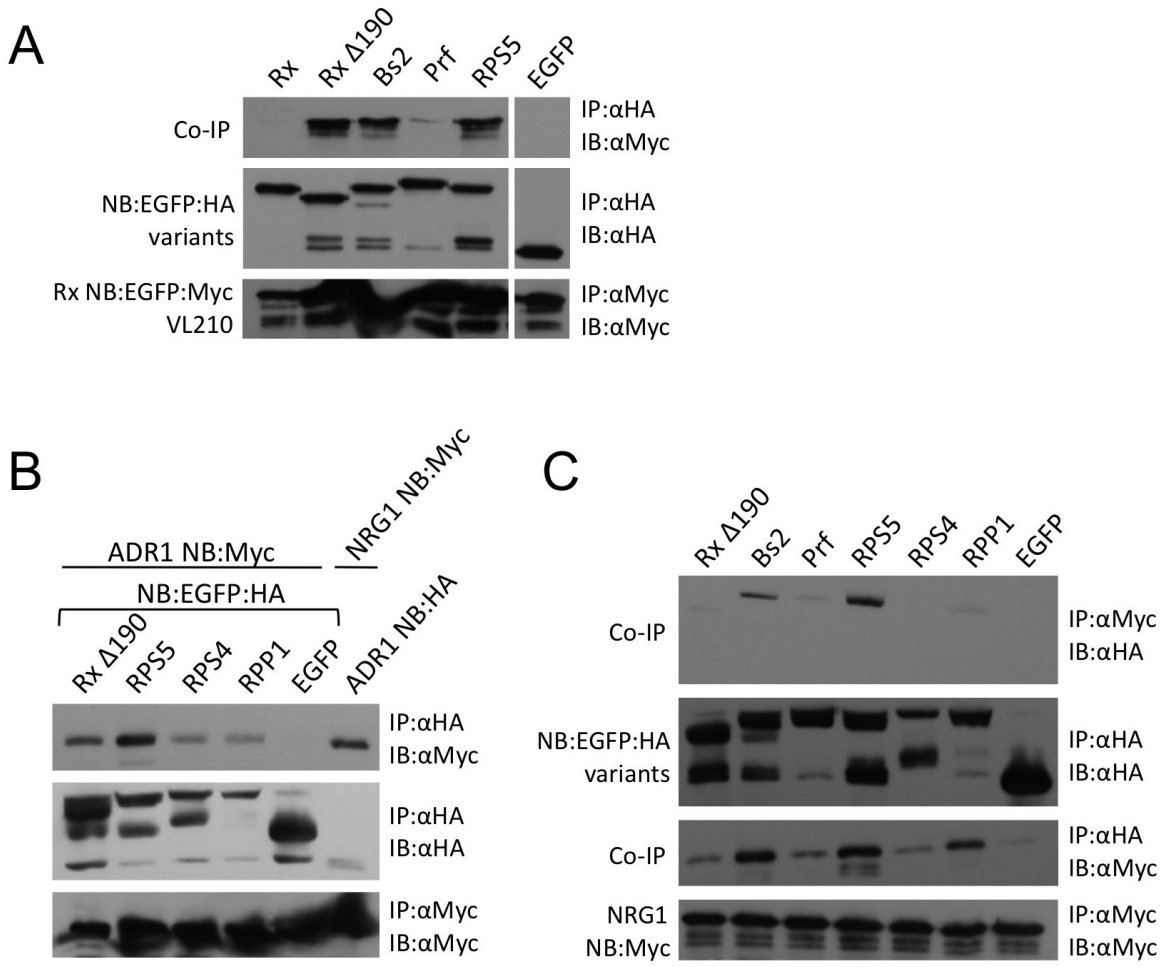


Figure 3.10 *Heterotypic oligomerization of NB-LRR NB domains*. **A,** Interaction between the NB domains of other NB-LRR proteins and that of Rx. NB:EGFP:HA fusion proteins from the indicated NB-LRR proteins, as well as EGFP:HA, were co-expressed with Rx NB:EGFP:Myc-VL210,214AA in *N. benthamiana* leaves. After 36 hours protein samples were collected and subjected to immunoprecipitation and immunoblotting as indicated. **B,** Interaction between various NB-LRR NB domains and the NB domain of ADR1. NB:EGFP:HA fusion proteins derived from the indicated NB-LRR proteins, or EGFP:HA, were co-expressed with ADR1 NB:Myc in *N. benthamiana* leaves. In the right-most lane ADR1 NB:HA was co-expressed with NRG1 NB:Myc. Two days after agroinfiltration protein samples were collected and subjected to immunoprecipitation and immunoblotting as indicated. **C,** Interaction between various NB-LRR NB domains and the NRG1 NB domain. NB:EGFP:HA fusion proteins derived from the indicated NB-LRR proteins, or EGFP:HA, were co-expressed with NRG1 NB:Myc in *N. benthamiana*. Two days after agroinfiltration protein samples were collected and subjected to immunoprecipitation and immunoblotting as indicated.

We further investigated the ability of the EGFP-stabilized NB domains of various NB-LRR proteins to interact with the NB domains of CC_R-NB-LRR proteins ADR1 of potato and NRG1 of *N. benthamiana*. As will be described further in Chapter 4, this unique and highly conserved class of NB-LRR protein is of particular interest due to its signaling-competent N-terminal domain. The NB domains of both TIR-NB-LRR and CC-NB-LRR protein classes were observed to be capable of interaction with the NB domains of ADR1 and NRG1, as were ADR1 and NRG1 NB domains with each other (Figure 3.10B and 3.10C). Furthermore, as the NB domains of ADR1 and NRG1 are reasonably stable withoug EGFP fusion, we were able to rule out the possibility that observed NB oligomerization might actually be the result of EGFP – EGFP interaction.

As there appears to be an intriguing yet poorly understood functional relationship between NRG1-like proteins and TIR-NB-LRR proteins (see Chapter 4), we were curious whether this relationship would be evident in the binding affinity between CC-NB-LRR / TIR-NB-LRR proteins vs. ADR1 / NRG1 proteins. While NB:EGFP fusion proteins derived from CC-NB-LRR proteins interact stably with both ADR1 and NRG1 NB domains (Figure 3.10B and 3.10C), NB:EGFP fragments derived from the TIR-NB-LRR proteins RPS4 and RPP1 displayed relatively low levels of interaction with ADR1 (Figure3.10B). However, TIR-NB-LRR-derived fragments were also somewhat less stable than those of CC-NB-LRR proteins. A similar pattern of low accumulation coupled with reduced ADR1-binding was also observed for a third TIR-NB-LRR-class NB:EGFP fusion protein, derived from the tomato (*Solanum lycopersicum*) Bs4 protein (data now shown). Thus we are currently unable to

establish whether selectivity in binding exists, due to the compromised stability of TIR-NB-LRR NB:EGFP fusion proteins.

Proteins able to interact with the Rx NB domain represent potential candidates for NB-LRR protein downstream signaling partners. Because the signaling-competent CC_R domain of CC_R-NB-LRR proteins makes them particularly intriguing in this role, we undertook further examination of the interaction between Rx and ADR1 NB domains using the previously described Rx NB:EGFP alanine mutation variants. As with Rx – Rx interaction, all variants were found to coimmunoprecipitate with the ADR1 NB domain, with active variants binding more weakly than inactive variants (Figure 3.11A). Additionally, *SGT1* silencing was found to have no significant impact on binding between the Rx and ADR1 NB domains (Figure 3.11B), suggesting that although SGT1 is required for total protein function, its presence is not likely necessary for oligomerization between NB domains. Mutations in the P-loop (PL) of the Rx kinase-1a motif do not compromise binding, either to the wild-type Rx NB domain or to that of ADR1 (Figure 3.11C), suggesting that oligomerization is independent of nucleotide binding, and occurs through a different mechanism than does the intramolecular CC – NB-ARC-LRR interaction, which is abrogated by the P-loop mutation (Moffett et al. 2002). Thus insensitivity to nucleotide binding status is a trait common to both NB-mediated HR induction and NB – NB oligomerization.

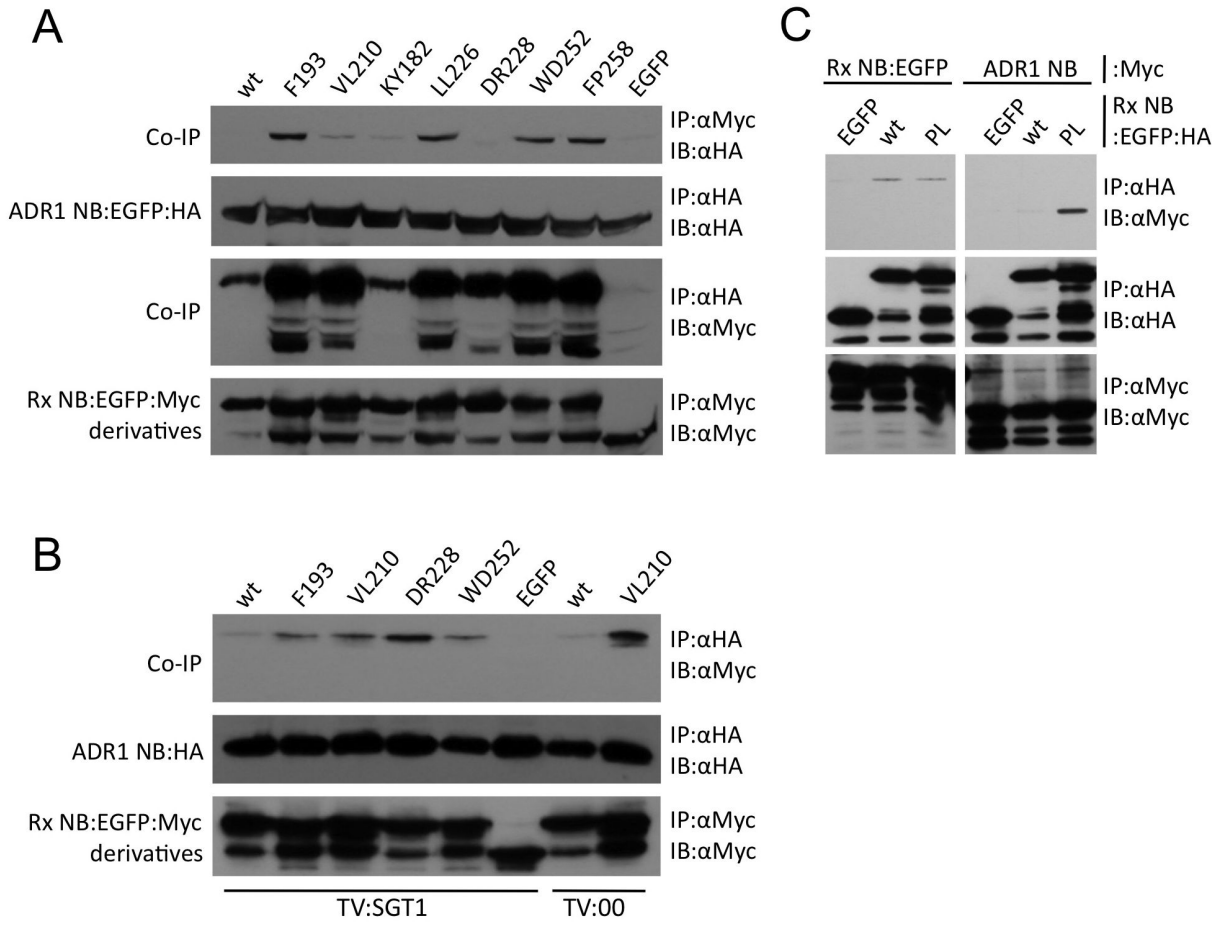


Figure 3.11 *Impact of mutations within the Rx NB domain on heterotypic oligomerization.* **A,** Interaction of Rx NB:EGFP derivatives with ADR1 NB:EGFP. The indicated Rx NB:EGFP:Myc derivatives, or EGFP:Myc, were co-expressed with ADR1 NB:EGFP:HA in *N. benthamiana*. 36 hours after agroinfiltration protein samples were collected and subjected to immunoprecipitation and immunoblotting as indicated. **B,** Effect of *SGT1* silencing on NB oligomerization. The indicated Rx NB:EGFP:Myc derivatives, or EGFP:Myc, were co-expressed with ADR1 NB:HA in *N. benthamiana* which had been infected with TRV empty vector (TV:00) or TRV carrying and insert from *N. benthamiana SGT1* (TV:SGT1), as indicated, three weeks previously. Two days after agroinfiltration protein samples were collected and subjected to immunoprecipitation and immunoblotting as indicated. **C,** Effect of P-loop mutation on Rx NB homotypic and heterotypic oligomerization. Wild-type Rx NB:EGFP:HA, the same fusion protein carrying a P-loop mutation (PL), or EGFP:HA were co-expressed with Rx NB:EGFP:Myc and ADR1 NB:Myc, as indicated, in *N. benthamiana* leaves. After 36 hours protein samples were collected and subjected to immunoprecipitation and immunoblotting as indicated.

DISCUSSION

Over-expression of the CC-NB fragment of Rx as well as the CC-NB-ARC fragments of RPS2 and RPS5 have been previously shown to induce HR, while over-expression of their respective CC domains does not (Tao et al. 2000; Bendahmane et al. 2002; Ade et al. 2007). Here, we report that HR induction by the N-terminal portion of Rx derives from the NB, rather than the CC domain. This could only be observed, however, after stabilization of the Rx NB domain by fusion to EGFP (Figure 3.1). Inherent instability of NB domains (see Figure 3.3) may explain why over-expression of NB domains has not been reported for other NB-LRR proteins. Alternately, the activity of longer NB-containing fragments of other NB-LRR proteins may reflect a role for additional domains in stabilizing or enhancing the activity of the NB domain. For example, interaction with the LRR domain is a prerequisite for an Rx CC-NB-ARC fragment containing an autoactivating mutation to be competent for Avr-independent signaling (Moffett et al. 2002). Thus although over-expression of the Rx NB domain allows for deregulated activity, signaling may be more readily accomplished in more complete context(s). Such a scenario might in part explain why the NB domains of only Rx, Rx2, and Gpa2, compared to the similarly stable derivatives of a number of other NB-LRR proteins, were able to independently induce an HR (Figure 3.3). The NB domains of these proteins might be uniquely predisposed to autoactivity, such that while other NB-LRR proteins are perfectly capable of signal initiation in their full-length context, their NB domains are more tightly regulated against the autoactivity displayed by those of Rx, Rx2, and Gpa2.

The distinction between the autoactive phenotype of the isolated Rx NB domain and natural signal initiation by the full-length molecule may also explain why a functional nucleotide binding pocket is not necessary for the isolated NB domain to induce an HR. An autoactive mutant of the I-2 NB-LRR protein is inactivated by the incorporation of an additional mutation within the P-loop (Tameling et al. 2006), indicating that nucleotide binding and/or hydrolysis likely plays a role in allowing the full-length protein to adopt a signaling-competent state. In the isolated NB domain these restraints appear to be overcome, and thus nucleotide binding is no longer necessary. Furthermore, this suggests that an NB motif other than those involved in nucleotide binding is likely to interact with downstream signaling factors.

Through mutational analysis we have identified the region of the Rx NB domain defined by helices α11h and α12h as a likely origin of signal initiation (Figures 3.5 through 3.8). The VL210 and DR228 alanine mutations located in this region abolish NB-mediated HR induction, whereas WD252 and FP258 mutations located in and around helix α13h interfere not only with HR induction but also with CC binding (Figure 3.6). Thus the portion of the Rx NB domain housing helix α13h appears to more broadly impact Rx conformation, while helices α11h and α12h are thus-far observed to affect only signaling. Notably, this signaling-relevant region of Rx corresponds to the regions of Apaf-1 and CED-4 comprised of helices α11 and α12, which mediate the oligomerization necessary for apoptosome formation (Figure 3.8) (Qi et al. 2010).

Oligomerization has long been suspected among plant NB-LRR proteins, given the resemblance of their NB-ARC regions to those of oligomerizing STAND proteins such as Apaf-1 and CED-4. To date, however, only three cases of NB-LRR protein

oligomerization have been reported, and all of these interactions appear to be mediated by N-terminal domains rather than NB-ARC regions. Prf has been reported to exist in multimeric complexes, formation of which is mediated by its novel extended N-terminal domain (Gutierrez et al. 2010). CC domains of the NB-LRR proteins encoded by barley (*Hordeum vulgare*) *MLA* alleles have been found to dimerize and to thereby facilitate oligomerization of the full-length protein, irrespective of MLA activation state. This ability to dimerize, however, correlates with functional MLA resistance signaling capacity (Maekawa et al. 2011). Finally, upon elicitation by its cognate Avr protein, N is also observed to oligomerize. Interestingly, although the TIR domain of N is independently capable of homotypic oligomerization, oligomerization of the full-length molecule is dependent on a functional P-loop (Mestre and Baulcombe 2006), possibly indicating that recognition-dependent oligomerization requires the protein to first achieve a signaling-competent state. Here, we have presented the first report of interaction between NB domains of NB-LRR proteins (Figures 3.9, 3.10, and 3.11). While the Rx NB domain is able to interact with the isolated NB domains of various other NB-LRR proteins, the same phenomenon is not apparent among full-length proteins. There is no observable oligomerization between unelicited full-length Rx molecules (Moffett et al. 2002), nor does the isolated Rx NB domain interact with full-length Rx (Figure 3.9). It should be noted, however, that the rapid onset of HR following Rx elicitation precludes any such studies involving active Rx.

While the coincidence in location between regions likely involved with Rx signal initiation and those mediating Apaf-1 and CED-4 oligomerization is striking, we have found no clear functional correlation between Rx signaling and oligomerization activities. Rather, we were surprised to find that active variants of the Rx NB domain

seemed to engage in both homotypic and heterotypic interactions less robustly than inactive variants (Figures 3.9 and 3.11). This might simply be due to the onset of HR, or may alternately suggest that some pre-existing oligomeric interaction is in fact abolished upon NB-LRR protein activation. It is also conceivable that the ability of isolated NB domains to oligomerize might simply be a residual function, bespeaking their common ancestry with other STAND proteins but having little or no relevance to NB-LRR protein function. Future experiments that more precisely target the likely oligomerization interface will hopefully shed more light on the role of NB-mediated oligomerization in NB-LRR protein function.

The fact that HR induced by the isolated Rx NB domain mirrors that of the full-length protein in its requirement for SGT1 (Figure 3.2) can be taken as confirmation that NB-mediated HR represents a true resistance signal. This finding is also somewhat perplexing, however, as the role of SGT1 in NB-LRR protein function is most likely that of a chaperone, providing stability or proper folding through interaction with LRR domains (Bieri et al. 2004; Leister et al. 2005; Kadota et al. 2010). If Rx likely requires SGT1 to stabilize its LRR domain, why then would the isolated NB domain be sensitive to a reduction in SGT1 levels? One possible explanation could be the existence of additional NB-LRR proteins downstream from Rx signaling, themselves also relying on SGT1 for stabilization. Precedent for such a relationship is set by the TIR-NB-LRR protein N, which requires the presence of the CC_R-NB-LRR protein NRG1 in order to function. Defense response induction by both N and NRG1 is ultimately SGT1-dependent (Peart et al. 2005). It is thus tempting to speculate that the phenomenon of one NB-LRR protein having a functional reliance on another similar protein might not be exclusive to the N – NRG1 system, but might rather be a common occurrence among NB-LRR proteins. This scenario is of particular interest

given the ability of Rx and other NB domains to oligomerize with those of ADR1 and NRG1, both of which possess signaling competent N-terminal domains. The unique properties of CC_R-NB-LRR proteins, and their possible implications for the function of canonical NB-LRR proteins, will be discussed in Chapter 4.

MATERIALS AND METHODS

Cloning and plasmid construction

The pBIN61-based vectors expressing HA and six tandem c-Myc (Myc) epitope-tagged proteins Rx CC:HA, Rx CC-NB:HA, Rx CC-NB:EGFP:HA, Rx CC-NB:EGFP(PL):HA, Rx NB:EGFP:HA, Rx NB:EGFP(PL):HA, Rx NB:HA, Rx NB-ARC-LRR:Myc, and FL Rx:Myc have been previously described (Moffett et al. 2002; Rairdan and Moffett 2006; Rairdan et al. 2008). All additional constructs were created by directional cloning of the indicated fragment into the XbaI and BamHI sites of either pBIN61-EGPF:HA, pBIN61-HA, or pBIN61-Myc. All site-directed mutations and NAAIRS substitutions described in this study were generated by extension-overlap PCR (Vallejo et al. 2003). Nucleotide sequences of all primers used in this study are listed in Table 3.2. The Rx Δ190 NB:EGFP truncation arose spontaneously during cloning, and hence has no corresponding forward primer. PCR was performed with KOD high-fidelity polymerase (Novagen) and/or the Expand high fidelity PCR system (Roche). All new constructs were sequence verified.

Table 3.2 Primer sequences

Description	Nucleotide Sequence (5′ – 3′)
Cloning	
XbaI-Rx(2),Gpa2-NB F	CGTATCTAGAACCACCATGGAGAATATAATGGTTGG
BamHI-Rx(2),Gpa2-NB R	ATGCGGATCCCATGAGGCGCATGTGATGAG
XbaI-Bs2-NB F	CGTATCTAGAACCACCATGAAGAACAATATGGTTGGACG
BamHI-Bs2-NB R	CGTAGGATCCTTGATCCATGAAGCTCATCCGC
XbaI-Prf-NB F	ATGCTCTAGAACCACCATGAACGAAGAAATGGAGGGTTTTC
BamHI-Prf-NB R	ATGCGGATCCGAATAAACGAAGATGATGGGG
XbaI-RPS2-NB F	CGTATCTAGAACCACCATGATCAAGTCCGTTGTCGG
BamHI-RPS2-NB R	CGTAGGATCCCTTCTCCAGAAACTCCACTCTCAAC
AvrII-RPS5-NB F	ATCCCCTAGGACCACCATGTTTCAACCCACAATTGTTGGTC
BglII-RPS5-NB R	ATGCAGATCTTTTGAGACAGCTAACTTCCATCGG
XbaI-HRT-NB F	ATGCTCTAGAACCACCATGAGCGATCTTGTCGGGGTGG
BamHI-HRT-NB R	ATGCGGATCCAAGGATGCTTGCTCTAAAAGTTAAAC
XbaI-N-NB F	ATGCTCTAGAACCACCATGCAAAACATTGTTGGAATAGATACTC
BamHI-N-NB R	ATGCGGATCCGTAGTGCAGTCACCTCATATATTATATC
XbaI-RPS4-NB F	CGTATCTAGAACCACCATGGTCGTGGGTGCTTTAGGT
BamHI-RPS4-NB R	CGTAGGATCCAGTATCATCAACCAAACC
XbaI-RPP1-NB F	CGTATCTAGAACCACCATGTTAGTAGGGATGAGAGCT
BamHI-RPP1-NB R	CGTAGGATCCAACATGATTGATCCCATG
XbaI-Bs4-NB F	CGTATCTAGAACCACCATGATTGTGGGAATAAATGCT
BamHI-Bs4-NB R	CGTAGGATCCTATCGCATCATCTTTCTC
XbaI-ADR1-NB F	CGTATCTAGAACCACCATGGGAGGAGGGTGTTTAGGG
BamHI-ADR1-NB R	CGTAGGATCCTAACAACTCTAAATCATA
XbaI-NRG1-NB F	CGTATCTAGAACCACCATGTTTGGTAGCACAAATGGG
BamHI-NRG1-NB R	CGTAGGATCCCAAAAGATTTAATTTATA
XbaI-Rx-Δ159 F	CGTATCTAGAACCACCATGGGAGGAAGGGAACTAGAAG
XbaI-Rx-Δ170 F	CGTATCTAGAACCACCATGGGGATGGGAGGCATCGG
Site-directed mutagenesis of Rx	
RNBS-A1 NAAIRS F	AACGCAGCAATCCGGAGTGCAACTGTTTCACAAGAGTATTGTG
RNBS-A1 NAAIRS R	ACTCCGGATTGCTGCGTTTCGAGACATAATGCACGGATCACTATAG
RNBS-A2 NAAIRS F	AACGCAGCAATCCGGAGTCAAGAGTATTGTGTGAGAAATGTAC
RNBS-A2 NAAIRS R	ACTCCGGATTGCTGCGTTACGAATATCAAATCGAGACATAATGC
RNBS-A3 NAAIRS F	AACGCAGCAATCCGGAGTTGTGTGAGAAATGTACTCCTAGGCC
RNBS-A3 NAAIRS R	ACTCCGGATTGCTGCGTTTGCTTTTGCACGAATATCAAATCG
LXXLL NAAIRS F	AACGCAGCAATCCGGAGTTTGACAAGTGATGAACCTGATG
LXXLL NAAIRS R	ACTCCGGATTGCTGCGTTTACATTTCTCACACAATACTCTTGTG
RNBS-B1 NAAIRS F	AACGCAGCAATCCGGAGTACTACTCGGAATGTGGAAGTGGCTG
RNBS-B1 NAAIRS R	ACTCCGGATTGCTGCGTTATTATAACAGTCTGGGAAACATAG
RNBS-B2 NAAIRS F	AACGCAGCAATCCGGAGTGTGGAAGTGGCTGAATATGCTAGTTC
RNBS-B2 NAAIRS R	ACTCCGGATTGCTGCGTTTATTCTGCTTCCATTATAACAGTCTGG
KY182,184AA F	CTTTGGCTACAGCACTCGCTAGTGATCCG
KY182,184AA R	CGGATCACTAGCGAGTGCTGTAGCCAAAG

Table 3.2 (Continued)

F193A F	CCGTGCATTATGTCTCGAGCTGATATTCGTGCAAAAGC
F193A R	GCTTTTGCACGAATATCAGCTCGAGACATAATGCACGG
VL210,214AA F	GTGTGAGAAATGCACTCCTAGGCGCTCTTTCTTTGAC
VL210,214AA R	GTCAAAGAAAGAGCGCCTAGGAGTGCATTTCTCACAC
LL226,230AA F	GATCAGGCAGCGGACCGAGCGCAAAAGCATCTGAAAGGC
LL226,230AA R	GCCTTTCAGATGCTTTTGCGCTCGGTCCGCTGCCTGATC
DR228,229AA F	CAGCTAGCGGCCGCACTGCAAAAGCATC
DR228,229AA R	GATGCTTTTGCAGTGCGGCCGCTAGCTG
WD252,253AA F	CTACAGAAGCTGCGGCTGATATAAAAC
WD252,253AA R	GTTTTATATCAGCCGCAGCTTCTGTAG
FP258,259AA F	ATAAAACTATGTGCCGCAGACTGTTATAATGG
FP258,259AA R	CCATTATAACAGTCTGCGGCACATAGTTTTAT
TT271,272AA F	AGAATACTCCTGGCTGCTCGGAATGTGGAAGTGGC
TT271,272AA R	CCACATTCCGAGCAGCCAGGAGTATTCTGCTTCC
R273A F	CTCCTGACTACTGCGAATGTGGAAGTGGCTGAATATGC
R273A R	CCACTTCCACATTCGCAGTAGTCAGGAGTATTCTGC
TTR271,272,273AAA F	CTCCTGGCTGCTGCGAATGTGGAAGTGGCTGAATATGC
TTR271,272,273AAA R	ACATTCGCAGCAGCCAGGAGTATTCTGCTTCCATTATAACAG

Transient expression and protein analysis

Binary vectors were transformed into *Agrobacterium tumefaciens* strain C58C1 carrying the virulence plasmid pCH32. *Agrobacterium*-mediated transient expression (agroexpression) was performed as previously described (Bendahmane et al. 2000). Briefly, *Agrobacterium* suspensions in 10mM $MgCl_2$ were syringe infiltrated (agroinfiltration) into *N. tabacum* or *N. benthamiana* leaves at $OD_{600} = 0.2$. All proteins were expressed under the 35S promoter of the pBIN61 vector (Bendahmane et al. 2002). HR phenotypes generally presented one to two days after agroinfiltration and were photographed two to three days post-agroinfiltration. Total protein extracts were collected by grinding leaf disks in 8M urea, followed by the addition of SDS-PAGE loading buffer and subsequent boiling. Protein extraction for

coimmunoprecipitation, as well as immunoprecipitation and immunoblotting, was performed as previously described (Rairdan and Moffett 2006).

Virus-induced gene silencing

Three leaves each of three-week-old *N. benthamiana* plants, either wild-type or transgenic for *Rx* (Lu et al. 2003), were infiltrated with *Agrobacterium tumefaciens* carrying plasmids pBINTra6 and a pTV vector (Ratcliff et al. 2001) containing fragments of *N. benthamiana EDS1* or *SGT1* (Peart et al 02 BOTH), the *GUS* gene (Tameling and Baulcombe 2007), or no insert. Plants were used for protein expression experiments three to four weeks later.

Sequence alignment and structural modeling

Amino acid alignments were performed by ClustalW with Lasergene 8.0 MegAlign software (DNASTAR, Inc.). The alignment in Figure 3.8A was edited manually to resemble that of Qi et al. (2010) as closely as possible to allow comparison of secondary structure. In order to verify an accurate representation of tertiary structure, modeling of the Rx NB-ARC region was performed by two different methods, each utilizing automated prediction servers. For Figure 3.7, HHpred (Söding 2005) was used to identify and create alignments with optimal templates CED-4 (PDB code 2a5y_B) and Apaf-1 (PDB code 1z6t_A), from which a three-dimensional model was constructed using MODELLER (Sali et al. 1995). For Figure 3.8, template identification and modeling was performed by I-TASSER (Roy et al. 2010), which identified and utilized CED-4, Apaf-1, and the Apaf-1-containing apoptosome-procaspase-9 CARD complex (PDB code 3iyt_A) as top templates. RMSD values were calculated using Swiss-PdbViewer (Guex and Peitsch 1997).

CHAPTER 4

Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein*

ABSTRACT

Plant genomes encode large numbers of nucleotide binding, leucine-rich repeat (NB-LRR) proteins, many of which are active in pathogen detection and defense response induction. NB-LRR proteins fall into two broad classes: those with a TIR (Toll and Interleukin-1 receptor) domain at their N-terminus, and those with a coiled-coil (CC) domain at the N-terminus. Within CC-NB-LRR-encoding genes, one basal clade is distinguished by having CC domains resembling the *Arabidopsis thaliana* RPW8 protein, which we refer to as CC_R domains. Here, we show that CC_R-NB-LRR-encoding genes are present in the genomes of all higher plants surveyed, and that they comprise two distinct subgroups – one typified by the *Nicotiana benthamiana* NRG1 protein, and the other typified by the Arabidopsis ADR1 protein. We further report that, in contrast to CC-NB-LRR proteins, the CC_R domains of both NRG1- and ADR1-like proteins are sufficient for the induction of defense responses, and that this activity appears to be SGT1-independent. Additionally, we report the apparent absence of

* Originally published as: Collier, S.M., Hamel, L-P., and Moffett, P. (2011) Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Molecular Plant-Pathogen Interactions* (In press, DOI: 10.1094/MPMI-03-11-0050). All data presented in this chapter was generated by S.M.C. with the exception of Figures 4.1 and 4.9B, which were generated by L-P.H.

both *NRG1* homologs and TIR-NB-LRR-encoding genes from the dicot *Aquilegia coerulea* and the dicotyledonous order Lamiales as well as from monocotyledonous species. This strong correlation in occurrence is suggestive of a functional relationship between these two classes of NB-LRR proteins.

INTRODUCTION

To defend themselves against pathogens, plants possess a sophisticated multi-level immune system. At a basal level, plant pattern recognition receptors are able to detect conserved pathogen-associated molecular patterns (PAMPs) and initiate a low-impact resistance responses known as PAMP-triggered immunity (PTI), which involves MAP kinase signaling, callose deposition, and production of reactive oxygen species (Chisholm et al. 2006). Should a pathogen overcome or circumvent this line of defense, it may yet be intercepted by a second resistance mechanism termed effector-triggered immunity (ETI) or gene-for-gene resistance (Flor 1971), in which the protein products of plant resistance (*R*) genes specifically recognize cognate pathogen avirulence (*Avr*) gene products and trigger a stronger resistance response. The majority of *R* genes encode NB-LRR proteins, which are defined by the presence of a central nucleotide binding (NB) domain and a C-terminal leucine-rich repeat (LRR) domain. Located between the NB and LRR lies a region known as the ARC (Apaf-1, R proteins, and CED4 homology) domain, which together with the NB constitutes a nucleotide binding pocket bearing homology to members of the STAND (signal transduction ATPases with numerous domains) class of ATPases (Tameling et al. 2002; Leipe et al. 2004). While both the NB and ARC domains possess numerous conserved,

functionally critical motifs, the LRR is more variable – an attribute which facilitates its role in determining recognition specificity (Collier and Moffett 2009).

Over 70 NB-LRR proteins of known specificity have been cloned (Sacco and Moffett 2009), and thousands more NB-LRR-encoding sequences are listed in GenBank (available online from the National Center for Biotechnology Information). Analysis of conserved NB-ARC sequence reveals two major classes within NB-LRR proteins (Meyers et al. 1999), a division also supported by their N-terminal structures. Most members of the first class possess a domain homologous to Toll and Interleukin-1 receptors (TIR) at their N-terminus, whereas much greater variability exists among N-terminal domains of the second class. Because many of these variable N-terminal regions are predicted to form coiled coil (CC) structures, members of this class are collectively referred to as CC-NB-LRR proteins. However, the CC domain is sometimes joined and/or replaced by additional structures, such as predicted BED DNA-binding and Solanaceae domains (Collier and Moffett 2009). Despite an overall lack of sequence similarity, essentially all characterized canonical CC domains – whether joined by additional domains or not – possess a small "EDVID" motif (Rairdan et al. 2008), and are thus described as belonging to the CC_{EDVID} subclass (Collier and Moffett 2009). A less abundant subclass of CC domain also exists, previously termed CC_{RPW8} (Collier and Moffett 2009), in which no EDVID motif is present and which bears closest sequence similarity to RPW8 – a non-NB-LRR R protein from *Arabidopsis thaliana* which confers broad-spectrum resistance against powdery mildew (*Erysiphe* spp.) (Xiao et al. 2001). Interestingly, NB-LRR proteins that possess an RPW8-like CC domain (hereafter referred to as CC_R) are also found to form a distinct subclass within the CC-NB-LRR proteins based on their NB-ARC sequence. This CC_R-NB-LRR subclass has variously been referred to as non-TIR N4 and CNL-A,

and appears to be one of the most ancestral of major CC-NB-LRR clades (Cannon et al. 2002; Meyers et al. 2003; McHale et al. 2006). Few CC_R-NB-LRR proteins have been cloned or functionally characterized to date although they have been reported to possess characteristic motifs that distinguish them from other NB-LRR proteins (Chini and Loake 2005). The best-studied members of this group include N-required gene 1 (NRG1) of *Nicotiana benthamiana* and activated disease resistance gene 1 (ADR1) of Arabidopsis. NbNRG1 was identified through a gene silencing screen as being required for the function of the tobacco (*N. glutinosa*) TIR-NB-LRR protein N, apparently acting downstream of effector recognition by N (Peart et al. 2005). ADR1 was identified in an activation-tagging screen, where it was observed that ADR1 over-expression leads to constitutive defense activation as well as drought tolerance (Grant et al. 2003; Chini et al. 2004). It is interesting to note that neither NRG1 nor ADR1 have been reported to act in the classical gene-for-gene manner of canonical R proteins, perhaps suggesting that CC_R-NB-LRR proteins function somewhat differently than the more common CC_{EDVID}-NB-LRR proteins.

While the expansion of NB-LRR-encoding genes has been remarkably prolific, with most plant genomes containing several hundred family members (Sacco and Moffett 2009), they also exhibit great plasticity. Orthologous relationships can be difficult to determine due to the gene duplication, loss, recombination, and/or rapid diversification frequently observed within many NB-LRR lineages, and the relative abundance of various lineages may also vary greatly between different species (McDowell and Simon 2006; McHale et al. 2006; Velasco et al. 2010). In an extreme example, TIR-NB-LRR-encoding genes are completely absent from all monocot genomes thus far queried (Meyers et al. 1999; Pan et al. 2000; Bai et al. 2002; Tarr and Alexander 2009). In contrast, genes encoding CC_R-NB-LRR proteins constitute a

relatively conservative clade: they have been identified in numerous monocot and dicot species, as well as in gymnosperms, and there is currently no indication that this gene family has expanded disproportionately in any plant lineage (this report). Certain CC_R-NB-LRR-encoding genes, including certain Arabidopsis homologs, have been found to display particularly low levels of sequence variation both within and between species (Grant et al. 2003; Bakker et al. 2006; Jermstad et al. 2006). Furthermore, in sharp contrast to typical NB-LRR proteins, the LRR domains of ADR1 and its three paralogs *ADR1-L1*, *-L2*, and *-L3*, display higher levels of sequence conservation than do their corresponding NB-ARC domains (Chini and Loake 2005). Taken together, these observations would seem to suggest that CC_R-NB-LRR proteins might play a role that is both conserved and unique among NB-LRR proteins.

NB-LRR proteins are collectively able to recognize a broad array of Avr proteins, and while the pathogens recognized are diverse, the response initiated upon NB-LRR activation is remarkably homogenous. NB-LRR-mediated resistance shares many components of basal resistance (Jones and Dangl 2006), but it is often also accompanied by programmed cell death at the site of infection, termed the hypersensitive response (HR). Additionally, the efficacy of NB-LRR-mediated resistance is not restricted to the class of pathogen initially recognized, but can extend across kingdoms (Tobias et al. 1999; Sohn et al. 2007; Rentel et al. 2008). While specific pathogen detection is thought to result from the combined action of the LRR and N-terminal domains, the origin of NB-LRR signal initiation remains unclear. Several TIR domains, including those of RPS4, RPP1, and L10 (Frost et al. 2004; Swiderski et al. 2009; Krasileva et al. 2010) have been shown to be sufficient for inducing defense responses upon over-expression but no TIR-interacting signaling proteins have been identified. To date a similar activity has not been demonstrated for any isolated CC

domain. Rather, the NB domain of the potato (*Solanum tuberosum*) Rx CC-NB-LRR protein was found to be sufficient to induce defense responses (Rairdan et al. 2008). Whether this behavior extends to other CC-NB-LRR proteins remains to be determined.

On investigating the potential for signaling capacity of the CC domains of various characterized solanaceous NB-LRR proteins, we find that only the CC_R domain of NRG1 is capable of independently inducing defense responses. We also demonstrate that CC_R-mediated resistance signaling is common to both NRG1- and ADR1-like proteins of solanaceous species and Arabidopsis, and that this function does not appear to depend on SGT1. We further demonstrate that NRG1- and ADR1-like proteins are involved in the resistance response mediated by a canonical CC-NB-LRR protein. Additionally, we report that NRG1 and ADR1 represent two distinct sub-clades of CC_R-NB-LRR proteins, and that these lineages show a degree of conservation and retention through speciation not seen with genes encoding other classes of NB-LRR proteins. Curiously, we note that while *ADR1*-like genes are present in every higher plant genome investigated, the occurrence of NRG1 family members mirrors that of TIR-NB-LRR-encoding genes, suggesting an evolutionary and functional relationship between these proteins.

RESULTS

NRG1 CC_R-mediated HR induction

To better understand whether the lack of demonstrated CC-mediated HR is indicative of contrasting modes of signaling by TIR-NB-LRR vs. CC-NB-LRR proteins, we examined the CC domains of a number of solanaceous CC-NB-LRR proteins,

including NRG1. We transiently expressed, by agroinfiltration, seventeen CC domains derived from solanaceous NB-LRR proteins, fourteen of which had not previously been so tested. A representative experiment, including the CC domains of I2, R3a, RB, Bs2 and Rx, is shown in Figure 4.1. None of these CC domains induced an HR in tobacco (*N. tabacum*) leaves, in contrast to the N-terminus of NRG1, which induced a strong HR, visible after 24 hours (Figure 4.1A).

Because the N-terminal region of CC_R-NB-LRR proteins cannot be aligned with those of CC_{EDVID}-NB-LRR sequences, we initially utilized an NRG1 clone encompassing the first 225 amino acids, terminating just before the P-loop motif of the NB domain. While ensuring inclusion of the entire CC_R domain, this clone also contained the extreme N-terminal portion of the NB domain. To better delineate the region involved in signal initiation, we performed a deletion analysis of the NRG1 N-terminus (Figure 4.1B). We constructed a clone terminating just before the NB domain (1-182), as well as N- and C-terminal truncations within the CC_R domain. Among these was a deletion of amino acids 1-13, which in RPW8 are thought to be part of a transmembrane and/or targeting motif (Xiao et al 2001). While removal of NB remnants had no affect on HR induction, all deletions within the CC_R domain failed to produce HR upon transient expression (Figure 4.1C). In the cases of NRG (1-147) and (13-147) lack of activity could be due to compromised stability, as protein accumulation for these two variants was somewhat less than that of active versions. NRG1 (13-225), however, showed high levels of protein accumulation, and yet was unable to induce HR (Figure 4.1C). We therefore conclude that the entire CC_R domain is necessary and sufficient to induce an HR, and that amino acids 1-13 are likely critical for this function.

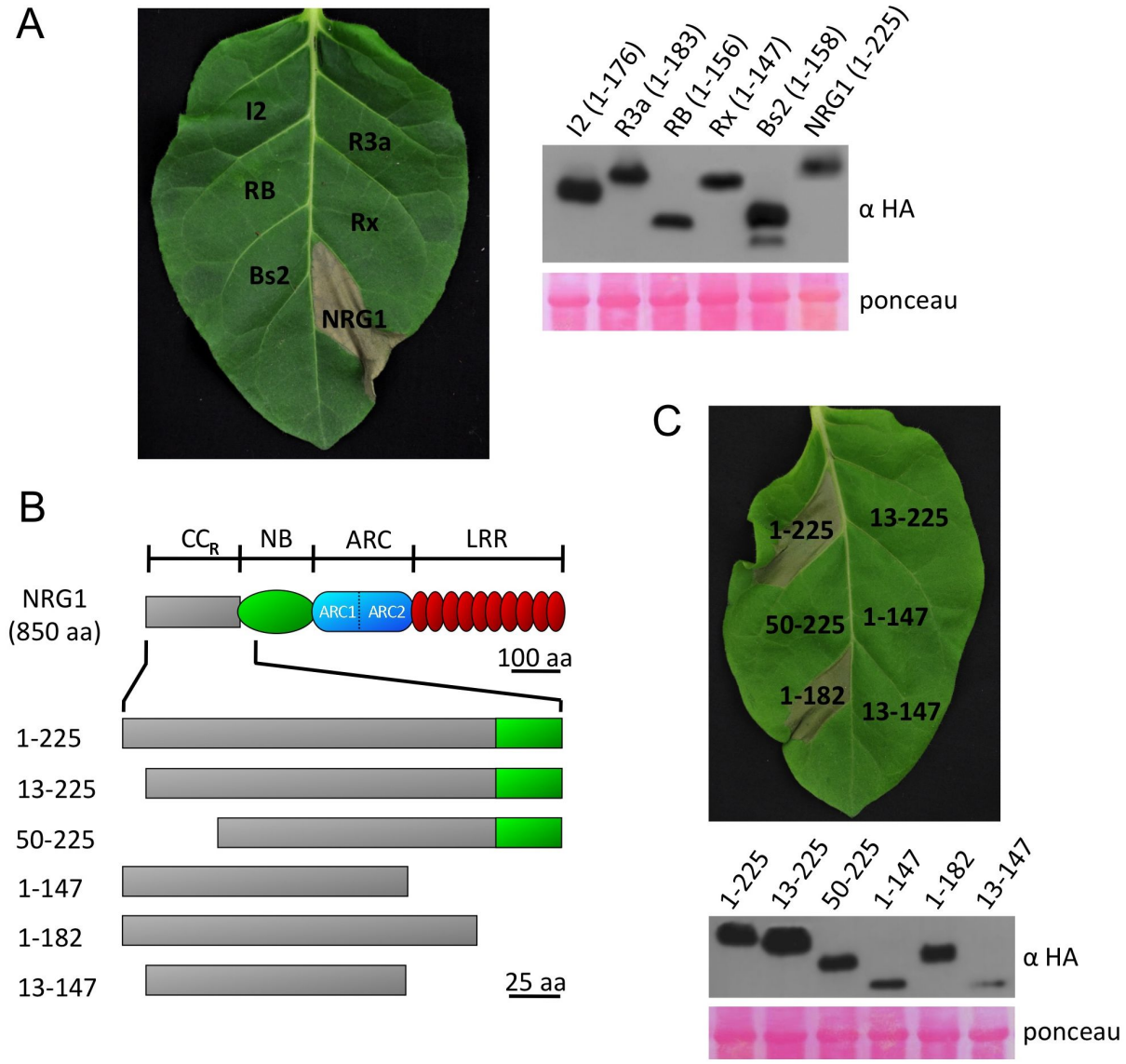


Figure 4.1. *NRG1 CC$_R$-mediated HR induction.* **A,** Transient expression of representative CC domains in *N. tabacum*. Leaves were agroinfiltrated with CC domains of I2, R3a, RB, Rx, Bs2, and NRG1 as indicated and photographed three days later. Proteins extracts from agroinfiltrated patches were subjected to anti-HA immunoblotting after 48 hours (24 hours for NRG1) (right panel). **B,** Schematic of NRG1 N-terminus truncations. **C,** HR induction after agroinfiltration of indicated NRG1 derivatives (top panel) and anti-HA immunoblotting of protein extracts prepared 24 hours after agroinfiltration (bottom panel).

NRG1 and ADR1 describe distinct CC_R-NB-LRR sub-clades

ADR1 is among Arabidopsis genes showing highest sequence similarity to NRG1. We therefore undertook a more extensive phylogenetic analysis of CC_R-NB-LRR-encoding genes to determine the distribution and complexity of this lineage. This was accomplished by a thorough search of fully-sequenced genomes, aiming not to identify every CC_R-NB-LRR-encoding sequence or sequence fragment currently available, but rather with the specific purpose of ascertaining the distribution patterns of CC_R-NB-LRR-encoding genes through examination of this limited but largely complete data set. CC_R-NB-LRR-encoding genes were identified by protein-protein BLAST (BLASTP) against the predicted proteomes of all 22 plants with sequenced and annotated genomes contained in the Phytozome database (available online from the DOE Joint Genome Institute and UC Berkley Center for Integrative Genomics). AtADR1 was used as a query, and sequences from each proteome were compiled down through the first non-CC_R-NB-LRR hit returned (determined by bit-score cutoff and subsequent phylogenetic confirmation), thus ensuring a reasonably complete representation of each genome's *CC_R-NB-LRR* complement (Table 4.1). While no *CC_R-NB-LRR* sequences were readily identifiable in *Volvox, Chlamydomonas*, *Physcomitrella*, or *Selaginella*, every higher plant surveyed contained at least one CC_R-NB-LRR-encoding gene.

Full-length CC_R-NB-LRR amino acid sequences were aligned by ClustalW. Phylogeny construction based on the NB-ARC region revealed two distinct, strongly supported clades, one containing NRG1 and the other containing ADR1 (Figure 4.2). Similar groupings were obtained from trees based on the LRR and CC_R domains, as well as from full-length sequences (Figures 4.3 and 4.8). Because some

Table 4.1 Results of ADR1 BLAST search of Phytozome-curated proteomes

Species	Defline	Bit Score	E-value	%Identity	Group[a]	Coverage[b]	Used?[c]
Aquilegia coerulea	AcoGoldSmith_v1.000926m	688.3	0	45.6	A	CNL	Y
Aquilegia coerulea	AcoGoldSmith_v1.014723m	90.5	2.8e-18	27.8	O		
Arabidopsis lyrata	473526	1478.8	0	91.1	A	CNL	
Arabidopsis lyrata	487251	1002.7	0	62.7	A	CNL	
Arabidopsis lyrata	944602	939.1	0	59.3	A	CNL	
Arabidopsis lyrata	496910	317.8	1.1e-86	31.3	N	CNL	
Arabidopsis lyrata	951540	251.9	7.3e-67	29.3	N	(C)NL	
Arabidopsis lyrata	918482	207.6	2e-53	33.7	N	(N)L	
Arabidopsis lyrata	488633	113.2	4.4e-25	26.2	O		
Arabidopsis thaliana	ADR1	1608.6	0	100	A	CNL	Y
Arabidopsis thaliana	ADR1-L2	1003.0	0	62.7	A	CNL	Y
Arabidopsis thaliana	ADR1-L1	936.8	0	58.9	A	CNL	Y
Arabidopsis thaliana	ADR1-L3	798.5	0	64.0	A	NL	
Arabidopsis thaliana	At5g66910 (NRG1.2)	336.3	4.2e-92	31.8	N	CNL	Y
Arabidopsis thaliana	At5g66900 (NRG1.1)	311.6	1.1e-84	31.3	N	CNL	Y
Arabidopsis thaliana	At5g66890	212.2	9.7e-55	33.6	N	(N)L	
Arabidopsis thaliana	TAO1	114.8	2.0e-25	24.4	O		
Brachypodium distachyon	Bradi4g03230	599.0	3.8e-171	41.1	A	CNL	Y
Brachypodium distachyon	Bradi4g28190	91.3	2.6e-18	22.0	O		
Carica papaya	ems64.11	281.2	7.7e-76	42.7	A	C(N)	(Y)
Carica papaya	ems6.400	243.8	1.3e-64	36.0	N	(N)L	(Y)
Carica papaya	ems6.402	154.5	1.2e-37	37.2	N	C(N)(L)	(Y)
Carica papaya	ems64.12	135.6	6.3e-32	48.9	A	(L)	(Y)
Carica papaya	ems234.9	134.0	1.5e-31	25.5	O		
Cucumis sativus	Cucsa.123410	690.3	0	46.1	A	CNL	Y
Cucumis sativus	Cucsa.133510	290.8	1.8e-78	28.2	N	CNL	Y
Cucumis sativus	Cucsa.102240	287.3	2.0e-77	31.5	N	CNL	
Cucumis sativus	Cucsa.189390	88.6	1.4e-17	23.4	O		

Table 4.1 (Continued)

Glycine max	Glyma17g36400	666.8	0	45.5	A	CNL	Y
Glycine max	Glyma14g08710	661.0	0	45.8	A	CNL	
Glycine max	Glyma17g36420	660.2	0	44.5	A	CNL	
Glycine max	Glyma14g08700	651.7	0	43.7	A	CNL	
Glycine max	Glyma17g21130	337.0	3.4e-92	33.5	N	NL	
Glycine max	Glyma05g09440	326.6	5.7e-89	32.1	N	CNL	Y
Glycine max	Glyma05g17470	324.3	2.5e-88	31.1	N	NL	
Glycine max	Glyma17g21240	316.2	7.8e-86	33.3	N	CNL	
Glycine max	Glyma17g21200	311.2	2.1e-84	30.8	N	NL	
Glycine max	Glyma11g06260	308.1	2.1e-83	34.5	N	CNL	
Glycine max	Glyma01g39010	301.6	2e-81	33.7	N	CNL	
Glycine max	Glyma05g17460	297.4	3.9e-80	32.2	N	CNL	
Glycine max	Glyma01g39000	291.2	2.4e-78	31.9	N	CNL	
Glycine max	Glyma17g21470	284.6	2.0e-76	27.5	N	CNL	
Glycine max	Glyma05g09430	240.4	4.5e-63	29.0	N	(N)L	
Glycine max	Glyma17g20900	209.1	1.2e-53	30.7	N	(N)(L)	
Glycine max	Glyma11g06270	188.3	2.5e-47	28.9	N	(N)(L)	
Glycine max	Glyma08g16380	184.5	3.6e-46	28.8	N	(N*)L	
Glycine max	Glyma16g25080	121.7	2.5e-27	23.5	O		
Manihot esculenta	cassava12942	765.0	0	50.1	A	CNL	Y
Manihot esculenta	cassava11923	354.4	1.4e-97	33.7	N	CNL	Y
Manihot esculenta	cassava46524	333.2	3.5e-91	30.4	N	CNL	
Manihot esculenta	cassava4857	109.4	7.8e-24	25.5	O		
Medicago truncatula	Medtr1g025080	694.9	0	45.6	A	CNL	Y
Medicago truncatula	Medtr1g025090	625.9	2.1e-179	43.8	A	CNL	
Medicago truncatula	Medtr5g018040	347.4	1.5e-95	31.2	N	CNL	Y
Medicago truncatula	Medtr5g018210	335.5	5.6e-92	33.8	N	CNL	
Medicago truncatula	Medtr8g096790	321.2	1.1e-87	32.3	N	CNL	
Medicago truncatula	Medtr8g096780	320.1	2.5e-87	32.9	N	CNL	
Medicago truncatula	Medtr5g018100	319.3	3.9e-87	34.1	N	CNL	
Medicago truncatula	Medtr5g018050	315.8	3.7e-86	33.9	N	C(N)L	
Medicago truncatula	Medtr8g096810	307.4	1.4e-83	34.1	N	CNL	
Medicago truncatula	Medtr8g096700	305.8	4.3e-83	32.1	N	CNL	
Medicago truncatula	Medtr8g096590	395.8	4.9e-83	32.3	N	C*NL	

Table 4.1 (Continued)

Medicago truncatula	Medtr8g096820	301.2	1.0e-81	32.7	N	CNL	
Medicago truncatula	Medtr8g096690	293.5	2.1e-79	30.8	N	C*NL	
Medicago truncatula	Medtr8g096720	250.0	3.0e-66	30.7	N	C(N)L	
Medicago truncatula	Medtr8g096600	193.4	3.5e-49	38.1	N	L	
Medicago truncatula	Medtr8g096750	181.8	9.5e-46	39.8	N	L	
Medicago truncatula	Medtr8g096740	136.0	5.6e-32	30.4	N	CN	
Medicago truncatula	Medtr6g074000	104.0	2.6e-22	25.6	O		
Mimulus gattatus	mgf005353	590.9	7.5e-169	48.2	A	(C)NL	Y
Mimulus gattatus	mgf024588	93.2	5.5e-19	23.9	O		
Oryza sativa	Os12g39620.1	556.2	3.0e-158	40.0	A	CNL	Y
Oryza sativa	Os08g12740.2	117.1	5.1e-26	25.1	O		
Populus trichocarpa	POPTR_0002s13070	758.1	0	48.9	A	CNL	Y
Populus trichocarpa	POPTR_0014s03500	754.2	0	49.5	A	CNL	
Populus trichocarpa	POPTR_0007s11510	366.3	4.2e-101	32.1	N	CNL	Y
Populus trichocarpa	POPTR_0007s11500	365.2	1.0e-100	31.9	N	CNL	
Populus trichocarpa	POPTR_0007s11480	344.0	2.3e-94	29.5	N	C*NL	
Populus trichocarpa	POPTR_0007s11550	331.6	1.2e-90	32.8	N	C*NL	
Populus trichocarpa	POPTR_0007s11530	285.4	1.1e-76	39.9	N	(N)L	
Populus trichocarpa	POPTR_0005s03070	132.1	1.5e-30	26.3	O		
Prunus persica	ppa001461m	723.8	0	48.6	A	CNL	Y
Prunus persica	ppa001610m	373.6	1.9e-103	31.9	N	CNL	Y
Prunus persica	ppa016036m	364.0	1.4e-100	35.0	N	(C)NL	
Prunus persica	ppa023198m	358.6	7.3e-99	31.6	N	CNL	
Prunus persica	ppa1027190m	354.4	1.1e-97	31.1	N	CNL	
Prunus persica	ppa024835m	354.0	1.5e-97	32.2	N	CNL	
Prunus persica	ppa001497m	347.1	2.2e-95	30.6	N	CNL	
Prunus persica	ppa001530m	343.2	2.9e-94	30.1	N	CNL	
Prunus persica	ppa001498m	337.0	1.8e-92	31.2	N	CNL	
Prunus persica	ppa016232m	330.9	1.4e-90	30.5	N	CNL	
Prunus persica	ppa014998m	318.2	9.0e-87	33.8	N	CNL	
Prunus persica	ppa025461m	163.3	4.2e-40	38.8	N	L	
Prunus persica	ppa016391m	123.2	5.1e-28	26.1	O		
Ricinus communis	30170.m013933	710.3	0	48.8	A	CNL	Y

Table 4.1 (Continued)

Ricinus communis	29676.m001640	368.6	5.8e-102	31.4	N	CNL	
Ricinus communis	29841.m002829	273.9	1.6e-73	32.0	N	CNL	
Ricinus communis	29676.m001639	246.1	3.6e-65	27.3	N	CNL	Y
Ricinus communis	30190.m011052	108.6	8.9e-24	24.8	O		
Setaria italica	siPROV001834m	609.0	3.8e-174	49.2	A	CNL	
Setaria italica	siPROV027793m	122.1	1.3e-27	23.2	O		
Sorghum bicolor	Sb08g019690	614.4	7.1e-176	48.9	A	CNL	Y
Sorghum bicolor	Sb05g006170	116.7	4.2e-26	24.4	O		
Vitis vinifera	GSVIVT00014747001	748.4	0	48.2	A	CNL	Y
Vitis vinifera	GSVIVT00003901001	398.3	6.6e-111	32.6	N	CNL	Y
Vitis vinifera	GSVIVT00010062001	389.8	2.4e-108	32.6	N	CNL	
Vitis vinifera	GSVIVT00003147001	387.5	1.1e-107	31.7	N	CNL	
Vitis vinifera	GSVIVT00004974001	360.9	1.2e-99	34.0	N	(C)NL	
Vitis vinifera	GSVIVT00003149001	337.4	1.6e-92	34.6	N	(C)NL	
Vitis vinifera	GSVIVT00009711001	335.9	4.5e-92	33.4	N	(C)NL	
Vitis vinifera	GSVIVT00004841001	327.8	1.1e-89	37.4	N	(N)L	
Vitis vinifera	GSVIVT00003151001	320.9	1.3e-87	33.9	N	(C)N*L*	
Vitis vinifera	GSVIVT00008450001	272.7	5e-73	31.4	N	CNL*	
Vitis vinifera	GSVIVT00004838001	243.8	2.3e-64	39.5	N	(N)L	
Vitis vinifera	GSVIVT00004842001	217.2	2.4e-56	28.7	N	CN	
Vitis vinifera	GSVIVT00012615001	198.0	1.2e-50	29.7	N	CN	
Vitis vinifera	GSVIVT00001806001	123.6	3.7e-28	26.4	O		
Zea mays	GRMZM2G044724	616.7	2.0e-176	41.6	A	CNL	Y
Zea mays	GRMZM2G443525	604.4	1.0e-172	48.5	A	CNL	
Zea mays	GRMZM2G005452	106.3	8.3e-23	25.7	O		

[a] A = ADR1-like, N = NRG1-like, O = outgroup.

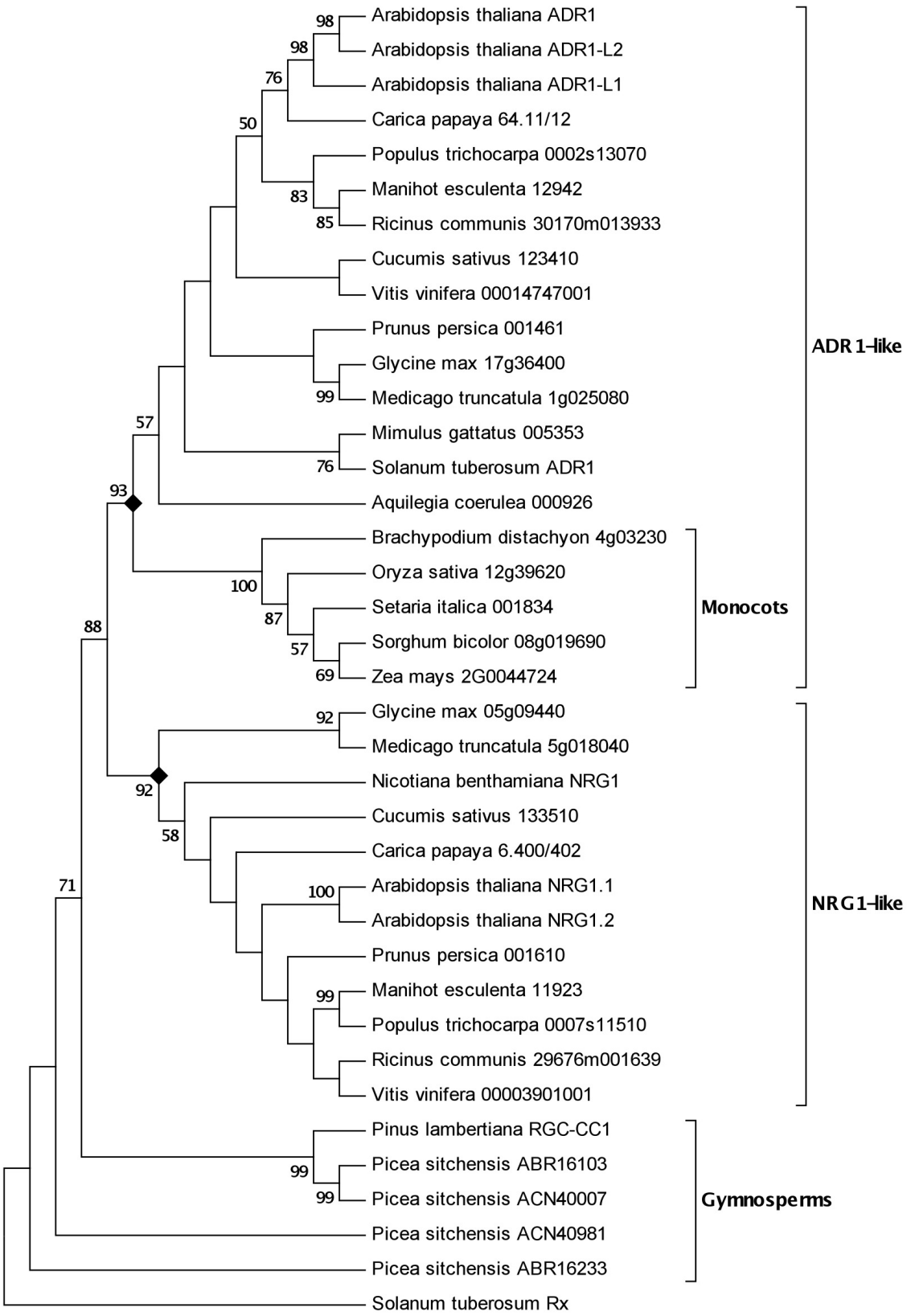
[b] Domain coverage of sequence obtained. C = CC_R, N = NB-ARC, L = LRR, () = partial sequence for domain, * = significant divergence from consensus. Coverage not shown for outgroup sequences.

[c] Y = included in phylogeny construction, (Y) = used only after manual re-annotation

species contained multiple homologs per clade, we limited our final analysis in Figure 4.2 to a single representative per clade per species, except in the case of gymnosperm sequences (see below). *A. lyrata* sequences were also excluded from analysis because of excessive redundancy with *A. thaliana*. One striking observation from this analysis is that, with few exceptions, every flowering plant queried encodes both ADR1-like and NRG1-like proteins (Figure 4.2, Table 4.1).

Although no fully sequenced gymnosperm genomes are currently available, we were interested in examining members of this lineage as well. We therefore performed a BLASTP search of the GenBank Coniferophyta protein collection with full-length AtADR1 as a query, and from this obtained five complete or nearly complete CC_R-NB-LRR sequences, shown in Figure 4.2. One of these (*P. lambertiana* RGC-CC1) has been previously studied and found to possess remarkably low levels of sequence variation between unrelated individuals (Jermstad et al. 2006). Although the five gymnosperm proteins all group as being most similar to one another, the limited number of sequences available were insufficient to determine whether conifers indeed possess just a single class of CC_R-NB-LRR protein, or whether, due to chance, sequences from only one of two clades had been recovered. In order to resolve this uncertainty, we also performed a search of conifer ESTs in the GenBank collection. To ensure representation of both ADR1-like and NRG1-like sequences, if present, we used the NB-ARC domains of AtADR1, NbNRG1, Rx, and N (the latter two as representative CC-NB-LRR and TIR-NB-LRR proteins, respectively) as queries, and compiled the top ten hits from each search. While searches with Rx and N each returned ten unique hits, hits from ADR1- and NRG1-based searches were somewhat overlapping. Phylogenetic analysis of the resulting sequences confirmed that while conifer CC_R-NB-LRR-encoding sequences are clearly distinct from TIR-NB-LRR and

Figure 4.2. *Phylogenetic analysis of CC_R-NB-LRR proteins.* Bootstrap consensus tree constructed using the Neighbor Joining (NJ) method from the aligned NB-ARC domains of NRG1-like and ADR1-like predicted amino acid sequences from plant species with fully sequenced genomes. Additional CC_R-NB-LRR proteins used elsewhere in this study as well as available gymnosperm sequences are also included. For species with multiple NRG1 or ADR1 homologs, a single full-length sequence was selected at random as a representative for each group. 1000 replicate trees were analyzed, and resultant bootstrap percentage values above 50 are shown at branch nodes. Putative proteins are labeled by genus, species, and (alpha)numerical identifier. Black diamonds demark strongly supported NRG1-like and ADR1-like clades (also indicated by brackets), and the location of monocot- and gymnosperm-derived sequences are indicated by brackets. The tree is rooted with Rx


Arabidopsis thaliana ADR1
Arabidopsis thaliana ADR1-L2
Arabidopsis thaliana ADR1-L1
Carica papaya 64.11/12
Populus trichocarpa 0002s13070
Manihot esculenta 12942
Ricinus communis 30170m013933
Cucumis sativus 123410
Vitis vinifera 00014747001
Prunus persica 001461
Glycine max 17g36400
Medicago truncatula 1g025080
Mimulus gattatus 005353
Solanum tuberosum ADR1
Aquilegia coerulea 000926
Brachypodium distachyon 4g03230
Oryza sativa 12g39620
Setaria italica 001834
Sorghum bicolor 08g019690
Zea mays 2G0044724
Glycine max 05g09440
Medicago truncatula 5g018040
Nicotiana benthamiana NRG1
Cucumis sativus 133510
Carica papaya 6.400/402
Arabidopsis thaliana NRG1.1
Arabidopsis thaliana NRG1.2
Prunus persica 001610
Manihot esculenta 11923
Populus trichocarpa 0007s11510
Ricinus communis 29676m001639
Vitis vinifera 00003901001
Pinus lambertiana RGC-CC1
Picea sitchensis ABR16103
Picea sitchensis ACN40007
Picea sitchensis ACN40981
Picea sitchensis ABR16233
Solanum tuberosum Rx
ADR1-like
Monocots
NRG1-like
Gymnosperms

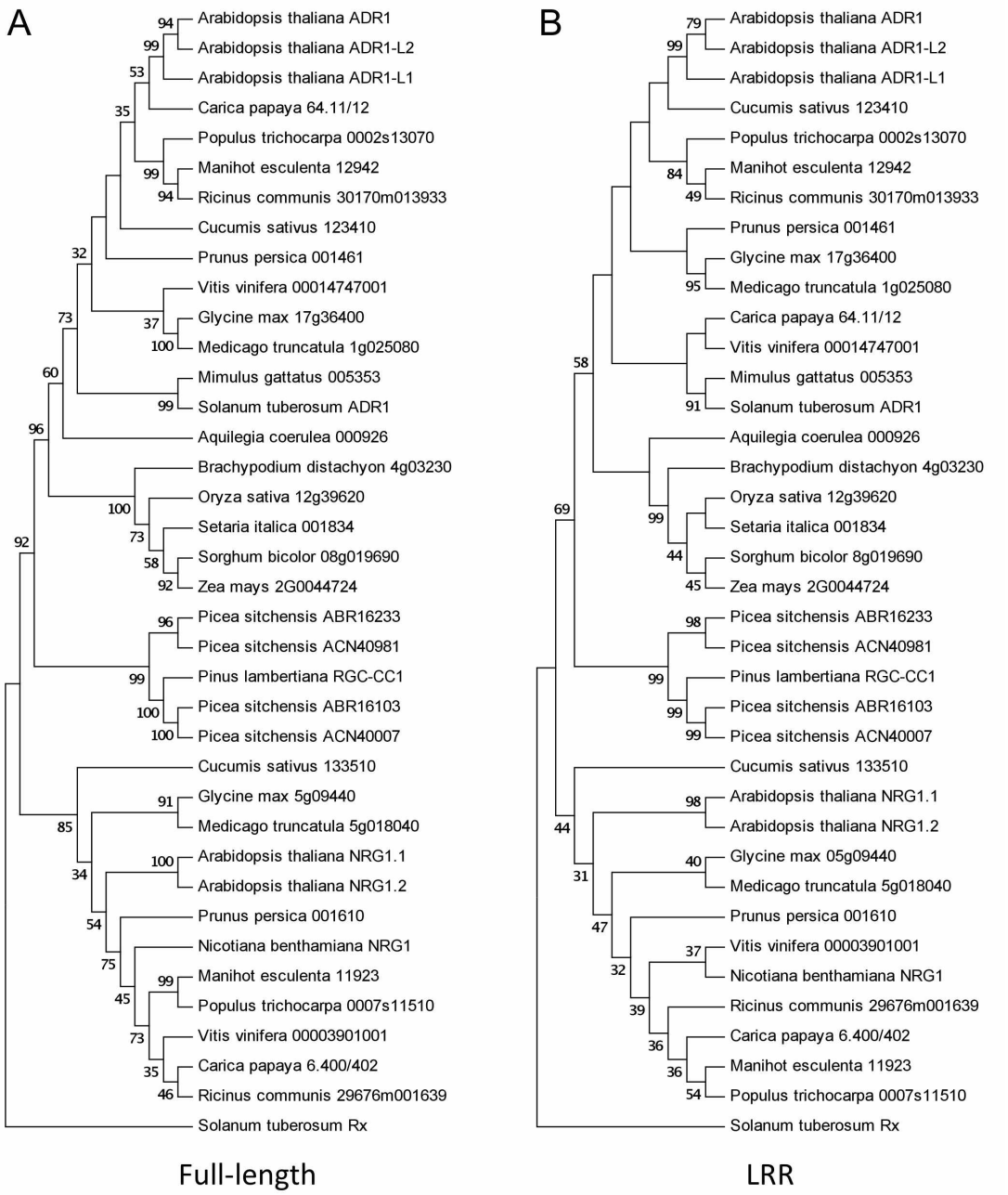


Figure 4.3. *Phylogenetic analysis of CC_R-NB-LRR proteins based on full-length and LRR sequences.* NJ trees of NRG1-like and ADR1-like predicted amino acids sequences from plant species with fully sequenced genomes. Also included are all CC_R-NB-LRR proteins used elsewhere in this study, and available gymnosperm sequences. For species with multiple *NRG1* or *ADR1* homologs, a single full-length sequence was selected as a representative for each group. 1000 replicate trees were analyzed, and resultant bootstrap percentage values above 30 are shown at branch nodes. **A,** Tree based on aligned full-length sequences and rooted with Rx. **B,** Tree based on aligned LRR domain sequences and rooted with Rx.

canonical CC-NB-LRR sequences, grouping most closely with angiosperm-derived CC_R-NB-LRR sequences, they form their own sub-clade without clear membership to either the ADR1 or NRG1 group (Figure 4.4). We can therefore be reasonably confident that conifers possess only a single class of CC_R-NB-LRR protein, and that the five full-length sequences obtained provide a fair representation. Although they are shown to occupy a basal position in Figure 4.2, gymnosperm-derived sequences can be seen alternately to group with ADR1-like sequences depending on the protein domain analyzed (Figure 4.3). It thus remains to be determined whether gymnosperm CC_R-NB-LRR-encoding genes bear more affinity to one group or whether they in fact originated before the divergence of ADR1 and NRG1 lineages.

Correlation of NRG1 and TIR-NB-LRR occurrence

While ADR1-like sequences were identified in the genomes of every angiosperm species investigated, there are certain exceptions concerning the NRG1 group. Most notable among these is the absence of NRG1-like sequences from the genomes of every monocot species examined (Figure 4.2, Table 4.1). Interestingly, this trend correlates with the absence of TIR-NB-LRR proteins among monocots (Meyers et al. 1999; Bai et al. 2002; McHale et al. 2006). We were also unable to identify NRG1 homologs in the genomes of the dicots monkey flower (*Mimulus guttatus*) and columbine (*Aquilegia coerulea*), both in our original BLASTP searches as well as in subsequent TBLASTN searches using NbNRG1 as a query (data not shown). Both to confirm this finding and to investigate whether NRG1 and TIR-NB-LRR occurrence might be correlated in these genomes as well, we undertook further investigation of these species' NB-LRR complements. BLASTP searches of the Phytozome-curated *M.*

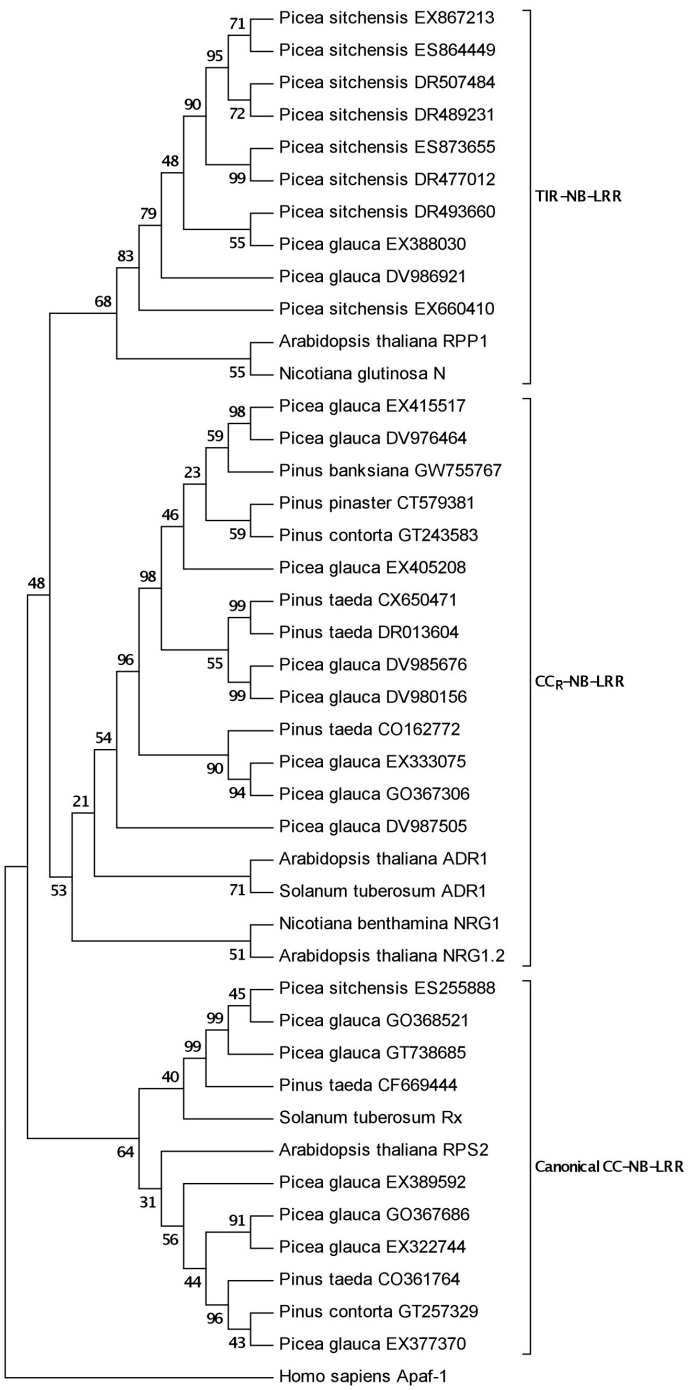


Figure 4.4. *Phylogenetic analysis of gymnosperm NB-LRR ESTs*. Bootstrap consensus tree constructed by the NJ method of gymnosperm ESTs described in the text, based on aligned NB-ARC protein sequence corresponding to Rx amino acids 240-397. StADR1, AtADR1, NbNRG1, and AtNRG1.2 are included for visualizing CC_R-NB-LRR structure. N and Arabidopsis RPP1 sequences are included to identify the TIR-NB-LRR class, and Rx and Arabidopsis RPS2 are included to identify the canonical CC-NB-LRR class. Bootstrap values from 1000 trials are displayed at branch nodes. The tree is rooted with human Apaf-1.

guttatus proteome (*Mimulus* Genome Project, DOE Joint Genome Institute) using Rx and N as queries revealed an abundance of CC-NB-LRR proteins, but no TIR-NB-LRR proteins (data not shown). As such a condition is unlikely to be confined to a single species, we extended our investigation to the order Lamiales, of which monkey flower is a member. This was accomplished by searching GenBank Lamiales ESTs, based on the NB-ARC regions of StADR1 (see below), NbNRG1, Rx, and N, all of which are derived from the Solanales, a sister group to the Lamiales (Chase and Reveal 2009). For each of the four searches, the top hit from each of the first six occurring species or varieties was collected, such that six sequences were compiled per search (Table 4.2). Although definitive conclusions cannot be made in the absence of additional fully sequenced genomes, there are nonetheless approximately 540,000 Lamiales ESTs currently available in GenBank. We were able to identify three Lamiales CC_R-NB-LRR ESTs, all of which grouped with ADR1 rather than NRG1 (Figure 4.5A). Furthermore, searches performed with N as a query returned exclusively CC-NB-LRR sequences. Thus all Lamiales EST sequences analyzed belonged either to the canonical CC-NB-LRR or the ADR1-like groups (Figure 4.5A, Table 4.2). As such, our findings are consistent with an absence of both *NRG1* homologs and TIR-NB-LRR-encoding genes in the order Lamiales. These findings are also consistent with those of McHale *et al.* (2006), whose analysis of available NB-ARC sequences, though not exhaustive, failed to identify any Lamiales-derived *TIR-NB-LRR* members.

In the case of columbine, minimal sequence resources available for related species precluded analysis at the family or order level. We therefore confined our examination to the Phytozome-curated *A. coerulea* proteome (DOE Joint Genome Institute). Having already determined columbine's CC_R-NB-LRR complement we

Table 4.2 Results of Lamiales and *A. coerulea* BLAST searches

Query	Speices	Identifier[a]	Group[b]
Lamiales			
N	*Mimulus guttatus*	GR052998	ADR1
N	*Antirrhinum majus*	AJ568282	ADR1
N	*Triphysaria pusilla*	EY157110	ADR1
N	*Mimulus guttatus var. nastus*	GO954238	CC
N	*Striga hermonthica*	FS493128	CC
N	*Mimulus lewisii*	GR192604	CC
Rx	*Mimulus guttatus var. nastus*	GO954238	CC
Rx	*Mimulus guttatus*	GR152322	CC
Rx	*Triphysaria pusilla*	EY182905	CC
Rx	*Striga hermonthica*	FS446680	CC
Rx	*Olea europaea*	GO244368	CC
Rx	*Antirrhinum majus*	AJ796877	CC
NRG1	*Antirrhinum majus*	AJ568282	ADR1
NRG1	*Mimulus guttatus*	GR052998	ADR1
NRG1	*Triphysaria pusilla*	EY157110	ADR1
NRG1	*Mimulus gutattus var. nastus*	GO955158	CC
NRG1	*Olea europaea*	GO244368	CC
NRG1	*Striga hermonthica*	FS444876	CC
StADR1	*Mimulus guttatus*	GR052998	ADR1
StADR1	*Antirrhinum majus*	AJ568282	ADR1
StADR1	*Triphysaria pusilla*	EY157110	ADR1
StADR1	*Striga hermonthica*	FS444876	CC
StADR1	*Mimulus gutattus var. nastus*	GO955158	CC
StADR1	*Olea europaea*	GO244368	CC
Aquilegia coerulea			
N	*A. coerulea Goldsmith*	002643m	CC
N	*A. coerulea Goldsmith*	025091m	CC
N	*A. coerulea Goldsmith*	022597m	CC
N	*A. coerulea Goldsmith*	019262m	CC
N	*A. coerulea Goldsmith*	027686m	CC
Rx	*A. coerulea Goldsmith*	019145m	CC
Rx	*A. coerulea Goldsmith*	016085m	CC
Rx	*A. coerulea Goldsmith*	013701m	CC
Rx	*A. coerulea Goldsmith*	023629m	CC
Rx	*A. coerulea Goldsmith*	027189m	CC

[a] Identifier for Lamiales = GenBank accession number, for *A. coerulea* = transcript name (AcoGoldSmith_v1.)

[b] "ADR1" = ADR1-like, "CC" = canonical CC-NB-LRR

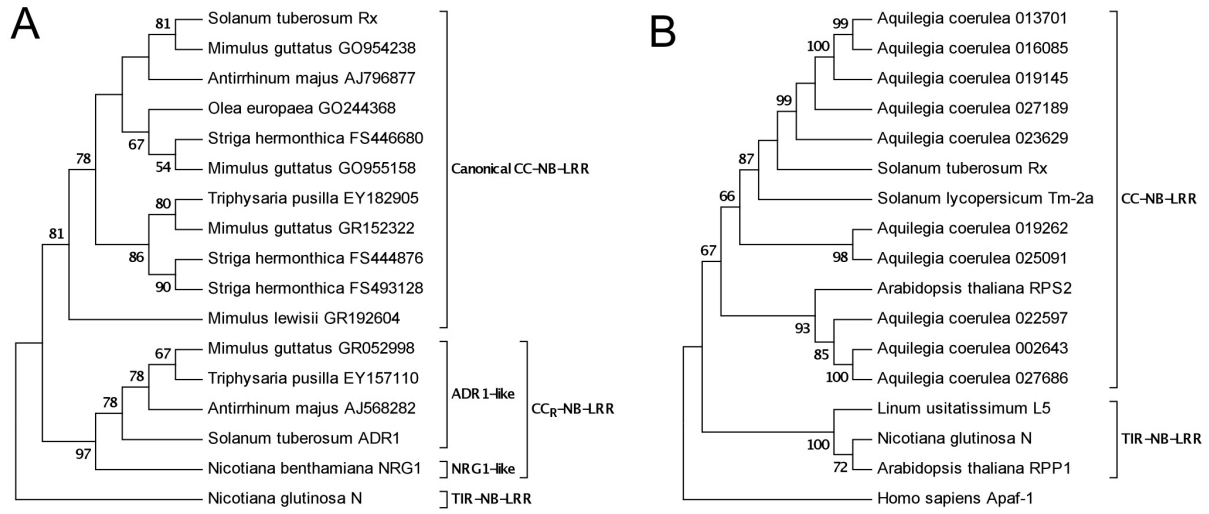


Figure 4.5. *Absence of TIR-NB-LRR sequences from Lamiales and* Aquilegia coerulea. **A,** Phylogenetic analysis of Lamiales EST sequences compiled from BLAST searches querying with StADR1, NbNRG1, Rx, and N sequences. Query sequences are also included as points of reference. The tree was constructed by NJ, from an alignment of internal NB-ARC sequence corresponding to NbNRG1 amino acids 110-176. **B,** Phylogenetic analysis of *A. coerulea* NB-LRR sequences compiled by querying the genome with Rx and N sequences. Sequences from CC-NB-LRR proteins Rx, *A. thaliana RPS2*, and *S. lycopersicum* Tm-2a, and from TIR-NB-LRR proteins N, *A. thaliana* RPP1, and *Linum usitatissimum* L5 are also included as points of reference. The tree was constructed by NJ and based on aligned NB-ARC domain sequences, and is rooted with human Apaf-1. For both trees bootstrap values are based on 1000 replicates, and percent values over 50 are displayed at branch points.

simply performed BLASTP searches with the Rx and N NB-ARC domains as queries, compiling the top five hits from each search (see Table 4.2). Phylogenetic analysis of this dataset revealed that all sequences recovered grouped with CC-NB-LRR proteins, indicating an absence of TIR-NB-LRR-encoding genes from the *A. coerulea* genome (Figure 4.5B). Thus combined data from monocots, Lamiales, and columbine indicates a strong association between the occurrence of NRG1-like and TIR-NB-LRR-encoding genes.

Sequence analysis of the CC_R domain

Although conserved domain searches of CC_R-NB-LRR proteins identify an RPW8 domain (RPW8 superfamily – PF05659, Pfam database available online from the Sanger Institute) at the N-terminus, this similarity has received relatively little attention. Using alignments of all angiosperm sequences represented in Figure 4.2, we constructed separate NRG1-like and ADR1-like CC_R domain consensus sequences. We also constructed an RPW8 consensus, based on Arabidopsis Ms-0 RPW8.1 and RPW8.2 and Col-0 HR2 and HR3. We then aligned these consensi, along with four representative sequences from each group, in order to examine the nature of sequence similarity between canonical RPW8 homologs and CC_R domains (Figure 4.6). Positions where residues are functionally similar between the three groups are evenly distributed over most of the sequence length (Figure 4.6, pink). These are interspersed with residues conserved solely among CC_R domains, or specifically among either ADR1 or NRG1 homologs (Figure 4.6, blue). Similarity largely ends at a position corresponding to RPW8.2 amino acid 144. After this point there is no similarity between RPW8 proteins and CC_R domains, and very little similarity preserved among

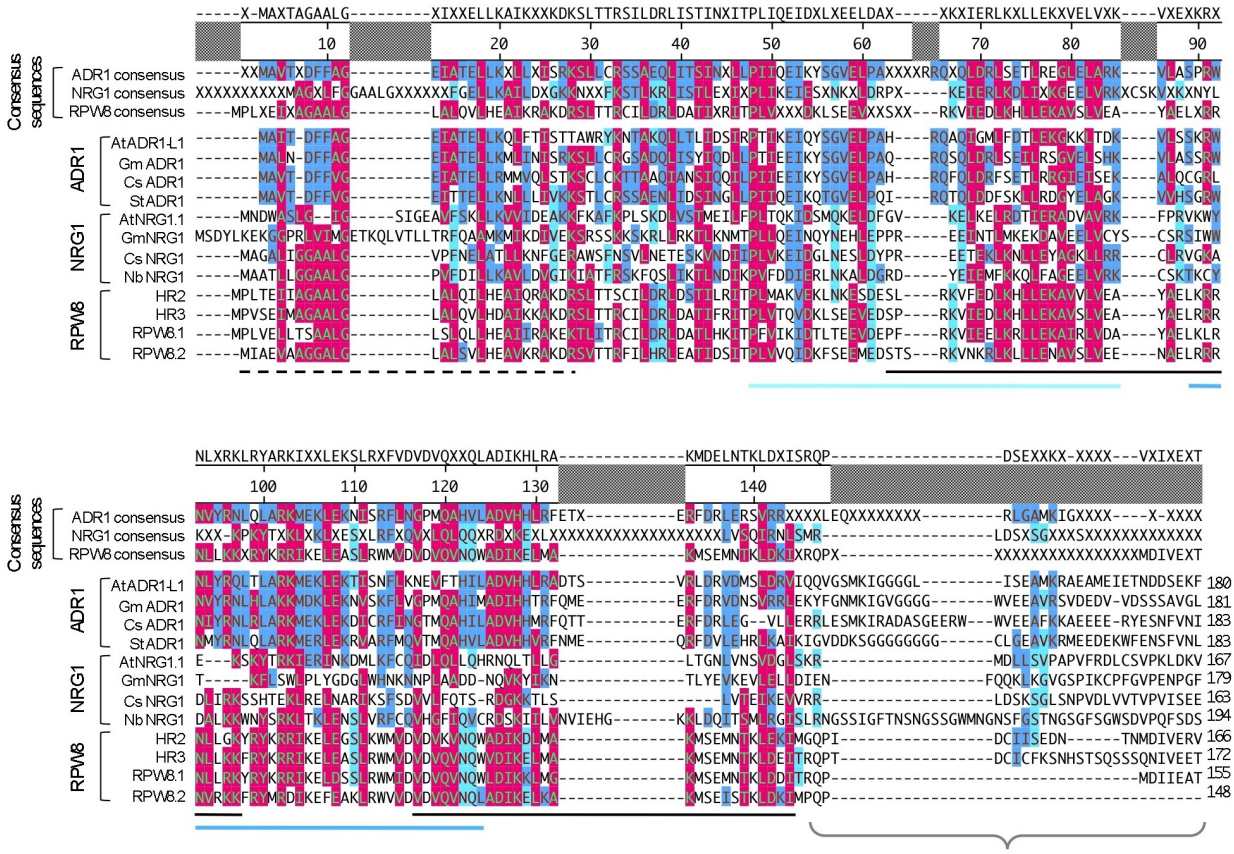


Figure 4.6 *Multiple sequence alignment of CC_R domains.* Consensus sequences for ADR1, NRG1, and RPW8 protein groups are displayed at top, with four representatives from each group displayed below (Gm – *Glycine max*, Cs – *Cucumis sativus*, sequences are the same as those used in Fig. 2). The ruler at top is of RPW8.2. Where a majority of residues are functionally similar to the RPW8 consensus sequence, those residues are highlighted in pink. Additional regions of conservation corresponding to the ADR1 and NRG1 consensus sequences are highlighted in shades of blue. Black underlines correspond to the two predicted coiled coils of RPW8.2, dashed black underline corresponds to the RPW8.2 predicted transmembrane domain / signal peptide. Blue underlines indicate locations of the most strongly predicted coiled-coils from NRG1-like and ADR1-like groups. Gray bracket indicates the location of probable CC_R – NB linker sequence.

CC_R domains. As this poor sequence conservation is at the end of the functionally defined CC_R domain, continuing until the beginning of the NB domain, it likely represents a linker region between the two domains (Figure 4.6, gray bracket).

RPW8 is described as having two C-terminal coiled-coils (Figure 4.6, black underlines) (Xiao *et al.* 2001). We analyzed each of the 12 representative sequences from Figure 4.6 with the COILS prediction program (Lupas et al. 1991) (available online from the Swiss Institute of Bioinformatics EMBnet), and indicate the location of the best-predicted coiled-coils for NRG1-like and ADR1-like groups as blue underlines in Figure 4.6. However, coiled-coil predictions for CC_R domains were rarely as strong as for RPW8 proteins (Figure 4.7). Given this lack of consistent prediction, it is uncertain whether coiled-coils are a relevant structural feature of CC_R domains. Sequence similarity between RPW8 proteins and CC_R domains extends to the extreme N-terminus, which in RPW8 has been proposed to constitute a transmembrane (TM) domain. We were unable to identify any similarly conserved TM predictions among CC_R-NB-LRR proteins. While RPW8.1, RPW8.2, HR2, and HR3 are all predicted by ConPredII (available online from Hirosaki University) to possess TM domains at their N-termini, only NbNRG1, AtNRG1.2, and *Populus trichocarpa* 0007s11510 (an NRG1 group member) out of all sequences included in Figure 4.2 receive similar TM predictions. It is therefore uncertain whether the N-terminal sequence conservation between RPW8 proteins and CC_R domains is indicative of functional conservation.

It is interesting to note that while the majority of CC_R domains do not possess a predicted TM domain, those that do all belong to the NRG1 rather than the ADR1 subclass. To further elucidate the relationship between NRG1 and ADR1 CC_R domains and RPW8 domains, we constructed a phylogeny from the CC_R domains of

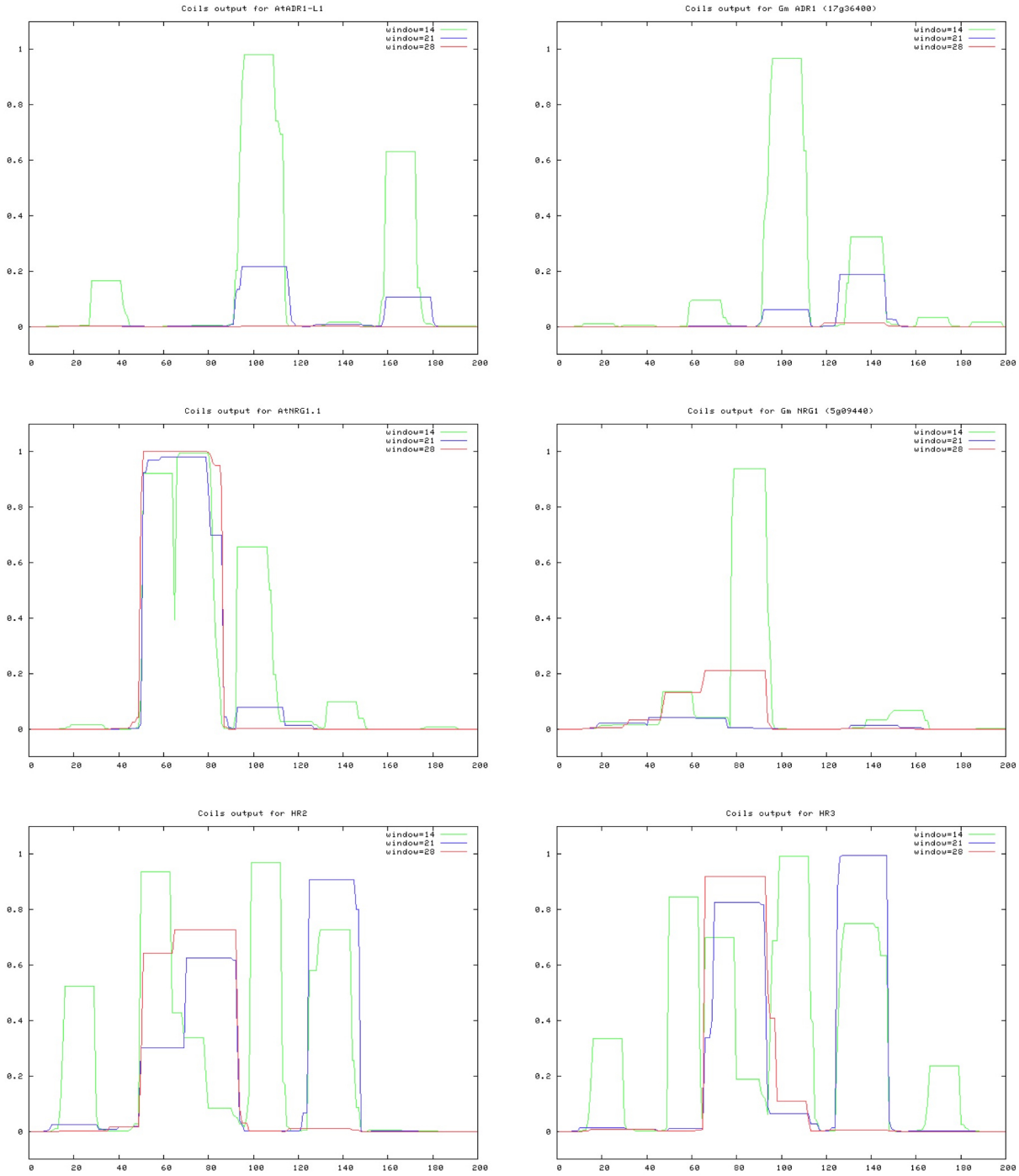
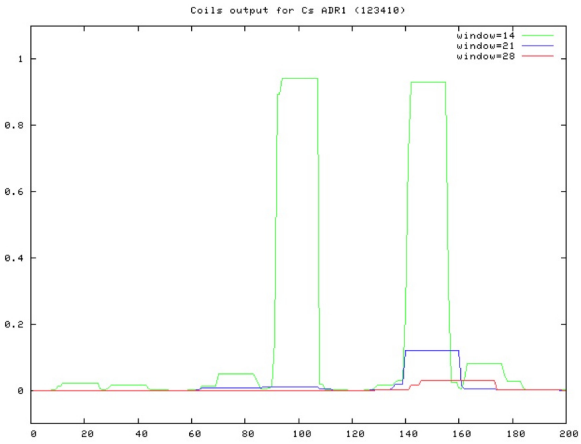
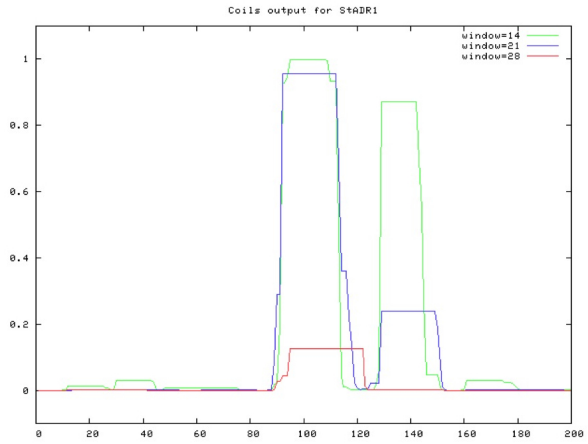
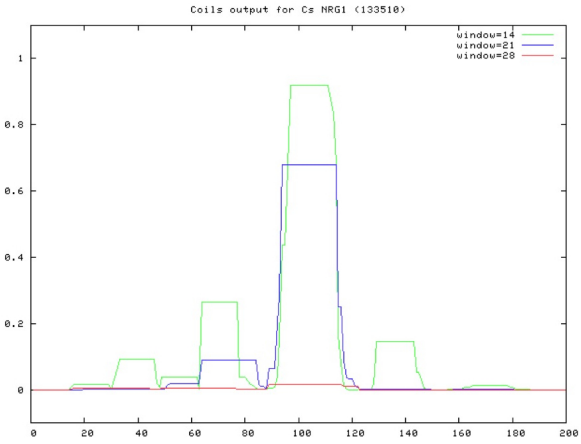


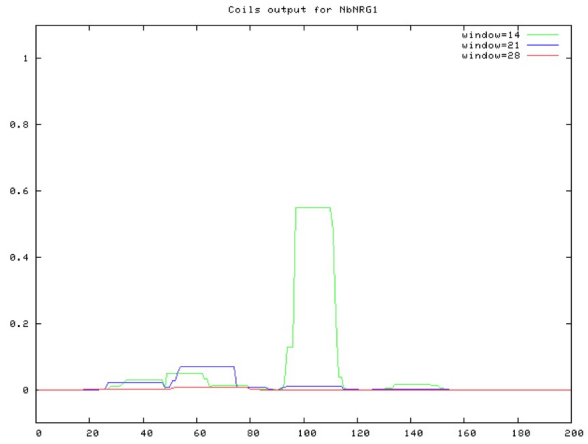
Figure 4.7 *Coiled-coil prediction for representative CC_R-NB-LRR N-termini.* The COILS program was used to analyze the first 200 amino acids of representative ADR1, NRG1, and RPW8 homologs (Gm – *Glycine max*, Cs – *Cucumis sativus*, sequences are the same as those used in Fig. 2) using the MTIDK matrix, unweighted. Different colored lines represent the length, in amino acids, of the scanning window used by the algorithm. Values on the y-axis indicate the probability of forming coiled-coils for each algorithm. Values on the x-axis indicate sequence position in amino acids.

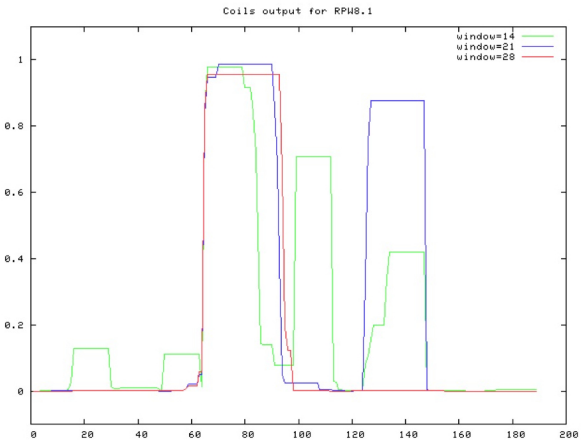
Figure 4.7 (Continued)

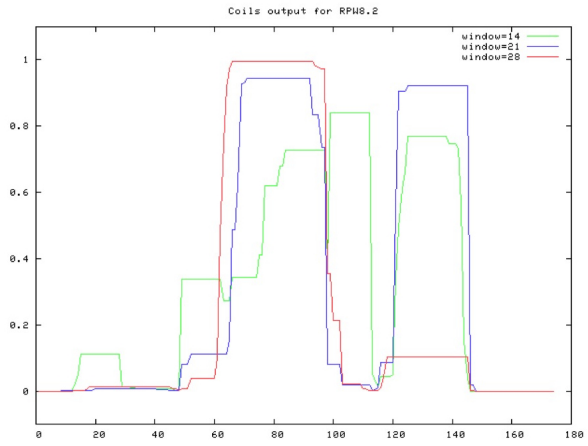












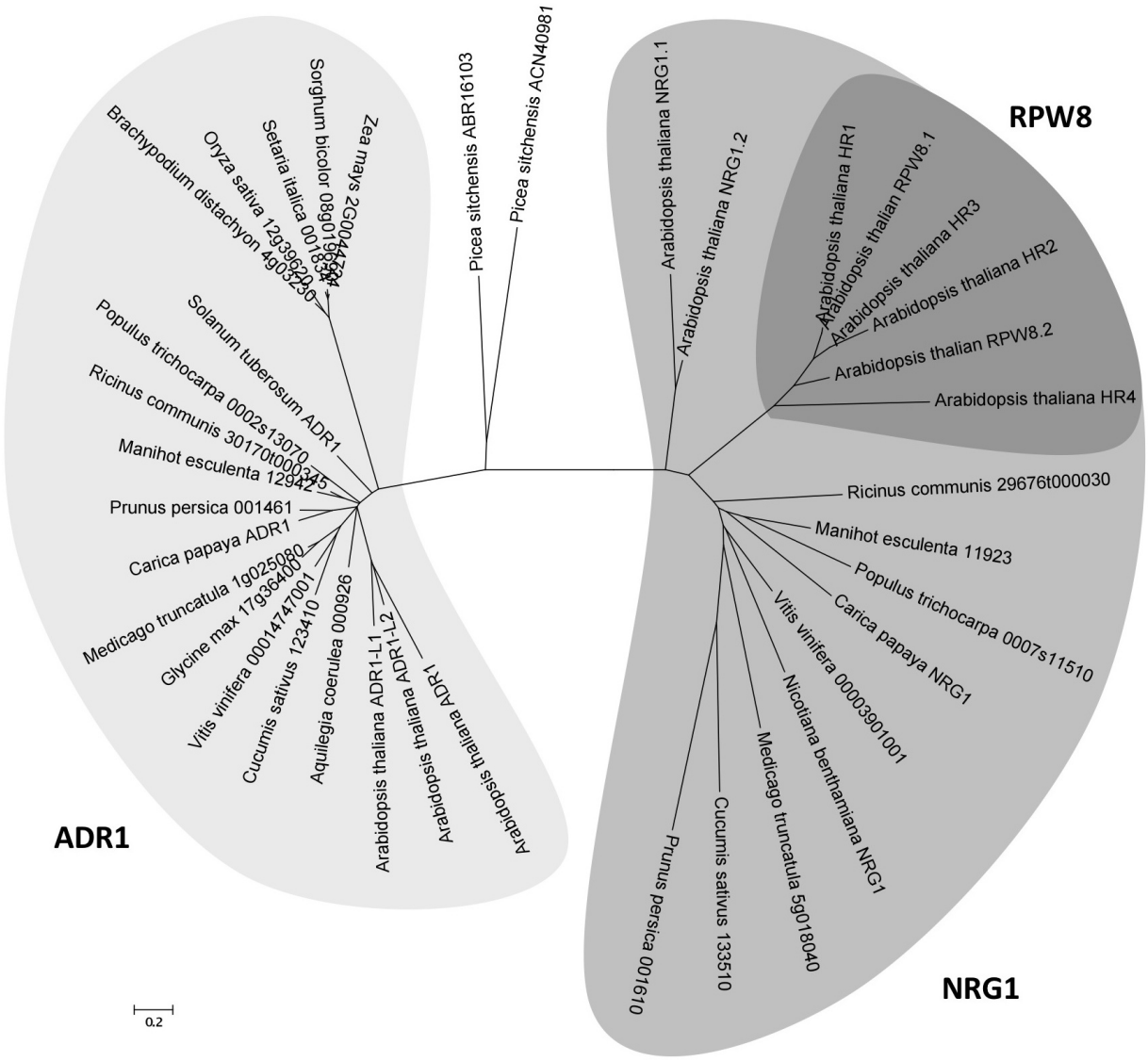


Figure 4.8. *Phylogenetic analysis of CC_R and RPW8 domains.* Bootstrap consensus tree from the aligned CC_R and RPW8 internal region corresponding to RPW8.2 positions 27-135. CC_R domains are from sequences included in Fig. 2 after removal of any sequences with incomplete N-terminal coverage. The tree was constructed by NJ and is drawn to scale, with branch lengths corresponding to evolutionary distance. The scale bar is in units of amino acid substitutions per site. ADR1-like, NRG1-like, and RPW8 groups are indicated by background shading.

sequences included in Figure 4.2, as well as RPW8 domains of Arabidopsis RPW8.1, RPW8.2, HR1, HR2, HR3, and HR4 (Figure 4.8). This analysis again illustrates a clear division between NRG1 and ADR1 groups, and also suggests that RPW8 homologs share more similarity with NRG1 rather than ADR1 CC_R domains. Remarkably, RPW8 homologs are in fact placed *within* the NRG1 group, albeit distinguished from other group members by a fairly long branch.

Similar defense response induction by NRG1 and ADR1 homologs

To expand our functional analysis to ADR1-like proteins, we cloned an *ADR1* homolog from potato, hereafter referred to as *StADR1* (GenBank accession number HQ906887), based on available potato genome sequence. We found that transient expression of the first 155 amino acids (terminating a few residues into the linker region) was sufficient to induce a rapid HR in tobacco and *N. benthamiana* (Figure 4.9A). We also cloned the CC_R-encoding domains of Arabidopsis *ADR1* (At1g33560), *ADR1-L1* (At4g33300), *ADR1-L2* (At5g04720), At5g66900, and At5g66910. We refer to the latter two genes as *AtNRG1.1* and *AtNRG1.2*, respectively, based on their similarity to *N. benthamiana NRG1*. Of these, the AtADR1 and AtNRG1.2 CC_R domains consistently induced a rapid and strong HR upon transient expression, while those of AtADR1-L1, AtADR1-L2, and AtNRG1.1 induced mild chlorosis (Figure 4.9A). Thus CC_R-mediated HR induction is a common property of both CC_R-NB-LRR clades and is conserved among different plant families. Interestingly, while solanaceous CC_R domains produced a strong HR in tobacco, their effect was weaker in *N. benthamiana,* whereas some Arabidopsis-derived CC_R domains showed an opposite pattern. In a striking example, NRG1.2 produced a strong HR in *N. benthamiana*, but yielded only

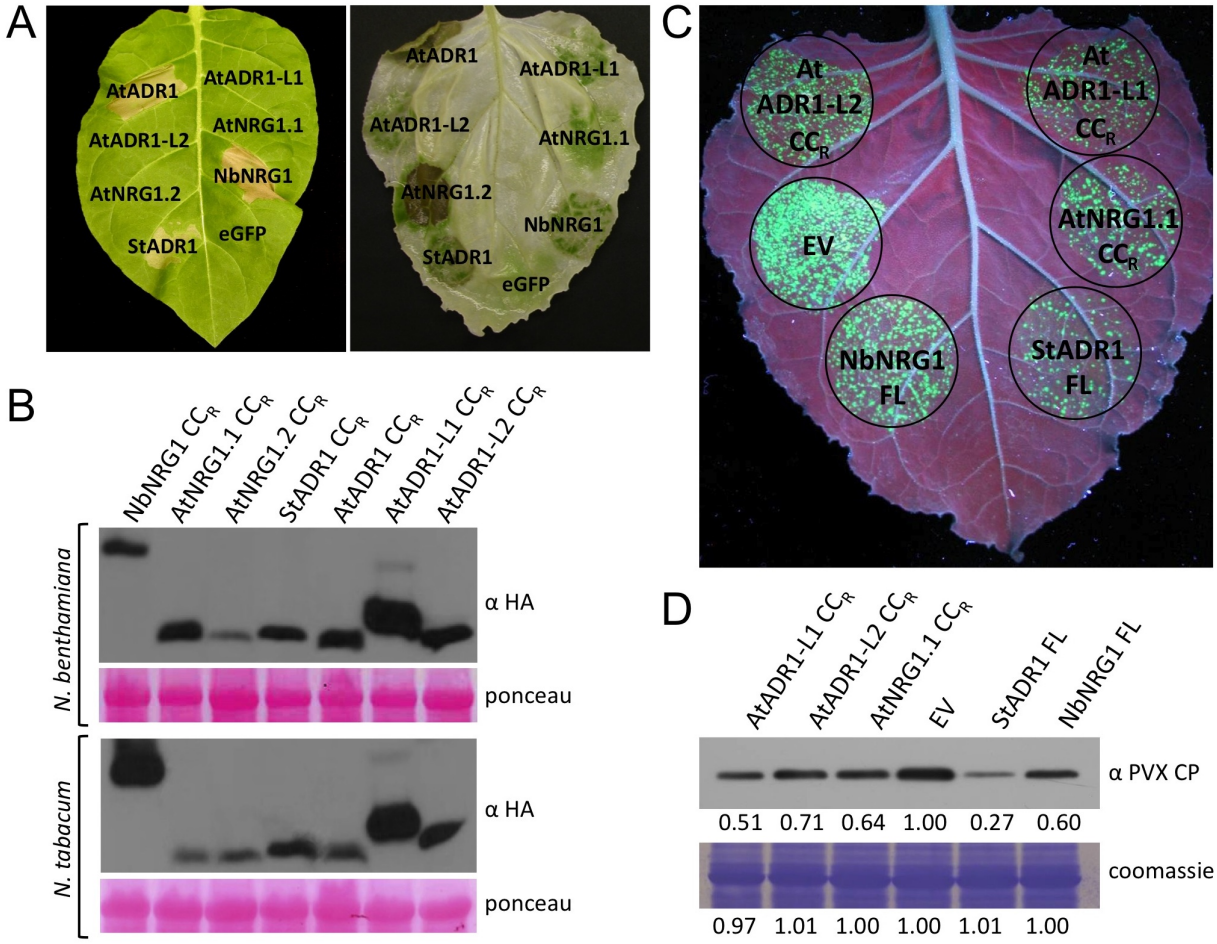


Figure 4.9 *Defense response induction by CC_R domains.* **A,** NRG1 and ADR1 CC_R-mediated HR in tobacco (left panel) and *N. benthamiana* (right panel), the latter having been ethanol cleared to enhance visibility. Leaves were photographed three days after agroinfiltration with the indicated HA-tagged Arabidopsis or solanaceous CC_R domains. Note the NbNRG1 CC_R corresponds to the construct 1-225. **B,** Protein extracts from agroinfiltrated patches of *N. benthamiana* or *N. tabacum* leaves, as indicated, in (**A**) were subjected to anti-HA immunoblotting after 24 hours. Equal loading was verified by ponceau staining. **C,** PVX:GFP accumulation after co-infiltration with non-HR-inducing CC_R-NB-LRR variants. GFP was visualized five days after agroinfiltration by handheld UV lamp. **D,** Five days after agroinfiltration, protein extracts from (**C**) were subjected to immunoblotting with antibody against the PVX coat protein (CP) (top panel). Equal loading was verified by coomassie staining (bottom panel). For quantification of protein levels (listed below each band), average gray values from non-saturated exposures (PVX CP) or total protein bands (coomassie) were measured and converted to a scale where the EV measurement = 1.0. All experiments were performed at least three times with similar results.

mild chlorosis in tobacco (Figure 4.9A). These differences were not due to differences in protein expression however, as CC_R domains were expressed at similar levels in both species (Figure 4.9B). These observations suggest that the ability of a given CC_R domain to induce an HR may be affected by genetic background.

Over-expression of full-length NRG1 has been reported to induce defense responses and suppress viral accumulation in the absence of an HR (Peart et al. 2005). We tested whether this activity was common to ADR1 as well by co-expressing full-length NbNRG1 and StADR1 with a version of *Potato virus X* expressing green florescent protein (PVX:GFP). Indeed, StADR1 reduced the amount of GFP accumulation from PVX:GFP as judged both visually and by anti-PVX coat protein (CP) immunoblotting (Figure 4.9C and 4.9D), indicating that, like NbNRG1, it is able to induce anti-viral defense responses in the absence of a visible HR. We similarly tested non-HR inducing Arabidopsis CC_R domains and found that their expression also caused reduced viral replication (Figure 4.9C and 4.9D), indicating that these CC_R-NB-LRR proteins also have the capacity to induce a genuine resistance response.

CC_R signaling is SGT1-independent

Many CC-NB-LRR proteins require SGT1 to function, due to a dependence on SGT1-containing chaperone complexes for proper folding and accumulation (Kadota et al. 2010). The suppression of viral replication associated with full-length NbNRG1 over-expression is compromised in SGT1-silenced plants (Peart et al. 2005). We investigated whether this SGT1 functional dependency extends to CC_R-mediated signaling as well. As expected, HR induction by both a constitutively active full-length CC-NB-LRR protein (Rx D460V) and by the Rx NB domain (NB:eGFP), both of

which induce an HR within the same timeframe as the NRG1 CC_R domain, was abolished in plants where SGT1 levels had been reduced by virus-induced gene silencing (VIGS) (Figure 4.10). In contrast, HRs induced by NbNRG1 and StADR1 CC_R domains were not affected by SGT1 silencing, nor was cell death induced by the mouse Bax protein used as a control (Figure 4.10). Thus, while full-length CC_R-NB-LRR proteins may require SGT1, the ability of CC_R domains to initiate signaling does not appear to require SGT1, at least not at the levels required for other strong HR inducers.

CC_R-NB-LRR requirement in CC-NB-LRR-mediated disease resistance

Given that NRG1 is required for the function of at least one NB-LRR protein, we were interested in whether this might represent a larger trend, and whether ADR1 might also be prerequisite for the activities of certain NB-LRR proteins. We therefore constructed VIGS vectors to silence either *NRG1*, *ADR1*, or both *NRG1* and *ADR1* combined, using *NRG1* sequence from *N. benthamiana* and *ADR1* sequence from *N. tabacum*. *N. benthamiana* possesses a single functional copy of *NRG1* (Peart et al. 2005) and we were able to identify only single *ADR1*-like sequences from tobacco, tomato (*S. lycopersicum*), and potato – suggestive of single-copy status within the Solanaceae. Thus, *Nicotiana* provides a more tractable system than Arabidopsis, which possesses multiple copies of both *NRG1* and *ADR1*, some arrayed in tandem. *NRG1* VIGS has previously been reported to efficiently break N- but not Rx-mediated virus resistance. However, Rx induces an exceptionally strong resistance response which may be difficult to break with gene knockdown techniques such as VIGS (Sacco et al. 2007). *Rx2* is a paralog of *Rx*, with similar recognition specificity for the PVX CP, but with weaker resistance-inducing activity compared to *Rx* (Bendahmane et al. 2000). We

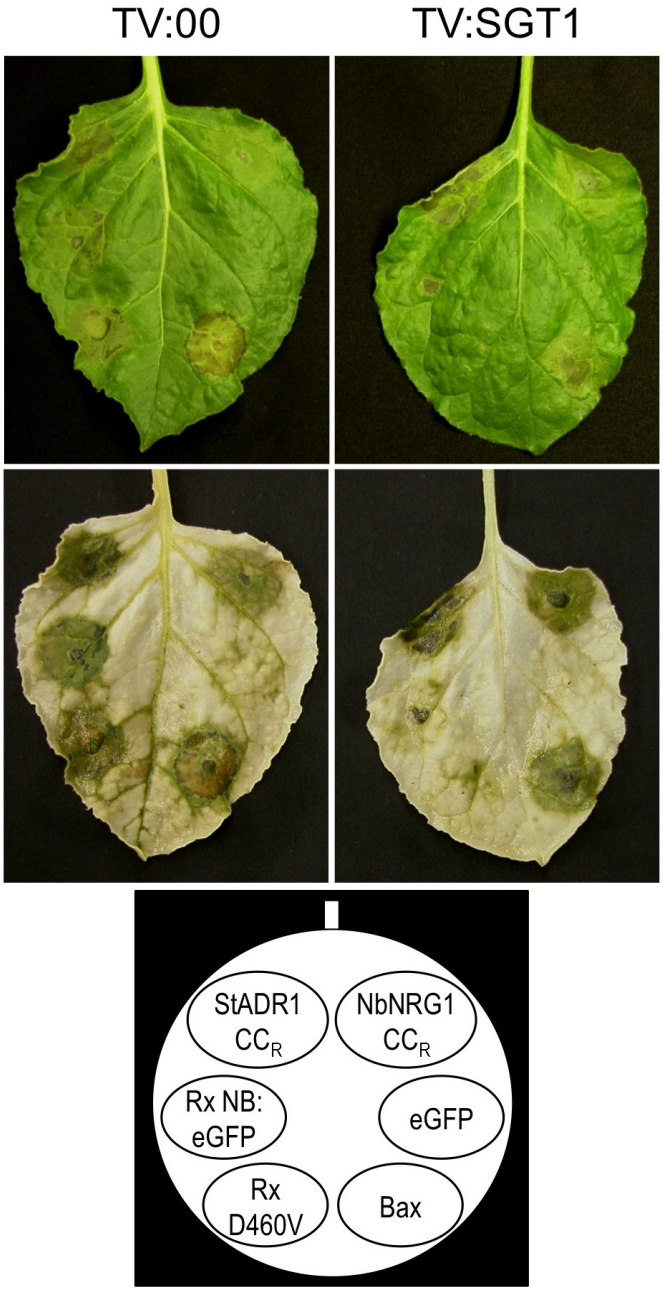


Figure 4.10 *CC_R-mediated HR induction is SGT1-independent.* Plants silenced for SGT1 (TV:SGT1, right) or infected with tobacco rattle virus (TRV) empty vector control (TV:00, left) were agroinfiltrated with the constructs indicated (bottom panel). Leaves were photographed five days later before (top panels) and after (middle panels) ethanol clearing.

therefore utilized *N. benthamiana* expressing *Rx2* as a transgene in VIGS experiments (Bhattacharjee et al. 2009). Silenced plants were rub-inoculated with PVX:GFP and subsequently monitored for breakage of resistance, which manifested as trailing necrosis – the result of virus escape followed by delayed defense response induction (Figure 4.11B). VIGS using the empty vector leads to a low level of spontaneous resistance breaking, and silencing of *NRG1* or *ADR1* alone had little additional effect on Rx2-mediated resistance. However, silencing of *NRG1* and *ADR1* together resulted in a consistent increase in the frequency of Rx2 breakage to nearly fourfold that of the empty vector control (Figure 4.11A), suggesting some involvement of CC_R-NB-LRR proteins in Rx2-mediated resistance.

DISCUSSION

We show here that ADR1 and NRG1 homologs comprise a family of proteins whose N-termini possess activities not observed for canonical CC-NB-LRR proteins. These homologs are easily identifiable between distantly related plant species, and phylogenetic analysis suggests that these proteins have been retained for conserved function(s), in contrast to other NB-LRR protein families, which have undergone a great deal of diversification and amplification. As such, we suggest that this family of proteins represents a unique set of NB-LRR proteins and propose that they be referred to hereafter as CC_R-NB-LRR proteins in order to differentiate them from other NB-LRR proteins.

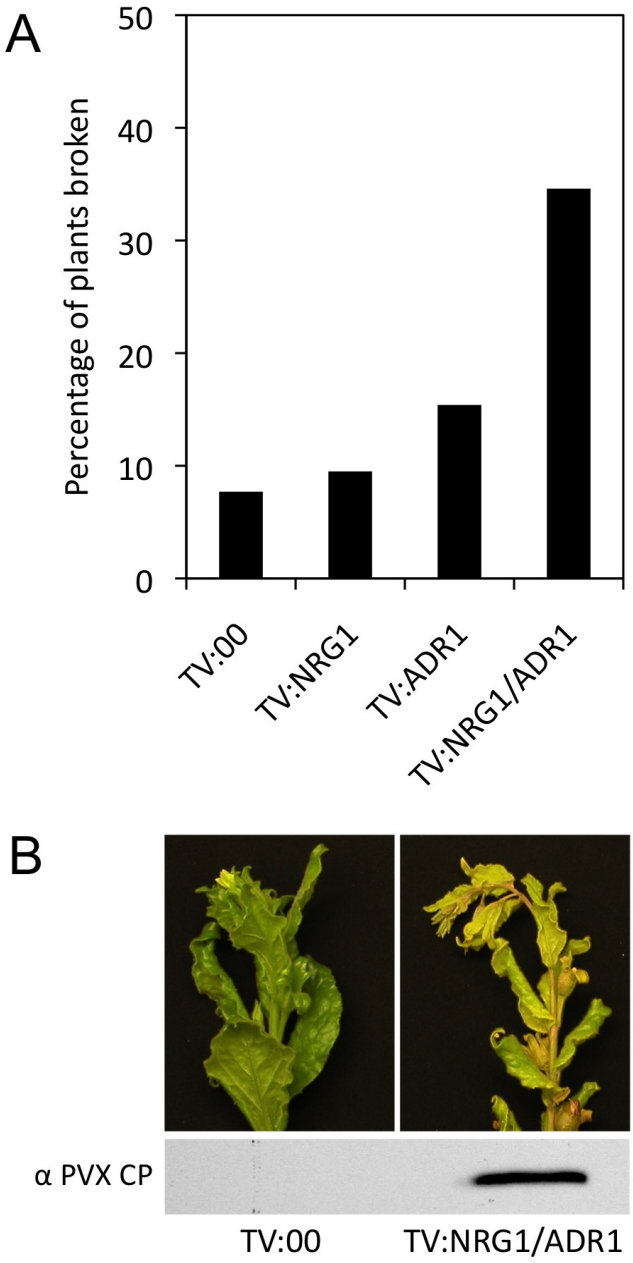


Figure 4.11 *Rx2-mediated resistance is partially compromised by combined NRG1 and ADR1 silencing*. **A,** Percentage of plants displaying trailing necrosis following PVX:GFP inoculation, having previously been infected with the TRV empty vector (TV:00) or silenced for NRG1, ADR1, or NRG1 + ADR1 by VIGS. **B,** Appearance of healthy (left) and diseased (right) plants (top panels) following PVX:GFP inoculation. All instances of resistance breaking were confirmed by anti-PVX CP immunoblotting, a representative example of which is shown (bottom panel).

The CC_R domain

The CC_R domains of both NRG1- and ADR1-like proteins bear similarity to RPW8 across the majority of their sequence, interspersed with regions of conservation shared only among NRG1 and ADR1 homologs (Figure 4.6). While the proteins encoded by RPW8 homologs are strongly predicted to form coiled-coil structures, coiled-coil predictions for CC_R domains are typically weaker and less consistent (Figure 4.7). Whether a coiled-coil structure has any relevance to CC_R function is therefore uncertain. It is striking that the similarity between CC_R domains and RPW8 extends to the extreme N-terminus, which in RPW8 is predicted to contain a transmembrane domain and signal peptide, thought to be critical to its targeting to the extrahaustorial membrane during powdery mildew infection (Xiao et al. 2001; Wang et al. 2009). As similar transmembrane and targeting functional predictions are uncommon among CC_R sequences, however, the relevance of this sequence conservation remains to be determined. Indeed, it is noteworthy that NbNRG1 (1-225) migrates more slowly than NbNRG1 (13-225) during SDS-PAGE (Figure 4.1A), which would not be the case if a canonical signal peptide had been cleaved from the former during translocation to the endoplasmic reticulum. An additional question raised by this analysis is whether RPW8 is derived from CC_R-NB-LRR proteins or vice-versa, or whether they arose from a common precursor. Notably, while we have found CC_R-NB-LRR genes to be present in every seed plant species examined, RPW8 homologs have thus-far been identified only among members of the Brassicaceae (Xiao et al. 2004), arguing against an ancient origin for the RPW8 gene family. This observation combined with our placement of RPW8 homologs within the NRG1 clade (Figure 4.8)

suggests that RPW8 homologs may have originated as an offshoot of CC_R-NB-LRR proteins, likely deriving from the NRG1 group.

Avr-independent HR induction by fragments of NB-LRR proteins is a useful tool for determining the domains involved in signal initiation (Rairdan et al. 2008; Swiderski et al. 2009; Krasileva et al. 2010). We have demonstrated that CC_R domains of both NRG1 and ADR1 homologs are able to trigger HR. We have further shown this to be part of a bona fide resistance response, capable of restricting viral accumulation (Figure 4.9). As CC_R-mediated defense response induction is found not only among solanaceous homologs but also among NRG1 and ADR1 homologs from Arabidopsis, it is likely that this is a common feature of CC_R domains. Thus the N-termini of CC_R-NB-LRR proteins appear to function quite differently from those of canonical CC-NB-LRR proteins, which by themselves have only been shown to mediate recognition (Collier and Moffett 2009).

The function of many CC-NB-LRR proteins, including full-length NRG1, depend on the presence of SGT1 for proper folding and accumulation, presumably through interaction with the LRR domain (Bieri et al. 2004; Leister et al. 2005; Peart et al. 2005; Zhang et al. 2008; Kadota et al. 2010). Furthermore, signal initiation by the Rx NB domain retains its SGT1-dependence (Rairdan et al. 2008), suggesting a similar requirement for SGT1 by some protein(s) which function downstream of Rx activation. It is therefore noteworthy that the HRs induced by NbNRG1 and StADR1 CC_R domains appear to be SGT1-independent (Figure 4.10), likely because SGT1 is not required for the folding of the relatively simple CC_R domain alone. It is thus tempting to speculate that either CC_R domains signal through different pathways than canonical CC-NB-LRR proteins, or that they function downstream of SGT1-dependent factors.

CC_R-NB-LRR evolution and evidence for functional requirement

The NB-LRR gene family is noted for its plasticity and rapid evolution, frequently featuring both gene duplication and gene loss during speciation (McDowell and Simon 2006; Sakai and Itoh 2010; Xu et al. 2011). Representation of various NB-LRR lineages can vary greatly between species, and indeed much variation is seen even at the broadest level of classification, with ratios of CC-NB-LRR- to TIR-NB-LRR-encoding genes ranging from greater than 10:1 in cocoa (*Theobroma cacao*) to less than 0.6:1 in Arabidopsis (Velasco et al. 2010; Argout et al. 2011). It is therefore remarkable that at least one ADR1 homolog is present in every higher plant species examined, and that the ADR1 lineage has not undergone a sizable expansion in any of the species studied (Arabidopsis and soybean (*Glycine max*) have the largest *ADR1* gene families, with four members each (Table 4.1)). Such a conservative evolutionary history suggests that ADR1 may perform some critical and specific function. While this function might conceivably be the recognition of an uncommonly ubiquitous pathogen, possibly through the highly conserved LRR domain, it is perhaps more likely to play a more general role in disease resistance. CC_R-NB-LRR proteins might act directly downstream of canonical NB-LRR proteins. At the same time, it is tempting to speculate that they might behave in a manner similar to the animal NOD-like receptor NLRP3, which responds to perturbations in cellular ion concentrations and reactive oxygen species caused indirectly by infection (Tschopp and Schroder 2010). Were such the case, CC_R-NB-LRR proteins might sense initial changes in cellular homeostasis and subsequently amplify resistance signaling capacity.

A specific interplay between CC_R-NB-LRR proteins and canonical NB-LRR proteins is further suggested by the close association between *NRG1* homologs and *TIR-NB-LRR* genes. While NbNRG1 was originally reported to be required

exclusively by N, the only other resistance proteins examined in a resistance assay for NRG1 requirement, Rx and Pto/Prf, were of the CC-NB-LRR rather than TIR-NB-LRR class (Peart et al. 2005). The correlation between the presence of both *NRG1* homologs and *TIR-NB-LRR* genes presented here suggests that the functional relationship between NRG1 and N may represent a broader trend rather than an isolated instance. NRG1-encoding sequences are absent from all monocot species examined (Figure 4.2), mirroring the observation that monocots lack TIR-NB-LRR proteins (Meyers et al. 1999; Pan et al. 2000; Bai et al. 2002; Tarr and Alexander 2009). This absence was, until now, thought to be exclusive to monocots. However, we have found that *NRG1* homologs as well as TIR-NB-LRR-encoding genes are also absent from the genomes of the dicots *M. guttatus* and *A. coerulea* (Figure 4.5). Furthermore, in the case of *M. guttatus* these absences are likely not limited to the species level, but extend to other members of the order Lamiales as well (Table 4.2, Figure 4.5).

The majority of fully sequenced dicot genomes belong to the rosid clade of core eudicots, with only a single basal eudicot (*A. coerulea*) and a single member of the asterid clade of core eudicots (*M. guttatus*). At first glance, one might speculate that the absence of *NRG1* homologs and TIR-NB-LRR-encoding genes observed in monocots, Lamiales, and *A. coerulea* represent single gene loss events. However, this scenario is highly unlikely, particularly considering that the Solanales – sister group to the Lamiales and fellow member of the asterid clade (Chase and Reveal 2009) – possesses members of both the *TIR-NB-LRR* and *NRG1* lineages. Thus any model of *TIR-NB-LRR* evolution must account for the presence of this lineage in gymnosperms, rosids, and the Solanales, and a single gene loss event would necessarily correspond to TIR-NB-LRR-encoding genes having arisen independently on three occasions. Much more likely is an evolutionary path described by a single gene acquisition in a

common ancestor of gymnosperms and angiosperms, followed by independent gene loss events in each of the three affected lineages (Figure 4.12A). In considering how such a large gene family could be entirely lost from at least three independent lineages, it is interesting to note that preliminary analysis of the cocoa genome has identified only eight TIR-NB-LRR-encoding genes, less than one tenth the number of CC-NB-LRR-encoding genes identified (Argout et al. 2011). It is possible that cocoa represents a plant lineage en route to eliminating its TIR-NB-LRR-encoding genes, perhaps offering a glimpse of the gene family at a midpoint in the process of decline. At the same time however, there is at least one *NRG1* homolog present in the cocoa genome (locus Tc01_g004490), which if functional would make it unlikely that the loss of *TIR-NB-LRR* genes is precipitated by the loss of *NRG1.*

The evolutionary pattern of *NRG1* homologs is somewhat more ambiguous than that of TIR-NB-LRR-encoding genes, owing largely to the uncertainty surrounding gymnosperm CC_R-NB-LRR-encoding genes. While our analysis of gymnosperm-derived sequences clearly indicates that they form a single group, displaying no NRG1 vs. ADR1 division, we were unable to resolve definitively whether this group predates the division of NRG1 and ADR1, or whether it might belong to a single clade (Figures 4.2, 4.4, and 4.8), thus calling into question the placement of NRG1 / ADR1 divergence. Assuming that gymnosperms possess an ancestral form of a *CC_R-NB-LRR* gene, the simple scenario whereby *NRG1* arises in the ancestor of core eudicots and is subsequently lost from the Lamiales lineage would explain the observed pattern of occurrence. However, this fails to address the fact that the split between NRG1 and ADR1 is, as suggested by Figure 4.2 and Figure 4.8, deeper than the split between monocots and dicots. Thus NRG1 and ADR1 presumably diverged before the divergence of monocots and dicots. The most

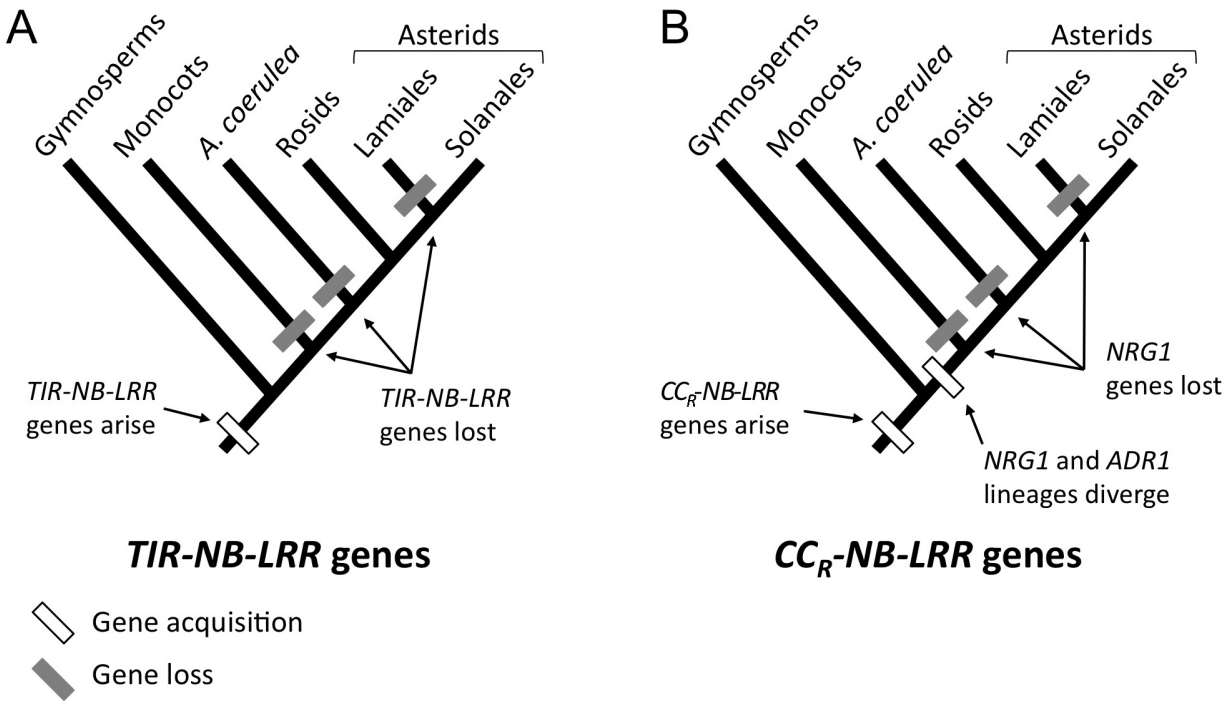


Figure 4.12 *Evolutionary models for* TIR-NB-LRR *and* CC_R-NB-LRR *lineages.* **A,** A probable evolutionary path leading to the observed patterns of *TIR-NB-LRR* occurrence. *TIR-NB-LRR* genes, having arisen before the divergence of gymnosperms and angiosperms, are subsequently lost from monocot, *A. coerulea*, and Lamiales lineages. **B,** A probable evolutionary path which produces the observed patterns of *NRG1* / *ADR1* occurrence. *CC_R-NB-LRR* genes having arisen before the divergence of gymnosperms and angiosperms, distinct *NRG1* and *ADR1* variants emerge within the angiosperm lineage and *NRG1* genes are subsequently lost from monocot, *A. coerulea*, and Lamiales lineages.

plausible scenario is therefore similar to that described for *TIR-NB-LRR* genes, the only difference being differentiation of *NRG1* and *ADR1* lineages after the divergence of angiosperms and gymnosperms (Figure 4.12B). Were the NRG1 / ADR1 divergence to have occurred in a common ancestor of gymnosperms and angiosperms, the single *CC_R-NB-LRR* clade present in gymnosperms could only be explained by the loss of either *NRG1* or *ADR1* homologs. This not only introduces additional evolutionary steps, but is unsupported by evidence either a) of any plant species losing the *ADR1* lineage, or b) of TIR-NB-LRR proteins existing in angiosperms in the absence of NRG1 proteins. Thus it is probable that gymnosperms, while possessing both CC-NB-LRR and TIR-NB-LRR proteins, possess an ancestral form of CC_R-NB-LRR proteins, and hence the dynamics between canonical NB-LRR proteins – especially of the TIR-NB-LRR class – and CC_R-NB-LRR proteins might be expected to be somewhat different in gymnosperms than in flowering plants.

Our phylogenetic analyses strongly suggest a critical role for ADR1-like proteins and a strong functional relationship between NRG1 and TIR-NB-LRR proteins. However, the corollary to this relationship, namely a requirement for ADR1-like proteins by CC-NB-LRR proteins, while intriguing, is not strictly supported by our data. Rather, we were surprised to find that while silencing of *ADR1* or *NRG1* alone had no apparent impact on Rx2-mediated resistance, the combined silencing of *NRG1* and *ADR1* resulted in partially compromised resistance (Figure 4.11). This suggests that while the Rx2 CC-NB-LRR protein does rely to at least some extent on CC_R-NB-LRR proteins, the reliance is not specific to a particular clade. We speculate therefore that while there may be an absolute requirement by TIR-NB-LRR proteins for NRG1-

like proteins, CC-NB-LRR proteins may be able to utilize either NRG1- or ADR1-like proteins as necessary.

It is not yet clear how CC_R-NB-LRR proteins fit into the larger picture of NB-LRR protein recognition and signaling pathways. Nonetheless we have identified multiple unique properties of this class, with regards both to their evolutionary history and N-terminal domains, which will have significant implications for our understanding of NB-LRR evolution and function. While the extent to which CC_R-NB-LRR proteins may be required by other classes of NB-LRR proteins is not yet fully known, our findings indicate that this unique NB-LRR lineage has an important role to play in plant resistance responses.

MATERIALS AND METHODS

Sequence analysis

Database searches were performed between May and November 2010. All searches were performed with either BLASTP or TBLASTN on default settings. An initial cutoff of bit score 200 was used to estimate CC_R-NB-LRR homology in Phytozome proteome searches, and after phylogenetic confirmation additional sequences were added or discarded to ensure the data set contained all CC_R-NB-LRR homologs and one outgroup sequence. Both the papaya *NRG1* and *ADR1* homologs lack sequence coverage in some areas, resulting in segmented and incomplete protein predictions. Genomic sequence was therefore manually analyzed to obtain sequences that are as complete as possible.

CC_R alignment for sequence comparison was performed by ClustalW with Lasergene 8.0 MegAlign software (DNASTAR, Inc.) on default settings. All alignments for phylogenetic analysis were performed in MEGA 4.0 (Tamura et al. 2007) using ClustalW on default settings. Phylogenies were constructed by the Neighbor Joining method (Saitou and Nei 1987) using the PAM matrix (Schhwarz and Dayhoff 1979), and all trees shown are bootstrap consensus trees derived from 1000 replicates.

Cloning and plasmid construction

All new constructs used for protein expression were created by inserting the sequence of interest into the XbaI and BamHI sites of the pBIN61 binary vector with a C-terminal HA tag (Bendahmane et al. 2002). I2, R3a, and RB CC domains were cloned by RT-PCR from tomato (I2) and potato (R3a and RB) cDNA. An ADR1 homolog was identified on potato BAC RH134N13 (GenBank accession AC233501.1) by BLAST search of genomic sequence available from the SOL Genomics Network (SGN) (Mueller et al. 2005), and was used for primer design. Full-length StADR1 was cloned from potato cultivar Eva cDNA (GenBank accession HQ906887), and served as template for construction of StADR1 CC_R. Full-length and truncated NRG1 variants were cloned from 35S-NRG1 (Peart et al. 2005). CC_R domains of Arabidopsis ADR1 (amino acids 1-146), ADR1-L1 (1-155), ADR1-L2 (1-153), NRG1.1 (1-149), and NRG1.2 (1-151) were cloned from *A. thaliana* ecotype Columbia by RT-PCR. Cloning of Rx and Bs2 CC domains (Moffett et al. 2002), Rx NB:eGFP (Rairdan et al. 2008), Rx D460V (Bendahmane et al. 2002), PVX:GFP (Peart et al. 2002a), and 35S:Bax (del Pozo et al. 2004) have been previously described.

To construct pTV:NRG1, a 493bp fragment of NRG1 from a similar position as in the previously described TRV:NRG1 clone (Peart et al. 2005) was cloned from *N. benthamiana* cDNA and inserted into the EcoRI and XhoI sites of the pGIITV vector (Ratcliff et al. 2001). pTV:ADR1 consists of a 543bp fragment of ADR1 cloned from *N. tabacum* inserted into the XmaI and XhoI sites of pGIITV. *Nicotiana* ADR1 primers were designed based on processed tobacco genome sequence c14852 identified through SGN. pTV:NRG1/ADR1 consists of the aforementioned NRG1 and ADR1 fragments, separated by an XhoI site, inserted into the XmaI and EcoRI sites of pGIITV. TV:SGT1 (Peart et al. 2002b) and TV:00 (Ratcliff et al. 2001) have been previously described.

Nucleotide sequences of all primers used in this study are listed in Table 4.3. RNA was extracted using TRIzol reagent (Invitrogen) and cDNA first strand synthesis was performed using SuperScript III reverse transcriptase (Invitrogen). PCR was performed with KOD high-fidelity polymerase (Novagen). All new constructs were sequence verified.

Transient expression and protein analysis

All binary vectors for protein expression were transformed into *A. tumafaciens* strain C58C1, carrying virulence plasmid pCH32. Agroinfiltrations were performed as previously described (Bendahmane et al. 2002) at OD_{600} = 0.2. For PVX:GFP co-expression, *A. tumafaciens* strain GV3101 carrying the PVX:GFP-expressing binary vector pGR208, along with virulence plasmid pSoup (Peart et al. 2002b), was agroinfiltrated at OD_{600} = 0.001. Protein analysis by immunoblotting was performed as previously described (Bhattacharjee et al. 2009), with the exception that HA-tagged proteins were ground in Laemmli buffer (Laemmli 1970). PVX CP was visualized as

Table 4.3 Primers used in this study

Description[a]	Nucleotide Sequence (5' - 3')
Cloning for protein expression	
XbaI-Kozak-I2 F	ATCTATCTAGAGCCACCATGGAGATTGGCTTAGCAGTTGG
BamHI-I2-CC R	AGACGGGATCCTTCGCTCTGCCTACCAAAGATATC
XbaI-Kozak-R3a F	ATCTATCTAGAGCCACCATGGAGATTGGCTTAGCAGTTGG
BamHI-R3a-CC-R	AGACGGGATCCATCATTCTGCCTTCCAAAGATATC
XbaI-Kozak-RB	ATCTAACTAGTGCCACCATGGCTGAAGCTTTCATTCAAGTTC
BamHI-RB-CC R	AGACGGGATCCTTTGTCTCTTCCATAAACCTGCGG
XbaI-Kozak-NRG1 F	CGTATCTAGAACCACCATGGCCGCAACTCTGTTG
XbaI-Kozak-NRG1-13 F	ATCTATCTAGAGCCACCATGGTTTTTGATATCCTACTCAAAGCC
XbaI-Kozak-NRG1-50 F	ATCTATCTAGAGCCACCATGGATGACATCGAGAGGCTAAACAAAG
BamHI-NRG1 R	CGTAGGATCCGATAAACAATGATTTGTGCAG
BamHI-NRG1-225 R	AGACGGGATCCACAGCCAGCAGGAGCAGAAAG
BamHI-NRG1-182 R	AGACGGGATCCAAATCCACTCCCATTTGTGCTACC
BamHI-NRG1-147 R	AGACGGGATCCTCTCAACATGGAGGTAATCTGATCC
XbaI-Kozak-StADR1 F	CGTATCTAGAACCACCATGGCGGTGACGGACTTTTTCGT
BamHI-StADR1 R	CGTAGGATCCCTGAGATAGCCAATCAAGAGTATAGC
BamHI-StADR1-155 R	CGTAGGATCCTCCACCACCACCACCACTT
XbaI-Kozak-ADR1 F	CGTATCTAGAACCACCATGGCTTCGTTCATAGATCTT
BamHI-ADR1-CCR R	CGTAGGATCCATCATTCCGCTCAGTCAACAG
XbaI-Kozak-ADR1L1 F	CGTATCTAGAACCACCATGGCCATCACCGATTTTTTC
BamHI-ADR1L1-CCR R	CGTAGGATCCTCCCCCAATTTTCATGGAACC
XbaI-Kozak-ADR1L2 F	CGTATCTAGAACCACCATGGCAGATATAATCGGCGGC
BamHI-ADR1L2-CCR R	CGTAGGATCCTCCCCTGAGTTTCATAGAACC
XbaI-Kozak-AtNRG1.1 F	CGTATCTAGAACCACCATGAACGATTGGGCTAGTTTG
BamHI-AtNRG1.1-CCR R	CGTAGGATCCACTCAAAAGGTCCATTCTTTTAC
XbaI-Kozak-AtNRG1.2 F	CGTATCTAGAACCACCATGGTCGTGGTCGATTGGCTT
BamHI-AtNRG1.2-CCR R	CGTAGGATCCACCACTCAAACGGTCGAGCTT
VIGS cloning	
EcoRI-NRG1-VIGS F	CGTAGAATTCTCGATGACATCGAGAG
XhoI-NRG1-VIGS R	CGTACTCGAGTCTTTCTCCTCAAGCA
XhoI-NtADR1-VIGS F	CGTACTCGAGTTAGGTGCTGGAATTGAGTTGGG
XmaI-NtADR1-VIGS R	CGTACCCGGGATAATTGAAACCACGCGGAATGG

[a] where no domain or amino acid position is specified, assume full-length for protein expression primers only

previously described (Bhattacharjee et al. 2009) and HA epitope tagged proteins were visualized using a monocolonal anti HA-peroxidase antibody (Sigma). Protein quantification by image analysis was performed using ImageJ (Abramoff et al. 2004), with images converted to an 8-bit grayscale.

Virus-induced gene silencing

Three leaves each of three week old Rx2 transgenic *N. benthamiana* (Bhattacharjee et al. 2009) seedlings were co-infiltrated with *Agrobacterium* carrying plasmids pBINTra6 (TRV RNA1 cDNA) and RNA2 constructs derived from pTV:00 (Ratcliff et al. 2001) at concentrations of OD_{600} = 0.1 and 0.2, respectively. Three weeks later, three to five leaves per plants were dusted with carborundum powder and rub-inoculated with sap containing PVX:GFP virions, prepared as previously described (Bhattacharjee et al. 2009). Systemic tissue was monitored for necrosis for up to 50 days following PVX:GFP inoculation. TV:00, TV:ADR1, and TV:NRG1/ADR1 results were each derived from 26 individuals distributed over three replicates, and TV:NRG1 results were derived from 21 individuals distributed over four replicates respectively.

CHAPTER 5

Conclusion

This body of work had its origins in the widespread expectation that the N-terminal domains of plant nucleotide binding, leucine-rich repeat (NB-LRR) disease resistance (R) proteins are responsible for signal initiation upon pathogen detection. However, detailed investigation of the N-terminal coiled-coil (CC) domain of the potato (*Solanum tuberosum*) Rx NB-LRR protein revealed no signaling activity, suggesting rather that the two primary roles of this domain are mediating an intramolecular interaction with the NB-ARC (Apaf-1, R proteins, and CED-4 homology)-LRR portion of the protein, and binding to the postulated recognition co-factor RanGAP2 ((Rairdan et al. 2008) and Chapter 2). Through analysis of various mutations throughout the Rx CC domain, we have been able to better understand the CC – RanGAP2 interface, and to see how juxtaposed sites of RanGAP2 and NB-ARC-LRR binding offer a possible mechanism for relaying information regarding pathogen – co-factor interactions to the rest of the R protein.

One reason the Rx CC domain was originally suspected to function in signaling was the previously demonstrated ability of the Rx CC-NB fragment to induce a hypersensitive response (HR), typical of R protein-mediated resistance responses, upon over-expression (Bendahmane et al. 2002). Thus it would seem that either the CC or NB domain of Rx is responsible for signaling the induction of defense

responses. Having found no such behavior ascribable to the CC domain, we also examined the NB domain, and discovered that the Rx NB domain is independently capable of inducing an HR upon over-expression, suggesting that it is the NB rather than CC domain of Rx from which resistance signaling originates (Chapter 3). This activity was only apparent when the inherently unstable NB domain was stabilized by fusion to enhanced green florescent protein (EGFP), offering a possible explanation of why such activity has not been previously identified. We found similar activity for the NB domains of the closely related Rx2 and Gpa2 R proteins (Chapter 3).

In exploring the capacity of signal initiation by the Rx NB domain, we identified a structural feature, consisting of two α-helices, that seemed to be most closely involved with NB-mediated signal initiation (Chapter 3). Intriguingly, this region of Rx corresponded precisely to the regions of metazoan Apaf-1 and CED-4 proteins responsible for mediating their homotypic oligomerization upon formation of the cell-death promoting apoptosome (Qi et al. 2010). We therefore investigated whether there might be a connection between NB-LRR signaling and oligomerization, and while we were unable to pinpoint any particular involvement in signal initiation, we discovered a previously unknown ability of NB-LRR NB domains to oligomerize, both homotypically and heterotypically (Chapter 3).

In contemplating the relevance of NB-mediated oligomerization, particularly as a potential mechanism of NB-LRR protein signal initiation, one question that immediately presents itself is: what's the point? That is, if the NB domain is the only region from which defense signaling originates (at least in the case of Rx), what would be the purpose of oligomerizing with other similar proteins? How could downstream signaling components be engaged if NB-LRR signaling components interact exclusively with one another? Such an NB-mediated oligomerization-based scenario

would suggest that either an interacting structure is able to activate downstream signaling whereas monomeric NB-LRR proteins are not, or that through its NB domain Rx might interact with the NB domain of some other, different protein able to signal in a manner different from canonical NB-LRR proteins. Regarding this latter possibility, TIR-NB-LRR proteins immediately come to mind, as TIR domains, unlike CC domains, have been found capable of inducing defense responses in a number of cases (Frost et al. 2004; Swiderski et al. 2009; Krasileva et al. 2010). One might hypothesize that upon activation CC-NB-LRR proteins oligomerize with TIR-NB-LRR proteins and thus initiate signaling via TIR domains. This mechanism, however, is made highly unlikely by the absence of *TIR-NB-LRR* genes from the genomes of all monocot and some dicot species, contrasted with the abundance and functionality of CC-NB-LRR proteins in these species (Meyers et al. 1999; Bai et al. 2002; Tarr and Alexander 2009) (Chapter 4). As an alternate to TIR-NB-LRR proteins, however, canonical CC-NB-LRR proteins might be hypothesized to oligomerize with CC_R-NB-LRR proteins, the unique N-terminal CC_R domains of which we have found to be capable of signal initiation (Chapter 4). Indeed, the NB domains of Rx and a number of other NB-LRR proteins are able to interact with those of CC_R-NB-LRR proteins (Chapter 3).

As noted above, we have herein characterized the unique and highly conserved CC_R-NB-LRR clade of CC-NB-LRR proteins. CC_R-NB-LRR-encoding genes can be further categorized into NRG1-like and ADR1-like sub-clades (Chapter 4). Among the most notable features of CC_R-NB-LRR proteins is the ability of CC_R domains to induce defense responses upon over-expression, as well as the striking correlation in occurrence between *TIR-NB-LRR* genes and NRG1-encoding genes across plant genomes (Chapter 4). Coupled with the previously described requirement by the TIR-

NB-LRR protein N for NRG1 (Peart et al. 2005), the founding member of the NRG1-like class, this strongly suggests a conserved and widespread functional relationship between TIR-NB-LRR and NRG1-like proteins. We have further identified a similar, although possibly less complete, requirement by canonical CC-NB-LRR proteins for CC_R-NB-LRR proteins.

Whether the functional requirement between canonical NB-LRR proteins and CC_R-NB-LRR proteins is manifested through oligomerization remains to be determined. Although we have identified a general ability of NB domains to oligomerize, there was no apparent specificity in interactions between NB domains derived from CC-NB-LRR, TIR-NB-LRR, NRG1-like, and ADR1-like proteins (Chapter 3). However, this behavior is unlikely reflective of the behavior of full-length proteins, as no oligomerization is evident between full-length Rx molecules (Moffett et al. 2002), in contrast to the oligomerization observed between Rx NB domains (Chapter 3). It would thus seem that oligomerization between full-length proteins, if present, is a more selective process than that observed between isolated NB domains. At this time, an apparent recalcitrance of CC_R-NB-LRR proteins to extraction precludes investigation of full-length oligomerization, despite use of the same methodology as originally described for the extraction of NRG1 (Peart et al. 2005). Oligomerization between N and NRG1 has previously been tested and not observed to occur (Mestre and Baulcombe 2006), and the extension of such studies to include additional NB-LRR proteins and sub-classes would be greatly informative.

An additional point of interest is the relationship between signaling activity and SGT1 dependence for various NB-LRR protein classes. Many NB-LRR proteins require SGT1 for proper folding and accumulation, presumably through interaction with the LRR domain (Bieri et al. 2004; Leister et al. 2005; Zhang et al. 2008; Kadota et

al. 2010). We have found that while Rx NB-mediated HR induction retains SGT1 dependence (Chapter 3), signaling mediated by CC_R domains is SGT1 independent (Chapter 4). This is in contrast to the dependence of full-length NRG1 activity on SGT1. Thus it would seem that while some component downstream of the Rx NB domain retains SGT1 dependence, CC_R-mediated signaling resides fully downstream of SGT1.

Were canonical NB-LRR proteins such as Rx to signal directly via CC_R-NB-LRR proteins, the dependence of Rx NB-mediated signaling on SGT1 might be explained by involvement of SGT1 in the LRR folding of CC_R-NB-LRR protein(s) downstream of Rx signaling. Alternately, CC_R-NB-LRR proteins may serve simply to amplify NB-LRR-mediated signaling, rather having direct involvement, and the SGT1-dependence of Rx NB-mediated HR may have an alternate explanation. Under either of these scenarios, signal transduction from canonical NB-LRR proteins to CC_R-NB-LRR proteins could take place via NB-mediated oligomerization, or via an alternate mechanism yet to be determined.

APPENDIX A

The Rx NBLet has no detectable effect on Rx-mediated resistance[*]

ABSTRACT

Genes encoding plant nucleotide binding, leucine-rich repeat (NB-LRR) proteins are often found to have small regions of duplicated sequence appearing in nearby genomic locations. Here, we investigate whether such a duplicated element is functionally relevant to the potato NB-LRR-encoding gene *Rx*.

[*] This data originally appeared in: Mazourek, M., Cirulli, E.T., Collier, S.M., Landry, L.G., Kang, B-C., Quirin, E.A., Bradeen, J.M., Moffett, P., and Jahn, M.M. (2009) The fractionated orthology of *Bs2* and *Rx*/*Gpa2* supports shared synteny of disease resistance in the Solanaceae. *Genetics* 182: 1351-1364. All data presented in this appendix was generated by S.M.C.

INTRODUCTION

Plant nucleotide binding, leucine-rich repeat (NB-LRR) proteins often occur in tandemly duplicated clusters (Meyers et al. 1998; Grube et al. 2000; Mazourek et al. 2009). In some cases, small truncated pieces of NB-LRR-encoding sequence, frequently from the 5′ end of the gene, are found to be duplicated up- or downstream of their associated NB-LRR-encoding gene. Such "NBLets" have been found to exist in proximity to the *Bs2*, *Rx*, *Mi 1.2*, *R1*, *RB*, and *Tm2*^{2} NB-LRR genes (Lanfermeijer et al. 2003; Mazourek et al. 2009). Tandem duplications have been implicated in affecting gene expression through the generation of small interfering RNAs, and the *RPP5* NB-LRR locus of *Arabidopsis thaliana* has been demonstrated to be sensitive to endogenous RNA silencing (Yi and Richards 2007). Interestingly, in contrast to *Rx*, the *Gpa2* NB-LRR-encoding gene which resides in the same potato (*Solanum tuberosum*) gene cluster as *Rx* (van der Vossen et al. 2000) is not accompanied by any detectable NBLet feature (Mazourek et al. 2009). As the resistance mediated by Rx is markedly stronger than that of Gpa2, despite the two proteins sharing over 88% amino acid identity (Bendahmane et al. 1999; van der Vossen et al. 2000; Sacco et al. 2009), we were curious whether the presence or absence of NBLets might impact NB-LRR efficacy. We therefore investigated whether the duplicated region of *Rx* has any bearing on Rx-mediated resistance.

RESULTS AND DISCUSSION

To determine whether the *Rx* NBLet plays a functional role in Rx-mediated resistance to *Potato Virus X* (PVX), we compared Rx protein accumulation and Rx-

mediated resistance when *Rx* was expressed from its native genomic context with and without its associated NBLet. A 300bp region encompassing the Rx 3′ NBLet duplication was deleted from a binary vector containing the *Rx* promoter, coding region fused to four HA epitope tags, and 3′ sequence. These *Rx* variants were co-expressed with a version of PVX fused to green florescent protein (PVX:GFP) in *Nicotiana benthamiana* leaves, such that Rx efficacy is inversely correlated with the amount of GFP florescence observed. Deletion of the *Rx* NBLet had no notable effect on the ability of *Rx* to confer resistance to PVX or on the level of Rx protein expressed (Figure A.1). These results rule out the possibility that the *Rx* NBLet expresses a protein fragment required for Rx function, and suggests that the NBLet does not alter the level of Rx protein expression or accumulation. Thus the difference in resistance mediated by Rx and Gpa2 is likely not related to the presence or absence of NBLet features. However, while *Rx* transcription, mRNA stability, translation, and Rx protein function have no apparent dependence on the Rx NBLet, it remains possible that NBLets may affect chromatin structure an endogenous context, which could in turn affect NB-LRR expression (Friedman and Baker 2007).

MATERIALS AND METHODS

Rx followed by four HA epitope tags (Rx:4HA) was constructed in the pB1 binary vector containing the *Rx* promoter and 3′ sequence as described (Bendahmane et al. 2002; Peart et al. 2002b). The NBLet sequence was deleted by overlapping PCR to create RX:4HAΔNBLet. Preliminary cloning was performed by G. Rairdan. Binary vectors were transformed into the *Agrobacterium tumefaciens* strain C58C1 carrying the virulence plasmid pCH32. PVX:GFP was expressed from the pGR208 binary vector in

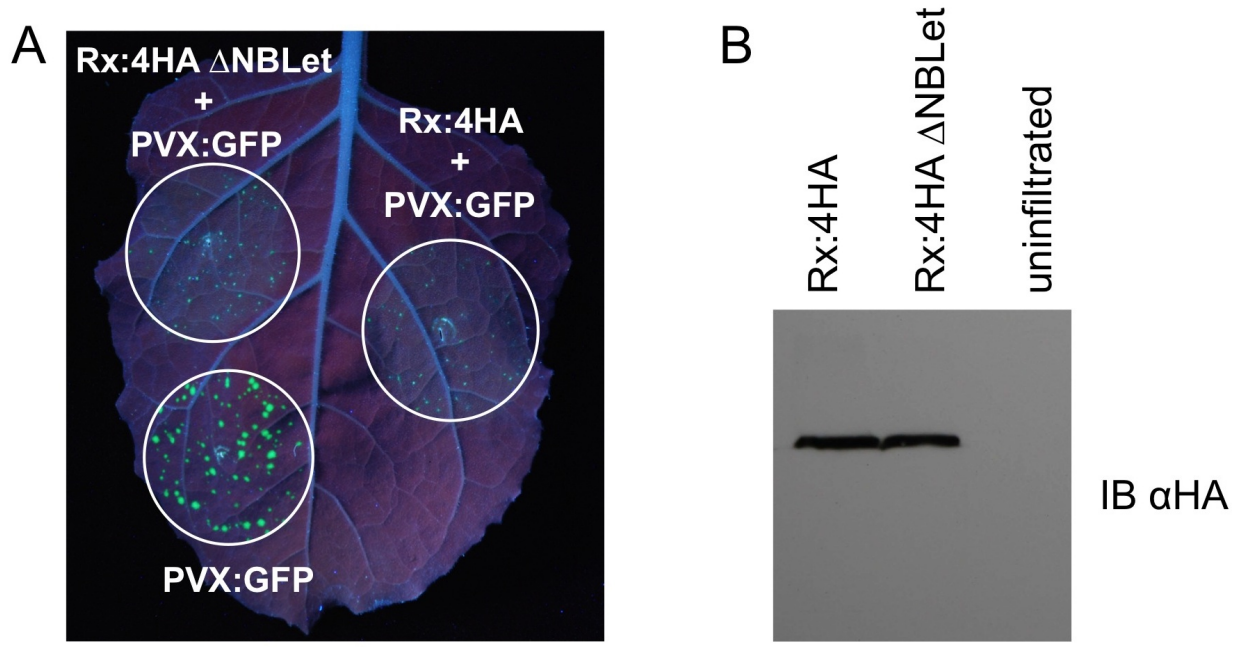


Figure A.1 *Effect of the Rx NBLet on Rx function.* **A,** PVX resistance conferred by Rx in its complete genomic context (Rx:4HA) and without the Rx NBLet (Rx:4HA ΔNBLet). Rx constructs were co-expressed in *Nicotiana benthamiana* leaves with an infectious PVX:GFP clone via agroinfiltration. PVX:GFP accumulation was monitored as GFP fluorescence under UV light five days after infiltration. PVX:GFP infiltrated in the absence of Rx is shown at bottom left for comparison. **B,** Immunoblot analysis showing Rx accumulation when expressed with and without NBLet. *N. benthamiana* leaves were infiltrated with *Agrobacterium* containing the indicated *Rx* constructs, and samples were collected two days later. Rx protein levels were detected by anti-HA immunoblotting.

A. tumefaciens strain GV3101 carrying the virulence plasmid pSoup (Peart et al. 2002b). Agroinfiltration was performed as previously described (Bendahmane et al. 2002) at OD_{600}=0.2 for pB1 constructs and OD_{600}=0.001 for PVX:GFP. GFP florescence was evaluated five days later using a hand-held UV lamp. Protein extraction and immunoblotting were performed essentially as described (Rairdan and Moffett 2006).

APPENDIX B

Effect of mutations within the PVX coat protein on Rx2-mediated resistance*

ABSTRACT

Amino acids occupying positions 121 and 127 of the *Potato Virus X* (PVX) coat protein (CP) have previously been found to be critical in determining whether PVX is recognized by the Rx resistance protein of potato. A series of mutations at these positions have previously been tested for their effect on Rx-mediated resistance in *N. benthamiana*. However, in this system Rx-mediated resistance was found to be exceptionally robust, possibly masking the effects of minor alterations. Here, we have tested a subset of CP mutants against resistance mediated by the weaker Rx2 resistance protein, in an effort to obtain a more detailed understanding of the impact of these mutations.

* All data presented in this appendix was generated by S.M.C.

INTRODUCTION

The amino acids present at positions 121 and 127 of the *Potato Virus X* (PVX) coat protein (CP) have previously been found to be of particular importance in determining whether PVX is recognized by the potato (*Solanum tuberosum*) Rx NB-LRR (nucleotide binding, leucine-rich repeat) protein, or whether the virus is able to evade detection. The avirulent CP4 strain of PVX possesses a threonine and lysine at positions 121 and 127, respectively, while the same positions in the CP of the virulent HB strain are occupied by lysine and arginine, respectively (Moreira and Jones 1980; Querci et al. 1993; Bendahmane et al. 1999). Baurés et al. (2008) recently examined the roles of residues at positions 121 and 127, utilizing a series of mutations at these locations. In that experiment, *Nicotiana benthamiana* plants carrying *Rx* as a transgene were infected with CP-mutant PVX variants, and the relative success of infection was recorded. All variants with lysine at position 127 (K127) were unable to mount an infection, having been controlled by Rx-mediated extreme resistance (ER), wherein the pathogen is eliminated before the onset of a visible hypersensitive response (HR) – the form of programmed cell death which often typifies NB-LRR-mediated resistance. Baurés et al. (2008) further found that a PVX variant with arginine at both positions 121 and 127 (PVX_{RR}) of the CP, while also unable to mount a successful infection, did cause HR lesions in infected tissue, indicating reduced efficiency of recognition by Rx. A PVX_{PR} variant was able to induce HR lesions in both infected and systemic tissue, indicating limited and temporary escape of detection by Rx, and the resistance-breaking HB strain (PVX_{KR}) produced trailing necrosis throughout systemic tissue. Such trailing necrosis symptoms indicate that although the virus is able to evade initial

detection and spread throughout the plant tissue, latent recognition does occur, resulting in widespread necrosis due to HR induction.

It is interesting to note that the HB strain resulted in trailing necrosis in *Rx N. benthamiana*, as it was originally identified as being completely resistance-breaking (infection without necrosis) in *Rx* potato cultivars (Moreira and Jones 1980). This suggests that Rx-mediated resistance might be somewhat more efficient in PVX_{KR} detection in *N. benthamiana* as compared to potato. Thus, while the investigations of Baurés et al. (2008) suggest that in the presence of R127 R121, P121, and K121 provide successively increasing levels of Rx evasion, the heightened efficiency of Rx in this system may also have masked subtler effects of various substitutions at each position. We therefore examined a subset of CP mutations in *N. benthamiana* carrying *Rx2* as a transgene. Similar to Rx, Rx2 confers PVX resistance through recognition of the CP. However, Rx2-mediated resistance is slightly weaker than that of Rx, and while *Rx* transgenic *N. benthamiana* confer ER to avirulent PVX, control of PVX spread by *Rx2* transgenic *N. benthamiana* involves the formation of HR lesions (Bendahmane et al. 1999; Bendahmane et al. 2000; Bhattacharjee et al. 2009). We reasoned that the weaker resistance mediated by Rx2 might provide enhanced resolution regarding the impacts of CP mutations.

RESULTS AND DISCUSSION

We challenged wild-type, *Rx*-transgenic, and *Rx2*-transgenic *N. benthamiana* plants with PVX variants with position 121/127 combinations TK, SK, KK, VR, SR, RR, PR, and KR. All variants were generally capable of systemically infecting wild-type plants, as expected, although the CP4 strain (PVX_{TK}) was somewhat unusual in

sporadically inducing minor lesions on wild-type leaves (Table B.1). PVX_{PR} was able to induce limited systemic necrosis in *Rx* plants (Table B.1, Figure B.1), confirming its ability to partially overcome Rx-mediated resistance.

The effective evasion of resistance by PVX_{PR} was even more evident in *Rx2* plants, where in all trials inoculation with PVX_{PR} caused no HR lesions yet led to systemic infection with no accompanying necrosis (Table B.1, Figures B.1A and B.1B). Thus PVX_{PR} is able to fully evade Rx2-mediated resistance. PVX_{RR} inoculation led to uniquely mild HR lesions on inoculated leaves of *Rx2* plants (Figure B.1A), and in half of trials was able to escape Rx2-mediated resistance and infect systemic tissue with no necrotic symptoms. In the other half of trials, PVX_{RR} caused discrete necrotic lesions in systemic tissue, indicating a moderate degree of detection by Rx2. PVX_{VR} was similar to PVX_{RR} in that it was sometimes able to infect systemic tissue without necrosis, although in other cases PVX_{VR} induced systemic necrosis or failed to infect systemic tissue at all. PVX_{KK} was consistently able to escape Rx2-mediated resistance to infect systemic tissue, but systemic infection always resulted in trailing necrosis. PVX_{SR} was able to escape Rx2-mediated resistance and spread systemically in only one out of six trials (Table B.1). VR, KK, and SR PVX variants all typically induced strong HR lesions in inoculated leaves (Figure B.1A). PVX_{TK} and PVX_{SK} were consistently unable to infect *Rx2* plants (Figure B.1A and B.1B). While inoculation with PVX_{SK} was consistently met with ER, PVX_{TK} inoculation induced HR lesions in the inoculated leaves in two of six trials (Table B.1).

Our results are generally in agreement with those of Baurés et al. (2008), while at the same time providing greater resolution of the effects of substitution at positions 121 and 127. The observation that PVX_{SK} was consistently met with ER responses in *Rx2*-transgenic *N. benthamiana*, to a greater extent even than the CP4 strain, suggests

Table B.1 Summary of PVX CP mutant inoculations in *N. benthamiana*

Rep. / AA[a]	HR lesions, inoculated leaves[b]													PVX infection symptoms, systemic leaves[c]										
	wild-type				Rx2						Rx			wild-type				Rx2						Rx		
	A	B	C	D	A	B	C	D	E	F	B	C	D	A	B	C	D	A	B	C	D	E	F	B	C	D
TK	(+)	-	(-)	(-)	-	-	+	+	-	-	-	-	-	I	I	I	I	-	-	-	-	-	-	-	-	-
SK	-	-	-	-		-	-	-	-	-	-	-	-	I	-	I	I		-	-	-	-	-	-	-	-
KK	-	-	-	-	+	+	-	+	+	(+)	-	-	-	I	I	I	I	N	N	N	N	N	N	-	-	-
VR	-	-	-	-	+	+	+	+	+	+	-	-	+	I	I	I	I	N	-	-	I	I	I	-	-	-
SR	-	-	-	-	+	+	+	+	+	+	-	-	-	I	I	I	I	-	-	-	-	I	-	-	-	-
RR	-	-	-	-	+	(+)	(+)	(+)	(+)	(+)	+	+	+	I	I	I	I	I	I	I	I/N	I/N	I/N	-	-	-
PR	-	-	-	-	-	-	-	-	-	-	+	+	+	I	I	I	I	I	I	I	I	I	I	I/N	-	N
Θ	-	-				-		-	-		-			-	-				-		-	-		-		

[a] Amino acids at positions 121 and 127, respectively (TK = CP4, Θ = sap prepared from mock-inoculated plants)
[b] Response of various plant genotypes in inoculated leaves 4 days after inoculation; + = clear necrotic HR lesions, (+) = sizeable lesions with mild necrosis, (-) = minor lesions, - = no lesions, blank = no data
[c] Response of various plant genotypes in systemic leaves, 10-15 days after inoculation; - = no infection, I = systemic infection, I/N = systemic infection with minor necrotic flecks, N = systemic necrosis, blank = no data

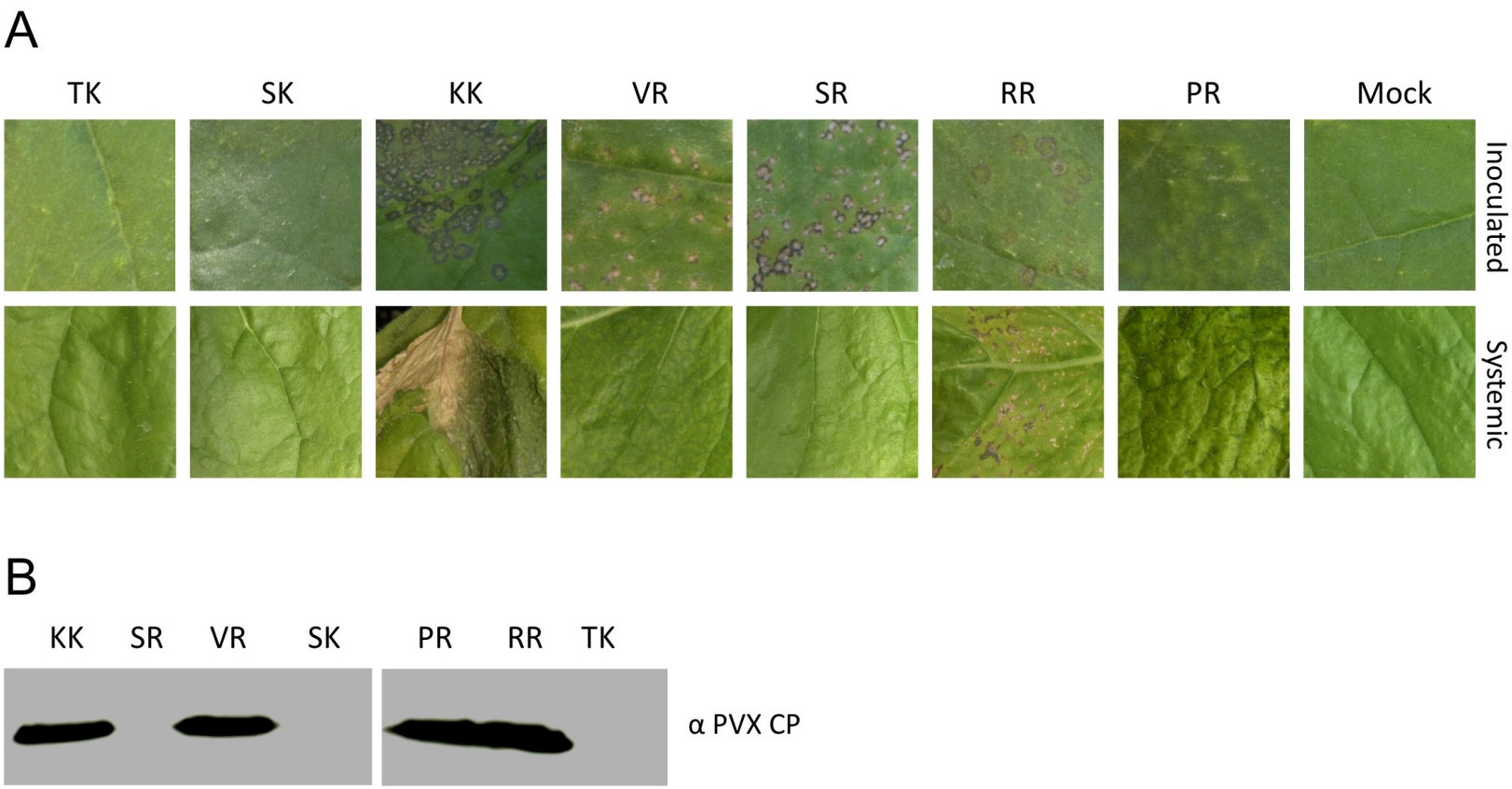


Figure B.1 *Response of* Rx2-*transgenic* N. benthamiana *to inoculation with PVX CP mutant variants.* **A,** Representative photographs of inoculated (top panels) and systemic (bottom panels) leaves of *Rx2*-transgenic *N. benthamiana* 4 and 10 days, respectively, following inoculation with the indicated PVX CP 121/127 variants. **B,** Representative anti-(α)PVX CP immunoblot of protein samples collected from systemic tissue of *Rx2 N. benthamiana* 13 days after inoculation with the indicated PVX CP 121/127 variants.

that S121 enhances recognition by Rx2. This is in accordance with the finding that of R127 variants, PVX_{SR} was least able to systemically infect *Rx2* plants (Table B.1). The partial resistance-breaking ability of PVX_{KK} confirms that K121 does indeed confer a heightened ability to evade detection. These results also indicate that K127 does not automatically rule out the ability to break resistance, contrary to previous findings with Rx (Baurés et al. 2008).

MATERIALS AND METHODS

All PVX variants used in this study have been previously described (Baurés et al. 2008). *Rx*- and *Rx2*-transgenic *N. benthamiana* lines have been described previously (Lu et al. 2003; Bhattacharjee et al. 2009). For virus inoculation, three leaves each of four-week-old *N. benthamiana* plants were dusted with carborundum powder and rub-inoculated with sap containing PVX virions, prepared as previously described (Bhattacharjee et al. 2009). Symptoms on inoculated leaves were recorded 4 days after inoculation, and infection of systemic tissue was monitored 10-15 days after inoculation. Protein extraction and immunoblotting was performed as previously described (Bhattacharjee et al. 2009).

APPENDIX C

A yeast two-hybrid screen for proteins interacting with the Rx NB domain*

ABSTRACT

Having identified the Rx nucleotide binding (NB) domain as the probable source of Rx resistance signal initiation, we undertook a yeast two-hybrid screen for proteins able to interact with the Rx NB domain, hypothesizing that such interacting proteins may represent downstream signaling components. This process involved the construction of an Rx NB bait fusion protein, a screen for interacting proteins, identification of candidate interactors, confirmation of interaction and functional assessment of candidate interactors. While bait construction was successful, screening resulted in only four likely interaction candidates, and while some were highly intriguing given their previously established biological functions, none could be confirmed as interacting with Rx *in planta*. Additional functional analysis of candidate interactors, relying primarily on virus-induced gene silencing, returned consistently ambiguous results, and the project was eventually discontinued. The intent of this appendix is therefore not to present conclusive or detailed results, but rather to provide a succinct narrative of this lengthy yet ultimately inconclusive undertaking, for the primary purpose of informing any future investigations into the same or a similar subject.

* All data presented in this appendix was generated by S.M.C.

INTRODUCTION

We have previously identified the nucleotide binding (NB) domain of the Rx NB-LRR (leucine-rich repeat) protein as being sufficient for the induction of defense response signaling. We found that when stabilized by fusion to enhanced green florescent protein (EGFP), over-expression of the Rx NB fusion protein induces a hypersensitive response (HR) indicative of resistance signaling (Chapter 3). Over-expression of the isolated NB domain presumably relieves negative regulation such that the signaling region(s) is exposed independent of whole-protein activation upon pathogen detection.

NB-LRR protein NB domains participate in nucleotide binding and hydrolysis, which are critical for overall NB-LRR protein function (Tameling et al. 2002). However, we have found that signal initiation by the Rx NB domain is independent of its nucleotide binding function (Chapter 3). Thus the means by which the Rx NB domain carries out its presumed role in signaling is currently unknown. We reasoned that proteins able to interact with the active Rx NB domain represent possible downstream signaling factors, and so undertook a yeast two-hybrid (Y2H) screen for Rx NB-interacting proteins. While full-length Rx does not lend itself to such interaction screens because it must first be elicited by recognition of its cognate pathogen avirulence (Avr) protein in order to adopt a signaling conformation, formats containing only the NB domain presumably allow the signaling region to be exposed or available in the absence of Avr elicitation. Additionally, while full-length Rx may undergo multiple protein – protein interactions (for example between its CC domain and RanGAP2, see Chapter 2), the NB minimal signaling region provides an ideal tool for exclusively isolating Rx signaling interactors.

RESULTS

Construction of a signaling-competent bait fusion protein

As a first step towards performing an interaction screen with the Rx NB domain, an NB-containing bait construct was engineered in the pEG202 vector of the LexA Y2H system (Gyuris et al. 1993). The complete bait protein consisted of the Rx NB domain fused N-terminally to the LexA DNA binding domain and C-terminally to EGFP and an HA epitope tag (LexA:NB:EGFP:HA). EGFP was included to provide stability, as a simple LexA:NB fusion was not stable when expressed *in planta* from the pBIN61 binary vector (Figure C.1A). The LexA:NB:EGFP:HA fusion protein, however, behaves similarly to Rx NB:EGFP:HA in HR induction and *in planta* protein accumulation when expressed from the 35S promoter-drive pBIN61 vector (Figure C.1A).

Bait construct quality control and Y2H screening was performed essentially as described (Golemis et al. 1999). To confirm proper function of LexA:NB:EGFP:HA in yeast cells, the pEG202 NB:EGFP:HA (pEG202-bait) was co-transformed into yeast (*Saccharomyces cerevisiae*) strain EGY48 with either pSH18-34 or pJK101 reporter plasmids. Six independent pEG202-bait + pSH18-34 colonies were grown in overnight culture and subjected to protein extraction and immunoblotting to confirm stability of the LexA:NB:EGFP:HA fusion protein in yeast cells (Figure C.1B). To confirm that LexA:NB:EGFP:HA does not independently activate the *lacZ* reporter gene, pSH18-34 + pEG202-bait, pSH18-34 + pSH17-4 (activation positive control), and pSH18-34 + pRFHM1 (activation negative control) colonies were streaked onto appropriate growth media, subsequently lifted onto filter paper, and treated with an Xgal solution.

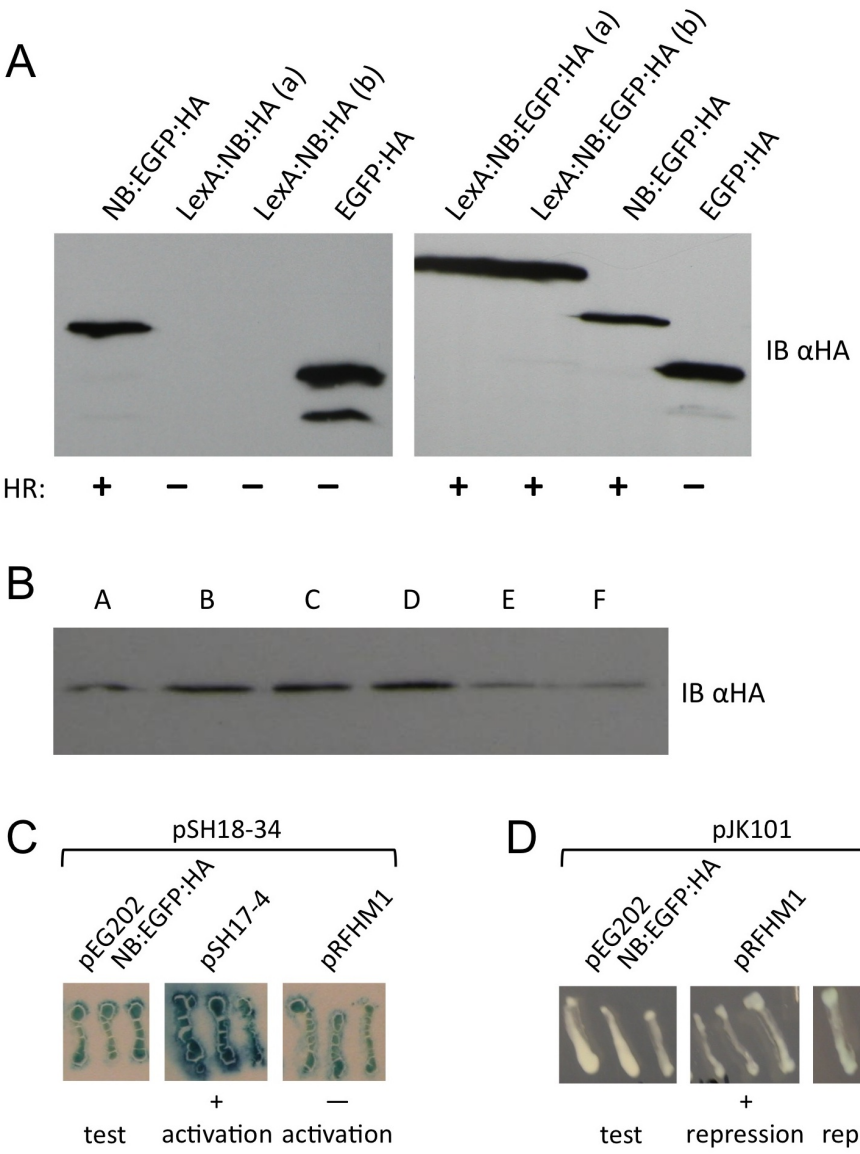


Figure C.1 *Characterization of bait protein activity.* **A,** Expression of bait proteins *in planta*. LexA:NB:HA and LexA:NB:EGFP:HA bait proteins were expressed from pBIN61 in *N. benthamiana* leaves. Protein samples were collected 36 hours later and subjected to anti-(α)HA immunoblotting (top panel). HR induction (+/-) was recorded after 2-3 days (bottom panel). **B,** Bait protein expression and stability in yeast cells. Six replicate EGY48 colonies carrying pSH18-34 and pEG202-bait plasmids were grown overnight and subjected to protein extraction and anti-HA immunoblotting. **C,** *LacZ* activation assay. Yeast colonies transformed with the indicated plasmids were lifted from a master plate onto filter paper, treated with an Xgal solution, and photographed one day later. **D,** Repression assay. Yeast colonies carrying the indicated plasmids were streaked onto growth media containing Xgal, incubated at 28°, and photographed approximately 20 hours later.

β-galactosidase activity was monitored after 20 minutes, 100 minutes, and 1 day (Figure C.1C). Leucine requirement of EGY48 pEG202-bait + pSH18-34 was also confirmed (data not shown). To confirm that the LexA:NB:EGFP:HA fusion protein is capable of binding *LexA* operator sequences, a repression assay was performed with colonies transformed with pJK101 + pEG202-bait, pJK101 + pRFHM1 (repression positive control), or pJK101 alone (repression negative control). Colonies were streaked onto appropriate growth media containing Xgal and monitored for β-galactosidase activity (Figure C.1D).

Library screening

EGY48 carrying pEG202-bait and pSH18-34 was transformed as previously described (Yamada et al. 1998) with a cDNA library constructed from a mixture of RNA derived from mock- and hrcC mutant *Pseudomonas syringae*-inocultated tomato (*Solanum lycopersicum*) tissue, contained in the pYESTrp2 prey vector. Prey proteins are thus expressed from the galactose-inducible *GAL1* promoter as fusions to the B42 activation domain. Library transformation yielded approximately 1.5×10^6 primary transformants, which were subsequently handled as previously described (Gyuris et al. 1993). The library was initially screened for leucine independence, yielding 160 colonies, 89 of which additionally displayed *lacZ* activation. DNA was extracted from colonies displaying both leucine independence and *lacZ* activation and transformed into KC8 *Escherichia coli* for subsequent sequencing of pYESTrp2 inserts. Sequences obtained from 44 colonies that had displayed positive interaction in both assays belonged to 12 unique sources (Table C.1). Five of these not known to commonly appear as false positives in Y2H screens were re-transformed into yeast to confirm

their leucine independence and *lacZ* activation. While all five retained leucine-independence, Malonyl CoA and VTC2 homologs failed to activate *lacZ* and were therefore excluded from further consideration, while Bax Inhibitor 1 (BI-1), "BUK" (encoding an unknown protein having received numerous hits, hence "Big UnKnown"), and "LUK" (encoding an unknown protein having a smaller number of hits, hence "Little UnKnown") displayed interaction in both re-transformation assays (Table C.1, Figure C.2).

Table C.1 Rx NB-interacting proteins from Y2H screen

Sequence homology[a]	**# Hits**[b]	**-leu**[c]	***lacZ***[d]
20S Proteasome α6	11	nd	nd
Bax Inhibitor 1 (BI-1)	1	+	+
"Big Unknown" (BUK)	15	+	+
Chlorophyll a-b Binding Protein	1	nd	nd
"Little Unknown" (LUK)	6	+	+
Malonyl CoA	1	+	-
Ribosomal Protein S15A	2	nd	nd
Ribosomal Protein S27	1	nd	nd
Unknown Protein, Solanaceae	1	nd	nd
Unknown Protein, Plantae	1	nd	nd
Pyrabactin Resistance 1 (PYR1)	3	nd	nd
Vitamin C Deficient 2 (VTC2)	1	+	-

[a] closest sequence homology as determined by BLAST
[b] number of isolations from independent positive interaction colonies
[c] leucine independence (+/-), nd = no data
[d] *lacZ* activation by β-galactosidase assay (+/-), nd = no data

Sequence analysis of interaction candidates

BI-1, BUK, LUK, as well as PYR1 candidate interactors received further analysis. The 15 independent hits to BUK included at least four distinct cDNA fragments, comprising a total of 602-663bp, with variable regions at both 5′ and 3′

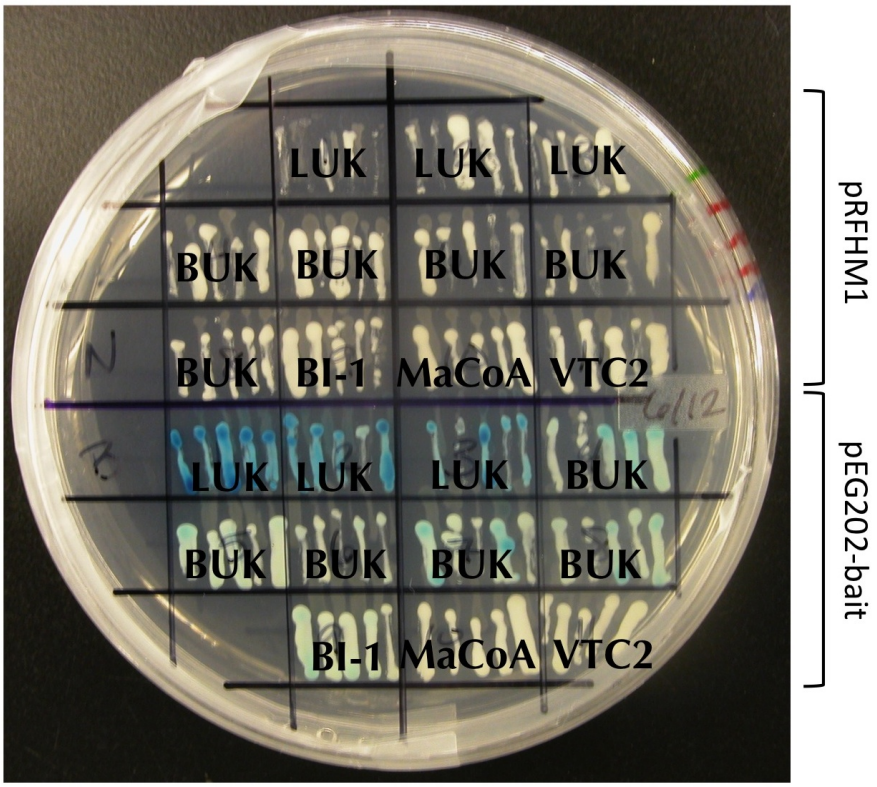


Figure C.2 *Re-transformation test of interaction candidates.* pYESTrp2 cDNA library plasmids derived from independent colonies of potential interacting proteins (3x LUK, 5x BUK, 1x BI-1, 1x Malonly CoA, 1x VTC2) were transformed into EGY48 + pSH18-34 additionally containing either pEG202-bait (test) or pRFHM1 (negative control) as indicated, and subsequently streaked onto Xgal growth media for assaying *lacZ* activation (blue).

ends, encoding a 145 amino acid in its open reading frame, and with a possible true start codon 168bp 3′ of the longest insertion's 5′ end. BLASTN of the TIGR plant transcript assembly database revealed a perfect match with tomato TA49378_4081, for which no function has been assigned. No homologous sequence was identified in *Nicotiana* or *Arabidopsis* species. 5′ and 3′ RACE of the *BUK* gene from tomato confirmed the internal start codon and indicated that the protein it encodes would be only 89 amino acids in length.

The six independent Y2H hits for *LUK* contained four distinct cDNAs with total sequence length of 598-812bp, including some variable 3′ regions, before a polyA tail. However, the open reading frame only encodes 67 amino acids following the LexA fusion, after which no open reading frame of significant length is evident before the polyA. Although BLASTP and TBLASTX searches revealed no close protein matches for LUK, BLASTN of the TIGR plant transcript assembly database revealed a number of close nucleotide matches between *LUK* and tomato and potato (*S. tuberosum*) transcripts (Table C.2), although none of these regions of sequence similarity exceeded 200bp. Nucleotide matches to the first ~150bp of LUK were also identified from a number of non-solanaceous species of various functions. It is possible therefore that the 5′ portion of *LUK* represents some sort of regulatory or repeated element rather than coding sequence, and it is unclear if any significant coding sequence exists in the 3′ portion of *LUK*.

The three positive interacting colonies containing *PYR1* prey sequence all appear to be derived from the same cDNA insert corresponding to TIGR tomato

Table C.2 Top returns from LUK BLASTN search

Species	Plant TA Accession	Percent Identity	Match Length[a]	E-value	5′ *LUK*[b]	# Rep[c]	Annotation
Solanum tuberosum	CN462174	73	170	3.7E-09	N	1	none
Solanum lycopersicum	BP901764	69	151	5.2e-06	Y	1	OSJNBb0046P18.2 protein related cluster
Bruguiera gymnorrhiza	DB995465	73	71	0.00014	Y/N	2	none
Solanum tuberosum	CN465398	71	142	0.00018	N	1	none
Solanum lycopersicum	TA56259_4081	72	131	0.00023	Y	1	none
Solanum lycopersicum	BI206811	74	124	0.00038	Y	1	Uncharacterized protein specific for M.kandleri, MK-14 family related cluster
Citrus reshni	CX306246	70	146	0.00062	Y	1	Hydroxyproline-rich glycoprotein dz-hrgp related cluster
Gossypium arboreum	AW726892	69	130	0.0013	Y	3	CG2839-PA related cluster
Solanum pennellii	AW618075	70	160	0.0014	N	1	Hypothetical protein related cluster
Brassica rapa	CV544506	67	139	0.0022	Y	1	ERD1 protein, chloroplast precursor related cluster
Beta vulgaris	EG550057	69	143	0.0036	Y	1	none
Citrus reshni	DY306470	70	142	0.0041	Y	5	none
Glycine max	TA67443_3847	67	138	0.0043	Y	1	F28L1.12 *A. thaliana* ribosomal protein L34e superfamily protein
Vitis vinifera	TA45204_29760	72	131	0.0048	Y	1	none

[a] match with highest bit score where multiple matches present
[b] Y = match region corresponds to *LUK* nucleotides ~1-150, N = match is elsewhere in *LUK* sequence
[c] number of distinct regions of identity to *LUK* in subject sequence

transcript assembly TA55271_4081, which encodes a protein homologous to the *A. thaliana* abscisic acid (ABA)-interacting protein PYR1 (At4g17870). Tomato and *A. thaliana* homologs share 75% sequence identity at the amino acid level. Notably, sequencing of the pYESTrp2-*PYR1* plasmid obtained from our Y2H screen revealed a single nucleotide deletion at position 385, resulting in an immediate premature stop codon and presumably resulting in a protein of 128 instead of the full 228 amino acids predicted for the tomato transcript.

The single hit to BI-1 consisted of a cDNA of 902bp plus a polyA tail. This sequence corresponded to the C-terminal 224 (out of 247) amino acids of tomato BI-1 (GenBank accession AY380778), but was presented out of frame. In animals, BI-1 is an inhibitor of Bax-induced cell death. BI-1 homologs are also found in plants, despite the apparent absence of Bax homologs in plants (Watanabe and Lam 2004).

Protein analysis of interaction candidates

Candidate interactor sequences were cloned from pYESTrp2 and inserted into pBIN61 with a C-terminal FLAG epitope tag for examining protein expression *in planta*. BUK was clones such that it would begin at its "internal" start codon as confirmed by RACE (see above), BI-1 was inserted in-frame in both its truncated and full-length form (obtained from tomato cDNA rather than the pYESTrp2 prey plasmid), and PYR1 was left truncated as in the Y2H prey construct (PYR1-tr), and also cloned in its full-length form from *Nicotiana benthamiana* (NbPYR1) and tomato (SlPYR1). FLAG-tagged candidate proteins were expressed via *Agrobacterium tumefaciens* (agroexpression) in *N. benthamiana* plants, from which protein samples were collected and subjected to immunoblotting. BUK showed little if any protein accumulation, while truncated BI-1 (BI-1-tr) was abundant compared to full-length BI-

1 (BI-1-FL), which we were unable to detect by immunoblot (Figure C.3A). PYR1:FLAG and LUK:FLAG were both reasonably stable, although they appeared to run slightly higher in SDS-PAGE than their expected sizes of 15.9 kDa (Figure C.3A). Interestingly, while agroexpression of BUK, LUK, PYR1-tr, and BI-1-FL had no visible effect on plant tissue, agroexpression of BI-1-tr results in mild but consistent cell death at the site of infiltration (Figure C.3B and data not shown).

We examined the ability of full-length and truncated BI-1 to inhibit the cell death induced by both Bax and Rx NB:EGFP:HA. While BI-1 variants had little if any impact on NB-mediated cell death, co-expression of full-length BI-1 with Bax reduced the incidence and severity of Bax-mediated cell death compared to treatments of Bax alone or Bax co-expressed with truncated BI-1 (Figure C.3C). Neither BUK, LUK, PYR1-tr, NbPYR1, SlPYR1, BI-1-FL, nor BI-1-tr were observed to interact with Rx NB:EGFP:HA in coimmunoprecipitation experiments, either with or without the presence of MG132 cross-linking agent (data now shown), although it should be noted that BUK protein was never sufficiently stable to visualize.

Functional analysis of interaction candidates

In order to test for a functional requirement for candidate interactors by NB-LRR proteins, we employed a virus-induced gene silencing (VIGS) approach. VIGS vectors targeting candidate interactors were inoculated onto *N. benthamiana* seedlings carrying *Rx2* as a transgene. Rx2 is closely related to Rx, and has the same recognition specificity as Rx for the coat protein of *Potato Virus X* (PVX). True downstream signal adaptors would likely be critical to effective disease resistance, and consequently reduction in their abundance by gene silencing would be predicted to compromise Rx2-mediated resistance. To assess the impact of candidate silencing on Rx2-mediated

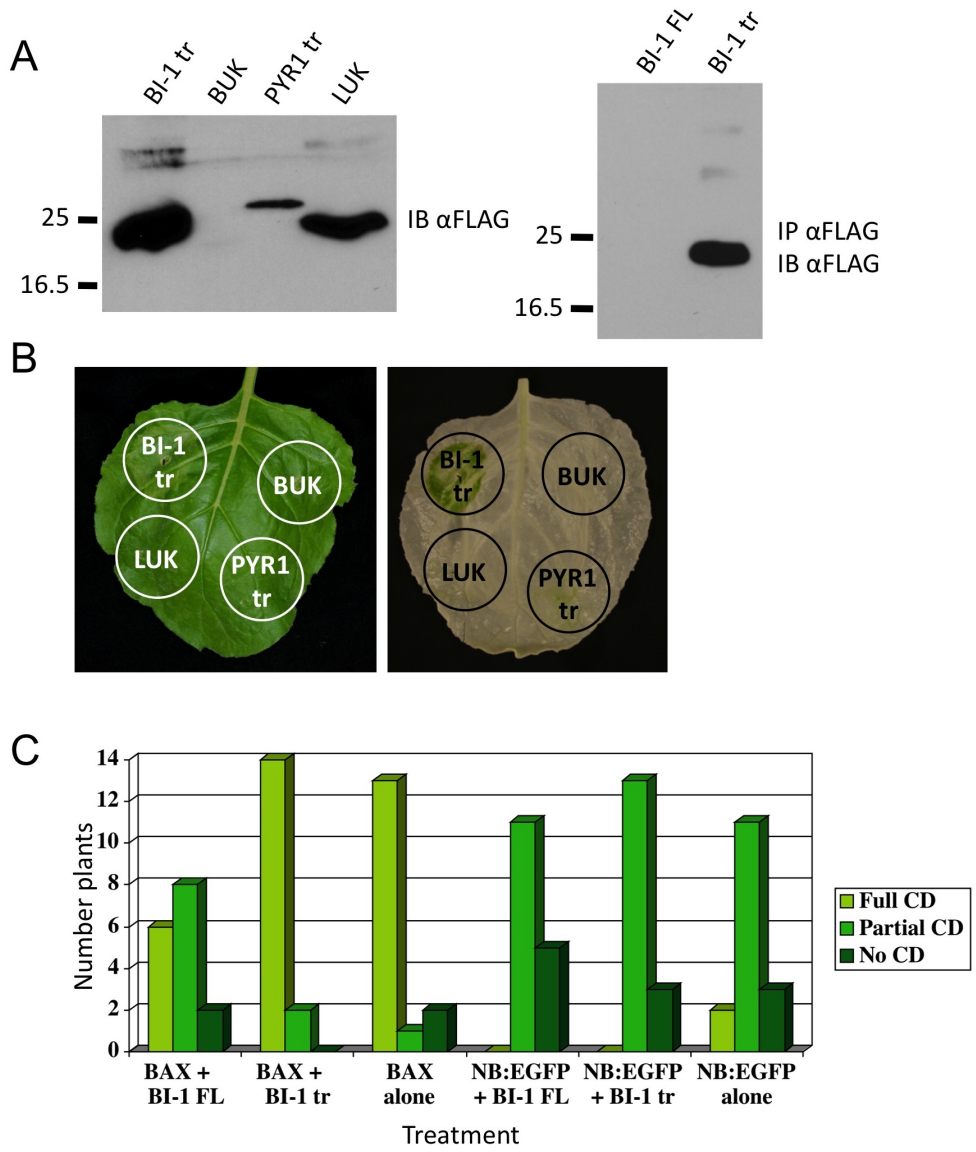


Figure C.3 *Candidate protein stability and activity* in planta. **A,** Stability of candidate proteins. LUK, BUK (with "internal" start), PRY1-tr, and BI-1 (full-length (FL) or truncated (tr)) were agroexpressed in *N. benthamiana* leaves, from which protein samples were collected two days later, as indicated. Samples in the left panel were subjected to SDS-PAGE followed by anti-(α)HA immunoblotting. Samples in the right panel were subjected to anti-HA immunoprecipitation prior to SDS-PAGE and anti-HA immunoblotting. **B,** Cell-death inducing activity of candidate proteins. The indicated proteins were agroexpressed in *N. benthamiana*. Leaves were photographed three days after agroinfiltration (left panel), cleared in ethanol, and photographed again (right panel). **C,** Impact of BI-1 on Bax- and NB-mediated cell death. Bax or Rx NB:EGFP were co-infiltrated with either full-length or truncated BI-1, or alone, as indicated in *N. benthamiana* and *N. tabacum* leaves. Severity and incidence of HR/cell death was recorded four days later (Full CD = complete cell collapse, Partial CD = chlorosis, some cell collapse, No CD = no visible cell death).

resistance, therefore, silenced plants were inoculated with PVX fused to a green florescent protein tag (PVX:GFP), and monitored for the development of disease and/or resistance symptoms.

In the construction of candidate VIGS vectors, we utilized sequence from *N. benthamiana* homologs wherever possible. Thus *BI-1* and *PYR1* VIGS vectors contained sequence derived from *N. benthamiana BI-1* and *PYR1* homologs, while *BUK* and *LUK* VIGS vectors contained tomato-derived sequence, as we were unable to identify homologous *N. benthamiana* sequences. While a single VIGS construct was used for the silencing each of *BUK* and *LUK*, *TUK* silencing was attempted with three overlapping regions of the gene, and *BI-1* silencing was carried out with combinations of five and two overlapping regions of two distinct homologs or alleles. In these cases where *N. benthamiana* sequence was available, RT-PCR analysis demonstrated that while silencing was not 100% effective, it did lower transcript abundance of target genes to some extent, with some VIGS constructs being more effective than others (Figure C.4A and C.4B).

Aside from a handful of isolated incidents, silencing of candidate genes did not result in a crippling of Rx2-mediated resistance significant enough to allow PVX:GFP to spread throughout the plant following inoculation (data not shown). However, candidate silencing did in some cases appear to impact the development of HR lesions at the site of PVX:GFP inoculation (Figure C.5). Most notably, silencing of *PYR1* in some cases appeared to result in an increased severity of HR lesions, although differences between *PYR1*-silenced and *GUS*-insert negative control treatments were slight, and not evident in all replicates (Figure C.5 and data now shown).

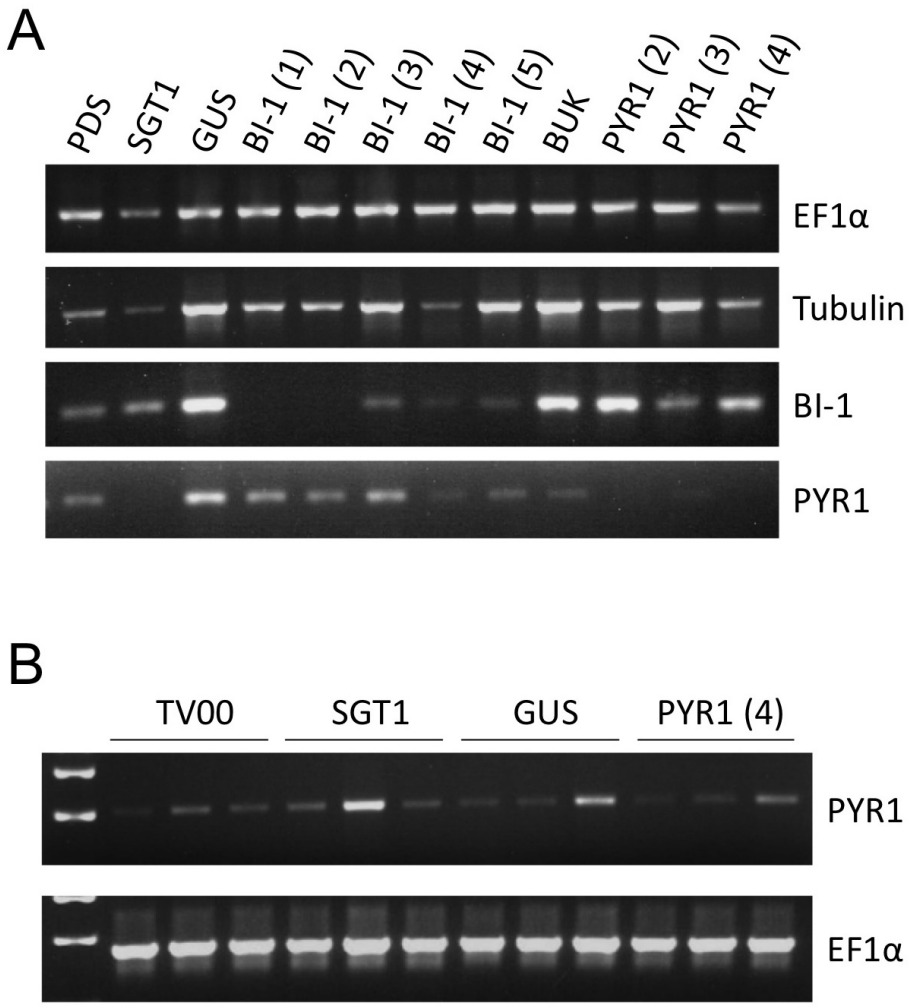


Figure C.4 *Efficacy of candidate silencing*. cDNA was collected from *Rx2 N. benthamiana* plants three weeks after treatment with the indicated VIGS vectors, and subjected to RT-PCR with the primers indicated at right. Samples in **A,** and **B,** are from different silencing replicates.

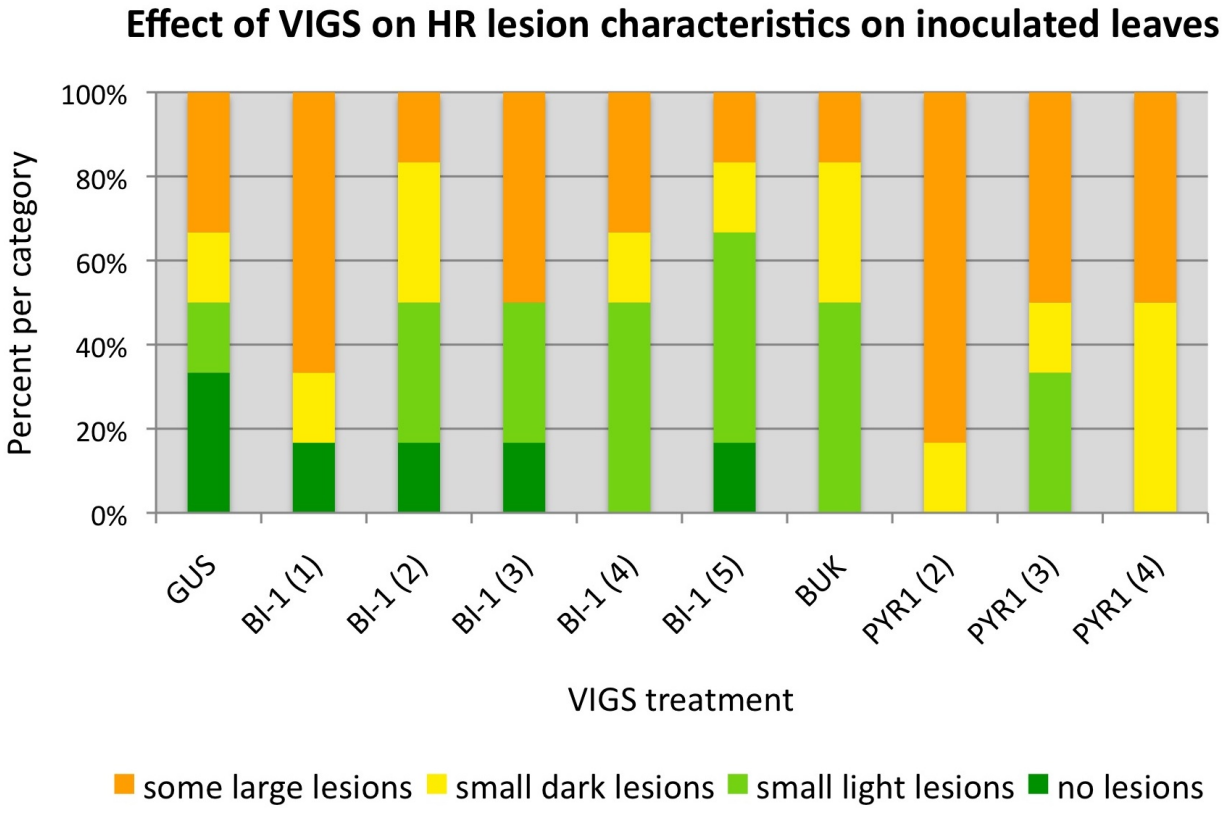


Figure C.5 *HR lesion development in silenced plants.* Three weeks after the indicated silencing treatments, *Rx2 N. benthamiana* plants were rub-inoculated with PVX:GFP and monitored for the development of HR lesions on inoculated leaves. Lesion characteristics were rated by their size and intensity. Results are based on six plants per VIGS treatment.

We also investigated the responses of silenced plants to additional HR/cell death-inducing treatments. PVX CP, Rx NB:EGFP, or Bax were agroexpressed in the leaves of silenced *Rx2 N. benthamiana* plants, and the development of HR/cell death was monitored. There was no clear effect of candidate silencing on Rx-mediated HR, although some *BI-1* silencing treatments did result in an increase in the severity of Bax-induced cell death, as might be expected (Figure C.6). In these and other experiments, we also noted a slight trend of general resistance to cell death in *BUK*-silenced plants. *BUK*-silenced plants were also noted to consistently possess a "crinkled" leaf phenotype, and so this moderate difference in cell death responses may alternately have been due to leaf morphology rather than any specific involvement of BUK in plant cell death responses. As we were unable to identify a *BUK* homolog in *N. benthamiana* and were therefore also unable to assess the efficacy of *BUK* silencing, we also tested *BUK* silencing in "VFNT" tomato plants possessing the *Tm-2a Tobacco Mosaic Virus* (TMV) resistance gene. *BUK* silencing in tomato resulted in a similarly clear phenotype as in *N. benthamiana*, with silenced plants exhibiting moderately curling and stunted top-growth for approximately the fourth through eighth true leaves (Figure C.7). Three weeks after seedlings were inoculated with VIGS vectors, top leaves were rub-inoculated with TMV sap, and monitored for symptom development. No TMV infection occurred, indicating that Tm-2a-mediated resistance was not compromised (data now shown).

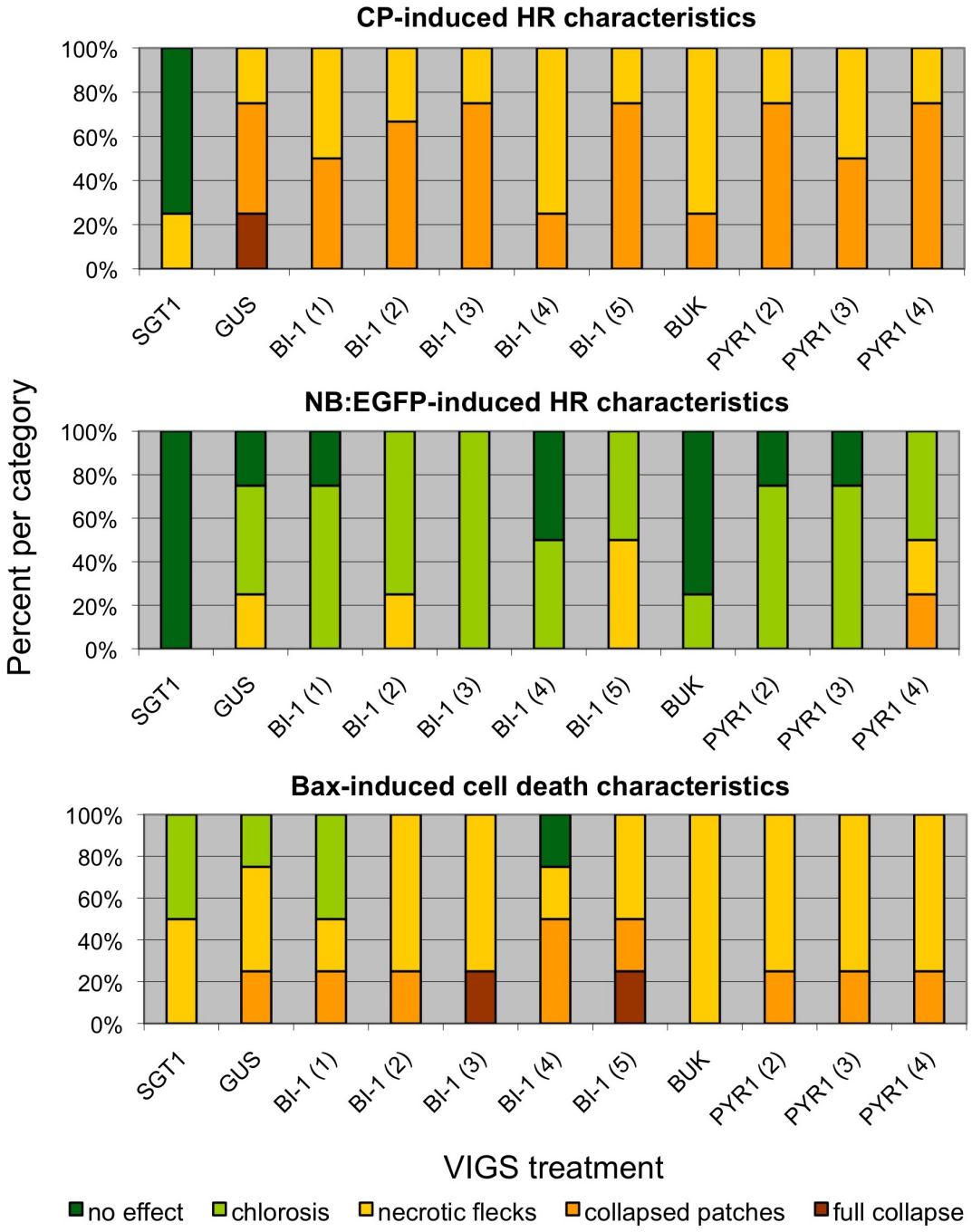


Figure C.6 *Effect of candidate silencing on cell death development.* Approximately three weeks after the indicated silencing treatments, *Rx2 N. benthamiana* leaves were infiltrated with *Agrobacterium* expressing either the PVX CP (top panel), Rx NB:EGFP (middle panel), or Bax (bottom panel), and cell death intensity was recorded nine days later. Results are based on four replicate infiltrations per cell death inducer per VIGS treatment.

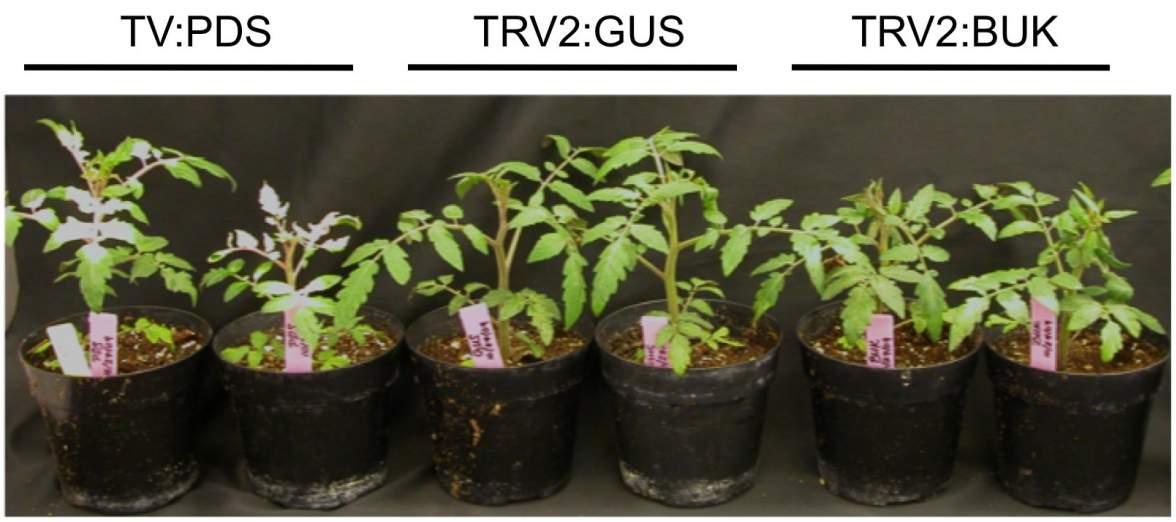


Figure C.7 BUK *silencing in tomato.* VFNT tomato seedlings were inoculated with *Agrobacterium* containing TRV VIGS vectors with inserts of either *PDS*, *GUS*, or *BUK*, and photographed three weeks later.

DISCUSSION

Candidate Rx NB-interacting proteins obtained from our Y2H screen, though few, were nonetheless intriguing. While LUK bore no clear homology to any particular gene or transcript, having a rather piecemeal appearance to its sequence architecture, BUK was clearly derived from a true tomato transcript. The repeated isolation of multiple BUK-encoding cDNAs further enhanced its credibility. However, as NB-LRR protein NB domains are highly conserved, as is the resistance response which results from NB-LRR activation, downstream signaling components are likely also broadly conserved. Thus our inability to identify any *BUK* homologs in *Arabidopsis* or other non-solanaceous genomes reduces the its probability of being a downstream signal adaptor.

ABA is a plant hormone involved in many aspects of plant development and abiotic stress responses. There is also accumulating evidence of an overlap between abiotic and biotic stress signaling networks, including ABA activity. ABA appears primarily to act as a negative regulator of disease resistance, although the exact means and extent of ABA involvement in plant pathogen responses remains unclear (Adie et al. 2007; Ton et al. 2009). *Arabidopsis* PYR1 was recently identified as an ABA-binding and regulatory protein (Park et al. 2009), suggesting it is ideally situated to influence ABA signaling networks, and making it an intriguing candidate for a point of entry that NB-LRR proteins might utilize to access plant signaling pathways. However, lack of observable binding between PYR1 and Rx NB:EGFP makes this scenario somewhat doubtful. Lack of an effect in VIGS experiments is perhaps less informative, as *PYR1* belongs to a highly redundant gene family, with *Arabidopsis* containing 14 *PYR1* homologs, at least seven of which are able to bind ABA (Park et al. 2009). It is thus

probable that our silencing treatments left other, possibly functionally redundant, *PYR1* homologs unaffected, masking any requirement for PYR1 in resistance responses.

Programmed cell death in plants and animals share many similar characteristics, and the ability of animal cell death regulators (such as Bax) to function in plants further suggests shared cell death signaling mechanisms. However, core components of the animal apoptosis pathway are apparently absent from plants (Watanabe and Lam 2004). It is therefore remarkable that plants possess BI-1 proteins, and that these plant proteins are able to suppress cell death in both plant and animal systems (Sanchez et al. 2000; Bolduc et al. 2003; Hückelhoven et al. 2003; Eichmann et al. 2004; Watanabe and Lam 2006). Regulation of plant BI-1-encoding genes has also been found to be responsive to pathogen infection, and reduction in BI-1 levels has been observed to result in accelerated cell death (Sanchez et al. 2000; Bolduc and Brisson 2002; Eichmann et al. 2004). The close relationship of BI-1 with plant cell death and pathogen responses, combined with the unknown nature of its role(s) in plant cell death, makes BI-1 and extremely appealing candidate NB-LRR signal adaptor protein. It is additionally intriguing that we have found that a truncated version of BI-1 induces, rather than represses, cell death (Figure C.3B). As BI-1, and particularly the truncated variant retrieved from your Y2H screen, is active in cell death pathways, it is conceivable that the out-of-frame presentation allowed significantly reduced translation levels of a protein that would otherwise be lethal to the yeast host cells. Regardless, as we were unable to observe either binding between BI-1 and Rx or an effect of *BI-1* silencing on Rx2-mediated resistance, it is unclear whether plant BI-1 has any involvement in NB-LRR protein signaling pathways.

MATERIALS AND METHODS

Cloning and plasmid construction

For sequences of primers used in this study, see table C.3. Rx NB:HA and NB:EGFP:HA were cloned into EcoRI and XhoI sites of pEG202. All sequences for protein expression *in planta* were cloned into XbaI and BamHI sites of pBIN61 with a C-terminal HA or FLAG tag (Bendahmane et al. 2002). Sequences for VIGS were cloned into XbaI and BamHI sites of pTRV2 (Liu et al. 2002), with additional *BI-1* homolog sequences cloned into BamHI and SacI sites of the same vector. PYR1 (2), (3), and (4) inserts correspond to NtPYR1 primer combinations F2/R1, F1/R2, and F2/R2, respectively. BI-1 inserts (1), (2), (3), (4), and (5) correspond to NbBI-1 primer combinations F1/R, F2/R, F3/R, F1/R1, and F2/R1, respectively. PVX:GFP (Peart et al. 2002a) PVX CP (Bendahmane et al. 2000), Rx NB:EGFP (Rairdan et al. 2008), and Bax (del Pozo et al. 2004) clones have been described previously. RNA was extracted using TRIzol reagent (Invitrogen) and cDNA first strand synthesis was performed using SuperScriptIII reverse transcriptase (Invitrogen). PCR was performed with KOD high-fidelity polymerase (Novagen). RACE was performed according to manufacturer instructions (GeneRacer, Invitrogen)

Yeast two-hybrid screen

Yeast two-hybrid screening, from bait quality control through isolation of interacting proteins, was performed essentially as described (Golemis et al. 1999).

Table C.3 Primer sequences

Description	**Sequence (5′ – 3′)**
EcoRI Rx NB F	GTACGGAATTCTTTGAGATGATGCTGGATC
XhoI pBIN61 R	CTTGTG TGCTCGAGAGAGACTGGTGATTTCAG
BUK RACE R (5′)	CTTGCCAGTCGTTGCCTATTGCATCT
BUK RACE F (3′)	CGAGGATCTCCGGTGGATAGCATCT
XbaI BUK Met F	CGTATCTAGAACCACCATGGAAGTGAGAAGCTCC
BamHI BUK R	CGTAGGATCCACTGTCATTATTGCCGAAATTTATGG
XbaI NbBI-1 tr F	CGTATCTAGAACCACCATGAACTTCCGCCAGATCTTACC
BamHI NbBI-1 R	CGTAGGATCCGTTTCTCCTCTTCTTCTTCTTCTCTTCC
BamHI LUK R	CGTAGGATCCCTCGCTGATAGGGCTGCTGGTTTGG
BamHI SlPYR1 tr R	CGTAGGATCCACGTTCGGTACTGGTAGCGGC
XbaI pYESTrp2 F	CGTATCTAGAACCACCATGAACGGCCGCCAGTGTGCTGG
BamHI SlBI-1 R	CGTAGGATCCGTTTCTCCTCTTCTTCTTCTTC
XbaI SlBI-1 F	CGTATCTAGAAGAAGAAGAAGAGAACAATGGAAGG
XbaI Nt BI-1 FL F	CGTATCTAGAGAAGCAAAGAGAGAGAAATGGAGTC
XbaI SlPYR1 F	CGTATCTAGAACCACCATGGATAATAAACCGGAAACG
BamHI SlPYR1 R	CGTAGGATCCCCTGTGACTCGCATCACG
XbaI NtPYR1 F1	CGTATCTAGAACCACCATGCCTCCTAGTTCTCCAG
BamHI NtPYR1 R1	CGTAGGATCCAGAATCTTTTCTTCTGTTCACG
XbaI NtPYR1 F2	CGTATCTAGATTCCATACTCACCCAGTGG
BamHI NtPYR1 R2	CGTAGGATCCAACAACGTAGGATTCAACG
BamHI NbBI-1 R	CGTAGGATCCGTTTCTCCTCTTCTTCTTCTTCTC
BamHI NbBI-1 R1	CGTAGGATCCAGGATGCAAAGTGCAA
XbaI NbBI-1 F1	CGTATCTAGACGCTGGAGTTACGATTCTCTT
XbaI NbBI-1 F2	CGTATCTAGACTGCTGGGGCTTACCTTCAC
XbaI NbBI-1 F3	CGTATCTAGAGAGGATAGCACTTCTGATGGCAG
SacI BI-1h R1	CGTAGAGCTCAGGAGAACTCAACAAAAGC
BamHI BI-1h F1	CGTAGGATCCATGGAAGGTTTTACCTCG
SacI BI-1h R2	CGTAGAGCTCGCATGATGATCAGAATACG
BamHI BI-1h F1	CGTAGGATCCCTCTGTTGGTTGCAACTG
PYR1 RT R	GATCTGCGATAACGATGTGAG
PYR1 RT F	TTACAAAGGCGTACTATGCCG
BI-1 RT R	GAAGCAACCAAAAGCCACAGC
BI-1 RT F	TTCTTCAATTCACAGTCGGCG

Protein expression and analysis

For expression of proteins *in planta*, binary vectors were transformed into *A. tumefaciens* strain C58C1 carrying the virulence plasmid pCH32. *Agrobacterium*-mediated transient expression was performed as previously described (Bendahmane et al. 2000). Proteins were expressed under the 35S promoter of the pBIN61 vector (Bendahmane et al. 2002). Total protein extracts were collected by grinding leaf disks in 8M urea followed by addition of SDS-PAGE loading buffer and boiling. Coimmunoprecipitation was performed as previously described (Rairdan and Moffett 2006). For yeast protein extraction, 5mL overnight cultures were spun down and resuspended in SDS-PAGE loading buffer, boiled for 10 minutes, and then loaded for SDS-PAGE.

Virus-induced gene silencing

Three leaves each of three-week-old *Rx2 N. benthamiana* (Bhattacharjee et al. 2009) seedlings were co-infiltrated with *Agrobacterium* carrying plasmids pBINTra6 (TRV RNA1 cDNA) (Ratcliff et al. 2001) and RNA2 constructs derived from pTRV2 (Liu et al. 2002) at concentrations of OD_{600} = 0.1 and 0.2, respectively. Plants were used 3-5 weeks later for either resistance breaking assays or for protein expression. For resistance breaking assays, approximately three leaves per plant were dusted with carborundum powder and rub-inoculated with sap containing PVX:GFP virions, prepared as previously described (Bhattacharjee et al. 2009). Systemic tissue was monitored for signs of viral infection and/or systemic necrosis for up to 40 days following inoculation.

REFERENCES

Aarts, N., Metz, M., Holub, E., Staskawicz, B. J., Daniels, M. J. and Parker, J. E. (1998) Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 10306-11.

Abramoff, M. D., Magelhaes, P. J. and Ram, S. J. (2004) Image processing with ImageJ. *Biophotonics International* **11**: 36-42.

Ade, J., Deyoung, B. J., Golstein, C. and Innes, R. W. (2007) Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 2531-2536.

Adie, B. A. T., Pérez-Pérez, J., Pérez-Pérez, M. M., Godoy, M., Sánchez-Serrano, J.-J., Schmelz, E. A. and Solano, R. (2007) ABA is an essential signal for plant resistance to pathogens affecting JA biosynthesis and the activation of defenses in *Arabidopsis*. *The Plant Cell* **19**: 1665-1681.

Albrecht, M. and Takken, F. L. (2006) Update on the domain architectures of NLRs and R proteins. *Biochemical and Biophysical Research Communications* **339**: 459-62.

Alcázar, R., García, A. V., Parker, J. E. and Reymond, M. (2009) Incremental steps toward incompatibility revealed by Arabidopsis epistatic interactions modulating salicylic acid pathway activation. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 334-339.

Argout, X., Salse, J., Aury, J.-M., Guiltinan, M. J., Droc, G., Gouzy, J., Allegre, M., et al. (2011) The genome of *Theobroma cacao*. *Nature Genetics* **43**: 101-108.

Axtell, M. J. and Staskawicz, B. J. (2003) Initiation of *RPS2*-specified disease resistance in *Arabidopsis* is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**: 369-377.

Bai, J., Pennill, L. A., Ning, J., Lee, S. W., Ramalingam, J., Webb, C. A., Zhao, B., et al. (2002) Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. *Genome Research* **12**: 1871-1884.

Bakker, E. G., Toomajian, C., Kreitman, M. and Bergelson, J. (2006) A genome-wide survey of *R* gene polymorphisms in *Arabidopsis*. *The Plant Cell* **18**: 1803-1818.

Balmuth, A. and Rathjen, J. P. (2007) Genetic and molecular requirements for function of the Pto/Prf effector recognition complex in tomato and *Nicotiana benthamiana*. *The Plant Journal* **51**: 978-990.

Baurés, I., Candresse, T., Leveau, A., Bendahmane, A. and Sturbois, B. (2008) The *Rx* gene confers resistance to a range of *Potexviruses* in transgenic *Nicotiana* plants. *Molecular Plant-Microbe Interactions* **21**: 1154-1164.

Belkhadir, Y., Nimchuk, Z., Hubert, D. A., Mackey, D. and Dangl, J. L. (2004) Arabidopsis RIN4 Negatively Regulates Disease Resistance Mediated by RPS2 and RPM1 Downstream or Independent of the NDR1 Signal Modulator and Is Not Required for the Virulence Functions of Bacterial Type III Effectors AvrRpt2 or AvrRpm1. *The Plant Cell* **16**: 2822-35.

Bendahmane, A., Farnham, G., Moffett, P. and Baulcombe, D. C. (2002) Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. *The Plant Journal* **32**: 195-204.

Bendahmane, A., Kanyuka, K. and Baulcombe, D. C. (1999) The *Rx* gene from potato controls separate virus resistance and cell death responses. *The Plant Cell* **11**: 781-791.

Bendahmane, A., Querci, M., Kanyuka, K. and Baulcombe, D. C. (2000) *Agrobacterium* transient expression system as a tool for the isolation of disease resistance genes: application to the *Rx2* locus in potato. *The Plant Journal* **21**: 73-81.

Bernoux, M., Timmers, T., Jauneau, A., Briere, C., De Wit, P. J., Marco, Y. and Deslandes, L. (2008) RD19, an *Arabidopsis* cysteine protease required for RRS1-R-mediated resistance, is relocalized to the nucleus by the *Ralstonia solanacearum* PopP2 effector. *The Plant Cell* **20**: 2252-2264.

Bhattacharjee, S., Zamora, A., Azhar, M. T., Sacco, M. A., Lambert, L. H. and Moffett, P. (2009) Virus resistance induced by NB-LRR proteins involves Argonaute4-dependent translational control. *The Plant Journal* **58**: 940-951.

Bieri, S., Mauch, S., Shen, Q. H., Peart, J., Devoto, A., Casais, C., Ceron, F., et al. (2004) RAR1 positively controls steady state levels of barley MLA resistance proteins and enables sufficient MLA6 accumulation for effective resistance. *The Plant Cell* **16**: 3480-95.

Bolduc, N. and Brisson, L. F. (2002) Antisense down regulation of Nt BI-1 in tobacco BY-2 cells induces accelerated cell death upon carbon starvation. *FEBS Letters* **532**: 111-114.

Bolduc, N., Ouellet, M., Pitre, F. and Brisson, L. F. (2003) Molecular characterization of two plant BI-1 homologues which suppress Bax-induced apoptosis in human 293 cells. *Planta* **216**: 377-386.

Bomblies, K. and Weigel, D. (2007) Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nature Reviews Genetics* **8**: 382-393.

Cannon, S. B., Zhu, H., Baumgarten, A. M., Spangler, R., May, G., Cook, D. R. and Young, N. D. (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *Journal of Molecular Evolution* **54**: 548-62.

Caplan, J. L., Mamillapalli, P., Burch-Smith, T. M., Czymmek, K. and Dinesh-Kumar, S. P. (2008) Chloroplastic protein NRIP1 mediates innate immune receptor recognition of a viral effector. *Cell* **132**: 449-462.

Chase, M. W. and Reveal, J. L. (2009) A phylogenetic classification of the land plants to accompany APG III. *Botanical Journal of the Linnean Society* **161**: 122-127.

Chini, A., Grant, J. J., Seki, M., Shinozaki, K. and Loake, G. J. (2004) Drought tolerance established by enhanced expression of the *CC-NBS-LRR* gene, *ADR1*, requires salicylic acid, EDS1, and ABI1. *The Plant Journal* **38**: 810-822.

Chini, A. and Loake, G. J. (2005) Motifs specific for the ADR1 NBS-LRR protein family in Arabidopsis are conserved among NBS-LRR sequences from both dicotyledonous and monocotyledonous plants. *Planta* **221**: 597-601.

Chisholm, S. T., Coaker, G., Day, B. and Staskawicz, B. J. (2006) Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**: 803-814.

Collier, S. M. and Moffett, P. (2009) NB-LRRs work a "bait and switch" on pathogens. *Trends in Plant Science* **14**: 521-529.

Dangl, J. L. and Jones, J. D. (2001) Plant pathogens and integrated defense responses to infection. *Nature* **411**: 826-33.

Day, B., Dahlbeck, D., Huang, J., Chisholm, S. T., Li, D. and Staskawicz, B. J. (2005) Molecular basis for the RIN4 negative regulation of RPS2 disease resistance. *The Plant Cell* **17**: 1292-1305.

del Pozo, O., Pedley, K. F. and Martin, G. B. (2004) MAPKKK_ is a positive regulator of cell death associated with both plant immunity and disease. *The EMBO Journal* **23**: 3072-82.

Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M., Boucher, C., Somssich, I., et al. (2003) Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 8024-8029.

Dodds, P. N., Lawrence, G. J., Catanzariti, A. M., Teh, T., Wang, C. I., Ayliffe, M. A., Kobe, B., et al. (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 of the United States of America* **103**: 8888-93.

Eichmann, R., Schultheiss, H., Kogel, K.-H. and Hückelhoven, R. (2004) The barley apoptosis suppressor homologue Bax Inhibitor-1 compromises nonhost penetration resistance of barley to the inappropriate pathogen *Blumeria graminis* f. sp. *tritici*. *Molecular Plant-Microbe Interactions* **17**: 484-490.

Ellis, J., Dodds, P. N. and Lawrence, G. J. (2007a) Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annual Review of Phytopathology* **45**: 289-306.

Ellis, J. G., Lawrence, G. J. and Dodds, P. N. (2007b) Further analysis of gene-for-gene disease resistance specificity in flax. *Molecular Plant Pathology* **8**: 103-109.

Farnham, G. and Baulcombe, D. C. (2006) Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 18828-18833.

Flor, H. H. (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology* **9**: 275-296.

Friedman, A. R. and Baker, B. J. (2007) The evolution of resistance genes in multi-protein plant resistance systems. *Current Opinion in Genetics and Development* **17**: 493-499.

Fritz, J. H., Ferrero, R. L., Philpott, D. J. and Girardin, S. E. (2006) Nod-like proteins in immunity, inflammation and disease. *Nature Immunology* **7**: 1250-1257.

Frost, D., Way, H., Howles, P., Luck, J., Manners, J., Hardham, A., Finnegan, J., et al. (2004) Tobacco transgenic for the flax rust resistance gene *L* expresses allele-specific activation of defense responses. *Molecular Plant-Microbe Interactions* **17**: 224-232.

Gimenez-Ibanez, S., Hann, D. R., Ntoukakis, V., Petutschnig, E., Lipka, V. and Rathjen, J. P. (2009) AvrPtoB targets the LysM receptor kinase CERK1 to promote bacterial virulence on plants. *Current Biology* **19**: 423-429.

Göhre, V., Spallek, T., Häweker, H., Mersmann, S., Mentzel, T., Boller, T., de Torres, M., et al. (2008) Plant pattern-recognition receptor FLS2 is directed for degradation by the bacterial ubiquitin ligase AvrPtoB. *Current Biology* **18**: 1824-1832.

Golemis, E. A., Serebriiskii, I., Finley, R. L., Kolonin, M. G., Gyuris, J. and Brent, R. (1999) Interaction trap/two hybrid system to identify interacting proteins. *Current Protocols in Molecular Biology* **20.1**: 1-40.

Grant, J. J., Chini, A., Basu, D. and Loake, G. J. (2003) Targeted activation tagging of the Arabidopsis NBS-LRR gene, ADR1, conveys resistance to virulent pathogens. *Molecular Plant-Microbe Interactions* **16**: 669-80.

Grube, R. C., Radwanski, E. R. and Jahn, M. (2000) Comparative genetics of disease resistance within the Solanaceae. *Genetics* **155**: 873-87.

Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protien modeling. *Electrophoresis* **18**: 2714-2723.

Gutierrez, J. R., Balmuth, A. L., Ntoukakis, V., Mucyn, T. S., Gimenez-Ibanez, S., Jones, A. M. and Rathjen, J. P. (2010) Prf immune complexes of tomato are oligomeric

and contain multiple Pto-like kinases that diversify effector recignition. *The Plant Journal* **61**: 507-518.

Gyuris, J., Golemis, E., Chertkov, H. and Brent, R. (1993) Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **75**: 791-803.

Holt, B. F., Belkhadir, Y. and Dangl, J. L. (2005) Antagonistic control of disease resistance protein stability in the plant immune system. *Science* **309**: 929-932.

Hückelhoven, R., Dechert, C. and Kogel, K.-H. (2003) Over-expression of barley BAX inhibitor 1 induces breakdown of *mlo*-mediated penetration resistance to *Blumeria graminis*. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 5555-5560.

Jermstad, K. D., Sheppard, L. A., Kinloch, B. B., Delfina-Mix, A., Ersoz, E. S., Krutovsky, K. V. and Neale, D. B. (2006) Isolation of a full-length CC-NBS-LRR resistance gene analog candidate from sugar pine showing low nucleotide diversity. *Tree Genetics and Genomics* **2**: 76-85.

Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P. and Valent, B. (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO Journal* **19**: 4004-14.

Jones, J. D. and Dangl, J. L. (2006) The plant immune system. *Nature* **444**: 323-329.

Kadota, Y., Shirasu, K. and Guerois, R. (2010) NLR sensors meet at the SGT1-HSP90 crossroad. *Trends in Biochemical Sciences* **35**: 199-207.

Kim, M. G., Geng, X., Lee, S. Y. and Mackey, D. (2009) The *Pseudomonas syringae* type III effector AvrRpm1 induces significant defenses by activating Arabidopsis nucleotide-binding leucine-rich repeat protein RPS2. *The Plant Journal* **57**: 645-653.

Krasileva, K. V., Dahlbeck, D. and 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. *The Plant Cell* **22**: 2444-2458.

Lanfermeijer, F. C., Dijkhuis, J., Sturre, M. J., de Haan, P. and Hille, J. (2003) Cloning and characterization of the durable tomato mosaic virus resistance gene Tm-2^2 from *Lycopersicon esculentum*. *Plant Molecular Biology* **52**: 1037-49.

Leipe, D. D., Koonin, E. V. and Aravind, L. (2004) STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *Journal of Molecular Biology* **343**: 1-28.

Leister, R. T., Dahlbeck, D., Day, B., Li, Y., Chesnokova, O. and Staskawicz, B. J. (2005) Molecular genetic evidence for the role of *SGT1* in the intramolecular complementation of Bs2 protein activity in *Nicotiana benthamiana*. *The Plant Cell* **17**: 1268-1278.

Liu, Y., Schiff, M., Marathe, R. and Dinesh-Kumar, S. P. (2002) Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. *The Plant Journal* **30**: 415-29.

Lu, R., Malcuit, I., Moffett, P., Ruiz, M. T., Peart, J., Wu, A. J., Rathjen, J. P., et al. (2003) High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *The EMBO Journal* **22**: 5690-9.

Lupas, A., Van Dyke, M. and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science* **252**: 1162-1164.

Mackey, D., Holt, B. F., Wiig, A. and Dangl, J. L. (2002) RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* **108**: 743-754.

Maekawa, T., Cheng, W., Spiridon, L. N., Töller, A., Lukasik, E., Saijo, Y., Liu, P., et al. (2011) Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal function module for triggering cell death. *Cell Host and Microbe* **9**: 187-199.

Mazourek, M., Cirulli, E. T., Collier, S. M., Landry, L. G., Kang, B.-C., Quirin, E. A., Bradeen, J. M., et al. (2009) The fractionated orthology of *Bs2* and *Rx*/*Gpa2* supports shared synteny of disease resistance in the Solanaceae. *Genetics* **182**: 1351-1364.

McDowell, J. M., Dhandaydham, M., Long, T. A., Aarts, M. G., Goff, S., Holub, E. B. and Dangl, J. L. (1998) Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *The Plant Cell* **10**: 1861-74.

McDowell, J. M. and Simon, S. A. (2006) Recent insight into *R* gene evolution. *Molecular Plant Pathology* **7**: 437-448.

McHale, L., Tan, X., Koehl, P. and Michelmore, R. W. (2006) Plant NBS-LRR proteins: adaptable guards. *Genome Biology* **7**: 212.

Mestre, P. and Baulcombe, D. C. (2006) Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *The Plant Cell* **18**: 491-501.

Meyers, B. C., Chin, D. B., Shen, K. A., Sivaramakrishnan, S., Lavelle, D. O., Zhang, Z. and Michelmore, R. W. (1998) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *The Plant Cell* **10**: 1817-32.

Meyers, B. C., Dickerman, A. W., Michelmore, R. W., Sivaramakrishnan, S., Sobral, B. W. and Young, N. D. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *The Plant Journal* **20**: 317-32.

Meyers, B. C., Kozik, A., Griego, A., Kuang, H. and Michelmore, R. W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell* **15**: 809-34.

Moffett, P., Farnham, G., Peart, J. and Baulcombe, D. C. (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *The EMBO Journal* **21**: 4511-4519.

Moffett, P. and Klessig, D. F. (2008). Plant Resistance to Viruses: Natural Resistance Associated with Dominant Genes. Encyclopedia of Virology, 3rd Edition. B. W. J. Mahy and van Regenmortel, M. H. V. Oxford, Elsevier. **4:** 170-177.

Moreira, A. and Jones, R. A. C. (1980) Properties of a resistance-breaking strain of potato virus X. *Annals of Applied Biology* **95**: 93-103.

Mosher, R. A., Durrant, W. E., Wang, D., Song, J. and Dong, X. (2006) A comprehensive structure-function analysis of *Arabidopsis* SNI1 defines essential regions and transcriptional repressor activity. *The Plant Cell* **18**: 1750-1765.

Mucyn, T., Wu, A. J., Balmuth, A. L., Maryan Arasteh, J. and Rathjen, J. P. (2009) Regulation of tomato Prf by Pto-like protein kinases. *Molecular Plant-Microbe Interactions* **22**: 391-401.

Mucyn, T. S., Clemente, A., Andriotis, M. E., Balmuth, A. L., Oldroyd, E. D., Staskawicz, B. J. and Rathjen, J. P. (2006) The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity. *The Plant Cell* **18**: 2792-2806.

Mueller, L. A., Solow, T. H., Nicolas, T., Skwarecki, B., Buels, R., Binns, J., Lin, C., et al. (2005) The SOL genomics network: a comparative resource for Solanaceae biology and beyond. *Plant Physiology* **138**: 1310-1317.

Nombela, G., Williamson, V. M. and Muniz, M. (2003) The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. *Molecular Plant-Microbe Interactions* **16**: 645-9.

Oldroyd, E. D. and Staskawicz, B. J. (1998) Genetically engineered broad-spectrum disease resistance in tomato. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 10300-10305.

Pan, Q., Wendel, J. and Fluhr, R. (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *Journal of Molecular Evolution* **50**: 203-213.

Park, S.-Y., Fung, P., Nishimura, N., Jensen, D. R., Fujii, H., Zhao, Y., Lumba, S., et al. (2009) Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* **324**: 1068-1071.

Peart, J. R., Cook, G., Feys, B. J., Parker, J. E. and Baulcombe, D. C. (2002a) An EDS1 orthologue is required for N-mediated resistance against tobacco mosaic virus. *The Plant Journal* **29**: 569-79.

Peart, J. R., Lu, R., Sadanandom, A., Malcuit, I., Moffett, P., Brice, D. C., Schauser, L., et al. (2002b) Ubiquitin ligase-associated protein SGT1 is required for host and

nonhost disease resistance in plants. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 10865-9.

Peart, J. R., Mestre, P., Lu, R., Malcuit, I. and Baulcombe, D. C. (2005) NRG1, a CC-NB-LRR protein, together with N, a TIR-NB-LRR protein, mediates resistance against tobacco mosaic virus. *Current Biology* **15**: 968-973.

Porter, B., Paidi, M., Ming, R., Alam, M., Nishijima, W. and Zhu, Y. (2009) Genome-wide analysis of *Carica papaya* reveals a small NBS resistance gene family. *Molecular Genetics and Genomics* **281**: 609-626.

Qi, S., Pang, Y., Hu, Q., Liu, Q., Li, H., Zhou, Y., He, T., et al. (2010) Crystal structure of the *Caenorhabditis elegans* apoptosome reveals an octameric assembly of CED-4. *Cell* **141**: 446-457.

Qu, S., Liu, G., Zhou, B., Bellizzi, M., Zeng, L., Dai, L., Han, B., et al. (2006) The broad-spectrum blast resistance gene *Pi9* encodes a nucleotide-binding site-leucine-rich repeat protein and is a member of a multigene family in rice. *Genetics* **172**: 1901-14.

Querci, M., van der Vlugt, R., Goldbach, R. and Salazar, L. F. (1993) RNA sequence of Potato virus X strain HB. *Journal of General Virology* **74**: 2251-2255.

Rairdan, G. and Moffett, P. (2007) Brothers in arms? Common and contrasting themes in pathogen perception by plant NB-LRR and animal NACHT-LRR proteins. *Microbes and Infection* **9**: 677-686.

Rairdan, G. J., Collier, S. M., Sacco, M. A., Baldwin, T. T., Boettrich, T. and Moffett, P. (2008) Novel roles for the CC and NB domains in signaling and recognition by the Rx NB-LRR disease resistance protein. *The Plant Cell* **20**: 739-751.

Rairdan, G. J. and Moffett, P. (2006) Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. *The Plant Cell* **18**: 2082-2093.

Ratcliff, F., Martin-Hernandez, A. M. and Baulcombe, D. C. (2001) Tobacco rattle virus as a vector for analysis of gene function by silencing. *The Plant Journal* **25**: 237-245.

Rentel, M. C., Leonelli, L., Dahlbeck, D., Zhao, B. and Staskawicz, B. J. (2008) Recognition of the *Hyaloperonospora parasitica* effector ATR13 triggeres resistance against oomycete, bacterial, and viral pathogens. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 1091-1096.

Riedl, S. J., Li, W., Chao, Y., Schwarzenbacher, R. and Shi, Y. (2005) Structure of the apoptotic protease-activating factor 1 bound to ADP. *Nature* **434**: 926-933.

Rosebrock, T. R., Zeng, L., Brady, J. J., Abramovitch, R. B., Xiao, F. and Martin, G. B. (2007) A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature* **448**: 370-374.

Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E. and Williamson, V. M. (1998) The nematode resistance gene *Mi* of tomato confers resistance against

the potato aphid. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 9750-4.

Roy, A., Kucukaral, A. and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**: 725-738.

Sacco, M. A., Koropacka, K., Grenier, E., Jaubert, M. J., Blanchard, A., Goverse, A., Smant, G., et al. (2009) The cyst nematode SPRYSEC protein RBP-1 elicits Gpa2- and RanGAP2-dependent plant cell death. *PLoS Pathogens* **5**: e1000564.

Sacco, M. A., Mansoor, S. and Moffett, P. (2007) A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *The Plant Journal* **52**: 82-93.

Sacco, M. A. and Moffett, P. (2009). Disease resistance genes: form and function. Molecular Plant-Microbe Interactions. K. Bouarab, Brisson, N. and Daayf, F. Wallingford, UK, CABI**:** 94-141.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**: 406-425.

Sakai, H. and Itoh, T. (2010) Massive gene losses in Asian cultivated rice unveiled by comparative genome analysis. *BMC Genomics* **11**: 121.

Sali, A., Potterton, L., Yuan, F., van Viljmen, H. and Karplus, M. (1995) Evaluation of comparative protien modeling by MODELLER. *Proteins* **23**: 318-326.

Sanchez, P., Zabala, M. D. T. and Grant, M. (2000) AtBI-1, a plant homologue of Bax Inhibitor-1, suppresses Bax-induced cell death in yeast and is rapidly upregulated during wounding and pathogen challenge. *The Plant Journal* **21**: 393-399.

Schwartz, R. and Dayhoff, M. (1979). Matrices for detecting distant relationships. Atlas of protein sequences. M. Dayhoff, National Biomedical Research Foundation**:** 353-358.

Seah, S., Telleen, A. C. and Williamson, V. M. (2007) Introgression and endogenous *Mi-1* gene clusters in tomato differ by complex rearrangements in flanking sequences and show sequence exchange and diversifying selection among homologues. *Theoretical and Applied Genetics* **114**: 1289-1302.

Shan, L., He, P., Li, J., Heese, A., Peck, S. C., Nürnberter, T., Martin, G. B., et al. (2008) Bacterial effectors target the common signaling partner BAK1 to disrupt multiple MAMP receptor-signaling complexes and impede plant immunity. *Cell Host and Microbe* **4**: 17-27.

Shao, F., Golstein, C., Ade, J., Stoutemyer, M., Dixon, J. E. and Innes, R. W. (2003) Cleavage of *Arabidopsis* PBS1 by a bacterial type III effector. *Science* **301**: 1230-1233.

Shen, Q. H., Saijo, Y., Mauch, S., Biskup, C., Bieri, S., Keller, B., Seki, H., et al. (2007) Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science* **315**: 1098-103.

Shen, Q. H., Zhou, F., Bieri, S., Haizel, T., Shirasu, K. and Schulze-Lefert, P. (2003) Recognition specificity and RAR1/SGT1 dependence in barley Mla disease resistance genes to the powdery mildew fungus. *The Plant Cell* **15**: 732-744.

Söding, J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* **21**: 951-960.

Sohn, K. H., Lei, R., Nemri, A. and Jones, J. (2007) The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in *Arabidopsis thaliana*. *The Plant Cell* **19**: 4077-4090.

Sun, Q., Collins, N. C., Ayliffe, M., Smith, S. M., Drake, J., Pryor, T. and Hulbert, S. H. (2001) Recombination between paralogues at the Rp1 rust resistance locus in maize. *Genetics* **158**: 423-38.

Swiderski, M. R., Birker, D. and Jones, J. D. (2009) The TIR domain of the TIR-NB-LRR resistance protein is a signaling domain involved in cell death induction. *Molecular Plant-Microbe Interactions* **22**: 157-165.

Takahashi, H., Miller, J., Nozaki, Y., Takeda, M., Shah, J., Hase, S., Ikegami, M., et al. (2002) RCY1, an Arabidopsis thaliana RPP8/HRT family resistance gene, conferring resistance to cucumber mosaic virus requires salicylic acid, ethylene and a novel signal transduction mechanism. *The Plant Journal* **32**: 655-67.

Takken, F. L., Albrecht, M. and Tameling, W. I. (2006) Resistance proteins: molecular switches of plant defence. *Current Opinion in Plant Biology* **9**: 383-90.

Takken, F. L. W. and Tameling, W. I. L. (2009) To nibble at plant resistance proteins. *Science* **324**: 744-746.

Tameling, W. I., Elzinga, S. D., Darmin, P. S., Vossen, J. H., Takken, F. L., Haring, M. A. and Cornelissen, B. J. (2002) The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *The Plant Cell* **14**: 2929-39.

Tameling, W. I., Vossen, J. H., Albrecht, M., Lengauer, T., Berden, J. A., Haring, M. A., Cornelissen, B. J., et al. (2006) Mutations in the NB-ARC domain of I-2 that impair ATP hydrolysis cause autoactivation. *Plant Physiology* **140**: 1233-45.

Tameling, W. I. L. and Baulcombe, D. C. (2007) Physical association of the NB-LRR resistance protein Rx with a Ran GTPase-activating protein is required for extreme resistance to potato virus X. *The Plant Cell* **19**: 1682-1694.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599.

Tao, Y., Yuan, F., Leister, R. T., Ausubel, F. M. and Katagiri, F. (2000) Mutational analysis of the Arabidopsis nucleotide binding site-leucine-rich repeat resistance gene *RPS2*. *The Plant Cell* **12**: 2541-2554.

Tarr, D. E. K. and Alexander, H. M. (2009) TIR-NBS-LRR genes are rare in monocots: evidence from diverse monocot orders. *BMC Research Notes* **2**: 197.

Tobias, C. M., Oldroyd, G. E., Chang, J. H. and Staskawicz, B. J. (1999) Plants expressing the Pto disease resistance gene confer resistance to recombinant PVX containing the avirulence gene AvrPto. *The Plant Journal* **17**: 41-50.

Ton, J., Flors, V. and Mauch-Mani, B. (2009) The multifaceted role of ABA in disease resistance. *Trends in Plant Science* **14**: 310-317.

Traut, T. W. (1994) The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *European Journal of Biochemistry* **222**: 9-19.

Tschopp, J. T. and Schroder, K. (2010) NLRP3 inflammasome activation: the convergence of multiple signalling pathways on ROS production? *Nature Reviews Immunology* **10**: 210-215.

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., et al. (2006) The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* **313**: 1596-604.

Ueda, H., Yamaguchi, Y. and 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.

Vallejo, A. N., Pogulis, R. J. and Pease, L. R. (2003). Mutagenesis and Synthesis of Novel Recombinant Genes using PCR. PCR primer : a laboratory manual. Cold Spring Harbor, N.Y. :, Cold Spring Harbor Laboratory Press**:** 467-474.

Van der Biezen, E. A. and Jones, J. D. (1998) Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochemical Science* **23**: 454-6.

van der Biezen, E. A. and Jones, J. D. (1998) The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Current Biology* **8**: R226-7.

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

van der Vossen, E. A., van der Voort, J. N., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D. C., Bakker, J., et al. (2000) Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *The Plant Journal* **23**: 567-76.

van Ooijen, G., Mayr, G., Albrecht, M., Cornelissen, B. J. and Takken, F. L. W. (2008a) Transcomplementation, but not physical association of the CC-NB-ARC and LRR

domains of tomato R protein Mi-1.2 is altered by mutations in the ARC2 subdomain. *Molecular Plant* **1**: 401-410.

van Ooijen, G., Mayr, G., Kasiem, M. M. A., Albrecht, M., Cornelissen, B. J. and Takken, F. L. W. (2008b) Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany* **59**: 1383-1397.

van Ooijen, G., van den Burg, H. A., Cornelissen, B. J. and Takken, F. L. (2007) Structure and function of resistance proteins in Solanaceous plants. *Annual Review of Phytopathology* **45**: 43-72.

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontanta, P., et al. (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). *Nature Genetics* **41**: 833-841.

Wang, W., Wen, Y., Berkey, R. and Xiao, S. (2009) Specific targeting of the *Arabidopsis* resistance protein RPW8.2 to the interfacial membrane encasing the fungal haustorium renders broad-spectrum resistance to powdery mildew. *The Plant Cell* **21**: 2898-2913.

Watanabe, N. and Lam, E. (2004) Recent advance in the study of caspase-like proteases and Bax inhibitor-1 in plants: their possible roles as regulator of programmed cell death. *Molecular Plant Pathology* **5**: 65-70.

Watanabe, N. and Lam, E. (2006) *Arabidopsis* Bax inhibitor-1 functions as an attenuator of biotic and abiotic types of cell death. *The Plant Journal* **45**: 884-894.

Wilson, I. A., Haft, D. H., Getzoff, E. D., Tainer, J. A., Lerner, R. A. and Brenner, S. (1985) Identical short peptide seuqences in unrelated protiens can have different conformations: a testing ground for theories of immune recognition. *Proceedings of the National Academy of Sciences of the United States of America* **82**: 5255-5259.

Xiang, T., Zhong, N., Zou, Y., Wu, Y., Zhang, J., Xing, W., Li, Y., et al. (2008) *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Current Biology* **18**: 74-80.

Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M. and Turner, J. G. (2001) Broad-spectrum mildew resistance in Arabidopsis thaliana mediated by RPW8. *Science* **291**: 118-20.

Xu, Q., Biswas, M. K., Lan, H., Zeng, W., Liu, C., Xu, J. and Deng, X. (2011) Phylogenetic and evolutionary analysis of NBS-encoding genes in Rutaceae fruit crops. *Molecular Genetics and Genomics* **285**: 151-161.

Xu, X. M., Zhao, Q., Rodrigo-Peiris, T., Brkljacic, J., He, C. S., Müller, S. and Meier, I. (2008) RanGAP1 is a continuous marker of the *Arabidopsis* cell division plane. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 18637-18642.

Yamada, K., Wang, J.-C., Osawa, H., Scott, D. K. and Granner, D. K. (1998) Efficient large-scale transformation of yeast. *BioTechniques* **24**: 596-600.

Yi, H. and Richards, E. J. (2007) A cluster of disease resistance genes in *Arabidopsis* is coordinately regulated by transcriptional activation and RNA silencing. *The Plant Cell* **19**: 2929-2939.

Zhang, M., Botör, M., Li, K., Kadota, Y., Panaretou, B., Prodromou, C., Shirasu, K., et al. (2008) Structural and functional coupling of Hsp90- and Sgt1-centered multi-protein complexes. *The EMBO Journal* **27**: 2789-2798.